

PIERRE-ANDRÉ NOËL

**DYNAMIQUES STOCHASTIQUES SUR  
RÉSEAUX COMPLEXES**

Thèse présentée  
à la Faculté des études supérieures et postdoctorales de l'Université Laval  
dans le cadre du programme de doctorat en physique  
pour l'obtention du grade de Philosophiæ Doctor (Ph.D.)

DÉPARTEMENT DE PHYSIQUE, DE GÉNIE PHYSIQUE ET D'OPTIQUE  
FACULTÉ DE SCIENCES ET DE GÉNIE  
UNIVERSITÉ LAVAL  
QUÉBEC

2012



# Résumé

Cette thèse a pour but d'élaborer et d'étudier des modèles mathématiques reproduisant le comportement de systèmes composés de plusieurs éléments dont les interactions forment un réseau complexe. Le corps du document est découpé en trois parties ; un chapitre introductif et une conclusion récapitulative complètent la thèse.

La partie I s'intéresse à une dynamique spécifique (propagation de type susceptible-infectieux-retiré, SIR) sur une classe de réseaux également spécifique (modèle de configuration). Ce problème a entre autres déjà été étudié comme un processus de branchement dans la limite où la taille du système est infinie, fournissant une solution probabiliste pour l'état final de ce processus stochastique. La principale contribution originale de la partie I consiste à modifier ce modèle afin d'introduire des effets dus à la taille finie du système et de permettre l'étude de son évolution temporelle (temps discret) tout en préservant la nature probabiliste de la solution.

La partie II, contenant les principales contributions originales de cette thèse, s'intéresse aux processus stochastiques sur réseaux complexes en général. L'état du système (incluant la structure d'interaction) est partiellement représenté à l'aide de motifs, et l'évolution temporelle (temps continu) est étudiée à l'aide d'un processus de Markov. Malgré que l'état ne soit que partiellement représenté, des résultats satisfaisants sont souvent possibles. Dans le cas particulier du problème étudié en partie I, les résultats sont exacts. L'approche se révèle très générale, et de simples méthodes d'approximation permettent d'obtenir une solution pour des cas d'une complexité appréciable.

La partie III cherche une solution analytique exacte sous forme fermée au modèle développé en partie II pour le problème initialement étudié en partie I. Le système est réexprimé en terme d'opérateurs et différentes relations sont utilisées afin de tenter de le résoudre. Malgré l'échec de cette entreprise, certains résultats méritent mention, notamment une généralisation de la relation de Sack, un cas particulier de la relation de Zassenhaus.



# Abstract

The goal of this thesis is to develop and study mathematical models reproducing the behaviour of systems composed of numerous elements whose interactions make a complex network structure. The body of the document is divided in three parts; an introductory chapter and a recapitulative conclusion complete the thesis.

Part I pertains to a specific dynamics (susceptible-infectious-removed propagation, SIR) on a class of networks that is also specific (configuration model). This problem has already been studied, among other ways, as a branching process in the infinite system size limit, providing a probabilistic solution for the final state of this stochastic process. The principal original contribution of part I consists of modifying this model in order to introduce finite-size effects and to allow the study of its (discrete) time evolution while preserving the probabilistic nature of the solution.

Part II, containing the principal contributions of this thesis, is interested in the general problem of stochastic processes on complex networks. The state of the system (including the interaction structure) is partially represented through motifs, then the (continuous) time evolution is studied with a Markov process. Although the state is only partially represented, satisfactory results are often possible. In the particular case of the problem studied in part I, the results are exact. The approach turns out to be very general, and simple approximation methods allow one to obtain a solution for cases of considerable complexity.

Part III searches for a closed form exact analytical solution to the the model developed in part II for the problem initially studied in part I. The system is re-expressed in terms of operators and different relations are used in an attempt to solve it. Despite the failure of this enterprise, some results deserve mention, notably a generalization of Sack's relationship, a special case of the Zassenhaus relationship.



# Remerciements

Je tiens d'abord à remercier le professeur Louis J. Dubé, mon directeur de recherche au cours de ces travaux de doctorat. Son support constant, autant sur les plans scientifique, financier et moral, a été essentiel à l'achèvement de cette thèse. Si j'ai un jour la chance de diriger des étudiants gradués, je saurai de qui prendre exemple.

Je remercie les Instituts de recherche en santé du Canada (IRSC) pour la « Bourse de recherche au doctorat des IRSC » ayant financé mes travaux pendant la période 2008–2010. Je reconnais également le support financier de monsieur Babak Pourbohloul, directeur du Division of Mathematical Modeling au British Columbia Centre for Disease Control, pour la période allant de l'automne 2006 au printemps 2008.

J'ai eu le privilège de travailler auprès d'Antoine Allard, de Laurent Hébert-Dufresne et de Vincent Marceau, mes collègues au sein du sous-groupe Dynamica::Réseaux. Cet environnement stimulant de collaboration et de camaraderie a été complété par la branche Dynamica::*μcav*, composée de Guillaume Painchaud-April, de Denis Gagnon et de Julien Poirier. Autour d'un café ou d'une bière, tous les problèmes du monde ont été résolus.

À travers leur support, plusieurs autres personnes ont contribué à l'achèvement de cette thèse. Amis, famille immédiate et élargie, ils sont trop nombreux pour être nommés ici. Je ne peux cependant pas taire le nom d'Andréanne, ma compagne au cours de ces années de doctorat. Elle a été présente dans les moments les plus difficiles, et a partagé avec moi les aléas d'une rédaction plus longue que prévue.





# Avant-propos

Cette thèse présente mes travaux de doctorat effectués entre l'automne 2006 et le printemps 2012. Trois articles dont je suis le principal auteur y sont insérés : j'ai effectué la majeure partie des travaux de recherche associés et j'ai été responsable de la rédaction des manuscrits. La thèse est séparée en trois parties, chacune étudiant un aspect différent d'un problème commun. Cette structure est illustrée à la page suivante.

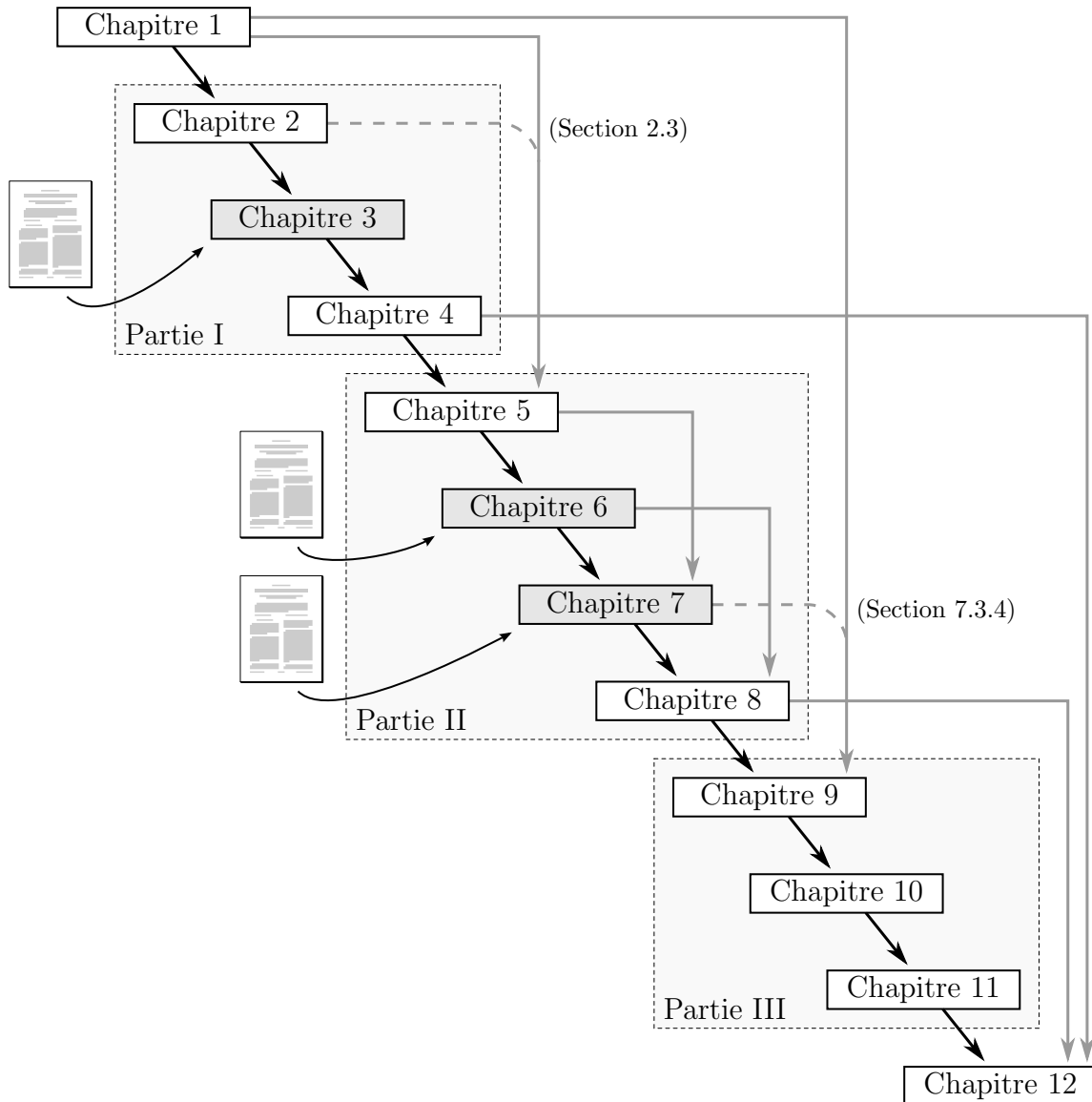
La partie I, composée des chapitres 2–4, présente des travaux effectués jusqu'au printemps 2008. Le chapitre 2 introduit des concepts spécifiques à cette partie et décrit rapidement l'approche utilisée dans l'article [79] (P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé et B. Pourbohloul, publié dans *Physical Review E* 79 : 026101), qui est intégralement repris au chapitre 3. Le chapitre 4 traite de certains points additionnels et discute de ces travaux avec le recul acquis depuis.

La partie II, composée des chapitres 5–8, présente des travaux effectués à la suite de ceux présentés en partie I et constitue la principale contribution de cette thèse. Après une rapide introduction au chapitre 5, les chapitres 6 et 7 présentent respectivement les contributions [76] (P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau et L. J. Dubé, publié dans *Physical Review E* 85 : 031118) et [77] (P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau et L. J. Dubé, soumis pour publication). Le chapitre 8 fournit quelques détails supplémentaires.

La partie III comprend les chapitres 9–11 et présente des travaux effectués de façon concomitante avec ceux de la partie II. Le chapitre 9 présente un problème spécifique, le chapitre 10 mentionne quelques outils jugés pertinents dans le contexte de ce problème, alors que le chapitre 11 tente d'y obtenir une solution. Malgré que ces travaux non publiés se soldent par un constat d'échec, certaines contributions méritant mention y sont présentées.

Le chapitre 1 donne une introduction globale à ces trois parties alors que le chapitre 12 relève les contributions originales et les perspectives futures qui y sont faites.

Si la majeure partie de ce document est écrite en français, les trois articles y sont insérés dans leur langue originale anglaise. Outre l'ajustement de l'orthographe en anglais canadien (par exemple *neighbour* plutôt que *neighbor*) et des éditions mineures de forme, les textes ont été gardés aussi près des originaux que possible. Lorsque nécessaires, des commentaires additionnels sont ajoutés dans les chapitres précédant et/ou suivant ces insertions.



Structure de la thèse et dépendance des chapitres. Le cours normal du texte est indiqué par les flèches noires; quelques alternatives sont présentées en gris. Les deux flèches grises quittant le chapitre 1 peuvent nécessiter quelques lectures supplémentaires, indiquées en pointillé gris. Les insertions d'articles sont faites aux chapitres 3, 6 et 7.

# Table des matières

Résumé	iii
Abstract	v
Remerciements	vii
Avant-propos	ix
Table des matières	xi
Liste des tableaux	xvii
Liste des figures	xix
Liste des contributions	xxi
Lexique	xxv
Glossaire	xxvii
<b>1 Introduction : dynamique et réseaux</b>	<b>1</b>
1.1 Le problème . . . . .	1
1.2 Réseaux . . . . .	2
1.2.1 Terminologie . . . . .	3
1.2.2 Représentations . . . . .	6
1.2.3 Ensembles de réseaux . . . . .	7
1.2.4 Réseaux complexes . . . . .	10
1.2.5 Caractéristiques des réseaux complexes . . . . .	12
1.2.6 Note sur les graphes . . . . .	14
1.3 Dynamique du réseau et sur réseau . . . . .	15
1.3.1 États intrinsèques . . . . .	16
1.3.2 Propagation et modèles épidémiologiques . . . . .	17
1.3.3 Stochasticité . . . . .	19
1.3.4 Modèles et réalité . . . . .	19

<b>I</b>	<b>Processus de branchement et évolution dans le temps</b>	<b>21</b>
<b>2</b>	<b>Préparation aux processus de branchements</b>	<b>23</b>
2.1	Percolation . . . . .	23
2.2	Transmissibilité . . . . .	25
2.3	Fonctions génératrices . . . . .	26
2.4	Prélude à « Time evolution of epidemic disease on finite and infinite networks » . . . . .	28
<b>3</b>	<b>Time evolution of epidemic disease on finite and infinite networks</b>	<b>31</b>
3.1	Introduction . . . . .	32
3.2	Formalism for infinite networks . . . . .	34
3.2.1	Basic generating functions . . . . .	35
3.2.2	Generations . . . . .	36
3.2.3	Phase-space representation . . . . .	37
3.3	Formalism for finite networks . . . . .	40
3.3.1	Evolution of the degree distribution of susceptibles . . . . .	41
3.3.2	Additional loss of transmissions . . . . .	45
3.3.3	Phase-space representation in finite-size networks . . . . .	46
3.4	Further Results and Discussion . . . . .	48
3.4.1	Behaviour around percolation threshold . . . . .	49
3.4.2	Importance of the variety of outcomes . . . . .	51
3.4.3	Effective reproduction number . . . . .	53
3.4.4	Clustering and continuous time evolution . . . . .	54
3.5	Conclusion . . . . .	55
3.A	The ratio $\rho_{s'm'} = \langle \langle \tilde{m} \rangle / m \rangle_{s'm'}$ . . . . .	55
3.A.1	Direct approach . . . . .	56
3.A.2	Continuous approach . . . . .	57
3.B	The average value $\langle n_R \rangle_{s'm'}$ . . . . .	59
3.C	Some dynamical models . . . . .	60
3.C.1	Volz (2008): A network-centric approach . . . . .	60
3.C.2	Moreno <i>et al.</i> (2004): A compartmental approach . . . . .	61
3.C.3	An improved compartmental approach . . . . .	62
<b>4</b>	<b>Temps continu et autres considérations</b>	<b>63</b>
4.1	Corrélations dues à la période infectieuse . . . . .	63
4.2	Temps continu . . . . .	68
4.3	Nouvelle perspective . . . . .	70

<b>II</b>	<b>Processus de Markov et représentation partielle</b>	<b>71</b>
<b>5</b>	<b>Préparation aux processus de Markov</b>	<b>73</b>
5.1	Modéliser un processus de Markov à l'aide d'un processus de Markov plus...	73
5.2	Prélude à « Propagation on networks : an exact alternative perspective »	75
5.3	Prélude à « Spreading dynamics on complex networks : a general... »	76
<b>6</b>	<b>Propagation on networks : an exact alternative perspective</b>	<b>79</b>
6.1	Introduction	80
6.2	The exact model	81
6.2.1	Networks	81
6.2.2	Propagation	81
6.2.3	Equivalent stochastic process	82
6.2.4	Comparison to numerical simulations	84
6.3	Asymptotically large systems	86
6.3.1	Gaussian approximation	86
6.3.2	Branching process approximation	88
6.3.3	Comparison to numerical simulations	91
6.4	Conclusion	93
6.4.1	Generalization	93
6.4.2	Summary and perspective	94
<b>7</b>	<b>Spreading dynamics on complex networks : a general stochastic approach</b>	<b>97</b>
7.1	Introduction	98
7.2	General modelling scheme	99
7.2.1	Networks	100
7.2.2	Motifs	101
7.2.3	Birth-death stochastic processes	102
7.3	Application to spreading dynamics	103
7.3.1	Pair-based SIS model	104
7.3.2	First neighbourhood SIS model	109
7.3.3	First neighbourhood SIR model	111
7.3.4	First neighbourhood on-the-fly SIR model	112
7.4	Discussion	115
7.4.1	Accuracy of the results	115
7.4.2	Pair-based models	117
7.4.3	First (and higher) neighbourhood models	118
7.4.4	On-the-fly models	118
7.4.5	Complicated states <i>vs</i> complex assumptions	119

7.4.6	Additional inference tools	120
7.5	Conclusion: general applicability	120
7.A	More on motifs	121
7.A.1	Pair motifs	122
7.A.2	Triple motifs	123
7.A.3	Triangle motifs and other small subnetworks	123
7.A.4	Clique motifs	123
7.A.5	Degreed motifs	124
7.A.6	$n$ -th neighbourhood motifs	124
7.A.7	Coarse-grained degree and/or neighbourhood	125
7.B	Deterministic solution of on-the-fly SIR model	125
<b>8</b>	<b>Modélisation alternative et solution directe</b>	<b>129</b>
8.1	Processus de branchement	130
8.2	Maximisation d'entropie	133
8.3	Solution directe des systèmes linéaires	135
<b>III</b>	<b>Approche opérationnelle</b>	<b>137</b>
<b>9</b>	<b>Énoncé du problème</b>	<b>139</b>
9.1	Définitions	139
9.2	Traduction de (7.10)	141
9.3	Forme compacte	143
9.4	Traitement de la normalisation	145
9.5	Traitement alternatif de la normalisation	147
9.6	Forme « demi-liens infectieux »	148
<b>10</b>	<b>Outillage</b>	<b>151</b>
10.1	Algèbres de Lie : un aperçu	151
10.2	Exponentielles d'opérateurs	152
10.3	Méthode de Magnus	154
10.4	Une généralisation de l'identité de Sack	157
<b>11</b>	<b>Application au problème</b>	<b>159</b>
11.1	Désintégration spontanée d'un élément	159
11.2	Annihilation de deux éléments identiques	161
11.3	Cas mixte désintégration-annihilation	161
11.4	Dégradation spontanée d'un élément	163
11.5	Cas mixte dégradation-annihilation	166
11.6	Processus de branchement	168
11.7	Constat d'échec	169

<b>12 Conclusion</b>	<b>171</b>
12.1 Processus de branchement et évolution dans le temps . . . . .	171
12.2 Processus de Markov et représentation partielle . . . . .	172
12.3 Approche opérationnelle . . . . .	172
<b>Bibliographie</b>	<b>175</b>
<b>Index</b>	<b>183</b>





# Liste des tableaux

3.1	Forme de $\theta(s)$ pour quelques distributions de degré courantes. . . . .	43
7.1	Number of motifs in selected classes. . . . .	122



# Liste des figures

1.1	Terminologie relative aux liens. . . . .	3
1.2	Terminologie relative aux degrés. . . . .	4
1.3	Terminologie relative aux chemins. . . . .	5
1.4	Terminologie relative aux composantes. . . . .	5
1.5	Représentation par matrice d'adjacence. . . . .	6
1.6	Modèle de Watts et Strogatz. . . . .	9
1.7	Club de karaté de Zachary. . . . .	12
3.1	Représentation de l'espace de phase $(s, m)$ pour un réseau de taille <i>infinie</i> . . . . .	39
3.2	Projection $p_s^g$ sur l'axe $s$ de l'espace de phase $(s, m)$ pour un réseau de taille <i>infinie</i> . . . . .	40
3.3	Évolution temporelle de quelques distributions en degrés typiques pour les nœuds susceptibles. . . . .	44
3.4	Représentation de l'espace de phase $(s, m)$ pour un réseau de taille <i>finie</i> . . . . .	47
3.5	Projection $p_s^g$ sur l'axe $s$ de l'espace de phase $(s, m)$ pour un réseau de taille <i>finie</i> . . . . .	48
3.6	Nombre moyen d'infectés en fonction de la transmissibilité pour différents modèles. . . . .	50
3.7	Interprétation en terme du nombre de reproduction effectif. . . . .	53
4.1	Distributions de périodes infectieuses (corrélation). . . . .	65
4.2	Distribution de nouvelles infections (corrélation). . . . .	66
4.3	Distribution du nombre d'infections (corrélation). . . . .	67
6.1	Illustration du processus <i>à-la-volée</i> . . . . .	83
6.2	Comparaison de l'évolution temporelle du modèle à-la-volée avec des simulations numériques ( $N = 30$ ). . . . .	85
6.3	Comparaison de l'évolution temporelle du modèle à-la-volée avec des simulations numériques ( $N = 300$ ). . . . .	86
6.4	État final du processus à-la-volée ( $N = 3000$ ). . . . .	92
7.1	Distribution finale pour une dynamique SIS sur un réseau régulier. . . . .	106

7.2	Évolution temporelle du nombre de nœuds infectieux pour une dynamique SIS sur un réseau régulier. . . . .	107
7.3	Distribution de probabilité à différent temps pour le nombre de nœuds infectieux. . . . .	107
7.4	Évolution temporelle du nombre de nœuds infectieux pour une dynamique SIS sur un réseau d'Erdős-Rényi. . . . .	109
7.5	Évolution temporelle du nombre de nœuds infectieux pour une dynamique SIS sur un modèle de configuration. . . . .	111
7.6	Évolution temporelle du nombre de nœuds infectieux et retirés pour une dynamique SIR sur un modèle de configuration. . . . .	114
7.7	Distribution de probabilité à différent temps pour le nombre de nœuds infectieux et retirés. . . . .	114
8.1	Réseau biparti d'éléments et de motifs. . . . .	131

# Liste des contributions

## Articles

- O. P. Günther, G. Ogilvie, M. Naus, E. Young, D. M. Patrick, S. Dobson, B. Duval, **P.-A. Noël**, F. Marra, D. Miller, R. C. Brunham, et B. Pourbohloul. Protecting the Next Generation : What Is the Role of the Duration of Human Papillomavirus Vaccine-Related Immunity? *J. Infect. Dis.*, 197 :1653-1661, 2008.
- **P.-A. Noël**, B. Davoudi, R. C. Brunham, L. J. Dubé, et B. Pourbohloul. Time evolution of epidemic disease on finite and infinite networks. *Phys. Rev. E*, 79 :026101, 2009. Repris au chapitre 3 de cette thèse.
- A. Allard, **P.-A. Noël**, L. J. Dubé, et B. Pourbohloul. Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. *Phys. Rev. E*, 79 :036113, 2009.
- A. Allard, **P.-A. Noël**, et L. J. Dubé. Des ponts d'Euler à la grippe aviaire. *Accromaxth (publié par l'Institut des sciences mathématiques, Montréal QC, Canada)*, 4 (Hiver-printemps) :24, 2009.
- L. Hébert-Dufresne, **P.-A. Noël**, V. Marceau, A. Allard, et L. Dubé. Propagation dynamics on networks featuring complex topologies. *Phys. Rev. E*, 82, 036115, 2010.
- V. Marceau, **P.-A. Noël**, L. Hébert-Dufresne, A. Allard, et L. Dubé. Adaptive networks : Coevolution of disease and topology. *Phys. Rev. E*, 82 :036116, 2010.
- V. Marceau, **P.-A. Noël**, L. Hébert-Dufresne, A. Allard, et L. Dubé. Modeling the dynamical interaction between epidemics on overlay networks. *Phys. Rev. E*, 84 :026105, 2011.
- L. Hébert-Dufresne, A. Allard, V. Marceau, **P.-A. Noël**, et L. Dubé. Structural Preferential Attachment : Network Organization beyond the Link *Phys. Rev. Lett.*, 107 :158702, 2011.
- L. Hébert-Dufresne, A. Allard, V. Marceau, **P.-A. Noël**, et L. Dubé. Structural preferential attachment : Stochastic process for the growth of scale-free, modular, and self-similar systems. *Phys. Rev. E*, 85 :026108, 2012.
- **P.-A. Noël**, A. Allard, L. Hébert-Dufresne, V. Marceau, et L. J. Dubé. Propagation on networks : an exact alternative perspective. *Phys. Rev. E*, 85 :031118, 2012. Repris au chapitre 6 de cette thèse.

- **P.-A. Noël**, A. Allard, L. Hébert-Dufresne, V. Marceau, et L. J. Dubé. Spreading dynamics on complex networks : a general stochastic approach. *arXiv :1201.0296*, 2012. Soumis pour publication. Repris au chapitre 7 de cette thèse.
- A. Allard, L. Hébert-Dufresne, **P.-A. Noël**, V. Marceau, et L. Dubé. Bond percolation on a class of correlated and clustered random graphs. *J. Phys. A*, 45 :405005, 2012.
- A. Allard, L. Hébert-Dufresne, **P.-A. Noël**, V. Marceau, et L. Dubé. Exact solution of bond percolation on small arbitrary graphs. *EPL*, 98 :16001, 2012.
- L. Hébert-Dufresne, V. Marceau, **P.-A. Noël**, A. Allard, et L. Dubé. The Social Zombie : Modelling undead outbreaks on social networks. Robert Smith? (Ed.) University of Ottawa Press. Sous presse.

## Présentations

- **P.-A. Noël**, B. Davoudi, R. C. Brunham, L. J. Dubé, et B. Pourbohloul. Time Evolution of Epidemic Dynamics on Finite and Infinite Networks. Congrès de l'ACP, Québec QC, Canada, 2008.
- A. Allard, **P.-A. Noël**, L. J. Dubé, et B. Pourbohloul. Heterogeneous Bond Percolation on Complex Networks : Application to Epidemiology. Congrès de l'ACP, Québec QC, Canada, 2008.
- **P.-A. Noël**, A. Allard, et L. J. Dubé. Time Evolution of Epidemics on Complex Networks. SIAM Conference on Applications of Dynamical Systems, Snowbird UT, US, 2009.
- **P.-A. Noël**, A. Allard, et L. J. Dubé. Exact(?) SIR dynamics on networks of finite size. NetSci 2010, Boston MA, US, 2010.
- A. Allard, **P.-A. Noël**, L. Hébert-Dufresne, V. Marceau, et L. Dubé. Modular multitype networks as a model of clustered social networks. NetSci 2010, Boston MA, US, 2010.
- V. Marceau, **P.-A. Noël**, L. Hébert-Dufresne, A. Allard, et L. Dubé. Adaptative networks : Coevolution of disease and topology. NetSci 2010, Boston MA, US, 2010.
- L. Hébert-Dufresne, **P.-A. Noël**, V. Marceau, A. Allard, et L. Dubé. Propagatory dynamics on networks featuring complex topology. NetSci 2010, Boston MA, US, 2010.
- **P.-A. Noël**, A. Allard, L. Hébert-Dufresne, V. Marceau, et L. Dubé. A study of network representations for Markov dynamics modelling. NetSci 2011, Budapest, Hongrie, 2011.
- V. Marceau, **P.-A. Noël**, L. Hébert-Dufresne, A. Allard, et L. Dubé. Modeling the dynamical interaction between epidemics on overlay networks. NetSci 2011, Budapest, Hongrie, 2011.

- 
- L. Hébert-Dufresne, A. Allard, V. Marceau, **P.-A. Noël**, et L. Dubé. Structural preferential attachment : network organisation beyond the link. NetSci 2011, Budapest, Hongrie, 2011.
  - **P.-A. Noël**, A. Allard, L. Hébert-Dufresne, V. Marceau, et L. Dubé. Stochastic Network Models. ISSTDR 2011, Québec QC, Canada, 2011.





# Lexique

Ce lexique fournit les équivalents francophones et anglophones de quelques termes utilisés au cours de cette thèse.

---

Terme français	Terme anglais	Page	Remarque
arbre	<i>tree(-like)</i>	6	
arête	<i>edge</i>	14	Lien (terme mathématique).
boucle	<i>(self-)loop</i>	4	
chemin	<i>path</i>	4	
composante	<i>component</i>	6	
cycle	<i>cycle</i>	4	
demi-lien	<i>stub</i>	4	Traduction libre.
degré	<i>degree</i>	4	
flambée	<i>outbreak</i>	18	J'utilise très peu le terme français.
gelé	<i>quenched</i>	9	Traduction libre. Alternative : trempé.
graphe	<i>graph</i>	14	Réseau (terme mathématique).
lien	<i>link</i>	2	
liens répétés	<i>repeated links</i>	4	
mouvant	<i>annealed</i>	10	Traduction libre. Alternative : recuit.
nœud	<i>node</i>	2	
réseau	<i>network</i>	2	
sommet	<i>vertex</i>	14	Nœud (terme mathématique).
voisin	<i>neighbour</i>	4	

---



# Glossaire

Ce glossaire mathématique décrit certaines expressions utilisées au cours de cette thèse.

## Objets généraux : réseaux.

$k$	Degré d'un nœud (nombre de voisins).
$n_k$	Nombre de nœuds de degré $k$ . Doit respecter $\langle n_k \rangle = Np_k$ (cas particulier $\sum_k n_k = N$ ) et la quantité $\sum_k kn_k$ doit être un nombre pair.
$N$	Nombre de nœuds dans le système.
$p_k$	Probabilité que le degré d'un nœud sélectionné au hasard soit $k$ . La séquence $\{p_0, p_1, p_2, \dots\}$ est la distribution en degré.

## Objets généraux : épidémiologie.

$E$	Exposé (porte l'infection mais ne peut pas encore la transmettre).
$g$	Génération (distance du patient zéro en terme de nombre de transmissions).
$I$	Infectieux (portant l'infection et pouvant la transmettre).
$R$	Retiré (a déjà porté l'infection, et ne peut pas l'avoir à nouveau).
$R_0$	Nombre de reproduction de base (nombre moyen de nouvelles infections découlant directement d'un nœud infectieux). Une épidémie est possible si $R_0 > 1$ .
$S$	Susceptible (non infecté, mais pourrait le devenir).
$T$	Transmissibilité (probabilité qu'un nœud susceptible acquière l'infection d'un voisin infectieux).
$\alpha$	Taux auquel les nœuds cessent d'être infectieux. Cas $SIR$ : taux de retrait. Cas $SIS$ : taux de resusceptibilisation.
$\beta$	Taux d'infection.

**Objets généraux : probabilités.**

$P(x)$	Probabilité que $x$ survienne. La notation utilisée dans cette thèse ne fait pas de distinction entre une variable aléatoire et la valeur que peut prendre cette variable (et donc $P(x) = P(X = x)$ ).
$P(x, y)$	Probabilité jointe que $x$ et $y$ surviennent.
$P(x y)$	Probabilité conditionnelle que $x$ survienne sachant $y$ .
$\langle x \rangle$	Moyenne de $x$ , donnée par $\langle x \rangle = \sum_x xP(x) = E(x)$ .
$\langle x^2 \rangle$	Second moment de $x$ , donné par $\langle x^2 \rangle = \sum_x x^2P(x) = E(x^2)$ .
$\text{cov}(\mathbf{x})$	Matrice de covariance de $\mathbf{x}$ , donnée par $\text{cov}(\mathbf{x}) = \langle (\mathbf{x} - \langle \mathbf{x} \rangle) \cdot (\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle$ .

**Objets spécifiques au chapitre 3.**

$G_0(x)$	Fonction génératrice de la distribution en degré, donnée par $G_0(x) = \sum_k p_k x^k$ .
$G_1(x)$	Fonction génératrice du degré excédentaire d'un nœud atteint en suivant un lien (excluant ce dernier lien lors du calcul du degré, d'où "degré excédentaire"), donnée par $G_1(x) = \frac{1}{z_1} G'_0(x) = \sum_k q_k x^k$ .
$G_g(x)$	Fonction génératrice du nombre de voisins pouvant potentiellement être infectés à la génération $g$ .
$H_0(x)$	Fonction génératrice du nombre final d'infections causées par un nœud infectieux choisi au hasard.
$H_1(x)$	Fonction génératrice du nombre final d'infections causées par un nœud rendu infectieux en suivant un lien.
$l$	Nombre de nouvelles infections causées par un nœud spécifique.
$m$	Nombre d'infections ayant lieu à une génération donnée.
$p_s^g$	Probabilité que le nombre d'infections entre les générations 0 et $g$ soit $s$ .
$p_m^g$	Probabilité que le nombre d'infections à la génération $g$ soit $m$ .
$q_k$	Probabilité que le degré excédentaire d'un nœud atteint en suivant un lien soit $k$ .
$s$	Nombre d'infections ayant eu lieu jusqu'ici.
$S$	Fraction des nœuds affectés par une épidémie.
$z_1$	Degré moyen, donné par $z_1 = \langle k \rangle = \sum_k k p_k = G'_0(1)$ .
$z_2$	Nombre moyen de second voisins, donné par $z_2 = G''_0(1)$ .
$\theta(s)$	Fonction utilisée pour obtenir les quantités spécifiques aux nœuds susceptibles (par exemple $p_k^S(s) = p_k \frac{N-1}{N-s} [\theta(s)]^k$ ).
$\rho_{sm}$	Fonction tenant compte de la perte additionnelle de transmissions due à la taille finie (à travers $\tilde{T}_{sm} = T \rho_{sm}$ ).

$\psi_{sm}^g$	Probabilité que $s$ nœuds soient infectés à la génération $g$ , $m$ d'entre eux venant tout juste d'être infectés.
$\Psi_0^g(x, y)$	Fonction génératrice de $\psi_{sm}^g$ , donnée par $\Psi_0^g(x, y) = \sum_{sm} \psi_{sm}^g x^s y^m$ .
$\square^S$	Tel que pour $\square$ , mais spécifique aux nœuds susceptibles (par exemple $p_k^S(s)$ ).
$\square^I$	Tel que pour $\square$ , mais spécifique aux nœuds infectieux (par exemple $G_0^I(x; s, m)$ ).
$\tilde{\square}$	Tel que pour $\square$ , mais tenant compte de la taille finie (par exemple $\tilde{\Psi}_0^g(x, y)$ ).

## Objets spécifiques au chapitre 6.

$\mathbf{a}(\mathbf{x})$	Contribution déterministe à l'évolution de $\mathbf{x}$ , donnée par $\mathbf{a}(\mathbf{x}) = \sum_j \mathbf{r}^j q_j(\mathbf{x})$ .
$B(\mathbf{x})$	Contribution stochastique à l'évolution de $\mathbf{x}$ , donnée par $B_i^j(\mathbf{x}) = r_i^j \sqrt{q_j(\mathbf{x})}$ .
$f_n(\xi, \zeta)$	Fonction génératrice définie en (6.19).
$g_0(\xi)$	Fonction génératrice de la distribution en degré, donnée par $g_0(\xi) = \sum_k p_k \xi^k$ .
$g_1(\xi)$	Fonction génératrice du degré excédentaire d'un nœud atteint en suivant un lien, donnée par $g_1(\xi) = g'_0(\xi)/g'_0(1)$ .
$h_n(\xi, \zeta)$	Fonction génératrice définie en (6.20).
$\tilde{h}_n(\xi, \zeta)$	Fonction génératrice définie en (6.21).
$i$	Index des éléments de $\mathbf{x}$ , contraint par $-1 \leq i \leq k_{\max}$ .
$I_0$	Nombre de nœuds initialement infectieux.
$j$	Index spécifiant un type d'événements pouvant survenir : $j = -1$ correspond à l'appariement de deux demi-liens infectieux alors que $j = k \geq 0$ correspond à l'appariement d'un demi-lien appartenant à un nœud de degré $k$ avec un demi-lien infectieux.
$k_{\max}$	Degré maximal permis dans le réseau.
$q_j(\mathbf{x})$	Taux auquel les événements de type $j$ surviennent lorsque le système est dans l'état $\mathbf{x}$ .
$\mathbf{r}^j$	Effet sur $\mathbf{x}$ d'un événement de type $j$ ( $\mathbf{x}$ devient $\mathbf{x} + \mathbf{r}^j$ ).
$u$	Probabilité qu'une infection n'entraîne au final qu'un nombre fini d'infections, donnée par le plus petit $u \geq 0$ satisfaisant $u = g_1(u)$ .
$\mathbf{W}(t)$	Processus de Wiener (mouvement brownien standard), de même dimension que $\mathbf{x}$ .
$\mathbf{x}$	Vecteur d'état $\mathbf{x} = [x_{-1} \ x_0 \ x_1 \ x_2 \ \cdots \ x_{k_{\max}}]^T$ . L'élément $x_{-1}$ spécifie le nombre de demi-liens non assignés alors que les éléments $x_k$ pour $k \geq 0$ spécifient le nombre de nœuds susceptibles de degré $k$ .
$\theta$	« Paramètre temporel » utilisé dans la solution déterministe de $\boldsymbol{\mu}(t)$ .
$\lambda(\mathbf{x})$	Nombre de demi-liens infectieux non assignés, soit $\lambda(\mathbf{x}) = x_{-1} - \sum_{k=0}^{k_{\max}} k x_k$ .

$\lambda_0$	Nombre initial de demi-liens infectieux, soit $\lambda_0 = \lambda(\mathbf{x}(0))$ .
$\boldsymbol{\mu}(t)$	Approximation déterministe de $\mathbf{x}$ , soit $\boldsymbol{\mu}(t) \approx \langle \mathbf{x}(t) \rangle$ .
$\boldsymbol{\nu}(t)$	Perturbation stochastique de $\mathbf{x}$ par rapport à sa moyenne, soit $\boldsymbol{\nu}(t) \approx \mathbf{x}(t) - \langle \mathbf{x}(t) \rangle$ .

## Objets spécifiques au chapitre 7.

$i$	Index spécifiant un type de motif (spécifie les éléments de $\mathbf{x}$ ).
$j$	Index spécifiant un type d'événement.
$\mathcal{L}$	Nombre d'états intrinsèques accessibles aux liens.
$\mathcal{N}$	Nombre d'états intrinsèques accessibles aux nœuds.
$q_j^+(\mathbf{x}, Y)$	Taux auquel les événements de type $j$ surviennent en sens direct.
$q_j^-(\mathbf{x}, Y)$	Taux auquel les événements de type $j$ surviennent en sens inverse.
$\mathbf{r}^j$	Effet d'un événement de type $j$ : $\mathbf{x}$ devient $\mathbf{x} + \mathbf{r}^j$ (sens direct) ou $\mathbf{x} - \mathbf{r}^j$ (sens inverse).
$V$	Règles gouvernant le système complet.
$W$	Règles gouvernant le système simplifié.
$\mathbf{x}$	Vecteur d'état, l'élément $x_i$ spécifiant le nombre de motifs de type $i$ .
$X$	État du système simplifié.
$Y$	Information préalable, utilisée lors de l'inférence $P(Z X, Y)$ .
$Z$	État du système complet.
$\ell$	Index caractérisant l'état intrinsèque d'un lien ( $\mathcal{L}$ valeurs possibles).
$\nu$	Index caractérisant l'état intrinsèque d'un nœud ( $\mathcal{N}$ valeurs possibles).
<i>(Motifs.)</i>	
$\nu$	Motif de type <i>nœud</i> , composé d'un nœud d'état intrinsèque $\nu$ .
$\nu-\nu'$	Motif de type <i>lien</i> , composé d'un nœud d'état intrinsèque $\nu$ lié à un nœud d'état intrinsèque $\nu'$ .
$\nu\Gamma_1(\mathbf{k})$	Motif de premier voisinage, composé d'un nœud de type $\nu$ et de ses voisins immédiats (spécifiés par $\mathbf{k}$ ).
$\nu\Gamma_2(\mathbf{K})$	Motif de second voisinage, composé d'un nœud de type $\nu$ , de ses voisins, et des voisins de ces voisins (spécifiés par $\mathbf{K}$ ).
$\nu\tilde{\Gamma}_1(\mathbf{K})$	Motif de premier voisinage centré sur le même nœud que $\nu\Gamma_2(\mathbf{K})$ .
$\nu\Lambda_1(\kappa)$	Motif <i>on-the-fly</i> (à-la-volée), composé d'un nœud de type $\nu$ possédant $\kappa$ demi-liens non assignés.

## Objets spécifiques aux chapitres 9 et 11.

$\hat{A}_m^n \cdots \hat{F}_m^n$	Opérateurs compacts, voir (9.21).
$\hat{a}, \hat{a}^\dagger$	Opérateurs d'annihilation et de création affectant $\lambda$ .
$\hat{a}_k, \hat{a}_k^\dagger$	Opérateurs d'annihilation et de création affectant le sous-espace $i_k$ .
$\hat{b}_k, \hat{b}_k^\dagger$	Opérateurs d'annihilation et de création affectant le sous-espace $s_k$ .
$\hat{c}_k, \hat{c}_k^\dagger$	Opérateurs d'annihilation et de création affectant le sous-espace $r_k$ .
$i_k$	Nombre de nœuds infectieux possédant $k$ demi-liens non assignés.
$\hat{L}$	Opérateur d'évolution. Fournit $ \psi(t)\rangle = e^{t\hat{L}} \psi(0)\rangle$ .
$r_k$	Nombre de nœuds retirés possédant $k$ demi-liens non assignés.
$s_k$	Nombre de nœuds susceptibles possédant $k$ demi-liens non assignés.
$v_k$	Variable remplaçant l'opérateur $\hat{b}_k$ dans $\hat{L}'_\lambda$ .
$w$	Variable remplaçant l'opérateur $\hat{\Omega}_m$ dans $\hat{L}_w, \hat{L}'_w, \hat{L}''_w, \hat{L}'_\lambda$ et $\hat{L}''_\lambda$ .
$\mathbf{x}$	Vecteur d'état dont les entrées spécifient tous les $s_k, i_k$ et $r_k$ .
$ \mathbf{x}\rangle$	Ket associé à l'état $\mathbf{x}$ . Base de l'espace vectoriel où vit $ \psi(t)\rangle$ .
$\lambda$	Nombre de demi-liens infectieux non assignés (pour les variantes $\hat{L}'_\lambda$ et $\hat{L}''_\lambda$ ).
$ \lambda\rangle$	Ket associé à l'état où $\lambda$ demi-liens infectieux sont non assignés.
$ \psi(t)\rangle$	Ket spécifiant la distribution de probabilité pour tous les états accessibles.
$\hat{\Omega}_m$	Opérateur de normalisation, voir (9.11).
$[\hat{X}, \hat{Y}]$	Commutateur des opérateurs $\hat{X}$ et $\hat{Y}$ , donné par $[\hat{X}, \hat{Y}] = \hat{X}\hat{Y} - \hat{Y}\hat{X}$ .

(Différentes formes de l'opérateur d'évolution.)

$$\hat{L} = \alpha(\hat{E}_0^0 - \hat{A}_0^0) + \left[ (\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01})\hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 + \hat{C}_0^1 - 1)\hat{A}_0^1 \right] \hat{\Omega}_1 \quad (9.20)$$

$$\hat{L}' = \left[ (\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 - 1)\hat{A}_0^1 \right] \hat{\Omega}_1 \quad (9.33)$$

$$\hat{L}_w = w(\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01})\hat{A}_1^{01} + \alpha(\hat{E}_0^0 - \hat{A}_0^0) - \hat{A}_0^1 \quad (9.32a)$$

$$\hat{L}'_w = w(\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - \hat{A}_0^1 \quad (9.32b)$$

$$\hat{L}''_w = w\hat{A}_1^{01}\hat{A}_1^{01} - \hat{A}_0^1 \quad (9.32c)$$

$$\hat{L}'_\infty = (\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 - 1)\hat{A}_0^1 \quad (9.34a)$$

$$\hat{L}''_\infty = \hat{A}_1^{01}\hat{A}_1^{01} - (\hat{A}_0^1 - 1)\hat{A}_0^1 \quad (9.34b)$$

$$\hat{L}'_\lambda = w \left( \sum_k k v_k (\hat{a}^\dagger)^{k-1} + \hat{a} \right) \hat{a} - \hat{a}^\dagger \hat{a} \quad (9.36a)$$

$$\hat{L}''_\lambda = w \hat{a}^2 - \hat{a}^\dagger \hat{a} \quad (9.36b)$$





# Chapitre 1

## Introduction : dynamique et réseaux

On présente la problématique de façon très générale puis on introduit des termes et concepts qui seront utilisés tout au long de cette thèse : réseaux, complexité, dynamique, stochasticité *etc.*

### 1.1 Le problème

La physique statistique étudie comment des règles relativement simples régissant les interactions de nombreux éléments (particules, atomes, molécules, ...) se traduisent à l'échelle globale d'un système. Dans la plupart des cas traditionnellement considérés, les interactions entre deux éléments constitutifs donnés dépendent fortement de la distance spatiale les séparant. Ainsi, qu'il s'agisse de la diffusion d'un composé dans un liquide ou de l'orientation de spins dans une structure cristalline, la dimensionnalité du système est souvent l'un des paramètres importants du problème.

Cependant, il existe une vaste classe de systèmes d'intérêt pour lesquels les interactions entre les éléments dépendent d'une *structure complexe* n'ayant pas d'équivalent géométrique simple : leur étude relève de la *science des réseaux* [9, 11, 16, 28, 68, 74]. Cette jeune discipline au caractère fondamentalement multidisciplinaire s'attarde autant aux systèmes technologiques (Internet, téléphones cellulaires, distribution d'électricité, ...) biologiques (écosystèmes, neurones, interactions protéine-protéine, ...) ou sociaux (connaissances, rencontres face-à-face, relations sexuelles, ...). La grande diversité de ces domaines d'applications est justifiée par des observations empiriques : les structures

complexes sous-jacentes à ces systèmes possèdent souvent des similarités surprenantes.

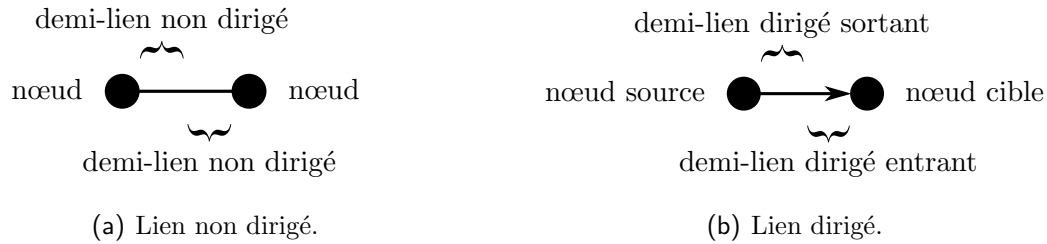
À l’instar de la physique statistique, on recherche comment le comportement des éléments constitutifs affecte le système dans son ensemble. Puisque le but est le même, plusieurs outils de physique statistique peuvent être récupérés. Bien entendu, certaines adaptations sont nécessaires : les concepts de géométrie et de dimensionnalité sont remplacés par des caractéristiques propres aux structures complexes considérées et les effets de taille finie sont souvent plus importants. D’autre part, les éléments fondamentaux eux même seront souvent plus complexes : par exemple, les règles gouvernant les décisions d’un être humain sont difficilement comparables à celles régissant le mouvement d’un électron. Néanmoins, l’expérience révèle que même des modèles basés sur des approximations simplistes peuvent aider à prédire, voire même contrôler, le comportement de tels systèmes.

Voilà donc la ligne directrice de cette thèse : produire ou améliorer des outils de modélisation qui sont adaptés aux problèmes rencontrés en science des réseaux. Lorsque possible, on préférera récupérer les techniques pertinentes de physique statistique ou d’autres domaines. Lorsque nécessaire, de nouvelles approches seront envisagées.

Le reste de ce chapitre introductif présente certains concepts fondamentaux qui s’avéreront utiles dans cette démarche. Ainsi, la section 1.2 définit la terminologie permettant de qualifier la structure complexe des systèmes étudiés ; quelques-unes des caractéristiques « universelles » typiquement observées en science des réseaux sont également présentées. De façon similaire, la section 1.3 s’intéresse aux processus pouvant prendre place dans ces systèmes. Entre autre, on introduit quelques concepts de base en modélisation épidémiologique ; cette classe de problèmes s’avère suffisamment complète pour servir d’exemple d’application tout au long de cette thèse.

## 1.2 Réseaux

Un *réseau* est un ensemble de *nœuds* et de *liens*. La *taille* d’un réseau correspond au nombre de nœuds le composant. Dans un contexte typique de modélisation, chacun des nœuds du réseau représente l’un des éléments constitutifs du système modélisé ; la présence d’un lien entre deux nœuds indique une interaction entre les deux éléments correspondants du système. La signification exacte du terme « interaction » est spécifique au modèle : il pourrait par exemple s’agir d’un câble réseau reliant deux ordinateurs ou encore de deux personnes se croisant dans un ascenseur.



**Figure 1.1** – Terminologie relative aux liens.

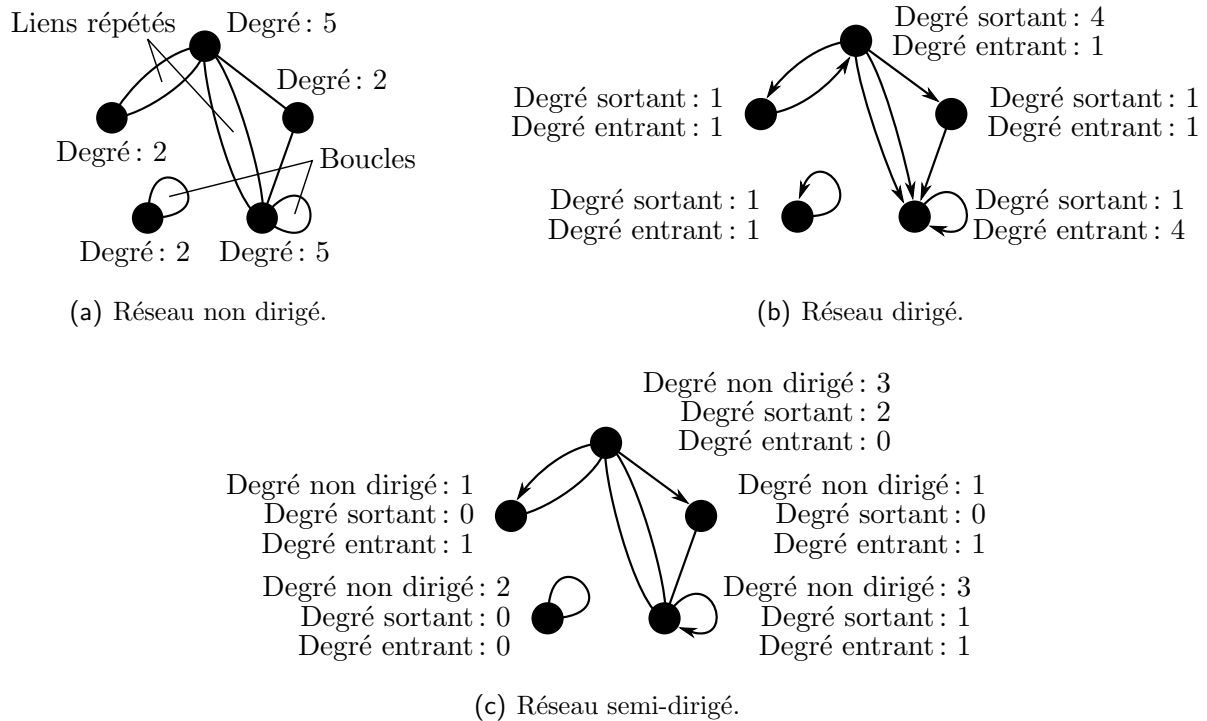
La section 1.2.1 introduit quelques termes généraux relatifs aux réseaux, la section 1.2.2 présente quelques façons de représenter un réseau, la section 1.2.3 introduit les ensembles de réseaux, la section 1.2.4 précise ce que l'on entend par « réseaux complexes », la section 1.2.5 énumère quelques caractéristiques spécifiques aux réseaux complexes et la section 1.2.6 adresse quelques points destinées aux lecteurs issus du milieu mathématique.

### 1.2.1 Terminologie

Un *lien non dirigé* exprime une relation (*a priori*) symétrique entre les nœuds qu'il unit. Schématiquement, on dépeint un lien non dirigé par une ligne joignant deux points, ces derniers représentant des nœuds. Par opposition, un *lien dirigé* unit un *nœud source* à un *nœud cible* dans sa direction *sortante*, de même qu'il unit le nœud cible au nœud source dans sa direction *entrante*<sup>1</sup>. Schématiquement, on représente un lien dirigé par une flèche quittant le nœud source et pointant vers le nœud cible. Chacun des nœuds unis par un lien non dirigé sont *voisins non dirigés* l'un de l'autre ; le nœud cible (resp. source) d'un lien dirigé est *voisin sortant* (resp. *voisin entrant*) du nœud source (resp. cible). Un lien non dirigé est composé de deux *demi-liens non dirigés*, chacun étant rattaché à l'un des nœuds unis par ce lien ; un lien dirigé est composé d'un *demi-lien sortant* et d'un *demi-lien entrant*, le premier étant rattaché au nœud source et le second au nœud cible. La figure 1.1 illustre ces définitions.

Le *degré non dirigé*, le *degré sortant* et le *degré entrant* d'un nœud sont respectivement donnés par le nombre de demi-liens non dirigés, de demi-liens sortants et de demi-liens entrants rattachés à ce nœud. Un réseau *non dirigé* (resp. *dirigé*) ne contient que des liens non dirigés (resp. dirigés) ; un réseau *semi-dirigé* peut à la fois contenir

1. Dans le reste de cette section, on considère une brisure complète de la symétrie entre les deux directions d'un lien dirigé : le nœud source peut avoir un effet causal sur le nœud cible, mais pas l'inverse. Il est cependant à noter qu'une brisure partielle de symétrie est possible dans un contexte général de modélisation.

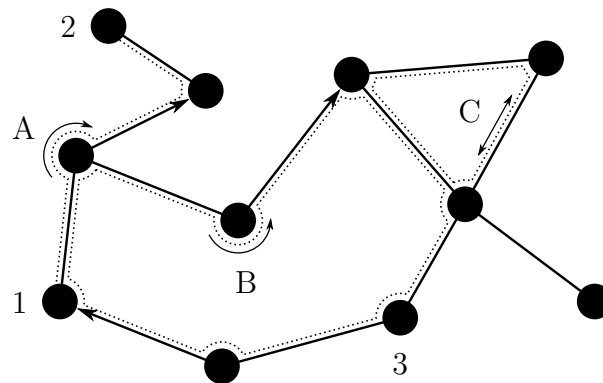


**Figure 1.2** – Terminologie relative aux degrés.

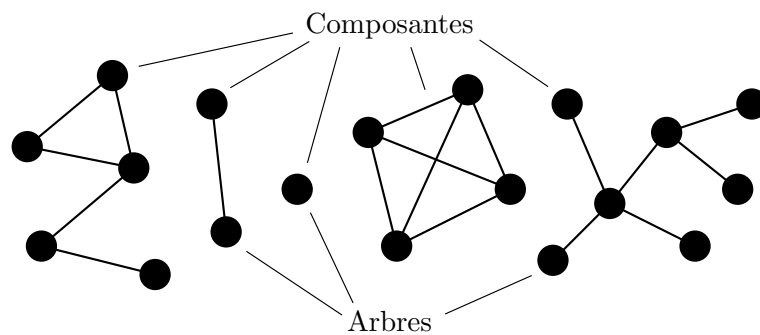
des liens non dirigés et dirigés. Lorsqu'il est clair que seulement des liens non dirigés sont possibles, il est commun d'omettre le terme « non dirigé » qualifiant ces termes : on parle alors simplement de réseaux, de liens, de voisins, de demi-liens et de degrés. Des *liens répétés* joignent les mêmes deux nœuds ; une *boucle* est un lien joignant un nœud à lui-même. En absence de liens répétés et de boucles, les degrés non dirigé, sortant et entrant d'un nœud correspondent respectivement au nombre de voisins non dirigés, sortants et entrants de ce nœud. La figure 1.2 illustre ces définitions.

Un *chemin* entre un *nœud de départ* et un *nœud d'arrivée* est une séquence de nœuds débutant par le nœud de départ et se terminant par le nœud d'arrivée telle qu'un nœud est voisin non dirigé ou voisin sortant du nœud le précédent dans la séquence ; soustraire un au nombre d'éléments dans cette séquence fournit la *longueur* du chemin. La *distance* entre deux nœuds est la longueur du plus court chemin permettant de les joindre ; la distance entre un nœud et lui-même est de zéro et on définit comme infinie la distance entre deux nœuds ne pouvant pas être joints par un chemin. Un *cycle* est un chemin de longueur non nulle tel que son nœud de départ coïncide avec son nœud d'arrivée<sup>2</sup> ; un *triangle* est un cycle de longueur trois. La figure 1.3 illustre ces définitions.

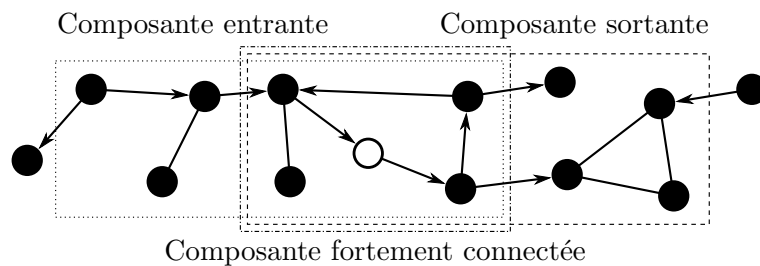
2. Il est à noter que, lors du calcul de la longueur d'un chemin, le nœud de départ est également compté comme nœud d'arrivée.



**Figure 1.3** – Terminologie relative aux chemins. Le chemin A, de longueur 3, quitte le nœud de départ 1 et atteint le nœud d'arrivée 2; puisqu'il s'agit du plus court chemin entre ces nœuds la distance est de 3. Il n'existe pas de chemin ayant comme point de départ le nœud 2 et comme point d'arrivée le nœud 1 : la distance est infinie. La distance entre 1 et 3 est de 5; la distance entre 3 et 1 est de 2. Le chemin B forme un cycle de longueur 7 dans le sens horaire seulement. Le chemin C forme un cycle peu importe la direction du parcours; puisque sa longueur est de 3, il s'agit également d'un triangle.

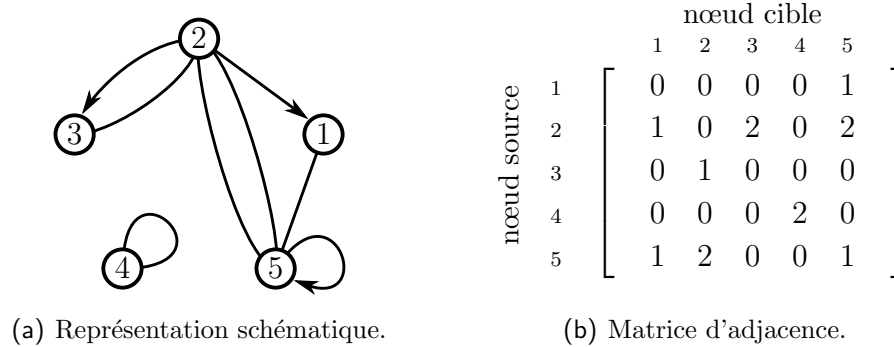


(a) Réseau non dirigé.



(b) Réseau semi-dirigé. Le nœud de référence est indiqué en blanc.

**Figure 1.4** – Terminologie relative aux composantes.



**Figure 1.5** – Représentation schématique et par matrice d'adjacence d'un même réseau.

La *composante sortante* (resp. *composante entrante*) d'un *nœud de référence* est l'ensemble des nœuds pouvant être le point d'arrivée (resp. de départ) d'un chemin  $y$  débutant (resp. s'y terminant); la *composante fortement connectée* d'un nœud de référence est l'intersection de sa composante entrante et sortante. Puisque chacun des nœuds de la composante fortement connectée d'un nœud de référence donné possède la même composante fortement connectée, on peut omettre de mentionner ce nœud de référence. Dans un réseau non dirigé, les composantes entrante, sortante et fortement connectée d'un nœud de référence correspondent toutes au même ensemble de nœuds, que l'on nomme tout simplement *composante*. La *taille* d'une composante correspond au nombre de nœuds qui la compose. Un *arbre* est une composante sans cycle. La figure 1.4 illustre ces définitions.

Due à la nature « bilingue » de cette thèse, un lexique (*cf.* page xxv) fournit les équivalents anglophones de ces termes.

## 1.2.2 Représentations

Outre la représentation schématique employée jusqu'ici, il existe plusieurs façons de représenter la structure d'un réseau. L'une d'entre elles, la *matrice d'adjacence*, spécifie le nombre total de demi-liens sortants et non dirigés entre chacune des combinaisons de nœud source et de nœud cible possible. La somme des éléments d'une même ligne (resp. colonne) de la matrice d'adjacence donne la somme du degré non dirigé et du degré sortant (resp. entrant) du nœud correspondant. Chaque boucle dirigée (resp. non dirigée) contribue 1 (resp. 2) à l'entrée correspondante sur la diagonale; tous les éléments diagonaux sont nuls pour un réseau sans boucles. Les éléments sont contraints aux valeurs  $\{0, 1\}$  en absence de liens répétés et de boucles non dirigées. La matrice d'adjacence d'un réseau non dirigé est symétrique. La figure 1.5 donne un exemple de

matrice d'adjacence.

Malgré le fait qu'une représentation par matrice d'adjacence puisse être utile dans certaines situations analytiques (*cf.* section 8.2), les algorithmes numériques utilisés dans cette thèse ont plutôt recours à une *liste d'adjacence* : un vecteur de vecteurs énumérant pour chaque nœud source l'index des entrées non nulles (répétées  $n$  fois pour une entrée  $n$ ) sur la ligne correspondante dans la matrice d'adjacence. La liste d'adjacence correspondant au réseau schématisé en figure 1.5(a) est

$$\left\{ \overset{1}{\underbrace{\{5\}}}, \overset{2}{\underbrace{\{1, 3, 3, 5, 5\}}}, \overset{3}{\underbrace{\{2\}}}, \overset{4}{\underbrace{\{4, 4\}}}, \overset{5}{\underbrace{\{1, 2, 2, 5\}}} \right\} . \quad (1.1)$$

Numériquement, cette représentation s'avère consommer beaucoup moins de mémoire pour des réseaux faiblement connectés (c'est-à-dire des réseaux pour lesquels la matrice d'adjacence aurait principalement été constituée d'entrées nulles).

Il est à noter que, telles qu'introduites, ni une matrice d'adjacence ni une liste d'adjacence ne permettent de distinguer un lien non dirigé de deux liens dirigés répétés en sens opposés<sup>3</sup>. Une solution simple consiste à utiliser une matrice d'adjacence différente pour les liens non dirigés et pour les liens dirigés. Quoique l'on puisse procéder de la même façon avec deux listes d'adjacence correspondantes, il est souvent plus avantageux (en terme de temps d'exécution) d'utiliser trois listes d'adjacence : une pour les liens non dirigés, une pour les liens sortants et une pour les liens entrants.

### 1.2.3 Ensembles de réseaux

Dans un contexte de modélisation d'un système réel complexe, pour des raisons de nature technique, monétaire ou même éthique, il est fréquent de n'avoir qu'une connaissance partielle de la structure du réseau. Un problème semblable est prédominant en physique statistique : le nombre de mesures disponibles est typiquement très petit face au nombre de degrés de liberté du système. Dans ce dernier contexte, un principe de base fort simple s'est avéré à la fois puissant et prolifique : considérer l'ensemble des états possibles du système permettant d'expliquer les observables connues. Il s'avère que l'on peut procéder de la même façon pour le problème énoncé précédemment.

Un *ensemble de réseaux* est une collection de réseaux, chacun pondéré par une certaine probabilité. On spécifie un ensemble de réseaux à travers les contraintes qui

---

3. Cette situation est souvent acceptable lorsque l'on suppose une brisure complète de symétrie en ce qui a trait à la direction des liens dirigés (voir note 1 en page 3).

lui sont imposées : une *contrainte forte* doit être respectée par tous les réseaux de l'ensemble alors qu'une *contrainte faible* doit seulement être statistiquement valide sur l'ensemble pondéré. Par analogie avec la physique statistique (particules  $\Leftrightarrow$  nœuds et énergie  $\Leftrightarrow$  liens), on parle d'un *ensemble microcanonique* lorsque les nombres de nœuds et de liens sont des contraintes fortes, d'*ensemble canonique* lorsque les nombres de nœuds et de liens sont respectivement des contraintes forte et faible, et d'*ensemble grand-canonique* lorsque les deux quantités sont des contraintes faibles. Il est à noter que le cas grand-canonique est beaucoup moins fréquent que les deux autres.

À titre d'exemple, le *modèle d'Erdős et Rényi* [29] définit un ensemble de réseaux (non dirigés) à partir de deux paramètres seulement. Le premier, toujours une contrainte forte, est le nombre  $N$  de nœuds dans les réseaux de l'ensemble. Deux variantes existent pour le second paramètre : on spécifie soit le nombre  $M$  de liens dans chacun des réseaux de l'ensemble (cas microcanonique), soit la probabilité  $p$  que deux nœuds aléatoirement sélectionnés dans un même réseau  $y$  soient joints par un lien (cas canonique). Dans ce dernier cas, la contrainte faible du nombre moyen de liens correspond à  $\langle M \rangle = \frac{1}{2}pN(N-1)$ .

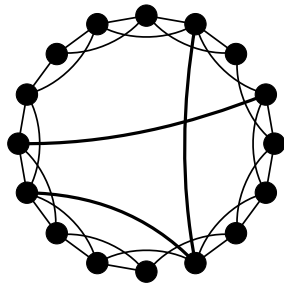
Un autre exemple d'importance, le *modèle de configuration* [64, 75], est obtenu en contraignant le degré des nœuds. Dans sa variante microcanonique, la *séquence de degrés*<sup>4</sup>  $\{n_0, n_1, n_2, \dots\}$  spécifie le nombre  $n_k$  de nœuds de degré  $k$  devant se trouver dans chacun des réseaux de l'ensemble ; les nombres de nœuds  $N = \sum_k n_k$  et de liens  $M = \frac{1}{2} \sum_k kn_k$  sont tous les deux fortement contraints. Dans le cas canonique, on donne spécifiquement le nombre de nœuds  $N$  et la *distribution*<sup>5</sup> *en degré*  $\{p_0, p_1, p_2, \dots\}$  donne la probabilité  $p_k$  qu'un nœud choisi au hasard dans un réseau de l'ensemble soit de degré  $k$  ; le nombre moyen de nœuds de degré  $k$  est  $\langle n_k \rangle = Np_k$  et le nombre moyen de liens est  $\langle M \rangle = \frac{1}{2}N \sum_k kp_k$ .

Le *modèle de Watts et Strogatz* [101] est un ensemble de réseaux obtenu en rendant légèrement aléatoire une grille préalablement régulière. Algorithmiquement, on génère d'abord la grille régulière en assignant un index unique  $i \in \{1, 2, \dots, N\}$  à chacun des  $N$  nœuds du réseau et en joignant d'un lien les nœuds  $i$  et  $i'$  si  $i' - i = 1 \pmod{N}$  ou si  $i' - i = 2 \pmod{N}$ . Par la suite, on retire  $m$  de ces liens sélectionnés au hasard (cas microcanonique) ou retire chacun des liens avec une probabilité  $p$  (cas canonique). Chacun des liens retirés est remplacé par un lien que l'on ajoute entre deux nœuds sélectionnés aléatoirement. La figure 1.6(a) donne un exemple pour  $N = 16$  et  $m = 3$ . Le processus se généralise aisément pour différentes grilles (figure 1.6(b)). Une variante

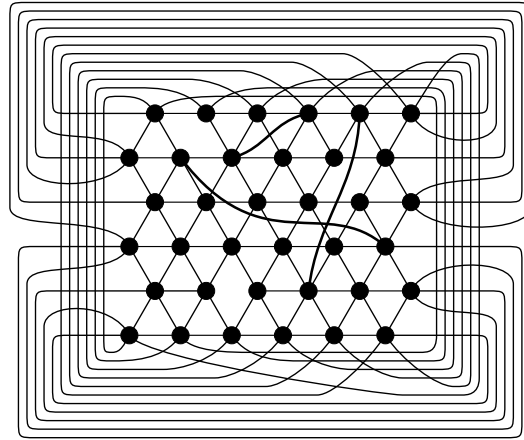
4. Ce terme représente souvent un objet légèrement différent.

5. Il est à noter qu'on emploie le terme « distribution » pour toute distribution de probabilité, qu'elle soit discrète (fonction de masse) ou continue (fonction de densité).





(a) Grille standard.



(b) Grille triangulaire bidimensionnelle.

**Figure 1.6** – Deux réseaux produits par le modèle de Watts et Strogatz. La grille décrite dans le texte pour  $N = 16$  nœuds (a) et une grille triangulaire bidimensionnelle avec conditions aux limites périodiques (b) servent de grille de départ. Dans les deux cas,  $m = 3$  liens sont retirés au hasard et le même nombre est ajouté aléatoirement.

assez courante exige que le degré de chaque nœud soit conservé : on brise alors les liens en paires de demi-liens (plutôt que de les retirer) puis, à la fin du processus, on apparie aléatoirement ces demi-liens pour former des liens (plutôt que d’ajouter de nouveaux liens).

Les ensembles précédents peuvent aisément être généralisés à des réseaux dirigés (ou semi-dirigés) et une multitude d’autres ensembles peuvent être définis de façon semblable. Il est à noter que des contraintes « qualitatives » peuvent également être utilisées : on peut par exemple interdire la présence de boucles et/ou de liens dirigés dans tous les réseaux d’un ensemble.

On parle parfois d’un « processus dynamique sur un ensemble de réseaux » : il s’agit là d’un abus de langage impliquant la réalisation du processus dynamique sur chacun des réseaux de l’ensemble puis la pondération des résultats par la probabilité de réaliser chacun des réseaux correspondants dans l’ensemble. Afin de spécifier que la structure du réseau demeure la même tout au long d’un même processus, on dit du réseau qu’il est *gelé* (« *quenched* »). Alternativement, tout comme en physique statistique, il est possible que l’ensemble spécifie plutôt un équilibre dynamique : la structure du réseau est en constant changement tout au long d’un même processus, des changements préservant les propriétés statistiques de l’ensemble. Dans de tels cas, il est normalement nécessaire de spécifier explicitement les règles régissant ces changements de la structure du réseau. Il existe cependant une exception d’importance, soit le cas sans mémoire : la structure du réseau change si rapidement par rapport au processus dynamique que l’on

peut complètement négliger la corrélations entre la structure du réseau sur des temps infinitésimalement rapprochés. On dit alors que le réseau est *mouvant* (« *annealed* »)<sup>6</sup>.

### 1.2.4 Réseaux complexes

On définit souvent les *réseaux complexes* comme des structures possédant certaines caractéristiques qui seraient jugées atypiques dans un réseau aléatoire ou régulier, mais qui sont fréquemment rencontrées dans cette nouvelle classe dite « complexe ». Malgré que cette définition soit relativement vague, on peut en tirer deux points principaux : 1. il ne s'agit pas d'une collection de cas isolés mais bien d'une classe de problèmes à part entière, et 2. les approches standard (présumant régularité ou désordre) échouent face à ces problèmes.

Afin d'obtenir une définition plus claire, on s'intéresse à la *complexité de Kolmogorov* [50, 93] d'une séquence, c'est-à-dire la longueur du plus court programme (informatique ou autre) retournant cette séquence en sortie. À fins d'exemples, soient les séquences

$$\begin{aligned} A &= 01 \dots \\ B &= 001001000011111101110101010001000100001011010001100001000 \dots \end{aligned} \quad (1.2)$$

Le court programme « Répéter "01" indéfiniment. » permet de reproduire la séquence  $A$  ; il ne s'agit pas nécessairement du plus court programme possible, mais son existence établit une borne supérieure à la complexité de Kolmogorov de la séquence  $A$ . Par opposition, on voit difficilement comment reproduire la séquence  $B$  autrement que par un programme du genre « Retourner 00100100001111110... ». Puisque ce programme est aussi long que la séquence elle-même, ceci semble indiquer une haute complexité de Kolmogorov pour la séquence  $B$ .

Une façon simple de définir la complexité de Kolmogorov d'un réseau est donc de demander que la séquence de sortie du programme corresponde à une représentation (*c.f.* section 1.2.2) donnée de ce réseau. Ainsi, une grille bidimensionnelle régulière peut aisément être générée par un court programme, en accord avec notre intuition qu'il ne s'agit pas là d'un réseau bien complexe. Par opposition, il semble beaucoup moins probable qu'un court programme puisse reproduire la structure d'interactions d'une société humaine.

L'étude du cas d'un réseau aléatoire produit par le modèle d'Erdős et Rényi semble

---

6. Je n'ai pas trouvé d'équivalents francophones acceptés pour « *quenched* » et « *annealed* » ; je trouve les termes « gelé » et « mouvant » plus évocateurs que les traductions littérales métallurgiques « trempé » et « recuit ».

révéler une faille importante de cette définition. Clairement, ce réseau ne devrait pas être considéré comme complexe. Cependant, pour un réseau typique produit de cette façon, aucun court programme ne permet de générer une représentation de sa structure. En fait, il s'agit du cas limite supérieur : la complexité de Kolmogorov d'une séquence aléatoire est maximale.

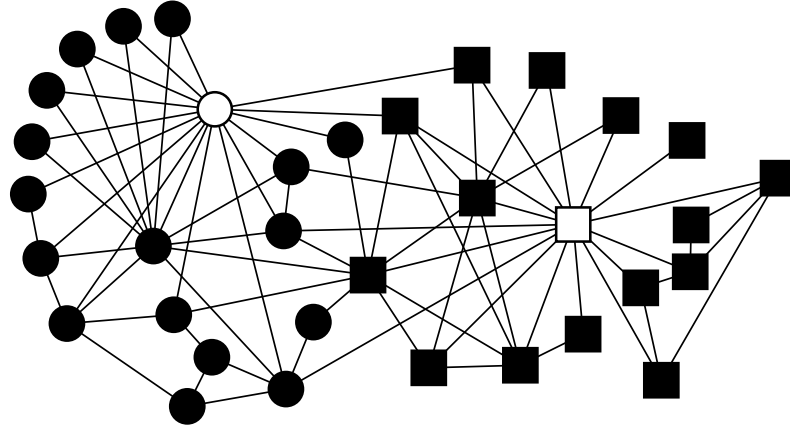
À l'instar de la première définition donnée dans cette section, on pourrait déclarer un réseau comme « complexe » si sa complexité de Kolmogorov est intermédiaire entre celle d'un réseau régulier et celle d'un réseau aléatoire. Cependant, la présente thèse préfère une solution alternative : appliquer la définition de la complexité de Kolmogorov à des ensembles de réseaux. Ainsi, on demande que les réseaux produits par un nombre arbitrairement élevé d'exécutions du programme minimal possèdent les mêmes propriétés statistiques que l'ensemble de réseaux dont on souhaite connaître la complexité<sup>7</sup>.

Ce changement de définition produit l'effet escompté dans le cas de l'ensemble de réseaux générés par le modèle d'Erdős et Rényi : un court programme peut facilement reproduire les propriétés statistiques de cet ensemble, indiquant une faible complexité de Kolmogorov. Pour appliquer la définition à une société humaine, on peut par exemple considérer l'ensemble de réseaux comme la structure d'interactions pour plusieurs villes différentes, ou encore s'intéresser à la structure d'une même population en des temps différents. Dans les deux cas, il est difficile d'imaginer un court programme reproduisant les propriétés statistiques de ces ensembles : des structures intuitivement complexes demeurent complexes.

Il s'avère instructif de s'intéresser à une autre faiblesse apparente de cette définition : sa subjectivité. En effet, il est souvent impossible de savoir si le plus court programme que l'on ait pu imaginer est vraiment le plus court possible (à moins, bien sûr, d'exécuter un à un tous les programmes alternatifs). Ainsi, deux personnes aux passés différents estimant la complexité de Kolmogorov d'un même ensemble de réseaux peuvent arriver à des résultats complètement discordants si l'un d'entre eux a songé à un court programme ayant échappé à l'autre. Cependant, il en est de même pour notre définition intuitive de la complexité : ce qui est complexe pour une personne peut ne pas l'être pour une autre. Néanmoins, on s'attend à ce que deux personnes raisonnables puissent, après discussion, s'entendre sur la complexité d'un système. Par exemple, informer quelqu'un que la séquence  $B$  de l'équation (1.2) est l'expression binaire des décimales de  $\pi$  risque de drastiquement réduire son opinion sur sa complexité, autant au sens intuitif qu'au

---

7. On doit également modifier le problème pour que deux exécutions répétées du même programme puissent produire une sortie différente. Ceci peut par exemple être fait en équipant le langage de programmation d'une commande générant des nombres aléatoires ou encore en passant au programme un paramètre lui indiquant le nombre de fois qu'il a été exécuté par le passé.



**Figure 1.7** – Réseau d'amitié dans un club de karaté [107] (connu sous le nom de « Zachary's karate club »). Les liens du réseau indiquent les personnes se rencontrant également à l'extérieur des activités du club (critère d'amitié). Suite à un conflit sur la valeur des cotisations, le club s'est scindé en deux clubs (cercles et carrés) dirigés par le professeur (cercle blanc) et par le président (carré blanc) du club initial. Ce réseau est utilisé à titre d'exemple dans la section 1.2.5.

sens de Kolmogorov. La subjectivité de cette définition n'est donc pas une simple faille, mais bien une difficulté inhérente au problème.

Voici donc la définition de « réseau complexe » qui est utilisée dans le reste de cette thèse : un réseau qui, au meilleur de notre connaissance actuelle, fait partie d'un ensemble dont la complexité de Kolmogorov est élevée. En ce sens, aucun modèle reproduisant certaines caractéristiques de réseaux complexes à l'aide d'un petit nombre de paramètres ne peut générer de réseaux complexes puisque le modèle lui-même peut être perçu comme un programme relativement court. Néanmoins, rien n'empêche de tels modèles d'adéquatement reproduire certains aspects de systèmes réels complexes à des fins prédictives.

### 1.2.5 Caractéristiques des réseaux complexes

Tel que mentionné précédemment, de nombreuses observations empiriques révèlent que certaines caractéristiques sont surreprésentées dans les réseaux complexes dépeignant la structure de plusieurs systèmes réels (par rapport à ce que l'on s'attendrait dans un réseau régulier ou aléatoire). Cette section relève quelques unes de ces caractéristiques.

Pour la même raison que le nombre de coordination<sup>8</sup> et la densité jouent des rôles

8. Nombre de sites voisins les plus proches.

souvent essentiels dans une grille régulière et dans un gaz aléatoire, respectivement, il est typiquement important de considérer le nombre de voisins que possèdent les nœuds d'un réseau à travers sa distribution en degrés. Dans un réseau complexe réel, la queue ( $k$  grand) de cette distribution mérite une attention particulière : la décroissance de la probabilité en fonction de  $k$  est typiquement plus lente que celle d'un réseau aléatoire. En fait, dans plusieurs cas [8], la distribution suit approximativement une loi de puissance sur un certain intervalle de  $k$ . Par conséquent, le comportement global d'un réseau peut être fortement influencé par la faible fraction de nœuds dont le degré est largement supérieur à la moyenne (« *hubs* »). La présence de tels nœuds de haut degré est visible dans l'exemple présenté en figure 1.7. La version canonique du modèle de configuration correspond à une situation où seule la distribution en degrés et le nombre de nœuds sont connus.

Sur une note semblable, il est fréquent d'observer dans des réseaux réels une *corrélation entre les degrés* des nœuds voisins. Le *coefficient de corrélation de Pearson*

$$r = \frac{\text{cov}(k, k')}{\sqrt{\text{var}(k) \text{var}(k')}} \quad (\text{donc } -1 \leq r \leq 1) \quad (1.3)$$

entre les degrés  $k$  et  $k'$  de deux nœuds voisins permet de quantifier cette caractéristique : on observe typiquement  $r > 0$  dans les réseaux sociaux et  $r < 0$  dans les réseaux technologiques et biologiques [69]. Ainsi, les connaissances d'une personne hautement connectée auront en moyenne un plus grand nombre de connaissances que le degré moyen du réseau. Par opposition, l'un des serveurs rattaché à un nœud de haut degré de l'Internet risque de n'être qu'un petit distributeur. Il est à noter qu'il ne s'agit là que d'une tendance générale ; malgré le fait que le réseau présenté en figure 1.7 soit de nature sociale, on y mesure  $r = -0.476$  [39].

On appelle *communauté* un groupe de nœuds partageant entre eux significativement plus de liens que ce que le hasard pourrait expliquer. Les communautés peuvent parfois être des effets résiduels de la nature géométrique du système : des éléments occupant un même voisinage spatial peuvent avoir plus de chances d'interagir entre eux qu'avec d'autres éléments plus éloignés. Cependant, si des gens habitant le même quartier ont plus de chances de se connaître, il en est de même pour des collègues de travail ou pour des gens partageant un passe-temps commun. Peu importe la cause, on dit d'un réseau contenant un nombre significatif de communautés qu'il a une *structure communautaire*. L'issue du conflit discuté en figure 1.7, c'est-à-dire quelle faction chacun des membres choisira, peut être prédite en étudiant la structure communautaire de ce réseau [34].

L'adage « l'ami de mon ami est mon ami » reflète le concept d'*agrégation* (ou de *transitivité*) : deux nœuds voisins d'un même nœud ont souvent plus de chances d'être également voisins l'un de l'autre que ce que le hasard seul pourrait expliquer. Le *coef-*

*ficient d'agrégation*

$$C = \frac{[\text{nombre de cycles de longueur 3 (triangles) dans le réseau}]}{[\text{nombre de chemins de longueur 2 dans le réseau}]} \quad (1.4)$$

permet de quantifier cette caractéristique<sup>9</sup>. La valeur de  $C$  est limitée à l'intervalle  $[0, 1]$ ;  $C = 0$  indique l'absence de triangles dans le réseau tandis que  $C = 1$  indique que chacune des paires de voisins de chacun des nœuds du réseau sont eux même voisins. On mesure  $C = 0.256$  [39] pour le réseau présenté en figure 1.7. Si la structure communautaire peut causer de l'agrégation, l'agrégation peut également avoir une origine indépendante. Par exemple, un triangle est formé lorsque deux amis sont présentés par un ami commun, contribuant ainsi à augmenter la valeur du coefficient d'agrégation.

On dit d'un réseau qu'il a une *structure hiérarchique* si une certaine « gradation » des nœuds (souvent en terme de leur nature, de leur degré ou de leur « position » dans le réseau) permet d'expliquer en partie la structure du réseau. Ce que l'on entend exactement par ce concept relativement flou doit normalement être spécifié dans les cas d'application. Il est à noter que la structure communautaire peut elle aussi être hiérarchique, par exemple un groupe d'amis dans une classe dans une école dans une ville...

La *distance moyenne* entre chacune des paires de nœuds d'un réseau complexe<sup>10</sup> est souvent de l'ordre du logarithme du nombre de nœuds dans le réseau (ou inférieure), soit beaucoup plus faible que celle d'une grille régulière. Cependant, la présence de cette caractéristique seule n'est pas vraiment une signature de la complexité puisque les réseaux aléatoires répondent habituellement à ce critère. Ainsi, on parle d'*effet small-world* lorsque la faible distance moyenne coexiste avec une certaine forme de localité, soit par exemple un coefficient d'agrégation élevé ou la présence de structure communautaire. Le modèle de Watts et Strogatz (*c.f.* section 1.2.3) a été conçu pour illustrer un simple mécanisme permettant l'effet small-world : l'ajout d'un petit nombre de liens aléatoires diminue grandement la distance moyenne entre les nœuds du réseau [101].

## 1.2.6 Note sur les graphes

Les termes « *graphe* », « *sommet* » et « *arête* » employés dans la littérature mathématique [12, 21, 102] désignent respectivement les mêmes objets que les termes

9. Il est à noter que le terme « coefficient d'agrégation » est parfois associé à des définitions différentes.

10. On prend ici pour acquis que le réseau ne comporte qu'une seule composante, ou encore que l'on ne s'intéresse qu'à une seule composante du réseau.

« réseau », « nœud » et « lien » préférés dans cette thèse. Comme pour la plupart des théories mathématiques modernes, la *théorie des graphes* est formellement définie en termes d'ensembles. Ainsi, un graphe  $G = (V, E)$  est une paire ordonnée d'un ensemble de sommets  $V$  et d'un ensemble d'arrêtes  $E$ . Dans le cas d'un graphe non dirigé (resp. dirigé), les éléments de  $E$  sont des ensembles  $\{v_1, v_2\} \in E$  (resp. des paires ordonnées  $(v_1, v_2) \in E$ ) de sommets  $v_1, v_2 \in V$ . On suppose normalement qu'un sommet ne peut être une arrête du même graphe, et donc  $V \cap E = \emptyset$ .

La théorie des graphes englobe la théorie des réseaux : grossièrement, un réseau est un graphe où la présence d'une arrête  $\{v_1, v_2\}$  (ou  $(v_1, v_2)$ ) dans  $E$  implique une « interaction » entre  $v_1$  et  $v_2$ . La distinction entre ces domaines d'étude est cependant plus grande que cette définition ne le laisse paraître : la théorie des réseaux est normalement plus proche de l'application, les structures étudiées sont souvent plus « complexes », l'approche est typiquement moins formelle et les questions posées sont généralement différentes.

Ainsi, la présente thèse préconise clairement une approche relevant de la théorie des réseaux et de sa littérature. Outre les différences terminologiques, on se permet d'être beaucoup moins rigoureux que ce n'est typiquement le cas en mathématiques. Comme il est souvent le cas en science, le verdict est prononcé lors de la confrontation à l'expérience ou, dans notre cas, à « l'expérience numérique ».

## 1.3 Dynamique du réseau et sur réseau

La *dynamique* d'un système correspond à son évolution dans le temps. Lorsque le système d'intérêt est représenté par un réseau, on peut discerner la *dynamique du réseau*, soit le changement de la structure même du réseau, de la *dynamique sur le réseau*, soit l'évolution de l'*état intrinsèque* des constituants (nœuds et liens) du réseau. Dans le cadre de cette thèse, on s'intéresse particulièrement au second type de dynamique (structure fixe) tout en gardant en tête l'application possible des méthodes développées à des structures elles-mêmes changeantes.

La section 1.3.1 raffine notre définition de l'état d'un réseau, la section 1.3.2 introduit les processus de propagation à travers des exemples de modélisation épidémiologique, la section 1.3.3 spécifie ce que l'on entend par « stochasticité » et la section 1.3.4 précise le niveau de modélisation visé dans cette thèse ainsi que son rapport avec la réalité physique.

### 1.3.1 États intrinsèques

On modélise l'*état complet* d'un système donné à l'aide du réseau spécifiant la structure d'interactions de ses éléments constitutifs, auquel on juxtapose des états intrinsèques aux nœuds et/ou aux liens le composant. *A priori*, ces états intrinsèques peuvent correspondre à n'importe quel ensemble de caractéristiques — fixes ou dynamiques — caractérisant ces éléments : le genre, l'âge et/ou l'opinion d'une personne ; la capacité maximale et la charge actuelle d'une station de transformation d'énergie ou d'une ligne de transmission ; *etc.*

Dans le cadre de cette thèse, on ne considère que des cas présentant un nombre limité d'états intrinsèques discrets. Cette restriction est beaucoup moins contraignante qu'on pourrait le croire : s'il est impossible de spécifier l'âge exacte d'une personne (puisqu'il s'agit d'un nombre réel), on peut tout de même préétablir un nombre limité d'intervalles d'âges et spécifier à quel intervalle appartient l'âge d'une personne. De plus, on peut spécifier dans un même état plusieurs caractéristiques intrinsèques différentes à l'aide d'un produit cartésien. Par exemple, 12 états intrinsèques permettent de spécifier le genre (*M*âle et *F*emelle,  $\{M, F\}$ ), la tranche d'âge ((*a*) de 18 à 25 ans, (*b*) de 25 à 60 ans, ou (*c*) 60 ans et plus ;  $\{a, b, c\}$ ) et l'opinion face à une proposition (en *A*ccord ou en *D*ésaccord,  $\{A, D\}$ ) d'un nœud du réseau

$$\begin{aligned} \{M, F\} \times \{a, b, c\} \times \{A, D\} = \\ \{MaA, MaD, MbA, MbD, McA, McD, FaA, FaD, FbA, FbD, FcA, FcD\} \quad . \quad (1.5) \end{aligned}$$

Dans le contexte de cet exemple, on pourrait n'étudier strictement que la dynamique *du* réseau en s'intéressant à l'évolution de la structure du réseau et en supposant que l'état intrinsèque de chaque nœud ne change pas au cours du temps. De même, on pourrait n'étudier que la dynamique *sur* le réseau en supposant une structure fixe et en s'intéressant à l'évolution au cours du temps de l'état intrinsèque des nœuds. Dans ce dernier cas, seule l'opinion des gens risque de dépendre de leurs interactions, donc de la structure du réseau (l'âge de chaque personne augmente au même taux indépendamment de la structure et on peut probablement négliger les changements de genres). Enfin, on parle de *coévolution des processus* lorsqu'à la fois la structure du réseau et l'état intrinsèque des éléments peuvent changer.



### 1.3.2 Propagation et modèles épidémiologiques

On définit un *processus de propagation* sur réseau comme la transmission de « quelque chose » entre les nœuds voisins du réseau : maladie infectieuse, rumeur, virus informatique, publicité virale, *etc.* Dans tous les cas, on emprunte le vocabulaire de la modélisation épidémiologique afin de décrire ces processus de propagation. Ainsi, qu'il s'agisse d'une caractéristique désirable ou non, on appelle *infection* le « quelque chose » pouvant se propager. On dit d'un nœud portant l'infection et pouvant la transmettre qu'il est *infectieux* ; un nœud ne portant pas l'infection mais pouvant l'acquérir est dit *susceptible* ; et un nœud ne portant pas l'infection ni ne pouvant l'acquérir est dit *retiré*. Ces caractéristiques intrinsèques aux nœuds sont appelées *états épidémiologiques*.

Dépendant des états épidémiologiques accessibles et des transitions possibles entre ces états, on peut classifier les processus de propagation en diverses catégories. Ainsi, un processus *SI* ne permet que deux états (Susceptible et Infectieux,  $\{S, I\}$ ) et qu'une seule transition : un nœud susceptible possédant un ou plusieurs voisins infectieux peut lui-même devenir infectieux. Il s'agit probablement là du plus simple processus de propagation possible ; ajouter des transitions et/ou des états accessibles permet de définir plusieurs autres processus d'intérêt. Par exemple, on obtient un processus *SIS* en permettant aux nœuds infectieux de redevenir susceptibles et un processus *SIR* en rendant plutôt Retirés (donc trois états possibles,  $\{S, I, R\}$ ) les nœuds cessant d'être infectieux. De même, un processus *SIRS* est obtenu à partir d'un processus *SIR* en permettant aux nœuds retirés de redevenir susceptibles.

Cette nomenclature tire son origine des *modèles épidémiologiques compartimentaux* [5, 20, 38, 47]. Dans ce paradigme, les éléments constitutifs d'un même état (compartiment) sont indiscernables entre eux et les interactions entre les éléments se font aléatoirement : la population est « bien mélangée ». Une pratique courante consiste alors à suivre l'évolution dans le temps du nombre moyen d'éléments de chaque état à l'aide d'un système d'équations différentielles. Par exemple, pour un système *SIR* où  $S(t)$ ,  $I(t)$  et  $R(t)$  donnent respectivement le nombre d'éléments susceptibles, infectieux et retirés au temps  $t$ , on a

$$\frac{dS(t)}{dt} = -\lambda S(t) \frac{I(t)}{S(t) + I(t) + R(t)} \quad (1.6a)$$

$$\frac{dI(t)}{dt} = \lambda S(t) \frac{I(t)}{S(t) + I(t) + R(t)} - \alpha I(t) \quad (1.6b)$$

$$\frac{dR(t)}{dt} = \alpha I(t) \quad . \quad (1.6c)$$

Le *taux de retrait*  $\alpha$  fixe la probabilité  $\alpha dt$  qu'un élément infectieux donné devienne retiré pendant un intervalle de temps  $[t, t + dt)$ . Par conséquent, chaque élément devenant infectieux a une probabilité  $\alpha e^{-\alpha\tau} d\tau$  de le rester pendant une période de temps comprise dans l'intervalle  $[\tau, \tau + d\tau)$ . Lorsqu'une distribution plus concentrée autour de sa moyenne est désirée, une solution fréquente consiste à discerner plusieurs stades infectieux : on parle alors d'un modèle  $SI^nR$ , permettant les états  $\{S, I_1, I_2, \dots, I_n, R\}$ , et la probabilité de passage d'un état infectieux au suivant (ou à l'état retiré pour le  $n$ -ième état infectieux) est de  $n\alpha dt$  pour l'intervalle de temps  $[t, t + dt)$ . Les cas particuliers  $SI^2R$  et  $SI^3R$  sont parfois respectivement notés  $SIIR$  et  $SIIR$ .

Le *taux d'infection*  $\lambda$  détermine à la fois la « virulence » de l'infection et la « densité » de la population étudiée. Afin d'isoler ces effets, on décompose  $\lambda$  en deux paramètres tels que  $\lambda = \beta \langle k \rangle$ . Ainsi,  $\langle k \rangle$  donne le nombre moyen d'interactions par éléments alors que  $\beta$ , également appelé *taux d'infection*, donne le nombre moyen  $\beta dt$  de nouvelles infections par interaction entre un élément susceptible et un élément infectieux pendant un intervalle de temps  $[t, t + dt)$ . Le terme  $\frac{I(t)}{S(t)+I(t)+R(t)}$  détermine la fraction des  $\langle k \rangle S(t)$  interactions des éléments susceptibles qui sont faites avec des éléments infectieux.

En termes de réseaux, la structure du système est régie par un modèle d'Erdős et Rényi canonique mouvant composé de  $N = S(t) + I(t) + R(t)$  nœuds ( $N$  constant) où chaque lien a une probabilité  $p = \langle k \rangle / (N - 1)$  d'exister. Dans ce contexte, un nœud susceptible possédant  $k_I$  liens vers des nœuds infectieux a une probabilité  $k_I \beta dt$  de devenir infectieux pendant un intervalle de temps  $[t, t + dt)$ . Aucune adaptation particulière n'est nécessaire pour le taux de retrait : un nœud infectieux a une probabilité  $\alpha dt$  de devenir retiré pendant un intervalle de temps  $[t, t + dt)$ .

Ainsi, les modèles épidémiologiques compartimentaux peuvent être perçus comme un cas particulier de propagation sur des réseaux très simples, c'est-à-dire peu complexe au sens décrit en section 1.2.4. Une généralisation directe permet de définir des problèmes plus complexes, comme par exemple la propagation de type  $SIR$  sur un modèle de configuration (canonique ou micro-canonique) gelé. Dans ce cas, la dynamique régie par les paramètres  $\alpha$  et  $\beta$  demeurent la même ; seule la structure des réseaux change. Quoiqu'il ne s'agisse pas là du seul problème étudié dans cette thèse, cet exemple spécifique y joue un rôle particulièrement important sous sa forme stochastique. En particulier, on souhaite connaître la probabilité qu'une *épidémie* affecte une fraction importante de la population, ou au contraire si l'infection ne se limite qu'à un nombre restreint d'individus (situation appelée *flambée épidémique*, « *outbreak* »).

### 1.3.3 Stochasticité

Un processus dynamique est dit *stochastique* si les règles gouvernant son évolution sont de nature probabiliste. En modélisation, on réserve normalement le terme « stochastique » aux approches permettant, dans une certaine limite, de prévoir la variabilité des issues possibles pour le système modélisé. Quoique ces deux concepts soient fortement reliés, il est parfois tout à fait justifiable d'utiliser un modèle déterministe afin de prévoir le comportement d'un système stochastique, ou encore de modéliser un système déterministe de façon stochastique.

La nécessité d'utiliser un modèle stochastique pour représenter un système stochastique dépend principalement de l'ampleur de la variabilité des issues possibles. En effet, obtenir l'évolution dans le temps d'une valeur moyenne peut souvent suffire à nos besoins si l'état du système fluctue très peu autour de cette valeur. Si ces fluctuations sont typiquement plus importantes dans les petits systèmes, elles ne sont toutefois pas toujours négligeables dans les grands. Par exemple, dans le cas d'un processus *SIR* sur un grand réseau dont un seul des nœuds est initialement infectieux, il est possible que ce nœud devienne retiré avant même de pouvoir transmettre l'infection, mettant immédiatement un terme à ce qui aurait pu être une épidémie dévastatrice.

Dans le cas de problèmes intrinsèquement déterministes, l'utilisation de modèles stochastiques peut être nécessaire lorsque l'on ne connaît pas tout du système à modéliser. Tel que présenté en section 1.2.3, contraindre un ensemble statistique de réseaux à respecter les quelques caractéristiques qui nous en sont connues permet de tenir compte de ces informations partielles sans avoir à se commettre sur les points nous échappant. Or, l'évolution temporelle d'un processus dynamique peut parfois fortement varier d'un réseau à l'autre d'un tel ensemble. Ainsi, notre incertitude de la structure d'interactions d'un processus intrinsèquement déterministe peut, de notre point de vue, le rendre stochastique. En fait, même à l'échelle d'une seule interaction *a priori* déterministe, notre ignorance du détail de cette interaction peut lui donner une apparence stochastique.

### 1.3.4 Modèles et réalité

La section précédente a établi une distinction entre la stochasticité d'un problème et celle du modèle utilisé pour prévoir son comportement. De façon générale, on distingue dans cette thèse au moins trois niveaux de complexité dans notre étude d'un système : la *réalité*, une *idéatisation* de cette réalité et une *simplification* de cette idéatisation.

En notant  $A > B$  le fait que  $A$  est plus complexe que  $B$ , on obtient l'échelle

$$\text{Réalité} > \begin{array}{c} \text{Idéalisation} \\ \text{(modèle numérique)} \end{array} > \begin{array}{c} \text{Simplification} \\ \text{(modèle analytique)} \end{array} .$$

On entend par « réalité » le système physique dont le comportement nous intéresse, par exemple la propagation d'une maladie infectieuse donnée dans une population humaine. Dans la plupart des cas, notre compréhension de la réalité n'en sera toujours qu'un modèle : on appelle « idéalisation » toute description incomplète de la réalité. Il est normalement possible de traduire cette idéalisation en une simulation informatique (modèle numérique). Cependant, cette approche peut demander une quantité considérable de ressources informatiques et, de plus, la complexité de tels modèles peut embrouiller les liens de causalité, limitant notre compréhension du processus. C'est pourquoi on recherche des approximations permettant d'obtenir un modèle suffisamment simplifié pour être approché analytiquement.

Si chaque étape reproduit suffisamment bien le comportement de la précédente, il peut alors être justifié de tirer des conclusions à propos du système complexe réel à partir du modèle analytique simplifié. Dans cette thèse, on s'intéresse avant tout à la seconde étape, c'est-à-dire l'obtention de modèles simplifiés fiables pour des systèmes idéalisés donnés. Ainsi, nos résultats analytiques sont comparés à des « expériences numériques » simulant ces idéalizations, plutôt qu'à une mesure directe de la réalité. Le choix des systèmes idéalisés servant dans ces comparaisons est cependant fortement inspiré de la réalité, puisque notre but ultime est de concevoir des outils permettant d'étudier des systèmes réels.

# Première partie

## Processus de branchement et évolution dans le temps



# Chapitre 2

## Préparation aux processus de branchements

Ce chapitre introduit la partie I de cette thèse, partie s'intéressant essentiellement aux processus de propagation *SIR* ayant lieu sur les réseaux d'un modèle de configuration canonique gelé non dirigé. On établit d'abord un parallèle entre la propagation et la percolation, puis étudie cette dernière comme un processus de branchement. Afin de limiter la redondance, certaines discussions sont reportées à des sections futures. En particulier, la section 2.4 met en contexte le chapitre 3 sans détailler la démonstration des résultats énoncés.

La section 2.1 introduit les concepts de composantes géantes et de percolation ; ces notions sont reliées aux processus de propagation à travers le concept de transmissibilité, présenté en section 2.2. La section 2.3 expose quelques bases concernant les fonctions génératrices alors que la section 2.4 explique rapidement comment cet outil permet de considérer la percolation sur réseaux. Cette dernière section présente également en quoi l'approche présentée au chapitre 3 diffère des précédentes.

### 2.1 Percolation

Soit un ensemble de réseaux dont tous les paramètres ont déjà été fixés, à l'exception de la taille  $N$  des réseaux produits. Dans la limite  $N \rightarrow \infty$ , la plus grande composante (*c.f.* page 6) des réseaux de cet ensemble est appelée *composante géante* si sa taille  $s$  est telle que le ratio  $s/N$  tend vers une valeur  $S \in (0, 1]$ , alors appelée *taille de la composante*

*géante*. Dans les cas qui nous intéressent, il s'agit alors de la seule composante pour laquelle  $s/N \neq 0$ ; toutes les autres composantes, dites *petites*, peuvent avoir une taille arbitrairement élevée, mais aucune d'entre elles ne compose à elle seule une fraction commensurable du réseau<sup>1</sup>.

On définit maintenant un second ensemble, paramétrisé par  $N$  et  $T$ , où chacun des réseaux est construit à l'aide du processus suivant :

1. tirer un réseau de  $N$  nœuds du premier ensemble,
2. conserver chacun des liens de ce réseau avec une probabilité  $T$ , les retirant donc avec une probabilité  $1 - T$ .

Lorsqu'il existe, on définit le *seuil de percolation* comme la valeur  $T_c \in [0, 1)$  telle que cet ensemble possède une composante géante pour  $T > T_c$ . Il s'agit là en fait de *percolation de liens*; la *percolation de nœuds* (ou de sites) est définie de façon semblable, à la différence que « retirer un nœud » correspond à retirer tous les liens qui lui sont rattachés.

Traditionnellement, la percolation est utilisée sur des réseaux réguliers. Par exemple, on peut modéliser le passage d'un liquide dans un matériel poreux (comme de l'eau dans du café moulu) à l'aide d'une grille régulière tridimensionnelle à laquelle on retire un certain nombre de liens, pour représenter les obstacles. Au dessus du seuil de percolation, le liquide peut traverser le matériel en passant par la composante géante.

Dans le cas d'un processus *SIR* sur un réseau complexe, on peut interpréter  $T$  comme la *transmissibilité*, soit la probabilité qu'un nœud susceptible acquière l'infection d'un voisin infectieux avant que ce dernier ne soit retiré. Dans ce contexte, le seuil de percolation correspond donc au *seuil épidémique*, seuil au dessus duquel une épidémie est possible, et la taille de la composante géante correspond à la proportion de la population qui serait affectée en cas d'épidémie. Si une telle correspondance existe réellement, on devrait pouvoir décrire  $T$  en terme des paramètres du processus *SIR*.

---

1. Supposons que ce soit le cas : la taille de la seconde plus grande composante est  $s^*$  et le ratio  $s^*/N$  tend vers la valeur  $S^* \in (0, 1 - S]$ . Ceci implique l'absence explicite de  $ss^* = SS^*N^2 = \mathcal{O}(N^2)$  liens unissant ces deux composantes (qui n'en formeraient une seule de taille  $s + s^*$  autrement). Cette situation pathologique a une probabilité nulle de se produire dans la limite  $N \rightarrow \infty$  pour un modèle d'Erdős et Rényi, pour un modèle de configuration, et pour tous ensembles nous intéressant ici.



## 2.2 Transmissibilité

Suivant l'approche de Mark E. J. Newman [70], on s'intéresse à la probabilité  $T_{ij}$  qu'un nœud  $i$  demeurant infectieux un temps  $\tau_i$  (la *période infectieuse*) puisse transmettre l'infection à un nœud  $j$  dont il est voisin. Si la probabilité de transmission est de  $\beta_{ij}dt$  pour l'intervalle de temps  $[t, t + dt)$  (avec  $0 \leq t < \tau_i$ ), la probabilité qu'il *n'y ait pas* de transmission est

$$1 - T_{ij} = \lim_{\Delta t \rightarrow 0} (1 - \beta_{ij}\Delta t)^{\tau_i/\Delta t} = e^{-\beta_{ij}\tau_i} \quad , \quad (2.1)$$

ce qui fournit  $T_{ij} = 1 - e^{-\beta_{ij}\tau_i}$ . Supposant maintenant que  $\tau_i$  et  $\beta_{ij}$  sont respectivement tirés de distributions de probabilités indépendantes  $f(\tau_i)$  et  $g(\beta_{ij})$ , on obtient alors

$$\langle T_{ij} \rangle = 1 - \int_0^\infty \int_0^\infty f(\tau_i)g(\beta_{ij}) e^{-\beta_{ij}\tau_i} d\beta_{ij}d\tau_i \quad . \quad (2.2)$$

Dans le cas d'un système *SIR* où chaque nœud a le même taux de retrait  $\alpha$  et d'infection  $\beta$ , on a  $f(\tau_i) = \alpha e^{-\alpha\tau_i}$  et  $g(\beta_{ij}) = \delta(\beta_{ij} - \beta)$  et donc

$$\langle T_{ij} \rangle = 1 - \int_0^\infty \alpha e^{-\alpha\tau_i} e^{-\beta\tau_i} d\tau_i = \frac{1}{1 + \alpha/\beta} \quad . \quad (2.3)$$

Jusqu'ici, toutes ces étapes sont justifiées. L'erreur commise dans [70], et relevée dans [49], consiste à supposer que si tous les nœuds et liens du réseau tirent leurs  $\tau_i$  et  $\beta_{ij}$  des mêmes distributions indépendantes  $f(\tau_i)$  et  $g(\beta_{ij})$ , alors la transmissibilité  $T = \langle T_{ij} \rangle$  est identique pour chaque paire de voisins  $i$  et  $j$ , rendant la percolation de liens équivalente à un processus de propagation *SIR* sur réseaux. Le problème est le suivant : les probabilités d'infection  $T_{ij}$  et  $T_{ij'}$  de deux nœuds susceptibles  $j$  et  $j'$  voisins d'un même nœud infectieux  $i$  sont corrélées à travers  $\tau_i$ , et devraient donc être tirées une seule fois<sup>2</sup> de la distribution  $f(\tau_i)$ .

Cependant, le chapitre 3 ne s'attarde pas trop à ces questions ; on relègue le reste de cette discussion à la section 4.1. Pour l'instant, on considère simplement que chacun des voisins d'un nœud infectieux a une probabilité  $T$  indépendante d'être infecté (niveau « idéalisation » de la section 1.3.4) et on souhaite reproduire ce comportement de façon analytique (niveau « simplification »).

---

2. Par opposition,  $\beta_{ij}$  et  $\beta_{ij'}$  sont indépendamment tirés de la distribution  $g(\beta_{ij})$ .

## 2.3 Fonctions génératrices

Pour paraphraser Herbert S. Wilf [105], une *fonction génératrice* est une corde à linge permettant de suspendre une séquence de nombres<sup>3</sup>. Par exemple, la série de puissances

$$f(x) = \sum_k a_k x^k \quad (2.4)$$

représente la séquence  $\{a_0, a_1, a_2, \dots\}$  sous la forme d'une fonction  $f(x)$ ; on dit de la fonction génératrice  $f(x)$  qu'elle *génère* la séquence  $\{a_0, a_1, a_2, \dots\}$ , et chacun des  $a_n$  peut être récupéré à l'aide de la relation

$$a_n = \frac{1}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=0}, \quad (2.5)$$

lorsque  $f(x)$  est analytique dans un domaine incluant l'origine. Une fonction génératrice générant ainsi les coefficients d'une série de puissance est dite *ordinaire*; à moins de mention du contraire, toutes les fonctions génératrice de cette thèse sont ordinaires. Le concept est cependant beaucoup plus général : le lecteur intéressé peut consulter [105] pour plus de détails.

Plusieurs propriétés de ces séries de puissances permettent de manipuler de façon compacte les séquences générées tout en profitant de la puissance des outils analytiques. Par exemple, la somme des éléments générés par la fonction  $f(x)$  est, lorsqu'elle existe, donnée par

$$\sum_k a_k = \sum_k a_k 1^k = f(1) \quad . \quad (2.6)$$

De même, d'autres sommes intéressantes comme

$$\sum_k \frac{k!}{(k-n)!} a_k = \left. \frac{d^n f(x)}{dx^n} \right|_{x=1} \quad \text{et} \quad \sum_k k^n a_k = \left[ \left( x \frac{d}{dx} \right)^n f(x) \right]_{x=1} \quad (2.7)$$

peuvent être obtenues par différenciation.

On appelle *fonction génératrice de probabilités* une fonction générant les éléments  $\{p_0, p_1, \dots\}$  d'une distribution de probabilité. Dans ce cas particulier, les relations précédentes donnent le critère de normalisation  $f(1) = 1$ , la moyenne  $\langle k \rangle = f'(1)$  et la variance  $\langle k^2 \rangle - \langle k \rangle^2 = f''(1) - f'(1)[f'(1) - 1]$ .

---

3. « A generating function is a clothesline on which we hang up a sequence of numbers for display. »

La convolution de deux séquences est générée par le produit de leurs fonctions génératrices. Ainsi, étant donné

$$f(x) = \sum_k a_k x^k, \quad g(x) = \sum_l b_l x^l \quad \text{et} \quad h(x) = \sum_m c_m x^m, \quad (2.8a)$$

on a

$$f(x)g(x) = \sum_{kl} a_k b_l x^{k+l} = \sum_m \underbrace{[\sum_k a_k b_{m-k}]}_{c_m} x^m = h(x) \quad . \quad (2.8b)$$

Ce résultat se généralise aisément : le produit de plusieurs fonctions génératrices génère la convolution des séquences respectives. En particulier,  $[f(x)]^n$  génère la séquence obtenue en convoluant  $n$  séquences  $\{a_0, a_1, \dots\}$  entre elles.

Pour des  $n$  un tant soit peu élevés, l'évaluation numérique de  $a_n$  en utilisant directement la relation (2.5) est sujette à de grandes erreurs d'arrondissement. On préfère donc plutôt la forme intégrale [75]

$$a_n = \frac{1}{2\pi i} \oint \frac{f(z)}{z^{n+1}} dz = r^{-n} \int_0^1 e^{-2\pi i n \rho} f(r e^{2\pi i \rho}) d\rho \quad (2.9)$$

obtenue du théorème de Cauchy en choisissant un parcours d'intégration complexe circulaire de rayon  $r > 0$  autour de l'origine,  $r$  devant être suffisamment petit pour que  $f(z)$  soit analytique dans le domaine circonscrit.

Une approximation de cette expression devient particulièrement utile lorsque la séquence décroît rapidement avec  $n$ . On discrétise d'abord l'intégrale en évaluant la fonction en  $M$  points  $\{f_0, f_1, \dots, f_{M-1}\}$  également espacés tels que  $f_m = f(r e^{2\pi i \frac{m}{M}})$  [15, 56]

$$a_n \approx \frac{1}{Mr^n} \sum_{m=0}^{M-1} f_m e^{-2\pi i n \frac{m}{M}} \quad . \quad (2.10)$$

Puisque ces  $f_m$  ne dépendent pas de  $n$ , les même points peuvent être utilisés pour évaluer différents  $a_n$ . En fait, la somme s'avère être une transformée de Fourier discrète et la relation

$$\{Ma_0, Mr^2 a_1, Mr^2 a_2, \dots, Mr^{M-1} a_{M-1}\} \approx \mathcal{F}^{-1}\{f_0, f_1, f_2, \dots, f_{M-1}\} \quad (2.11)$$

permet de calculer  $\{a_0, a_1, \dots, a_{M-1}\}$  en une seule passe d'un algorithme de transformée de Fourier rapide inverse<sup>4</sup>.

---

4. L'opération  $\mathcal{F}^{-1}$  suppose ici une définition de la transformée inverse sans normalisation, comme celle utilisée dans l'algorithme FFTW [31]. Par opposition, supposant  $r = 1$ , la commande `ifft` de MATLAB donne directement  $\{a_0, a_1, \dots, a_{M-1}\}$  à partir des  $\{f_0, f_1, \dots, f_{M-1}\}$  puisqu'elle normalise la séquence en la multipliant par  $\frac{1}{M}$ .

Si  $f(x)$  est un polynôme d'ordre  $N$  tel que  $N < M$ , l'équation (2.11) devient exacte. En effet, la discrétisation effectuée en (2.10) représente exactement l'intégrale (2.9) dans la limite  $M \rightarrow \infty$  et, par le théorème d'échantillonnage de Nyquist-Shannon, le résultat de (2.11) pour  $0 \leq n \leq N$  ne dépend pas de  $M$  en autant que  $N < M$ . Lorsque  $N \geq M$ , un phénomène de crénelage (« *aliasing* ») se produit et les  $a_n$  pour  $n \geq M$  sont repliés sur les valeurs plus basses, causant de l'erreur sur ces valeurs. Cependant, même pour une séquence infinie, il est souvent possible de choisir un  $M$  suffisamment grand pour que l'erreur commise soit acceptable, à condition que les  $a_n$  décroissent suffisamment rapidement.

Même lorsque l'équation (2.11) est exacte d'un point de vue analytique, l'imprécision numérique peut parfois devenir importante. Afin de minimiser cette erreur, le choix du rayon  $r$  du parcours d'intégration peut nécessiter certains ajustements :  $r$  doit être le plus grand possible, tout en étant suffisamment petit pour éviter les singularités et leur voisinage. Cavers [15] recommande d'utiliser un  $r$  de l'ordre de 90% de la distance à l'origine du pôle qui en est le plus près<sup>5</sup> ; le lecteur intéressé peut consulter cet ouvrage pour plus de détails sur la précision de l'approximation.

## 2.4 Prélude à « Time evolution of epidemic disease on finite and infinite networks »

Le chapitre 3 généralise le formalisme de percolation sur modèles de configuration canoniques gelés non dirigés de Newman [70] afin de tenir compte : 1. de l'évolution dans le temps à travers des générations d'infection et 2. de l'effet de la taille finie du système. Tout comme dans [70, 75], la séquence  $\{p_0, p_1, p_2, \dots\}$  spécifiant la probabilité  $p_k$  qu'un nœud sélectionné aléatoirement soit de degré  $k$  est générée par la fonction génératrice de probabilités  $G_0(x)$  (définie en (3.1)). De même, pour un nœud atteint en suivant un lien choisi au hasard, la fonction  $G_1(x)$  (définie en (3.3)) génère la distribution de probabilité pour le nombre d'*autres* liens que possède ce nœud, soit son *degré excédentaire*.

Le processus de percolation que l'on souhaite modéliser est le suivant : pour chacun de ses voisins susceptibles, chaque nœud infectieux a une probabilité  $T$  de le rendre infectieux. Puisque ce  $T$ , la transmissibilité, correspond à la probabilité qu'un nœud susceptible acquière l'infection d'un voisin infectieux avant que ce dernier ne soit retiré (*c.f.* 2.1), ce test ne doit être effectué qu'une seule fois par voisin susceptible d'un nœud

---

5. En fait, la coordonnée  $z$  de Cavers correspond à  $z^{-1}$  dans notre notation. Ainsi, il recommande un  $r$  qui est 10% *supérieur* à la distance à l'origine du pôle le plus *éloigné*.

infectieux. Tous les nœuds du réseau sont initialement susceptibles à l'exception d'un seul infectieux, le *patient zéro*, que l'on choisit au hasard. Au final, on s'intéresse au nombre de nœuds portant l'infection ou l'ayant déjà portée, c'est-à-dire à la somme du nombre de nœuds infectieux et retirés : on appelle ce nombre la *taille de l'infection*.

On s'intéresse d'abord au nombre d'infections directement causées par le patient zéro. Comme ce nœud a été choisi au hasard, sa distribution en degré est générée par  $G_0(x)$  et, puisqu'il s'agit initialement du seul nœud infectieux, chacun de ses voisins est susceptible d'acquérir l'infection avec une probabilité  $T$ . Toujours comme dans [70], on obtient en équation (3.5) que le nombre d'infections directement causées par le patient zéro est générée par  $G_0(1 + [x - 1]T)$ .

Dans un premier temps, on considère que le système est de taille infinie. Ainsi, puisque les liens du modèle de configuration sont assignés de façon aléatoire, la probabilité qu'un nœud infectieux soit lié à un autre nœud infectieux que celui dont il a acquit l'infection tend vers zéro : la propagation de l'infection forme donc un arbre, soit une composante sans cycles, et l'étude de la structure de cette arbre correspond à un *processus de branchement*. Comme un nœud infectieux autre que le patient zéro est atteint en suivant un lien, son degré excédentaire est donné par  $G_1(x)$ , et le nombre d'infections directement causées par ce nœud est généré par  $G_1(1 + [x - 1]T)$ .

Se tournant maintenant vers Marder [56], on introduit le concept de *génération d'infection*, un nombre entier  $g \geq 0$ . La génération 0 est uniquement composée du patient zéro, les nœuds directement infectés par le patient zéro forment la génération 1, et les nœuds directement infectés par la génération  $g$  forment la génération  $g + 1$ . On peut également percevoir la génération comme une mesure discrète du temps : au temps  $g$ , les nœuds de génération  $g$  sont infectieux et ceux de génération inférieure à  $g$  sont retirés. La taille de l'infection au temps  $g$  correspond donc à la somme des nombres de nœuds dans les générations inférieures ou égales à  $g$ .

Dans [56], Marder obtient les fonctions générant la taille de l'infection à la génération  $g$  par une composition de fonctions

$$x \text{ à la génération } 0, \quad (2.12a)$$

$$x G_0(1 + [x - 1]T) \text{ à la génération } 1, \quad (2.12b)$$

$$x G_0\left(1 + [x G_1(1 + [x - 1]T) - 1]T\right) \text{ à la génération } 2, \quad (2.12c)$$

$$x G_0\left(1 + \left[x G_1\left(1 + [x G_1(1 + [x - 1]T) - 1]T\right) - 1\right]T\right) \text{ à la génération } 3, \quad (2.12d)$$

...

pouvant s'écrire  $x G_0(1 + [F_{g-1}(x) - 1]T)$  pour la génération  $g > 0$  à l'aide de la

réursion<sup>6</sup>

$$F_g(x) = x G_1(1 + [F_{g-1}(x) - 1]T) \quad \text{et} \quad F_0(x) = x \quad . \quad (2.13)$$

Newman [70] saute ces étapes et définit plutôt la fonction générant le nombre total d'infections découlant de la transmission de l'infection le long d'un lien

$$H_1(x) = x G_1(1 + [H_1(x) - 1]T) \quad (2.14)$$

de façon à ce qu'elle soit cohérente avec elle-même. Il obtient ensuite la fonction génératrice de probabilité pour la taille finale de l'infection

$$H_0(x) = x G_0(1 + [H_1(x) - 1]T) \quad , \quad (2.15)$$

ce qui correspond à l'expression de Marder dans la limite  $g \rightarrow \infty$ . Ainsi, l'approche de Newman ne permet pas d'étudier l'évolution temporelle du processus, même de façon discrète. Il n'en demeure pas moins que l'équation (2.14) mène à des relations très intéressantes pour l'état final du système, comme par exemple la fraction totale occupée par la composante géante en cas d'épidémie

$$S = 1 - G_0(1 + [u - 1]T) \quad , \quad \text{où} \quad u = G_1(1 + [u - 1]T) \quad . \quad (2.16)$$

Les relations (2.13) et (2.14) de Marder et Newman sont toutes deux définies « à rebours » par rapport au processus de propagation : la contribution d'une génération  $g$  dépend de celle de la génération  $g + 1$ . Grâce à la fonction génératrice de deux variables  $\Psi_0^g(x, y)$  (définie en (3.9)), informant non seulement de la taille de l'infection à la génération  $g$  mais aussi du nombre de nouvelles infections à cette génération, l'approche présentée au chapitre 3 permet une relation de récurrence (3.13) allant « vers l'avant » : la contribution d'une génération  $g$  dépend de celle de la génération  $g - 1$ .

Cette distinction s'avère très importante lorsque l'on souhaite considérer un système de taille finie  $N$ . En effet, puisque le processus itératif a accès à l'état de la situation à la génération  $g - 1$ , il est possible de le modifier afin de tenir compte de la présence de cycles ou, plus concrètement, du fait qu'une fraction des voisins d'un nœud infectieux peuvent eux-même avoir déjà acquis l'infection. Cette démarche mène à la relation (3.36), la principale contribution du chapitre 3, permettant d'étudier de façon stochastique l'évolution temporelle discrète de la propagation sur un réseau fini tiré d'un modèle de configuration canonique gelé non dirigé.

---

6. Ces définitions ont été légèrement adaptées de celles de [56] pour fins de comparaison.

# Chapitre 3

## Time evolution of epidemic disease on finite and infinite networks <sup>1</sup>

P.-A. Noël,<sup>†</sup> B. Davoudi,<sup>\*</sup> R. C. Brunham,<sup>\*</sup> L. J. Dubé<sup>†‡</sup> and B. Pourbohloul<sup>\*§</sup>

<sup>\*</sup>University of British Columbia Centre for Disease Control, Vancouver

<sup>†</sup>Département de Physique, de Génie Physique et d'Optique, Université Laval, Québec

<sup>‡</sup>Laboratoire de Chimie Physique-Matière et Rayonnement, Université Pierre et Marie Curie, Paris

<sup>§</sup>Department of Health Care and Epidemiology, University of British Columbia, Vancouver

Deux approches générales permettent de concevoir des modèles mathématiques en principe solubles de façon analytique. La première, souvent connue sous le nom de modélisation compartimentale, traite l'évolution temporelle de la propagation de l'infection au coût d'une simplification des interactions menant à la transmission. La seconde approche, fondée sur la théorie des réseaux, tient compte de la structure des contacts entre les individus mais néglige toutes considérations temporelles. Jusqu'à maintenant, la seule alternative permettant l'intégration simultanée de ces deux aspects *tout en préservant la variété des issues possibles* a été d'abandonner l'approche analytique et de se tourner vers des simulations informatiques. Nous offrons un nouveau cadre analytique incorporant à la fois la complexité d'une structure de contacts en réseaux et la progression dans le temps de l'étalement de l'infection. De plus, nous démontrons que cette approche est tout aussi efficace pour des réseaux de taille finie que pour ceux de taille « infinie ». Ce formalisme peut également être appliqué à des processus similaires de percolation sur réseaux dans d'autres secteurs des sciences et des technologies.

Mathematical models of infectious diseases, which are in principle analytically tractable, use two general approaches. The first approach, generally

---

1. Originellement publié dans [79].

known as compartmental modelling, addresses the time evolution of disease propagation at the expense of simplifying the pattern of transmission. The second approach uses network theory to incorporate detailed information pertaining to the underlying contact structure among individuals while disregarding the progression of time during outbreaks. So far, the only alternative that enables the integration of both aspects of disease propagation simultaneously *while preserving the variety of outcomes* has been to abandon the analytical approach and rely on computer simulations. We offer a new analytical framework, which incorporates both the complexity of contact network structure and the time progression of disease spread. Furthermore, we demonstrate that this framework is equally effective on finite- and “infinite”-size networks. This formalism can be equally applied to similar percolation phenomena on networks in other areas of science and technology.

### 3.1 Introduction

The spread of communicable diseases is a dynamical process and as such, understanding and controlling infectious disease outbreaks and epidemics is pertinent to the temporal evolution of disease propagation. Historically, this aspect of disease transmission has been studied using coarse-grained dynamical representation of populations, known as compartmental models [5, 13, 20, 27, 38]. In these models, a population is divided into a number of epidemiological states (or classes) and the time evolution of each is described by a differential equation.

Although this approach, and its more complex variants, has been instrumental in understanding several features of infectious diseases over the past 3 decades, it comes with a major simplification. The simplifying assumption states that the population is “well mixed”, *i.e.*, every infectious individual has an equal opportunity to infect others. This assumption may be valid in the broader context of population biology. Human populations, however, tend to contact each other in a heterogeneous manner based on one’s age, profession, socio-economic status or behaviour, and thus, the well-mixed approximation cannot portray an accurate image of disease spread among humans [58].

Recent advances in network- and percolation-theories, have paved the way for physicists to bring a new perspective to understanding disease spread. Over the past decade, seminal works by Watts and Strogatz on small-world networks [100, 101], Barabasi *et al.* on scale-free networks [1] and Dorogovtsev, Mendes [22], Pastor-Satorras and Vespignani [86], among others, on the dynamics of networks, have shed light on a number



of intriguing aspects of epidemiological processes. In particular, groundbreaking work by Newman *et al.* [70, 72, 75] has provided a strong foundation for the formulation of epidemiological problems using tools developed by physicists.

Various dynamical processes that propagate from neighbour to neighbour on complex (natural or artificial) networks, *e.g.*, a crawler (or spider) browsing the World Wide Web or rumours spreading in a population, reveals interesting similarities with the spread of epidemics in human population [65, 66]. In the present work, we specifically focus on disease propagation as the dynamical phenomenon and use the associated terminology. Our methodology is however quite general and can be applied *mutatis mutandis* to other processes that manifest similar dynamical properties.

We are primarily interested in diseases where infected individuals are eventually removed from the dynamics of the system (*i.e.*, infection is followed by naturally-acquired immunity or death), implying that the same person cannot be infected more than once. At any given time, we call an individual “susceptible” if he has never been exposed to the disease; “exposed” if he has acquired the infection but not currently able to pass on the disease to another person; “infectious” if he is currently able to transmit the infection to others; “removed” if he became immune or succumbed to death after acquiring the infection; and finally, “infected” if he has been exposed to the infectious agent at least once in the past, regardless of his current state (*e.g.*, exposed, infectious, removed).

Network analysis using the generating function formalism, developed by Newman *et al.*, has proven to be a powerful tool when analyzing the spread of disease within networks [70]. Without directly addressing the question of “when the transmission occurred?”, it provides reliable results on the final size of an outbreak/epidemic by addressing the question of “whether transmission occurred?”. The first question, *i.e.*, the time evolution of the system, is presently beyond the formalism as originally derived.

Several researchers have recognized the importance of incorporating the notion of time into the generating function formalism that describes percolation dynamics on networks. In broaching this issue, many notable advances have been made. Numerous contributions (for a recent review, see [11]) have used an approach closely related to compartmental models to assign a higher strength of infection to nodes of higher degree [67, 83, 84]. More recently, Marder [56] has calculated the probability distribution of outbreak sizes as a function of time for infinite-size networks, while Volz [99] has addressed the finite-size effect by deriving estimates of the mean size of a large-scale epidemic, rather than the probability distribution. Despite these advances, one is required to develop a truly integrated analytical framework that simultaneously encompasses the time progression of disease, the network finite-size effects and the wide

variety of possible outcomes. The major steps towards an integrative formalism are the principal achievement of this paper.

The outline of the paper is as follows. In section 3.2 we define the type of dynamics that will be studied and recall some of the tools used by Newman [70] for infinite networks, to which we add the concept of generations and phase-space representation. In section 3.3 we extend these tools for uses on finite-size networks and section 3.4 presents some results of our analysis, compares them with some existing models, and discuss possible extensions of the formalism. In section 3.5 we give our conclusions and Appendices complete the analysis of section 3.3 and 3.4.

## 3.2 Formalism for infinite networks

We map a system of  $N$  individuals to a *network* in which each individual is represented by a *node* (or vertex) and the connection between each pair of individuals is represented by a *link* (or edge). Two nodes are *neighbours* if they are joined by a link. Contrary to compartmental approaches, a network representation of a system takes into account that each node does not have the same probability of interacting with every other node; in fact, one interacts only with its topological neighbours. Each node has a *degree*  $k_i$  (number of neighbours) and the set  $\{k_i\}$  (called *degree sequence*) partially defines the network. Similarly, the set of probabilities  $\{p_k\}$  that a random node has degree  $k$  is the *degree distribution*.

In many practical situations, the degree distribution is, together with the size  $N$  of the network, the only available information on the network structure. We consider the ensemble of all possible networks obtained by drawing a degree sequence from the provided degree distribution and then, for every node  $i \in \{1, 2, \dots, N\}$ , randomly connect each of its  $k_i$  links to those of other nodes (no *self-loops*) [70]. The quantities obtained through this paper are averages on this ensemble. We add to this maximum entropy definition the restriction that two nodes cannot share more than one link (no *repeated links*). For sparse graphs — in which the number of links scales linearly with the number of nodes — the probability for such an event decreases as  $1/N$  and can be neglected for large networks [75].

To perform Monte Carlo simulations of epidemic propagations on a network, one requires an explicit knowledge of that network structure. We have used the following method [70] to produce a network belonging to the ensemble described above:

- (i) generate a random degree sequence  $\{k_i\}$  of length  $N$  subjected to the degree

distribution  $\{p_k\}$ ;

- (ii) make sure that  $\sum_i k_i$  is an even number since a link is composed of 2 “stubs”;
- (iii) for each  $i$ , produce a node with  $k_i$  stubs;
- (iv) randomly choose a pair of unconnected stubs and connect them together. Repeat until all unconnected stubs are exhausted;
- (v) test for the presence of self-loops and repeated links. Remove the faulty stubs by randomly choosing a pair of connected stubs and rewire them to the former stubs. Repeat until no self-loop and/or repeated links are found.

### 3.2.1 Basic generating functions

With knowledge of the degree sequence of the physical or social network of interest, we can obtain the corresponding degree distribution  $\{p_k\}$ . Following Newman *et al.* [75] and Newman [70], we define the probability generating function (*pgf*) [105] for a random node as

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k \quad , \quad (3.1)$$

respecting the normalization  $G_0(1) = \sum_{k=0}^{\infty} p_k = 1$ . The average degree,  $z_1$ , can be easily obtained from

$$z_1 = \langle k \rangle = \sum_{k=0}^{\infty} k p_k = G'_0(1) \quad , \quad (3.2)$$

where the prime denotes the derivative with respect to the argument. The probability,  $q_k$ , that  $k$  nodes could be reached from the node we arrived at by following a random link (excluding this link from the count), can also be derived as

$$G_1(x) = \sum_{k=0}^{\infty} q_k x^k = \frac{\sum_k (k+1) p_{k+1} x^k}{\sum_k (k+1) p_{k+1}} = \frac{1}{z_1} G'_0(x) \quad . \quad (3.3)$$

While  $G_0(x)$  and  $G_1(x)$  contain information about the structure of the physical network linking nodes within the epidemiological system, they do not contain any information about the risk of disease transmission between two neighbouring nodes. However, with the additional knowledge of the transmissibility,  $T$  — the probability that an infectious node will infect one of its neighbours — we can determine the probability of infecting  $l$  out of  $k$  neighbours as

$$p_{l|k} = \binom{k}{l} T^l (1-T)^{k-l} \quad . \quad (3.4)$$

The *pgf* for the number of infections directly caused by the initially infected node (“patient zero”) is then

$$\sum_{l=0}^{\infty} \sum_{k=l}^{\infty} p_k p_{l|k} x^l = \sum_{k=0}^{\infty} \sum_{l=0}^k p_k \binom{k}{l} T^l (1-T)^{k-l} x^l = G_0 (1 + (x-1)T) \quad . \quad (3.5)$$

Similarly, the probability distribution for the number of infections directly caused by a node reached by following a random link is generated by  $G_1(1 + (x-1)T)$ .

We can continue in the same vein and obtain informative results about the final state of the population in an infinite network, after the outbreak/epidemic has ended [58, 70]; however, this approach in itself does not yield any information about the duration of the epidemic, its speed of propagation or other time-related quantities.

### 3.2.2 Generations

To study the progression of the outbreak/epidemic over the network, we adopt an approach based on *generations of infection*. We define *generation 0* as the initial infected node of the outbreak/epidemic; nodes of generation  $g$  are those who acquire the disease from a member of generation  $g-1$ .

There is clearly a causality link among generations and we would expect nodes of higher generations to become infected, on average, later than those in generations closer to the initial infected node. In a future contribution, we will look more closely into the relationship between generations and continuous time evolution.

We now extend the generating function formalism to introduce a new *pgf* for an arbitrary generation  $g$

$$G_g(x) = \begin{cases} G_0(x) & (g = 0) \\ G_1(x) & (g \geq 1) \end{cases} \quad . \quad (3.6)$$

As in equation (3.5), the *pgf* for the number of nodes that acquire infection directly from a single node of generation  $g$  is given by  $G_g(1 + (x-1)T)$ . This assumption holds when the total number of infected nodes in the current and previous generations is small compared to the size,  $N$ , of the network. In such a case, the probability of infecting a node that is already infected is proportional to  $1/N$ ; we assume that a node cannot be infected twice and the propagation of the disease follows a tree-like structure (without a closed loop). This condition is fulfilled in large networks either when there is no giant

component or when we limit ourselves to the first few generations. Section 3.3 removes these limitations to some extent.

From the properties of the  $pgf$ 's, the expected number of secondary infections caused directly by an infected node in generation  $g$  is given by

$$\langle l_g \rangle = \left. \frac{dG_g(1 + (x-1)T)}{dx} \right|_{x=1} = \begin{cases} TG'_0(1) & (g=0) \\ TG'_1(1) & (g \geq 1) \end{cases} . \quad (3.7)$$

Notice that  $\langle l_g \rangle$  is identical for every generation, except the first one. It corresponds to a fundamental quantity in epidemiology, the *basic reproductive number* [5, 20]

$$R_0 = \langle l_1 \rangle = \langle l_2 \rangle = \dots = TG'_1(1) = T \frac{z_2}{z_1} , \quad (3.8)$$

where  $z_2 = G''_0(1)$  is the expected number of second neighbours for a randomly chosen node.  $R_0 < 1$  implies that the expected number of infectious nodes decreases in consecutive generations, leading to the extinction of the disease. Conversely,  $R_0 > 1$  implies that the expected number of infectious nodes increases and can potentially lead to an epidemic, which is a *giant component* of occupied links in the language of percolation theory [95]. It is worth noting that  $R_0 > 1$  alone does not guarantee the occurrence of an epidemic; indeed, some realizations may have a number of new infections below the mean value  $R_0$ , and therefore lead to the extinction of the disease.

### 3.2.3 Phase-space representation

To proceed further, we define the quantity  $\psi_{sm}^g$  as the probability of having  $s$  infected nodes by the end of the  $g$ -th generation, of which  $m$  became infected during the  $g$ -th generation. This probability is generated by

$$\Psi_0^g(x, y) = \sum_{s,m} \psi_{sm}^g x^s y^m . \quad (3.9)$$

Each element of the (triangular) matrix,  $\psi_{sm}^g$ , can be seen as a possible “state of infection” where the  $s$  and  $m$  dimensions provide information about the “position” (number of infected) and “momentum” (new infections) in the infection space, respectively.

We know from the previous sections that the probability distribution for the number of nodes that acquire infection directly from a single node of generation  $g-1$  is generated by  $G_{g-1}(1 + (x-1)T)$ . Moreover, the  $pgf$  for the sum of two independent quantities is given by the product of their  $pgf$ 's [75, 105]. Therefore, the probability  $P(m|s', m')$

that each state  $(s', m')$  of generation  $g - 1$  leads to  $m$  new infections in generation  $g$  is generated by

$$\sum_m P(m|s', m')x^m = [G_{g-1}(1 + (x-1)T)]^{m'} \quad (3.10)$$

Also, the state  $(s', m')$  has probability  $\psi_{s'm'}^{g-1}$  at generation  $g - 1$  and thus makes a contribution  $\psi_{s'm'}^{g-1}P(m|s', m')$  to the state  $(s = s' + m, m)$  at generation  $g$ . Hence, we obtain the recurrence relation

$$\sum_{s', m'} \psi_{s'm'}^{g-1} x^{s'} [G_{g-1}(1 + (xy-1)T)]^{m'} = \sum_{s, m} \psi_{sm}^g x^s y^m \quad (3.11)$$

with the initial condition  $\psi_{sm}^0 = \delta_{s1}\delta_{m1}$  ( $\delta_{ij}$  is the Kronecker delta). The states for which  $m' = 0$  are absorbing states; the region above the main diagonal ( $m' > s'$ ) is forbidden; and the main diagonal ( $m' = s'$ ) is only accessible for the initial condition.

Finally, inserting equation (3.11) into equation (3.9) provides the (*forward*) recurrence relation (for  $g \geq 1$ )

$$\Psi_0^g(x, y) = \sum_{s', m'} \psi_{s'm'}^{g-1} x^{s'} [G_{g-1}(1 + (xy-1)T)]^{m'} \quad (3.12)$$

with the initial condition  $\Psi_0^0(x, y) = xy$  or  $\psi_{sm}^0 = \delta_{s1}\delta_{m1}$ , where  $\delta_{ij}$  is the Kronecker delta. Equation (3.12) implies further the recurrence

$$\Psi_0^g(x, y) = \Psi_0^{g-1}(x, G_{g-1}(1 + (xy-1)T)) \quad (3.13)$$

Note that  $\Psi_0^g(x, 1)$  generates the probabilities

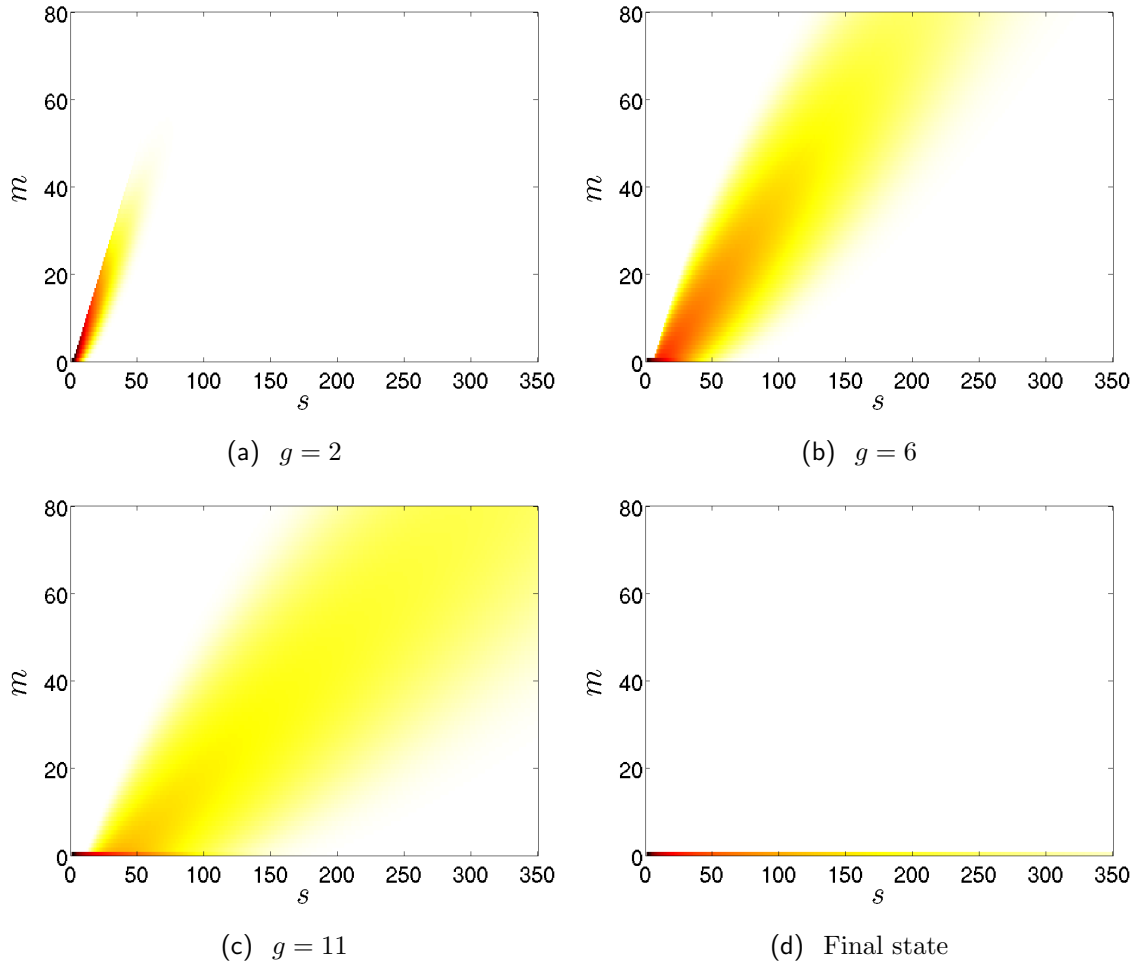
$$p_s^g = \sum_m \psi_{sm}^g \quad (3.14)$$

that  $s$  nodes are infected at generation  $g$ , independent of the number of new infections. Similarly,  $\Psi_0^g(1, y)$  generates the probabilities

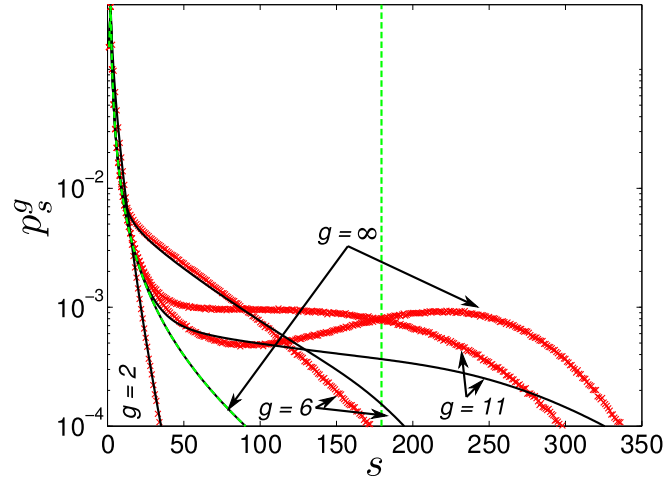
$$p_m^g = \sum_s \psi_{sm}^g \quad (3.15)$$

that  $m$  nodes are infected during generation  $g$ .

Figure 3.1 illustrates some results of this method for a network of  $N = 1\,000$  nodes with a power law distribution  $p_k \propto k^{-\tau} e^{-k/\kappa}$  with  $\tau = 2$ ,  $\kappa = 5$  and the transmissibility  $T = 0.8$ . The phase-space representation for generations 2, 6 and 11 as well as the final state are shown. Figure 3.2 provides the corresponding projection  $p_s$  on the  $s$  axis (solid curves). The numerical results (crosses), theoretical “infinite-size” outbreak



**Figure 3.1:** Representation of  $(s, m)$  phase space for the *infinite-size* network algorithm:  $s$  denotes the number of infected nodes by the end of the  $g$ -th generation and  $m$  denotes the number of new infections that occurred since the last generation. The degree distribution of the  $N = 1\,000$  nodes follows a power-law  $p_k \propto k^{-\tau} e^{-k/\kappa}$  with  $\tau = 2$  and  $\kappa = 5$  and the probability of transmission along an edge is  $T = 0.8$ . The phase-space representations, equation (3.12), are displayed for generations 2, 6, 11 and the final state.



**Figure 3.2:** Projection  $p_s^g$  on the  $s$  axis of the  $(s, m)$  phase space for the *infinite-size* network algorithm:  $s$  denotes the number of infected nodes by the end of the  $g$ -th generation and  $p_s^g$  denotes the probability for  $s$  to occur. The degree distribution and transmissibility of figure 3.1 are used and the generations shown (2, 6, 11 and final state) are also the same. The value of  $p_s^g$  is plotted against  $s$  (solid curves). The numerical results (dots), theoretical “infinite-size” outbreak distribution (dashed curves) and theoretical “infinite-size” epidemic size (vertical dashed lines) are also displayed. Numerical results are obtained by creating an ensemble of  $10^3$  equivalent graphs, each of which is used to run  $10^5$  simulations, performing  $10^8$  epidemic simulations in total. The discrepancies are explained and corrected in section 3.3.

distribution (dashed curves) and theoretical “infinite-size” epidemic size (vertical dashed lines) are also displayed. Numerical results are obtained by creating an ensemble of  $10^3$  equivalent graphs, each of which was used to run  $10^5$  simulations, performing  $10^8$  epidemic simulations in total.

Figure 3.2 clearly demonstrates that apart from the small-scale outbreaks, the results from the infinite-size formalism may not correctly predict the outbreak/epidemic size distribution for *finite-size* networks when the fraction of the network that has been infected is no longer negligible. The remedy to this shortcoming is offered in the next section.

### 3.3 Formalism for finite networks

As long as one is only interested in the *initial stage* of an outbreak, the finite-size of a network has negligible effects on the dynamics of disease spread. However, the impact of finite-size effects becomes important when a sizable fraction of the network has been affected. While the size of small outbreaks is mostly governed by stochastic fluctuations,



the size of the giant component (when one exists) is limited by two principal finite-size effects: the evolution over time of the degree distribution of susceptible nodes and the failure of transmission due to the impossibility of re-infection.

Since equation (3.12) is *exact* in the infinite limit, we can search for a similar form where the finite-size effects are introduced as a dependency in  $s'$  and  $m'$  of the degree distribution and/or of its parameters. In the following, we describe these effects and how they are introduced into the formalism described in the previous section.

It is worth noting that the finite-size effects considered here, affecting directly the dynamics on the network, should be distinguished from those that alter the structure of the network, for instance through a cutoff in the degree distribution as done in [85].

### 3.3.1 Evolution of the degree distribution of susceptibles

As the disease progresses across the network, susceptible nodes with a higher degree of connectivity are more likely to acquire the disease than those with fewer connections. If we focus only on the degree distribution of susceptible cases, the distribution will vary over time; the portion representing high-degree susceptibles will decrease and the segment representing low-degree susceptibles will increase, to comply with normalization requirements. This variability over time has a direct effect on the ratio  $z_2/z_1$  and can lower the reproduction number,  $R_0$ , below the threshold value of 1. These effects can potentially cause the extinction of the disease although a high number of susceptible nodes is still remaining. This is particularly important for degree distributions in which some nodes have a degree much higher than the mean degree distribution (*e.g.*, power-law distribution): the removal of these nodes will significantly impact the connectivity of the network.

To take this effect into account we define the generating function for the degree distribution of the remaining susceptibles for the current size,  $s$ , of the outbreak/epidemic

$$G_0^S(x; s) = \sum_k p_k^S(s) x^k \quad . \quad (3.16)$$

The mean number of susceptibles of degree  $k$  is thus given by  $(N - s)p_k^S(s)$ . However, the actual number of susceptibles in a network characterized by  $G_0^S(x; s)$  will in general be different from the mean value. Nonetheless, the difference becomes negligible in a *sufficiently large* population. In this limit, each  $p_k^S(s)$  can be treated as a continuous function of its parameter  $s$ . The assumption of a large population is less restrictive than it appears at first glance and, by comparison with the results of numerical simulations

presented at the end of this section (figure 3.3), we can say that it holds for reasonably small populations (*e.g.*,  $N = 1\,000$  for figure 3.3).

Since  $p_k^S(s)$  must be normalized ( $\sum_k p_k^S(s) = 1$ ) and because the susceptibles of degree  $k$  have a probability  $k$  times greater of being newly infected than those of degree 1, we can derive a differential equation system for the evolution of  $p_k^S(s)$

$$\frac{dp_k^S(s)}{ds} = \frac{p_k^S(s)}{N-s} \left( 1 - \frac{k}{z_1^S(s)} \right) \quad , \quad (3.17)$$

where the average degree is defined by

$$z_1^S(s) = \sum_k k p_k^S(s) \quad . \quad (3.18)$$

In the present dynamics, the first infection targets a random susceptible node and thus, does not affect the degree distribution. Therefore, we use the initial condition  $p_k^S(1) = p_k$  (with  $p_k$  being the degree distribution of the whole network) in equation (3.17) to get the solution

$$p_k^S(s) = p_k \frac{N-1}{N-s} [\theta(s)]^k \quad , \quad (3.19)$$

with  $\theta(s)$  given by

$$\theta(s) = \exp \left( - \int_1^s \frac{ds'}{(N-s')z_1^S(s')} \right) \quad . \quad (3.20)$$

The normalization of  $p_k^S(s)$  leads to the convenient expression

$$\sum_k p_k [\theta(s)]^k = G_0(\theta(s)) = \frac{N-s}{N-1} \quad . \quad (3.21)$$

Moreover, equation (3.19) allows us to express  $G_0^S(x; s)$  in terms of the original  $G_0(x)$  as

$$G_0^S(x; s) = \frac{N-1}{N-s} G_0(x\theta(s)) \quad . \quad (3.22)$$

For example, using the Poisson distribution

$$p_k = e^{-z} z^k / k! \quad (3.23)$$

in equation (3.21) gives

$$\frac{N-s}{N-1} = e^{-z} \sum_{k=0}^{\infty} \frac{[z\theta(s)]^k}{k!} = e^{-z} e^{z\theta(s)} \quad , \quad (3.24)$$

**Table 3.1:** Expression for  $\theta(s)$  for some commonly used degree distributions.

	Degree distribution	Expression for $\theta(s)$
Poisson	$p_k = \frac{e^{-z}}{k!} z^k$	$\theta(s) = \frac{1}{z} \left[ z + \ln \left( \frac{N-s}{N-1} \right) \right]$
Binomial	$p_k = \binom{N}{k} p^k (1-p)^{N-k}$	$\theta(s) = \frac{1}{p} \left[ \left( \frac{N-s}{N-1} \right)^{1/N} + p - 1 \right]$
Exponential	$p_k = (1 - e^{-1/\kappa}) e^{-k/\kappa}$	$\theta(s) = \frac{N-1 - (s-1)e^{1/\kappa}}{N-s}$
Power law	$p_k = \frac{k^{-\tau} e^{-k/\kappa}}{\text{Li}_\tau(e^{-1/\kappa})}$ for $k \geq 1$	$\text{Li}_\tau(e^{-1/\kappa} \theta(s)) = \frac{N-s}{N-1} \text{Li}_\tau(e^{-1/\kappa})$

from which we can easily isolate

$$\theta(s) = \frac{1}{z} \left[ z + \ln \left( \frac{N-s}{N-1} \right) \right] . \quad (3.25)$$

It follows that the degree distribution of the *susceptibles* is also a Poisson distribution

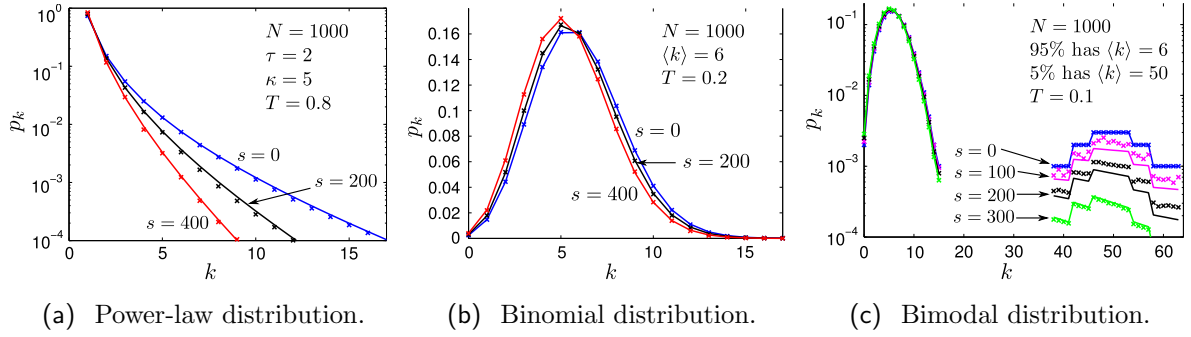
$$p_k^S(s) = e^{-z_1^S(s)} [z_1^S(s)]^k / k! \quad (3.26)$$

with an average degree given by

$$z_1^S(s) = z + \ln \left( \frac{N-s}{N-1} \right) . \quad (3.27)$$

Inspection of this last expression reveals that  $z_1^S(s)$  becomes negative when  $N-s < (N-1)e^{-z}$ . This limitation is due to the fact that the large population assumption is no longer respected: it implies the presence of  $(N-1)e^{-z}$  nodes of degree zero, but these nodes cannot be infected by the process leading to equation (3.17). Nevertheless, the probability of an epidemic reaching such high values of  $s$  typically vanishes in most realistic cases. Table 3.1 compiles the analytic forms of  $\theta(s)$  for some typical distributions. When a closed form satisfying equation (3.21) cannot be found, the quantity  $p_k^S(s)$  can nevertheless be derived numerically for each pair of  $k$  and  $s$ .

Figure 3.3 shows the variation of the degree distributions of susceptible nodes for 3 networks: (a) the power-law distribution used in figure 3.1; (b) a binomial distribution with  $p = 6/N$  and  $N = 1\,000$ ; and (c) a bimodal distribution, in which the vast majority of nodes (95%) has an average degree of 6 and the rest has an average degree of 50.



**Figure 3.3:** Time evolution of typical degree distributions representing susceptible nodes. The evolution of the degree distribution of the susceptibles is shown for 3 networks: (a) the power-law distribution used in figure 3.1; (b) a binomial distribution with  $p = 6/N$  and  $N = 1\,000$ ; and (c) a bimodal distribution, in which the vast majority of nodes (95%) has an average degree of 6 and the rest has an average degree of 50. The analytical (curves) and numerical (crosses,  $10^3$  equivalent graphs with  $10^3$  simulations per graph) results are given for different outbreak/epidemic sizes  $s$ . The transmissibility values  $T$  were solely used to produce numerical results.

This latter network is particularly interesting as it may correspond to realistic settings such as hospitals, schools or shopping malls. Although slight deviations are caused by the underlying assumptions, there is very good agreement between numerical and analytical results.

Once  $G_0^S(x; s)$  is known, we can show that the degree distribution of the susceptibles in the previous generation is given by  $G_0^S(x; s - m)$ . Again using mean-value considerations, we obtain the degree distribution,  $G_0^I(x; s, m)$ , of those that became infectious in the last generation, *i.e.*,

$$mG_0^I(x; s, m) = (N - (s - m))G_0^S(x; s - m) - (N - s)G_0^S(x; s) \quad , \quad (3.28)$$

as

$$G_0^I(x; s, m) = (N - 1) \left[ \frac{G_0(x\theta(s - m)) - G_0(x\theta(s))}{m} \right] \quad . \quad (3.29)$$

The excess degree of the currently infectious nodes is therefore generated by

$$\tilde{G}_g(x; s, m) = \begin{cases} G_0(x) & (g = 0) \\ \frac{G_0^I(x; s, m)}{x} & (g \geq 1) \end{cases} \quad . \quad (3.30)$$

This distribution can then be used in equation (3.12) as a substitute for  $G_g(x)$  when finite-size effects cannot be neglected.

### 3.3.2 Additional loss of transmissions

For networks of finite size, it is no longer possible to completely neglect the effect of closed loops on the dynamics of outbreaks. Indeed, it is possible that some of the neighbours of a newly infected node have previously been infected and, for dynamics where re-infection is impossible, this implies fewer new infections than would have been predicted in an infinite network. Similarly, links between two infectious nodes or links from more than one infectious node to the same susceptible node also reduce the number of new infections.

Furthermore, since the pair  $s'$  and  $m'$  completely characterizes the “state” of the system in an infinite network, it still carries a lot of information about the corresponding state in finite-size networks. We thus make the assumption that  $s'$  and  $m'$  are a sufficient basis to incorporate the finite-size effects, *i.e.*, the loss of transmissions in the finite network that would have occurred in an infinite one.

Our main step is to seek a “mean field” approximation (where every parameter other than  $s'$  and  $m'$  is assumed to take its mean value) to the ratio

$$\rho_{s'm'} = \left\langle \frac{\langle \tilde{m} \rangle}{m} \right\rangle_{s'm'} \quad (3.31)$$

of mean number of transmissions  $\tilde{m}$  that *actually happen* in the *finite* network to the number of transmissions  $m$  that *would have occurred* in an *infinite* one. If each of the  $m$  infinite network transmissions is treated as having the *independent* probability  $1 - \rho_{s'm'}$  of being lost (a probability  $\rho_{s'm'}$  of occurring) in the finite network, there is then a probability

$$\binom{m}{\tilde{m}} (\rho_{s'm'})^{\tilde{m}} (1 - \rho_{s'm'})^{m-\tilde{m}} \quad (3.32)$$

that  $\tilde{m}$  transmissions occur in the finite network. Similarly, a subgroup  $l$  of the  $m$  infinite network transmissions will contribute  $\tilde{l}$  to the  $\tilde{m}$  finite network transmissions with probability

$$\binom{l}{\tilde{l}} (\rho_{s'm'})^{\tilde{l}} (1 - \rho_{s'm'})^{l-\tilde{l}} \quad (3.33)$$

Although these transmissions are not exactly independent events, the independence assumption is a good approximation when the system has many degrees of freedom, *i.e.*, it holds when  $m$  is small compared to the number of susceptibles of nonzero degree. The probability distributions for  $\tilde{m}$  (equation (3.32)) and  $\tilde{l}$  (equation (3.33)) therefore

hold unless most nodes of the network have been infected, quite unlikely for example in realistic epidemiological applications. One of the first effects of a non-negligible correlation would typically be a reduction of the variance of the distributions of  $\tilde{m}$  and  $\tilde{l}$ .

Appendix 3.A provides an expression for  $\rho_{s'm'}$  under assumptions of continuity (a differential equation approach similar to the previous section) for a sizable population, large enough to obtain meaningful mean values. All but one of these mean values are relatively easy to calculate, the remaining one is derived in Appendix 3.B to complete the task.

### 3.3.3 Phase-space representation in finite-size networks

We now combine the degree distribution of infectious nodes

$$\tilde{G}_{g-1}(x; s', m') = \sum_k \tilde{p}_k(s', m') x^k \quad (3.34)$$

obtained in section 3.3.1 with the distribution of  $\tilde{l}$  to obtain the finite-size network counterpart of equation (3.5)

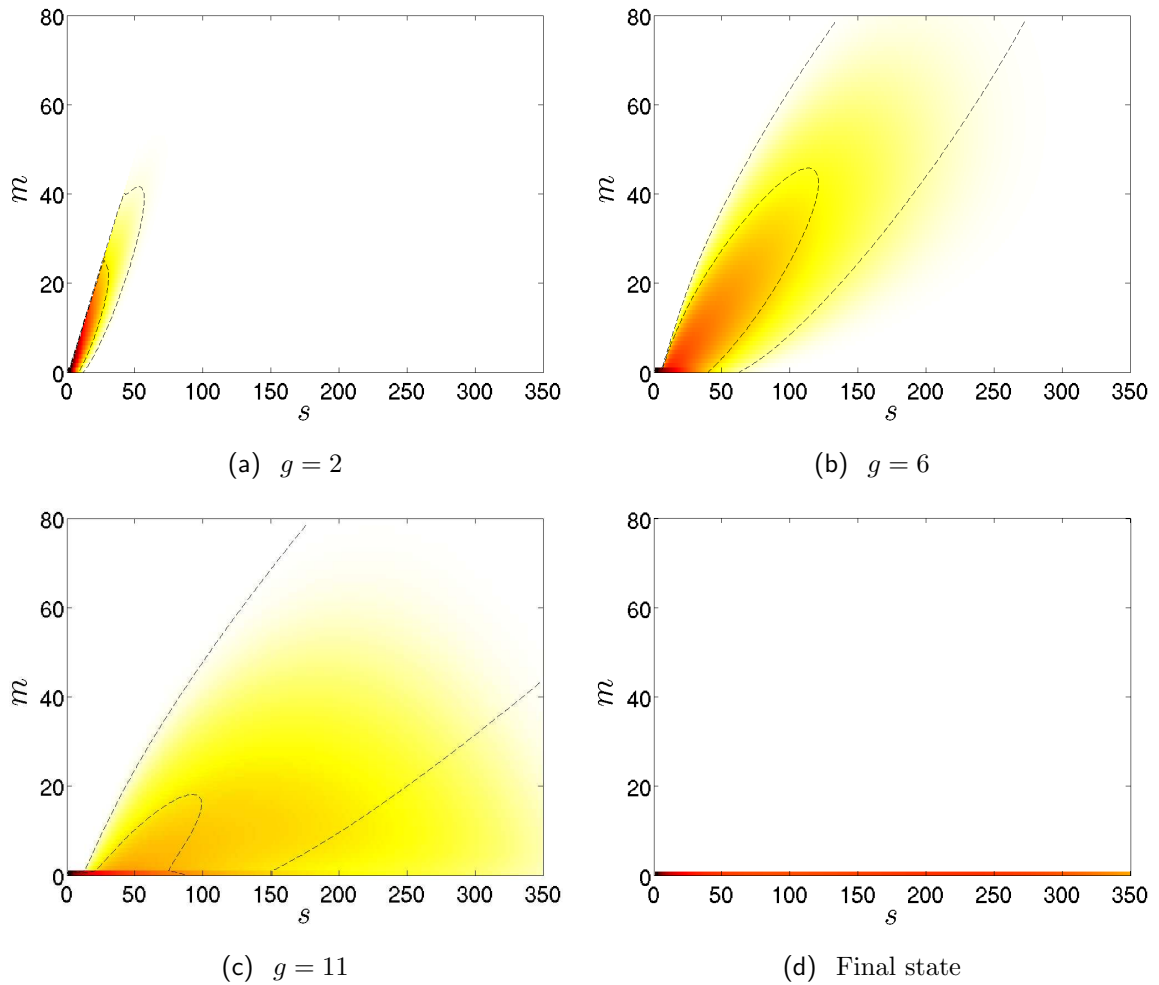
$$\begin{aligned} \sum_{\tilde{l}=0}^{\infty} \sum_{l=\tilde{l}}^{\infty} \sum_{k=l}^{\infty} \tilde{p}_k(s', m') \binom{k}{l} T^l (1-T)^{k-l} \binom{l}{\tilde{l}} (\rho_{s'm'})^{\tilde{l}} (1-\rho_{s'm'})^{l-\tilde{l}} x^{\tilde{l}} \\ = \tilde{G}_{g-1}(1 + (x-1)T\rho_{s'm'}; s', m') \quad . \quad (3.35) \end{aligned}$$

This is now the generating function for the number of new infections caused by a single infectious node in a finite network when the state of the disease is characterized by  $s'$  and  $m'$  (at generation  $g$ ). Defining an effective transmissibility  $\tilde{T}_{s'm'} = T\rho_{s'm'}$ , the forward recurrence relation, equation (3.12), is then replaced by a new expression of the same structure

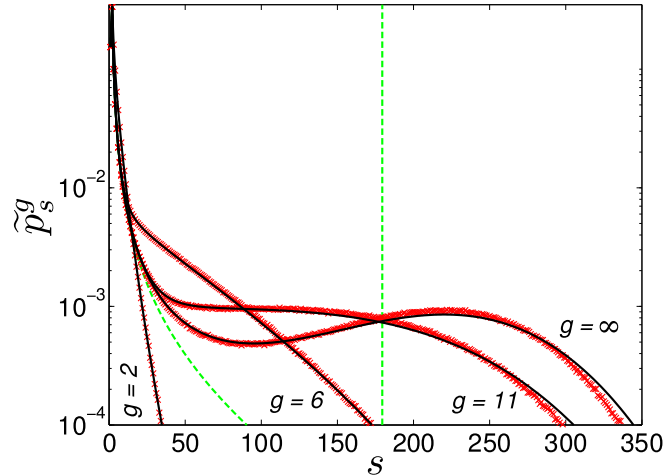
$$\tilde{\Psi}_0^g(x, y) = \sum_{s', m'} \tilde{\psi}_{s'm'}^{g-1} x^{s'} \left[ \tilde{G}_{g-1} \left( 1 + (xy-1)\tilde{T}_{s'm'}; s', m' \right) \right]^{m'} \quad , \quad (3.36)$$

where the evolution of the degree distribution is taken into account by the change  $G_{g-1}(x) \rightarrow \tilde{G}_{g-1}(x; s', m')$  while the additional losses of transmissions are introduced by the replacement  $T \rightarrow \tilde{T}_{s'm'}$ .

Figure 3.4 depicts the phase-space representation for the finite-size power-law network, using the finite-size algorithm of equation (3.36) instead of the infinite size approach of equation (3.12). The system studied is identical with that of figure 3.1 and for



**Figure 3.4:** Representation of  $(s, m)$  phase space for the *finite-size* network algorithm:  $s$  denotes the number of infected nodes by the end of the  $g$ -th generation and  $m$  denotes the number of new infections that occurred since the last generation. The degree distribution, transmissibility and symbols are the same as for figure 3.1 except that the finite-size algorithm (3.36) is used for the calculations. For comparison, the dashed curves are contour plots of the previous results for infinite-size network (figure 3.1, equation (3.12)).



**Figure 3.5:** Projection  $\tilde{p}_s^g$  on the  $s$  axis of the  $(s, m)$  phase space for the *finite-size* network algorithm:  $s$  denotes the number of infected nodes by the end of the  $g$ -th generation and  $\tilde{p}_s^g$  denotes the probability for  $s$  to occur. The degree distribution and transmissibility of figure 3.1 are used and the symbols used are those of figure 3.2. The results produced by the finite-size algorithm are in very good agreement with the numerical simulations ( $10^3$  equivalent graphs with  $10^5$  simulations per graph) over the entire range of possible outbreak/epidemic sizes.

comparison the infinite-size calculations are superimposed on the new results. Figure 3.5 presents the corresponding projections on the  $s$  axis. There is a major improvement in the agreement with the numerical simulations for all generations and the final state. Despite the approximations, and the small size of the network ( $N = 10^3$ ), this agreement makes us confident that we have captured the major part of the finite-size effects.

Furthermore, while the infinite-size formalism produces a single number,  $S$ , for the giant component size (represented by the dashed vertical line in figure 3.5), the finite-size formalism produces the *whole* probability distribution of sizes above the epidemic threshold. More on this in the next section.

### 3.4 Further Results and Discussion

In what follows, we examine in succession the results of the finite versus infinite formalism on and around the percolation threshold, make a comparison with recent dynamical models of disease propagation, establish the relationship with the reproductive number of epidemiology and briefly discuss possible avenues of improvements concerning clustering and continuous time evolution.



### 3.4.1 Behaviour around percolation threshold

We recall first that the current paper has been strongly influenced by the previous works of [70, 75]. Their formalism allows for the calculation of the whole distribution of outbreak/epidemic size and probability at the final state of the dynamics ( $g \rightarrow \infty$  in the language of this paper) for infinite networks. It is based on the *pgfs*  $H_0(x)$  and  $H_1(x)$  related to  $G_0(x)$  and  $G_1(x)$  of equations (3.1) and (3.3) by the relations

$$H_1(x) = xG_1(1 - (1 - H_1(x))T) \quad (3.37)$$

$$H_0(x) = xG_0(1 - (1 - H_1(x))T) \quad (3.38)$$

In particular, the *pgf*  $H_0(x)$  generates the distribution of size for outbreaks from which the average size of outbreak/epidemic for infinite networks can be expressed as

$$\langle s \rangle_\infty = H_0(1)H'_0(1) + N(1 - H_0(1))^2 \quad (3.39)$$

Note also that above threshold the fraction of nodes belonging to the giant component is given by  $S = 1 - H_0(1)$ ; below threshold  $H_0(1) = 1$  as it should.

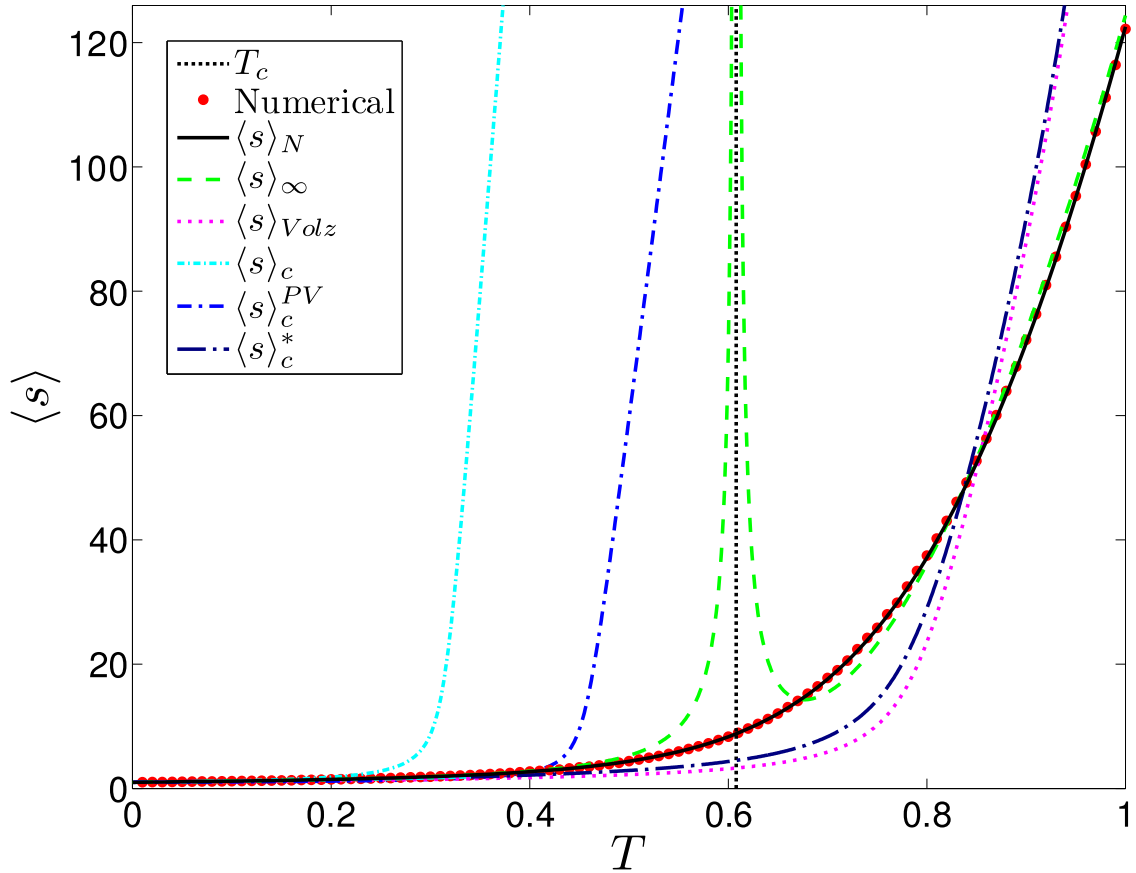
Since section 3.2 is a direct extension of this formalism to discrete time evolution, the values it provides all converge to that of [70] at the final state of the dynamics. Formally, this is no surprise since as  $g \rightarrow \infty$ , the recurrence relationship (3.12) approaches the self-consistency expressions (3.37) and (3.38). In this asymptotic limit,  $H_0(x) = \Psi_0^\infty(x, 1)$ . Similar remarks hold for the work of Marder [56], another extension of [70] to discrete time evolution.

By contrast, the formalism of section 3.3 takes into account finite-size effects. Instead of equation (3.39), the average size of an outbreak/epidemic is now provided by

$$\langle s \rangle_N = \sum_{s,m} s \tilde{\psi}_{sm}^\infty \quad (3.40)$$

since the distribution  $\tilde{\psi}_{sm}^\infty$  is always properly normalized with  $\tilde{p}_s^\infty = \sum_m \tilde{\psi}_{sm}^\infty$  and  $\sum_s \tilde{p}_s^\infty = 1$ .

An important feature of the new formalism is its ability to capture the finite size of an outbreak in the vicinity of the critical transmissibility,  $T_c$ , which separates small outbreak and large-scale epidemic zones. This notion is depicted in figure 3.6 where we consider the same power-law network for different transmissibilities  $T$  ( $T_c \simeq 0.6$ ). In an infinite-size network, we observe a “divergence” in the outbreak/epidemic size,  $\langle s \rangle_\infty$ , just around the transmissibility threshold. However, numerical simulations on finite-size networks (and similarly real-life outbreaks) never exhibit such divergence. Figure 3.6



**Figure 3.6:** Average size of outbreaks and/or epidemics for the power-law network of figure 3.1 ( $N = 1\,000$ ) as a function of transmissibility,  $T$ . The infinite-size formalism provides the position of critical transmissibility,  $T_c \simeq 0.6080$ , above which epidemics can occur (vertical dotted line), as well as the expected size of outbreaks or epidemics  $\langle s \rangle_\infty$  (equation (3.39), dashed curve). In contrast to the divergence at  $T_c$  in the infinite-size formalism, numerical simulations (circles,  $10^3$  equivalent graphs with  $10^3$  simulations per graph for 100 different values of  $T$ ) show a smooth monotonic increase of  $\langle s \rangle$ . The prediction of the new finite-size algorithm ( $\langle s \rangle_N$ , equation (3.40), solid line) is in perfect agreement with numerical results. Four other curves provide the results of the models presented in Sec 3.4.2.

demonstrates that the finite-size formalism can accurately retrieve the size of outbreaks on and around the transmissibility threshold. Away from threshold, infinite and finite methods agree with the simulations, but only the finite-size approach covers smoothly the complete range of transmissibilities without unrealistic divergence.

### 3.4.2 Importance of the variety of outcomes

It would be informative to compare our method with some other propagation approaches on a structured network. The chosen models all perform some “mean field” approximations that prevent them from taking into account the wide variety of possible outcomes, which can span the whole range from small outbreaks to large-scale epidemics. In effect, these models can only provide mean values and the quality of the resulting averages is sometimes difficult to estimate.

For the purpose of comparison, all the models are defined for the same network ensemble as the one presented in section 3.2 and the quantity that we will focus on is the average size of outbreaks/epidemics at the asymptotic final state of the dynamics.

#### **Volz (2008): A network-centric approach**

Close in spirit with the present contribution, but technically very different, is the recent approach derived by Volz [99]. The formalism is based on network-centric quantities (such as edges linking susceptible nodes to infectious nodes or to other susceptible nodes) in order to introduce time evolution and finite-size effects. Although it is also based on the *pgf*  $G_0(x)$  (equation (3.1)), it uses a set of differential equations to track the evolution of mean values. The network-centric quantities are later converted to node-centric quantities, such as the number of infected or susceptible nodes. The asymptotic ( $t \rightarrow \infty$ ) average size of an outbreak/epidemic we are interested in can be cast in the following form

$$\langle s \rangle_{\text{Volz}} = N (1 - G_0(\theta_V(\infty))) \quad (3.41)$$

where  $\theta_V$  and the full set of differential equations are defined in Appendix 3.C.1. It is interesting to note that  $\theta_V(t)$  is conceptually related to its counterpart  $\theta(s)$  (equation (3.20)) of the finite-size algorithm, although they are obtained in a completely different manner. Figure 3.6 displays  $\langle s \rangle_{\text{Volz}}$  as a function of  $T$  and shows a remarkable improvement over the infinite size results. The remaining discrepancy with the simulations and  $\langle s \rangle_N$  can be attributed to the mean field approximation used where only *average* values are evolved in time.

**Moreno *et al.* (2004): A compartmental approach**

A special class of models for the dynamical propagation on networks is the use of compartmental models where nodes of different degree are placed into different compartments [67, 83, 84]. We have integrated the SIR-like model of [67], whose description is detailed in Appendix 3.C.2, to construct the asymptotic average for a population  $N$

$$\langle s \rangle_c = N \sum_k R_k(\infty) \quad (3.42)$$

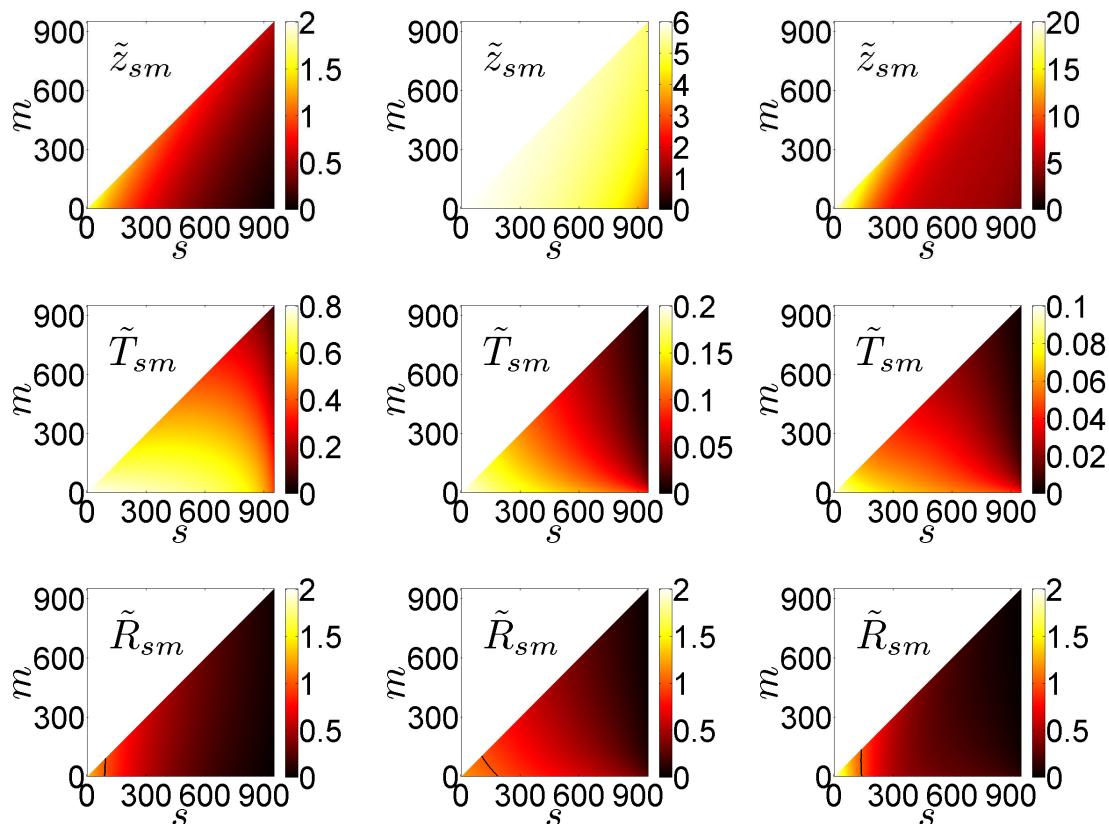
where  $R_k$  is the fraction of nodes that are recovered and of degree  $k$ . A slight correction to this model has been proposed in [86] and Appendix 3.C.2 explains how to obtain the corresponding asymptotic average  $\langle s \rangle_c^{PV}$ . The results seen in figure 3.6 show large deviations of both models long before the critical  $T_c$  for the network and appears valid only for small transmissibility. The source of failure in reproducing the simulations is quite subtle and has led us to propose an improved version.

**An improved compartmental approach**

Although the system (3.65) appears to be a natural choice, it neglects some important points. Firstly, all infectious nodes (except the first one) become infected from one of their neighbours. Since a link between two nodes will always remain attached to these two nodes, the “effective” degree of an infectious node is (at least) one less than its actual degree. We introduce this in the model by transferring nodes from  $S_k$  to  $I_{k-1}$  (instead of  $I_k$ ). Secondly, an infectious node that infects a previously susceptible node along a given link cannot infect again along that same link; we should reduce the degree of infectious nodes, by one, each time they cause new infections. Thirdly, the same procedure must be done for infectious nodes linked to infectious or removed nodes (and applied to these infectious or removed nodes, as well). To put it differently, the knowledge that no infection occurs carries as much information as the knowledge that an infection occurs and both of these scenarios must be taken into account when modelling the dynamics of the system.

We have implemented these improvements in the original differential equations (see Appendix 3.C.3) and the asymptotic result, say  $\langle s \rangle_c^*$ , is plotted in figure 3.6. The agreement is much better and essentially reproduces the result of Volz,  $\langle s \rangle_{\text{Volz}}$ . It seems that our modifications have nicely included the better part of the correct dynamics.

In contrast to the finite-size formalism that takes into account the full stochastic nature of the process, these three systems assume that the “average evolution” occurs



**Figure 3.7:** Effective reproduction number interpretation. For the 3 networks presented in figure 3.3 we show the expected excess degree of the infectious  $\tilde{z}_{sm}$ , the effective transmissibility  $\tilde{T}_{sm}$  and the corresponding effective reproduction number  $\tilde{R}_{sm}$  for each  $(s, m)$  state. The left, center, and right columns display the results of the power-law, the binomial, and the bimodal distributions respectively. The solid black line in the  $\tilde{R}_{sm} = \tilde{T}_{sm}\tilde{z}_{sm}$  plots corresponds to the “threshold” value  $\tilde{R}_{sm} = 1$ .

at each time. Such an approach seems to be more suitable for an endemic disease in a susceptible-infectious-susceptible (SIS) dynamics since the stochasticity is averaged when many nodes are infectious at the same time. When only a small amount of infectious nodes are present at some point in the evolution of the disease (*e.g.*, the very beginning), the stochastic nature of the dynamics can have tremendous impacts on the long-term behaviour of the system. In particular, it is those very stochastic effects that allow for small components even beyond the epidemic threshold. Systems solely tracking averages cannot take into account this important behaviour.

### 3.4.3 Effective reproduction number

In order to establish a link between the present formalism and the classical epidemiological models, it is worth revisiting the interpretation of the basic reproductive

number — a key parameter in classical epidemiology [5, 38]. We can derive the mean excess degree of the infectious nodes as  $\tilde{z}_{sm} = \tilde{G}_{g-1}(1; s, m)$ , and together with the effective transmissibility  $\tilde{T}_{sm}$ , obtain the corresponding *effective reproduction number*  $\tilde{R}_{sm} = \tilde{T}_{sm}\tilde{z}_{sm}$  for each  $(s, m)$  state. Figure 3.7 shows the dependency of  $\tilde{z}_{sm}$ ,  $\tilde{T}_{sm}$ , and  $\tilde{R}_{sm}$  on  $s$  and  $m$  for the networks introduced in figure 3.3. The behaviour of the power law distribution is dominated by the variability of  $\tilde{z}_{sm}$ , as  $\tilde{T}_{sm}$  remains relatively uniform in the vicinity of the epidemic threshold. However, the converse is true for the binomial distribution; it is the variability of  $\tilde{T}_{sm}$  that is responsible for the behaviour of this distribution. The bimodal distribution is seen as a mixture of these two behaviours.

### 3.4.4 Clustering and continuous time evolution

Although most real-world contact networks show a clustered structure, and ways to incorporate this property into the dynamics are rapidly developing [91, 92, 25, 60], our model does not at present explicitly account for clustering. In clustered networks, two neighbors of the same node  $i$  are more likely to be neighbours of one another than of any other random node. In fact these “triangles” (or short loops) are not forbidden in the network ensemble introduced in section 3.2, and over which our formalism is defined, but their number, however, is ruled by randomness alone and is less than what would be expected in a typical realistic human population for instance. Hence, since it has recently been shown that clustering tends to decrease the number of infections in epidemic dynamics [60], our model represents a worst case scenario for networks whose clustering properties are stronger than their random values. One promising way to improve upon the present framework, while keeping the generality of a *pgf*-based formalism and allowing for better treatment of the evolution on clustered networks, would be to generalize the model to a bipartite graph structure where nodes (one part of the network) are assigned to different groups (the other part of the network) much in the same spirit as originally studied in [71].

Another aspect that needs further development is the relationship between discrete (generational) and continuous time evolution. For systems where the underlying epidemiological dynamics is a discrete process in time (*e.g.*, there is a constant time interval  $\tau$  between successive infections), the generational formalism developed in section 3.2 directly represents the evolution of the system over time (infection of the  $g$ -th generation occurs at time  $g\tau$ ). However, in most natural situations, the underlying dynamics is a continuous process. We are actively pursuing the issue of a general continuous epidemiological dynamics; the analysis will be the subject of a forthcoming contribution.

## 3.5 Conclusion

The emergence and re-emergence of infectious diseases pose a great threat to public health. The potential spread of a new pandemic strain of influenza or other emerging infection, such as SARS, may have a devastating impact on human lives and economies. There is an urgent need to develop reliable quantitative tools that can be used to compare the impact of various intervention strategies in real time. These tools must be able to incorporate the detailed structure of contact networks responsible for disease spread, as well as compare various intervention outcomes during the time of crisis, in a relatively short time span. In addition, these tools should be as equally applicable to large-scale networks as to finite-size networks, seeing that many interventions must be implemented not only globally, but locally (*e.g.*, hospital settings, schools) as well.

In this paper, we have introduced and validated a theoretical framework that enables us to incorporate these two important aspects of disease outbreaks/epidemics, simultaneously. Specifically an extension to the existing formalism has been derived while keeping its appealing structure in terms of generating functions. With the introduction of the concept of generations and phase-space representation, equation (3.12) — for the infinite-size network with fixed transmissibility  $T$  — has been replaced by equation (3.36) to account for finite-size effects through a modified generating function and an effective transmissibility. This is our finite-size, discrete time algorithm. One of its important features is its ability to follow the complete diversity of outcomes, not only averages, from small clusters (outbreaks) to the giant component (epidemic) within each generation. A complete formalism including both finite-size and continuous time with or without correlations between transmissions is still lacking and its derivation is part of ongoing research.

### 3.A The ratio $\rho_{s'm'} = \langle \langle \tilde{m} \rangle / m \rangle_{s'm'}$ (equation (3.31))

It is important to take into account that the number  $\tilde{m}$  of new infections that actually occur in a finite network will typically be smaller than the number  $m$  of new infections that would have happened in an infinite network. As presented in section 3.3.2, the  $m$  links leaving infectious nodes and causing new infections in the infinite network can lead to either susceptible nodes, other infectious nodes or recovered nodes. Of these links, only those leading to susceptible nodes can actually result in new infections in the finite network and, moreover, a susceptible node targeted by more than one of these links contributes to only one transmission.

We note by  $n_S$ ,  $n_I$  and  $n_R$  the number of links that are *not forbidden* to join infectious nodes to susceptible nodes, other infectious nodes and recovered nodes respectively. Since there is no special restriction forbidding susceptible nodes to be linked to infectious ones,  $n_S$  is simply the sum of the degrees of all susceptibles. In the same way,  $n_I$  is the sum of the excess degree of the infectious nodes (since the sole restriction is that, for generations other than zero, there is at least one link from each infectious nodes to a recovered one). When the state of the infection is characterized by  $s'$  and  $m'$ , these considerations translate to the mean values

$$\langle n_S \rangle_{s'm'} = (N - s') G_0^{S'}(1; s') \quad (3.43)$$

$$\langle n_I \rangle_{s'm'} = m' \tilde{G}'_{g-1}(1; s', m') \quad (3.44)$$

$$\langle m \rangle_{s'm'} = T \langle n_I \rangle_{s'm'} \quad , \quad (3.45)$$

where  $G_0^S(x; s')$  and  $\tilde{G}_{g-1}(x; s', m')$  are defined in equation (3.22) and equation (3.30) respectively. More complicated constraints apply to  $n_R$  and its mean value  $\langle n_R \rangle_{s'm'}$  is obtained separately in Appendix 3.B.

We develop two methods to evaluate  $\rho_{s'm'} = \langle \langle \tilde{m} \rangle / m \rangle_{s'm'}$ , the ratio of the mean number of new infections occurring in the finite network to the number of new infections that would have occurred in an infinite one. Both methods are based on the assumption that the  $m$  links leading to infections in the infinite limit have the same *a priori* probability of targeting any of the  $n_S$ ,  $n_I$  and  $n_R$  targets, which is justified when  $n_S + n_I + n_R$  is large compared to the degree of the node of highest degree in the network and is typically satisfied unless most of the network has been infected.

The first of these methods is quick and simple. The second method is based on a differential equation approach similar to the one used in section 3.3.1. As both methods bring the same result, this second approach is an additional justification of the differential equation method for cases where an alternative approach is not known, as in Appendix 3.B for instance.

### 3.A.1 Direct approach

Since the  $m$  links causing infections in the infinite network can lead to  $n_S + n_I + n_R$  potential targets in the finite one, each link belonging to a susceptible node (*i.e.*, one of the  $n_S$  targets) has a probability

$$\lambda = \frac{m}{n_S + n_I + n_R} \quad (3.46)$$



to be one of these links. If that node is of degree  $k$ , there is a probability  $1 - (1 - \lambda)^k$  that *at least one* of the  $m$  links leads to it, causing a new infection. Using the degree distribution of the susceptibles obtained in section 3.3.1 and the fact that  $N - s'$  susceptibles are left when the state of the infection is characterized by  $s'$  and  $m'$ , the expected number of new infections in the finite network is then

$$\langle \tilde{m} \rangle = (N - s') \sum_k p_k^S(s') [1 - (1 - \lambda)^k] = (N - s') [1 - G_0^S(1 - \lambda; s')] \quad . \quad (3.47)$$

Therefore, with the use of equations (3.43)–(3.47), the ratio  $\rho_{s'm'} = \langle \langle \tilde{m} \rangle / m \rangle_{s'm'}$  is provided by

$$\rho_{s'm'} = \frac{1}{T} \frac{(N - s')}{m'} \frac{[1 - G_0^S(1 - \lambda_{s'm'}; s')]}{\tilde{G}'_{g-1}(1; s', m')} \quad (3.48)$$

where

$$\lambda_{s'm'} = \frac{\langle m \rangle_{s'm'}}{\langle n_S \rangle_{s'm'} + \langle n_I \rangle_{s'm'} + \langle n_R \rangle_{s'm'}} \quad , \quad (3.49)$$

and  $\langle n_R \rangle_{s'm'}$  is given in Appendix 3.B.

### 3.A.2 Continuous approach

In this approach, we relax the numerous constraints of the problem by changing the task to an “assignment process” evolving in time. By analogy, this process is similar to the system children use for their casual sports team when two leaders choose their team members one by one. It should be clear that the time dependency introduced in this manner is an artificial construct and that only the result at  $t = \infty$  (“assignments of teams is completed”) is relevant.

We define  $\alpha(t)$  as the pool of potential infections that occur in the infinite network and for which a target has yet to be assigned at time  $t$ . This definition implies the initial condition  $\alpha(0) = m$ ; we must also have  $\alpha(\infty) = 0$  (the process must assign a target to each potential infection). We also define  $\beta(t)$ , the pool of targets for which assignment of one of the potential infections does not lead to a new infection, yielding the initial condition  $\beta(0) = n_R + (n_I - m)$ . We finally define  $\gamma_k(t)$ , the pool of targets belonging to susceptible nodes of degree  $k$  and to whom the assignment of a potential infection leads to an actual new infection in the finite network. From mean value considerations,

$$\gamma_k(0) = (N - s') k p_k^S(s') \quad (3.50)$$

with  $\sum_k \gamma_k(0) = n_S$ .

Since the time dependence is arbitrary, we assign targets to the potential infections at a rate proportional to their current population  $\alpha(t)$ , which is a simple way to obtain  $\alpha(\infty) = 0$ . With the definition

$$\omega(t) = \alpha(t) + \beta(t) + \sum_k \gamma_k(t) \quad , \quad (3.51)$$

the probability of assigning a potential infection to other potential infections, other targets not leading to infections, or to targets leading to infections belonging to degree  $k$  is  $\alpha(t)/\omega(t)$ ,  $\beta(t)/\omega(t)$  and  $\gamma_k(t)/\omega(t)$ , respectively. Moreover, upon assignment of a potential infection to a target belonging to  $\gamma_k(t)$ , we transfer the remaining  $k-1$  targets to  $\beta(t)$  to ensure that later assignment to these targets will not lead to new infections. These considerations translate into a set of coupled nonlinear differential equations:

$$\begin{aligned} \frac{d\alpha(t)}{dt} &= -\alpha(t) \left[ 1 + \frac{\alpha(t)}{\omega(t)} \right] \\ \frac{d\beta(t)}{dt} &= -\alpha(t) \frac{\beta(t)}{\omega(t)} + \sum_k \alpha(t) \frac{\gamma_k(t)}{\omega(t)} (k-1) \\ \frac{d\gamma_k(t)}{dt} &= -\alpha(t) \frac{\gamma_k(t)}{\omega(t)} k \end{aligned} \quad (3.52)$$

for all  $k$ . By summing these terms, we note that  $d\omega(t)/dt = -2\alpha(t)$ , together with  $d\alpha(t)/dt$ , yields the solution

$$\frac{\alpha(t)}{\omega(t)} = \frac{\lambda}{\lambda + (1-\lambda)e^t} \quad (3.53)$$

where  $\lambda = \alpha(0)/\omega(0) = n_T/(n_S + n_I + n_R)$ . This result allows us to rewrite the equation governing  $\gamma_k(t)$  as

$$\frac{d\gamma_k(t)}{dt} = \frac{-k\gamma_k(t)}{1 + (\lambda^{-1} - 1)e^t} \quad , \quad (3.54)$$

which is completely decoupled from the rest of the system. This equation has the solution

$$\gamma_k(t) = \gamma_k(0) (\lambda e^{-t} + 1 - \lambda)^k \quad , \quad (3.55)$$

and produces the limit

$$\gamma_k(\infty) = \gamma_k(0) (1 - \lambda)^k \quad . \quad (3.56)$$

The number of nodes of degree  $k$  that are infected in the process is thus

$$\frac{\gamma_k(0) - \gamma_k(\infty)}{k} = \frac{\gamma_k(0)}{k} (1 - (1 - \lambda)^k) = (N - s') p_k^S(s') (1 - (1 - \lambda)^k) \quad . \quad (3.57)$$

Using the initial conditions, we obtain the average total number of infections in the finite network

$$\langle \tilde{m} \rangle = (N - s') \left( \sum_k p_k^S(s') - \sum_k p_k^S(s') (1 - \lambda)^k \right) = (N - s') [1 - G_0^S(1 - \lambda; s')] \quad . \quad (3.58)$$

Notice that this expression is the same as equation (3.47) and thus leads to the same  $\rho_{s'm'}$ , *i.e.*, equation (3.48).

### 3.B The average value $\langle n_R \rangle_{s'm'}$

While it is quite easy to obtain mean values or even distributions for  $n_I$ ,  $n_S$  and  $m$ , an independent method is required to evaluate the quantity  $n_R$ . Here, we use a continuous approach similar to that of Appendix 3.A.

Again, we design a differential equation governing the evolution of the continuous counterpart of  $n_R$ , denoted  $\eta(s)$ , as a function of the number of infections,  $s$ . If the  $k - 1$  excess degrees belonging to a newly infected node, of degree  $k$ , are linked to susceptible nodes, we expect a fraction  $1 - T$  of them to be unsuccessful in transmitting the infection to others and thus, contribute to  $\eta(s)$ . However, we expect a fraction,  $\eta(s)/\langle n_S(s) \rangle$ , of the  $k - 1$  excess degrees to be linking to recovered nodes, with  $\langle n_S(s) \rangle = (N - s)z_1^S(s)$ . These last links actually *reduce* the value of  $\eta(s)$ , because links between susceptible and recovered nodes are converted to links between recovered and infectious nodes. Defining

$$z_2^S(s) = \sum_k k(k - 1)p_k^S(s) = G_0^{S''}(1; s) \quad , \quad (3.59)$$

the differential equation considering these effects reads

$$\begin{aligned} \frac{d\eta(s)}{ds} &= \sum_k \frac{k p_k^S(s)}{z_1^S(s)} (k - 1) \left[ \left( 1 - \frac{\eta(s)}{\langle n_S(s) \rangle} \right) (1 - T) - \frac{\eta(s)}{\langle n_S(s) \rangle} \right] \\ &= \frac{z_2^S(s)}{z_1^S(s)} \left[ (1 - T) - (2 - T) \frac{\eta(s)}{\langle n_S(s) \rangle} \right] \quad . \end{aligned} \quad (3.60)$$

Together with the initial condition  $\eta(1) = (1 - T)z_1^S(1) = (1 - T)z_1$ , this differential equation can be integrated numerically to provide an expectation value for the number of links between recovered and susceptible nodes when the size of the outbreak/epidemic is  $s$ .

The quantity  $\langle n_R \rangle_{s'm'}$ , required to fix  $\lambda_{s'm'}$  in Appendix 3.A, is the mean number of links emerging from recovered nodes that are allowed to target infectious nodes when

we know the current state of the epidemic is characterized by  $s'$  and  $m'$ . As  $m'$  is the current number of infectious nodes, the size of the outbreak/epidemic was  $s' - m'$  when these infectious nodes were still susceptible nodes (in the previous generation). At that time, we expect  $\eta(s' - m')$  links to join susceptible and recovered nodes; this number is equal to the expectation value  $\langle n_R \rangle_{s'm'}$ .

### 3.C Some dynamical models

The following models are all defined for the same network ensemble as the one presented in section 3.2, *i.e.*, their degree distribution  $\{p_k\}$  is generated by  $G_0(x)$ .

#### 3.C.1 Volz (2008): A network-centric approach

At time  $t$ , let  $\theta_V(t)$  be the fraction of degree one nodes that remain susceptible,  $p_I(t)$ , the probability that a susceptible node be connected to an infectious one and  $p_S(t)$ , the probability that a susceptible node be connected to a susceptible one, then

$$\begin{aligned} \frac{d\theta_V(t)}{dt} &= -rp_I(t)\theta_V(t) \\ \frac{dp_I(t)}{dt} &= rp_S(t)p_I(t)\theta_V(t)\frac{G_0''(\theta_V(t))}{G_0'(\theta_V(t))} \\ &\quad - rp_I(t)(1 - p_I(t)) - \mu p_I(t) \\ \frac{dp_S(t)}{dt} &= rp_S(t)p_I(t)\left(1 - \theta_V(t)\frac{G_0''(\theta_V(t))}{G_0'(\theta_V(t))}\right) \end{aligned} \quad , \quad (3.61)$$

where  $r$  is the force of infection, the constant rate at which infectious nodes infect a neighbour and  $\mu$  the recovery rate at which infected nodes become recovered. For the purpose of calculations, we have chosen  $\mu = 1$  and a related transmissibility  $T = 1 - e^{-r}$  [70]. The initial conditions have been set to  $\theta_V(0) = 1 - \epsilon$ ,  $p_I(0) = \epsilon/(1 - \epsilon)$  and  $p_S(0) = (1 - 2\epsilon)/(1 - \epsilon)$  for  $\epsilon$  satisfying  $G_0(1 - \epsilon) = 1 - 1/N$ . The asymptotic ( $t \rightarrow \infty$ ) average size of outbreak and/or epidemic is then given by

$$\langle s \rangle_{\text{Volz}} = N (1 - G_0(\theta_V(\infty))) \quad . \quad (3.62)$$

### 3.C.2 Moreno *et al.* (2004): A compartmental approach

At time  $t$ , let  $S_k(t)$  be the fraction of nodes that are susceptible ( $I_k(t)$  and  $R_k(t)$  for the infectious and the removed nodes respectively) *and* of degree  $k$  subject to the normalization

$$\sum_k (S_k(t) + I_k(t) + R_k(t)) = 1 \quad . \quad (3.63)$$

With the probability that any given link pointing to an infected node is

$$\Theta(t) = \frac{\sum_k k I_k(t)}{\sum_{k'} k' (S_{k'}(t) + I_{k'}(t) + R_{k'}(t))} \quad , \quad (3.64)$$

and  $r$  defined as in section 3.C.1, the dynamics obeys the differential equations

$$\begin{aligned} \frac{dS_k(t)}{dt} &= -r \Theta(t) k S_k(t) \\ \frac{dI_k(t)}{dt} &= -I_k(t) + r \Theta(t) k S_k(t) \quad . \\ \frac{dR_k(t)}{dt} &= I_k(t) \end{aligned} \quad (3.65)$$

Solving this system for the initial conditions  $S_k(0) = (1 - 1/N)p_k$ ,  $I_k(0) = p_k/N$  and  $R_k(0) = 0$  leads to the asymptotic average size of outbreak and/or epidemic

$$\langle s \rangle_c = N \sum_k R_k(\infty) \quad . \quad (3.66)$$

Note that a slight correction to this model has been proposed in [86] in order to take into account that infectious nodes acquired their infection along one of their links and that new transmission is impossible along that link. The corresponding asymptotic average size of outbreak and/or epidemic  $\langle s \rangle_c^{PV}$  is obtained in the same way as for the compartmental approach except that the factor  $k$  in equation (3.64) has to be replaced by  $k - 1$ .

### 3.C.3 An improved compartmental approach

Incorporating in equation (3.65) the improvements described in section 3.4 results in a new differential system

$$\begin{aligned}
 \frac{dS_k^*(t)}{dt} &= -r \Theta^*(t) k S_k^*(t) \\
 \frac{dI_k^*(t)}{dt} &= -I_k^*(t) + r \Theta^*(t) (k+1) S_{k+1}^*(t) \\
 &\quad + r (\Theta^*(t) + 1) \left[ (k+1) I_{k+1}^*(t) - k I_k^*(t) \right] \\
 \frac{dR_k^*(t)}{dt} &= I_k^*(t) + r \Theta^*(t) \left[ (k+1) R_{k+1}^*(t) - k R_k^*(t) \right]
 \end{aligned} \tag{3.67}$$

$\Theta^*(t)$  is the corresponding expression for  $\Theta(t)$  (equation 3.64) and the integration is carried out under the same initial conditions. The asymptotic average size of outbreak and/or epidemic is then also

$$\langle s \rangle_c^* = N \sum_k R_k^*(\infty) \tag{3.68}$$

Both infinite differential systems (3.65) and (3.67) are truncated at  $k_{\max} = 50$  in numerical simulations.

# Chapitre 4

## Temps continu et autres considérations

On s'intéresse à la correspondance entre le processus de branchement présenté au chapitre 3 et un processus de propagation en temps continu. La section 4.1 évalue l'influence de la corrélation des périodes infectieuses, précédemment mentionnée en section 2.2; la section 4.2 réexprime le temps discret (génération) utilisé dans le chapitre 3 en un temps continu; et la section 4.3 explique pourquoi une perspective différente est préférée dans la partie II de cette thèse. Le matériel présenté dans les sections 4.1 et 4.2 provient de [78], une version préliminaire du chapitre 3.

### 4.1 Corrélations dues à la période infectieuse

On poursuit ici la discussion débutée en section 2.2 concernant le fait que la percolation de liens n'est pas tout à fait équivalente à un processus *SIR* sur réseau [49]. On rappelle que les probabilités d'infections de chacun des voisins d'un nœud infectieux  $i$  dépendent toutes du même temps  $\tau_i$  pendant lequel  $i$  demeure infectieux, entraînant des corrélations qui sont complètement négligées dans le processus de percolation décrit au chapitre 3. L'importance des effets de ces corrélations — autrement dit, à quel point la percolation est un bon outil pour étudier la propagation de type *SIR* — dépend principalement de la distribution  $f(\tau_i)$  de laquelle on tire la période infectieuse  $\tau_i$ .

Dans le cas extrême

$$f(\tau_i) = \delta(\tau_i - \tau) \tag{4.1}$$

(durée déterminée  $\tau$  pour toute les périodes infectieuses), la correspondance entre la propagation et la percolation est exacte. En effet, puisque la période d'infection est toujours  $\tau$ , il n'est pas important de savoir si deux nœuds susceptibles sont voisins d'un même nœud infectieux ou non. De façon plus générale, si le choix du  $\tau_i$  dans la distribution  $f(\tau_i)$  n'entraîne que de faibles fluctuations pour  $T_{ij}$ , alors on s'attend à ce qu'un processus de percolation soit une bonne approximation du processus de propagation correspondant. Par exemple, ce devrait être le cas pour une distribution  $f(\tau_i)$  gaussienne de faible écart-type par rapport à sa moyenne, soit une distribution s'apparentant grandement à (4.1).

À l'opposé, le cas présenté en exemple dans [49]

$$f(\tau_i) = p\delta(\tau_i - \tau) + (1 - p)\delta(\tau_i) \quad (4.2)$$

(probabilité  $p$  que la période infectieuse soit de durée  $\tau$ , probabilité  $1 - p$  qu'elle soit de durée nulle) entraîne des déviations très importantes entre les deux processus. En supposant que tous les nœuds soient initialement susceptibles, à l'exception d'un nœud infectieux de degré  $k$ , il y a une probabilité  $1 - p$  que ce seul infectieux soit immédiatement retiré. Dans une telle éventualité, aucun des  $k$  voisins susceptibles ne peut acquérir l'infection et le processus de propagation est terminé, faute d'infectieux. Si  $\tau_i$  avait été tiré de façon indépendante pour chacun des  $k$  voisins susceptibles, situation correspondant au processus de percolation de liens utilisant une transmissibilité unique, la probabilité que chacun d'entre eux donne une période infectieuse nulle aurait été  $(1 - p)^k$ . La corrélation a donc ici un effet considérable sur l'issue du processus lorsque  $k > 1$ .

Ainsi, s'il existe des cas où il est justifiable de considérer un processus de propagation comme un processus de percolation où  $T = \langle T_{ij} \rangle$ , ce n'est pas toujours le cas. Cependant, de façon semblable à ce qui est effectué dans [49], il est possible d'adapter le processus de branchement à l'endroit même où la transmissibilité est considérée, c'est-à-dire en remplaçant l'équation (3.4) par

$$p_{l|k} = \int_0^\infty f(\tau_i) \binom{k}{l} (T_{\tau_i})^l (1 - T_{\tau_i})^{k-l} d\tau_i \quad (4.3)$$

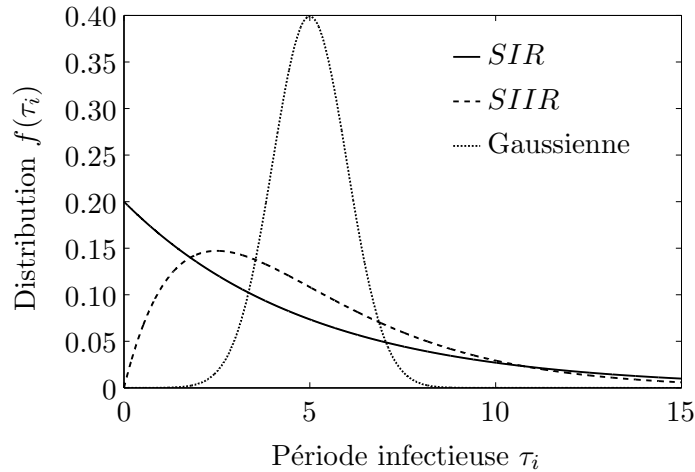
où on a défini la *transmissibilité sachant la durée de la période infectieuse*

$$T_{\tau_i} = 1 - \int_0^\infty g(\beta_{ij}) e^{-\beta_{ij}\tau_i} d\beta_{ij} \quad . \quad (4.4)$$

Ainsi, l'équation (3.12) devient simplement

$$\Psi_0^g(x, y) = \sum_{s', m'} \psi_{s'm'}^{g-1} x^{s'} \left[ \int_0^\infty f(\tau_i) G_{g-1}(1 + [xy - 1]T_{\tau_i}) d\tau_i \right]^{m'} \quad . \quad (4.5)$$



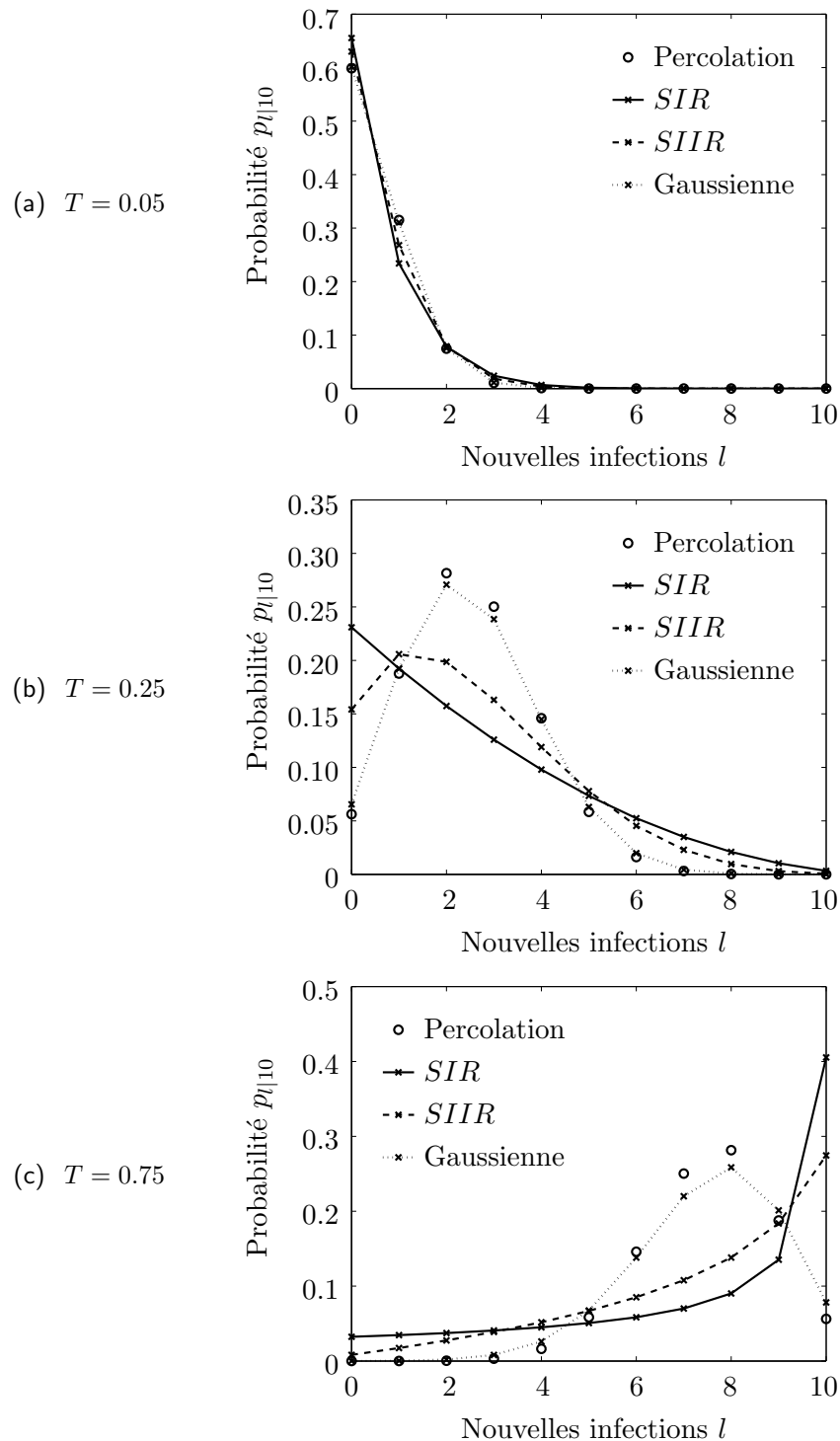


**Figure 4.1** – Distributions de périodes infectieuses utilisées dans les figures 4.2 et 4.3. Les paramètres de ces trois distributions sont choisis pour une période infectieuse de durée moyenne  $\langle \tau_i \rangle = 5$ . Trait plein :  $f(\tau_i)$  correspondant à un processus *SIR* standard. Trait hachuré :  $f(\tau_i)$  pour un processus *SIIR* (voir section 1.3.2). Trait pointillé :  $f(\tau_i)$  posé comme une distribution gaussienne d'écart-type 1 (et de moyenne 5).

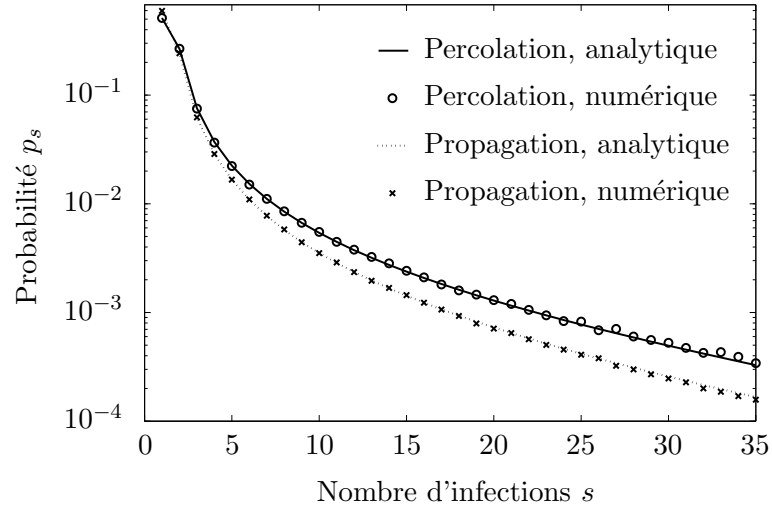
On souhaite maintenant étudier l'impact de la corrélation pour d'autres  $f(\tau_i)$  que les cas extrêmes donnés en équations (4.1) et (4.2) ; la figure 4.1 présente les cas choisis. Utiliser un taux de retrait  $\alpha$  constant, comme c'est le cas pour un processus *SIR* standard, donne la distribution exponentielle  $f(\tau_i) = \alpha e^{-\alpha\tau_i}$  (trait plein). Cependant, le processus *SIR* n'est qu'un modèle — souvent choisi pour sa simplicité — et les distributions réelles sont normalement plus concentrées autour de leurs moyennes. Ainsi, on s'intéresse également à la distribution de périodes infectieuses correspondant à un modèle *SIIR* (trait hachuré, voir section 1.3.2) et à une distribution gaussienne d'écart-type 1 (trait pointillé). Pour fins de comparaison, les paramètres de ces trois distributions sont choisis pour une période infectieuse de durée moyenne  $\langle \tau_i \rangle = 5$ .

La figure 4.2 montre la distribution  $p_{l|k}$  telle que produite par l'équation (4.3) pour les  $f(\tau_i)$  donnés en figure 4.1. La situation considérée correspond à un nœud infectieux possédant  $k = 10$  voisins susceptibles, une valeur suffisamment élevée pour amplifier l'effet de la corrélation (il n'y a aucun effet pour  $k = 1$ ). Les taux d'infection  $\beta$  sont choisis de façon à obtenir les transmissibilités équivalentes  $T = 0.05$ ,  $T = 0.25$  et  $T = 0.75$ . Pour fins de comparaison, les cercles présentent les résultats obtenus utilisant (3.4), soit en négligeant la corrélation (comme pour la percolation).

De façon générale, l'effet de la corrélation devient négligeable pour de très faibles transmissibilités. Pour de plus grandes valeurs, la distribution  $p_{l|k}$  obtenue pour la propagation est typiquement plus étalée que celle de la percolation : cet effet d'étalement est d'autant plus grand pour les distributions  $f(\tau_i)$  elles même étalées et/ou lorsque de



**Figure 4.2** – Distribution du nombre de nouvelles infections causées par un nœud infectieux possédant  $k = 10$  voisins susceptibles. Les résultats sont produit à l'aide de l'équation (4.3) pour les distributions présentées en figure 4.1 ; les taux d'infection  $\beta$  sont choisis de façon à obtenir les transmissibilités équivalentes (a)  $T = 0.05$ , (b)  $T = 0.25$ , et (c)  $T = 0.75$ . Pour fins de comparaison, les cercles présentent les résultats obtenus utilisant (3.4), soit en négligeant la corrélation (comme pour la percolation).



**Figure 4.3** – Distribution du nombre d'infections  $s$  pour un modèle  $SIR$ , soit  $f(\tau_i) = \alpha e^{-\alpha\tau_i}$ . Les résultats analytiques (courbes) sont comparés à des simulations numériques (symboles) pour l'état final d'un modèle de configuration en loi de puissance (taille  $N = 10^3$ ; distribution en degrés  $p_k \propto k^{-2}e^{-k/5}$ ; transmissibilité  $T = 0.4$ ). Le trait plein et les cercles (resp. le trait pointillé et les croix) correspondent à un processus de percolation (resp. de propagation) négligeant (resp. considérant) les corrélations entre les périodes infectieuses.

très faibles valeurs de  $\tau_i$  sont probables. À l'opposé, l'effet de la corrélation est minimal pour des distributions bien regroupées autour de leurs moyennes : on observe très peu de déviations entre la percolation et la propagation pour la distribution gaussienne d'écart-type 1 (trait pointillé). Ainsi, il n'est pas nécessaire d'atteindre le cas extrême (4.1) pour que les effets de la corrélation deviennent négligeables.

La figure 4.3 présente l'effet net de la corrélation sur l'état final du processus de propagation  $SIR$ , soit en utilisant  $f(\tau_i) = \alpha e^{-\alpha\tau_i}$ . On remarque que la distinction entre le processus de propagation (trait pointillé et croix) et de percolation (trait plein et cercles) est de beaucoup moins prononcée qu'en figure 4.2 : les contributions individuelles des infectieux « se contrebalancent » partiellement, amenuisant donc l'effet de la corrélation. De plus, les simulations numériques (symboles) confirment les prédictions analytiques (traits) : lorsque nécessaire, il est donc possible de tenir compte de la corrélation.

## 4.2 Temps continu

Le formalisme présenté au chapitre 3 permet d'étudier l'évolution temporelle de la propagation par rapport à un temps discret : la génération d'infection  $g$ . Cette section cherche à établir un parallèle entre ce temps discret et le temps continu « réel »  $t$ . Il est à noter que toutes considérations concernant la corrélation des périodes infectieuses sont ici négligées.

La méthode choisie consiste à obtenir une *transmissibilité efficace*  $T_g(t)$  pour chaque génération  $g$  au temps  $t$ . Étant donné un individu infectieux de génération  $g$  et l'un de ses voisins initialement susceptible,  $T_g(t)$  correspond à la probabilité que ce voisin ait été infecté par l'individu infectieux au temps  $t$ . En comparant cette définition à celle de la transmissibilité, on remarque que  $T_g(\infty) = T$ . De plus, une simple altération de la borne supérieure d'intégration de  $\tau_i$  dans l'équation (2.2) permet d'obtenir  $T_0(t)$

$$1 - \int_0^t \int_0^\infty f(\tau_i) g(\beta_{ij}) e^{-\beta_{ij}\tau_i} d\beta_{ij} d\tau_i \quad (4.6)$$

en considérant que le nœud de génération  $g = 0$  devient infectieux au temps  $t = 0$ .

Cependant, puisqu'on s'intéresse ici à l'évolution dans le temps continu  $t$ , on considère une situation légèrement plus générale où une *période latente* de durée  $\tau'_i$  est nécessaire avant qu'un élément ayant nouvellement acquis l'infection puisse la transmettre à son tour<sup>1</sup>. Si la probabilité que cette période latente se situe dans l'intervalle  $[\tau'_i, \tau'_i + d\tau'_i)$  est  $h(\tau'_i)d\tau'_i$ , alors l'expression précédente devient

$$T_0(t) = 1 - \int_0^t \int_0^{t-\tau'_i} \int_0^\infty f(\tau_i) g(\beta_{ij}) h(\tau'_i) e^{-\beta_{ij}\tau_i} d\beta_{ij} d\tau_i d\tau'_i \quad (4.7)$$

Il est maintenant utile de définir  $\phi_g(t)$  tel que

$$\phi_g(t) = \frac{1}{T} \frac{d}{dt} T_g(t) \quad \text{et} \quad T_g(t) = T \int_0^t \phi_g(t) dt \quad , \quad (4.8)$$

la seconde relation découlant de la première (puisque  $T_g(0) = 0$ ). On peut interpréter  $\phi_g(t)dt$  comme la probabilité qu'un nœud ait été infecté dans l'intervalle de temps  $[t, t + dt)$  par son voisin infectieux (de génération  $g$ ) sachant qu'il a éventuellement été

---

1. Les modèles compartimentaux reproduisent normalement une telle période latente en ajoutant un ou plusieurs compartiments *exposés* : un élément exposé a acquis l'infection mais ne peut pas encore la transmettre. On appelle processus *SEIR* un processus permettant les états  $\{S, E, I, R\}$ ; les généralisations *SE<sup>m</sup>IR* et *SE<sup>m</sup>I<sup>n</sup>R* considèrent plusieurs compartiments exposés.

infecté. Autrement dit, en sélectionnant au hasard un nœud de génération  $g + 1$ , la probabilité que ce nœud ait été infecté pendant l'intervalle  $[t, t + dt)$  est  $\phi_g(t)dt$ .

Puisque les distributions  $f(\tau_i)$ ,  $g(\beta_{ij})$  et  $h(\tau'_i)$  sont identiques pour tous les nœuds et stationnaires dans le temps, le délai entre l'infection d'un nœud et le moment où il transmet l'infection à l'un de ses voisins est une quantité distribuée de façon identique pour tous les nœuds. En fait, puisque le nœud de génération 0 acquiert l'infection au temps  $t = 0$ , la distribution de ce délai est donnée par  $\phi_0(t)$ . Ainsi, on obtient la récurrence

$$\phi_g(t) = \int_0^\infty \phi_{g-1}(t')\phi_0(t-t')dt' \quad (4.9)$$

qui, grâce au fait que  $\phi_g(t) = 0$  pour tout  $t < 0$ , peut s'écrire sous la forme d'une convolution

$$\phi_g(t) = \int_{-\infty}^\infty \phi_{g-1}(t')\phi_0(t-t')dt' = (\phi_{g-1} * \phi_0)(t) \quad (4.10)$$

En appliquant la récurrence  $g - 1$  fois, on obtient

$$\phi_g(t) = \underbrace{(\phi_0 * \phi_0 * \dots * \phi_0)}_{g+1 \text{ termes}}(t) \quad (4.11)$$

qui permet donc d'obtenir tous les  $T_g(t)$  par la relation (4.8).

Si la moyenne de  $\phi_0(t)$  est  $\mu$  et que sa variance est  $\sigma^2$ , les propriétés de la convolution impliquent que la moyenne de  $\phi_g(t)$  est  $(g + 1)\mu$  et que sa variance est  $(g + 1)\sigma^2$ . De plus, par le théorème central limite, les distributions  $\phi_g(t)$  pour de grandes valeurs de  $g$  sont approximativement gaussiennes.

Enfin, le nombre d'infections ayant eu lieu au temps  $t$  est généré par  $\Psi_0^\infty(x, 1)$  obtenu à l'aide de la récurrence

$$\Psi_0^g(x, y) = \sum_{s', m'} \psi_{s'm'}^{g-1} x^{s'} \left[ G_{g-1}(1 + [xy - 1]T_{g-1}(t)) \right]^{m'} \quad (4.12)$$

La qualité de la méthode dépend largement de  $\phi_0(t)$  : de bons résultats sont obtenus pour des distributions dont l'écart-type est faible par rapport à leur moyenne. Pour des distributions plus larges, le chevauchement des générations entraîne des effets de corrélation entre le temps d'infection d'une génération et de la suivante. Dans certains cas, les résultats s'apparentent plus à une interpolation entre les valeurs de chaque génération qu'à une véritable modélisation en temps continu.

Le texte original [78] discute également du cas non stationnaire ainsi que de l'effet de la corrélation des périodes infectieuses sur une telle projection des générations vers

un temps continu. Cependant, puisque ces traitements découlent assez directement des points traités dans cette section et que leurs applications sont limitées, il n'a pas été jugé pertinent de les reproduire ici.

### 4.3 Nouvelle perspective

Avec le recul, l'approche utilisée dans cette partie I est un peu maladroite. On débute avec un processus discret modélisant la situation pour un réseau de taille infinie. À l'aide d'une approximation continue (permettant d'obtenir  $\rho_{s'm'}$  et  $\tilde{G}_g(x; s', m')$ ), on raccommode le processus original pour lui permettre de tenir compte de la taille finie du système. Plus tard, on tente de projeter le temps discret vers son homologue continu. Pourquoi alors ne pas travailler en temps continu dès le début ?

Quoiqu'elle soit intéressante d'un point de vue théorique, la chimère qui en résulte est peu pratique du point de vue des applications. Par exemple, il est impossible d'obtenir des quantités équivalentes à  $\rho_{s'm'}$  et  $\tilde{G}_g(x; s', m')$  pour le cas non stationnaire puisqu'elles devraient alors dépendre du *parcours* dans l'espace  $(s, m)$  plutôt que de la position actuelle  $(s', m')$ . Par exemple, si l'on souhaitait considérer le scénario d'intervention « à la génération 5, la transmissibilité passe de  $T$  à  $T'$  suite à une campagne de sensibilisation », il faudrait non seulement connaître l'état  $(s', m')$  pour la génération actuelle  $g > 5$ , mais également savoir ce qu'était cet état à la génération 5. Des problèmes semblables surviennent lorsqu'on tente de considérer des structures de réseaux plus complexes ou des processus dynamiques différents. Enfin, les ressources informatiques demandées croissent rapidement avec la taille du réseau.

Malgré tout, quelques bons points ressortent de l'exercice. Ainsi, tout le chapitre 3 repose sur le fait que l'on représente un nombre incroyablement élevé de degrés de libertés à l'aide de deux quantités : une « position »  $s$  et une « impulsion »  $m$ . Malgré le fait que la représentation ne soit que partielle, des améliorations notables sont rendues possibles par cette simple procédure. Existe-t-il d'autres bases simples permettant une bonne représentation de la situation ?

Une réponse partielle est donnée en section 3.C.3 : en suivant le nombre de nœuds de degré donnés pour chaque état, on parvient à représenter la dynamique de façon fort acceptable. En fait, le chapitre 6 est simplement une version stochastique du modèle présenté en 3.C.3. L'idée de base derrière la partie II est donc la suivante : chercher une base minimale permettant de représenter convenablement l'état du système et considérer l'évolution temporelle (en temps continu) de cet état en terme d'un processus stochastique ne dépendant — en première approximation — que de cet état.

## Deuxième partie

### Processus de Markov et représentation partielle





# Chapitre 5

## Préparation aux processus de Markov

Ce chapitre introduit la partie II de cette thèse : on cherche à représenter de façon partielle l'état d'un système afin de définir un processus stochastique markovien reproduisant approximativement le comportement du système original. Afin de limiter la redondance, on ne donne que les grandes lignes ; plus de détails sont donnés dans les chapitres suivants.

La section 5.1 définit ce qu'est un processus de Markov puis décrit rapidement l'idée fondamentale à la base de la partie II. La section 5.2 introduit le chapitre 6, une application particulière de cette méthode, alors que la section 5.3 introduit le chapitre 7, ayant une perspective beaucoup plus générale du problème. Une discussion complémentaire est donnée au chapitre 8.

### 5.1 Modéliser un processus de Markov à l'aide d'un processus de Markov plus simple

Un *processus de Markov* est un processus stochastique qui, sauf pour son état actuel, n'a aucune mémoire de ses états antérieurs. De façon plus formelle, la distribution conditionnelle de probabilité des états futurs du système étant donné son état actuel ainsi que tous ses états antérieurs est identique à la distribution obtenue lorsque la probabilité n'est conditionnelle qu'à l'état actuel du système. On appelle ce critère *propriété de Markov* et on dit d'un processus le satisfaisant qu'il est *markovien*.

Même si deux processus présentent un comportement dynamique identique, il est possible que l'un d'entre eux soit markovien alors que l'autre ne le soit pas, dépendent de la façon dont leur « état » est défini. Afin d'illustrer ce fait, on considère l'exemple d'un outil risquant de défaillir à chaque fois qu'il est utilisé. Pour un processus dont l'état en un temps donné ne peut être que « fonctionnel » ou « défaillant », la propriété de Markov n'est pas respectée si, suite à une utilisation, la probabilité qu'un outil fonctionnel devienne défaillant dépend de son nombre d'utilisations antérieures<sup>1</sup>. Cependant, le même système peut être considéré comme markovien lorsque les états permis prennent plutôt la forme « défaillant », « fonctionnel et jamais utilisé », « fonctionnel après 1 utilisation », « fonctionnel après 2 utilisations », *etc.* De façon similaire, un dispositif électronique impliquant une longue ligne de transmission peut être qualifié de non markovien lorsque l'on considère cette ligne de transmission comme introduisant un délai dans le signal la traversant, ou encore de markovien lorsque l'on considère plutôt le champ électromagnétique tout au long de la ligne comme faisant partie de l'état du système.

Tel que discuté en section 1.3.4, on considère une idéalisation de la réalité pouvant être exprimée sous la forme d'une simulation informatique et l'on souhaite obtenir des résultats analytiques à son sujet. Cette idéalisation numérique peut toujours être perçue comme un processus de Markov : dans le pire des cas, l'état du système correspond au contenu total de la mémoire de l'ordinateur. L'approche choisie en partie II consiste à reproduire de façon approximative le comportement de ce processus markovien à l'aide d'un second processus markovien<sup>2</sup> dont l'état ne représente que de façon partielle l'état du premier. On peut ensuite utiliser ce second processus afin d'obtenir des résultats analytiques intéressants, à condition bien entendu que la correspondance entre ces modèles soit suffisante et que le second processus soit suffisamment simple.

Il est parfois possible d'obtenir des résultats exacts à partir d'un processus beaucoup plus simple que l'original ; le chapitre 6 s'intéresse à une sous-classe de problèmes pour laquelle on démontre que c'est le cas. De façon plus générale, il est nécessaire d'introduire certaines approximations afin d'obtenir un modèle suffisamment simple. Dans de tels cas, le chapitre 7 explique comment systématiser la procédure et étudie certains facteurs influençant la qualité des approximations.

---

1. Le système serait markovien si la probabilité de défaillance était indépendante du nombre d'utilisations antérieures.

2. Il s'agit là d'un choix de modélisation de notre part : le fait que le premier processus soit markovien ne garantit pas *a priori* qu'un processus reproduisant son comportement respectera la propriété de Markov.

## 5.2 Prélude à « Propagation on networks : an exact alternative perspective »

Pour une dynamique *SIR* sur un modèle de configuration gelé, la section 3.C.3 présente un système d'équations différentielles déterminant, pour chaque valeur de « degré efficace<sup>3</sup> » possible, le nombre moyen de nœuds susceptibles, infectieux et retirés se trouvant dans le réseau. Ce modèle a l'avantage de considérer la taille finie du système en plus d'offrir un traitement naturel de l'évolution continue du temps. Cependant, puisque seules des valeurs moyennes sont considérées, on néglige complètement tout effet que pourrait avoir la stochasticité du système et on ne peut pas juger de la variabilité des issues possibles.

S'inspirant de cette approche, le chapitre 6 considère un processus stochastique markovien représentant *exactement* le processus complet à reproduire. Plutôt que de suivre le nombre moyen de nœuds pour chaque « degré efficace », on s'intéresse à la distribution de probabilité pour ces nombres de nœuds. On profite ensuite de la simplicité du processus afin d'obtenir des approximations analytiques valides pour des réseaux suffisamment grands. Malgré que le cas *SI* soit exploré en plus de détails, l'approche s'applique à toute une classe de processus, incluant le cas *SIR*.

Deux méthodes complémentaires sont utilisées afin d'obtenir des solutions approximatives simples à ce processus stochastique exact. La première de ces méthodes utilise l'*approximation de van Kampen* afin de transformer le processus de naissance-mort (« *birth-death* »), pour lequel les états prennent la forme d'un vecteur de nombres entiers, sous la forme d'un processus d'*Ornstein-Uhlenbeck*, pour lequel les états prennent la forme d'un vecteur de nombres réels. Étant donnée une condition initiale déterministe, la solution prend la forme d'une distribution gaussienne et reproduit particulièrement bien la distribution de probabilité en cas d'épidémies (composantes géantes). Il est à noter qu'il s'agit là de l'application d'un outil standard à un problème particulier : on ne s'attarde donc pas trop aux détails de la méthode (comme par exemple l'approximation de van Kampen) mais on place plutôt l'emphase sur les nouvelles contributions.

La seconde méthode consiste à linéariser le processus au voisinage de la condition initiale, négligeant ainsi les effets dûs à la taille finie du système. La dynamique résultante peut être perçue comme un processus de branchement et, tel que vu en partie I, l'usage de fonctions génératrices permet de préserver la nature discrète des états. Le formalisme est légèrement amélioré : on obtient de nouvelles relations permettant de

---

3. On parle ici des indices inférieurs de l'équation (3.67).

« séparer » la distributions de probabilités des épidémies (composante géante) de celle des flambées épidémiques (petites composantes) en terme de générations d'infections. Puisque les effets de taille finie sont négligés, la distribution obtenue pour les épidémies n'est valide qu'au début de la propagation. Cependant, celle obtenue pour les flambées épidémiques reproduit bien le processus complet. Combiner ces résultats à ceux obtenus avec l'approximation de van Kampen permet donc une description adéquate de toutes les éventualités, en autant que le réseau soit suffisamment grand.

### 5.3 Prélude à « Spreading dynamics on complex networks : a general stochastic approach »

Le chapitre 7 reprend le thème central de cette thèse : étudier de façon analytique le comportement de systèmes composés de plusieurs éléments dont les interactions sont régies par un réseau complexe. Tout comme au chapitre 6, on peut percevoir le problème comme la recherche d'un processus de Markov simple reproduisant la dynamique d'un processus de Markov plus complexe. Cependant, on ne s'intéresse plus à une sous-classe particulière de processus mais plutôt à une méthode générale de modélisation applicable à ces problèmes.

L'approche choisie consiste à représenter de façon partielle la structure du réseau et l'état intrinsèque de ses composants à l'aide de petits sous-réseaux appelés *motifs*. On peut concevoir les motifs comme des « blocs de construction » pouvant être emboîtés afin de construire des réseaux : connaissant le nombre de blocs de construction de chaque type ainsi que les règles expliquant comment ils peuvent être emboîtés, on peut définir un ensemble (micro)canonique composé de toutes les dispositions possibles employant tous les blocs. Ainsi, on peut spécifier l'état du processus de Markov simplifié en énumérant les motifs qu'il contient, puis inférer l'état du processus de Markov original en supposant qu'il est tiré de l'ensemble de réseaux généré par ces motifs.

Pour une application particulière, le succès de l'approche repose sur l'identification des motifs à utiliser dans la définition de l'état du système simplifié : on souhaite utiliser le moins de motifs possible tout en représentant suffisamment bien l'état du système complet afin de reproduire adéquatement son comportement. Lorsqu'on choisit d'inclure un motif dans cette description, c'est typiquement parce qu'il correspond à une structure surreprésentée dans le réseau (par rapport à ce que le hasard pourrait expliquer) et/ou parce qu'il a une influence particulièrement importante sur la dynamique du système.

Une fois le système simplifié obtenu, le chapitre 7 fait un usage systématique de l'approximation gaussienne décrite au chapitre 6 afin d'obtenir une solution approximative pour la distribution des états du système. Cependant, il existe de nombreux autres outils permettant d'analyser de tels processus markoviens, sans compter que le même système peut aussi être perçu selon la perspective des processus de branchements ou encore de la physique statistique à l'équilibre. Ces deux dernières possibilités sont explorées au chapitre 8.



# Chapitre 6

## Propagation on networks : an exact alternative perspective<sup>1</sup>

P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau and L. J. Dubé

Département de Physique, de Génie Physique et d'Optique, Université Laval, Québec

En générant les spécificités de la structure d'un réseau seulement lorsque nécessaire (à-la-volée), on obtient un processus stochastique simple modélisant exactement l'évolution temporelle d'une dynamique susceptible-infectieux sur des réseaux de taille finie. Le petit nombre de variables dynamiques de ce processus de Markov de type naissance-mort simplifie grandement les calculs analytiques. On montre comment une description analytique duale, traitant les épidémies à grande échelle à l'aide d'une approximation gaussienne et les petites flambées épidémiques à l'aide d'un processus de branchement, fournit une approximation précise de la distribution même pour des réseaux relativement petits. L'approche offre également d'importants avantages numériques et se généralise à une vaste classe de systèmes.

By generating the specifics of a network structure only when needed (on-the-fly), we derive a simple stochastic process that exactly models the time evolution of susceptible-infectious dynamics on finite-size networks. The small number of dynamical variables of this birth-death Markov process greatly simplifies analytical calculations. We show how a dual analytical description, treating large scale epidemics with a Gaussian approximation and small outbreaks with a branching process, provides an accurate approximation of the distribution even for rather small networks. The approach also offers important computational advantages and generalizes to a vast class of systems.

---

1. Originellement publié dans [76].

## 6.1 Introduction

Real-world systems are often composed of numerous interacting elements. Complex network models prove to be valuable tools for systems where interactions are neither completely random nor completely regular [9, 11]. Among these systems, an important subclass concerns the propagation of *something* through interactions among the constituting elements. Examples include spreading of infectious diseases in populations [7, 18, 46, 54, 79, 106] as well as propagation of information [41, 45, 63], rumours [52, 98, 108], or viral marketing [36, 51] on social networks. We will hereafter call *infection* whatever is propagating.

Some modelling approaches are known to exactly reproduce the behaviour of propagation on networks in specific limiting cases. For example, branching processes [44, 70] may exactly predict the probability distribution for the final state of a system of infinite size. Similarly, heterogeneous mean field models [67, 84] may exactly predict the time evolution of relevant mean values for an infinite system that is annealed (*i.e.*, its structure changes at a rate arbitrarily faster than the propagation process). Finally, exact models are also possible for very specific network structures, *e.g.*, a linear chain [90].

In this paper, we present a stochastic process that exactly reproduces a propagation dynamics on quenched (fixed structure) configuration model networks of arbitrary size allowing for repeated links and self-loops (to be defined shortly). Section 6.2 defines the problem at hand and then presents our approach by comparing it to a computer simulation algorithm, which does not require a “network building” phase. However, this perspective is much more than an algorithmic trick saving computer resources: it changes a problem of propagation on a network into a Markov birth-death process, a momentous difference from an analytical point of view. In section 6.3, we assume a large system size and obtain analytical results for both the asymptotic behaviour of the “epidemics,” where an important fraction of the network gets infected, and for the probability distribution of the outbreaks, where a small number of nodes are affected. Our results compare advantageously to numerical simulations and account for finite-size effects. Finally, we show in section 6.4 how this approach generalizes to a vast class of systems and discuss possibilities for future improvements.



## 6.2 The exact model

### 6.2.1 Networks

A network model uses *nodes* to represent the elements composing the system of interest and assigns *links* between each pair of nodes corresponding to interacting elements. Two nodes sharing a link are said to be *neighbours*, and the *degree* of a node is its number of neighbours. The part of a link that is attached to a node is called a *stub*: there are two stubs per link and each node is attached to a number of stubs equal to its degree. A link with both ends leading to the same node is called a *self-loop*, and *repeated links* occur when more than one link joins the same pair of nodes.

We define the *configuration model* (CM) [64] specified by the vector  $\mathbf{n} = [n_0 \ n_1 \ \dots]^T$  as the (microcanonical) ensemble of networks such that each network of this ensemble contains, for each  $k$ , exactly  $n_k$  nodes of degree  $k$ . Clearly, each network of this ensemble has the same number of nodes  $N = \sum_k n_k$ . Since there are two stubs per link, the total number of stubs  $\sum_k kn_k$  must be even.

It is common practice to explicitly *forbid* self-loops and repeated links in CMs (CMF) since these structures are not observed in many real-world systems. However, it is often easier to study CMs *allowing* for self-loops and repeated links (CMA). Of importance is the fact that the distinction between CMF and CMA vanishes for large networks (the probability for a link in a CMA to be a self-loop or a repeated link goes as  $N^{-1}$ ). The knowledge acquired on CMAs can thus be translated to CMFs.

A simple way to build a CM network goes as follows.

- (i) For each  $k \in \{0, 1, \dots\}$ , create  $n_k$  nodes with  $k$  stubs.
- (ii) Randomly select a pair of unmatched stubs and match them to form a link. Special restriction for CMFs: if a self-loop or repeated link is created, discard the whole network and return to step *i*.
- (iii) Repeat *ii* until there are no unmatched stubs left.

### 6.2.2 Propagation

For the sake of demonstration, we first consider what may well be the simplest form of propagation on networks: the *susceptible-infectious* (SI) model. A node is said to

be *susceptible* if it does not carry the infection and *infectious* if it does. During an infinitesimal time interval  $[t, t + dt)$ , a susceptible node, neighbour to an infectious one, has a probability  $\beta dt$  to acquire the infection from the latter, hence becoming infectious. Once infectious, a node remains in this state forever.

For a given network structure, the following gives an algorithmic implementation of the SI model.

- (i) Set each node as either susceptible or infectious according to the initial conditions.
- (ii) Define small time intervals and start with the first one.
- (iii) For each infectious node, look up their susceptible neighbours. For each of them, randomly generate a number in the interval  $[0, 1)$  and test if it is lower than  $\beta dt$ . If yes, mark the corresponding node as infectious in the next time interval.
- (iv) Repeat *iii* for the next time interval.

Now consider the following change to step *iii*: perform the random number test for *each* neighbour of the infectious node, and, only when the test returns positive, verify if the corresponding neighbour is susceptible (if yes, mark it as infectious). This alternative algorithm is equivalent in all points to the original, except that the knowledge of who is the neighbour of an infectious node is not required until the very moment an infection may occur. Inspired by this seemingly benign observation, we will shortly present a stochastic process, equivalent to susceptible-infectious dynamics on CMA, that does not require an initial network construction step. Instead, the network will be built *on-the-fly*, concurrently with the propagation.

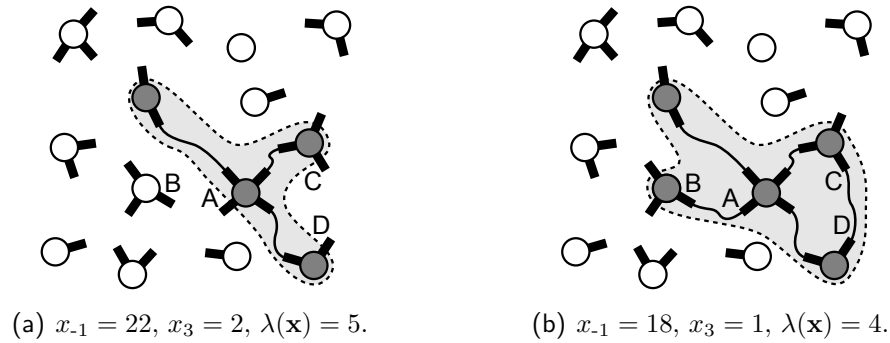
### 6.2.3 Equivalent stochastic process

CMA networks are built by randomly matching stubs together. In order to perform this match *on-the-fly*, we track the total number  $x_{-1}$  of unmatched stubs. All stubs belonging to susceptible nodes are unmatched. Denoting  $x_k$  the number of susceptible nodes of degree  $k$ , the total number of unmatched stubs belonging to infectious nodes is then

$$\lambda(\mathbf{x}) = x_{-1} - \sum_{k=0}^{k_{\max}} kx_k, \quad (6.1)$$

where  $\mathbf{x} = [x_{-1} \ x_0 \ x_1 \ \cdots \ x_{k_{\max}}]^T$  is the *state vector*.

During the interval  $[t, t + dt)$ , each of these  $\lambda(\mathbf{x})$  stubs has a probability  $\beta dt$  to infect the corresponding neighbouring node *under the condition* that it is currently susceptible. Since this infectious stub is currently unmatched, knowing which node is



**Figure 6.1:** Illustration of *on-the-fly* network construction. Susceptible (white circles) and infectious (grey circles) nodes each have a number of stubs equal to their respective degree. (a) At some point in the process, three links (thin black curves) have already been assigned. Future dynamics does not depend on how infectious nodes are linked (content of the grey zone) except for the total number  $\lambda(\mathbf{x})$  of unassigned stubs belonging to infectious nodes (stubs crossing the dashed border of the grey zone). (b) During any time interval  $[t, t + dt)$ , there is probability  $\beta\lambda(\mathbf{x})dt$  for an event to occur. Here, after many such time intervals, two new links have been assigned through an event of type  $j = 3$  (matching stubs A and B) and an event of type  $j = -1$  (matching stubs C and D). Again, other than for  $\lambda(\mathbf{x})$ , the future dynamics is not affected by how infectious nodes are linked.

at the other end simply requires to match it at random to one of the  $(x_{-1} - 1)$  other unassigned stubs. If a susceptible stub is chosen, the corresponding node is immediately infected and no matched susceptible stubs are created.

Since  $dt$  is infinitesimal, matching one of the  $\lambda(\mathbf{x})$  stubs has a probability  $\beta\lambda(\mathbf{x})dt$  to occur. In this case, the other stub selected for match has a probability  $[\lambda(\mathbf{x}) - 1](x_{-1} - 1)^{-1}$  to also be infectious, causing no new infection. Matching the two stubs amounts to decrease  $x_{-1}$  [and therefore  $\lambda(\mathbf{x})$ ] by 2. We refer to this class of events as a transition of type  $j = -1$ .

Alternatively, there is a probability  $kx_k(x_{-1} - 1)^{-1}$  for matching the infectious stub to a stub belonging to a susceptible node of degree  $k$ : it is marked as infectious by decreasing  $x_k$  by 1. Again,  $x_{-1}$  is decreased by 2 since two stubs have been matched together. This kind of event is referred to as a transition of type  $j = k$ .

Figure 6.1 illustrates the Markov stochastic process defined by these state vectors and transition rules. One may see the process from the infection's perspective: until it has crossed a link, it has no information concerning the node at the other end. More formally, the master equation (notation compatible with [32] section 7.5)

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{j=-1}^{k_{\max}} \left[ q_j(\mathbf{x} - \mathbf{r}^j) P(\mathbf{x} - \mathbf{r}^j, t) - q_j(\mathbf{x}) P(\mathbf{x}, t) \right], \quad (6.2)$$

governs the probability  $P(\mathbf{x}, t)$  to observe state  $\mathbf{x}$  at time  $t$ . For each transition type  $j$ , the function  $q_j(\mathbf{x})$  gives the probability rate at which this type of event occurs (given that the state of the system is currently  $\mathbf{x}$ ) while the vector  $\mathbf{r}^j$  gives the change caused by the transition (*i.e.*, the state becomes  $\mathbf{x} + \mathbf{r}^j$  after the transition). Translating the previous discussion in those terms, we obtain

$$q_j(\mathbf{x}) = \begin{cases} \beta\lambda(\mathbf{x}) \frac{\lambda(\mathbf{x}) - 1}{x_{-1} - 1} & \text{if } j = -1 \\ \beta\lambda(\mathbf{x}) \frac{jx_j}{x_{-1} - 1} & \text{if } j \geq 0 \end{cases} \quad (6.3)$$

for the rate at which transitions occur and

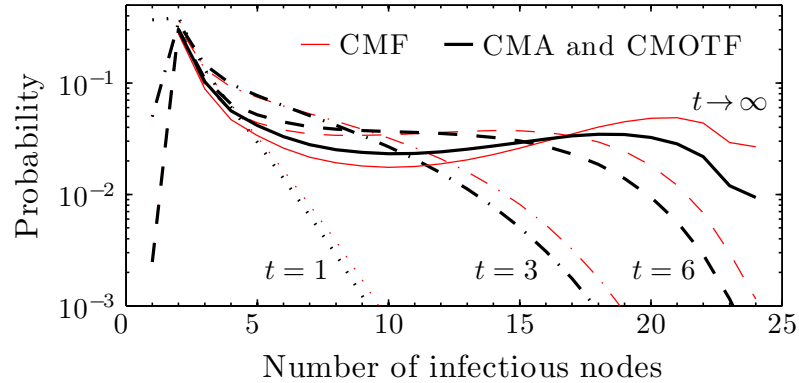
$$\mathbf{r}^j = \begin{cases} [-2 & 0 & 0 & 0 & \cdots]^T & \text{if } j = -1 \\ [-2 & -\delta_{0j} & -\delta_{1j} & -\delta_{2j} & \cdots]^T & \text{if } j \geq 0 \end{cases} \quad (6.4)$$

( $\mathbf{r}^{-1}$  has -2 at position -1 and 0 everywhere else, and  $\mathbf{r}^j$  with  $j \geq 0$  has an additional -1 at position  $j$ ) for the effect of such transitions. We use  $\beta = 1$  without loss of generality (scaling of time unit). Equations (6.2)–(6.4) define the stochastic process of the configuration model generated on-the-fly (CMOtF). A similar approach [19] has been developed independently for the rigorous proof that a specific spreading model proposed by Volz [99] holds true in the limit of large network size.

## 6.2.4 Comparison to numerical simulations

Figure 6.2 is obtained through direct Monte Carlo simulations for a network of  $N = 30$  nodes. Results for CMA and CMOtF are essentially identical (*i.e.*, the difference between them decreases as inversed square root of number of Monte Carlo simulations), in agreement with our claim that CMOtF exactly reproduces the behaviour of CMA. The effect of forbidding self-loops and repeated links accounts for the slight difference between results for CMF simulations and their CMOtF counterparts. Larger system sizes decrease further these differences ( $N = 300$  in figure 6.3) and therefore CMOtFs become excellent approximations of CMFs.

In terms of storage requirements, each CMOtF Monte Carlo simulation needs only to track the  $k_{\max} + 2$  integers composing the state vector  $\mathbf{x}$ . By comparison, a standard algorithm, first building the network then propagating the infections, must store the network structure as an adjacency list of  $Nz$  elements, where  $z$  is the average degree. Since  $k_{\max} \ll N$  for many networks of interest, the scaling of the memory requirements much favours CMOtF for large  $N$  (*e.g.*,  $N = 10^6$ ,  $z = 5$  and  $k_{\max} = 100$ ).



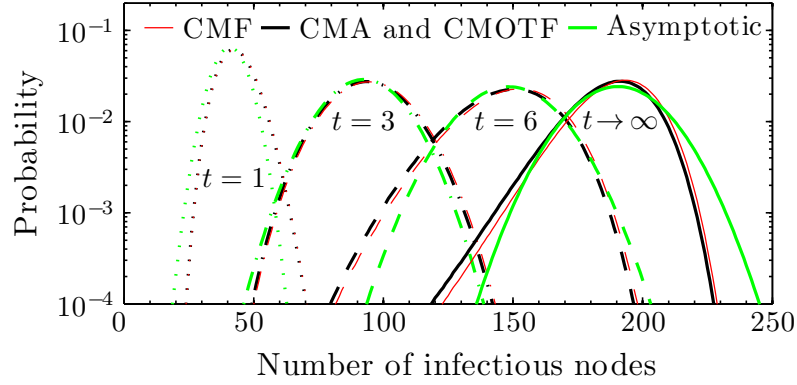
**Figure 6.2:** Snapshots at different times (line styles) of the probability distributions for the number of infectious nodes in three configuration models (line weight and colour). The *on-the-fly* process (CMOTF) and the configuration model allowing self-loops and repeated links (CMA) both give the same results. Even in such a small network ( $N = 30$ ), forbidding self-loops and repeated links (CMF) has minimal effect. Each distribution has been obtained through  $10^8$  Monte Carlo simulations. Degree sequence used:  $n_1 = 16$ ,  $n_2 = 8$ ,  $n_3 = 4$ , and  $n_4 = 2$ . All nodes are initially susceptible except for one infectious node of degree 1.

Moreover, CMOTF will usually run faster than a standard algorithm since it does not need to generate the parts of the network that are not affected by the infection. Hence, if CMA requires time  $\tau_{\text{build}}$  to generate the network and time  $\tau_{\text{spread}}$  to perform the SI simulation, CMOTF will approximately require time  $\rho \tau_{\text{build}} + \tau_{\text{spread}}$ , where  $\rho \in [0, 1]$  is the fraction of the links that only need to be allocated on-the-fly. At worst ( $\rho = 1$ ), the execution time will be similar.

For the sake of simplicity, the numerical algorithms for CMA, CMF and CMOTF were all presented in terms of infinitesimal time intervals. While this perspective is closer to our analytical work, these algorithms may be translated to a Gillespie-like [33] form that is faster and exact (to numerical precision). Here is how this translation is done.

In the case of CMA and CMF, the network construction is done as usual and the following algorithm is used.

- (i) Set each node as either susceptible or infectious according to the initial conditions.
- (ii) For each infectious node, draw a random number  $\Delta t > 0$  from the probability density function  $\beta e^{-\beta \Delta t}$  for each of its susceptible neighbour, and assign to this neighbour a clock that will ring at time  $\Delta t$ .
- (iii) Whenever a clock rings, check the state of the associated node. If it is susceptible, make the node infectious and proceed to step *iv*. If it is already infectious, ignore step *iv* and go to step *v*.



**Figure 6.3:** Probability distribution for the number of infectious nodes. The network is sufficiently large ( $N = 300$ ) for our asymptotic approximation to match the CM distributions around their peaks. Note that, especially at early times, the CMF results are very close to those of CMA and CMOTF (which is also an effect of larger network).  $10^8$  Monte Carlo simulations. Degree sequence:  $n_1 = 160$ ,  $n_2 = 80$ ,  $n_3 = 40$  and  $n_4 = 20$ . For each degree, 5% of the nodes are initially infectious.

- (iv) For each susceptible neighbours of the newly infectious node, draw a random number  $\Delta t > 0$  from the probability density function  $\beta e^{-\beta \Delta t}$  and assign to this neighbour a clock that will ring at time  $t + \Delta t$  (where  $t$  is the current time).
- (v) Return to step *iii* until no clocks remain.

In the case of CMOTF, the algorithm goes as follows.

- (i) Set  $\mathbf{x} = \mathbf{x}(0)$  (its initial condition) and  $t = 0$ .
- (ii) Draw a random number  $\Delta t > 0$  from the probability density function  $\beta \lambda(\mathbf{x}) e^{-\beta \lambda(\mathbf{x}) \Delta t}$ .
- (iii) Draw a random integer  $j \geq -1$  such that  $j = -1$  has probability  $(\lambda(\mathbf{x}) - 1)/(x_{-1} - 1)$  to occur while each  $j \geq 0$  occurs with probability  $j x_j / (x_{-1} - 1)$ .
- (iv) Increment  $t$  by  $\Delta t$  and  $\mathbf{x}$  by  $\mathbf{r}^j$ .
- (v) Return to step *ii* until  $\lambda(\mathbf{x}) = 0$ .

## 6.3 Asymptotically large systems

### 6.3.1 Gaussian approximation

The framework of a stochastic equation of the type defined by equations (6.2)–(6.4) offers the possibility of further simplification. Here, also, the analytical tractability is a consequence of the reduction to a state vector of dimension  $k_{\max} + 2$  and perhaps the most significant advantage of our approach. As long as all elements of  $\mathbf{x}(t)$  [and

$\lambda(\mathbf{x}(t))]$  are sufficiently “large”, equation (6.2) can be approximated by a stochastic differential equation (see [32] section 4.3.5)

$$d\mathbf{x} = \mathbf{a}(\mathbf{x})dt + B(\mathbf{x}) \cdot d\mathbf{W}, \quad (6.5)$$

where the vector  $\mathbf{W}(t)$  is a Wiener process while vector  $\mathbf{a}(\mathbf{x})$  and matrix  $B(\mathbf{x})$  are given in terms of  $q_j(\mathbf{x})$  and  $\mathbf{r}^j$  as

$$a_i(\mathbf{x}) = \sum_j r_i^j q_j(\mathbf{x}), \quad B_i^j(\mathbf{x}) = r_i^j \sqrt{q_j(\mathbf{x})}. \quad (6.6)$$

An approximate solution  $\mathbf{x}(t) \approx \boldsymbol{\mu}(t) + \boldsymbol{\nu}(t)$ , composed of a deterministic term  $\boldsymbol{\mu}(t)$  and a stochastic perturbation  $\boldsymbol{\nu}(t)$ , can be obtained when the noise term  $B(\mathbf{x}) \cdot d\mathbf{W}$  is much smaller than the deterministic term  $\mathbf{a}(\mathbf{x})dt$  [which implies that the value of  $\mathbf{x}(t)$  remains close to that of  $\boldsymbol{\nu}(t)$ ]. Using the initial conditions  $\boldsymbol{\mu}(0) = \mathbf{x}(0)$  and  $\boldsymbol{\nu}(0) = \mathbf{0}$ , the ordinary differential equation

$$\frac{d\boldsymbol{\mu}}{dt} = \mathbf{a}(\boldsymbol{\mu}). \quad (6.7)$$

governs the deterministic contribution. The approximation  $\mu_{-1} - 1 \approx \mu_{-1}$ , valid when  $\mu_{-1}$  remains large, gives

$$\frac{d\mu_{-1}}{dt} = -2\lambda(\boldsymbol{\mu}) \quad \frac{d\mu_k}{dt} = -\frac{k\mu_k\lambda(\boldsymbol{\mu})}{\mu_{-1}}. \quad (6.8)$$

One way to solve this system is to introduce a “time parameter”

$$\theta = \left[ \frac{\mu_{-1}}{x_{-1}(0)} \right]^{\frac{1}{2}} \quad \text{such that} \quad \frac{d\theta}{dt} = -\frac{\lambda(\boldsymbol{\mu})}{\theta x_{-1}(0)}. \quad (6.9)$$

We may then use equation (6.9) as a change of variable in equation (6.8), replacing the “actual time”  $t$  by  $\theta$ . Note that  $t = 0$  corresponds to  $\theta = 1$  and that  $\theta$  decreases with time. The resulting  $d\mu_j/d\theta$  differential equations are much simpler with solutions

$$\mu_{-1} = x_{-1}(0) \theta^2, \quad \mu_k = x_k(0) \theta^k \quad (6.10)$$

as a function of  $\theta$ . These can then be used in equation (6.9) to obtain an ordinary differential equation depending on  $\theta$  alone

$$\frac{d\theta}{dt} = \sum_{k=0}^{k_{\max}} \frac{x_k(0)}{x_{-1}(0)} k \theta^{k-1} - \theta. \quad (6.11)$$

Solving for  $\theta$  as a function of  $t$  will then provide  $\boldsymbol{\mu}(t)$  through equation (6.10). This is in agreement with previous results based on an heterogeneous mean field approach

[61, 62, 99]. The final state  $t \rightarrow \infty$  corresponds to the largest  $\theta \in [0, 1]$  such that  $x_{-1}(0)\theta^2 = \sum_{k=0}^{k_{\max}} kx_k(0)\theta^k$ .

The perturbation term  $\boldsymbol{\nu}$  can be obtained by solving the stochastic differential equation (see [32] section 6.2)

$$d\boldsymbol{\nu} = J_{\mathbf{a}}(\boldsymbol{\mu}) \cdot \boldsymbol{\nu} dt + B(\boldsymbol{\mu}) \cdot d\mathbf{W}, \quad (6.12)$$

where  $J_{\mathbf{a}}(\boldsymbol{\mu})$  is the Jacobian matrix of  $\mathbf{a}$  evaluated at  $\boldsymbol{\mu}$ . Initial conditions  $\langle \boldsymbol{\nu}(0) \rangle = \mathbf{0}$  and  $\text{cov}(\boldsymbol{\nu}(0)) = 0$  give the solutions  $\langle \boldsymbol{\nu} \rangle = \mathbf{0}$  and (see [32] section 4.4.9)

$$\text{cov}(\boldsymbol{\nu}) = \int_0^t \exp \left[ \int_{t'}^t J_{\mathbf{a}}(\boldsymbol{\mu}(t'')) dt'' \right] \cdot B(\boldsymbol{\mu}(t')) \cdot B(\boldsymbol{\mu}(t'))^T \cdot \exp \left[ \int_{t'}^t J_{\mathbf{a}}(\boldsymbol{\mu}(t''))^T dt'' \right] dt'. \quad (6.13)$$

Since  $\boldsymbol{\mu}$  and  $\boldsymbol{\nu}$  contribute exclusively to the mean and covariance of  $\mathbf{x}$ , respectively, we obtain

$$\langle \mathbf{x} \rangle = \boldsymbol{\mu}, \quad \text{cov}(\mathbf{x}) = \text{cov}(\boldsymbol{\nu}). \quad (6.14)$$

Specifically, the mean number of infectious nodes is given by  $N - \sum_{k=0}^{k_{\max}} \mu_k$  while its variance is  $\sum_{k,k'=0}^{k_{\max}} [\text{cov}(\boldsymbol{\nu})]_{kk'}$ . These values allow us to approximate the probability distribution for the number of infectious nodes by a Gaussian distribution.

### 6.3.2 Branching process approximation

In section 6.3.1, we have assumed that the probability distribution remains concentrated about its mean value. However, it is well known that this assumption is invalid when the initial condition contains a very small amount of infectious nodes. In fact, even if the parameters are such that the infection should initially grow *on average*, random events may cause an early end to the infection, thus splitting the probability distribution in two parts: small outbreaks and large scale epidemics.

In order to consider such eventualities, we focus on the initial behaviour of asymptotically large systems for which  $\lambda(\mathbf{x}(0)) \ll x_{-1}(0)$ . Since  $\mathbf{x}$  does not change much during these early times, we may treat as a constant the probability  $p_k$  for a random node to be of degree  $k$

$$p_k = \frac{x_k(0)}{\sum_{k'} k' x_{k'}(0)}. \quad (6.15)$$



The transition rates thus becomes

$$q_j(\mathbf{x}) \approx \beta \lambda(\mathbf{x}) j p_j \quad (6.16)$$

for events of type  $j \geq 0$ , and we may consider that events of type  $j = -1$  do not occur. In this form, the problem can be viewed as a branching process: an infection event of type  $j \geq 1$  directly causes  $j - 1$  future infection events, the probability for each of those future events to be of type  $j'$  being proportional to  $j' p_{j'}$ . We define *generations of infections* as follow: the nodes which begin as infectious at time  $t = 0$  are part of generation 0, and generation  $n$  contains all the nodes that have been infected by nodes of generation  $n - 1$ . Although some nodes of generation  $n$  may be infected at an earlier time than some nodes of generation  $n - 1$ , a higher generation usually implies a later time of infection.

Following previous work [75, 56], we model this branching process using *probability generating functions* (PGFs). For our purpose, we define a PGF as a power series, the coefficients of which are probabilities; see [105] for further details together with a more general perspective. A PGF *generates* its associated sequence of coefficients. Hence, the PGF

$$g_0(\xi) = \sum_k p_k \xi^k \quad (6.17)$$

generates the probability distribution for the degree of a random node, while the PGF

$$g_1(\xi) = \frac{g'_0(\xi)}{g'_0(1)} = \frac{\sum_k k p_k \xi^{k-1}}{\sum_{k'} k' p_{k'}} \quad (6.18)$$

generates the probability distribution for the excess degree of a node reached by following a random link (“excess” here means that the followed random link is excluded from the degree count). Alternatively, one may view the probability distribution generated by  $g_1(\xi)$  as the number of infections of generation  $n + 1$  that follow from a single infection of generation  $n$ .

PGFs allow for formal and/or analytical treatment of the generated sequences under the form of functions, often simplifying both the notation and the calculations [56, 75, 105]. For example, the composition  $g_1(g_1(\xi))$  generates the distribution of the number of infections of generation  $n + 2$  that follow from a single infection of generation  $n$ . Similarly,  $\xi g_1(\xi g_1(\xi))$  generates the total number of infections of generations  $n$ ,  $n + 1$  and  $n + 2$  that follow from a single infection of generation  $n$  (including that infection). The concept generalizes to more than one variable:  $\xi g_1(\xi g_1(\zeta))$  generates — through  $\xi$  — the total number of infections of generation  $n$  and  $n + 1$  and — through  $\zeta$  — the number of infections from generation  $n + 2$  that follows from a single infection of generation  $n$ .

As a slight generalization of the method presented in [56], we recursively introduce the two-variables PGFs

$$f_n(\xi, \zeta) = \xi g_1(f_{n-1}(\xi, \zeta)) \quad \text{with} \quad f_0(\xi, \zeta) = \zeta \quad (6.19)$$

such that  $f_n(\xi, \zeta)$  generates — through  $\xi$  — the total number of infections from generation 1 to  $n$  and — through  $\zeta$  — the number of infections of generation  $n + 1$  that follow from a single infection of generation 1. Hence, for an initial condition where all nodes are susceptible except for one randomly-chosen infectious node (generation 0), the PGF

$$h_n(\xi, \zeta) = \xi g_0(f_n(\xi, \zeta)) \quad (6.20)$$

generates — through  $\xi$  — the total number of infections from generation 0 to  $n$  and — through  $\zeta$  — the number of infections of generation  $n + 1$  that stem from these initial conditions; the results of [56] correspond to  $h_n(\xi, 1)$ . More generally, for an initial condition containing  $I_0$  initially infectious nodes and  $\lambda_0$  initially infectious stubs, the PGF becomes

$$\tilde{h}_n(\xi, \zeta; I_0, \lambda_0) = \xi^{I_0} [f_n(\xi, \zeta)]^{\lambda_0}. \quad (6.21)$$

We now seek to distinguish small outbreaks from large scale epidemics: the infinite-size propagation process terminates during an outbreak, while finite-size effects are required for an epidemic to end. In an infinite CM network [75], the probability for a single infection event to cause a terminating chain of infections (*i.e.*, it may cause infections that themselves cause infections, *etc.*, but the total number of infections caused this way is finite) is given by the lowest  $u \geq 0$  satisfying

$$u = g_1(u). \quad (6.22)$$

Hence,  $u < 1$  is the criteria for an epidemic to be possible. Noting  $m$  the number of infectious nodes in generation  $n + 1$ , the infinite-size infection process will terminate if and only if each one of the corresponding  $m$  infection events causes a terminating chain of events; this occurs with probability  $u^m$ . Therefore, the total number of infectious nodes from generation 0 to  $n$  that are part of outbreaks is generated by

$$\tilde{h}_n(\xi, u; I_0, \lambda_0) \quad (6.23)$$

in the general case [or by  $h_n(\xi, u)$  for a single random initially infectious node]. Since any remaining case leads to an epidemic, the total number of infectious nodes from generation 0 to  $n$  that are part of epidemics is generated by

$$\tilde{h}_n(\xi, 1; I_0, \lambda_0) - \tilde{h}_n(\xi, u; I_0, \lambda_0) \quad (6.24)$$

in the general case [or by  $h_n(\xi, 1) - h_n(\xi, u)$  for a single random initially infectious node]. Since  $f_n(1, u) = u$  for all  $n$ , one easily demonstrates that the total probability for an outbreak (or epidemic) is independent of the generation  $n$ .

Extracting the generated distribution (coefficients) from a PGF may be done numerically through a Cauchy integral [75] or, more efficiently, through a fast Fourier transform (FFT) [15, 56].

### 6.3.3 Comparison to numerical simulations

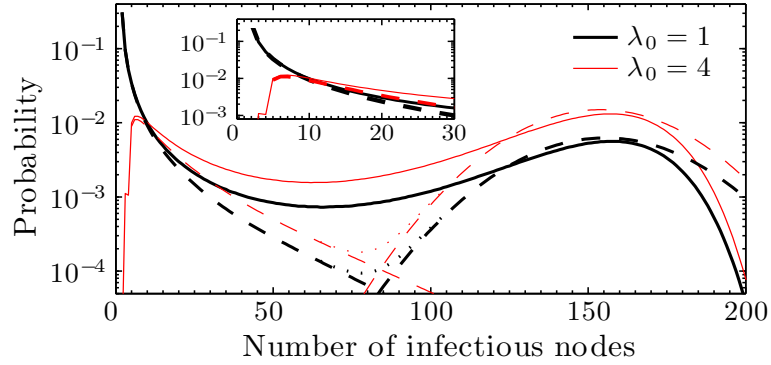
Applying the method of section 6.3.1 to the case studied in figure 6.3 shows that, although this Gaussian approximation of the exact dynamics [equations (6.2)–(6.4)] assumes an asymptotically large system, equation (6.14) provides reasonable results for networks as small as  $N = 300$ . In other words, for  $N$  sufficiently large, we can follow for all times the first two moments (mean and variance) of the exact dynamics. This size-independent Gaussian distribution becomes the universal limit for the underlying finite-size propagation model.

Part of this success is due to the fact that the initial condition contains  $\lambda_0 = 26$  infectious stubs (since for each degree 5% of the nodes are infectious), a sufficiently large value to (almost) guarantee that an epidemic will occur. In fact, using the method of section 6.3.2, we find  $u \approx 0.6375$ , which implies that the total probability for a small outbreak,  $u^{\lambda_0} \approx 8 \times 10^{-6}$ , is very unlikely. This explains why the complete neglect of the influence of small outbreaks provides accurate results in this case.

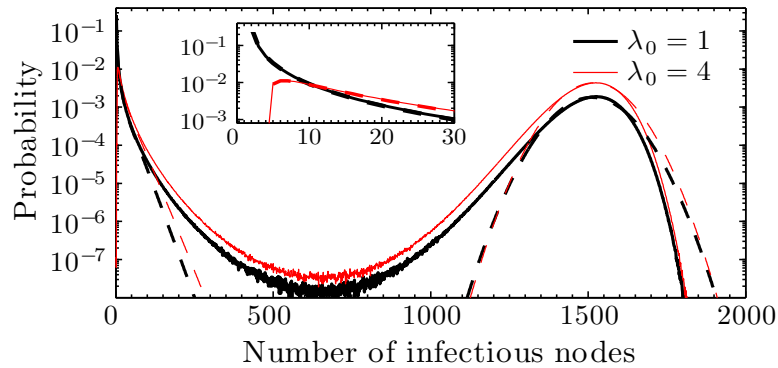
Figure 6.4 investigates the behaviour of the final distribution ( $t \rightarrow \infty$ ) when small outbreaks can not be neglected. Specifically, a single initially infectious node of degree 1 ( $\lambda_0 = 1$ ) or of degree 4 ( $\lambda_0 = 4$ ) is used for the same network as in figure 6.3 [ $N = 300$ , figure 6.4(a)] and for one with the same degree distribution with ten times as many nodes [ $N = 3000$ , figure 6.4(b)]. The distinction between the two limiting behaviours (outbreaks and epidemics) becomes clearer as  $N$  increases. Further comparisons may also be made with figure 6.2 for  $N = 30$  and  $\lambda_0 = 1$ .

For the small outbreaks, the branching process method of section 6.3.2 provides the final distribution for the small components with  $\tilde{h}_\infty(\xi, u; 1, \lambda_0)$ . These results are in good agreement with the numerical simulations for the small outbreaks, and increasing the network size improves this agreement. However, the same branching process method can not be used to predict the probability distribution for the epidemics in the limit  $n \rightarrow \infty$ : this distribution grows without bounds with  $n$  since finite-size effects are completely neglected. One result that does hold is that the total probability for an epidemic is  $1 - u^{\lambda_0}$ .

We also use the Gaussian approximation of section 6.3.1 to predict the shape of



(a)  $n_1 = 160$ ,  $n_2 = 80$ ,  $n_3 = 40$  and  $n_4 = 20$  ( $10^8$  simulations).



(b)  $n_1 = 1600$ ,  $n_2 = 800$ ,  $n_3 = 400$  and  $n_4 = 200$  ( $10^9$  simulations).

**Figure 6.4:** Probability distribution for the number of infectious nodes in the limit  $t \rightarrow \infty$ . Since the initial condition contains a small amount of infectious stubs ( $\lambda_0 = 1$  and  $\lambda_0 = 4$ ), the CMOtF probability distribution (plain curves) is roughly divided into two sub-distributions: small outbreaks and large scale epidemics. (a) While the separation between these sub-distributions is unclear for small networks ( $N = 300$ ), (b) the distinction becomes sharper as the size increase ( $N = 3000$ ). Analytical results (dashed curves) are obtained through branching processes (outbreaks) and Gaussian approximation (epidemics). Summing the contributions of these two limiting behaviours [dotted curve, only visible in (a) around the 80 infectious nodes mark] is insufficient to obtain the correct distribution for the outbreaks of intermediary size. However, such intermediate events gets less and less likely as the network size increase, thus making our two analytical distributions better approximations. Insets: zoom on the distributions for few infectious nodes.

the probability distribution for the outbreaks, then weight the whole distribution with a factor  $1 - u^{\lambda_0}$ . As seen on figure 6.4, the results are again in good agreement with the numerical simulations, and increasing the network size improves this agreement. It should be noted that, *a priori*, there was no guarantee for this simple approach to work: not only are the assumptions leading to the Gaussian approximation not met, but also the propagation processes that have a number of infectious stubs below the average are more likely to end early (outbreaks) than those that are above average. This introduces a bias in the distribution for the epidemics. Nonetheless, the global shape of the final distribution is quite stable under such early perturbations. While the early behaviour (and the initial conditions) is important for obtaining the total probability for epidemics, the final state of the epidemics is mainly governed by the finite-size effects. Combining the two methods thus provides a reliable estimate for the final distribution of the epidemics.

Although our analytical predictions are rather good, we systematically underestimate the value of the distribution for intermediate number of infections: the missing probabilities are being assigned to a larger number of infections. We may view such intermediate events as “small epidemics”: they would have led to “real epidemics” in a larger network, but finite-size effects caused the propagation to stop earlier, leading to a number of infections that may be comparable to those of outbreaks. Increasing  $N$  and/or  $\lambda_0$  decreases the probability of these events, and therefore improves the quality of the results of our dual approach.

Finally, even for large  $N$ , our Gaussian approximation for the distribution of epidemics shows systematic deviations: the distribution falls off faster than a Gaussian for large number of infections, and falls off slower for smaller-than-average epidemics. This is due to the fact that the finite-size effects become noticeable faster than predicted by our linear approximation [the Jacobian matrix  $J_{\mathbf{a}}(\boldsymbol{\mu})$ ]. Higher-order approximations should improve the description.

## 6.4 Conclusion

### 6.4.1 Generalization

The approach presented in this contribution heavily relies on the fact that the SI dynamics can be expressed under a form where, for each link, we *at most once* need to simultaneously know the state of the two nodes joined by that link. In fact, we

can generalize our exact approach to a vast class of systems for which this condition is respected. Indeed, given an arbitrary number of accessible node states (instead of “susceptible” and “infectious”), one could define a state vector  $\mathbf{x}$  such that its elements track the number of nodes with  $k$  unassigned stubs for each accessible node state and for each possible value of  $k$ .

As a concrete example, a *susceptible-infectious-removed* (SIR) system, *i.e.*, a susceptible-infectious where infectious nodes are removed at a constant probability rate, could be represented by the state vector

$$\mathbf{x} = [x_{S0} \ x_{I0} \ x_{R0} \ x_{S1} \ x_{I1} \ x_{R1} \ x_{S2} \ \cdots]^T$$

where  $x_{Sk}$ ,  $x_{Ik}$  and  $x_{Rk}$  stand for the number of susceptible, infectious and removed nodes with  $k$  unassigned stubs, respectively<sup>2</sup>. Since the simultaneous knowledge of the state of two neighbouring nodes is at most required once, we may perform on-the-fly neighbour assignment at the very time this knowledge is required, discarding the two stubs that were matched in the process.

An earlier version of this work [79] has made possible a recent contribution [54] which introduces a model for the deterministic (mean value) behaviour of two interacting SIR processes taking place on two partially overlaying networks. Even though the dynamics is quite complicated, the on-the-fly perspective allows us to accurately describe it at low computational cost. In this case as in many others, an exact stochastic version of the model could be implemented using the approach described in this paper.

## 6.4.2 Summary and perspective

We have presented a procedure that allows the construction of a network in a dynamical way on a need to know basis. This slight change of perspective has profound implications on the propagation dynamics on networks. It allows for a conceptual framework where the propagation is described exactly by a low-dimensional stochastic equation equivalent in all respects to the complete time evolution of the original problem. The low-dimensionality translates in large computational gains and, most importantly, it allows for analytical results through the use of standard tools from stochastic calculus. Perhaps the simplest of these tools allowed us to obtain a Gaussian approximation of the distribution for all times which becomes exact in the large network limit. Another simple tool, the branching process approach, allowed for a basic study of

---

2. A quick analysis reveals that tracking the sum  $\sum_k x_{Rk}$  instead of all the  $x_{Rk}$  would suffice for the same reason that tracking  $x_{-1}$  instead of the  $x_{Ik}$  was sufficient in the SI case.

---

the bimodal behaviour of the distribution (outbreaks and epidemics) that occurs when the initial condition does not guarantee a certain epidemic. Future contributions could improve the analytical description of intermediate events caused by early finite-size effects, and refine the distribution for the epidemics beyond the Gaussian assumption. Another interesting area of research concerns the application of the general method to other problems. Recent steps toward a general stochastic approach of the spreading dynamics on complex networks have already been taken [77].





# Chapitre 7

## Spreading dynamics on complex networks : a general stochastic approach <sup>1</sup>

P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau and L. J. Dubé

Département de Physique, de Génie Physique et d'Optique, Université Laval, Québec

On considère la dynamique sur réseaux du point de vue des processus stochastiques de Markov. On décrit partiellement l'état du système à l'aide de motifs et infère toutes données manquantes à l'aide de l'information disponible. Cette approche polyvalente est particulièrement bien adaptée à la modélisation des processus de propagation et/ou de la dynamique des populations. En particulier, la généralité de notre cadre systématique et le fait que ses suppositions soient explicitement données, suggère qu'il pourrait être utilisé comme une base commune pour comparer des modèles d'épidémies existants trop complexes pour une comparaison directe, tel que des simulations informatiques basées sur des agents. On fournit plusieurs exemples pour les cas particuliers des dynamiques susceptible-infectieux-susceptible (SIS) et susceptible-infectieux-retiré (SIR) (par exemple la propagation d'épidémies) et on observe plusieurs situations où des résultats précis peuvent être obtenus à un faible coût algorithmique. Notre perspective révèle un subtil équilibre entre les exigences complexes d'un modèle réaliste et ses suppositions de base.

Dynamics on networks is considered from the perspective of Markov stochastic processes. We partially describe the state of the system through network motifs and infer any missing data using the available information. This

---

1. Soumis pour publication.

versatile approach is especially well adapted for modelling spreading processes and/or population dynamics. In particular, the generality of our systematic framework and the fact that its assumptions are explicitly stated suggests that it could be used as a common ground for comparing existing epidemics models too complex for direct comparison, such as agent-based computer simulations. We provide many examples for the special cases of susceptible-infectious-susceptible (SIS) and susceptible-infectious-removed (SIR) dynamics (*e.g.*, epidemics propagation) and we observe multiple situations where accurate results may be obtained at low computational cost. Our perspective reveals a subtle balance between the complex requirements of a realistic model and its basic assumptions.

## 7.1 Introduction

Mathematical modelling has proven a valuable tool when addressing population dynamics problems, both for public health and ecological issues. The increase in availability of powerful computer resources has facilitated the use of agent-based models and other complex modelling approaches, all accounting for numerous parameters and assumptions [6, 14, 57]. Our confidence in these models may increase when they are shown to agree with empirical observations and/or with previously accepted models. However, when discrepancies appear, the complexity of these computer programs may obfuscate the effect of underlying assumptions, making it difficult to isolate the source of disagreement. While analytical approaches offer more insights on the underlying assumptions, their use is often restricted to simpler interaction structures and/or dynamics.

The purpose of this paper is to systematically model the global behaviour of stochastic systems composed of numerous elements interacting in a complex way. “Complex” here implies that interactions among the elements follow some nontrivial patterns that are neither perfectly regular nor completely random, as often seen in real-world systems. “Stochastic” implies that the system may not be completely predictable to us and that a probabilistic solution is sought.

To this end, we present (section 7.2) a general modelling scheme where network theory [9, 11] accounts for the interactions between the elements of the system and where a birth-death Markov process [32] models the stochastic dynamics. Since a tremendous amount of information may be required to store the state of the whole system, we seek the part of this information that is important for the problem at hand and then approximate the dynamics by tracking only this limited subset. Part of the discarded data may still affect, albeit weakly, the behaviour of the system. We fill this knowledge gap

by inferring the missing information such that it is consistent both with the information we follow and any other prior information that is available to us.

An important part of this paper (section 7.3) provides explicit examples to these general ideas. Keeping in mind that the generality of our approach is in the mathematical description and not in the studied cases, we focus our study to spreading processes such as the propagation of infectious diseases on contact networks [7, 18, 46]. Specifically, our examples correspond to either one of two standard epidemic models, the susceptible-infectious-susceptible (SIS) and the susceptible-infectious-removed (SIR) dynamics, for which we also adopt the relevant vocabulary. While our first examples study simpler cases, facilitating the understanding of our systematic method, the later models show how the same approach applies to more complex interaction structures.

We then compare and analyse the results of these examples (section 7.4). This reveals some general considerations for both the accuracy and the complexity of our modelling approach. We find that treating the inferences of missing information explicitly helps systematize the model development and highlights numerous possibilities for future developments. An important simplification occurs for SIR spreading processes and related dynamics, leading to an *exact* model with a small number of dynamical variables.

We conclude (section 7.5) on how our general approach may be applied to population dynamics in general, as well as to other spreading processes such as the cascading extinctions of species in food webs [10, 23, 88]. Returning to the problem of understanding the source of discrepancies in complex models, we explain how *modelling these models* with our method could help identifying important assumptions and isolating the source of disagreement. Mathematical details and further generalizations are presented in two separate Appendices.

## 7.2 General modelling scheme

We assume that the real-world system to be modelled is sufficiently well understood to implement a Monte Carlo computer simulation that approximately reproduces its behaviour. We refer to this hypothetical computer simulation as the *full system*:  $Z$  denotes the *state* of the full system (*i.e.*, all the data that would be stored by the computer program) while  $V$  denotes the *rules* governing the evolution of  $Z$  in time (*i.e.*, the program itself).

However, there are many situations where a direct implementation of the full system is impractical due to storage and/or computation considerations. We thus design a *simplified system* that aims at reproducing the behaviour of the full one, while requiring less resources.

The state  $X$  of this simplified system (*much* smaller than  $Z$ ) evolves in time according to the rules  $W$ . Moreover, we note  $Y$  any known *prior information* that is relevant in a Bayesian inference of  $Z$

$$P(Z|X, Y) = \frac{P(Z|Y) P(X|Y, Z)}{P(X|Y)} \quad . \quad (7.1)$$

This last point is crucial:  $Y$  often makes the difference between an accurate model and a useless one. It bridges the gap between the simplified representation of the state of the system (*i.e.*,  $X$ ) and the full one (*i.e.*,  $Z$ ).

Since we are interested in systems composed of many elements interacting through complex patterns, we express the previous quantities in terms of networks.

### 7.2.1 Networks

A *network* (graph) is a collection of *nodes* (vertices) and *links* (edges). Nodes model the elements of a system; links join nodes pairwise to represent interactions between the corresponding elements. Two nodes sharing a link are said to be *neighbours* and the *degree* of a node is its number of neighbours. The part of a link that is attached to a node is called a *stub*: there are two stubs per link and each node is attached to a number of stubs equal to its degree. A link with both ends leading to the same node is called a *self-loop* and *repeated links* occur when more than one link join the same two nodes.

Links may (or may not) be directed: an *undirected* link represents a bidirectional and symmetric interaction between the two linked nodes, while a *directed* link represents interactions that are either unidirectional or asymmetrical. A network that has only undirected links is said to be *undirected*, one that has only directed links is *directed* and one that has both is *semi-directed*.

There may be systems such that specifying its state  $Z$  exactly amounts to specifying the network structure. However, most systems are not purely structural: they are composed of elements that, by themselves, require additional information to be properly characterized. Hence, we assign to each node (resp. link) a *node state* (resp. *link state*)

that specifies the intrinsic properties of the corresponding element (resp. interaction) in the system. There is a total of  $\mathcal{N}$  discrete node states (resp.  $\mathcal{L}$  discrete link states) and the intrinsic state  $\nu$  of a given node (resp.  $\ell$  of a given link) may change in time to any of these accessible values; the special case  $\mathcal{N} = 1$  (resp.  $\mathcal{L} = 1$ ) corresponds to the situation where every nodes (resp. links) remain intrinsically indiscernible during the whole process. At any given time, the state  $Z$  of the full system is specified by the intrinsic state of all of its components (nodes and links) in addition to the network structure governing their interactions. While each node (or link) may be in only one intrinsic state at any given time, different pieces of information may be encoded in this single state through, *e.g.*, Cartesian product of sub-states.

## 7.2.2 Motifs

Specifying the complete structure of a complex network requires a tremendous amount of information. Since we want the state  $X$  of a simplified system to be of manageable size, approximations have to be made. A convenient way to do so, and one that has proven to give good results in the past [35, 37, 40, 44, 54, 55, 76], is to specify the network structure through its building blocks.

A network *motif* is a pattern that may appear a number of times in the network. For example, two linked nodes form a *pair motif* while three nodes all neighbours of one another form a *triangle motif*. Motifs may encode intrinsic node states or other relevant information; further details and examples are provided throughout section 7.3 as well as in Appendix 7.A.

We define the *state vector*  $\mathbf{x}$  of a system as a vector of integers specifying how many times different motifs appear in the network. These motifs may be attached together to form a network structure: the state vector  $X = \mathbf{x}$  enumerates the available building blocks while the prior information  $Y$  specifies how such blocks may be attached. There will usually be numerous valid ways to attach the blocks, some more probable than others. Given the available information, the resulting distribution is our best estimate for  $P(Z|\mathbf{x}, Y)$ .

By judiciously choosing the motifs enumerated in  $\mathbf{x}$  and by specifying informative prior information  $Y$ , one may hope for this probability distribution to be densely localized around the “real” value of  $Z$  in the full system. This mapping can then be used to convert the rules  $V$  of the full system to the rules  $W$  of the new simplified one. We approach this problem from the perspective of birth-death Markov processes.

### 7.2.3 Birth-death stochastic processes

In a *birth-death* process, the elements composing a system may be destroyed (death) while new ones may be created (birth). It is therefore natural to state the rules  $W$  of our simplified system in those terms: any change in the state vector  $\mathbf{x}$  may be perceived as an event where motifs are created and/or destroyed.

Quantitatively, a *forward transition event of type  $j$*  takes the system from state  $\mathbf{x}$  to state  $\mathbf{x} + \mathbf{r}^j$  and has probability  $q_j^+(\mathbf{x}, Y) dt$  to occur during the time interval  $[t, t + dt)$ . Similarly, a *backward transition event of type  $j$*  takes the system from state  $\mathbf{x}$  to state  $\mathbf{x} - \mathbf{r}^j$  and has probability  $q_j^-(\mathbf{x}, Y) dt$  to occur during the same time interval.

Specifying for each  $j$  the elements  $r_i^j$  of the *shift vector*  $\mathbf{r}^j$  together with the *rate functions*  $q_j^+(\mathbf{x}, Y)$  and  $q_j^-(\mathbf{x}, Y)$  thus completely define the rules  $W$  governing the simplified system. This Markov process is summarized in the master equation

$$\begin{aligned} \frac{dP(\mathbf{x}|Y, t)}{dt} = \sum_j \left[ q_j^+(\mathbf{x} - \mathbf{r}^j, Y) P(\mathbf{x} - \mathbf{r}^j|Y, t) - q_j^+(\mathbf{x}, Y) P(\mathbf{x}|Y, t) \right. \\ \left. + q_j^-(\mathbf{x} + \mathbf{r}^j, Y) P(\mathbf{x} + \mathbf{r}^j|Y, t) - q_j^-(\mathbf{x}, Y) P(\mathbf{x}|Y, t) \right] \end{aligned} \quad (7.2)$$

specifying the evolution of the probability  $P(\mathbf{x}|Y, t)$  to observe state  $\mathbf{x}$  at time  $t$  (notation compatible with [32] section 7.5).

We now consider two approximations that are often justified for large systems: the elements of  $\mathbf{x}$  may be treated as varying continuously and the probability distribution is strongly concentrated around its mean value. In such cases, the evolution of the mean value  $\boldsymbol{\mu}(t) = \sum_{\mathbf{x}} \mathbf{x} P(\mathbf{x}|Y, t)$  for the vector  $\mathbf{x}$  at time  $t$  is approximately given by

$$\frac{d\boldsymbol{\mu}(t)}{dt} = \mathbf{a}(\boldsymbol{\mu}(t)) \quad (7.3a)$$

$$a_i(\mathbf{x}) = \sum_j r_i^j [q_j^+(\mathbf{x}, Y) - q_j^-(\mathbf{x}, Y)] \quad (7.3b)$$

where we defined the *drift vector*  $\mathbf{a}(\mathbf{x})$  of elements  $a_i(\mathbf{x})$  (see [32] section 7.5.3 and 4.4.9).

In order to further refine our knowledge of  $P(\mathbf{x}|Y, t)$  in the vicinity of this deterministic solution, we define the *evolution matrix*  $A(t, t')$ , the *diffusion matrix*  $B(\mathbf{x})$  (of

elements  $B_{ii'}(\mathbf{x})$  and the *covariance matrix*  $C(t)$

$$A(t, t') = \exp \left[ \int_{t'}^t J_{\mathbf{a}}(\boldsymbol{\mu}(t'')) dt'' \right] \quad (7.4a)$$

$$B_{ii'}(\mathbf{x}) = \sum_j r_i^j r_{i'}^j \left[ q_j^+(\mathbf{x}, Y) + q_j^-(\mathbf{x}, Y) \right] \quad (7.4b)$$

$$C(t) = \int_0^t A(t, t') \cdot B(\boldsymbol{\mu}(t')) \cdot A(t, t')^T dt' \quad (7.4c)$$

where  $J_{\mathbf{a}}(\mathbf{x})$  is the Jacobian matrix of  $\mathbf{a}$  evaluated at  $\mathbf{x}$ . Noting  $d$  the size of the vector  $\mathbf{x}$ , the probability distribution may be approximated by a  $d$ -dimensional Gaussian

$$P(\mathbf{x}|Y, t) = \frac{\exp \left\{ -\frac{1}{2} [\mathbf{x}(t) - \boldsymbol{\mu}(t)]^T \cdot C(t)^{-1} \cdot [\mathbf{x}(t) - \boldsymbol{\mu}(t)] \right\}}{\sqrt{(2\pi)^d |C(t)|}} \quad (7.5)$$

where  $|C(t)|$  is the determinant of  $C(t)$ . Note that (7.2)–(7.5) are all textbook relationships.

Although many other tools are available for the analysis of stochastic systems, the simplicity, the generality and the straightforwardness of the Gaussian approximation make it an instrument of choice that will be used extensively in this article.

## 7.3 Application to spreading dynamics

Without prejudice to the generality of section 7.2, we now focus our study to spreading processes. An epidemiological terminology is used: whatever propagates among neighbouring nodes, be it desirable or not, is called an *infection*. We find that the basic SIS and SIR epidemiological models, both to be defined shortly, require little prior knowledge from the part of the reader while being sufficiently complex for the needs of the present study. For similar reasons, we assume that links are undirected and indiscernible from one another (*i.e.*,  $\mathcal{L} = 1$ ); see Appendix 7.A for discussions concerning directed links and more than one intrinsic link state.

At a given time, the intrinsic state of each node of an *SIS model* may be either of  $\mathcal{N} = 2$  accessible states: *Susceptible* (not carrying the infection) or *Infectious* (carrying the infection). The full system state  $Z$  hence specifies each node's intrinsic state together with the complete structure of the network. The rules  $V$  are simple: during any time interval  $[t, t + dt)$ , each infectious node may recover (*i.e.*, it becomes susceptible) with

probability  $\alpha dt$  and, for each of its susceptible neighbours, has probability  $\beta dt$  to transmit the infection (*i.e.*, the neighbour becomes infectious).

In addition to the susceptible and infectious intrinsic states, the nodes of an *SIR model* may also be *Removed* (once had the infection and can neither acquire nor transmit it ever again) and there are thus  $\mathcal{N} = 3$  accessible states. The rules  $V$  are the same than for the SIS model with respect to infection (*i.e.*, infectious nodes transmit to their susceptible neighbour with probability  $\beta dt$ ), but recovery is replaced by removal (*i.e.*, infectious nodes become removed with probability  $\alpha dt$ ).

The remainder of this section studies how different choices of state vector  $\mathbf{x}$  and prior information  $Y$  translate in the rules  $W$  of the simplified system. In each case,  $W$  is defined through a set of equations whose tags all share the same numeral, *e.g.*, (7.6a)–(7.6f). Although figures present results concomitantly with the specification of the corresponding models, all discussions are delayed to section 7.4.

### 7.3.1 Pair-based SIS model

Section 7.2.2 defined a pair motif as two linked nodes. Since the nodes of a SIS model are either susceptible or infectious, there are three possibilities for pair motifs: two linked susceptible nodes (noted  $S$ – $S$ ), two linked infectious nodes (noted  $I$ – $I$ ) and a susceptible node linked to an infectious one (noted  $S$ – $I$ ). Two nodes involved in a pair motif may have other neighbours.

Pair motifs are often used in conjunction with *node motifs*: the trivial structure that is one node. In the SIS model, there are two possibilities for a node motif: susceptible nodes (noted  $S$ ) and infectious nodes (noted  $I$ ). A state vector  $\mathbf{x}$  based on both node and pair motifs would thus be composed of five elements enumerating the amount of times each motif appears in the network:  $x_S$ ,  $x_I$ ,  $x_{S-S}$ ,  $x_{S-I}$  and  $x_{I-I}$ . However, additional assumptions about the structure of the network may cause some of these quantities to be redundant.

#### Degree-regular network

We first consider the simple case where the network is known to be a *n-regular network* of size  $N$ : there are  $N$  nodes in the network which all have  $n$  neighbours (degree  $n$ ). Such a network must respect the structural constraints  $x_S = N - x_I$ ,



$x_{S-S} = \frac{1}{2}(nx_S - x_{S-I})$  and  $x_{I-I} = \frac{1}{2}(nx_I - x_{S-I})$ . Hence, with the prior information  $Y$  specifying  $N$  and  $n$ , the state vector

$$\mathbf{x} = (x_I, x_{S-I}) \quad (7.6a)$$

suffices to obtain all the five node and pair motifs.

In those terms, the rules  $V$  specify that an infection has probability  $\beta x_{S-I} dt$  to occur during the time interval  $[t, t + dt)$  while a recovery has probability  $\alpha x_I dt$  to occur. Clearly, an infection translates to the destruction of a  $S$  motif and the creation of a new  $I$  one, and a recovery corresponds to the inverse process. However, pair motifs are also affected by such transitions since the affected node had neighbours. Hence, the effect on  $\mathbf{x}$  of the infection or recovery of a node depends on some information that is not directly tracked by  $\mathbf{x}$  — *i.e.*, what is the state of the infected or recovered node's neighbours — and we thus have to *infer* this information from the available data.

In order to facilitate this inference, we define the *first neighbourhood motif*  $ST_1(k_S, k_I)$  as a susceptible node that has  $k_S$  susceptible neighbours and  $k_I$  infectious neighbours. Similarly, the motif  $IT_1(k_S, k_I)$  corresponds to an infectious node with  $k_S$  susceptible neighbours and  $k_I$  infectious ones. In both cases, we qualify as *central* the node whose neighbours are explicitly stated. The other nodes of the first neighbourhood motif, *i.e.*, the neighbours of the central node, may have other neighbours of their own.

We can now define a forward transition event of type  $j \in \{0, 1, \dots, n\}$  as the infection of the central node of a  $ST_1(n - j, j)$  motif. In terms of node and pair motifs, this implies the destruction of one of the  $S$  motifs, of  $n - j$  of the  $x_{S-S}$  motifs and of  $j$  of the  $x_{S-I}$  motifs together with the creation of one new  $I$  motif, of  $n - j$  new  $x_{S-I}$  motifs and of  $j$  new  $x_{I-I}$  motifs. Since only  $x_I$  and  $x_{S-I}$  are tracked, the shift vectors are

$$\mathbf{r}^j = (1, n - 2j) \quad . \quad (7.6b)$$

This same vector also defines the backward transition events  $j \in \{0, 1, \dots, n\}$  which correspond to the recovery of the central node of a  $IT_1(n - j, j)$  motif.

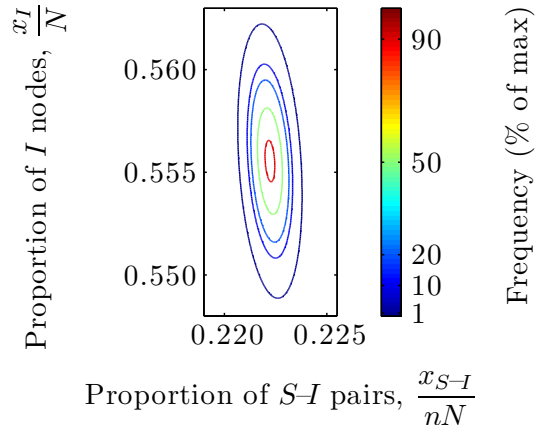
Looking back at the rules  $V$ , the corresponding forward and backward transition rate functions are

$$q_j^+(\mathbf{x}, Y) = \beta x_{S-I} P(ST_1(n - j, j) | S, j \geq 1, \mathbf{x}, Y) \quad (7.6c)$$

$$q_j^-(\mathbf{x}, Y) = \alpha x_I P(IT_1(n - j, j) | I, \mathbf{x}, Y) \quad (7.6d)$$

where two inference terms have been defined.

The inference term of (7.6d) gives the probability for a motif to be a  $IT_1(n - j, j)$  knowing that it has an infectious node at its center and that the current state vector



**Figure 7.1:** Distribution of post-transient ( $t \rightarrow \infty$ ) outcomes as predicted by the SIS model on regular network (7.6) using the approximations (7.3)–(7.4). The axes (proportion of  $I$  among node motifs vs proportion of  $S$ – $I$  among pair motifs), network structure ( $N = 10^5$  nodes, each of degree  $n = 5$ ) and parameters ( $\alpha = 0.1^2$  and  $\beta = 0.05$ ) are the same as for figure 2(c) of [17]. Frequencies (in percent) are used to facilitate comparison with [17]; they are simply obtained from the probability densities of (7.5) multiplied by 100.

is  $\mathbf{x}$  with prior information  $Y$ . For a sufficiently large network, this approximately corresponds to randomly drawing the  $n$  neighbours of the central infectious node among the pair motifs  $S$ – $I$  and  $I$ – $I$

$$P(I\Gamma_1(n-j, j)|I, \mathbf{x}, Y) = \binom{n}{j} \left(\frac{x_{S-I}}{n x_I}\right)^j \left(1 - \frac{x_{S-I}}{n x_I}\right)^{n-j}. \quad (7.6e)$$

The inference term of (7.6c) is very similar except that the central susceptible node is known to have at least one infectious neighbours since it acquired the infection through a  $S$ – $I$  motif

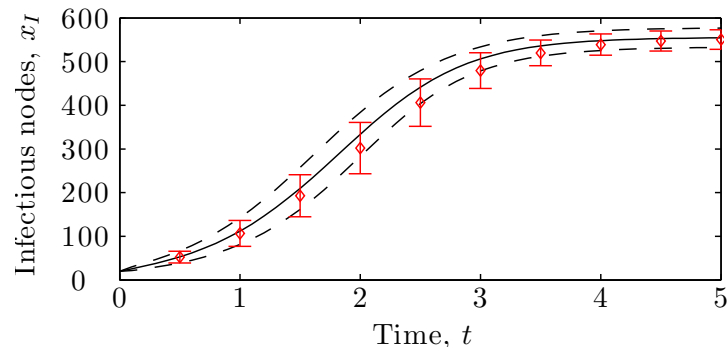
$$P(S\Gamma_1(n-j, j)|S, j \geq 1, \mathbf{x}, Y) = \binom{n-1}{j} \left(\frac{x_{S-I}}{n(N-x_I)}\right)^j \left(1 - \frac{x_{S-I}}{n(N-x_I)}\right)^{n-1-j} \quad (7.6f)$$

which complete the rules  $W$  for the pair-based SIS model on a  $n$ -regular network of size  $N$ .

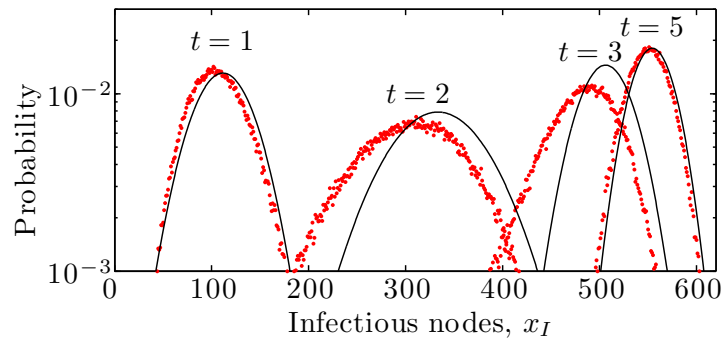
Figure 7.2 compares the results produced by this simplified model (defined by  $W$ ,  $\mathbf{x}$  and  $Y$ ) to the corresponding full one (defined by  $V$  and  $Z$ ). Figure 7.3 shows the probability distribution for the same data. Note that, although presented differently, this model corresponds to the one presented in [17]; figure 7.1 is provided for comparison with figure 2(c) of [17].

---

2. M.J. Keeling confirmed us that  $\alpha = 0.01$ , noted  $\gamma$  in [17], is a typo.



**Figure 7.2:** Time evolution of the number of infectious nodes  $x_I$  for SIS dynamics ( $\alpha = 2$  and  $\beta = 1$ ) on a regular network of  $N = 10^3$  nodes (20 initially infectious) of degree  $n = 5$ . Curves: results for the simplified system (7.6) approximated with (7.3)–(7.4). The continuous curve shows the mean value while the dashed curves delimit the range of one standard deviation above and below the mean. Symbols: averaged results of  $10^5$  numerical simulations of the full system. The parameters  $\alpha$  and  $\beta$  correspond to those of figure 7.1 after rescaling the time unit.



**Figure 7.3:** Probability distribution at different times for the number of infectious nodes  $x_I$ . All parameters are the same as in figure 7.2. Curves: Gaussian approximation for the simplified system. Symbols: binned results of  $10^5$  numerical simulations of the full system.

### Erdős-Rényi network

We now consider the case where the network is an Erdős-Rényi network: there are  $N$  nodes in the networks and  $M$  links are randomly assigned. This knowledge constrains two of the five node and pair motifs (*i.e.*,  $x_S = N - x_I$  and  $x_{S-S} = M - x_{S-I} - x_{I-I}$ ) and a state vector of three elements suffices

$$\mathbf{x} = (x_I, x_{S-I}, x_{I-I}) \quad . \quad (7.7a)$$

The method used in section 7.3.1 has to be adapted since the degree of each node is not constrained to a single value. Indeed, a susceptible node that gets infected may *a priori* be the center of any of the  $S\Gamma_1(k_S, k_I)$  motifs. Still, we could design a bijective mapping between the vector of integers  $\mathbf{k} = (k_S, k_I)$  and an event type  $j$ .

The details of the chosen mapping do not matter: we simply define the forward transition event of type  $\mathbf{k}$  as the infection of the central node of a  $S\Gamma_1(k_S, k_I)$  motif, which may conveniently be noted  $S\Gamma_1(\mathbf{k})$  instead. Similarly, the backward transition event of type  $\mathbf{k}$  is defined as the recovery of the central node to a  $I\Gamma_1(\mathbf{k})$  motif. The corresponding shift vector and rate functions are

$$\mathbf{r}^{\mathbf{k}} = (1, k_S - k_I, k_I) \quad (7.7b)$$

$$q_{\mathbf{k}}^+(\mathbf{x}, Y) = \beta x_{S-I} P(S\Gamma_1(\mathbf{k})|S, k_I \geq 1, \mathbf{x}, Y) \quad (7.7c)$$

$$q_{\mathbf{k}}^-(\mathbf{x}, Y) = \alpha x_I P(I\Gamma_1(\mathbf{k})|I, \mathbf{x}, Y) \quad (7.7d)$$

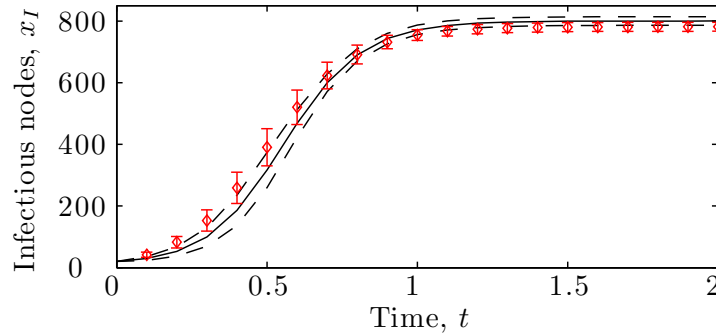
where the inference terms bear the same meaning as their previous counterpart. Again assuming large network size, these inference terms are obtained by evaluating the probability for each pair motif to include the considered node. Hence, the products of binomial distributions

$$\begin{aligned} & P(S\Gamma_1(\mathbf{k})|S, l \geq 1, \mathbf{x}, Y) \\ &= \binom{2x_{S-S}}{k_S} x_S^{-k_S} (1 - x_S^{-1})^{2x_{S-S} - k_S} \binom{x_{S-I} - 1}{k_I - 1} x_S^{-(k_I - 1)} (1 - x_S^{-1})^{x_{S-I} - k_I} \end{aligned} \quad (7.7e)$$

$$\begin{aligned} & P(I\Gamma_1(\mathbf{k})|I, \mathbf{x}, Y) \\ &= \binom{x_{S-I}}{k_S} x_I^{-k_S} (1 - x_I^{-1})^{x_{S-I} - k_S} \binom{2x_{I-I}}{k_I} x_I^{-k_I} (1 - x_I^{-1})^{2x_{I-I} - k_I} \end{aligned} \quad (7.7f)$$

complete the rules  $W$  for the pair-based SIS model on a Erdős-Rényi network of size  $N$  with  $M$  links.

Figure 7.4 compares the results produced by this simplified model to the corresponding full one. Section 7.4.2 discusses these results and provides further details concerning pair-based models.



**Figure 7.4:** Time evolution of the number of infectious nodes  $x_I$  for SIS dynamics ( $\alpha = 2$  and  $\beta = 1$ ) on an Erdős-Rényi network of  $N = 10^3$  nodes (20 initially infectious) and  $M = 5 \cdot 10^3$  links. The mean and range of one standard deviation above and below the mean are shown. Curves: simplified system. Symbols: full system ( $10^5$  simulations).

### 7.3.2 First neighbourhood SIS model

We consider a full model ( $V$  and  $Z$ ) for SIS dynamics on a *configuration model* (CM) network: given a sequence  $\{n_0, n_1, n_2, \dots\}$ , links are randomly assigned between nodes such that, for each degree  $\kappa$ , there are  $n_\kappa$  nodes of degree  $\kappa$ . In a computer simulation, we create  $n_\kappa$  nodes with  $\kappa$  stubs for each possible  $\kappa$  and then randomly pair stubs to form links. No particular mechanism is used to prevent the formation of repeated links and self-loops: this simplifies the analytical treatment and has little effect when the network size is sufficiently large.

Our simplified model handles the heterogeneity in node degree by enumerating every possible first neighbourhood motifs in its state vector

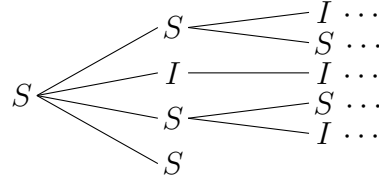
$$\mathbf{x} = (x_{S\Gamma(0,0)}, x_{S\Gamma(1,0)}, \dots, x_{I\Gamma(0,0)}, \dots) \quad . \quad (7.8a)$$

Although this vector should be infinite in the general case, it is not the case when, *e.g.*, the prior information  $Y$  states that no node has a degree superior to  $\mathcal{K}$ .

For the same reasons that models tracking node and pair motifs (section 7.3.1) had their transition events defined in terms of first neighbourhood motifs, the transition events are here defined in terms of *second neighbourhood motifs*: a central node, its neighbours and the neighbours of those neighbours. In the same way that we note  $\nu\Gamma_1(\mathbf{k})$  the first neighbourhood motif formed by a state  $\nu$  central node with neighbourhood specified by  $\mathbf{k}$ , we note  $\nu\Gamma_2(\mathbf{K})$  the second neighbourhood motif formed by a state  $\nu$  central node with neighbourhood specified by  $\mathbf{K}$ .

The elements of  $\mathbf{K}$  may be indexed with first neighbourhood motifs: the central node

has  $K_{\nu\Gamma(\mathbf{k}' )}$  state  $\nu'$  neighbours whose other neighbours (*i.e.*, excluding the central node) are specified by  $\mathbf{k}'$ . Hence, the second neighbourhood motif



is noted  $S\Gamma_2(\mathbf{K})$  with all elements of  $\mathbf{K}$  zero except for  $K_{S\Gamma(0,0)} = 1$ ,  $K_{I\Gamma(0,1)} = 1$  and  $K_{S\Gamma(1,1)} = 2$ . Note that the central node of this second neighbourhood motif is also the central node of the first neighbourhood motif  $S\Gamma_1(3,1)$ . In general, we note  $\nu\tilde{\Gamma}_1(\mathbf{K})$  the first neighbourhood motif that shares the same central node as the second neighbourhood motif  $\nu\Gamma_2(\mathbf{K})$ .

We digress further to introduce the *unit vector* notation  $\hat{\mathbf{e}}_{\mathcal{M}}$  where  $\mathcal{M}$  represents a motif; all the elements of this vector are zero except for the  $\mathcal{M}$ -th, which is one. The total number of elements in  $\hat{\mathbf{e}}_{\mathcal{M}}$  should be clear from the context. As a concrete example, the right hand side of (7.6b) could be noted  $\hat{\mathbf{e}}_I + (n - 2j)\hat{\mathbf{e}}_{S-I}$ .

Similarly to section 7.3.1, we define the forward transition event of type  $\mathbf{K}$  to be the infection of the central node of a  $S\Gamma_2(\mathbf{K})$  motif and the backward transition event of type  $\mathbf{K}$  as the recovery of the central node of a  $I\Gamma_2(\mathbf{K})$  motif. The corresponding shift vector is

$$\mathbf{r}^{\mathbf{K}} = \hat{\mathbf{e}}_{I\tilde{\Gamma}_1(\mathbf{K})} - \hat{\mathbf{e}}_{S\tilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu} \sum_{\mathbf{k}} K_{\nu\Gamma(\mathbf{k})} (\hat{\mathbf{e}}_{\nu\Gamma(\mathbf{k}+\hat{\mathbf{e}}_I)} - \hat{\mathbf{e}}_{\nu\Gamma(\mathbf{k}+\hat{\mathbf{e}}_S)}) \quad . \quad (7.8b)$$

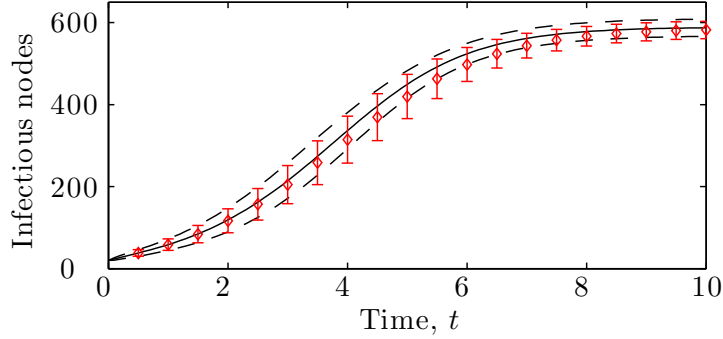
The first line shows the direct effect of a change of state in the central node while the second one handles the ‘‘collateral effect’’ on its immediate neighbours. Here the unit vector  $\hat{\mathbf{e}}_{\nu}$  has the same dimension as  $\mathbf{k}$  (*i.e.*, two) while  $\hat{\mathbf{e}}_{\nu\Gamma(\mathbf{k})}$  has the same dimension as  $\mathbf{x}$ . Sums are taken over all the accessible values of  $\nu$  and  $\mathbf{k}$ .

The corresponding rate functions are

$$q_{\mathbf{K}}^+(\mathbf{x}, Y) = \beta x_{S\tilde{\Gamma}_1(\mathbf{K})} \left( \sum_{\mathbf{k}} K_{I\Gamma(\mathbf{k})} \right) P \left( S\Gamma_2(\mathbf{K}) \middle| S\tilde{\Gamma}_1(\mathbf{K}), \mathbf{x}, Y \right) \quad (7.8c)$$

$$q_{\mathbf{K}}^-(\mathbf{x}, Y) = \alpha x_{I\tilde{\Gamma}_1(\mathbf{K})} P \left( I\Gamma_2(\mathbf{K}) \middle| I\tilde{\Gamma}_1(\mathbf{K}), \mathbf{x}, Y \right) \quad . \quad (7.8d)$$

Note that, unlike (7.6c)–(7.6d) and (7.7c)–(7.7d), the inference terms in (7.8c)–(7.8d) have the same form: the probability for a motif to be a  $\nu\Gamma_2(\mathbf{K})$  knowing (in addition to  $\mathbf{x}$  and  $Y$ ) that its central node is also the central node of a  $\nu\tilde{\Gamma}_1(\mathbf{K})$  motif. Again assuming



**Figure 7.5:** Time evolution of the number of infectious nodes for SIS dynamics ( $\alpha = 1$  and  $\beta = 1$ ) on a CM network for which the number of nodes of each degree is prescribed by the sequence  $\{0, 50, 200, 450, 300\}$  (total  $N = 10^3$  nodes) with 2% of the nodes of each degree initially infectious. The mean and range of one standard deviation above and below the mean are shown. Curves: simplified system. Symbols: full system ( $10^5$  simulations).

a large network size, they are provided by a product of multinomial distributions

$$\begin{aligned}
 P\left(\nu\Gamma_2(\mathbf{K}) \mid \nu\tilde{\Gamma}_1(\mathbf{K}), \mathbf{x}, Y\right) \\
 = \prod_{\nu'} \left(\sum_{\mathbf{k}} K_{\nu'\Gamma_1(\mathbf{k})}\right)! \prod_{\mathbf{k}} \frac{1}{(K_{\nu'\Gamma_1(\mathbf{k})})!} \left(\frac{(k_{\nu} + 1)x_{\nu'\Gamma_1(\mathbf{k} + \hat{\mathbf{e}}_{\nu})}}{\sum_{\mathbf{k}'} k'_{\nu'} x_{\nu'\Gamma_1(\mathbf{k}')}}\right)^{K_{\nu'\Gamma_1(\mathbf{k})}} \quad (7.8e)
 \end{aligned}$$

which complete the rules  $W$  for the first neighbourhood SIS model.

Figure 7.5 compares the results produced by this simplified model and the full one. Note that this is a stochastic version of the model presented in [55], except that the network structure is here static.

### 7.3.3 First neighbourhood SIR model

As in section 7.3.2, we consider a full network model where the network structure is specified solely by the degree of its nodes. However, this time we consider SIR epidemiological dynamics: the accessible node states are  $\nu \in \{S, I, R\}$ , infection is the same as in SIS but recovery is replaced by removal (see the introduction of section 7.3 for details).

We define the forward transition event of type  $\mathcal{IK}$  to be the  $\mathcal{I}$ nfection of the central node of a  $S\Gamma_2(\mathbf{K})$  motif while a forward transition event of type  $\mathcal{RK}$  is the  $\mathcal{R}$ emoval of the central node of a  $I\Gamma_2(\mathbf{K})$  motif. There is no backward transition events. The model

is specified by

$$\mathbf{x} = (\cdots, x_{S\Gamma_1(\mathbf{k})}, \cdots, x_{I\Gamma_1(\mathbf{k})}, \cdots, x_{R\Gamma_1(\mathbf{k})}, \cdots) \quad (7.9a)$$

$$\mathbf{r}^{\mathcal{IK}} = \hat{\mathbf{e}}_{I\tilde{\Gamma}_1(\mathbf{K})} - \hat{\mathbf{e}}_{S\tilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu} \sum_{\mathbf{k}} K_{\nu\Gamma_1(\mathbf{k})} (\hat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\hat{\mathbf{e}}_I)} - \hat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\hat{\mathbf{e}}_S)}) \quad (7.9b)$$

$$\mathbf{r}^{\mathcal{RK}} = \hat{\mathbf{e}}_{R\tilde{\Gamma}_1(\mathbf{K})} - \hat{\mathbf{e}}_{I\tilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu} \sum_{\mathbf{k}} K_{\nu\Gamma_1(\mathbf{k})} (\hat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\hat{\mathbf{e}}_R)} - \hat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\hat{\mathbf{e}}_I)}) \quad (7.9c)$$

$$q_{\mathcal{IK}}^+(\mathbf{x}, Y) = \beta x_{S\tilde{\Gamma}_1(\mathbf{K})} \left( \sum_{\mathbf{k}} K_{I\Gamma_1(\mathbf{k})} \right) P \left( S\Gamma_2(\mathbf{K}) \mid S\tilde{\Gamma}_1(\mathbf{K}), \mathbf{x}, Y \right) \quad (7.9d)$$

$$q_{\mathcal{RK}}^+(\mathbf{x}, Y) = \alpha x_{I\tilde{\Gamma}_1(\mathbf{K})} P \left( I\Gamma_2(\mathbf{K}) \mid I\tilde{\Gamma}_1(\mathbf{K}), \mathbf{x}, Y \right) \quad (7.9e)$$

$$q_{\mathcal{IK}}^-(\mathbf{x}, Y) = q_{\mathcal{RK}}^-(\mathbf{x}, Y) = 0 \quad (7.9f)$$

where the inference terms are the same as in (7.8e).

### 7.3.4 First neighbourhood on-the-fly SIR model

We take a different perspective to the problem considered in section 7.3.3 which requires to track much less elements in the state vector. Instead of considering “complete” first neighbourhood motifs, such as  $\nu\Gamma_1(\mathbf{k})$ , that specify the state of each of the central node’s neighbours, we define the  $\nu\Lambda_1(\kappa)$  motif as a central node of state  $\nu$  for which we know that it has  $\kappa$  neighbours *unknown to us*. This last statement is important: were we to learn the state of one of these neighbours, this would cease to be a  $\nu\Lambda_1(\kappa)$  motif and instead become a  $\nu\Lambda_1(\kappa - 1)$  one. As usual, the state vector tracks the number of such motifs

$$\mathbf{x} = (\cdots, x_{S\Lambda_1(\kappa)}, \cdots, x_{I\Lambda_1(\kappa)}, \cdots, x_{R\Lambda_1(\kappa)}, \cdots) . \quad (7.10a)$$

We recall from section 7.3.2 how a CM network is built in a computer simulation: for each  $\kappa$ ,  $n_{\kappa}$  nodes with  $\kappa$  stubs are created and the stubs are then randomly paired to form links. From this perspective,  $\nu\Lambda_1(\kappa)$  may be reinterpreted as a  $\nu$  state node with  $\kappa$  unpaired stubs: as stubs are removed once they are paired in the computer simulation, neighbours that were unknown are removed from these motifs once they become known to us. Hence,

$$\frac{\kappa x_{\nu\Lambda_1(\kappa)} - \delta_{\nu\nu'} \delta_{\kappa\kappa'}}{\sum_{\nu''} \sum_{\kappa''} \kappa'' x_{\nu''\Lambda_1(\kappa'')} - 1}$$

*exactly* gives the probability for an unknown neighbours of the central node of  $\nu'\Lambda_1(\kappa')$  to be the central node of  $\nu\Lambda_1(\kappa)$ . Note that the Kronecker deltas ( $\delta_{ii'} = \begin{cases} 1 & i=i' \\ 0 & i \neq i' \end{cases}$ ) in the numerator and the  $-1$  in the denominator both account for the stub of  $\nu'\Lambda_1(\kappa')$  that we are pairing with a random stub.



A typical computer simulation would first build the network and then perform the SIR propagation dynamics on this network. However, we do not want to have to store the network structure for later consultation, which would require additional space in  $\mathbf{x}$ . Instead, we delay the network construction, leaving the stubs unpaired, and start the propagation dynamics right away. Just when the state of an unknown neighbour is required do we pair the corresponding stub with a randomly selected one, hence building the network *on-the-fly*. Since the knowledge of stubs being matched will be lost in the future, this information must only be required at the very moment it is obtained if we want the resulting dynamics to *exactly* reproduce the behaviour of the full system.

We thus take a different, although equivalent, perspective on the infection dynamics where each link is “probed” at most once. Instead of considering a probability  $\beta dt$  of infection for each *susceptible* neighbours of infectious nodes, we consider the same probability for each of their *unknown* neighbours. Only when this probability returns true do we wonder about the state of the neighbour, whose state changes to infectious if and only if it was previously susceptible. In any case, we learned who were the neighbours of two nodes (*i.e.*, the infectious and its neighbour) and we must update the state vector accordingly.

Hence, we define the  $\mathcal{I}$ nfection transition event  $\mathcal{I}\nu\kappa\kappa'$  such that an infectious at the center of a  $I\Lambda_1(\kappa')$  motif attempts to infect the  $\nu$ -state node at the center of a  $\nu\Lambda_1(\kappa)$  motif. Of course, only  $\mathcal{I}S\kappa\kappa'$  transition events result in real infections. The more traditional transition event  $\mathcal{R}\kappa$  corresponds to the  $\mathcal{R}$ emoval of the infectious node at the center of a  $I\Lambda_1(\kappa)$  motif, thus becoming  $R\Lambda_1(\kappa)$ . The model is specified by

$$\mathbf{r}^{\mathcal{I}S\kappa\kappa'} = \widehat{\mathbf{e}}_{I\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{S\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \quad (7.10b)$$

$$\mathbf{r}^{\mathcal{I}I\kappa\kappa'} = \widehat{\mathbf{e}}_{I\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \quad (7.10c)$$

$$\mathbf{r}^{\mathcal{I}R\kappa\kappa'} = \widehat{\mathbf{e}}_{R\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{R\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \quad (7.10d)$$

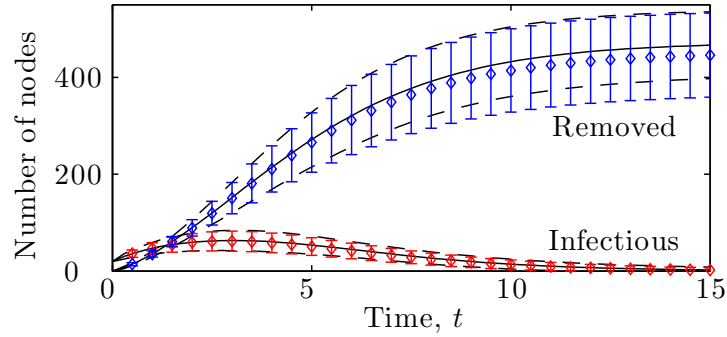
$$\mathbf{r}^{\mathcal{R}\kappa} = \widehat{\mathbf{e}}_{R\Lambda_1(\kappa)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa)} \quad (7.10e)$$

$$q_{\mathcal{I}\nu\kappa\kappa'}^+(\mathbf{x}, Y) = \beta \kappa' x_{I\Lambda_1(\kappa')} \frac{\kappa x_{\nu\Lambda_1(\kappa)} - \delta_{I\nu} \delta_{\kappa\kappa'}}{\sum_{\nu''} \sum_{\kappa''} \kappa'' x_{\nu''\Lambda_1(\kappa'')} - 1} \quad (7.10f)$$

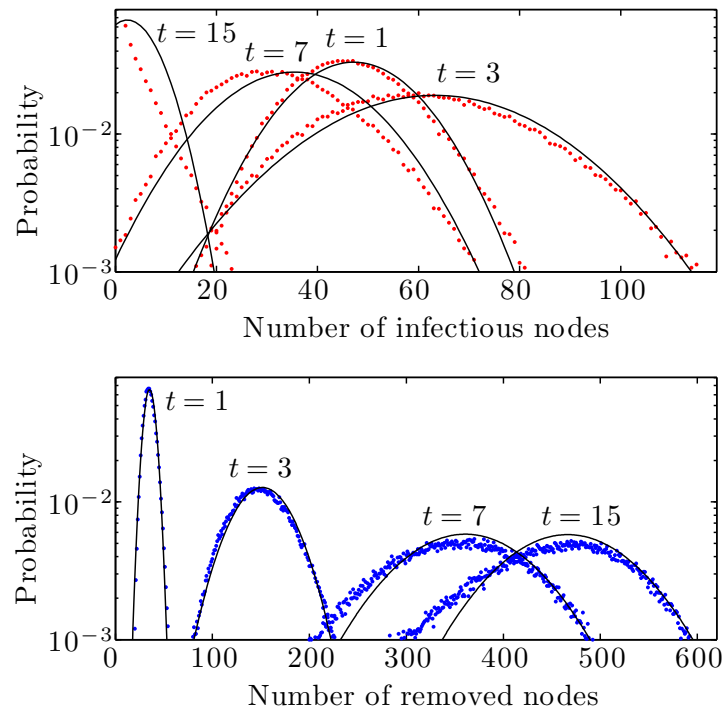
$$q_{\mathcal{R}\kappa}^+(\mathbf{x}, Y) = \alpha x_{I\Lambda_1(\kappa)} \quad (7.10g)$$

$$q_{\mathcal{I}\nu\kappa\kappa'}^-(\mathbf{x}, Y) = q_{\mathcal{R}\kappa}^-(\mathbf{x}, Y) = 0 \quad . \quad (7.10h)$$

The system (7.10) *exactly* reproduces the behaviour of the full system through the solution of (7.2). Since (7.3)–(7.4) are only approximations of (7.2), results obtained through these relationships are only approximative (figure 7.6 and figure 7.7). This model may be solved analytically for the mean value (see Appendix 7.B) and the results



**Figure 7.6:** Time evolution of the number of infectious and removed nodes for SIR dynamics ( $\alpha = 1$  and  $\beta = 1$ ) on a CM network for which the number of nodes of each degree is prescribed by the sequence  $\{0, 50, 200, 450, 300\}$  (total  $N = 10^3$  nodes) with 2% of the nodes of each degree initially infectious (all others are susceptible). The mean and range of one standard deviation above and below the mean are shown. Curves: simplified system. Symbols: full system ( $10^5$  simulations).



**Figure 7.7:** Probability distribution at different times for the number of infectious and removed nodes. The parameters are the same as in figure 7.6. Curves: simplified system. Symbols: full system ( $10^5$  simulations).

are in agreement with [61, 99]. This is a generalization to the case  $\alpha \neq 0$  of the model presented in [76]. Further details are discussed in section 7.4.4. We note that a conceptually similar approach has recently been developed independently [19] as a tool for a mathematically rigorous proof that a specific heterogeneous mean field model [99] holds in the limit of large network size.

## 7.4 Discussion

We now take a retrospective look at the results presented in section 7.3 and obtain from these special examples general considerations concerning our modelling approach.

### 7.4.1 Accuracy of the results

One of the aims of this paper is to obtain simplified models that accurately reproduce the behaviour of complex systems. Since approximations are usually involved, it is to be expected that the results of the simplified model only agree with those of the full system over some range of parameters, where the approximations were valid.

The parameters used in figure 7.2–7.7 were chosen in order to investigate the limits of our approximations: while there is no perfect correspondence between the results of the full and simplified systems, their agreement is probably sufficient for both qualitative and quantitative applications. We distinguish between two categories of approximations: those inherent to the use of (7.3)–(7.4) and those due to the imperfect representation of  $Z$  through  $\mathbf{x}$  and  $Y$ .

#### Gaussian approximation

Since (7.2) and (7.10) define a system that *exactly* reproduces the behaviour of the corresponding full system, any discrepancy in figure 7.6 must originate from the use of the Gaussian approximation (7.3)–(7.4). An important requirement for this approximation to be valid is that the size  $N$  of the system must be large.

Figures 7.2–7.7 all use networks of size  $N = 1000$ . As a rule of thumb, we found that (7.3)–(7.4) perform better for networks of at least a few hundred nodes, which is

the case of many relevant real-world systems. Note that, for very small systems (tens of nodes), one could also directly and completely solve (7.2).

While a large network size  $N$  is required to justify treating the elements of  $\mathbf{x}$  as real numbers, other phenomena may affect the validity of this approximation. For example, when the initial conditions are such that there is a single infectious node, the continuous approximation fails at considering the probability for that node to recover (or to be removed) before transmitting the infection to one of its neighbours. Figures 7.2–7.7 circumvent this problem by using an initial condition with 20 infectious nodes: the probability for all of them to recover (or to be removed) before transmitting the infection is very low.

It is worth noting that the plateaux seen on figure 7.2 and 7.4–7.6 reflect different dynamical behaviours for the SIS and SIR systems. Indeed, while the total number of removed nodes reaches a maximum in the SIR system because there are no infectious left to recover, the steady state observed at the later times for our SIS models corresponds to a constant flow of recovery and new infections. In the former case, the approximation errors performed at earlier times accumulate. In the later, the exact path taken to attain equilibrium is of lesser importance and errors do not accumulate the same way.

## Representation approximation

In general, the simplified system will not exactly reproduce the behaviour of the full system, even when using (7.2) instead of (7.3)–(7.4). This is the case of all our SIS models; while some of the discrepancy seen in figure 7.2–7.5 is explained by the Gaussian approximation, the imperfect representation of  $Z$  also contributes to the error.

Part of the problem can be understood as our failure to consider the correlation between the neighbours of a node and the time elapsed since this node has been in its present intrinsic state. For example, the neighbours of a susceptible node that has just recovered (*i.e.*, it was infectious a moment ago) may be much different than those of a susceptible node that has recovered a long time ago, while being similar to those of a node that is still infectious. Hence, one could hope to improve these SIS models through changes in  $Y$  alone (*i.e.*, with the same  $\mathbf{x}$ ): first estimate the probability distribution for the time since when each node has last changed state and then infer the neighbourhoods accordingly. An alternative that could be simpler to implement, at the cost of increasing the size of  $\mathbf{x}$ , would consist in tracking more exhaustive motifs (*e.g.*, second neighbourhoods instead of first ones in section 7.3.2).

However, there are more intricate consequences to the recovery of infectious nodes on a structure that is fixed in time: if at some point all the nodes of the same component (*i.e.*, a connected subnetwork that is disconnected from the rest of the network) are susceptible at the same time, then none of them may ever become infectious again. The connectivity of a network is strongly affected by the average degree of its nodes: our parameters correspond to an average degree of 5 for figure 7.2–7.4 (average degree of a neighbour also 5) and of 3 for figure 7.5 (average degree of a neighbour  $\approx 3.23$ ). When using smaller parameter values, this components-induced discrepancy becomes much larger since the simplified model then overestimates the number of infectious nodes. One could take the components into account by solving independent systems for each component (and merge the results afterwards) or by a clever adaptation of the inference process (see section 7.4.6 for possible directions). Note that these effects are usually much less important when the network structure changes over time.

## 7.4.2 Pair-based models

Compared to the other models presented in section 7.3, the two pair-based models of section 7.3.1 use very small state vectors (*i.e.*, two or three elements). This is an important advantage of pair-based models in general: there are usually much less pair and node motifs than, *e.g.*, first neighbourhood ones, and tracking them thus requires much smaller  $\mathbf{x}$ .

Although we limited our study of pair-based models to regular and Erdős-Rényi networks, more complex network structures could also be considered. In the same way that (7.6) and (7.7) differ mostly by their inference terms, obtaining good inference from the little information stored in  $\mathbf{x}$  is probably the principal challenge behind general and accurate pair-based stochastic models.

However, non-stochastic pair-based models are already possible on nontrivial network structures for SIR dynamics or, more generally, for processes such that a change in the state of one neighbour of a node can be treated as independent of that of another neighbour (SIS fails this assumption) [62]. Knowing (in  $Y$ ) that a system behaves in this manner greatly simplifies the inference process, and this is the main reason for the success of the SIR pair-based model for the evolution of mean values on CM networks that is presented in [61, 99]. Whether or not the same approach may be used to obtain stochastic results is an open question.

### 7.4.3 First (and higher) neighbourhood models

By opposition, sufficiently accurate inference terms for first neighbourhood models are often straightforward to obtain. Although (7.8e) may be difficult to appreciate at first sight, it is the only inference term used in both section 7.3.2 and section 7.3.3. In fact, (7.8e) may well be the only inference term needed for generic first neighbourhood models for CM network structures.

Although first neighbourhood motifs are a “natural language” for expressing dynamics taking place on CM networks, they could also be used in the presence of other complex structures. This may be done through changes in  $\mathbf{x}$  and/or  $Y$ ; see Appendix 7.A for details.

The generality and ease of design of first-neighbourhood models comes at a cost: the state vector  $\mathbf{x}$  is typically much larger than it would be in an equivalent pair-model. How large is  $\mathbf{x}$  strongly depends on the maximal node degree present in the network and on the total number of accessible intrinsic node states (see Appendix 7.A for details). For typical values of these quantities, this does not cause major problems for the evaluation of the mean: numerically solving (7.3a) requires an acceptable amount of resources even for an  $\mathbf{x}$  of dimension  $10^6$  and (7.3b) may often be simplified (*i.e.*, summed analytically).

However, evaluating the covariance matrix using (7.4) may cause problems: unless analytical simplifications are possible, solving this system scales as the square of the number of elements in  $\mathbf{x}$ . Future developments may decrease this bottleneck effect of the covariance matrix; see section 7.4.5 for details. In any case, the size of  $\mathbf{x}$  may be decreased by “coarse graining” the number of links between the central node and its neighbours; see Appendix 7.A.7 for details.

### 7.4.4 On-the-fly models

The on-the-fly model presented in section 7.3.4 for SIR dynamics on CM networks *exactly* reproduces the behaviour of the full system. This is even more remarkable when one considers that the size of the state vector in the on-the-fly model is much smaller than in the alternative first neighbourhood model of section 7.3.3. The reasons behind the success of the on-the-fly approach are similar to those discussed in section 7.4.2 for the pair models presented in [61, 62, 99]: it is encoded in  $Y$  that, for each link, we *at most once* need to simultaneously know the state of the two nodes joined by that link [76].

The inference term (7.8e) is of “general purpose” in the sense that its  $Y$  does not provide information on the dynamical properties of the system, but only on how the motifs in  $\mathbf{x}$  may be interconnected. This is why both (7.8) and (7.9) rely on (7.8e).

However, the inference terms of (7.10) have a specific character:  $Y$  contains information about (7.10) itself. Any change to the dynamics implies changes in the inference terms, with no guarantee that an acceptable solution exists. In fact, (7.10) was *designed* with this problem in mind. In other words, we obtained a simple and reliable model at the cost of “pre-computations” in the design process. Of all the possibilities in model-space, the information acquired by pointing at this specific one is what replaces the reduced size of the state vector. The same could be said of the deterministic SIR pair-based model on CM networks presented in [61, 99].

By contrast with the case discussed in section 7.4.3, the small size of the state vector here allows for evaluations of the covariance matrix through (7.4), even when relatively high degree nodes are present. Alternatively, one may take advantage of the fact that, even for more complicated dynamics, the state vector of on-the-fly models can remain of manageable size for mean values calculations; see the introduction to Appendix 7.A for the concrete example of [54].

### 7.4.5 Complicated states *vs* complex assumptions

Section 7.4.4 revealed an unexpected depth to  $Y$ : one may achieve models of similar levels of accuracy by trading off complexity in the assumptions for a reduction in the size of the state vector  $\mathbf{x}$ . As an extreme example, if  $Y$  already gives the full behaviour of the system, then there is no need for tracking any information in  $\mathbf{x}$ . Without reaching such extremes, our on-the-fly model and the deterministic SIR pair-based model presented in [61, 99] both demonstrate the benefits of investing some time in the assumptions of our models.

While these examples required case-by-case analysis, one may benefit from the same realization in a general context: a first simplified model ( $W$ ,  $\mathbf{x}$  and  $Y$ ) may generate the assumptions  $Y'$  to a different simplified model ( $W'$ ,  $\mathbf{x}'$  and  $Y'$ ). For example, when some dynamical process (*e.g.*, SIS or SIR) occurs on a network whose structure changes in time independently from this dynamics, one could obtain a first model for the structure alone and then feed the results to the second model, handling the remaining dynamics. Even more generally, one could compensate for the higher computational requirements of (7.4) by first solving (7.3) on an elaborate model then feeding the resulting mean values to a simpler model for the sole purpose of estimating the covariance matrix.

### 7.4.6 Additional inference tools

While we introduced  $Y$  as a direct application of Bayes' rule, we have now seen that useful assumptions may be obtained by other means, including the solution of another system of the form (7.2). The next step in this direction would be to improve our inference process using alternative tools and models available to network science.

For example, branching processes [75] may be used to infer information concerning the connectivity and the components of the network structure. As discussed in section 7.4.1, this point was a major shortfall of SIS models. This approach is even more interesting for the recently developed tools [4, 44] that are particularly compatible with the motifs and intrinsic node state approach presented in this paper.

Another tool of considerable interest are exponential random networks [81]. Indeed, these maximum entropy methods can simplify inferences that would have otherwise been prohibitively complex. Once again, this approach may be generalized to different kind of motifs and intrinsic node states.

## 7.5 Conclusion: general applicability

Although the examples of section 7.3 focus on simple SIS and SIR dynamics, any specificity that could be modelled through a standard epidemiological compartmental model may *a priori* be considered by our approach: genders, age groups, vaccination, incubation period, disease phases, *etc.* Furthermore, population dynamics considerations may be accounted for in a straightforward manner. Assuming first neighbourhood motifs, births and deaths of individuals correspond to events adding and removing motifs, respectively. Similarly, changes in interaction patterns amount to events replacing the affected motifs by new ones. In fact, from the model's perspective, there is no important distinction between a change in the interaction structure of the population and a change in the node states: both are events affecting motifs.

Beyond the propagation of infections or parasites, an additional class of spreading processes is particularly relevant to population biology: the cascading extinctions of species in food webs [10, 23, 88]. Our formalism is applicable to such problems by representing each species as a node and by using links to indicate feeding connections between species. While such cascades may require different rules than those of SIS or SIR dynamics, the general approach may still be adapted to this specific application.



This illustrates how our philosophy has the potential to be applicable to any dynamics as long as the relevant information can be encoded through the structure of a network and the intrinsic state of its components.

The generality of our systematic approach and the fact that its assumptions are explicitly stated suggests that it could be used as a common ground for comparing existing models too complex for direct comparison. Indeed, by considering such an existing model as the full system (specified by  $V_1$  and  $Z_1$ ), one may seek a simplified system (specified by  $W_1$ ,  $X_1$  and  $Y_1$ ) approximately reproducing the original model (over a sufficient range of parameters).

If some transition event (in  $W_1$ ) appears essential, this may reveal an important feature of the original model; the same holds true for motifs (in  $X_1$ ) and prior knowledge (in  $Y_1$ ). Moreover, assuming that this procedure has been done for a second existing model (specified by  $V_2$  and  $Z_2$ ), one may directly compare their simplified version in a common framework, which will help identify the assumptions required for their description. Note that this perspective is similar to a commutation diagram

$$\begin{array}{ccc}
 & \text{Difficult to compare} & \\
 (V_1, Z_1) & \not\leftrightarrow & (V_2, Z_2) \\
 \text{Equivalent behaviour} \updownarrow & & \updownarrow \text{Equivalent behaviour} \\
 (W_1, X_1, Y_1) & \longleftrightarrow & (W_2, X_2, Y_2) \quad .
 \end{array}$$

For example, if  $X_1 = X_2$  and  $Y_1 = Y_2$ , we know that the discrepancies between the two original models is imputable to the difference in the transition events. Finding a minimal set of changes to  $W_1$  and/or  $W_2$  causing both models to agree may then help identify the very cause of the discrepancies.

## 7.A More on motifs

We introduce new motifs and generalize those presented in the main text for intrinsic link types and directed (or semi-directed) networks; Table 7.1 summarizes the total number of motifs in each of these classes. We define  $\mathcal{K}$  as the highest accessible node degree and we set  $\mathcal{D} = 1$  for undirected,  $\mathcal{D} = 2$  for directed and  $\mathcal{D} = 3$  for semi-directed networks.

Of high practical importance is the fact that the entries of this table differ widely in their scaling behaviours. For example, the on-the-fly model [54] for two interacting SIR dynamics each propagating on their own network structure uses  $\mathcal{N} = 9$  (Cartesian

**Table 7.1:** Number of motifs in selected classes.

Motifs	Number
Node	$\mathcal{N}$
Pair	$\frac{1}{2}\mathcal{L}\mathcal{N}(\mathcal{D}\mathcal{N} + (\mathcal{D} - 2)^2)$
Triple	$\frac{1}{2}\mathcal{D}\mathcal{L}\mathcal{N}^2(\mathcal{D}\mathcal{L}\mathcal{N} + 1)$
Triangle	$\frac{1}{3}\mathcal{D}\mathcal{L}\mathcal{N} + \frac{1}{2}\mathcal{D}((\mathcal{D} - 2)\mathcal{L}\mathcal{N})^2 + \frac{1}{6}(\mathcal{D}\mathcal{L}\mathcal{N})^3$
Degreed node	$\mathcal{N} \binom{\mathcal{K} + \mathcal{D}\mathcal{L}}{\mathcal{K}}$
Degreed pair	$\frac{\mathcal{L}\mathcal{N}}{2} \binom{\mathcal{K} - 1 + \mathcal{D}\mathcal{L}}{\mathcal{K} - 1} \left( \mathcal{D}\mathcal{N} \binom{\mathcal{K} - 1 + \mathcal{D}\mathcal{L}}{\mathcal{K} - 1} + (\mathcal{D} - 2)^2 \right)$
First neighbourhood (node)	$\mathcal{N} \binom{\mathcal{K} + \mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K}}$
First neighbourhood pair	$\frac{\mathcal{L}\mathcal{N}}{2} \binom{\mathcal{K} - 1 + \mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K} - 1} \left( \mathcal{D}\mathcal{N} \binom{\mathcal{K} - 1 + \mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K} - 1} + (\mathcal{D} - 2)^2 \right)$
Second neighbourhood (node)	$\mathcal{N} \binom{\mathcal{K} + \mathcal{D}\mathcal{L}\mathcal{N} \binom{\mathcal{K} - 1 + \mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K} - 1}}{\mathcal{K}}$

product of the sub-states of two SIR dynamics),  $\mathcal{L} = 3$  (links of the first network alone, links of the second network alone and overlapping links) and  $\mathcal{D} = 1$  (undirected network). Looking up in table 7.1, we see that this requires of the order of  $\mathcal{K}^3$  on-the-fly (degreed node) motifs, where  $\mathcal{K}$  denotes the highest accessible node degree. By opposition, implementing a first neighbourhood version of this model would require of the order of  $\mathcal{K}^{27}$  motifs!

### 7.A.1 Pair motifs

We note  $\nu \overset{\ell}{-} \nu'$  (resp.  $\nu \overset{\ell}{\rightarrow} \nu'$ ) an undirected (resp. directed) pair motif formed of a state  $\nu$  node linked through a state  $\ell$  link to a state  $\nu'$  node. In the case of directed motifs, the direction of the arrow usually specifies the “strongest causal effect” of this asymmetric interaction (although this needs not be the case). All the undirected and directed motifs are possible in a semi-directed network. We may omit the index  $\ell$  over the links when  $\mathcal{L} = 1$ .

### 7.A.2 Triple motifs

We note  $\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''$  an undirected triple motif formed of a state  $\nu$  node linked through a state  $\ell$  link to a state  $\nu'$  node itself linked through a state  $\ell'$  link to a state  $\nu''$  node. As for pair motifs, each of these nodes may have other neighbours than those that are explicitly specified. The notation directly generalizes to directed (*e.g.*,  $\nu \overset{\ell}{\rightarrow} \nu' \overset{\ell'}{\leftarrow} \nu''$ ) and semi-directed (*e.g.*,  $\nu \overset{\ell}{-} \nu' \overset{\ell'}{\rightarrow} \nu''$ ) triple motifs.

The term “2-star” is often used to refer to a triple motif for which both extremities (*e.g.*, the nodes of state  $\nu$  and  $\nu''$  in the motif  $\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''$ ) are explicitly *forbidden* to be neighbours. In models that also use triangle motifs, 2-star motifs may explicit the absence of the last link that would form a triangle. Another common use of triple motifs comes in the inference process of (usually deterministic) pair based models.

### 7.A.3 Triangle motifs and other small subnetworks

Three nodes that are all neighbours of each other form a triangle motif. An horizontal bracket represents the additional link that would be missing in the analogous triple motif, *e.g.*,

$$\underbrace{\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''}_{\ell''}$$

for an undirected network.

Triangle motifs are usually considered in models that should account for clustering. Their number may either directly be tracked in the state vector  $\mathbf{x}$  [40] or their implicit presence (stated in  $Y$ , *e.g.*, through a clustering coefficient) may be accounted for in the inference process [48].

The same notation may be generalized to other motifs consisting of small subnetworks, *e.g.*, square motifs [40].

### 7.A.4 Clique motifs

A vague definition of a clique motif is “a subgroup of nodes that share more links among themselves than what could be expected otherwise for the same number of randomly selected nodes”. In applications, one may refine this definition according

to the specificities of the problem at hand, *e.g.*, “an Erdős-Rényi subnetwork (link probability  $p$ ) of  $n_S$  susceptible nodes and  $n_I$  infectious nodes”. Clique motifs are usually considered in models that should account for community structure [37].

### 7.A.5 Degreed motifs

A degreed motif is a motif for which we know the degree of all the nodes forming the motif: a degreed node motif is a node of specified state and degree; a degreed link motif is two nodes of specified state and degree that are known to be neighbours; *etc.* The in-degree and out-degree are both specified in directed and semi-directed networks; the latter cases also specify undirected degrees. Likewise, degrees pertaining to different types of links are specified independently.

In the same way that pair motifs are usually combined with node motifs, degreed pair motifs are usually combined with degreed node motifs [26]. Note that the on-the-fly motifs  $\nu\Lambda_1(\kappa)$  presented section 7.3.4 can be understood as a special case of degreed node motifs where the degree is replaced by the “degree to unknown nodes”.

### 7.A.6 $n$ -th neighbourhood motifs

Similarly to degreed motifs, an  $n$ -th neighbourhood motif is a motif for which we specify the state of all the  $n$ -th neighbours of the nodes forming the original motif. Hence, the notation  $\nu\Gamma_1(\mathbf{k})$  [resp.  $\nu\Gamma_2(\mathbf{K})$ ] of the main text corresponds to a first (resp. second) neighbourhood *node* motif. These concepts are directly generalizable to types of links and to directed or semi-directed networks.

First neighbourhood node motifs can be understood as tracking the correlation between the state of a node and the state of all its neighbours. By opposition, degreed pair motifs track the correlation between the state and degree of two neighbouring nodes. While similar information may be obtained from both motif classes, a model based on one may perform better than a model based on the other depending on the characteristics of the full system to be modelled.

### 7.A.7 Coarse-grained degree and/or neighbourhood

Not all entries of table 7.1 depend on the maximal degree  $\mathcal{K}$ , but those that do quickly increase for large  $\mathcal{K}$ . This is problematic since many real-world systems contain high degree nodes.

However, one may overcome this limitation by coarse-graining degrees through ranges: the range containing the degree of a node is specified instead of the degree itself. For example, given the ranges

$$\underbrace{[0, 0]}_{\text{range 0}}, \underbrace{[1, 1]}_{\text{range 1}}, \underbrace{[2, 3]}_{\text{range 2}}, \underbrace{[4, 7]}_{\text{range 3}}, \underbrace{[8, 15]}_{\text{range 4}}, \underbrace{[16, 31]}_{\text{range 5}} \text{ and } \underbrace{[32, 63]}_{\text{range 6}},$$

we would say of a degree 23 node that its degree lies within range 5. Hence, for the purpose of evaluating the number of motifs in table 7.1, one should here use  $\mathcal{K} = 6$  (one less than the total number of ranges) instead of  $\mathcal{K} = 63$  (highest representable degree).

While the previous example used powers of 2 for simplicity, a slower increase is probably desirable in most applications. However, since the number of neighbourhood and degreed motifs strongly depends on  $\mathcal{K}$ , even the slightest reduction in this number may be significant. Note that this coarse-graining method is of particular interest when the real-world data used to calibrate the model is already coarse-grained, which is commonly the case for census data.

## 7.B Deterministic solution of on-the-fly SIR model

This section provides the deterministic solution of (7.10). We first rewrite (7.3) for the specific case of (7.10)

$$\frac{d\boldsymbol{\mu}(t)}{dt} = \sum_{\nu} \sum_{\kappa} \sum_{\kappa'} \mathbf{r}^{\mathcal{I}\nu\kappa\kappa'} q_{\mathcal{I}\nu\kappa\kappa'}^+(\boldsymbol{\mu}(t), Y) + \sum_{\kappa} \mathbf{r}^{\mathcal{R}\kappa} q_{\mathcal{R}\kappa}^+(\boldsymbol{\mu}(t), Y) \quad (7.11)$$

then collect the contributions to  $\mu_{S\Lambda_1(\kappa)}$ ,  $\mu_{I\Lambda_1(\kappa)}$  and  $\mu_{R\Lambda_1(\kappa)}$  (dropping the functional dependencies for brevity)

$$\frac{d\mu_{S\Lambda_1(\kappa)}}{dt} = - \sum_{\kappa'} q_{IS\kappa\kappa'}^+ \quad (7.12a)$$

$$\frac{d\mu_{I\Lambda_1(\kappa)}}{dt} = \sum_{\kappa'} \left( q_{IS(\kappa+1)\kappa'}^+ + q_{II(\kappa+1)\kappa'}^+ - q_{II\kappa\kappa'}^+ \right) + \sum_{\nu} \sum_{\kappa'} \left( q_{I\nu\kappa'(\kappa+1)}^+ - q_{II\kappa'\kappa}^+ \right) - q_{R\kappa}^+ \quad (7.12b)$$

$$\frac{d\mu_{R\Lambda_1(\kappa)}}{dt} = \sum_{\kappa'} \left( q_{IR(\kappa+1)\kappa'}^+ - q_{IR\kappa\kappa'}^+ \right) + q_{R\kappa}^+ \quad (7.12c)$$

Using the definitions

$$\lambda = \sum_{\kappa} \kappa \mu_{I\Lambda_1(\kappa)} \quad \text{and} \quad \omega = \sum_{\nu} \sum_{\kappa} \kappa \mu_{\nu\Lambda_1(\kappa)} \quad (7.13)$$

where  $\lambda$  is the total number of stubs belonging to infectious nodes and  $\omega$  is the total number of stubs in the system, (7.12) becomes

$$\frac{d\mu_{S\Lambda_1(\kappa)}}{dt} = - \frac{\beta \lambda \kappa \mu_{S\Lambda_1(\kappa)}}{\omega} \quad (7.14a)$$

$$\frac{d\mu_{I\Lambda_1(\kappa)}}{dt} = \frac{\beta \lambda (\kappa + 1) \mu_{S\Lambda_1(\kappa+1)}}{\omega} - \alpha \mu_{I\Lambda_1(\kappa)} + \beta \left( 1 + \frac{\lambda}{\omega} \right) \left( (\kappa + 1) \mu_{I\Lambda_1(\kappa+1)} - \kappa \mu_{I\Lambda_1(\kappa)} \right) \quad (7.14b)$$

$$\frac{d\mu_{R\Lambda_1(\kappa)}}{dt} = \frac{\beta \lambda}{\omega} \left( (\kappa + 1) \mu_{R\Lambda_1(\kappa+1)} - \kappa \mu_{R\Lambda_1(\kappa)} \right) + \alpha \mu_{I\Lambda_1(\kappa)} \quad (7.14c)$$

Note that (7.10f) has been approximated by dropping the Kronecker delta in the numerator and the  $-1$  in the denominator.

We now consider the evolution of the total number of stubs  $\omega$  by summing the contributions from (7.14)

$$\frac{d\omega}{dt} = \sum_{\nu} \sum_{\kappa} \kappa \frac{d\mu_{\nu\Lambda_1(\kappa)}}{dt} = -2\beta \lambda \quad (7.15)$$

One may understand (7.15) as “during the time interval  $[t, t + dt)$ , each one of the  $\lambda$  stubs belonging to infectious nodes have probability  $\beta dt$  to be paired to another stub, thus causing a decrease by 2 of  $\omega$ . Noting  $\omega_0 = \omega(0)$  the total number of stubs in the initial condition, we introduce the change of variable

$$\theta = \sqrt{\frac{\omega}{\omega_0}} \quad \text{such that} \quad \frac{d\theta}{dt} = -\frac{\beta \lambda}{\theta \omega_0} \quad (7.16)$$

Notice that  $t = 0$  corresponds to  $\theta = 1$  and that  $\theta$  decreases with time. Using this change of variable in (7.14a) gives

$$\frac{d\mu_{S\Lambda_1(\kappa)}}{d\theta} = \frac{\kappa \mu_{S\Lambda_1(\kappa)}}{\theta} \quad (7.17)$$

which has the solution

$$\mu_{S\Lambda_1(\kappa)}(t) = x_{S\Lambda_1(\kappa)}(0) (\theta(t))^\kappa . \quad (7.18)$$

using the initial condition  $\mu_{S\Lambda_1(\kappa)}(0) = x_{S\Lambda_1(\kappa)}(0)$ .

For convenience, we define

$$f(\theta) = \sum_{\kappa} x_{S\Lambda_1(\kappa)}(0) \theta^\kappa . \quad (7.19)$$

Noticing that

$$\sum_{\kappa} \kappa(\kappa + 1) \mu_{S\Lambda_1(\kappa+1)} = \theta^2 \sum_{\kappa} (\kappa - 1) \kappa x_{S\Lambda_1(\kappa)}(0) \theta^{\kappa-2} = \theta^2 f''(\theta) , \quad (7.20)$$

we obtain the evolution of the total number  $\lambda$  of stubs belonging to infectious nodes by summing the contributions (7.14b)

$$\frac{d\lambda}{dt} = \sum_{\kappa} \kappa \frac{d\mu_{I\Lambda_1(\kappa)}}{dt} = \frac{\beta \lambda \theta^2 f''(\theta)}{\omega} - \beta \lambda \left(1 + \frac{\lambda}{\omega}\right) - \alpha \lambda . \quad (7.21)$$

Again using the change of variable (7.16), we get

$$\frac{d\lambda}{d\theta} = \frac{\lambda}{\theta} + \theta \omega_0 \left(1 + \frac{\alpha}{\beta}\right) - \theta f''(\theta) \quad (7.22)$$

which has the solution

$$\lambda = \theta^2 \omega_0 \left(1 + \frac{\alpha}{\beta}\right) - \theta \omega_0 \frac{\alpha}{\beta} - \theta f'(\theta) \quad (7.23)$$

for an initial condition without removed nodes [*i.e.*,  $\lambda(0) = \omega_0 - f'(1)$ ].

Using this solution in (7.16) gives

$$\frac{d\theta}{dt} = -\beta \theta + \alpha (1 - \theta) + \beta \frac{f'(\theta)}{\omega_0} \quad (7.24)$$

whose solution provides  $\theta(t)$ . Using

$$S(t) = f(\theta(t)) \quad (7.25a)$$

$$I(t) = N - S(t) - R(t) \quad (7.25b)$$

$$\frac{dR(t)}{dt} = \alpha I(t) , \quad (7.25c)$$

we finally obtain the total number  $S = \sum_{\kappa} \mu_{S\Lambda_1(\kappa)}$  of susceptibles,  $I = \sum_{\kappa} \mu_{I\Lambda_1(\kappa)}$  of infectious and  $R = \sum_{\kappa} \mu_{R\Lambda_1(\kappa)}$  of removed nodes at any given time  $t$ . A direct application of (7.18)–(7.19) provides (7.25a), conservation of the nodes provides (7.25b) and using the definitions of  $I(t)$  and  $R(t)$  in (7.14c) provides (7.25c). Although obtained differently, this solution corresponds to that of the pair-based SIR model presented in [99] and in [61].





# Chapitre 8

## Modélisation alternative et solution directe

Ce chapitre présente des discussions complémentaires aux chapitres 6 et 7. Dans un premier temps, on explore certains outils de modélisation particulièrement bien adaptés pour des systèmes dont l'état est partiellement décrit par les motifs le composant. On note deux usages importants de ces méthodes : fournir une perspective alternative ou complémentaire au processus de Markov simplifié (tel que fait en section 6.3.2) et obtenir des inférences qui sont difficiles à produire autrement (tel que mentionné en section 7.4.6). La section 8.1 explore en plus de détails l'utilisation des processus de branchements dans ce contexte alors que la section 8.2 préfère plutôt une perspective de maximisation d'entropie.

Sur une note différente, la section 8.3 présente comment les processus markoviens simplifiés obtenus aux chapitres 6 et 7 peuvent être directement et exactement résolus. Cependant, même si la forme générale des expressions analytiques est connue, obtenir les coefficients  $y$  apparaissant fait typiquement intervenir une équation de récurrence pour laquelle une solution fermée est inconnue. On doit alors se résigner à de petits systèmes et à des calculs numériques ; la partie III de cette thèse explore des méthodes alternatives dans le but de lever ces contraintes.

## 8.1 Processus de branchement

Les processus de branchement s'avèrent être un outil puissant dans l'étude des composantes d'un réseau et des quantités qui leurs sont rattachées. Deux raisons principales nous incitent à adapter ces techniques dans le contexte où un processus markovien simplifié est obtenu à l'aide d'une représentation par motifs : soit on cherche une méthode de solution alternative ou complémentaire pour la dynamique du système (comme c'est le cas en section 6.3.2), soit on cherche à effectuer des inférences qui résistent aux approches plus standard (tel que discuté en section 7.4.6). Dans les deux cas, il est d'abord nécessaire de réexprimer le problème en terme de processus de branchement.

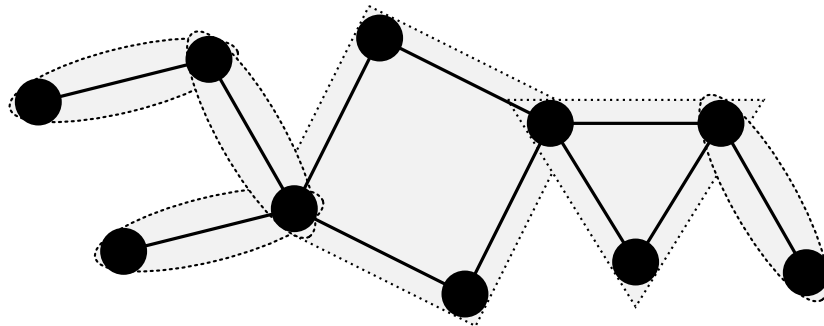
Dans la partie I de cette thèse, on s'intéresse au nombre de voisins qu'un nœud possède, aux voisins de ces voisins *etc.* afin de connaître le nombre de nœuds faisant partie de la même composante que le nœud d'origine. Afin de généraliser cette perspective à l'approche considérée en partie II, on reprend l'image des motifs comme des blocs de construction pouvant être joints selon certaines règles : on s'intéresse d'abord au nombre et au type de motifs pouvant être atteint à partir d'un motif donné, puis on obtient l'arborescence de tous les motifs pouvant être atteints de cette façon.

Même si la littérature emploie une terminologie différente, les processus de branchements y ont déjà été utilisés afin de sonder la structure d'un réseau partiellement défini à travers les motifs le composant. Par exemple, l'approche présentée dans [4] utilise des motifs de premier voisinage afin de décrire la structure d'un réseau non dirigé présentant plusieurs types de nœuds et un seul type de liens<sup>1</sup>. De même, les contributions [73, 59] considèrent la présence d'agrégation (« *clustering* », voir 1.2.5) à l'aide de motifs de premier voisinage explicitant la présence de liens entre ces premier voisins (formant donc des triangles).

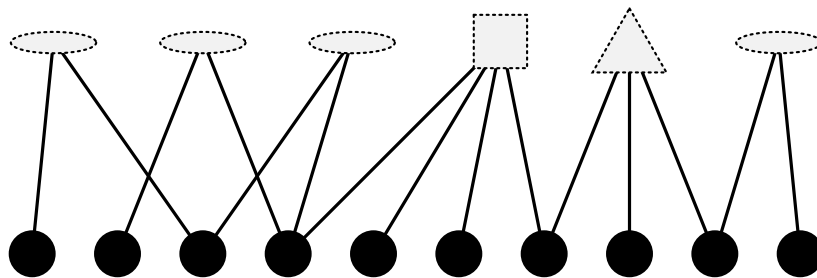
Il est possible de systématiser l'approche en représentant la structure sous la forme d'un réseau biparti<sup>2</sup> composé des éléments du système et des motifs utilisés dans la représentation : chacun des liens indique l'appartenance d'un élément à l'un des motifs (voir l'exemple en figure 8.1). Cette perspective facilite l'expression de la structure sous la forme d'un processus de branchement [44] et peut être généralisée pour des types arbitraires de nœuds, de liens et de motifs ainsi que pour des réseaux dirigés et semi-dirigés [2]. De plus, il est également possible de systématiser le calcul de la contribution des motifs [3].

1. L'approche est généralisable aux réseaux (semi-)dirigés et/ou présentant plus d'un type de liens.

2. Un réseau est dit *biparti* si les nœuds le composant peuvent être divisés en deux ensembles de nœuds tel qu'aucun des liens ne relie deux nœuds d'un même ensemble.



(a) Structure d'un réseau obtenu par la juxtaposition de motifs.



(b) Représentation de la même structure par un réseau biparti d'éléments de motifs.

**Figure 8.1** – Deux représentations d'un même réseau. La position horizontale des éléments du système (cercles noirs) est la même dans les deux cas. Des motifs (formes grises aux contours pointillés) de type « lien », « triangle » et « carré » sont utilisés. L'approche est généralisable pour des réseaux (semi-)dirigés possédant divers types de nœuds et/ou de liens.

Dans tous ces cas, la formulation mathématique du problème en terme de fonctions génératrices demande de résoudre une équation de la forme

$$\mathbf{f}(\boldsymbol{\xi}) = \mathbf{g}(\boldsymbol{\xi}, \mathbf{f}(\boldsymbol{\xi})) \quad (8.1)$$

où  $\boldsymbol{\xi}$  est un vecteur alors que  $\mathbf{f}(\boldsymbol{\xi})$  et  $\mathbf{g}(\boldsymbol{\xi}, \mathbf{f}(\boldsymbol{\xi}))$  sont des champs vectoriels tous deux de même dimension (mais la dimension de  $\boldsymbol{\xi}$  peut être différente). Normalement, le calcul de  $\mathbf{f}(\boldsymbol{\xi})$  n'est qu'une étape intermédiaire afin d'évaluer  $\mathbf{h}(\boldsymbol{\xi}, \mathbf{f}(\boldsymbol{\xi}))$ , le champ vectoriel générant les quantités qui nous sont d'intérêt. Sans entrer dans les détails, on peut se donner une bonne idée du rôle de chacun de ces objets en établissant le parallèle

$$\boldsymbol{\xi} \iff x \quad (8.2a)$$

$$\mathbf{f}(\boldsymbol{\xi}) \iff H_1(x) \quad (8.2b)$$

$$\mathbf{g}(\boldsymbol{\xi}, \mathbf{f}(\boldsymbol{\xi})) \iff x G_1(1 + [H_1(x) - 1]T) \quad (8.2c)$$

$$\mathbf{h}(\boldsymbol{\xi}, \mathbf{f}(\boldsymbol{\xi})) \iff x G_0(1 + [H_1(x) - 1]T) \quad (8.2d)$$

avec le cas simple étudié en (2.14)–(2.15). Plus d'informations sont disponibles dans [2, 4, 44, 59, 73]. Bien entendu, l'approche peut être directement adaptée afin de considérer des « générations ».

Jusqu'ici, on a perçu les processus de branchement d'un point de vue *structurel* : on parcourt le réseau en considérant les motifs comme des points d'embranchement. Cependant, il est également possible de considérer une perspective *évènementielle* : on s'intéresse au nombre et au type d'événements futurs qui sont directement causés par un type d'évènement donné, puis on étudie l'arborescence de tous les évènements<sup>3</sup> résultant d'un nombre donné d'évènements initiaux.

Il arrive que les deux perspectives soient tout à fait équivalentes d'un point de vue mathématique. Ainsi, malgré que la section 6.3.2 mette de l'avant une perspective évènementielle, il aurait été tout aussi justifié d'utiliser une perspective structurelle afin d'obtenir les mêmes équations. Cependant, la perspective évènementielle est plus générale puisqu'elle permet de considérer des structures changeantes ainsi que des phénomènes n'ayant rien à voir avec la structure. De plus, alors que la perspective structurelle écarte le concept de temps du problème, le fait que les évènements surviennent en des instants donnés facilite la formulation du problème en tant que processus de branchement en temps continu, un outil jusqu'ici sous utilisé dans la littérature des réseaux complexes.

---

3. Les évènements correspondent aux changements dans l'état du système (tel que défini en sections 6.2.3 et 7.2.3).

## 8.2 Maximisation d'entropie

Le concept d'ensembles de réseaux discuté en section 1.2.3 est directement inspiré de la physique statistique, et il semble donc logique de profiter d'un outil fondamental à ce domaine : la *fonction de partition*. Cependant, plutôt que d'utiliser la perspective classique « d'équilibre thermodynamique » (ce qui nécessiterait entre autres de donner un sens aux concepts de « température » et « d'énergie »), on préfère plutôt considérer la situation selon la perspective universelle des problèmes d'inférence. Il est à noter que cette section ne donne que les grandes lignes de la méthode ; plus de détails sur l'approche générale sont présentés dans [43] section 11.6 alors que l'application au cas des ensembles de réseaux est couverte dans [81].

On souhaite connaître la probabilité  $P(G)$  que le système soit dans l'état  $G$ , où  $G$  renseigne à la fois sur la structure du réseau ainsi que sur l'état intrinsèque de ses composants<sup>4</sup>. Certaines informations nous sont connues du système sous forme de contraintes fortes et de contraintes faibles, que l'on note

$$G \in \mathcal{G} \quad (8.3a)$$

$$X_i = \sum_{G \in \mathcal{G}} P(G) x_i(G) \quad . \quad (8.3b)$$

Les contraintes fortes sont formulées à travers l'ensemble  $\mathcal{G}$  spécifiant tous les états possibles (autrement dit, on sait avec certitude que  $G \notin \mathcal{G}$  est un état impossible). Pour les contraintes faibles, on spécifie un certain nombre de paires d'une valeur  $X_i$  et d'une fonction  $x_i(G)$  telles que la relation (8.3b) soit respectée. Normalement, on conçoit  $X_i$  comme la valeur moyenne de  $x_i(G)$  sur la distribution de probabilité  $P(G)$ . Cependant, n'importe quelle perspective en accord avec (8.3b) est valide et, par exemple, on peut de cette façon contraindre la covariance de deux quantités.

Outre les contraintes (8.3), notre ignorance de la situation est complète ; la distribution  $P(G)$  doit refléter ce fait. Afin de s'en assurer, on maximise l'*entropie de Shannon*

$$H = -k \sum_{G \in \mathcal{G}} P(G) \ln P(G) \quad (8.4)$$

(où la constante  $k > 0$  détermine les unités de  $H$  et n'affecte donc pas le résultat du processus de maximisation) qui quantifie cette ignorance [43]. Lors de ce processus, on introduit les contraintes faibles (8.3b) à l'aide des *multiplicateurs de Lagrange*  $\theta_i$  et, sachant également que la distribution de probabilité doit être normalisée, on obtient la

---

4. Ainsi,  $G$  joue un rôle très semblable au  $Z$  défini au chapitre 7. On préfère ici noter l'état  $G$  puisqu'il est d'usage d'utiliser  $Z$  pour la fonction de partition.

fonction de partition

$$Z = \sum_{G \in \mathcal{G}} \exp \left( - \sum_i \theta_i x_i(G) \right) \quad (8.5)$$

dans laquelle les multiplicateurs de Lagrange  $\theta_i$  doivent être fixés à l'aide de

$$-\frac{\partial \ln Z}{\partial \theta_i} = X_i \quad . \quad (8.6)$$

La distribution peut alors être obtenue à l'aide de

$$P(G) = \frac{1}{Z} \exp \left( - \sum_i \theta_i x_i(G) \right) \quad (8.7)$$

(pour  $G \in \mathcal{G}$ ). À partir de ce point, plusieurs méthodes standard de physique statistique peuvent être récupérées.

On ne peut trop insister sur le caractère général de l'approche : si on sait exprimer l'état du système sous une forme  $G$  ainsi que les contraintes sous la forme (8.3), alors l'approche s'applique. Dans le cadre de la description par motif mise de l'avant dans la partie II de cette thèse, la principale application envisagée pour cette méthode consiste à systématiser l'inférence des quantités qui nous sont inconnues (tel que discuté en section 7.4.6). En fait, le cas des quantités pouvant être exprimées sous la forme (8.3b) est particulièrement simple : ajouter cette quantité dans (8.5) (comme si on souhaitait la contraindre) à l'aide d'un multiplicateur de Lagrange  $\theta_i$ , obtenir la valeur de  $X_i$  à l'aide de (8.6), puis fixer  $\theta_i = 0$  afin de laisser  $X_i$  libre.

Cependant, à l'exception des cas les plus simples, il n'est normalement pas possible d'obtenir une solution analytique exacte sous une forme fermée. La méthode n'en est pas pour autant inutile puisqu'il existe de nombreuses façons d'obtenir des solutions approximatives de façon analytique, et que l'alternative numérique est normalement possible.

Des problèmes plus sérieux peuvent survenir lorsqu'il nous est impossible de correctement formuler l'information qui nous est accessible sous la forme des contraintes (8.3). Par exemple, sachant le nombre de 2-étoiles (« *two-star* », voir la section 7.A.3) et/ou de triangles présents dans un réseau, on peut tenter d'introduire ces informations dans le modèle en contraignant le nombre moyen de ces motifs avec (8.3b). Tel que démontré dans [80, 82], procéder ainsi peut générer un ensemble de réseaux n'ayant qu'une très faible probabilité de contenir le nombre de motifs désirés, la distribution contenant plus d'une valeur maximale. Bien entendu, on pourrait alors tenter de trouver de nouvelles

contraintes permettant d'obtenir un ensemble de réseaux montrant les propriétés recherchées. Cependant, avant de contourner le problème avec de telles solutions *ad hoc*, il serait peut-être plus sage de prendre un peu de recul et de se questionner sur la situation que l'on souhaite réellement modéliser.

### 8.3 Solution directe des systèmes linéaires

Une brève inspection des équations maîtresses (6.2) et (7.2) nous révèle qu'il ne s'agit dans les deux cas que d'équations différentielles linéaires d'ordre 1. Ainsi, on pourrait s'attendre à ce qu'il soit possible de résoudre ces systèmes de façon analytique.

Dans le cas du système (6.2), chaque transition implique que le nombre de demi-liens non assignés diminue de 2. Ainsi, si la condition initiale est telle que le système est initialement dans l'état  $\mathbf{x}(0)$  [donc  $P(\mathbf{x}(0), 0) = 1$  et toutes les autres probabilités sont initialement nulles], alors il est impossible que le système atteigne un état où le nombre de demi-liens non assignés soit supérieur à  $x_{-1}(0)$ . Ainsi, l'évolution de la probabilité que le système soit encore dans son état initial  $\mathbf{x}(0)$  au temps  $t$  est gouvernée par

$$\frac{d}{dt}P(\mathbf{x}(0), t) + \lambda(\mathbf{x}(0))P(\mathbf{x}(0), t) = 0 \quad (8.8)$$

puisque  $P(\mathbf{x}(0) - \mathbf{r}^j, t) = 0$  peu importe  $j$  et  $t$ . La solution analytique exacte est donc simplement

$$P(\mathbf{x}(0), t) = e^{-\lambda(\mathbf{x}(0))t} \quad (8.9)$$

On peut maintenant utiliser cette solution dans l'équation différentielle gouvernant l'évolution de la probabilité que le système soit dans un état  $\mathbf{x}$  tel que  $x_{-1} = x_{-1}(0) - 2$ .

En fait, pour tout  $\mathbf{x}$  donné, l'équation (6.2) peut être écrite sous la forme

$$\frac{d}{dt}P(\mathbf{x}, t) + \lambda(\mathbf{x})P(\mathbf{x}, t) = f(\mathbf{x}, t) \quad (8.10)$$

où  $f(\mathbf{x}, t)$  est une fonction ne pouvant dépendre que des états contenant deux demi-liens non assignés de plus que  $\mathbf{x}$ . Ainsi, la solution analytique exacte pour la probabilité que le système soit dans l'état  $\mathbf{x}$  au temps  $t$  prend la forme

$$P(\mathbf{x}, t) = \sum_{m,n} c_{mn}(\mathbf{x}) t^m e^{-nt} \quad (8.11)$$

où les  $c_{mn}(\mathbf{x})$  doivent être fixés par récurrence, en commençant par la condition initiale  $\mathbf{x}(0)$  puis en évaluant ces quantités pour des états contenant de moins en moins de

demi-liens non assignés. Il est à noter que les puissances de  $t$  apparaissent lorsqu'un état  $\mathbf{x}'$  contribuant à  $f(\mathbf{x}, t)$  est tel que  $\lambda(\mathbf{x}) = \lambda(\mathbf{x}')$ .

Malgré que la forme (8.11) soit exacte et que l'on puisse obtenir tout  $c_{mn}(\mathbf{x})$  donné par récurrence, aucune forme systématique fermée n'est connue pour un  $c_{mn}(\mathbf{x})$  quelconque. De plus, dû au grand nombre d'états  $\mathbf{x}$  possibles, une évaluation numérique directe des  $c_{mn}(\mathbf{x})$  n'est possible que pour de petits réseaux. Une légère simplification est possible lorsqu'on ne s'intéresse qu'à l'état final du système, permettant alors d'obtenir une forme analytique exacte de la distribution de probabilité sous forme de nombres rationnels pour des réseaux contenant de l'ordre de 30 nœuds. Cependant, même sous cette forme simplifiée, le nombre d'états accessible augmente trop rapidement pour que la méthode puisse être appliquée à de grands réseaux.

Dans la partie III de cette thèse, on simplifie d'abord le problème en le réexprimant en termes d'opérateurs, puis on tente de le résoudre en utilisant diverses approches. Malgré qu'aucune solution concluante n'ait été obtenue pour le problème complet, certains résultats intermédiaires méritent mention.



## Troisième partie

### Approche opérationnelle



# Chapitre 9

## Énoncé du problème

La partie III recherche une solution analytique au modèle  $SI$  présenté au chapitre 6 ainsi qu'à sa version  $SIR$  présentée en section 7.3.4. Ce chapitre énonce le problème à l'aide d'opérateurs régissant l'évolution temporelle du système. La section 9.1 donne un bref aperçu de l'approche puis définit les opérateurs fondamentaux qui sont utilisés en section 9.2 pour exprimer le système  $SIR$ . La section 9.3 reprend le même système sous une forme plus compacte et la section 9.4 simplifie le traitement de la normalisation avant d'introduire les versions  $SI$  et « minimale » du problème. Une méthode alternative pour traiter la normalisation est rapidement présentée en section 9.5. Enfin, la section 9.6 profite d'une perspective plus simple qui est applicable dans les versions  $SI$  et « minimale » du problème.

### 9.1 Définitions

On souhaite ré-exprimer l'évolution temporelle du système (7.10) sous la forme

$$\frac{d|\psi(t)\rangle}{dt} = \hat{L}|\psi(t)\rangle \quad (9.1)$$

où  $|\psi(t)\rangle$  représente la distribution de probabilité pour l'état du système au temps  $t$  et où  $\hat{L}$  est l'opérateur d'évolution. Le but ultime de cette procédure est d'obtenir l'évolution temporelle de  $|\psi(t)\rangle$  en profitant de la relation

$$|\psi(t)\rangle = e^{t\hat{L}}|\psi(0)\rangle \quad , \quad (9.2)$$

où  $|\psi(0)\rangle$  est l'état initial du système.

Pour ce faire, on associe un *ket*  $|\mathbf{x}\rangle$  à chacun des états  $\mathbf{x}$  puis on exprime la distribution de probabilité au temps  $t$  sous la forme

$$|\psi(t)\rangle = \sum_{\mathbf{x}} P(\mathbf{x}|Y, t) |\mathbf{x}\rangle \quad . \quad (9.3)$$

Afin de simplifier la discussion, on introduit la notation

$$|\mathbf{x}\rangle = \left| \begin{array}{cccccc} i_0 & i_1 & i_2 & i_3 & \cdots \\ s_0 & s_1 & s_2 & s_3 & \cdots \\ r_0 & r_1 & r_2 & r_3 & \cdots \end{array} \right\rangle \quad (9.4)$$

où les correspondances  $i_k = x_{I\Lambda_1(k)}$ ,  $s_k = x_{S\Lambda_1(k)}$  et  $r_k = x_{R\Lambda_1(k)}$  sont utilisées<sup>1</sup>. De façon semblable, on définit les *bras*  $\langle \mathbf{x} |$  et complète l'espace d'Hilbert avec le produit scalaire

$$\left\langle \begin{array}{cccccc} i'_0 & i'_1 & i'_2 & i'_3 & \cdots \\ s'_0 & s'_1 & s'_2 & s'_3 & \cdots \\ r'_0 & r'_1 & r'_2 & r'_3 & \cdots \end{array} \middle| \begin{array}{cccccc} i_0 & i_1 & i_2 & i_3 & \cdots \\ s_0 & s_1 & s_2 & s_3 & \cdots \\ r_0 & r_1 & r_2 & r_3 & \cdots \end{array} \right\rangle = \prod_k i_k! s_k! r_k! \delta_{i_k i'_k} \delta_{s_k s'_k} \delta_{r_k r'_k} \quad . \quad (9.5)$$

On introduit maintenant les *opérateurs d'annihilation*  $\hat{a}_k$ ,  $\hat{b}_k$  et  $\hat{c}_k$  à travers leur effet sur un ket

$$\hat{a}_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = i_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k - 1 & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle \quad (9.6a)$$

$$\hat{b}_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = s_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k - 1 & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle \quad (9.6b)$$

$$\hat{c}_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = r_k \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k - 1 & r_{k+1} & \cdots \end{array} \right\rangle \quad (9.6c)$$

et on procède de même pour les *opérateurs de création*  $\hat{a}_k^\dagger$ ,  $\hat{b}_k^\dagger$  et  $\hat{c}_k^\dagger$

$$\hat{a}_k^\dagger \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k + 1 & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle \quad (9.7a)$$

$$\hat{b}_k^\dagger \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = \left| \begin{array}{cccccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k + 1 & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle \quad (9.7b)$$

1. On utilise systématiquement  $k$  plutôt que  $\kappa$  pour les motifs à-la-vollée.

$$\hat{c}_k^\dagger \left| \begin{array}{cccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k & r_{k+1} & \cdots \end{array} \right\rangle = \left| \begin{array}{cccc} \cdots & i_{k-1} & i_k & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k + 1 & r_{k+1} & \cdots \end{array} \right\rangle . \quad (9.7c)$$

On définit le *commutateur* de deux opérateurs comme

$$[\hat{X}, \hat{Y}] = \hat{X}\hat{Y} - \hat{Y}\hat{X} . \quad (9.8)$$

De cette définition, on obtient les relations utiles

$$\begin{aligned} [\hat{X}, \hat{Y}] &= -[\hat{Y}, \hat{X}] , \quad [\alpha\hat{X}, \beta\hat{Y}] = \alpha\beta [\hat{X}, \hat{Y}] , \quad \left[ \sum_k \hat{X}_k, \sum_{k'} \hat{Y}_{k'} \right] = \sum_{k,k'} [\hat{X}_k, \hat{Y}_{k'}] \\ \text{et} \quad [\hat{W}\hat{X}, \hat{Y}\hat{Z}] &= [\hat{W}, \hat{Y}] \hat{X}\hat{Z} + \hat{W} [\hat{X}, \hat{Y}] \hat{Z} + \hat{Y} [\hat{W}, \hat{Z}] \hat{X} + \hat{Y}\hat{W} [\hat{X}, \hat{Z}] . \end{aligned} \quad (9.9)$$

Dans le cas des opérateurs de création et d'annihilation, on obtient les *relations de commutation*

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'} \quad [\hat{b}_k, \hat{b}_{k'}^\dagger] = \delta_{kk'} \quad [\hat{c}_k, \hat{c}_{k'}^\dagger] = \delta_{kk'} \quad (9.10a)$$

$$[\hat{a}_k, \hat{a}_{k'}] = [\hat{a}_k^\dagger, \hat{a}_{k'}^\dagger] = [\hat{b}_k, \hat{b}_{k'}] = [\hat{b}_k^\dagger, \hat{b}_{k'}^\dagger] = [\hat{c}_k, \hat{c}_{k'}] = [\hat{c}_k^\dagger, \hat{c}_{k'}^\dagger] = 0 \quad (9.10b)$$

$$[\hat{a}_k, \hat{b}_{k'}] = [\hat{a}_k^\dagger, \hat{b}_{k'}^\dagger] = [\hat{b}_k, \hat{c}_{k'}] = [\hat{b}_k^\dagger, \hat{c}_{k'}^\dagger] = [\hat{c}_k, \hat{a}_{k'}] = [\hat{c}_k^\dagger, \hat{a}_{k'}^\dagger] = 0 \quad (9.10c)$$

$$[\hat{a}_k, \hat{b}_{k'}^\dagger] = [\hat{a}_k^\dagger, \hat{b}_{k'}] = [\hat{b}_k, \hat{c}_{k'}^\dagger] = [\hat{b}_k^\dagger, \hat{c}_{k'}] = [\hat{c}_k, \hat{a}_{k'}^\dagger] = [\hat{c}_k^\dagger, \hat{a}_{k'}] = 0 . \quad (9.10d)$$

Finalement, on introduit l'*opérateur de normalisation*  $\hat{\Omega}_m$  dont l'effet sur un ket est

$$\hat{\Omega}_m \left| \begin{array}{cccc} i_0 & i_1 & i_2 & \cdots \\ s_0 & s_1 & s_2 & \cdots \\ r_0 & r_1 & r_2 & \cdots \end{array} \right\rangle = \frac{1}{\sum_k k(i_k + s_k + r_k) - m} \left| \begin{array}{cccc} i_0 & i_1 & i_2 & \cdots \\ s_0 & s_1 & s_2 & \cdots \\ r_0 & r_1 & r_2 & \cdots \end{array} \right\rangle . \quad (9.11)$$

Cet opérateur atypique est considéré de façon plus détaillée à la section 9.4.

## 9.2 Traduction de (7.10)

On souhaite maintenant obtenir  $\hat{L}$  à partir de l'équation (7.10) de telle sorte que l'équation (9.1) corresponde à l'équation (7.2).

On s'intéresse d'abord à la transition  $\mathcal{R}k$  correspondant au retrait des individus infectieux. Pour le terme de droite de l'équation (7.2), on cherche l'opérateur qui, lorsqu'appliqué sur  $|\mathbf{x}\rangle$ , retourne  $-q_{\mathcal{R}k}^+(\mathbf{x}, Y)|\mathbf{x}\rangle$ , soit  $-\alpha i_k |\mathbf{x}\rangle$ . On constate que  $-\alpha \hat{a}_k^\dagger \hat{a}_k$

répond à ces critères : le  $\hat{a}_k$  apporte le facteur  $i_k$  tandis que le  $\hat{a}_k^\dagger$  ramène l'état à sa valeur d'origine  $|\mathbf{x}\rangle$ . Pour le terme de gauche, on souhaite obtenir  $q_{\mathcal{R}k}^+(\mathbf{x}, Y) |\mathbf{x} + \mathbf{r}^{\mathcal{R}k}\rangle$ , soit

$$\alpha i_k \left\langle \begin{array}{cccccc} \cdots & i_{k-1} & i_k - 1 & i_{k+1} & \cdots \\ \cdots & s_{k-1} & s_k & s_{k+1} & \cdots \\ \cdots & r_{k-1} & r_k + 1 & r_{k+1} & \cdots \end{array} \right\rangle, \quad (9.12)$$

à partir de  $|\mathbf{x}\rangle$  : l'expression recherchée est  $\alpha \hat{c}_k^\dagger \hat{a}_k$ . La contribution totale du retrait à l'opérateur  $\hat{L}$  est donc

$$\alpha \sum_k \left( \hat{c}_k^\dagger \hat{a}_k - \hat{a}_k^\dagger \hat{a}_k \right). \quad (9.13)$$

Pour une transition  $\mathcal{I}Ik k'$ , on cherche l'expression qui produit  $-q_{\mathcal{I}Ik k'}^+(\mathbf{x}, Y) |\mathbf{x}\rangle$ , soit

$$-\beta k (i_k + \delta_{k(k'-1)} - \delta_{kk'}) k' i_{k'} \frac{1}{\sum_{k''} k'' (i_k + s_k + r_k) - 1} |\mathbf{x}\rangle, \quad (9.14)$$

lorsqu'elle est appliquée sur  $|\mathbf{x}\rangle$ . En procédant comme précédemment, on obtient

$$-\beta k \left( \hat{a}_k^\dagger \hat{a}_k + \delta_{k(k'-1)} - \delta_{kk'} \right) k' \hat{a}_{k'}^\dagger \hat{a}_{k'} \hat{\Omega}_1. \quad (9.15)$$

Pour  $q_{\mathcal{I}Ik k'}^+(\mathbf{x}, Y) |\mathbf{x} + \mathbf{r}^{\mathcal{I}Ik k'}\rangle$ , on obtient simplement  $\beta k \hat{a}_{k-1}^\dagger \hat{a}_k k' \hat{a}_{k'-1}^\dagger \hat{a}_{k'} \hat{\Omega}_1$  : les deltas de Kronecker ne sont pas nécessaires puisque  $\hat{a}_k$  agit après  $\hat{a}_{k-1}^\dagger \hat{a}_k$ . La contribution totale des transitions  $\mathcal{I}Ik k'$  est donc

$$\beta \left[ \sum_k (k+1) \hat{a}_k^\dagger \hat{a}_{k+1} \sum_{k'} (k'+1) \hat{a}_{k'}^\dagger \hat{a}_{k'+1} - \left( \sum_k k \hat{a}_k^\dagger \hat{a}_k - 1 \right) \sum_{k'} k' \hat{a}_{k'}^\dagger \hat{a}_{k'} \right] \hat{\Omega}_1 \quad (9.16)$$

où on a décalé de 1 deux des indices de sommation muets (termes de gauche) et on a fait usage de  $\sum_k (k \delta_{k(k'-1)} - k \delta_{kk'}) = -1$  (termes de droite).

De façon semblable, on obtient

$$\beta \left( \sum_k (k+1) \hat{a}_k^\dagger \hat{b}_{k+1} \sum_{k'} (k'+1) \hat{a}_{k'}^\dagger \hat{a}_{k'+1} - \sum_k k \hat{b}_k^\dagger \hat{b}_k \sum_{k'} k' \hat{a}_{k'}^\dagger \hat{a}_{k'} \right) \hat{\Omega}_1 \quad (9.17)$$

pour  $\mathcal{I}Sk k'$  et

$$\beta \left( \sum_k (k+1) \hat{c}_k^\dagger \hat{c}_{k+1} \sum_{k'} (k'+1) \hat{a}_{k'}^\dagger \hat{a}_{k'+1} - \sum_k k \hat{c}_k^\dagger \hat{c}_k \sum_{k'} k' \hat{a}_{k'}^\dagger \hat{a}_{k'} \right) \hat{\Omega}_1 \quad (9.18)$$

pour  $\mathcal{I}Rk k'$ . En combinant toutes ces contributions, on obtient finalement l'opérateur

d'évolution

$$\hat{L} = \alpha \sum_k \left( \hat{c}_k^\dagger \hat{a}_k - \hat{a}_k^\dagger \hat{a}_k \right) + \beta \left\{ \sum_k (k+1) \left( \hat{a}_k^\dagger \hat{a}_{k+1} + \hat{a}_k^\dagger \hat{b}_{k+1} + \hat{c}_k^\dagger \hat{c}_{k+1} \right) \sum_{k'} (k'+1) \hat{a}_{k'}^\dagger \hat{a}_{k'+1} \right. \\ \left. - \left[ \sum_k k \left( \hat{a}_k^\dagger \hat{a}_k + \hat{b}_k^\dagger \hat{b}_k + \hat{c}_k^\dagger \hat{c}_k \right) - 1 \right] \sum_{k'} k' \hat{a}_{k'}^\dagger \hat{a}_{k'} \right\} \hat{\Omega}_1 \quad (9.19)$$

qui peut être utilisé dans l'équation (9.2) pour obtenir l'évolution temporelle de  $|\psi(t)\rangle$ . Cependant, calculer l'exponentielle d'opérateurs ne commutant pas entre eux n'est pas une tâche triviale; le reste de ce chapitre cherche à simplifier l'expression de  $\hat{L}$  afin de minimiser ces difficultés.

### 9.3 Forme compacte

Une brève inspection de l'équation (9.19) révèle la présence de nombreuses redondances. Cette observation permet d'écrire l'opérateur d'évolution sous une forme plus compacte (en posant  $\beta = 1$  sans perte de généralité)

$$\hat{L} = \alpha (\hat{E}_0^0 - \hat{A}_0^0) + \left[ (\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01}) \hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 + \hat{C}_0^1 - 1) \hat{A}_0^1 \right] \hat{\Omega}_1 \quad (9.20)$$

à l'aide des définitions

$$\hat{A}_1^{01} = \hat{A}_1^0 + \hat{A}_1^1 \quad \hat{D}_1^{01} = \hat{D}_1^0 + \hat{D}_1^1 \quad \hat{C}_1^{01} = \hat{C}_1^0 + \hat{C}_1^1 \quad (9.21a)$$

$$\hat{A}_m^n = \sum_k k^n \hat{a}_k^\dagger \hat{a}_{k+m} \quad \hat{B}_m^n = \sum_k k^n \hat{b}_k^\dagger \hat{b}_{k+m} \quad \hat{C}_m^n = \sum_k k^n \hat{c}_k^\dagger \hat{c}_{k+m} \quad (9.21b)$$

$$\hat{D}_m^n = \sum_k k^n \hat{a}_k^\dagger \hat{b}_{k+m} \quad \hat{E}_m^n = \sum_k k^n \hat{c}_k^\dagger \hat{a}_{k+m} \quad \hat{F}_m^n = \sum_k k^n \hat{c}_k^\dagger \hat{b}_{k+m} \quad . \quad (9.21c)$$

Puisqu'ils préservent les nombres totaux de nœuds susceptibles, infectieux, et retirés, on qualifie d'*homogènes* les opérateurs  $\hat{A}_m^n$ ,  $\hat{B}_m^n$  et  $\hat{C}_m^n$ . Par opposition, on dit des opérateurs  $\hat{D}_m^n$ ,  $\hat{E}_m^n$  et  $\hat{F}_m^n$  qu'ils sont *inhomogènes*.

On peut calculer les relations de commutation de ces opérateurs en combinant les relations (9.9) et (9.10) à la formule du binôme. Par exemple, on a<sup>2</sup>

$$\left[ \hat{A}_m^n, \hat{A}_{m'}^{n'} \right] = \left[ \sum_k k^n \hat{a}_k^\dagger \hat{a}_{k+m}, \sum_{k'} k'^{n'} \hat{a}_{k'}^\dagger \hat{a}_{k'+m'} \right] = \sum_{kk'} k^n k'^{n'} \left[ \hat{a}_k^\dagger \hat{a}_{k+m}, \hat{a}_{k'}^\dagger \hat{a}_{k'+m'} \right] \\ = \sum_{kk'} k^n k'^{n'} \left( \hat{a}_k^\dagger \left[ \hat{a}_{k+m}, \hat{a}_{k'}^\dagger \right] \hat{a}_{k'+m'} + \hat{a}_{k'}^\dagger \left[ \hat{a}_k^\dagger, \hat{a}_{k'+m'} \right] \hat{a}_{k+m} \right)$$

2. Pour les besoins de ces calculs, on doit considérer  $0^0 = 1$  lorsque  $m = 0$  et/ou  $m' = 0$ .

$$\begin{aligned}
&= \sum_{kk'} \left( k^n k'^{n'} \delta_{k'(k+m)} \hat{a}_k^\dagger \hat{a}_{k'+m'} - k^n k'^{n'} \delta_{k(k'+m')} \hat{a}_{k'}^\dagger \hat{a}_{k+m} \right) \\
&= \sum_k \left( k^n (k+m)^{n'} - (k+m')^n k^{n'} \right) \hat{a}_k^\dagger \hat{a}_{k+m+m'} \\
&= \sum_j \left[ \binom{n'}{j} m^j - \binom{n}{j} m'^j \right] \sum_k k^{n+n'-j} \hat{a}_k^\dagger \hat{a}_{k+m+m'} \\
&= \sum_j \left[ \binom{n'}{j} m^j - \binom{n}{j} m'^j \right] \hat{A}_{m+m'}^{n+n'-j} \tag{9.22a}
\end{aligned}$$

et la même démarche permet d'obtenir

$$[\hat{B}_m^n, \hat{B}_{m'}^{n'}] = \sum_j \left[ \binom{n'}{j} m^j - \binom{n}{j} m'^j \right] \hat{B}_{m+m'}^{n+n'-j} \tag{9.22b}$$

$$[\hat{C}_m^n, \hat{C}_{m'}^{n'}] = \sum_j \left[ \binom{n'}{j} m^j - \binom{n}{j} m'^j \right] \hat{C}_{m+m'}^{n+n'-j} \quad . \tag{9.22c}$$

Les relations faisant intervenir les opérateurs inhomogènes peuvent être traités de façon semblable, mais seulement la moitié des termes sont présents

$$[\hat{D}_m^n, \hat{B}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{D}_{m+m'}^{n+n'-j} \quad [\hat{A}_m^n, \hat{D}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{D}_{m+m'}^{n+n'-j} \tag{9.22d}$$

$$[\hat{E}_m^n, \hat{A}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{E}_{m+m'}^{n+n'-j} \quad [\hat{C}_m^n, \hat{E}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{E}_{m+m'}^{n+n'-j} \tag{9.22e}$$

$$[\hat{F}_m^n, \hat{B}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{F}_{m+m'}^{n+n'-j} \quad [\hat{C}_m^n, \hat{F}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{F}_{m+m'}^{n+n'-j} \tag{9.22f}$$

$$[\hat{E}_m^n, \hat{D}_{m'}^{n'}] = \sum_j \binom{n'}{j} m^j \hat{F}_{m+m'}^{n+n'-j} \quad . \tag{9.22g}$$

La relation (9.22g) explique pourquoi on a défini  $\hat{F}_m^n$  malgré son absence dans (9.20). Enfin, toutes autres combinaisons de ces opérateurs commutent entre elles

$$[\hat{A}_m^n, \hat{B}_{m'}^{n'}] = [\hat{B}_m^n, \hat{C}_{m'}^{n'}] = [\hat{C}_m^n, \hat{A}_m^n] = 0 \tag{9.22h}$$

$$[\hat{D}_m^n, \hat{D}_{m'}^{n'}] = [\hat{E}_m^n, \hat{E}_{m'}^{n'}] = [\hat{F}_m^n, \hat{F}_m^n] = 0 \tag{9.22i}$$

$$[\hat{D}_m^n, \hat{C}_{m'}^{n'}] = [\hat{E}_m^n, \hat{B}_{m'}^{n'}] = [\hat{F}_m^n, \hat{A}_m^n] = 0 \tag{9.22j}$$

$$[\hat{F}_m^n, \hat{D}_{m'}^{n'}] = [\hat{F}_m^n, \hat{E}_{m'}^{n'}] = 0 \quad . \tag{9.22k}$$

Le cas où  $n$  et  $n'$  sont contraints aux valeurs 0 et 1 mérite une attention particulière. En effet, tous les opérateurs (excluant  $\hat{\Omega}_m$ , voir la section suivante) composant  $\hat{L}$



respectent ces contraintes et ils forment une algèbre fermée sous les opérations

$$\left[ \hat{A}_m^0, \hat{A}_{m'}^0 \right] = 0 \quad \left[ \hat{A}_m^0, \hat{A}_{m'}^1 \right] = m \hat{A}_{m+m'}^0 \quad \left[ \hat{A}_m^1, \hat{A}_{m'}^1 \right] = (m - m') \hat{A}_{m+m'}^1 \quad (9.23a)$$

(et les autres relations ne faisant intervenir que les opérateurs homogènes) ainsi que sous les opérations

$$\left[ \hat{A}_m^0, \hat{D}_{m'}^0 \right] = \hat{D}_{m+m'}^0 \quad \left[ \hat{A}_m^0, \hat{D}_{m'}^1 \right] = m \hat{D}_{m+m'}^0 + \hat{D}_{m+m'}^1 \quad \left[ \hat{A}_m^1, \hat{D}_{m'}^0 \right] = \hat{D}_{m+m'}^1 \quad (9.23b)$$

(et les autres relations impliquant les opérateurs inhomogènes où au moins un des  $n$  et  $n'$  a la valeur 0). Cependant, la relation

$$\left[ \hat{A}_m^1, \hat{D}_{m'}^1 \right] = m \hat{D}_{m+m'}^1 + \hat{D}_{m+m'}^2 \quad (9.23c)$$

(et d'autres relations impliquant des opérateurs inhomogènes où  $n = n' = 1$ ) fait intervenir des opérateurs à l'extérieur de cette algèbre restreinte. Cette constatation s'avère importante dans les cas où on tente d'éviter une « escalade » des indices  $n$ .

## 9.4 Traitement de la normalisation

L'opérateur de normalisation  $\hat{\Omega}_m$  diffère grandement des autres opérateurs considérés jusqu'ici et, à première vue, sa présence semble apporter des difficultés majeures. Cependant, une inspection plus approfondie permet de minimiser cet impact.

On remarque d'abord que l'inverse

$$\hat{\Omega}_m^{-1} = \sum_k k \left( \hat{a}_k^\dagger \hat{a}_k + \hat{b}_k^\dagger \hat{b}_k + \hat{c}_k^\dagger \hat{c}_k \right) - m = \hat{A}_0^1 + \hat{B}_0^1 + \hat{C}_0^1 - m \quad (9.24)$$

se trouve dans l'opérateur d'évolution

$$\hat{L} = \alpha (\hat{E}_0^0 - \hat{A}_0^0) + (\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01}) \hat{A}_1^{01} \hat{\Omega}_1 - \hat{\Omega}_1^{-1} \hat{A}_0^1 \hat{\Omega}_1 \quad . \quad (9.25)$$

De plus, on observe les propriétés

$$\hat{\Omega}_m \hat{a}_k = \hat{a}_k \hat{\Omega}_{m+k} \quad \hat{\Omega}_m \hat{b}_k = \hat{b}_k \hat{\Omega}_{m+k} \quad \hat{\Omega}_m \hat{c}_k = \hat{c}_k \hat{\Omega}_{m+k} \quad (9.26a)$$

$$\hat{\Omega}_m \hat{a}_k^\dagger = \hat{a}_k^\dagger \hat{\Omega}_{m-k} \quad \hat{\Omega}_m \hat{b}_k^\dagger = \hat{b}_k^\dagger \hat{\Omega}_{m-k} \quad \hat{\Omega}_m \hat{c}_k^\dagger = \hat{c}_k^\dagger \hat{\Omega}_{m-k} \quad (9.26b)$$

qui se traduisent sous la forme

$$\hat{\Omega}_m \hat{A}_{m'}^{n'} = \hat{A}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad \hat{\Omega}_m \hat{B}_{m'}^{n'} = \hat{B}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad \hat{\Omega}_m \hat{C}_{m'}^{n'} = \hat{C}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad (9.27a)$$

$$\hat{\Omega}_m \hat{D}_{m'}^{n'} = \hat{D}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad \hat{\Omega}_m \hat{E}_{m'}^{n'} = \hat{E}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad \hat{\Omega}_m \hat{F}_{m'}^{n'} = \hat{F}_{m'}^{n'} \hat{\Omega}_{m+m'} \quad (9.27b)$$

et permettent d'obtenir

$$\hat{L} = \underbrace{\left[ (\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01}) \hat{A}_1^{01} \right]}_{\hat{L}_2} \hat{\Omega}_1 + \underbrace{\left[ \alpha (\hat{E}_0^0 - \hat{A}_0^0) - \hat{A}_0^1 \right]}_{\hat{L}_0} . \quad (9.28)$$

En écrivant cette dernière relation sous la forme  $\hat{L} = \hat{L}_2 \hat{\Omega}_1 + \hat{L}_0$  et en notant que

$$\hat{\Omega}_m \hat{L}_0 = \hat{L}_0 \hat{\Omega}_m \quad \text{et} \quad \hat{\Omega}_m \hat{L}_2 = \hat{L}_2 \hat{\Omega}_{m+2} , \quad (9.29)$$

on obtient les puissances de  $\hat{L}$

$$\hat{L}^1 = \hat{L}_2 \hat{\Omega}_1 + \hat{L}_0 \quad (9.30a)$$

$$\hat{L}^2 = \hat{L}_2^2 \hat{\Omega}_3 \hat{\Omega}_1 + \hat{L}_2 \hat{L}_0 \hat{\Omega}_1 + \hat{L}_0 \hat{L}_2 \hat{\Omega}_1 + \hat{L}_0^2 \quad (9.30b)$$

$$\begin{aligned} \hat{L}^3 = & \hat{L}_2^3 \hat{\Omega}_5 \hat{\Omega}_3 \hat{\Omega}_1 + \hat{L}_2^2 \hat{L}_0 \hat{\Omega}_3 \hat{\Omega}_1 + \hat{L}_2 \hat{L}_0 \hat{L}_2 \hat{\Omega}_3 \hat{\Omega}_1 + \hat{L}_0 \hat{L}_2^2 \hat{\Omega}_3 \hat{\Omega}_1 \\ & + \hat{L}_2 \hat{L}_0^2 \hat{\Omega}_1 + \hat{L}_0 \hat{L}_2 \hat{L}_0 \hat{\Omega}_1 + \hat{L}_0^2 \hat{L}_2 \hat{\Omega}_1 + \hat{L}_0^3 \end{aligned} \quad (9.30c)$$

⋮

L'observation suivante est capitale : pour un terme donné contenant  $j$  opérateurs  $\hat{L}_2$ , peu importe l'ordre dans lequel apparaissent les  $\hat{L}_0$  et  $\hat{L}_2$ , il est possible de placer tous les opérateurs  $\hat{\Omega}_n$  à droite du terme de telle sorte qu'on aura  $\hat{\Omega}_{2j-1} \cdots \hat{\Omega}_7 \hat{\Omega}_5 \hat{\Omega}_3 \hat{\Omega}_1$ . Au final, les puissances de  $\hat{L}$  sont tout ce qui importe : ce sont ces termes qui composent l'évolution temporelle de  $|\psi(t)\rangle$  puisque

$$|\psi(t)\rangle = e^{t\hat{L}} |\psi(0)\rangle = \sum_k \frac{t^k}{k!} \hat{L}^k |\psi(0)\rangle . \quad (9.31)$$

On pose donc l'opérateur d'évolution

$$\hat{L}_w = w (\hat{A}_1^{01} + \hat{D}_1^{01} + \hat{C}_1^{01}) \hat{A}_1^{01} + \alpha (\hat{E}_0^0 - \hat{A}_0^0) - \hat{A}_0^1 \quad (9.32a)$$

où on a introduit la variable  $w$  (pour laquelle la commutation n'est pas une préoccupation) afin de « compter » le nombre de fois où l'opérateur  $\hat{L}_2$  apparaît dans le résultat final. Ainsi, on peut calculer  $e^{t\hat{L}_w}$  en utilisant le procédé qui nous convient le mieux puis, à la toute fin, on considère le résultat comme une fonction génératrice dépendant de  $w$ . Pour chacun des termes, on substitue  $w^j$  par son équivalent  $\hat{\Omega}_{2j-1} \cdots \hat{\Omega}_3 \hat{\Omega}_1$  que l'on place à la fin du terme<sup>3</sup>. Dans le cas particulier où  $|\psi(0)\rangle$  comporte uniquement des états pour lesquels  $\omega = \sum_k k(i_k + s_k + r_k)$  est le même, la procédure de substitution devient

$$w^j \implies \frac{(\omega - 2j - 1)!!}{(\omega - 1)!!} .$$

3. Sans une telle substitution, les résultats produits par  $\hat{L}_w$  sont non physiques.

Dans le cas d'un système  $SI$ , on pose  $\alpha = 0$  et on constate que l'action de  $\hat{C}_m^n$  sur un état sans nœuds retirés sera toujours nulle. Ainsi, l'opérateur d'évolution

$$\hat{L}'_w = w(\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - \hat{A}_0^1 \quad (9.32b)$$

suffit à l'étude d'un tel système. Dans la situation hypothétique où l'état initial ne contient aucun nœud susceptible, l'action de  $\hat{D}_m^n$  sera également nulle et on obtient alors l'opérateur d'évolution minimal

$$\hat{L}''_w = w\hat{A}_1^{01}\hat{A}_1^{01} - \hat{A}_0^1 \quad (9.32c)$$

D'un point de vue pratique, la dynamique modélisée par  $\hat{L}''_w$  n'est pas très intéressante : tous les nœuds sont initialement infectieux et ils le resteront à tout jamais. Néanmoins, l'étude de  $\hat{L}''_w$  constitue un bon « banc d'essais » pour tester des approches destinées à résoudre  $e^{t\hat{L}'_w}$  et/ou  $e^{t\hat{L}''_w}$ . Il serait en effet surprenant qu'une méthode échouant face à  $\hat{L}''_w$  puisse fonctionner dans les deux autres cas<sup>4</sup>.

## 9.5 Traitement alternatif de la normalisation

Outre la substitution de  $\hat{\Omega}_1$  par  $w$ , on considère une méthode alternative pour les cas  $SI$  et « minimal ». On reprend d'abord (9.20) dans le cas  $SI$

$$\hat{L}' = \left[ (\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 - 1)\hat{A}_0^1 \right] \hat{\Omega}_1 \quad (9.33)$$

et constate que  $\hat{\Omega}_1$  est présent dans tous les termes composant cet opérateur d'évolution. L'effet de  $\hat{\Omega}_1$  sur un état donné est simplement de multiplier cet état par une constante. Or, multiplier tous les termes de droite dans l'équation (9.1) par un facteur  $\lambda$  correspond à un changement d'échelle du temps  $t$  vers un temps  $\lambda^{-1}t'$  : l'effet de  $\hat{\Omega}_1$  n'est qu'un changement de l'échelle temporelle. Si l'on ne s'intéresse qu'à l'état final, on peut simplement définir les opérateurs d'évolution sans  $\hat{\Omega}_1$

$$\hat{L}'_\infty = (\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - (\hat{A}_0^1 + \hat{B}_0^1 - 1)\hat{A}_0^1 \quad (9.34a)$$

$$\hat{L}''_\infty = \hat{A}_1^{01}\hat{A}_1^{01} - (\hat{A}_0^1 - 1)\hat{A}_0^1 \quad (9.34b)$$

et obtenir l'état final avec  $\lim_{t \rightarrow \infty} e^{t\hat{L}'_\infty} |\psi(0)\rangle$  pour un système  $SI$  et  $\lim_{t \rightarrow \infty} e^{t\hat{L}''_\infty} |\psi(0)\rangle$  pour un système minimal.

---

4. Par contre, l'inverse n'est pas nécessairement vrai : il n'y a aucune garantie qu'une méthode résolvant  $e^{t\hat{L}'_w}$  sera également efficace face aux deux autres problèmes.

En réalité, la situation est légèrement plus compliquée : à deux transitions différentes peuvent être associées des échelles temporelles différentes. En temps normal, cette situation entraînerait des différences dans l'état final, rendant inutile les définitions (9.34). Ce n'est cependant pas le cas ici, en raison d'une interaction de plusieurs phénomènes : la linéarité du système, sa nature « dynamiquement acyclique<sup>5</sup> », et le fait que le facteur d'échelle apporté par  $\hat{\Omega}_1$  est le même pour tous les flots de probabilités quittant un état  $|\mathbf{x}\rangle$  donné<sup>6</sup>.

L'approche (9.34) n'est pas explorée en profondeur pour la simple raison que celle présentée en équation (9.32) est plus générale (autant pour l'évolution temporelle que pour le cas *SIR*) tout en présentant une forme *a priori* plus simple. Un certain nombre de circonstances pourraient tout de même donner avantage aux relations (9.34) par rapport aux (9.32), principalement dû au fait qu'elles sont « plus physiques ». Entre autres, (9.34b) peut être perçu comme la désexcitation mutuelle de deux particules à travers une interaction (chacune baissant d'un niveau d'énergie). De plus, le fait que les relations (9.32) ne préservent pas la normalisation avant la resubstitution des  $w^j$  pourrait interdire l'usage de certaines méthodes, alors que ce n'est pas le cas pour les relations (9.34).

## 9.6 Forme « demi-liens infectieux »

Jusqu'ici, on a considéré l'approche introduite en section (7.3.4) : on s'intéresse au nombre de demi-liens des nœuds de chaque état. Tel que vu au chapitre 6, il est possible d'amener le nombre de variables dynamiques au 2/3 de cette valeur dans le cas *SIR* et au 1/3 dans le cas *SI*. La présente section s'inspire de ce dernier cas pour introduire un opérateur d'évolution ne dépendant que d'une seule paire d'opérateurs de création et d'annihilation.

Le système *SI* du chapitre 6 suit l'évolution du nombre total de demi-liens non assignés  $\omega$  en plus du nombre de demi-liens des nœuds susceptibles. Ces quantités permettent d'obtenir le nombre  $\lambda$  de demi-liens infectieux non assignés par la simple relation  $\lambda = \omega - \sum_k k s_k$ . On considère ici l'approche inverse : suivre  $|\lambda; s_0, s_1, \dots\rangle$  puis déduire le nombre total de demi-liens non assignés par  $\omega = \lambda + \sum_k k s_k$ . L'opérateur

---

5. Le flot de probabilité quittant un état donné ne peut jamais y retourner. En effet, la quantité  $\sum_k k(i_k + s_k)$  décroît de 2 à chaque transition, rendant impossible tout retour à un état antérieur. Dans le cas *SIR*, soit  $\sum_k k(i_k + s_k + r_k)$  décroît de 2, soit  $\sum_k r_k$  augmente de 1.

6. Cette dernière condition n'est pas respectée dans le cas *SIR*, expliquant l'absence d'un  $\hat{L}_\infty$  en équation (9.34).

d'évolution pourrait alors être exprimé comme

$$\left( \sum_k k(\hat{a}^\dagger)^{k-1} \hat{b}_k + \hat{a} \right) \hat{\Omega}_0 \hat{a} - \hat{a}^\dagger \hat{a} \quad (9.35)$$

où il est entendu que  $\hat{a}$  et  $\hat{a}^\dagger$  agissent sur  $\lambda$ .

Deux remarques importantes s'imposent. Tout d'abord, l'approche utilisée en section 9.4 peut être utilisée à nouveau pour remplacer  $\hat{\Omega}_n$  par  $w$ . Une fois cette substitution faite, on constate que les  $\hat{b}_k$  commutent avec tous les objets restants : on indique ce fait en les substituant par  $v_k$ . Ainsi, on définit les opérateurs d'évolution

$$\hat{L}'_\lambda = w \left( \sum_k k v_k (\hat{a}^\dagger)^{k-1} + \hat{a} \right) \hat{a} - \hat{a}^\dagger \hat{a} \quad (9.36a)$$

$$\hat{L}''_\lambda = w \hat{a}^2 - \hat{a}^\dagger \hat{a} \quad (9.36b)$$

pour les cas  $SI$  ( $\hat{L}'_\lambda$ ) et minimal ( $\hat{L}''_\lambda$ ) agissant sur des états de la forme  $|\lambda\rangle$  à l'aide des relations

$$\hat{a}|\lambda\rangle = \lambda|\lambda-1\rangle \quad \hat{a}^\dagger|\lambda\rangle = |\lambda+1\rangle \quad [\hat{a}, \hat{a}^\dagger] = 1 \quad . \quad (9.37)$$

Si la condition initiale n'est composée que d'un seul état en terme des  $s_0, s_1, \dots, i_0, i_1, \dots$ , alors on peut utiliser les substitutions

$$w^j \implies \frac{(\omega - 2j - 1)!!}{(\omega - 1)!!} \quad v_k^j \implies \frac{s_k!}{(s_k - j)!} \quad (9.38)$$

une fois que  $e^{t\hat{L}'_\lambda}$  (ou  $e^{t\hat{L}''_\lambda}$ ) est obtenu.

Dans le contexte des travaux ayant menés à cette thèse, cette forme du problème a été envisagée bien après les deux précédentes. Tout comme pour  $\hat{L}'_\infty$ , l'approche n'est pas applicable dans le cas  $SIR$ . Cependant, payer ce coût semble cette fois beaucoup plus intéressant : il en résulte une simplification majeure du système. Si la solution du cas minimal  $\hat{L}''_\lambda$  est toujours aussi inintéressante, une solution exacte de  $\hat{L}'_\lambda$  aurait entre autres des applications bien réelles concernant la partition des composantes dans les réseaux du modèle de configuration.



# Chapitre 10

## Outillage

On présente quelques outils pouvant peut-être aider à la solution de problèmes de la forme  $e^{t\hat{L}}$ , tels que ceux formulés au chapitre 9. Puisqu'il ne s'agit pas là du sujet principal de cette thèse, l'emphase est placée sur la méthode plutôt que sur la démonstration.

Ce chapitre est centré autour de la relation (10.8) : la section 10.1 définit certains concepts nécessaires à sa compréhension, la section 10.2 introduit cette relation ainsi que des relations reliées ou alternatives et la section 10.3 présente rapidement la méthode qui est utilisée en section 10.4 afin d'obtenir (10.8). Bien qu'il soit possible que cette relation ait été obtenue par le passé, elle semble peu connue et n'a pas été rencontrée dans la littérature.

### 10.1 Algèbres de Lie : un aperçu

Une *algèbre de Lie* [30, 42] est composée d'un espace vectoriel  $\mathfrak{g}$  sur un corps commutatif  $\mathcal{F}$  et d'une application  $[\cdot, \cdot]$ , le *produit de Lie*, respectant

$$[ax + by, z] = a[x, z] + b[y, z] \quad [z, ax + by] = a[z, x] + b[z, y] \quad (10.1a)$$

$$[x, x] = 0 \quad (10.1b)$$

$$[[x, y], z] + [[z, x], y] + [[y, z], x] = 0 \quad (10.1c)$$

où  $x, y, z \in \mathfrak{g}$  et  $a, b \in \mathcal{F}$ . Dans le contexte de cette thèse, on peut simplement considérer le produit de Lie comme le commutateur défini en (9.8) et poser  $\mathcal{F} = \mathbb{R}$ .

Suivant grossièrement l'approche de Magnus [53] afin d'introduire le concept d'algèbre de Lie libre, on définit d'abord un anneau  $\mathcal{R}(x_1, x_2, \dots, x_n)$  à l'aide de  $n$  géné-

*rateurs libres*  $x_1, x_2, \dots, x_n$  : chaque élément de  $\mathcal{R}(x_1, \dots)$  correspond à une série de puissance (formelle) dont les coefficients sont tirés de  $\mathcal{F}$  et où l'ordre d'apparition des générateurs importe. Le *degré*  $d$  d'un terme d'une telle séquence correspond au nombre total de générateurs y apparaissant, chaque instance d'un même générateur contribuant 1 au total. Un terme de degré 0 est un coefficient tiré de  $\mathcal{F}$  multipliant 1. Un élément de  $\mathcal{R}(x_1, \dots)$  ne contenant que des termes de degré  $d$  est dit être un *élément homogène de degré*  $d$  de  $\mathcal{R}(x_1, \dots)$ .

On définit maintenant l'*algèbre de Lie libre*  $\mathfrak{f}(x_1, \dots)$  : tous les éléments homogènes de degré 1 de  $\mathcal{R}(x_1, \dots)$  sont éléments de  $\mathfrak{f}(x_1, \dots)$  et, de façon récursive, le produit de Lie de deux éléments de  $\mathfrak{f}(x_1, \dots)$  est un élément de  $\mathfrak{f}(x_1, \dots)$ . Dans le cas de  $\mathfrak{f}(x, y)$ , les éléments de degré  $\leq 3$  sont de forme

$$c_1x + c_2y + c_{12}(xy - yx) + c_{121}((xy - yx)x - x(xy - yx)) + c_{122}((xy - yx)y - y(xy - yx)) \quad (10.2)$$

où  $c_1, c_2, \dots \in \mathcal{F}$ . De façon générale, tous les éléments de  $\mathfrak{f}(x_1, \dots)$  sont dans  $\mathcal{R}(x_1, \dots)$ , mais l'inverse n'est pas vrai.

Malgré que le sujet n'ait qu'à peine été effleuré, ces quelques définitions concernant les algèbres de Lie suffisent aux besoins de cette thèse.

## 10.2 Exponentielles d'opérateurs

La *relation de Zassenhaus* [53] stipule que, pour  $x, y \in \mathfrak{g}$ , il existe des  $C_2, C_3, \dots \in \mathfrak{g}$  uniquement déterminés tels que

$$e^{t(x+y)} = e^{tx} e^{ty} e^{t^2 C_2} e^{t^3 C_3} e^{t^4 C_4} \dots \quad (10.3)$$

De plus, les  $C_n$  peuvent être obtenus en terme de commutateurs de  $x$  et  $y$ ; en fait, si  $\mathfrak{g} = \mathfrak{f}(x, y)$ , alors  $C_n$  est un élément homogène de degré  $n$ . Même si  $\mathfrak{g}$  est une algèbre non libre, il est possible d'obtenir les  $C_n$  pour le système alternatif  $\mathfrak{f}(x', y')$  puis de substituer  $x$  et  $y$  dans les relations obtenues. Ainsi, la solution pour une algèbre de Lie libre peut être utilisée pour calculer  $e^{t(\hat{X} + \hat{Y})}$  même si les opérateurs  $\hat{X}$  et  $\hat{Y}$  présentent des relations de commutation non triviales.

Dans le cas général  $\mathfrak{f}(x, y)$ , les premières valeurs de  $C_n$  sont

$$C_2 = -\frac{1}{2}[x, y] \quad (10.4a)$$

$$C_3 = -\frac{1}{3}[[x, y], y] - \frac{1}{6}[[x, y], x] \quad (10.4b)$$



Malheureusement, il n'existe pas de règle générale simple permettant d'obtenir  $C_n$  sous une forme fermée : toutes les relations connues demandent un calcul récursif explicite. Par exemple, la méthode de Suzuki [96, 97] prend la forme

$$C_n = \frac{1}{n!} \left( \frac{\partial^n}{\partial \lambda^n} e^{-\lambda^{n-1} C_{n-1}} \dots e^{-\lambda^2 C_2} e^{-\lambda y} e^{-\lambda x} e^{\lambda(x+y)} \right) \Big|_{\lambda=0}. \quad (10.5)$$

La section 10.3 présente une méthode alternative introduite par Magnus dans [53] ; l'évaluation numérique des  $C_n$  est considérée dans [104].

Toutefois, certains cas particuliers peuvent être traités exactement. Par exemple, le cas suivant est bien connu en mécanique quantique : étant donné  $[\hat{X}, \hat{Y}] = \alpha$ , on obtient

$$e^{t(\hat{X}+\hat{Y})} = e^{t\hat{X}} e^{t\hat{Y}} e^{-\frac{1}{2}\alpha t^2} \quad (10.6)$$

puisque tous les  $C_n$  pour  $n > 2$  impliquent des commutateurs de la forme  $[\alpha, \cdot] = 0$ . Un cas moins connu est appelé *identité de Sack* [89] : étant donné  $[\hat{X}, \hat{Y}] = \alpha \hat{Y}$ , on obtient

$$e^{t(\hat{X}+\hat{Y})} = \exp\left(\alpha^{-1}(e^{\alpha t} - 1)\hat{Y}\right) e^{t\hat{X}} = e^{t\hat{X}} \exp\left(\alpha^{-1}(1 - e^{-\alpha t})\hat{Y}\right). \quad (10.7)$$

Il est possible de généraliser cette identité : étant donné  $[\hat{X}, \hat{Y}_n] = \hat{Y}_{n+1}$  et  $[\hat{Y}_n, \hat{Y}_{n'}] = 0$ , on obtient

$$e^{t(\hat{X}+\hat{Y}_1)} = e^{t\hat{X}} \exp\left(-\sum_{n=1}^{\infty} \frac{(-t)^n}{n!} \hat{Y}_n\right) \quad (10.8)$$

où tous les termes de la seconde exponentielle commutent entre eux. Cette dernière relation est démontrée<sup>1</sup> à la section 10.4.

Souvent considérée comme la relation duale de la relation de Zassenhaus, la *relation de Baker-Campbell-Hausdorff* garantit l'existence d'un  $z \in \mathfrak{g}$  satisfaisant

$$z = \ln(e^x e^y) \quad (10.9)$$

pour  $x, y \in \mathfrak{g}$ . Les premiers termes de  $z$  prennent la forme

$$z = x + y + \frac{1}{2}[x, y] + \frac{1}{12}[[x, y], y] - \frac{1}{12}[[x, y], x] - \frac{1}{24}[[[x, y], x], y] + \dots \quad (10.10)$$

et les termes manquants impliquent toujours de plus en plus de commutateurs imbriqués. De façon générale, on a [24]

$$z = \sum_{n>0} \sum_{\substack{r_1+s_1>0 \\ \vdots \\ r_n+s_n>0}} \frac{(-1)^{m+n-1}}{mn r_1! s_1! \dots r_n! s_n!} \{y^{s_n} x^{r_n} \dots y^{s_1} x^{r_1}\} \quad (10.11)$$

1. Je n'ai pas rencontré (10.8) dans la littérature. Cependant, il ne serait pas surprenant qu'elle ait déjà été obtenue.

où il est compris que  $m = \sum_{i=1}^n (r_i + s_i)$ . L'opérateur  $\{\cdot\}$  est introduit en section 10.3. Une relation souvent associée à celle de Baker-Campbell-Hausdorff est l'identité

$$e^{-tx} y e^{tx} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \{yx^k\} \quad . \quad (10.12)$$

Enfin, la *relation de Trotter*

$$e^{t(x+y)} = \lim_{n \rightarrow \infty} \left( 1 + \frac{t}{n}(x+y) \right)^n \quad (10.13a)$$

$$e^{t(x+y)} = \lim_{n \rightarrow \infty} \left( e^{\frac{t}{n}x} e^{\frac{t}{n}y} \right)^n \quad (10.13b)$$

est couramment utilisée comme approximation dans les applications numériques, la forme (10.13b) convergeant beaucoup plus rapidement que la première. On peut d'avantage accélérer la convergence en utilisant plutôt la *relation de Suzuki-Trotter* [96, 97]

$$e^{t(x+y)} = \lim_{n \rightarrow \infty} \left( e^{\frac{t}{n}x} e^{\frac{t}{n}y} e^{\left(\frac{t}{n}\right)^2 C_2} \dots e^{\left(\frac{t}{n}\right)^m C_m} \right)^n \quad (10.13c)$$

faisant intervenir les mêmes  $C_2, C_3, \dots$  que ceux de la relation de Zassenhaus. Il est à noter qu'il est possible d'utiliser la relation de Trotter (10.13a) afin d'écrire  $e^{t\hat{L}}$  sous la forme d'une intégrale fonctionnelle (ou de parcours, « *path integral* ») [87].

Cependant, ni cette dernière formulation en intégrale fonctionnelle, ni aucune des relations (10.9)–(10.13), n'a permis d'obtenir des résultats méritant mention concernant la solution du problème tel que formulé au chapitre 9. En fait, ces relations ne sont énumérées ici que par soucis de commodité pour un lecteur qui souhaiterait s'attaquer au problème; elles ne sont nullement nécessaires à la compréhension du reste de cette thèse.

Le reste de ce chapitre vise à obtenir la relation (10.8). Malgré que cette dernière relation n'ait porté fruit que dans les cas les plus simples du problème formulé au chapitre 9, elle pourrait peut-être s'avérer utile dans un contexte différent. De plus, comprendre pourquoi (10.8) échoue à résoudre certains problèmes aide à comprendre ce qui rend ces problèmes difficiles.

## 10.3 Méthode de Magnus

On explicite ici les grandes lignes de l'approche de Magnus [53] dans le cas général  $f(x, y)$  : appliquer un opérateur  $\{\cdot\}$  des deux côtés de (10.3) puis apparier les éléments homogènes de même degré.

Ainsi, on définit l'opérateur  $\{\cdot\}$  dont l'action sur tout  $r \in \mathcal{R}(x_1, \dots)$  est spécifiée par les propriétés suivantes.

1. Pour  $r \in \mathcal{R}(x_1, \dots)$  et  $c \in \mathcal{F}$ , on a

$$\{cr\} = c\{r\} \quad . \quad (10.14a)$$

2. Pour  $r, r' \in \mathcal{R}(x_1, \dots)$ , on a

$$\{r + r'\} = \{r\} + \{r'\} \quad . \quad (10.14b)$$

3. Pour un produit de générateurs  $w_1 w_2 w_3 \cdots w_m$  (chacun des  $w_i$  correspond à l'un des générateurs  $x_1, \dots, x_n$ ), on a

$$\{w_1 w_2 w_3 \cdots w_m\} = [[\cdots [[w_1, w_2], w_3] \cdots], w_m] \quad , \quad \{w_1\} = w_1 \quad \text{et} \quad \{1\} = 0 \quad . \quad (10.14c)$$

De ces propriétés, on peut démontrer

$$\{h_d\} = dh_d \quad (10.15a)$$

$$\{h_d^2 r\} = 0 \quad (10.15b)$$

pour un élément  $h_d \in \mathfrak{f}(x_1, \dots)$  homogène de degré  $d$  et  $r \in \mathcal{R}(x_1, \dots)$ . La relation (10.15a) est obtenue dans [24, 94, 103] alors que (10.15b) l'est dans [103]; les deux relations sont rappelées dans [53].

On applique maintenant cet opérateur à la relation (10.3). Pour le côté gauche, on explicite l'exponentielle sous la forme d'une série de puissances puis on utilise les relations précédentes afin d'obtenir

$$\begin{aligned} \{e^{t(x+y)}\} &= \left\{ \sum_{k=0}^{\infty} \frac{t^k}{k!} (x+y)^k \right\} \stackrel{(10.14b)}{=} \sum_{k=0}^{\infty} \left\{ \frac{t^k}{k!} (x+y)^k \right\} \stackrel{(10.14a)}{=} \sum_{k=0}^{\infty} \frac{t^k}{k!} \{(x+y)^k\} \\ &\stackrel{(10.14c)}{=} \overbrace{\frac{t^0}{0!} \{1\}} + \frac{t^1}{1!} \{x+y\} + \overbrace{\sum_{k=2}^{\infty} \frac{t^k}{k!} \{(x+y)^2 (x+y)^{k-2}\}} \stackrel{(10.15b)}{=} t\{x+y\} \\ &\stackrel{(10.14b)}{=} t(\{x\} + \{y\}) \stackrel{(10.14c)}{=} t(x+y) \quad . \end{aligned} \quad (10.16)$$

Il est à noter que l'usage de la relation (10.15b) à la seconde ligne est permis puisque  $(x+y)$  est un élément homogène de  $\mathfrak{f}(x, y)$  et que  $(x+y)^{k-2} \in \mathcal{R}(x, y)$  pour  $k \geq 2$ .

Pour le côté droit, on a

$$\{e^{tx} e^{ty} e^{t^2 C_2} \cdots\} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \{x^k e^{ty} e^{t^2 C_2} \cdots\} = \{e^{ty} e^{t^2 C_2} \cdots\} + t\{x e^{ty} e^{t^2 C_2} \cdots\} \quad (10.17)$$

où les cas  $k \geq 2$  s'annulent par (10.15b) ; un argument semblable permet d'écrire

$$\{e^{ty} e^{t^2 C_2} \dots\} = \{e^{t^2 C_2} \dots\} + t\{y e^{t^2 C_2} \dots\} \quad (10.18)$$

et donc

$$\begin{aligned} \{e^{tx} e^{ty} e^{t^2 C_2} \dots\} & \quad (10.19) \\ & = \{e^{t^2 C_2} \dots\} + t\{y e^{t^2 C_2} \dots\} + t\{x e^{t^2 C_2} \dots\} + \sum_{k=1}^{\infty} \frac{t^{k+1}}{k!} \{xy^k e^{t^2 C_2} \dots\} \quad . \end{aligned}$$

On considère maintenant

$$\begin{aligned} e^{t^2 C_2} e^{t^3 C_3} e^{t^4 C_4} \dots & \\ & = \left(1 + t^2 C_2 + \frac{1}{2} t^4 C_2^2 + \dots\right) \left(1 + t^3 C_3 + \dots\right) \left(1 + t^4 C_4 + \dots\right) \left(1 + t^5 C_5 + \dots\right) \dots \\ & = 1 + t^2 C_2 + t^3 C_3 + t^4 \left(\frac{1}{2} C_2^2 + C_4\right) + t^5 (C_2 C_3 + C_5) + \dots \quad (10.20) \end{aligned}$$

où tous les éléments sont explicitement présentés jusqu'au degré 5. On utilise ce résultat pour calculer le premier terme de (10.19)

$$\{e^{t^2 C_2} \dots\} = t^2 2C_2 + t^3 3C_3 + t^4 4C_4 + t^5 (\{C_2 C_3\} + 5C_5) + \dots \quad (10.21a)$$

où on a obtenu  $\{1\} = 0$  par (10.14c),  $\{C_2^2\} = 0$  par (10.15b) et  $\{C_n\} = nC_n$  par (10.15a) ; le terme  $\{C_2 C_3\}$  doit être évalué explicitement par (10.14c). Pour les autres termes, on a

$$t\{y e^{t^2 C_2} \dots\} = ty + t^3 \{y C_2\} + t^4 \{y C_3\} + t^5 \left(\frac{1}{2} \{y C_2^2\} + \{y C_4\}\right) + \dots \quad (10.21b)$$

$$t\{x e^{t^2 C_2} \dots\} = tx + t^3 \{x C_2\} + t^4 \{x C_3\} + t^5 \left(\frac{1}{2} \{x C_2^2\} + \{x C_4\}\right) + \dots \quad (10.21c)$$

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{t^{k+1}}{k!} \{xy^k e^{t^2 C_2} \dots\} & = t^2 \{xy\} + t^3 \frac{1}{2} \{xy^2\} + t^4 (\{xy C_2\} + \frac{1}{6} \{xy^3\}) \\ & + t^5 (\{xy C_3\} + \frac{1}{2} \{xy^2 C_2\} + \frac{1}{6} \{xy^3 C_2\} + \frac{1}{24} \{xy^4\}) + \dots \quad (10.21d) \end{aligned}$$

où tous les éléments sont explicitement présentés jusqu'au degré 5.

On égalise maintenant le côté gauche (10.16) de l'équation de Zassenhaus aux quatre contributions au côté droit (10.21), puis on collecte tous les éléments d'un même degré. Les termes de degré 1 s'annulant, on obtient

$$2C_2 = -\{xy\} \quad (10.22a)$$

$$3C_3 = -\{y C_2\} - \{x C_2\} - \frac{1}{2} \{xy^2\} \quad (10.22b)$$

$$4C_4 = -\{y C_3\} - \{x C_3\} - \{xy C_2\} - \frac{1}{6} \{xy^3\} \quad (10.22c)$$

$$\begin{aligned} 5C_5 & = -\{C_2 C_3\} - \frac{1}{2} \{y C_2^2\} - \{y C_4\} - \frac{1}{2} \{x C_2^2\} - \{x C_4\} \\ & \quad - \{xy C_3\} - \frac{1}{2} \{xy^2 C_2\} - \frac{1}{6} \{xy^3 C_2\} - \frac{1}{24} \{xy^4\} \quad (10.22d) \\ & \quad \vdots \end{aligned}$$

En explicitant (10.22a) à l'aide de (10.14c), on peut isoler  $C_2$  et obtenir (10.4a). On peut ensuite substituer cette relation dans (10.22b), expliciter les commutateurs avec (10.14c), isoler  $C_3$  et obtenir (10.4b). La généralisation aux  $C_n$  pour  $n > 3$  est directe, quoique fastidieuse.

## 10.4 Une généralisation de l'identité de Sack

Cette section démontre la relation (10.8) à l'aide de la méthode de Magnus [53], présentée en 10.3. En supposant que tous les opérateurs  $\{\cdot\}$  aient été explicités en termes de commutateurs, on considère la forme que prendront ces termes sachant que

$$[\hat{X}, \hat{Y}_n] = \hat{Y}_{n+1} \quad \text{et} \quad [\hat{Y}_n, \hat{Y}_{n'}] = 0 \quad (10.23)$$

en substituant  $x \Rightarrow \hat{X}$  et  $y \Rightarrow \hat{Y}_1$ . Ainsi, on a

$$\{x\} = x \quad \Longrightarrow \quad \hat{X} \quad (10.24a)$$

$$\{y\} = y \quad \Longrightarrow \quad \hat{Y}_1 \quad (10.24b)$$

$$\{xyx^{n-2}\} = [[\cdots [[x, y], x] \cdots], x] \Longrightarrow [[\cdots [[\hat{X}, \hat{Y}_1], \hat{X}] \cdots], \hat{X}] = (-1)^n \hat{Y}_n \quad (10.24c)$$

$$\{yx^{n-1}\} = [[\cdots [[y, x], x] \cdots], x] \Longrightarrow [[\cdots [[\hat{Y}_1, \hat{X}], \hat{X}] \cdots], \hat{X}] = (-1)^{n-1} \hat{Y}_n \quad (10.24d)$$

où  $n \geq 2$ . De plus, on constate que tout opérateur  $\{\cdot\}$  contenant deux  $y$  ou plus annule automatiquement le résultat final : seuls les cas donnés en (10.24) peuvent contribuer aux  $C_n$ .

En fait, on constate que les éléments homogènes de degré  $n$  de  $\mathfrak{f}(x, y)$  se traduisent tous sous la forme  $\alpha \hat{Y}_n$  où  $\alpha \in \mathcal{F}$  (à l'exception des termes de degré 1, prenant la forme  $\alpha \hat{Y}_1 + \beta \hat{X}$  où  $\alpha, \beta \in \mathcal{F}$ ). Ainsi, comme les  $C_n$  sont des éléments homogènes de degré  $n$ , ils doivent prendre la forme  $C_n = c_n \hat{Y}_n$ , où les  $c_n$  sont des constantes à déterminer. En termes de  $\mathfrak{f}(x, y)$ , on a  $C_n = (-1)^{n-1} c_n \{yx^{n-1}\}$  : pour la même raison qu'un opérateur  $\{\cdot\}$  contenant deux  $y$  ou plus a une contribution nulle, sa contribution est également nulle s'il contient à la fois  $y$  et  $C_n$ , ou encore  $C_n$  et  $C_{n'}$ . Ainsi, les contributions aux

termes de droite de (10.21) deviennent

$$\{e^{t^2 C_2} \dots\} = \sum_{n=2}^{\infty} t^n n (-1)^{n-1} c_n \{yx^{n-1}\} \quad (10.25a)$$

$$t\{y e^{t^2 C_2} \dots\} = ty \quad (10.25b)$$

$$t\{x e^{t^2 C_2} \dots\} = tx + \sum_{n=2}^{\infty} t^{n+1} (-1)^{n-1} c_n \{x\{yx^{n-1}\}\} \quad (10.25c)$$

$$\sum_{k=1}^{\infty} \frac{t^{k+1}}{k!} \{xy^k e^{t^2 C_2} \dots\} = t^2 \{xy\} \quad (10.25d)$$

Afin d'évaluer  $\{x\{yx^{n-1}\}\}$ , on constate que tous les termes composant  $\{yx^{n-1}\}$  débutent par  $x$ , à l'exception de  $yx^{n-1}$ . Or,  $\{x^2 \dots\}$  est toujours nul et donc

$$\{x\{yx^{n-1}\}\} = \{xyx^{n-1}\} = -\{yx^n\} \quad (10.26)$$

On égale donc le côté gauche (10.16) de l'équation de Zassenhaus à ces quatres contributions (10.25) puis on collecte les éléments de même degré pour obtenir

$$c_2 = -\frac{1}{2} \quad \text{et} \quad c_n = -\frac{1}{n} c_{n-1} \quad (10.27)$$

ayant comme solution

$$c_n = \frac{(-1)^{n-1}}{n!} \quad \text{et donc} \quad C_n = \frac{(-1)^{n-1}}{n!} \hat{Y}_n \quad (10.28)$$

La relation (10.8) découle directement de ce résultat.

Dans le cas spécifique où  $[\hat{X}, \hat{Y}] = \alpha \hat{Y}$  et  $\hat{Y}_n = \alpha^{n-1} \hat{Y}$ , on constate que la condition

$$[\hat{X}, \hat{Y}_n] = \alpha^{n-1} [\hat{X}, \hat{Y}] = \alpha^n \hat{Y} = \hat{Y}_{n+1} \quad (10.29)$$

est bien vérifiée puis on utilise (10.8)

$$\begin{aligned} e^{t(\hat{X}+\hat{Y})} &= e^{t\hat{X}} \exp\left(-\sum_{n=1}^{\infty} \frac{(-t)^n}{n!} \alpha^{n-1} \hat{Y}\right) = e^{t\hat{X}} \exp\left[\alpha^{-1} \left(-\sum_{n=1}^{\infty} \frac{(-\alpha t)^n}{n!}\right) \hat{Y}\right] \\ &= e^{t\hat{X}} \exp\left(\alpha^{-1} (1 - e^{-\alpha t}) \hat{Y}\right) \end{aligned} \quad (10.30)$$

afin de retrouver l'identité de Sack (10.7).

# Chapitre 11

## Application au problème

On tente ici de résoudre le problème posé au chapitre 9 à l'aide des outils présentés au chapitre 10. Avant de s'attaquer au problème complet, on explore d'abord des problèmes connexes.

Ainsi, la section 11.1 présente un cas de désintégration pour lequel on obtient la solution classique, alors que l'on ne sait pas résoudre le processus d'annihilation présenté en section 11.2. Un cas intermédiaire est cependant résolu en section 11.3, et une généralisation de la désintégration est résolue en section 11.4. Le problème posé au chapitre 9 semble être un cas intermédiaire entre cette dernière généralisation de la désintégration et l'annihilation : la section 11.5 montre un cas pouvant être résolu, mais le problème d'intérêt échappe à nos moyens. La section 11.6 explique comment résoudre un processus de branchement en temps continu, puis démontre pourquoi la même méthode ne peut être utilisée pour le problème d'intérêt. Enfin, la section 11.7 constate l'échec de cette entreprise.

### 11.1 Désintégration spontanée d'un élément

Avant de tenter de résoudre les problèmes du chapitre 9, on s'attaque d'abord à des cas plus simples afin de montrer comment les outils présentés au chapitre 10 peuvent être utilisés.

La *désintégration* est l'un des plus simples processus de naissance-mort que l'on puisse imaginer : chacun des éléments du système peut, de façon individuelle, disparaître spontanément. On note souvent ce genre de processus sous la forme  $A \xrightarrow{1} \emptyset$ , que l'on lit

« tout élément de type  $A$  peut disparaître<sup>1</sup> à un taux<sup>2</sup> de 1 ».

En utilisant les opérateurs d'annihilation et de création agissant sur des états de la forme  $|\lambda\rangle$ ,  $\lambda$  représentant le nombre d'éléments, l'évolution dans le temps du système est gouverné par

$$\hat{L}_d = \hat{a} - \hat{a}^\dagger \hat{a} \quad . \quad (11.1)$$

À l'aide des définitions

$$\hat{X} = -\hat{a}^\dagger \hat{a} \quad \text{et} \quad \hat{Y} = \hat{a} \quad , \quad (11.2)$$

on observe

$$[\hat{X}, \hat{Y}] = [-\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a} = \hat{Y} \quad , \quad (11.3)$$

ce qui permet d'utiliser la relation de Sack (10.7) afin d'obtenir

$$e^{t\hat{L}_d} = e^{(e^t-1)\hat{a}} e^{-t\hat{a}^\dagger \hat{a}} \quad . \quad (11.4)$$

On explicite l'exponentielle de droite et constate que son effet sur l'état initial  $|\lambda\rangle$  est simplement de le multiplier par une constante

$$\begin{aligned} e^{t\hat{L}_d}|\lambda\rangle &= e^{(e^t-1)\hat{a}} \sum_{m=0}^{\infty} \frac{(-t)^m}{m!} (\hat{a}^\dagger \hat{a})^m |\lambda\rangle = e^{(e^t-1)\hat{a}} |\lambda\rangle \sum_{m=0}^{\infty} \frac{(-t)^m}{m!} \lambda^m \\ &= e^{(e^t-1)\hat{a}} |\lambda\rangle e^{-\lambda t} \quad . \end{aligned} \quad (11.5)$$

Dans le cas de l'exponentielle de gauche, plusieurs états différents sont obtenus

$$\begin{aligned} e^{t\hat{L}_d}|\lambda\rangle &= \sum_{\Delta=0}^{\infty} \frac{(e^t-1)^\Delta}{\Delta!} \hat{a}^\Delta |\lambda\rangle e^{-\lambda t} = \sum_{\Delta=0}^{\infty} |\lambda - \Delta\rangle \frac{(e^t-1)^\Delta}{\Delta!} \frac{\lambda!}{(\lambda - \Delta)!} e^{-\lambda t} \\ &= \sum_{\Delta=0}^{\infty} |\lambda - \Delta\rangle e^{-\lambda t} (e^t-1)^\Delta \binom{\lambda}{\lambda - \Delta} \quad . \end{aligned} \quad (11.6)$$

Le coefficient<sup>3</sup> multipliant  $|\lambda'\rangle$  correspond à la probabilité

$$P(\lambda', t|\lambda, 0) = e^{-\lambda t} (e^t-1)^{\lambda-\lambda'} \binom{\lambda}{\lambda'} \quad (11.7)$$

qu'il y ait  $\lambda'$  éléments au temps  $t$  sachant qu'il y en avait initialement  $\lambda$ . Il s'agit là de la solution exacte classique à ce simple problème.

1. Le  $\emptyset$  symbolise ici « rien » : là où il y avait  $A$ , il n'y a maintenant plus rien. D'un point de vue de modélisation, il peut rester un ou plusieurs éléments d'autres types, mais on ne s'y intéresse pas.

2. La probabilité qu'un élément donnée se désintègre pendant l'intervalle  $[t, t + dt)$  est de  $1 dt$ . Puisque chacun des éléments peut se désintégrer, la probabilité totale qu'il y ait une désintégration doit être multipliée par le nombre d'éléments restants.

3. En utilisant un produit de la forme (9.5), ce coefficient peut être noté  $\frac{1}{\lambda!} \langle \lambda' | e^{t\hat{L}_d} | \lambda \rangle$ .



## 11.2 Annihilation de deux éléments identiques

On s'intéresse ici à l'*annihilation* de deux éléments identiques, souvent notée  $A+A \xrightarrow{1} \emptyset$ . Puisque deux éléments doivent « se rencontrer », la probabilité qu'une annihilation ait lieu est proportionnelle au carré du nombre d'éléments (ou, plus précisément, à  $\lambda(\lambda-1)$ ). L'évolution dans le temps du système est gouverné par

$$\hat{L}_a = \hat{a}^2 - (\hat{a}^\dagger)^2 \hat{a}^2 \quad (11.8)$$

et on remarque que

$$[-(\hat{a}^\dagger)^2 \hat{a}^2, \hat{a}^2] = 4\hat{a}^\dagger \hat{a}^3 + 2\hat{a}^2 \quad . \quad (11.9)$$

Cette relation interdit l'utilisation de la règle de Sack, même sous sa forme généralisée (10.8).

En fait, aucune méthode connue ne permet d'obtenir une solution analytique exacte explicite à ce problème. Aussi, puisque les opérateurs  $\hat{L}'_\infty$  et  $\hat{L}''_\infty$  définis en section 9.5 ont une forme semblable, on ne s'attend pas à pouvoir les résoudre.

## 11.3 Cas mixte désintégration-annihilation

On considère un processus dans lequel les événements surviennent au même taux que dans un processus de désintégration, mais pour lequel la conséquence d'un tel événement est plutôt celle d'un processus d'annihilation. Ainsi, la probabilité que deux éléments disparaissent est proportionnelle au nombre d'éléments.

La combinaison des phénomènes  $A \xrightarrow{1} \gamma$  et  $A + \gamma \xrightarrow{\alpha} \emptyset$  fournit une représentation « physique » du problème dans la limite où  $\alpha$  est infini : les éléments  $A$  se désintègrent à un taux de 1, laissant derrière une « particule d'interaction »  $\gamma$  qui, au contact avec un élément  $A$ , cause leur annihilation mutuelle à un taux arbitrairement élevé. En notant  $\hat{g}^\dagger$  et  $\hat{g}$  les opérateurs de création et d'annihilation pour  $\gamma$ , l'évolution temporelle d'un tel système est gouvernée par

$$\hat{L}_{da} = (\hat{g}^\dagger - \hat{a}^\dagger)\hat{a} + \alpha(1 - \hat{a}^\dagger \hat{g}^\dagger)\hat{a}\hat{g} \quad . \quad (11.10)$$

Cependant, comme on ne sait pas résoudre les processus d'annihilation, cette approche n'a pas été explorée d'avantage.

Il s'avère que l'opérateur

$$\hat{L}''_{\lambda} = w\hat{a}^2 - \hat{a}^{\dagger}\hat{a} \quad (9.36b)$$

défini à la section 9.6 produit la même dynamique hybride désintégration-annihilation. Bien entendu, la correspondance n'est valide qu'après la substitution <sup>4</sup>

$$w^m \implies \frac{(\lambda - 2m - 1)!!}{(\lambda - 1)!!} \quad (11.11)$$

afin de préserver la normalisation. De plus, l'état initial  $|\lambda\rangle$  doit être tel que  $\lambda$  est un nombre pair positif.

En posant

$$\hat{X} = -\hat{a}^{\dagger}\hat{a} \quad \text{et} \quad \hat{Y} = w\hat{a}^2, \quad (11.12)$$

on observe

$$[\hat{X}, \hat{Y}] = [-\hat{a}^{\dagger}\hat{a}, w\hat{a}^2] = 2w\hat{a}^2 = 2\hat{Y} \quad (11.13)$$

ce qui permet d'utiliser la relation de Sack (10.7) pour obtenir

$$e^{t\hat{L}''_{\lambda}} = e^{\frac{1}{2}w(e^{2t}-1)\hat{a}^2} e^{-t\hat{a}^{\dagger}\hat{a}} \quad (11.14)$$

que l'on résout de la même façon que dans le cas de la désintégration

$$\begin{aligned} e^{t\hat{L}''_{\lambda}}|\lambda\rangle &= e^{\frac{1}{2}w(e^{2t}-1)\hat{a}^2} e^{-t\hat{a}^{\dagger}\hat{a}}|\lambda\rangle = \sum_m \frac{w^m(e^{2t}-1)^m}{2^m m!} \hat{a}^{2m}|\lambda\rangle e^{-\lambda t} \\ &= \sum_m |\lambda - 2m\rangle \frac{e^{-\lambda t}(e^{2t}-1)^m}{2^m m!} \frac{\lambda!}{(\lambda - 2m)!} w^m. \end{aligned} \quad (11.15)$$

On isole enfin la contribution à  $|\lambda'\rangle$  pour  $\lambda'$  pair et procède à la substitution de  $w^m$  afin d'obtenir

$$P(\lambda', t|\lambda, 0) = \frac{e^{-\lambda t}(e^{2t}-1)^{\frac{1}{2}(\lambda-\lambda')}}{2^{\frac{1}{2}(\lambda-\lambda')}(\frac{1}{2}(\lambda-\lambda'))!} \frac{\lambda! (\lambda' - 1)!!}{\lambda'! (\lambda - 1)!!} = e^{-\lambda t}(e^{2t}-1)^{\frac{1}{2}(\lambda-\lambda')} \begin{pmatrix} \frac{1}{2}\lambda \\ \frac{1}{2}\lambda' \end{pmatrix}. \quad (11.16)$$

On constate une forte ressemblance avec la solution obtenue pour la désintégration. En fait, les changements de variables  $\frac{1}{2}\lambda = \tilde{\lambda}$ ,  $\frac{1}{2}\lambda' = \tilde{\lambda}'$  et  $2t = \tilde{t}$  permettent de retrouver (11.7). Cette situation est due au fait que l'on aurait pu exprimer le problème sous la forme alternative  $(\hat{a}^{\dagger})^{-1}\hat{a} - \hat{a}^{\dagger}\hat{a}$ , puis considérer les éléments par paires comme un seul élément. La raison pour laquelle on n'a pas procédé de cette façon dès le début est de nature stœchiométrique. En effet, on désire obtenir une méthode pouvant être généralisée aux problèmes où  $\gamma$  interagit avec divers types d'éléments en fonction de leur nombre, comme c'est le cas avec  $\hat{L}'_w$  et  $\hat{L}'_{\lambda}$ .

4. Il s'agit d'une forme simplifiée de (9.38) valide seulement dans le cas  $\hat{L}''_{\lambda}$ .

## 11.4 Dégradation spontanée d'un élément

On entend ici par *dégradation* une chaîne d'événements de désintégration  $A_k \xrightarrow{k} A_{k-1}$  tel qu'un élément de type  $A_k$  se désintègre spontanément à un taux  $k$ , laissant derrière un élément de type  $A_{k-1}$  qui pourra à son tour se désintégrer de la même façon. L'évolution temporelle d'un tel système est gouvernée par

$$\hat{L}_D = \sum_k k(\hat{a}_{k-1}^\dagger - \hat{a}_k^\dagger)\hat{a}_k = \hat{A}_1^{01} - \hat{A}_0^1 \quad (11.17)$$

qui, à l'aide des définitions

$$\hat{X} = -\hat{A}_0^1 \quad \text{et} \quad \hat{Y} = \hat{A}_1^{01} = \hat{A}_1^0 + \hat{A}_1^1 \quad (11.18)$$

ainsi que de la relation

$$[\hat{X}, \hat{Y}] = [\hat{A}_1^0 + \hat{A}_1^1, \hat{A}_0^1] = [\hat{A}_1^0, \hat{A}_0^1] + [\hat{A}_1^1, \hat{A}_0^1] = \hat{A}_1^0 + \hat{A}_1^1 = \hat{Y} \quad , \quad (11.19)$$

peut être exprimée sous la forme

$$e^{t\hat{L}_D} = e^{(e^t - 1)\hat{A}_1^{01}} e^{-t\hat{A}_0^1} \quad (11.20)$$

à l'aide de la relation de Sack. Pour un  $|\mathbf{x}\rangle$  prenant la forme (9.4), la contribution de  $e^{-t\hat{A}_0^1}$  à  $e^{t\hat{L}_D}|\mathbf{x}\rangle$  équivaut simplement à une multiplication par une constante

$$\begin{aligned} e^{-t\hat{A}_0^1}|\mathbf{x}\rangle &= \sum_n \frac{(-t)^n}{n!} (\hat{A}_0^1)^n |\mathbf{x}\rangle = \sum_n \frac{(-t)^n}{n!} \left( \sum_k k \hat{a}_k^\dagger \hat{a}_k \right)^n |\mathbf{x}\rangle = |\mathbf{x}\rangle \sum_n \frac{(-t)^n}{n!} \left( \sum_k k i_k \right)^n \\ &= |\mathbf{x}\rangle e^{-t \sum_k k i_k} \quad . \end{aligned} \quad (11.21)$$

Cependant, le calcul du reste de l'expression n'est malheureusement pas aussi trivial. Afin de simplifier la tâche, on définit l'opérateur

$$(\hat{A}_1^{01})_l = \sum_{k=1}^l k \hat{a}_{k-1}^\dagger \hat{a}_k \quad (11.22)$$

tel que  $(\hat{A}_1^{01})_\infty = \hat{A}_1^{01}$ . Pour un  $l$  fixé, on définit

$$\hat{X} = (\hat{A}_1^{01})_{l-1} \quad \text{et} \quad \hat{Y}_n = \frac{l!}{(l-n)!} \hat{a}_{l-n}^\dagger \hat{a}_l \quad (11.23)$$

où il est sous-entendu que  $\hat{Y}_n = 0$  pour  $n \geq l$ . On constate que

$$[\hat{X}, \hat{Y}_n] = \sum_{k=1}^{l-1} k \frac{l!}{(l-n)!} \hat{a}_{k-1}^\dagger [\hat{a}_k, \hat{a}_{l-n}^\dagger] \hat{a}_l = \frac{l!}{(l-(n+1))!} \hat{a}_{l-(n+1)}^\dagger \hat{a}_l = \hat{Y}_n \quad (11.24)$$

et que  $[\hat{Y}_n, \hat{Y}_{n'}] = 0$ , rendant ainsi possible l'usage de la relation de Sack généralisée (10.8)

$$\begin{aligned} e^{\tau(\hat{A}_1^{01})_l} &= e^{\tau(\hat{A}_1^{01})_{l-1}} \exp \left( - \sum_{n=1}^l \frac{(-\tau)^n}{n!} \frac{l!}{(l-n)!} \hat{a}_{l-n}^\dagger \hat{a}_l \right) \\ &= e^{\tau(\hat{A}_1^{01})_{l-1}} \sum_m (-1)^m \left( \sum_{k=0}^{l-1} \binom{l}{k} (-\tau)^{l-k} \hat{a}_k^\dagger \right)^m \frac{\hat{a}_l^m}{m!} . \end{aligned} \quad (11.25)$$

On suppose ensuite l'existence d'un  $K$  arbitrairement élevé tel qu'aucun élément de degré supérieur à  $K$  n'est présent dans  $\mathbf{x}$ , ce qui implique  $\hat{A}_1^{01}|\mathbf{x}\rangle = (\hat{A}_1^{01})_K|\mathbf{x}\rangle$ . En notant  $|\mathbf{0}\rangle$  l'état tel que le système ne contient aucun élément, on peut écrire  $|\mathbf{x}\rangle = (\hat{a}_0^\dagger)^{i_0} \dots (\hat{a}_K^\dagger)^{i_K} |\mathbf{0}\rangle$ . En appliquant (11.25) sur un état de cette forme, on obtient

$$\begin{aligned} e^{\tau \hat{A}_1^{01}} |\mathbf{x}\rangle & \quad (11.26) \\ &= e^{\tau(\hat{A}_1^{01})_{K-1}} \sum_m (-1)^m \sum_{\mathbf{j}^K} \binom{m}{\mathbf{j}^K} \left[ \prod_{k=0}^{K-1} \binom{K}{k}^{j_k^K} (-\tau)^{(K-k)j_k^K} (\hat{a}_k^\dagger)^{i_k + j_k^K} \right] \frac{\hat{a}_K^m}{m!} (\hat{a}_K^\dagger)^{i_K} |\mathbf{0}\rangle \end{aligned}$$

où

$$\binom{m}{\mathbf{j}^K} = \frac{m!}{j_0^K! j_1^K! \dots j_{K-1}^K!} \quad (11.27)$$

est un coefficient multinomial et où le vecteur  $\mathbf{j}^K = [j_0^K \ j_1^K \ \dots \ j_{K-1}^K]^T$  respecte la contrainte  $\sum_{k=0}^{K-1} j_k^K = m$ . On cherche ensuite la contribution de (11.26) à l'état  $|\mathbf{x}'\rangle = (\hat{a}_0^\dagger)^{i'_0} \dots (\hat{a}_K^\dagger)^{i'_K} |\mathbf{0}\rangle$ , c'est-à-dire le terme pour lequel  $m = i_K - i'_K$

$$e^{\tau(\hat{A}_1^{01})_{K-1}} \sum_{\mathbf{j}^K} (-1)^{i_K - i'_K} \binom{i_K - i'_K}{\mathbf{j}^K} \left[ \prod_{k=0}^{K-1} \binom{K}{k}^{j_k^K} (-\tau)^{(K-k)j_k^K} (\hat{a}_k^\dagger)^{i_k + j_k^K} \right] \binom{i_K}{i'_K} . \quad (11.28)$$

On peut ensuite utiliser le même processus pour  $e^{\tau(\hat{A}_1^{01})_{K-1}}$ , puis  $e^{\tau(\hat{A}_1^{01})_{K-2}}, \dots$ , « épiluchant » ainsi l'opérateur  $\hat{A}_1^{01}$  terme à terme. On pose l'Ansatz

$$\begin{aligned} e^{\tau(\hat{A}_1^{01})_{L-1}} & \sum_{\mathbf{j}^K} \sum_{\mathbf{j}^{K-1}} \dots \sum_{\mathbf{j}^L} (-1)^{\sum_{k=L}^K (i_k - i'_k + \sum_{l=k+1}^K j_k^l)} \\ & \times \left\{ \prod_{k=L}^K \binom{i_k + \sum_{l=k+1}^K j_k^l}{i'_k} \binom{i_k - i'_k + \sum_{l=k+1}^K j_k^l}{\mathbf{j}^k} \left[ \prod_{l=k+1}^K \binom{l}{k}^{j_k^l} \right] (-\tau)^{\sum_{l=k+1}^K (l-k)j_k^l} \right\} \\ & \times \left\{ \prod_{k=0}^{L-1} \left[ \prod_{l=L}^K \binom{l}{k}^{j_k^l} \right] (-\tau)^{\sum_{l=L}^K (l-k)j_k^l} (\hat{a}_k^\dagger)^{i_k + \sum_{l=L}^K j_k^l} \right\} \end{aligned} \quad (11.29)$$

pour le résultat après  $K - L$  itérations et on vérifie qu'une itération supplémentaire retourne bel et bien (11.29) avec  $L$  remplacé par  $L - 1$ . On note d'abord que la dernière

ligne de (11.29)

$$\begin{aligned} & \left\{ \prod_{k=0}^{L-1} \left[ \prod_{l=L}^K \binom{l}{k}^{j_k^l} \right] (-\tau)^{\sum_{l=L}^K (l-k)j_k^l} (\hat{a}_k^\dagger)^{i_k + \sum_{l=L}^K j_k^l} \right\} \\ & = \left\{ \prod_{k=0}^{L-2} \dots \right\} \left[ \prod_{l=L}^K \binom{l}{L-1}^{j_{L-1}^l} \right] (-\tau)^{\sum_{l=L}^K (l-(L-1))j_{L-1}^l} (\hat{a}_{L-1}^\dagger)^{i_{L-1} + \sum_{l=L}^K j_{L-1}^l} \end{aligned} \quad (11.30)$$

fournit quelques uns des termes de la seconde ligne en plus des seuls opérateurs  $\hat{a}_{L-1}^\dagger$  présents dans l'expression. On utilise ensuite (11.25) afin de connaître la contribution de  $e^{\tau(\hat{A}_1^{01})_{L-1}}$  sur ces opérateurs (toujours en ne collectant que les termes contribuant à  $|\mathbf{x}'\rangle$ )

$$\begin{aligned} & e^{\tau(\hat{A}_1^{01})_{L-1}} (\hat{a}_{L-1}^\dagger)^{i_{L-1} + \sum_{l=L}^K j_{L-1}^l} \\ & = e^{\tau(\hat{A}_1^{01})_{L-2}} \sum_{\mathbf{j}^{L-1}} (-1)^{i_{L-1} - i'_{L-1} + \sum_{l=L}^K j_{L-1}^l} \binom{i_{L-1} - i'_{L-1} + \sum_{l=L}^K j_{L-1}^l}{\mathbf{j}^{L-1}} \\ & \quad \times \left[ \prod_{k=0}^{L-2} \binom{L-1}{k}^{j_k^{L-1}} (-\tau)^{((L-1)-k)j_k^{L-1}} (\hat{a}_k^\dagger)^{j_k^{L-1}} \right] \binom{i_{L-1} + \sum_{l=L}^K j_{L-1}^l}{i'_{L-1}} \end{aligned} \quad (11.31)$$

ce qui permet de retrouver (11.29) et de vérifier l'*Ansatz*.

En utilisant  $L = 1$ , on obtient

$$\begin{aligned} & \sum_{\mathbf{j}^K} \sum_{\mathbf{j}^{K-1}} \dots \sum_{\mathbf{j}^1} (-1)^{\sum_{k=1}^K (i_k - i'_k + \sum_{l=k+1}^K j_k^l)} \\ & \quad \times \left\{ \prod_{k=1}^K \binom{i_k + \sum_{l=k+1}^K j_k^l}{i'_k} \binom{i_k - i'_k + \sum_{l=k+1}^K j_k^l}{\mathbf{j}^k} \left[ \prod_{l=k+1}^K \binom{l}{k}^{j_k^l} \right] (-\tau)^{\sum_{l=k+1}^K (l-k)j_k^l} \right\} \\ & \quad \times \left\{ (-\tau)^{\sum_{l=1}^K l j_0^l} (\hat{a}_0^\dagger)^{i_0 + \sum_{l=1}^K j_0^l} \right\} . \end{aligned} \quad (11.32)$$

La contrainte sur les  $j_k^l$  prend la forme

$$\sum_{k=0}^{l-1} j_k^l = i_l - i'_l + \sum_{l'=l+1}^K j_l^{l'} , \quad \text{ce qui donne} \quad j_0^l = i_l - i'_l + \sum_{l'=l+1}^K j_l^{l'} - \sum_{k=1}^{l-1} j_k^l . \quad (11.33)$$

Ainsi, puisque seul le terme contenant  $(\hat{a}_0^\dagger)^{i_0}$  nous intéresse, on a

$$\begin{aligned} i'_0 & = i_0 + \sum_{l=1}^K j_0^l = i_0 + \sum_{l=1}^K \left( i_l - i'_l + \sum_{l'=l+1}^K j_l^{l'} - \sum_{k=1}^{l-1} j_k^l \right) \\ & = i_0 + \sum_{l=1}^K (i_l - i'_l) \end{aligned} \quad (11.34)$$

ce qui est en accord avec la conservation des éléments  $\sum_k (i_k - i'_k) = 0$ . L'expression prend alors la forme légèrement simplifiée

$$\sum_{\mathbf{j}^K} \sum_{\mathbf{j}^{K-1}} \dots \sum_{\mathbf{j}^1} (-1)^{\sum_{k=0}^K \sum_{l=k+1}^K j_k^l} (-\tau)^{\sum_{l=1}^K l j_0^l} \\ \times \left\{ \prod_{k=1}^K \binom{i_k + \sum_{l=k+1}^K j_k^l}{i'_k} \binom{i_k - i'_k + \sum_{l=k+1}^K j_k^l}{\mathbf{j}^k} \left[ \prod_{l=k+1}^K \binom{l}{k}^{j_k^l} \right] (-\tau)^{\sum_{l=k+1}^K (l-k) j_k^l} \right\} \quad (11.35)$$

et, en retournant au problème complet, on obtient la solution

$$P(\mathbf{x}', t | \mathbf{x}, 0) = e^{-t \sum_{k=1}^K k i_k} \sum_{\mathbf{j}^K} \sum_{\mathbf{j}^{K-1}} \dots \sum_{\mathbf{j}^1} (-1)^{\sum_{k=0}^K \sum_{l=k+1}^K j_k^l} (1 - e^{-t})^{\sum_{l=1}^K l j_0^l} \\ \times \left\{ \prod_{k=1}^K \binom{i_k + \sum_{l=k+1}^K j_k^l}{i'_k} \binom{i_k - i'_k + \sum_{l=k+1}^K j_k^l}{\mathbf{j}^k} \left[ \prod_{l=k+1}^K \binom{l}{k}^{j_k^l} \right] (1 - e^{-t})^{\sum_{l=k+1}^K (l-k) j_k^l} \right\}. \quad (11.36)$$

Outre la conversion de ces multiples sommes sous forme de séries hypergéométriques, il ne semble pas possible de simplifier l'expression davantage.

## 11.5 Cas mixte dégradation-annihilation

De la même façon que la section 11.3 présente un cas hybride entre la désintégration et l'annihilation, on s'intéresse ici au cas mixte dégradation-annihilation où  $A_k \xrightarrow{k} A_{k-1} + \gamma$  et  $A_k + \gamma \xrightarrow{k\alpha} A_{k-1}$  dans la limite  $\alpha$  infini. L'évolution temporelle du processus est donc gouvernée par

$$\hat{L}_{Da} = (\hat{g}^\dagger \hat{A}_1^{01} - \hat{A}_0^1) + \alpha (\hat{A}_1^{01} - \hat{A}_0^1 \hat{g}^\dagger) \hat{g} \quad . \quad (11.37)$$

Pour les mêmes raisons qu'en section 11.3, on considère plutôt l'opérateur d'évolution

$$\hat{L}_w'' = w (\hat{A}_1^{01})^2 - \hat{A}_0^1 \quad . \quad (9.32c)$$

Les définitions

$$\hat{X} = -\hat{A}_0^1 \quad \text{et} \quad \hat{Y} = w (\hat{A}_1^{01})^2 \quad (11.38)$$

ont pour commutateur

$$[\hat{X}, \hat{Y}] = \left[ -\hat{A}_0^1, w (\hat{A}_1^{01})^2 \right] = 2w (\hat{A}_1^{01})^2 = 2\hat{Y} \quad (11.39)$$

et ils permettent donc d'utiliser l'identité de Sack (10.7)

$$e^{t\hat{L}'_w} = \exp\left(\frac{1}{2}w(e^{2t}-1)(\hat{A}_1^{01})^2\right)e^{-t\hat{A}_0^1} \quad . \quad (11.40)$$

L'effet du terme  $e^{-t\hat{A}_0^1}$  sur un état  $|\mathbf{x}\rangle$  a déjà été obtenu en section 11.4. De plus, il est possible d'exprimer  $e^{(\hat{A}_1^{01})^2}$  en terme de  $e^{\hat{A}_1^{01}}$ , ce dernier ayant également été calculé dans cette même section. En effet, pour un  $\hat{X}$  quelconque, on observe que

$$\begin{aligned} \cosh(\sqrt{\tau w'}\hat{X}) &= \frac{e^{\sqrt{\tau w'}\hat{X}} + e^{-\sqrt{\tau w'}\hat{X}}}{2} = \sum_l \frac{(\tau w')^{\frac{l}{2}}}{l!} \hat{X}^l \frac{[1 + (-1)^l]}{2} \\ &= \sum_k \frac{\tau^k w'^k}{(2k)!} \hat{X}^{2k} \end{aligned} \quad (11.41a)$$

devient

$$e^{\tau\hat{X}^2} = \sum_k \frac{\tau^k}{k!} \hat{X}^{2k} \quad (11.41b)$$

après la transformation

$$w'^m \implies \frac{(2m)!}{m!} \quad . \quad (11.41c)$$

On peut donc exprimer  $e^{(\tau\hat{A}_1^{01})^2}$  comme  $\cosh(\sqrt{\tau w'}\hat{A}_1^{01})$  et compléter l'évaluation exacte et explicite de  $e^{t\hat{L}'_w}|\mathbf{x}\rangle$ .

Les choses se compliquent pour

$$\hat{L}'_w = w(\hat{A}_1^{01} + \hat{D}_1^{01})\hat{A}_1^{01} - \hat{A}_0^1 \quad , \quad (9.32b)$$

que l'on peut interpréter comme  $A_k \xrightarrow{k} A_{k-1} + \gamma$ ,  $A_k + \gamma \xrightarrow{k\alpha} A_{k-1}$  et  $B_k + \gamma \xrightarrow{k\alpha} A_{k-1}$  dans la limite où  $\alpha$  est infini. En effet, l'ajout du terme  $\hat{D}_1^{01}\hat{A}_1^{01}$  interdit l'utilisation de l'identité de Sack, ou de sa généralisation, afin de séparer ces trois opérateurs. À ce jour, toutes tentatives visant à « épilucher » ces opérateurs terme à terme (tel que fait pour  $\hat{A}_1^{01}$  en section 11.4) sans d'abord les séparer les uns des autres se sont soldées par un échec. Sous la forme  $\hat{L}'_w$ , il est difficile de savoir si, comme dans le cas de l'annihilation, on ne peut le résoudre. Tel que discuté dans la section suivante, les choses sont plus claires avec  $\hat{L}'_\lambda$ .

## 11.6 Processus de branchement

Un processus de branchement en temps continu prend la forme  $A \xrightarrow{\alpha_j} A^j$  pour  $j \neq 1$  (le cas  $j = 1$  gardant le système inchangé). L'évolution du processus est donnée par

$$\hat{L}_b = \sum_j \alpha_j (\hat{a}^\dagger)^j \hat{a} \quad (11.42)$$

où on demande typiquement que  $\alpha_1 = -\sum_{j \neq 1} \alpha_j$  afin d'obtenir  $\sum_{j \neq 1} \alpha_j ((\hat{a}^\dagger)^j - \hat{a}^\dagger) \hat{a}$ . Cependant, afin de ne pas interdire le « truc » de substitution de  $w^m$ , on n'impose pas une telle contrainte sur  $\alpha_1$ .

Plutôt que de tenter de séparer  $\hat{L}_b$  à l'aide de la relation de Sack, on se tourne plutôt vers une méthode plus standard fondée sur les correspondances  $\hat{a}^\dagger \Leftrightarrow z$  et  $\hat{a} \Leftrightarrow \frac{\partial}{\partial z}$  que l'on vérifie à l'aide de la fonction test  $f(z)$

$$\left[ \frac{\partial}{\partial z}, z \right] f(z) = \frac{\partial}{\partial z} z f(z) - z \frac{\partial}{\partial z} f(z) = f(z), \quad \text{donc} \quad \left[ \frac{\partial}{\partial z}, z \right] = 1 \quad . \quad (11.43)$$

La correspondance supplémentaire  $|0\rangle \Leftrightarrow 1$  permet d'écrire tout état  $|\lambda\rangle = (\hat{a}^\dagger)^\lambda |0\rangle$  comme  $z^\lambda$ . Ainsi, la fonction génératrice  $\phi(z, t)$  doit répondre à l'équation différentielle partielle

$$\frac{\partial \phi(z, t)}{\partial t} = \left( \sum_j \alpha_j z^j \right) \frac{\partial \phi(z, t)}{\partial z} \quad ; \quad (11.44)$$

la condition initiale fixe  $\phi(z, 0)$  et la solution au temps  $t$  peut être reconvertie en terme de kets à l'aide de  $\phi(\hat{a}^\dagger, t)|0\rangle$ . Puisque l'équation est hyperbolique, on peut utiliser la méthode des caractéristiques afin de la résoudre.

On introduit donc un paramètre  $\tau$  permettant d'écrire  $\tilde{\phi}(\tau) = \phi(\tilde{z}(\tau), \tilde{t}(\tau))$  où les  $\tilde{z}(\tau)$ ,  $\tilde{t}(\tau)$  et  $\tilde{\phi}(\tau)$  respectent les équations caractéristiques

$$\frac{d\tilde{z}(\tau)}{d\tau} = \sum_j \alpha_j (\tilde{z}(\tau))^j \quad (11.45a)$$

$$\frac{d\tilde{t}(\tau)}{d\tau} = -1 \quad (11.45b)$$

$$\frac{d\tilde{\phi}(\tau)}{d\tau} = 0 \quad . \quad (11.45c)$$

Deux de ces équations sont triviales à résoudre

$$\tilde{t}(\tau) = \tilde{t}(0) - \tau \quad \tilde{\phi}(\tau) = \tilde{\phi}(0) \quad (11.46)$$



et la constance de  $\tilde{\phi}(\tau)$  permet d'écrire

$$\phi(\tilde{z}(0), \tilde{t}(0)) = \tilde{\phi}(0) = \tilde{\phi}(\tau) = \phi(\tilde{z}(\tau), \tilde{t}(\tau)) \quad . \quad (11.47)$$

En fixant les constantes d'intégration  $\tilde{z}(0) = z$  et  $\tilde{t}(0) = t$ , on peut évaluer  $\phi(z, t)$  en « suivant » la courbe caractéristique correspondante jusqu'à la condition initiale  $\phi(z, 0)$

$$\phi(z, t) = \phi(\tilde{z}(t), 0) \quad . \quad (11.48)$$

Résoudre un processus de branchement se résume donc à intégrer l'équation différentielle ordinaire (11.45a) avec la condition initiale  $\tilde{z}(0) = z$  afin d'évaluer  $\phi(\tilde{z}(t), 0)$ .

Une inspection rapide de l'équation

$$\hat{L}'_{\lambda} = w\hat{a}^2 + \sum_k kwv_k(\hat{a}^{\dagger})^{k-1}\hat{a} - \hat{a}^{\dagger}\hat{a} \quad (9.36a)$$

révèle que  $\hat{L}'_{\lambda}$  a *presque* la forme d'un processus de branchement, si ce n'était du terme  $w\hat{a}^2$ . Cette simple différence est cependant importante puisqu'elle rend l'équation différentielle partielle correspondante parabolique

$$\frac{\partial\phi(z, t)}{\partial t} = w\frac{\partial^2\phi(z, t)}{\partial z^2} + \left(\sum_k wv_k z^{k-1} - z\right)\frac{\partial\phi(z, t)}{\partial z} \quad , \quad (11.49)$$

interdisant l'usage de la méthode des caractéristiques. Lorsque le degré des éléments ne peut dépasser  $K = 2$ , on retrouve un cas semblable à l'hybride désintégration-annihilation et on peut résoudre le système. Cependant, ce cas est inintéressant puisque le réseau correspondant ne peut former que des chaînes ou des boucles. En fait, le premier cas intéressant correspondrait à  $K = 3$ .

## 11.7 Constat d'échec

On a vu en sections 11.1 et 11.2 qu'il était trivial de résoudre un problème de désintégration, mais que l'on ne savait pas résoudre un problème d'annihilation. Cependant, la section 11.3 montre qu'un cas mixte désintégration-annihilation peut être résolu. Puisque le problème présenté au chapitre 9 semble être de cette dernière catégorie, il est *a priori* justifié d'y chercher une solution analytique exacte et explicite.

Malgré que l'on soit capable de résoudre  $\hat{L}''_w$  et  $\hat{L}''_{\lambda}$ , ces cas n'ont pas vraiment d'applications pratiques dans le cadre du problème présenté au chapitre 9. Lorsqu'on

tente d'appliquer les mêmes méthodes aux cas plus complets  $\hat{L}'_w$  et  $\hat{L}'_\lambda$ , des problèmes surgissent.

Le plus simple cas « intéressant » que l'on puisse imaginer est celui du processus *SI* avec condition initiale

$$\left\langle \begin{array}{cccc} 0 & i_1 & 0 & 0 \\ 0 & s_1 & 0 & s_3 \\ 0 & 0 & 0 & 0 \end{array} \right\rangle . \quad (11.50)$$

Le processus se limite donc aux transitions

$$A_1 \xrightarrow{1} \gamma \quad A_2 \xrightarrow{2} A_1 + \gamma \quad B_1 + \gamma \xrightarrow{\alpha} \emptyset \quad B_3 + \gamma \xrightarrow{3\alpha} A_2 \quad (11.51)$$

dans la limite où  $\alpha$  est infini<sup>5</sup>, et la plus simple représentation de ce processus que l'on connaisse est

$$\hat{L}_\lambda^* = w (v_1 + 3v_3(\hat{a}^\dagger)^2 + \hat{a}) \hat{a} - \hat{a}^\dagger \hat{a} \quad (11.52)$$

où  $w$ ,  $v_1$  et  $v_3$  doivent être substitués à la fin du calcul (voir section 9.6). Or, cette expression ne peut pas être séparée par la règle de Sack, et on ne sait pas résoudre l'équation différentielle partielle résultante. De plus, tenter de construire un carré parfait afin d'utiliser la relation (11.41) ne semble pas pouvoir aider à la situation. Si un seul des deux termes  $(\hat{a}^\dagger)^2 \hat{a}$  et  $\hat{a}^2$  était absent, on saurait résoudre le problème. Cependant, sous la forme (11.52), la solution nous échappe.

Puisqu'il s'agit du plus simple cas d'intérêt que l'on puisse imaginer, la tentative de solution du problème posé au chapitre 9 est donc un échec. Néanmoins, certains des résultats intermédiaires et/ou méthodes de solution pourraient être utiles dans le cadre d'un problème différent.

---

5. On peut écrire  $B_1 + \gamma \xrightarrow{\alpha} \emptyset$  plutôt que  $B_1 + \gamma \xrightarrow{\alpha} A_0$  puisqu'on ne s'intéresse pas aux  $A_0$ .

# Chapitre 12

## Conclusion

On effectue un bref retour sur les contributions originales présentées dans les trois parties de cette thèse et donne quelques perspectives futures relatives à ces travaux.

### 12.1 Processus de branchement et évolution dans le temps

- La partie I de cette thèse présente un modèle de propagation sur réseaux considérant
- l'hétérogénéité du degré des nœuds composant le réseau (distribution en degré) ;
  - la nature stochastique du processus de propagation et/ou de la structure du réseau ;
  - le nombre fini de nœuds composant le réseau ;
  - l'évolution dans le temps du processus (de façon discrète au chapitre 3 et de façon continue en section 4.2) ; et
  - les corrélations dues aux périodes infectieuses (section 4.1).

Tel que discuté en section 4.3, les moyens pris pour parvenir à ces fins n'étaient peut-être pas les meilleurs. Malgré tout, l'expérience acquise lors de ces travaux a eu une influence positive importante sur la partie II de ce document ainsi que sur certaines contributions ne faisant pas explicitement partie de cette thèse (voir page xxi).

## 12.2 Processus de Markov et représentation partielle

La partie II constitue la contribution principale de cette thèse. En plus de pouvoir tenir compte des mêmes phénomènes qu'en partie I, la méthode générale qui est décrite au chapitre 7 permet

- la représentation d'une grande variété de caractéristiques typiques des réseaux complexes (à l'aide de motifs) ;
- la description de processus dynamiques affectant aussi bien la structure même du réseau que l'état intrinsèque de ses composantes ;
- l'élaboration d'une procédure systématique fournissant une approximation de la distribution de probabilités ; et
- l'usage de méthodes alternatives d'inférence et/ou de modélisation (sections 8.1 et 8.2)

Des avancées plus spécifiques ont également été effectuées, notamment l'étude du processus à-la-volée (chapitre 6) ainsi que l'introduction d'une méthode permettant de séparer la distribution de probabilité pour les épidémies de celle des flambées épidémiques (section 6.3.2).

Une étape importante dans la continuation de ces travaux consiste à appliquer ces méthodes générales à des problèmes spécifiques biens réels. Les modèles qui en résulteront devraient aider à mieux comprendre ces systèmes complexes, permettant entre autres d'orienter la prise de décisions les concernant. Au cours de cette démarche, de nouveaux outils devront également être développés afin de mieux tenir compte de la structure et de la dynamique particulière de ces systèmes, poursuivant ainsi notre quête de réalisme et d'efficacité.

## 12.3 Approche opérationnelle

La partie III vise l'obtention d'une solution analytique exacte et explicite du processus à-la-vollée présenté aux chapitres 6 et 7. Malgré que l'entreprise se soit soldé par un échec, plusieurs contributions<sup>1</sup> ont tout de même été réalisées :

- la généralisation (10.8) de la règle de Sack ;

---

1. Puisque la partie III fait appel à plusieurs concepts ne relevant pas du sujet principal de cette thèse, la littérature correspondante n'a pas été couverte avec autant d'attention que pour les parties I et II. Par conséquent, il se peut que certaines de ces contributions ne soient pas inédites.

- l'utilisation à la section 11.4 de la règle de Sack généralisée afin « d'éplucher » un opérateur ;
- le « truc » de substitution (par exemple de  $w^m$ ) introduit à la section 9.4 ;
- la relation 11.41 permettant d'exprimer l'exponentielle d'un opérateur en terme de l'exponentielle de la racine de cet opérateur ;
- l'élaboration aux sections 11.3 et 11.5 d'une méthode de solution pour des cas mixtes désintégration-annihilation ; et
- la mise en évidence des difficultés inhérentes à la solution exacte du modèle à-la-vollée.

L'impasse rencontrée limite les répercussions futures que pourraient avoir ces travaux. Cependant, il est tout de même possible que les outils et méthodes développés puissent servir à résoudre un système différent. De plus, la technique utilisée afin d'obtenir l'identité de Sack généralisée (section 10.4) pourrait peut-être fournir d'autres relations du même genre.



# Bibliographie

- [1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74 :47–97, 2002.
- [2] A. Allard, L. Hébert-Dufresne, P.-A. Noël, V. Marceau, and L. Dubé. Bond percolation on a class of correlated and clustered random graphs. *J. Phys. A*, 45 :405005, 2012.
- [3] A. Allard, L. Hébert-Dufresne, P.-A. Noël, V. Marceau, and L. Dubé. Exact solution of bond percolation on small arbitrary graphs. *EPL*, 98 :16001, 2012.
- [4] A. Allard, P.-A. Noël, L. J. Dubé, and B. Pourbohloul. Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. *Phys. Rev. E*, 79 :036113, 2009.
- [5] R. M. Anderson and R. M. May. *Infectious Diseases of Humans : Dynamics and Control*. Oxford University Press, Oxford, 1991.
- [6] A. H. Auchincloss and A. V. Diez Roux. A new tool for epidemiology : The usefulness of dynamic-agent models in understanding place effects on health. *Am. J. Epidemiol.*, 168(1) :1–8, 2008.
- [7] S. Bansal, B. Grenfell, and L. Meyers. When individual behaviour matters : homogeneous and network models in epidemiology. *J. R. Soc. Interface*, 4 :879–891, 2007.
- [8] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439) :509–512, 1999.
- [9] A. Barrat, M. Barthélemy, and A. Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, New York, 2008.
- [10] J. Bascompte and D. B. Stouffer. The assembly and disassembly of ecological networks. *Phil. Trans. R. Soc. B*, 364(1524) :1781–1787, 2009.
- [11] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks : Structure and dynamics. *Phys. Rep.*, 424 :175–308, 2006.
- [12] B. Bollobas. *Random Graphs (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 2nd edition, 2001.

- [13] F. Brauer and C. Castillo-Chávez. *Mathematical Models in Population Biology and Epidemiology*. Springer-Verlag, New York, 2001.
- [14] W. Broeck, C. Gioannini, B. Goncalves, M. Quaggiotto, V. Colizza, and A. Vespignani. The gleamviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infect. Dis.*, 11(1) :37, 2011.
- [15] J. K. Cavers. On the fast Fourier transform inversion of probability generating functions. *IMA J. Appl. Math.*, 22(3) :275–282, 1978.
- [16] R. Cohen and S. Havlin. *Complex Networks : Structure, Robustness and Function*. Cambridge University Press, 2010.
- [17] C. E. Dangerfield, J. V. Ross, and M. J. Keeling. Integrating stochasticity and network structure in an epidemic model. *J. R. Soc. Interface*, 6 :761–774, 2009.
- [18] L. Danon, A. Ford, T. House, C. Jewell, M. Keeling, G. Roberts, J. Ross, and M. Vernon. Networks and the epidemiology of infectious disease. *Interdiscip. Perspect. Infect. Dis.*, 2011 :284909–1–28, 2011.
- [19] L. Decreusefond, J. Dhersin, P. Moyal, and C. V. Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, 22 :541–575, 2012.
- [20] O. Diekmann and J. A. P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases : Model Building, Analysis and Interpretation*. John Wiley & Sons Ltd, 2000.
- [21] R. Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, 3rd edition, 2006.
- [22] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks : From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford, 2003.
- [23] J. A. Dunne and R. J. Williams. Cascading extinctions and community collapse in model food webs. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 364 :1711–1723, 2009.
- [24] E. B. Dynkin. Calculation of the coefficients in the campbell-hausdorff formula. *Doklady Akad. Nauk USSR (N.S.)*, 57 :367–376, 1947. Traduction anglaise dans Evgeniï Borisovich Dynkin, G. M. Seitz et A. L. Onishchik. *Selected papers of E.B. Dynkin with commentary*. American Mathematical Society, 2000.
- [25] K. T. D. Eames. Modelling disease spread through random and regular contacts in clustered populations. *Theo. Pop. Bio.*, 73 :104–111, 2008.
- [26] K. T. D. Eames and M. J. Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *PNAS*, 99 :13330–13335, 2002.
- [27] D. J. D. Earn, P. Rohani, B. M. Bolker, and B. T. Grenfell. A simple model for complex dynamical transitions in epidemics. *Science*, 287 :667–670, 2000.



- [28] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets : Reasoning About a Highly Connected World*. Cambridge University Press, 2010.
- [29] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6 :290–297, 1959.
- [30] K. Erdmann and M. Wildon. *Introduction to Lie Algebras*. Springer undergraduate mathematics series. Springer, 2006.
- [31] M. Frigo and S. G. Johnson. The design and implementation of FFTW3. *Proc. IEEE*, 93(2) :216–231, 2005.
- [32] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer-Verlag, 2004.
- [33] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25) :2340–2361, 1977.
- [34] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *PNAS*, 99(12) :7821–7826, 2002.
- [35] J. P. Gleeson. High-accuracy approximation of binary-state dynamics on networks. *Phys. Rev. Lett.*, 107 :068701, 2011.
- [36] A. Grabowski and R. A. Kosiński. Life span in online communities. *Phys. Rev. E*, 82 :066108, 2010.
- [37] L. Hébert-Dufresne, P.-A. Noël, V. Marceau, A. Allard, and L. J. Dubé. Propagation dynamics on networks featuring complex topologies. *Phys. Rev. E*, 82(3) :036115, 2010.
- [38] H. S. Hethcote. The mathematics of infectious diseases. *SIAM Rev.*, 42(4) :599–653, 2000.
- [39] P. Holme. Efficient local strategies for vaccination and network attack. *Europhys. Lett.*, 68(6) :908, 2004.
- [40] T. House, G. Davies, L. Danon, and M. J. Keeling. A motif-based approach to network epidemics. *Bull. Math. Biol.*, 71 :1693–1706, 2009.
- [41] L. Huang, K. Park, and Y.-C. Lai. Information propagation on modular networks. *Phys. Rev. E*, 73(3) :035103, 2006.
- [42] F. Iachello. *Lie Algebras and Applications*. Lecture Notes in Physics. Springer, 2006.
- [43] E. T. Jaynes. *Probability Theory : The Logic of Science*. Cambridge University Press, 2003.
- [44] B. Karrer and M. E. J. Newman. Random graphs containing arbitrary distributions of subgraphs. *Phys. Rev. E*, 82(6) :066118, 2010.
- [45] M. Karsai, M. Kivela, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world : How network topology and burstiness slow down spreading. *Phys. Rev. E*, 83(2) :025102, 2011.

- [46] M. Keeling and K. Eames. Networks and epidemic models. *J. R. Soc. Interface*, 2(4) :295–307, 2005.
- [47] M. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.
- [48] M. J. Keeling, D. A. Rand, and A. J. Morris. Correlation models for childhood epidemics. *Proc. R. Soc. B*, 264(1385) :1149–1156, 1997.
- [49] E. Kenah and J. M. Robins. Second look at the spread of epidemics on networks. *Phys. Rev. E*, 76 :036113, 2007.
- [50] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Int. J. Comput. Math.*, 2(1-4) :157–168, 1968. Originellement publié dans *Problemy Peredachi Informatsii*, 1(1) :3–11, 1965.
- [51] R. Lambiotte, S. Thurner, and R. Hanel. Unanimity rule on networks. *Phys. Rev. E*, 76(4) :046101, 2007.
- [52] P. Lind, L. da Silva, J. Andrade, and H. Herrmann. Spreading gossip in social networks. *Phys. Rev. E*, 76(3) :036117, 2007.
- [53] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.*, 7(4) :649–673, 1954.
- [54] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, and L. Dubé. Modeling the dynamical interaction between epidemics on overlay networks. *Phys. Rev. E*, 84(2) :026105, 2011.
- [55] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé. Adaptive networks : Coevolution of disease and topology. *Phys. Rev. E*, 82(3) :036116, 2010.
- [56] M. Marder. Dynamics of epidemics on random networks. *Phys. Rev. E*, 75 :066103, 2007.
- [57] A. J. McLane, C. Semeniuk, G. J. McDermid, and D. J. Marceau. The role of agent-based models in wildlife ecology and management. *Ecol. Model.*, 222(8) :1544 – 1556, 2011.
- [58] L. A. Meyers, B. Pourbohloul, M. E. J. Newman, D. M. Skowronski, and R. C. Brunham. Network theory and SARS : predicting outbreak diversity. *J. Theor. Biol.*, 232 :71–81, 2005.
- [59] J. C. Miller. Percolation and epidemics in random clustered networks. *Phys. Rev. E*, 80 :020901, 2009.
- [60] J. C. Miller. Spread of infectious disease through clustered populations. *J. R. Soc. Int.*, 6(41) :1121–1134, 2009.
- [61] J. C. Miller. A note on a paper by Erik Volz : SIR dynamics in random networks. *J. Math. Biol.*, 62(3) :349–358, 2010.

- [62] J. C. Miller, A. C. Slim, and E. M. Volz. Edge-based compartmental modeling for infectious disease spread. *J. R. Soc. Interface*, 2011. doi :10.1098/rsif.2011.0403.
- [63] G. Miritello, E. Moro, and R. Lara. Dynamical strength of social ties in information spreading. *Phys. Rev. E*, 83(4) :045102, 2011.
- [64] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Rand. Struct. Algor.*, 6(2-3) :161–180, 1995.
- [65] Y. Moreno, M. Nekovee, and A. F. Pacheco. Dynamics of rumor spreading in complex networks. *Phys. Rev. E*, 69 :066130, 2004.
- [66] Y. Moreno, M. Nekovee, and A. Vespignani. Efficiency and reliability of epidemic data dissemination in complex networks. *Phys. Rev. E*, 69 :055101, 2004.
- [67] Y. Moreno, R. Pastor-Satorras, and A. Vespignani. Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B*, 26 :521–529, 2002.
- [68] M. Newman. The physics of networks. *Phys. Today*, 61(11) :33–38, 2008.
- [69] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89(20) :208701, 2002.
- [70] M. E. J. Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 66 :016128, 2002.
- [71] M. E. J. Newman. Properties of highly clustered networks. *Phys. Rev. E*, 68 :026121, 2003.
- [72] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2) :167–256, 2003.
- [73] M. E. J. Newman. Random graphs with clustering. *Phys. Rev. Lett.*, 103 :058701, 2009.
- [74] M. E. J. Newman. *Networks : An Introduction*. Oxford University Press, 2010.
- [75] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64 :026118, 2001.
- [76] P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau, and L. J. Dubé. Propagation on networks : an exact alternative perspective. *Phys. Rev. E*, 85 :031118, 2012.
- [77] P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau, and L. J. Dubé. Spreading dynamics on complex networks : a general stochastic approach. *arXiv :1201.0296*, 2012. Soumis pour publication.
- [78] P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé, and B. Pourbohloul. Time evolution of epidemic disease on finite and infinite networks. *arXiv :0804.1807v1*, 2008. Version préliminaire de [79].
- [79] P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé, and B. Pourbohloul. Time evolution of epidemic disease on finite and infinite networks. *Phys. Rev. E*, 79(2) :026101, 2009.

- [80] J. Park and M. E. J. Newman. Solution of the two-star model of a network. *Phys. Rev. E*, 70 :066146, 2004.
- [81] J. Park and M. E. J. Newman. Statistical mechanics of networks. *Phys. Rev. E*, 70 :066117, 2004.
- [82] J. Park and M. E. J. Newman. Solution for the properties of a clustered network. *Phys. Rev. E*, 72 :026136, 2005.
- [83] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14) :3200–3203, 2001.
- [84] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E*, 63 :066117(1–8), 2002.
- [85] R. Pastor-Satorras and A. Vespignani. Immunization of complex networks. *Phys. Rev. E*, 65 :036104(1–8), 2002.
- [86] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet*. Cambridge University Press, Cambridge, 2004.
- [87] L. Peliti. Path integral approach to birth-death processes on a lattice. *Journal de Physique*, 46 :1469–1483, 1985.
- [88] E. L. Rezende, J. E. Lavabre, P. R. Guimaraes, P. Jordano, and J. Bascompte. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, 448 :925–928, 2007.
- [89] R. A. Sack. Taylor’s theorem for shift operators. *Philos. Mag.*, 3(29) :497–503, 1958.
- [90] G. M. Schütz, M. Brandau, and S. Trimper. Exact solution of a stochastic susceptible-infectious-recovered model. *Phys. Rev. E*, 78 :061132, 2008.
- [91] M. A. Serrano and M. Boguñá. Clustering in complex networks. I. General formalism. *Phys. Rev. E*, 74 :056114, 2006.
- [92] M. A. Serrano and M. Boguñá. Percolation and epidemic thresholds in clustered networks. *Phys. Rev. Lett.*, 97 :088701, 2006.
- [93] R. Solomonoff. A formal theory of inductive inference. Part I. *Inform. Control*, 7(1) :1 – 22, 1964.
- [94] W. Specht. Die linearen beziehungen zwischen höheren kommutatoren. *Mathematische Zeitschrift*, 51 :367–376, 1948.
- [95] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor & Francis, revised second edition, 1994.
- [96] M. Suzuki. Generalized trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Comm. Math. Phys.*, 51 :183–190, 1976.

- [97] M. Suzuki. On the convergence of exponential operators—the zassenhaus formula, bch formula and systematic approximants. *Comm. Math. Phys.*, 57 :193–200, 1977.
- [98] D. Trpevski, W. K. S. Tang, and L. Kocarev. Model for rumor spreading over networks. *Phys. Rev. E*, 81(5) :056102, 2010.
- [99] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.*, 56 :293–310, 2008.
- [100] D. J. Watts. *Small Worlds*. Princeton University Press, Princeton, NJ, 1999.
- [101] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393 :440–442, 1998.
- [102] D. B. West. *Introduction to Graph Theory*. Prentice Hall, 2nd edition, 2000.
- [103] F. Wever. Operatoren in lieschen ringen. *Journal für die reine und angewandte mathematik*, 187 :44–55, 1947.
- [104] M. Weyrauch and D. Scholz. Computing the Baker-Campbell-Hausdorff series and the Zassenhaus product. *Comput. Phys. Commun.*, 180 :1558–1565, 2009.
- [105] H. S. Wilf. *Generatingfunctionology*. A K Peters, Ltd., 3rd edition, 2006.
- [106] R. Yaesoubi and T. Cohen. Generalized markov models of infectious disease spread : A novel framework for developing dynamic health policies. *Eur. J. Oper. Res.*, 215(3) :679–687, 2011.
- [107] W. W. Zachary. An information flow model for conflict and fission in small groups. *J. Anthropol. Res.*, 33(4) :452–473, 1977.
- [108] D. Zanette. Dynamics of rumor propagation on small-world networks. *Phys. Rev. E*, 65(4) :041908, 2002.



# Index

- à-la-volée, 82, 113, 118
- agrégation, **13**, 54, 130
  - coefficient d', **14**
- aliasing*, voir crénelage
- analytique, fonction, 26, 27
- analytique, modèle, 7, 20, 26, 31, 33
- annealed*, voir mouvant
- annihilation, 161
  - opérateur d', **140**
- arête, **14**, voir aussi lien
- arbre, **6**, 29, 36
- avant, récurrence vers l', **30**, 38
  
- backward recurrence*, voir rebours, récurrence à
- Baker-Campbell-Hausdorff, relation de, **153**
- birth-death process*, voir naissance-mort, processus de
- boucle, **4**, 6, 9, 34, 80
- bra, **140**
- branchement, 23, **29**, 63, 80, 88, 120, 130, 168
  
- café, 24
- canonique, voir ensemble canonique
- Cauchy, théorème de, 27
- chemin, **4**, 14
- coévolution des processus, **16**
- communauté, **13**
- commutateur, **141**, 151
- compartiment, **17**
- compartimental, modèle, **17**, 31, 68
- composante, **6**, 23, 29
  - entrante, **6**
  - fortement connectée, **6**
  - géante, **23**, 37, 75
    - taille de la, **24**, 30
  - petite, **24**, 76
  - sortante, **6**
  - taille d'une, **6**, 23
- configuration, modèle de, **8**, 13, 18, 23, 28, 30, 34, 81, 109
- continu, temps, 54, 68, 70, 75
- contrainte, 8, 19
  - faible, **8**, 133
  - forte, **8**, 133
  - qualitative, **9**
- convolution, 27, 69
- création, opérateur de, **140**
- crénelage, 28
- cycle, **4**, 6, 14, 29, 30, 36
  
- degré, **4**, 6, 8, 9, 28, 34
  - corrélation des, **13**
  - distribution en, **8**, 13, 28, 34, 41
  - entrant, **3**
  - excédentaire, **28**, 35
  - non dirigé, **3**
  - séquence de, **8**, 34
  - sortant, **3**
- dégradation, 163
- degree*, voir degré
- demi-lien, 3, **4**, 9, 35, 82
  - entrant, **3**
  - non dirigé, **3**
  - sortant, **3**
- désintégration, 159
- déterministe, **19**

- différentielle, équation, 17, 32, 42, 51, 75  
 discret, temps, **29**, 30, 36, 68, 70  
 distance (entre deux nœuds), **4**  
   moyenne, **14**  
 dynamique, **15**, 28, 30, 31  
   de propagation, voir propagation  
   du réseau, **15**  
   sur le réseau, **15**  
   sur un ensemble de réseaux, **9**
- edge*, voir arête et lien  
 élément, voir nœud  
 ensemble, **7**, 11, 19, 23, 34, 76  
   canonique, **8**, 13, 18, 23, 28, 30  
   grand-canonique, **8**  
   microcanonique, **8**, 18, 81  
 entropie, 34, 120, 133  
 épidémie, **18**, 24, 30, 32, 37, 75, 88  
 épidémiologique, modèle, **17**, 33  
 Erdős et Rényi, modèle d', **8**, 10, 18, 24, 108  
 état, 7, 16, 19, 30, 37, 70, 73, 76  
   complet, **16**, 99  
   épidémiologique, **17**, 32  
   final, 30, 33, 36  
   intrinsèque, **15**, 100  
   vecteur d', 82  
 évènement, **83**, 102, 132  
 exposé, 33, **68**
- FFT*, voir Fourier, transformée de  
 flambée épidémique, **18**, 32, 76, 88  
   taille de la, voir infection, taille de l'  
 fonction de partition, 133  
 fonction génératrice, **26**, 33, 132  
   composition, 29, 89  
   de probabilité, **26**, 28, 35, 89  
   évaluation numérique, **27**  
   ordinaire, **26**  
   produit, **27**, 37  
   puissance, **27**
- forward recurrence*, voir avant, récurrence vers l'  
 Fourier, transformée de, 27, 91  
 gaussienne, distribution, 64, 69, 75, 77, 86, 103, 115  
 gelé, **9**, 18, 23, 28, 30, 75, 80, 117  
 génération (d'infection), **29**, 36, 68, 76, 89  
 générer, voir fonction génératrice  
 Gillespie, algorithme de type, 85  
 grand-canonique, voir ensemble grand-canonique  
 graphe, **14**, voir aussi réseau  
 grille, voir réseau régulier
- Hilbert, espace d', **140**  
*hub*, **13**
- idéalisation (de la réalité), **19**, 25  
 infectieux, **17**, 24, 28, 33  
 infection, **17**, 24, 28, 31, 80, 103  
   taille de l', **29**, 30, 33, 36, 38  
   taux d', **18**, 25  
 information partielle, 7, 19, 20, 34  
 intégrale fonctionnelle, 154  
 interaction, voir lien
- karaté, voir Zachary, club de karaté de ket, **140**  
 Kolmogorov, complexité de, **10**
- lattice*, voir réseau régulier  
 Lie, algèbre de, **151**  
   libre, **152**  
   produit, voir commutateur  
 lien, **2**, 4, 24, 34  
   dirigé, **3**, 7, 9  
   entrant, **3**  
   non dirigé, **3**, 6, 7  
   répété, **4**, 7, 34, 80  
   sortant, **3**, 6



- limite, théorème central, 69  
*link*, voir lien  
 liste d'adjacence, 7  
 loi de puissance, 13, 32, 38  
 longueur (d'un chemin), 4  
*loop(self-)*, voir boucle
- Magnus, méthode de, 154  
 maîtresse, équation, 83, 102  
 Markov, processus de, 73, 75, 76, 80, 83, 102  
*master equation*, voir maîtresse, équation  
 matrice d'adjacence, 6  
 mélangé, bien, 17, 32  
 microcanonique, voir ensemble microcanonique  
 Monte Carlo, 34  
 motif, 76, 101, 130  
 mouvant, 10, 18, 80
- naissance-mort, processus de, 75, 80, 102  
*neighbour*, voir voisin  
*network*, voir réseau  
*node*, voir nœud  
 nœud, 2, 24, 34
  - cible, 3, 6
  - d'arrivée, 4
  - de départ, 4
  - de référence, 6
  - source, 3, 6
- normalisation, opérateur de, 141  
 numérique, modèle, 7, 15, 20, 31, 34  
 Nyquist-Shannon, théorème d'échantillonnage de, 28
- on-the-fly*, voir à-la-volée  
 Ornstein-Uhlenbeck, processus de, 75  
*outbreak*, voir flambée
- path*, voir chemin  
*path integral*, voir intégrale fonctionnelle  
 patient zéro, 29, 36
- Pearson, coefficient de corrélation de, 13  
 percolation, 23, 28, 31, 33
  - de nœuds, 24
  - de liens, 24, 25
  - de sites, voir percolation de nœuds
  - seuil de, 24, 49
- période infectieuse, 25, 63
  - corrélation, 25, 63
- période latente, 68
- pgf*, voir fonction génératrice de probabilités
- physique statistique, 2, 8, 133
- power law*, voir loi de puissance
- processus, voir dynamique
- produit cartésien, 16
- propagation, 17, 18, 20, 23, 25, 30, 31, 63, 80, 81, 103
- $\Psi_0^g(x, y)$ , 30, 37, 69  
 $\tilde{\Psi}_0^g(x, y)$ , 46
- quenched*, voir gelé
- réalité, 19, 74
- rebours, récurrence à, 30
- représentation, 6, 10
  - partielle, 70, 74, 76, 100, 116
- reproduction, nombre de, 37, 53
- réseau, 2, 4, 18, 19, 23, 31, 34
  - aléatoire, 8–10, 12, 14, 17
  - biparti, 130
  - communautaire, 13
  - complexe, 1, 10, 12, 24, 33, 98
    - caractéristiques d'un, 12–14
  - dirigé, 3, 15
  - fini, 2, 28, 30, 31, 33, 40, 46, 70, 75, 80, 91, 115
  - hiérarchique, 14
  - infini, 29, 31, 33, 34, 70, 88
  - non dirigé, 3, 6, 8, 15, 23, 28, 30
  - régulier, 1, 8, 10, 12, 24
  - semi-dirigé, 3

- taille d'un, **2**, 23, 34, 36  
 retiré, **17**, 24, 29, 33  
 retrait, taux de, **18**, 25  
  
 Sack, identité de, **153**, 160  
     généralisée, **153**, 157, 164  
*scale free*, voir loi de puissance  
*SEIR*, **68**  
*SE<sup>m</sup>I<sup>n</sup>R*, **68**  
*SE<sup>m</sup>IR*, **68**  
 série de puissances, voir fonction généra-  
     trice ordinaire  
 Shannon, entropie de, voir entropie  
*SI*, **17**, 75, 81  
*SIIIR*, **18**  
*SIIR*, **18**, 65  
 simplification (de la réalité), **19**, 25  
 simulation, voir numérique, modèle  
*SI<sup>n</sup>R*, **18**  
*SIR*, **17**, 23, 24, 33, 63, 75, 104  
*SIRS*, **17**  
*SIS*, **17**, 103  
*small world*, **14**, 32  
 sommet, **14**, voir aussi nœud  
 stochastique, **19**, 30, 31, 34, 51, 73, 98  
     équation différentielle, 87  
 structure, voir réseau  
*stub*, voir demi-lien  
 susceptible, **17**, 24, 28, 33  
 Suzuki-Trotter, relation de, **154**  
  
 taille d'un réseau, voir réseau, taille d'un  
 taille d'une composante, voir composante,  
     taille d'une  
 taille d'une composante géante, voir com-  
     posante géante, taille d'une  
 taille de l'infection, voir infection, taille  
     de l'  
 temps, voir dynamique  
 transition, **17**  
 transitivité, voir agrégation  
  
 transmissibilité, **24**, 28, 35, 64, 68  
*tree(-like)*, voir arbre  
 triangle, **4**, 14, 101  
 Trotter, relation de, **154**  
  
 van Kampen, approximation de, 75  
 variété des issues, voir stochastique  
 vecteur d'état, **101**  
*vertex*, voir sommet et nœud  
 voisin, **4**, 13, 25, 28, 34  
     entrant, **3**  
     non dirigé, **3**, 4  
     sortant, **3**, 4  
  
 Watts et Strogatz, modèle de, **8**, 14  
 Wiener, processus de, 87  
  
 Zachary, club de karaté de, **12**  
 Zassenhaus, relation de, **152**