

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES À FINALITÉ DIDACTIQUE

Construction d'hypergraphes aléatoires et leur dynamique

MARCIPONT, Julie

Award date:
2023

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Construction d'hypergraphes aléatoires et leur dynamique

Julie Marcipont

Août 2023



Mémoire présenté pour l'obtention du grade académique
de master en « Sciences mathématiques à finalité didactique »

Promoteur : Timotéo Carletti

Résumé

La théorie des hypergraphes permet de généraliser la théorie des graphes en introduisant la notion d'hyperarêtes. L'intérêt de cette théorie est de pouvoir considérer des interactions à plus de deux corps, facilitant l'analyse de systèmes associés à des interactions sociales, des réseaux neuronaux ou d'autres interactions corporelles. Bien qu'il existe de nombreuses manières de construire des hypergraphes, dans la première partie de ce mémoire, nous allons nous intéresser à une classe de modèles d'hypergraphes aléatoires basée sur la probabilité de connexion entre les noeuds et les hyperarêtes. Nous nous pencherons ensuite sur l'analyse de différentes mesures intéressantes telles que le degré des noeuds ou la taille des hyperarêtes. Dans la seconde partie, nous étudierons la dynamique de systèmes associés à des hypergraphes arbitraires. Nous pourrons ensuite appliquer cette étude à un processus de diffusion sur les systèmes associés aux hypergraphes des classes de modèles présentés dans la première partie.

Mots-clés : Hypergraphe, Noeud, Hyperarête, Processus de diffusion, Stabilité linéaire.

Abstract

Hypergraph theory generalises graph theory by introducing the notion of hyperedges. The interest of this theory is to be able to consider interactions with more than two bodies, facilitating the analysis of systems associated with social interactions, neural networks or other bodily interactions. Although there are many ways of constructing hypergraphs, in the first part of this thesis we will focus on a class of random hypergraph models based on the probability of connection between nodes and hyperedges. We will then look at the analysis of various interesting measures such as the degree of the nodes or the size of the hyperedges. In the second part, we will study the dynamics of systems associated with arbitrary hypergraphs. We can then apply this study to a diffusion process on the systems associated with the hypergraphs of the model classes presented in the first part.

Keywords : Hypergraph, Node, Hyperedge, Diffusion process, Linear stability.

Remerciements

Nous adressons nos remerciements les plus sincères à toutes les personnes qui nous ont aidés à réaliser ce travail et plus particulièrement :

A Timotéo Carletti qui a accepté de superviser ce travail : pour ses conseils judicieux, sa rigueur, sa grande disponibilité durant la réalisation du travail ;

A Mathieu Bastin, maître-assistant en mathématiques à la HELHa, pour les séances de rencontres qu'il m'a consacrées, permettant de structurer ma démarche de travail.

Nos plus grandes gratitudes à :

Madame Annick Castiaux, rectrice à l'Université de Namur, ainsi qu'à tous les professeurs de l'université ; nous adressons notre reconnaissance pour la richesse de leur enseignement.

A Monsieur Ledoux, professeur émérite en sciences actuarielles, de l'Université catholique de Louvain pour ses conseils judicieux ainsi que son soutien stimulant.

A Laurence pour l'aide apportée dans la rédaction de ce travail.

Merci à ma famille et mes amis pour leur soutien et encouragement tout au long de mes études et de ce travail de fin d'études.

Table des matières

Introduction	1
1 Les graphes	2
1.1 Notions théoriques	2
1.2 Modèle d'Erdős-Renyi	6
1.3 Attachement Préférentiel	7
2 Notions théoriques	10
2.1 Qu'est-ce qu'un hypergraphe?	10
2.2 Comment construire un hypergraphe?	13
2.2.1 Une classe de modèles d'hypergraphes	13
2.2.2 Hypergraphes aléatoires simples	16
2.2.3 Attachement préférentiel pour les hypergraphes	22
2.2.4 Hypergraphes spatiaux aléatoires	24
3 Analyse de système dynamique d'hypergraphes	34
3.1 Analyse théorique de la stabilité linéaire	36
3.2 Application de l'étude de stabilité sur les modèles construits précédemment.	39
3.2.1 Modèle 1 : hypergraphes simples	40
3.2.2 Modèle 2 : hypergraphes spatiaux - cas exponentiel	44
3.2.3 Modèle 3 : hypergraphes spatiaux - cas géométrique	48
Conclusion et perspectives	50

Introduction

Les graphes et les hypergraphes sont des structures fondamentales utilisées pour modéliser et analyser des systèmes complexes dans de nombreux domaines. Alors que les graphes se concentrent sur des relations binaires entre les noeuds, les hypergraphes permettent de considérer des relations entre des ensembles de noeuds, ce qui les rendent particulièrement adaptés pour représenter des interactions complexes et multidimensionnelles.

Ce mémoire se penche essentiellement sur une classe de modèles d'hypergraphes basée sur la probabilité de connexion entre un noeud et une hyperarête, dont nous étudierons différentes mesures intéressantes. Ce mémoire s'intéresse également à la dynamique des systèmes associées à ces différents hypergraphes.

Dans le premier chapitre, nous allons présenter la théorie des graphes en mettant en avant les concepts clés qui nous permettront d'exposer et de comprendre la théorie des hypergraphes.

Dans le deuxième chapitre, nous allons nous baser sur la théorie des graphes afin d'étendre certaines notions aux hypergraphes, qui nous seront utiles pour la compréhension de la suite du travail. Une fois ce cadre théorique effectué, nous nous intéresserons à la classe de modèles d'hypergraphes exposée dans l'article « *A class of models for random hypergraphs* » de M. Barthelemy [3]. Nous expliciterons la manière dont ces différents hypergraphes sont construits pour ensuite pouvoir étudier différentes mesures intéressantes telles que le degré des noeuds ou la taille des hyperarêtes.

Dans le dernier chapitre, nous analyserons de manière théorique la stabilité des systèmes dynamiques associés à des hypergraphes arbitraires en nous basant sur l'article « *Phase transitions and stability of dynamical processes on hypergraphs* » de G. Ferraz de Arruda, M. Tizzani et Y. Moreno [10]. Nous appliquerons ensuite cette étude de stabilité dynamique à un processus de diffusion sur les systèmes associés aux différents hypergraphes construits dans le deuxième chapitre. Nous étudierons ainsi le rôle joué par les différents paramètres des modèles considérés sur la convergence de la solution des systèmes associés.

Chapitre 1

Les graphes

Dans ce chapitre, nous exposons les définitions et concepts de base concernant les graphes, afin de pouvoir définir le modèle d'Erdős-Renyi. Ce modèle nous sera utile dans la suite du travail puisqu'il servira de base et d'exemple pour la construction d'hypergraphes. Les définitions et concepts associés aux graphes serviront de base pour définir les hypergraphes. En effet, nous verrons par la suite que de nombreux liens peuvent être faits entre ces deux notions et que les hypergraphes sont en fait une extension de la notion de graphe. Nous verrons ensuite le modèle d'attachement préférentiel qui nous servira dans le chapitre suivant afin de construire des hypergraphes.

1.1 Notions théoriques

Les graphes sont, de manière générale, des structures permettant de représenter des interactions par paires entre différents objets. Plus précisément, un graphe est constitué d'un ensemble de noeuds, que l'on note V , ainsi que d'un ensemble d'arêtes reliant ces noeuds, que l'on note E . Chaque arête est ainsi associée à un noeud ou une paire de noeuds par une fonction dite d'incidence, que l'on note ψ . Un graphe est donc un triplet (V, E, ψ) , dont la structure est caractérisée par ces trois éléments : l'ensemble des noeuds, l'ensemble des arêtes et les relations entre ces deux ensembles.

Un exemple de graphe est illustré à la figure 1.1. Celui-ci est caractérisé par un ensemble de noeuds

$$V_G = \{u, v, w, x, y\},$$

un ensemble d'arêtes

$$E_G = \{a, b, c, d, e, f, g, h\},$$

ainsi que par une fonction d'incidence ψ_G qui est telle que, entre autre,

$$\begin{aligned}\psi_G(a) &= uv, \\ \psi_G(b) &= uu, \\ \psi_G(c) &= vw, \\ &\dots \\ \psi_G(g) &= ux.\end{aligned}$$

On notera donc $G = (V_G, E_G, \psi_G)$ ce graphe illustré.

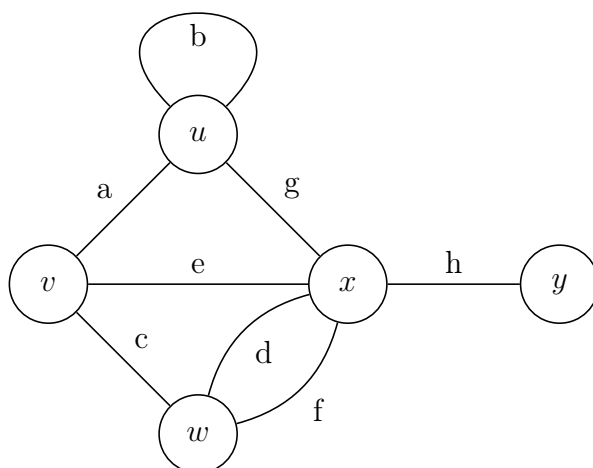


FIGURE 1.1 – Schéma d'un graphe tiré de [17]

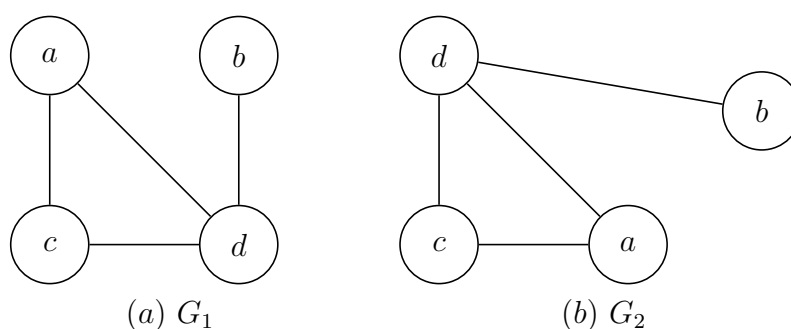


FIGURE 1.2 – Schéma de deux graphes isomorphes tiré de [13]

Nous pouvons à présent nous intéresser à quelques notions et propriétés importantes des graphes. Notamment, on dit que deux nœuds sont adjacents s'ils sont incidents à la même arête. Dans l'exemple de la figure 1.1, les nœuds u et v sont adjacents, car ils sont reliés par la même arête a ; on peut également prendre les nœuds x et y , incidents à l'arête h ; et ainsi de suite pour les autres nœuds de ce graphe. Dans le cas particulier où une arête est incidente à un seul nœud, nous dirons que celle-ci est une boucle, comme illustré par l'arête b du graphe de la figure 1.1 qui est uniquement incidente au nœud u . Si nous sommes dans le cas où le graphe ne présente ni de boucle, ni d'arête multiple, alors nous dirons que ce graphe est simple. Le graphe de la figure 1.1 qui présente une boucle incidente au nœud u et des arêtes multiples entre les nœuds w et x , n'est pas un graphe simple. Par contre, les graphes représentés à la figure 1.2 ne présentent ni boucle ni arête multiple et sont donc des graphes simples.

Nous pouvons également énoncer la notion de degré d'un nœud, qui représente le nombre d'arêtes qui lui sont incidentes. Pour reprendre l'exemple de la figure 1.1, le nœud v est de degré 3, le nœud x est de degré 5, et ainsi de suite pour les autres nœuds. Pour le nœud u , nous sommes dans le cas particulier d'une boucle. Celle-ci comptant pour deux, le degré du nœud u est de 4. Dans la suite de ce chapitre, nous utiliserons comme notations $\gamma(G)$ le nombre de nœuds du graphe G et $\epsilon(G)$ le nombre de ces arêtes.

La représentation du graphe sous forme de diagramme est très utile pour comprendre et analyser la structure du système. Toutefois, il est important de noter que celle-ci peut s'avérer trompeuse. En effet, comme illustré à la figure 1.2, un même graphe peut être représenté de diverses façons tout en gardant la même structure et donc les mêmes

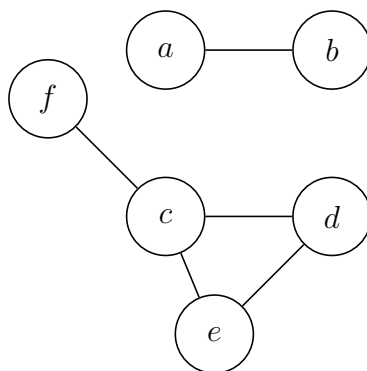


FIGURE 1.3 – Schéma d'un graphe non connexe tiré de [13]

propriétés. Dans ce cas, ces graphes sont dits isomorphes. Formellement, deux graphes $G_1 = (V_1, E_1, \psi_1)$ et $G_2 = (V_2, E_2, \psi_2)$ sont isomorphes s'il existe des bijections $f : V_1 \rightarrow V_2$ et $g : E_1 \rightarrow E_2$ qui sont telles que $\psi_1(e) = uv$ si et seulement si $\psi_2(g(e)) = f(u)f(v)$.

De plus, deux graphes isomorphes possèdent la même matrice d'adjacence. Cette matrice, de dimension $\gamma \times \gamma$, permet de recenser le nombre d'arêtes reliant deux noeuds. Plus précisément, l'élément ij de cette matrice est le nombre d'arêtes reliant les noeuds v_i et v_j . Dans le cas d'un graphe simple, cette matrice sera donc composée uniquement de 0 et de 1. Pour illustrer ce concept, nous pouvons prendre le graphe illustré à la figure 1.3 dont la matrice d'adjacence est la suivante :

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

Comme c'est un graphe simple, sa matrice d'adjacence est bien composée uniquement de 0 et de 1, avec les éléments diagonaux nuls. Par contre, si nous prenons le graphe illustré à la figure 1.4, celui-ci comporte une boucle et des arêtes multiples. Sa matrice d'adjacence, sera alors composée d'entiers :

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 2 & 0 \end{pmatrix}.$$

Cette matrice ne permet cependant pas de juger de la connexité d'un graphe. En effet, un graphe est dit "connexe" s'il existe un parcours qui relie chaque paire de points. Un parcours est une suite

$$v_0 e_1 v_1 e_2 \dots e_n v_n,$$

où v_0, v_1, \dots sont des noeuds et e_1, e_2, \dots sont des arêtes reliant les noeuds consécutifs du parcours. Dans le cas où le noeud d'origine v_0 et le noeud d'arrivée v_n sont égaux, le parcours est dit fermé. Si tous les noeuds sont distincts, on parlera alors de chemin. Un exemple de graphe non-connexe est illustré à la figure 1.3.

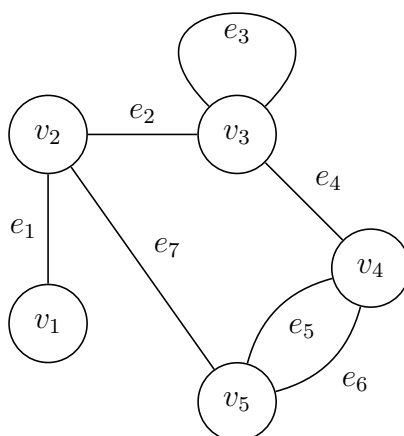


FIGURE 1.4 – Schéma d'un graphe tiré de [17]

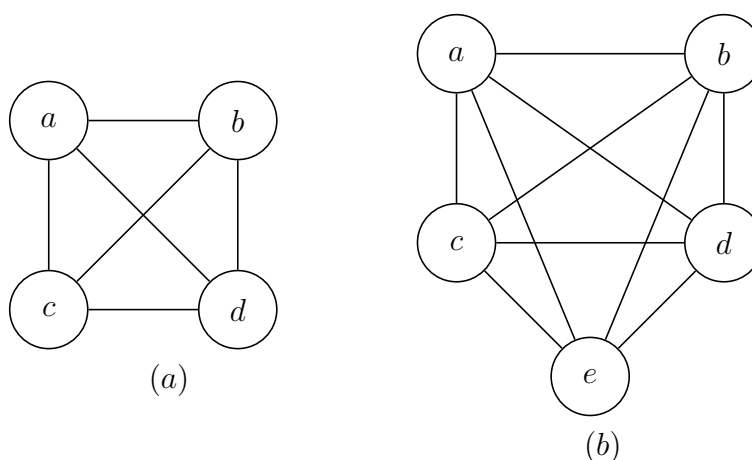


FIGURE 1.5 – Schéma d'un graphe planaire en (a) et schéma d'un graphe non planaire en (b), tirés de [13]

Une autre matrice permettant de caractériser un graphe est la matrice d'incidence. Celle-ci, de dimension $\gamma \times \epsilon$, permet de recenser les liens entre les arêtes et les noeuds. Plus précisément, l'élément ij de cette matrice est le nombre de fois que le noeud v_i est incident à l'arête e_j , c'est-à-dire 0, 1 ou 2 dans le cas d'une boucle. Pour illustrer ce concept, nous pouvons prendre le graphe illustré à la figure 1.4 dont la matrice d'incidence est la suivante :

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Grâce à la représentation d'un graphe sous la forme d'un diagramme, nous pouvons introduire la notion de graphe planaire. Celle-ci désigne un graphe pouvant être représenté sans que des arêtes ne s'intersectent. La figure 1.5 permet d'illustrer ce concept en montrant la différence entre un graphe planaire et un graphe non-planaire.

Une autre notion intéressante qui sera utile pour l'étude de la stabilité est celle de la matrice Laplacienne combinatoire d'un graphe [5]. C'est une matrice carrée qui décrit

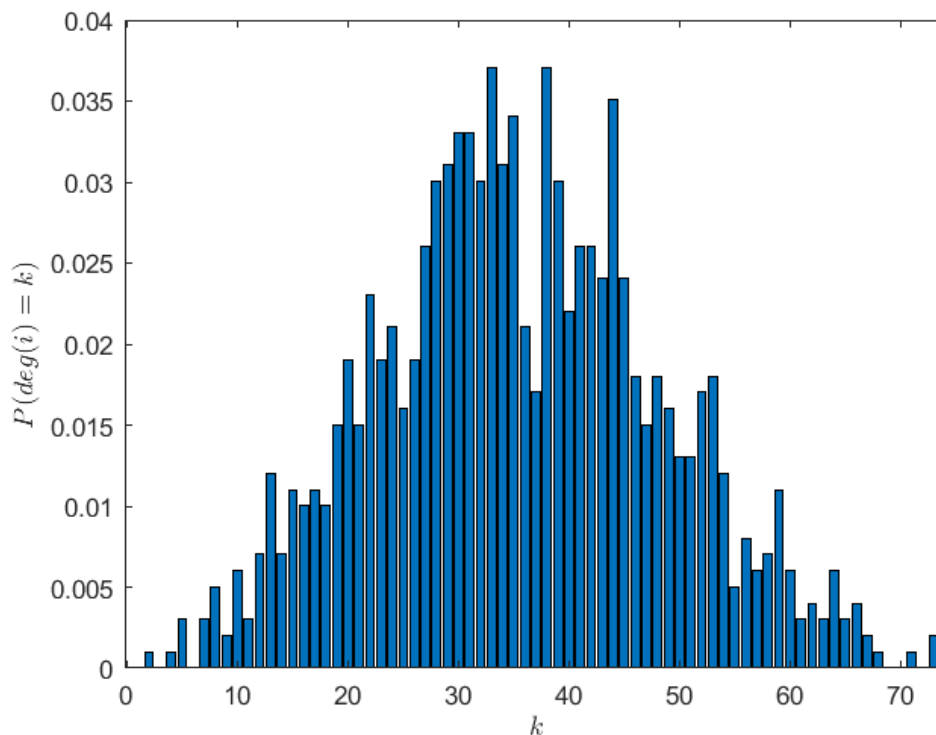


FIGURE 1.6 – Probabilité qu’un noeud soit de degré k pour un grappe d’Erdős-Rényi.

les relations entre les noeuds d’un graphe. Elle est définie comme la différence entre la matrice diagonale des degrés D et la matrice d’adjacence A du graphe. Formellement, elle est définie par

$$\mathcal{L} = D - A,$$

où la matrice A est la matrice d’adjacence définie précédemment, et la matrice D est une matrice diagonale dont les éléments diagonaux, (d_1, \dots, d_n) , correspondent aux degrés de chaque noeud (d_i est le degré du noeud i). La différence entre ces deux matrices est donc une matrice qui représente les relations entre les noeuds en termes de degrés de connexion. La matrice Laplacienne est une notion importante en théorie des graphes et en algèbre linéaire, car elle est utilisée pour étudier les propriétés spectrales des graphes. En particulier, les valeurs propres de la matrice Laplacienne sont souvent utilisées pour caractériser les propriétés structurelles des graphes, telles que leur connectivité.[18]

1.2 Modèle d’Erdős-Rényi

Le modèle d’Erdős-Rényi [9] est un modèle permettant de générer des graphes aléatoirement en fixant le nombre de noeuds et en considérant une certaine probabilité de connexion entre une paire de noeuds.

Soient n un entier strictement positif et p un paramètre tel que $p \in [0, 1]$. Soit $(X_{i,j})_{1 \leq i < j \leq n}$ une famille de variables aléatoires de Bernoulli indépendantes de paramètre p , définies par

$$X_{i,j} = \begin{cases} 1 & \text{si } \{i, j\} \text{ est une arête du graphe } G \\ 0 & \text{sinon} \end{cases}$$

et telles que $P[X_{ij} = 1] = 1 - P[X_{ij} = 0] = p$.

On appelle alors graphe aléatoire d'Erdős-Rényi, et on note $G(n, p)$, le graphe dont les noeuds sont les entiers $i \in [1, n]$ et l'ensemble des arêtes est $\{(i, j) | i < j, X_{i,j} = 1\}$.

Dès lors, chaque arête possible entre les noeuds du graphe G existe avec une certaine probabilité p , indépendamment des autres arêtes. [20] [23]

Pour de tels graphes, on peut déterminer la probabilité qu'un noeuds i soit de degré k . Soit $\binom{n}{k}$ le nombre de façons possibles de choisir k noeuds parmi n , et soit p^k la probabilité qu'ils aient des arêtes. Alors $\binom{n}{k}p^k$ est la probabilité qu'un noeud donné possède des arêtes le reliant à k autres noeuds et $(1 - p)^{n-k}$ est la probabilité que ce noeud ne soit par relié aux $n - k$ noeuds restants. Par conséquent, la probabilité qu'un noeud i soit de degré k est définie par [25]

$$P[\text{deg}(i) = k] = \binom{n}{k} p^k (1 - p)^{n-k}.$$

La figure 1.6 permet de donner une idée de la forme de cette distribution. Cette figure a été obtenue en construisant un graphe d'Erdős-Rényi contenant 1000 noeuds et dont la probabilité qu'un noeud i soit connecté à un noeud j est de $p = 0.2$.

1.3 Attachement Préférentiel

L'attachement préférentiel est un phénomène communément observé dans les réseaux et graphes qui décrit la tendance des noeuds ou des sommets à se lier davantage aux autres noeuds qui ont déjà un grand nombre de connexions. En d'autres termes, les noeuds ont tendance à se lier aux noeuds les plus populaires, qui sont déjà fortement connectés au réseau. Afin d'approfondir cette notion, nous allons nous intéresser au modèle de Barabasi-Albert [2].

Contrairement au modèle d'Erdős-Rényi où le nombre de noeuds N est fixé à l'avance et où chaque paire de noeuds a la même probabilité p de former un lien entre eux de manière aléatoire et indépendante, le modèle de Barabasi-Albert propose une dynamique de croissance préférentielle. Dans leur modèle, le réseau commence avec un petit nombre de noeuds et, à chaque étape, un nouveau noeud est ajouté au réseau. Celui-ci est ensuite connecté à m noeuds déjà présents dans le réseau, en fonction de leur degré de connectivité. Plus précisément, la probabilité qu'un nouveau noeud se connecte à un noeud existant est proportionnelle à son degré de connectivité.

Le modèle permet de prendre en compte deux aspects génériques qui ne sont pas pris en compte dans le modèle d'Erdős-Rényi. Premièrement, le modèle de Barabasi-Albert tient compte du fait que la plupart des réseaux du monde réel sont ouverts, c'est-à-dire qu'ils continuent de croître et de s'étendre au fil du temps. En effet, ce modèle permet l'ajout continu de nouveaux noeuds au réseau, au fur et à mesure de son évolution. Cela reflète la réalité de nombreux réseaux réels tels que les réseaux sociaux, les réseaux de collaboration scientifique ou les réseaux de citations.

Deuxièmement, le modèle de Barabasi-Albert reconnaît que la connectivité dans les réseaux réels présente une préférence. Selon ce mécanisme, les nouveaux noeuds ont tendance à se connecter aux noeuds déjà fortement connectés, ce qui renforce leur degré de connectivité. Cela conduit à une distribution des degrés des noeuds qui suit une loi de puissance.

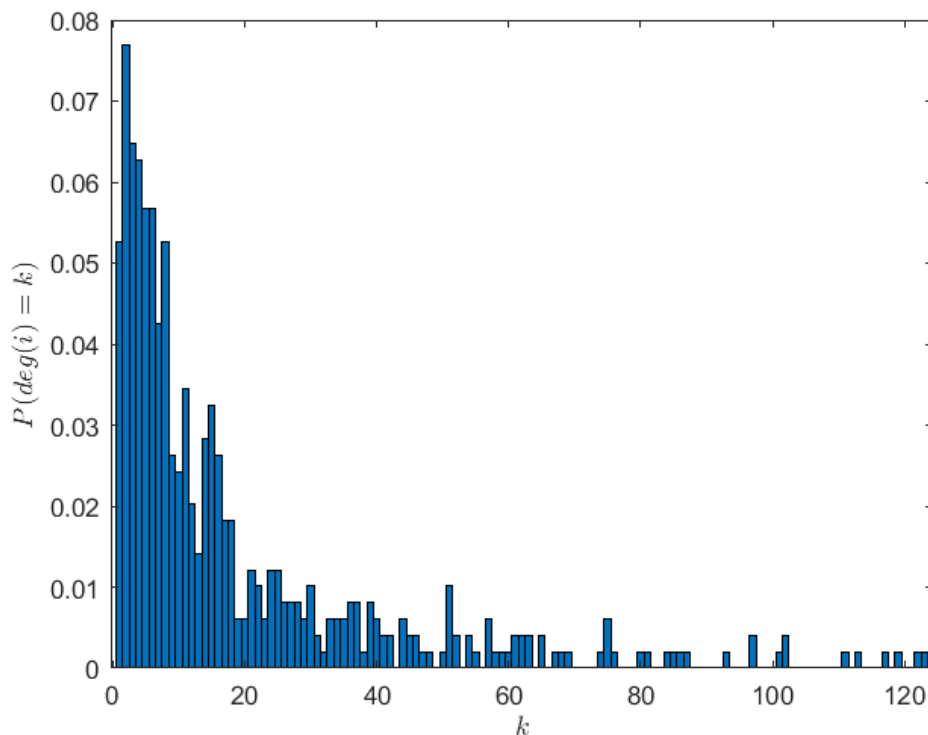


FIGURE 1.7 – Probabilité qu’un noeud soit de degré k pour un graphe construit via le processus d’attachement préférentiel.

L’article [2] décrit comment le modèle de Barabasi-Albert est défini. Premièrement, le modèle commence avec un petit nombre de noeuds M , généralement deux, qui sont connectés par une arête. Ensuite, à chaque étape de croissance, un nouveau noeud est ajouté au réseau. Ce nouveau noeud est connecté à m noeuds déjà présents dans le réseau où $m < M$. La sélection des noeuds existants se fait de manière aléatoire, mais avec une probabilité proportionnelle à leur degré. Plus précisément, pour intégrer l’attachement préférentiel dans le modèle de Barabasi-Albert, on suppose que la probabilité π qu’un nouveau noeud soit connecté au noeud i dépend de la connectivité k_i de ce noeud, de sorte que $\pi(k_i) = \frac{k_i}{\sum_j k_j}$, où k_i est le degré du noeud i et $\sum_j k_j$ est la somme des degrés de tous les noeuds du réseau. Finalement, les étapes de croissance sont répétées jusqu’à ce que le réseau atteigne la taille souhaitée ou le nombre de noeuds souhaité.

En utilisant cette probabilité d’attachement préférentiel, le modèle de Barabasi-Albert génère des réseaux qui présentent une distribution de degré suivant une loi de puissance, ce qui est une caractéristique courante dans de nombreux réseaux réels. La figure 1.7 permettant d’illustrer cette distribution est obtenue en appliquant le processus d’attachement préférentiel afin d’obtenir un graphe de 500 noeuds à partir d’un graphe contenant initialement 5 noeuds. Afin de mettre en évidence le fait que la distribution des degrés des noeuds suit une loi de puissance, nous pouvons afficher ces données en échelle logarithmique. La figure 1.8 représente donc la distribution des degrés des noeuds avec les mêmes paramètres que pour la figure 1.7, à savoir que le graphe contient initialement 5 noeuds et le graphe final en contient 500. On peut remarquer la présence de bruit numérique lors de la représentation des données. Cela peut être dû à des erreurs de calculs ou à des imprécisions lors du calcul de la probabilité de connexion d’un noeud.

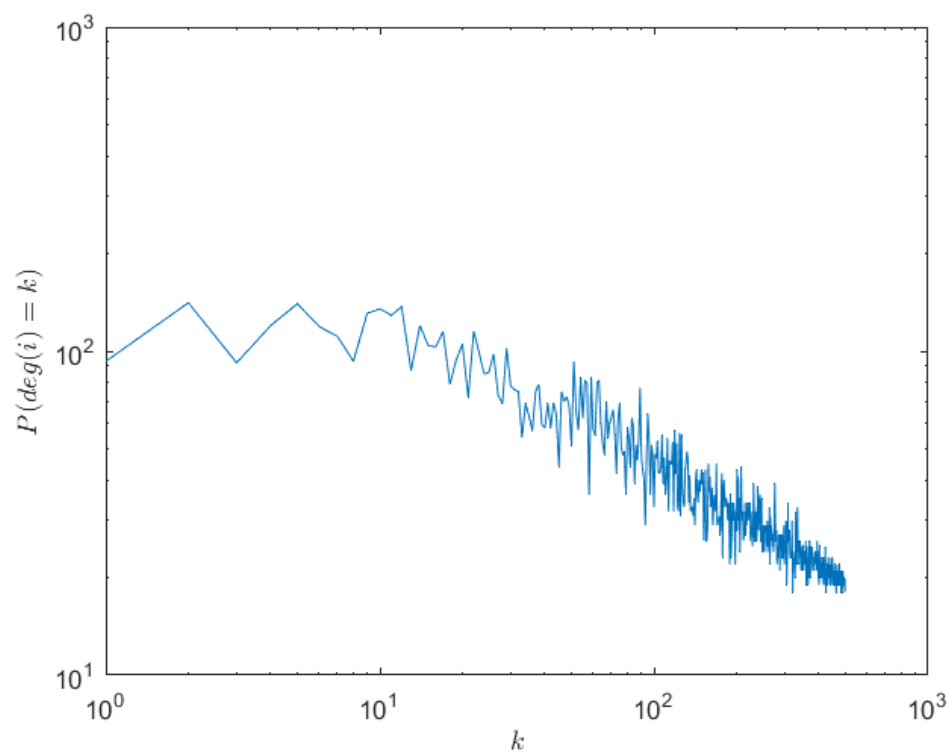


FIGURE 1.8 – Probabilité qu'un noeud soit de degré k pour un graphe construit via le processus d'attachement préférentiel en échelle logarithmique.

Chapitre 2

Notions théoriques

Dans ce chapitre, nous allons définir ce qu'est un hypergraphe en faisant le lien avec les définitions et concepts associés aux graphes qui ont été présentés dans la section 1.1. Ensuite, nous nous pencherons sur la construction d'hypergraphes selon différents modèles présentés dans l'article [3] afin d'étudier les différentes mesures intéressantes de ces modèles.

2.1 Qu'est-ce qu'un hypergraphe ?

Les hypergraphes sont une généralisation des graphes qui permettent de considérer des interactions multi-corps et ainsi de traiter des problèmes plus "complexes" tels que les interactions sociales.

Les hypergraphes permettent donc de généraliser la notion de graphe en introduisant la notion d'hyperarête, étendant ainsi la notion d'arête de graphe en n'imposant pas de restriction sur le nombre de noeuds. Une hyperarête étant ainsi définie comme pouvant contenir plusieurs noeuds.

Plus précisément, un hypergraphe, noté $H = (V, E)$, est caractérisé par les deux ensembles suivants : [5][8]

- $V = \{v_1, v_2, \dots, v_n\}$ un ensemble fini dont les éléments sont appelés les noeuds,
- $E = \{e_1, e_2, \dots, e_m\}$ une famille des parties de V dont les éléments sont appelés les hyperarêtes.

Un exemple d'hypergraphe est illustré à la figure 2.1, où nous avons six noeuds et quatre hyperarêtes, agencés tels que

$$\begin{aligned}e_1 &= \{v_1, v_2, v_3\} \\e_2 &= \{v_2, v_3\} \\e_3 &= \{v_3, v_5, v_6\} \\e_4 &= \{v_4\}\end{aligned}$$

Nous pouvons dès lors caractériser la simplicité d'un hypergraphe en regardant ses hyperarêtes. Plus précisément, un hypergraphe est dit simple si :

$$\forall e_i, e_j \in E, e_i \subseteq e_j \Rightarrow i = j,$$

c'est-à-dire que si aucune hyperarête de H n'en contient une autre, alors l'hypergraphe est simple. Dans le cas contraire, l'hypergraphe est dit multiple. Cela permet d'appuyer le

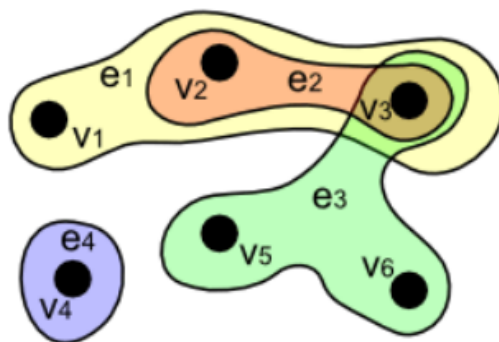


FIGURE 2.1 – Schéma d'un hypergraphe tiré de [8]

fait que les hypergraphes généralisent la notion de graphe. En effet, "un graphe simple est un hypergraphe simple, dont toutes les hyperarêtes sont de cardinalité 2 et un multigraphe est un hypergraphe, dont toutes les hyperarêtes sont de cardinalité au plus 2" [8].

Cette notion de cardinalité permet de caractériser la complétude d'un hypergraphe. En effet, un hypergraphe H est dit r -complet s'il est l'ensemble des parties de V qui sont de cardinalité r .

Nous pouvons également généraliser la notion d'adjacence de noeuds et d'arêtes. Pour un hypergraphe donné, deux noeuds v_i et v_j sont dits adjacents s'il existe une hyperarête e_k les contenant tous les deux, c'est-à-dire telle que $v_i \in e_k$ et $v_j \in e_k$.

Deux hyperarêtes e_k et e_l sont dites adjacentes si leur intersection est non vide, c'est-à-dire telle que $e_k \cap e_l \neq \emptyset$.

Nous pouvons également énoncer la notion de degré d'un noeud, qui représente le nombre d'hyperarêtes auquel il appartient. De façon analogue, le degré d'une hyperarête est le nombre de noeuds qu'elle contient [16].

Le rang d'un hypergraphe H , noté $r(H)$, est le nombre maximum de noeuds qu'une hyperarête peut contenir. Formellement, le rang est défini de la manière suivante :

$$r(H) = \max_{e_i \in E} |e_i|.$$

Inversément, l'anti-rang d'un hypergraphe H , noté $ar(H)$, est le nombre minimum de noeuds qu'une hyperarête peut contenir. Formellement, il est défini de la manière suivante :

$$ar(H) = \min_{e_i \in E} |e_i|.$$

De manière triviale, nous savons que $r(H) \geq ar(H)$. Dans le cas particulier où nous avons l'égalité, l'hypergraphe H est dit uniforme. De plus, H est dit r -uniforme s'il est de rang r . Nous pouvons reprendre la figure 2.1 afin d'illustrer ces concepts. Dans cet exemple, l'hypergraphe possède un rang de 3 et un anti-rang de 1. Dès lors, l'hypergraphe n'est pas uniforme.

Tout comme les graphes, les hypergraphes peuvent d'une part être caractérisés par une matrice d'incidence, notée J_{ji} de taille $m \times n$. Celle-ci permet de caractériser l'appartenance des noeuds à des hyperarêtes de la manière suivante [21] :

$$J_{ij} = \begin{cases} 1 & \text{si } v_i \in e_j, \\ 0 & \text{sinon.} \end{cases} \quad (2.1)$$

Dans l'exemple de la figure 2.1, la matrice d'incidence associée à l'hypergraphe est la matrice suivante :

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

D'autre part, les hypergraphes peuvent également être caractérisés par une matrice d'adjacence, notée A_{ij} , de taille $m \times m$. Nous pouvons reprendre la définition donnée dans l'article [10] afin de la définir :

$$A_{ik} = \sum_{\substack{e_j \in E \\ i, k \in e_j \\ i \neq k}} \frac{1}{|e_j| - 1}.$$

L'article [10] propose une autre méthode permettant de la définir. Pour cela, on utilise la matrice d'incidence pondérée définie par :

$$B_{ij} = \begin{cases} (|e_j| - 1)^{-1/2} & \text{si } i \in e_j \\ 0 & \text{sinon} \end{cases},$$

ainsi que la matrice diagonale

$$\hat{D}_{ii} = \sum_{e_j \cap i \neq \emptyset} \frac{1}{|e_j| - 1}.$$

La matrice d'adjacence est alors définie par l'équation suivante :

$$A = BB^T - \hat{D}. \tag{2.2}$$

Cette dernière définition permet de représenter les hypergraphes sous forme de graphes pondérés comme illustré à la figure 2.2. En (a), nous avons la projection de l'hypergraphe qui se base sur la matrice d'incidence, en (b), nous avons les hypergraphes originaux et en (c) nous avons la projection pondérée qui se base sur la matrice d'adjacence. En (a) et (c), le poids des arêtes est représenté par les couleurs. Ceux-ci sont déterminés grâce à l'équation 2.2, où on a :

$$B = \begin{pmatrix} (3-1)^{-1/2} & 0 & 0 & 0 \\ (3-1)^{-1/2} & (2-1)^{-1/2} & (2-1)^{-1/2} & 0 \\ (3-1)^{-1/2} & (2-1)^{-1/2} & 0 & (3-1)^{-1/2} \\ 0 & 0 & (2-1)^{-1/2} & (3-1)^{-1/2} \end{pmatrix},$$

$$\hat{D} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 \\ 0 & 5/2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 3/2 \end{pmatrix} \text{ et } A = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 3/2 & 0 & 1 \\ 1/2 & 3/2 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1 & 1/2 & 1/2 & 0 \end{pmatrix}.$$

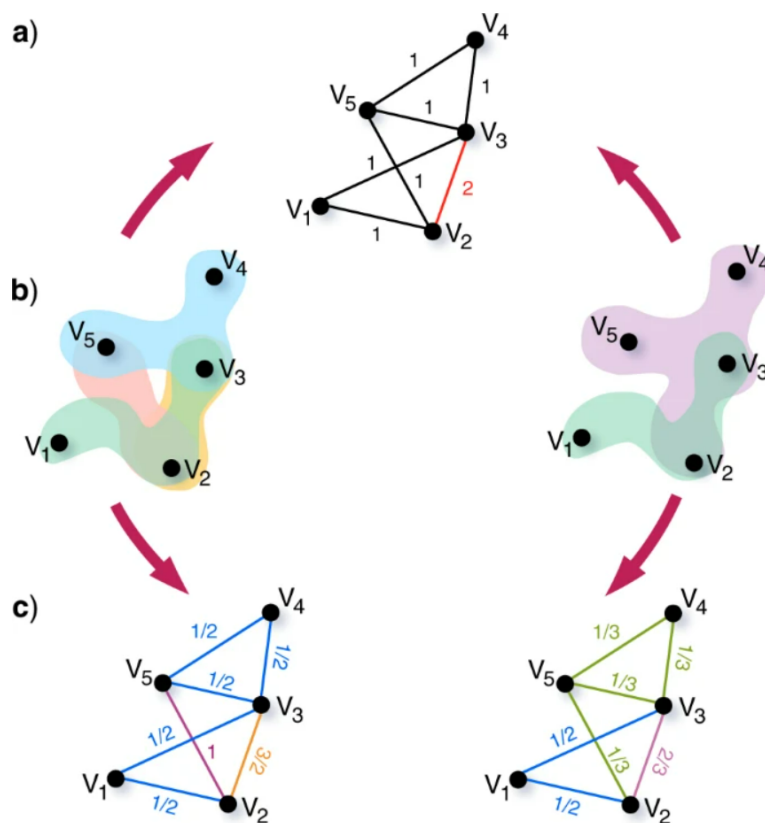


FIGURE 2.2 – Projection de deux graphes tirée de [10]

2.2 Comment construire un hypergraphe ?

Dans cette section, nous allons travailler sur la construction d’hypergraphes en nous basant sur l’article [3], dont nous reprenons la structure afin de faciliter la lecture. Dans celui-ci, Marc Barthelemy propose des modèles permettant de générer des hypergraphes aléatoires avec différentes propriétés, en se basant sur la probabilité de connexion entre un noeud et une hyperarête.

2.2.1 Une classe de modèles d’hypergraphes

Le concept central introduit pour la construction d’hypergraphes est celui de probabilité de connexion

$$P(v \in e).$$

qu’un noeud v appartienne à une hyperarête e . Ce concept est directement lié à la matrice d’incidence de l’hypergraphe puisque celle-ci vaut 1 si le noeud appartient à l’hyperarête et 0 si non (équation 2.1).

Cette distribution de probabilité peut prendre en compte plusieurs propriétés de l’hypergraphe telles que sa taille, sa position dans l’espace ou ses propriétés topologiques. Dès lors, on peut noter cette probabilité sous la forme :

$$P(v \in e) = F(e, d(v, e), \dots), \tag{2.3}$$

où F est une fonction qui dépend de l’hyperarête e , de ses noeuds, ou d’une certaine distance entre v et e .

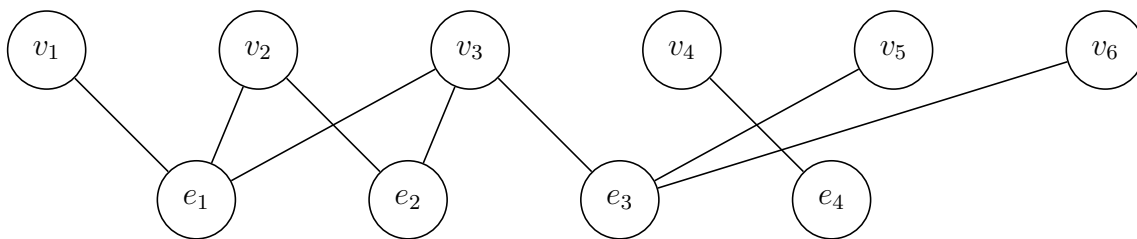


FIGURE 2.3 – Schéma de l'hypergraphe de la figure 2.1 représenté par un graphe biparti

Dans la suite, nous allons considérer le cas constant $F = p \in [0, 1]$, qui correspond à une généralisation du modèle d'Erdős-Rényi pour les graphes, où la probabilité de connexion entre deux noeuds est constante et égale à p . Dans le cas des hypergraphes, la probabilité de connexion dépend de la présence ou non du noeud dans l'hyperarête, mais reste constante.

Ensuite, nous allons considérer le cas de l'attachement préférentiel, où la fonction F dépend de la taille de l'hyperarête. Cela correspond à un mécanisme où les noeuds ont tendance à se connecter aux hyperarêtes qui sont déjà grandes.

Enfin, nous allons considérer le cas où les noeuds et les hyperarêtes sont situés dans un espace en deux dimensions, et où la fonction F dépend de la distance entre le noeud v et l'hyperarête e . Cette approche permet de prendre en compte la structure spatiale de l'hypergraphe, en modélisant les interactions entre les noeuds et les hyperarêtes en fonction de leur position relative.

Dans la suite, nous allons désigner un hypergraphe par un couple $H = (V, E)$, où V est un ensemble de noeuds, et E est un ensemble d'hyperarêtes tel que chaque hyperarête est un sous-ensemble non vide de V . L'ordre d'un hypergraphe, noté

$$N = |V|,$$

est le nombre de ses noeuds et le degré d'un hypergraphe, noté

$$M = |E|,$$

est le nombre de ses hyperarêtes. Nous pouvons reprendre l'exemple de la figure 2.1 afin de fixer ces notions. Pour faciliter la suite de l'analyse, nous pouvons représenter cet hypergraphe par un graphe biparti entre les noeuds et les hyperarêtes, comme à la figure 2.3. Cet hypergraphe compte 6 noeuds et 4 hyperarêtes ; dès lors, son ordre est de $N = |V| = 6$ et son degré est de $M = |E| = 4$.

Pour tous les modèles considérés, nous supposons que le nombre d'hyperarêtes E est donné. Nous allons alors construire les hypergraphes à partir d'un ensemble initial d'hyperarêtes, composé de noeuds choisis aléatoirement, et nous allons itérer sur les noeuds en appliquant la règle de l'équation 2.3.

Pour de tels modèles, certaines mesures sont importantes telles que le degré d'un noeud ou la taille d'une hyperarête. La taille k_i d'un noeud i est simplement le nombre d'hyperarêtes auxquelles il appartient. Nous pouvons également définir le degré d'un noeud en utilisant la matrice d'incidence de l'hypergraphe (équation 2.1) puisque celle-ci indique 1 si le noeud appartient à l'hyperarête considérée ; dans le cas contraire elle indique 0. Il suffit alors de sommer les éléments de la matrice d'incidence du noeud considéré [22][4] :

$$k_i = \sum_j J_{ij}.$$

La taille $m_l = |e_l|$ de l'hyperarête l correspond à son cardinal, c'est-à-dire au nombre de noeuds qui la composent.

Sachant que le degré k_i d'un noeud i est le nombre d'hyperarêtes auxquelles il est connecté, la somme des degrés de tous les noeuds de l'hypergraphe compte chaque apparition d'un noeud dans une hyperarête exactement une fois. Cela signifie que la somme des degrés de tous les noeuds est égale au nombre total de fois où les noeuds apparaissent dans les hyperarêtes, c'est-à-dire à la somme des tailles de toutes les hyperarêtes. Ainsi, on peut écrire :

$$L = \sum_i k_i = \sum_j |e_j|, \quad (2.4)$$

où L désigne le nombre de liens.

C'est une simple égalité entre la somme des degrés de tous les noeuds et la somme des tailles de toutes les hyperarêtes dans un hypergraphe qui montre la dépendance entre le degré moyen et la taille moyenne. Si nous reprenons l'exemple de la figure 2.3, nous constatons que

$$\begin{array}{lll} k_1 = 1 & k_3 = 3 & k_5 = 1 \\ k_2 = 2 & k_4 = 1 & k_6 = 1, \end{array}$$

où k_1 est le degré du noeud v_1 , k_2 le degré du noeud v_2 et ainsi de suite pour tous les noeuds. Cette information sur le degré des noeuds peut également être obtenue en sommant les éléments de la ligne correspondante de la matrice d'incidence de l'hypergraphe. Pour rappel, la matrice associée à l'hypergraphe de la figure 2.3 est

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

Le degré du noeud k_i est alors obtenu en sommant les éléments de la ligne i :

$$k_i = \sum_j J_{ij}.$$

Concernant la taille des hyperarêtes de cet hypergraphe, nous constatons que

$$\begin{array}{ll} m_1 = 3 & m_3 = 3 \\ m_2 = 2 & m_4 = 1, \end{array}$$

où m_1 est la taille de l'hyperarête e_1 , m_2 la taille de l'hyperarête e_2 et ainsi de suite pour toutes les hyperarêtes.

Nous pouvons dès lors vérifier la propriété de l'équation 2.4 :

$$\begin{aligned} \sum_i k_i &= 1 + 2 + 3 + 1 + 1 + 1 = 9, \\ \sum_j m_j &= 3 + 2 + 3 + 1 = 9, \end{aligned}$$

et donc,

$$L = \sum_i k_i = \sum_j m_j = 9.$$

Les deux expressions de l'équation 2.4 reflètent le fait que le nombre total de liens dans l'hypergraphe est égal à la somme des degrés de tous les noeuds (première expression) et à la somme des tailles de toutes les hyperarêtes (deuxième expression).

En utilisant cette équation, on peut établir une relation entre le degré moyen des noeuds $\langle k \rangle$ et la taille moyenne des hyperarêtes $\langle m \rangle$. En effet, en moyenne, chaque hyperarête connecte $\langle m \rangle$ noeuds, ce qui donne une contribution $\langle m \rangle$ au nombre total de liens L . De même, chaque noeud est connecté en moyenne aux $\langle k \rangle$ hyperarêtes, ce qui donne une contribution $\langle k \rangle$ au nombre total de liens L . On a donc :

$$N\langle k \rangle = E\langle m \rangle, \quad (2.5)$$

où N est le nombre total de noeuds dans l'hypergraphe, $\langle k \rangle$ est le degré moyen des noeuds et $\langle m \rangle$ est la taille moyenne des hyperarêtes.

Nous pouvons également illustrer cette propriété grâce à l'exemple de l'hypergraphe à la figure 2.3. En effet, dans notre cas,

$$\begin{aligned} N &= 6 & E &= 4 \\ \langle k \rangle &= \frac{\sum_i k_i}{N} = \frac{9}{6} & \langle m \rangle &= \frac{\sum_j m_j}{E} = \frac{9}{4}, \end{aligned}$$

et donc la propriété de l'équation 2.5 est bien vérifiée :

$$\begin{aligned} & N\langle k \rangle = E\langle m \rangle, \\ \iff & 6 \cdot \frac{9}{6} = 4 \cdot \frac{9}{4}, \\ \iff & 9 = 9. \end{aligned}$$

La relation de l'équation 2.5 montre que le degré moyen des noeuds est proportionnel à la taille moyenne des hyperarêtes, avec un facteur de proportionnalité qui dépend du nombre total de noeuds dans l'hypergraphe. En d'autres termes, plus les hyperarêtes sont grandes en moyenne, plus les noeuds auront en moyenne un degré élevé. Cette relation entre le degré moyen et la taille moyenne est importante, car elle permet de contrôler la structure globale de l'hypergraphe. En fixant la taille moyenne des hyperarêtes, on peut ainsi contrôler le degré moyen des noeuds et donc la densité des connexions dans l'hypergraphe.

2.2.2 Hypergraphes aléatoires simples

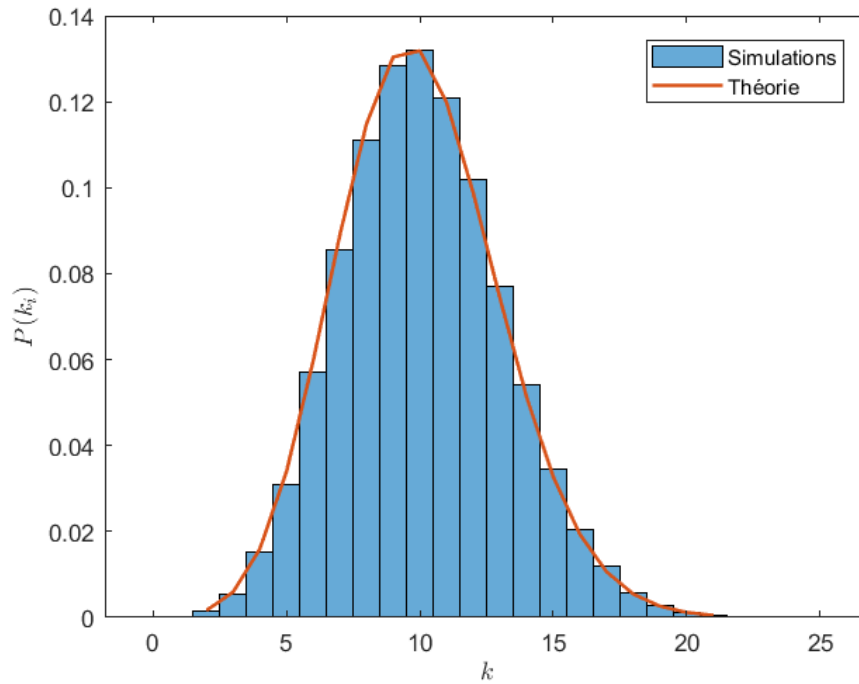
Dans ce modèle à un seul paramètre, la probabilité de connexion est fixée à

$$P(v \in e) = p, \quad (2.6)$$

où $p \in [0, 1]$. Autrement dit, chaque noeud v est choisi indépendamment avec une probabilité uniforme p d'appartenir à chaque hyperarête e .

Pour construire un hypergraphe avec ce modèle, on commence avec un ensemble donné N de noeuds. Ensuite, on choisit un entier E compris entre 1 et $\binom{N}{k}$, qui représente les noeuds initiaux, choisis arbitrairement. Ces E noeuds choisis vont former les hyperarêtes de départ. Pour chacune d'entre elles, on ajoute de manière récursive des noeuds à celles-ci en suivant la règle décrite par l'équation 2.6.

Nous pouvons à présent exposer différentes propriétés de ce modèle. La distribution du degré k peut être définie en se basant sur la probabilité qu'un noeud i soit de degré k dans

FIGURE 2.4 – Distribution du degré des noeuds k pour $p = 0.1$.

le modèle d'Erdős-Rényi (section 1.2). Dès lors, la probabilité qu'un noeud appartienne à k hyperarêtes est donnée par :

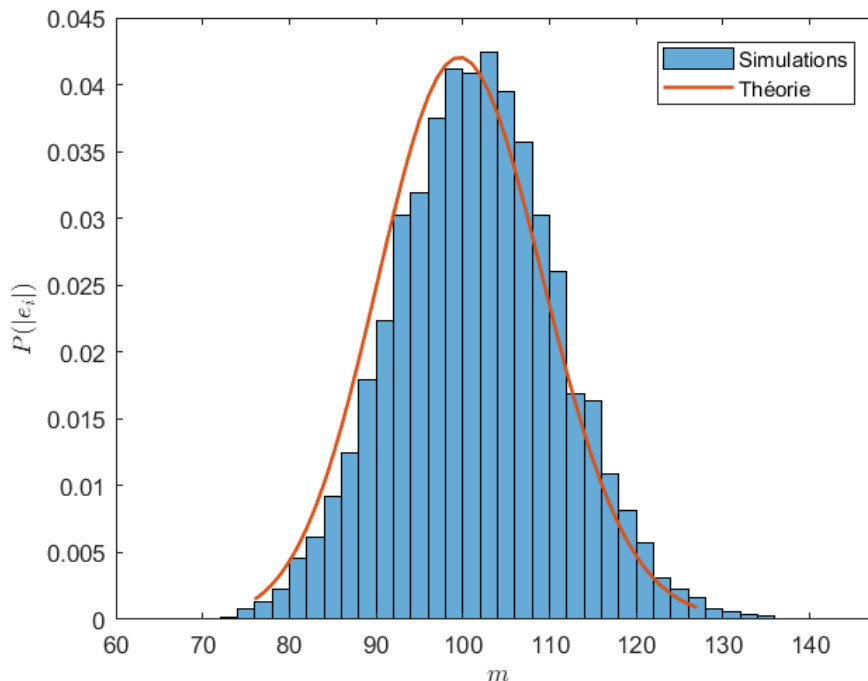
$$P(k_i = k) = \binom{E}{k} p^k (1-p)^{E-k}, \quad (2.7)$$

où E est le nombre total d'hyperarêtes dans l'hypergraphe et p est la probabilité qu'un noeud soit inclus dans une hyperarête donnée. Cette distribution décrit donc la probabilité que le noeud soit de degré k . De manière analogue, on peut définir la distribution de la taille m d'une hyperarête comme étant :

$$P(|e_i| = m) = \binom{N}{m} p^m (1-p)^{N-m}, \quad (2.8)$$

où N est le nombre total de noeuds dans l'hypergraphe. Cette distribution décrit donc la probabilité qu'une hyperarête donnée soit de taille m .

Les figures 2.4 et 2.5 permettent d'illustrer les distributions de la probabilité qu'un noeud appartienne à k hyperarêtes (équation 2.7) et la probabilité qu'une hyperarête donnée soit de taille m (équation 2.8), respectivement. Ces graphiques ont été obtenus en considérant que le nombre de noeuds N est fixé à 1000 et le nombre d'hyperarêtes E est fixé à 100. De plus, on considère que la probabilité p qu'un noeud appartienne à une hyperarête est de 0.1 et on fait cela pour 100 configurations. Les graphiques obtenus permettent d'illustrer le fait que le degré k d'un noeud et la taille m d'une hyperarête sont des variables dont les distributions respectives sont des lois binomiales de paramètres $E = 100$ et $p = 0.1$, pour le premier, et $N = 1000$ et $p = 0.1$ pour le second, les probabilités respectives étant décrites par les équations 2.7 et 2.8.


 FIGURE 2.5 – Distribution de la taille des hyperarêtes m pour $p = 0.1$.

Le degré moyen d'un noeud, noté $\langle k \rangle$, est la moyenne des degrés de tous les noeuds, à savoir $\sum_k kP(k)$. Ainsi, $\langle k \rangle$ est le produit de la probabilité p que ce noeud appartienne à une hyperarête donnée et du nombre total d'hyperarêtes E dans l'hypergraphe. Formellement,

$$\langle k \rangle = pE.$$

De même, la taille moyenne d'une hyperarête, notée $\langle m \rangle$, est la moyenne des tailles de toutes les hyperarêtes, à savoir $\sum_m mP(m)$. Ainsi, $\langle m \rangle$ est le produit de la probabilité p qu'un noeud soit inclus dans cette hyperarête et du nombre total de noeuds N dans l'hypergraphe. Formellement,

$$\langle m \rangle = pN.$$

Si les noeuds sont dans l'espace, nous pouvons également déterminer l'étendue spatiale moyenne, notée $s(e_i)$ pour une hyperarête e_i . Cette mesure permet d'évaluer la séparation des noeuds au sein de l'hyperarête e_i . Plus les noeuds sont séparés les uns des autres, plus l'étendue spatiale est grande. Il faut donc associer des coordonnées aux noeuds en choisissant arbitrairement un rayon et un angle afin que ces noeuds se trouvent dans un disque de rayon r_0 . On peut ensuite calculer l'étendue spatiale en prenant les coordonnées de toutes les paires de noeuds v_j et v_l dans l'hyperarête e_i , et en calculant la distance Euclidienne entre ces noeuds dans l'espace :

$$d_E(v_j, v_l) = \sqrt{\sum_{k=1}^d (v_{j,k} - v_{l,k})^2},$$

où d représente la dimension de l'espace dans lequel se trouvent les noeuds v_j et v_l , $v_{j,k}$ est la $k^{\text{ème}}$ coordonnée du point v_j et $v_{l,k}$ est la $k^{\text{ème}}$ coordonnée du point v_l . Dès lors,

l'étendue spatiale moyenne peut être définie comme étant :

$$s(e_i) = \frac{1}{m_i(m_i - 1)} \sum_{v_j, v_l \in e_i} d_E(v_j, v_l).$$

Concernant les hypergraphes d'Erdős-Rényi générés selon ce modèle, les noeuds sont distribués uniformément dans les hyperarêtes. Dès lors, l'étendue spatiale moyenne est calculée sur des noeuds choisis arbitrairement dans le rayon initial r_0 . Cette quantité, calculée dans [11]¹ est donnée par

$$s = s_0 = \frac{128r_0}{45\pi}.$$

Le modèle étudié permet aux hyperarêtes d'avoir des noeuds en commun, ce qui signifie qu'elles peuvent s'intersecter. Si on considère deux hyperarêtes e_i et e_j , on peut définir leur intersection I comme étant [6]

$$I = |e_i \cap e_j|, \quad (2.9)$$

où $|\cdot|$ représente le cardinal d'un ensemble. En d'autres termes, I est le nombre d'éléments en commun entre e_i et e_j . Plus précisément, si on considère deux hyperarêtes e_i et e_j contenant respectivement m_i et m_j noeuds, on peut déterminer la taille de leur intersection en trouvant le nombre de noeuds qui sont présents dans les deux hyperarêtes, c'est-à-dire en calculant $|e_i \cap e_j|$. Cette quantité peut varier de 0, si les deux hyperarêtes n'ont aucun noeud en commun, à $\min(m_i, m_j)$, si une hyperarête est contenue dans l'autre.

Dans ce cas, on peut déterminer l'intersection moyenne, qui est le nombre moyen de fois qu'un noeud est présent dans deux hyperarêtes distinctes, ce qui est égal à la probabilité qu'un noeud soit présent dans deux hyperarêtes, multipliée par le nombre total de noeuds. Sachant que la probabilité qu'un noeud appartienne à une hyperarête est p , la probabilité qu'un noeud appartienne à deux hyperarêtes est donnée par le produit des probabilités qu'il soit dans chacune des deux hyperarêtes, à savoir p^2 . Le nombre total de noeuds étant N , l'intersection moyenne est donnée par

$$\langle I \rangle = p^2 N. \quad (2.10)$$

La figure 2.6, qui affiche le nombre moyen d'intersections $\langle I \rangle$ en fonction du nombre de noeuds N , permet de vérifier cette relation. En effet, la ligne rouge représente la droite d'équation $f(N) = p^2 N$, pour N variant de 0 à 2000, alors que les ronds bleus représentent les simulations numériques effectuées en considérant une probabilité p de 0.1, un nombre d'hyperarêtes E fixé à 100, un nombre de noeuds N variant de 0 à 2000 et cela pour 100 configurations.

Une mesure du nombre d'intersections d'hyperarêtes contenant un noeud donné est exprimée par la distribution de probabilité de la variable aléatoire I :

$$P(I = n) = \binom{N}{n} p^{2n} (1 - p^2)^{N-n}. \quad (2.11)$$

Cette équation est un produit de trois termes. $\binom{N}{n}$ est le nombre de façons de choisir n éléments à partir d'un ensemble de N éléments. Cela correspond aux différentes manières

1. Le lecteur intéressé par la détermination de cette étendue spatiale peut se référer à [11]

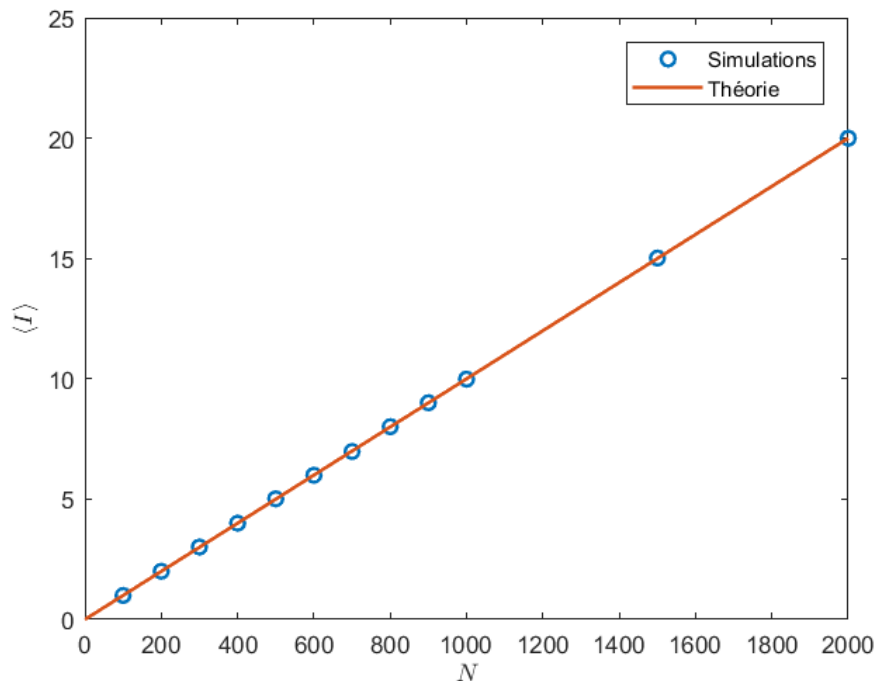


FIGURE 2.6 – Nombre d’intersection moyenne $\langle I \rangle$ en fonction du nombre de noeuds N avec $p = 0.1$.

de choisir les noeuds qui appartiennent à la fois à la première hyperarête et à la seconde. Ensuite, p^{2n} désigne la probabilité que les n noeuds choisis appartiennent aux deux hyperarêtes. Comme chaque hyperarête a une probabilité p de contenir un noeud donné, la probabilité que deux hyperarêtes contiennent le même noeud est p^2 . Finalement, le terme $(1 - p^2)^{N-n}$ désigne la probabilité que les $N - n$ noeuds restants n’appartiennent à aucune des deux hyperarêtes considérées. Cela correspond à la probabilité qu’un noeud donné n’appartienne à aucune des deux hyperarêtes, qui est $(1 - p^2)$, et nous avons $N - n$ noeuds qui doivent tous satisfaire cette condition. Par conséquent, la probabilité totale est donnée par $(1 - p^2)^{N-n}$.

La figure 2.7 permet d’illustrer la distribution de la probabilité d’intersection entre des hyperarêtes (équation 2.11). Ce graphique a été obtenu en utilisant les mêmes données de l’article, c’est-à-dire qu’on considère que le nombre de noeuds N est fixé à 1000, le nombre d’hyperarêtes E est fixé à 100, la probabilité p qu’un noeud appartienne à une hyperarête est de 0.1 et on considère cela pour 100 configurations. Le graphique obtenu permet d’illustrer le fait que l’intersection I de deux hyperarêtes est une variable dont la distribution est une loi binomiale de paramètres $N = 1000$ et $p^2 = 0.01$, et dont la probabilité est décrite par l’équation 2.11.

Nous pouvons aller plus loin en définissant l’intersection I_{jl} entre deux hyperarêtes de tailles j et l , respectivement. Plus précisément, I_{jl} est le nombre d’éléments en commun entre deux hyperarêtes e_j et e_l de taille j et l [6]

$$I_{jl} = |e_j \cap e_l|.$$

On peut donc voir I_{jl} comme une variable aléatoire qui prend des valeurs dans $\{0, 1, \dots, \min(j, l)\}$. La probabilité associée à cette intersection est alors donnée par $N_{jl}(k)$. Ce coefficient mul-

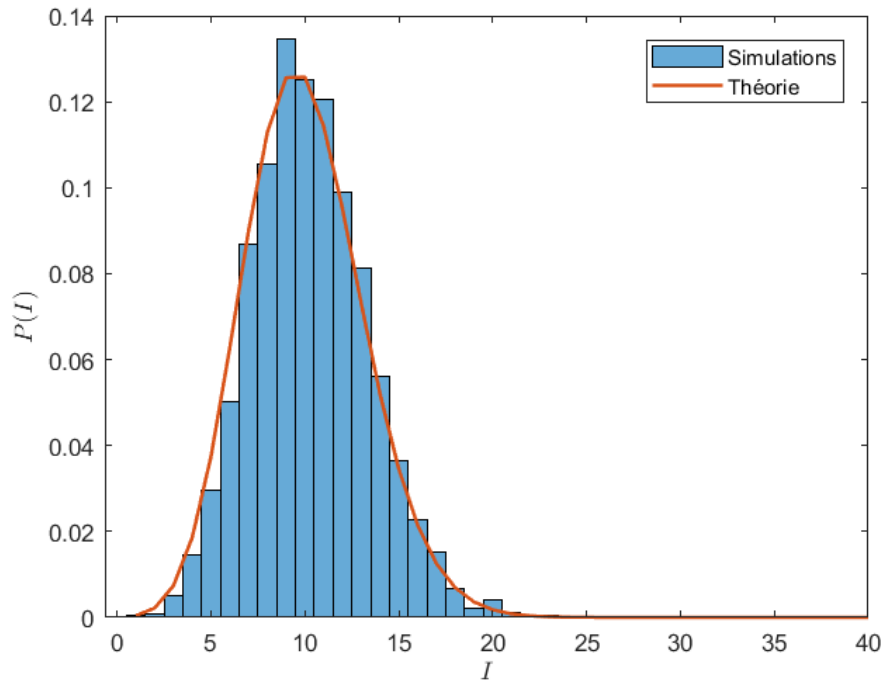


FIGURE 2.7 – Distribution des intersections des hyperarêtes I pour $p = 0.1$.

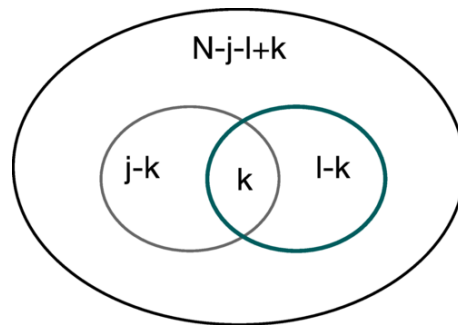


FIGURE 2.8 – Schéma de l’intersection de deux hyperarêtes de taille j et l , respectivement, tiré de [3]

tinomial représente le nombre de paires d’hyperarêtes de taille j et l ayant une intersection de taille k . Plus précisément, sachant qu’on a N noeuds, on fait quatre groupe de noeuds, un qui en contient k , un autre qui en contient $(j - k)$, un autre qui en contient $(l - k)$ et le dernier qui en contient $(N - j - l + k)$, comme illustré à la figure 2.8. La probabilité associée à l’intersection I_{jl} est alors définie par :

$$P(I_{jl} = k) = N_{jl}(k) = \frac{N!}{k!(j - k)!(l - k)!(N - j - l + k)!},$$

qui représente le nombre de paires d’hyperarêtes de taille j et l ayant une intersection de taille k .

Comme pour les graphes aléatoires d’Erdős-Rényi (section 1.2), nous pouvons étudier le comportement de la composante géante en fonction de la variation de la probabilité p qui contrôle la formation d’hyperarêtes dans l’hypergraphe aléatoire. Pour définir la

composante géante, on suppose que tous les nœuds appartenant à la même hyperarête sont connectés entre eux (équivalant à chaque hyperarête étant un graphe complet) et que deux hyperarêtes sont connectées si leur intersection est supérieure ou égale à un (c'est-à-dire si elles ont au moins un nœud en commun). Il existe d'autres définitions de composante géante, notamment en utilisant la notion de marche d'hypergraphe d'ordre supérieur ("high-order hypergraphs walk") [1].

Dans un tel contexte, la probabilité qu'il y ait au moins une intersection entre deux hyperarêtes est donnée par :

$$P(I \geq 1) = 1 - (1 - p^2)^N, \\ \approx p^2 N.$$

En effet, la probabilité qu'aucune intersection n'existe entre deux hyperarêtes est $(1 - p^2)^N$, donc la probabilité qu'il y ait au moins une intersection est égale à 1 moins la probabilité qu'il n'y en ait aucune.

De plus, si le nombre d'hyperarêtes E est suffisamment grand, il y aura une composante géante dans l'hypergraphe aléatoire si et seulement si les hyperarêtes constituent une composante géante ; en d'autres termes, si on peut trouver une composante géante dans l'hypergraphe aléatoire.

Le résultat classique pour les graphes aléatoires d'Erdős-Rényi (section 1.2) énonce qu'il y a une composante géante qui apparaît lorsque le degré moyen est égal à 1. On peut dès lors adapter cette idée au cas des hypergraphes aléatoires en considérant que la condition pour avoir une composante géante dans le sous-graphe induit par les hyperarêtes est $p_c^2 NE = 1$, où p_c est la probabilité critique pour laquelle la composante géante apparaît dans l'hypergraphe aléatoire et N est le nombre de nœuds dans l'hypergraphe. En utilisant l'approximation $P(I \geq 1) \approx p^2 N$, on peut exprimer la probabilité critique p_c en fonction de N et E comme suit [15] :

$$p_c \sim \frac{1}{\sqrt{NE}},$$

ce qui montre que la probabilité critique diminue lorsque le nombre d'hyperarêtes ou le nombre de nœuds dans l'hypergraphe augmente.

2.2.3 Attachement préférentiel pour les hypergraphes

Dans cette section, nous allons étendre la notion d'attachement préférentiel vue pour les graphes dans la section 1.3 aux hypergraphes. Pour rappel, ce modèle d'attachement préférentiel pour les graphes déclare que la probabilité qu'un nouveau nœud se connecte à un nœud existant est proportionnelle au degré de ce dernier.

Afin d'étendre cette notion d'attachement préférentiel aux hypergraphes, on construit un modèle qui, selon l'article [24] cité dans [3] comprend deux éléments principaux. Premièrement, à chaque étape de la construction d'un hypergraphe, une hyperarête est créée, soit en choisissant aléatoirement un nœud et en lui ajoutant m nouveaux nœuds afin de créer une nouvelle hyperarête, soit en sélectionnant un nombre aléatoire de nœuds pour former une hyperarête et en y ajoutant un nouveau nœud. Deuxièmement, on introduit un mécanisme d'attachement préférentiel pour les hyperarêtes, où la probabilité qu'une hyperarête se connecte à un nœud existant est proportionnelle au degré de ce dernier.

Plusieurs approches sont possibles afin d'introduire l'attachement préférentiel dans la construction d'hypergraphe. Dans le cadre de notre analyse, nous allons choisir de décrire

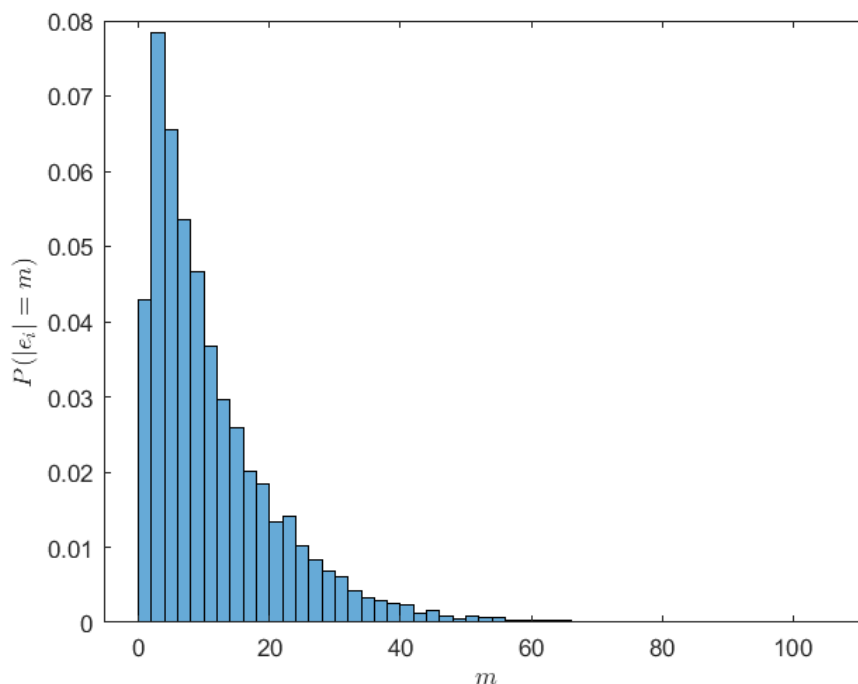


FIGURE 2.9 – Probabilité qu’une hyperarête soit de taille m dans le modèle d’attachement préférentiel.

la probabilité de connexion comme dans l’article [3], c’est-à-dire que celle-ci est définie comme une fonction de la taille de l’hyperarête de taille $m = |e|$

$$P(v \in e) = F[|e|].$$

La fonction F utilisée dans ce cadre est la fonction linéaire suivante :

$$P(v \in e_i) = \frac{|e_i|}{\sum_j |e_j|}. \quad (2.12)$$

Dans cette équation, $|e_i|$ représente la taille de l’hyperarête e_i , c’est-à-dire le nombre de noeuds qu’elle contient, et $\sum_j |e_j|$ est la somme des tailles de toutes les hyperarêtes existantes dans l’hypergraphe. Cette équation indique que la probabilité qu’un noeud v appartienne à une hyperarête e_i est proportionnelle à la taille de cette hyperarête par rapport à la somme des tailles de toutes les hyperarêtes existantes. En d’autres termes, plus une hyperarête est grande par rapport aux autres, plus elle aura de chances d’attirer de nouvelles connexions de noeuds.

L’article [3] énonce le fait que la distribution limite pour ce modèle d’hypergraphes est une distribution multinomiale de Dirichlet

$$P_t(m_1, m_2, \dots, m_E | \alpha_1, \alpha_2, \dots, \alpha_E) = \frac{\Gamma(\alpha_0) \Gamma(t+1)}{\Gamma(t + \alpha_0)} \prod_{j=1}^E \frac{\Gamma(m_j + \alpha_j)}{\Gamma(\alpha_j) \Gamma(m_j + 1)},$$

où $m_j = |e_j|$, α_j désigne la taille initiale de l’hyperarête j , $\alpha_0 = \sum \alpha_m$ et $\Gamma(x)$ est la fonction gamma [3].

La figure 2.9 permet d’illustrer la distribution de la probabilité qu’une hyperarête soit de taille m . Ce graphique a été obtenu en considérant un nombre de noeuds N de 1000, un nombre d’hyperarêtes E de 100, avec une probabilité calculée avec l’équation 2.12 et cela pour 100 configurations.

2.2.4 Hypergraphes spatiaux aléatoires

Cette section décrit une extension du modèle d'hypergraphes aléatoires précédemment présenté, en incluant des informations spatiales dans la génération de l'hypergraphe. Le modèle consiste à ajouter des noeuds à l'aide d'un processus de croissance itératif, tout en prenant en compte les coordonnées spatiales de chaque noeud et la distance entre les noeuds pour déterminer la probabilité de connecter un nouveau noeud à des noeuds existants.

Plus précisément, la probabilité qu'un nouveau noeud soit connecté à un noeud existant dépend à la fois de la distance spatiale entre les deux noeuds et de leur degré. Cette approche permet de générer des hypergraphes qui présentent des propriétés topologiques et spatiales similaires à celles observées dans de nombreux systèmes réels, tels que les réseaux de transport ou les réseaux neuronaux. En effet, dans de nombreux systèmes, tels que la propagation d'une maladie infectieuse, les réseaux de communication ou les réseaux neuronaux, l'espace joue un rôle important en raison de la proximité nécessaire pour les interactions de plus haut niveau.

En considérant que les N noeuds sont distribués uniformément sur un disque de rayon r_0 , on peut définir la probabilité qu'un nouveau noeud v soit connecté à une hyperarête e comme

$$P(v \in e) = F[d(v, e)],$$

qui dépend de la distance spatiale $d(v, e)$ entre le nouveau noeud v et l'hyperarête e .

Plusieurs choix sont possibles afin de définir la distance entre un noeud et une hyperarête. Dans un premier temps, nous allons considérer que F est une fonction de décroissance exponentielle, $F(d) \sim \exp(-d/r_c)$, où r_c est une échelle de distance caractéristique. Dans un second temps, nous allons considérer un modèle proche du modèle du graphe aléatoire géométrique, qui est un modèle classique de génération de graphes dans lequel les noeuds sont distribués uniformément dans l'espace et sont connectés s'ils sont situés à une distance inférieure à un certain seuil [12].

Cas exponentiel

Dans cette partie, nous allons considérer le cas où la fonction F est une fonction exponentielle décroissante avec la distance. L'objectif est de capturer l'idée intuitive selon laquelle il est moins probable qu'un noeud fasse partie d'une hyperarête éloignée. Cette intuition est basée sur le fait qu'il est généralement plus difficile pour les noeuds éloignés de se connecter dans un contexte spatial. Ainsi, la probabilité de connexion entre un noeud v et une hyperarête e est donnée par

$$P(v \in e) = pe^{-d(v,e)/r_c},$$

où p est un paramètre compris entre 0 et 1, et r_c est la portée de la décroissance exponentielle. La distance $d(v, e)$ mesure la distance entre le noeud v et l'hyperarête e .

Le paramètre r_c contrôle l'étendue spatiale sur laquelle la probabilité de connexion diminue exponentiellement. Plus r_c est grand, plus cette décroissance est lente, ce qui signifie que les noeuds ont une plus grande probabilité de se connecter même s'ils sont éloignés spatialement. À l'inverse, lorsque r_c est petit, la probabilité de connexion diminue rapidement avec la distance, ce qui rend moins probable la connexion de noeuds éloignés.

Concernant la mesure de distance $d(v, e)$, nous allons considérer les trois définitions de

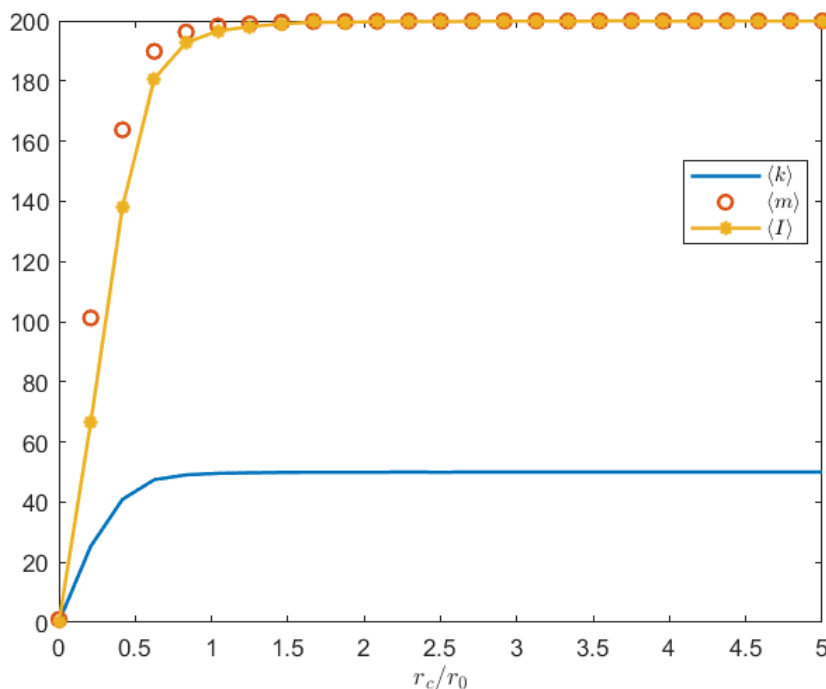


FIGURE 2.10 – Taille moyenne des hyperarêtes, degré moyen et intersection moyenne, en fonction de r_c/r_0 pour $p = 0.1$ avec la distance moyenne.

distances proposées dans l'article [3]. Premièrement, nous allons prendre la distance euclidienne entre le noeud v et chaque noeud w de l'hyperarête e , et ensuite prendre la moyenne sur tous les noeuds de l'hyperarête :

$$d(v, e) = \frac{1}{m} \sum_{w \in e} d_E(v, w), \quad (2.13)$$

où m est la taille de l'hyperarête e et $d_E(v, w)$ est la distance euclidienne entre v et w .

Selon l'article [3], lorsque $r_c \gg r_0$, c'est-à-dire lorsque la portée de décroissance exponentielle r_c est beaucoup plus grande que le rayon initial de distribution des noeuds r_0 , le modèle d'hypergraphe aléatoire sans contrainte spatiale est retrouvé. Cela est dû au fait que, dans de telles conditions, le terme exponentiel est proche de 1 et que par conséquent, la notion d'espace n'est plus pertinente. En d'autres termes, cela signifie essentiellement que l'influence de l'espace devient négligeable et que le modèle se comporte comme un hypergraphe aléatoire simple.

Nous pouvons dès lors étudier le comportement des quantités suivantes en fonction de la portée de décroissance exponentielle r_c : le degré moyen $\langle k \rangle$, la taille moyenne des hyperarêtes $\langle m \rangle$ et l'intersection moyenne des hyperarêtes $\langle I \rangle$. Les figures 2.10 et 2.11 illustrent l'évolution de ces quantités en fonction de r_c/r_0 , où l'étendue spatiale est représentée sur une figure séparée en vue de son ordre de grandeur très petit par rapport aux autres quantités. Ces figures ont été générées avec les paramètres suivants : le nombre de noeuds N a été fixé à 200, le nombre d'hyperarêtes E à 50, la probabilité p qu'un noeud appartienne à une hyperarête est de 0.1, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations. Cette figure permet difficilement d'interpréter le comportement des différentes courbes en raison de leurs différentes échelles associées.

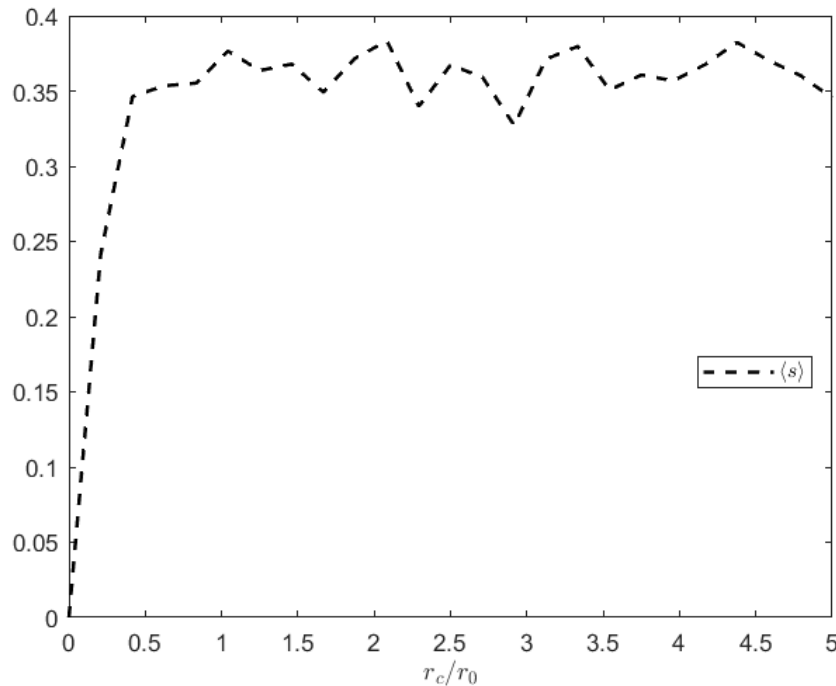


FIGURE 2.11 – Étendue spatiale moyenne en fonction de r_c/r_0 pour $p = 0.1$ avec la distance moyenne.

Pour pallier ce problème, nous pouvons normaliser les différentes quantités pour les mettre à l'échelle en fonction de p , E et N .

La figure 2.12 illustre le comportement de ces différentes quantités après une mise à l'échelle. Pour ce faire, nous avons multiplié le degré moyen $\langle k \rangle$ par pE , où E est le nombre moyen d'hyperarêtes. Concernant la taille moyenne des hyperarêtes $\langle m \rangle$, nous l'avons multipliée par pN , où N est le nombre de noeuds. Finalement, nous avons multiplié l'intersection moyenne $\langle I \rangle$ par Np^2 . Cette figure a été obtenue en considérant les paramètres suivants : le nombre de noeuds N a été fixé à 200, le nombre d'hyperarêtes E à 50, la probabilité p qu'un noeud appartienne à une hyperarête est de 0.1, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations.

Grâce à la figure 2.12, nous pouvons constater que les différentes quantités augmentent à mesure que la portée de décroissance exponentielle augmente, avec un ralentissement de la croissance lorsque le rapport r_c/r_0 devient grand. Cela signifie, pour le degré, que les noeuds ont tendance à se connecter à plus d'hyperarêtes ; pour la taille, que les hyperarêtes ont tendance à inclure un plus grand nombre de noeuds ; et pour les intersections, que les hyperarêtes ont tendance à partager davantage de noeuds. De plus, nous pouvons observer que les courbes représentant le degré moyen et la taille moyenne se chevauchent, ce qui se justifie par le fait que le nombre total de liens est donné par $\sum_j k_j = \sum_i |e_i|$, ce qui implique que

$$\frac{\langle m \rangle}{pN} = \frac{\langle k \rangle}{pE}.$$

Deuxièmement, nous pouvons considérer la distance minimale définie par

$$d(v, e) = \min_{w \in e} d_E(v, w), \quad (2.14)$$

où $d_E(v, w)$ est la distance euclidienne entre v et w . La figure associée est la figure 2.13 qui

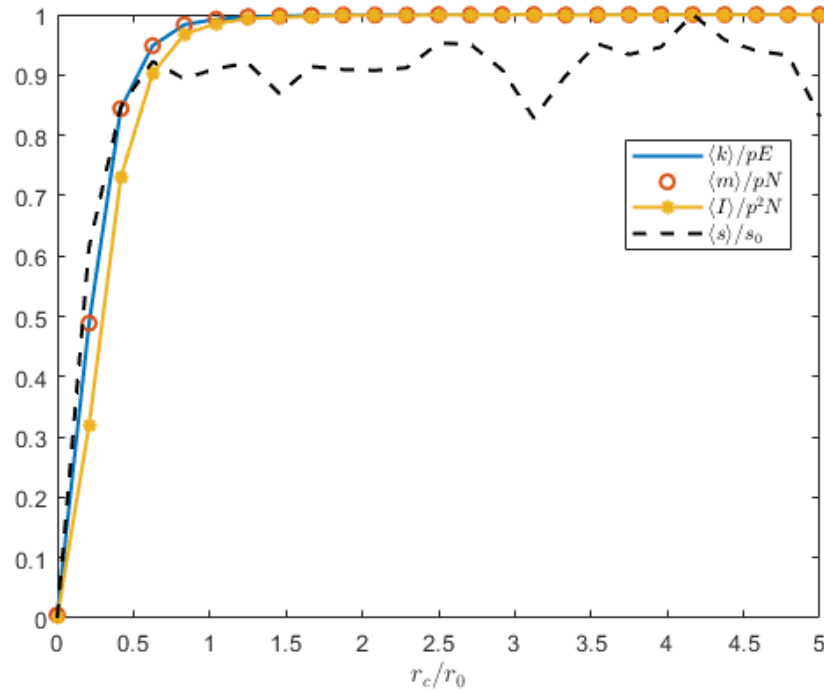


FIGURE 2.12 – Taille moyenne des hyperarêtes, degré moyen et intersection moyenne, normalisés, en fonction de r_c/r_0 pour $p = 0.1$ avec la distance moyenne.

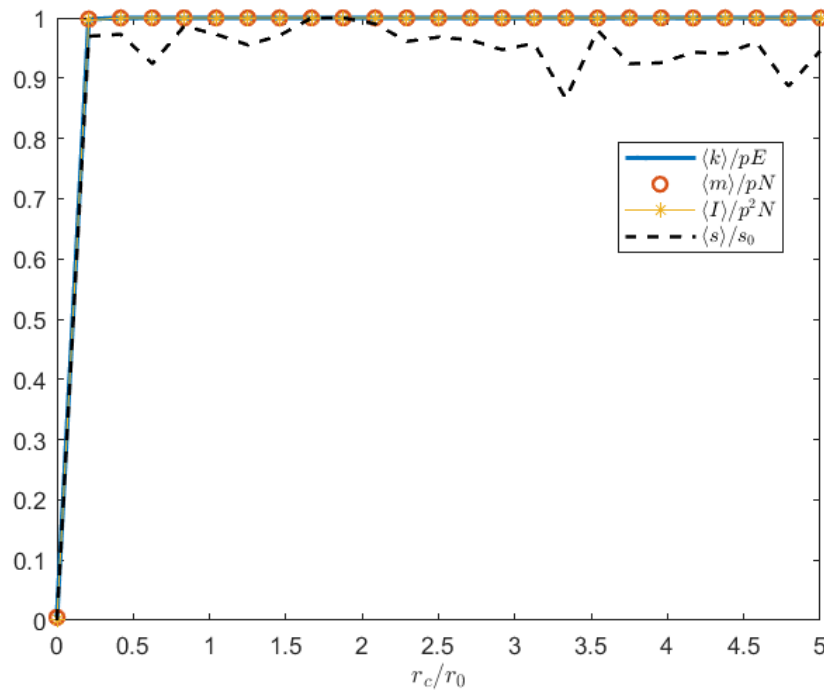


FIGURE 2.13 – Taille moyenne des hyperarêtes, degré moyen et intersection moyenne, normalisés, en fonction de r_c/r_0 pour $p = 0.1$ avec la distance minimale.

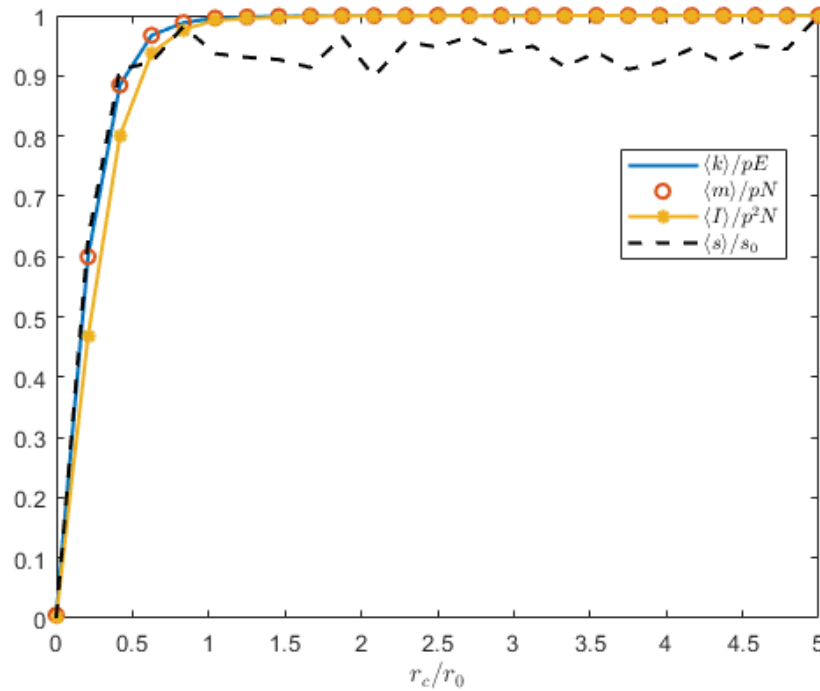


FIGURE 2.14 – Taille moyenne des hyperarêtes, degré moyen et intersection moyenne, normalisés, en fonction de r_c/r_0 pour $p = 0.1$ avec la distance par rapport à un centroïde.

permet d'illustrer le comportement des mêmes quantités normalisées que pour la distance précédente. Cette figure a été obtenue avec les mêmes paramètres que pour la première définition de distance, c'est-à-dire que le nombre de noeuds N est fixé à 200, le nombre d'hyperarêtes E est fixé à 50, la probabilité p qu'un noeud appartienne à une hyperarête est de 0.1, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations.

Troisièmement, nous pouvons prendre la distance par rapport à un centroïde qui est définie par

$$d(v, e) = d(c, c(e)), \quad (2.15)$$

où $c(e)$ est le centroïde qui a pour coordonnées :

$$(x_c(e), y_c(e)) = \frac{1}{|e|} \sum_{i \in e} (x(i), y(i)).$$

La figure correspondante est la figure 2.14 qui a été obtenue avec les mêmes paramètres que pour les deux définitions de distances précédentes.

Les graphiques obtenus pour les trois distances étant assez similaires, nous pouvons étudier la convergence de l'étendue spatiale moyenne pour ces trois définitions. Pour cela, nous avons la figure 2.15 qui représente un moins la moyenne de l'étendue spatiale normalisée, c'est-à-dire $1 - \langle s \rangle / s_0$ en fonction de r_c/r_0 . Pour rappel, l'étendue spatiale est définie par

$$s(e_i) = \frac{1}{m_i(m_i - 1)} \sum_{v_j, v_l \in e_i} d_E(v_j, v_l).$$

Dans notre cas, nous considérons que la distance est calculée avec les trois définitions vues précédemment 2.13, 2.14 et 2.15. Cette figure a été obtenue en considérant les paramètres

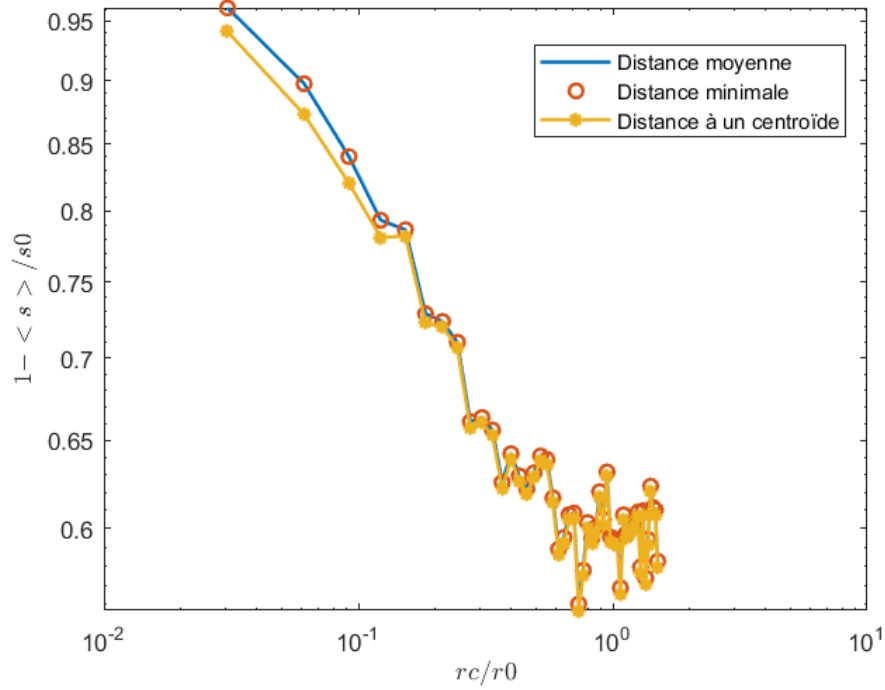


FIGURE 2.15 – Étendue spatiale moyenne normalisée en fonction de r_c/r_0 pour $p = 0.1$ avec les trois distances.

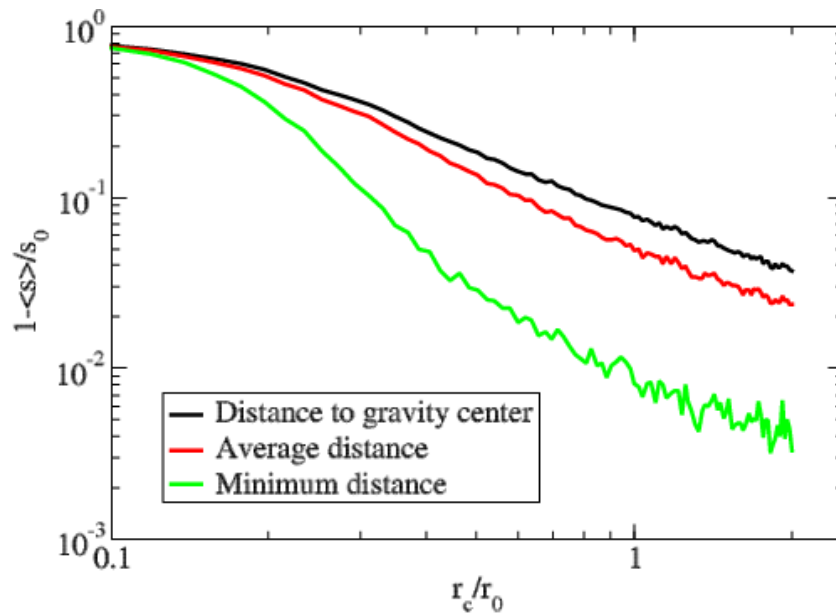


FIGURE 2.16 – Étendue spatiale moyenne normalisée en fonction de r_c/r_0 tiré de [3].

suyvants : le nombre de noeuds N est fixé à 200, le nombre d'hyperarêtes E est fixé à 50, la probabilité p qu'un noeud appartienne à une hyperarête est de 0.1, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations.

Grâce à la figure 2.15, nous pouvons constater que les étendues spatiales moyennes associées aux distances 2.13 et 2.15 se comportent de manière très similaire alors que celle associée à la distance 2.14 semble converger plus rapidement que les deux autres. L'étendue spatiale moyenne semble donc sensible au choix de la distance entre un noeud et une hyperarête qui est choisie pour définir la probabilité que ce noeud appartienne à l'hyperarête.

Cependant, nous pouvons constater que les courbes semblent osciller autour de 0.6. Or, si on compare avec la figure 2.16 de l'article [3], on se rend compte que les comportements ne sont pas similaires. En effet, sur la figure 2.16, les courbes semblent osciller autour de 0.06. Le facteur 10 entre ces deux résultats peut venir de plusieurs facteurs. Soit cela est issu d'une erreur de calcul ou de construction des hypergraphes, soit les paramètres utilisés ne sont pas les mêmes. De fait, dans nos simulations, nous avons par exemple fixé le rayon initial à $r_0 = 1$ alors que dans l'article il n'est pas précisé la valeur de ce dernier.

Hypergraphe aléatoire géométrique

Dans cette partie, nous allons étendre la notion de graphe aléatoire géométrique introduite par E. N. Gilbert dans l'article [12] aux hypergraphes. Dans cet article, les graphes aléatoires géométriques sont composés de noeuds, avec des coordonnées dans le plan en deux dimensions, qui sont connectés entre eux si la distance qui les sépare est inférieure à un certain seuil r_c , c'est-à-dire s'ils sont suffisamment proches.

Dans ce cadre, la densité de noeuds dans une certaine région A , définie par

$$\rho = \frac{N}{A}, \quad (2.16)$$

nous permet de définir le degré moyen des noeuds par

$$\langle k \rangle = \rho \pi r_c^2,$$

où on considère que les hyperarêtes sont des disques dont la surface est donnée par πr_c^2 .

Pour étendre cette notion aux hypergraphes, nous allons également considérer un seuil r_c en dessous duquel les noeuds sont connectés aux hyperarêtes. Pour cela, nous allons définir la probabilité de connexion entre un noeud v et une hyperarête e par

$$P(v \in e) = \Theta(r_c - d(v, e)), \quad (2.17)$$

où $d(v, e)$ est la distance entre le noeud v et l'hyperarête e , et où $\Theta(x)$ est la fonction de Heaviside [7]

$$\Theta(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0. \end{cases}$$

Dans la section précédente (Section 2.2.4,) nous avons travaillé avec trois définitions différentes pour la distance entre un noeud et une hyperarête. Pour notre étude des hypergraphes aléatoires géométriques, nous allons en considérer une quatrième qui est la distance maximale entre un noeud v et une hyperarête e

$$d(v, e) = \max_{w \in e} d(v, w).$$

En combinant cette définition de distance avec la probabilité définie par l'équation 2.17, on construit les hypergraphes aléatoires géométriques avec l'idée intuitive qu'un noeud se connectera à une hyperarête s'il est suffisamment proche de tous les noeuds déjà connectés à cette hyperarête, et se trouvant à une distance inférieure à un certain seuil r_c .

Comme pour l'hypergraphe aléatoire simple, nous allons exposer différentes propriétés de ce modèle telles que le degré moyen, la taille et les intersections entre les hyperarêtes. Tout d'abord, nous pouvons évaluer la probabilité qu'un noeud appartienne à une hyperarête, dans le cas où les noeuds sont distribués uniformément dans un disque de rayon r_0 . Cette probabilité est déterminée en comparant l'aire du cercle de rayon r_0 , qui représente la zone où les noeuds sont distribués uniformément, avec l'aire d'un cercle de rayon r_c , où r_c est le seuil qui détermine si deux noeuds sont suffisamment proches pour être connectés à la même hyperarête. La probabilité qu'un noeud donné appartienne à une hyperarête est donc donnée par $\pi r_c^2 / \pi r_0^2$.

Dès lors, en considérant que $r_0 = 1$, nous pouvons définir le degré moyen d'un noeud par

$$\langle k \rangle = E r_c^2,$$

où on a multiplié la probabilité qu'un noeud appartienne à une hyperarête par le nombre total d'hyperarêtes E .

De manière similaire, nous pouvons définir la taille moyenne d'une hyperarête comme étant le nombre de noeuds présents à une distance inférieure à un certain seuil r_c . Pour cela, nous pouvons considérer la densité de noeuds présents dans cette hyperarête, comme utilisée pour les graphes aléatoires géométriques (équation 2.16),

$$\langle m \rangle \approx \rho r_c^2 \sim N r_c^2.$$

Grâce à ces définitions, nous pouvons énoncer la même propriété que dans la section 2.2.2, qui est

$$\frac{\langle k \rangle}{E} = \frac{\langle m \rangle}{N}. \quad (2.18)$$

Cette relation souligne la corrélation entre le degré moyen des noeuds et la taille moyenne des hyperarêtes. De plus, ils révèlent une dépendance proportionnelle à r_c^2 .

Afin de vérifier ces définitions, nous pouvons construire des hypergraphes aléatoires géométriques en utilisant la probabilité de l'équation 2.17. En calculant alors le degré moyen des noeuds et la taille moyenne des hyperarêtes, on peut se rendre compte, grâce à la figure 2.17, que la relation 2.18 est bien vérifiée. En effet, les courbes pour le degré moyen des noeuds et la taille moyenne des hyperarêtes, respectivement normalisés par le nombre total de noeuds et le nombre total d'hyperarêtes, se superposent et ont un comportement quadratique proche de celui de $a r_c^2$, où $a = 0.4$. Cette figure a été obtenue en considérant les paramètres suivants : le nombre de noeuds N a été fixé à 200, le nombre d'hyperarêtes E à 50, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations.

Nous pouvons ensuite nous intéresser à l'intersection entre deux hyperarêtes. Pour quantifier cette intersection, nous allons utiliser l'aire, notée $A(r, l)$, de l'intersection entre deux disques de même rayon r et se trouvant à une distance l l'un de l'autre. Cette aire, représentée à la figure 2.18, est alors définie par

$$A(r, l) = \begin{cases} 2 \cos^{-1}\left(\frac{l}{2r}\right)r^2 - \frac{l}{2}\sqrt{4r^2 - l^2} & , \text{ si } l \leq 2r \\ 0 & , \text{ si } l > 2r \end{cases}$$

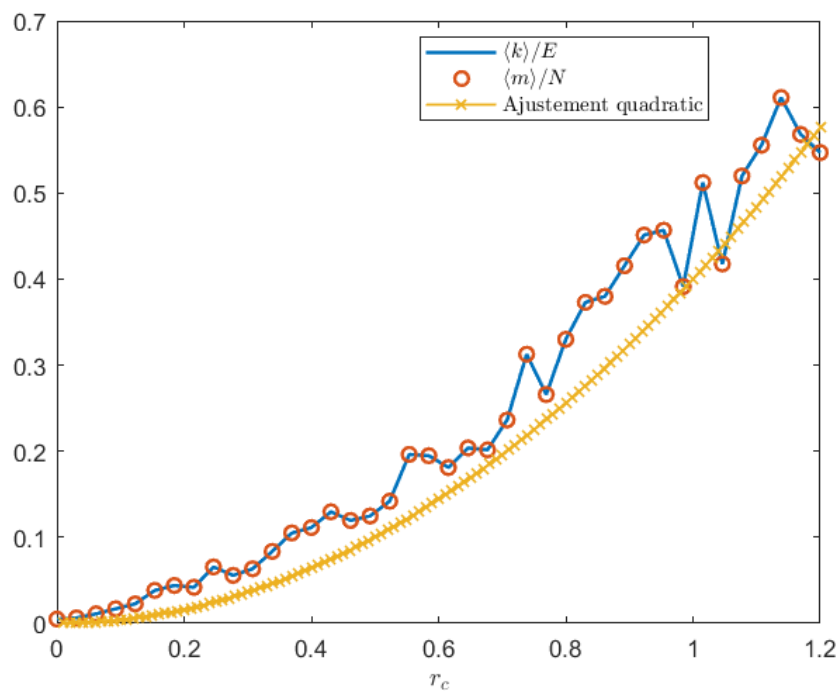


FIGURE 2.17 – Taille moyenne des hyperarêtes et degré moyen des noeuds, normalisés, en fonction de r_c .

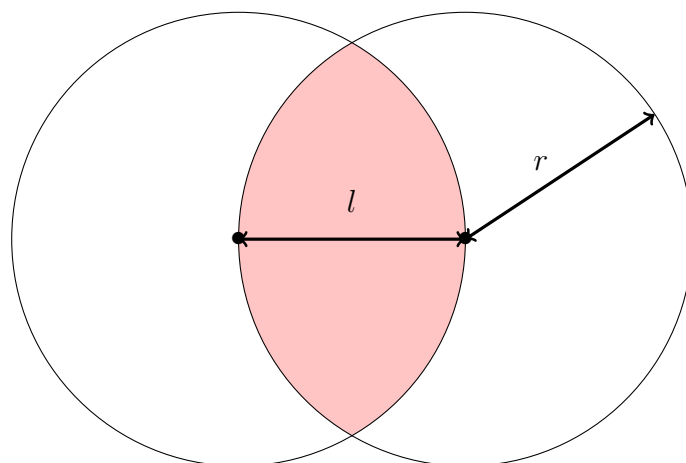
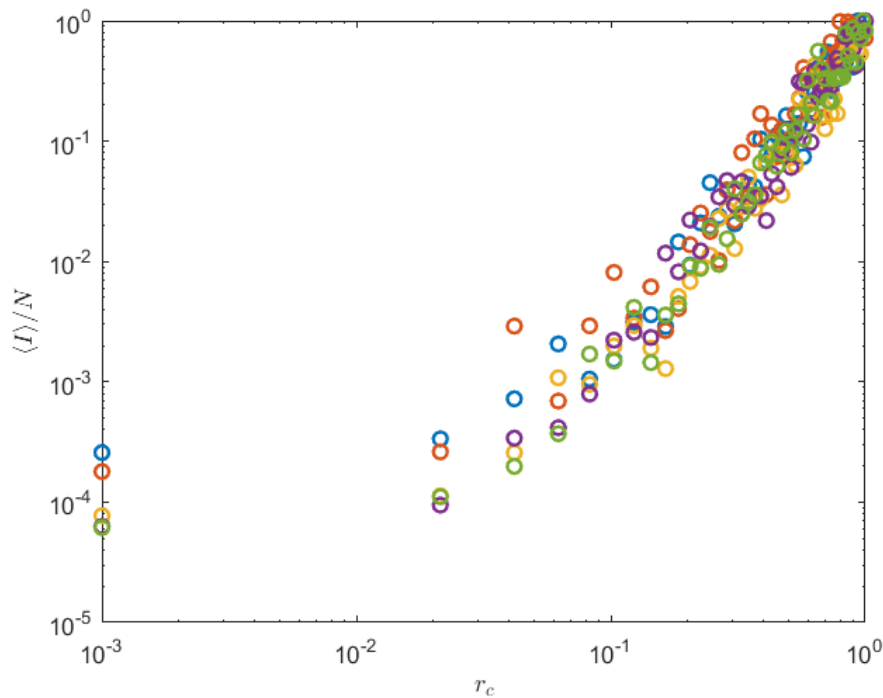


FIGURE 2.18 – Aire de l'intersection de deux disques de rayon r se trouvant à une distance l l'un de l'autre.

FIGURE 2.19 – Intersection moyenne normalisée en fonction de r_c .

Dès lors, nous pouvons définir la probabilité d'intersection de deux hyperarêtes séparées par une distance $l \sim \frac{1}{\sqrt{E}}$ par

$$p_I \approx \frac{A(r_c, 1/\sqrt{E})}{\pi r_0^2}$$

De manière analogue à l'hypergraphe simple (Section 2.2.2), nous pouvons définir l'intersection moyenne entre deux hyperarêtes par

$$\langle I \rangle = p_I N.$$

La figure 2.19 permet d'illustrer le comportement de l'intersection moyenne $\langle I \rangle$, normalisée par le nombre total de noeuds N en fonction du rayon r_c , représentant le seuil de connexion des noeuds aux hyperarêtes. Nous pouvons observer que pour un seuil $r_c \leq \frac{1}{\sqrt{E}}$, il n'y a presque pas d'intersections. Cependant, à mesure que r_c augmente, le nombre d'intersections moyennes augmente proportionnellement à Nr_c^2 .

Cette figure a été obtenue en considérant les paramètres suivants : le nombre de noeuds N prend les valeurs $\{200, 400, 600, 800, 1000\}$, le nombre d'hyperarêtes E a été fixé à 50, le rayon du disque initial r_0 est de 1 et on considère cela pour 100 configurations.

Chapitre 3

Analyse de système dynamique d'hypergraphes

Dans ce chapitre, nous allons analyser le système dynamique linéaire défini sur les hypergraphes. Dans un premier temps, nous allons nous baser essentiellement sur les articles [10], [19] et [14] afin d'étudier la stabilité de systèmes associés à des hypergraphes arbitraires. Dans un deuxième temps, nous allons transposer notre étude aux systèmes associés aux différents modèles construits et étudiés dans le chapitre précédent. De plus, nous étudierons le rôle des paramètres de ces différents modèles sur la convergence des solutions associées aux différents systèmes.

Nous allons reprendre les notations du chapitre précédent afin de définir la structure des hypergraphes. Un hypergraphe $H = (V, E)$ est donc défini par un ensemble de noeuds $V = \{v_i\}$, pour $i = 1, \dots, N$, avec $N = |V|$ le nombre de noeuds et un ensemble d'hyperarêtes $E = \{e_j\}$, où e_j est un sous-ensemble de V de cardinalité arbitraire $|e_j|$, pour $j = 1, \dots, m$.

La matrice d'adjacence peut dès lors être définie comme une matrice carrée de taille $N \times N$. Les éléments de cette matrice indiquent les connexions entre les noeuds de l'hypergraphe. Si les noeuds i et j appartiennent à une hyperarête commune dans l'hypergraphe, alors l'élément A_{ij} de la matrice est positif et dépend de la taille de l'hyperarête qui les contient. Si les noeuds i et j n'appartiennent pas à une hyperarête commune, alors l'élément A_{ij} est nul. La matrice d'adjacence est donc une représentation mathématique de la structure de l'hypergraphe. Formellement, on peut la définir de la manière suivante :

$$A_{ik} = \sum_{\substack{e_j \in E \\ i, k \in e_j \\ i \neq k}} \frac{1}{|e_j| - 1}. \quad (3.1)$$

Dans cette définition, on parcourt toutes les hyperarêtes e_j contenant les noeuds i et j , et on calcule la contribution de chacune de ces hyperarêtes au poids de l'hyperarête reliant i et j . Cette contribution est donnée par le terme $\frac{1}{|e_j| - 1}$, où $|e_j|$ est le nombre de noeuds dans l'hyperarête e_j . Un exemple d'un tel graphe est illustré à la figure 3.1. Nous pouvons également définir l'équivalent pondéré de cette matrice d'adjacence qui permet d'attribuer un poids à chaque hyperarête. C'est une matrice carrée de taille $N \times N$ qui peut être définie de la manière suivante :

$$W_{ik} = \sum_{\substack{e_j \in E \\ i, k \in e_j}} \frac{w_{ik}(|e_j|)}{|e_j| - 1}, \quad (3.2)$$

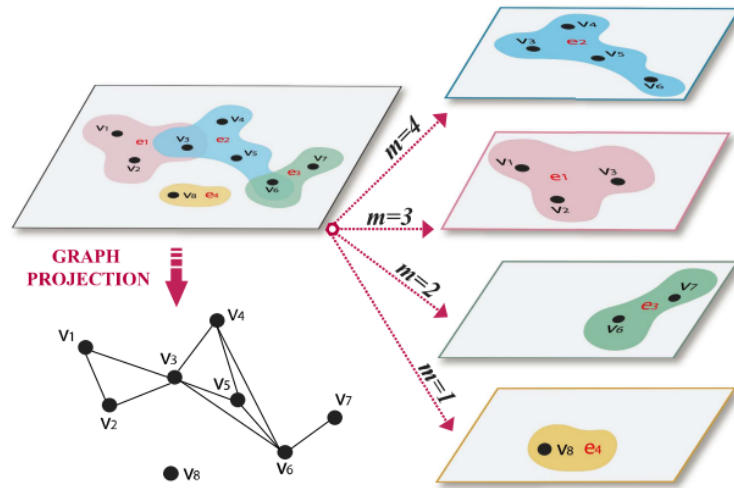


FIGURE 3.1 – Représentation graphique d'un hypergraphe, la décomposition de sa structure et sa projection tirées de [10]

où $|e_j|$ est le nombre de noeuds de l'hyperarête e_j , et $w(|e_j|)$ est une fonction de poids qui prend en compte la taille de l'hyperarête e_j . La pondération des hyperarêtes peut être utilisée pour représenter différentes caractéristiques des relations entre noeuds, telles que la distance, la force de l'interaction, etc. Elle permet également de prendre en compte des variations de poids selon la taille des hyperarêtes.

De plus, nous pouvons définir la matrice Laplacienne de manière analogue à celle des graphes 1.1. C'est une matrice carrée de taille $N \times N$ définie par

$$\mathcal{L} = D - A,$$

où A est la matrice d'adjacence définie ci-dessus, et D est une matrice diagonale dont les éléments diagonaux correspondent aux degrés des noeuds de l'hypergraphe. Formellement, la matrice D est définie par $D = \text{diag}(k_i)$, où le degré k_i est donné par $k_i = \sum_{j=1}^N A_{ij}$.

Afin de faciliter l'analyse du système dynamique linéaire défini sur l'hypergraphe, nous allons le décomposer en m -hypergraphes uniformes. Un m -hypergraphe est un hypergraphe dans lequel toutes les hyperarêtes ont la même taille m . Autrement dit, un m -hypergraphe uniforme est un hypergraphe dans lequel chaque hyperarête est composée de m noeuds. Pour décomposer l'hypergraphe, on définit la matrice d'adjacence du m^e hypergraphe, A^m , dans laquelle deux noeuds sont reliés si et seulement si ils appartiennent à une même hyperarête de taille m . Formellement, cette matrice est définie par

$$A_{ik}^m = \sum_{\substack{e_j \in E \\ |e_j|=m \\ i, k \in e_j \\ i \neq k}} \frac{1}{|e_j| - 1}. \quad (3.3)$$

Nous avons donc que la matrice d'adjacence de l'hypergraphe entier peut être écrite par $A = \sum_m A^m$, comme illustré à la figure 3.1. De par cette définition, on peut remarquer que cette matrice peut prendre des valeurs réelles autres que 0 et 1, contrairement à la matrice d'adjacence pour les graphes. Dans cet exemple, l'hypergraphe de départ est décomposé en quatre hypergraphes uniformes. Le premier est un 4-hypergraphe uniforme puisque son hyperarête e_2 est de taille 4 (elle contient quatre noeuds, v_3, v_4, v_5, v_6). Le deuxième

sous-hypergraphe est un 3-hypergraphe uniforme puisque son hyperarête e_1 est de taille 3 (elle contient 3 noeuds, v_1, v_2, v_3). Les deux autres sous-hypergraphes sont construits de manière analogue, le troisième étant un 2-hypergraphe uniforme et le quatrième un 1-hypergraphe uniforme.

Nous pouvons ensuite appliquer la même procédure de dérivation de la matrice Laplacienne combinatoire [5] pour l'hypergraphe original à chaque m-hypergraphe uniforme, ce qui permettra, par la suite, d'analyser la stabilité des processus dynamiques sur des sous-ensembles de l'hypergraphe. Nous avons donc

$$\mathcal{L}^m = D^m - A^m,$$

où la matrice diagonale est définie par $D^m = \text{diag}(k_i)$, avec le degré k_i qui est donné par $k_i = \sum_{j=1}^N A_{ij}^m$. Finalement, nous pouvons définir l'équivalent pondéré de la matrice d'adjacence de chaque m-hypergraphe uniforme par

$$W_{ik}^m = \sum_{\substack{e_j \in E \\ |e_j|=m \\ i, k \in e_j \\ i \neq k}} \frac{w(|e_j|)}{|e_j| - 1}.$$

3.1 Analyse théorique de la stabilité linéaire

Dans cette section, nous analysons de manière théorique la stabilité linéaire de la dynamique d'un hypergraphe autour d'un point fixe $x_i = x^*$. Pour commencer, un processus dynamique général sur un hypergraphe peut être formulé comme suit :

$$\frac{dx_i}{dt} = f_i(x_i) + \sum_{e_j \in E_i} g_j(x_{\{e_j\}}), \quad (3.4)$$

où x_i est l'état du noeud i , $f_i : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction qui ne dépend que de l'état du noeud i et qui représente la dynamique interne du noeud ; et $g_j : \mathbb{R}^{|e_j|} \rightarrow \mathbb{R}$ est une fonction qui calcule la contribution de l'hyperarête e_j , notée $x_{\{e_j\}}$, par rapport à tous les états des noeuds x_i ; cette fonction représente donc l'interaction externe, exprimée comme la contribution des hyperarêtes.

Nous pouvons à présent exprimer la version linéarisée de l'équation 3.4 autour du point fixe $x_i = x^*$ de la manière suivante :

$$\frac{d\epsilon_i}{dt} = \left. \frac{df_i(x_i)}{dx_i} \right|_{x_i=x_i^*} \epsilon_i + \sum_{e_j \in E_i} \sum_{k \in e_j} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{x=x^*} \epsilon_k, \quad (3.5)$$

où x est un vecteur dont les composantes sont x_i et où $\epsilon_i = x_i - x^*$ est la perturbation autour du point fixe x^* .

Comme nous avons choisi de décomposer l'hypergraphe en m-hypergraphes, nous pouvons reformuler l'équation 3.5 sous forme matricielle grâce aux m-hypergraphes uniformes, qui constituent l'hypergraphe général, de la manière suivante :

$$\frac{d\epsilon}{dt} = \left(F(x^*) + \sum_{m=1}^{\max(|e_j|)} G^m(x^*) \right) \epsilon = M\epsilon, \quad (3.6)$$

où

$$F_{ii}(x^*) = \begin{cases} \left. \frac{df_i(x_i)}{dx_i} \right|_{x_i=x_i^*} & , \text{ si } i = j, \\ 0 & \text{sinon ,} \end{cases} \quad (3.7)$$

$$G_{ik}^m(x^*) = \sum_{\substack{e_j \in E_i \\ |e_j|=m}} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{x=x^*} \quad (3.8)$$

et

$$M(x^*) = F(x^*) + G(x^*).$$

Nous pouvons dès lors énoncer la condition de stabilité suivante : " le point fixe x^* est stable si

$$\Lambda_i(M) < 0, \forall i \in \{1, 2, \dots, N\}, \quad (3.9)$$

où $\Lambda_i(M)$ est la i^e valeur propre du Jacobien M " [10].

Nous pouvons à présent travailler sur la fonction représentant l'interaction externe g_j . Nous allons considérer deux cas ; dans le premier, nous considérons que la fonction est séparable et dans le second, nous considérons des fonctions en interaction avec des dérivées symétriques.

1^{er} cas :

Supposons, pour le premier cas, que la fonction g_j est séparable, nous pouvons alors la réécrire de la façon suivante :

$$g_j(x_{\{e_j\}}) = \frac{1}{|e_j| - 1} \sum_{\substack{k \in e_j \\ k \neq i}} (g(x_i) - g(x_k)). \quad (3.10)$$

Supposons également que la dynamique interne f_i est la même pour tous les noeuds c'est-à-dire que $f_i(x_i) = f_j(x_j), \forall i \neq j$. Dans ce cas, le point fixe est symétrique ; formellement, nous avons que

$$x_i^* = x_j^*, \text{ pour toute paire } i, j.$$

Nous pouvons donc réécrire l'équation 3.8, en remplaçant la fonction g par cette expression, comme

$$G_{ik}^m(x^*) = \sum_{\substack{e_j \in E_i \\ |e_j|=m}} \frac{1}{|e_j| - 1} \sum_{\substack{k \in e_j \\ k \neq i}} \left. \partial_{x_k} (g_j(x_i) - g_j(x_k)) \right|_{x=x^*}$$

Comme nous avons que la somme impose $k \neq i$, nous devons également définir les éléments diagonaux de G^m , c'est-à-dire

$$G_{ii}^m(x^*) = \sum_{\substack{e_j \in E_i \\ |e_j|=m}} \frac{1}{|e_j| - 1} \left. \partial_{x_i} (g_j(x_i) - g_j(x_k)) \right|_{x=x^*}$$

Dès lors, grâce à la définition de matrice d'adjacence A 3.1 en notant $I \in \mathbb{R}^{N \times N}$ la matrice identité et en supposant que toutes les dérivées de g_j sont les mêmes au point fixe, on peut poser

$$\alpha = \left. \frac{df(x)}{dx} \right|_{x=x^*} \quad \text{et} \quad \beta = \left. \frac{dg(x)}{dx} \right|_{x=x^*}. \quad (3.11)$$

Dès lors, nous pouvons réécrire les équations 3.7 et 3.8 de la façon suivante :

$$F = \alpha I \text{ et } G = \beta \mathcal{L}. \quad (3.12)$$

Par conséquent, l'équation 3.5 peut se réécrire comme

$$\frac{d\epsilon_i}{dt} = \left(\alpha + \sum_{e_j \in E_i} \sum_{k \in e_j} \frac{\beta}{|e_j| - 1} \right) \epsilon_i - \sum_{e_j \in E_i} \sum_{k \in e_j} \frac{\beta \epsilon_k}{|e_j| - 1}, \quad (3.13)$$

ou sous forme matricielle, comme

$$\frac{d\epsilon}{dt} = (\alpha I + \beta \mathcal{L})\epsilon. \quad (3.14)$$

Dans ce cas, la condition de stabilité pour le point fixe x^* devient

$$\alpha + \beta \mu_i < 0, \forall i, \quad (3.15)$$

où les μ_i sont les valeurs propres de la matrice \mathcal{L} .

En considérant que la matrice Laplacienne \mathcal{L} est semi-définie positive et que la dérivée $\frac{dg(x)}{dx} \neq 0$ ¹, nous pouvons également énoncer une condition sur la plus grande valeur propre de \mathcal{L} , notée μ_n qui est

$$\frac{1}{\mu_n} > - \left[\frac{dg(x)}{dx} / \frac{df(x)}{dx} \right]_{x_i=x_i^*}. \quad (3.16)$$

2^e cas :

Supposons maintenant, pour le deuxième cas, que la fonction g_j est exprimée de la façon suivante :

$$g_j^*(x_{\{e_j\}}) = \frac{1}{|e_j| - 1} g_j(x_{\{e_j \setminus \{v_i\}\}}), \quad (3.17)$$

où la fonction ne dépend que de x_k , où $k \in e_j$ et $k \neq i$, pour un noeud i donné.

Nous pouvons donc réécrire l'équation 3.8, en remplaçant la fonction g par cette expression, comme

$$G_{ik}^m(x^*) = \sum_{\substack{e_j \in E_j \\ |e_i|=m}} \frac{1}{|e_j| - 1} \partial_{x_k} g_j(x_{\{e_j \setminus \{v_i\}\}}) \Big|_{x=x^*}$$

De manière similaire au cas d'une fonction séparable, nous pouvons dès lors supposer que la fonction f_i est la même pour tous les noeuds ($f_i(x_i) = f_j(x_j)$) et que toutes les dérivées de g_j^* sont les mêmes au point fixe. Cela nous permet alors de réécrire l'équation 3.5 comme

$$\frac{d\epsilon_i}{dt} = \alpha \epsilon_i + \sum_{e_j \in E_i} \sum_{k \in e_j} \frac{\beta \epsilon_k}{|e_j| - 1}, \quad (3.18)$$

où nous avons considéré que nous pouvons écrire F et G sous la même forme que pour les équations 3.12.

Sous forme matricielle, cette équation s'écrit

$$\frac{d\epsilon}{dt} = (\alpha I + \beta A)\epsilon. \quad (3.19)$$

1. Une dérivée nulle ne permettrait pas d'isoler μ dans l'équation 3.15.

De façon analogue au cas précédent, nous pouvons énoncer la condition de stabilité suivante, où nous notons λ_i les valeurs propres de la matrice d'adjacence A :

$$\frac{1}{\lambda_n} < - \left[\frac{dg(x)}{dx} / \frac{df(x)}{dx} \right]_{x_i=x_i^*} < \frac{1}{\lambda_1}. \quad (3.20)$$

Pour conclure cette analyse où nous considérons deux types de fonctions g_j différents, nous pouvons souligner le fait que l'équation 3.9 est également valable pour une dynamique générale sur un hypergraphe arbitraire, contrairement aux autres dérivations précédentes qui sont spécifiques aux deux cas traités.

Le contexte utilisé pour ce début d'analyse de stabilité est général, il peut donc être appliqué à de nombreux processus dynamiques et peut aussi être généralisé à des processus ne sollicitant pas que des interactions par paires.

Deux exemples d'applications étudiés dans l'article [10] sont la contagion sociale et le processus de diffusion. Dans la suite, nous allons appliquer cette analyse dynamique aux modèles construits dans le chapitre précédent (Chapitre 2).

3.2 Application de l'étude de stabilité sur les modèles construits précédemment.

Dans cette section, nous allons appliquer l'analyse dynamique vue à la section 3.1 aux modèles que nous avons étudiés dans le chapitre précédent (Chapitre 2) dans le cadre d'un processus de diffusion.

Dans la continuité de la section 3.1, nous allons nous baser sur la démarche suivie dans l'article [10]. Dès lors, nous allons prendre une fonction d'interaction externe g_j qui est séparable et qui est telle que $g(x_i) = x_i$. Dans ce cas, la dérivée de g vaut 1 et $\beta = \frac{dg(x)}{dx} = \frac{dx}{dx} = 1$. Par conséquent, la dynamique du processus de diffusion peut être exprimée par

$$\frac{dx}{dt} = -D\mathcal{L}x, \quad (3.21)$$

où D est une constante de diffusion² et où le signe moins est issu de la condition de stabilité exprimée à l'équation 3.16. Cette équation traduit la façon dont la quantité présente dans le noeud x évolue vers les autres noeuds et la matrice \mathcal{L} joue un rôle de médiateur en essayant de compenser les différences locales.

Nous allons également décomposer l'hypergraphe en m -hypergraphes uniformes de façon à faciliter l'analyse. Par conséquent, le Laplacien peut aussi être décomposé en deux, avec d'une part les composantes d'ordre supérieur et d'autre part les composantes d'ordre inférieur :

$$\mathcal{L} = \mathcal{L}_O(\theta) + \mathcal{L}_R(\theta),$$

avec

$$\mathcal{L}_O(\theta) = \sum_{m=1}^{\theta} \mathcal{L}^m \quad (3.22)$$

2. L'article [10] ne précisant pas la manière dont cette constante est déterminée, nous allons considérer que $D = 1.1$ dans la suite de notre analyse.

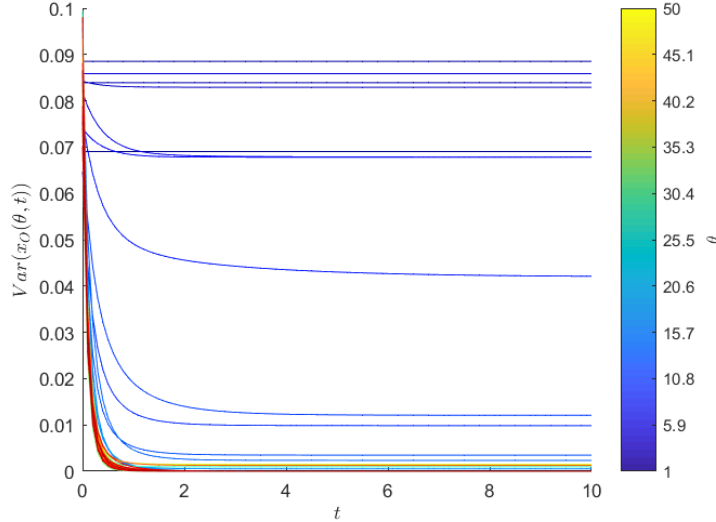


FIGURE 3.2 – Variance de la solution 3.24 en fonction du temps pour θ qui varie de 1 à 50, dans le modèle d'hypergraphes simples.

et

$$\mathcal{L}_R(\theta) = \sum_{m=\theta+1}^{\max(|e_j|)} \mathcal{L}^m, \quad (3.23)$$

où θ permet de distinguer le Laplacien d'ordre inférieur et celui d'ordre supérieur. Cette décomposition va nous permettre d'étudier si la contribution d'ordre inférieur est suffisante pour décrire la dynamique complète du système associé à l'hypergraphe considéré.

Afin de mettre en évidence l'importance de la structure des hypergraphes ainsi que l'importance des différents paramètres qui interviennent dans leur construction, nous allons nous intéresser à la variance de la solution du système d'équations différentielles associé au Laplacien d'ordre inférieur. Ainsi, nous allons considérer le système décrit par l'équation 3.21, avec le Laplacien \mathcal{L}_O défini par l'équation 3.22. La solution associée à ce système est donnée par

$$x_O(\theta, t) = e^{-D\mathcal{L}_O(\theta)t}x(0). \quad (3.24)$$

Dans la suite de notre analyse, nous allons considérer le même intervalle d'intégration pour tous les modèles, à savoir $[0, 10]$ avec un pas d'intégration de 0.1.

3.2.1 Modèle 1 : hypergraphes simples

Pour rappel, ce modèle est construit en considérant la probabilité qu'un noeud v appartienne à une hyperarête e :

$$P(v \in e) = p,$$

où $p \in [0, 1]$.

Dans un premier temps, nous allons étudier le comportement de la variance de la solution 3.24 comme dans l'article [10]. Pour les figures qui suivent, nous avons considéré que θ varie de 1 à 50. La figure 3.2 permet d'illustrer cette quantité en fonction du temps et pour différentes valeurs de θ . Pour la générer, nous avons considéré que le nombre de noeuds N est fixé à 100, le nombre d'hyperarêtes E à 50 et la probabilité de connexion p

à 0.1.

Nous pouvons observer que, au fur et à mesure que le temps augmente et pour une valeur de θ suffisamment grande, la variance diminue et converge vers la solution générale de l'équation 3.21 qui coïncide avec la moyenne. De plus, on peut aussi observer que, si θ n'est pas suffisamment grand, la variance décroît lentement voire pas du tout avec le temps. Grâce à cette première figure, nous pouvons mettre en évidence le fait que la décomposition du Laplacien joue un rôle important sur la convergence de la variance.

Dans un second temps, nous allons nous intéresser à l'influence des différents paramètres du modèle sur la convergence de la solution 3.24.

Premièrement, la figure 3.3 illustre l'évolution de la variance en fonction du temps pour différentes valeurs de la probabilité de connexion $p \in \{0.1, 0.3, 0.5, 0.8\}$. Cette figure a été générée en fixant le nombre de noeuds N à 100 et le nombre d'hyperarêtes E à 50. Bien que la valeur de θ joue un rôle important sur la convergence de la solution 3.24, cette figure permet de mettre en lumière un autre paramètre dont le rôle est important. En effet, on peut observer que si on prend une probabilité de connexion $p = 0.1$, la variance converge plus rapidement que si la probabilité est fixée à $p = 0.3$ ou $p = 0.5$. De plus, nous remarquons que lorsque la probabilité est de 0.8, la variance ne converge pas du tout sur l'intervalle de temps considéré. Plus la probabilité p est grande, plus les possibilités de connexion entre les noeuds et les hyperarêtes augmentent. Le système nécessite donc plus d'informations sur la structure de l'hypergraphe, et donc un θ plus grand, pour pouvoir étudier la dynamique de ce système en conservant un maximum d'informations concernant la convergence de la solution vers la moyenne.

Deuxièmement, nous pouvons étudier le comportement de la variance pour des hypergraphes construits avec un nombre de noeuds ou d'hyperarêtes différent. Les figures 3.4 et 3.5 ont été obtenues en considérant que la probabilité de connexion était fixée à 0.1. Dans la figure 3.4, on fixe le nombre d'hyperarêtes à 50 et on considère que le nombre de noeuds prend des valeurs dans l'ensemble $\{100, 200, 300, 400\}$. Dans la figure 3.5, on fixe le nombre de noeuds à 100 et on fait varier le nombre d'hyperarêtes dans l'ensemble $\{20, 40, 60, 80\}$.

Tout comme la figure 3.2, ces graphiques mettent en évidence l'importance de la décomposition du Laplacien. En effet, on retrouve l'observation selon laquelle la variance converge plus rapidement lorsque θ est grand.

Bien que la variation du nombre d'hyperarêtes ne semble pas avoir un impact significatif sur la convergence de la solution de ce modèle, ce n'est pas le cas pour le nombre de noeuds. En effet, on peut observer que la convergence est plus lente et que celle-ci nécessite un θ plus important lorsque N est grand. Cela peut notamment être dû au fait que plus le nombre de noeuds augmente, plus la structure de l'hypergraphe est grande. Par conséquent, le système associé à un tel hypergraphe nécessite plus d'informations (θ grand) pour pouvoir étudier la dynamique globale de l'hypergraphe.

Ce premier modèle, qui concerne les hypergraphes simples, est donc sensible à la probabilité de connexion permettant de construire de tels hypergraphes. De plus, le nombre de noeuds a un impact sur la convergence de la variance et donc sur la dynamique du système associé aux hypergraphes de ce modèle. L'étude de ce premier modèle selon la dynamique décrite précédemment permet de mettre en évidence l'intérêt de décomposer le Laplacien en deux. En effet, lorsque la solution converge, les θ associés correspondent à différentes échelles de temps. Cependant, on peut constater que même en considérant un

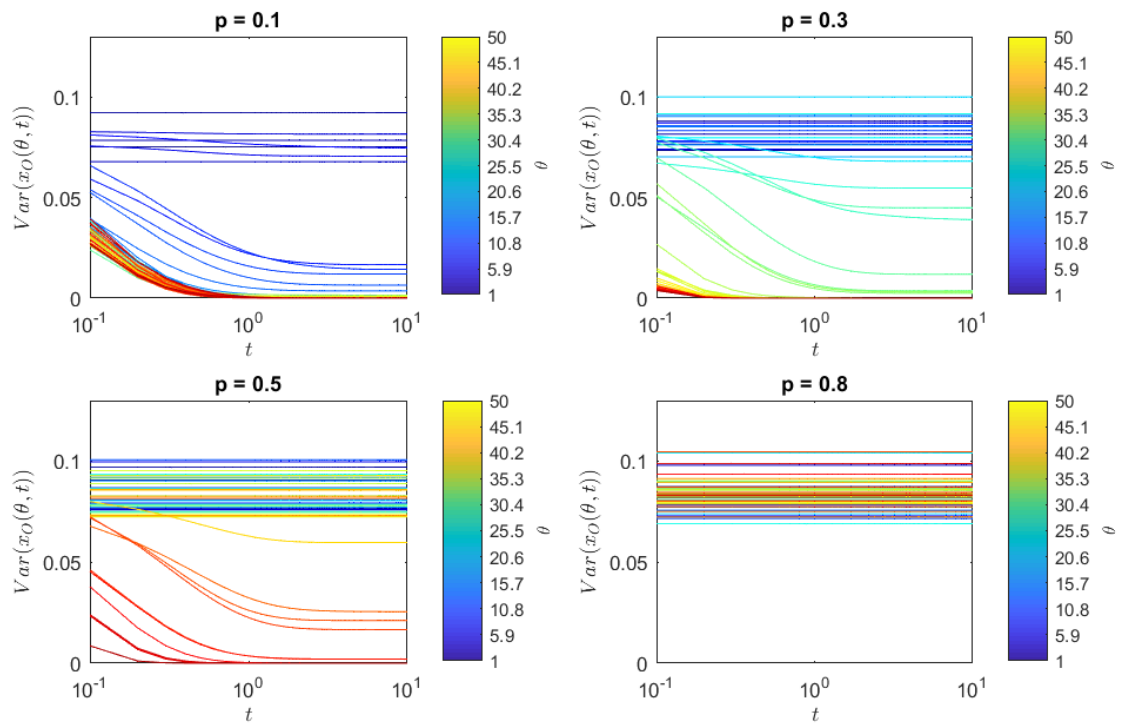


FIGURE 3.3 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de p , dans le modèle d'hypergraphes simples.

Laplacien plus petit, c'est-à-dire une structure plus petite (avec moins d'informations), on peut représenter la dynamique de l'hypergraphe entier et le processus de diffusion associé, pour peu que θ soit suffisamment grand.

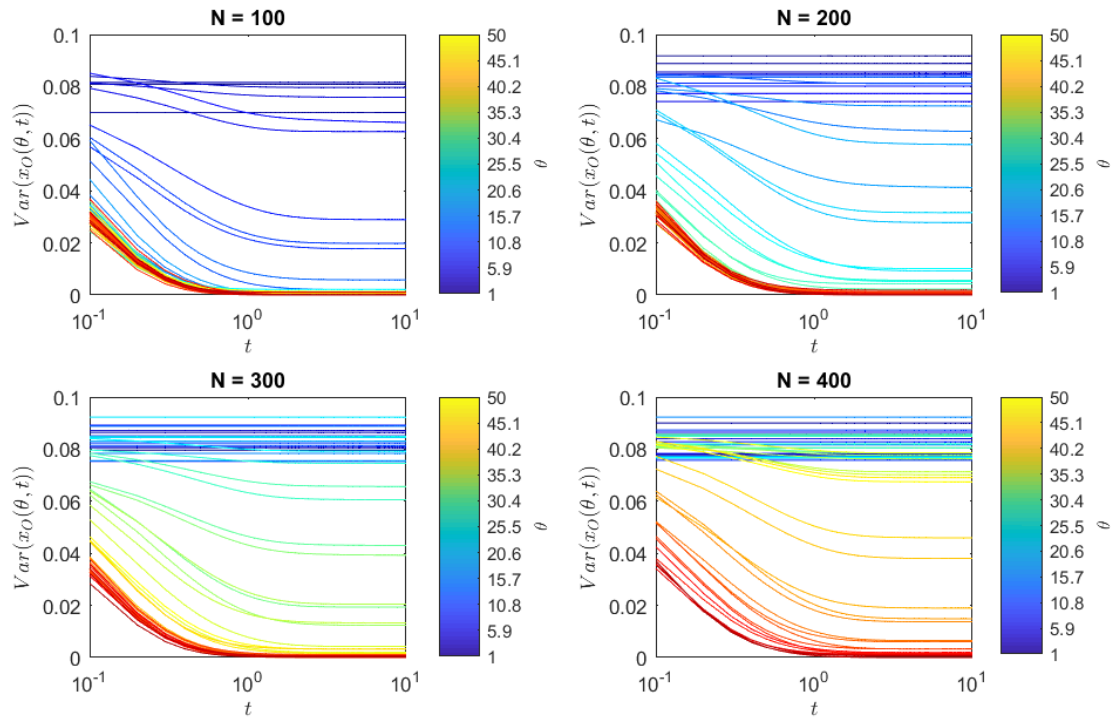


FIGURE 3.4 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de N , dans le modèle d'hypergraphes simples.

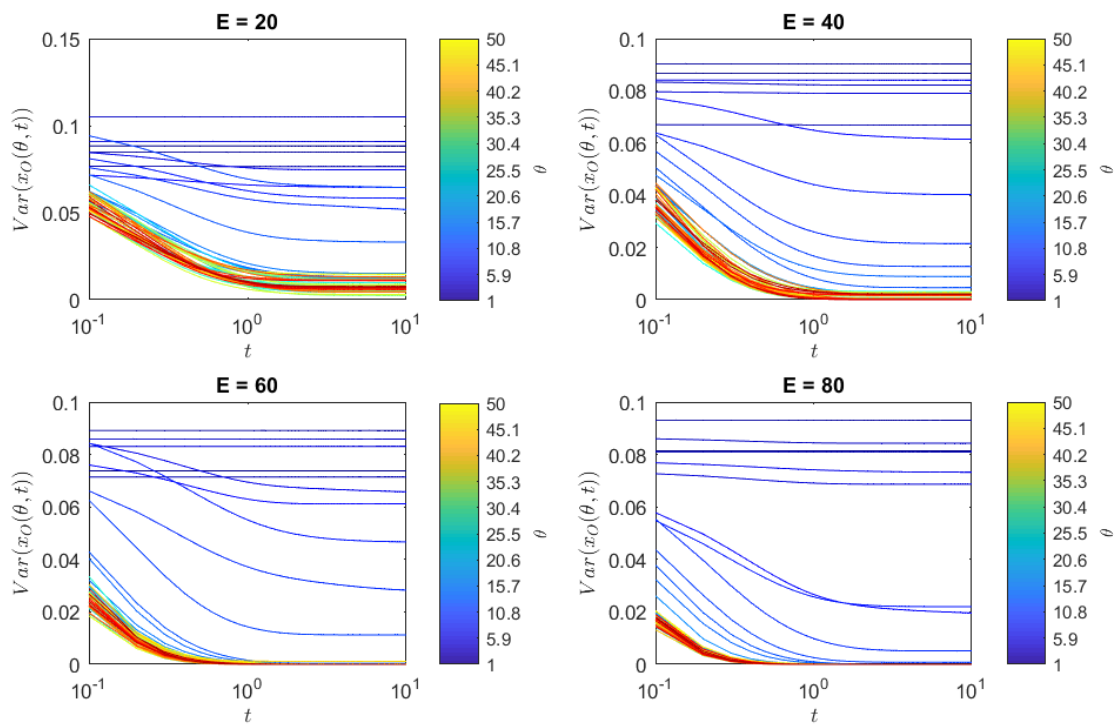


FIGURE 3.5 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de E , dans le modèle d'hypergraphes simples.

3.2.2 Modèle 2 : hypergraphes spatiaux - cas exponentiel

Dans ce modèle, on considère que la probabilité de connexion entre un noeud v et une hyperarête e est donnée par

$$P(v \in e) = pe^{-d(v,e)/r_c},$$

où $p \in [0, 1]$ et r_c est la portée de la décroissance exponentielle.

Dans la section 2.2.4, nous avons vu trois manières différentes de définir la distance $d(v, e)$: la distance moyenne, la distance minimale et la distance à un centroïde. Pour notre analyse, nous allons travailler avec la distance moyenne

$$d(v, e) = \frac{1}{m} \sum_{w \in e} d_E(v, w),$$

où m est la taille de l'hyperarête e et $d_E(v, w)$ est la distance euclidienne entre v et w .

Comme pour le modèle précédent, nous allons, dans un premier temps, étudier le comportement de la solution 3.24 en fonction du temps pour différentes valeurs de θ . Pour les figures qui suivent, nous avons considéré que θ varie de 1 à 50 et que le rayon r_c varie de 0 à 2. La figure 3.6 permet d'illustrer cette quantité pour les mêmes paramètres que le premier modèle, à savoir : le nombre de noeuds N vaut 100, le nombre d'hyperarêtes E vaut 50 et la probabilité de connexion vaut 0.1.

Nous pouvons observer que, au fur et à mesure que le temps augmente, la variance diminue et converge vers la solution générale de l'équation 3.21 qui coïncide avec la moyenne. Si nous comparons nos résultats avec ceux obtenus dans le modèle précédent (figure 3.2), nous pouvons noter que, bien que les résultats soient très proches, la variance nécessite des valeurs de θ plus grande pour ce modèle d'hypergraphes spatiaux. On peut également notifier que, si θ n'est pas suffisamment grand, la variance ne décroît pas avec le temps. Ce modèle permet donc également de montrer l'importance du rôle de la décomposition du Laplacien sur la convergence de la variance.

Cette nécessité de considérer un θ plus grand peut notamment être dû à la manière dont la probabilité de connexion entre un noeud et une hyperarête est définie. En effet, celle-ci dépend de la distance entre ces deux composantes et prend donc en compte la structure (spatiale) de l'hypergraphe, contrairement aux hypergraphes simples où la probabilité est une constante. Le modèle d'hypergraphes spatiaux, dans le cas exponentiel, demande donc un θ plus grand afin de prendre en compte plus d'informations sur la structure de l'hypergraphe, nécessaire à l'étude dynamique de son système associé.

Dans un second temps, nous allons nous intéresser à l'influence des différents paramètres du modèle sur la convergence de la solution 3.24.

Premièrement, la figure 3.7 illustre l'évolution de la variance en fonction du temps pour différentes valeurs de la probabilité de connexion $p \in \{0.1, 0.3, 0.5, 0.8\}$. Cette figure a été générée en fixant le nombre de noeuds N à 100 et le nombre d'hyperarêtes E à 50. Bien que le rôle de θ reste important, cette figure nous permet de constater que la probabilité de connexion ne joue pas un rôle aussi important que dans le modèle des hypergraphes simples. De manière cohérente avec la figure 3.6, on peut observer que lorsque $p = 0.1$ il y a un plus faible taux de convergence que pour les hypergraphes simples. Pour les valeurs supérieures de p (0.3, 0.5, 0.8), la variance nécessite un θ plus grand pour converger. De plus, pour $p = 0.8$, on constate que la variance ne converge pas vers la moyenne sur l'intervalle de temps d'intégration considéré ($[0, 10]$ avec un pas d'intégration de 0.1.). Cette

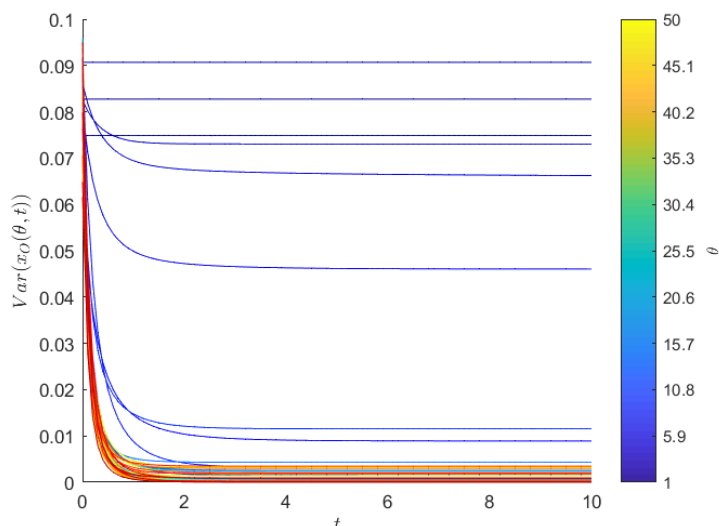


FIGURE 3.6 – Variance de la solution 3.24 en fonction du temps pour θ qui varie de 1 à 50, dans le modèle d'hypergraphes spatiaux (cas exponentiel).

observation est due au fait que, lorsque la probabilité p augmente, le produit entre celle-ci et le terme exponentiel ($e^{-d(v,e)}$) augmente. La probabilité de connexion entre un noeud et une hyperarête étant alors plus grande, on retrouve la même constatation que pour le modèle d'hypergraphes simples. C'est-à-dire que le système nécessite plus d'informations sur la structure de l'hypergraphe afin de pouvoir étudier sa dynamique.

Deuxièmement, nous pouvons étudier le comportement de la variance pour des hypergraphes construits avec un nombre de noeuds ou d'hyperarêtes différents. Les figures 3.4 et 3.5 ont été obtenues en considérant que la probabilité de connexion était fixée à 0.1. Dans la figure 3.4, on fixe le nombre d'hyperarêtes à 50 et on considère que le nombre de noeuds prend des valeurs dans l'ensemble $\{100, 200, 300, 400\}$. Dans la figure 3.5, on fixe le nombre de noeuds à 100 et on fait varier le nombre d'hyperarêtes dans l'ensemble $\{20, 40, 60, 80\}$.

De manière similaire au modèle d'hypergraphes simples, nous pouvons observer que lorsque N est grand, la convergence de la variance nécessite un θ plus important. Cela peut également être dû au fait que plus le nombre de noeuds augmentent, plus la structure de l'hypergraphe est grande. Par conséquent, le système associé à un tel hypergraphe nécessite plus d'informations (θ grand) pour pouvoir étudier la dynamique globale de l'hypergraphe.

En revanche, on peut observer que le nombre d'hyperarêtes joue un rôle plus important dans ce modèle. En effet, lorsque ce nombre E est grand, la variance converge plus rapidement pour des θ plus petits. Par exemple, lorsque $E = 20$ la variance demande un θ plus grand pour converger que lorsque $E = 80$. La définition de probabilité de connexion permet d'expliquer ce phénomène. Comme celle-ci prend en compte la distance entre un noeud et une hyperarête, s'il existe un grand nombre d'hyperarêtes, alors on peut supposer que la structure de l'hypergraphe peut être représentée par une plus petite structure (avec moins d'informations). Il est important de noter que cela est d'autant plus vrai si le quotient E/N est proche de 1. En effet, plus le nombre d'hyperarêtes est proche du nombre de noeuds, plus on peut étudier la dynamique avec un Laplacien d'ordre inférieur défini avec un petit θ .

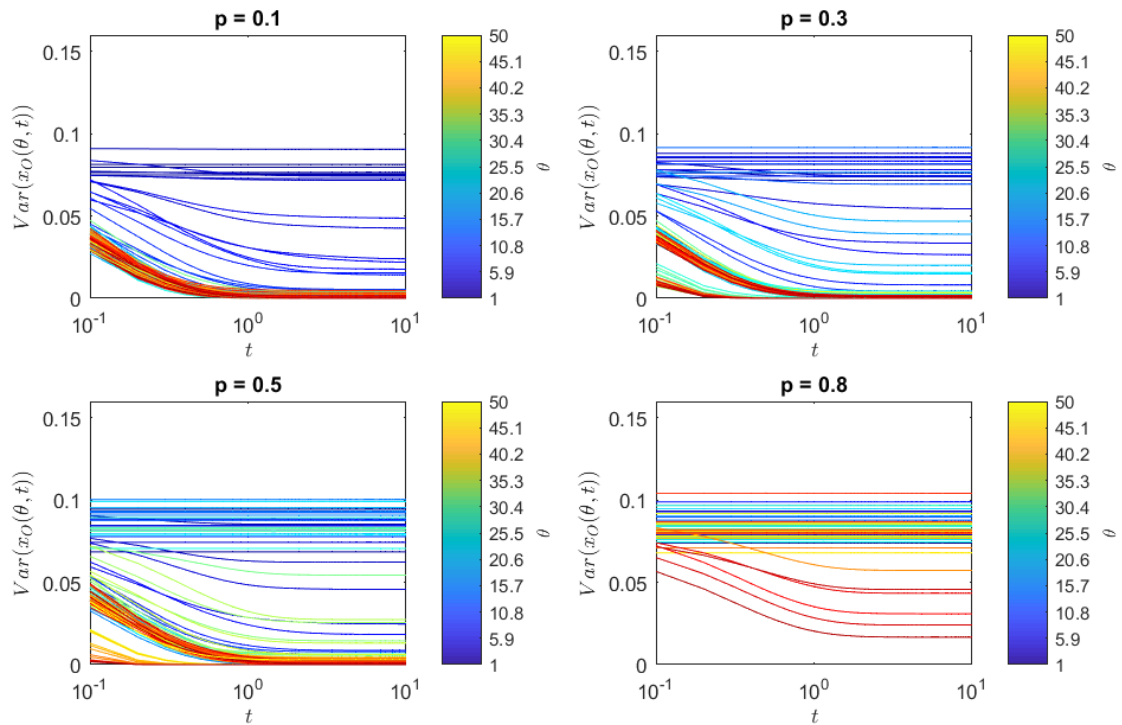


FIGURE 3.7 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de p , dans le modèle d’hypergraphes spatiaux (cas exponentiel).

Troisièmement, nous pouvons étudier la convergence de la variance pour des hypergraphes construits sur un rayon initial r_0 différent dans le but de voir si le rapport r_0/r_c joue un rôle important sur la dynamique. Dans le chapitre 2, nous avons considéré un rayon initial $r_0 = 1$, cependant il peut être intéressant de faire varier ce rayon afin d’étudier son impact sur la dynamique du système associé à l’hypergraphe. La figure 3.10 a été obtenue en considérant que la probabilité de connexion était fixée à 0.1, le nombre de noeuds à 100 et le nombre d’hyperarêtes à 50. Ces graphiques, obtenus pour un rayon r_0 prenant des valeurs dans l’ensemble $\{0.5, 1, 1.5, 2\}$ illustrent que, plus le rayon initial est proche de l’étendue de l’exponentielle r_c , c’est-à-dire que le rapport r_0/r_c est proche de 1, plus la variance converge rapidement pour un θ moins grand. Cela rentre en adéquation avec le fait que si le rayon initial est proche du rayon final (étendue de l’exponentielle), plus l’hypergraphe peut être représenté par une plus petite structure et donc l’étude de la dynamique peut se faire avec un Laplacien d’ordre inférieur avec un θ plus faible.

Ce second modèle, qui concerne les hypergraphes spatiaux dans le cas exponentiel, est donc sensible à la probabilité de connexion permettant de construire de tels hypergraphes ainsi qu’à la proportion de noeuds et d’hyperarêtes qui les composent. Tout comme le premier modèle, cette étude de la dynamique sur les systèmes des hypergraphes spatiaux permet de mettre en évidence l’intérêt de décomposer le Laplacien. En effet, en fonction des différents paramètres qui définissent ces hypergraphes, il est possible d’étudier leur dynamique en considérant des plus petites structures (avec moins d’informations) tout en conservant la dynamique globale du processus de diffusion associé à l’hypergraphe entier.

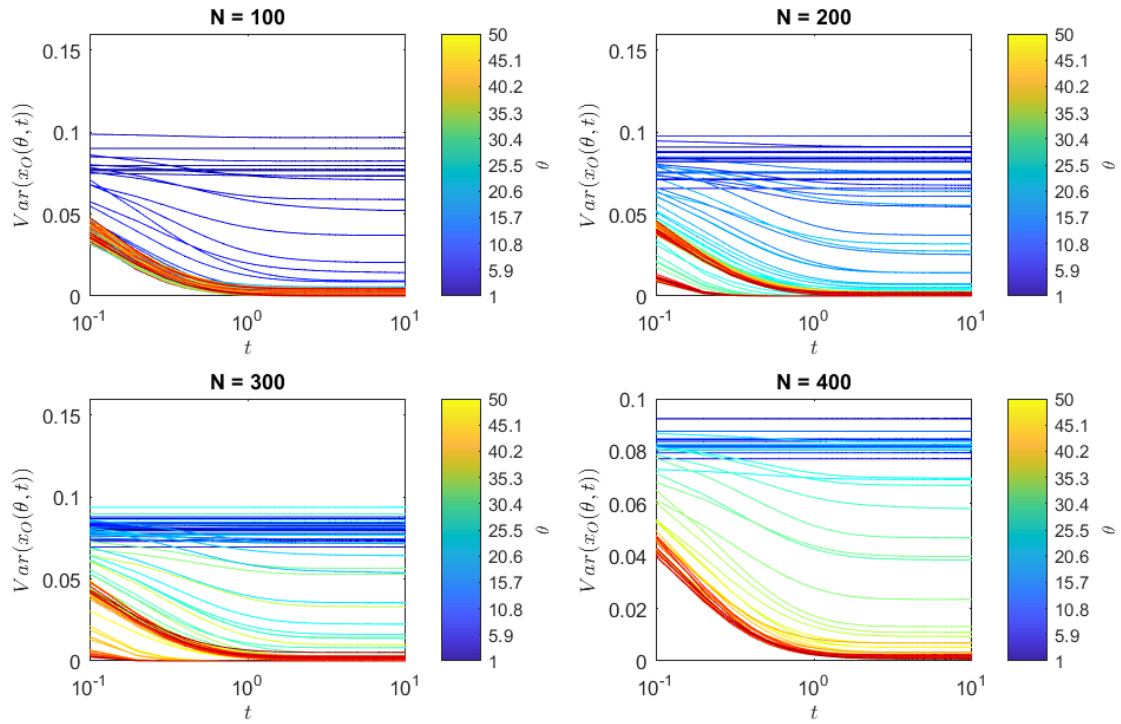


FIGURE 3.8 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de N , dans le modèle d'hypergraphes spatiaux (cas exponentiel).

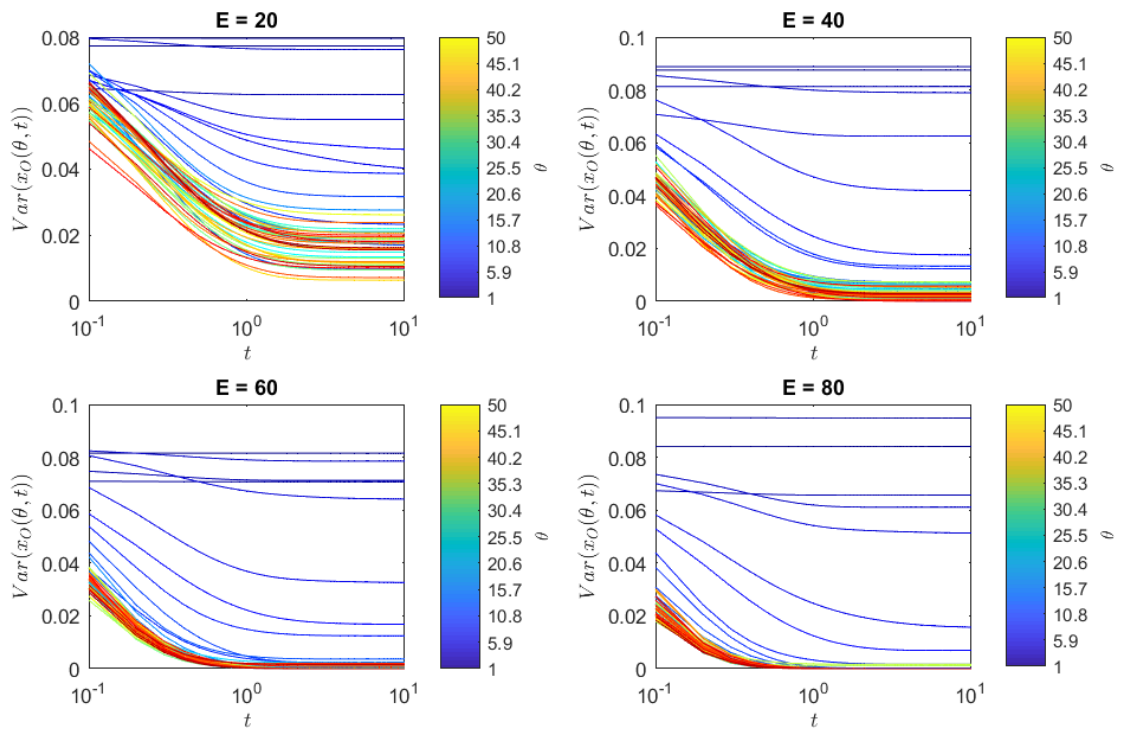


FIGURE 3.9 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de E , dans le modèle d'hypergraphes spatiaux (cas exponentiel).

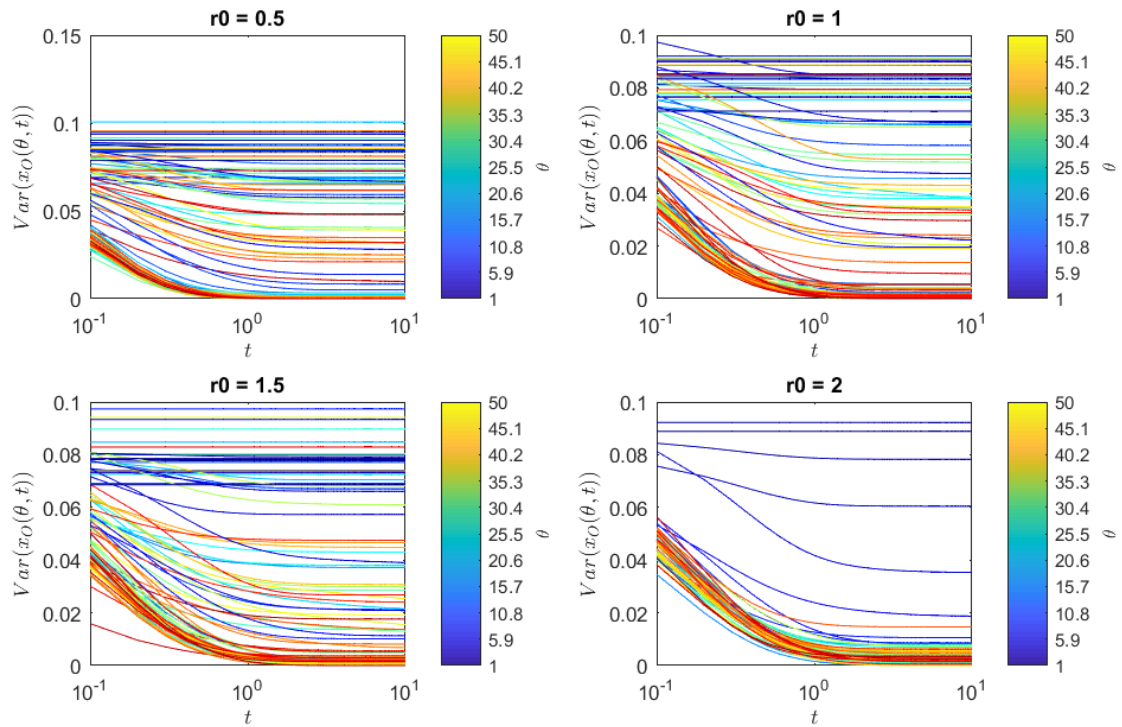


FIGURE 3.10 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de r_0 , dans le modèle d'hypergraphes spatiaux (cas exponentiel).

3.2.3 Modèle 3 : hypergraphes spatiaux - cas géométrique

Pour ce modèle, nous avons utilisé la fonction de Heaviside $\Theta(x)$ [7]

$$\Theta(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0. \end{cases}$$

afin de définir la probabilité de connexion entre un noeud v et une hyperarête e , qui est donnée par

$$P(v \in e) = \Theta(r_c - d(v, e)), \quad (3.25)$$

où $d(v, e)$ est la distance maximale défini par

$$d(v, e) = \max_{w \in e} d(v, w).$$

De manière similaire aux deux modèles étudiés précédemment, nous allons, dans un premier temps, étudier le comportement de la solution 3.24 en fonction du temps pour différentes valeurs de θ . Pour les figures qui suivent, nous avons considéré que θ varie de 1 à 50 et que le rayon r_c varie de 0 à 2. La figure 3.11 permet d'illustrer cette quantité pour les mêmes paramètres que les deux premiers modèles, à savoir : le nombre de noeuds N vaut 100, le nombre d'hyperarêtes E vaut 50 et la probabilité de connexion vaut 0.1. En comparant cette figure avec les figures 3.2 et 3.6 des modèles précédents, nous pouvons observer que la convergence de la variance nécessite un θ plus grand. Cela est dû au fait que la probabilité de connexion entre un noeud et une hyperarête se fait en utilisant la fonction Heaviside 3.25. De ce fait, la probabilité ne peut prendre des valeurs que dans l'ensemble $\{0, 1\}$. On retrouve alors la même constatation que pour l'étude du premier modèle lorsque p varie, à savoir que le nombre de possibilités de connexion augmente. Dès

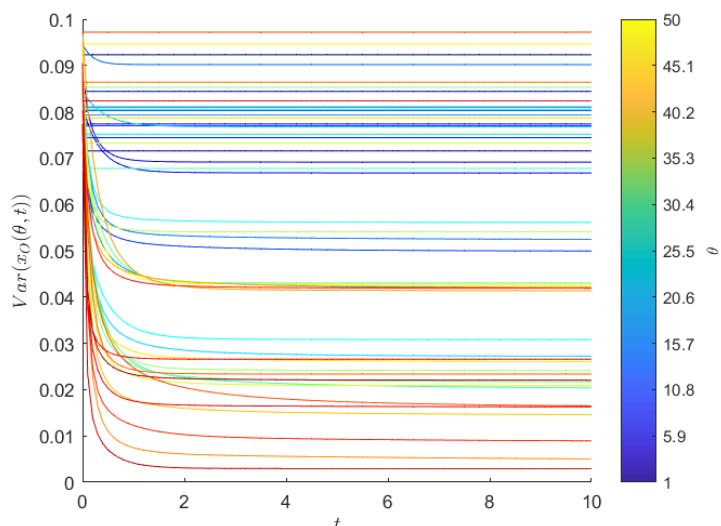


FIGURE 3.11 – Variance de la solution 3.24 en fonction du temps pour θ qui varie de 1 à 50, dans le modèle d’hypergraphes spatiaux (cas géométrique).

lors, le système nécessite plus d’informations sur la structure de l’hypergraphe, et donc un θ plus grand, pour pouvoir étudier sa dynamique.

Bien que la définition utilisée pour la probabilité de connexion aie un impact considérable sur la convergence de la variance vers la moyenne, nous allons tout de même étudier l’impact des autres paramètres qui entre en jeu dans la construction de ce modèle d’hypergraphes spatiaux, dans le cas géométrique.

Dans un second temps, nous allons donc nous intéresser à l’influence des différents paramètres du modèle sur la convergence de la solution 3.24.

Si nous comparons les différentes figures avec celles obtenues pour le modèle d’hypergraphes spatiaux dans le cas exponentiel, nous pouvons observer les mêmes comportements concernant la variation des paramètres. Cependant, comme énoncé précédemment, la variance nécessite, dans tous les cas, un θ plus grand pour converger vers la moyenne. De plus, nous pouvons remarquer que le temps de convergence est également plus grand et que la variance ne converge pas nécessairement sur l’intervalle de temps considéré.

Premièrement, la figure 3.12 illustre l’évolution de la variance en fonction du temps pour un nombre de noeuds différents. Cette figure a été générée en fixant le nombre d’hyperarêtes E à 50, le rayon initial r_0 à 1 et pour un nombre de noeuds N prenant des valeurs dans l’ensemble $\{100, 200, 300, 400\}$.

De manière similaire au modèle d’hypergraphes spatiaux, dans le cas exponentiel, nous pouvons observer que lorsque N est grand, la convergence de la variance nécessite un θ plus important. Cela signifie que le système associé à un tel hypergraphe nécessite plus d’informations (θ grand) pour pouvoir étudier la dynamique globale de l’hypergraphe.

Deuxièmement, nous pouvons étudier le comportement de la variance pour des hypergraphes construits avec un nombre d’hyperarêtes différents. La figure 3.13 a été obtenue en considérant que le nombre de noeuds N vaut 100, le rayon initial r_0 vaut 1 et le nombre d’hyperarêtes E prend des valeurs dans l’ensemble $\{20, 40, 60, 80\}$.

Tout comme pour le modèle précédent, on peut observer que le nombre d’hyperarêtes

joue un rôle plus important au vu de la convergence de la variance plus rapide lorsque le nombre d'hyperarêtes E est grand. Là encore, la définition de probabilité de connexion joue un rôle important sur la convergence puisqu'elle prend en compte la distance entre un noeud et une hyperarête. Dès lors, s'il existe un grand nombre d'hyperarêtes, alors on peut supposer que la structure de l'hypergraphe peut être représentée par une plus petite structure (avec moins d'informations).

Troisièmement, nous pouvons étudier la convergence de la variance pour des hypergraphes construits sur un rayon initial r_0 différent dans le but de voir si le rapport r_0/r_c joue un rôle important sur la dynamique. La figure 3.14 a été obtenue en considérant que le nombre de noeuds était fixé à 100, le nombre d'hyperarêtes à 50 et un rayon initial r_0 dont les valeurs sont prises dans l'ensemble $\{0.5, 1, 1.5, 2\}$. Contrairement au modèle précédent, nous pouvons observer que la variation du rayon initial n'a pas un impact significatif sur la convergence de la variance, que le rapport r_0/r_c soit proche de 1 ou non. Cela est notamment dû à la manière dont la probabilité de connexion est définie. En effet, le rayon initial va avoir un impact sur la distance $d(v, e)$ entre un noeud et une hyperarête. Cependant, la fonction de Heaviside ne pouvant prendre comme valeurs que 0 ou 1, le choix du rayon initial a un impact réduit comparé au modèle précédent.

Ce dernier modèle, qui concerne les hypergraphes spatiaux dans le cas géométrique, nécessite donc θ plus grand, c'est-à-dire un système contenant plus d'informations sur la structure de l'hypergraphe, en raison de l'utilisation de la fonction de Heaviside pour sa construction. Tout comme les deux autres modèles, cette étude de la dynamique sur les systèmes des hypergraphes spatiaux permet de mettre en évidence l'intérêt de décomposer le Laplacien. En effet, en fonction des différents paramètres qui définissent ces hypergraphes, il est possible d'étudier leur dynamique en considérant des plus petites structures (avec moins d'informations) tout en conservant la dynamique globale du processus de diffusion associé à l'hypergraphe entier. Cependant, il est important de noter que, comme ce modèle nécessite un θ plus grand, l'intérêt de décomposer le Laplacien en deux est moins significatif que pour les deux autres modèles.

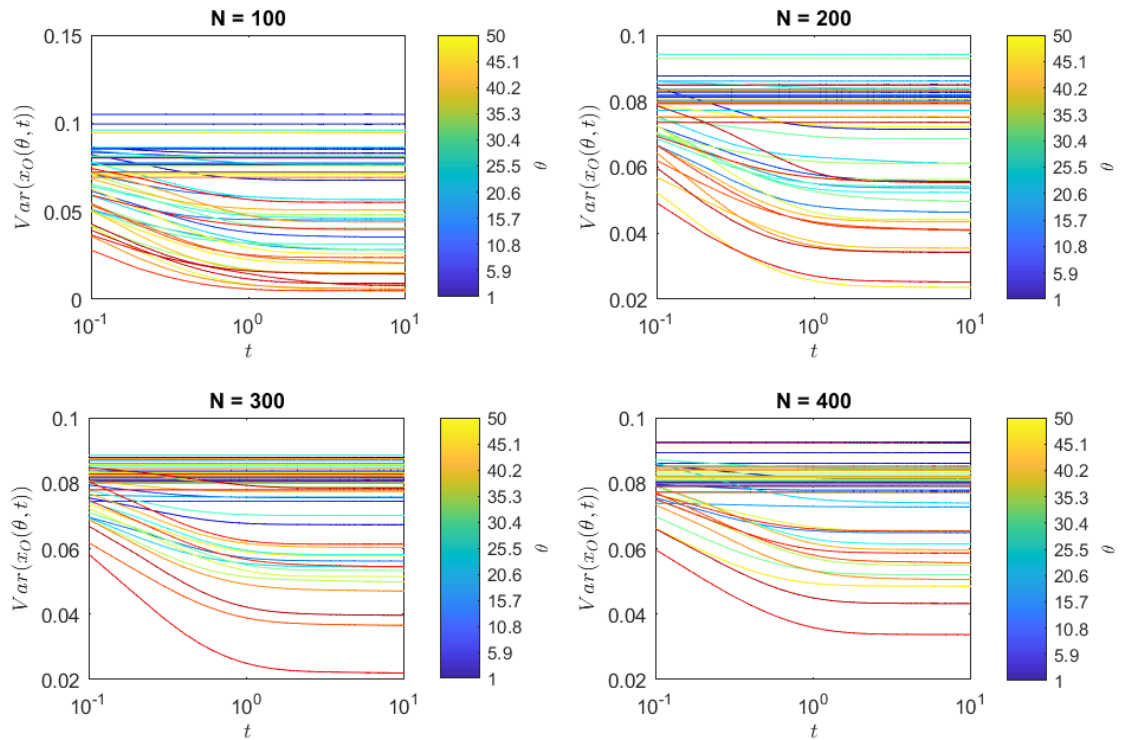


FIGURE 3.12 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de N , dans le modèle d'hypergraphes spatiaux (cas géométrique).

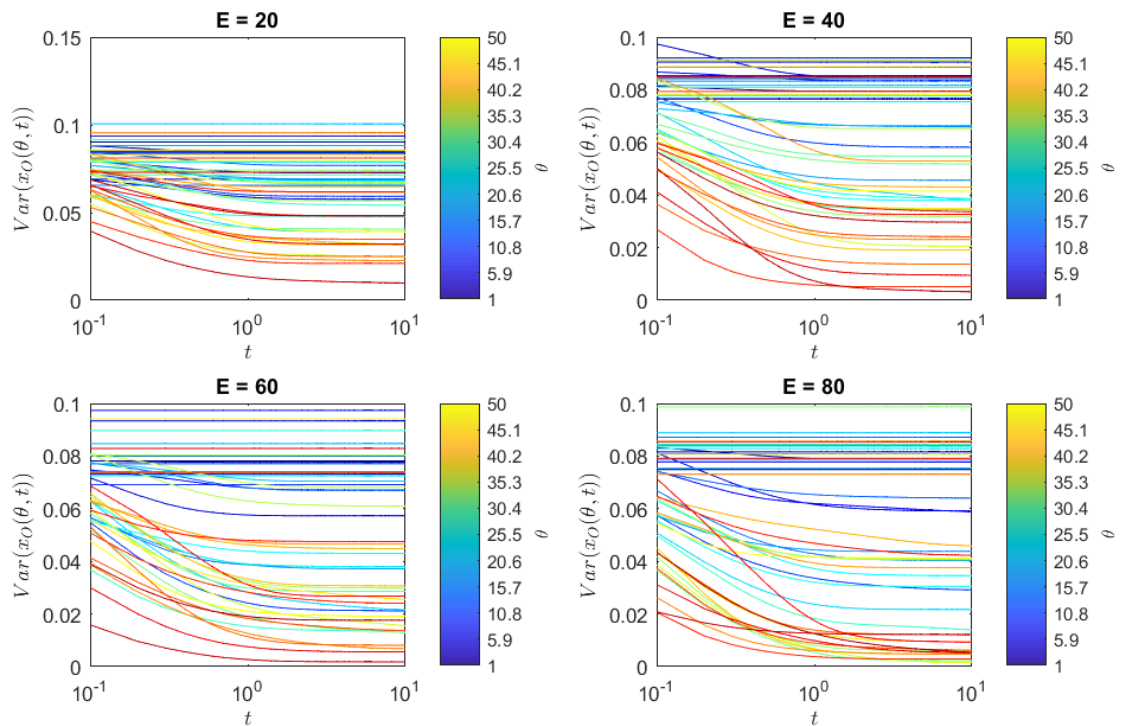


FIGURE 3.13 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de E , dans le modèle d'hypergraphes spatiaux (cas géométrique).

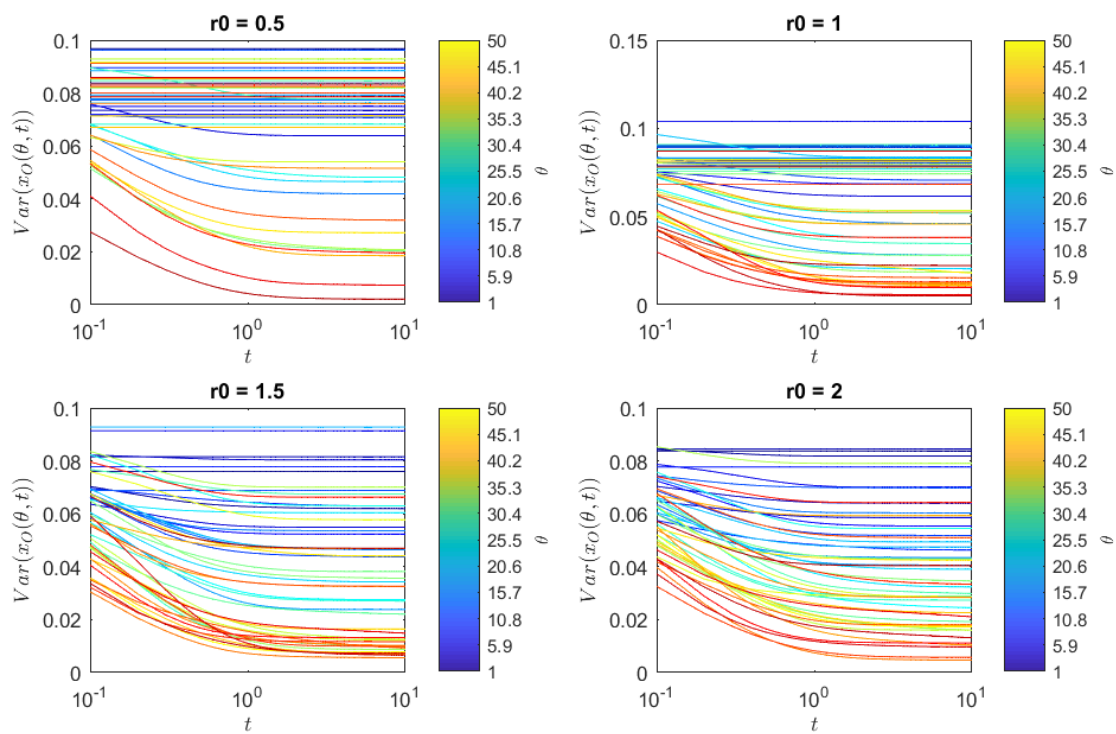


FIGURE 3.14 – Variance de la solution 3.24 en fonction du temps pour différentes valeurs de r_0 , dans le modèle d'hypergraphes spatiaux (cas géométrique).

Conclusion et perspectives

Dans ce mémoire, nous avons étendu la notion de graphes à celles d'hypergraphes en introduisant le concept d'hyperarêtes, permettant ainsi de représenter des interactions entre des groupes de noeuds. Nous nous sommes ensuite basés sur l'article « *A class of models for random hypergraphes* » de M. Barthelemy [3] afin d'étudier des modèles de construction d'hypergraphes établis sur la probabilité de connexion entre un noeud et une hyperarête .

Dans le cas où on considère que cette probabilité est une constante choisie arbitrairement dans l'intervalle $[0, 1]$, nous obtenons des hypergraphes aléatoires que l'on peut considérer comme étant des extensions des graphes d'Erdős-Rényi. En étendant les propriétés associées des graphes, nous avons pu montrer que la distribution des degrés des noeuds et des tailles des hyperarêtes sont des lois binomiales de paramètres $Bin(E, p)$ et $Bin(N, p)$ respectivement. Nous nous sommes ensuite intéressés aux intersections entre les hyperarêtes. Les simulations numériques ont permis de montrer que le nombre moyen d'intersections entre les hyperarêtes peut être obtenu en multipliant le nombre de noeuds par la probabilité de connexion. De plus, elles ont permis de montrer que la distribution des intersections est une loi binomiale de paramètres $Bin(N, p^2)$.

Nous avons ensuite pu étendre la notion d'attachement préférentiel aux hypergraphes en définissant la probabilité de connexion entre un noeud et une hyperarête; pour ce faire nous nous sommes basés sur la taille des hyperarêtes composant l'hypergraphe. Pour ce modèle, nous avons simplement illustré le fait que la distribution de la taille des hyperarêtes est une distribution (multinomiale) de Dirichlet. Nous avons choisi de nous concentrer sur les trois autres modèles présentés dans l'article de référence [3]. Dès lors, des perspectives d'action existent concernant ce modèle. Plus particulièrement, comme celui-ci est construit sur base de la taille des hyperarêtes, une étude plus poussée peut être faite concernant la structure des hypergraphes. En effet, la distribution de la taille des hyperarêtes étant une distribution multinomiale de Dirichlet, celle-ci dépend fortement de la structure initiale de l'hypergraphe.

Les deux derniers modèles ont été construits en rajoutant la dimension spatiale aux hypergraphes, c'est-à-dire en attribuant des coordonnées aux noeuds. Pour le premier, le cas exponentiel, nous avons défini la probabilité de connexion en considérant trois définitions de distance entre un noeud et une hyperarête : la distance moyenne, la distance minimale et la distance à un centroïde. Nous avons alors pu étudier la convergence de l'étendue spatiale des hyperarêtes pour ces trois définitions. Nos simulations numériques nous ont permis de conclure que la distance minimale permettait une convergence plus rapide. Plus généralement, l'étendue spatiale est très sensible au choix de la définition de distance, et donc au choix de la probabilité de connexion. Pour le second, le cas géométrique, nous avons étendu la notion de graphes géométriques aléatoires introduite par E.N. Gilbert aux hypergraphes. Pour cela, nous avons défini la probabilité de connexion

entre un noeud et une hyperarête en nous basant sur la distance maximale. En étudiant le degré moyen des noeuds et la taille moyenne des hyperarêtes, nous avons pu montrer une égalité entre ces deux quantités normalisées. N'ayant pu réaliser l'étude concernant la composante géante, une perspective d'action pour ce modèle serait de développer ce point afin de prouver numériquement la présence d'une valeur critique pour le rayon des hyperarêtes en $r_c^* \sim r_0/\sqrt{E}$.

Après avoir étudié différents modèles de construction d'hypergraphes, nous avons eu l'occasion de nous pencher en partie sur l'étude dynamique de leurs systèmes associés. Nous avons commencé par l'étude théorique de la stabilité linéaire de systèmes associés à des hypergraphes arbitraires en nous basant sur l'article « *Phase transitions and stability of dynamical processes on hypergraphs* » de G. Ferraz de Arruda, M. Tizzani et Y. Moreno [10]. Cette étude nous a permis d'extraire deux conditions de stabilité sur la solution des systèmes. La première concerne les valeurs propres de la matrice Laplacienne et la deuxième concerne celles de la matrice d'adjacence. Nous avons ensuite pu appliquer cette étude de stabilité à un processus de diffusion sur les systèmes associés aux hypergraphes construits dans le deuxième chapitre. Pour cela, nous avons scindé le Laplacien en deux et nous avons étudié la convergence de la solution associée au Laplacien d'ordre inférieur en faisant varier les différents paramètres des modèles respectifs. De manière générale, les trois modèles ont permis d'appuyer l'intérêt de décomposer le Laplacien. De fait, ils ont montré qu'il est possible d'étudier la dynamique des hypergraphes associés en considérant des structures plus petites, c'est-à-dire avec moins d'informations. Concernant les modèles en particulier, celui concernant les hypergraphes simples a montré une grande sensibilité au choix de la probabilité de connexion ainsi qu'au nombre de noeuds. Le modèle d'hypergraphes spatiaux dans le cas exponentiel a également montré une grande sensibilité au choix de ces deux paramètres. Le dernier modèle, qui concerne les hypergraphes spatiaux dans le cas géométrique a, quant à lui, montré une plus grande sensibilité que les deux autres modèles au choix du nombre de noeuds et d'hyperarêtes. Pour les deux modèles d'hypergraphes spatiaux, nous avons également étudié l'impact du rapport r_0/r_c sur la convergence de la solution associée. Bien que nous ayons pu constater un certain impact sur la convergence, une perspective d'action pourrait être de pousser cette étude en décryptant plus précisément l'impact du choix de ces deux rayons ainsi que de leur rapport. De plus, il est possible de réaliser cette analyse sur des hypergraphes de plus grandes envergures, c'est-à-dire avec un nombre de noeuds et d'hyperarêtes plus conséquent, moyennant des ressources informatiques plus importantes.

Nous pouvons à présent proposer de nouvelles pistes de recherche comme perspective de ce mémoire. Nous avons eu l'occasion de travailler sur certains modèles d'hypergraphes étant des extensions de modèles de graphes, tels que le modèle d'Erdős-Rényi ou le modèle géométrique. Cependant, il existe d'autres modèles de graphes qui pourraient être étendus aux hypergraphes tels que les graphes de Gabriel ou les graphes " β -squelettes". Il pourrait donc être intéressant d'étudier différentes mesures associées à ces modèles ou encore d'étudier la dynamique de leurs systèmes associés en fonction des différents paramètres qui les composent. De plus, une étude dynamique de leurs systèmes associés pourrait également faire l'objet de perspectives d'actions.

Bibliographie

- [1] S. AKSOY AND AL., *Hypernetwork science via high-order hypergraph walks*, EPJ data science, 9 (2020).
- [2] A.-L. BARABASI AND R. ALBERT, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–512.
- [3] M. BARTHELEMY, *A class of models for random hypergraphs*, Physical Review E, 106 (2022), p. 064310.
- [4] F. BATISTON, G. CENCETTI, AND ALS., *Networks beyond pairwise interactions : Structure and dynamics*, Physics Reports, 874 (2020), pp. 1–92.
- [5] T. CARLETTI, D. FANELLI, AND S. NICOLETTI, *Dynamical systems on hypergraphs*, Journal of Physics : Complexity, 1 (2020).
- [6] P. CHODROW AND A. MELLOR, *Annotated hypergraphs : models and applications*, Applied Network Science, 5 (2020).
- [7] T. CLUZEAU, *Mathématiques pour l'ingénieur*, École Nationale Supérieure d'Ingénieurs de Limoges, 2015.
- [8] V. DELAITRE, *Hypergraphes et matroïdes*, Ecole Normale Supérieure de Lyon, 2009.
- [9] P. ERDÖS AND A. RÉNYI, *On the evolution of random graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5 (1960), pp. 17–61.
- [10] G. FERRAZ DE ARRUDA, M. TIZZANI, AND Y. MORENO, *Phase transitions and stability of dynamical processes on hypergraphs*, Communications Physics, 24 (2021).
- [11] R. GARCIA-PELAYO, *Distribution of distance in the spheroid*, Journal of Pphysics A : Mathematical and general, 38 (2005), p. 3475–3482.
- [12] E. GILBERT, *Random graphs*, The Annals of Mathematical Statistics, 30 (1959), pp. 1141–1144.
- [13] C. GRASLAND, *Analyse spatiale et modélisation des phénomènes géographiques. Chapitre 2 : L'analyse des graphes et des réseaux*, Université Paris VII, 2000.
- [14] T.-H. HUBERT CHAN AND Z. LIANG, *Generalizing the hypergraph laplacian via a diffusion process with mediators*, Theoretical Computer Science, 806 (2020), pp. 416–428.
- [15] T. JOHANSSON, *The giant component of the random bipartite graph*, Master's thesis, Department of Mathematical Sciences - Chalmers University of Technology, 2012.
- [16] S. JOULI AND S. TABBONE, *Indexation de graphes à partir d'une structure d'hypergraphe*, 2011.
- [17] J. KUNEGIS, *Théorie des graphes*, Université de Namur, 2017.
- [18] C. MATIAS, *Notes de cours : Analyse statistique de graphes*, Université Pierre et Marie Curie.
- [19] R. MULAS, C. KUEHN, AND J. JOST, *Coupled dynamics on hypergraphs : Master stability of steady states and synchronization*, Physical review, E-101 (2020).

- [20] P.-L. MÉLIOT, *Graphes aléatoires d'Erdős-Rényi*, Université Paris-Saclay.
- [21] M. NIDHAL JELASSI, *Etude, représentation et applications des traverses minimales d'un hypergraphe*, Master's thesis, Université de Jean Monnet, 2014.
- [22] F. SARACCO, G. PETRI, AND AL., *Entropy-based random models for hypergraphs*, ArXiv, abs/2207.12123 (2022).
- [23] P. THEVENIN, *Graphes d'erdős-rényi et grandes déviations - introduction au domaine de recherche*, Master's thesis, Département de mathématiques et applications, 2016.
- [24] J.-W. WANG, L.-L. RONG, Q.-H. DENG, AND J.-Y. ZHANG, *Evolving hypernetwork model*, The European Physical Journal B, 77 (2010), pp. 493–498.
- [25] J. ZHANG AND D. GOLDBURT, *Lecture 01 : Erdős-Renyi Model*, Cornell University, 2006.