

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**Université A.MIRA-BÉJAIA**  
Faculté des Sciences Exactes  
Département de Recherche Opérationnelle



## *Mémoire*

Présentée par :

***ALEM Lala Maghnia***

Pour l'obtention du diplôme de

***Magister***

Filière : Mathématiques Appliquées

Option : Modélisation Mathématique et Techniques de Décision

## *Thème*

**Inégalités Stochastiques des Systèmes de Files d'Attente  
avec Rappels à Communication Bidirectionnelle**

Soutenu publiquement le

Devant le jury composé de :

Mr Smail ADJABI	Professeur	Univ. de Béjaïa	<b>Président</b>
Mr Djamil AÏSSANI	Professeur	Univ. de Béjaïa	<b>Rapporteur</b>
Mr Mohamed BOUALEM M.C.A		Univ. de Béjaïa	<b>Co-Rapporteur</b>
Mr Khaled KHALDI	M.C.A	Univ. de Boumerdès	<b>Examineur</b>
Mr Karim ABBAS	M.C.A	Univ. de Béjaïa	<b>Examineur</b>
Mme Louiza LEKADIR	M.C.B	Univ. de Béjaïa	<b>Invitée</b>



*Louange A Dieu, le miséricordieux, sans Lui rien de tout cela  
n'aurait pu être.*

*Je tiens à remercier vivement mon encadreur M<sup>r</sup> D. AISSANI (Professeur à l'Université de Béjaia) d'avoir accepté de m'encadrer et pour l'honneur qu'il m'a fait en assurant le suivi scientifique et technique du présent travail. Qu'il trouve ici l'expression de mes respectueuse reconnaissance.*

*Je tiens à remercier également M<sup>r</sup> S. ADJABI (Professeur à l'Université de Béjaia) pour l'honneur qu'il me fait en acceptant de présider le jury de ce mémoire.*

*Mes remerciements s'adressent également à M<sup>r</sup> K. KHALDI (M.C.A à l'Université de Boumerdès) , M<sup>r</sup> K. ABBAS (M.C.A à l'Université de Béjaia) et M<sup>r</sup> M. BOUALEM (M.C.A à l'Université de Béjaia) de m'avoir honoré en acceptant d'examiner ce travail et M<sup>me</sup> L. LEKADIR (M.C.B à l'Université de Béjaia) de m'avoir honoré en acceptant l'invitation.*

*Je remercie tous ceux qui ont participé de près ou de loin à l'élaboration de ce travail particulièrement mon mari.*

*Je tiens également à remercier ma famille, tous mes professeurs, tous les membres de LA-MOS, et tous mes amis(es).*



*À l'être le plus cher au monde : mon mari MOH.*

*À mes sources de courage, à ceux que j'ai de plus chers : ma mère et mon frère loupi.*

*À mon père et sa femme Nora*

*À mon adorable sœurs : Tchatcha et sa petite famille.*

*À mes frères : Nadir et sa petite famille, Fares et sa femme Assia.*

*À ma belle famille*

*À celle qui ma toujours soutenu et aidé dans les moments difficiles :Fafa.*

*À tous mes amis(es), en particulier : Mouloud , Baya, Djouder, Radia ,Saida, .*

*À mes collègues : Sofiane, Wahid, Mohamed .*

*À tout mes amis(es) de la R.O.*

# Table des matières

<b>Table des matières</b>	<b>1</b>
<b>Table des Figures</b>	<b>3</b>
<b>Liste des Tableaux</b>	<b>1</b>
<b>Introduction générale</b>	<b>2</b>
<b>Chapitre 1 Systèmes de files d'attente</b>	<b>5</b>
Introduction . . . . .	5
1.1 Description du modèle d'attente classique . . . . .	6
1.2 Analyse mathématique d'un système de files d'attente . . . . .	6
1.3 Classification des systèmes d'attente . . . . .	7
1.4 Notation de Kendall [?] . . . . .	7
1.5 Les différentes disciplines de service . . . . .	8
1.6 Mesures de performance d'une file d'attente . . . . .	8
1.7 Les files d'attente markoviennes . . . . .	9
1.7.1 Système d'attente $M/M/1$ . . . . .	9
1.7.2 La file d'attente $M/M/m$ . . . . .	10
1.7.3 La file d'attente $M/M/m/K$ . . . . .	11
1.7.4 La file d'attente $M/M/\infty$ . . . . .	12
1.7.5 Le modèle d'attente $M^{[X]}/M/1$ [?, ?] . . . . .	12
1.8 Les files d'attente non markoviennes . . . . .	13
1.8.1 Système d'attente $M/G/1$ . . . . .	14
1.8.2 Système d'attente $G/M/1$ . . . . .	16
Conclusion . . . . .	18

<b>Chapitre 2</b>	<b>Systèmes d'attente avec rappels et priorité</b>	<b>19</b>
Introduction	. . . . .	19
2.1	Systèmes d'attente avec rappels . . . . .	20
2.1.1	Description du modèle d'attente avec rappels . . . . .	20
2.1.2	Politiques d'accès au serveur à partir de l'orbite . . . . .	22
2.2	Quelques cas modélisés par des systèmes de files d'attente avec rappels . . . . .	23
2.2.1	Problème de réservation . . . . .	23
2.2.2	Système informatique à temps réel . . . . .	23
2.2.3	Réseaux locaux <i>CSMA</i> . . . . .	24
2.3	Files d'attentes avec rappels et priorité . . . . .	24
Notes bibliographiques	. . . . .	25
2.4	Quelques cas modélisés par des systèmes de files d'attente avec rappels et priorité [?] . . . . .	27
2.5	Quelques modèles d'attente avec rappels et priorité . . . . .	28
2.5.1	File d'attente $M_1, M_2/G/1$ avec rappels et priorité [?] . . . . .	28
2.5.2	File d'attente $M_1, M_2/G/1$ avec rappels et perte géométrique : . . . . .	30
2.5.3	File $M_1, M_2/G/1/K + 1$ avec rappels et file prioritaire à capacité finie . . . . .	30
2.5.4	La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels . . . . .	31
2.5.5	La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels et feedback . . . . .	31
2.5.6	La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidi- rectionnelle . . . . .	31
2.6	Conclusion . . . . .	32
<b>Chapitre 3</b>	<b>Généralités sur la théorie des inégalités stochastiques</b>	<b>33</b>
Introduction	. . . . .	33
3.1	Propriétés générales des ordres partiels . . . . .	34
3.1.1	Ordre stochastique . . . . .	35
3.1.2	Ordre convexe . . . . .	36
3.1.3	Ordre en transformée de Laplace . . . . .	38
3.1.4	Ordre en fonctions génératrices . . . . .	40
3.1.5	Relations entre les ordres partiels . . . . .	40
3.1.6	Propriétés de monotonie . . . . .	40
3.2	Comparabilité et monotonie des processus markoviens homogènes . . . . .	42
3.2.1	Opérateurs monotones et comparables . . . . .	42
3.2.2	Conditions de monotonie et de comparabilité . . . . .	44
3.3	Distributions non-paramétriques . . . . .	44

3.3.1	Relation avec les distributions paramétriques . . . . .	46
3.3.2	Relation entre les classes de distributions non-paramétriques . . . . .	46
	Conclusion . . . . .	46
<b>Chapitre 4</b>	<b>Inégalités stochastiques pour le système d'attente <math>M_1, M_2/G_1, G_2/1</math> avec rappels à communication bidirectionnelle</b>	<b>47</b>
	Introduction . . . . .	47
4.1	Analyse du système $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidirectionnelle [2] . . . . .	49
4.1.1	Description du modèle . . . . .	49
4.1.2	Chaîne de Markov induite [2] : . . . . .	51
4.1.3	Approche par les processus régénératifs . . . . .	54
4.2	Inégalités stochastiques pour le système $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidirectionnelle . . . . .	55
4.2.1	Inégalités préliminaires . . . . .	55
4.2.2	Monotonie de la chaîne de Markov incluse . . . . .	62
4.2.3	Inégalités stochastiques des distributions stationnaires du nombre de clients dans le système . . . . .	68
4.3	Application numérique . . . . .	70
	Conclusion . . . . .	72
	<b>Conclusion générale</b>	<b>1</b>
	<b>Conclusion générale</b>	<b>2</b>
	<b>Bibliographie</b>	<b>4</b>

# Table des figures

1.1	Un système de files d'attente . . . . .	6
2.1	Système d'attente avec rappels . . . . .	21
2.2	Schéma d'un réseau local . . . . .	25
3.1	Relations entre les classes de distributions d'âge . . . . .	46
4.1	File d'attente avec rappels à communication bidirectionnelle. . . . .	50
4.2	Les transitions entre les états du modèle. . . . .	51
4.3	Les transitions entre les états $i > 0$ . . . . .	52
4.4	Comparaison des distributions $\{k_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique, pour $\lambda^{(1)} = 0.25, \lambda^{(2)} = 0.5, \theta^{(1)} = 0.7, \theta^{(2)} = 0.6$ . . . . .	58
4.5	Comparaison des distributions $\{k_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique, pour $\lambda^{(1)} = \lambda^{(2)} = 0.5$ et $\theta^{(1)} = \theta^{(2)} = 0.6$ . . . . .	58
4.6	Comparaison des distributions $\{\bar{k}_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre convexe, pour $\lambda^{(1)} = 0.25, \lambda^{(2)} = 0.5, \theta^{(1)} = 0.7$ et $\theta^{(2)} = 0.6$ . . . . .	60
4.7	Comparaison des distributions $\{\bar{k}_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre convexe, pour $\lambda^{(1)} = \lambda^{(2)} = 0.5$ et $\theta^{(1)} = \theta^{(2)} = 0.6$ . . . . .	60
4.8	Comparaison des probabilités de transition $\bar{p}_{n,m}$ et $\bar{p}_{n-1,m}, \forall n, m$ . . . . .	63
4.9	Comparaison des probabilités de transition $\{p_{nm}^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique. . . . .	67
4.10	Comparaison des probabilités de transition $\{p_{nm}^{(i)}\}, i = 1, 2$ , par rapport à l'ordre convexe . . . . .	68
4.11	Comparaison des probabilités stationnaires des systèmes . . . . .	71

# Liste des tableaux

4.1	Les taux de transition de la file d'attente avec rappels à communication bidirectionnelle . . . . .	51
4.2	Différentes situations prises en considération lors de simulation . . . . .	71

# Introduction générale

La théorie des files d'attente est une théorie mathématique relevant du domaine des probabilités, né en 1917, des travaux de l'ingénieur danois Erlang sur la gestion des réseaux téléphoniques de Copenhague. Entre 1909 et 1920, elle a étudié notamment les systèmes d'arrivée dans une queue, les différentes priorités de chaque nouvel arrivant, ainsi que la modélisation statistique des temps d'exécution. C'est grâce aux apports des mathématiciens Khintchine, Palm, Kendall, Pollaczek et Kolmogorov que la théorie s'est vraiment développée. La théorie de files d'attente est aujourd'hui largement utilisée et ses applications sont multiples.

Vue l'apparition d'autres systèmes réels de plus en plus complexes, tel que les systèmes téléphoniques où les abonnés répétaient leurs appels en recomposant le numéro plusieurs fois jusqu'à l'obtention de la communication, des chercheurs tels que Kosten [?] et Wilkinson ont mis en évidence les limites de la théorie classique des files d'attente qui ne permettait pas d'expliquer le comportement stochastique de ce type de système.

Ce phénomène de répétition de demandes du service a poussé certains chercheurs à étendre le modèle d'attente classique à celui dit avec rappels [?]. Ce type de systèmes de files d'attente avec rappels peut être appliqué pour résoudre les problèmes pratiques, tels que l'analyse du comportement des abonnés dans les réseaux téléphoniques, l'analyse du temps d'attente pour accéder à la mémoire sur les disques magnétiques [?, ?]. Ce type de modèles se rencontre également dans la modélisation de protocoles spécifiques de communication, tels que *CSMA* (Carrier Sens Multiple Access) ou encore les disciplines Auto-Repeat, Ring-Back-When-Free, Repeat-Last-Number [?, ?, ?]. Les progrès récents dans ce domaine sont résumés dans les articles de synthèse de Falin (1990) [?], Aïssani (1994) [?], Kulkarni et Liang (1997) [?], Templeton (1999) [?] et dans les monographies de Falin et Templeton (1997) [?] et Artalejo et Gómez (2008) [?].

Les files d'attente avec rappels ont été largement utilisés pour modéliser de nombreux problèmes dans les systèmes de commutation téléphoniques, informatiques, des réseaux locaux et des situations de la vie quotidienne. Dans la plupart des publications sur les files d'attente avec rappels, le serveur ne fournit que le service aux arrivées entrantes effectuées par les clients

réguliers. Cependant, il existe des situations réelles par exemple : les centres d'appels où un opérateur non seulement sert les appels entrants, mais il effectue aussi des appels sortants vers l'extérieur lorsque le serveur est libre. Cette fonction est connue sous le nom de files d'attente avec rappels à communication bidirectionnelle [2].

L'étude des modèles de files d'attente avec rappels à communication bidirectionnelle remonte aux travaux de Falin (1979) [?]. Ce qui a ouvert les portes à diverses publications de plusieurs auteurs, à savoir Artalejo et Resing (2010) [?], Artalejo et Phung-Duc (2011) [1], Artalejo et Phung-Duc (2013) [2].

La théorie analytique des modèles d'attente avec rappels s'avère d'une portée limitée en raison de la complexité des résultats connus. En effet, dans la majorité des cas, on se retrouve confronté à des systèmes d'équations dont la résolution est complexe ou possédant des solutions qui ne sont pas facilement interprétable afin que le praticien puisse en bénéficier. Par ailleurs, on peut citer le degré de difficulté pour l'obtention de certaines caractéristiques dans quelques modèles tels que les modèles de files d'attente avec rappels et vacances, avec rappels et priorité, avec rappels à communication bidirectionnelle, avec rappels de distribution générale ayant deux types de clients. Cette difficulté réside essentiellement dans l'utilisation des inverses des transformées de Laplace-Stieljes et des distributions marginales. Pour pallier à toutes ces difficultés, les chercheurs ont recouru aux méthodes d'approximation telle que les méthodes de comparaison stochastique [?], qui permettent d'avoir des estimations qualitatives pour certaines mesures de performance. L'idée générale de cette méthode est de borner un système complexe par un autre système, plus simple à résoudre et fournissant des bornes qualitatives pour ces mesures de performance [?, ?, ?].

Le but de notre travail est d'appliquer les méthodes de comparaison stochastique, pour étudier les propriétés de monotonie du modèle  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle relativement à l'ordre stochastique, l'ordre convexe et l'ordre en transformée de Laplace afin d'obtenir des bornes simples pour la distribution stationnaire de la chaîne de Markov induite liée à ce modèle [5, ?, ?, 6].

Ce mémoire est constitué d'une introduction générale, de quatre chapitres, d'une conclusion générale et d'une bibliographie.

Le premier chapitre comprend les concepts et techniques de base de la théorie de files d'attente classiques.

Dans le second chapitre, on donne une synthèse sur les systèmes de files d'attente avec rap-

pels et priorité.

Le troisième chapitre est consacré à l'étude des ordres partiels usuels (ordre stochastique, convexe et de Laplace), ainsi que des éléments sur la théorie de comparabilité des processus stochastiques. En particulier, on définit la monotonie interne et externe d'un processus stochastique. On présente aussi les classes de distributions d'âge issues de la théorie de la fiabilité.

Dans le quatrième chapitre, nous obtenons des inégalités stochastiques pour le modèle  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle. Nous donnons les conditions pour lesquelles l'opérateur de transition de la chaîne de Markov incluse est monotone par rapport aux ordres stochastique et convexe. On étudie la comparabilité des opérateurs de transition pour les chaînes de Markov incluses de deux systèmes  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, ainsi que la comparabilité des distributions stationnaires respectives de nombre de clients dans les deux systèmes. En dernier lieu, on montre que la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle si la distribution de service est *NBUE* (New Better than Used in Expectation) (respectivement *NWUE*-New Worse than Used in Expectation).

# 1

## Systemes de files d'attente

### Introduction

La théorie des files d'attente, ou queues, et des réseaux de files d'attente sont des outils analytiques les plus puissants pour la modélisation de systèmes logistiques et de communication. En quelques mots, cette théorie a pour objet l'étude des systèmes où des entités, appelées *clients*, cherchant à accéder à des ressources, généralement limitées, afin d'obtenir un service. La demande concurrente d'une même ressource par plusieurs clients engendre des délais dans la réalisation des services et la formation de files de clients désirant accéder à une ressource indisponible. L'analyse théorique de tels systèmes permet d'établir à l'avance les performances de l'ensemble, d'identifier les éléments critiques ou, encore, d'appréhender les effets d'une modification des conditions de fonctionnement.

Les systèmes de files d'attente sont très étudiés et une abondante littérature couvre ce sujet (voir [?, ?, ?]).

## 1.1 Description du modèle d'attente classique

Une *file d'attente* ou *queue* est un système stochastique composé d'un certain nombre (fini ou non) de *places d'attente* d'un ou plusieurs *serveurs* et bien sûr de *clients* qui arrivent, attendent, se font servir selon des *règles de priorité* données et quittent le système. La description précédente d'une file d'attente, dont une représentation schématique est donnée en figure (1.1), ne saurait capturer toutes les caractéristiques des différents modèles que compte la littérature, mais elle identifie les éléments principaux permettant la classification de la grande majorité des files d'attente simples.



FIGURE 1.1 – Un système de files d'attente

## 1.2 Analyse mathématique d'un système de files d'attente

L'étude mathématique d'un système de files d'attente se fait généralement par l'introduction d'un processus stochastique, défini de façon appropriée. On s'intéresse principalement au nombre de clients  $X(t)$  se trouvant dans le système à l'instant  $t$  ( $t \geq 0$ ).

En fonction des quantités qui définissent le système, on cherche à déterminer :

- Les probabilités d'état  $P_n(t) = P(X(t) = n)$ , qui définissent le régime transitoire du processus stochastique  $\{X(t), t \geq 0\}$ . Il est évident que les fonctions  $P_n(t)$  dépendent de l'état initial ou de la distribution initiale du processus.
- Le régime stationnaire du processus stochastique est défini par :

$$\pi_n = \lim_{t \rightarrow \infty} P_n(t) = P(X(+\infty) = n) = P(X = n), \quad (n = 0, 1, 2, \dots),$$

où,  $\{\pi_n\}_{n \geq 0}$  est appelée distribution stationnaire du processus  $\{X(t), t \geq 0\}$ .

Le calcul explicite du régime transitoire s'avère généralement pénible, voire impossible, pour la plupart des modèles donnés. On se contente donc de déterminer le régime stationnaire.

### 1.3 Classification des systèmes d'attente

Pour identifier un système d'attente, on a besoin des spécifications suivantes :

- La nature stochastique du processus des arrivées, qui est défini par la distribution des intervalles séparant deux arrivées consécutives ;
- La distribution du temps aléatoire de service ;
- Le nombre  $m$  de serveurs (stations de service) qui sont montées en parallèle. On admet généralement que les temps de service correspondants suivent la même distribution et que les clients qui arrivent forment une seule file d'attente ;
- La capacité  $N$  du système. Si  $N < \infty$ , la file d'attente ne peut dépasser une longueur de  $N-m$  unités. Dans ce cas, certains clients arrivant vers le système n'ont pas la possibilité d'y entrer.

### 1.4 Notation de Kendall [?]

Un modèle de file d'attente est totalement décrit selon la notation de Kendall. Dans sa version étendue, un modèle est spécifié par une suite de six symboles :

$$A/B/s/N/M/D$$

La signification de chacun de ces symboles est :

- $A$  : nature du processus des arrivées ;
- $B$  : nature du processus de service ;
- $s$  : nombre de serveurs ;
- $N$  : capacité d'accueil de la file d'attente ;
- $M$  : taille de la population ;
- $D$  : discipline de la file.

Dans la description des processus d'arrivée et de service, les symboles les plus courants sont :

- $M$  : loi Exponentielle (memoryless) ;
- $E$  : loi d'Erlang ;
- $\Gamma$  : loi Gamma ;
- $D$  : loi Déterministe (temps d'inter-arrivées ou de service constant) ;
- $G$  : loi Générale (quelconque).

La forme abrégé :  $A/B/s$  signifie que  $N$  et  $M$  sont infinies.

### 1.5 Les différentes disciplines de service

La discipline de service décrit l'ordre avec lequel les arrivées dans le système vont accéder au service. Ces disciplines peuvent être :

- **FIFO (First In First Out)** : Le premier arrivé est le premier servi ;
- **LIFO (Last In First Out)** : Le dernier arrivé sera le premier servi ;
- **Random (aléatoire)** : Les clients accèdent au serveur de manière aléatoire, indépendamment de l'ordre des arrivées ;
- **Priorité relative** : Un client accède au service selon sa priorité. La file est gérée par ordre de priorité de la plus forte à la plus faible ;
- **Priorité absolue** : Le service d'un client est interrompu lorsqu'un client de priorité supérieure se présente devant la file d'attente. Le client dont ce service est interrompu est remis en tête de la file.

### 1.6 Mesures de performance d'une file d'attente

L'étude d'une file d'attente ou d'un réseau de files d'attente a pour but de calculer ou d'estimer les performances d'un système dans des conditions de fonctionnement données. Ce calcul se fait le plus souvent pour le régime stationnaire uniquement, et les mesures les plus fréquemment utilisées sont :

- $\bar{N} = E(X)$  : nombre moyen de clients dans le système ;
- $\bar{Q}$  : nombre moyen de clients dans la file d'attente ;
- $\bar{T}$  : temps moyen de séjour d'un client dans le système ;
- $\bar{W}$  : temps moyen d'attente d'un client dans la file ;
- $\bar{U}$  : taux d'utilisation de chaque serveur ;
- $\bar{S}$  : le temps moyen de service ;
- $\bar{A}$  : le temps moyen entre deux arrivées.

Ces valeurs ne sont pas indépendantes les unes des autres, mais sont liées par les relations suivante :

- $\bar{N} = \lambda \bar{T}$  (Formule de Little), où  $\lambda$  représente le taux d'arrivées ;
- $\bar{Q} = \lambda \bar{W}$  ;
- $\bar{T} = \bar{W} + \frac{1}{\mu}$ , où  $\mu$  représente le taux de service ;
- $\bar{N} = \bar{Q} + \rho$  ;

De manière générale, une file est *stable* si et seulement si le nombre moyen d'arrivées de clients par unité de temps, noté  $\lambda$ , est inférieur au nombre moyen de clients pouvant être servis par unité de temps. Si chaque serveur peut traiter  $\mu$  clients par unité de temps et si le nombre de serveurs est  $m$ , une file est stable si et seulement si

$$\lambda < m\mu \iff \rho = \lambda/m\mu < 1,$$

où,  $\rho$  est appelé l'intensité du trafic.

### 1.7 Les files d'attente markoviennes

Ils caractérisent les systèmes dans lesquels les deux quantités stochastiques principales qui sont le temps des inter-arrivées et la durée de service sont des variables aléatoires indépendantes exponentiellement distribuées (modèle  $M/M/1$ ). La propriété d'absence de mémoire de la loi exponentielle facilite l'étude de ces modèles. L'étude mathématique de tels systèmes se fait par l'introduction d'un processus stochastique approprié. Ce processus est souvent le processus  $\{X(t), t \geq 0\}$  défini comme étant le nombre de clients dans le système à l'instant  $t$ . L'évolution temporelle du processus markovien  $\{X(t), t \geq 0\}$  est complètement définie grâce à la propriété d'absence de mémoire.

#### 1.7.1 Système d'attente $M/M/1$

Le système de files d'attente  $M/M/1$  est le système le plus élémentaire de la théorie des files d'attente. Le flot des arrivées est poissonnien de paramètre  $\lambda$  et la durée de service est exponentielle de paramètre  $\mu$ .

#### Régime transitoire

Soit  $X(t)$  le nombre de clients présents dans le système à l'instant  $t$  ( $t \geq 0$ ). Grâce aux propriétés fondamentales du processus de Poisson et de la loi exponentielle,  $X(t)$  est un processus markovien homogène.

Les probabilités d'état  $P_n(t) = P[X(t) = n]$  peuvent être calculées par les équations

différentielles de Kolmogorov ci-dessous, connaissant les conditions initiales du processus.

$$\begin{cases} P'_n(t) = -(\lambda + \mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t), \\ \text{et} \\ P_0(t) = -\lambda P_0(t) + \mu P_1(t). \end{cases}$$

### Régime stationnaire

Sous la condition d'ergodicité du système  $\rho = \frac{\lambda}{\mu} < 1$ , pour laquelle le régime stationnaire existe, il est aisé d'obtenir les probabilités stationnaires

$$\pi_n = \lim_{t \rightarrow \infty} P_n(t) = (1 - \rho)\rho^n, \quad \forall n \in \mathbb{N}. \quad (1.1)$$

$\pi = \{\pi_n\}_{n \geq 0}$  est appelé distribution stationnaire, elle suit une loi géométrique.

### Caractéristiques du système

- Le nombre moyen de clients dans le système est :

$$\bar{N} = \mathbb{E}(X) = \sum_{n \geq 0} n\pi_n = (1 - \rho) \sum_{n \geq 0} n\rho^n.$$

D'où :

$$\bar{N} = \frac{\rho}{1 - \rho}. \quad (1.2)$$

- Le nombre moyen de clients dans la file :

$$\bar{Q} = \sum_{n \geq 1} (n - 1)\pi_n = \frac{\rho^2}{1 - \rho}. \quad (1.3)$$

Le temps moyen de séjour dans le système  $\bar{T}$  et le temps moyen d'attente dans la file  $\bar{W}$  sont obtenus à partir des formules de Little, ou des distributions du système :

- Le temps moyen de séjour dans le système :

$$\bar{T} = \frac{\rho}{\lambda(1 - \rho)}. \quad (1.4)$$

- Le temps moyen d'attente dans la file :

$$\bar{W} = \frac{\rho^2}{\lambda(1 - \rho)}. \quad (1.5)$$

### 1.7.2 La file d'attente M/M/m

Dans ce modèle,  $m$  serveurs identiques et indépendants partagent les mêmes places d'attente. Les arrivées suivent un processus de Poisson de paramètre  $\lambda$  et la durée de chaque service est

une variable exponentielle de paramètre  $\mu$ . caractéristiques de ce système sont données par les relations suivantes :

- La probabilité qu'il y ait  $n$  clients dans le système à l'instant d'entrée est

$$p_n = \begin{cases} \frac{(\lambda/\mu)^n}{n!} p_0, & \text{si } n \leq m, \\ \frac{(\lambda/\mu)^n}{m!m^{n-m}} p_0, & \text{si } n \geq m, \end{cases}$$

où

$$p_0 = \left[ \sum_{n=0}^m \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^{m+1}}{m!(m - \lambda/\mu)} \right]^{-1}.$$

$$\zeta = P(\text{attente}) = P(X \geq m) = \frac{p_m}{1 - \rho}.$$

- Le taux d'utilisation de chaque serveur est :

$$U = \rho = \frac{\lambda}{m\mu}.$$

- Le nombre moyen de clients présents et en attente sont :

$$\bar{N}m\rho + \frac{\rho\zeta}{1 - \rho}, \quad \bar{Q} = \frac{\rho\zeta}{1 - \rho}.$$

- Le temps moyen de réponse et d'attente sont

$$\bar{T} = \frac{1}{\mu} \left( 1 + \frac{\zeta}{m(1 - \rho)} \right), \quad \bar{W} = \frac{\zeta}{m\mu(1 - \rho)}.$$

### 1.7.3 La file d'attente M/M/m/K

La file M/M/m/K est une file markovienne composée de  $m$  serveurs et disposant de  $K$  places au total. Le nombre maximal de clients en attente est donc  $K-m$ . Si un client arrive alors que le système est plein, il ne peut y entrer et doit repartir. Elle est donc toujours stable quel que soit l'intensité du trafic  $\rho = \lambda/m\mu < 1$ .

Le taux de service de cette file est :

$$\mu_k = \begin{cases} k\mu, & k = 1, 2, \dots, m - 1; \\ m\mu, & k = m, m + 1 \dots K. \end{cases}$$

Comme tout client arrivant alors que le système est plein doit repartir, le taux effectif d'arrivées dans le système n'est pas  $\lambda$  mais  $\lambda' = \sum_{k=0}^{K-1} \lambda p_k = \lambda(1 - p_K)$ .

Où,  $p_k$  est la probabilité qu'il y a  $k$  clients dans le système.

Ayant calculé

$$\bar{N} = \sum_{k=1}^K k p_k,$$

et

$$\bar{Q} = \sum_{k=m+1}^K (k - m) p_k,$$

c'est ce taux effectif  $\lambda'$  qu'il faut utiliser pour calculer  $\bar{T}$  et  $\bar{W}$  à l'aide de la formule de Little.

### 1.7.4 La file d'attente M/M/ $\infty$ [?]

Cette file est composée d'une infinité de stations de service identiques. Il est évident qu'aucune file d'attente ne se forme ; chaque client est servi dès son entrée. Ce système possède non seulement un intérêt théorique, mais il permet des études approximatives de phénomène d'attente de type M/M/s ou M/M/s/s comprenant un grand nombre de stations en parallèle.

La distribution stationnaire de ce système d'attente est :

$$p_n = \left(\frac{\lambda}{\mu}\right)^n \frac{e^{-\lambda/\mu}}{n!} \quad (n = 0, 1, 2, \dots)$$

En ce qui concerne les caractéristiques du système, on a

$$\bar{N} = \frac{\lambda}{\mu} \quad \text{et} \quad \bar{T} = \frac{1}{\mu},$$

tandis que

$$\bar{Q} = \bar{W} = 0.$$

### 1.7.5 Le modèle d'attente $M^{[X]}/M/1$ [?, ?]

La plupart des travaux sur les modèles d'attente classiques traitent des systèmes dans lesquels les clients arrivent un par un. Cependant, dans plusieurs situations pratiques, on rencontre souvent des cas où les clients arrivent par groupes et non pas séparément. Ces situations d'attente peuvent être représentées par des modèles appelés **modèles avec arrivées par groupe**.

Ce modèle est caractérisé par les éléments suivants :

- Les groupes arrivent selon un processus de Poisson de taux  $\lambda$  ;
- Les clients sont servis individuellement, les durées des services étant indépendantes et identiquement distribuées suivant une loi exponentielle de moyenne  $\frac{1}{\mu}$  ;
- La politique de service est *FIFO* ;
- Le nombre de clients par groupe est une variable aléatoire  $X$  strictement positive. On pose :

$$P(X = x) = c_x.$$

Les arrivées des clients ne forment pas un processus de Naissance et de Mort. Cependant, le système est Markovien, puisque le comportement future du système dépend uniquement de la situation présente.

### 1.8 Les files d'attente non markoviennes

En l'absence de l'exponentialité ou plutôt lorsque l'on s'écarte de l'hypothèse d'exponentialité de l'une des deux quantités stochastiques : le temps des inter-arrivées et la durée de service, ou en prenant en compte certaines spécificités des problèmes par introduction de paramètres supplémentaires, on aboutit à un modèle non markovien. La combinaison de tous ces facteurs rend l'étude mathématique du modèle très délicate. On essaye alors de se ramener à un processus de Markov judicieusement choisi à l'aide de l'une des méthodes d'analyse suivantes [?, ?] :

**Méthode des étapes d'Erlang :** Son principe est d'approximer toute loi de probabilité ayant une transformée de Laplace rationnelle par une loi de Cox (mélange de lois exponentielles), cette dernière possède la propriété d'absence de mémoire par étapes.

**Méthode de la chaîne de Markov induite :** Cette méthode, élaborée par Kendall [?], est souvent utilisée. Elle consiste à choisir une séquence d'instantants  $1, 2, 3, \dots, n$  (déterministes ou aléatoires) telle que la chaîne induite  $\{X_n, n \geq 0\}$ , où  $X_n = X(n)$ , soit markovienne et homogène.

**Méthode des variables auxiliaires :** Elle consiste à compléter l'information sur le processus  $\{X(t), t \geq 0\}$  de telle manière à lui donner le caractère markovien. Ainsi, on se ramène à l'étude du processus  $\{X(t), A(t_1), A(t_2), \dots, A(t_n)\}$ . Les variables  $A(t_k), k \in \{1, 2, \dots, n\}$  sont dites auxiliaires.

**Méthode des événements fictifs :** Le principe de cette méthode est d'introduire des événements fictifs qui permettent de donner une interprétation probabiliste aux transformées de Laplace et aux variables aléatoires décrivant le système étudié.

**Simulation :** C'est un procédé d'imitation artificielle d'un processus réel donné sur ordinateur.

Elle nous permet d'étudier les systèmes les plus complexes, de prévoir leurs comportements et de calculer leurs caractéristiques. Les résultats obtenus ne sont qu'approximatifs, mais peuvent être utilisés avec une bonne précision. Cette technique se base sur la génération de variables aléatoires [?] suivant les lois gouvernant le système.

### 1.8.1 Système d'attente $M/G/1$

Le flot des arrivées dans le système  $M/G/1$  est poissonnien de paramètre  $\lambda$  et la durée de service est distribuée selon une loi générale  $G$  de moyenne  $1/\mu$ . La particularité de ce système est que, contrairement au cas  $M/M/1$ , le processus  $X(t)$  n'est pas markovien.

#### Chaîne de Markov induite et probabilités de transition

Soit  $X_n$  : le nombre de clients dans le système  $M/G/1$  à la fin de service du  $n^{\text{ième}}$  client. Notons par  $G(s)$  la distribution de la durée de service et par  $\lambda$  le paramètre de la distribution exponentielle régissant la durée entre deux arrivées consécutives.

Le processus  $\{X_n, n \geq 0\}$  est une chaîne de Markov, d'opérateur de transition  $P = [p_{ij}]_{i,j \geq 0}$ , où :

$$p_{ij} = \begin{cases} p_j, & \text{si } i = 0; \\ p_{j-i+1}, & \text{si } i \geq 1; \end{cases} \quad (1.6)$$

avec

$$p_k = \int \frac{e^{-\lambda s} (\lambda s)^k}{k!} dG(s), \quad k = 0, 1, 2, \dots$$

En effet, si  $A_n$  est le nombre de clients qui entrent dans le système pendant le  $n^{\text{ième}}$  service, on a :

$$X_{n+1} = X_n - \delta_n + A_{n+1}, \quad \text{avec } \delta_n = \begin{cases} 1, & \text{si } X_n > 0; \\ 0, & \text{si } X_n = 0. \end{cases} \quad (1.7)$$

Ceci montre que  $X_{n+1}$  ne dépend que de  $X_n$  et de  $A_{n+1}$  et non pas de  $X_{n-1}, X_{n-2}, \dots$ . Ce qui signifie que la suite  $\{X(t), t \geq 0\}$  est markovienne, où  $X(t)$  est le nombre de clients dans le système à l'instant  $t$ .

Par ailleurs,  $P(A_n = \frac{k}{t}) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}$ , car le nombre de clients  $A_n$  qui entrent dans le système, est distribué suivant une loi de Poisson de paramètre  $(\lambda t)$ . Et d'après le théorème des probabilités totales,  $P(A_n = k) = P_k = \int \frac{e^{-\lambda t} (\lambda t)^k}{k!} dG(t)$ , où  $P_k > 0$ , ( $k = 1, 2, \dots$ ).

### Régime stationnaire

Le régime stationnaire du système existe et est identique à l'état stationnaire de la chaîne de Markov induite  $X_n$ , si  $\rho = \lambda/\mu < 1$ . Il ne sera généralement pas possible de trouver la distribution stationnaire  $\pi = (\pi_0, \pi_1, \dots)$ . Cependant, nous pouvons calculer la fonction génératrice correspondante  $\Pi(z)$  (voir [?]) :

$$\Pi(z) = G^*(\lambda - \lambda z) \frac{(1 - \rho)(1 - z)}{G^*(\lambda - \lambda z) - z}, \quad (1.8)$$

où  $G^*$  représente la transformée de Laplace de la densité de probabilité du temps de service, et  $z$  est un nombre complexe vérifiant  $|z| \leq 1$ . La formule (1.8) est appelée formule de *Pollaczek-Khintchine*.

### Caractéristiques du système

On note  $\lambda$  le taux d'arrivée des clients. Cela signifie que l'espérance de la durée séparant deux arrivées successives est  $E(X) = 1/\lambda$ .

On note  $\mu$  le taux de service des clients. Cela signifie que l'espérance de la durée de service est  $E(Y) = 1/\mu$ .

L'intensité du trafic s'exprime de la manière suivante :

$$\rho = \frac{\lambda}{\mu} = \frac{E(Y)}{E(X)},$$

où  $X$  est la loi des inter-arrivées et  $Y$  est la loi de service.

- Le nombre moyen de clients dans le système :

Cette quantité peut être déterminée, en régime stationnaire, en utilisant la relation :

$$\mathbb{E}(X) = \lim_{z \rightarrow 1} \Pi'(z).$$

Néanmoins, ce calcul s'avère compliqué.

Par contre, elle peut être obtenue aisément en utilisant la relation (1.7).

$$\bar{N} = \mathbb{E}(X_n) = \rho + \frac{\rho^2 + \lambda^2 \mathbb{V}(Y)}{2(1 - \rho)}. \quad (1.9)$$

Où  $\mathbb{V}(Y)$  est la variance de la variable aléatoire  $Y$ .

- Le nombre moyen de clients dans la file est :

$$\bar{Q} = \frac{\rho^2 + \lambda^2 \mathbb{V}(Y)}{2(1 - \rho)}. \quad (1.10)$$

En utilisant la formule de Little, on obtient :

- Le temps moyen de séjour dans le système :

$$\bar{T} = \frac{1}{\mu} + \lambda \left( \frac{\mathbb{V}(Y) + 1/\mu^2}{2(1-\rho)} \right). \quad (1.11)$$

- Le temps moyen d'attente dans la file :

$$\bar{W} = \bar{T} - \frac{1}{\mu} = \lambda \left( \frac{\mathbb{V}(Y) + 1/\mu^2}{2(1-\rho)} \right). \quad (1.12)$$

### 1.8.2 Système d'attente $G/M/1$

Le système de files d'attente  $G/M/1$  constitue le "dual" du système  $M/G/1$ . Dans ce cas, les temps des inter-arrivées des clients est une suite de variables aléatoires distribuées selon une loi générale commune  $F$ , de moyenne  $1/\lambda$ , et les temps de service sont indépendants et identiquement distribués selon une loi exponentielle de paramètre  $\mu$ .

#### Chaîne de Markov induite

On peut montrer que le processus bidimensionnel  $\{X(t), \delta(t)\}$ , où  $\delta(t)$  représente le temps écoulé depuis la dernière arrivée avant  $t$ , est un processus markovien. Comme dans le cas  $M/G/1$ , ce processus peut être simplifié à un processus unidimensionnel en le considérant à des instants particuliers. En effet, en choisissant les instants  $t_n$  de l'arrivée du  $n^{\text{ième}}$  client, il est clair que les variables aléatoires  $\delta(t_n)$  sont nulles. On aura donc à étudier la chaîne de Markov à temps discret  $X_n = X(t_n) =$  "nombre de clients dans le système juste avant l'arrivée du  $n^{\text{ième}}$  client".

Soit  $\pi_k = \lim_{t \rightarrow \infty} P(X(t) = k)$ , et  $p_k = \lim_{n \rightarrow \infty} P(X_n = k)$  la probabilité stationnaire de la chaîne  $X_n$ . Contrairement au cas du système  $M/G/1$ , dans le système de files d'attente  $G/M/1$ , l'égalité entre  $\pi_k$  et  $p_k$ ,  $k \in \mathbb{N}$ , n'est généralement pas réalisée. Par ailleurs, la condition d'ergodicité de la chaîne de Markov induite  $X_n$  est la même que du système  $M/G/1$  ( $\rho = \lambda/\mu < 1$ ).

#### Probabilités de transition

Soient  $P_{ij} = P[X_{n+1} = j / X_n = i]$ , les probabilités de transition de la chaîne de Markov induite  $X_n$ . Il est aisé de vérifier que la matrice de transition  $P = [P_{ij}]_{i,j \in \mathbb{N}}$ , a la forme suivante :

$$P = \begin{pmatrix} \beta_1 & \alpha_0 & 0 & 0 & 0 & \cdots \\ \beta_2 & \alpha_1 & \alpha_0 & 0 & 0 & \cdots \\ \beta_3 & \alpha_2 & \alpha_1 & \alpha_0 & 0 & \cdots \\ \beta_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 & \cdots \\ & & \ddots & & & \ddots \end{pmatrix}$$

où

$$\beta_{i+1} = \int_0^\infty \sum_{k=i+1}^\infty e^{-\mu x} \frac{(\mu x)^k}{k!} dF(x), i \geq 0,$$

et

$$\alpha_k = \int_0^\infty e^{-\mu x} \frac{(\mu x)^k}{k!} dF(x), k \geq 0.$$

### Régime stationnaire

Nous sommes à présent en mesure de trouver les probabilités ergodiques,  $p_k$  ( $k \in \mathbb{N}$ ), de la chaîne de Markov induite  $X_n$ . La condition d'ergodicité étant  $\rho < 1$ . On peut facilement vérifier que :

$$p_k = (1 - \sigma)\sigma^k, \quad k = 0, 1, 2, \dots$$

où  $\sigma$  est l'unique solution de l'équation :

$$\sigma = F^*(\mu - \mu\sigma) = \int_0^\infty e^{-(\mu - \mu\sigma)x} dF(x),$$

$F^*$  étant la transformée de Laplace de la densité de probabilité des temps entre les arrivées des clients. On peut montrer que  $0 < \sigma < 1$ . Ainsi, le nombre de clients dans le système  $G/M/1$  à l'instant d'occurrence d'une arrivée, est distribué selon une loi géométrique.

Si  $\pi_k = \lim_{t \rightarrow \infty} P(X(t) = k)$ , alors on peut aisément vérifier les relations suivantes :

$$\pi_k = \rho p_{k-1}, \quad k = 1, 2, \dots \quad \text{et} \quad \pi_0 = 1 - \rho.$$

Ces relations confirment l'efficacité de la méthode de la chaîne de Markov induite. On constate que toutes les caractéristiques stationnaires du système de files d'attente  $G/M/1$ , peuvent être déduites à partir des caractéristiques stationnaires de la chaîne de Markov  $X_n$  (bien que les probabilités stationnaires des deux processus  $X(t)$  et  $X_n$  soient différentes).

### Caractéristiques du système

On remarque que l'étude du système de files d'attente  $G/M/1$  est plus simple que celle du système  $M/G/1$ . Dans ce cas, il suffit de trouver la valeur de  $\sigma$  pour déduire toutes les caractéristiques de ce système. En effet :

- Le nombre moyen de clients dans le système s'obtient facilement par la formule :

$$\begin{aligned} L = \mathbb{E}(X) &= \sum_{k \geq 0} k \pi_k = \sum_{k \geq 0} k \rho p_{k-1} = \sum_{k \geq 0} \rho(k+1)p_k \\ &= \rho \sum_{k \geq 0} k p_k + \rho \sum_{k \geq 0} p_k = \frac{\rho}{1 - \sigma}. \end{aligned}$$

- Le nombre moyen de clients dans la file est donné par :

$$L_q = \frac{\rho\sigma}{1 - \sigma}. \quad (1.13)$$

- Le temps moyen de séjour dans le système est donné, en utilisant la formule de Little par :

$$W = \frac{L}{\lambda} = \frac{1}{(1 - \sigma)\mu}. \quad (1.14)$$

- Le nombre moyen de clients dans le système à l'arrivée d'un client :

On peut obtenir également le nombre moyen de clients dans le système que trouve un client à son arrivée ( $L_a$ ). Cette quantité, contrairement au cas  $M/G/1$ , est différente de  $L$ .

$$L_a = \sum_{k \geq 1} k p_k = \frac{\sigma}{1 - \sigma}. \quad (1.15)$$

On remarque que  $L_a/L = \sigma/\rho$ . Donc, on ne peut pas à priori donner aucune comparaison entre ces deux valeurs (vu que  $\rho < 1$  et  $\sigma < 1$ ).

## Conclusion

Dans ce chapitre, nous avons rappelé et présenté les concepts et techniques de base de la théorie de files d'attente classiques. Plus précisément, on a exposé quelques modèles d'attente particuliers et on a donné leurs principales caractéristiques, tout en abordant les files markoviennes et les files non markoviennes.

Les modèles d'attente développés ces dernières décennies tentent de prendre en considération des phénomènes de répétition de demandes de service et de priorité à la fois. Ces phénomènes affectent les caractéristiques de performance des systèmes réels. Ces systèmes d'attente avec rappels et priorité feront l'objet du chapitre suivant.

# 2

## Systemes d'attente avec rappels et priorité

### Introduction

Beaucoup de situations de files d'attente ont la particularité que les clients qui arrivent et trouvent la zone de service occupé quittent le système temporairement et se joignent à un groupe de clients insatisfaits, mais ils répètent leur demande après un certain temps aléatoire. Entre les essais du client est dit être en orbite. Ces modèles de files d'attente apparaissent dans la modélisation stochastique de nombreux protocoles de communication, des réseaux locaux et des situations de la vie quotidienne. L'exemple le plus simple et évident est fourni par une personne qui effectue un appel téléphonique, si la ligne est occupée, alors il n'est pas en mesure de faire la queue, mais il tente à nouveau sa chance après un certain temps.

Pour des raisons de commodité, nous classons les files d'attente avec rappels en deux catégories selon le nombre d'appels différents, modèle avec un type unique d'appels et modèles dans lequel il y'a deux flux d'arrivée [2].

### 2.1 Systèmes d'attente avec rappels

La modélisation des tentatives répétées a été l'objet de nombreuses recherches portant sur l'influence du phénomène de rappel sur la qualité du service et l'analyse de la performance des systèmes avec rappels [?, ?].

Une classification bibliographique sur les systèmes avec rappels est donnée par Artalejo [?, ?, ?]. Pour illustrer le rôle actif des files d'attente avec rappels au cours des dernières années, nous mentionnons certains articles récents [7, ?, 8, ?, ?, ?].

#### 2.1.1 Description du modèle d'attente avec rappels

Un système d'attente avec rappels (Retrial Queue) est un système composé de  $s$  ( $s \geq 1$ ) serveurs identiques et indépendants, d'un buffer de capacité  $N - s$  ( $N \geq s$ ) et d'une orbite de capacité  $M$ . À l'arrivée d'un client, s'il y a un ou plusieurs serveurs libres et en bon état, le client sera servi immédiatement et quittera le système à la fin de son service. Sinon, s'il y a des positions d'attente libres dans le buffer, le client le rejoindra. Par ailleurs, si un client arrive et trouve tous les serveurs et toutes les positions d'attente du buffer occupés, il quittera le système définitivement avec la probabilité  $1 - H_0$ , ou bien entre en orbite avec la probabilité  $H_0$  et devient une source d'appels répétés et tentera sa chance après une durée de temps aléatoire.

Les clients qui reviendront et rappelleront pour le service sont dits en "orbite". Cette dernière peut être finie ou infinie. Dans le cas d'une orbite à capacité finie, si elle est pleine, un client qui trouve tous les serveurs et les positions d'attente du buffer occupés, sera obligé de quitter le système définitivement sans être servi.

Chaque client en orbite appelé aussi «*client secondaire*», est supposé rappeler pour le service à des intervalles de temps suivant une loi de probabilité et une intensité de rappels bien définie (rappels constants, rappels classiques, ou bien rappels linéaires, ...). Chacun de ces clients secondaires est traité comme un *client primaire* c'est-à-dire un nouveau client qui arrive de l'extérieur du système. S'il trouve un serveur libre, il sera servi immédiatement puis quittera le système. Sinon, s'il y a des positions d'attente disponibles dans le buffer, il le rejoindra. Par contre, si tous les serveurs et les positions d'attente sont encore occupés, le client quittera le système pour toujours avec la probabilité  $1 - H_k$  (si c'est le  $k^{\text{ème}}$  rappel sans succès) ou bien entre en orbite avec la probabilité  $H_k$  si l'orbite n'est pas pleine.

Le schéma général d'un système d'attente avec rappels est donné par la figure 2.1.

**Remarque 2.1.** 1. Le modèle d'attente avec rappels décrit ci-dessus est un modèle général.

Plusieurs systèmes de files d'attente avec rappels peuvent être considérés comme des cas particuliers tels que : les systèmes sans buffer, les systèmes à un seul serveur, ...

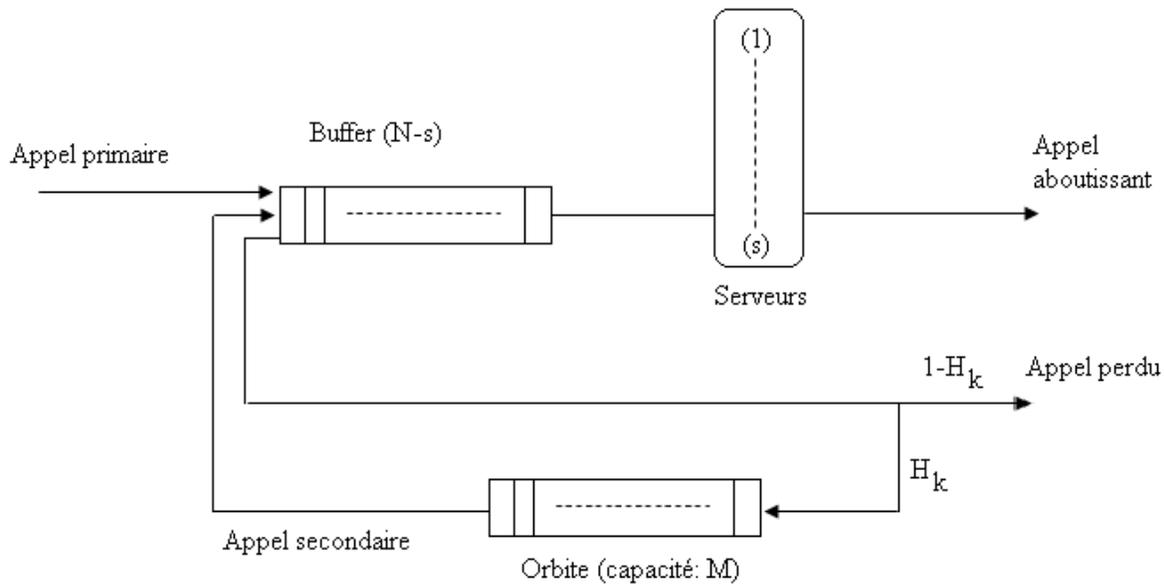


FIGURE 2.1 – Système d'attente avec rappels

- La description d'un système de files d'attente ordinaire (classique) se fait avec ses éléments principaux : le processus d'arrivées, le mécanisme de service (disponibilité et nombre de serveurs) et la discipline d'attente. Pour un système avec rappels, on doit ajouter un élément décrivant la loi des répétitions d'appels. En fonction du modèle considéré, on pourra introduire d'autres éléments décrivant la fiabilité du serveur, les types de priorité, ...
- Les clients primaires ou secondaires qui arrivent durant un temps de service, entrent en orbite sans aucune influence sur le processus de service.

### Notation : [?]

En utilisant la notation de Kendall, un modèle de files d'attente avec rappels est noté comme suit :  $A/B/s/N/M/H$ , où

- $A$  : décrit la distribution des temps des inter-arrivées des clients,
- $B$  : décrit la distribution du temps de service de chaque client,
- $s$  : le nombre de serveurs dans le système,
- $N$  : la capacité du système,
- $M$  : la taille de la population (source) de clients,
- $H$  : la fonction de persévérance qui permet de définir le comportement du client devant une situation de blocage (serveurs occupés),

$H$  peut être décrite par un vecteur  $H = (H_0, H_1, H_2, \dots, H_k, \dots)$ , où  $H_k$  est la probabilité qu'après que la  $k^{\text{ème}}$  tentative échoue, un abonné rappelle pour la  $(k + 1)^{\text{ème}}$  fois,

- Quand  $H_k = 1$  pour  $k \geq 0$ , le système devient un système sans perte. Ainsi, chaque client reçoit éventuellement le service si  $M$  est infinie. Dans ce cas,  $H = NL$  (sans perte),
- Quand  $H_k = \alpha < 1$  pour  $k \geq 0$ , le système est dit un système à perte géométrique et  $H = GL$  (Geometric Loss).

### 2.1.2 Politiques d'accès au serveur à partir de l'orbite

La définition du protocole de rappels est en effet un sujet de controverse (voir Falin (1990) [?]) et concerne l'aspect modélisation du système sous étude. Le protocole le plus décrit dans la théorie classique des files d'attente avec rappels est la politique de rappels classiques dans laquelle chaque source dans l'orbite rappelle après un temps exponentiellement distribué avec un paramètre  $\alpha$ . Donc, il y a une probabilité  $n\alpha dt + o(dt)$  d'un nouveau rappel dans le prochain intervalle  $(t, t + dt)$  sachant que  $n$  clients sont en orbite à l'instant  $t$ . Une telle politique a été motivée par des applications dans la modélisation du comportement des abonnés dans les réseaux téléphoniques depuis les années 1940.

Dans les années précédentes, la technologie a considérablement évoluée. La littérature de files d'attente avec rappels décrit différents protocoles de rappels spécifiques à certains réseaux informatiques et de communication modernes dans lesquels le temps inter-rappels est contrôlé par un dispositif électronique et par conséquent, est indépendant du nombre d'unités demandant le service. Dans ce cas, la probabilité d'un rappel durant  $(t, t + dt)$ , sachant que l'orbite est non vide, est  $\nu dt + o(dt)$ . Ce type de discipline de rappels est appelé politique de rappels constants. Le premier travail dans cette direction est celui de Fayolle [?] qui considère une file d'attente  $M/M/1$ , où uniquement le client en tête de la file en orbite peut demander un service après un temps de rappels exponentiellement distribué avec un taux constant. Cette sorte de politique de contrôle de rappels est bien connue pour le protocole ALOHA dans les systèmes de communication. Certains autres travaux décrivent des applications aux réseaux locaux, protocole de communication, systèmes mobiles et autres (Choi (1992) [?], Dudin et al. (2004) [?], Li et Zhao (2005) [?], Shikata (1999) [?]). Artalejo et Gómez-Corral (1997) [?] traitent les deux cas d'une manière unifiée en définissant une politique de rappels linéaires pour laquelle la probabilité d'un rappel durant  $(t, t+dt)$  sachant que  $n$  client sont en orbite à l'instant  $t$  est  $(\nu (1 - \delta_{0n}) + n \alpha) dt + o(dt)$ . On mentionne aussi l'existence d'une autre politique dite politique de rappels quadratiques [?].

### 2.2 Quelques cas modélisés par des systèmes de files d'attente avec rappels

Il existe aujourd'hui des centaines de publications sur les systèmes avec rappels où des exemples concrets ont été cités (Yang et Templeton (1987) [?], Aïssani (1994) [?], Falin et Templeton (1997) [?], Artalejo et Gómez (2008) [?], Amador et Artalejo (2009) [?]) en rapport avec les nouveaux développements technologiques dont l'intérêt porté s'accroît de jour en jour. Nous présentons quelques exemples de problèmes (extraits de [?]) pouvant être modélisés par ces systèmes. Ceux-ci vont du cas le plus simple de réservation à d'autres cas plus complexes comme les réseaux locaux *CSMA*.

#### 2.2.1 Problème de réservation

C'est l'exemple le plus simple d'un client qui sollicite une réservation par téléphone dans un restaurant. Il y a une ligne unique qui est consacrée à répondre aux requêtes des réservations. Ainsi, si un client appelle et trouve la ligne occupée, il renouvellera sa tentative après une certaine période de temps aléatoire avec la probabilité  $H_k$  qui, en pratique, est strictement inférieure à 1 car le client ne peut rappeler indéfiniment.

Cet exemple peut être modélisé par une file d'attente  $M/G/1$  avec rappels et avec perte en considérant que le processus d'arrivée des appels est poissonnien. L'étude de ce genre de problèmes permet de prédire le temps d'attente du client, le nombre de clients perdus dû à ce blocage, ...

#### 2.2.2 Système informatique à temps réel

Dans un système informatique à temps réel, on trouve  $M$  terminaux et  $S$  canaux de transmission tels que  $M > S$ . Pour qu'un terminal soit connecté à l'ordinateur, il suffit d'un canal de transmission libre. L'illustration de ce genre de système est le centre de calcul où arrive un étudiant pour utiliser l'ordinateur pendant une période de temps aléatoire. Celui-ci doit d'abord trouver un terminal libre pour se connecter. S'il n'y a aucun terminal disponible, il retentera sa chance après un temps aléatoire. Sinon, il envoie sa demande au commutateur central pour se connecter à l'ordinateur. Le terminal est alors connecté selon que le canal serait disponible ou pas. Dans ce dernier cas, la demande est mise dans la file par le commutateur en attente de libération d'un canal.

Ce système peut être modélisé par une file  $G/G/S$  avec rappels, avec un tampon (espace d'attente) de capacité  $M$  et une orbite de taille infinie, où les canaux de transmission correspondent aux serveurs et les terminaux au tampon.

### 2.2.3 Réseaux locaux *CSMA*

Dans les réseaux locaux se partageant un bus unique, l'un des protocoles de communication le plus généralement utilisé est appelé protocole non-persistant *CSMA* (Carrier Sense Multiple Access), c'est une méthode d'accès à un réseau local.

Un réseau local simple est composé de stations ou de terminaux interconnectés par un bus unique, qui est le canal de communication. Ainsi, les stations communiquent les unes avec les autres via le bus qui peut être utilisé par une seule station à la fois. Une telle architecture de réseau d'ordinateurs local est appelée architecture en bus.

Des messages de longueurs variables arrivent aux stations du monde extérieur. En recevant le message, la station le découpe en un nombre fini de paquets de longueur fixe, et consulte immédiatement le bus pour voir s'il est occupé ou bien libre. Si le bus est libre, l'un de ces paquets est transmis via ce bus à la station de destination et les autres paquets sont stockés dans le tampon pour une transmission ultérieure. Par contre, si le bus n'est pas libre, tous les paquets sont stockés dans le tampon (positions d'attente) et la station peut reconsulter le bus après une certaine période aléatoire.

Ce problème peut être modélisé comme un système d'attente avec rappels à un seul serveur, qui est le bus, et les tampons des stations représentent l'orbite.

Ce système est décrit dans la Figure 2.2.

Les caractéristiques du système  $M/G/1$  avec rappels sont données dans l'article de Yang et Templeton [?].

## 2.3 Files d'attentes avec rappels et priorité

Dans ce type de système, il est assez naturel d'envisager une file d'attente avec rappels à deux types d'appels où, une fois bloqué, les appels de type I sont en attente et vont être servis dès que le serveur soit libre, et les appels de type II entrent en orbite pour retenter leur chance, de sorte que les appels de type I ont la priorité sur les appels de type II. Les files d'attente avec rappels avec deux types d'appels sont essentiellement plus difficiles que les files d'attente avec rappels à un seul type d'appels.

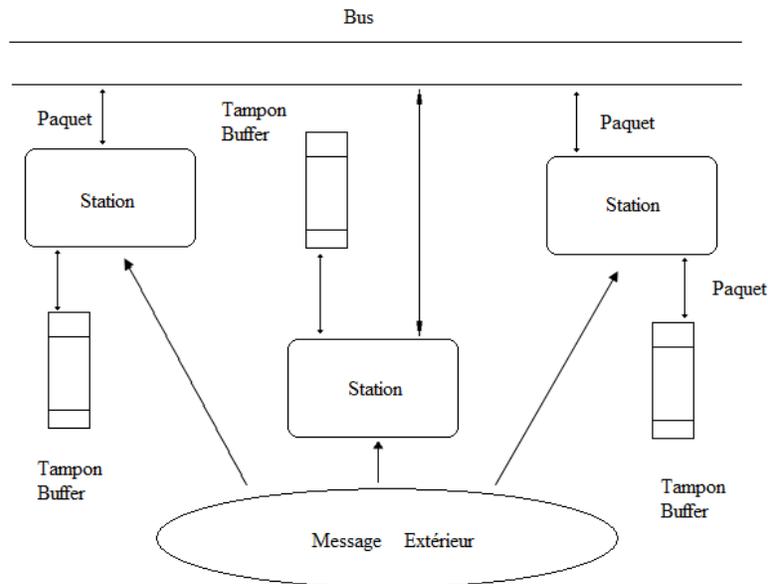


FIGURE 2.2 – Schéma d'un réseau local

### Notes bibliographiques

Les files d'attente avec rappels à un seul serveur avec appels prioritaires ont été étudiés par plusieurs auteurs [?, ?, ?, ?].

Choi et Park (1990) [?] optent pour la méthode de la variable supplémentaire pour étudier une file d'attente  $M_1, M_2/G/1$  avec rappels à deux types d'appels, avec une file d'attente prioritaire infinie pour le type d'appels I et une orbite infinie de type d'appels II, et ont obtenu la fonction génératrice commune des longueurs de la file d'attente, alors que le temps d'attente virtuelle de type d'appel II dans ce modèle a été obtenu par Choi et al (1993) [?]. Ce modèle s'est avéré être essentiellement identique à la file d'attente avec rappels à communication bidirectionnelle et une source d'appels sortants infinie étudiés par Falin (1990) [?]. Ces mêmes mesures de performance ont été obtenues par Moutzoukis et Langaris (1995) [?] pour un modèle avec deux types d'appels et priorité pré-emptive. Et en (1996) ils ont étendu le modèle du Choi et Park au modèle dans lequel il existe plusieurs types d'appels, arrivées par lots et vacances du serveur et ont obtenu la fonction génératrice des longueurs de la file d'attente, alors que Falin et al. (1993) [9] ont étendu les résultats du Choi et Park [?] du cas où deux types d'appels peuvent avoir différentes distributions du temps de service.

En utilisant la méthode de la matrice analytique, Choi et al. (1995) [?] ont étudié une file d'attente  $M_1, M_2/G/1/K + 1$  avec rappels et obtiennent la fonction génératrice commune des longueurs de la file d'attente. Différentes mesures de performance du système de file d'attente  $M_1, M_2/G_1, G_2/1/1$  avec rappels constant tel que, les distributions du nombre d'appels de type I perdu, le nombre de tentatives de rappels bloqués, la période d'occupation du serveur, le temps d'attente virtuelle, et la fonction génératrice commune des longueurs de la file d'attente ont été obtenues par Martin et Artalejo (1995) [?]. Han et Lee (1996) [?] ont traité une file d'attente  $M/G/1$  avec rappels et politique de contrôle des taux de rappels où les temps de rappel sont distribués exponentiellement avec un taux  $v/n$ , quand il y'a  $n$  clients en orbite. Ils obtiennent explicitement la fonction génératrice commune des longueurs de la file d'attente. Choi et Kim (1997) [?] ont analysé une file d'attente  $Geo_1, Geo_2/G/1$  avec rappels, ce qui est une contrepartie du  $M_1, M_2/G/1$  avec rappels. Ils ont obtenu la fonction génératrice commune de la taille de la file d'attente et de leurs moments.

Pour les files d'attente avec rappels à deux types d'appels et aucune priorité, Kulkarni (1986) [?] a étudié une file d'attente  $M_1^X, M_2^X/G_1, G_2/1$  où les temps de service et les taux de rappels sont différents pour chaque type et a obtenu la longueur moyenne de la file d'attente dans le système, le temps moyen d'attente, le nombre moyen de rappels pour chaque type d'appel. Ces résultats ont été étendu par Falin (1988) [?] au cas de plus de deux types d'appels. Grishchkin (1992) [?] a utilisé la théorie des processus de branchement pour analyser la file d'attente avec rappels à plusieurs types d'arrivées par lots. Il a obtenu la fonction génératrice de la taille de la file d'attente, le temps d'attente virtuelle, et le nombre moyen de rappels. Shin et Pearce ont considéré la file  $M_1, M_2/G_1, G_2/1$  avec rappels avec deux types d'appels où la capacité de l'orbite est infinie et l'autre orbite a une capacité finie, ils ont obtenu la matrice de transition et ont présenté un algorithme d'approximation de la distribution stationnaire du nombre de clients dans la file.

Dans la littérature ancienne, Falin (1979) [?] formula les dérivées intégrales pour les fonctions génératrices partielles et l'expression explicite de la valeur attendue de certaines caractéristiques de performance de la file d'attente  $M/G/1$  avec rappels à communication bidirectionnelle en supposant que les appels entrants et sortants ont la même distribution du temps de service. Cependant, cette hypothèse est restrictive dans la pratique, car les différents types de clients présentent généralement des comportements différents et, par conséquent, ils doivent avoir des durées de services différentes. Artalejo et Resing (2010) [?] optent pour l'analyse de la valeur moyenne technique pour obtenir des valeurs attendues relatives au temps d'attente dans une file d'attente  $M/G/1$  avec rappels à communication bidirectionnelle et les différentes distributions du temps de service des appels entrants et sortants.

Récemment, Artalejo et Phung-Duc (2011) [1] ont effectué une étude détaillée de la file d'attente  $M/M/1$  avec rappels à communication bidirectionnelle. Ils ont obtenu des expressions explicites pour la distribution stationnaire du nombre de clients en orbite, ainsi que les moments partiels.

La plupart des formules explicites dans [1] sont exprimées en termes de séries hypergéométrique, en accord avec le rôle particulier joué par ces fonctions spéciales dans le calcul de solutions analytiques pour beaucoup d'autres files d'attente avec rappels. En (2013) [2] ils ont étudié le comportement d'un état d'équilibre d'une file  $M/G/1$  avec rappels dans laquelle il y'a deux flux d'arrivée entrants à savoir les appels effectués par des clients réguliers et les appels sortants effectués par le serveur lorsqu'il est inactif. Ils ont effectué une analyse stationnaire du système, y compris la condition de stabilité, la chaîne de Markov induite, la distribution stationnaire de l'état du serveur, le nombre de clients en orbite et le calcul des premiers moments. Ils ont aussi obtenu les résultats asymptotiques pour le nombre de clients en orbite.

### 2.4 Quelques cas modélisés par des systèmes de files d'attente avec rappels et priorité [?]

**Exemple 2.1. Téléphone au restaurant :** Considérons un téléphone dans un restaurant qui est utilisé à la fois pour accepter une commande à l'avance pour les réservations des clients à l'extérieur et d'effectuer un appel par les clients dans le restaurant. Les clients du restaurant attendaient devant le téléphone dans une ligne pour effectuer un appel et utiliser le téléphone dès que il est libre.

Ce système peut être modélisé comme une file d'attente  $M_1, M_2/G/1$  avec rappels dont les clients de l'extérieur sont considérés comme des appels de type II et les clients à l'intérieur comme des appels de type I.

**Exemple 2.2. Modules de ligne d'abonné :** Dans les échanges téléphoniques modernes, des lignes d'abonnés sont généralement connectés à ce qu'on appelle des modules de ligne d'abonné qui servent à la fois les appels entrants et sortants. Une différence importante entre ces deux types d'appels réside dans le cas de blocage dû à l'ensemble des canaux occupés, les appels sortants peuvent être en file d'attente dans le tampon infini, alors que les appels entrants sont rejetés et doivent s'engager pour rétablir la connexion. Dès que l'un des canaux est libre, le cas échéant dans le tampon, un appel sortant l'occupe immédiatement. Ainsi, les appels entrants ne peuvent pas établir la connexion aussi longtemps qu'il y'a des appels sortants dans le tampon. Cette règle implique que les appels sortants sont prioritaires par rapport au appels entrants.

Ce système peut être modélisé comme une file d'attente avec rappels et priorité à deux types d'appels où le type I sont les appels sortants et le type II sont les appels entrants.

**Exemple 2.3. Réseaux locaux CSMA :** Dans les réseaux locaux se partageant un bus unique (LAN), l'un des protocoles de communication communément utilisés est le protocole CSMA.

Il existe plusieurs versions de CSMA dont on pourra citer **CSMA non persistant** et **CSMA persistant**. En CSMA non persistant, si l'utilisateur a un paquet à transmettre et trouve le canal occupé, alors il réessaie la transmission après une période du temps aléatoire. En CSMA persistant, l'utilisateur qui détecte le canal et estime qu'il est occupé, continue d'attendre et transmet son paquet dès que le canal est inactif, de sorte que le canal est toujours utilisé s'il y'a un utilisateur avec un paquet.

Supposons qu'un réseau dispose de deux types d'utilisateurs (utilisateurs non persistants et utilisateurs persistants) reliés par un bus unique (ou canal).

La communication entre les utilisateurs est réalisé par le bus. Les utilisateurs persistants sont contrôlés par le système central de sorte que, dès que le canal est inactif, le système central permet à un utilisateur persistant d'occuper le canal de commande, pour envoyer un message s'il n'y a aucun utilisateur persistant avec un paquet. L'utilisateur non persistant avec un paquet réessaie la transmission indépendamment après un certain temps aléatoire.

Ce système peut être modélisé comme une file d'attente  $M_1, M_2/G/1$  avec rappels et priorité, le serveur correspond au canal, les utilisateurs persistants dans le système central sont considérés comme des appels de type I et les utilisateurs non persistants dans le système non central sont considérés comme des appels de type II.

## 2.5 Quelques modèles d'attente avec rappels et priorité

### 2.5.1 File d'attente $M_1, M_2/G/1$ avec rappels et priorité [?]

Dans la plupart des files d'attente avec rappels à un seul type d'appel, il est supposé que les caractéristiques des appels arrivant au système, comme, le temps d'inter-arrivées, le temps de service et le temps d'inter-rappels sont homogènes. Cependant, dans la pratique, il existe plusieurs types d'appels ayant généralement des caractéristiques différentes.

#### Ergodicité

Choi et park [?] ont montré que :  $\rho = \rho_1 + \rho_2 < 1$  est la condition de stabilité d'une file  $M_1, M_2/G/1$  avec rappels.

La condition d'ergodicité d'une  $M_1, M_2/G_1, G_2/1$  a été obtenue dans [9].

### Distribution du nombre de clients dans les deux files

Choi et Park [?] ont obtenu la distribution commune des longueurs des deux files en utilisant la méthode de la variable supplémentaire où la variable aléatoire supplémentaire est le temps de service restant d'un appel en service.

Soient :

- $N_1(t)$  : est le nombre d'appels de Type I en file d'attente prioritaire (sauf l'appel en service) à l'instant  $t$ ,
- $N_2(t)$  : est le nombre d'appels de Type II dans l'orbite à l'instant  $t$ ,

$$\xi(t) = \begin{cases} 0, & \text{si le serveur est libre à l'instant } t ; \\ 1, & \text{si le serveur est occupé par un appel de type I à l'instant } t ; \\ 2, & \text{si le serveur est occupé par un appel de type II à l'instant } t. \end{cases}$$

Le nombre moyen de clients dans la file d'attente prioritaire et en orbite sont donnés par :

$$E[N_1, \xi = 1] = \frac{\lambda_1 \lambda b^{(2)}}{2(1 - \rho_1)}, \quad (2.1)$$

$$E[N_2, \xi = 0] = \frac{\lambda \rho_2}{\nu}, \quad (2.2)$$

$$E[N_2, \xi = 1] = \frac{\lambda_2}{1 - \rho} \left( \frac{(\lambda b)^2}{\nu} + \frac{\lambda b^2}{2(1 - \rho_1)} \right), \quad (2.3)$$

$$E[N_1] = \frac{\lambda_1 \lambda b^{(2)}}{2(1 - \rho_1)}, \quad (2.4)$$

$$E[N_2] = \frac{\lambda_2 \lambda b^{(2)}}{2(1 - \rho_1)(1 - \rho)} + \frac{\lambda_2}{\nu} \frac{\rho}{1 - \rho}. \quad (2.5)$$

où :

- $\lambda_1$  :taux d'arrivée d'appel du type I ;
- $\lambda_2$  :taux d'arrivée d'appel du type II ;
- $\nu$  : taux de rappel ;

### Cas Particuliers :

1. Lorsque  $\lambda_1 = 0$ , ce modèle devient la file d'attente  $M/G/1$  avec rappels à un seul type d'appels.
2. Lorsque  $\lambda_2 = 0$ , ce modèle devient alors la file d'attente ordinaire  $M/G/1$ .

### 2.5.2 File d'attente $M_1, M_2/G/1$ avec rappels et perte géométrique :

Dans la pratique, un appel répété au bout de quelques tentatives de rappels infructueuses ne peut plus répéter son appel et décide d'abandonner le système. Soit  $H_j$  la probabilité qu'un appel fera la  $(j + 1)$ ème tentative après les  $j$ ème échecs. En particulier, nous considérons le système de tirage au sort géométrique  $p \equiv H_0$  et  $q \equiv H_i$  ( $i \geq 1$ ). Le cas  $q < 1$  et  $q = 1$  donnent des solutions fondamentalement différentes du problème.

Choi et Chang [?] ont étudié une file  $M_1, M_2/M/c$  avec rappels, où le temps de service a une distribution exponentielle et l'orbite de capacité finie  $K$ , et ont obtenu les résultats suivants : le nombre moyen de client dans le système, temps d'attente, la période d'occupation et le nombre de rappels par clients.

On a une file d'attente  $M_1, M_2/G/1$  avec rappels et perte géométrique,

Sous la condition d'ergodicité  $\rho_1 + p\rho_2 < 1$ , le nombre moyen de clients dans la file prioritaire et en orbite sont donnés par :

$$E[N_1, \xi = 1] = \frac{\lambda_1 \lambda b^{(2)}}{2(1 - \rho_1)(1 + \rho_2(1 - p))}, \quad (2.6)$$

$$E[N_2, \xi = 0] = \frac{\lambda}{\nu} \frac{p\rho_2}{1 + \rho_2(1 - p)}, \quad (2.7)$$

$$E[N_2, \xi = 1] = \frac{p\lambda_2}{(1 - \rho_1 - p\rho_2)(1 + \rho_2(1 - p))} \left( \frac{(\lambda b)^2}{\nu} + \frac{\lambda b^{(2)}}{2(1 - \rho_1)} \right), \quad (2.8)$$

$$E[N_1] = \frac{\lambda_1 \lambda b^{(2)}}{2(1 - \rho_1)(1 + \rho_2(1 - p))}, \quad (2.9)$$

$$E[N_2] = \frac{p\lambda_2 \lambda}{1 - \rho_1 - p\rho_2} \left( \frac{b}{\nu} + \frac{b^{(2)}}{2(1 - \rho_1)(1 + \rho_2(1 - p))} \right). \quad (2.10)$$

### 2.5.3 File $M_1, M_2/G/1/K + 1$ avec rappels et file prioritaire à capacité finie

Choi et al. [?] ont analysé la une file d'attente  $M_1, M_2/G/1/K + 1$  avec rappels et une file d'attente prioritaire d'une capacité finie  $K$ . En général, l'analyse de la file d'attente avec une capacité finie est plus complexe qu'un modèle correspondant à une capacité infinie. Falin [?] a étudié le cas particulier ( $K = 0$ ) et a obtenu la condition de stabilité et la fonction génératrice du nombre de clients dans la file d'attente. Lorsque  $K \geq 1$ , Choi et al. [?] ont obtenue la fonction génératrice du nombre de clients dans la file d'attente en utilisant la méthode de la variable supplémentaire est le temps de service restant de l'appel en service (résiduels).

### 2.5.4 La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels

Falin et al. [9] ont étendu le modèle de Choi et Park au modèle où les deux types d'appels ont des temps de service différents. Ils ont utilisé la méthode de la variable supplémentaire où complémentaire où la variable supplémentaire est le temps de service écoulé. Ils ont obtenu les mesures de performance suivantes pour ce système :

- (1) Probabilité que le serveur est occupé par un type d'appel I (resp. type II) est  $\rho_1$  (resp.  $\rho_2$ ).
- (2) Le nombre moyen d'appels de type I dans la file d'attente prioritaire est :

$$E[N_1] = \frac{\lambda_1(\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)})}{2(1 - \rho_1)}. \quad (2.11)$$

- (3) Le nombre moyen d'appels de type II en orbite :

$$E[N_2] = \frac{\lambda_2(\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)})}{2(1 - \rho_1)(1 - \rho_2)} + \frac{\lambda_2 \rho_2}{\nu(1 - \rho_2)}. \quad (2.12)$$

### 2.5.5 La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels et feedback

Choi et Lee ont étudié la file d'attente avec rappels et feedback à deux types d'appels où, après que le type d'appels I (resp. Type II) soit servi il rejoint la file prioritaire (resp. l'orbite) ou quitte le système définitivement.

Le phénomène de feedback dans le système de files d'attente avec rappels peut se produire dans beaucoup de situations pratiques, par exemple les systèmes de telecommunication, où un message d'erreur s'est avéré à destination et est à nouveau envoyé. En file d'attente avec rappels et feedback le type d'appels I reçu quitte le système avec une probabilité  $1 - \delta_1$  ou retourne à la file d'attente prioritaire pour être servi à nouveau avec une probabilité  $\delta_1$ . Le type d'appels II qui a reçu le service quitte le système avec probabilité  $1 - \delta_2$  ou rejoint l'orbite avec probabilité  $\delta_2$ .

Ils ont dérivé la fonction génératrice de la taille de la file d'attente en utilisant la méthode de la variable supplémentaire.

### 2.5.6 La file d'attente $M_1, M_2/G_1, G_2/1$ avec rappels à communication bi-directionnelle

Artalejo et Phung-Duc [2] ont étudié le comportement d'un état d'équilibre d'une  $M/G/1$  – *type* avec rappels dans laquelle il y'a deux flux d'arrivée entrants à savoir les appels effectués par des clients réguliers et les appels sortants effectués par le serveur lorsqu'il est inactif. Ils ont effectué une analyse stationnaire vaste du système considéré, y compris la condition de stabilité, la

chaîne de Markov induite, la distribution stationnaire de l'état du serveur et le nombre de clients dans l'orbite et le calcul des premiers moments. Ils ont aussi obtenu les résultats asymptotiques pour le nombre de clients en orbite.

### 2.6 Conclusion

Les modèles d'attente développés ces dernières décennies tentent de prendre en considération des phénomènes de répétition de demandes de service et de priorité à la fois. Dans ce chapitre, nous avons montré l'intérêt et les applications des modèles d'attente avec rappels et priorité.

# 3

## Généralités sur la théorie des inégalités stochastiques

### Introduction

Beaucoup d'efforts ont été consacrés à l'obtention de mesures de performance, tels que la longueur de la file d'attente, le temps d'attente, la distribution de la période d'occupation, ...

Toutefois, ces caractéristiques de performances ont été obtenues par des méthodes d'approximation qui ont donné des expressions très complexes et les résultats obtenus ne peuvent pas être utilisés en pratique.

Lors de la dernière décennie, il y a eu une tendance à la recherche d'approximations et les limites, et les propriétés qualitatives des modèles stochastiques constituent une importante base théorique des méthodes d'approximation.

La monotonie et la comparabilité sont les plus importantes propriétés qualitatives et méthodes d'approximation qui peuvent être étudié à l'aide de la théorie générale des ordres stochastiques,

voir [?, ?, ?, ?].

Ordonnement stochastique est utile pour étudier les changements internes de la performance en raison de variations de paramètres, à comparer des systèmes distincts, afin de se rapprocher d'un système complexe à un autre système simple, et pour obtenir des bornes supérieures et inférieures pour les principales mesures de performance d'un système.

### 3.1 Propriétés générales des ordres partiels

On appelle un ordre partiel, noté " $\prec$ ", une relation binaire définie sur un ensemble  $\mathcal{D}$  d'éléments  $a, b, c, \dots$ , satisfaisant les trois axiomes :

- (i)  $a \prec a$  (réflexivité),
- (ii) si  $a \prec b$  et  $b \prec c$  alors  $a \prec c$  (transitivité),
- (iii) si  $a \prec b$  et  $b \prec a$  alors  $a = b$  (antisymétrie).

Notons que  $a \prec b$  est équivalent à dire que  $b \succ a$ .

Cette section est consacrée à quelques propriétés de l'ordre partiel défini sur l'ensemble  $\mathcal{D}$  de toutes les fonctions de répartition de variables aléatoires réelles (ou bien l'un de ses sous-ensembles).

Pour les deux variables aléatoires  $X$  et  $Y$  de fonctions de répartition  $F$  et  $G$  (respectivement) on a par convention :

$$F \prec G \Leftrightarrow X \prec Y.$$

On suppose que deux variables aléatoires  $X$  et  $Y$  sont définies sur le même espace de probabilité, alors leurs fonctions de répartition respectives  $F$  et  $G$  peuvent satisfaire la propriété d'antisymétrie (iii) sans pour autant avoir  $X = Y$ .

Lorsque les variables aléatoires sont dégénérées, certaines propriétés des ordres partiels définies sur  $\mathcal{D}$  découlent directement des propriétés de l'ordre des nombres réels. Pour cela, on utilisera la distribution de Dirac, notée par  $\Theta_c(\cdot)$ , définie pour tous les nombres réels comme suit :

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

**Définition 3.1.** Soit un ordre partiel donné "  $\prec$  " défini sur (un sous ensemble de) l'espace  $\mathcal{D}$  des fonctions de répartition.

On dit que cet ordre possède la propriété :

- **(R)** : si  $\forall a, b \in \mathbb{R}$  tels que  $a \leq b$ , alors  $\Theta_a \prec \Theta_b$ .
- **(E)** : si  $F \prec G$ , alors  $m_F \leq m_G$  lorsque les moyennes existent.
- **(M)** : si  $F \prec G$ , alors  $F^c \prec G^c$ ,  $\forall c > 0$ , où  $F^c(x) = F(x/c)$ ,  $\forall x$ .
- **(C)** : si  $F_1 \prec F_2$  alors  $F_1 * G \prec F_2 * G$ , où  $(F_i * G)(x) = \int_{-\infty}^{+\infty} F_i(x-y)dG(y)$ ,  $i = 1, 2$ .
- **(W)** : si  $F_n$  et  $G_n$  convergent faiblement vers  $F$  et  $G$  (respectivement) alors :

$$\forall n, F_n \prec G_n \Rightarrow F \prec G.$$

**Remarque 3.1.** Pour les deux variables aléatoires  $X$  et  $Y$  :

La propriété (M) assure que :

$$X \prec Y \Leftrightarrow cX \prec cY \text{ pour tout } c \in ]0, +\infty[.$$

La propriété (C) assure que :

$$X_1 \prec X_2 \Rightarrow X_1 + Y \prec X_2 + Y,$$

où  $Y$  est une variable aléatoire indépendante de  $X_1$  et  $X_2$ .

La propriété (E) assure que :

$$X \prec Y \Rightarrow E(X) \leq E(Y).$$

On remarque que la propriété (E) découle des autres propriétés

**Proposition 3.1.** Un ordre partiel  $\prec$  sur un ensemble (ou bien sur un sous ensemble de)  $\mathcal{D}$  qui vérifie les propriétés (R), (M), (C) et (W), vérifie aussi la propriété (E).

**Définition 3.2.** Pour une classe de fonctions réelles  $\mathfrak{S}_\prec$ , l'ordre partiel  $\prec$  défini sur l'ensemble (ou sur le sous ensemble de)  $\mathcal{D}$  est dit généré par  $\mathfrak{S}_\prec$  si :

$$F \prec G \Leftrightarrow \int_{-\infty}^{+\infty} f(x)dF(x) \leq \int_{-\infty}^{+\infty} f(x)dG(x),$$

pour toute fonction  $f$  dans  $\mathfrak{S}_\prec$ , telle que les intégrales existent.

### 3.1.1 Ordre stochastique

**Définition 3.3.** On dit que la variable aléatoire  $X$  de fonction de répartition  $F$ , est stochastiquement inférieure (ou bien inférieure en distribution) à la variable aléatoire  $Y$  de fonction de répartition  $G$ , et on note  $F \leq_{st} G$ , lorsque

$$F(x) \geq G(x), \quad \forall x \in \mathbb{R}.$$

On écrit aussi  $X \leq_{st} Y$  ( $\leq_{st}$  noté aussi par l'ordre  $\leq_d$ ).

Dans le cas où  $X$  et  $Y$  sont des variables aléatoires discrètes prenant des valeurs sur l'ensemble des entiers relatifs  $\mathbb{Z}$ , et en notant par  $P_i^{(1)} = P\{X = i\}$  et  $P_i^{(2)} = P\{Y = i\}$  pour  $i \in \mathbb{Z}$ , alors

$$X \leq_{st} Y \Leftrightarrow \sum_{j=-\infty}^i P_j^{(1)} \geq \sum_{j=-\infty}^i P_j^{(2)}, \quad i \in \mathbb{Z},$$

ce qui est équivalent à :

$$\sum_{j=i}^{+\infty} P_j^{(1)} \leq \sum_{j=i}^{+\infty} P_j^{(2)}, \quad i \in \mathbb{Z}.$$

Remarquons que l'ordre stochastique  $\leq_{st}$  satisfait les axiomes de l'ordre partiel  $\prec$ .

**Proposition 3.2.** Si  $F_1 \leq_{st} F_2$ , alors il existe deux variables aléatoires  $X_1$  et  $X_2$  définies sur le même espace de probabilité  $(\Omega, \mathcal{A}, \mathcal{P})$  pour lesquelles

$$X_1(\omega) \leq X_2(\omega), \quad \forall \omega \in \Omega,$$

et

$$P(\omega : X_k(\omega) \leq x) = F_k(x) \quad \text{pour } k = 1, 2.$$

Notons par  $\mathfrak{R}_{st}(\mathbb{R})$  la classe des fonctions réelles non décroissantes, alors la classe  $\mathbb{R}_{\leq_{st}}$  des fonctions  $\leq$ -monotones est confondue avec la classe  $\mathfrak{R}_{st}(\mathbb{R})$ , c'est-à-dire  $\mathbb{R}_{\leq_{st}} = \mathfrak{R}_{st}(\mathbb{R})$ .

**Théorème 3.1.** L'inégalité

$$\int_{-\infty}^{+\infty} f(t) dF_1(t) \leq \int_{-\infty}^{+\infty} f(t) dF_2(t), \quad (3.1)$$

est vérifiée pour toute fonction  $f$  appartenant à  $\mathfrak{R}_{st}(\mathbb{R})$  pour laquelle l'intégrale existe, si et seulement si  $F_1 \leq_{st} F_2$ . Pour une fonction  $f$  donnée, l'inégalité (3.1) est vérifiée pour tout  $F_1$  et  $F_2$  telles que  $F_1 \leq_{st} F_2$  uniquement si  $f$  est non décroissante.

### 3.1.2 Ordre convexe

On note par  $x_+ = \max(0, x)$ .

**Définition 3.4.** On dit que la variable aléatoire  $X$ , de fonction de répartition  $F$ , est inférieure en moyenne de vie résiduelle à la variable aléatoire  $Y$ , de fonction de répartition  $G$ , et on écrit

$X \leq_v Y$ , ou bien  $F \leq_v G$  si et seulement si :

$$\begin{aligned} E((X - x)_+) &= \int_x^{+\infty} (t - x) dF(t) = \int_x^{+\infty} (1 - F(t)) dt \\ &\leq \int_x^{+\infty} (1 - G(t)) dt = E((Y - x)_+), \end{aligned} \quad (3.2)$$

lorsque les espérances (ou bien les intégrales) sont bien définies.

Dans le cas discret, on a :

$$X \leq_v Y \Leftrightarrow \sum_{i=k}^{\infty} \sum_{j=i}^{\infty} P_j^{(1)} \leq \sum_{i=k}^{\infty} \sum_{j=i}^{\infty} P_j^{(2)}.$$

Une conséquence immédiate de cette définition :

$$\text{si } F \leq_{st} G \text{ et } E(Y_+) < \infty \text{ alors } F \leq_v G.$$

**Théorème 3.2.** 1. L'inégalité

$$\int_{-\infty}^{+\infty} f(t) dF_1(t) \leq \int_{-\infty}^{+\infty} f(t) dF_2(t), \quad (3.3)$$

est vérifiée pour toute fonction  $f$  appartenant à  $\mathfrak{R}_v(\mathbb{R})$  pour laquelle les intégrales sont bien définies, si et seulement si  $F_1 \leq_v F_2$ .

2. Pour une fonction donnée  $f$ , l'inégalité (3.3) a lieu pour toutes les fonctions  $F_1$  et  $F_2$  telles que  $F_1 \leq_v F_2$  uniquement si  $f$  est une fonction convexe et non décroissante.
3. Si  $F_1 \leq_v F_2$  et leurs moyennes existent et sont égales, alors l'inégalité (3.3) est vérifiée pour toute fonction convexe  $f$  donnée.

**Corollaire 3.1.** Pour deux variables aléatoires  $X$  et  $Y$  non négatives telles que  $X \leq_v Y$  on a

$$E(X^r) \leq E(Y^r), \quad (r \geq 1),$$

lorsque les espérances existent.

En général, pour des variables aléatoires  $X$  et  $Y$  telles que

$$E(X) = E(Y), \text{ et } X \leq_v Y,$$

alors,

$$E(X^r) \leq E(Y^r), \quad (r = 2, 4, 6, \dots).$$

Il est intéressant de remarquer que pour deux variables aléatoires telles que  $X$  et  $Y$  sont non négatives et  $X \leq_v Y$ , alors l'égalité  $E(X^r) = E(Y^r)$  pour tout  $r \geq 1$  implique l'égalité  $X =_{st} Y$ .

En effet

$$E(X^r) = \int_0^{+\infty} r x^{r-1} (1 - F(x)) dx = \int_0^{+\infty} r(r-1) x^{r-2} dx \int_x^{+\infty} (1 - F(y)) dy.$$

Cette propriété est l'analogie de la propriété suivante pour l'ordre stochastique

$$X \leq_{st} Y \text{ et } E(X) = E(Y) \Rightarrow X =_{st} Y.$$

**Proposition 3.3.** Supposons que les suites de variables aléatoires  $X_n$  et  $Y_n$  convergent faiblement vers  $X$  et  $Y$  (respectivement).

Si

$$\begin{aligned} E(X_+) \text{ et } E(Y_+) \text{ sont finis,} \\ E((X_n)_+) \longrightarrow E(X_+) \text{ quand } n \longrightarrow +\infty, \\ E((Y_n)_+) \longrightarrow E(Y_+) \text{ quand } n \longrightarrow +\infty, \end{aligned}$$

et si  $X_n \leq_v Y_n$ , alors

$$X \leq_v Y.$$

### 3.1.3 Ordre en transformée de Laplace

#### Transformée de Laplace

Lorsque la variable aléatoire  $X$  est du type continu, sa distribution peut être caractérisée par la transformée de Laplace de la densité  $f(x)$  :

$$\hat{f}(s) = E(e^{-sX}) = \int_0^{\infty} f(x) e^{-sx} dx,$$

où  $s$  est une variable complexe. Cette intégrale est définie au moins pour  $Re(s) \geq 0$ . La transformée de Laplace est notée aussi  $L[f(x)]$ .

#### Propriétés

- Si  $X$  et  $Y$  sont indépendantes, la transformée de Laplace de  $X + Y$  est le produit des transformées de Laplace de  $X$  et de  $Y$ ,
- $L[f'(x)] = s\hat{f}(s) - f(0)$ ,

- $L[f''(x)] = s^2 \hat{f}(s) - sf(0) - f'(0)$ ,
- $L[\int_0^x f(u)du] = \frac{\hat{f}(s)}{s}$ ,
- Si  $F(x)$  est la fonction de répartition de  $X$  et si  $R(x) = 1 - F(x)$  alors

$$\lim_{s \rightarrow 0} \hat{R}(s) = \int_0^{\infty} R(x)dx.$$

**Définition 3.5.** Pour deux variables aléatoires non négatives  $X$  et  $Y$  de fonctions de répartition  $F$  et  $G$  (respectivement),  $F$  est dite inférieure par rapport à l'ordre laplacien à  $G$ , et on note  $F \leq_L G$ , si pour tout  $s$  positif on a l'inégalité suivante

$$E(\exp(-sX)) = \int_0^{+\infty} \exp(-sX)dF(x) \geq \int_0^{+\infty} \exp(-sX)dG(x) = E(\exp(-sY)).$$

Il est clair que l'ordre en transformée de Laplace est réflexif, transitif et antisymétrique.

**Théorème 3.3.** Soit une fonction  $f$  strictement monotone, alors  $F \leq_L G$  implique

$$\int_0^{+\infty} f(t)dF(t) \leq \int_0^{+\infty} f(t)dG(t).$$

**Corollaire 3.2.** 1. Pour deux variables aléatoires  $X$  et  $Y$  non négatives, de fonctions de répartition  $F$  et  $G$  respectivement, telles que  $F \leq_L G$  alors, on a l'inégalité suivante :

$$\frac{1 - E(\exp(-sX))}{s} \leq \frac{1 - E(\exp(-sY))}{s}, \quad \forall s > 0.$$

2. Lorsqu'on fait tendre  $s$  vers 0, on obtient le résultat suivant :

$$F \leq_L G \Rightarrow E(X) \leq E(Y),$$

lorsque les espérances existent.

Le résultat qui suit donne une caractérisation de l'ordre en transformée de Laplace.

**Théorème 3.4.** Soient  $X$  et  $Y$  deux variables aléatoires quelconques de fonctions de répartition  $F$  et  $G$  respectivement, alors :

$$F \leq_L G \Leftrightarrow E(f(x)) \leq E(f(y)),$$

pour toute fonction  $f$  strictement monotone, telle que les espérances existent.

### 3.1.4 Ordre en fonctions génératrices

Soient  $X$  et  $Y$  deux variables aléatoires non négatives discrètes de fonctions de répartition  $F$  et  $G$  respectivement. On dit que  $X$  est inférieure à  $Y$  par rapport à l'ordre en fonctions génératrices, et on note  $F \leq_g G$ , si et seulement si :

$$E(z^X) \geq E(z^Y),$$

où,

$$E(z^X) = \sum_{n=0}^{+\infty} P(X = n)z^n \text{ et } E(z^Y) = \sum_{n=0}^{+\infty} P(Y = n)z^n, \quad |z| < 1.$$

Cet ordre peut-être déduit de l'ordre laplacien en posant  $s = -\ln z$ .

### 3.1.5 Relations entre les ordres partiels

Soient  $X$  et  $Y$  deux variables aléatoires de fonctions de répartition  $F$  et  $G$  respectivement. Alors, on a les relations suivantes :

- Si  $F \leq_{st} G$  et  $E(Y_+) < \infty \Rightarrow F \leq_v G$ .
- Si  $E(X) = E(\max(0, -x)) < \infty \Rightarrow F \leq_{cv} G$ .
- Si  $E(X) = E(Y)$ , alors  $F \leq_{cv} G \Leftrightarrow G \leq_v F$ .
- $F \leq_{st} G \Rightarrow F \leq_L G \Rightarrow F \leq_g G$ .
- $F \leq_{cv} G \Rightarrow F \leq_L G \Rightarrow F \leq_g G$ .
- Si  $E(X) = E(Y)$  et  $F \leq_v G \Rightarrow G \leq_L F \Rightarrow G \leq_g F$ .
- $F \leq_L G \Rightarrow F \leq_g G$ .

### 3.1.6 Propriétés de monotonie

Étudier mathématiquement les modèles stochastiques, c'est d'obtenir des estimations des quantités qui, pour un modèle  $\Sigma$  donné, avec une structure spécifique et des distributions  $F_i$  des  $X_i, \dots$ , décrivent son comportement.

Soit  $c_\Sigma$  une caractéristique dans  $\Sigma$  et soit  $C_\Sigma$  l'ensemble des valeurs possibles de  $c_\Sigma$ .

Pour une structure donnée et une distribution initiale  $U$ ,  $c_\Sigma$  dépend uniquement des  $F_i$ , et on écrit

$$c_\Sigma = c_\Sigma(F_1, F_2, \dots) \in C_\Sigma.$$

Pour des modèles simples, on peut déduire une expression explicite de  $c_\Sigma$ . Cependant, dans plusieurs situations, cela n'est pas possible et les calculs mathématiques peuvent mener à des

formules compliquées qui ne peuvent pas être exploitées en pratique.

De telles circonstances nous suggèrent de rechercher les propriétés qualitatives de  $c_\Sigma$  par rapport aux  $F_i$ , i.e, la manière avec laquelle  $c_\Sigma$  est affectée par les changements en  $F_i$ . Parmi les propriétés qualitatives importantes des modèles stochastiques on trouve la monotonie (i.e, si les  $F_i$  croissent dans un certain sens, alors  $c_\Sigma$  croît aussi).

#### Monotonie interne :

Soit  $\Sigma$  un modèle stochastique constitué de distributions paramétriques  $(U, F_1, F_2, \dots) \equiv (U, F)$ , où  $U$  est la distribution initiale.

On note par  $c_\Sigma$  les indices de performance du système  $\Sigma$ .

Par exemple, pour un système de files d'attente  $\Sigma$ ,  $c_\Sigma$  peut-être le temps moyen d'attente virtuel à l'instant  $t$ , ou la distribution de probabilité du nombre de clients dans le système à l'instant  $t$ , ou bien ses mesures de performance seront calculées à une suite d'instants  $(t_n)_{n \in \mathbb{N}^*}$  (déterministe ou aléatoire).

D'une manière plus précise, on peut exprimer  $c_\Sigma$  comme suit :

$$\begin{aligned} c_\Sigma(t) &= c_\Sigma(t, U, \{F_i\}), \text{ ou bien} \\ c_\Sigma(t_n) &= c_\Sigma(n) = c_\Sigma(n, U, \{F_i\}). \end{aligned}$$

On note par  $\prec$  l'ordre partiel défini sur  $C_\Sigma$ .

**Définition 3.6.** L'indice de performance  $c_\Sigma(\cdot)$  est non décroissant (resp. non croissant) par rapport à la distribution initiale  $U$  si pour tout  $t < u$ , on a :

$$t < u \Rightarrow c_\Sigma(t) \prec c_\Sigma(u) \text{ (resp. } c_\Sigma(t) \succ c_\Sigma(u)), \quad (3.4)$$

ou bien pour les entiers  $m < n$  :

$$m < n \Rightarrow c_\Sigma(m) \prec c_\Sigma(n) \text{ (resp. } c_\Sigma(m) \succ c_\Sigma(n)).$$

Cette propriété est appelée monotonie interne. D'autres appellations sont utilisées telles que, monotonie temporelle ou intrinsèque. Celles-ci découlent du fait que cette monotonie ne dépend en aucun cas des distributions paramétriques  $\{F_i\}$ , mais seulement peut-être de la distribution initiale.

#### Monotonie externe :

On note par  $\mathcal{D}_k$  l'ensemble des distributions  $F_k$  partiellement ordonnées par l'ordre " $<_k$ " (qui est l'ordre de la  $k^{\text{ème}}$  distribution paramétrique), et soit " $<_c$ " l'ordre partiel défini sur  $C_\Sigma$ .

**Définition 3.7.** L'indice de performance  $c_\Sigma$  est non décroissant sur  $\mathcal{D}_k$  par rapport à l'ordre  $<_k$  si pour tout  $F$  et  $G$  dans  $\mathcal{D}_k$  et toute autre distribution paramétrique constante, on a :

$$F <_k G \Rightarrow c_\Sigma(F_1, \dots, F_{k-1}, F, F_{k+1}, \dots) <_c (F_1, \dots, F_{k-1}, G, F_{k+1}, \dots).$$

Cette propriété est appelée monotonie externe.

Lorsque un système possède la propriété de monotonie externe, les indices de performance des modèles stochastiques, possédant la même structure avec des distributions paramétriques comparables mais différentes, sont comparables.

On peut interpréter la monotonie externe comme suit :

Soient  $\Sigma_1$  et  $\Sigma_2$  deux modèles stochastiques ayant la même structure et la même distribution initiale. On dira que ces modèles possèdent la propriété de monotonie externe lorsque pour deux distributions paramétriques  $F$  et  $G$  dans  $\Sigma_1$  et  $\Sigma_2$  respectivement, on a :

$$F \prec G \Rightarrow c_\Sigma(F) <_c c_\Sigma(G),$$

pour l'indice de performance  $c_\Sigma$ .

La monotonie externe est un outil d'une grande importance dans la construction des bornes pour les mesures de performance d'un système donné. Ainsi, la distribution paramétrique  $F_k$  peut être bornée par les distributions  $G_1$  et  $G_2$  appartenant à l'ensemble  $\mathcal{D}_k$  pour lesquelles :

$$G_1 <_k F_k <_k G_2,$$

alors pour les mesures de performance correspondantes, on obtient :

$$c_\Sigma(G_1) <_k c_\Sigma(F) <_k c_\Sigma(G_2),$$

lorsque les systèmes ont la propriété de la monotonie externe.

## 3.2 Comparabilité et monotonie des processus markoviens homogènes

### 3.2.1 Opérateurs monotones et comparables

Soient  $(\mathbb{E}, \mathbb{M})$  un espace probabilisable et  $P_{\mathbb{M}}$  l'ensemble de toutes les mesures de probabilité définies sur  $\mathbb{M}$ . Soient aussi les opérateurs  $\mathcal{T}$ ,  $\mathcal{T}^{(1)}$  et  $\mathcal{T}^{(2)}$  définis de  $P_{\mathbb{M}}$  dans  $P_{\mathbb{M}}$  et l'ordre partiel " $\prec$ " défini sur  $P_{\mathbb{M}}$ .

**Définition 3.8.** Un opérateur  $\mathcal{T}$  est dit  $\prec$ -monotone si pour toutes mesures de probabilités  $p^{(1)}$ ,  $p^{(2)}$  appartenant à  $P_{\mathbb{M}}$  telles que  $p^{(1)} \prec p^{(2)}$ , on a

$$\mathcal{T}p^{(1)} \prec \mathcal{T}p^{(2)}.$$

L'opérateur  $\mathcal{T}^{(1)}$  est inférieur à  $\mathcal{T}^{(2)}$  si  $\mathcal{T}^{(1)}p \prec \mathcal{T}^{(2)}p$  pour tout  $p \in P_{\mathbb{M}}$  et on écrit,

$$\mathcal{T}^{(1)} \prec \mathcal{T}^{(2)}.$$

Pour des applications aux processus de Markov homogènes, on s'intéresse à la comparabilité des distributions  $p_n^{(1)}$  et  $p_n^{(2)}$  définies par

$$p_n^{(k)} = (\mathcal{T}^{(k)}p^{(k)})_n, \quad k = 1, 2 \text{ et } n \in \mathbb{N}^*,$$

pour deux distributions initiales  $p^{(k)}$  et les opérateurs  $\mathcal{T}^{(k)}$ , pour  $k = 1, 2$ .

**Théorème 3.5.** Soient  $\mathcal{T}^{(1)}$ ,  $\mathcal{T}^{(2)}$  deux opérateurs définis sur  $P_{\mathbb{M}}$  et  $p^{(1)}$ ,  $p^{(2)}$  deux mesures de probabilité définies sur  $\mathbb{M}$ , alors

$$p^{(1)} \prec p^{(2)} \quad \text{implique} \quad p_n^{(1)} \prec p_n^{(2)}, \quad \forall n \in \mathbb{N}^*,$$

s'il existe un opérateur  $\mathcal{T}$   $\prec$ -monotone défini sur  $P_{\mathbb{M}}$  tel que

$$\mathcal{T}^{(1)} \prec \mathcal{T} \prec \mathcal{T}^{(2)}.$$

Remarquons que ce théorème reste vrai, en général, pour les opérateurs définis dans un espace partiellement ordonné.

À présent, on considère les opérateurs de transition d'une chaîne de Markov homogène  $(X_n)_{n \geq 1}$  d'espace d'état  $(\mathbb{E}, \mathbb{M})$ . Les opérateurs de transition sont décrits par leurs fonctions de transition  $p(x, \mathcal{B})$ ,

$$p(x, \mathcal{B}) = P(X_{n+1} \in \mathcal{B} / X_n = x), \quad x \in \mathbb{E} \text{ et } \mathcal{B} \in \mathbb{M},$$

ou bien, dans le cas où les processus sont à valeurs réelles, par leurs distributions de transition

$$p(x, y) = P(X_{n+1} < y / X_n = x), \quad x, y \in \mathbb{E} \subseteq \mathbb{R}.$$

Maintenant, on donne des conditions sur les fonctions de transition, qui assurent la monotonie ou la comparabilité des opérateurs de transition.

**Théorème 3.6.** Les opérateurs de transition  $\mathcal{T}^{(1)}$  et  $\mathcal{T}^{(2)}$  satisfont l'inégalité  $\mathcal{T}^{(1)} \prec \mathcal{T}^{(2)}$  si et seulement si leurs fonctions  $p^{(1)}$  et  $p^{(2)}$  satisfont

$$p^{(1)}(x, \cdot) \prec p^{(2)}(x, \cdot), \quad \forall x \in \mathbb{E}.$$

### 3.2.2 Conditions de monotonie et de comparabilité

Pour l'étude de la monotonie et de la comparabilité des chaînes de Markov homogènes, on peut énoncer les deux théorèmes suivants qui constituent un outil important pour prouver la monotonie interne et/ou externe de ces modèles stochastiques.

**Théorème 3.7.** Une chaîne de Markov homogène  $(X_n)_{n \geq 1}$ , de fonction de transition  $p$ , est non décroissante (resp. non croissante) par rapport à l'ordre partiel " $\prec$ " si :

$$X_1 \prec X_2 \quad (\text{resp. } X_2 \prec X_1),$$

et si  $p$  est  $\prec$ -monotone.

**Théorème 3.8.** Deux chaînes de Markov homogènes  $(X_n)_{n \geq 1}$  et  $(Y_n)_{n \geq 1}$ , de fonctions de transition  $p^{(1)}$  et  $p^{(2)}$  respectivement, satisfont l'inégalité :

$$X_n \prec Y_n, \quad \forall n \in \mathbb{N}^*,$$

si  $X_1 \prec X_2$  et s'il existe une fonction de transition  $p$   $\prec$ -monotone telle que :

$$p^{(1)}(x, \cdot) \prec p(x, \cdot) \prec p^{(2)}(x, \cdot), \quad \forall x \in \mathbb{E}.$$

## 3.3 Distributions non-paramétriques

Les notions de vieillissement et de relations d'ordre entre variables aléatoires sont étroitement liées. Nous présentons les principaux ordres permettant de comparer des variables aléatoires puis les notions de vieillissement. Cette présentation sera cependant partielle car l'activité scientifique sur ces sujets est importante. Il est donc difficile de prétendre faire un exposé exhaustif. L'un des états de l'art les plus récents sur ce sujet est [?, ?], mais on peut citer aussi [?, ?, ?].

En théorie de fiabilité, les classes de distributions nous renseignent sur la notion de jeunesse ou de vieillesse du système du point de vue de sa durée de vie résiduelle connaissant son âge (propriété qualitative). La connaissance de la classe (d'âge) de la loi de fiabilité d'un équipement permet une aide à la décision.

Dans cette section, sont présentées aussi les principales classes de distributions de survie recensées dans la littérature de fiabilité ces dernières années.

Les distributions non-paramétriques ont été introduites pour l'étude de certains problèmes en relation avec la théorie de fiabilité. Elles permettent ainsi de modéliser et caractériser des

propriétés qualitatives telles que le vieillissement et le rajeunissement du système.

Ces distributions sont utilisées actuellement dans divers domaines de la modélisation stochastique : analyse de survie (médecine), files d'attente, ordonnancement, théorie de décision, économie, gestion des stocks [?].

**Définition 3.9.** Soient  $X$  et  $X_\tau$  des variables aléatoires représentant respectivement la durée de vie et la durée de vie résiduelle d'un élément, et soient  $F$  et  $F_\tau$  leurs distributions respectives. On dit que  $F$  est :

- ◆ *NBU* (New Better than Used), si  $F_\tau \leq_{st} F$ , ( $0 < \tau < \infty$ ).
- ◆ *NWU* (New Worse than Used), si  $F \leq_{st} F_\tau$ , ( $0 < \tau < \infty$ ).
- ◆ *NBUE* (New Better than Used in Expectation), si  $E(X_\tau) \leq E(X)$ , ( $0 < \tau < \infty$ ).
- ◆ *NWUE* (New Worse than Used in Expectation), si  $E(X) \leq E(X_\tau)$ , ( $0 < \tau < \infty$ ).
- ◆ *IFR* (Increasing Failure Rate), si  $F_y \leq_{st} F_x$ , ( $0 \leq x < y < \infty$ ).
- ◆ *IFRA* (Increasing Failure Rate in Average), si  $(-1/t) \log(1 - F(t))$  croissante,  $t \geq 0$ .
- ◆ *DFR* (Decreasing Failure Rate), si  $F_x \leq_{st} F_y$ , ( $0 \leq x < y < \infty$ ).
- ◆ *DFRA* (Decreasing Failure Rate in Average), si  $(-1/t) \log(1 - F(t))$  décroissante,  $t \geq 0$ .
- ◆ *IMRL* (Increasing Mean Residual Life), si

$$E(X_\tau) = \frac{1}{1 - F(x)} \int_{\tau}^{+\infty} (1 - F(u)) du, \quad \text{croissante} \quad (0 < \tau < \infty).$$

- ◆ *DMRL* (Decreasing Mean Residual Life), si

$$E(X_\tau) = \frac{1}{1 - F(x)} \int_{\tau}^{+\infty} (1 - F(u)) du, \quad \text{décroissante} \quad (0 < \tau < \infty).$$

**Proposition 3.4.** Soit la variable aléatoire  $X$  de fonction de répartition  $F$  ayant une moyenne finie  $m$ .

1. Si  $F$  est *NBU* (resp. *NWU*), alors :

$$F \leq_{st} \exp(\lambda), \quad (\text{resp. } F \geq_{st} \exp(\lambda)),$$

pour un certain  $\lambda \leq m^{-1}$  (resp.  $\lambda \geq m^{-1}$ ), avec la possibilité d'avoir une égalité seulement si  $F = \exp(m^{-1})$ .

2. Si  $F$  est *NBUE* (resp. *NWUE*), alors

$$F \leq_v \exp(m^{-1}), \quad (\text{resp. } F \geq_v \exp(m^{-1})).$$

### 3.3.1 Relation avec les distributions paramétriques

- La loi d'Erlang  $E_k$  est *IFR*.
- La loi de Weibull  $W(a)$ , pour  $a > 1$  (paramètre de la forme), est *IFR*.
- La loi de Weibull  $W(a)$ , pour  $a \leq 1$ , est *DFR*.
- La loi Gamma  $\Gamma(a)$ , avec  $0 \leq a < 1$ , est *DFR*.
- La loi exponentielle est à la fois *IFR* et *DFR*.
- La distribution Hyper-exponentielle  $H$  est *DFR*.

### 3.3.2 Relation entre les classes de distributions non-paramétriques

La figure 3.1 illustre les relations d'implication existantes entre certaines classes de distributions non-paramétriques.

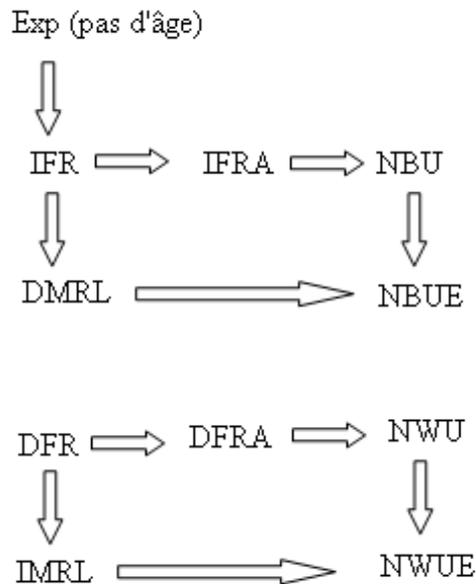


FIGURE 3.1 – Relations entre les classes de distributions d'âge

## Conclusion

On a présenté quelques concepts de base de la théorie des ordres stochastiques et la monotonie des processus stochastiques. On a donné aussi les classes de distributions d'âge issues de la théorie de la fiabilité. La méthode de comparaison stochastique sera appliquée, dans le chapitre suivant, au modèle d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle.

# 4

## Inégalités stochastiques pour le système d'attente $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidirectionnelle

### Introduction

Vu la complexité des réseaux et des services d'aujourd'hui (en terme de nombre de noeuds, d'événements : pannes/réparation, arrivée/services), il devient difficile d'effectuer une analyse quantitative de ces systèmes. Il est ainsi indispensable de faire appel à des méthodes efficaces d'évaluation des performances permettant de faire face à la complexité de ces systèmes. C'est pourquoi une partie importante de ce chapitre est focalisée sur l'application des méthodes de comparaison stochastiques pour étudier les propriétés de monotonie du modèle d'attente avec rappels à communication bidirectionnelle en s'inspirant des résultats obtenus par Boualem et al. [6, 7] Stoyan [?], Schantikumar [?], Khalil et Falin [?], Liang [?] et Liang et Kulkarni [?].

Khalil et Falin (1994) [?] ont dérivé certaines propriétés de monotonie d'une file d'attente  $M/G/1$  avec rappels avec des temps de rappels exponentiels et politique de rappels classique par rapport aux ordres stochastique, convexe et en transformée de Laplace. Ils ont montré que le nombre de clients diminue stochastiquement lorsque le taux d'arrivée décroît avec l'augmentation du taux de rappels et de diminuer stochastiquement ou par rapport à l'ordre convexe le temps de service. Les inégalités sont établies pour les caractéristiques moyennes de la période d'occupation et le nombre de clients servis durant cette période.

Liang et Kulkarni (1993) [?] ont étudié les propriétés de monotonie des files d'attente avec rappels, afin d'étudier comment la répartition des temps de rappels affecte le comportement du système. Ils supposent que les temps de rappels ont une distributions de type phase et montrent que les systèmes avec un temps de rappels plus grand, par rapport à la  $K$ -dominance, crée plus de clients dans le système et dans l'orbite. A partir de ces résultats, ils ont tiré des propriétés de monotonie de plusieurs mesures du performance.

Shin et Kim (2000) [?] considèrent les files d'attente Markovienne avec rappels en montrant la relation par rapport à l'ordre stochastique entre deux processus à deux variables représentant le nombre de clients dans le service et en orbite par la construction des processus équivalents sur un espace de probabilité commun.

Shin (2006) [?] considère plusieurs modèles de files d'attente multi-serveurs avec rappels exponentiels comme  $A^X/G/c/K$  avec rappels, la file d'attente à deux nœuds  $A^X/G/c_1/K_1 \rightarrow .G/c_2/K_2$ ,  $MAP_1$ ,  $MAP_2/M/c$  avec rappels et  $M/M/c/c$  avec rappels et arrivées négatives. Il a montré que l'opérateur de transition est monotone si le taux de rappels dans un système est borné par le taux de rappels dans le second. Les résultats de monotonie sont appliqués pour montrer la convergence des systèmes tronqués généralisées à l'origine.

Boualem et al. (2009) [5] ont étudié les conditions de comparabilité pour certaines mesures de performance d'un système de files d'attente  $M/G/1$  avec rappels et vacance du serveur, en utilisant la théorie générale des ordres stochastiques.

Taleb et Aissani (2010) [?] ont étudié la propriété de monotonie pour une file d'attente non fiable  $M/G/1$  avec une distribution générale du temps inter-rappels, perte géométrique et de discipline de l'orbite FIFO. Ils ont montré la monotonie de l'opérateur de transition de la chaîne de Markov induite par rapport aux ordres stochastique et convexe. De plus, ils ont obtenus des conditions de comparabilité des deux opérateurs de transition et des conditions de comparabilité de nombre de clients dans le système. Les inégalités sont établies pour les caractéristiques moyennes de la période d'occupation, le nombre de clients servis au cours d'une période d'occupation, le nombre de clients en orbite. Les inégalités sont également obtenues pour les proba-

bilités stationnaires.

Boualem et al. (2012) [7] ont utilisé la méthode de comparaison stochastique de chaînes de Markov pour obtenir quelques approximations qualitatives pour le système  $M/G/1$  avec rappels et Bernoulli feedback. L'objectif principal est d'employer les techniques d'ordres stochastiques pour établir différents résultats de monotonie par rapport aux taux d'arrivée, distributions du temps de service et le paramètre de rappel.

### 4.1 Analyse du système $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidirectionnelle [2]

Nous classons les files d'attente avec rappels en deux catégories selon le nombre d'appels différents, modèle avec un type unique d'appels et le modèle dans lequel il y'a deux flux d'arrivées entrants à savoir les appels effectués par des clients réguliers et les appels sortants effectués par le serveur lorsqu'il est inactif.

L'exemple le plus simple et évident est fourni par un centre d'appels où les serveurs ont une chance d'effectuer des appels téléphoniques sortants alors qu'ils ne sont pas engagés dans une conversation. Cette fonction est connue sous le nom de files d'attente à commutation couplé ou des modèles de deux communications à sens unique ou à communication bidirectionnelle.

#### 4.1.1 Description du modèle

Dans notre travail, nous considérons le modèle  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle [2]. Le choix de ce modèle est motivé par des applications réelles, par exemple : les centres d'appels où un opérateur non seulement sert les appels entrants, mais il effectue aussi des appels sortants vers l'extérieur lorsque le serveur est libre.

Nous considérons un système de files d'attente à un seul serveur auquel les clients primaires entrants arrivent selon un processus de Poisson de taux  $\lambda$ . En outre, si le serveur est libre, alors il génère un appel sortant dans un temps exponentiellement distribué avec un taux  $\alpha$ . Nous supposons que les appels entrants et les appels sortants reçoivent des temps de service différents.  $B_1(x)$  ( $B_1(0) = 0$ ) représente la distribution du temps de service d'un appel entrant, alors que  $B_2(x)$  ( $B_2(0) = 0$ ) désigne la distribution du temps de service d'un appel sortant. Un appel entrant qui trouve le serveur occupé rejoint l'orbite et retente d'entrer dans la zone de service selon une distribution exponentielle avec un taux  $\mu$ . Si le nombre de clients en orbite à l'instant  $t$  ( $N(t) = j$ ), alors le taux de rappels est  $j\mu$  (rappels classique). Également désigner la transformée

de Laplace-Stieltjes et le  $k^{\text{ième}}$  moment de  $B_l(x)$  comme  $\beta_l(s)$  et  $\beta_l^k$  (resp.), pour  $l = 1, 2$  et  $k \in \mathbb{Z}_+$ , où  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ . L'arrivée de flux d'appels entrants et sortants, temps de service et les inter-rappels mutuellement indépendants.

L'état du système à l'instant  $t$  peut être décrit par le processus  $Y(t) = (C(t), N(t), \xi(t))$ ,  $t \geq 0$ , où :

$$C(t) = \begin{cases} 0, & \text{si le serveur est inactif à l'instant } t; \\ 1, & \text{si le serveur est occupé par un appel entrant à l'instant } t; \\ 2, & \text{si le serveur lance un appel vers l'exterieur à l'instant } t. \end{cases}$$

Si  $C(t) \in \{1, 2\}$ , alors  $\xi(t)$  représente le temps écoulé du service de l'appel (entrant ou sortant) en cours. Dans ce qui suit, nous négligeons  $\xi(t)$  et ne considérons que la paire  $X(t) = (C(t), N(t))$  dont l'espace est  $S = \{0, 1, 2\} \times \mathbb{Z}_+$ .

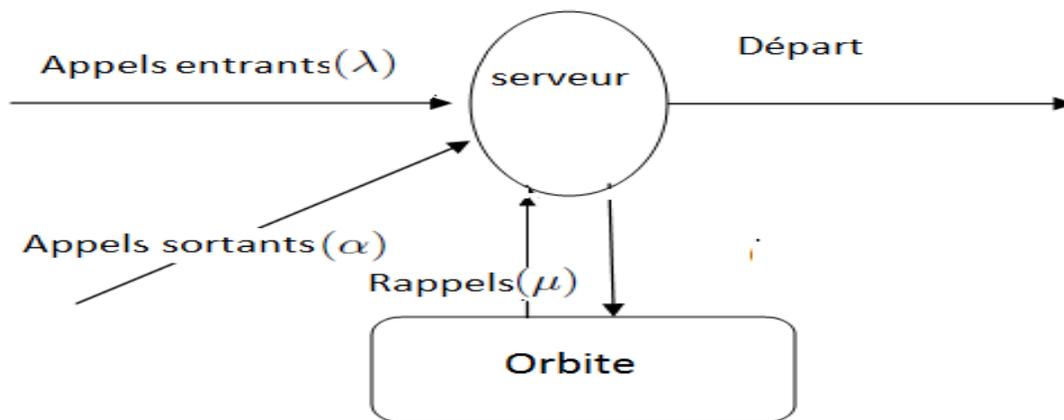


FIGURE 4.1 – File d'attente avec rappels à communication bidirectionnelle.

Les transitions du modèle considéré, sont représentées dans la Figure 4.2. Nous décrivons la dynamique du modèle en détail dans le Tableau 4.1.

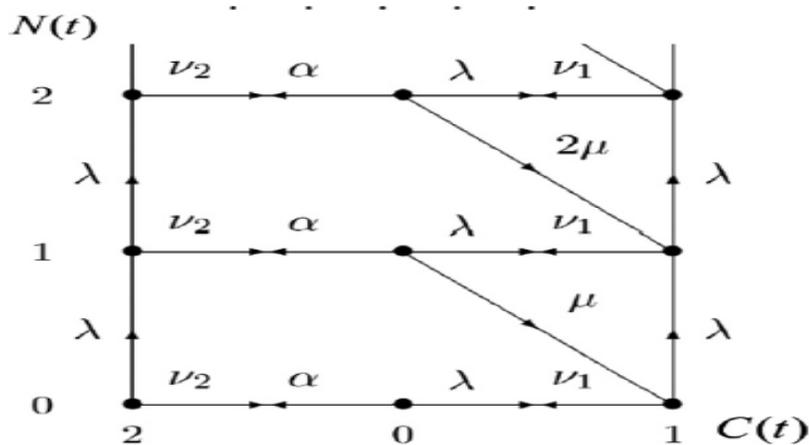


FIGURE 4.2 – Les transitions entre les états du modèle.

Evénements	Transitions	Taux
Appels entrants non bloqués	$(0, j) \rightarrow (1, j)$	$\lambda$
Appels entrants bloqués	$(i, j) \rightarrow (i, j + 1), i = 1, 2$	$\lambda$
Appels sortants	$(0, j) \rightarrow (2, j)$	$\alpha$
Service d'un appel entrant	$(1, j) \rightarrow (0, j)$	$\nu_1$
Service d'un appel sortant	$(2, j) \rightarrow (0, j)$	$\nu_2$
Rappels	$(0, j) \rightarrow (1, j - 1), j > 0$	$j\mu$

TABLE 4.1 – Les taux de transition de la file d'attente avec rappels à communication bidirectionnelle

### 4.1.2 Chaîne de Markov induite [2] :

Nous considérons d'abord les époques auxquelles le temps de service demandé par un appel arbitraire est expiré.

Soit  $\eta_n$  l'instant de fin de service du  $n^{\text{ième}}$  client.

La suite  $Z_n = N(\eta_n^+)$  forme une chaîne de Markov qui est le processus de renouvellement du processus de Markov à temps continu  $Y(t)$ . On remarque que  $\{Z_n\}_{n=1}^\infty$  vérifie l'équation fondamentale suivante :

$$Z_n = Z_{n-1} - W_n + V_n,$$

où  $V_n$  est le nombre d'arrivées entrantes pendant le  $n^{\text{ième}}$  temps de service , et

$$W_n = \begin{cases} 1, & \text{si le } n^{\text{ième}} \text{ client en service provient de l'orbite,} \\ 0, & \text{sinon.} \end{cases}$$

Il est clair que  $\{Z_n\}_{n=1}^{\infty}$  est une chaîne de Markov à temps discret qui est irréductible et apériodique (CMTD).

$\rho = \lambda\beta_1^1 < 1$  est la condition nécessaire et suffisante d'ergodicité.

La matrice de transition  $P = (p_{ij})$  a une structure de la matrice d'une file  $M/G/1$ , voir la Figure 4.3, avec les éléments :

$$p_{ij} = \begin{cases} \frac{i\mu}{\lambda+\alpha+i\mu} k_0^1, & \text{si } i \geq 1, j = i - 1, \\ \frac{\lambda}{\lambda+\alpha+i\mu} k_{j-i}^1 + \frac{\alpha}{\lambda+\alpha+i\mu} k_{j-i}^2 + \frac{i\mu}{\lambda+\alpha+i\mu} k_{j-i+1}^1, & \text{si } 0 \leq i \leq j. \end{cases}$$

$$k_j^l = \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^j}{j!} dB_l(x), \quad l = 1, 2, \quad j \in \mathbb{Z}_+,$$

où  $k_j^l$  représente la probabilité qu'il y ait  $j$  arrivées entrantes pendant un temps de service avec la fonction de distribution  $B_l(x)$ ,  $l = 1, 2$ . La dérivation des expressions ci-dessus pour les probabilités de transition en une seule étape  $p_{ij}$  est fondée sur un argument de la première étape dans lequel on conditionne sur l'identité du  $n^{\text{ième}}$  client en service.

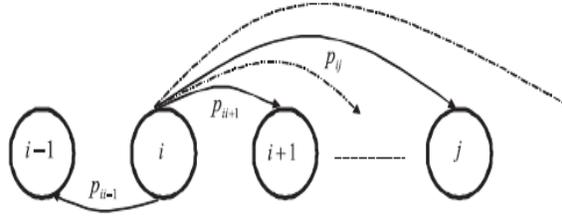


FIGURE 4.3 – Les transitions entre les états  $i > 0$ .

Maintenant, nous supposons que  $\rho < 1$  pour garantir que les états sont récurrents non nuls. Ainsi, les probabilités limites :

$$\pi_j = \lim_{n \rightarrow \infty} \{Z_n = j\}, \quad j \in \mathbb{Z}_+$$

existent et sont positives.

**Théorème 4.1.** La fonction génératrice partielle  $\Pi(z) = \sum_{j=0}^{\infty} z^j \pi_j$  est donnée par :

$$\begin{aligned} \Pi(z) &= \frac{1 - \rho}{\lambda(1 + \sigma) + \alpha(1 - \rho)} \frac{\lambda(1 - z)\beta_1(\lambda - \lambda z) + \alpha(\beta_1(\lambda - \lambda z) - z\beta_2(\lambda - \lambda z))}{\beta_1(\lambda - \lambda z) - z} \\ &\quad \times \exp \left\{ \frac{-1}{\mu} \int_z^1 \frac{\lambda(1 - \beta_1(\lambda - \lambda u)) + \alpha(1 - \beta_2(\lambda - \lambda u))}{\beta_1(\lambda - \lambda u) - u} du \right\}. \end{aligned} \quad (4.1)$$

**Preuve.** De la matrice de transition  $P = (p_{ij})_{i,j \geq 0}$ , nous avons

$$\pi_n = \sum_{m=0}^n \pi_m \frac{\lambda k_{n-m}^1 + \alpha k_{n-m}^2}{\lambda + \alpha + m\mu} + \sum_{m=1}^{n+1} \pi_m \frac{m\mu}{\lambda + \alpha + m\mu} k_{n-m+1}^1, \quad n \geq 0. \quad (4.2)$$

l'équation (4.2) peut être transformée à l'aide des fonctions génératrices auxiliaires suivantes :

$$\begin{aligned} \Phi(z) &= \sum_{j=0}^{\infty} z^j \frac{\pi_j}{\lambda + \alpha + j\mu}, \\ K_l(z) &= \sum_{j=0}^{\infty} z^j k_j^l, \quad l = 1, 2, \end{aligned}$$

où  $K_l(z) = \beta_l(\lambda - \lambda z)$ , pour  $l = 1, 2$ . En introduisant les fonctions génératrices, l'équation (4.2) devient :

$$\Pi(z) = (\lambda\beta_1(\lambda - \lambda z) + \alpha\beta_2(\lambda - \lambda z))\Phi(z) + \mu\beta_1(\lambda - \lambda z)\Phi'(z). \quad (4.3)$$

D'autre part, nous avons aussi la relation :

$$\Pi(z) = (\lambda + \alpha)\Phi(z) + \mu z\Phi'(z). \quad (4.4)$$

En combinant (4.3) et (4.4), on obtient l'équation différentielle suivante pour la fonction génératrice  $\Phi(z)$  :

$$\mu(\beta_1(\lambda - \lambda z) - z)\Phi'(z) = (\lambda(1 - \beta_1(\lambda - \lambda z)) + \alpha(1 - \beta_2(\lambda - \lambda z)))\Phi(z), \quad (4.5)$$

et

$$\Pi(z) = \frac{\lambda(1 - z)\beta_1(\lambda - \lambda z) + \alpha(\beta_1(\lambda - \lambda z)) - z\beta_2(\lambda - \lambda z)}{\beta_1(\lambda - \lambda z) - z}\Phi(z). \quad (4.6)$$

Comme  $\Pi(1) = 1$ , l'équation (4.6) nous donne :

$$\Phi(1) = \frac{1 - \rho}{\lambda(1 + \sigma) + \alpha(1 - \rho)}.$$

La résolution de l'équation différentielle (4.5) donne :

$$\Phi(z) = \frac{1 - \rho}{\lambda(1 + \sigma) + \alpha(1 - \rho)} \exp \left\{ \frac{-1}{\mu} \int_z^1 \frac{\lambda(1 - \beta_1(\lambda - \lambda u)) + \alpha(1 - \beta_2(\lambda - \lambda u))}{\beta_1(\lambda - \lambda u) - u} du \right\}. \quad (4.7)$$

Après quelques manipulations algébriques, on obtient le résultat donné dans le corollaire suivant :

**Corollaire 4.1.** Le nombre moyen de clients en orbite aux instants de fin de service est :

$$\Pi'(1) = \sum_{j=1}^{\infty} j\pi_j = \frac{\lambda^2\beta_1^2}{2(1-\rho)} + \frac{2\lambda(\rho+\sigma) + \lambda^2\alpha(\beta_2^2 - \beta_1^2)}{2(\lambda(1+\sigma) + \alpha(1-\rho))} + \frac{\lambda(\rho+\sigma)}{(1-\rho)\mu}.$$

Il convient de souligner que les probabilités limites  $\{\pi_j\}_{j=0}^{\infty}$  peuvent être calculées à partir de l'équation réursive (4.2).

### 4.1.3 Approche par les processus régénératifs

Plusieurs processus stochastiques survenant, par exemple, dans les systèmes de files d'attente et les systèmes de gestion de stock possèdent la propriété de "régénération" en certains instants. Alors, le comportement futur du processus après ces instants devient une réplique, c'est-à-dire, le comportement futur du processus après ces instants possède exactement la loi de probabilité qu'il aurait eu s'il avait commencé à l'instant zéro. De tels processus sont appelés "processus régénératifs" [?, ?, ?].

#### Distributions limites

On a :

$$P_{ij} = \lim_{t \rightarrow \infty} P\{(C(t), N(t)) = (i, j)\}, \quad (i, j) \in S.$$

**Théorème 4.2.** Sous la condition de stabilité  $\rho < 1$ ,

(i) Les probabilités limites sont données par :

– Lorsque le serveur est inactif :

$$P_{0j} = \left(\lambda + \frac{\alpha(1-\rho)}{1+\sigma}\right) \frac{\pi_j}{\lambda + \alpha + j\mu}, \quad j \in \mathbb{Z}_+. \quad (4.8)$$

– Lorsque le serveur est occupé par un appel entrant :

$$P_{1j} = \left(\lambda + \frac{\alpha(1-\rho)}{1+\sigma}\right) \left( \sum_{n=0}^j \frac{\lambda}{\lambda + \alpha + n\mu} \pi_n \widehat{k}_{j-n}^1 + \sum_{n=1}^{j+1} \frac{n\mu}{\lambda + \alpha + n\mu} \pi_n \widehat{k}_{j-n+1}^1 \right), \quad j \in \mathbb{Z}_+. \quad (4.9)$$

– Lorsque le serveur est occupé par un appel sortant :

$$P_{2j} = \left(\lambda + \frac{\alpha(1-\rho)}{1+\sigma}\right) \sum_{n=0}^j \frac{\alpha}{\lambda + \alpha + n\mu} \pi_n \widehat{k}_{j-n}^2, \quad j \in \mathbb{Z}_+. \quad (4.10)$$

où les quantités  $\widehat{k}_j^l$  sont données par :

$$\widehat{k}_j^l = \frac{1}{\lambda} \left( 1 - \sum_{i=0}^j k_i^l \right), \quad l = 1, 2, \quad j \in Z_+, \quad (4.11)$$

(ii) Les fonctions génératrices partielles des probabilités limites :

$$P_i(z) = \sum_{j=0}^{\infty} z^j P_{ij}, \quad \text{pour } i \in \{0, 1, 2\},$$

sont données par

$$P_0(z) = \left( \lambda + \frac{\alpha(1-\rho)}{1+\sigma} \right) \Phi(z), \quad (4.12)$$

$$P_1(z) = \frac{1 - \beta_1(\lambda - \lambda z)}{\beta_1(\lambda - \lambda z) - z} \left( 1 + \frac{\alpha(1 - \beta_2(\lambda - \lambda z))}{\lambda(1 - z)} \right) P_0(z), \quad (4.13)$$

$$P_2(z) = \alpha(1 - \beta_2(\lambda - \lambda z))\lambda(1 - z)P_0(z). \quad (4.14)$$

## 4.2 Inégalités stochastiques pour le système $M_1, M_2/G_1, G_2/1$ avec rappels à communication bidirectionnelle

Dans cette section, on utilise la théorie générale des ordres partiels pour l'étude des propriétés de monotonie du système d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle relativement aux ordres : stochastique " $\leq_{st}$ ", laplacien " $\leq_L$ " et convexe " $\leq_v$ ".

On introduit les notations suivantes :

Soient  $\Sigma_1$  et  $\Sigma_2$  deux modèles d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle de paramètres, pour  $i = 1, 2$  :

$\lambda^{(i)}$  : taux d'arrivées entrantes dans  $\Sigma_i$ .

$\mu^{(i)}$  : taux de rappels dans  $\Sigma_i$ .

$\alpha^{(i)}$  : taux d'appels sortants dans  $\Sigma_i$

$B_1^{(i)}(x)$  : distribution du temps de service d'un appel entrant dans  $\Sigma_i$ .

$B_2^{(i)}(x)$  : distribution du temps de service d'un appel sortant dans  $\Sigma_i$ .

$k_n^{(i)}$  : le nombre de nouvelles arrivées durant le service du  $n^{\text{ème}}$  client dans  $\Sigma_i$ .

$\pi_n^{(i)}$  : la distribution stationnaire du nombre de clients dans le système  $\Sigma_i$ .

### 4.2.1 Inégalités préliminaires

Les lemmes suivants donnent les conditions, sur les paramètres des deux systèmes, sous lesquelles les probabilités du nombre de clients arrivant durant le service d'un appel entrant ou

sortant de deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle  $\{k_n^{(i)}, i = 1, 2 \text{ et } n \in \mathbb{N}\}$  sont comparables suivant les ordres partiels : stochastique, convexe et en transformée de Laplace.

**Lemme 4.1.** Soient  $\Sigma_1$  et  $\Sigma_2$  deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle,

$$\text{si } \lambda^{(1)} \leq \lambda^{(2)} \text{ et } B_l^{(1)} \leq_{st} B_l^{(2)} \text{ alors } \{k_n^{(1)}\} \leq_{st} \{k_n^{(2)}\},$$

où,

$$k_n^{(i)} = P(X = n) = \int_0^{+\infty} \frac{(\lambda^{(i)}t)^n}{n!} e^{-\lambda^{(i)}t} dB_l^{(i)}(t), \quad i = 1, 2, \quad l = 1, 2.$$

**Preuve.** Supposons que  $\lambda^{(1)} \leq \lambda^{(2)}$  et  $B_l^{(1)} \leq_{st} B_l^{(2)}, l = 1, 2.$

Par définition de l'ordre stochastique  $\leq_{st}$ , on a pour une loi discrète, les équivalences suivantes :

$$\begin{aligned} \{k_n^{(1)}\} \leq_{st} \{k_n^{(2)}\} &\Leftrightarrow \bar{k}_n^{(1)} = \sum_{m=n}^{+\infty} k_m^{(1)} \leq \sum_{m=n}^{+\infty} k_m^{(2)} = \bar{k}_n^{(2)} \\ &\Leftrightarrow \sum_{m=n}^{+\infty} \int_0^{+\infty} \frac{(\lambda^{(1)}x)^m}{m!} \exp\{-\lambda^{(1)}x\} dB_l^{(1)}(x), \\ &= \int_0^{+\infty} \sum_{m=n}^{+\infty} \frac{(\lambda^{(1)}x)^m}{m!} \exp\{-\lambda^{(1)}x\} dB_l^{(1)}(x), \\ &\leq \int_0^{+\infty} \sum_{m=n}^{+\infty} \frac{(\lambda^{(2)}x)^m}{m!} \exp\{-\lambda^{(2)}x\} dB_l^{(2)}(x), \quad l = 1, 2. \end{aligned} \quad (4.15)$$

Pour prouver l'inégalité numérique (4.15), on doit prouver que la fonction

$$f_n(x, \lambda) = \sum_{m=n}^{+\infty} \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\},$$

est non décroissante.

On a :

$$\begin{aligned}
 \frac{\partial}{\partial x} f_n(x, \lambda) &= \sum_{m=n}^{+\infty} \left[ m \lambda \frac{(\lambda x)^{m-1}}{m!} \exp\{-\lambda x\} - \lambda \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\} \right] \\
 &= \sum_{m=n}^{+\infty} \lambda \frac{(\lambda x)^{m-1}}{(m-1)!} \exp\{-\lambda x\} - \sum_{m=n}^{+\infty} \lambda \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\} \\
 &= \sum_{m=n+1}^{+\infty} \lambda \frac{(\lambda x)^{m-1}}{(m-1)!} \exp\{-\lambda x\} - \sum_{m=n}^{+\infty} \lambda \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\} + \lambda \frac{(\lambda x)^{n-1}}{(n-1)!} \exp\{-\lambda x\} \\
 &= \lambda \frac{(\lambda x)^{n-1}}{(n-1)!} \exp\{-\lambda x\} > 0, \quad \forall x \geq 0.
 \end{aligned}$$

Donc  $f_n(x, \lambda)$  est une fonction croissante en  $x$ .

La dérivée par rapport à  $\lambda$  s'écrit comme suit :

$$\begin{aligned}
 \frac{\partial}{\partial \lambda} f_n(x, \lambda) &= \sum_{m=n}^{+\infty} m x \frac{(\lambda x)^{m-1}}{m!} \exp\{-\lambda x\} - \sum_{m=n}^{+\infty} x \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\} \\
 &= x \frac{(\lambda x)^{n-1}}{(n-1)!} \exp\{-\lambda x\} + \sum_{m=n+1}^{+\infty} x \frac{(\lambda x)^{m-1}}{(m-1)!} \exp\{-\lambda x\} - \sum_{m=n}^{+\infty} x \frac{(\lambda x)^m}{m!} \exp\{-\lambda x\} \\
 &= x \frac{(\lambda x)^{n-1}}{(n-1)!} \exp\{-\lambda x\} > 0.
 \end{aligned}$$

On remarque que la dérivée est positive pour toutes les valeurs positives que peut prendre le paramètre  $\lambda$ . Alors, la fonction  $f_n(x, \lambda)$  est croissante par rapport aux valeurs du paramètre  $\lambda$ .

En effet,

$$\frac{\partial}{\partial \lambda} f_n(x, \lambda) = x \exp\{-\lambda x\} \frac{(\lambda x)^{n-1}}{(n-1)!} > 0.$$

Comme  $f_n(x, \lambda)$  est une fonction croissante en  $x$  et  $B_l^{(1)} \leq_{st} B_l^{(2)}$ , alors d'après le Théorème 3.1 énoncé dans le chapitre 3, l'inégalité suivante est vérifiée :

$$\int_0^{+\infty} f_n(x, \lambda^{(1)}) dB_l^{(1)}(x) \leq \int_0^{+\infty} f_n(x, \lambda^{(1)}) dB_l^{(2)}(x), \quad l = 1, 2. \quad (4.16)$$

D'autre part, puisque la fonction  $f_n(x, \lambda)$  est monotone par rapport à  $\lambda$  et que  $\lambda^{(1)} \leq \lambda^{(2)}$ , on a

$$\int_0^{+\infty} f_n(x, \lambda^{(1)}) dB_l^{(2)}(x) \leq \int_0^{+\infty} f_n(x, \lambda^{(2)}) dB_l^{(2)}(x), \quad l = 1, 2. \quad (4.17)$$

Par conséquent, des inégalités (4.16) et (4.17), l'inégalité (4.15) est vérifiée par transitivité.

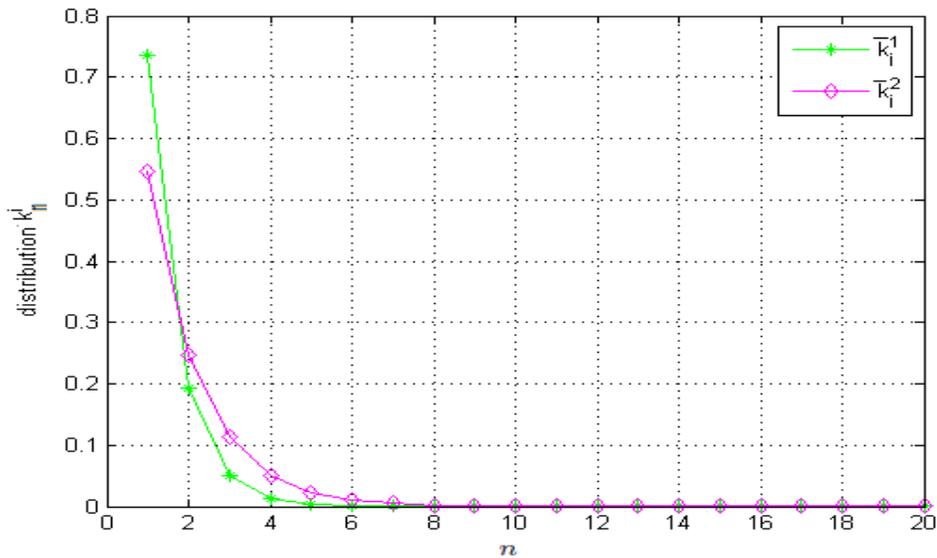


FIGURE 4.4 – Comparaison des distributions  $\{k_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique, pour  $\lambda^{(1)} = 0.25, \lambda^{(2)} = 0.5, \theta^{(1)} = 0.7, \theta^{(2)} = 0.6$

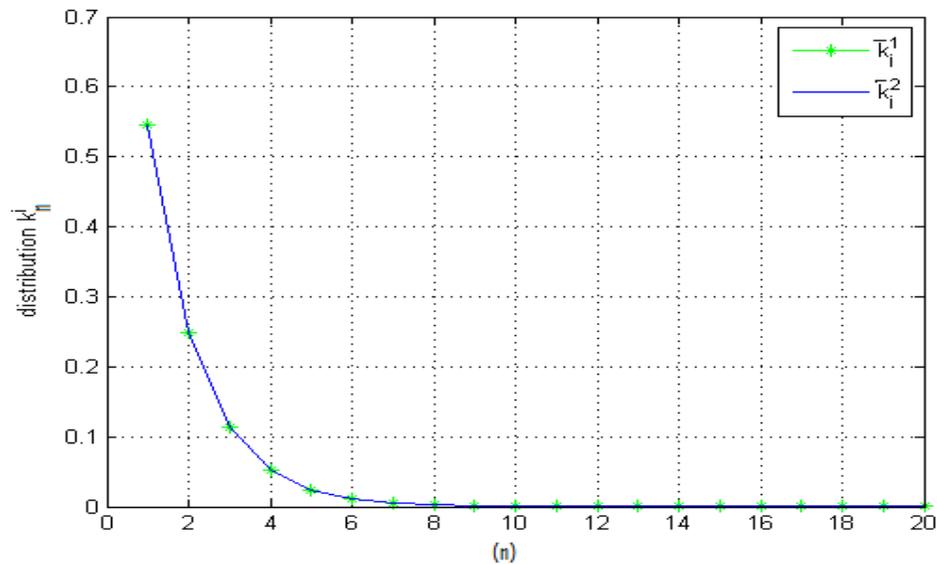


FIGURE 4.5 – Comparaison des distributions  $\{k_n^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique, pour  $\lambda^{(1)} = \lambda^{(2)} = 0.5$  et  $\theta^{(1)} = \theta^{(2)} = 0.6$

Selon les deux Figures ci-dessus, on remarque que les probabilités du nombre de clients

arrivant durant le service du  $n^{\text{ième}}$  client, diminue stochastiquement lorsque le taux d'arrivée décroît et en diminuant stochastiquement le temps de service.

**Lemme 4.2.** Soient  $\Sigma_1$  et  $\Sigma_2$  deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle,

$$\text{si } \lambda^{(1)} \leq \lambda^{(2)} \text{ et } B_l^{(1)} \leq_v B_l^{(2)}, l = 1, 2 \text{ alors } \{k_n^{(1)}\} \leq_v \{k_n^{(2)}\}.$$

**Preuve.** Par définition de l'ordre convexe  $\leq_v$  on a :

$$\begin{aligned} \{k_n^{(1)}\} \leq_v \{k_n^{(2)}\} &\Leftrightarrow \bar{k}_n^{(1)} = \sum_{m=n}^{+\infty} \bar{k}_m^{(1)} \leq \sum_{m=n}^{+\infty} \bar{k}_m^{(2)} = \bar{k}_n^{(2)} \\ &\Leftrightarrow \int_0^{+\infty} \sum_{m=n}^{+\infty} \sum_{l=m}^{+\infty} \frac{(\lambda^{(1)}x)^l}{l!} \exp\{-\lambda^{(1)}x\} dB_l^{(1)}(x) \quad l = 1, 2 \\ &\leq \int_0^{+\infty} \sum_{m=n}^{+\infty} \sum_{l=m}^{+\infty} \frac{(\lambda^{(2)}x)^l}{l!} \exp\{-\lambda^{(2)}x\} dB_l^{(2)}(x), l = 1, 2 \quad (4.18) \\ &\Leftrightarrow \int_0^{+\infty} \sum_{m=n}^{+\infty} f_m(x, \lambda^{(1)}) dB_l^{(1)}(x) \leq \int_0^{+\infty} \sum_{m=n}^{+\infty} f_m(x, \lambda^{(2)}) dB_l^{(2)}(x), \end{aligned}$$

avec,

$$f_m(x, \lambda^{(i)}) = \sum_{l=m}^{+\infty} \frac{(\lambda^{(i)}x)^l}{l!} \exp\{-\lambda^{(i)}x\}.$$

Les fonctions  $f_m(x, \lambda)$  sont croissantes par rapport à  $\lambda$ , alors la fonction définie par :

$\bar{f}_n(x, \lambda) = \sum_{m=n}^{+\infty} f_m(x, \lambda)$  l'est aussi. D'autre part, on a

$$\frac{\partial^2}{\partial x^2} \bar{f}_n(x, \lambda) = \lambda \frac{\partial}{\partial x} f_{n-1}(x, \lambda) = \lambda^2 \left( \frac{(\lambda x)^{n-2}}{(n-2)!} \right) \exp\{-\lambda x\} > 0.$$

Par conséquent,  $\bar{f}_n(x, \lambda)$  est croissante et convexe par rapport à la variable  $x$ .

D'après le théorème 3.2, on obtient l'inégalité suivante :

$$\int_0^{+\infty} \bar{f}_n(x, \lambda^{(1)}) dB_l^{(1)}(x) \leq \int_0^{+\infty} \bar{f}_n(x, \lambda^{(1)}) dB_l^{(2)}(x). \quad (4.19)$$

Et on obtient, grâce à la monotonie de la fonction  $\bar{f}_n(x, \lambda)$  par rapport à  $\lambda$ , l'inégalité :

$$\int_0^{+\infty} \bar{f}_n(x, \lambda^{(1)}) dB_l^{(2)}(x) \leq \int_0^{+\infty} \bar{f}_n(x, \lambda^{(2)}) dB_l^{(2)}(x). \quad (4.20)$$

Finalement, l'inégalité (4.18) est vérifiée par transitivité des inégalités (4.19) et (4.20).

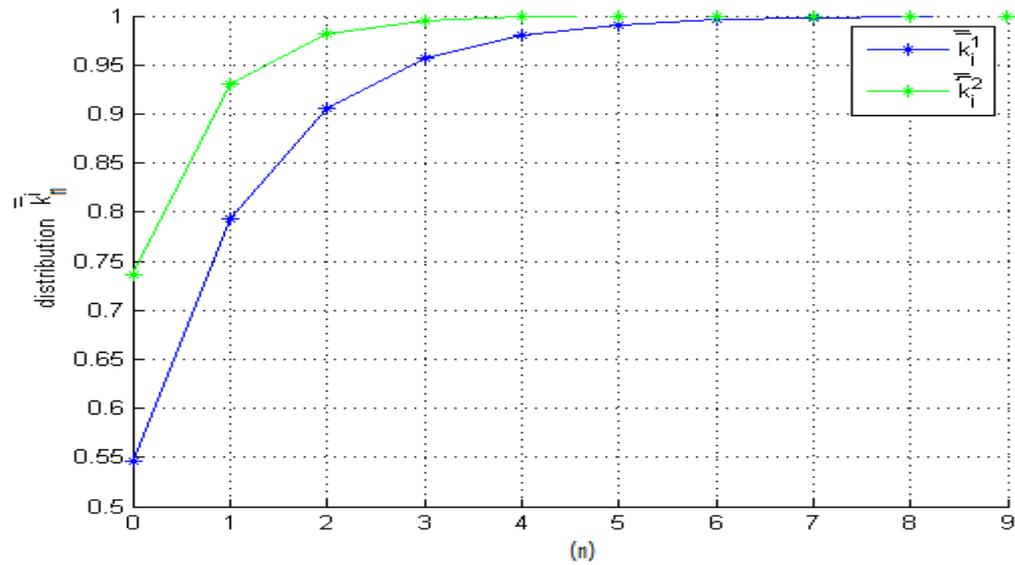


FIGURE 4.6 – Comparaison des distributions  $\{\bar{k}_n^{(i)}\}$ ,  $i = 1, 2$ , par rapport à l'ordre convexe, pour  $\lambda^{(1)} = 0.25$ ,  $\lambda^{(2)} = 0.5$ ,  $\theta^{(1)} = 0.7$  et  $\theta^{(2)} = 0.6$ .

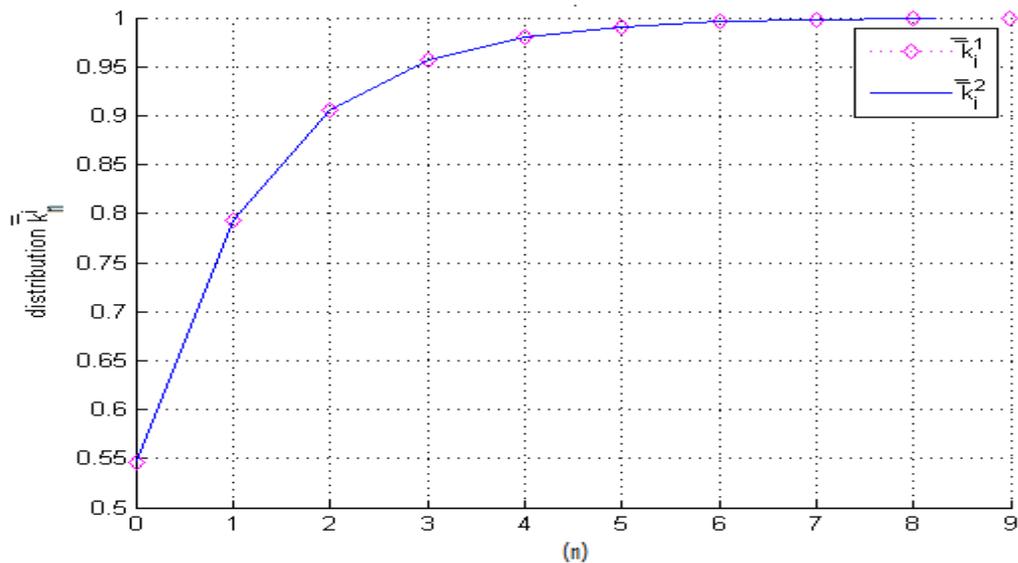


FIGURE 4.7 – Comparaison des distributions  $\{\bar{k}_n^{(i)}\}$ ,  $i = 1, 2$ , par rapport à l'ordre convexe, pour  $\lambda^{(1)} = \lambda^{(2)} = 0.5$  et  $\theta^{(1)} = \theta^{(2)} = 0.6$ .

Les deux Figures ci-dessus, montrent que les probabilités du nombre de clients arrivant durant

le service du  $n^{\text{ième}}$  client, diminue par rapport à l'ordre convexe lorsque le taux d'arrivée décroît et en diminuant par rapport à l'ordre convexe le temps de service.

**Lemme 4.3.** Soient  $\Sigma_1$  et  $\Sigma_2$  deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle,

$$\text{si } \lambda^{(1)} \leq \lambda^{(2)} \text{ et } B_l^{(1)} \leq_L B_l^{(2)} \text{ alors } \{k_n^{(1)}\} \leq_L \{k_n^{(2)}\}.$$

**Preuve.** Pour prouver que l'inégalité  $\{k_n^{(1)}\} \leq_L \{k_n^{(2)}\}$  a lieu, il suffit d'établir l'inégalité suivante, pour les fonctions génératrices correspondantes

$$k^{(1)}(z) \geq k^{(2)}(z),$$

ce qui est équivalent à montrer que

$$\tilde{B}_l^{(1)}(\lambda^{(1)}(1-z)) \geq \tilde{B}_l^{(2)}(\lambda^{(2)}(1-z)),$$

c'est-à-dire montrer l'équivalence suivante :

$$\{k_n^{(1)}\} \leq_L \{k_n^{(2)}\} \Leftrightarrow \tilde{B}_l^{(1)}(\lambda^{(1)}(1-z)) \geq \tilde{B}_l^{(2)}(\lambda^{(2)}(1-z)). \quad (4.21)$$

Par définition, on a :

$$\begin{aligned} k(z) = \sum_{n \geq 0} k_n z^n &= \sum_{n \geq 0} \int_0^{+\infty} \frac{(\lambda x)^n}{n!} \exp\{-\lambda x\} z^n dB(x) \\ &= \int_0^{+\infty} \sum_{n \geq 0} \frac{(\lambda x z)^n}{n!} \exp\{-\lambda x\} dB(x) \\ &= \int_0^{+\infty} \exp\{-\lambda x(1-z)\} dB(x) = \tilde{B}(\lambda(1-z)). \end{aligned}$$

De plus,

$$B_l^{(1)} \leq_L B_l^{(2)} \Rightarrow \tilde{B}_l^{(1)}(s) \geq \tilde{B}_l^{(2)}(s), \quad \forall s \geq 0, \quad l = 1, 2.$$

En particulier, pour  $s = \lambda^{(1)}(1-z)$ , on a :

$$\tilde{B}_l^{(1)}(\lambda^{(1)}(1-z)) \geq \tilde{B}_l^{(2)}(\lambda^{(1)}(1-z)). \quad (4.22)$$

Puisque toute transformée de Laplace est une fonction décroissante, l'inégalité  $\lambda^{(1)} \leq \lambda^{(2)}$  implique l'inégalité suivante :

$$\tilde{B}_l^{(2)}(\lambda^{(1)}(1-z)) \geq \tilde{B}_l^{(2)}(\lambda^{(2)}(1-z)), \quad l = 1, 2. \quad (4.23)$$

Par conséquent, l'inégalité (4.21) découle des inégalités (4.22) et (4.23).

### 4.2.2 Monotonie de la chaîne de Markov incluse

Les probabilités de transition en un pas de la chaîne de Markov incluse pour le système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle sont données par la formule suivante :

$$p_{n,m} = \begin{cases} \frac{n\mu}{\lambda+\alpha+n\mu} k_0^1, & \text{si } m = n - 1 \text{ et } n \geq 1, \\ \frac{\lambda}{\lambda+\alpha+n\mu} k_{m-n}^1 + \frac{\alpha}{\lambda+\alpha+n\mu} k_{m-n}^2 + \frac{n\mu}{\lambda+\alpha+n\mu} k_{m-n+1}^1, & \text{si } 0 \leq n \leq m. \end{cases} \quad (4.24)$$

Soit l'opérateur de transition  $\tau$  de la chaîne de Markov incluse. Pour chaque distribution  $p = (p_n)_{n \geq 0}$ , on associe une distribution  $\tau p = q = (q_m)_{m \geq 0}$  telle que

$$q_m = \sum_{n \geq 0} p_n p_{nm}.$$

Les deux théorèmes suivants donnent la condition sous laquelle l'opérateur de transition  $\tau$  est monotone par rapport aux ordres stochastique et convexe.

**Théorème 4.3.** Si l'inégalité  $B_2 \leq_{st} B_1$  a lieu, alors l'opérateur de transition  $\tau$  est monotone, par rapport à l'ordre stochastique. C'est-à-dire, pour deux distributions quelconques  $p^{(1)}$  et  $p^{(2)}$ , l'inégalité  $p^{(1)} \leq_{st} p^{(2)}$  implique la suivante :  $\tau p^{(1)} \leq_{st} \tau p^{(2)}$ .

**Preuve.** Un opérateur est monotone par rapport à l'ordre stochastique si et seulement si on a l'inégalité suivante :

$$\bar{p}_{n-1,m} \leq \bar{p}_{n,m}, \quad \forall n, m, \quad (4.25)$$

avec,

$$\begin{aligned} \bar{p}_{n,m} &= \sum_{k=m}^{+\infty} p_{n,k} = \sum_{k=m}^{+\infty} \left[ \frac{\lambda}{\lambda+\alpha+n\mu} k_{k-n}^1 + \frac{\alpha}{\lambda+\alpha+n\mu} k_{k-n}^2 + \frac{n\mu}{\lambda+\alpha+n\mu} k_{k-n+1}^1 \right] \\ &= \frac{\lambda}{\lambda+\alpha+n\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda+\alpha+n\mu} \bar{k}_{m-n}^2 + \frac{n\mu}{\lambda+\alpha+n\mu} \bar{k}_{m-n+1}^1 \\ &= \frac{\lambda+n\mu}{\lambda+\alpha+n\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda+\alpha+n\mu} \bar{k}_{m-n}^2 - \frac{n\mu}{\lambda+\alpha+n\mu} k_{m-n}^1 \\ &= \frac{\lambda+n\mu}{\lambda+\alpha+n\mu} \bar{k}_{m-n+1}^1 + \frac{\alpha}{\lambda+\alpha+n\mu} \bar{k}_{m-n}^2 + \frac{\lambda}{\lambda+\alpha+n\mu} k_{m-n}^1, \end{aligned}$$

et,

$$\bar{p}_{n-1,m} = \frac{\lambda+(n-1)\mu}{\lambda+\alpha+(n-1)\mu} \bar{k}_{m-n+1}^1 + \frac{\alpha}{\lambda+\alpha+(n-1)\mu} \bar{k}_{m-n+1}^2 - \frac{(n-1)\mu}{\lambda+\alpha+(n-1)\mu} k_{m-n+1}^1.$$

D'où :

$$\begin{aligned} \bar{p}_{n,m} - \bar{p}_{n-1,m} &= \frac{\alpha\mu}{(\lambda+\alpha+n\mu)(\lambda+\alpha+(n-1)\mu)} [\bar{k}_{m-n+1}^1 - \bar{k}_{m-n+1}^2] \\ &\quad + \frac{\lambda}{\lambda+\alpha+n\mu} k_{m-n}^1 + \frac{(n-1)\mu}{\lambda+\alpha+(n-1)\mu} k_{m-n+1}^1 + \\ &\quad \frac{\alpha}{\lambda+\alpha+n\mu} k_{m-n}^2 \geq 0. \end{aligned}$$

Par conséquent, puisqu'on a l'inégalité  $B_2 \leq_{st} B_1$ , alors l'inégalité (4.25) est vérifiée.

En conclusion, l'opérateur  $\tau$  est monotone par rapport à l'ordre stochastique.

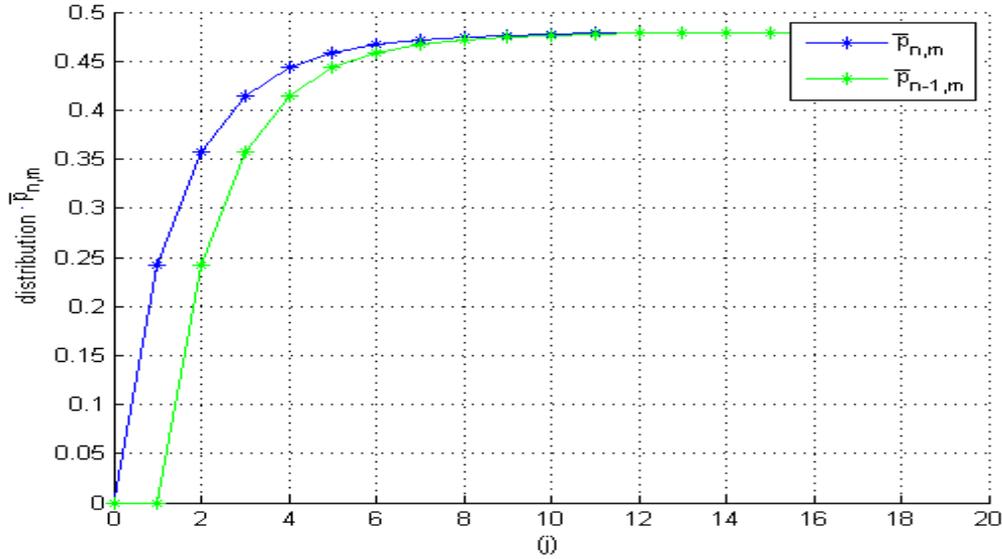


FIGURE 4.8 – Comparaison des probabilités de transition  $\bar{p}_{n,m}$  et  $\bar{p}_{n-1,m}$ ,  $\forall n, m$ .

**Remarque 4.1.** En particulier, si les deux distributions  $B_1$  et  $B_2$  possèdent la même loi de probabilité, c'est-à-dire,  $B_2 \equiv_{st} B_1$ , alors l'opérateur de transition est aussi monotone.

**Théorème 4.4.** Si l'inégalité  $B_1 \equiv_v B_2$  a lieu, alors l'opérateur de transition  $\tau$  est monotone, par rapport à l'ordre convexe. C'est-à-dire, pour deux distributions quelconques  $p^{(1)}$  et  $p^{(2)}$ , l'inégalité  $p^{(1)} \leq_v p^{(2)}$  implique la suivante :  $\tau p^{(1)} \leq_v \tau p^{(2)}$ .

**Preuve.** L'opérateur  $\tau$  est monotone par rapport à l'ordre convexe si et seulement si :

$$2\bar{\bar{p}}_{n,m} \leq \bar{\bar{p}}_{n-1,m} + \bar{\bar{p}}_{n+1,m}, \quad \forall n, m, \quad (4.26)$$

où,

$$\begin{aligned} \bar{\bar{p}}_{n,m} &= \sum_{k=m}^{+\infty} \bar{p}_{n,k} = \sum_{k=m}^{+\infty} \left[ \frac{\lambda}{\lambda + \alpha + n\mu} \bar{k}_{k-n}^1 + \frac{\alpha}{\lambda + \alpha + n\mu} \bar{k}_{k-n}^2 + \frac{n\mu}{\lambda + \alpha + n\mu} k_{k-n+1}^1 \right] \\ &= \frac{\lambda}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^2 + \frac{n\mu}{\lambda + \alpha + n\mu} \bar{k}_{m-n+1}^1 \\ &= \frac{\lambda + n\mu}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^2 - \frac{n\mu}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^1 \\ &= \frac{\lambda + n\mu}{\lambda + \alpha + n\mu} \bar{k}_{m-n+1}^1 + \frac{\alpha}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^2 + \frac{\lambda}{\lambda + \alpha + n\mu} \bar{k}_{m-n}^1, \end{aligned}$$

$$\bar{p}_{n-1,m} = \frac{\lambda + (n-1)\mu}{\lambda + \alpha + (n-1)\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda + \alpha + (n-1)\mu} \bar{k}_{m-n}^2 - \left( \frac{\lambda + (n-1)\mu}{\lambda + \alpha + (n-1)\mu} \right) \bar{k}_{m-n}^1 + \frac{(n-1)\mu}{\lambda + \alpha + (n-1)\mu} k_{m-n}^1 - \frac{\alpha}{\lambda + \alpha + (n+1)\mu} \bar{k}_{m-n}^2,$$

et

$$\bar{p}_{n+1,m} = \frac{\lambda + (n+1)\mu}{\lambda + \alpha + (n+1)\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} \bar{k}_{m-n}^2 + \frac{\lambda}{\lambda + \alpha + (n+1)\mu} \bar{k}_{m-n}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} \bar{k}_{m-n}^2 + \frac{\lambda}{\lambda + \alpha + n\mu} k_{m-n-1}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^2.$$

On a :

$$\begin{aligned} \bar{p}_{n-1,m} + \bar{p}_{n+1,m} - 2\bar{p}_{n,m} &= \left[ \frac{\lambda + (n-1)\mu}{\lambda + \alpha + (n-1)\mu} + \frac{\lambda + (n+1)\mu}{\lambda + \alpha + (n+1)\mu} - \frac{\lambda + n\mu}{\lambda + \alpha + n\mu} - \frac{\lambda + n\mu}{\lambda + \alpha + n\mu} \right] \bar{k}_{m-n}^1 \\ &+ \left[ \frac{\alpha}{\lambda + \alpha + (n-1)\mu} + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} - \frac{\alpha}{\lambda + \alpha + n\mu} - \frac{\alpha}{\lambda + \alpha + n\mu} \right] \bar{k}_{m-n}^2 \\ &+ \left[ \frac{\lambda}{\lambda + \alpha + (n+1)\mu} - \frac{\lambda + 2(n-1)\mu}{\lambda + \alpha + (n-1)\mu} + \frac{2n\mu}{\lambda + \alpha + n\mu} \right] \bar{k}_{m-n}^1 \\ &+ \left[ \frac{\alpha}{\lambda + \alpha + (n+1)\mu} - \frac{\alpha}{\lambda + \alpha + (n-1)\mu} \right] \bar{k}_{m-n}^2 \\ &+ \frac{(n-1)\mu}{\lambda + \alpha + (n-1)\mu} k_{m-n}^1 + \frac{\lambda}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^2. \end{aligned}$$

$$\begin{aligned} \bar{p}_{n-1,m} + \bar{p}_{n+1,m} - 2\bar{p}_{n,m} &= \frac{2\alpha\mu^2}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + n\mu)(\lambda + \alpha + (n-1)\mu)} \left[ \bar{k}_{m-n}^2 - \bar{k}_{m-n}^1 \right] \\ &+ \frac{2\mu(\lambda\alpha + \alpha^2 + \alpha n\mu + \lambda\mu + \alpha\mu)}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + (n-1)\mu)(\lambda + \alpha + n\mu)} \bar{k}_{m-n}^1 \\ &- \frac{2\alpha\mu}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + (n-1)\mu)} \bar{k}_{m-n}^2 \\ &+ \frac{(n-1)\mu}{\lambda + \alpha + (n-1)\mu} k_{m-n}^1 + \frac{\lambda}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^2. \end{aligned}$$

$$\begin{aligned} \bar{p}_{n-1,m} + \bar{p}_{n+1,m} - 2\bar{p}_{n,m} &= \frac{2\alpha\mu^2}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + n\mu)(\lambda + \alpha + (n-1)\mu)} \left[ \bar{k}_{m-n+1}^2 - \bar{k}_{m-n+1}^1 \right] \\ &+ \frac{2\alpha\mu^2}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + n\mu)(\lambda + \alpha + (n-1)\mu)} \left[ \bar{k}_{m-n}^2 - \bar{k}_{m-n}^1 \right] \\ &+ \frac{2\alpha\mu(\lambda + \alpha + n\mu)}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + (n-1)\mu)(\lambda + \alpha + n\mu)} \left[ \bar{k}_{m-n}^1 - \bar{k}_{m-n}^2 \right] \\ &+ \frac{2\mu^2(\lambda + \alpha)}{(\lambda + \alpha + (n+1)\mu)(\lambda + \alpha + n\mu)} \bar{k}_{m-n}^1 \\ &+ \frac{(n-1)\mu}{\lambda + \alpha + (n-1)\mu} k_{m-n}^1 + \frac{\lambda}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^1 + \frac{\alpha}{\lambda + \alpha + (n+1)\mu} k_{m-n-1}^2 \geq 0. \end{aligned}$$

En conclusion, si on a un système de file d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle et l'inégalité  $B_1 \equiv_v B_2$ , alors l'opérateur de transition  $\tau$  est monotone, par rapport à l'ordre convexe.

Maintenant, on considère  $\Sigma_1$  et  $\Sigma_2$  deux modèles d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle ayant les paramètres suivants :  $\lambda^{(1)}, \mu^{(1)}, \alpha^{(1)}, B_1^{(1)}(x), B_2^{(1)}(x)$  et  $\lambda^{(2)}, \mu^{(2)}, \alpha^{(2)}, B_1^{(2)}(x), B_2^{(2)}(x)$  respectivement. Notons par  $\tau^{(1)}, \tau^{(2)}$  les opérateurs de transition associés aux chaînes de Markov incluses de chaque système.

Les deux théorèmes suivants donnent les conditions de comparabilité de ces opérateurs par rapport aux ordres partiels : stochastique et convexe.

**Théorème 4.5.** Soient  $\Sigma_1$  et  $\Sigma_2$  deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle,

$$\text{si } \lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \geq \mu^{(2)}, \alpha^{(1)} \leq \alpha^{(2)} \text{ } B_1^{(1)} \leq_{st} B_1^{(2)} \text{ et } B_2^{(1)} \leq_{st} B_2^{(2)} \text{ alors } \tau^{(1)} \leq_{st} \tau^{(2)},$$

c'est-à-dire que pour une distribution quelconque  $p$  on a  $\tau^{(1)}p \leq_{st} \tau^{(2)}p$ .

**Preuve.** D'après le théorème 3.6, nous devons vérifier les inégalités suivantes pour l'ordre stochastique,

$$\bar{p}_{nm}^{(1)} \leq \bar{p}_{nm}^{(2)}, \quad \forall 0 \leq n \leq m,$$

Ce qui revient à montrer :

$$\begin{aligned} & \frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{1(1)}_{m-n+1} + \frac{\alpha^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{2(1)}_{m-n} - \frac{n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} k^{1(1)}_{m-n} \leq \\ & \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{1(2)}_{m-n+1} + \frac{\alpha^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{2(2)}_{m-n} - \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} k^{1(2)}_{m-n} \end{aligned} \quad (4.27)$$

D'après le lemme 4.1 on a :

$$\{k_n^l(1)\} \leq_{st} \{k_n^l(2)\}. \quad (4.28)$$

On a  $\lambda^{(1)} \leq \lambda^{(2)}, \alpha^{(1)} \leq \alpha^{(2)}$  et  $\mu^{(1)} \geq \mu^{(2)}$ , implique que :

$$\frac{\lambda^{(1)} + \alpha^{(1)}}{\mu^{(1)}} \leq \frac{\lambda^{(2)} + \alpha^{(2)}}{\mu^{(2)}},$$

et puisque la fonction :  $\frac{m}{x+m}$  est décroissante alors :

$$\frac{n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \geq \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}}, \quad (4.29)$$

on a aussi :  $\frac{\lambda^{(1)}}{\mu^{(1)}} \leq \frac{\lambda^{(2)}}{\mu^{(2)}}$

et du fait que la fonction  $\frac{x}{x+m}$  est croissante alors :

$$\frac{\lambda^{(1)}}{\lambda^{(1)} + n\mu^{(1)}} \leq \frac{\lambda^{(2)}}{\lambda^{(2)} + n\mu^{(2)}},$$

d'où :

$$\lambda^{(1)} + n\mu^{(1)} \geq \lambda^{(2)} + n\mu^{(2)},$$

par conséquent :

$\alpha^{(1)} \leq \alpha^{(2)}$  et  $\lambda^{(1)} + n\mu^{(1)} \geq \lambda^{(2)} + n\mu^{(2)}$  nous donnent :

$$\frac{\lambda^{(1)} + n\mu^{(1)}}{\alpha^{(1)}} \geq \frac{\lambda^{(2)} + n\mu^{(2)}}{\alpha^{(2)}},$$

et comme, la fonction  $\frac{x}{x+m}$  est croissante, alors :

$$\frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \geq \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}}. \quad (4.30)$$

Des inégalités (4.28), (4.29) et (4.30) on obtient :

$$\begin{aligned} & \frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{1(1)}_{m-n} - \frac{n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} k^{1(1)}_{m-n} - \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{1(2)}_{m-n} \\ & + \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} k^{1(2)}_{m-n} \leq \frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{1(1)}_{m-n} - \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} k^{1(2)}_{m-n} \\ & - \frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{1(1)}_{m-n} + \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} k^{1(2)}_{m-n} = 0. \text{ De plus, on a :} \end{aligned}$$

$$\frac{\lambda + \alpha + n\mu}{\lambda + \alpha + n\mu} = 1 \Rightarrow \frac{\lambda + n\mu}{\lambda + \alpha + n\mu} = 1 - \frac{\alpha}{\lambda + \alpha + n\mu},$$

donc

$$\frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \geq \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}},$$

devient :

$$1 - \frac{\alpha^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \geq 1 - \frac{\alpha^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}},$$

ceci implique que :

$$\frac{\alpha^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \leq \frac{\alpha^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}}. \quad (4.31)$$

par conséquent, si  $B_1^{(1)} \leq_{st} B_1^{(2)}$  et  $B_2^{(1)} \leq_{st} B_2^{(2)}$ , alors l'inégalité (4.27) est vérifiée.

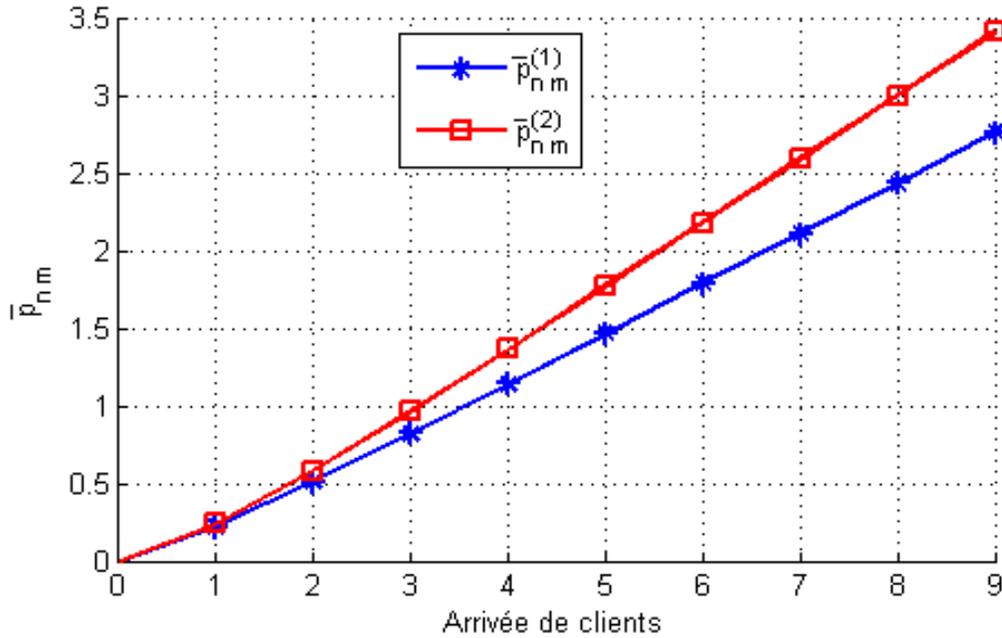


FIGURE 4.9 – Comparaison des probabilités de transition  $\{p_{nm}^{(i)}\}, i = 1, 2$ , par rapport à l'ordre stochastique.

**Théorème 4.6.** Soient  $\Sigma_1$  et  $\Sigma_2$  deux systèmes d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle,

si  $\lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \geq \mu^{(2)}, \alpha^{(1)} \leq \alpha^{(2)}, B_1^{(1)} \leq_v B_1^{(2)}$  et  $B_2^{(1)} \leq_v B_2^{(2)}$  alors  $\tau^{(1)} \leq_v \tau^{(2)}$ ,

c'est-à-dire que pour une distribution quelconque  $p$  on a  $\tau^{(1)}p \leq_v \tau^{(2)}p$ .

**Preuve.** D'après le Théorème 3.6, nous devons vérifier les inégalités suivantes pour l'ordre convexe,

$$\bar{p}_{nm}^{(1)} \leq \bar{p}_{nm}^{(2)}, \quad \forall 0 \leq n \leq m.$$

C'est à dire :

$$\frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{(1)}_{m-n+1} + \frac{\alpha^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{(1)}_{m-n} - \frac{n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \bar{k}^{(1)}_{m-n} \leq \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{(2)}_{m-n+1} + \frac{\alpha^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{(2)}_{m-n} - \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \bar{k}^{(2)}_{m-n}. \quad (4.32)$$

En effet : D'après le lemme (4.2) on a :

$$\{k_n^l\} \leq_v \{k_n^l\} \quad \text{pour } l = 1, 2. \quad (4.33)$$

et d'après les inégalité, (4.29), (4.30) et (4.31) on a :

$$\frac{\lambda^{(1)} + n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \overline{k}^{(1)}_{m-n+1} + \frac{\alpha^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \overline{k}^{(1)}_{m-n} - \frac{n\mu^{(1)}}{\lambda^{(1)} + \alpha^{(1)} + n\mu^{(1)}} \overline{k}^{(1)}_{m-n} \leq \frac{\lambda^{(2)} + n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \overline{k}^{(2)}_{m-n+1} + \frac{\alpha^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \overline{k}^{(2)}_{m-n} - \frac{n\mu^{(2)}}{\lambda^{(2)} + \alpha^{(2)} + n\mu^{(2)}} \overline{k}^{(2)}_{m-n}. \quad (4.34)$$

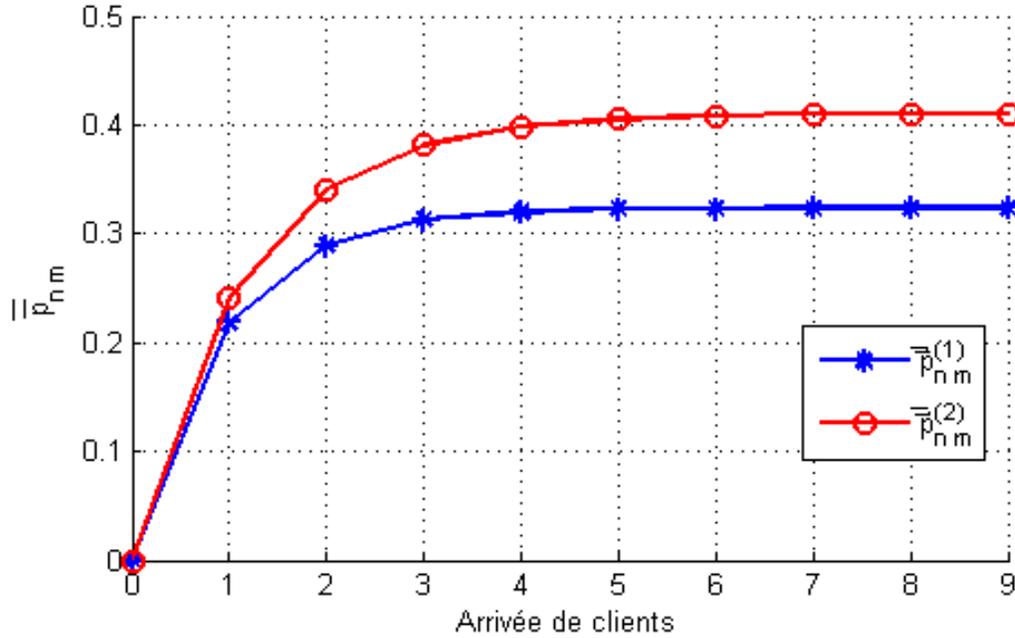


FIGURE 4.10 – Comparaison des probabilités de transition  $\{p_{nm}^{(i)}, i = 1, 2\}$ , par rapport à l'ordre convexe

### 4.2.3 Inégalités stochastiques des distributions stationnaires du nombre de clients dans le système

Les deux théorèmes suivants donnent les conditions de comparabilité des distributions stationnaires du nombre de clients, pour deux systèmes de files d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, par rapport aux ordres partiels : stochastiques et convexe.

**Théorème 4.7.** On considère  $\Sigma_1, \Sigma_2$  deux systèmes de files d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle ayant les paramètres  $\lambda^{(1)}, \mu^{(1)}, \alpha^{(1)}, B_1^{(1)}(x), B_2^{(1)}(x)$  et  $\lambda^{(2)}, \mu^{(2)}, \alpha^{(2)}, B_1^{(2)}(x), B_2^{(2)}(x)$  respectivement, et soient  $\pi_n^{(1)}, \pi_n^{(2)}$  les distributions stationnaires

du nombre de clients dans chaque système, alors si les inégalités suivantes ont lieu :

$$\lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \geq \mu^{(2)}, \alpha^{(1)} \leq \alpha^{(2)}, B_1^{(1)} \leq_{so} B_1^{(2)}, B_2^{(1)} \leq_{so} B_2^{(2)} \text{ et } B_2^{(2)} \leq_{st} B_1^{(2)} \text{ (resp. } B_1^{(2)} \leq_v B_2^{(2)}),$$

on a aussi les inégalités suivantes sur les distributions stationnaires

$$\{\pi_n^{(1)}\} \leq_{so} \{\pi_n^{(2)}\}, \quad \text{où } so = st \text{ (ou } v).$$

**Preuve.** D'après le théorème 4.6, les inégalités  $\lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \geq \mu^{(2)}, \alpha^{(1)} \leq \alpha^{(2)}, B_1^{(1)} \leq_v B_1^{(2)}$  et  $B_2^{(1)} \leq_v B_2^{(2)}$ , impliquent  $\tau^{(1)} \leq_{so} \tau^{(2)}$ , c'est-à-dire, pour une distribution quelconque  $p$  on a l'inégalité suivante :

$$\tau^{(1)} p \leq_{so} \tau^{(2)} p. \quad (4.35)$$

Par hypothèse, on a  $B_2^{(2)} \leq_{st} B_1^{(2)}$  (resp.  $B_1^{(2)} \leq_v B_2^{(2)}$ ), alors d'après le théorème 4.4 (resp. le théorème 4.6), l'opérateur  $\tau^{(2)}$  associé à la chaîne de Markov incluse, du deuxième système, est monotone. C'est-à-dire, pour deux distributions quelconques  $p_1^{(2)}, p_2^{(2)}$  telles que  $p_1^{(2)} \leq_{so} p_2^{(2)}$ , on a

$$\tau^{(2)} p_1^{(2)} \leq_{so} \tau^{(2)} p_2^{(2)}. \quad (4.36)$$

Cependant, de l'inégalité (4.35), on obtient

$$\tau^{(1)} p^{(1)} \leq_{so} \tau^{(2)} p^{(1)}. \quad (4.37)$$

Il existe une probabilité  $p_1^{(2)}$  telle qu'on ait l'inégalité suivante

$$\tau^{(2)} p^{(1)} \leq_{so} \tau^{(2)} p_1^{(2)}. \quad (4.38)$$

En combinant les inégalités (4.36)-(4.38), on obtient le résultat suivant

$$\tau^{(1)} p^{(1)} \leq_{so} \tau^{(2)} p^{(2)}, \quad (4.39)$$

pour deux distributions quelconques  $p^{(1)}, p^{(2)}$ .

L'inégalité (4.39) peut être réécrite de la manière suivante

$$\begin{aligned} \tau^{(1)} p_n^{(1)} &= P(Z_k^{(1)} = n) = P(Z_k^{(1)} = n) \\ &\leq_{so} P(Z_k^{(2)} = n) = P(Z_k^{(2)} = n) = \tau^{(2)} p_n^{(2)}. \end{aligned}$$

Quand  $k \rightarrow \infty$ , on a  $\{\pi_n^{(1)}\} \leq_{so} \{\pi_n^{(2)}\}$ .

**Théorème 4.8.** Si pour le modèle  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, la distribution de temps de service est *NBUE* (New Better than Used in Expectation) (respectivement *NWUE*- New Worse than Used in Expectation), et si de plus  $B_2^{(1)} \leq_v B_2^{(2)}$ ,

$B_1^* = B_1^{(2)} \leq_v B_2^{(2)}$ , alors la distribution stationnaire du nombre de clients dans ce système est inférieure (respectivement supérieure), par rapport à l'ordre convexe, à la distribution stationnaire du nombre de clients dans le système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle.

**Preuve.** Considérons un système de files d'attente  $M_1, M_1/M_1, M_2/1$  avec rappels à communication bidirectionnelle avec les mêmes paramètres : taux d'appels entrants  $\lambda$ , taux de rappels  $\mu$ , taux d'appels sortants  $\alpha$  temps moyen de service  $\beta_1^1$  et  $\beta_2^1$ , que le système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, mais avec des temps de service exponentiellement distribués avec des taux  $\theta_1 = \frac{1}{\beta_1^1}$  et  $\theta_2 = \frac{1}{\beta_2^1}$ .

$$B_1^*(x) = \begin{cases} 1 - e^{-\frac{x}{\beta_1^1}}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

$$B_2^*(x) = \begin{cases} 1 - e^{-\frac{x}{\beta_2^1}}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

D'après la proposition (3.4), Si  $B_1(x)$  est *NBUE* (respectivement *NWUE*), alors

$$B_1(x) \leq_v B_1^*(x), \quad (\text{respectivement } B_1(x) \geq_v B_1^*(x)).$$

Et comme  $B_2^{(1)} \leq_v B_2^{(2)}$  et  $B_1^* \leq_v B_2^{(2)}$ , alors d'après le théorème 4.7, on déduit que la distribution stationnaire du nombre de clients dans le système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle est inférieure (respectivement supérieure) à la distribution stationnaire du nombre de clients dans le système  $M_1, M_1/M_1, M_1/1$  avec rappels à communication bidirectionnelle.

### 4.3 Application numérique

Après avoir élaboré un simulateur, sous environnement Matlab, décrivant le comportement du modèle  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, nous avons estimé les probabilités stationnaires d'un tel système lorsque la distribution des temps de service est *NBUE* ( resp. *NWUE*). Ensuite, les comparer à celles du système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle afin de valider le résultat obtenu dans le Théorème 4.8. Pour ce faire, on a choisi deux lois de probabilités de type *NBUE* (une distribution de Weibull ( $Wbl(a, b)$ , avec  $a > 1$ ) et deux autres lois de type *NWUE* (une distribution de Weibull ( $Wbl(a, b)$ , avec  $a \leq 1$ ) et une distribution Gamma ( $\Gamma(a, b)$ , avec  $0 \leq a < 1$ ) pour les temps

de service. Ainsi que nous avons fixé le taux d'arrivées entrantes  $\lambda = 0.3$ , le taux d'arrivées sortantes  $\alpha = 0.2$ , le taux de rappels  $\mu = 1$ , le temps de simulation  $T_{max} = 1000$  unités de temps et  $n = 100$  (le nombre de replications) et prenant en considération les différentes lois citées dans le tableau suivant :

$B_1(x)$			
loi NBUE	loi exp	loi NWUE	
$Wbl(2, 4)$	$exp(2.00)$	$Wbl(0.5, 0.3699)$	$\Gamma(0.6, 4)$
$B_2(x)$			
loi NBUE	loi exp	loi NWUE	
$Wbl(3.3098, 4)$	$exp(3.00)$	$Wbl(0.5, 0.3333)$	$\Gamma(0.5, 6.0)$

TABLE 4.2 – Différentes situations prises en considération lors de simulation

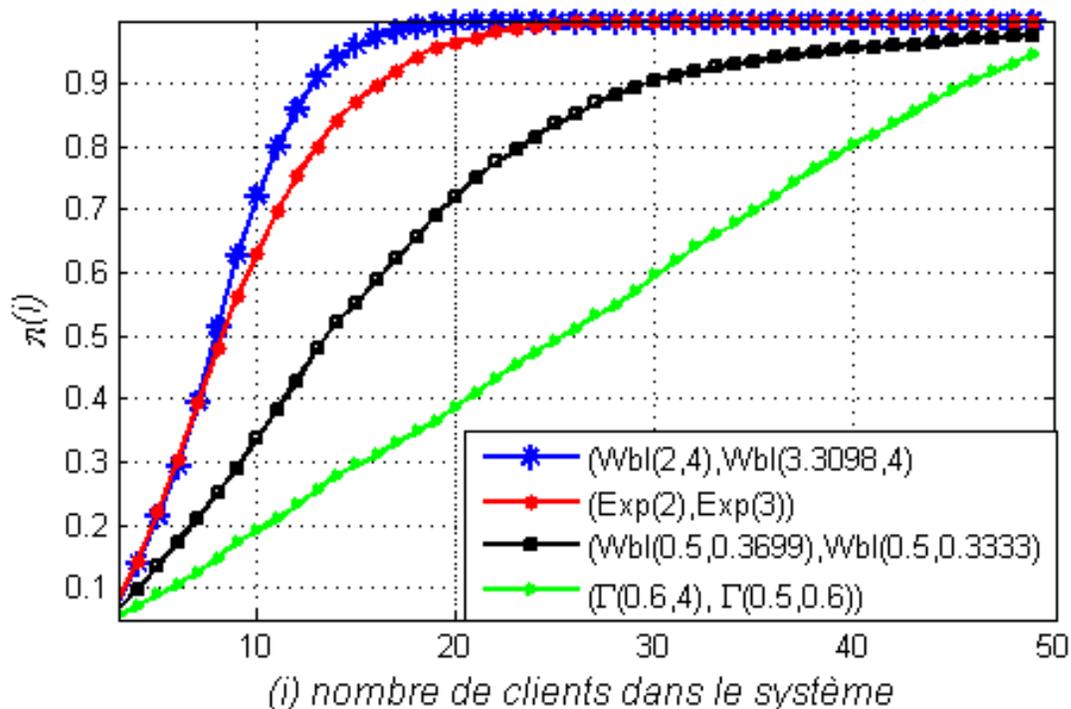


FIGURE 4.11 – Comparaison des probabilités stationnaires des systèmes

D'après cette Figure, on constate que :

La distribution stationnaire du nombre de clients dans le système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle est inférieure à la distribution stationnaire du nombre de clients

dans le système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle, lorsque la distribution des temps de service des appels entrants (resp. appels sortants) est *NBUE* c'est à dire  $Wbl(2,4)$  (resp.  $Wbl(3.3098,4)$ ).

La distribution stationnaire du nombre de clients dans le système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle est supérieure à la distribution stationnaire du nombre de clients dans le système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle avec distribution exponentielle des temps de service des appels entrants ( $exp(2.00)$ ) (resp. appels sortants ( $exp(3.00)$ )), lorsque la distribution des temps de service des appels entrants (resp. appels sortants) est *NWUE* ( $Wbl(0.5, 0.3699)$  ou  $\Gamma(0.6, 4)$ ) (resp. ( $Wbl(0.5, 0.3333)$  ou  $\Gamma(0.5, 6.0)$ )).

Cette Figure nous montre que les résultats théoriques obtenus dans le Théorème 4.8 sont confirmés par les résultats de simulation (une bonne concordance entre les résultats analytiques et ceux issus de la simulation). alors les bornes stochastiques en question sont une bonne approximation pour les probabilités stationnaires du modèle considéré dans notre étude et ce quelque soit la distribution des temps de service  $G$  (*NBUE* ou *NWUE*). Par conséquent, les performances d'un tel système (nombre moyen de clients dans le système, le temps d'attente moyen,...) peuvent être estimés par celles du système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle.

## Conclusion

Dans ce chapitre, nous avons étudié les propriétés de monotonie d'une file  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle [?] en utilisant la théorie générale des ordres stochastiques. Nous avons montré la monotonie de l'opérateur de transition de la chaîne de Markov incluse par rapport à l'ordre stochastique et convexe. De plus, nous avons obtenu des conditions de comparabilité des deux opérateurs de transition. Ainsi qu'on a montré que la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un système  $M_1, M_1/M_1, M_2/1$  avec rappels à communication bidirectionnelle, si la distribution des temps de service est *NBUE* (respectivement *NWUE*). Et enfin, on a confirmé les résultats théoriques obtenus par une application numérique.

# Conclusion générale

# Conclusion générale

Dans ce mémoire, nous avons montré, dans un premier temps, l'intérêt et les applications des systèmes d'attente (classiques, avec rappels et priorité). Les modèles classiques ne prennent pas en considération le phénomène de répétition de demandes de service qui exerce une influence non négligeable sur les caractéristiques de performance de certains systèmes réels, tels que les systèmes de télécommunication. Ce phénomène de répétition de demandes du service est étudié par la théorie de files d'attente avec rappels dont nous avons actualisé une synthèse des résultats connus. Nous nous sommes ensuite intéressés aux modèles avec rappels et à communication bidirectionnelle (cas particulier des systèmes d'attente avec rappels et priorité).

Nous avons ensuite étudié quelques problèmes de comparabilité pour l'analyse du système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle en utilisant la méthode de comparaison stochastique. L'avantage de ce type de méthodes d'approximation réside dans le fait que des résultats explicites puissent être obtenus pour des situations relativement complexes où les méthodes numériques et les expériences de simulation constituaient souvent la seule alternative.

Certaines caractéristiques de ce modèle sont en fait connues, mais ici, on considère deux problèmes :

1. La monotonie de la chaîne de Markov induite.
2. La comparabilité des modèles de ce type, mais ayant des paramètres différents.

Pour cela, on a établi des conditions de comparabilité et de monotonie sur les paramètres d'un système de files d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, qui assurent la monotonie de l'opérateur de transition associé à la chaîne de Markov incluse. On a aussi établi les conditions pour lesquelles les opérateurs de transition ainsi que les distributions stationnaires du nombre de clients, de deux chaînes de Markov incluses associées à deux systèmes  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle ayant la même

## Conclusion générale

---

structure mais avec des paramètres différents, sont comparables au sens des ordres stochastique et convexe. On a montré aussi que la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle, est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle, si la distribution des temps de service est NBUE (respectivement NWUE). De plus, on a conçu un simulateur pour illustrer les résultats théoriques obtenues (une bonne concordance entre les résultats analytiques et ceux issus de la simulation).

Parmi les perspectives de recherche, citons :

- Extension de ce travail en ajoutant une file d'attente prioritaire .
- Analyse mathématique du modèle d'attente  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle et distribution générale des temps inter-rappels.
- Étude de la V-stabilité du système  $M_1, M_2/M_1, M_2/1$  avec rappels à communication bidirectionnelle après perturbation de service afin d'obtenir les bornes de stabilité et de les comparer à ceux obtenues dans ce mémoire

# Bibliographie

- [1] J.R. Artalejo, T. Phung-Duc (2011). *Markovian single server retrial queues with two way communication*, in : Proceedings of the 6<sup>th</sup> International Conference on Queueing Theory and Network Applications, Seoul, pp. 1–7.
- [2] J. R. Artalejo and T. Phung-Duc (2013). *Single server retrial queues with two way communication*, Applied Mathematical Modelling. **37**, 1811-1822.
- [3] M. Boualem (2003). *Inégalités pour les systèmes d'attente avec rappels*, Thèse de Magister en Mathématiques Appliquées, Université des Sciences et de la Technologie Houari Boumediene, U.S.T.H.B., Alger.
- [4] M. Boualem (2013). *Insensitive Bounds for the Stationary Distribution of a Single Server Retrial Queue with Server Subject to Active Breakdowns*, Advances in Operations Research, revised.
- [5] M. Boualem and A. Aïssani (2004). *Inégalités pour les systèmes de files d'attente avec rappels et vacances du serveur*, Actes du Colloque International MSS'2004 (Modélisation Statistique et Stochastique), U.S.T.H.B., Alger, 52-57, ISBN : 978-9947-0-173.
- [6] M. Boualem, N. Djellab and D. Aïssani (2009). *Stochastic inequalities for  $M/G/1$  retrial queues with vacations and constant retrial policy*, International Journal MCM (Mathematical and Computer Modelling), Elsevier Ed, **50**, 207-212.
- [7] M. Boualem, N.Djellab and D. Aïssani (2012). *Stochastic approximations and monotonicity of a single server feedback retrial queue*, Mathematical Problems in Engineering, Vol.2012, Article ID 536982, 13 pages.
- [8] M. Boualem, N.Djellab and D. Aïssani (2013). *Stochastic bounds for a single server queue with general retrial times*, Bulletin of the Iranian mathematical Society, (in Press).
- [9] G. I. Falin and J. R. Artalejo and M. Martin (1993). *On the single server retrial queue with priority Customers*, Queueing Systems, **14**, 439-455.

## Résumé

Dans ce travail, nous avons présenté une synthèse sur les systèmes avec rappels et priorité. Le système étudié en est un cas particulier. Il s'agit du modèle d'attente avec rappels à communication bidirectionnelle. Ce type de systèmes diffère des systèmes classiques par l'existence de deux paramètres supplémentaires : rappels et nombre d'appels différent.

Nous avons étudié les propriétés de monotonie d'une file  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle en utilisant la théorie générale des ordres stochastiques. Nous avons d'abord montré la monotonie de l'opérateur de transition de la chaîne de Markov incluse par rapport à l'ordre stochastique et convexe. Par la suite, nous avons obtenu des conditions de comparabilité des deux opérateurs de transition. Nous avons ensuite montré que la distribution stationnaire du nombre de clients dans un système  $M_1, M_2/G_1, G_2/1$  avec rappels à communication bidirectionnelle est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un système  $M_1, M_1/M_1, M_2/1$  avec rappels à communication bidirectionnelle, si la distribution des temps de service est *NBUE* (respectivement *NWUE*). Enfin, nous avons confirmé les résultats théoriques obtenus par une application numérique.

**Mots-clés :** Files d'attente avec rappels ; Communication bidirectionnelle ; Chaîne de Markov induite ; Distribution stationnaire ; Ordre stochastique ; Bornes stochastiques ; Monotonie.

## ملخص

في هذه المذكرة، قدمنا ملخص عن انظمة الطوابير مع إعادة الطلب و الأولوية، النظام المدروس في هذا العمل عبارة عن حالة خاصة من هذه الأنظمة. يتعلق بنظام الطوابير مع إعادة الطلب مزدوج الإتصال. هذا النوع من الأنظمة يختلف عن باقي الأنظمة في وجود عامل إعادة الطلب و عدد الإتصالات المختلفة.

درسنا خصائص الإنتظام لنظام الطوابير مع إعادة الطلب ومزدوج الإتصال بإستعمال نظرية الأنظمة العشوائية. اولاً برهنا إنتظام عامل الإنتقال لسلسلة ماركوف المحتواة مقارنة مع النظام العشوائي و المحدب. بعدها، تحصلنا على شروط مقارنة عاملاً الإنتقال. ثم برهنا بأن التوزيع الثابت لعدد الزبائن في النظام المدروس اللامركوبي محدود من الأعلى -محدود من الأسفل على التوالي- بالتوزيع الثابت لعدد الزبائن في النظام المدروس المركوبي. و أخيراً أكدنا النتائج المتحصل عليها بتطبيق عددي.

**المصطلحات:** انظمة الطوابير مع إعادة الطلب، إتصال مزدوج، سلسلة ماركوف المحتواة، التوزيع الثابت، الأنظمة العشوائية، الحدود العشوائية، الإنتظام.