



**HAL**  
open science

# Processus de risque : modélisation de la dépendance et évaluation du risque sous des contraintes de convexité

Manel Kacem

► **To cite this version:**

Manel Kacem. Processus de risque : modélisation de la dépendance et évaluation du risque sous des contraintes de convexité. Mathématiques générales [math.GM]. Université Claude Bernard - Lyon I, 2013. Français. NNT : 2013LYO10051 . tel-01085025

**HAL Id: tel-01085025**

**<https://theses.hal.science/tel-01085025>**

Submitted on 20 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Processus de risque : modélisation de la dépendance et évaluation du risque sous des contraintes de convexité

## THÈSE

présentée et soutenue publiquement le 20 mars 2013

pour l'obtention du

**Doctorat de l'Université Claude Bernard Lyon I**  
(mathématiques appliquées)

par

Manel KACEM

### Composition du jury

*Rapporteurs :* Hélène COSSETTE, Professeure à l'Université de Laval  
Sana LOUHICHI, Professeure à l'Université Joseph Fourier de Grenoble

*Examineurs :* Claude LEFÈVRE, Professeur à l'Université Libre de Bruxelles  
Stéphane LOISEL, Professeur à l'Université Lyon 1 (co-directeur de thèse)  
Véronique MAUME-DESCHAMPS, Professeure à l'Université Lyon 1 (directrice de thèse)  
Christian-Yann ROBERT, Professeur à l'Université Lyon 1 (président du jury)



## Remerciements

Je tiens à remercier mes directeurs de thèse Véronique Maume-Deschamps et Stéphane Loisel pour leurs suivis précieux et pour les travaux que nous avons réalisés ensemble. Que le soutien permanent de Véronique, la confiance qu'elle m'a accordée tout au long de ces années ainsi que sa disponibilité trouvent ici l'expression de ma plus vive gratitude. Je remercie également Stéphane pour les échanges enrichissants que nous avons pu avoir ainsi que pour ses conseils et orientations toujours pertinents. Qu'ils trouvent ici le témoignage de mon respect profond.

Je suis très honorée qu'Hélène Cossette et Sana Louhichi aient accepté d'être rapporteurs de ma thèse. Merci à eux pour le temps consacré, leurs suggestions et les remarques précieuses qui m'ont permis d'améliorer ce travail. Je leur en suis très reconnaissante.

Je tiens à remercier également Christian-Yann Robert et Claude Lefèvre d'avoir accepté de faire partie de mon jury. Je remercie Christian pour ses remarques constructives et pour sa disponibilité. La collaboration avec Claude m'a été particulièrement enrichissante et ses conseils, d'une grande aide.

Merci à toutes les personnes qui composent l'équipe pédagogique de l'ISFA. Merci pour les discussions autour du café et leurs conseils durant ces années de thèse. Merci à Jean Claude Augros pour sa confiance et son soutien dès le début de ma thèse. Merci à Didier Rullière pour sa gentillesse et pour les échanges enrichissants qu'il m'a accordés. Merci à Areski Cousin pour les nombreux e-mails dont l'objet est " articles intéressants pour ta thèse ". Merci à Béatrice Rey-Fournier pour ses encouragements. Je tiens également à remercier Esterina Masiello et Pierre Ribereau pour leurs conseils, leurs convivialité et bienveillance.

Je tiens à remercier Christian Mazza pour les discussions qu'il m'a accordées et ses réflexions éclairées.

Merci à toute l'équipe administrative de l'ISFA pour leur disponibilité.

Je tiens également à souligner toute mon amitié à mes collègues doctorants et ATER Elena, Soffana, Anisa, Sawssen, Abdou, Mathieu, Romain, Alain, Christophe, Xavier, Yahia, Julien A., Mehdi, Julien V., Jean-Charles, Andrés, Rudy, Elias, Khader, Said, Ibrahima, Huy et à ceux que j'aurais oubliés. Comme le disait Romain " chacun par sa personnalité, a su amener l'élément indispensable à une ambiance, décontractée, qu'il est agréable de retrouver au quotidien ". Je n'oublie pas d'adresser un merci à Thomas Laloë.

Merci à Samy pour sa gentillesse, pour sa bonne humeur et pour son amitié. Je le remercie pour tous les repas de midi partagés au RU !

Un immense merci à mes deux grands amis Elena et Abdou qui m'ont accompagné durant toute ma thèse. Elena, je n'ai pas trouvé mieux que ton expression " tu es une personne-cadeau du ciel !" Tu m'es une amie formidable. Abdou, ton amitié m'est très chère ! Je salue la bonté du grand sage que tu es ! Vous étiez toujours là pour m'écouter, me con-

seiller et m'encourager ! Merci pour votre humeur chaleureuse et joyeuse !

Je ne crois pas qu'un simple merci suffirait pour remercier mes deux grandes amies Rym et Mouna. Votre amitié m'est très importante ! Merci pour tous les bons moments passés ensemble ! Merci de votre soutien au quotidien ! Tout simplement merci pour votre fraternité !

Merci à ma sœur Ines, à son mari et à leur petit garçon, merci à mes deux frères Hamdi et Aymen qui m'ont tous apporté l'équilibre nécessaire à mon épanouissement.

Du fond du cœur, je remercie mes chers parents. Je suis grandement reconnaissante pour tout ce qu'ils m'ont apporté ! Merci d'être toujours à mes côtés pour m'encourager et me soutenir durant mes nombreuses années d'études.

à ma mère *Mounira*,  
à mon père *Habib*,  
à ma grand-mère *Khadija*.



## Résumé

Ce travail de thèse porte principalement sur deux problématiques différentes mais qui ont pour point commun, la contribution à la modélisation et à la gestion du risque en actuariat.

Dans le premier thème de recherche abordé dans cette thèse, on s'intéresse à la modélisation de la dépendance en assurance et en particulier, on propose une extension des modèles à facteurs communs qui sont utilisés en assurance.

Dans le deuxième thème de recherche, on considère les distributions discrètes décroissantes et on s'intéresse à l'étude de l'effet de l'ajout de la contrainte de convexité sur les extrema convexes. Des applications en liaison avec la théorie de la ruine motivent notre intérêt pour ce sujet.

Dans l'introduction générale seront exposées en détail chacune de nos deux problématiques ainsi que nos principaux résultats. La thèse comportera deux parties chacune d'entre elles sera consacrée à l'étude de l'un des deux thèmes énoncés précédemment.

La première partie de la thèse concerne les modèles à facteurs pour la modélisation de la dépendance en assurance. Une propriété intéressante de ces modèles est que les variables sont conditionnellement indépendantes par rapport à un facteur. Dans la plupart des travaux trouvés dans la littérature le conditionnement se fait généralement par rapport à la dernière observation du facteur, on cite par exemple l'article de Cossette *et al.* (2002)[33]. On propose un nouveau modèle dans lequel le conditionnement se fait par rapport à toute la mémoire passée du facteur. La loi de la somme des variables aléatoires modélisant le risque a un grand intérêt en actuariat puisqu'elle permet de donner une vision sur le comportement global des risques. Dans le cadre de notre modèle de dépendance la problématique centrale est de montrer quelques conditions sous lesquelles la loi de la somme des variables aléatoires converge vers une loi normale.

Une première approche, dans laquelle nous avons utilisé les fonctions caractéristiques, nous a amené à conclure qu'une hypothèse de mélange conditionnel du processus de risque permet de répondre à la question centrale. De ce fait il était intéressant d'étudier en premier lieu les propriétés de mélange des processus conditionnellement indépendants par rapport à un facteur dynamique. C'est l'objet du premier chapitre dans lequel on considère que le conditionnement est fait par rapport à tout le passé du facteur qui est lui-même mélangeant. Le résultat principal de ce chapitre est énoncé dans le Théorème 1.2.1. Dans ce théorème, on donne quelques propriétés de mélange pour le processus de risque sous des conditions nécessaires liées aux propriétés de mélange du facteur et de mélange conditionnel du processus du risque.

Notons que dans ce chapitre nous avons donné une définition particulière pour les processus mélangeants qui est une variante de la définition donnée par Doukhan et Louhichi

(1999)[67] pour les processus faiblement dépendants. Rappelons que leur définition rend explicite l'indépendance asymptotique entre des fonctions du futur et du passé d'une séquence de variables aléatoires.

En général, la variance de la somme des variables aléatoires n'est pas connue en pratique. Pour que le théorème central limite soit appliqué nous avons besoin de l'estimer. L'objet du deuxième chapitre de la première partie sera donc de donner un estimateur de la variance pour les processus vérifiant notre propriété de mélange et donc donner un théorème central limite auto-normalisé. Nous commençons d'abord par donner des inégalités de covariance et de moments adaptées avec notre type particulier de mélange. Les résultats sont donnés dans la Proposition 2.2.1 et le Lemme 2.2.2. Par la suite nous énonçons le Théorème 2.3.1 dans lequel on donne un théorème central limite adapté au type de mélange que nos processus satisfont. Dans le Théorème 2.3.2 on montre la faible consistance de notre estimateur proposé. Notons que cet estimateur est une variante de l'estimateur de Peligrad et Shao (1995)[149]. Aussi, on donne le Théorème 2.3.3 dans lequel on prouve la normalité asymptotique de notre estimateur de la variance. A la fin de chaque chapitre, nous présentons des modèles pertinents dans le contexte d'assurances qui vérifient nos hypothèses.

Dans la deuxième partie de la thèse on considère la classe des distributions discrètes dont la fonction de masse de probabilité (f.m.p.) est décroissante dans un support  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Les extrema convexes dans cette classe de distribution sont connus. Notre but est de souligner comment les contraintes de forme supplémentaires de type convexité modifient ces extrema. Les deux cas suivants sont considérés : la f.m.p. est globalement convexe dans  $\mathbb{N}$  ou elle est convexe uniquement à partir d'un point  $m$  fixé. Les extrema convexes correspondants sont calculés en utilisant une simple propriété de croisement entre les f.m.p. Dans le troisième chapitre on donne les principaux résultats liés à cette partie de la thèse. Plusieurs illustrations en théorie de la ruine sont présentées. Ces résultats fournissent un complément aux deux travaux récents de Lefèvre et Loisel (2010,2012)[119; 120].

## Abstract

In this thesis we focus on two different problems which have as common point the contribution to the modelling and to the risk management in insurance. In the first research theme, we are interested by the modelling of the dependence in insurance. In particular we propose an extension to model with common factor. In the second research theme we consider the class of nonincreasing discrete distributions and we are interested in studying the effect of additional constraint of convexity on the convex extrema. Some applications in ruin theory motivate our interest to this subject. In the general introduction will be exposed in detail each of our two research themes as well as our main results. Afterward, the thesis contains two parts each of them will be devoted to the study of one of the two themes mentioned above.

The first part of this thesis is concerned with factor models for the modeling of the dependency in insurance. An interesting property of these models is that the random variables are conditionally independent with respect to a factor. In most works found in the literature the conditioning is with respect to the last observation of the factor (see for example Cossette *et al.* (2002)[33]). We propose a new model in which the conditioning is with respect to the entire memory of the factor. The law of the sum of random variables has a great interest in actuarial science. In fact this law allows to give an overview on the overall behavior of risk. In our framework our main problem is to give some conditions under which the law of the sum of random variables converges to a normal distribution.

We use a first approach based on characteristic functions. This led us to conclude that the additional hypothesis of a conditional mixing risk process can give us a response to the main question. Therefore, when the conditioning is with respect to a dynamic factor, it was interesting to give some mixing properties of the risk process. This is the subject of the first chapter in which we consider that the conditioning is with respect to a unbounded memory of a factor. The main result of this chapter is stated in Theorem 1.2.1. In this last we give some mixing properties of risk process under conditions related to the mixing properties of the factor process and to the conditional mixing risk process. Note that in this chapter we have given a specific definition for mixing processes which is a variant of the definition introduced by Doukhan and Louhichi (1999)[67] for weakly dependent processes. We Recall that their definition makes explicit the asymptotic independence between the functions of the future and the past of a sequence of random variables.

In general, the variance of the sum of random variables is not known in practice. We need to estimate it in order to make the central limit theorem applicable. The purpose of the second chapter of the first part is to give an estimate for the variance of processes satisfying our mixing condition. We begin by giving suitable moments and covariance inequalities witch are adapted to our particular kind of mixing. The results are given in Proposition 2.2.1 and Lemma 2.2.2.

Subsequently we state our main results which are Theorems 2.3.1, 2.3.2 and 2.3.3. In

Theorem 2.3.1 we provide a central limit theorem adapted with our structure of mixing. In Theorem 2.3.2 we derive the consistency of our proposed estimator. Note that this estimator is a variant of the estimator of Peligrad and Shao (1995)[149]. In Theorem 2.3.3 we prove the asymptotic normality of our estimator of the variance. At the end of each chapter, we present some relevant models in insurance that satisfy our assumptions.

In the second part of the thesis we consider the class of discrete distributions whose probability mass functions (p.m.f.) are nonincreasing on a support  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Convex extrema in that class of distributions are well-known. Our purpose is to point out how additional shape constraints of convexity type modify these extrema. Two cases are considered: the p.m.f. is globally convex on  $\mathbb{N}$  or it is convex only from a given positive point  $m$ . The corresponding convex extrema are derived by using a simple crossing property between two distributions. Several applications to some ruin problems are presented for illustration. These results provide a complement to the two recent works by Lefèvre and Loisel (2010,2012)[119; 120].

*Ecole ISFA - 50, av. Tony Garnier - 69007 Lyon - France*

# Table des matières

---

---

## Introduction générale

---

---

1	Modèles de dépendance incluant des processus de mémoire non bornée .	5
1.1	Introduction . . . . .	5
1.2	Problématiques de recherche . . . . .	7
1.3	Modèles classiques de la théorie de la ruine . . . . .	8
1.4	Modélisation de la dépendance en théorie du risque : une revue .	10
1.5	Modèles à facteurs commun : notre extension . . . . .	16
1.6	Notion de mélange et ses principales propriétés . . . . .	19
1.6.1	Suites mélangeantes : une revue . . . . .	19
1.6.2	Suites conditionnellement mélangeantes : notre contexte	22
1.7	Inégalités de covariance . . . . .	23
1.8	Inégalités des moments . . . . .	25
1.8.1	Inégalités de type Marcinkiewicz-Zygmund . . . . .	25
1.8.2	Inégalités de type Rosenthal . . . . .	26
1.9	T.C.L. pour des suites de variables aléatoires conditionnellement indépendantes par rapport à un facteur . . . . .	28
1.10	Principaux résultats asymptotiques des processus mélangeants : une revue . . . . .	29
1.10.1	T.C.L. pour des processus mélangeants . . . . .	30
1.10.2	Estimation de la variance et T.C.L. autonormalisé . . . .	32
2	Extrema convexes pour des distributions discrètes décroissantes . . . . .	35
2.1	Introduction . . . . .	35

2.2	Problématiques de recherche . . . . .	36
2.3	Littérature : quelques résultats de base . . . . .	38
2.4	Principaux résultats . . . . .	38

---



---

**Partie I Modèle à facteur commun pour la modélisation de la dépendance en théorie du risque**

---



---

**Chapitre 1**

**Quelques propriétés de mélange pour des processus C.I.**

1.1	Introduction . . . . .	43
1.2	Definitions and statement of results. . . . .	44
1.2.1	Mixing sequences . . . . .	44
1.2.2	Conditional independence . . . . .	46
1.2.3	Conditionally mixing sequences given an unbounded memory . . . . .	47
1.2.4	Transition kernel . . . . .	48
1.2.5	Main result . . . . .	50
1.3	Proof of the main result . . . . .	52
1.4	Examples and some applications . . . . .	53
1.4.1	Mixing properties for sequence controlled by a bounded memory of the factor . . . . .	54
1.4.2	Mixing properties for sequences controlled by a factor with an unbounded memory . . . . .	54
1.4.2.1	Example in a discrete case . . . . .	55
1.4.2.2	Example in an absolutely continuous case . . . . .	57
1.4.2.3	Models for applications of the absolutely continuous case . . . . .	59
1.4.2.4	Mixed cases . . . . .	60

**Chapitre 2**

**Estimation de la variance pour des processus mélangeants**

2.1	Introduction . . . . .	65
2.2	Main tools : moment inequalities, covariance inequality and uniform integrability . . . . .	67

---

2.2.1	Covariance inequality . . . . .	67
2.2.2	Moment inequalities . . . . .	68
2.2.3	Uniform integrability . . . . .	70
2.3	Self Normalized Central Limit Theorem for stationary $\alpha$ mixing sequences	71
2.4	Proof of the main results . . . . .	74
2.5	Asymptotic stationarity and conditionally independent processes . . . .	94
2.6	Applications . . . . .	95
2.6.1	Example in a discrete case . . . . .	95
2.6.2	Examples in an absolutely continuous case . . . . .	100

---



---

**Partie II Extrema convexes pour des distributions discrètes décroissantes: effets de la contrainte de convexité**

---



---

**Chapitre 3**

**Extrema convexes pour des distributions discrètes décroissantes**

3.1	Introduction . . . . .	113
3.2	Crossing properties and convex ordering . . . . .	114
3.3	For decreasing p.m.f. that are convex on $\mathbb{N}$ . . . . .	116
3.3.1	The upper bound . . . . .	116
3.3.2	The lower bound . . . . .	119
3.4	For decreasing p.m.f. that are convex in $[1, +\infty)$ . . . . .	120
3.4.1	The upper bound . . . . .	120
3.4.2	The lower bound . . . . .	122
3.4.2.1	Case : $\frac{1}{3} \leq \nu \leq \frac{1}{2}$ . . . . .	122
3.4.2.2	Case : $\xi < 3\nu \leq \xi + 1 + \frac{2}{(\xi+4)}$ where $3\nu - \xi > \frac{2}{\xi+3}$ for all $\xi \geq 1$ and $n \geq \xi + 1$ . . . . .	122
3.5	For decreasing p.m.f. that are convex in $[m, +\infty)$ for fixed $m > 1$ . . . .	124
3.5.1	The upper bound . . . . .	124
3.5.2	The lower bound . . . . .	125
3.5.2.1	Case : $0 \leq \xi < 2\nu \leq (\xi + 1) \leq m$ . . . . .	125

3.5.2.2	Case : $m < \xi \leq 3\nu \leq \xi + 1$ , $\frac{m(m+1)}{m+\xi+2} \leq 3\nu - \xi$ and $n \geq \xi + 1$ . . . . .	126
3.6	Some numerical illustration . . . . .	126
3.6.1	Convex extrema for decreasing convex distribution globally on $\mathbb{N}$ . . . . .	128
3.6.2	Convex extrema for decreasing distributions that are concave until 2 and convex on $[1, \infty)$ . . . . .	129
3.6.3	Decreasing distribution globally on $\mathbb{N}$ and convex on $[m, +\infty)$ for fixed $m \geq 1$ . . . . .	129

<b>Conclusion et perspectives</b>
-----------------------------------

<b>Appendix A</b>
<b>Detailed proofs of some results of the first part</b>

A.1	Proof of equations 31 and 32 of the introduction . . . . .	138
A.2	Model of application in a continuous case . . . . .	139
A.3	Complement to the proof of Theorem 1.2.1 . . . . .	142
A.4	Complement to the proof of Proposition 1.4.2 . . . . .	144
A.5	Proof of equations (2.97) and (2.95) . . . . .	146

<b>Appendix B</b>
<b>Simulations and some examples related to the first part</b>

B.1	Choice of $x$ in Theorems 2.3.2 and 2.3.3 . . . . .	150
B.2	Discrete case: the factor $V$ is a Bernoulli r.v. with parameter $q$ . . . . .	151
B.3	Application when the factor $V$ is a Markov chain . . . . .	153

<b>Appendix C</b>
<b>Proofs of some results of the second part</b>

C.1	Proof of Proposition 3.4.1 . . . . .	156
C.2	Proof of Proposition 3.4.3 . . . . .	159
C.3	Proof of Proposition 3.5.1 . . . . .	161
C.4	Proof of Proposition 3.5.4 . . . . .	162
C.5	Proof of Proposition 3.5.5 . . . . .	164

<b>Bibliographie</b>
----------------------

# Table des figures

2.1	Normfit and QQplot of the normalized sum $S_n$ : the case of Bernoulli sample data with $m = 1000$ , $n = 20000$ , $q = 0.6$ , $x = 60$ . . . . .	98
2.2	Normfit and QQplot of the normalized sum $S_n$ in the Pareto model, $m = 1000$ , $n = 10000$ , $q = 0.5$ , $x = 50$ , $\beta = 4$ . . . . .	101
2.3	Normfit and QQplot of the normalized sum $S_n$ in the Pareto model, $m = 1000$ , $n = 10000$ , $q = 0.5$ , $x = 50$ , $\beta = 5$ . . . . .	102
3.1	Convex extrema in the class of discrete r.v.'s with decreasing p.m.f. on a support $\mathcal{D}_n$ , $\nu = 0.69$ and $n = 10$ . . . . .	114
3.2	An example where $S^-(P_Y - P_X) = 2$ and $S^-(F_Y - F_X) = 1$ . . . . .	115
3.3	Crossing situation, $j = s = 4$ and $n = 25$ . . . . .	117
3.4	The upper bound (3.9) vs the upper bound (3.3), $\nu = 0.69$ and $n = 10$ . . . . .	118
3.5	The lower bound (3.11) vs the lower bound (3.4), $\nu = 0.69$ and $n = 10$ . . . . .	119
3.6	Convex extrema in the class of decreasing p.m.f. that are convex on $\mathbb{N}$ with $\nu = 0,69$ et $n = 10$ . . . . .	121
3.7	The upper bound (3.15) vs the upper bound (3.9) with $\nu = 0.69$ , $n = 10$ . . . . .	122
3.8	The lower bound (3.19) vs the lower bound (3.11), $\nu = 0.69$ , $n = 10$ . . . . .	123
3.9	The lower bound (3.19) vs the upper bound (3.15), $\nu = 0.69$ , $n = 10$ . . . . .	124
3.10	The upper bound (3.21), $m = 8$ , $\nu = 3.8$ , $n = 10$ . . . . .	125
3.11	The lower bounds for decreasing p.m.f. that are convex on $[m, +\infty)$ , $\nu = 3.8$ , $n = 10$ for $m = 2$ and $m = 1$ , $m > 7$ . . . . .	127
3.12	$(\tilde{N}_{\max})$ for differents $m$ and $n$ , $\nu = 0.37$ . . . . .	131
B.1	Estimation of $\sigma$ of the sum $S_k$ over time k where $k = 35, \dots, 1000$ for different values of $x$ . . . . .	150
B.2	Behaviour of $\text{Var}(I_i)$ , $i = 1 \dots, 30$ . . . . .	151
B.3	Behaviour of $\text{Cov}(I_i, I_j)$ , $i$ is fixed, $j > i$ , $q = 0.9$ . . . . .	152
B.4	Behaviour of $\text{Cov}(I_1, I_i)$ , $i = 2, \dots, 12$ , $q = 0.9$ , $q = 0.6$ , $q = 0.3$ . . . . .	152
B.5	Normfit and QQplot of the normalized sum $S_n$ of the Bernoulli sample data, $m = 1000$ , $n = 1000$ , (the factor is a Markov chain) . . . . .	153
C.1	The upper bound (3.15) vs the case 1.1, $\nu = 0.69$ , $n = 10$ . . . . .	157
C.2	The upper bound (3.15) vs the case 1.2, $\nu = 0.69$ , $n = 10$ . . . . .	157
C.3	The upper bound (3.15) vs the case 2.1, $\nu = 0.69$ , $n = 10$ . . . . .	158
C.4	The lower bound (3.19) vs the case 1, $\nu = 0.69$ , $n = 10$ . . . . .	160
C.5	The lower bound (3.19) vs the case 2, $\nu = 0.69$ , $n = 10$ . . . . .	160
C.6	The upper bound (3.21) vs the case 2.1, $\nu = 0.69$ , $n = 10$ . . . . .	162

*Table des figures*

---

C.7	The lower bound (3.23) vs the case 1.1 , $m = 8, \nu = 3.8, n = 10$ . . . . .	163
C.8	The lower bound (3.23) vs the case 1.2. , $m = 8, \nu = 3.8, n = 10$ . . . . .	163
C.9	The lower bound (3.23) vs the case 2., $m = 8, \nu = 3.8, n = 10$ . . . . .	164
C.10	The lower bound (3.25) vs the case 1, $m = 2, \nu = 3.8, n = 14$ . . . . .	165
C.11	The lower bound (3.25) vs case 2, $m = 2, \nu = 3.8, n = 14$ . . . . .	166

# Introduction générale



## Présentation de la thèse

Cette thèse se compose de deux parties principales. La première partie porte sur les modèles à facteur commun pour la modélisation de la dépendance en assurance. La deuxième partie porte sur l'évaluation du risque sous des contraintes de convexité des fonctions de masse de probabilité.

Dans la première partie on considère un modèle de risque en temps discret dans lequel les variables aléatoires sont dépendantes mais conditionnellement indépendantes par rapport à un facteur commun. Dans ce cadre de dépendance on introduit un nouveau concept pour la modélisation de la dépendance temporelle entre les risques d'un portefeuille d'assurance. En effet, notre modélisation inclut des processus de mémoire non bornée. Plus précisément, le conditionnement se fait par rapport à un vecteur aléatoire de longueur variable au cours du temps. On note que dans notre étude asymptotique c'est plutôt le temps qui tend vers l'infini que le nombre de risques. On donne des résultats asymptotiques pour le processus agrégé, ce qui permet de donner une approximation du risque d'une compagnie d'assurance lorsque le temps tend vers l'infini.

Notre contexte de dépendance temporelle généralise les définitions de modèles à facteur puisque la mémoire du conditionnement peut augmenter d'une période à la suivante. Cette propriété rend complexe notre étude. Sous des conditions de mélange du facteur et d'une structure de mélange conditionnel, nous avons obtenu des propriétés de mélange pour les processus non conditionnels. Avec ces résultats on peut obtenir des propriétés asymptotiques intéressantes. Nous verrons dans la suite de cette thèse que le type de mélange satisfait par nos processus est plus général que les formes de mélange classiques. Ainsi les résultats asymptotiques qui existent dans la littérature sur les processus mélangeants ne peuvent pas être appliqués directement. Des exemples d'applications en théorie du risque de notre structure de dépendance et de nos résultats asymptotiques sont présentés.

Dans cette première partie nous supposons disposer des informations statistiques complètes pour les distributions de pertes. En assurance et en finance les risques et les pertes sont difficiles à estimer et nous pouvons parfois faire face à des situations dans lesquelles on dispose uniquement d'une connaissance statistique incomplète sur certaines composantes des modèles. Par exemple, les informations partielles à notre disposition sur les distributions de sinistres peuvent être les premiers moments ou différentes contraintes de forme comme la décroissance ou la convexité.

Dans ce contexte, la deuxième partie porte sur l'effet de la contrainte de convexité sur les extrema convexes dans la classe des distributions discrètes dont les fonctions de masse de probabilité (f.m.p.) sont décroissantes sur un support  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Les extrema convexes dans cette classe de distributions sont bien connus (voir Denuit *et al.* (1999)[57]). Notre but est de souligner comment les contraintes de forme supplémentaires de type convexité modifient ces extrema. Deux cas sont considérés pour la f.m.p. :

1. Globalement convexe sur  $\mathbb{N}$  ;
2. Convexe seulement à partir d'un point positif donné  $m$ .

Les extrema convexes correspondants sont calculés en utilisant de simples propriétés de

croisement entre deux distributions. Plusieurs illustrations en théorie de la ruine sont présentées. Ces résultats fournissent un complément aux deux travaux récents de Lefèvre et Loisel (2010,2012)[119; 120].

Dans la suite de cette introduction nous annonçons pour chacune des deux parties de ce travail les problématiques, une revue de la littérature des principaux résultats qui y sont liés ainsi que nos principales contributions en théorie du risque.

# 1 Modèles à facteur en théorie du risque : indépendance conditionnelle, dépendance temporelle et quelques propriétés asymptotiques

## 1.1 Introduction

La théorie du risque en général et la théorie de la ruine en particulier sont traditionnellement considérées comme faisant partie des mathématiques de l'assurance. Elles ont pour objectif de fournir et de développer différents outils mathématiques pour l'évaluation et la quantification des risques souscrits par les organismes d'assurance et de réassurance. Depuis les travaux de Lundberg (1903,1926)[127; 128] et Cramèr (1930)[40] et jusqu'à aujourd'hui la recherche dans ce domaine est très active. Des modèles mathématiques sont sans cesse développés pour une meilleure évaluation et compréhension des risques. Voici une liste non exhaustive d'ouvrages de référence qui donnent un aperçu général des résultats et des fondements de la théorie du risque : Gerber (1979)[81], De Vylder (1996)[44], Rolski *et al.* (1999)[160], Asmussen (2000)[7], Goovaerts *et al.* (2001)[83], McNeil *et al.* (2005)[135], Denuit et Charpentier (2004)[50], Denuit et Charpentier (2005)[51], Denuit *et al.* (2005)[53], Dickson (2005)[62], Klugman *et al.* (2008)[115], Marceau (2013)[129], Mikosch (2009)[138] et Asmussen *et al.* (2010)[8].

Dans le modèle classique de la théorie de la ruine l'hypothèse d'indépendance entre les différents risques d'un portefeuille d'assurance joue un rôle important. En réalité, cette hypothèse est très restrictive puisque dans de nombreuses situations, les risques sont corrélés. En effet, il est connu que la gestion des risques en assurance fait face à une multitude de facteurs de risques qui peuvent introduire la dépendance entre les risques assurés. Cette dépendance a un impact sur la mutualisation des risques. Désormais, pour une meilleure prise en compte du risque, il est crucial de considérer la dépendance dans la modélisation des risques assurés.

En sciences actuarielles et en théorie du risque, on trouve une littérature abondante sur les sujets qui traitent la structure de la dépendance entre les risques. Plusieurs modèles ont été développés tels que les modèles fondés sur le concept de copule, les modèles à chocs communs, les modèles dans un environnement markovien et les modèles avec mélange commun. Notons que les modèles avec mélange commun en théorie du risque sont connus sous la dénomination de modèle à facteur en finance. En effet, l'approche factorielle a eu un grand succès dans la modélisation de la dépendance d'un portefeuille de crédit. Wilson (1997)[183], Figer (2001)[76], Gordy (2003)[88], Frey et McNeil (2003)[79] et Embrechts *et al.* (2005)[74] ont notamment contribué au développement de cette approche et à sa mise en œuvre dans le cadre de la réglementation Bâle II (voir aussi les références bibliographiques de la thèse de Cousin (2008)[39]).

En assurance, dans les modèles avec mélange commun on considère que la dépendance est introduite via un ou plusieurs facteurs qui peuvent affecter l'ensemble ou des parties du portefeuille. Ces facteurs peuvent représenter des situations économiques, juridiques ou politiques, des lieux géographiques, des conditions climatiques ou environnementales ou tout autres facteurs externes. Ce type de dépendance peut être introduit par un mécanisme externe qui doit influencer sur le portefeuille d'assurance entier. Soit  $v$  une réalisation d'une

variable aléatoire positive  $V$  discrète ou continue qui représente le facteur. Le mécanisme externe peut être représenté par la réalisation  $v$  qui représente l'influence du facteur externe. Une propriété importante de ces modèles est la suivante : les variables aléatoires sont dépendantes mais conditionnellement indépendantes par rapport au facteur. Considérant par exemple le cas d'un portefeuille d'assurance vie, la dépendance entre les risques individuels peut découler de l'état de l'environnement socio-économique lorsque tous les risques sont exposés au même environnement. De plus, si on considère un portefeuille d'assurance maladie alors la relation de dépendance peut être présente pour l'ensemble du portefeuille via le risque inflation par exemple.

Inspirés des modèles avec mélange commun classiques, l'idée dans cette thèse est de considérer que l'ensemble des risques d'un portefeuille d'assurance sont influencés par une mémoire de longueur non bornée d'un facteur externe. Plus précisément on considère que le risque observé à un instant donné est influencé par tout le passé du facteur jusqu'à la date d'observation du risque. De ce fait la mémoire du facteur grossit chaque fois que le temps avance d'où notre appellation de mémoire non bornée. Cette structure particulière de modélisation de la dépendance temporelle peut donc être qualifiée de dynamique. Notre motivation pour de tels modèles est expliquée par leur intérêt dans la modélisation des processus de risques en présence d'une dépendance temporelle. A notre connaissance, les processus de variables aléatoires conditionnellement indépendantes par rapport à un vecteur aléatoire de longueur variable au cours du temps ne sont pas considérés en littérature.

Dans la littérature actuarielle, il est intéressant de mesurer le risque global d'un portefeuille d'assurance. Soit  $(X_i, i = 1, \dots, n)$  une suite de variables aléatoires modélisant le montant total de sinistres dans une période fixée  $i$ , dans ce cas  $S_n = X_1 + \dots + X_n$  modélise la somme agrégée des risques sur  $n$  périodes. Dans ce travail on s'intéresse au comportement asymptotique de la somme  $S_n$  lorsque les variables aléatoires  $(X_i)_{i \in \mathbb{N}}$  sont telles qu'elles étaient décrites dans le paragraphe précédent. Rappelons que dans ces modèles les variables aléatoires sont considérées conditionnellement indépendantes par rapport à des vecteurs qui s'intersectent. De ce fait, afin de mener notre étude il nous était nécessaire de mettre des conditions supplémentaires sur le comportement du processus du facteur.

En statistiques, l'un des concepts de dépendance le plus largement utilisé est la notion de mélange fort présentée pour la première fois par Rosenblatt (1956)[161], qui est une mesure de la baisse de la dépendance entre deux événements séparés par des laps de temps plus longs. Les processus mélangeants ont des propriétés probabilistes intéressantes principalement dans le cas où ils sont stationnaires. En effet, la propriété de mélange pour une suite de variables aléatoires dépendantes est un outil utile pour étendre certains résultats limites qui existent sous l'hypothèse classique d'indépendance. En effet, par un simple contrôle de la vitesse avec laquelle la dépendance entre le passé et le futur d'une suite de variables aléatoires décroît, il est possible de donner des résultats type théorème central limite (T.C.L.) et loi forte des grands nombres pour des suites satisfaisant ce type de dépendance faible.

La loi des grands nombres et le théorème central limite sont fréquemment évoqués lorsqu'on parle d'assurabilité. Charpentier (2011)[30] a étudié l'application de ces résultats en assurance lorsque les risques sont mutualisés c'est-à-dire lorsque le nombre de risque

composant le portefeuille d'assurance est assez grand. L'interprétation de ce théorème est simple : lorsque le nombre de risques  $n$  tend vers l'infini, la loi de la charge totale peut-être approximée par une loi normale si certaines conditions sont vérifiées. Dans notre travail on s'intéresse à l'application de ces résultats limites lorsque le temps tend vers l'infini. Dans ce cas, le T.C.L. permet en particulier de donner une estimation lorsque le temps  $n$  tend vers l'infini, de la somme, sur  $n$  périodes, des risques d'un portefeuille d'assurance.

En mathématiques actuarielles le T.C.L. permet de contrôler l'écart entre la charge totale et la prime pure et indique que la dispersion autour de la prime pure est distribuée selon une loi normale et que l'écart-type de cette loi est pris en compte dans l'estimation des chargements de sécurité des primes.

En actuariat comme en finance, la base de la mesure du risque est la volatilité. Elle est par définition une mesure des amplitudes des variations (positives ou négatives) par rapport à la moyenne. La volatilité historique est déterminée par la mesure de dispersion « l'écart-type ». Cette mesure reste parmi les mesures les plus utilisées en finance et en actuariat vue la simplicité de son calcul.

L'argument principal, en faveur de cette mesure, repose sur la normalité de la distribution des risques. En effet, dans ce cadre restreint, l'espérance et la variance sont suffisantes pour caractériser le risque. On dit que la variance traduit la notion d'incertitude. Plus la variance est élevée et plus la valeur du risque est susceptible de s'éloigner de sa moyenne. Suite au travail de Markowitz (1952)[130], la variance a eu un grand succès dans la théorie de la gestion des portefeuilles financiers. En actuariat, il est possible que le capital de solvabilité requis soit lié à la volatilité du risque. Dans ce cas, le T.C.L. trouve son intérêt dans l'estimation de la marge de risque et donc du besoin en capital.

Dans le paragraphe suivant nous allons préciser les problématiques de notre étude.

## 1.2 Problématiques de recherche

La première problématique explorée dans cette première partie de la thèse correspond à l'étude des propriétés de mélange des processus conditionnellement indépendants. Notre investigation englobe bien évidemment le cas où le conditionnement est par rapport à des vecteurs de longueur variable au cours du temps. Aussi on présente quelques exemples de modèles de risque qui vérifient nos hypothèses. Cette problématique fera l'objet du premier chapitre.

En continuité avec la première problématique, la deuxième problématique consiste à donner des résultats asymptotiques pour le processus de risque agrégé  $S_n$ . Nous étudions la convergence de  $S_n$  vers une loi normale et nous proposons un estimateur pour la variance asymptotique de  $S_n$ . On note que la particularité des coefficients de mélange auxquels obéissent nos processus de risque rend notre étude complexe. En effet, nos coefficients de mélange sont des fonctions exponentielles du passé et du futur. Ainsi les résultats limites de la théorie des processus mélangeants classiques doivent être adaptés au type de mélange considéré. Cette deuxième problématique fera l'objet du développement présenté dans le deuxième chapitre.

Dans la suite nous développons les points évoqués ci-dessous. D'abord on introduit le modèle classique de la théorie de la ruine ainsi que les principaux modèles classiques utilisés en actuariat pour la modélisation des risques. Par la suite on présente une revue de quelques modèles de dépendance pour les processus de risque, en particulier ceux à partir desquels est inspirée notre extension. Puis, on donne une formalisation mathématique de nos problématiques de recherche. Ensuite, on présente le concept des processus mélangeants qui est fondamental dans la formalisation de nos résultats. Enfin, on présente une revue des principaux résultats sur lesquels s'est appuyée cette thèse ainsi que nos principaux résultats.

### 1.3 Modèles classiques de la théorie de la ruine

Le risque engendré par un portefeuille de contrats d'assurance est difficile à évaluer, mais il est néanmoins important de tenter de le faire afin d'assurer la viabilité d'une opération d'assurance. Le projet de réforme européenne Solvabilité II sur les réglementations des compagnies d'assurances et de réassurances devrait entrer en vigueur en janvier 2014. Ce projet de réforme s'appuie principalement sur la prise en compte de l'exposition aux risques et exige que le capital de solvabilité requis (SCR) soit calculé sur un horizon d'une année. Bien que la réglementation exige un horizon d'une année, les modèles internes des organismes assureurs sont développés de manière à pouvoir contrôler le risque sur le moyen et le long terme. De ce fait, il est important de considérer une approche multi-périodique dans laquelle la richesse d'un contrat, d'un portefeuille ou d'une société sont suivis au cours du temps. L'utilisation la plus commune de cette approche est la théorie de la ruine, dans laquelle on s'intéresse au montant du surplus également appelé montant des réserves. Notons que la ruine se produit lorsque le niveau de réserve devient négatif. Pour suivre l'évolution du surplus il ne faut pas se restreindre à la modélisation des montants de sinistres uniquement mais il faut aussi inclure les primes, les revenus des placements, les frais ainsi que tout autre élément qui a un impact sur les flux de trésorerie. Afin de maintenir la simplicité mathématique, le modèle qu'on présente ici est un modèle basique.

Notre étude concerne le comportement asymptotique du processus de la perte totale  $S_n = X_1 + \dots + X_n$  dans un cadre particulier de dépendance qui sera présenté dans la suite de cette introduction. Notre travail peut être vu comme une étape dans l'étude des processus de risque conditionnellement indépendants.

#### Modèle en temps discret

Notons par  $\{R(t) : t = 0, 1, \dots\}$ , le processus du risque en temps discret qui mesure le surplus d'un portefeuille d'assurance à la fin de chaque période  $t$ . On suppose qu'à l'instant zéro, la compagnie d'assurance dispose d'une réserve initiale  $u$  pour absorber un éventuel excès de la sinistralité et que la compagnie collecte des cotisations de ses assurés, aussi appelées primes. Pour plus de simplicité, on suppose que le montant de primes sur une période est égale à une constante  $c$ . On considère le montant des réserves de la compagnie d'assurances à la fin d'un certain nombre de périodes de longueur égales. Par conséquent soient  $\{X_1, X_2, \dots\}$  les montants de la perte totale sur chaque période. Les  $(X_i)_{i \geq 1}$  sont des variables aléatoires positives supposées indépendantes et identiquement distribuées (i.i.d.). Ainsi le montant de la réserve après la  $t$  ème période est donné par

$$R(t) = u + c \times t - S_t, \tag{1}$$

avec  $S_t = \sum_{i=1}^t X_i$ . Il est insuffisant de décrire le processus  $\{X_i : i = 0, 1, \dots\}$  en spécifiant la distribution de  $X_i$  pour  $i$  arbitraire puisque dans de nombreux processus on a une corrélation entre les valeurs observées à des moments différents.

Nous avons présenté le modèle en temps discret puisque c'est le cadre général dans lequel on se place. On n'a pas présenté le modèle en temps continu puisque ça ne relève pas de ce qu'on fait.

## Modèle collectif et modèle individuel

Traditionnellement, le calcul ou l'approximation de la fonction de répartition du montant global de sinistres est l'un des points centraux des mathématiques de l'assurance. Afin d'étudier la distribution du montant global des risques, il est habituel de considérer l'un des deux modèles suivants : le modèle individuel ou le modèle collectif.

### Modèle individuel

Dans ce modèle largement utilisé dans la modélisation des contrats d'assurances de personnes (vie, santé, risque de crédit), les contrats d'une même classe sont individualisés. Soient  $(X_1, \dots, X_n)$  les risques individuels d'un portefeuille d'assurance observés sur une période de temps donnée. On suppose que les risques individuels sont positifs, indépendants mais non nécessairement identiquement distribués. Soit  $F_{X_i}$  la fonction de distribution de  $X_i$  définie par  $F_{X_i} = (1 - \theta_i)\delta_0 + \theta_i F_{B_i}$  avec  $0 < \theta_i \leq 1$  et  $F_{B_i}$  est la distribution de la variable aléatoire positive  $B_i$ ,  $i = 1, \dots, n$  et  $\delta_0$  est la fonction de Dirac telle que

$$\delta_d(x) = \begin{cases} 1, & \text{si } x \geq d, \\ 0, & \text{si non.} \end{cases} \quad (2)$$

$\theta_i$  correspond à la probabilité que le risque  $i$  produit un sinistre positif d'un montant  $B_i$ . On peut trouver cette construction dans Bowers *et al.* (1997)[20], Klugman *et al.* (1998)[115] et Rolski (1999)[160]. Sauf pour le dernier, les auteurs définissent  $X_i = B_i \times I_i$ . Notons que dans le modèle individuel classique les  $I_i$  sont supposées indépendantes des  $B_i$  avec  $I_i$  sont des variables aléatoires de Bernoulli. Dans la première partie de la thèse, les exemples de modèles qu'on présente rentrent dans le cadre d'une modélisation individuelle du risque.

Notons que l'étude de quelques exemples de modèles collectifs qui satisfont nos hypothèses peut être envisageable en perspective.

### Modèle collectif

Soit un portefeuille de risque composé de  $n$  contrats anonymes qu'on n'observe pas séparément. Supposons que le nombre total  $N$  de sinistres se produisant dans une période donnée est aléatoire. En outre, les montants de sinistres  $X_i$  sont strictement positifs et forment une suite de variables aléatoires indépendantes et identiquement distribuées. On considère également que la suite  $(X_i)_{i \geq 1}$  est indépendante du nombre de sinistres  $N$ . Typiquement  $N$  a une distribution de Poisson, Binomiale ou Binomiale Négative, mais d'autres choix sont possibles. L'idée est de rapprocher le modèle individuel par un modèle collectif convenablement choisit si la taille  $n$  du portefeuille est assez grande.

Pour une discussion plus approfondie sur le modèle de risque collectif étudié en mathématiques de l'assurance nous renvoyons le lecteur vers Gerber (1979)[81], Heilmann (1988)[91], Panjer and Willmot (1992)[144], Straub (1988)[172] and Sundt (1993)[173], Rolski *et al.* (1999)[160], Rotar (2006)[163] et Marceau (2013)[129].

Dans cette thèse, nous avons commencé par l'étude des propriétés du processus  $(X_i)_{i \geq 1}$  dans un cadre particulier de dépendance des variables conditionnellement indépendantes. Les résultats que nous avons trouvé ouvrent plusieurs directions de recherche. L'une de ces directions est de donner des résultats asymptotiques pour le processus  $(S_t)_{t \geq 1}$  lorsque le temps  $t$  tend vers l'infini dans notre cadre généralisé de dépendance. Cette piste de recherche est étudiée dans le deuxième chapitre de cette thèse. Bien évidemment, la construction du modèle de processus du risque (1), a pour but de déterminer si le portefeuille va survivre au fil du temps. Ainsi, l'une des principales perspectives de recherche ouvertes dans le prolongement de nos résultats est d'étudier l'impact du type de dépendance que nous considérons sur la probabilité de ruine et aussi sur d'autres indicateurs de risques en particulier à un horizon de temps fini. Cette perspective peut être exploitée dans des travaux futurs.

#### 1.4 Modélisation de la dépendance en théorie du risque : une revue

De nombreux résultats classiques en théorie de la ruine reposent sur l'hypothèse d'indépendance entre les montants de sinistres d'une part et d'indépendance entre les montants de sinistres et les temps inter sinistres d'autre part. Cependant, la dépendance entre risques n'est pas un cas exceptionnel et certains exemples de processus de risque en assurance imposent de prendre en compte un certain degré de dépendance, en particulier la dépendance entre les montants de sinistres dans les modèles à environnement markovien et la dépendance entre montants de sinistres et temps inter sinistres lorsqu'il s'agit du risque « catastrophes naturelles ». En effet, si on prend le risque tremblement de terre, alors l'arrivée d'un séisme dévastateur peut être influencée par le temps écoulé depuis le dernier tremblement arrivé avec une forte intensité. La dépendance entre montants de sinistres et temps inter sinistres ne sera pas étudiée dans cette thèse, on renvoie le lecteur vers la liste suivante non exhaustive de références qui traite ce type de dépendance : Albrecher et Boxma (2004)[2], Albrecher et Teugels (2006)[4], Boudreault *et al.* (2006)[19], Meng *et al.* (2008)[136] et Biard *et al.* (2011)[11].

Dans cette thèse on s'intéresse à la modélisation de la dépendance entre les risques. Ces derniers peuvent être dépendants sous différents angles. On mentionne que dans les modèles classiques de risque, la dépendance peut intervenir entre les contrats d'un même portefeuille de risque. Elle se traduit par une corrélation entre les montants de sinistres. Par exemple, deux contrats Multirisques habitations dans la même zone géographique peuvent être touchés par le même événement naturel. On peut aussi observer une dépendance entre des risques de différentes branches. Parmi les travaux traitant ce type de dépendance on trouve Ambagaspitiya (1998)[5], Wang (1998)[181], Baurele et Müller (1998)[9], Cossette *et al.* (2000)[36] et Frostig (2003)[80].

En actuariat, dans le modèle individuel, un certain nombre de stratégies sont proposées pour modéliser la structure de dépendance entre les risques. Ces modèles peuvent être clas-

sifiés en deux groupes : ceux dans lesquels la dépendance est décrite en termes de variables aléatoires comme les modèles avec mélange commun et les modèles dans un environnement markovien et ceux dans lesquels la relation entre des risques est exprimée en termes des fonctions de distribution comme les modèles basés sur les copules et les modèles avec choc commun.

Puisque notre extension concerne les modèles avec mélange commun, nous allons commencer par les présenter. Aussi, dans un souci de clarté, nous présentons les modèles avec choc commun afin de les distinguer des modèles avec mélange commun et aussi les modèles dans un environnement markovien pour montrer en quoi ces derniers sont différents des modèles avec mélange commun.

Vue l'existence d'une vaste littérature, nous ne prétendons pas faire ici une présentation exhaustive des travaux de recherche consacrés à chaque modèle de dépendance. Néanmoins, pour bien situer notre travail par rapport à la littérature nous allons décrire d'une manière simpliste les modèles évoqués dans le paragraphe précédent et nous allons tracer leurs principales particularités.

Le lecteur intéressé par les applications des copules pour la modélisation de la dépendance en actuariat peut consulter Frees et Valdez (1998)[78], Denuit *et al.* (2005)[53], McNeil *et al.* (2005)[135], Marceau (2013)[129]. Pour un traitement technique avancé sur la théorie des copules le lecteur peut consulter les livres de Drouet-Mari and Kotz (2001)[71], Joe (1997)[107] et Nelsen (1999)[140].

### Modèle avec mélange commun

Dans ce paragraphe, on présente le modèle appelé par Wang (1998)[181] « modèle avec mélange commun ». Ces modèles impliquent l'indépendance conditionnelle des variables aléatoires et forment la base des modèles de crédibilité. L'intuition derrière ces modèles est la suivante : on considère que la dépendance est introduite par le biais d'un facteur externe qui peut impacter le portefeuille d'assurance entier ou des parties du portefeuille. Ce facteur commun peut modéliser des situations économiques, des lieux géographiques, des conditions climatiques ou tout autres facteurs externes communs. La structure de dépendance avec mélange commun peut être adoptée dans un contexte financier dans l'analyse du risque de crédit ou en assurance lorsque certains facteurs externes peuvent affecter les différentes variables du modèle de risque. En assurance, on peut considérer un contrat de prévoyance collective dont le bénéficiaire est un ensemble d'employés travaillant dans une mine. Dans ce cas une explosion peut influencer le risque de tout le groupe. On rajoute qu'une catastrophe représente un cas extrême de dépendance et que lorsqu'elle se produit, la relation de dépendance peut être présente pour différentes branches d'affaire composant le portefeuille d'assurance. D'autres exemples sont étudiés par Wang et Dhaene (1998)[182], Wang (1989)[181] et Bauerle et Muller (1998)[9] et Cossette *et al.* (2002)[33]. On rappelle la définition suivante des variables aléatoires dépendantes par mélange commun donnée dans Denuit *et al.* (2005)

**Définition 1.** *Un vecteur  $X$  de variables aléatoires est dépendant par mélange commun si*

sa fonction de distribution jointe peut s'écrire de la manière suivante

$$F_X(x) = \int_{-\infty}^{+\infty} \left( \prod_{i=1}^n F_i(x_i|v) \right) dF_V(v), \quad x \in \mathbb{R}^n,$$

où  $F_1(\cdot|v), \dots, F_n(\cdot|v)$  et  $F_V$  sont des fonctions de distributions.

Notons que  $F_i(\cdot|v)$  est la fonction de distribution de  $(X_i|V = v)$ ,  $i = (1, 2, \dots, n)$ . La distribution non conditionnelle de  $X_i$ ,  $i = (1, 2, \dots, n)$  est donc donnée par

$$F_i(x_i) = \int_{-\infty}^{+\infty} F_i(x_i|v) dF_V(v).$$

Le processus de perte globale peut être considéré comme un processus en deux étapes : d'abord, le paramètre  $V = v$  externe est tiré à partir de sa fonction de distribution, ensuite pour chaque risque individuel, le montant de sinistre correspond à une génération d'une variable aléatoire à partir de la fonction de distribution conditionnelle de  $X_i$  sachant  $V = v$ .

Un cas particulier des modèles avec mélange commun est le cas où sachant  $V = v$ , les montants de sinistres  $X_i$  sont dégénérés en  $x_i$ , avec les  $x_i = x_i(v)$  des fonctions non-décroissantes de  $v$ . Un tel modèle est en quelque sorte une forme extrême d'un modèle avec mélange, puisque dans ce cas le paramètre externe  $V = v$  détermine complètement le montant global de sinistres. Dans ce cas les risques  $X_1, \dots, X_n$  sont comonotones (voir Goovaerts *et al.* (2001)[83] pour le principe de comonotonicité).

Cossette *et al.* (2002)[33] considèrent le modèle individuel classique de risque et proposent de modéliser la structure de dépendance entre les fréquences de sinistres par un modèle avec mélange commun. Soit un portefeuille d'assurance composé de  $n$  risques individuels dans lequel le montant total de sinistre du  $i$  éme contrat est modélisé par la variable aléatoire  $X_i$ , ( $i = 1, \dots, n$ ) de l'équation (5). Chaque variable aléatoire  $I_i$ , ( $i = 1, \dots, n$ ) est influencée par les valeurs possibles de la variable aléatoire  $V$  qui représente le facteur. Dans ce cas  $(I_i|V = v)$ , ( $i = 1, \dots, n$ ) sont aussi des variables de Bernoulli tels que pour tout ( $i = 1, \dots, n$ )

$$\mathbb{P}(I_i = 1|V = v) = r_i^v \text{ et } Pr(I_i = 0|V = v) = 1 - r_i^v, \quad (3)$$

avec  $r_i^v$  est le paramètre de base de la distribution conditionnelle de  $(I_i|V = v)$ . On note que dans leur modèle la probabilité conditionnelle d'occurrence sachant  $V = v$  est une fonction croissante de  $v$ .

Notons qu'en raison de la dépendance entre les variables aléatoires  $I_i$ , ( $i = 1, 2, \dots, n$ ), les composantes du vecteur aléatoire  $X = (X_1, \dots, X_n)$  sont elles aussi dépendantes même si les variables aléatoires  $B_i$  sont indépendantes des variables aléatoires  $I_i$ .

Frostig et Denuit (2008) [54] considèrent un portefeuille de risque de crédit dans lequel la dépendance entre les probabilités de défaut individuels est entraînée par un certain nombre de facteurs environnementaux. Plus précisément, soit  $(I_1, \dots, I_n)$  un ensemble de variables aléatoires de Bernoulli. Chaque  $I_i$  prend la valeur 1 si le débiteur  $i$  fait défaut et 0 si non. Sachant le vecteur des facteurs macro-économiques  $\Theta = \theta$ , la probabilité conditionnelle de défaut est  $p_i(\theta) = P(I_i = 1|\Theta = \theta)$  et la probabilité de défaut est  $p_i = E(p_i(\Theta))$ .

Dans cet article les auteurs ont étudié l'effet d'une augmentation de la force de la dépendance entre les facteurs macro-économiques sur les indicateurs de défaut dans le modèle de risque de crédit standard. En particulier, dans un modèle à un seul facteur ils ont établi que la variabilité du facteur conduit l'intensité de la dépendance entre les variables de Bernoulli : plus le facteur est variable, plus les  $I_i$  sont positivement dépendants. Dans un modèle avec plus d'un facteur, ils montrent que plus les facteurs sont dépendants, plus les  $I_i$  sont dépendants. Frostig et Denuit (2009)[55] ont étendu les résultats obtenus dans [54] pour des variables aléatoires plus générales que les variables aléatoires de Bernoulli.

Dans notre étude nous considérons les modèles de dépendance avec mélange commun qui peuvent vérifier une condition de dépendance faible en statistique. En littérature actuarielle on trouve des travaux qui concernent la notion de dépendance faible qui est l'association. Dans le paragraphe suivant nous exposons brièvement l'un des résultats liés à la modélisation de la dépendance avec mélange commun lorsque la suite du facteur satisfait la condition d'association.

### Dépendance avec mélange commun et association

En statistiques, le concept d'association est un concept de dépendance qui a été introduit par Esary *et al.* (1967)[75]. Rappelons sa définition

**Définition 2.** Soit  $f$  et  $g$  deux fonctions non décroissantes pour lesquelles la covariance existe. Les variables aléatoires  $X_1, X_2, \dots, X_n$  sont dites associées si

$$\text{Cov}(f(X_1, X_1, \dots, X_n), g(X_1, X_1, \dots, X_n)) \geq 0.$$

Notons que le concept d'association est une notion de dépendance faible. En science actuarielle le concept d'association a été considéré par Norberg (1989)[141] dans la quantification des conséquences d'une dépendance possible entre les montants des primes liées aux contrats d'assurance vie multivies. Comme Shaked et Spizzichino (1998)[167] et Ahmed *et al.* (1981)[1] l'ont montré, la dépendance par mélange commun est une notion de dépendance positive sous des hypothèses supplémentaires sur les marginales conditionnelles. On reporte le résultat suivant introduit et montré par Ahmed *et al.* (1981)[1] et redémontré par Denuit *et al.* (2005)

**Proposition 1.** Soit  $X$  un vecteur aléatoire de dimension  $n$  et soit  $V$  un vecteur aléatoire de dimension  $p$ . Supposons que

- $(X_1, \dots, X_n)$  sont conditionnellement indépendants sachant  $V$  ;
- $X_i$  est stochastiquement monotone en fonction de  $V$  ;
- $V$  est associé.

Alors  $X$  est associé.

Dans ce cas il est clair que la structure de dépendance entre les  $X_i$  dérive de celle entre les  $V_i$ . Cette proposition montre que si le facteur est associé et si les marginales conditionnelles satisfont certaines propriétés de monotonie stochastique via le facteur alors, on peut en déduire la structure de dépendance qui découle des modèles avec mélange commun. Pour d'autres résultats et propriétés qui traitent la dépendance avec mélange pour des suites associées voir Denuit *et al.* (2002)[61], Denuit *et al.* (2005)[53] et Frostig et Denuit (2009)[55].

Dans le premier chapitre de cette thèse nous donnons un résultat similaire au résultat de la Proposition 1 mais pour des variables mélangeantes (voir Définition 6). En effet, nous considérons que pour tout  $i \in \mathbb{N}$  les variables aléatoires  $X_1, \dots, X_i$  sont conditionnellement indépendantes sachant  $\underline{V}_i = (V_1, \dots, V_i)$  et on suppose que la suite du facteur  $V_1, \dots, V_i$  est une suite mélangeante. Nous avons montré que sous d'autres conditions suffisantes la suite  $(X_1, \dots, X_i)$  est mélangeante (voir Théorème 1.2.1). La motivation actuarielle pour donner un tel résultat est développée dans la section 1.5.

Les modèles avec mélange commun sont souvent assimilés aux modèles avec choc commun. On mentionne que les deux modèles sont des modèles avec indépendance conditionnelle. Alors que dans le premier modèle la dépendance est introduite via un facteur externe, dans le deuxième la dépendance est introduite via un choc commun appliqué directement aux marginales des variables aléatoires modélisant les risques. Nous allons décrire les modèles à choc commun dans le paragraphe suivant.

### Modèle avec choc commun

Les modèles avec choc commun sont bien adaptés à l'assurance non vie et aux modèles multivariés de risque. En effet, ces modèles permettent de modéliser la dépendance entre branches : ils consistent à appliquer un même choc à différentes branches, ce qui a pour effet de créer la dépendance entre ces dernières. Ces modèles sont présentés dans Marshall et Olkin (1967,1988)[131; 132] et dans Kocherlakota et Kocherlakota (1992)[116]. Parmi les modèles à choc commun on trouve ceux qui portent sur les lois de Poisson modélisant le nombre de sinistres. Les modèles de Poisson à choc commun ont été proposés par Wang (1998)[181] et utilisés par Cossette *et al.* (2000)[36] qui ont défini la probabilité de ruine sur un horizon fini et infini pour ces types de modèles. Leurs résultats ont été généralisés par Frostig (2003)[80]. Voici un exemple simpliste d'application de ces modèles : soit une classe d'assurés qui ont souscrit des contrats d'assurance automobile et d'assurance habitation. Dans le cas où une catastrophe se produit, un tremblement de terre par exemple, les dommages ne peuvent pas être indépendants. Pour introduire une relation de dépendance entre les classes d'affaires on considère un modèle de Poisson à choc commun tel que :

$$\begin{cases} N_1 = N_{10} + N_0 \\ N_2 = N_{20} + N_0, \end{cases} \quad (4)$$

avec  $N_1, N_2$  deux variables aléatoires de Poisson qui modélisent respectivement le nombre de sinistres de la branche assurance automobile et de la branche habitation,  $N_0$  est le choc commun qui représente la fréquence des sinistres touchant les deux contrats simultanément ;  $N_{10}, N_{20}$  sont aussi des variables de Poisson. Dans ce type de modèle on suppose qu'une partie du montant total des sinistres pour le risque ( $i = 1, 2$ ) provient de la source individuelle  $N_{i0}$ , ( $i = 1, 2$ ) associée au risque  $i$  et une autre partie provient de la source commune  $N_0$ . Dans le modèle (4), la variable aléatoire  $N_{10}$  conditionnée par rapport à  $N_0 = n_0$  est indépendante de la variable aléatoire  $N_{20}$  conditionnée par rapport à  $N_0 = n_0$ .

Ces modèles présentent l'inconvénient de mêler ensemble la structure de dépendance et les marginales et de ne prendre en compte que des marginales de même nature.

### Dépendance avec choc commun et association

Soit un portefeuille d'assurance composé de  $n$  risques individuels dans lequel le montant total de sinistre du  $i$  ème contrat est modélisé par la variable aléatoire  $X_i$ , ( $i = 1, \dots, n$ ) tel que

$$X_i = \begin{cases} B_i, & I_i = 1 \\ 0, & I_i = 0, \end{cases} \quad (5)$$

avec  $I_i$  une variable de Bernoulli de paramètre  $q_i$  qui modélise la fréquence des sinistres et  $B_i$  une variable aléatoire positive de fonction cumulative de distribution  $F_{B_i}$  qui modélise le montant de sinistre. Cossette *et al.* (2002)[34] proposent une construction qui permet d'introduire la dépendance entre les risques individuels  $X_i$  dans laquelle les fréquences  $I_i$  sont influencées par un facteur de risque individuel et un facteur de risque collectif. Ce dernier représente un choc commun qui peut produire des sinistres pour chaque contrat. Spécifiquement pour tout  $i = (1, 2, \dots, n)$  on a

$$I_i = \min(J_i + J_0, 1),$$

avec  $J_i$ ,  $i = (0, 1, 2, \dots, n)$  des variables aléatoires de Bernoulli indépendantes. On mentionne que grace à cette construction les  $I_i$  sont des variables aléatoires associées. En effet, pour toute fonctions  $f$  et  $g \mathbb{R}^n \rightarrow \mathbb{R}$ , il existe deux fonctions  $\tilde{f}$  et  $\tilde{g} \mathbb{R}^n \rightarrow \mathbb{R}$  tel que

$$Cov(f(I_1, I_1, \dots, I_n), g(I_1, I_1, \dots, I_n)) = Cov(\tilde{f}(J_0, J_1, \dots, J_n), \tilde{g}(J_0, J_1, \dots, J_n)),$$

qui n'est pas négative puisque les variables aléatoires  $J_i$ ,  $i = (0, 1, 2, \dots, n)$  sont indépendantes, et donc associées (voir Denuit *et al.* (2005)[53]).

Afin de souligner la différence entre les modèles avec mélange commun lorsque le mécanisme externe est modélisé par une chaîne de Markov et les modèles dans un environnement markovien, nous présentons dans le paragraphe suivant les principales propriétés de ces derniers et nous décrivons brièvement l'une des différences principales de ces modèles avec ceux que nous allons généralisés c'est-à-dire les modèles avec mélange commun.

### Modèle dans un environnement markovien

Les processus de risque modulés par un environnement markovien sont largement utilisés dans le cas d'une dépendance liée à un contexte environnemental. En effet, lorsque différentes variables du modèle sont influencées par un environnement externe modélisé par un processus de Markov ou une chaîne de Markov (lorsqu'il s'agit d'un processus de Markov avec un nombre fini d'états) on dit que le modèle est dans un environnement markovien.

Rappelons que notre extension concerne des modèles de dépendance en temps discret. En actuariat plusieurs auteurs se sont intéressés aux modèles de dépendance dans un environnement markovien en temps discret, parmi d'autres on cite Lehtonen et Nyrhinen (1992)[124], Cai et Dickson (2004) [28], Cossette *et al.* (2004)[35], Yuen and Guo (2006) [187].

Nous commençons par présenter le modèle Binomial composé défini dans un environnement markovien de Cossette *et al.* (2004)[35] dans lequel le facteur externe est modélisé par une chaîne de Markov.

Soit  $S_n = Y_1 + \dots + Y_n$  le montant total de sinistres durant  $n$  périodes tels que

$$Y_j = \begin{cases} B_j, & I_j = 1 \\ 0, & I_j = 0, \end{cases} \quad (6)$$

avec au maximum un sinistre par période,  $\bar{I} = \{I_j, j \in \mathbb{N}\}$  est une suite de variables aléatoires de Bernoulli i.i.d. et  $\bar{B} = \{B_j, j \in \mathbb{N}\}$  est une suite de variables aléatoires positives et i.i.d. Les deux suites de variables aléatoires sont supposées indépendantes l'une de l'autre, ce qui implique que les variables aléatoires  $Y_j$  sont aussi i.i.d.

Dans le modèle Binomial composé défini dans un environnement markovien, les variables aléatoires  $Y_j$  sont toujours identiquement distribuées mais ils ne sont plus indépendantes. En effet, les deux processus  $\bar{B}$  et  $\bar{I}$  sont définis en fonction d'une chaîne de Markov sous-jacente  $\bar{\Theta} = \{\Theta_j, j \in \mathbb{N}\}$ , définie dans un espace d'état fini  $(\theta^1, \dots, \theta^m)$ . Dans une période donnée, la probabilité d'avoir un sinistre dépend de l'état de  $\bar{\Theta}$  dans cette période. Étant donné  $\Theta_j = \theta^i$ , les variables de Bernoulli  $I_j$  sont distribuées avec une moyenne  $\alpha_i$  ( $\alpha_i \in [0, 1]$ ) pour tout  $i \in \{1, \dots, m\}$ . Sachant  $\bar{\Theta}$ , les variables aléatoires  $I_j$  sont indépendantes et le processus  $\bar{I}$  est un processus de Bernoulli défini dans un environnement markovien.

Considérons le modèle Binomial composé avec mélange commun dans lequel seules les variables de Bernoulli  $I_i$  dépendent du facteur  $V$  qui a l'espace d'état  $\bar{v}$  de la précédente chaîne de Markov i.e  $\bar{v} = (\theta^1, \dots, \theta^m)$ . Dans ce cas la distribution de  $V$  correspond à la loi stationnaire de  $\bar{\Theta}$  qui représente l'environnement markovien.

Notons que dans le modèle Binomial composé avec mélange commun, la probabilité d'avoir un sinistre ne dépend pas de l'état du facteur et demeure inchangée dans le temps. Ceci n'est pas le cas du modèle Binomial composé défini dans un environnement markovien dans lequel la probabilité d'avoir un sinistre change dans chaque période selon l'état de la chaîne de Markov sous-jacente. Dans notre cadre de dépendance, qui est une extension des modèles avec mélange commun, le conditionnement est fait par rapport à un vecteur de longueur qui varie au cours du temps. Ainsi les probabilités d'avoir un sinistre changent d'une période à l'autre à cause de la longueur de vecteur du conditionnement. On rajoute que dans les exemples de modèles qu'on considère, les processus de variables aléatoires sont supposés asymptotiquement stationnaires. De ce fait, lorsque le temps devient assez grand, les probabilités d'avoir un sinistre deviennent inchangées dans le temps.

## 1.5 Modèles à facteurs commun : notre extension

Dans ce travail on propose d'étendre des modèles utilisés par les praticiens de l'assurance à la modélisation de la dépendance d'un portefeuille de risque. Notre extension est inspirée des modèles avec mélange communs présentés précédemment.

On a vu que, dans ces modèles, la structure de dépendance entre les risques peut provenir d'un facteur commun qui représente l'évolution de l'environnement socio-économique, naturel ou autre, dans lequel évolue la compagnie d'assurance. Dans le but d'introduire notre projet d'extension on revient brièvement sur le modèle de Cossette *et al.* (2002)[33] en formalisant leur cadre particulier d'indépendance conditionnelle de la manière suivante : soit  $\underline{V}_n$  le vecteur aléatoire  $(V_1, \dots, V_n)$  correspondant à l'évolution du facteur jusqu'au temps  $n \geq 1$ . Pour toutes suites d'indices  $(i_1, \dots, i_u)$ ,  $i_1 < \dots < i_u$  et pour tout entier

$n \geq i_u$ , on a

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_n) = \int_{x_{i_1}, \dots, x_{i_u}} f(x_{i_1}, \dots, x_{i_u}) \prod_{j=1}^u N^{i_j}(dx_{i_j}, V_{i_j}), \quad (7)$$

avec  $N^i(dx_i, V_i)$  le noyau de transition de  $X_i$  sachant  $V_i$  ce qui signifie qu'à l'instant  $i$  la loi conditionnelle de la variable aléatoire  $X_i$  dépend de l'état du facteur observé uniquement à l'instant  $i$ .

La notion du noyau de transition est très utilisé dans le premier chapitre. Nous rappelons sa définition (Revuz (1984)[156]).

**Définition 3.**  $\mathcal{B}(\mathbb{R}^k)$  désigne la tribu de borélien de  $\mathbb{R}^k$ . Une application  $\nu : \mathcal{B}(\mathbb{R}^d) \times \mathbb{R}^p \rightarrow [0, 1]$  est appelée noyau de transition si elle satisfait les propriétés suivantes :

- i)  $\forall x \in \mathbb{R}^p$ ,  $\nu(\cdot, x)$  est une mesure de probabilité sur  $\mathbb{R}^d$  ;
- ii)  $\forall A \in \mathcal{B}(\mathbb{R}^d)$ , l'application  $\nu(A, \cdot)$  est  $\mathcal{B}(\mathbb{R}^p)$ - mesurable.

Soit  $X$  et  $Y$  deux variables aléatoires à valeur dans  $\mathbb{R}^p$  et  $\mathbb{R}^d$ . On appelle la loi conditionnelle de  $Y$  sachant  $X$  un noyau de transition  $\nu$  sur  $\mathbb{R}^d \times \mathbb{R}^p$  tel que  $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \times \nu$ .

En raison des délais de règlement de sinistres, des normes comptables et des phénomènes d'accumulation du risque dans le temps, toute la mémoire du facteur joue souvent un rôle important dans l'évolution du processus de risque en assurance. Il est donc important de considérer de tels modèles dans la modélisation de la dépendance entre les risques. On mentionne que les processus de variables aléatoires conditionnellement indépendantes par rapport à un vecteur aléatoire de longueur variable au cours du temps n'ont pas encore été considérés dans la littérature, de ce fait il est important d'examiner certaines propriétés de ces modèles. Notre idée est donc de considérer des processus de variables aléatoires conditionnellement indépendantes par rapport à toute la trajectoire du facteur. Formellement notre notion d'indépendance conditionnelle est la suivante : si on note par  $N^i(dx_i, \underline{V}_i)$  le noyau de transition de  $X_i$  sachant  $\underline{V}_i$  alors on suppose que pour tout multi indices  $(i_1, \dots, i_u)$ ,  $i_1 < \dots < i_u$  et pour tout entier  $n \geq i_u$ , le processus  $(X_i)_{i \in \mathbb{N}}$  satisfait la condition suivante

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_n) = \int_{x_{i_1}, \dots, x_{i_u}} f(x_{i_1}, \dots, x_{i_u}) \prod_{j=1}^u N^{i_j}(dx_{i_j}, \underline{V}_{i_j}). \quad (8)$$

La particularité de cette représentation réside en plus dans le fait que les suites  $(\underline{V}_i)_{i \in \mathbb{N}}$  sont des suites qui s'intersectent.

Comme dans les modèles avec mélange commun, les montants de sinistres, la fréquence ou les deux ensembles peuvent être influencés par le facteur. Ce facteur est observable et peut être une variable aléatoire continue ou discrète et peut aussi être une chaîne de Markov.

De plus, le processus  $(X_i)_{i \in \mathbb{N}}$  peut être un processus absolument continu qui modélise le montant total de sinistre, un processus discret qui modélise la fréquence de sinistre ou un processus mixte qui permet de prendre en considération une éventuelle influence du facteur sur le montant et la fréquence des sinistres simultanément.

Dans les modèles avec indépendance conditionnelle, la probabilité conditionnelle de la réalisation d'un sinistre dépend des valeurs prises par le vecteur aléatoire du facteur. Dans le cas où le conditionnement est fait par rapport à une mémoire bornée du facteur, la probabilité non conditionnelle demeure inchangée et le processus  $(X_i)_{i \in \mathbb{N}}$  est stationnaire si le facteur l'est. Dans le cas où la mémoire du facteur est non bornée, la probabilité non conditionnelle change au cours du temps et donc le processus  $(X_i)_{i \in \mathbb{N}}$  n'est pas stationnaire.

Dans la proposition ci-dessous nous remarquons que la covariance des fonctions des  $X_i$  coïncide avec la covariance des espérances conditionnelles.

**Proposition 2.** *Soit  $u, v$  et  $r$  trois entiers et soit  $f$  et  $g$  deux fonctions réelles bornées. On considère la suite d'indices  $i_1 < i_2 \dots < i_u < i_u + r \leq j_1 < j_2 \dots < j_v$ . Si l'équation (8) est satisfaite alors la relation suivante est vraie*

$$\begin{aligned} & Cov(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v})) \\ &= Cov(\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v})). \end{aligned} \quad (9)$$

Ce résultat signifie que la covariance de  $f(X_{i_1}, \dots, X_{i_u})$  et  $g(X_{j_1}, \dots, X_{j_v})$  est égale à la covariance de leur espérance conditionnelle chaque fois que le conditionnement est effectué respectivement par rapport à  $\underline{V}_{i_u}$  et  $\underline{V}_{j_v}$ .

La première partie de cette thèse a deux objectifs principaux.

Le premier objectif consiste à fournir de nouveaux outils et concepts pour la modélisation de la dépendance temporelle en théorie du risque qui inclut des processus de mémoire non bornées. Ici on parle principalement des processus conditionnellement indépendants par rapport à un vecteur de longueur variable au cours du temps. Ces processus présentent un fort intérêt en théorie du risque.

Le deuxième objectif correspond à donner des résultats asymptotiques pour les processus agrégés dans le cadre des modèles incluant des processus à mémoire non bornée. La question qu'on se pose est donc la suivante : sous quelles conditions une approximation normale du processus agrégé est possible lorsque les variables sont conditionnellement indépendantes par rapport à un vecteur de longueur variable au cours du temps ?

Un article d'O'Brein (1974)[142] a marqué le début de notre étude de cette dernière question. Dans cet article, l'auteur considère un processus de variables aléatoires  $(X_i)_{i \in \mathbb{N}}$  défini dans un environnement markovien telle que la relation (7) est satisfaite. Il a donné un théorème central limite (T.C.L.) pour le processus  $(X_i)_{i \in \mathbb{N}}$  sous réserve que le processus de la chaîne de Markov satisfait l'hypothèse de Doeblin. Pour les chaînes de Markov cette condition est équivalente à la condition de  $\Phi$ -mélange introduite par Ibragimov (1962)[98]. Dans le même article l'auteur montre que lorsque la chaîne de Markov est fortement mélangeante alors le processus  $(X_i)_{i \in \mathbb{N}}$  l'est aussi. Ce dernier résultat généralise celui de Davydov (1969)[42] qui montre que la chaîne de Markov est elle-même fortement mélangeante. La notion de mélange est un outil très performant pour obtenir des résultats asymptotiques en statistique des séries chronologiques. Son avantage principal est de résumer les propriétés de dépendance de très nombreux modèles via le comportement d'une suite de coefficients.

A part les processus mélangeants, il existe de nombreux autres exemples de modèles satisfaisant ce type de dépendance tels que, les processus associés, les fonctions de processus gaussien. Doukhan et Louhichi (1999)[67] ont proposé une condition unifiée pour la notion de faible dépendance qui inclut tous les modèles précédemment cités et qui rend explicite l'indépendance asymptotique entre des fonctions, qui appartiennent à une classe convenable, du passé et du futur de la suite.

Beaucoup de référence portent sur les applications et les propriétés des processus mélangeants voir par exemple Doukhan (1994)[64], Rio (2000)[159], Bradley (1986,2005,2007)[21; 22; 23] et Dedecker *et al.* (2007)[46].

Il est donc connu que les processus mélangeants ont des propriétés asymptotiques intéressantes (voir e.g [45; 46; 47; 67; 101; 146; 158] et plusieurs autres). En effet, il est connu par un résultat d'Ibragimov(1962)[98] que si le processus stationnaire  $(X_i)_{i \in \mathbb{N}}$  est fortement mélangeant alors le processus agrégé a une distribution dans le domaine d'attraction d'une loi stable. Si de plus, d'autres conditions sont satisfaites par le processus alors on peut prouver la convergence asymptotique de  $S_n$  normalisée vers une loi normale. Des résultats à propos du T.C.L. des processus mélangeants sont présentés dans Ibragimov (1962)[98], Ibragimov et Linnik (1971)[101] et Iosifescu-Theodorescu (1969) [104] et résumés dans Iosifescu (1980)[103] et Peligrad (1986)[146].

Partant de ces idées, nous allons dans un premier temps étudier les propriétés de mélange des processus conditionnellement indépendants lorsque le conditionnement est fait par rapport à une mémoire non bornée. Nous verrons plus loin que le type de mélange vérifié par nos processus n'est pas classique. Dans un deuxième temps nous allons donner quelques résultats asymptotiques vérifiés par le processus agrégé des suites de variables aléatoires sous notre condition de mélange. En particulier on donne un T.C.L. autonormalisé.

Dans ce qui précède nous avons présenté le questionnement de notre thèse qui émerge d'un contexte actuariel. Notamment, nous avons commencé par exposer les principaux modèles de dépendance existant dans la littérature actuarielle en particulier, ceux qui sont en lien avec le modèle de dépendance que nous prétendons généraliser et nous avons par la suite formalisé notre extension. Dans ce qui suit nous allons présenter les outils et les résultats statistiques qui forment la base de notre cheminement pour donner des réponses à nos problématiques initiales.

## 1.6 Notion de mélange et ses principales propriétés

En statistiques, le concept du mélange, introduit pour la première fois par Rosenblatt (1956) [161], est une notion de dépendance qui décrit l'indépendance asymptotique des variables aléatoires indexées par le temps lorsque la différence entre leurs indices tend vers l'infini.

### 1.6.1 Suites mélangeantes : une revue

Les suites mélangeantes de variables aléatoires sont des suites pour lesquelles le passé et le futur sont asymptotiquement indépendants. Rappelons ici les différentes notions de mélange pour une suite  $(X_n)_{n \in \mathbb{Z}}$ , elles s'expriment en terme de coefficient de mélange qui

sont liées à la mesure de la dépendance entre les tribus engendrées par le futur et le passé de la suite. Bradley et Bryc (1985)[24] ont donné une présentation unifiée pour certains coefficients de mélange.

**Définition 4.** Soit  $(X_n)_{n \in \mathbb{Z}}$  une suite de variables aléatoires dans un espace de probabilité  $(\Omega, \mathbb{F}, \mathbb{P})$ ,  $s \in \mathbb{Z}$ ,  $F_{-\infty}^s := \sigma(X_k, k \leq s)$  et  $F_s^{+\infty} := \sigma(X_k, k \geq s)$ .

Pour chaque  $n \geq 1$ , on définit les coefficients de dépendance suivants

1. La suite  $(X_n)_{n \in \mathbb{Z}}$  est fortement mélangée si  $\alpha(n) \rightarrow 0$  lorsque  $n \rightarrow \infty$  avec

$$\alpha(n) := \sup_{s \in \mathbb{Z}} \left\{ |P(A \cap B) - P(A)P(B)| : A \in F_1^s, B \in F_{s+n}^{\infty}, k \geq 1 \right\}.$$

Ce coefficient a été introduit par Rosenblatt (1956)[161].

2. La suite  $(X_n)_{n \in \mathbb{Z}}$  est  $\rho$ -mélangée si  $\rho(n) \rightarrow 0$  lorsque  $n \rightarrow \infty$ . Cette classe a été étudiée par Kolmogorov and Rozanov (1960)[117]. Le coefficient de dépendance classique qui génère cette classe est fondé sur la corrélation maximale  $\rho(n)$ , définie par

$$\rho(n) := \sup_{s \in \mathbb{Z}} \left\{ \text{corr}(X, Y), X \in L_2(F_1^s), Y \in L_2(F_{s+n}^{\infty}), k \geq 1 \right\}.$$

Ce coefficient a été introduit par Hirschfeld (1935)[94].

3. La suite  $(X_n)_{n \in \mathbb{Z}}$  est  $\psi$ -mélangée si  $\psi(n) \rightarrow 0$  lorsque  $n \rightarrow \infty$  avec

$$\psi(n) := \sup_{s \in \mathbb{Z}} \left\{ \frac{|P(A \cap B) - P(A)P(B)|}{P(A)P(B)} : A \in F_1^s, B \in F_{s+n}^{\infty}, P(A)P(B) \neq 0, k \geq 1 \right\}.$$

Ce coefficient a été introduit par Blum, Hanson et Koopmans (1963)[14].

4. La suite  $(X_n)_{n \in \mathbb{Z}}$  est  $\phi$ -mélangée si  $\phi_n \rightarrow 0$  lorsque  $n \rightarrow \infty$  avec

$$\phi(n) := \sup_{s \in \mathbb{Z}} \left\{ |P(B|A) - P(B)| : A \in F_1^s, B \in F_{s+n}^{\infty}, P(A) \neq 0, k \geq 1 \right\}.$$

On dit que la suite satisfait la condition de mélange uniforme introduite par Ibragimov (1959)[97]. Cette condition a été étudiée par Cobgurn (1960)[31].

Voici une liste non exhaustive des ouvrages qui traitent les propriétés de ces coefficients : Doob (1953)[63], Billingsley (1968)[12], Ibragimov et Linnik (1971)[101], Hall et Heyde (1980)[90], Roussas et Ioannides (1987)[164] et Bradley (2007)[23].

Il est facile de voir qu'on a la relation suivante pour les propriétés de mélange des processus :

$$\Psi \text{ mélange} \implies \Phi \text{ mélange} \implies \rho \text{ mélange} \implies \alpha \text{ mélange}.$$

On remarque que lorsque deux variables aléatoires sont indépendantes alors elles sont  $\psi$ -mélangées avec un coefficient de mélange nul. Notons que les mesures de dépendance utilisées dans les définitions des coefficients de mélange de la Définition 4 s'inscrivent dans un cadre général qui utilise des normes de la covariance (voir Bradley (1986,2005)[21; 22] et Bulinskii (1987)[27]).

Cependant beaucoup de classes de séries temporelles ne satisfont pas de conditions de mélange e.g. Doukhan (1994)[64] de plus l'évaluation des coefficients de mélange reste souvent difficile. Partant de ces constatations Doukhan et Louhichi (1999)[67] ont introduit une notion de dépendance faible qui rend explicite l'indépendance asymptotique entre le passé et le futur, ce qui veut dire que le passé est progressivement oublié. Dans leur nouvelle notion l'écart n'est plus considéré entre les tribus engendrées par le passé et le futur des variables aléatoires mais il est plutôt entre des fonctions du passé et du futur de la suite. Plus précisément ils supposent que la covariance  $\text{Cov}(f(\text{passé}),g(\text{futur}))$  est petite lorsque la distance entre le passé et le futur est suffisamment grande. Formellement

**Définition 5** (Doukhan and Louhichi (1999)[67]). *La suite  $(X_1, \dots, X_n)$  de variables aléatoires est dite  $(\theta, H, \psi)$ - faiblement dépendante, s'il existe une classe  $H$  de fonctions à valeurs réelles, une suite de nombre réels  $\theta = (\theta_r)_{r \in \mathbb{N}}$  décroissante vers zéro lorsque  $r$  tend vers l'infini et une fonction  $\psi$  définie sur  $H^2 \times \mathbb{N}^2$  avec des arguments  $(f, g, u, v)$  tels que pour tout  $r \in \mathbb{N}$  et tout  $i_1 \leq \dots \leq i_u < i_u + r \leq j_1 \leq \dots \leq j_v$  on a*

$$|\text{Cov}(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq \psi(f, g, u, v)\theta_r, \quad (10)$$

pour toute fonctions  $f, g \in H$  définies respectivement sur  $\mathbb{R}^u$  et  $\mathbb{R}^v$ .

On note que l'article de Doukhan et Louhichi (1999)[67] est un article marquant dans notre étude. Pour cette raison il est intéressant de présenter avec plus de détail la définition précédente.

### Un regard sur la définition 5 de Doukhan et Louhichi (1999) [67] :

Mentionnons tout d'abord que si une suite est fortement mélangeante au sens de Rosenblatt (1956) alors elle est  $(\theta, L^\infty, \psi)$ - faiblement dépendantes où  $L^\infty$  est la classe des fonctions réelles bornées dans un espace de la forme  $\mathbb{R}^p$ ,  $\theta_r = \alpha_r$  et  $\psi(f, g, u, v) = 4 \|f\|_p \|g\|_q$  avec  $\|\cdot\|_\infty$  la norme usuelle sur  $L^\infty$ . Ce type de dépendance contient des exemples pertinents et est utilisé dans de nombreuses situations telles que les T.C.L. empiriques (e.g Doukhan et Louhichi (1999)[67] et Doukhan *et al.* (2009)[65]), l'estimation fonctionnelle (e.g Coulon-Prieur et Doukhan (2000)[37]) et les T.C.L. autonormalisés (e.g Doukhan *et al.* (2010)[66]). Dans cette définition selon les différents choix de la classe  $H$  on définit différents types de mélange faible. On note  $\Gamma$  l'ensemble de toutes les fonctions lipschitziennes définies et bornées sur  $\mathbb{R}^u$ ,  $u \in \mathbb{N}^*$ . Soit  $h : \mathbb{R}^u \rightarrow \mathbb{R}$  une fonction de  $\Gamma$ , son module de lipschitz est défini par

$$\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1},$$

avec

$$\|(x_1, \dots, x_u)\|_1 = |x_1| + \dots + |x_u|.$$

En particulier si  $f$  et  $g$  sont deux fonctions réelles de  $\Gamma_1$  avec  $\{\Gamma_1 = h \in \Gamma \mid \|h\|_\infty \leq 1\}$  et si  $\psi(f, g, u, v) = (u\text{Lip}(f) + v\text{Lip}(g) + uv\text{Lip}(f)\text{Lip}(g))$  alors la suite est dite  $\lambda$ - faiblement dépendante. Cette notion est appropriée pour un grand nombre d'exemples (pour plus de détails sur ce type de mélange et sur ses applications voir Dedecker *et al.* (2007) [46] et Doukhan *et al.* (2010)[66]). Plusieurs autres classes de fonction lipschitziennes peuvent être définies à la place de la classe  $\Gamma$  ce qui conduit à différentes variantes de la Définition 5

(voir Doukhan et Louhichi (1999)[64] et Dedecker (2007)[46]).

### 1.6.2 Suites conditionnellement mélangeantes : notre contexte

Les exemples de processus que nous considérons dans notre cadre de mélange sont caractérisés par des coefficients de mélange qui sont exponentiels en  $u$  et  $v$ . Ceci n'est pas le cas des exemples présentés dans Doukhan et Louhichi (1999)[67] et Dedecker (2007)[46]. De plus la classe de fonction qu'on considère dans nos exemples est la classe des fonctions bornées à valeurs réelles. Dans le premier chapitre de cette thèse nous introduisons la définition suivante qui est une variante de la définition introduite dans Doukhan et Louhichi (1999)[67]. Comme c'est le même nom de mélange qui est en jeu, nous avons choisi une notation étoilée des coefficients de mélange pour distinguer la notion de mélange qu'on considère de la notion de mélange classique.

**Définition 6.** *Soit  $u$  et  $v$  deux entiers. Un processus aléatoire  $(X_1, \dots, X_n)$  est dit  $\eta_{(u,v)}^*$  mélangeant, s'il existe une fonction  $r \mapsto \eta^*(r)$  décroissante vers zéro lorsque  $r$  tend vers l'infini et une constante  $C(u, v) > 0$  telle que pour toute fonction à valeur réelles bornées  $f$  et  $g$ , pour tout entier  $u, v$  et pour tout multi indices satisfaisant la relation  $(\star)$  :*

$$i_1 < \dots < i_u < i_u + r \leq j_1 < \dots < j_v, \quad (\star)$$

on a

$$\sup |Cov(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq C(u, v)\eta^*(r)\|f\|_a\|g\|_b, \quad (11)$$

où le sup est calculé sur toutes les suites  $(i_1, \dots, i_u)$  et  $(j_1, \dots, j_v)$  satisfaisant  $(\star)$  et  $r \leq j_1 - i_u$  est le trou de temps entre le futur et le passé.

$\|\cdot\|_a$  and  $\|\cdot\|_b$  sont les normes des fonctions bornées (ou des sous-espaces de fonctions bornées). Selon ces normes, nous avons différents types de mélange tel que introduits dans les définitions classiques du mélange :

1. Si  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_\infty$ , nous dirons que le processus est  $\alpha_{(u,v)}^*$ -mélangeant et nous écrirons  $\alpha^*(r)$  au lieu de  $\eta^*(r)$ . Si le processus est  $\alpha_{(u,v)}^*$ -mélangeant pour tout  $u, v$  et si  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  alors le processus est  $\alpha$ -mélangeant au sens de Rosenblatt (1956)[161] et nous écrirons  $\alpha(r)$  au lieu de  $\eta^*(r)$ .
2. Si  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_2$ , nous dirons que le processus est  $\rho_{(u,v)}^*$ -mélangeant et nous écrirons  $\rho^*(r)$  au lieu de  $\eta^*(r)$ . Si le processus est  $\rho_{(u,v)}^*$ -mélangeant pour tout  $u, v$  et si  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  alors le processus est  $\rho$ -mélangeant au sens de Kolmogorov et Rozanov (1960)[117] et nous écrirons  $\rho(r)$  au lieu de  $\eta^*(r)$ .
3. Si  $\|\cdot\|_a = \|\cdot\|_1$  et  $\|\cdot\|_b = \|\cdot\|_\infty$ , nous dirons que le processus est  $\Phi_{(u,v)}^*$ -mélangeant et nous écrirons  $\Phi^*(r)$  au lieu de  $\eta^*(r)$ . Si le processus est  $\Phi_{(u,v)}^*$ -mélangeant pour tout  $u, v$  et si  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  alors le processus est  $\Phi$ -mélangeant au sens d'Ibragimov (1962)[98] et nous écrirons  $\Phi(r)$  au lieu de  $\eta^*(r)$ .
4. Si  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_1$ , nous dirons que le processus est  $\Psi_{(u,v)}^*$ -mélangeant et nous écrirons  $\Psi^*(r)$  au lieu de  $\eta^*(r)$ . Si le processus est  $\Psi_{(u,v)}^*$ -mélangeant pour tout  $u, v$  et si  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  alors le processus est  $\Psi$ -mélangeant au sens de Blum et al. (1963)[14] et nous écrirons  $\Psi(r)$  au lieu de  $\eta^*(r)$ .

La particularité de l'indépendance conditionnelle liée à l'ensemble du vecteur  $\underline{V}_i$  implique une dépendance qui peut être persistante pour le processus  $(X_i)_{i \in \mathbb{N}}$ . En particulier, l'indépendance conditionnelle des variables aléatoires des processus de risque qu'on considère et la propriété de mélange du processus du facteur ne sont pas suffisantes pour obtenir des propriétés de mélange pour le processus de  $(X_i)_{i \in \mathbb{N}}$  (voir exemple 1.2.1 du premier chapitre pour une illustration). En analysant (9) on voit bien que par un simple contrôle de la covariance des espérances conditionnelles des fonctions du passé et du futur du processus  $(X_i)_{i \in \mathbb{N}}$  on arrive à contrôler la covariance des fonctions du passé et du futur du même processus. Nous remarquons que la condition d'indépendance conditionnelle avec la condition de mélange conditionnel implique le mélange pour la suite de variables aléatoires  $(X_i)_{i \in \mathbb{N}}$ . Il est donc nécessaire d'étudier sous quelles conditions est obtenue la propriété du mélange conditionnel.

Partant de la Définition 6, on introduit la définition suivante des processus conditionnellement mélangeants :

**Définition 7.** Soit  $u, v$  et  $r$  trois entiers,  $(i_1, \dots, i_u)$  et  $(j_1, \dots, j_v)$  satisfaisant  $(\star)$  et soit  $\|\cdot\|_a$  et  $\|\cdot\|_b$  des normes des fonctions bornées (ou des sous-espaces de fonctions bornées). La suite des variables aléatoires  $(X_n)_{n \in \mathbb{N}}$  est dite conditionnellement  $\eta_{(u,v)}^*$ -mélangeante par rapport au facteur  $\underline{V}_n$  s'il existe une suite positive  $\eta^*(r) \rightarrow 0$  lorsque  $r \rightarrow \infty$  et une constante  $C(u, v) > 0$  telle que l'inégalité suivante est vraie pour toute fonction bornée à valeurs réelles  $f : \mathbb{R}^u \rightarrow \mathbb{R}$  et  $g : \mathbb{R}^v \rightarrow \mathbb{R}$ ,

$$\begin{aligned} & \sup |Cov(\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v}))| \\ & \leq C(u, v) \eta^*(r) \|f\|_a \|g\|_b, \end{aligned} \quad (12)$$

où le sup est calculé par rapport à toutes suites  $(i_1, \dots, i_u)$  et  $(j_1, \dots, j_v)$  satisfaisant  $(\star)$ .

La terminologie du mélange conditionnel a été définie et étudiée dans Rao (2009)[154], Dziubdziela (1986)[72] et Veijanen (1990)[176] mais dans des contextes différents du nôtre. Lorsque le conditionnement est par rapport à une mémoire non bornée du facteur, le choix d'un facteur qui satisfait une propriété de mélange n'implique pas le mélange conditionnel du processus  $(X_i)_{i \in \mathbb{N}}$ . On peut voir ceci directement d'après les Lemmes 1.3.1 et 1.3.2 de la première partie de la thèse. En effet, il est nécessaire que le noyau de transition vérifie certaines conditions de régularité.

Le résultat principal du premier chapitre est le Théorème 1.2.1. Ce résultat met l'accent sur le lien entre les propriétés de mélange conditionnel du processus  $(X_i)_{i \in \mathbb{N}}$ , les propriétés du mélange du processus  $(V_i)_{i \in \mathbb{N}}$  du facteur et la régularité des noyaux de transition. Ce résultat est un résultat clé dans cette thèse puisque nous allons pouvoir donner des inégalités de moments, de covariance et des théorèmes limites pour les processus qui satisfont notre type de mélange y compris les processus conditionnellement indépendants par rapport à une mémoire non bornée du facteur.

## 1.7 Inégalités de covariance

### Revue des inégalités de covariance

Les inégalités de covariance des suites mélangeantes jouent un rôle important dans le calcul des inégalités de moments des sommes partielles de des variables aléatoires. Soit  $(\Omega, \mathbb{F}, \mathbb{P})$

un espace de probabilité. Soit  $F_1^s := \sigma(X_k, 1 \leq k \leq s)$  et  $F_{s+n}^{+\infty} := \sigma(X_k, k \geq s)$  deux tribus  $\forall s \in \mathbb{Z}$ . On considère deux variables aléatoires  $X$  et  $Y$  mesurables respectivement par rapport à  $F_1^s$  et  $F_{s+n}^{+\infty}$ . Dans la littérature on trouve plusieurs inégalités de covariance qui diffèrent selon le type de mélange considéré. Nous en rappelons quelques unes :

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq 2\phi^{1/q}(n) \|X\|_p \|Y\|_q \text{ avec } p, q \geq 1 \text{ et } \frac{1}{q} + \frac{1}{p} = 1; \quad (13)$$

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq 12\alpha^{1/r}(n) \|X\|_p \|Y\|_q \text{ avec } p, q, r \geq 1 \text{ et } \frac{1}{q} + \frac{1}{p} + \frac{1}{r} = 1; \quad (14)$$

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq \psi(n) \|X\|_1 \|Y\|_1; \quad (15)$$

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq 2 \int_0^{2\alpha} Q_X(u)Q_Y(u)du; \quad (16)$$

où  $Q_X(u) = \inf \{t : \mathbb{P}(|X| > t) \leq u\}$  est la fonction quantile.

L'inégalité (13) est valable pour des suites  $\phi$ -mélangeantes, (on la trouve par exemple dans Billingsley (1968)[12]) et elle est due à Ibragimov (1962)[98] (voir preuve de 13 dans [98]). L'inégalité (14) est valable pour des suites  $\alpha$ -mélangeantes (on la trouve dans Davydov (1968)[41]). Notons que  $|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq 4C_1C_2\alpha(n)$ , si  $|X| \leq C_1$  et  $|Y| \leq C_2$  (voir l'inégalité (17.2.2) dans Ibragimov et Linnik (1971)[101]). On trouve l'inégalité (15) dans Blum, Hanson et Koopmans (1963)[14] et elle est valable pour des suites  $\psi$ -mélangeantes. L'inégalité (16) est celle de Rio (1993)[157] qui s'écrit en utilisant les fonctions quantiles et est valable pour des suites  $\alpha$ -mélangeantes. On note que Dedecker et Doukhan [45] ont donné une inégalité de covariance qui est équivalente à celle de Rio (1994)[157] dans le cas des suites  $\alpha$  mélangeantes. On trouve d'autres inégalités dans Lin et Lu [125](2006) et dans Doukhan (1994)[64].

## Notre inégalité de covariance

Dans le théorème suivant nous avons adapté l'inégalité de Davydov (1968)[41] à notre structure de mélange

**Théorème 1.** *Soit  $(X_i)_{i \in \mathbb{N}}$  une suite stationnaire  $\alpha_{(u,v)}^*$  mélangeante. Soit  $\|f\|_p < \infty$  et  $\|g\|_q < \infty$ , avec  $1 < p, q \leq \infty$  et  $\frac{1}{p} + \frac{1}{q} < 1$ , alors*

$$|Cov(f(X_{i1}, \dots, X_{iu}), g(X_{j1}, \dots, X_{jv}))| \leq (C(u, v) + 8)(\mathbb{E}|f|^p)^{\frac{1}{p}}(\mathbb{E}|g|^q)^{\frac{1}{q}}\alpha^*(r)^{1 - \frac{1}{p} - \frac{1}{q}}. \quad (17)$$

Nous utiliserons cette inégalité pour donner des inégalités des moments et aussi dans les preuves de nos théorèmes limites.

## 1.8 Inégalités des moments

Les inégalités des moments des sommes partielles jouent un rôle très important dans diverses preuves des théorèmes limites. Soit  $X_1, \dots, X_n$  des variables aléatoires arbitraires (non nécessairement indépendantes) et soit  $S_n = \sum_{i=1}^n X_i$  alors

$$\mathbb{E}|S_n|^q \leq \sum_{k=1}^n \mathbb{E}|X_k|^q \quad \text{si } 0 < q \leq 1,$$

et

$$\mathbb{E}|S_n|^q \leq n^{q-1} \sum_{k=1}^n \mathbb{E}|X_k|^q \quad \text{si } q > 1.$$

Si de plus  $X_1, \dots, X_n$  sont des variables aléatoires indépendantes alors on a l'inégalité suivante dite inégalité de type Rosenthal

$$\mathbb{E}|X_1 + \dots + X_n|^q \leq C_q \left( \sum_{i=1}^n \mathbb{E}|X_i|^q + \left( \sum_{i=1}^n \mathbb{E}|X_i|^2 \right)^{q/2} \right), \quad (18)$$

où  $C_q$  est une constante qui dépend uniquement de  $q$ , (voir la preuve dans Petrov (1995)[151]). Rappelons également l'inégalité Marcinkiewicz-Zygmund (voir aussi Petrov (1995)[151])

$$\mathbb{E}|X_1 + \dots + X_n|^q = O(n^{\frac{q}{2}}). \quad (19)$$

### 1.8.1 Inégalités de type Marcinkiewicz-Zygmund

Cette borne a été étudiée pour différentes classes de variables aléatoires  $\{X_j, j \geq 1\}$ . Pour  $q$  un réel positif et  $K < \infty$  une constante, l'inégalité suivante de type Marcinkiewicz-Zygmund

$$\mathbb{E} \left| \sum_{k=1}^n X_k \right|^q \leq K n^{q/2}, \quad (20)$$

est vraie si  $\{X_j, j \geq 1\}$  est une suite

- de variables aléatoires mutuellement indépendantes ( $q \geq 2$ ) (voir Brillinger (1962)[25]);
- de Markov strictement stationnaire satisfaisant une condition de Doeblin (Voir Doob (1953)[63])
- de différence de martingale (voir Stout (1974)[171]);
- $\phi$ -mélangeantes au sens de Rosenblatt (1956)[161] strictement stationnaire (voir Ibragimov (1962)[98]);
- $\alpha$ -mélangeantes strictement stationnaire (voir Yokoyama (1980)[185]);
- $\rho$ -mélangeantes strictement stationnaire (voir Peligrad (1987)[147]);
- de des variables aléatoires associées strictement stationnaire (voir Birkel(1988)[13]).

Ce type d'inégalité est d'une grande utilité pour l'obtention de plusieurs types de lois limites notamment le T.C.L. et la loi forte des grands nombres.

Doukhan et Louhichi (1999)[67] ont obtenu une borne pour  $\mathbb{E}|S_n^q|$  de type (19) pour tout  $q \geq 2$  avec  $q$  un entier. Leur preuve s'appuie essentiellement sur le travail de Doukhan et Portal (1983)[69]. On reporte la définition suivante de Doukhan et Louhichi (1999)[67] qui est utile pour leur résultat énoncé dans le Théorème 2.

**Définition 8** (Doukhan et Louhichi (1999) [67]). *Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de des variables aléatoires centrées. Pour un entier positif  $r$ , on définit par  $(C_{r,q})_{q \geq 2}$  la suite non décroissante tel que*

$$C_{r,q} := \sup |Cov(X_{t_1}, \dots, X_{t_m}, X_{t_{m+1}}, \dots, X_{t_q})|, \quad (21)$$

où le sup est pris sur tout  $t_1, \dots, t_q$  tel que  $1 \leq t_1 \leq \dots \leq t_q$  et  $t_{m+1} - t_m = r$ .

Notons que le terme  $C_{r,q}$  donné dans (21) est utilisée dans l'énoncé de certain de nos résultats du Chapitre 2.

Le résultat de Doukhan et Louhichi (1999)[67] est le suivant :

**Théorème 2** (Doukhan et Louhichi (1999) [67]). *Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires centrées vérifiant pour  $q \in \mathbb{N}$ ,  $q \geq 2$*

$$C_{r,q} = O(r^{-q/2}) \text{ lorsque } r \rightarrow \infty.$$

Alors il existe une constante positive  $K$  qui ne dépend pas de  $n$  pour laquelle l'inégalité (20) est vraie.

Soit la suite strictement stationnaire  $(X_i)_{i \in \mathbb{N}}$  tel que  $\mathbb{E}(X_0^2) < \infty$ . Rio (1993)[157] a établi la majoration suivante de la variance de  $S_n$  :

$$Var(S_n) \leq 4n \int_0^1 \alpha^{-1}(t) Q^2(t) dt,$$

avec  $Q$  la fonction quantile de  $|X_0|$  et  $f^{-1}$  la fonction inverse càdlag de  $f$  où  $f$  une fonction non décroissante. Si de plus

$$\int_0^1 \alpha^{-1}(t) Q^2(t) dt < \infty, \quad (22)$$

alors  $Var(S_n) < Cn$ . Soit  $S_n^* = \sup_{j \leq n} |S_j|$ . Rio (1994)[158] a donné un résultat qui assure l'inégalité de Marcinkiewicz-Zygmund :  $\mathbb{E}(S_n^{*q}) < Cn^{q/2}$ . Notons que pour les variables  $\alpha$ -mélangantes, Doukhan et Louhichi (1999)[67] mentionnent que leur inégalité de Marcinkiewicz-Zygmund est atteinte pour tout entier  $q$  sous les conditions de Rio.

### 1.8.2 Inégalités de type Rosenthal

On trouve des inégalités type Rosenthal dans Doukhan et Portal (1983) [69], Doukhan (1994)[64], Rio (1993,1994)[158; 157], Doukhan et Louhichi (1999) [67], Louhichi [126], Rio (2000)[159] et Dedecker (2007)[46]. Pour les moments d'ordre  $q$  de  $S_n$  avec  $q \in \{2, 4\}$  nous allons utiliser les inégalités type Rosenthal de Doukhan et Louhichi (1999) [67] pour en déduire des inégalités adaptées à notre type de mélange. Soit  $S_n = \sum_{i=1}^n X_i$ , on reporte les moments d'ordre 2 et 4 énoncés dans Doukhan et Louhichi (1999)[67]

**Lemme 1** (Doukhan et Louhichi (1999)[67]). *Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires centrées alors*

$$\mathbb{E}(S_n^2) \leq 2n \sum_{r=0}^{n-1} C_{r,2}, \quad \mathbb{E}(S_n^4) \leq 4! \left\{ \left( n \sum_{r=0}^{n-1} C_{r,2} \right)^2 + n \sum_{r=0}^{n-1} (r+1)^2 C_{r,4} \right\}. \quad (23)$$

Ce lemme a été montré dans Bryc et Smolenski (1993)[26] mais pour des variables  $\rho$  mélangeantes (on rappelle que le  $\rho$ -mélange est plus restrictif que le  $\alpha$  mélange).

En se basant sur (23) et sur l'inégalité de covariance (17) on obtient pour  $q$  un entier dans  $\{2, 4\}$  des inégalités des moments qui vont nous servir dans les preuves de certains théorèmes limites.

Pour des suites  $\alpha$  mélangeantes, Shao et Yu (1996)[168] ont donné des inégalités type Rosenthal pour  $q$  un réel avec  $q > 2$ . Nous rappelons ici leur résultat, étant donnée que nous allons l'adapter à notre cas (voir Chapitre 2).

**Théorème 3** (Shao et Yu (1996)[168], Théorème 4.1). *Soit  $2 < p < r < \infty$ ,  $2 < v \leq r$  et  $(X_n)_{n \in \mathbb{N}}$  une suite stationnaire de variables aléatoires  $\alpha$ -mélangeantes avec*

$$\mathbb{E}X_n = 0, \quad \|X_n\|_r = \mathbb{E}(|X_n|^r)^{1/r} < \infty. \quad (24)$$

Supposons de plus que pour certain  $C > 0$  et  $\theta > 0$

$$\alpha(n) \leq Cn^{-\theta}. \quad (25)$$

Alors, pour tout  $\varepsilon > 0$  il existe  $A = A(\varepsilon, r, p, v, \theta, C) < \infty$  tel que

$$E|S_n|^p \leq A \left( (nC_n)^{p/2} \max_{i \leq n} \|X_i\|_v^p + n^{(p-(r-p)\theta/r) \vee (1+\varepsilon)} \max_{i \leq n} \|X_i\|_r^p \right), \quad (26)$$

avec

$$C_n = \sum_{i=0}^n (i+1)^{\frac{2}{v-2}} \alpha(i)^{\frac{v-2}{v}}.$$

En particulier, pour tout  $\varepsilon > 0$ , si  $\theta \geq (p-1)r/(r-p)$  et  $\theta > v/(v-2)$  alors

$$E|S_n|^p \leq A \left( n^{p/2} \max_{i \leq n} \|X_i\|_v^p + n^{1+\varepsilon} \max_{i \leq n} \|X_i\|_r^p \right), \quad (27)$$

et si  $\theta \geq \frac{pr}{2(r-p)}$  alors

$$E|S_n|^p \leq An^{p/2} \max_{i \leq n} \|X_i\|_r^p. \quad (28)$$

Sous qu'elles conditions la loi asymptotique des variables aléatoires conditionnellement indépendantes par rapport à un facteur de mémoire non bornée est une loi normale ? Pour répondre à cette question nous avons exploré un premier chemin fondé sur les fonctions caractéristiques. A partir de cette piste nous étions persuadé de l'intérêt de considérer des suites conditionnellement mélangeantes. Dans le paragraphe suivant nous explicitons brièvement cette piste.

## 1.9 T.C.L. pour des suites de variables aléatoires conditionnellement indépendantes par rapport à un facteur

Dans la littérature plusieurs auteurs se sont intéressés aux T.C.L. pour des processus de variables aléatoires conditionnellement indépendantes (C.I.) vérifiant (7), notamment lorsque la variable de contrôle est une chaîne de Markov à espace d'état fini. Pour des résultats connexes, voir Grigorescu et Oprisan (1976)[89], O'Brein (1974)[142] et Bokuchava (1978)[15]. Dans ce dernier article, l'auteur a prouvé un T.C.L pour  $f(V_i)$  avec  $f(\cdot)$  une fonction réelle de l'état de la chaîne. Dans ce cas un T.C.L pour la suite  $(X_i)_{i \in \mathbb{N}}$  de des variables aléatoires conditionnellement indépendantes est aussi prouvé.

Le choix d'une suite d'un facteur fortement mélangeant, ergodique ou  $m$ -dépendante conduit à prouver un T.C.L. pour la somme  $S_n = \sum_{i=1}^n X_i$ . Pour donner quelques références, on cite les travaux de Bokuchava (1978)[15] dans le cas d'un facteur ergodique, Bokuchava (1984)[18] lorsque le facteur satisfait une condition de mélange fort, Kvatadze *et al.* (1988)[118] pour des suites  $m$ -dépendantes contrôlées par un facteur ergodique.

Notons que les résultats donnés dans Bokuchava (1978,1982) [15; 16] sont généralisés par Bokuchava *et al.*(1987)[17] et Kvatadze *et al.* (1988)[118]. On cite aussi les travaux de Maxwell et Woodroffe (1996)[133] et de Shervashidze (1997)[169] qui ont donné un T.C.L. local pour les chaînes de Markov cachées.

Dans sa preuve du T.C.L. Bokuchava (1978)[15] a considéré la représentation suivante de la somme :  $S'_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ .

$$\begin{aligned} S'_n &= \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j - \mathbb{E}(X_j)) \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j - \mathbb{E}(X_j|V_j) + \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}(X_j|V_j) - \mathbb{E}(X_j) \\ &= S_n^* + S_n^{**}, \end{aligned} \quad (29)$$

avec

$$S_n^* = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j - \mathbb{E}(X_j|V_j), \quad S_n^{**} = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}(X_j|V_j) - \mathbb{E}(X_j). \quad (30)$$

En utilisant les fonctions caractéristiques, l'auteur a montré que les deux sommes partielles  $S_n^*$  et  $S_n^{**}$  sont asymptotiquement indépendantes et que chaque somme converge vers une loi normale.

Soit  $\underline{V}_n = (V_1, \dots, V_n)$  le vecteur aléatoire correspondant à l'évolution du facteur jusqu'à l'instant  $n > 1$ . Soit  $(X_i)_{i \in \mathbb{N}}$  une suite de variables aléatoires conditionnellement indépendantes sachant tout le passé du facteur  $\underline{V}_n = (V_1, \dots, V_n)$  satisfaisant (8). On a

$$S_n^* = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j - \mathbb{E}(X_j|\underline{V}_j), \quad S_n^{**} = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j). \quad (30)$$

Soit  $\Psi_{S_n^*}(t)$  (res  $\Psi_{S_n^{**}}(t)$ ) la fonction caractéristique de  $S'_{n_1}$  (res  $S'_{n_2}$ ). On peut montrer que

$$\Psi_{S_n^*}(t) = \mathbb{E}_{V_n}(\Psi_{S'_n|V_n}(t)) = \mathbb{E}_{V_n} \left\{ e^{-\frac{t^2}{2n} \sum_{j=1}^n \text{Var}(X_j|V_j)} \right\} \times e^{(no(\frac{t^2}{n}))}, \quad (31)$$

et que

$$\Psi_{S_n^{**}}(t) = \mathbb{E}_{V_n} \left( e^{itS'_{n_2}} \right) = e^{-\frac{t^2}{2n} \mathbb{E}_{V_n} \left\{ \sum_{j=1}^n [\mathbb{E}(X_j|V_j) - \mathbb{E}(X_j)] \right\}^2} e^{o(\frac{t^2}{n})}. \quad (32)$$

voir Annexe (A.1) pour plus de détail sur (31) and (32).

Dans ce cas afin de prouver un T.C.L. pour  $S'_n$  il est nécessaire que  $\tau$  et  $\tau_1$  convergent presque sûrement vers des valeurs positives non nulles avec

$$\tau = \mathbb{E}_{V_n} \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n [\mathbb{E}(X_j|V_j) - \mathbb{E}(X_j)] \right\}^2 \quad (33)$$

$$= \frac{1}{n} \sum_{j=1}^n \text{Var}(\mathbb{E}(X_j|V_j)) + \frac{2}{n} \sum_{j=1}^{n-1} \sum_{i>j}^n \text{Cov}(\mathbb{E}(X_j|V_j), \mathbb{E}(X_i|V_i)), \quad (34)$$

et

$$\tau_1 = \frac{1}{n} \sum_{j=1}^n \text{Var}(X_j|V_j). \quad (35)$$

On considère que le processus du facteur est mélangeant. Les vecteurs de conditionnement s'interceptent, alors dans ce cas le mélange du facteur ne permet pas de conclure que la suite des variables conditionnelles  $(X_i|V_i)$  est elle aussi mélangeante. De plus d'après l'expression de  $\tau$  il est clair qu'on a besoin de contrôler les covariances des espérances conditionnelles des variables aléatoires  $X_i$ , d'où l'intérêt de rajouter une condition supplémentaire de mélange conditionnel.

Dans le premier chapitre nous montrons, sous des conditions de mélange du facteur et de mélange conditionnel, que les processus satisfaisant (9) vérifient la condition de mélange (11). Dans le deuxième chapitre nous étudions le comportement asymptotique de la somme. D'une part nous donnons un estimateur convergent de la variance et d'autre part nous donnons un T.C.L. autonormalisé.

### 1.10 Principaux résultats asymptotiques des processus mélangeants : une revue

Pour prouver des théorèmes limites pour une suite de variables aléatoires satisfaisant une condition de mélange fort, il est souvent utile d'approximer la suite par une autre suite aléatoire vérifiant certaines propriétés. Une première méthode consiste à approximer la suite par une suite de différence de martingales voir [86; 179; 180; 178; 87; 90; 153; 12; 99]. Une deuxième méthode consiste à approximer la suite par une autre suite de variables aléatoires indépendantes. Cette méthode qui est introduite par Bernstein (1927)[10], est utilisée fréquemment pour prouver des T.C.L. pour des suites  $\alpha$ -mélangeantes et a eu un succès plus grand que la première approche. Les principaux T.C.L. prouvés par cette méthode sont ceux donnés dans Ibragimov (1959,1962)[97; 98], Ibragimov et Linnik (1971)[101] et dans Iosifescu-Theodorescu (1969)[104] pour les processus stationnaires et dans Wither (1981)[184] pour des suites non stationnaires. Une bonne revue des principaux T.C.L.

pour des suites mélangées est donnée dans Li et Lu (1996)[125].

Dans cette section nous mettons l'accent sur les principaux résultats qui nous intéressent dans notre preuve du théorème central limite autonormalisé présenté dans le Chapitre 2.

### 1.10.1 T.C.L. pour des processus mélangés

Au départ on commence par expliquer la technique de Bernstein qui a l'avantage de réduire le cas des variables aléatoires dépendantes au cas des variables aléatoires indépendantes. L'idée est de représenter  $S_n = \sum_{i=1}^n X_i$  comme une somme de variables aléatoires (gros blocs) alternant avec d'autres termes (petits blocs) dont la somme est négligeable en vertu d'une normalisation appropriée. Soit  $S_n = \sum_{i=1}^n X_i$ ,  $p = p(n)$ ,  $q = q(n)$  et  $k = k(n)$  avec  $1 \geq p \geq n$ ,  $q = o(p)$ ,  $k = [n/(p+q)]$ , où  $[.]$  denote la partie entière. Dans ce cas la somme  $S_n$  se présente comme suit :

$$S_n = \sum_{j=1}^k \xi_j + \sum_{j=1}^k \eta_j$$

avec

$$\xi_j = \sum_{i=(j-1)(p+q)+1}^{jp+(j-1)q} X_i, \quad \eta_j = \sum_{i=jp+(j-1)q+1}^{j(p+q)} X_i, \quad j = 1, \dots, k, \quad \eta_{k+1} = \sum_{i=k(p+q)+1}^n X_i.$$

Si  $q = o(p)$  est assez grand, alors  $\sum_{j=1}^k \eta_j$  est négligeable par rapport à  $S_n$ . Par ailleurs le comportement asymptotique de  $\sum_{j=1}^k \xi_j$  est mené en utilisant la propriété de mélange.

Ainsi la méthode de Bernstein permet de considérer la somme des variables mélangées comme des variables indépendantes.

Supposons que la suite  $(X_i)_{i \in \mathbb{N}}$  satisfait une condition de mélange fort. En utilisant cette méthode Ibragimov (1962)[98] a prouvé que  $(S_n - B_n)/(A_n)$  converge en distribution vers une distribution stable lorsque  $n \rightarrow \infty$ , avec  $A_n$  et  $B_n$  deux constantes de normalisation. En particulier, si  $(S_n - B_n)/(A_n)$  converge vers une loi normale alors  $(X_i)_{i \in \mathbb{N}}$  satisfait un T.C.L.

Ibragimov (1959,1962)[97; 98] est le premier à avoir donné des conditions nécessaires et suffisantes pour qu'une suite  $\alpha$  mélangée admette un T.C.L.

Une condition nécessaire et suffisante pour que le T.C.L. soit satisfait et qui est plus simple que ceux données dans Ibragimov (1962)[98] est donnée dans le travail de Denker (1985)[49]. En effet, il a été montré dans cet article que pour des variables aléatoires fortement mélangées, le T.C.L. est vrai si la somme normalisée  $\left\{ \frac{S_n^2}{\sigma_n^2}, n \geq 1 \right\}$  est uniformément intégrable (U.I.). On note que le premier résultat qui relie l'uniforme intégrabilité (U.I.) de  $\left\{ \frac{S_n^2}{\sigma_n^2}, n \geq 1 \right\}$  avec le T.C.L. est obtenu par Volkonskiï et Rozanov (1959)[177] (voir aussi Ibragimov et Linnik (1971) Théorème (18.4.2)[101]). Un autre résultat relatif est donné par Dehling, Denker et Philipp (1986)[48].

Rosenblatt (1956)[161] était le premier à avoir donné des conditions suffisantes pour que le T.C.L. et le principe d'invariance faible soient valides. Ibragimov (1962)[98] a par la

suite discuté ce sujet et a obtenu des nouveaux résultats. On reporte ci-dessous le T.C.L. qui a été prouvé dans Ibragimov (1962)[98] (voir aussi Gordin (1969)[86]).

**Théorème 4** (Ibragimov(1962)[98]). *Si le processus aléatoire centré  $(X_i)_{i \in \mathbb{N}}$  est stationnaire et satisfait la condition de mélange fort avec  $\mathbb{E}|X_j|^{2+\delta} < \infty$  pour  $\delta > 0$  et si*

$$\sum_{r=1}^{\infty} \alpha(r)^{\frac{\delta}{2+\delta}} < \infty, \quad (36)$$

alors

$$\sigma^2 := \mathbb{E}(X_1^2) + 2 \sum_{j=2}^{\infty} \mathbb{E}(X_1 X_j) < \infty.$$

De plus, si  $\sigma \neq 0$  alors

$$\frac{S_n}{\sigma \sqrt{n}} \xrightarrow{d} N(0, 1).$$

Dans le deuxième chapitre de la thèse nous avons adapté ce théorème au type de mélange que nos processus satisfont (voir Théorème 2.3.1, Chapitre 2).

**Remarque 1.** *Si les variables aléatoires  $\{X_i, i \in \mathbb{N}\}$  sont bornées c'est-à-dire  $\delta = \infty$ , alors la condition (36) est réduite à*

$$\sum_{n=1}^{\infty} \alpha(n) < \infty. \quad (37)$$

Gordin (1969)[86] a obtenu des conditions suffisantes pour prouver l'asymptotique normalité des sommes partielles des suites mélangeantes. Sa preuve est basée sur les martingales avec des différences stationnaires. Le lecteur peut trouver une explication de cette méthode dans Merlevède *et al.* (2006)[137]. Doukhan, Massart et Rio (1994)[68] ont donné une extension du Théorème d'Ibragimov (1962)[98].

Nous allons maintenant examiner le principe d'invariance faible. Soit  $W_n(t) = \frac{S_{[nt]}}{\sigma_n}$  où  $[x]$  désigne la partie entière la plus grande de  $x$ . La fonction  $w \rightarrow W_n(\cdot, w)$  est une application mesurable d'un espace de probabilité  $(\omega, K) \rightarrow (D, \mathcal{B})$  avec  $D$  l'ensemble de toutes les fonctions dans l'intervalle  $[0, 1]$  qui ont des limites à gauche et qui sont continues à droite et  $\mathcal{B}$  la sigma-algèbre de Borel dans  $D^1$ . Par définition on dit que la suite  $(X_n)_{n \in \mathbb{N}}$  satisfait le T.C.L. fonctionnel ou encore le principe d'invariance faible si

$$W_n \Rightarrow W \text{ lorsque } n \rightarrow \infty, \quad (38)$$

où  $W_n \Rightarrow W$  est utilisée pour la convergence faible de la distribution de  $W_n$  vers la mesure standard de Wiener  $W$ . Davydov (1970)[43] est le premier qui a généralisé le T.C.L. au principe d'invariance pour des variables bornées mais la condition (36) est renforcée par  $\sum_{n=1}^n \alpha(n)^{\frac{1}{2}} < \infty$ . Oodaira et Yoshihara (1972)[143] ont amélioré le résultat de Davydov (1970)[43] et ont obtenu un T.C.L. fonctionnel pour des suites stationnaires sous les mêmes conditions que le T.C.L. (i.e  $\sum_{n=1}^n \alpha(n) < \infty$ ). Herrndorf (1984)[92] a étendu ce résultat en remplaçant l'hypothèse de stationnarité par une hypothèse sur les moments. (voir Théorème suivant)

---

<sup>1</sup> Voir les sections 9 et 10 du livre de Billingsley (1968)[12] pour la définition de la mesure de Wiener et pour le T.C.L. fonctionnel de Donsker pour les suites de variables aléatoires indépendantes et voir sa section 14 pour la convergence faible dans l'espace de Skorohod  $D([0,1])$ .

**Théorème 5** ( Herrndorf (1984)[92]). *Soit la suite  $(X_n)_{n \in \mathbb{N}}$  telle que*

$$\mathbb{E}(X_1) = 0, \quad \mathbb{E}|X_n|^2 < \infty, \quad \text{pour } n \in \mathbb{N}, \quad (39)$$

*et*

$$\frac{\sigma_n^2}{n} \rightarrow \sigma^2, \quad \text{lorsque } n \rightarrow \infty. \quad (40)$$

*S'il existe  $\beta > 0$  tel que*

$$\sup_{n \in \mathbb{N}} \mathbb{E}|X_n|^\beta < \infty, \quad \sum_{k \in \mathbb{N}} \alpha(k)^{1-\frac{2}{\beta}} < \infty, \quad (41)$$

*alors*

$$W_n \Rightarrow W. \quad (42)$$

Un résultat plus général est donné dans Herrndorf (1985)[93].

Doukhan, Massart et Rio (1994)[68] ont discuté le T.C.L. fonctionnel pour des suites  $\alpha$ -mélangeantes strictement stationnaires via la fonction quantile de  $|X_1|$ . Ils ont donné une autre condition suffisante comme suit : Soit  $(X_n)_{n \in \mathbb{N}}$  une suite stationnaire de variables aléatoires  $\alpha$ -mélangeantes centrées telle que la condition (22) est satisfaite. Alors la série  $\sum_{n=1}^{\infty} Cov(X_1, X_n)$  est absolument convergente et  $Z_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} X_k$  converge vers une loi normale. Il convient de noter qu'un grand nombre de travaux concernant aussi le T.C.L. fonctionnel pour des suites dépendantes, on cite ici une liste non exhaustive de ces travaux : Doukhan et Witenberger (2007)[70], Philipp (1986)[152], Mcleish (1975)[134], Philipp et Stout (1975)[153], Rio (2000)[159] Peligrad (1986)[146], Merlevède (2006)[137] et Bradley (2007)[23].

### 1.10.2 Estimation de la variance et T.C.L. autonormalisé

En pratique la variance est inconnue et elle doit être estimée. Dans un T.C.L. si la variance est remplacée par un estimateur à partir des données alors l'estimateur est dit auto-normalisateur, dans ce cas on dit qu'on a un T.C.L. autonormalisé.

Pour les suites dépendantes la méthode la plus naturelle est de remplacer  $\sigma_n^2$  par un estimateur convergent. Ce sujet a été marqué par les travaux de Shao et Peligrad (1994)[148] pour des suites faiblement dépendantes, Jiang and Hahn (2008)[106] et Shao et Peligrad (1995,1996)[149; 150] pour des suites  $\rho$ -mélangeantes, Peligrad et Suresh (1995)[145] pour des suites associées, Carlstein (1986)[29], Shi (2000)[170] pour des suites  $\alpha$ -mélangeantes et par Doukhan *et al.* (2010)[66] pour des suites  $\lambda$ -faiblement dépendantes (voir section 1.6.1 pour plus de détail sur ce type de dépendance faible). Notons que l'estimateur de Shi (2000)[170] n'est que celui de Shao et Peligrad (1995)[148] mais adapté au cas des suites  $\alpha$ -mélangeantes.

De nombreuses techniques d'estimation de la variance ont été développées dans le cas des suites i.i.d. de variables aléatoires. Pour une description de ces techniques voir Efron (1982)[73].

Partant du travail de Carlstein (1986)[29], Peligrad et Shao (1995)[149] et Doukhan *et al.* (2010)[66] ont construit des estimateurs qui se basent sur la division en blocs de la série des données de manière à ce que l'ensemble de tout les blocs forme un échantillon dont la distribution est proche d'une distribution normale  $N(0, \sigma^2)$ .

1. Estimateur de Peligrad et Shao (1995)[149]

Soit  $\{X_n, n \in \mathbb{N}\}$  une suite de variables aléatoires centrées. L'article de Peligrad et Shao (1995)[149] traite les suites de variables aléatoires  $\rho$  mélangeantes. Le découpage en blocs utilisé dans la construction de l'estimateur de la variance est tel que :

$$\tilde{\Delta}_{i,m} = \frac{1}{\sqrt{m}} \sum_{j=i+1}^{i+m} X_j.$$

Les blocs sont d'une longueur  $m = m(n)$ . Le nombre de blocs, noté  $n - m(n)$ , est beaucoup plus grand que  $N(n) = \frac{n}{m}$  et  $\mathbb{E}(F(\Delta))$  est estimée par :

$$\tilde{F}_n = \frac{1}{n - m(n)} \sum_{i=1}^{n-m(n)} F(\tilde{\Delta}_{i,m(n)}).$$

2. Estimateur de Doukhan *et al.* (2010)[66]

Soit  $\{X_n, n \in \mathbb{N}\}$  une suite de variables aléatoires centrées. Notons que l'article de Doukhan *et al.* (2010)[66] traite des vecteurs aléatoires  $\lambda$ - faiblement dépendants. Dans la construction de cet estimateur les blocs qu'ils utilisent sont représentés par :

$$\hat{\Delta}_{i,m} = \frac{1}{\sqrt{m}} \sum_{j=(i-1)(m+\ell)+1}^{(i-1)(m+\ell)+m} X_j. \quad (43)$$

De plus le fait que les blocs définis dans (43) ne s'interceptent pas, implique la nécessité de choisir un autre paramètre  $\ell$  qui lui aussi dépend de  $n$ . Le nombre de blocs est noté par  $N(n) = \frac{n}{m}$  et  $\mathbb{E}(F(\Delta))$  est estimée par :

$$\hat{F}_n = \frac{1}{N(n)} \sum_{i=1}^{N(n)} F(\hat{\Delta}_{i,m(n)}).$$

Dans ces deux cas, la fonction  $F = x'x$  est naturellement la plus adaptée pour l'estimation de la variance de la somme. Notons que dans nos exemples  $C(u, v)$  est une fonction exponentielle de  $u$  et  $v$ . On mentionne que le coefficient de mélange pour des suites de variables aléatoires  $\lambda$ - faiblement dépendantes est une fonction polynomiale de  $u$  et  $v$  (voir Section 1.6.1). Ainsi cet estimateur n'est pas applicable directement aux exemples que nous visons.

Dans notre contexte, on propose l'estimateur de la variance suivant qui est convergent pour des suites de variables aléatoires  $\alpha_{(u,v)}$  mélangeantes. Notons que cet estimateur est une version modifiée de l'estimateur de Peligrad et Shao [149] : Soit  $\ell = \ell_n$

$$B_{n,2}^2 = \frac{1}{n - \ell + 1} \sum_{j=0}^{n-\ell} \left( \frac{S_j(\ell) - \bar{X}_\ell}{\sqrt{\ell}} \right)^2,$$

et

$$\widehat{B}_{n,2}^2 = \frac{1}{n - \ell + 1} \sum_{j=0}^{n-\ell} \left( \frac{S_j(\ell) - \ell\mu}{\sqrt{\ell}} \right)^2,$$

avec

$$S_j(\ell) = \sum_{k=j+1}^{j+\ell} X_k \text{ and } \bar{X}_\ell = \frac{1}{n - \ell + 1} \sum_{j=0}^{n-\ell} S_j(\ell).$$

Nous étudions le comportement asymptotique de cet estimateur en particulier on prouve la faible consistance de cet estimateur (voir Théorème 2.3.2, Chapitre 2) ainsi que sa normalité asymptotique (voir Théorème 2.3.3, Chapitre 2). Notons que le choix de  $\ell$ , qui est une fonction de  $n$ , dépend du comportement de  $C(u, v)$  en fonction de  $u$  et  $v$ . Supposons qu'il existe  $K > 0$ , dans ce cas on a les différents exemples suivants

- Cas 1 :  $C(u, v) \leq K^{u+v}$  alors pour

$$\ell_n = \ell, \ell = [\ln(n)/5 \ln(K)] + x,$$

on peut prouver sous des conditions des moments supplémentaires que  $\widehat{B}_{n,2}^2$  est un estimateur convergent et asymptotiquement normale.

- Cas 2 :  $C(u, v) \leq K$  alors pour  $\ell_n = o(n)$  on peut prouver sous des conditions des moments supplémentaires que  $\widehat{B}_{n,2}^2$  est un estimateur convergent et asymptotiquement normale.
- Cas 3 :  $C(u, v) \leq K(u+v)^\beta$  i.e  $C(u, v)$  polynomial en  $u$  et  $v$  alors pour  $\ell_n = n^s$  avec  $s < \frac{1}{4}$  si  $\beta < 4$  et  $s < \frac{1}{\beta+1}$  si  $\beta \geq 4$  on peut prouver sous des conditions des moments supplémentaires que  $\widehat{B}_{n,2}^2$  est un estimateur convergent et asymptotiquement normale.

Nos preuves sont fondées sur des hypothèses d'uniforme intégrabilité d'une puissance adéquate de la somme normalisée  $|\frac{S_n}{\sqrt{n}}|$ . Cet argument est le même utilisé dans Peligrad et Shao (1995)[149] et Shi (2000)[170].

## 2 Extrema convexes pour des distributions discrètes décroissantes : effets de la contrainte de convexité

### 2.1 Introduction

Commençons par mentionner que nous considérons dans cette partie des distributions décroissantes au sens large. En général, la comparaison des variables aléatoires est fondée sur des fonctions qui caractérisent leurs distributions de probabilité. En sciences actuarielle et financière les comparaisons fondées sur les mesures statiques, tel que l'écart-type, sont dans certains cas limitées. Il est donc nécessaire d'avoir certaines mesures de risque dynamiques qui servent à la comparaison des modèles complexes et à l'analyse des problèmes actuariels divers.

Les ordres stochastiques sont des outils de comparaison importants et permettent d'obtenir des bornes dans de nombreux cas. Ils fournissent des informations précieuses sur le comportement des modèles stochastiques complexes. Ces outils ont été largement utilisés, au cours de ces dernières années, dans des domaines de probabilités et de statistiques variés. Pour une revue des propriétés de base, des concepts et des résultats liés à la théorie des ordres stochastiques, on renvoie le lecteur vers les livres de Goovaerts *et al.* (1998, 1990)[85; 84], Ross (1996)[162], Müller et Stoyan (2002)[139] et Shaked et Shanthikumar (1994,2007)[165; 166].

Dans cette partie de la thèse on s'intéresse à la comparaison de la variabilité des risques d'un portefeuille d'assurance. Dans la littérature, l'ordre le plus important et le plus commun permettant de comparer la variabilité des variables aléatoires est l'ordre convexe.

**Définition 9.** Soient  $X$  et  $Y$  deux variables aléatoires tel que

$$E[f(X)] \leq E[f(Y)] \text{ pour toute fonction convexe } f. \quad (44)$$

alors  $X$  est dite plus petite que  $Y$  au sens de l'ordre convexe et on note  $X \leq_{cx} Y$ .

Ceci implique que  $E(X) = E(Y)$  et  $Var(X) \leq Var(Y)$ , par conséquent, l'ordre convexe est parfois appelé l'ordre de variabilité. Généralement, les fonctions convexes sont des fonctions qui prennent leurs valeurs les plus élevées sur les régions de la forme  $(-\infty, a) \cup (b, \infty)$  pour  $a < b$ . Ainsi, si (44) est vraie alors la variable aléatoire  $Y$  a plus de chance d'avoir des grandes valeurs aux extrêmes que la variable aléatoire  $X$ . On mentionne qu'il est suffisant de considérer des fonctions convexes uniquement sur l'union des supports des variables aléatoires  $X$  et  $Y$  et non pas sur toute la droite des nombres réels.

Dans la littérature (voir e.g Denuit et Lefèvre (1997)[56]), on trouve une définition plus générale. Elle concerne l'ordre  $s$  convexe avec  $s \in \mathbb{N}_0$  tel que  $N_0 = \{1, 2, \dots\}$ . En effet, on dit que  $X$  est plus petite que  $Y$  au sens de l'ordre  $s$ -convexe et on note  $X \leq_{s-cx} Y$  si la relation (45) est vraie

$$E[f(X)] \leq E[f(Y)] \text{ pour toute fonction } s\text{-convexe } f. \quad (45)$$

*i.e.* pour toute fonction  $f \in \mathcal{F}_s$  tel que  $\mathcal{F}_s = \{f : \mathcal{D}_n \rightarrow \mathbb{R} \mid \Delta^s(f(i)) \geq 0, \forall i = 0, 1, \dots, n-s\}$ , avec  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$  et  $\Delta^s(f(i)) = \Delta^{s-1}(f(i+1)) - \Delta^{s-1}(f(i))$ . Notons que, puisque les fonctions  $f(j) = j^i$  et  $f(j) = -j^i$  appartiennent à  $\mathcal{F}_s$  pour tout  $i = 0, 1, \dots, s-1$  et

$j \in \mathbb{N}$  alors les premiers  $s - 1$  moments de  $Y$  et  $X$  sont nécessairement égaux. C'est dans ce sens ou on peut remarquer que le problème d'optimization  $s$ - convexe correspond à un problème traditionnel de moments (voir e.g. Hoeffding (1955)[95], Karlin et Studden (1966)[112], Kemperman (1968)[113], Utev (1985)[174], Prékopa (1990)[155] et Hurlimann (1999)[96]). L'ordre  $\leq_{1-cx}$  est tout simplement l'ordre stochastique connu dans la littérature économique par la dominance stochastique de premier ordre. Cet ordre est basé sur la comparaison des fonctions de répartition. Si  $s = 2$ , alors  $\leq_{2-cx}$  est l'ordre convexe usuel.

L'une des difficultés souvent rencontrées dans les probabilités appliquées réside dans le manque d'une connaissance statistique complète sur certaines composantes des modèles en question, notamment dans le cas de l'assurance et de la finance. En effet, les risques et les pertes sont souvent difficiles à estimer. Par exemple, pour une distribution de sinistres, les informations partielles à disposition peuvent être liées à la forme de la distribution (par exemple croissante, décroissante) les premiers moments et certaines contraintes de forme (par exemple concave, convexe). Dans notre travail, on s'intéresse aux distributions de probabilité discrètes et décroissantes. Dans le contexte des distributions discrètes, les ordres stochastiques ont été examinés par Lefèvre et Picard (1993)[121], Fishburn et La-valle (1995)[77], Lefèvre et Utev (1995)[122], Lefèvre et Utev (1996)[123], Denuit et Lefèvre (1997)[56] et Denuit *et al.* (1999b,1999c)[57; 60]. Notons que notre intérêt pour les distributions décroissantes provient de certains exemples en assurance où il est logique d'imposer une telle propriété sur certaines distributions de risque (voir e.g Denuit *et al.* (2000)[59]).

## 2.2 Problématiques de recherche

Dans un contexte actuariel, les risques extrêmes représentent les risques les plus et les moins dangereux. Leurs connaissances peuvent donner des bornes inférieures et supérieures sur des quantités diverses de risques (comme la prime). De plus, dans le nouveau projet Solvabilité II, l'exigence de marge de solvabilité doit être calculée selon les nouvelles règles, qui sont fondées sur l'analyse des scénarios de risque. Depuis, les assureurs font de plus en plus attention au risque extrême. Dans le cas discret, des formules explicites pour les extrema convexes d'ordre  $s = 1, 2, 3, 4$  ont été obtenus dans Denuit et Lefèvre (1997)[56], Denuit *et al.* 1999b,1999c)[57; 60] et Courtois *et al.* (2006)[38]. Dans le cas des variables réelles, les extrema convexes ont été étudiés par Jansen *et al.* (1986)[105] De Vylder (1996)[44] et Denuit *et al.* (1998,1999a)[58; 52].

Dans notre travail, nous considérons uniquement les extrema 2- convexes que nous appellerons les extrema convexes et que nous noterons  $\leq_{cx}$ . Considérons la classe des distributions discrètes dont les fonctions de masse de probabilité sont décroissantes sur un support  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Les extrema convexes dans cette classe de distributions (voir (3.3) et (3.4) dans Denuit *et al.*(1999b)[57]). Soit  $N$  une variable aléatoire arbitraire avec une fonction de masse de probabilité (f.m.p.) décroissantes sur un support  $\mathcal{D}_n$  de moyenne  $\nu = E(N)$ . Dans ce cas on a  $N_{min} \leq_{cx} N \leq_{cx} N_{max}$  avec

$$N_{max} = \begin{cases} 0 & \text{avec probabilité } 1 - 2\nu/(n + 1), \\ 1, \dots, n & \text{avec probabilités } 2\nu/n(n + 1), \end{cases}$$

et

$$N_{min} = \begin{cases} 0, \dots, \xi & \text{avec des probabilités égales } 2(\xi + 1 - \nu)/(\xi + 1)(\xi + 2), \\ \xi + 1 & \text{avec des probabilités } (2\nu - \xi)/(\xi + 2). \end{cases}$$

tel que  $\xi$  est un entier dans  $[0, n - 1]$  satisfaisant  $\xi < 2\nu \leq \xi + 1$ . Dans certaines situations, il est possible de disposer de quelques informations supplémentaires sur la forme de la fonction de masse de probabilité. La question qu'on se pose est la suivante : comment les extrema précédents sont-ils modifiés sous des contraintes de forme supplémentaires de type convexité ? En effet, dans ce travail on considère l'hypothèse selon laquelle la f.m.p. est non seulement décroissante sur  $\mathcal{D}_n$ , mais aussi convexe globalement sur  $\mathbb{N}$  ou sur une partie de  $\mathbb{N}$ .

Deux cas sont considérés pour la f.m.p. :

1. globalement convexe sur  $\mathbb{N}$  ;
2. convexe seulement à partir d'un point positif donné  $m$ .

De nombreuses illustrations en théorie de ruine sont présentées.

Les résultats qu'on a trouvés fournissent un complément aux deux travaux récents de Lefèvre et Loisel (2010,2012)[119; 120]. Effectivement, Lefèvre et Loisel (2010)[119] ont donné une nouvelle approche unificatrice pour construire les extrema  $s$ -convexes pour des distributions discrètes ou continues qui sont  $t$ -monotones. On rappelle la définition suivante des distributions discrètes  $t$ -monotones

**Définition 10.** Une fonction  $f(j)$ ,  $j \geq 0$  est  $t$  monotone,  $t \geq 1$  si

$$(-1)^k \Delta^k f(j) \geq 0, j \geq 0 \text{ for } k = 1, \dots, t.$$

avec

$$\Delta f(j) = f(j + 1) - f(j).$$

En d'autres termes une variable aléatoire discrète est dite  $t$ -monotone si les signes des premières différences  $t$  de la fonction de masse de probabilité alternent comme suit  $(-, +, \dots, (-1)^t)$ .

En particulier, pour  $t = 2$ , la f.m.p. est dite 2-monotone et correspond au cas où elle est décroissante et convexe.

La méthode de Lefèvre et Loisel (2010)[119] est fondée sur l'opérateur stationnaire des excès et sur ses itérations. Dans cet article, les auteurs prouvent que cet opérateur transforme un ordre  $s + 1$ -convexe en un ordre  $s$ -convexe pour des distributions décroissantes. Notons que pour  $s = 2$ , les bornes dérivées dans Lefèvre et Loisel (2010)[119] sont valables pour des variables aléatoires qui ont des f.m.p. décroissantes convexes uniquement sur leurs supports et pas globalement sur  $\mathbb{N}$ . Lefèvre et Loisel (2012)[120] considèrent une caractérisation des distributions par un mélange d'un nombre minimum de fonctions uniformes. Ils ont prouvé que l'ordre  $s$ -convexe de la distribution du mélange implique un ordre  $s$ -convexe de ces distributions. Notons que pour  $s = 2$  les bornes dérivées dans Lefèvre et Loisel (2012)[120] sont valables pour des variables aléatoires qui ont des f.m.p. décroissantes convexes globalement sur  $\mathbb{N}$ .

La méthode que nous utiliserons dans les preuves de nos résultats est plus simple. Elle est fondée sur des simples propriétés de croisement entre deux distributions. Dans le paragraphe suivant, on donne un aperçu sur les principaux résultats de la littérature qui seront utilisés dans nos preuves.

### 2.3 Littérature : quelques résultats de base

Définissons par  $S^-$  l'opérateur qui, lorsqu'il est appliqué à une fonction compte le nombre de changements de signe de cette fonction sur son domaine. On dit que deux f.m.p.  $P_X$  et  $P_Y$  se croisent  $k$  fois,  $k \in \mathbb{N}$ , si  $S^-(P_X - P_Y) = k$ . Le résultat suivant montre que si nous avons deux variables aléatoires avec des moyennes égales alors nous avons au moins deux points de passage entre leurs f.m.p. (voir Denuit et Lefèvre (1997)[56]).

**Proposition 3.** (Denuit et Lefèvre (1997)[56])

$$\text{Si } E(X) = E(Y), \text{ alors } S^-(P_Y - P_X) \geq 2. \quad (46)$$

Le résultat suivant donne une condition de type croisement qui implique un ordre convexe.

**Proposition 4.** (Denuit et Lefèvre (1997)[56])

$$\text{Si } E(X) = E(Y) \text{ et } S^-(P_Y - P_X) = 2 \text{ avec } P_Y \geq P_X \text{ proche de } n, \text{ alors } X \leq_{cx} Y. \quad (47)$$

Plus clairement cette proposition montre que  $X \leq_{cx} Y$  si la f.m.p de  $Y$  a des valeurs plus lourdes sur les extrêmes. Mentionnons que toutes nos preuves des extrema convexes sont fondées sur ces deux propositions démontrées par Denuit et Lefèvre (1997)[56] que nous avons redémontrées en proposant des preuves plus simples.

### 2.4 Principaux résultats

Dans chacune des trois situations énumérées précédemment nous avons donné les bornes convexes supérieure et inférieure. Nous avons également comparé nos résultats avec ceux donnés dans Lefèvre et Loisel (2010,2012)[119; 120]. Ces résultats seront utilisés pour évaluer la variabilité d'un risque donné et pour l'estimation du capital de solvabilité requis. Nous avons repris les exemples de Lefèvre et Loisel (2010,2012)[119; 120] pour pouvoir comparer numériquement nos résultats.

## Première partie

# Modèle à facteur commun pour la modélisation de la dépendance en théorie du risque



## Chapitre 1

# Quelques propriétés de mélange pour des processus conditionnellement indépendants (C.I.)

## Résumé

Dans ce premier chapitre on considère des processus conditionnellement indépendants par rapport à un facteur dynamique. Plus précisément on suppose que pour tout  $i \in \mathbb{N}$ , les variables aléatoires  $X_1, \dots, X_i$  sont conditionnellement indépendantes par rapport à un vecteur  $\underline{V}_i = (V_1, \dots, V_i)$ . Nous étudions les propriétés de mélange du processus  $(X_i)_{i \in \mathbb{N}}$  lorsque le conditionnement est donné par rapport à une mémoire du facteur non bornée. Notre travail est motivé par des exemples concrets liés à la théorie du risque. En statistiques, le concept du mélange qui est introduit pour la première fois par Rosenblatt (1956) est une notion de dépendance faible connue aussi par dépendance de courte mémoire, il permet de décrire l'asymptotique indépendance entre le passé et le futur d'une suite de variables aléatoires lorsque le trou entre le passé et le futur est assez grand. Après une brève introduction, on commence par donner quelques définitions des processus mélangeants, des processus conditionnellement mélangeants et des processus conditionnellement indépendants. Par la suite on indique notre principal résultat dans le Théorème (1.2.1). Ce résultat montre que si le processus du facteur est mélangeant et si la suite satisfait une structure de mélange conditionnel donnée alors le processus des v.a. conditionnellement indépendantes satisfont certaines propriétés de mélange. Dans la dernière partie de ce chapitre nous introduisons quelques exemples de processus en théorie du risque satisfaisant nos hypothèses de mélange. Ce chapitre est un article soumis intitulé « Some mixing properties of conditionally independent processes » co-écrit avec Stéphane Loisel et Véronique Maume-Deschamps.

## 1.1 Introduction

In risk theory the structure of dependence between individual risks may come from a time-varying common factor which represents the evolution of socio-economic and natural environment. Markov modulated models have been applied to risk theory for the first time by Asmussen (1989)[2]. Cossette *et al.* (2002)[11] studied a Poisson risk model with common mixture on which the factor is represented by a discrete random variable. More recently, Cossette *et al.* (2004)[12] proposed a compound binomial model modulated by a Markovian environment as an extension to the compound binomial model proposed by Gerber (1988)[17].

Denote by  $\underline{V}_n$  the random vector  $(V_1, \dots, V_n)$  corresponding to the evolution of the factor up to time  $n \geq 1$ . Cossette *et al.* (2002)[11] considered a particular case of conditional independence: for any multi-indices  $(i_1, \dots, i_u)$ ,  $i_1 < \dots < i_u$ , any integer  $n \geq i_u$  and any measurable bounded function  $f$ :

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_n) = \int_{x_{i_1}, \dots, x_{i_u}} f(x_{i_1}, \dots, x_{i_u}) \prod_{j=1}^u N^{i_j}(dx_{i_j}, V_{i_j}), \quad (1.1)$$

where the measure  $N^i(dx_i, V_i)$  denotes the transition kernel of  $X_i$  given  $V_i$ .

Due to claim settlement delays, accounting standards and risk time-accumulation phenomena, the whole memory of the factor often plays an important role in the risk process evolution in insurance.

Past crises, important events, and other factors contribute both to claim frequency and amounts and may have long-lasting effects, because such events modify the prices and the contracts offered by the insurer and by competitors, which in turn changes the characteristics of the remaining and new policyholders of the company. Consequently, it is not always enough to rely on a finite memory, even if recent events often have more importance than old ones. One may wonder whether these features affect significantly the risk faced by the insurer in a multi-year framework. To our knowledge, conditionally independent variables with respect to a random vector with time-varying length have not yet been considered in the literature. The aim of this paper is to extend the dependence frame introduced in Cossette *et al.* (2002)[11] by considering that dependence among risks may be explained by an unbounded memory of the common factor. This factor is not necessarily a Markov chain and may influence claim amounts and/or frequencies. Our work is then a first step toward the modeling of long memory effects of a factor. We study the dependence properties of our model but more work has to be done in order to provide, for example, asymptotic behavior of sums and quantify asymptotic diversification.

In this paper, we consider that random variables (r.v's)  $(X_n)_{n \in \mathbb{N}}$  are conditionally independent given the entire trajectory of the factor. More precisely, if we denote by  $N^j(dx_j, \underline{V}_j)$  the transition kernel of  $X_j$  given  $\underline{V}_j$ , we assume that for any multi-indices  $(i_1, \dots, i_u)$ ,  $i_1 < \dots < i_u$  and any integer  $n \geq i_u$

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_n) = \int_{x_{i_1}, \dots, x_{i_u}} f(x_{i_1}, \dots, x_{i_u}) \prod_{j=1}^u N^{i_j}(dx_{i_j}, \underline{V}_{i_j}). \quad (1.2)$$

Note that in this last case sequences  $(\underline{V}_j)_{(j \in \mathbb{N})}$  overlap. We introduce our definition of mixing sequences and we derive some mixing properties for random processes subject to

our time dependence model. Mixing processes have been initially studied by Rosenblatt (1956) [33], Kolmogorov and Rosanov (1960)[25], Ibragimov (1962)[21] and Ibragimov and Linnik (1971)[22].

Mixing conditions on sequences of random variables express asymptotic independence between past and distant future events. This means that past is progressively forgotten. A simple way to check mixing conditions is to provide a precise control of the covariance as the distance between past and future goes to infinity or is sufficiently large. For a unified presentation of the main mixing coefficients and their properties we refer to Bradley and Bryc (1985)[6], Peligrad (1986)[29], Bradley (1986,2005)[7; 8], Merlevède *et al.* (2006)[27] and Billingsley (1968)[4]. A large number of references concerning mixing coefficients and related dependence measures are given in Bradley (2006)[8] and Doukhan (1994)[14]. We follow the approach of Dedecker *et al.* (2007)[13] where the concepts and properties of weakly dependent processes are introduced.

It is well known that mixing properties are a key tool to derive moment conditions and asymptotic properties of aggregated processes. This is very important in economics and in actuarial science. In order to study asymptotic behavior of mixing sequences, one may either approximate mixing sequences with independent ones following Bernstein's method (1927)[3], or use Gordin's method (1969)[18] and approximate mixing sequences by a sequence of martingale differences (see e.g [19; 20; 30; 26] for further details).

The paper is organized as follows. In Section 1.2 we give a definition of mixing sequences thanks to a generalization of the one introduced in Doukhan and Louhichi (1999) for weakly dependent r.v.'s, as well as the definitions of conditionally independent r.v.'s and conditionally mixing sequences. Our main result is then stated in Theorem 1.2.1 in which we provide some mixing properties of the process  $(X_i)_{i \in \mathbb{N}}$  under necessary conditions on the mixing properties of the factor process  $(V_i)_{i \in \mathbb{N}}$  and on the conditional mixing structure. Section 1.3 is devoted to the proof of our main theorem. Section 1.4 deals with some examples that satisfy our assumptions. These examples are relevant in risk theory contexts.

## 1.2 Definitions and statement of results.

We first recall some standard definitions of mixing processes. Then we state our standing assumptions on conditional mixing and provide some basic properties. Finally, we state our main result (Theorem 1.2.1) which relates the mixing properties of conditionally independent processes to the mixing properties of the factor process and to the structure of the conditional independence.

### 1.2.1 Mixing sequences

Doukhan and Louhichi (1999)[15] introduced a unified weak dependence condition, more general than mixing condition, in which they make explicit the asymptotic independence between functions of the past and of the future of a sequence of random variables (and not between the past and the future of a sigma algebra as in Rosenblatt (1956)[33]).

**Definition 1.2.1.** *Doukhan and Louhichi (1999)[15]*

*The sequence  $(X_1, \dots, X_n)$  of random variables is called  $(\theta, H, \psi)$ - weakly dependent, if there exist a class  $H$  of real-valued functions, a sequence  $\theta = (\theta_r)_{r \in \mathbb{N}}$  decreasing to 0 at*

infinity, and a function  $\psi$  with arguments  $(f, g, u, v) \in H^2 \times N^2$  such that for all  $r \in \mathbb{N}$  and any  $i_1 \leq \dots \leq i_u < i_u + r \leq j_1 \leq \dots \leq j_v$  one has

$$|Cov[f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v})]| \leq \psi(f, g, u, v)\theta_r, \quad (1.3)$$

for all functions  $f, g \in H$  that are defined respectively on  $\mathbb{R}^u$  and  $\mathbb{R}^v$ . Note that  $\theta$  depends on the class  $H$  and on  $\psi$ . The function  $\psi$  may depend on  $u$  and  $v$ , which is not the case for classical strongly mixing sequences.

Note that if a sequence  $(X_1, \dots, X_n)$  is strongly mixing in the sense of Rosenblatt (1956)[33], then it is  $(\theta, L^\infty, \psi)$ -weakly dependent where  $L^\infty$  denotes the class of real-valued and bounded functions on some space of the form  $\mathbb{R}^p$ . Following the definition introduced by Doukhan and Louhichi (1999)[15] we introduce the following definition.

Note that, in order to distinguish the mixing notion that we consider from the classical notion of mixing, we choose a studied notation for our mixing coefficients.

**Definition 1.2.2.** Consider  $\|\cdot\|_a$  and  $\|\cdot\|_b$  two norms on bounded functions. Let  $u, v$  be integers, a random process  $(X_1, \dots, X_n)$  is said to be  $\eta_{(u,v)}^*$ -mixing, if there exist a function  $r \mapsto \eta^*(r)$  decreasing to 0 as  $r$  goes to infinity and a constant  $C(u, v) > 0$  such that for any real valued bounded functions  $f$  and  $g$  on  $\mathbb{R}^u$  and  $\mathbb{R}^v$  respectively, for any multi-indices satisfying the relation  $(\star)$ :

$$i_1 < \dots < i_u < i_u + r \leq j_1 < \dots < j_v, \quad (\star)$$

and for any  $r > 0$ , we have

$$|Cov[f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v})]| \leq C(u, v)\eta^*(r)\|f\|_a\|g\|_b. \quad (1.4)$$

The most used norms are the  $L^p$  ones ( $p \in \mathbb{N}^* \cup \{\infty\}$ ) denoted by  $\|\cdot\|_p$ , related to the law of  $(X_i)_{i \in \mathbb{N}^2}$ . Depending on the norms used in the definition of mixing, we have various kind of mixing as referred to the classical strongly mixing definitions, more precisely,

1. If  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_\infty$ , we shall say that the process is  $\alpha_{(u,v)}^*$ -mixing and we shall write  $\alpha^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\alpha_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\alpha$ -mixing in the sense of Rosenblatt (1956)[33].
2. If  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_2$ , we shall say that the process is  $\rho_{(u,v)}^*$ -mixing and we shall write  $\rho^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\rho_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\rho$ -mixing in the sense of Kolmogorov and Rozanov (1960)[25].
3. If  $\|\cdot\|_a = \|\cdot\|_1$  and  $\|\cdot\|_b = \|\cdot\|_\infty$ , we shall say that the process is  $\Phi_{(u,v)}^*$ -mixing and we shall write  $\Phi^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\Phi_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\Phi$ -mixing in the sense of Ibragimov (1962).
4. If  $\|\cdot\|_a = \|\cdot\|_b = \|\cdot\|_1$ , we shall say that the process is  $\Psi_{(u,v)}^*$ -mixing and we shall write  $\Psi^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\Psi_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\Psi$ -mixing in the sense of Blum et al. (1963)[5].

<sup>2</sup>If  $f$  is a bounded function on  $\mathbb{R}^u$ , we shall denote for  $p \in \mathbb{N}$ ,  $\|f\|_p^p = \mathbb{E}(|f|^p(X_{i_1}, \dots, X_{i_u}))$  which depends on the index  $(i_1, \dots, i_u)$

For classical strongly mixing definitions, it is well-known that every  $\Psi$ -mixing sequence is  $\Phi$ -mixing, every  $\Phi$ -mixing sequence is  $\rho$ -mixing and every  $\rho$ -mixing sequence is  $\alpha$ -mixing. The same proof remains valid to get the chain

$$\Psi_{(u,v)}^* \text{-mixing} \implies \Phi_{(u,v)}^* \text{-mixing} \implies \rho_{(u,v)}^* \text{-mixing} \implies \alpha_{(u,v)}^* \text{-mixing}.$$

Of course, sequences of independent random variables and  $m$ -dependent sequences satisfy mixing conditions. Various examples of processes satisfying mixing conditions are presented in the books of Doukhan (1994)[14] and Bradley (2007)[9].

**Remark 1.2.1.** For other choices of  $\|\cdot\|_a, \|\cdot\|_b$  (for example Lipschitz norms or bounded variation norms), we shall recover the notions of weak dependence (see Dedecker et al. (2007) [13]).

Note that concerning the classical strongly mixing definitions,  $C(u, v)$  is uniform in  $u$  and  $v$ .

Let us now introduce the main notions of conditional independence that we shall use.

## 1.2.2 Conditional independence

We shall make the following standing assumptions.

- H1: For all  $i \in \mathbb{N}$ , r.v's  $X_1, \dots, X_i$  are conditionally independent given  $\underline{V}_i = (V_1, \dots, V_i)$ .
- H2: The conditional law of  $X_i | \underline{V}_n$  is the same as the one of  $X_i | \underline{V}_i$ , for all  $n \geq i$ .

Assumptions H1 and H2 are equivalent to:

For any multi index  $i_1 \leq i_2 \dots \leq i_u$ , for any  $n \geq i_u$ , for any bounded measurable function  $f$  on  $\mathbb{R}^u$ , we have

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u} | \underline{V}_n)) = \int_{x_{i_1}, \dots, x_{i_u}} f(x_{i_1}, \dots, x_{i_u}) \prod_{j=1}^u N^{i_j}(dx_{i_j}, \underline{V}_{i_j}). \quad (1.5)$$

where the measure  $N^{i_j}(dx_{i_j}, \underline{V}_{i_j})$  denotes the transition kernel of  $X_{i_j}$  given  $\underline{V}_{i_j}$ .

**Proposition 1.2.1.** Let  $u, v$  be integers and let  $f$  and  $g$  be real valued bounded measurable functions on  $\mathbb{R}^u$  and  $\mathbb{R}^v$  respectively. Consider the multi-indices  $i_1 < i_2 \dots < i_u < i_u + r \leq j_1 < j_2 \dots < j_v$ . If H1 and H2 hold then

$$\begin{aligned} & Cov[f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v})] \\ &= Cov[\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v})]. \end{aligned} \quad (1.6)$$

This means that the covariance of  $f(X_{i_1}, \dots, X_{i_u})$  and  $g(X_{j_1}, \dots, X_{j_v})$  is equal to the covariance of their conditional expectation whenever conditioning is done respectively with respect to  $\underline{V}_{i_u}$  and  $\underline{V}_{j_v}$ .

*Proof.* Denote by  $X_{i_1}^{i_u} = (X_{i_1}, \dots, X_{i_u})$  and by  $X_{j_1}^{j_v} = (X_{j_1}, \dots, X_{j_v})$ . For any multi-indices  $i_1 < i_2 < \dots < i_u < j_1 < j_2 < \dots < j_v$  and for any bounded functions  $f$  and  $g$  on  $\mathbb{R}^u$  and  $\mathbb{R}^v$  respectively, we have:

$$\begin{aligned} & \text{Cov} \left[ f(X_{i_1}^{i_u}), g(X_{j_1}^{j_v}) \right] \\ &= \mathbb{E} \left[ f(X_{i_1}^{i_u}) \cdot g(X_{j_1}^{j_v}) \right] - \mathbb{E} \left[ f(X_{i_1}^{i_u}) \right] \mathbb{E} \left[ g(X_{j_1}^{j_v}) \right] \\ &= \mathbb{E} \left[ \mathbb{E}(f(X_{i_1}^{i_u}) | \underline{V}_{i_u}) \cdot \mathbb{E}(g(X_{j_1}^{j_v}) | \underline{V}_{j_v}) \right] - \mathbb{E} \left[ \mathbb{E}(f(X_{i_1}^{i_u}) | \underline{V}_{i_u}) \right] \mathbb{E} \left[ \mathbb{E}(g(X_{j_1}^{j_v}) | \underline{V}_{j_v}) \right] \\ &= \text{Cov} \left[ \mathbb{E}(f(X_{i_1}^{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}^{j_v}) | \underline{V}_{j_v}) \right]. \end{aligned}$$

□

### 1.2.3 Conditionally mixing sequences given an unbounded memory

In our dependence framework we propose the following definition for conditionally strongly mixing sequences.

**Definition 1.2.3.** Let  $u, v$  and  $r$  be integers and let  $\| \cdot \|_a$  and  $\| \cdot \|_b$  be norms on bounded functions (or on subspaces of bounded functions).

The sequence of  $r.v$ 's  $(X_n)_{n \in \mathbb{N}}$  is called conditionally  $\eta_{(u,v)}^*$ -mixing with respect to  $\underline{V}_n$  if there exist a positive sequence  $\eta^*(r) \rightarrow 0$  as  $r \rightarrow \infty$  and a constant  $C(u, v) > 0$  such that the following inequality holds for any bounded functions  $f : \mathbb{R}^u \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^v \rightarrow \mathbb{R}$  and for any multi-indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$

$$\begin{aligned} & \left| \text{Cov} \left[ \mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v}) \right] \right| \\ & \leq C(u, v) \eta^*(r) \|f\|_a \|g\|_b. \end{aligned} \tag{1.7}$$

**Remark 1.2.2.** Note that Definition 1.2.3 can be seen as a generalization of Definition 1.2.1 of Doukhan and Louhichi (1999)[15] by replacing in (1.7),  $\underline{V}_{i_u}$  with  $\underline{X}_{i_1}^{i_u}$  and  $\underline{V}_{j_v}$  with  $\underline{X}_{j_1}^{j_v}$ .

**Remark 1.2.3.** In the same spirit as for Definition 1.2.2, with respect to the choice of the norms  $\| \cdot \|_a$  and  $\| \cdot \|_b$ , we shall say that the process is  $\alpha_{(u,v)}^*$ <sup>3</sup>, or  $\rho_{(u,v)}^*$ <sup>4</sup>, or  $\Phi_{(u,v)}^*$ <sup>5</sup> or  $\Psi_{(u,v)}^*$ <sup>6</sup> conditionally mixing with respect to  $(\underline{V}_n)$ .

Let us mention that the terminology *conditionally mixing* has been used with different meanings by Parakasa Rao (2009)[31], Dziubdziela (1986)[16] and Veijanen (1990)[34]. In particular, the work of Veijanen is concerned with partially observable random fields. The author considers that  $(X_i, \dots, X_j)$  is conditionally independent of  $(V_r, r \notin [i, j])$  given  $(V_i, \dots, V_j)$  for any positive integers  $i$  and  $j$ , which is not our case here.

Below, we give an example showing that the conditional independence related to the whole vector  $\underline{V}_i$  implies a dependence that may be persistent for the process  $(X_i)_{i \in \mathbb{N}}$ . In particular, our conditional independence condition and some mixing condition on the factor process  $(V_i)_{i \in \mathbb{N}}$  are not sufficient to get mixing properties for the process  $(X_i)_{i \in \mathbb{N}}$ .

<sup>3</sup>  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_\infty$

<sup>4</sup>  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_2$

<sup>5</sup>  $\| \cdot \|_a = \| \cdot \|_1, \| \cdot \|_b = \| \cdot \|_\infty$

<sup>6</sup>  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_1$

**Example 1.2.1.** Consider a r.v  $I_i$  such that,  $\forall i \geq 1$ ,  $P(I_i = 1 | \underline{V}_i) = \frac{1}{2}(V_1 + V_i)$ . Let  $(V_i)_{i \in \mathbb{N}}$  be a sequence of identically distributed r.v's with  $\mathbb{E}(V_1) = p$ . Then we have  $P(I_i = 1) = \mathbb{E}(V_1) = p$  and  $P(I_r = 1) = \mathbb{E}(V_1) = p$ . Remark that:

$$P(I_1 = 1, I_r = 1 | \underline{V}_r) = P(I_1 = 1 | V_1)P(I_r = 1 | \underline{V}_r) = \frac{1}{2}V_1(V_1 + V_r).$$

So that, we have  $P(I_1 = 1, I_r = 1) = \frac{1}{2}\mathbb{E}[V_1 \times (V_1 + V_r)]$ . We thus get

$$\begin{aligned} & P(I_1 = 1, I_r = 1) - P(I_1 = 1)P(I_r = 1) \\ &= \frac{1}{2} [\mathbb{E}(V_1^2) + \mathbb{E}(V_1 V_r) - \mathbb{E}(V_1)(\mathbb{E}(V_1) + \mathbb{E}(V_r))] \\ &= \frac{1}{2} [\mathbb{E}(V_1^2) - \mathbb{E}(V_1)^2 + \underbrace{\mathbb{E}(V_1 V_r) - \mathbb{E}(V_1)\mathbb{E}(V_r)}_{(**)}]. \end{aligned}$$

The factor process  $(V_i)_{i \in \mathbb{N}}$  is independent so that  $(**) = 0$  and then,

$$P(I_1 = 1, I_r = 1) - P(I_1 = 1)P(I_r = 1) = \frac{1}{2}p(p - 1)$$

for all  $r \in \mathbb{N}$ .

Therefore, this example shows that the conditional independence condition is not sufficient to have mixing properties, even if the factor is independent. This is due to the fact that all the factor's sequence is taken into account in the conditioning.

Proposition 1.2.1 proves that if the process  $(X_i)_{i \in \mathbb{N}}$  is conditionally independent and conditionally  $\eta_{(u,v)}^*$ -mixing with respect to  $(V_i)_{i \in \mathbb{N}}$  then it is  $\eta_{(u,v)}^*$ -mixing, with the same kind of  $\eta^*$ -mixing.

Before we state our main result, we give some definitions and basic facts about transition kernels.

#### 1.2.4 Transition kernel

Recall that, in our context, a transition kernel (see Revuz (1984)) is a mapping

$$N : \mathcal{B}(\mathbb{R}^d) \times \mathbb{R}^k \longrightarrow \mathbb{R}^+$$

such that

- for any  $v \in \mathbb{R}^k$ ,  $N(\cdot, v)$  is a probability measure,
- for any  $A \in \mathcal{B}(\mathbb{R}^d)$  the mapping  $v \rightsquigarrow N(A, v)$  is measurable.

**Example 1.2.2.** In the example of Section 1.4.2.1, the transition kernel of  $X_i$  given  $\underline{V}_i$ , where  $X_i = I_i \times B_i$ , is:

$$N^i(\cdot, \underline{v}_i) = \sum_{a \in \{0,1\}} \delta_{\{a\}} \mu^a \mathbb{P}(I_i = a | \underline{V}_i = \underline{v}_i),$$

where  $\mu^1$  is a probability measure and  $\mu^0 \equiv 1$ .

With a slight abuse of notation, we shall define some measures associated to transition kernels.

- Let  $N(\cdot, \cdot)$  be a transition kernel on  $\mathcal{B}(\mathbb{R}^d) \times \mathbb{R}^k$ , if  $Y$  is a random vector in  $\mathbb{R}^k$ , having distribution function  $F_Y$ , then  $\mathbb{E}_Y(N)$  is defined by

$$\mathbb{E}_Y(N)(A) = \int N(A, dF_Y) = \int N(A, y) dF_Y(y),$$

we shall also denote it by  $\|N_Y\|_1$ , it defines a probability measure on  $\mathcal{B}(\mathbb{R}^d)$ . Of course, if  $N$  denotes the conditional law of  $X$  given  $Y$  then  $\mathbb{E}_Y(N)$  is the law of  $X$ .

- Let  $N(\cdot, \cdot)$  and  $M(\cdot, \cdot)$  be two transition kernels on  $\mathcal{B}(\mathbb{R}^d) \times \mathbb{R}^k$  and  $\mathcal{B}(\mathbb{R}^p) \times \mathbb{R}^q$ , for  $Y$  and  $Z$  random vectors of  $\mathbb{R}^k$  and  $\mathbb{R}^q$ , the expectation of  $N \times M$  is defined as: for any  $A \in \mathcal{B}(\mathbb{R}^d)$  and  $B \in \mathcal{B}(\mathbb{R}^p)$ ,

$$\mathbb{E}_{Y,Z}(N \times M)(A \times B) = \int N(A, y) \times M(B, z) dF_{(Y,Z)}(y, z).$$

If  $N$  (resp.  $M$ ) denotes the conditional law of  $X$  (resp.  $W$ ) given  $Y$  (resp.  $Z$ ) then  $\mathbb{E}_{Y,Z}(N \times M)$  is the law of the couple  $(X, W)$ .

- The covariance of two transition kernels is then the signed measure:

$$\text{Cov}_{Y,Z}(N, M)(A \times B) = \mathbb{E}_{Y,Z}(N \times M)(A \times B) - \mathbb{E}_Y(N)(A) \times \mathbb{E}_Z(M)(B).$$

In the same spirit, one would like to associate to each norm on bounded functions of  $\mathbb{R}^k$  a new measure: let  $\|\cdot\|_a$  be a norm on bounded functions of  $\mathbb{R}^k$ , for any  $A \in \mathcal{B}(\mathbb{R}^d)$ , let

$$\|N\|_a(A) = \|N(A, \cdot)\|_a.$$

In general, this does not define a measure (because of the lack of additivity).

**Definition 1.2.4.** We shall say that the transition kernel  $N(\cdot, \cdot)$  is  $\|\cdot\|_a$ -dominated if it exists a finite measure  $\mathcal{N}_a$ , on  $\mathcal{B}(\mathbb{R}^d)$  such that for any  $A \in \mathcal{B}(\mathbb{R}^d)$ ,

$$\|N(A, \cdot)\|_a \leq \mathcal{N}_a(A).$$

In particular, we shall work with  $\|\cdot\|_\infty$ -dominated transition kernels. Remark that if the transition kernel  $N$  is  $\|\cdot\|_\infty$ -dominated by a bounded measure  $\mathcal{N}_\infty$ , then for any bounded function  $f$  on  $\mathbb{R}^d$ ,

$$\int f(x) N(dx, y) dF_Y(y) \leq \mathcal{N}_\infty(f).$$

All these notions may extend to more general positive kernels, namely to mappings

$$E : \mathcal{B}(\mathbb{R}^d) \times \mathbb{R}^k \longrightarrow \mathbb{R}^+$$

such that

- for any  $v \in \mathbb{R}^k$ ,  $E(\cdot, v)$  is a finite measure,
- for any  $A \in \mathcal{B}(\mathbb{R}^d)$  the mapping  $v \rightsquigarrow E(A, v)$  is measurable.

The transition kernel considered in Example 1.2.2 is  $\|\cdot\|_\infty$ -dominated by the measure:

$$\sum_{a \in \{0,1\}} \delta_{\{a\}} \mu^a.$$

We conclude this section by noting that if the process  $(V_i)_{i \in \mathbb{N}}$  is  $\eta_{(u,v)}^*$ -mixing with respect to norms  $\|\cdot\|_a$  and  $\|\cdot\|_b$  and  $N$  and  $M$  are two transition kernels on  $\mathcal{B}(\mathbb{R}^u) \times \mathbb{R}^k$  and  $\mathcal{B}(\mathbb{R}^v) \times \mathbb{R}^q$  that are  $\|\cdot\|_a$  (resp.  $\|\cdot\|_b$ )-dominated by a finite measure  $\mathcal{N}_a$  (resp.  $\mathcal{M}_b$ ), then for any  $A \in \mathcal{B}(\mathbb{R}^u)$  and  $B \in \mathcal{B}(\mathbb{R}^v)$ , for any multi-indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$  and  $r \leq j_1 - i_u$

$$\text{Cov}_{\underline{V}_{i_u}, \underline{V}_{j_v}}(N, M)(A \times B) \leq C(u, v) \eta^*(r) \mathcal{N}_a(A) \mathcal{M}_b(B).$$

### 1.2.5 Main result

Our main result emphasizes the relation between the conditional mixing properties of the process  $(X_i)_{i \in \mathbb{N}}$ , the mixing properties of the process  $(V_i)_{i \in \mathbb{N}}$  and the regularity of the transition kernels. Recall Equation (1.2): if we denote by  $N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})$  the transition kernel of the random vector  $(X_{i_1}, \dots, X_{i_u})$  given  $\underline{V}_{i_u} = (V_1, \dots, V_{i_u})$ , then the conditional independence condition implies

$$N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u}) = \prod_{k=1}^u N^{i_k}(\cdot, \underline{V}_{i_k}).$$

Consider multi indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$ , satisfying  $(\star)$ , let  $\ell$  be an integer such that  $0 < \ell \leq j_1 - i_u$ . Given  $\underline{v}_{j_v} \in \mathbb{R}^{j_v}$ , let

$$\tilde{v}_{i_u, j_v, \ell} = (0, \dots, 0, v_{i_u + \ell + 1}, \dots, v_{j_1}, \dots, v_{j_v}) \in \mathbb{R}^{j_v}$$

and denote

$$D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell}) = N^{j_1, \dots, j_v}(\cdot, \underline{v}_{j_v}) - N^{j_1, \dots, j_v}(\cdot, \tilde{v}_{i_u, j_v, \ell}).$$

where  $D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})$  is a signed measure on  $\mathcal{B}(\mathbb{R}^v)$ .

We shall make the following assumptions on  $D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})$  and  $N^{j_1, \dots, j_v}$ .

**Assumption 1.2.1.**  $K_v$  and  $K'_v$  denote positive constants depending on  $v$ ,  $(\kappa(\ell))_{\ell \in \mathbb{N}}$  denotes a decreasing to zero sequence and  $\tilde{N}^{j_1, \dots, j_v}$  is a positive finite measure on  $\mathbb{R}^v$  having total mass  $K_v$ . Let  $F_{\underline{V}}$  be the distribution function of the vector of r.v.'s  $\underline{V}$ . Let  $k = j_1 - i_u - \ell$  with  $0 < \ell \leq j_1 - i_u$ . For multi indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$ , we consider the following assumptions.

The positive kernel

$$|D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| \text{ is } \|\cdot\|_\infty\text{-dominated by } K_v \kappa(k) \|N^{j_1, \dots, j_v}\|_1. \quad (1.8)$$

The positive kernel

$$|D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| \text{ is } \|\cdot\|_\infty\text{-dominated by } K_v \kappa(k) \tilde{N}^{j_1, \dots, j_v}. \quad (1.9)$$

$$\int |D(dx_{j_1} \dots dx_{j_v}, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| dF_{\underline{V}_{j_v}}(\underline{v}_{j_v}) \leq \kappa(k) K_v. \quad (1.10)$$

$$N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v}) \text{ is } \|\cdot\|_{\infty}\text{-dominated by } K'_v \|N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v})\|_1. \quad (1.11)$$

$$N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v}) \text{ is } \|\cdot\|_{\infty}\text{-dominated by a finite measure } \mathcal{N}_{\infty}^{j_1, \dots, j_v} \quad (1.12)$$

and  $\mathcal{N}_{\infty}^{j_1, \dots, j_v}$  has total mass  $K'_v$ .

**Theorem 1.2.1.** *Assume that  $(X_i)_{i \in \mathbb{N}}$  is conditionally independent with respect to  $(V_i)_{i \in \mathbb{N}}$  and that  $(V_i)_{i \in \mathbb{N}}$  is mixing in the sense of Definition 1.2.2, we shall denote by  $C^V(u, v)$  and  $\eta^{*V}$  the constants and mixing coefficients. Depending on the kind of mixing satisfied by  $(V_i)_{i \in \mathbb{N}}$  and on the condition of Assumption 1.2.1, we obtain that the process  $(X_i)_{i \in \mathbb{N}}$  satisfies some kind of mixing properties, we will denote by  $C^X(u, v)$  and  $\eta^{*X}$  the constants and mixing coefficients. The results are presented in Table 1.1.*

Table 1.1:

Kind of mixing for $(V_i)$	Satisfied equation of Assumption 1.2.1	Result: type of mixing for $(X_i)$	Mixing coefficient for $(X_i)$
$\Psi_{(u,v)}^*$	(1.8)	$\Psi_{(u,v)}^*$	$\Psi^{*X}(r) = \kappa(\frac{r}{2}) + \Psi^{*V}(\frac{r}{2})$ , $C^X(u, v) = \max(2K_v, C^V(u, v))$
	(1.9)	$\Phi_{(u,v)}^*$	$\Phi^{*X}(r) = \kappa(\frac{r}{2}) + \Psi^{*V}(\frac{r}{2})$ , $C^X(u, v) = \max(2K_v, C^V(u, v))$
$\Phi_{(u,v)}^*$	(1.9), (1.12)	$\Phi_{(u,v)}^*$	$\Phi^{*X}(r) = \kappa(\frac{r}{2}) + \Phi^{*V}(\frac{r}{2})$ , $C^X(u, v) = \max(2K_v, K'_v C^V(u, v))$
$\alpha_{(u,v)}^*$	(1.8), (1.11)	$\Psi_{(u,v)}^*$	$\Psi^{*X}(r) = \kappa(\frac{r}{2}) + \alpha^{*V}(\frac{r}{2})$ , $C^X(u, v) = \max(2K_v, K'_v K'_u C^V(u, v))$
	(1.10), (1.11)	$\Phi_{(u,v)}^*$	$\Phi^{*X}(r) = \kappa(\frac{r}{2}) + \alpha^{*V}(\frac{r}{2})$ , $C^X(u, v) = K'_u \max(2K_v, K'_v C^V(u, v))$
	(1.10), (1.12)	$\alpha_{(u,v)}^*$	$\alpha^{*X}(r) = \kappa(\frac{r}{2}) + \alpha^{*V}(\frac{r}{2})$ , $C^X(u, v) = K'_v \max(2K_v, K'_u C^V(u, v))$ ,

As an example, one can read the first line of Table 1.1 as follows: if the process  $(V_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing and if (1.8) of Assumption 1.2.1 is satisfied then the process  $(X_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing.

It is worth pointing out that with some additional condition, the process  $(X_i)_{i \in \mathbb{N}}$  may satisfy a stronger mixing property than that satisfied by the factor process  $(V_i)_{i \in \mathbb{N}}$ . For example, if the process  $(V_i)_{i \in \mathbb{N}}$  is  $\alpha_{(u,v)}^*$ -mixing and (1.8) and (1.11) are satisfied then the process  $(X_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing. It should be noticed that the rate of mixing of  $(X_i)_{i \in \mathbb{N}}$  depends not only on the rate of mixing of  $(V_i)_{i \in \mathbb{N}}$ , but also on the sequence  $\kappa(k)$  of Assumption 1.2.1.

**Remark 1.2.4.** *Let us remark that in the proof of Theorem 1.2.1, we prove that  $(X_i)_{i \in \mathbb{N}}$  is conditionally mixing with respect to  $(V_i)_{i \in \mathbb{N}}$  in the sense of Definition 1.2.3. Then Proposition 1.2.1 gives the result.*

### 1.3 Proof of the main result

The proof of Theorem 1.2.1 is based on a decomposition of the covariance.

**Lemma 1.3.1.** *Assume that  $(V_i)_{i \in \mathbb{N}}$  is  $\eta_{(u,v)}^*$ -mixing in the sense of Definition 1.2.2 and that the transition kernels  $N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})$  (resp.  $N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v})$ ) are  $\|\cdot\|_a$ -dominated (resp.  $\|\cdot\|_b$ -dominated) by finite measures  $\mathcal{N}_a^{i_1, \dots, i_u}$  (resp.  $\mathcal{N}_b^{j_1, \dots, j_v}$ ) and  $|D(\cdot, \cdot, \cdot)|$  is  $\|\cdot\|_\infty$ -dominated by a finite measure  $\mathcal{D}$ . For any multi indices  $i_1 < \dots < i_u < j_1 < \dots < j_v$  and  $0 < \ell \leq j_1 - i_u$  and bounded functions  $f$  and  $g$ , we have*

$$\begin{aligned} & \left| \text{Cov} \left[ \mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v}) \right] \right| \\ & \leq 2 \|f\|_1 \int |g| \cdot \mathcal{D}(dx_{j_1} \dots dx_{j_v}) + \\ & \quad \eta_{(u,v)}^{*V}(\ell) \int |fg| \mathcal{N}_a^{i_1, \dots, i_u}(dx_{i_1} \dots dx_{i_u}) \cdot \mathcal{N}_b^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}). \end{aligned}$$

*Proof.* We remark that

$$\begin{aligned} & \left| \text{Cov} \left[ \mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v}) \right] \right| \\ & = \left| \text{Cov} \left[ \int f \cdot N^{i_1, \dots, i_u}(dx_{i_1} \dots dx_{i_u}, \underline{V}_{i_u}), \int g \cdot N^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}, \underline{V}_{j_v}) \right] \right| \\ & = \left| \int fg \cdot \text{Cov}_{\underline{V}_{i_u}, \underline{V}_{j_v}}(N^{i_1, \dots, i_u}, N^{j_1, \dots, j_v})(dx_{i_1}, \dots, dx_{i_u}, dx_{j_1}, \dots, dx_{j_v}) \right| \end{aligned}$$

and

$$\begin{aligned} & \left| \text{Cov}_{\underline{V}_{i_u}, \underline{V}_{j_v}}(N^{i_1, \dots, i_u}, N^{j_1, \dots, j_v}) \right| \\ & \leq \left| \text{Cov}_{\underline{V}_{i_u}, \underline{V}_{j_v}}(N^{i_1, \dots, i_u}, D) \right| + \left| \text{Cov}_{\underline{V}_{i_u}, \tilde{V}_{j_v}}(N^{i_1, \dots, i_u}, N^{j_1, \dots, j_v}) \right|. \end{aligned}$$

Now, because  $(V_i)_{i \in \mathbb{N}}$  satisfies the  $\eta_{(u,v)}^*$ -mixing property,

$$\left| \text{Cov}_{\underline{V}_{i_u}, \tilde{V}_{j_v}}(N^{i_1, \dots, i_u}, N^{j_1, \dots, j_v}) \right| \leq \eta_{(u,v)}^{*V}(\ell) \mathcal{N}_a^{i_1, \dots, i_u} \mathcal{N}_b^{j_1, \dots, j_v},$$

where  $\eta_{(u,v)}^{*V}(\ell) = C^V(u, v) \times \eta^{*V}(\ell)$ . On the other hand,

$$\int |f| \cdot N^{i_1, \dots, i_u}(dx_{i_1} \dots dx_{i_u}, dF_{\underline{V}_{i_u}}) = \|f\|_1.$$

This concludes the proof. □

**Lemma 1.3.2.** *Assume that  $(V_i)_{i \in \mathbb{N}}$  is  $\eta_{(u,v)}^*$ -mixing in the sense of Definition 1.2.2 and that the transition kernels  $N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})$  (resp.  $N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v})$ ) are  $\|\cdot\|_a$ -dominated (resp.*

$\|\cdot\|_b$ -dominated) by finite measures  $\mathcal{N}_a^{i_1, \dots, i_u}$  (resp.  $\mathcal{N}_b^{j_1, \dots, j_v}$ ). Assume also that the transition kernels  $N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})$  are  $\|\cdot\|_\infty$ -dominated by finite measures  $\mathcal{N}_\infty^{i_1, \dots, i_u}$ . For any multi indices  $i_1 < \dots < i_u < j_1 < \dots < j_v$  and  $0 < \ell \leq j_1 - i_u$  and bounded functions  $f$  and  $g$  on  $\mathbb{R}^u$  and  $\mathbb{R}^v$  respectively, we have

$$\begin{aligned}
 & | \text{Cov} [\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v})] | \\
 & \leq 2 \int |f| \mathcal{N}_\infty^{i_1, \dots, i_u}(dx_{i_1} \dots dx_{i_u}) \times \|g\|_\infty \int |D(dx_{j_1} \dots dx_{j_v}, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| dF_{\underline{V}_{j_v}} + \\
 & \quad \eta_{(u,v)}^{*V}(\ell) \int |fg| \mathcal{N}_a^{i_1, \dots, i_u}(dx_{i_1} \dots dx_{i_u}) \cdot \mathcal{N}_b^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}).
 \end{aligned}$$

*Proof.* The proof is done in the same way as the previous one.  $\square$

*Proof of Theorem 1.2.1.* We shall give the complete proof of the fact that

If  $(V_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing and (1.8) of Assumption 1.2.1 is satisfied then  $(X_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing.

The other cases are treated similarly. Note that when (1.8) or (1.9) are satisfied, we use Lemma 1.3.1 while when (1.10) is fulfilled, we use Lemma 1.3.2.

From Proposition 1.2.1, in order to prove that  $(X_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing, it suffices to prove that it is conditionally  $\Psi_{(u,v)}^*$ -mixing with respect to  $(V_i)_{i \in \mathbb{N}}$ .

Consider multi indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$ . Let  $r = j_1 - i_u$  and  $0 < \ell < r$ . From Lemma 1.3.1, (1.8) of Assumption 1.2.1 and the  $\Psi_{(u,v)}^*$ -mixing of  $(V_i)_{i \in \mathbb{N}}$ , we have

$$\begin{aligned}
 & | \text{Cov} [\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v})] | \\
 & \leq 2 \|f\|_1 \kappa(r - \ell) K_v \int g \|N^{j_1, \dots, j_v}\|_1(dx_{i_1} \dots dx_{i_u}) + \\
 & \quad \Psi_{(u,v)}^{*V}(\ell) \int fg \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \cdot \|N^{j_1, \dots, j_v}\|_1(dx_{j_1} \dots dx_{j_v}) \\
 & = 2 K_v \kappa(r - \ell) \|f\|_1 \|g\|_1 + \Psi_{(u,v)}^{*V}(\ell) \|f\|_1 \|g\|_1.
 \end{aligned}$$

where for the random variable  $V$  we denote by  $\Psi_{(u,v)}^{*V}(\ell) = \Psi^*V(\ell) \times C^V(u, v)$ . We choose  $\ell = \lfloor \frac{r}{2} \rfloor + 1$  to conclude.  $\square$

## 1.4 Examples and some applications

In this section, we provide some examples of conditionally independent random variables  $(X_n)_{n \in \mathbb{N}}$  given a factor  $(V_1, \dots, V_n)$  that satisfy conditions of Theorem 1.2.1. First of all, we remark that if the conditional mixing has bounded memory - roughly speaking if the conditional law  $X_i | \underline{V}$  is the same as  $X_i | (V_i, \dots, V_{i-\ell})$  - then we derive the mixing properties of  $(X_i)_{i \in \mathbb{N}}$  from the mixing properties of  $(V_i)_{i \in \mathbb{N}}$ ; no additional properties on the structure of the conditioning are required. Secondly, we provide examples with unbounded memory that satisfy Assumption 1.2.1.

### 1.4.1 Mixing properties for sequence controlled by a bounded memory of the factor

Note that bounded memory means here memory of order  $\ell \in \mathbb{N}$ . In classical risk theory related with conditionally independent r.v's  $(X_n)_{n \in \mathbb{N}}$  given a factor  $(V_1, \dots, V_n)$  it is very useful to consider that the conditional distribution of  $X_i$  depends only on the value of  $V_i$ . This case corresponds to  $\ell = 0$  and is considered for example in Cossette *et al.* (2004)[12]. In connection with mixing types of dependence the reader is referred to Obrein's (1974)[28] work where the factor is modulated by a Markov chain.

Next we give the following proposition satisfied for all  $\ell \geq 0$ .

Let  $(u, v)$  be a couple of integers. Write  $\underline{V}_{i,\ell} = (V_i, \dots, V_{i-\ell})$ ,  $\underline{V}_{i_1, i_u, \ell} = (V_{i_u}, \dots, V_{i_1-\ell})$  and  $\underline{V}_{j_1, j_v, \ell} = (V_{j_v}, \dots, V_{j_1-\ell})$ .

**Proposition 1.4.1.** *Let  $(X_i)_{i \in \mathbb{N}}$  be a sequence of conditionally independent random variables given  $(\underline{V}_{i,\ell})_{i \in \mathbb{N}}$ . Let  $p, q \in [1, \infty)$ , assume that  $(V_i)_{i \in \mathbb{N}}$  is  $\eta_{(u,v)}^*$ -mixing with respect to the  $L^p$  and  $L^q$  norms, with mixing coefficients denoted by  $\eta^{*V}(r)$ . Then  $(X_i)_{i \in \mathbb{N}}$  satisfies also the  $\eta_{(u,v)}^*$ -mixing condition with respect to the  $L^p$  and  $L^q$  norms, with mixing coefficients  $\eta^{*X}(r) = \eta^{*V}(r - \ell)$ .*

*Proof.* With the notation

$$\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_1, i_u, \ell}) = h_1(\underline{V}_{i_1, i_u, \ell}),$$

and

$$\mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_1, j_v, \ell}) = h_2(\underline{V}_{j_1, j_v, \ell}),$$

since  $f, g$  are two bounded real functions, we have

$$\begin{aligned} & \left| \text{Cov} \left[ \mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_1, i_u, \ell}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_1, j_v, \ell}) \right] \right| \\ &= \left| \text{Cov} \left[ h_1(\underline{V}_{i_1, i_u, \ell}), h_2(\underline{V}_{j_1, j_v, \ell}) \right] \right| \\ &\leq C^V(u, v) \eta^{*V}(r - \ell) \|h_1\|_p \|h_2\|_q \leq C^V(u, v) \eta^{*V}(r - \ell) \|f\|_p \|g\|_q, \end{aligned}$$

using Jensen's inequality. □

### 1.4.2 Mixing properties for sequences controlled by a factor with an unbounded memory

It should be noted that by unbounded memory of the factor we mean that the conditional independence is with respect to a length varying factor vector. Such a type of conditioning has not been considered in the literature to our knowledge. In addition to this fact, our study is motivated by the interest of such a type of conditioning in describing models with time dependent risks in finance and insurance. Indeed, we wish to compare the aggregate risk of an insurance portfolio of conditionally independent risks when the conditioning is done with respect to a bounded memory of the factor or with respect to unbounded memory. Hence we obtain additional information about the magnitude of risk.

Hereafter, we give examples in discrete, absolutely continuous and mixed cases.

### 1.4.2.1 Example in a discrete case

Consider the process  $(X_i)_{(i \in N)}$  such that  $X_i = I_i \times B_i$ , where

- $I_i$ 's are Bernoulli r.v.'s conditionally to  $\underline{V}_i$  and conditionally independent with respect to  $\underline{V}_i$ ,
- $B_i$ 's are considered independent and identically distributed and independent of the  $I_i$ 's and of  $\underline{V}_i$ . We shall denote by  $\mu_B$  the common law.
- $(V_i)_{(i \in N)}$  is a mixing sequence of Bernoulli random variables.

This example is inspired from insurance risk theory: if  $I_i = 1$  then there is a claim and  $B_i$  is the claim amount.

We shall consider that the conditional law of  $I_i$  has the following structure:

$$\mathbb{P}(I_i = 1 | \underline{V}_i) = \sum_{j=1}^i \frac{h(V_j)}{2^{i-j}},$$

where  $h$  is a measurable, non negative and bounded function:  $0 < \rho \leq h(v) \leq \frac{\kappa}{2}$ , with  $\kappa < 1$  to ensure that  $\mathbb{P}(I_i = 1 | \underline{V}_i) < 1$ . We have:

$$\mathbb{P}(I_i = 0 | \underline{V}_i) = \sum_{j=1}^i \frac{\xi_0^i(V_j)}{2^{i-j}}$$

with  $\xi_0^i(v) = \frac{1}{2(1 - \frac{1}{2^i})} - h(v)$ .

Denoting  $\xi_1^i(v) = h(v)$ , we may write for  $a \in \{0, 1\}$

$$\mathbb{P}(I_i = a | \underline{V}_i) = \sum_{j=1}^i \frac{\xi_a^i(V_j)}{2^{i-j}}.$$

Note that  $\xi_a^i$  is bounded,  $0 < \tau_1 \leq \xi_a^i(v) \leq \tau_2$ . The conditional transition kernel is  $N^i(\cdot, \underline{V}_i) = \delta_{\{0\}} \mathbb{P}(I_i = 0 | \underline{V}_i) + \delta_{\{1\}} \mu_B \mathbb{P}(I_i = 1 | \underline{V}_i)$ , which rewrites

$$N^i(\cdot, \underline{V}_i) = \sum_{a \in \{0,1\}} \delta_{\{a\}} \mu_B^a \mathbb{P}(I_i = a | \underline{V}_i)$$

where  $\mu_B^1 = \mu_B$  and  $\mu_B^0 = 1$  and  $\delta_{\{a\}}$  is the dirac measure.

**Lemma 1.4.1.** *The transition kernel  $N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})$  satisfies conditions (1.8) and (1.11) of Assumption 1.2.1.*

*Proof.* We have

$$\begin{aligned} N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u}) &= \prod_{k=1}^u \sum_{a_k \in \{0,1\}} \delta_{a_k} \mu_B^{a_k} \mathbb{P}(I_{i_k} = a_k | \underline{V}_{i_k}) \\ &\leq \prod_{k=1}^u \sum_{a_k \in \{0,1\}} \delta_{a_k} \mu_B^{a_k} \\ &= \sum_{a_1, \dots, a_u \in \{0,1\}^u} \delta_{a_1, \dots, a_u} \mu_B^{a_1} \cdots \mu_B^{a_u}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \|N^{i_1, \dots, i_u}\|_1 &= \int \prod_{k=1}^u \sum_{a_k \in \{0,1\}} \delta_{a_k} \mu_B^{a_k} \mathbb{P}(I_{i_k} = a_k | \underline{V}_{i_k}) dF_{\underline{V}_{i_k}} \\ &= \sum_{a_1, \dots, a_u \in \{0,1\}^u} \delta_{a_1, \dots, a_u} \mu_B^{a_1} \cdots \mu_B^{a_u} \mathbb{P}(I_{i_1} = a_1 \dots I_{i_u} = a_u). \end{aligned}$$

We remark that

$$\begin{aligned} \mathbb{P}(I_{i_1} = a_1 \dots I_{i_u} = a_u) &= \int \prod_{k=1}^u \sum_{j=1}^{i_k} \frac{\xi_{a_k}^{i_k}(V_j)}{2^{i_k-j}} dF_{\underline{V}_{i_k}} \\ &\geq \tau_1^u. \end{aligned}$$

Finally,  $N^{i_1, \dots, i_u}$  is  $\|\cdot\|_\infty$ -dominated by

$$\mathcal{N}_\infty^{i_1, \dots, i_u} := \left(\frac{1}{\tau_1}\right)^u \|N^{i_1, \dots, i_u}\|_1.$$

This proves that condition (1.11) is satisfied.

Now, we consider multi indices such that  $i_1 < \dots < i_u < i_u + \ell \leq j_1 < \dots < j_v$

$$\begin{aligned} &D(\cdot, \underline{V}_{j_v}, \tilde{V}_{i_u, j_v, \ell}) \\ &= \sum_{(a_1, \dots, a_v) \in \{0,1\}^v} \delta_{a_1, \dots, a_v} \mu_B^{a_1} \times \cdots \times \mu_B^{a_v} \times \\ &\quad \left[ \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \underline{V}_{j_v}) - \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \tilde{V}_{j_v}) \right]. \end{aligned}$$

Denoting  $\tilde{V}_{j_v} = (\tilde{V}_1, \dots, \tilde{V}_{j_v})$ ,  $\tilde{V}_j = 0$  for  $j = 1, \dots, i_u + \ell$  and  $\tilde{V}_j = V_j$  for  $j > i_u + \ell$ . We have,

$$\begin{aligned} &\left| \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \underline{V}_{j_v}) - \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \tilde{V}_{j_v}) \right| \\ &= \left| \sum_{\substack{p=1, \dots, v \\ k_p=1, \dots, j_p}} \frac{1}{2^{j_1-k_1}} \cdots \frac{1}{2^{j_v-k_v}} [\xi_{a_1}^{j_1}(V_{k_1}) \cdots \xi_{a_v}^{j_v}(V_{k_v}) - \xi_{a_1}^{j_1}(\tilde{V}_{k_1}) \cdots \xi_{a_v}^{j_v}(\tilde{V}_{k_v})] \right| \\ &\leq \binom{v}{0} 2^v \sup \left| [\xi_{a_1}^{j_1}(V_{k_1}) \cdots \xi_{a_v}^{j_v}(V_{k_v}) - \xi_{a_1}^{j_1}(\tilde{V}_{k_1}) \cdots \xi_{a_v}^{j_v}(\tilde{V}_{k_v})] \right| + \\ &\quad \binom{v}{1} \sup |\xi_a^i|^v \frac{2}{2^{j_1-i_u-\ell}} \times 2^{v-1} + \\ &\quad \binom{v}{2} \sup |\xi_a^i|^v \frac{2^2}{2^{j_1+j_2-2(i_u+\ell)}} \times 2^{v-2} + \\ &\quad \vdots \\ &\quad \binom{v}{v-1} \sup |\xi_a^i|^v \frac{2^{v-1}}{2^{\sum_{k=1}^{v-1} j_k - (v-1)(i_u+\ell)}} \times 2 + \\ &\quad \binom{v}{v} \sup |\xi_a^i|^v \sum_{\substack{k_p \leq (i_u+\ell) \\ p=1, \dots, v}} \frac{1}{2^{j_1-k_1}} \cdots \frac{1}{2^{j_v-k_v}}, \end{aligned}$$

and finally,

$$\begin{aligned} \left| \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \underline{V}_{j_v}) - \mathbb{P}(I_{j_1} = a_1 \dots I_{j_v} = a_v | \tilde{V}_{j_v}) \right| &\leq \sum_{i=1}^v \binom{v}{i} \frac{(2\tau_2)^v}{2^{i(j_1 - i_u - \ell)}} \\ &\leq \frac{(4\tau_2)^v}{2^k}, \end{aligned} \quad (1.13)$$

with  $k = j_1 - i_u - \ell$ . We have proved that  $|D|$  is  $\|\cdot\|_\infty$ -dominated by

$$\frac{1}{2^k} \left( \frac{4\tau_2}{\tau_1} \right)^v \|N^{i_1, \dots, i_u}(\cdot, \underline{V}_{i_u})\|_1,$$

thus (1.8) is satisfied.  $\square$

**Remark 1.4.1.** *If the function  $h$  is not bounded from below,  $0 < h(v) \leq \frac{\kappa}{2}$ , then the functions  $\xi_a^i$  are not bounded from below and (1.8) and (1.11) are not satisfied anymore but (1.9) and (1.12) are satisfied. So that by using Theorem 1.2.1 we derive mixing properties of the process  $(X_i)_{i \in \mathbb{N}}$ .*

#### 1.4.2.2 Example in an absolutely continuous case

This subsection is devoted to the case where the conditional law  $X_i | \underline{V}_i$  is absolutely continuous.

**Example 1.4.1.** *Consider that for all  $i \in \mathbb{N}$ , the conditional law  $X_i | \underline{V}_i$  is Pareto( $\alpha, \theta_i$ ) where  $\alpha > 2$  is the shape parameter and  $\theta_i > 0$  is the scale parameter: the conditional density of  $X_i$  has the form*

$$f_{\underline{V}_i}^i(x_i; \alpha, \theta_i) = \alpha \times \frac{\theta_i^\alpha}{x_i^{\alpha+1}} \text{ for } x_i \geq \theta_i,$$

where the scale parameter  $\theta_i$  is a random variable depending on  $\underline{V}_i$  and on  $\alpha$  with the following structure:

$$\theta_i^\alpha = \sum_{j=1}^i \frac{h(V_j)}{2^{i-j}}, \quad (1.14)$$

and where the function  $h$  satisfies  $0 < h(v) \leq \tau_2$ .

It is interesting to mention that this model has the following useful representation: let  $X_i$  be distributed as the following conditional distribution given  $\underline{V}_i$ :

$$X_i | \underline{V}_i = Z_i \times \theta_i,$$

where the parameter  $\theta_i$  is depending on  $\underline{V}_i$  and where  $Z_i$  is Pareto( $\alpha, 1$ ) distributed and is independent of  $\theta_i$  then in this case  $X_i | \underline{V}_i$  is Pareto( $\alpha, \theta_i$ ) distributed.

**Remark 1.4.2.** *Conditional Pareto models have been proposed and used in some actuarial contexts in Albrecher et al. (2011)[1] and Constantinescu et al. (2011)[10]. In these papers,  $X_i = \theta_i \times Z_i$  with  $Z_i$  as in Example 1.4.1 and  $\theta_i$  being  $V_i$ -measurable. In our framework, this means that the  $X_i$ 's are conditionally Pareto with respect to  $V_i$  i.e. the model is with bounded memory with respect to  $(V_i)_{i \in \mathbb{N}}$ . Our model is more general because it allows long term memory.*

If the process  $(X_i)_{i \in \mathbb{N}}$  is as in Example 1.4.1, the conditional density of  $(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}$  is

$$f_{x_{i_1}, \dots, x_{i_u}}(\underline{V}_{i_u}) = \frac{\alpha^u}{\prod_{k=1}^u x_{i_k}^{\alpha+1}} \prod_{k=1}^u \theta_{i_k}^\alpha \mathbf{1}_{\{x_{i_k} \geq \theta_{i_k}\}}.$$

**Proposition 1.4.2.** *If  $0 < \tau_1 \leq h(v) \leq \tau_2$  for any  $v$  then the process of Example 1.4.1 satisfies (1.10) and (1.12) of Assumption 1.2.1.*

*Proof.* The fact that (1.12) is satisfied is straightforward from the expression of  $f_{x_{i_1}, \dots, x_{i_u}}(\underline{V}_{i_u})$  and the fact that  $\tau_1 \leq h(v) \leq \tau_2$  implies:  $\tau_1 \leq (\theta_i)^\alpha \leq 2\tau_2$ .

On one other hand,

$$\begin{aligned} & \left| f_{x_{j_1}, \dots, x_{j_v}}(v_1, \dots, v_{j_v}) - f_{x_{j_1}, \dots, x_{j_v}}(0, \dots, 0, v_{i_u+\ell+1}, \dots, v_{j_v}) \right| \\ &= \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} \left| \prod_{k=1}^v \theta_{j_k}^\alpha \mathbf{1}_{\{x_{j_k} \geq \theta_{j_k}\}} - \prod_{k=1}^v \tilde{\theta}_{j_k}^\alpha \mathbf{1}_{\{x_{j_k} \geq \tilde{\theta}_{j_k}\}} \right|, \end{aligned}$$

where

$$\tilde{\theta}_i = \sum_{j=1}^i \frac{h(\tilde{V}_j)}{2^{i-j}}$$

with  $\tilde{V}_j = 0$  for  $j = 1, \dots, i_u + \ell$  and  $\tilde{V}_j = V_j$  for  $j \geq i_u + \ell + 1$ . Proceeding as in Subsection 1.4.2.1, we have that

$$\left| \prod_{p=1}^v \theta_{i_p}^\alpha - \prod_{p=1}^v \tilde{\theta}_{i_p}^\alpha \right| \leq \frac{(4\tau_2)^v}{2^p}$$

with  $p = j_1 - i_u - \ell$ . Also,

$$|\theta_i^\alpha - \tilde{\theta}_i^\alpha| \leq \frac{2\tau_2}{2^p},$$

and

$$|\theta_i - \tilde{\theta}_i| \leq \frac{2\tau_2}{2^p} \frac{1}{\alpha \tau_1^\alpha}, \quad (1.15)$$

by applying Taylor's equality. Let us denote by  $\Theta_i$  the interval  $[\min(\theta_i, \tilde{\theta}_i), \max(\theta_i, \tilde{\theta}_i)]$ . We get:

$$\begin{aligned} D(\cdot, \underline{V}_{j_v}, \tilde{V}_{i_u, j_v, \ell}) &\leq \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} \frac{(4\tau_2)^v}{2^p} \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \geq \tau_1^{\frac{1}{\alpha}}\}} \\ &\quad + \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} (2\tau_2)^v \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \in \Theta_{j_k}\}}. \end{aligned}$$

So, we conclude that (1.10) is fulfilled by using that the Lebesgue measure of  $\Theta_i$  is  $|\theta_i - \tilde{\theta}_i|$  and using (A.6) and Fubini's Theorem.  $\square$

**Remark 1.4.3.** Note that the case of an Exponential conditional law (instead of a Pareto conditional law) could also be considered. In particular if we consider an Exponential law with parameter  $(1/\theta_i)$  where  $\theta_i$  is defined as in (1.14) with  $\alpha = 1$  and  $0 < \tau_1 \leq h(v) \leq \tau_2$  for any  $v$ , then (1.10) and (1.12) of Assumption 1.2.1 are satisfied.

### 1.4.2.3 Models for applications of the absolutely continuous case

It is interesting to explain further how our conditionally independent model may have applications in insurance or financial contexts.

Life insurance companies as well as financial institutions use hedging strategies with different portfolio adjustment frequencies: hedging is more frequent for banks, often weekly for life insurance companies offering variable annuities, and more monthly or quarterly for classical life insurance companies, even if some financial lines of business of some insurance companies correspond to financial institutions. All those firms face transaction costs, which prove to be very expensive when things go wrong, both due to liquidity issues and due to other conditions like negative gamma situations. Optimal hedging strategies in presence of liquidity risk and transaction costs have been studied by various authors (see e.g. Kabanov (2009) and references therein). Here, we do not aim at studying those optimal strategies; if we rather adopt a P& C loss model approach, we would try to study the impact of correlations between elementary transaction costs on the aggregated transaction cost. Correlations might come from factors like liquidity and market conditions that vary over time but present some memory. We consider that the aggregated transaction cost is the sum of infinitely many elementary, correlated transaction costs. Assume that  $(Z_i)_{i \geq 1}$  is a sequence of i.i.d. random variables representing random transaction costs in a steady environment (without liquidity shocks or bad market conditions), and that the  $V_j, j \geq 1$  are random variables which contribute to create liquidity shocks and market conditions during some time when they take large values. Consider for each  $i \geq 1$ ,

$$X_i = \theta_i \cdot Z_i,$$

where

$$\theta_i = \left( \sum_{j=1}^i \frac{1}{2^{i-j}} h(V_j) \right)^{1/\alpha},$$

corresponds to a multiplicative factor that takes into account liquidity and market conditions at time  $i$ , the function  $h$  is as before,  $0 < \tau_1 \leq h(v) \leq \tau_2$ . The  $Z_i, i \geq 1$  are independent from the  $V_j, j \geq 1$ . If  $Z_i$  is *Pareto* $(\alpha, 1)$  distributed then  $X_i$  is *Pareto* $(\alpha, \theta_i)$  distributed and the  $X_i$ 's are as in Example 1.4.1.

If the variables  $V_i$  are bounded and

$$\theta_i = \left( \sum_{j=1}^i \frac{1}{2^{i-j}} (1 + V_j) \right)^{1/\alpha},$$

then,

$$\theta_{i+1}^\alpha = \frac{1}{2} \theta_i^\alpha + 1 + V_{i+1},$$

which enables one to understand the correlation structure between  $\theta_i$  and  $\theta_{i+1}$ , as  $V_{i+1}$  is independent from  $\theta_i$ .

As in the previous sections, in that case also, (1.10) and (1.12) of Assumption 1.2.1 are satisfied. The use of Theorem 1.2.1 then leads to mixing properties of  $(X_i)_{i \in \mathbb{N}}$ .

#### 1.4.2.4 Mixed cases

Other examples that have interest in insurance applications are so called mixed cases where the claim amount  $X_i$  has the following form

- $X_i = I_i \times B_i$ ,
- $I_i$ 's are Bernoulli r.v.'s conditionally to  $\underline{V}_i$  and conditionally independent with respect to  $\underline{V}_i$ ,
- $B_i$ 's are conditionally independent and conditionally independent of the  $I_i$ 's.

If the  $I_i$ 's are as in Subsection 1.4.2.1 and the  $B_i$ 's are as in Example 1.4.1, then following the ideas of the previous examples leads to prove that Conditions (1.10) and (1.12) of Assumption 1.2.1 are satisfied.

## Concluding remarks

It is well known in the literature that processes satisfying some mixing properties have interesting statistical properties (see [13; 15; 21; 27] and many others) such as moment inequalities and central limit theorem. The mixing properties satisfied by our models, namely  $\eta_{(u,v)}^*$ -mixing properties, are not standard ones (mainly because the constants  $C(u, v)$  depend on  $u$  and  $v$ ) and some more work is needed to get these statistical properties. This is the aim of another work ([24]), in which we prove a self normalized central limit theorem and study further applications to insurance risk processes. We note that in order to have limit theorem of this kind, some stationarity on the  $(X_i)_{i \in \mathbb{N}}$  is required. We may prove that if there exists a limiting kernel then the process  $(X_i)_{i \in \mathbb{N}}$  converges to a stationary process, roughly speaking, it is required that there exists a transition kernel  $N(\cdot, \underline{V}_\infty)$  on  $\mathcal{B}(\mathbb{R}) \times \mathbb{R}^{\mathbb{N}}$  such that,

$$N^i(\cdot, \underline{V}_i) \xrightarrow{i \rightarrow \infty} N(\cdot, \underline{V}_\infty).$$

The examples considered in Subsections 1.4.2.1 and 1.4.2.2 satisfy this condition and thus have asymptotic stationary laws. It would also be interesting to study the impact of this kind of dependence on ruin probabilities and other risk and profit indicators. This is left for further research.

## Bibliography

- [1] Albrecher, H., Constantinescu, C. and Loisel, S.: Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics*. **2**, (2011), 265–270.
- [2] Asmussen, S.: Risk theory in a markovian environment. *Scandinavian Actuarial Journal*. **2**, (1989), 69–100.
- [3] Bernstein, S.N.: Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.* (1927), 1–59.

- 
- [4] Billingsley, P.: Convergence of probability measures. *Wiley*. (1968).
- [5] Blum, J. R., Hanson, D.L. and Koopmans., L. H.: On the strong law of large number for a class of stochastic processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*. **2**, (1963), 1–11.
- [6] Bradley, R.C. and Bryc, W.: Multilinear forms and measures of dependence between random variables. *Journal of Multivariate Analysis*. **16**, (1985), 335–367.
- [7] Bradley, R.C.: Basic properties of strong mixing conditions. Dependence in probability and statistics. *Progr. Probab. Statist.* **11**, (1986), 165–192.
- [8] Bradley, R.C.: Basic properties of strong mixing conditions. A survey and some open questions. Update of, and a supplement to, the 1986 original. *Probability Surveys*. **2**,(2005), 107–144.
- [9] Bradley, R.C.: Introduction to Strong Mixing Conditions. *Kendrick Press, Heber City, Utah*. **1,2,3**, (2007).
- [10] Constantinescu, C., Hashorva, E. and Ji L.: Archimedean copulas in finite and infinite dimensions - with application to ruin problems. *Insurance: Mathematics & Economics*. **49**,(2011), 487–495.
- [11] Cossette, H., Gaillardetz, P. and Marceau, E.: Common mixture in the individual risk model. *Mitteilungen der Schweiz. Aktuarvereinigung*. **2**, (2002), 131–157.
- [12] Cossette, H., Landriault, D. and Marceau, E.: Compound binomial risk model in a markovian environment. *Insurance: Mathematics & Economics*. **35** (2004), 425–443.
- [13] Dedecker, J., Doukhan, P., Lang, G., León R, J. R., Louhichi, S. and Prieur, C.: Weak Dependence: With Examples and Applications. *Lecture Notes in Statistics*, **190**. Springer, New York. (2007).
- [14] Doukhan, P.: Mixing properties and examples. *Lecture note in statistics*. **85**. (1994).
- [15] Doukhan, P. and Louhichi S.: A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*. **84**, (1999), 313–342.
- [16] Dziubdziela, W.: Limit laws for  $k$ th order statistics from conditionally mixing arrays of random variables. *Journal of Applied Probability*. **23**, (1986), 679–687.
- [17] Gerber, H. U.: Mathematical fun with the compound binomial process. *ASTIN Bull.* **18**, (1988), 161–168.
- [18] Gordin, M. I.: On the central limit theorem for stationary random processes. *Dokl. Akad. Nauk. SSSR*.(1969), 188, 739–741.
- [19] Gordin, M. I.: A remark on the martingale method for proving the central limit theorem for stationary sequences. *Journal of Mathematical Sciences*. **133**, (2004).
- [20] Hall, P. and Heyde, C. C.: Martingale limit theory and its application. *Probability and Mathematical Statistics*. Academic Press, Inc, New York-London. (1980).

- [21] Ibragimov, I.A.: Some limit theorems for stationary processes. *Theory of Probability and its Applications*. **7**, (1962), 349–382.
- [22] Ibragimov, I.A. and Linnik, Y. U.: Independent and stationary sequences of random variables. *Wolters-Noordhoff Publishing, Groningen*. (1971), 443.
- [23] Kabanov, Y. M.: Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics*. **2**, (2009), 237–248.
- [24] Kacem, M., Maume-Deschamps V.: Self Normalised Central Limit Theorem for some mixing processes. Work in progress.
- [25] Kolmogorov, A. N. and Rozanov, Ju. A.: On a strong mixing condition for stationary Gaussian processes. *Theory Of Probability and Its Applications*. **V**, (1960), 204–208.
- [26] Mcleish, D. L.: Invariance principle for dependent variables. *Z. Wahrscheinlichkeits-theorie und Verw. Gebiete* **32**,(1975) 165–178.
- [27] Merlevède, F., Peligrad, M. and Utev, S.: Recent advances in invariance principles for stationary sequences. *Probability Surveys*. **3**, (2006), 1–36.
- [28] Obrein, G.L.: Limit theorems for sums of chain-dependent processes. *Journal of applied probability*. **11**, (1974), 582–587.
- [29] Peligrad, M.: Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. *Dependence in probability and statistics (Oberwolfach,1985)*. *Progress Probab. Statist.* (1986), **11**, 193–223.
- [30] Philipp, W. and Stout, W.F.: Almost sure invariance principles for partial sums of weakly dependent random variables. *Memoirs of the American Mathematical Society*. **2**, (1975), iv+140.
- [31] Rao, P.: Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*. **61**, (2009), 441–460.
- [32] Revuz, D. *Markov chains*. Second edition. North-Holland Mathematical Library, 11. North-Holland Publishing Co., Amsterdam, (1984).
- [33] Rosenblatt, M.: A central limit theorem and a strong mixing condition. *Proceeding national academy of Sciences of the United States of America*. (1956), **42**, 427–433.
- [34] Veijanen, A.: On conditionally mixing processes. *Statistics & Probability Letters*. **9**, (1990), 59–65.

## Chapitre 2

# Estimation de la variance pour des processus mélangeants

## Résumé

Ce chapitre concerne les suites de variables aléatoires qui sont  $\eta_{(u,v)}^*$  mélangeantes (voir Chapitre 1). Notre but est de fournir un théorème central limite auto-normalisé adapté au type de mélange considéré. Rappelons que ce type de mélange est vérifié par des processus de risque explicités au Chapitre 1. Nous obtenons un estimateur de la variance limite qui admet un théorème centrale limite. Nous établissons également des inégalités de covariance type Davydov et des inégalités de moment type Rosenthal qui seront utilisées dans la preuve de nos théorèmes limites. Notre travail est illustré par quelques applications à des processus de risque en assurance. Ce chapitre est organisé de la manière suivante. Après l'introduction nous présentons dans la Section 2.2, les outils nécessaires dans les preuves de nos principaux résultats. Dans la Section 2.3 nous énonçons nos principaux résultats qui sont principalement les Théorèmes 2.3.1, 2.3.3 et 2.3.2. Par la suite, dans la Section 2.4 nous présentons les preuves de nos principaux résultats. Enfin, dans la dernière section, nous présentons quelques exemples d'applications. Ce chapitre est un article en cours intitulé « Self Normalised Central Limit Theorem for some mixing processes » co-écrit avec Véronique Maume-Deschamps.

## 2.1 Introduction

It is well known that asymptotic theorems allow to obtain approximations of various laws. One of these theorem, which is fundamental in probability theory, is the central limit theorem (C.L.T.). In risk theory, the C.L.T. achieves, with a certain level of confidence, quantitative information on the capital which is required to ensure solvency. In fact, the C.L.T. leads to control the difference between the aggregate claim amounts and the pure premium. It also indicates that the dispersion around the pure premium is distributed according to a gaussian law and that the standard-deviation of this law is taken into account to estimate the safety loading premium.

In a previous work [24] we have considered processes of dependent random variables (r.v's) that are conditionally independent (C.I.) given a factor. In particular, we have assumed that the conditioning is with respect to an unbounded memory of the factor. These processes are of interest in risk theory. Our main result consisted on deriving some mixing properties for such processes. This is the starting point of this work on which we investigate the asymptotic normality of the aggregate claim amount when time goes to infinity. In fact, mixing processes have very interesting asymptotic properties especially in the stationary case.

Note that the variance is a measure of risk which find its aim argument on the normality of the distribution. In fact, when a r.v. is normally distributed it is well characterized by its mean and variance. In practice, the variance of the aggregate sum of random variables is not known. If we want a C.L.T. to be applicable then the variance must be estimated.

Note that when the real variance in a C.L.T. is replaced by an estimate from the given data then the estimator is called a self-normalizer and we say that a self normalized central limit theorem (S.N.C.L.T.) holds. While for independent and identically distributed r.v's it is very simple to estimate the variance, this is not the case for dependent r.v's for which we should construct an estimator which takes into account the dependence structure. The aim of this work is to provide a self-normalizer which is adapted to our kind of mixing and whence a (S.N.C.L.T.) for mixing processes. Our result are useful for C.I. processes under some stationarity conditions.

We restate here the following definition of mixing provided in [24] which makes explicit the asymptotic independence between functions of the past and of the future of a sequence of r.v's. The notion of  $\eta_{(u,v)}^*$ -mixing is related to  $\eta_{(\infty,\infty)}^*$ -mixing, but is a weaker condition as it measures the dependence between  $(u+v)$  random variables and not the entire tails.

**Definition 2.1.1.** *Let  $u, v$  be integers, a sequence  $(X_n)_{n \in \mathbb{N}}$  of r.v's is said to be  $\eta_{(u,v)}^*$ -mixing, if there exist a function  $r \mapsto \eta^*(r)$  decreasing to 0 as  $r$  goes to infinity and a constant  $C(u, v) > 0$  such that for any real valued bounded functions  $f$  and  $g$ , for any integers  $u, v$  and for any multi-indices satisfying the relation  $(\star)$ :*

$$i_1 < \dots < i_u < i_u + r \leq j_1 < \dots < j_v, \quad (\star)$$

we have

$$\sup |Cov(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq C(u, v) \eta^*(r) \|f\|_a \|g\|_b, \quad (2.1)$$

where the supremum is taken over all the sequences  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$  and  $r \leq j_1 - i_u$  is the gap of time between past and future.

In the previous definition  $\| \cdot \|_a$  and  $\| \cdot \|_b$  are norms on bounded functions (or on subspaces of bounded functions). With respect to these norms, we have various kind of mixing.

1. If  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_\infty$ , we shall say that the process is  $\alpha_{(u,v)}^*$ -mixing and we shall write  $\alpha^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\alpha_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\alpha$ -mixing in the sense of Rosenblatt (1956)[40] and we shall write  $\alpha(r)$  instead of  $\eta^*(r)$ .
2. If  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_2$ , we shall say that the process is  $\rho_{(u,v)}^*$ -mixing and we shall write  $\rho^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\rho_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\rho$ -mixing in the sense of Kolmogorov and Rozanov (1960)[25] and we shall write  $\rho(r)$  instead of  $\eta^*(r)$ .
3. If  $\| \cdot \|_a = \| \cdot \|_1$  and  $\| \cdot \|_b = \| \cdot \|_\infty$ , we shall say that the process is  $\Phi_{(u,v)}^*$ -mixing and we shall write  $\Phi^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\Phi_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\Phi$ -mixing in the sense of Ibragimov (1962)[21] and we shall write  $\Phi(r)$  instead of  $\eta^*(r)$ .
4. If  $\| \cdot \|_a = \| \cdot \|_b = \| \cdot \|_1$ , we shall say that the process is  $\Psi_{(u,v)}^*$ -mixing and we shall write  $\Psi^*(r)$  instead of  $\eta^*(r)$ . If the process is  $\Psi_{(u,v)}^*$ -mixing for all  $u, v$  and if  $\sup_{u,v \in \mathbb{N}} C(u, v) \leq C < \infty$  then the process is  $\Psi$ -mixing in the sense of Blum *et al.* (1963)[2] and we shall write  $\Psi(r)$  instead of  $\eta^*(r)$ .

**Remark 2.1.1.** Let  $h : \mathbb{R}^u \rightarrow \mathbb{R}$ , be an arbitrary function and denote his lipschitz modulus by  $Lip(h) = \sup_{x \neq y} |h(x) - h(y)| / \|x - y\|_1$ . If

$$\sup |Cov(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq C(u, v)\eta^*(r)Lip(f)Lip(g), \quad (2.2)$$

where  $C(u, v) = (uLip(f) + vLip(g) + uvLip(f)Lip(g))$  then the process is said  $\lambda$ -weakly dependent. For more details on this type of mixing and on its applications we refer to Dedecker *et al.* (2007)[10].

For the examples of processes that we have in mind the coefficient  $C(u, v)$  are of exponential kind. This is the main motivation for Definition 2.1.1.

Basic property of mixing condition can be found in Bradley (1986),(2005) ([3],[4]). In the case of classical mixing sequences (see i.e. [10; 12; 2; 25; 40; 21; 22]), mixing coefficients are uniform on  $u$  and  $v$ . Clearly this is not the case for mixing sequences satisfying our mixing condition (2.1).

Many authors are interested by investigating the classical central limit theorem (C.L.T.) for the sum  $S_n = \sum_{i=1}^n X_i$  where  $X_i$  are r.v.'s satisfying mixing conditions. This was studied in [10; 2; 25; 40; 21].

For  $\alpha$ -mixing sequences, Shi (2000) [41] took over the class of estimator introduced in Peligrad and Shao (1995)[33] for  $\rho$ -mixing sequences and showed that this class is available

for strongly mixing sequences of r.v.'s. For stationary sequence of associated r.v.'s we refer to Peligrad and Suresh (1995) [34] in which the self-normalizer arises from the classical Bernstein-Block-Technique (see Bernstein [1]).

Doukhan *et al.* [14] dealt with random vectors of  $\lambda$ -weakly dependent sequences. They use nonoverlapping blocks denoted by

$$\Delta_{i,m} = \frac{1}{\sqrt{m}} \sum_{j=(i-1)(m+\ell)+1}^{(i-1)(m+\ell)+m} X_j,$$

in the construction of their estimator. The fact that blocks are nonoverlapping leads to choose two parameter  $m$  and  $\ell$  while for Peligrad and Shao's estimator the blocks denoted by

$$\tilde{\Delta}_{i,m} = \frac{1}{\sqrt{m}} \sum_{j=i+1}^{i+m} X_j,$$

are overlapping and only the parameter  $m$  must be chosen.

Choosing nonoverlapping blocks enables Doukhan *et al.* [14] to provide not only an estimator of the variance but also an estimator of the variance of the estimator. In our examples  $C(u, v)$  is an exponential function of  $u$  and  $v$  where the mixing coefficient for  $\lambda$ -weakly dependent r.v.'s is a polynomial function of  $u$  and  $v$  and of lipschitz functions (see Remark 2.1.1). Hence this estimator is not adapted with examples that we consider.

In our work we provide a self-normalizer which is a modified version of the one obtained in Peligrad and Shao (1995)[33]. As in Shi (2000) [41] we adapt this estimator to our kind of mixing. We investigate its weak consistency as well as its asymptotic normality. In this frame we derive useful tools such as covariance inequality of Davydov's kind and moment inequalities of Rosenthal's kind which are adapted with our structure of mixing.

This work is organized as follows. In Section 2.2 we give our main preliminary results: moment inequalities, covariance inequality and uniform integrability property. These are the main tools for the proofs of our results. In Section 2.3 we derive a C.L.T. for mixing processes which is adapted with our type of mixing and we provide an estimator of the variance. In particular we derive the weak consistency and the asymptotic normality of our estimator. Section 2.4 is devoted to the proofs and the last section is devoted to present some examples of risk models for which our results are applicable.

## 2.2 Main tools: moment inequalities, covariance inequality and uniform integrability

### 2.2.1 Covariance inequality

In the following proposition we provide a covariance inequality which is adapted with our structure of mixing. This inequality is a modified version of the one provided in Theorem 3 in Doukhan (1994)[12] which is itself due to Davydov's (1968)[8] inequality and is extensively used in proofs of limit theorems.

**Proposition 2.2.1.** *Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(u,v)}^*$ -mixing sequence. Let  $\|f\|_p < \infty$  and  $\|g\|_q < \infty$ , where  $1 < p, q \leq \infty$  and  $\frac{1}{p} + \frac{1}{q} < 1$ , then*

$$|\text{Cov}(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq (C(u, v) + 8)(\mathbb{E}|f|^p)^{\frac{1}{p}}(\mathbb{E}|g|^q)^{\frac{1}{q}}\alpha^*(r)^{1 - \frac{1}{p} - \frac{1}{q}}, \quad (2.3)$$

where  $j_v - i_u = r$  where  $r$  is a positive integer.

**Remark 2.2.1.** *Note that if  $(X_n)_{n \in \mathbb{N}}$  is a sequence of complex valued r.v's then by separating the real and imaginary parts of the complex number, we obtain the inequality (2.3) with  $4 * (C(u, v) + 8)$  instead of  $(C(u, v) + 8)$ .*

Note that Rio [38; 39] provided covariance inequality based on marginal laws of r.v's.

### 2.2.2 Moment inequalities

Moment inequalities of partial sums play a very important role in various proofs of limit theorems.

For independent r.v's we have the following Rosenthal type inequality

$$\mathbb{E}|X_1 + \dots + X_n|^q \leq C_q \left( \sum_{i=1}^n \mathbb{E}|X_i|^q + \left( \sum_{i=1}^n \mathbb{E}|X_i|^2 \right)^{q/2} \right), \quad (2.4)$$

and the Marcinkiewicz-Zygmund inequality

$$\mathbb{E}|X_1 + \dots + X_n|^q = O(n^{\frac{q}{2}}). \quad (2.5)$$

For dependent r.v's one refers to Yokoyama (1980) [43] who first obtained the bound (2.5) for strictly stationary strongly mixing sequences when  $q > 2$ . For strongly mixing sequences, Rio (1994) [39] obtained this bound under assumptions which are given in terms of quantile function of the sequence  $(X_n)_{n \in \mathbb{N}}$ . Doukhan and Louhichi (1999) [13] proposed a more precise bound (see the Appendix, equation (2.99)) for (2.5) for which Rio's (1994) [39] assumptions are reached. Also we refer to Kim (1994)[26] who provided a unified approach for establishing moment bounds of order  $q \geq 2$  for sums of mixing sequences satisfying a stationary condition. Note that these moment inequalities for strongly mixing sequences are not adequate with our proofs related to empirical processes.

Shao and Yu (1996)[19] provided moment inequalities of order  $p > 2$  for empirical processes satisfying strongly mixing conditions in the sense of Rosenblatt (1956)[40]. Their proof, is based on martingale difference argument and on Bernstein decomposition. Doukhan and Louhichi (1999) [13] provided moment inequalities for weakly dependent r.v's and for associated sequences which can be easily derived under our mixing condition. Their results is extended by Doukhan *et al.* [14] for  $\lambda$ -weakly dependent sequences.

The aim of this section is to provide Rosenthal type inequalities for mixing sequence satisfying  $\alpha_{(u,v)}^*$ -mixing condition given by inequality (2.1). The main results are stated in Lemmas 2.2.1 and 2.2.2. Recall the definition of the the coefficients of weak dependence introduced in Doukhan and Louhichi (1999) [13].

**Definition 2.2.1.** For positive integer  $r$ , define the coefficients of weak dependence as the non-decreasing sequence  $(C_{r,q})_{(q \geq 2)}$  such that

$$C_{r,q} := \sup |Cov(X_{t_1} \times \dots \times X_{t_m}, X_{t_{m+1}} \times \dots \times X_{t_q})|,$$

where the supremum is taken over the multi-indices  $1 \leq t_1 \leq \dots \leq t_q$  with  $t_{m+1} - t_m = r$ .

We give the following lemma for bounded r.v's satisfying  $\alpha_{(u,v)}^*$ -mixing condition.

**Lemma 2.2.1.** Consider uniformly bounded r.v's  $X_i$  such that  $\sup_{i \in \mathbb{N}} |X_i| \leq M$ . In addition assume that  $(X_n)_{n \in \mathbb{N}}$  satisfies  $\alpha_{(m,q-m)}^*$ -mixing where  $1 < m < q$  then for positive integer  $r$ , we have

$$C_{r,q} \leq M^q C(m, q - m) \alpha^*(r).$$

For non bounded r.v's we derive the following lemma which follows from Proposition 2.2.1 and Lemma 2.6.1 in Dedecker *et al.* (2007)[10] which is proved in the Appendix for completeness.

**Lemma 2.2.2.** Assume that the sequence of r.v's  $(X_n)_{n \in \mathbb{N}}$  is centered and let  $S_n = \sum_{i=1}^n X_i$ . If the sequence  $(X_n)_{n \in \mathbb{N}}$  is  $\alpha_{(1,1)}^*$ -mixing and for some  $\delta > 0$ ,  $\sup_i \|X_i\|_{2+\delta} < \infty$  then

$$\mathbb{E}(S_n^2) \leq 2nK_2 \sum_{r=0}^{n-1} \alpha^*(r)^{\frac{\delta}{\delta+2}}. \quad (2.6)$$

If the sequence  $(X_n)_{n \in \mathbb{N}}$  is  $\alpha_{(3,3)}^*$ -mixing such that for some  $\delta > 0$ ,  $\sup_i \|X_i\|_{4+\delta} < \infty$  then

$$\mathbb{E}(S_n^4) \leq 4!n^2K_2^2 \left( \sum_{r=0}^{n-1} \alpha^*(r)^{\frac{\delta}{\delta+2}} \right)^2 + nK_4 \sum_{r=0}^{n-1} (r+1)^2 \alpha^*(r)^{\frac{\delta}{\delta+2}}, \quad (2.7)$$

where we denote by

$$K_2 = (C(1, 1) + 8) \sup_i \|X_i\|_{2+\delta}^2, \quad (2.8)$$

and by

$$K_4 = (C(3, 3) + 8) \sup_i \|X_i\|_{4+2\delta}^4. \quad (2.9)$$

In our proof of the limit theorem provided in Section 2.3 we need an estimation of the moments of the sum of order  $p$  greater than 4 where  $p$  is not necessary an integer. For this aim we give the following theorem which is a modified version of Theorem 4.1 provided in Shao and Yu (1996) [19]. We define

$$D_{n,2,2} = \sum_{r=0}^{n-1} \alpha^*(r)^{\frac{\delta}{\delta+2}}. \quad (2.10)$$

**Theorem 2.2.1.** Let  $2 < p < r < \infty$ ,  $2 < v \leq r$  and  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(m,2)}^*$ -mixing sequence of r.v's where  $1 < m < [\frac{n}{2}] + 1$ . Let  $\mathbb{E}X_n = 0$  and  $\|X_n\|_r = \mathbb{E}(|X_n|^r)^{1/r} < \infty$ . Assume that for some  $C > 0$  and  $\theta > 0$

$$\alpha^*(n) \leq Cn^{-\theta}. \quad (2.11)$$

Then, for any  $\epsilon > 0$  there exists  $A = A(\epsilon, r, p, \theta, C) < \infty$  such that

$$E|S_n|^p \leq A \left( n^{p/2} (K_2 D_{n,2,2})^{p/2} + n^{(p-(r-p)\theta/r)\vee(1+\epsilon)} \max_{i \leq n} \|X_i\|_r^p \right). \quad (2.12)$$

In particular, for any  $\epsilon > 0$ , if  $\theta \geq (p-1)r/(r-p)$  then

$$E|S_n|^p \leq A \left( n^{p/2} (K_2 D_{n,2,2})^{p/2} + n^{(1+\epsilon)} \max_{i \leq n} \|X_i\|_r^p \right), \quad (2.13)$$

and if  $\theta \geq \frac{pr}{2(r-p)}$  then

$$E|S_n|^p \leq A n^{p/2} \left( (K_2 D_{n,2,2})^{p/2} + \max_{i \leq n} \|X_i\|_r^p \right). \quad (2.14)$$

### 2.2.3 Uniform integrability

The concept of uniform integrability (U.I.) is essential in the proof of some limit theorems. Recall that a family of r.v.'s  $(X_n)_{n \in \mathbb{N}}$  is uniformly integrable if  $\lim_{A \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{E}(|X_n| 1_{|X_n| > A}) = 0$ .

Let us mention that if  $(X_n)_{n \in \mathbb{N}}$  is dominated by  $Y$  in  $L^1$  (i.e.  $|X_n| < Y$  a.e. and  $\mathbb{E}(|Y|) < +\infty$ ) then  $(X_n)_{n \in \mathbb{N}}$  is uniformly integrable. Lemma 2.2.2 leads to the following corollary which gives sufficient conditions to have U.I. condition of order  $p$  where  $p \in \{2, 4\}$  under mixing condition.

**Corollary 2.2.1.** *Choose  $p$  an integer in  $\{2, 4\}$ . Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(p-1, p-1)}^*$  mixing sequence. Without loss of generality consider that  $\mathbb{E}(X_i) = 0$  and assume that  $\mathbb{E}(|X_i|^{p+\delta}) < \infty$ . For each integer  $s$  in  $\{2, p\}$ , we define*

$$D_{n,s,p} = \left( \sum_{r=0}^{n-1} (1+r)^{p-s} \alpha^*(r)^{\frac{\delta}{\delta+s}} \right)^{\left(\frac{p-s}{2}+1\right)}. \quad (2.15)$$

If in addition, for each  $s$  in  $\{2, p\}$ ,

$$D_{\infty,2,s} < \infty,$$

then

$$\left\{ \left| \frac{S_n}{\sqrt{n}} \right|^p, n \geq 1 \right\},$$

is uniformly integrable.

Corollary 2.2.1 implies that if  $\mathbb{E}(|X_i|^{2+\delta}) < \infty$  and if  $D_{\infty,2,2} < \infty$  then  $\left\{ \left| \frac{S_n}{\sqrt{n}} \right|^2, n \geq 1 \right\}$  is uniformly integrable.

Also it shows that if  $\mathbb{E}(|X_i|^{4+\delta}) < \infty$ ,  $D_{\infty,2,2} < \infty$  and  $D_{\infty,2,4} < \infty$  then  $\left\{ \left| \frac{S_n}{\sqrt{n}} \right|^4, n \geq 1 \right\}$  is uniformly integrable.

**Remark 2.2.2.** *Note that  $D_{\infty,2,2} < \infty$  may be replaced by  $\alpha^*(r) \leq Cr^{-\frac{\delta+2}{\delta}}$  where  $C$  is a constant.*

**Remark 2.2.3.** Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(m,2)}^*$ -mixing sequence of random variables, where  $1 < m < [\frac{n}{2}] + 1$ . Assume that for some,  $C > 0$ ,  $\delta > 0$ ,  $p > 2$ ,  $\theta > \frac{p(p+\delta)}{2\delta}$ ,

$$\alpha^*(n) \leq Cn^{-\theta}, \quad E|X_1|^{p+\delta} < \infty,$$

thus immediately from Theorem 2.2.1 we have that

$$\left\{ \left| \frac{S_n - \mathbb{E}(S_n)}{\sqrt{n}} \right|^p, n \geq 1 \right\},$$

is uniformly integrable.

## 2.3 Self Normalized Central Limit Theorem for stationary $\alpha$ mixing sequences

In the literature, to prove limit theorems for strongly mixing sequences of r.v's, one can approximate the strongly mixing sequence by another sequences satisfying certain properties. Gordin (1969)[16] considers a direct approximation of mixing sequence by a sequence of martingale differences (for more details we refer to Dalibor Volný (1988,1993)[6; 7], Hall and Heyde (1980)[18], Gordin (2004)[17], and (1975)[36]). Another method known by Bernstein's method (1927)[1] is to consider a direct approximation of mixing sequences by a sequence of independent r.v's This is the method used in the proof of the central limit theorem derived in Ibragimov and Linnik (1971)[23].

Invariance principle has been extensively investigated for weakly dependent and mixing sequences (see i.e. [30; 27; 32; 36; 15; 8; 9; 13; 31]). In particular McLeish (1986)[28] derived a weak invariance principle for dependent r.v's, Herrndrof (1985)[20] extended this previous result and derived a general invariance principle for  $\alpha$ -mixing sequence of r.v's with flexible moment conditions. We refer to [5; 23; 10; 12] for a survey on C.L.T. for mixing sequences.

**Remark 2.3.1.** Assume that  $(X_n)_{n \in \mathbb{N}}$  is a stationary centered sequence. If  $\sum_{j=1}^{\infty} \mathbb{E}(X_1 X_j) < \infty$  then

$$\sigma^2 = \mathbb{E}(X_1^2) + 2 \sum_{j=1}^{\infty} \mathbb{E}(X_1 X_j) < \infty, \quad (2.16)$$

and

$$\sigma_n^2 = \mathbb{E}\left(\sum_{j=1}^n X_j\right)^2 = n\mathbb{E}(X_1^2) + 2 \sum_{j=1}^n (n-j) \text{Cov}(X_1, X_j) = n\sigma^2(1 + o(1)). \quad (2.17)$$

If in addition,  $\sigma > 0$  then  $\sigma_n^2 \rightarrow \infty$  as  $n \rightarrow \infty$ .

The aim of this section is to provide a self normalized central limit theorem for  $\alpha_{(u,v)}^*$ -mixing sequence of r.v's. First of all we give the following theorem for  $\alpha_{(u,v)}^*$ -mixing process for all integers  $(u, v) \in \mathbb{N}^*$ . This theorem is a modified version of the classical C.L.T. derived in Ibragimov (1962)[21] for strongly mixing sequences.

**Theorem 2.3.1.** Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(u,v)}^*$ -mixing process for all  $(u, v) \in \mathbb{N}^*$ . Let  $\mathbb{E}(X_1) = \mu$  and

$$\mathbb{E}|X_1|^{2+\delta} < \infty, \text{ for some } \delta > 0, \quad (2.18)$$

be satisfied. If

$$\sum_{r=1}^{\infty} \alpha^*(r)^{\frac{\delta}{2+\delta}} < \infty. \quad (2.19)$$

holds, then (2.16) is true. If in addition there exist  $M > 0$  and  $K > 0$  such that for any  $u, v$ , and  $r$ ,

$$\begin{cases} C(u, v) \leq K^{u+v}, \\ \alpha^*(r) \leq \frac{M}{r^\delta}, \quad \delta > \max\{2 \log(K), 2\}, \end{cases} \quad (2.20)$$

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0, 1). \quad (2.21)$$

**Remark 2.3.2.** Note that if  $C(u, v) \leq K(u+v)^\beta$  (i.e.  $C(u, v)$  is a polynomial function of  $u$  and  $v$ ) then the Theorem 2.3.1 holds if we replace (2.20) by the following condition

$$\begin{cases} C(u, v) \leq K(u+v)^\beta, \\ \alpha^*(r) \leq \frac{M}{r^\delta}, \quad \delta > 2, \end{cases} \quad (2.22)$$

where  $M$  is a constant.

Now we introduce the following estimator which is a modified version of some estimators introduced in Shao *et al.* (1995) [33] (write  $\ell = \ell_n$ ).

$$B_{n,2}^2 = \frac{1}{n-\ell+1} \sum_{j=0}^{n-\ell} \left( \frac{S_j(\ell) - \bar{X}_\ell}{\sqrt{\ell}} \right)^2,$$

and

$$\hat{B}_{n,2}^2 = \frac{1}{n-\ell+1} \sum_{j=0}^{n-\ell} \left( \frac{S_j(\ell) - \ell\mu}{\sqrt{\ell}} \right)^2,$$

with

$$S_j(\ell) = \sum_{k=j+1}^{j+\ell} X_k \text{ and } \bar{X}_\ell = \frac{1}{n-\ell+1} \sum_{j=0}^{n-\ell} S_j(\ell).$$

Note that as  $n \rightarrow \infty$ , (2.17) may be written as follows

$$\frac{\text{Var}(S_n)}{n} := \frac{\sigma_n^2}{n} \rightarrow \sigma^2. \quad (2.23)$$

We want to prove that  $B_{n,2} \rightarrow \sigma$  in  $L^2$  as  $n \rightarrow \infty$  where  $\sigma^2$  is defined in (2.16). Our second result is given in Theorem 2.3.2. This theorem shows the weak consistency of  $B_{n,2}$  to  $\sigma$ .

**Theorem 2.3.2.** Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(\ell,\ell)}^*$ -mixing process of random variables where (2.24) and (2.25) are satisfied.

$$\mathbb{E}|X_1|^{2+\delta} < \infty, \text{ for some } \delta > 0, \quad (2.24)$$

$$\sum_{r=1}^{\infty} \alpha^*(r)^{\frac{\delta}{2+\delta}} < \infty. \quad (2.25)$$

In addition assume that

$$C(u, v) \leq K^{u+v}, \quad (2.26)$$

and consider

$$\ell_n = \ell, \quad \ell = [\ln(n)/5 \ln(K)] + x, \quad (2.27)$$

where  $[y]$  denotes the integer part of  $y$  and  $x$  is an integer. If  $0 < \sigma < \infty$ , then

$$B_{n,2} \rightarrow \sigma \text{ in } L^2. \quad (2.28)$$

**Remark 2.3.3.** Note that a consequence of (2.21) and (2.28) is that

$$\frac{S_n - n\mu}{\sqrt{n}B_{n,2}} \xrightarrow{d} N(0, 1). \quad (2.29)$$

This proves a S.N.C.L.T. for mixing processes satisfying our  $\alpha_{u,v}^*$ -mixing condition.

We recall that from Corollary 2.2.1 conditions (2.24) and (2.25) can be replaced by the U.I. condition of  $\left\{ \left| \frac{S_\ell - \ell\mu}{\sqrt{\ell}} \right|^2, n \geq 1 \right\}$ .

Our third principal result is given in Theorem 2.3.3 and shows the asymptotic normality of  $\widehat{B}_{n,2}$ .

**Theorem 2.3.3.** Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary  $\alpha_{(u,v)}^*$ -mixing process for all  $(u, v) \in \mathbb{N}^*$  with  $\mathbb{E}(X_1) = \mu$ . Denote by  $S(\ell) = \sum_{k=1}^{\ell} X_k$  and let  $\ell$  be as in (2.27). Assume that

$$\left\{ \left( \frac{S_n - n\mu}{\sqrt{n}} \right)^{4+2\delta}, n \geq 1 \right\}, \quad (2.30)$$

is uniformly integrable with  $\delta > 0$ . If in addition (2.23) holds with  $\sigma > 0$  and if there exist  $M > 0$  and  $K > 0$  such that for any  $u, v$ , and  $r$ ,

$$\begin{cases} C(u, v) \leq K^{u+v}, \\ \alpha^*(r) \leq \frac{M}{r^\delta}, \quad \delta > \max\{2 \log(K), 2\}, \end{cases} \quad (2.31)$$

then

$$\sqrt{\frac{n}{\ell}} \left( \widehat{B}_{n,2} - \mathbb{E} \left( \left| \frac{S(\ell) - \ell\mu}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{3}\right). \quad (2.32)$$

**Remark 2.3.4.** Note that the choice of  $\ell$  depends on the behavior of  $C(u, v)$ . Assume that there is  $K > 0$  such that  $\forall (u, v)$  we have

1. Case 1:  $C(u, v) \leq K$  we may choose  $\ell_n = o(n)$  and Theorems 2.3.2 and 2.3.3 hold.
2. Case 2:  $C(u, v) \leq K(u+v)^\beta$  i.e.  $C(u, v)$  is a polynomial on  $u$  and  $v$  then we can choose  $\ell_n = n^s$  with  $s < \frac{1}{4}$  if  $\beta < 4$  and  $s < \frac{1}{\beta+1}$  if  $\beta \geq 4$  and Theorems 2.3.2 and 2.3.3 hold.

**Remark 2.3.5.** Note that from Theorem 2.2.1 the uniform integrability condition in Theorem 2.3.3 may be replaced by the following two conditions:  $\mathbb{E}|X|^{4+3\delta} < \infty$  and  $\alpha^*(r) \leq Cr^{-\theta}$  where  $\theta > \frac{(4+2\delta)(4+3\delta)}{2\delta}$ . Consider  $\theta = \frac{4(2+\delta)(1+\delta)}{\delta}$  this implies that  $\sum \alpha^*(r)^{\frac{\delta}{2+\delta}} < \infty$  where  $C$  is a constant.

The following corollary states that the same result provided in Theorem 2.3.3 holds for  $B_{n,2}$  instead of  $\widehat{B}_{n,2}$ .

**Corollary 2.3.1.** If the sequence  $(X_n)_{n \in \mathbb{N}}$  satisfies the same conditions as in Theorem 2.3.3 then the result of Theorem 2.3.3 holds for  $B_{n,2}$  instead of  $\widehat{B}_{n,2}$ , that is

$$\sqrt{\frac{n}{\ell}} \left( B_{n,2} - \mathbb{E} \left( \left| \frac{S(\ell) - \ell\mu}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \right) \xrightarrow{d} N(0, \frac{\sigma^2}{3}). \quad (2.33)$$

## 2.4 Proof of the main results

*Proof of Proposition 2.2.1.* For convenience, we denote  $f$  for  $f(X_{i_1}, \dots, X_{i_u})$  (resp.  $g$  for  $g(X_{j_1}, \dots, X_{j_v})$ ). We introduce the random variables  $f_N, \widehat{f}_N$  defined by

$$f_N = \begin{cases} f, & \text{if } |f| \leq N \\ N, & \text{if } |f| > N, \end{cases}$$

$$\widehat{f}_N = f - f_N,$$

and  $g_N, \widehat{g}_N$  similarly defined. Then

$$|\text{Cov}(f, g)| \leq |\text{Cov}(f_N, g_N)| + |\text{Cov}(f - f_N, g_N)| + |\text{Cov}(f_N, g - g_N)| + |\text{Cov}(f - f_N, g - g_N)|.$$

Using the  $\alpha_{(u,v)}^*$  mixing property, we have  $|\text{Cov}(f_N, g_N)| \leq \alpha^*(r)C(u, v)N^2$ , and using Markov inequality we obtain

$$|\text{Cov}(\widehat{f}_N, g_N)| = |\mathbb{E}(\widehat{f}_N, g_N) - \mathbb{E}(\widehat{f}_N)\mathbb{E}(g_N)| \leq 2N\mathbb{E}|\widehat{f}_N| \leq 2N \frac{\mathbb{E}|f|^p}{N^{p-1}},$$

and

$$|\text{Cov}(f_N, \widehat{g}_N)| \leq 2N\mathbb{E}|\widehat{g}_N| \leq 2N \frac{\mathbb{E}|g|^q}{N^{q-1}}.$$

Indeed, take  $1 < s \leq \infty$  then

$$\mathbb{E}|\widehat{f}_N| = \int |f - f_N| \mathbf{1}_{\{|f| > N\}} \leq \int \frac{|f|^s}{|f|^{s-1}} \mathbf{1}_{\{|f| > N\}} \leq \int \frac{|f|^s}{N^{s-1}} = \frac{\mathbb{E}|f|^s}{N^{s-1}}.$$

By the inequalities  $\mathbb{E}|\widehat{f}_N| \leq N \frac{\mathbb{E}|f|^s}{N^s}$  and  $|\text{Cov}(\widehat{f}_N, \widehat{g}_N)| \leq 2 \left\| \widehat{f}_N \right\|_{\frac{p+q}{p}} \left\| \widehat{g}_N \right\|_{\frac{p+q}{q}}$  we obtain

$$|\text{Cov}(\widehat{f}_N, \widehat{g}_N)| \leq N^2 \left( \frac{\mathbb{E}|f|^p}{N^p} \right)^{\frac{q}{q+p}} \left( \frac{\mathbb{E}|g|^q}{N^q} \right)^{\frac{p}{q+p}} \leq N^2 \frac{(\mathbb{E}|f|^p)^{\frac{q}{q+p}} (\mathbb{E}|g|^q)^{\frac{p}{q+p}}}{N^{\frac{2pq}{p+q}}}.$$

Hence

$$|\text{Cov}(f, g)| \leq \alpha^*(r)C(u, v)N^2 + 2N^2 \frac{\mathbb{E}|f|^p}{N^p} + 2N^2 \frac{\mathbb{E}|g|^q}{N^q} + 2N^2 \frac{(\mathbb{E}|f|^p)^{\frac{q}{q+p}} (\mathbb{E}|g|^q)^{\frac{p}{q+p}}}{N^{\frac{2pq}{p+q}}}.$$

We choose  $N$  such that

$$N^2 \alpha^*(r) \leq N^2 \frac{\mathbb{E}(|f|^p)}{N^p}, \quad (2.34)$$

and that

$$N^2 \alpha^*(r) \leq N^2 \frac{\mathbb{E}(|g|^q)}{N^q}, \quad (2.35)$$

then

$$N \leq \left( \frac{\mathbb{E}(|f|^p)}{\alpha^*(r)} \right)^{\frac{1}{p}},$$

and

$$N \leq \left( \frac{\mathbb{E}(|g|^q)}{\alpha^*(r)} \right)^{\frac{1}{q}}.$$

By (2.34) and (2.35) we have that

$$N^2 \leq \left( \frac{\mathbb{E}(|f|^p)}{\alpha^*(r)} \right)^{\frac{1}{p}} \left( \frac{\mathbb{E}(|g|^q)}{\alpha^*(r)} \right)^{\frac{1}{q}},$$

whence

$$N^2 \alpha^*(r) \leq \alpha^*(r)^{1-\frac{1}{p}-\frac{1}{q}} \mathbb{E}(|g|^q)^{\frac{1}{q}} \mathbb{E}(|f|^p)^{\frac{1}{p}}. \quad (2.36)$$

so that

$$2\alpha^*(r) \leq \left( \frac{\mathbb{E}(|g|^q)}{N^q} + \frac{\mathbb{E}(|f|^p)}{N^p} \right).$$

By Young's inequality (see Young (1912)[44]) we have

$$\frac{(\mathbb{E}|f|^p)^{\frac{q}{q+p}} (\mathbb{E}|g|^q)^{\frac{p}{q+p}}}{N^{\frac{2pq}{p+q}}} \leq \frac{p}{p+q} \frac{\mathbb{E}(|g|^q)}{N^q} + \frac{q}{p+q} \frac{\mathbb{E}(|f|^p)}{N^p} \leq \frac{\mathbb{E}(|g|^q)}{N^q} + \frac{\mathbb{E}(|f|^p)}{N^p},$$

hence

$$|\text{Cov}(f, g)| \leq \alpha^*(r) C(u, v) N^2 + 4N^2 \left( \frac{\mathbb{E}|f|^p}{N^p} + \frac{\mathbb{E}|g|^q}{N^q} \right).$$

In addition choose  $N$  such

$$\left( \frac{\mathbb{E}(|g|^q)}{N^q} + \frac{\mathbb{E}(|f|^p)}{N^p} \right) = 2\alpha^*(r), \quad (2.37)$$

then by (2.37) and (2.36) we have

$$|\text{Cov}(f, g)| \leq (C(u, v) + 8) (\mathbb{E}|f|^p)^{\frac{1}{p}} (\mathbb{E}|g|^q)^{\frac{1}{q}} \alpha^*(r)^{1-\frac{1}{p}-\frac{1}{q}}. \quad (2.38)$$

□

*Proof of Theorem 2.2.1.* This proof is an adaptation to our case of the proof of Theorem 4.1 in Shao and Yu (1996)[19]. In fact in our case the mixing coefficients depend on  $u$  and  $v$ . We shall prove inequality (2.12) by induction on  $n$ . Consider

$$E|S_k|^p \leq A \left( k^{p/2} (K_2 D_{n,2,2})^{p/2} + k^{p+\frac{\theta(p-r)}{r} \vee (1+\epsilon)} \|X\|_r^p \right). \quad (2.39)$$

Note that (2.40) is obvious for  $k = 1$  and assume that (2.40) is true for each  $1 < k < n$ .

$$E|S_k|^p \leq A \left( k^{p/2} (K_2 D_{n,2,2})^{p/2} + k^{p+\frac{\theta(p-r)}{r} \vee (1+\epsilon)} \|X\|_r^p \right). \quad (2.40)$$

Define by

$$\xi_i = \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} X_j, \eta_i = \sum_{j=1+(2i-1)m}^{n \wedge 2im} X_j, \quad (2.41)$$

where  $1 \leq i \leq k_n := 1 + \lceil n/(2m) \rceil$  and  $m < \lceil n/2 \rceil - 1$ . Note that for fixed  $i$  there exists a gap of length  $m$  between  $\xi_i$  and  $\eta_i$ . Let  $a$  be a constant such that  $0 < a < \frac{1}{2}$  and take  $m = \lceil an \rceil + 1$ . By Minkowski inequality we write

$$E|S_n|^p \leq 2^{p-1} (\mathbb{E} \left| \sum_{i=1}^{k_n} \xi_i \right|^p + \mathbb{E} \left| \sum_{i=1}^{k_n} \eta_i \right|^p) := 2^{p-1} (I_1 + I_2).$$

The proof of Theorem 2.2.1 proceeds as in Shao and Yu (1996)[19]. We begin by recalling a lemma useful for our proof.

**Lemma 2.4.1** ([19], Theorem 4.1). *Let  $\xi_i, 1 \leq i \leq n$  be a sequence of random variables and define  $F_i$  by  $F_i = \sigma(\xi_j, j \leq i)$ . Then for any  $p \geq 2$  there is a constant  $D = D(p)$  such that*

$$\begin{aligned} \mathbb{E} \left| \sum_{i=1}^n \xi_i \right|^p &\leq D \left( \left( \sum_{i=1}^n \mathbb{E}(\xi_i^2) \right)^{\frac{p}{2}} + \sum_{i=1}^n \mathbb{E}|\xi_i|^p + n^{p-1} \sum_{i=1}^n \mathbb{E}|\mathbb{E}(\xi_i|F_{i-1})|^p \right. \\ &\quad \left. + n^{\frac{p}{2}-1} \times \left| \sum_{i=1}^n \mathbb{E}|\mathbb{E}(\xi_i^2|F_{i-1}) - \mathbb{E}(\xi_i^2)|^{\frac{p}{2}} \right| \right). \end{aligned}$$

Replacing  $n$  by  $k_n$  in Lemma 2.4.1 yields to

$$I_1 \leq D \left( \underbrace{\left( \sum_{i=1}^{k_n} \mathbb{E}|\xi_i|^2 \right)^{\frac{p}{2}}}_{I_{1,1}} + \sum_{i=1}^{k_n} \mathbb{E}|\xi_i|^p + \underbrace{k_n^{p-1} \sum_{i=1}^{k_n} \mathbb{E}|\mathbb{E}(\xi_i|F_{i-1})|^p}_{I_{1,2}} + \underbrace{k_n^{\frac{p}{2}-1} \sum_{i=1}^{k_n} \mathbb{E}|\mathbb{E}(\xi_i^2|F_{i-1}) - \mathbb{E}(\xi_i^2)|^{\frac{p}{2}}}_{I_{1,3}} \right).$$

We begin with the bound for  $I_{1,1}$ . From (2.6) and (2.15) we deduce that

$$I_{1,1} \leq (k_n 2m K_2 D_{m,2,2})^{\frac{p}{2}} \leq 2^{\frac{p}{2}} (k_n m)^{\frac{p}{2}} (K_2 D_{m,2,2})^{\frac{p}{2}} \leq 2^p K_2^{\frac{p}{2}} (n D_{n,2,2})^{\frac{p}{2}}.$$

where  $k_n m < 2n$ .

Now we turn our attention to  $I_{1,3}$ . Define  $Y_i = \mathbb{E}(\xi_i^2|F_{i-1}) - \mathbb{E}(\xi_i^2)$  and denote by  $F_{i-1} = \sigma(\xi_1, \dots, \xi_{i-1})$ .  $Y_i$  is a measurable function with respect to  $F_{i-1}$ . Write  $\mathbb{E}(|Y_i|^{p/2}) = \mathbb{E}|Y_i|^{(p/2)-1} \text{sgn}(Y_i) Y_i$  then we have

$$\begin{aligned} \mathbb{E}(|Y_i|^{p/2}) &= \mathbb{E}(|Y_i|^{(p/2)-1} \text{sgn}(Y_i) (\xi_i^2 - \mathbb{E}(\xi_i^2))) \\ &\leq \sup_{j,l} \sum_{2(i-1)m < j, l \leq n \wedge (2i-1)m} \text{Cov} \left( |Y_i|^{(p/2)-1}, X_j X_l - \mathbb{E}(X_j X_l) \right). \end{aligned}$$

Consider that

$$f(X_{1+2(i-1)m}, \dots, X_{n \wedge (2i-1)m}) = |Y_i|^{p/2-1} \text{ and } g(X_j, X_l) = |X_j X_l - \mathbb{E}(X_j X_l)|. \quad (2.42)$$

Choose  $p$  in (2.3) equals to  $\frac{p}{p-2}$  and  $q = \frac{r}{2}$  such that  $\frac{2}{r} + \frac{p-2}{p} < 1$ , in this case by Proposition 2.2.1 we have

$$\left| \text{Cov} \left( |Y_i|^{p/2-1}, X_j X_l - \mathbb{E}(X_j X_l) \right) \right| \leq (C(m, 2) + 8) (\mathbb{E}|Y_i|^{\frac{p}{2}})^{\frac{p-2}{p}} \|X_j X_l\|_{\frac{2}{r}} \alpha^*(m)^{1-\frac{r}{2}-\frac{p-2}{p}}.$$

Then

$$\mathbb{E}|Y_i|^{p/2} \leq (C(m, 2) + 8) m^2 (\mathbb{E}|Y_i|^{\frac{p}{2}})^{\frac{p-2}{p}} \|X\|_r^2 \alpha^*(m)^{\frac{2}{p}-\frac{2}{r}}.$$

Let  $B$  be a constant. Hence

$$\mathbb{E}|Y_i|^{p/2} \leq (C(m, 2) + 8)^{\frac{p}{2}} m^p \|X\|_r^p \alpha^*(m)^{1-\frac{p}{r}}. \quad (2.43)$$

By inequality (2.11) and as  $\theta > 0$  and  $p < r$  we write

$$\begin{aligned} I_{1,3} &\leq k_n^{\frac{p}{2}} (C(m, 2) + 8)^{\frac{p}{2}} m^p \|X\|_r^p \alpha^*(m)^{1-\frac{p}{r}} \\ &\leq 2^p (C(m, 2) + 8)^{\frac{p}{2}} n^p \alpha^*(m)^{1-\frac{p}{r}} \|X\|_r^p \\ &\leq 2^p B (C(m, 2) + 8)^{\frac{p}{2}} n^p m^{\theta \frac{(p-r)}{r}} \|X\|_r^p \\ &\leq 2^p B (C(m, 2) + 8)^{\frac{p}{2}} n^{p+\theta \frac{(p-r)}{r}} a^{\theta \frac{(p-r)}{r}} \|X\|_r^p \\ &\leq 2^p B (C(m, 2) + 8)^p n^{(p+\theta \frac{(p-r)}{r}) \vee (1+\varepsilon)} a^{\theta \frac{(p-r)}{r}} \|X\|_r^p \end{aligned}$$

then

$$I_{1,3} \leq 2^p B (C(m, 2) + 8)^p n^{(p+\theta \frac{(p-r)}{r}) \vee (1+\varepsilon)} a^{\theta \frac{(p-r)}{r}} \|X\|_r^p.$$

Now we focus on  $I_{1,2}$ . Define  $Z_i = \mathbb{E}(\xi_i | F_{i-1})$  and denote by

$$f(X_{1+2(i-1)m}, \dots, X_{n \wedge (2i-1)m}) = |Y_i|^p \text{ and } g(X_j) = |X_j - \mathbb{E}(X_j)|. \quad (2.44)$$

Now choose  $p$  in (2.3) equals to  $\frac{p}{p-1}$  and  $q = r$  where  $\frac{1}{r} + \frac{p-1}{p} < 1$ . Then in this case we have

$$\begin{aligned} \mathbb{E}(|Z_i|^p) &= \mathbb{E}(|Z_i|^{p-1} \text{sgn}(Z_i)(\xi_i)) \leq \sup \sum_{2(i-1)m < j \leq n \wedge (2i-1)m} \text{Cov}(|Z_i|^{p-1}, X_j) \\ &\leq (C(m, 1) + 8) m (\mathbb{E}|Z_i|^p)^{\frac{p-1}{p}} \|X\|_r \alpha^*(m)^{\frac{1}{p}-\frac{1}{r}}. \end{aligned}$$

Hence

$$\mathbb{E}(|Z_i|^p) \leq (C(m, 1) + 8)^p m^p \|X\|_r^p \alpha^*(m)^{1-\frac{p}{r}}. \quad (2.45)$$

By inequality (2.11) and as  $\theta > 0$  and  $p < r$  we have

$$\begin{aligned} I_{1,2} &\leq k_n^p (C(m, 2) + 8)^p m^p \|X\|_r^p \alpha^*(m)^{1-\frac{p}{r}} \\ &\leq 2^p C (C(m, 2) + 8)^p n^{(p+\theta \frac{(p-r)}{r}) \vee (1+\varepsilon)} a^{\theta \frac{(p-r)}{r}} \|X\|_r^p, \end{aligned}$$

whence

$$\begin{aligned} I_1 &\leq D \left( \sum_{i=1}^{k_n} \mathbb{E}|\xi_i|^p + 2^p K_2^{\frac{p}{2}} (n D_{n,2,2})^{\frac{p}{2}} \right. \\ &\quad \left. + 2^{p+1} B (C(m, 2) + 8)^p n^{(p+\theta \frac{(p-r)}{r}) \vee (1+\varepsilon)} a^{\theta \frac{(p-r)}{r}} \|X\|_r^p \right). \end{aligned}$$

where  $C$  is a constant. Similarly for

$$I_2 \leq D \left( \sum_{i=1}^{k_n} \mathbb{E}|\eta_i|^p + 2^p K_2^{\frac{p}{2}} (nD_{n,2,2})^{\frac{p}{2}} + 2^{p+1} B(C(m, 2) + 8)^p n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} a^{\theta\frac{(p-r)}{r}} \|X\|_r^p \right).$$

Consequently, we have

$$\mathbb{E}|S_n|^p \leq 2^{p-1} D \left( \sum_{i=1}^{k_n} \mathbb{E}|\xi_i|^p + \sum_{i=1}^{k_n} \mathbb{E}|\eta_i|^p + 2^{p+1} K_2^{\frac{p}{2}} (nD_{n,2,2})^{\frac{p}{2}} + 2B(C(m, 2) + 8)^p 2^{p+1} n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} a^{\theta\frac{(p-r)}{r}} \|X\|_r^p \right). \quad (2.46)$$

Remark that  $\sum_{i=1}^{k_n} \mathbb{E}|\xi_i|^p + \sum_{i=1}^{k_n} \mathbb{E}|\eta_i|^p \leq 2k_n \mathbb{E}|\xi_1|^p = 2k_n \mathbb{E}|S_m|^p$ . Take  $a = (2^{p+4}D)^{-\frac{1}{\epsilon} - \frac{2}{p-2}}$ . Assume that (2.40) is true for all  $k < n$  where  $A = 2^{p+1}D \left[ 2^p + 2^{p+1}C(C(m, 2) + 8)^p a^{\theta\frac{(p-r)}{r}} \right]$ . In this case by (2.46) and (2.40) we get

$$\begin{aligned} \mathbb{E}|S_n|^p &\leq 2^p D k_n A \left[ m^{p/2} (K_2 D_{m,2,2})^{p/2} + m^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right] \\ &\quad + \frac{A}{2} \left( n^{\frac{p}{2}} (K_2 D_{n,2,2})^{\frac{p}{2}} + n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} \|X\|_r^p \right) \\ &\leq 2^p D \frac{n}{m} A \left( m^{p/2} (K_2 D_{n,2,2})^{p/2} + m^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right) \\ &\quad + \frac{A}{2} \left( n^{\frac{p}{2}} (K_2 D_{n,2,2})^{\frac{p}{2}} + n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} \|X\|_r^p \right) \\ &\leq \frac{A}{2} 2^{p+1} D \frac{n}{m} \left( m^{p/2} (K_2 D_{n,2,2})^{p/2} + m^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right) \\ &\quad + \frac{A}{2} \left( n^{\frac{p}{2}} (K_2 D_{n,2,2})^{\frac{p}{2}} + n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} \|X\|_r^p \right) \quad (2.47) \\ &\leq \frac{A}{2} 2^{p+1} D \left[ a^{\frac{p-2}{2}} n^{p/2} (K_2 D_{n,2,2})^{p/2} + a^\epsilon n^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right] \\ &\quad + \frac{A}{2} \left( n^{\frac{p}{2}} (K_2 D_{n,2,2})^{\frac{p}{2}} + n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} \|X\|_r^p \right) \\ &\leq \frac{A}{2} \left[ n^{p/2} (K_2 D_{n,2,2})^{p/2} + n^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right] \\ &\quad + \frac{A}{2} \left( n^{\frac{p}{2}} (K_2 D_{n,2,2})^{\frac{p}{2}} + n^{(p+\theta\frac{(p-r)}{r})\vee(1+\epsilon)} \|X\|_r^p \right) \\ &= A \left[ n^{p/2} (K_2 D_{n,2,2})^{p/2} + n^{p+\theta\frac{(p-r)}{r}\vee(1+\epsilon)} \|X\|_r^p \right], \end{aligned}$$

Finally we conclude that (2.40) remains valid for  $k = n$ .  $\square$

*Proof of Theorem 2.3.1.* To prove this theorem we follow the proof of Theorem 18.5.3 in Ibragimov and Linnik (1971)[23] (see also Theorem 1.7 in Ibragimov (1962)[21]) for strongly mixing sequences. It is convenient to mention that in their proof, these authors deal with the case of bounded r.v's, to which the general theorem for non bounded r.v's may be reduced under a truncation technique. We concentrate on points that we must readjust to our framework. Recall that in our case, mixing coefficients may be exponentially bounded

with respect to  $u$  and  $v$  such that  $C(u, v) \leq K^{u+v}$  where  $K$  is a positive constant. By a careful analysis of the proof of Ibragimov (1962)[21], which is based on Bernstein's method, we note that Theorem 2.3.1 holds if these conditions are satisfied:

1. There exist  $p = o(n)$ ,  $q = \lfloor n/p \rfloor$  and  $k = \lfloor n/(p+q) \rfloor$  such that the distribution function  $F_p(z) = P(X_1 + \dots + X_p < z)$  satisfies

$$\lim_{n \rightarrow \infty} \frac{n}{p\sigma_n^2} \int_{|z| > \epsilon\sigma_n} z^2 dF_p(z) = 0, \quad (2.48)$$

and

- 2.

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \frac{1}{\sigma_n} \sum_{i=0}^k \eta_i \right)^2 = 0, \quad (2.49)$$

where

$$\eta_i = \sum_{(i+1)p+iq+1}^{(i+1)p+(i+1)q} X_j, \quad (0 \leq i \leq k-1).$$

Using Markov's inequality the first condition (2.48) may be written as follows

$$\frac{n}{p\sigma_n^2} \int_{|z| > \epsilon\sigma_n} z^2 dF_p(z) \leq \frac{n}{\epsilon^2 p \sigma_n^4} \int_{\infty}^{\infty} z^4 dF_p(z) = \frac{n}{\epsilon^2 p \sigma^4 n^2} \int_{\infty}^{\infty} z^4 dF_p(z), \quad (2.50)$$

where  $\int_{\infty}^{\infty} z^4 dF_p(z) = E(\sum_{i=1}^p X_i)^4$ . Consider that  $|X_i| \leq C_0 < \infty$ . Under the condition  $\sum_{j=1}^{\infty} \alpha^*(j) < \infty$  and from the inequality  $|\mathbb{E}(X_0, X_j)| \leq C(1, 1)C_0^2 \alpha^*(j)$  it follows that  $\sigma^2 < \infty$  and that  $\sigma_n^2 = \sigma^2 n(1 + o(1))$ . Hence

$$\lim_{n \rightarrow \infty} \frac{n}{p\sigma_n^2} \int_{|z| > \epsilon\sigma_n} z^2 dF_p(z) = \lim_{n \rightarrow \infty} \frac{n E(\sum_{i=1}^p X_i)^4}{\epsilon^2 p \sigma^4 n^2}. \quad (2.51)$$

Ibragimov and Linnik (1971)[23] have proved in Lemma 18.5.2 that  $E(\sum_{i=1}^n X_i)^4 = o(n^3)$ . By taking their proof we derive that  $E(\sum_{i=1}^n X_i)^4 = O(n^2)$  if  $\sum_{i=1}^{\infty} j \alpha^*(j) < \infty$ . This last condition is satisfied if we assume that  $\alpha^*(j) \leq \frac{M}{j^\delta}$  where  $\delta > 2$  and  $M$  is a constant. Hence (2.50) is reduced to

$$\lim_{n \rightarrow \infty} \frac{np^2}{\epsilon^2 p \sigma^4 n^2} = \lim_{n \rightarrow \infty} \frac{p}{\epsilon^2 \sigma^4 n} = O\left(\frac{p}{n}\right) = 0.$$

This is true for example for  $p = \lfloor \frac{n}{\log(n)} \rfloor$ . For the second condition (2.49) we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \frac{1}{\sigma_n} \sum_{i=0}^k \eta_i \right)^2 = \frac{1 + o(1)}{\sigma^2 n} \left( k \mathbb{E}(\eta_0)^2 + 2 \sum_{j=1}^k (k-j) \mathbb{E}(\eta_0 \eta_j) + 2 \sum_{i=0}^{k-1} \mathbb{E}(\eta_i \eta_k) + \mathbb{E}(\eta_k)^2 \right).$$

Define,  $p = \lfloor \frac{n}{\log(n)} \rfloor$ ,  $q = \lfloor \log(n) \rfloor$  and  $k = \lfloor n/(p+q) \rfloor \approx \frac{n}{\log(n)}$ . Note that when  $n \rightarrow \infty$ , the number of term in  $\eta_k$  is less than  $q$ . It is clear that  $\frac{k}{n} \mathbb{E}(\eta_0)^2 = o(1)$  and  $\mathbb{E}(\eta_k)^2 = o(1)$ . So it suffices to prove that

$$\left| n^{-1} \sum_{j=1}^{k-1} (k-j) \mathbb{E}(\eta_0 \eta_j) \right| \rightarrow 0. \quad (2.52)$$

By (2.3) we deduce that  $|\mathbb{E}(\eta_i \eta_j)| \leq C_0^2 q^2 (C(q, q) + 8) \alpha^*(p|i-j)$ . Therefore if  $C(q, q) \leq K^{2q}$  then

$$\left| n^{-1} \sum_{j=1}^{k-1} (k-j) \mathbb{E}(\eta_i \eta_j) \right| \leq \frac{k C_0^2 q^2 (C(q, q) + 8) \sum_{j=1}^{k-1} \alpha^*(pj)}{n} \leq \frac{k C_0^2 q^2 (K^{2q} + 8) \sum_{j=1}^{k-1} \alpha^*(pj)}{n}.$$

Then if we replace  $q$  and  $k$  by their values we obtain

$$\left| n^{-1} \sum_{j=1}^{k-1} (k-j) \mathbb{E}(\eta_i \eta_j) \right| \leq B \times \log(n) (K^{2 \log(n)} + 8) \sum_{j=1}^{k-1} \alpha^*(pj).$$

where  $B$  is a constant. Then in this case by the fact that  $\alpha(j) \leq \frac{M}{j^\delta}$  we have

$$\sum_{j=1}^{k-1} \alpha^*(pj) \leq \frac{\log^\delta(n)}{n^\delta} \underbrace{\sum_{j=1}^{k-1} \frac{M}{j^\delta}}_{< Cte},$$

where  $Cte$  is a constant.

Hence

$$B \times \log(n) K^{2 \log(n)} \sum_{j=1}^{k-1} \alpha^*(pj) \leq B \times \log(n) K^{2 \log(n)} \frac{\log^\delta(n)}{n^\delta} = B \times \frac{\log^{\delta+1}(n) K^{2 \log(n)}}{n^\delta}.$$

where  $B$  is a constant. Note that

$$\frac{\log^{\delta+1}(n) K^{2 \log(n)}}{n^\delta} = O\left(\frac{K^{2 \log(n)}}{n^\delta}\right).$$

If  $\delta > 2 \log(K)$ , then  $O\left(\frac{K^{2 \log(n)}}{n^\delta}\right) = o(1)$ . Hence (2.52) holds if  $\alpha^*(j) < \frac{M}{j^\delta}$  with  $\delta > \max\{2, 2 \log(K)\}$ .  $\square$

**Remark 2.4.1.** Note that if  $C(u, v) \leq K(u+v)^\beta$  (i.e.  $C(u, v)$  is a polynomial function of  $u$  and  $v$ ) and if

$$\alpha^*(r) \leq \frac{M}{r^\delta}, \quad \delta \geq 2,$$

where  $M$  is a constant then

$$\left| n^{-1} \sum_{j=1}^{k-1} (k-j) \mathbb{E}(\eta_i \eta_j) \right| \leq B' \frac{k}{n} (2q)^\beta \sum_{j=1}^{k-1} \alpha^*(pj) \leq B'' \log(n)^{\delta+\beta-1} \frac{1}{n^\delta} \rightarrow 0,$$

where  $B'$  and  $B''$  are constants. Hence results of Theorem 2.3.1 holds.

Before proceeding to the proof of Theorem 2.3.2, we give the following lemma which is useful in the proof of this theorem. This lemma is a variant of Lemma 2.3 provided in Shi (2000)[41].

**Lemma 2.4.2.** Let  $(X_n, n \geq 1)$  be  $\alpha_{(\ell_n, \ell_n)}^*$ -mixing sequence where  $1 \leq \ell_n \leq n$ . Let  $f$  be a real function on  $\mathbb{R}^{\ell_n}$  and put  $Z_i = f(X_{i+1}, \dots, X_{i+\ell_n})$ . Consider  $\mathbb{E}|Z_k|^{2+\delta} < \infty$  for some  $\delta > 0$  and  $\sup_{j \in \mathbb{N}} \|Z_j\|_{2+\delta}^2 < \infty$ . Then we have

$$\text{Var} \left( \sum_{k=1}^n Z_k \right) \leq 2Kn\ell_n \sup_{j \in \mathbb{N}} \|Z_j\|_{2+\delta}^2 (C(\ell_n, \ell_n) + 8) \sum_{r=1}^{\lfloor \frac{n}{\ell_n} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}}.$$

where  $K$  is a constant.

*Proof of Lemma 2.4.2.* Let  $f$  be a real function on  $\mathbb{R}^{\ell_n}$  and denote by  $Z_i = f(X_{i+1}, \dots, X_{i+\ell_n})$ . Let  $\lfloor \frac{n}{\ell_n} \rfloor$  be the integer part of  $\frac{n}{\ell_n}$ , then by Proposition 2.2.1 we have

$$\begin{aligned} \text{Var} \left( \sum_{k=1}^n Z_k \right) &\leq \text{Var} \left( \sum_{j=1}^{\ell_n} \sum_{i=0}^{\lfloor \frac{n}{\ell_n} \rfloor} Z_{i\ell_n+j} \right) \\ &\leq \ell_n^2 \max_j \text{Var} \left( \sum_{i=0}^{\lfloor \frac{n}{\ell_n} \rfloor} Z_{i\ell_n+j} \right) \\ &\leq 2 \left( \lfloor \frac{n}{\ell_n} \rfloor + 1 \right) \ell_n^2 \sup_{j \in \mathbb{N}} \|Z_{j,1}\|_{2+\delta}^2 \sum_{r=1}^{\lfloor \frac{n}{\ell_n} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} (C(\ell_n, \ell_n) + 8) \\ &\leq 2Kn\ell_n \sup_{j \in \mathbb{N}} \|Z_{j,1}\|_{2+\delta}^2 (C(\ell_n, \ell_n) + 8) \sum_{r=1}^{\lfloor \frac{n}{\ell_n} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}}, \end{aligned}$$

□

Note that the idea in the proof of Theorem (2.3.2) (resp. of Theorem 2.3.3) is to adapt the proof of Theorem 1.1 (resp. Theorem 1.2) derived in Shi (2000)[41]. This is due to our mixing condition on which mixing coefficient may have an exponential behaviour with respect to  $u$  and  $v$ .

*Proof of Theorem 2.3.2.* Without loss of generality consider that  $\mu = 0$ . It derives from inequality (2.25) that  $n^{\frac{2}{\delta}} \alpha^*(n) \leq \alpha^*(n)^{\frac{\delta}{2+\delta}}$ . In fact condition (2.25) implies that for large  $n$  we have  $\alpha^*(n)^{\frac{\delta}{\delta+2}} \leq \frac{1}{n}$  whence  $\alpha^*(n)n^{\frac{2}{\delta}} \leq \alpha^*(n)^{\frac{\delta}{\delta+2}}$  for  $n$  large enough. First we prove that

$$\left| B_{n,2} - \widehat{B}_{n,2} \right| \rightarrow 0 \text{ in } L^2. \quad (2.53)$$

By Minkowski inequality we have

$$\left( \sum_{j=0}^{n-\ell} \left| \frac{S_j(\ell) - \overline{X}_\ell}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \leq \left( \sum_{j=0}^{n-\ell} \left| \frac{S_j(\ell)}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} + \left( \sum_{j=0}^{n-\ell} \left| \frac{\overline{X}_\ell}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}}.$$

Hence,

$$\left| B_{n,2} - \widehat{B}_{n,2} \right| \leq \frac{1}{\sqrt{n-\ell+1}} \left\{ \left( \sum_{j=0}^{n-\ell} \left| \frac{S_j(\ell)}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} + \left( \sum_{j=0}^{n-\ell} \left| \frac{\overline{X}_\ell}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} - \left( \sum_{j=0}^{n-\ell} \left| \frac{S_j(\ell)}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \right\}.$$

We obtain

$$E \left| B_{n,2} - \widehat{B}_{n,2} \right|^2 \leq \frac{1}{n-\ell+1} E \left| \left( \sum_{j=0}^{n-\ell} \left| \frac{\overline{X}_\ell}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \right|^2 = E \left| \frac{\overline{X}_\ell^2}{\ell} \right|.$$

Moreover,

$$E \left| \frac{\overline{X}_\ell^2}{\ell} \right| = \frac{1}{\ell} \text{Var}(\overline{X}_\ell) = \frac{1}{\ell(n-\ell+1)^2} \text{Var} \left( \sum_{j=0}^{n-\ell} S_j(\ell) \right).$$

From Lemma 2.4.2

$$\text{Var} \left( \sum_{j=0}^{n-\ell} S_j(\ell) \right) \leq 2A' \ell(n-\ell+1) \sum_{r=0}^{\lfloor \frac{n-\ell}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} (C(\ell, \ell) + 8) \sup_j \|S_j(\ell)\|_{2+\delta}^2, \quad (2.54)$$

where  $A'$  is a constant. Then

$$\frac{1}{\ell(n-\ell+1)^2} \text{Var} \left( \sum_{j=0}^{n-\ell} S_j(\ell) \right) \leq \frac{2A'}{(n-\ell+1)} \sum_{r=0}^{\lfloor \frac{n-\ell}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} (C(\ell, \ell) + 8) \sup_j \|S_j(\ell)\|_{2+\delta}^2.$$

The choice of  $\ell_n$  depends on the behavior of  $C(u, v)$  with respect to  $u$  and  $v$ . Take  $\ell = \ell_n$  and assume that

$$C(u, v) \leq K^{u+v}, \quad \ell = \left\lceil \frac{\ln(n)}{5 \ln(K)} \right\rceil + x, \quad (2.55)$$

where  $x$  is a fixed integer then in this case

$$C(\ell, \ell) \leq K^{2x} n^{\frac{2}{5}}. \quad (2.56)$$

Hence

$$\frac{2A'}{(n-\ell+1)} \sum_{r=0}^{\lfloor \frac{n-\ell}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} (K^{2x} n^{\frac{2}{5}} + 8) \sup_j \|S_j(\ell)\|_{2+\delta}^2 = O\left(\frac{1}{n^{\frac{3}{5}}}\right) = o(1).$$

Hence (2.53) holds. In order to prove that  $B_{n,2} \rightarrow \sigma$  as  $n \rightarrow \infty$  it suffices to prove that as  $n \rightarrow \infty$

$$\widehat{B}_{n,2} \rightarrow \sigma \text{ in } L^2. \quad (2.57)$$

Now we first prove that

$$E \left| \widehat{B}_{n,2}^2 - \mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2 \right| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (2.58)$$

Define

$$Z_{j,1} = \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 I \left\{ \frac{|S_j(\ell)|}{\sqrt{\ell}} \leq \left( \frac{n}{\ell} \right)^{1/8} \right\} \quad \text{and} \quad Z_{j,2} = \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 I \left\{ \frac{|S_j(\ell)|}{\sqrt{\ell}} > \left( \frac{n}{\ell} \right)^{1/8} \right\}.$$

We have

$$\begin{aligned} \mathbb{E} \left| \widehat{B}_{n,2}^2 - \mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2 \right| &= E \left| \frac{1}{n-\ell+1} \sum_{j=0}^{n-\ell} \frac{S_j^2(\ell)}{\ell} - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right| \\ &= \frac{1}{n-\ell+1} \mathbb{E} \left| \sum_{j=0}^{n-\ell} (Z_{j,1} - \mathbb{E}(Z_{j,1})) + \sum_{j=0}^{n-\ell} (Z_{j,2} - \mathbb{E}(Z_{j,2})) \right| \\ &\leq \frac{1}{n-\ell+1} \left( \mathbb{E} \left| \sum_{j=0}^{n-\ell} Z_{j,1} - \mathbb{E}(Z_{j,1}) \right| + \mathbb{E} \left| \sum_{j=0}^{n-\ell} Z_{j,2} - \mathbb{E}(Z_{j,2}) \right| \right) \\ &\leq K' n^{-1} \left( \left( \mathbb{E} \left| \sum_{j=0}^{n-\ell} Z_{j,1} - \mathbb{E}(Z_{j,1}) \right|^2 \right)^{\frac{1}{2}} + n \mathbb{E} |Z_{0,2}| \right) \\ &\leq K' \left( \left( n^{-2} \text{Var} \left( \sum_{j=0}^{n-\ell} Z_{j,1} \right) \right)^{\frac{1}{2}} + \mathbb{E} \frac{|S^2(\ell)|}{\ell} I \left\{ \frac{|S(\ell)|}{\sqrt{\ell}} > \left( \frac{n}{\ell} \right)^{1/8} \right\} \right). \end{aligned} \tag{2.59}$$

where  $K'$  is a constant. From Lemma 2.4.2 we obtain

$$n^{-2} \text{Var} \left( \sum_{j=0}^{n-\ell} Z_{j,1} \right) \leq \frac{2A'_1 \ell (n-\ell+1)}{n^2} (C(\ell, \ell) + 8) \sup_j \|Z_{j,1}\|_{2+\delta}^2 \sum_{r=0}^{\lfloor \frac{n-\ell}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} \tag{2.60}$$

where  $A'_1$  is a constant. Once again, with our choice of  $\ell = \lfloor \frac{\ln(n)}{5 \ln(K)} \rfloor + x$  if  $C(u, v) \leq K^{u+v}$  then

$$\frac{\ell}{n} C(\ell, \ell) \leq K^{2x} K^{\frac{2 \ln(n)}{5 \ln(K)}} \left( \frac{\ln(n)}{5 \ln(K)} + x \right) \frac{1}{n} \leq K^{2x} \frac{n^{\frac{2}{5}}}{n} \left( \frac{\ln(n)}{5 \ln(K)} + x \right) = \frac{\ln(n) K^{2x} + 5x K^{2x} \ln(K)}{5 \ln(K) n^{\frac{3}{5}}},$$

where as  $n \rightarrow \infty$ ,  $\ln(n)/n^{\frac{3}{5}} \rightarrow 0$ . Hence as  $n \rightarrow \infty$ ,

$$\frac{2A'_1 \ell (n-\ell+1)}{n^2} (C(\ell, \ell) + 8) = o(1).$$

The uniform integrability of  $\left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2$  implies that

$$\mathbb{E} \left\{ \left( \frac{|S(\ell)|}{\sqrt{\ell}} \right)^2 I \left\{ \frac{|S(\ell)|}{\sqrt{\ell}} > \left( \frac{n}{\ell} \right)^{1/8} \right\} \right\} \rightarrow 0, \tag{2.61}$$

as  $n \rightarrow \infty$ . Directly from equations (2.60) and (2.61) the relation (2.58) holds.

By the uniform integrability of  $\left\{ \frac{S^2(\ell)}{\ell} \right\}$  we have as  $R \rightarrow \infty$

$$\mathbb{E} \left\{ \frac{S^2(\ell)}{\ell} I_{\left| \frac{S(\ell)}{\sqrt{\ell}} \right| \leq R} \right\} \rightarrow \mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2,$$

so that from Remark 2.3.1 we obtain

$$\mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2 \rightarrow \sigma^2 \text{ as } \ell \rightarrow \infty. \quad (2.62)$$

From (2.58) it follows that

$$\mathbb{E} |\widehat{B}_{n,2}^2 - \sigma^2| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (2.63)$$

Finally, we use the fact that for any  $x, y \geq 0$ , for any  $p \geq 1$ .

$$|x - y|^2 \leq K' |x^p - y^p|^{\frac{2}{p}},$$

where  $K'$  is a constant, to get that

$$\widehat{B}_{n,2} \rightarrow \sigma \text{ in } L^2 \text{ as } n \rightarrow \infty. \quad (2.64)$$

Finally by (2.53) and (2.64) we have  $B_{n,2} \rightarrow \sigma$  in  $L^2$  as  $n \rightarrow \infty$ .  $\square$

**Remark 2.4.2.** Note that if  $C(u, v)$  is a polynomial function of  $u$  and  $v$  such that  $C(u, v) \leq K(u + v)^\beta$  and if we choose  $\ell = n^s$  such that  $s < \frac{1}{(1+\beta)}$  then the result of this proof holds.

In the proof of Theorem 2.3.3 we need the following lemma which is a variant of Lemma 2.4 of Shi (2000)[41] which is itself a variant of Theorem 2.1 from Peligrad and Utev (1997)[35].

**Lemma 2.4.3.** Let  $(X_n)_{n \geq 1}$  be  $\alpha_{(u,v)}^*$ -mixing sequence of r.v.'s for every  $1 \leq u \leq v \leq n$ . Assume that

$$\mathbb{E} |X_k|^2 < \infty \quad \inf_k \text{Var} X_k > 0, \quad \sum_{n=1}^{\infty} n^{\frac{2}{\delta}} \alpha^*(n) < \infty.$$

where  $\delta > 0$ . Consider  $r = o(n)$  and assume that  $\{|X_k|^{2+\delta}\}$  is a uniformly integrable family and

$$\text{Var} \left( \sum_{k=1}^n \sqrt{\frac{r}{n}} X_k \right) \rightarrow \sigma^2.$$

Then

$$\sum_{k=1}^n \sqrt{\frac{r}{n}} X_k \xrightarrow{d} N(0, \sigma^2), \quad (2.65)$$

as  $n \rightarrow \infty$ .

In the proof of Lemma 2.4.3 we need the following lemma which is a variant of Lemma 3.2 in Peligrad and Utev (1997)[35]. In this lemma we will use the notion of triangular array. We recall that a triangular array is defined as follows: the first row is a weakly increasing sequence of integers of length  $n$ , say  $a_1 \leq a_2 \leq \dots \leq a_n$ . The second row is a sequence of  $n - 1$  integers  $b_1 \leq b_2 \leq \dots \leq b_{n-1}$  between the first sequence, i.e., such that  $a_i \leq b_i \leq a_{i+1}$  and so on.

**Lemma 2.4.4.** *Let  $(X_n)_{n \in \mathbb{N}}$  be an  $\alpha^*_{(1,1)}$ -mixing sequence with  $\mathbb{E}|X_k|^{2+\delta} < \infty$  for some  $\delta > 0$ . Without loss of generality assume that  $E(X_k) = 0$ . Let  $\{a_{nk}; 1 \leq k \leq n\}$  be a triangular array of real number such that*

$$\sum_{k=1}^{\infty} a_{nk}^2 < \infty \text{ and } \max_{1 \leq k \leq n} |a_{nk}| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (2.66)$$

Assume that  $\sum_r \alpha^*(r)^{\frac{\delta}{2+\delta}} < \infty$  then

$$\text{Var} \left( \sum_{k=a}^b a_{nk} X_k \right) \leq C_1 \sum_{k=a}^b a_{nk}^2, \quad (2.67)$$

where  $C_1 = 2(C(1, 1) + 8) \sup_{i \in \mathbb{N}} \|X_i\|_{2+\delta}^2 \sum_{r=0}^{\infty} \alpha^*(r)^{\frac{\delta}{\delta+2}}$ .

*Proof of Lemma 2.4.4.* We have

$$\text{Var} \left( \sum_{k=a}^b a_{nk} X_k \right) = \sum_{k=a}^b a_{nk}^2 \text{Var}(X_k) + 2 \sum_{j=a}^{b-1} \sum_{k=j+1}^b a_{nk} a_{nj} \text{Cov}(X_k, X_j).$$

By the fact that  $\{a_{nk}; 1 \leq k \leq n\}$  is a triangular array we may write

$$\begin{aligned} & \sum_{k=a}^b a_{nk}^2 \text{Var}(X_k) + 2 \sum_{j=a}^{b-1} \sum_{k=j+1}^b a_{nk} a_{nj} \text{Cov}(X_k, X_j) \\ & \leq \sum_{k=a}^b a_{nk}^2 \text{Var}(X_k) + 2 \sum_{k=a}^b a_{nk}^2 \sum_{j=a, j \neq k}^b |\text{Cov}(X_k, X_j)| \\ & \leq 2 \sum_{k=a}^b a_{nk}^2 \sum_{k, j=a}^b |\text{Cov}(X_k, X_j)|, \end{aligned}$$

whence the result follows by Lemma 2.2.2.  $\square$

Now we provide the proof of Lemma 2.4.3

*Proof of Lemma 2.4.3.* We use a truncation technic. Let  $N$  be a constant, we write

$$X'_i = X_i I(|X_i| \leq N) - \mathbb{E}(X_i I(|X_i| \leq N)),$$

$$X''_i = X_i I(|X_i| > N) - \mathbb{E}(X_i I(|X_i| > N)),$$

From Lemma 2.4.4 and by the uniform integrability condition of  $|X_i|^{2+\delta}$  we have

$$\lim_{N \rightarrow \infty} \sup_n \text{Var} \left( \sum_{i=1}^n \sqrt{\frac{r}{n}} X''_{ni} \right) = 0.$$

so to prove this lemma it suffices to prove that

$$\frac{\sum_{i=1}^n \sqrt{\frac{r}{n}} X'_i}{\sqrt{\text{Var}(\sum_{i=1}^n \sqrt{\frac{r}{n}} X'_i)}} \xrightarrow{d} N(0, 1),$$

At this step we need Theorem B provided in Peligrad and Utev (1997) [35] which is itself derived by a careful analysis of Theorem 4.1 provided in Utev (1990)[42]. For convenience we restate Theorem B [35] in the Appendix. In our proof we should use the result of Theorem B [35] for  $u = (c - a + 1)^2$ . Condition (2.100) requires that the variance of the partial sum  $\sum_{k=1}^n \sqrt{\frac{r}{n}} X'_k$  is uniformly bounded which follows by Lemma 2.4.4. Condition (2.101) and the Lindberg condition (2.103) follows immediately from the uniform integrability of  $|X_i|^{2+\delta}$  and by the fact that  $r = o(n)$ . So it remains to verify that condition (2.102) holds. Following Peligrad and Utev (1997) [19] we have

$$\begin{aligned} & \left| \text{Cov} \left( \left[ \exp \left\{ it \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\} \right], \left[ \exp \left\{ it \sum_{j=b+u}^c \sqrt{\frac{r}{n}} X'_j \right\} \right] \right) \right| \\ &= \left| \text{Cov} \left( \left[ \exp \left\{ it \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\} - 1 \right], \left[ \exp \left\{ it \sum_{j=b+u}^c \sqrt{\frac{r}{n}} X'_j \right\} - 1 \right] \right) \right| \\ &\leq 4(C(b-a+1, c-b-u+1) + 8)t^2 \alpha^*(u)^{\frac{\delta}{\delta+1}} \left\| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\|_{2+2\delta} \left\| \sum_{j=b+u}^c \sqrt{\frac{r}{n}} X'_j \right\|_{2+2\delta}. \end{aligned}$$

By Lemma 2.4.4 and by Hölder inequality, we find

$$\begin{aligned} E \left| \sum_{i=a}^b \sqrt{\frac{r}{n}} X'_j \right|^{2+2\delta} &= E \left( \left| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right|^2 \left| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right|^{2\delta} \right) \\ &\leq E \left| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right|^2 \left\| \left| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right|^{2\delta} \right\|_{\infty} \\ &\leq C_1 \sum_a^b \frac{r}{n} \left( \sum_a^b \frac{r}{n} \right)^{\delta} (2N)^{2\delta} (b-a+1)^{\delta}. \end{aligned}$$

Hence

$$\left\| \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\|_{2+2\delta} \leq C_1^{\frac{1}{2(\delta+1)}} (b-a+1)^{\frac{\delta}{2(\delta+1)}} \left( \sum_a^b \frac{r}{n} \right)^{\frac{1}{2}} (2N)^{\frac{\delta}{\delta+1}}. \quad (2.68)$$

Thus

$$\begin{aligned}
 & \left| \text{Cov} \left( \left[ \exp \left\{ it \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\} \right], \left[ \exp \left\{ it \sum_{j=b+u}^c \sqrt{\frac{r}{n}} X'_j \right\} \right] \right) \right| \\
 & \leq 4(C(b-a+1, c-b-u+1) + 8) t^2 \alpha^*(u)^{\frac{\delta}{\delta+1}} (2N)^{\frac{2\delta}{1+\delta}} \times \\
 & (b-a+1)^{\frac{\delta}{2(\delta+1)}} (c-b-u+1)^{\frac{\delta}{2(\delta+1)}} \left( \sum_a^c \frac{r}{n} \right) \\
 & = 4(C(b-a+1, c-b-u+1) + 8) t^2 \alpha^*(u)^{\frac{\delta}{\delta+1}} C_1^{\frac{1}{1+\delta}} (2N)^{\frac{2\delta}{1+\delta}} \times \\
 & (c-a+1)^{\frac{2\delta}{(\delta+1)}} (C_1(c-a+1))^{\frac{1}{\delta+1}} \frac{r}{n} \\
 & = \left( \text{Var} \sum_{j=a}^c X'_j \right)^{\frac{1}{\delta+1}} \times 4(C(b-a+1, c-b-u+1) + 8) t^2 \times \\
 & (\alpha^*(u)(c-a+1)^2)^{\frac{\delta}{\delta+1}} (2N)^{\frac{2\delta}{1+\delta}} \frac{r}{n}.
 \end{aligned}$$

Denote by  $h_t(u) = t^2 (C(b-a+1, c-b-u+1) + 8) \times (\alpha^*(u)(c-a+1)^2)^{\frac{\delta}{\delta+1}}$ . We must prove that  $\sum_i h_t(2^i) < \infty$ . The result depends on the form of  $C(u, v)$ . Take the example where  $C(u, v) \leq K^{u+v}$ . From previous we have that  $\alpha(u) \leq Cu^{-\frac{\delta+2}{\delta}}$ . Take  $u = (c-a+1)^2$  then in one hand we have

$$(\alpha^*(u) \times u)^{\frac{\delta}{1+\delta}} \leq Cu^{\frac{-2}{1+\delta}}$$

and on the other hand we have

$$C(b-a+1, c-b-u+1) \leq K^{c-a+1+1-u} = K^{\sqrt{u}-u+1}.$$

If  $u = 2^i$  then in this case  $h_t(2^i) \leq t^2 (K^{\sqrt{2^i}-2^i+1} + 8) C 2^{i \frac{-2}{1+\delta}}$ , whence  $\sum_i h_t(2^i) < \infty$ . Thus

$$\begin{aligned}
 & \left| \text{Cov} \left( \left[ \exp \left\{ it \sum_{j=a}^b \sqrt{\frac{r}{n}} X'_j \right\} \right], \left[ \exp \left\{ it \sum_{j=b+u}^c \sqrt{\frac{r}{n}} X'_j \right\} \right] \right) \right| \times (2N)^{\frac{2\delta}{1+\delta}} \\
 & \leq \left( \text{Var} \sum_{j=a}^c X'_j \right)^{\frac{1}{\delta+1}} 4(2 \times K^{\sqrt{u}-u+1} + 8) t^2 (2N)^{\frac{2\delta}{1+\delta}} C u^{\frac{-2}{1+\delta}},
 \end{aligned}$$

where  $\sum_i 2^{\frac{-2}{1+\delta}} < \infty$  and  $\sum h_t(2^i) < \infty$ . This concludes the proof.

□

**Remark 2.4.3.** Note that if  $C(u, v) \leq K(u+v)^\beta$  i.e.  $C(u, v)$  is a polynomial function of  $u$  and  $v$  and if  $\alpha^*(u) \leq Cu^{-\frac{4(\delta+2)(1+\delta)}{\delta}}$  then (2.102) still holds. We remark that this condition made on  $\alpha^*(u)$  is not very restrictive. In fact Theorem 2.3.3 holds under the uniform integrability condition (2.30) which itself holds under this condition made on  $\alpha^*(u)$ .

*Proof of Theorem 2.3.3.* Assume without loss of generality that  $\mathbb{E}(X_1) = 0$ ,  $r = o(n)$  and  $\ell = o(r)$ . Define

$$s^2 = \text{Var} \left( \sqrt{\frac{n}{\ell}} \left( \widehat{B}_{n,2}^2 - \mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2 \right) \right), \quad (2.69)$$

$$\begin{aligned}\xi_{m,n} &= \sum_{j=m(2\ell+r)}^{m(2\ell+r)+r-1} \left( \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right), \\ \eta_{m,n} &= \sum_{j=m(2\ell+r)+r}^{(m+1)(2\ell+r)-1} \left( \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right),\end{aligned}$$

where  $m = 0, 1, \dots, k_n := \lfloor (n - \ell + 1)/(2\ell + r) \rfloor - 1$ ,  $\xi$  is a partial sum of  $r$  terms and  $\eta$  is a partial sum of  $2\ell$  terms. It is easy to see that

$$\begin{aligned}\sqrt{\frac{n}{\ell}} \left( \widehat{B}_{n,2}^2 - \mathbb{E} \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2 \right) &= \sqrt{\frac{n}{\ell}} \frac{1}{(n - \ell + 1)} \sum_{j=0}^{n-\ell} \left( \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right) \\ &= \sqrt{\frac{n}{\ell}} \frac{1}{(n - \ell + 1)} \sum_{m=0}^{k_n} \xi_{m,n} + \sqrt{\frac{n}{\ell}} \frac{1}{(n - \ell + 1)} \sum_{m=0}^{k_n} \eta_{m,n} \\ &+ \sqrt{\frac{n}{\ell}} \frac{1}{(n - \ell + 1)} \sum_{j=(k_n+1)(2\ell+r)}^{n-\ell} \left( \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right) \\ &= J_{1,n} + J_{2,n} + J_{3,n}.\end{aligned}$$

In this case by Lemma 2.4.2 we have

$$\begin{aligned}\text{Var}(J_{3,n}) &= \frac{n}{\ell} \frac{1}{(n - \ell + 1)^2} \text{Var} \left( \sum_{j=(k_n+1)(2\ell+r)}^{n-\ell} \left( \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 - \mathbb{E} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right) \right) \\ &\leq \frac{2A'_2 n(2\ell + r)}{(n - \ell + 1)^2} (C(\ell, \ell) + 8) \sup_{j \in \mathbb{N}} \left\| \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right\|_{2+\delta}^2 \sum_{r=1}^{\lfloor \frac{n}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}},\end{aligned}$$

where  $A'_2$  is a constant. Also by Lemma 2.4.2 as  $k_n \leq \frac{n}{r}$  we have

$$\begin{aligned}\text{Var}(J_{2,n}) &= \frac{n}{(n - \ell + 1)^2 \ell} \text{Var} \left( \sum_{m=0}^{k_n} \eta_{m,n} \right) \\ &\leq \frac{n}{(n - \ell + 1)^2 \ell} \text{Var} \left( \sum_{m=0}^{k_n} \eta_{m,n} \right) \\ &\leq \frac{n}{(n - \ell + 1)^2 \ell} \frac{n}{r} \text{Var}(\eta_{m,n}) \\ &\leq \frac{n}{(n - \ell + 1)^2 \ell} \frac{n}{r} \text{Var} \left( \sum_{j=m(2\ell+r)+r}^{(m+1)(2\ell+r)-1} \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right) \\ &\leq \frac{2\ell n A'_3 (C(\ell, \ell) + 8)}{(n - \ell + 1)^2 \ell} \frac{n\ell}{r} \sup_{j \in \mathbb{N}} \left\| \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right\|_{2+\delta}^2 \sum_{r=1}^{\lfloor \frac{n}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}} \\ &= \frac{2\ell n^2 A'_3 (C(\ell, \ell) + 8)}{r(n - \ell + 1)^2} \sup_{j \in \mathbb{N}} \left\| \left( \frac{S_j(\ell)}{\sqrt{\ell}} \right)^2 \right\|_{2+\delta}^2 \sum_{r=1}^{\lfloor \frac{n}{\ell} \rfloor + 1} \alpha^*(r)^{\frac{\delta}{\delta+2}},\end{aligned}$$

where  $A'_3$  is a constant. Assume that (2.55) holds and consider  $r = \sqrt{n}$ . By this choice of  $r$  and by (2.56), we have, as  $n \rightarrow \infty$ ,  $\frac{(2\ell+r)}{n}C(\ell, \ell) \leq K^{2x} \frac{(2\ell+r)}{n} n^{\frac{2}{5}} = o(1)$  and  $\frac{2\ell}{r}C(\ell, \ell) \leq K^{2x} \frac{2\ell}{r} n^{\frac{2}{5}} = o(1)$ .

Remark that if  $C(u, v) \leq K(u+v)^\beta$  and  $\ell = n^s$  then  $\text{Var}(J_{2,n}) = o(1)$  and  $\text{Var}(J_{3,n}) = o(1)$  if and only if  $s < \frac{1}{4}$ . Hence as  $n \rightarrow \infty$ ,  $\text{Var}\left(\sqrt{\frac{n}{\ell}}\left(\widehat{B}_{n,2}^2 - \mathbb{E}\left(\frac{S(\ell)}{\sqrt{\ell}}\right)^2\right)\right)$  is equal to  $\text{Var}(J_{1,n})$ . Therefore to complete the proof of Theorem 2.3.3 it remains to show that

$$\text{Var}(J_{1,n}) \xrightarrow{d} N\left(0, \frac{4\sigma^2}{3}\right), \quad (2.70)$$

or equivalently,

$$\frac{1}{\sqrt{n\ell}} \sum_{m=0}^{k_n} \xi_{m,n} \xrightarrow{d} N\left(0, \frac{4\sigma^2}{3}\right). \quad (2.71)$$

We have

$$\frac{1}{\sqrt{n\ell}} \sum_{m=0}^{k_n} \xi_{m,n} = \sqrt{\frac{r}{n}} \sqrt{\frac{1}{r\ell}} \sum_{m=0}^{k_n} \xi_{m,n} = \sqrt{\frac{r}{n}} \sum_{m=0}^{k_n} \zeta_{m,n}, \quad (2.72)$$

where we denote by  $\zeta_{m,n} = \sqrt{\frac{1}{r\ell}} \xi_{m,n}$ . Assume that

$$\left|S_j(\ell)/\sqrt{\ell}\right|^{4+2\delta} \text{ is U.I.} \quad (2.73)$$

In addition with our choice of  $\ell$  and  $r$  we have

$$\left(\frac{r}{\ell}\right)^{1+\frac{\delta}{2}} \mathbb{E} \left| \left(\frac{S(\ell)}{\sqrt{\ell}}\right)^2 \right|^{2+\delta} I \left\{ \left(\frac{S(\ell)}{\sqrt{\ell}}\right)^2 \geq \frac{1}{2} \sqrt{\frac{r}{\ell}} \right\} = o(1). \quad (2.74)$$

By stationarity and (2.74) we have

$$\begin{aligned} \sup_m \mathbb{E} |\zeta_{m,n}|^{2+\delta} I \left\{ |\zeta_{m,n}| > \frac{r}{\ell} \right\} &= \sup_m \mathbb{E} \left| \frac{1}{\sqrt{r\ell}} \xi_{m,n} \right|^{2+\delta} \cdot I \left\{ |\xi_{m,n}|^{2+\delta} > \left(\frac{r}{\ell} \sqrt{r\ell}\right)^{2+\delta} \right\} \\ &= (r\ell)^{-\frac{2+\delta}{2}} \mathbb{E} |\xi_{0,n}|^{2+\delta} I \left\{ |\xi_{0,n}|^{2+\delta} > \left(\frac{r^3}{\ell}\right)^{\frac{2+\delta}{2}} \right\} \\ &\leq 2^{\delta+1} \left(\frac{r}{\ell}\right)^{1+\frac{\delta}{2}} \mathbb{E} \left| \left(\frac{S(\ell)}{\sqrt{\ell}}\right)^2 \right|^{2+\delta} I \left\{ \left(\frac{S(\ell)}{\sqrt{\ell}}\right)^2 \geq \frac{1}{2} \sqrt{\frac{r}{\ell}} \right\} \\ &= o(1) \text{ as } n \rightarrow \infty. \end{aligned} \quad (2.75)$$

which means that  $\left\{|\zeta_{m,n}|^{2+\delta}\right\}$  is uniformly integrable. Now we will prove that

$$\text{Var} \left( \sqrt{\frac{r}{n}} \sum_{m=0}^{k_n} \zeta_{m,n} \right) \rightarrow \frac{4\sigma^2}{3}. \quad (2.76)$$

For this we can use the same proof as in Peligrad and Shao (1995) [33]. We have

$$\text{Var} \left[ \sqrt{\frac{r}{n}} \sum_{m=0}^{k_n} \zeta_{m,n} \right] = \frac{r}{n} \sum_{m=0}^{k_n} \mathbb{E}(\zeta_{m,n}^2) + 2 \sum_{i=0}^{k_n-1} \sum_{j=i+1}^{k_n} \mathbb{E} \left( \sqrt{\frac{r}{n}} \zeta_{j,n} \sqrt{\frac{r}{n}} \zeta_{i,n} \right). \quad (2.77)$$

For  $r$  large enough we have  $\alpha^*(r)^{\frac{\delta}{\delta+2}} \leq \frac{1}{r}$ , then as  $\alpha^*(r) \leq 1$  and as  $\delta/(\delta+2) \leq 1$  we have  $\alpha^*(r)^{\frac{\delta}{\delta+2}} \leq \alpha^*(t)^{\frac{\delta}{\delta+2}} \forall t > r$ . We have  $\forall j > i$

$$\begin{aligned} \mathbb{E}\left(\sqrt{\frac{r}{n}}\zeta_{i,n}\sqrt{\frac{r}{n}}\zeta_{j,n}\right) &= \frac{1}{n\ell}\mathbb{E}(\xi_{i,n}\xi_{j,n})\frac{1}{n\ell^3}\text{Cov}\left(\sum_{k=i(2\ell+r)}^{i(2\ell+r)+r-1}S_k^2(\ell),\sum_{s=j(2\ell+r)}^{j(2\ell+r)+r-1}S_s^2(\ell)\right) \\ &\leq \frac{A'_4}{n\ell^3}\sum_{k=i(2\ell+r)}^{i(2\ell+r)+r-1}\sum_{s=j(2\ell+r)}^{j(2\ell+r)+r-1}(C(\ell,\ell)+8)\|S_0^2(\ell)\|_{2+\delta}^2\alpha^*(s-k)^{\frac{\delta}{\delta+2}}. \end{aligned} \quad (2.78)$$

By  $\alpha^*_{(\ell,\ell)}$ -mixing condition, (2.3) and with our choice of  $\ell = \lceil \ln(n)/5 \ln(K) \rceil + x$  we obtain

$$\begin{aligned} \mathbb{E}\left(\sqrt{\frac{r}{n}}\zeta_{i,n}\sqrt{\frac{r}{n}}\zeta_{j,n}\right) &\leq \frac{A'_4\left(K^{2x}n^{\frac{2}{5}}+8\right)}{n\ell^3}\sum_{k=i(2\ell+r)}^{i(2\ell+r)+r-1}\sum_{s=j(2\ell+r)}^{j(2\ell+r)+r-1}\|S_0^2(\ell)\|_{2+\delta}^2\alpha^*(s-k)^{\frac{\delta}{\delta+2}} \\ &\leq \frac{A'_4\left(K^{2x}n^{\frac{2}{5}}+8\right)r^2}{n\ell^3}\|S_0^2(\ell)\|_{2+\delta}^2\alpha^*\left((j-i)(2\ell+r)+r\right)^{\frac{\delta}{\delta+2}} \\ &\leq \frac{A'_4\left(K^{2x}n^{\frac{2}{5}}+8\right)r^2}{n\ell^3}\|S_0^2(\ell)\|_{2+\delta}^2\alpha^*\left((j-i+1)(2\ell+r)\right)^{\frac{\delta}{\delta+2}}, \end{aligned} \quad (2.79)$$

where  $A'_4$  is a constant, with our choice of  $r = \sqrt{n}$  we have

$$\mathbb{E}\left(\sqrt{\frac{r}{n}}\zeta_{i,n}\sqrt{\frac{r}{n}}\zeta_{j,n}\right) \leq \frac{A'_4\left(n^{\frac{2}{5}}+8\right)}{\ell^3}\|S_0^2(\ell)\|_{2+\delta}^2\alpha^*\left((j-i+1)(2\ell+r)\right)^{\frac{\delta}{\delta+2}}. \quad (2.80)$$

Hence as  $n \rightarrow \infty$

$$\begin{aligned} &\lim\sum_{i=0}^{k_n-1}\sum_{j=i+1}^{k_n}\mathbb{E}\left(\sqrt{\frac{r}{n}}\zeta_{j,n}\sqrt{\frac{r}{n}}\zeta_{i,n}\right) \\ &= \lim\frac{A'_4\left(n^{\frac{2}{5}}+8\right)}{\ell^3}\|S_0^2(\ell)\|_{2+\delta}^2\sum_{i=0}^{k_n-1}\sum_{j=i+1}^{k_n}\alpha^*\left((j-i+1)(2\ell+r)\right)^{\frac{\delta}{\delta+2}} \\ &= \lim\frac{A_4n^{\frac{2}{5}}}{\ell^3}\sum_{i=0}^{k_n-1}\sum_{j=i+1}^{k_n}\alpha^*\left((j-i+1)(2\ell+r)\right)^{\frac{\delta}{\delta+2}}. \end{aligned} \quad (2.81)$$

where  $A_4$  is a constant. Consider

$$\sum_{j=i+1}^{k_n}\alpha^*\left((j-i+1)(2\ell+r)\right)^{\frac{\delta}{\delta+2}} = \sum_{w=2}^{k_n-i-1}\alpha^*(w(2\ell+r))^{\frac{\delta}{\delta+2}}.$$

Note that the assumption of uniform integrability (2.30) implies that  $\alpha^*(j) \leq \frac{C''}{j^\theta}$  where  $\theta > \frac{(4+2\delta)(4+3\delta)}{2\delta}$  (see Remark 2.3) and where  $C''$  is a constant. Hence

$$\sum_{w=2}^{k_n-i-1}\alpha^*(w(2\ell+r))^{\frac{\delta}{\delta+2}} \leq (2\ell+r)^{-\theta\delta/(\delta+2)}\sum_{w=2}^{k_n-i-1}\frac{1}{w^{\frac{\theta\delta}{\delta+2}}} \leq (2\ell+r)^{-\theta\delta/(\delta+2)}C''$$

where  $C'$  is a constant. Then

$$\sum_{i=0}^{k_n-1} \frac{n^{\frac{2}{5}}}{\ell^3(2\ell+r)^{\theta\delta/(\delta+2)}} \leq \frac{n^{\frac{2}{5}+1}}{r\ell^3 n^{\theta\delta/(\delta+2)}} = \frac{n^{\frac{2}{5}+\frac{1}{2}}}{\ell^3 n^{\theta\delta/(\delta+2)}} = \frac{n^{\frac{9}{10}}}{\ell^3 n^{\theta\delta/(2(\delta+2))}},$$

such that  $\frac{\theta\delta}{(2(\delta+2))} > \frac{9}{10}$ . Then as  $n \rightarrow \infty$  we conclude that by the uniform integrability condition (2.73) and by our choice of  $r$  and  $\ell$ , the following equation holds

$$2 \sum_{i=0}^{k_n-1} \sum_{j=i+1}^{k_n} \mathbb{E}(\sqrt{\frac{r}{n}} \zeta_{i,n} \sqrt{\frac{r}{n}} \zeta_{j,n}) = o(1). \quad (2.82)$$

**Remark 2.4.4.** Note that if  $C(u, v) \leq K(u+v)^\beta$  i.e.  $C(u, v)$  is polynomial in  $u$  and  $v$  and if  $\ell = n^s$  then by the same reasoning (2.82) holds.

Therefore (2.77) and (2.82) imply that  $\frac{r}{n} \sum_{m=0}^{k_n} \mathbb{E}(\zeta_{m,n}^2) = \frac{r}{n} \sum_{m=0}^{k_n} \mathbb{E}(\zeta_{0,n}^2)$  may be an approximation of  $\text{Var}[\sqrt{\frac{r}{n}} \sum_{m=0}^{k_n} \zeta_{m,n}]$  as  $n \rightarrow \infty$ . Recall that by our choice of  $\ell$  and  $r$  we have  $\ell = o(r)$ . As  $r = o(n)$  it follows that  $k_n = [(n-\ell+1)/(2\ell+r)] = o(n/r)$ . Then it follows that as  $n \rightarrow \infty$ ,  $\frac{r}{n} \sum_{m=0}^{k_n} \mathbb{E}(\zeta_{0,n}^2) = \mathbb{E}(\zeta_{0,n}^2)$ .

We have

$$\begin{aligned} \mathbb{E}(\zeta_{0,n}^2) &= \mathbb{E} \left( \frac{1}{\sqrt{r\ell}} \sum_{j=0}^{r-1} \frac{S_j^2(\ell)}{\sqrt{\ell}} - \mathbb{E} \left( \frac{S_j^2(\ell)}{\sqrt{\ell}} \right) \right)^2 \\ &= \frac{1}{r\ell^3} \sum_{j=0}^{r-1} \text{Var}(S_j^2(\ell)) + \frac{2}{r\ell^3} \sum_{j=0}^{r-2} \sum_{i=j+1}^{r-1} \text{Cov}(S_j^2(\ell), S_i^2(\ell)) \\ &= \underbrace{\frac{1}{\ell^3} \text{Var}(S^2(\ell))}_{T_{1,n}} + \underbrace{\frac{2}{r\ell^3} \sum_{j=r-\ell}^{r-2} \sum_{i=j+1}^{r-1} \text{Cov}(S_j^2(\ell), S_i^2(\ell))}_{T_{2,n}} + \\ &\quad \underbrace{\frac{2}{r\ell^3} \sum_{j=0}^{r-\ell-1} \sum_{i=j+1}^{j+\ell-1} \text{Cov}(S_j^2(\ell), S_i^2(\ell))}_{T_{3,n}} + \underbrace{\frac{2}{r\ell^3} \sum_{j=0}^{r-\ell-1} \sum_{i=j+\ell}^{r-1} \text{Cov}(S_j^2(\ell), S_i^2(\ell))}_{T_{4,n}} \\ &= T_{1,n} + T_{2,n} + T_{3,n} + T_{4,n}. \end{aligned} \quad (2.83)$$

With our choices of  $\ell$  and  $r$  it is easy to see that  $T_{1,n} \rightarrow 0$  as  $\ell \rightarrow \infty$ . By Lemma 2.2.2 we have

$$T_{2,n} \leq \frac{AC(\ell, \ell)}{r\ell^2} \rightarrow 0 \text{ as } \ell \rightarrow \infty.$$

where  $A$  is a constant. We also have

$$\begin{aligned} T_{4,n} &\leq \frac{2}{r\ell^3} \sum_{j=\ell}^{r-1} (j-\ell+1) \alpha^*(j) (C(\ell, \ell) + 8) \|S(\ell)\|_{2+\delta}^2 \\ &\leq \frac{A * C(\ell, \ell)}{r\ell^3} \sum_{j=\ell}^{r-1} j \alpha^*(j), \end{aligned} \quad (2.84)$$

where by (2.25) we have  $\sum_{j=\ell}^{r-1} j\alpha^*(j) = o(1)$  then

$$\frac{A * C(\ell, \ell)}{r\ell^3} \sum_{j=\ell}^{r-1} j\alpha^*(j) \rightarrow 0 \text{ as } \ell \rightarrow \infty.$$

Note that

$$T_{3,n} \approx 2 \frac{1}{\ell} \sum_{i=1}^{\ell-1} \text{Cov} \left( \left( \frac{S_0(\ell)}{\sqrt{\ell}} \right)^2, \left( \frac{S_i(\ell)}{\sqrt{\ell}} \right)^2 \right).$$

In this step we need Lemma 2.6.2 given in the Appendix. This lemma is another writing of Lemma 2.5 of Shi (2000)[41] which itself derived by a careful analysis of Theorem 1.2 in Peligrad and Shao (1995) [33]. By Lemma 2.6.2 we get

$$\lim_{\ell \rightarrow \infty} 2 \frac{1}{\ell} \sum_{i=1}^{\ell-1} \text{Cov} \left( \left( \frac{S_0(\ell)}{\sqrt{\ell}} \right)^2, \left( \frac{S_i(\ell)}{\sqrt{\ell}} \right)^2 \right) = \frac{4\sigma^4}{3}. \quad (2.85)$$

Finally as it was proved in Shi (2000) [41] we have

$$\frac{\widehat{B}_{n,2} - \mathbb{E} \left( \left| \frac{S(\ell) - \ell\mu}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}}}{\widehat{B}_{n,2}^2 - \mathbb{E} \left( \left( \frac{S(\ell) - \ell\mu}{\sqrt{\ell}} \right)^2 \right)} \rightarrow \frac{1}{2\sigma^2} \text{ in probability.} \quad (2.86)$$

and by (2.105) we get

$$\frac{n}{\ell} \text{Var} \left( \widehat{B}_{n,2} - \mathbb{E} \left( \left| \frac{S(\ell) - \ell\mu}{\sqrt{\ell}} \right|^2 \right)^{\frac{1}{2}} \right) \rightarrow \frac{\sigma^2}{3}. \quad (2.87)$$

In our case we replace the sum  $\sum_{j=a}^b \sqrt{\frac{\ell}{n}} X_j$  in Lemma 2.4.3 by  $\sqrt{\frac{\ell}{n}} \sum_{m=0}^{k_n} \zeta_{m,n}$ . We have  $\alpha_\zeta^*(k) \leq \alpha^*((k-1)(2\ell+r) + \ell)$ , where  $\alpha_\zeta^*(k)$  denotes the mixing coefficients of random sequence  $\{\zeta_{n,m}, n \geq 1\}$ , whence by (2.25) one get  $\sum_{n=1}^{\infty} n^{\frac{2}{3}} \alpha_\zeta^*(n) < \infty$ . In addition by the uniform integrability condition of  $\zeta_{m,n}$  (see (2.75)) and by (2.105) we get  $\sqrt{\frac{\ell}{n}} \sum_{m=0}^{k_n} \zeta_{m,n} \xrightarrow{d} N(0, 4\sigma^2/3)$ . Finally the result of Theorem 2.3.3 is a consequence of (2.87) and of (2.65).  $\square$

*Proof of Corollary 2.3.1.* Note that, in this proof we follow Peligrad and Shao (1995)[33] and Shi (2000)[41]. Without loss of generality, assume that  $\mu = 0$ . Recall that

$$\bar{X}_\ell = \frac{1}{(n-\ell+1)} \sum_{j=0}^{n-\ell} S_j(\ell).$$

To prove this corollary it suffices to prove that

$$\sqrt{\frac{n}{\ell}} \mathbb{E} \left| \widehat{B}_{n,2}^2 - B_{n,2}^2 \right| \rightarrow 0$$

as  $n$  goes to infinity. We have

$$\widehat{B}_{n,2}^2 - B_{n,2}^2 = \frac{1}{(n-\ell+1)} \sum_{j=0}^{n-\ell} \frac{\overline{X}_\ell^2}{\ell} - \frac{2\overline{X}_\ell S_j(\ell)}{\ell}.$$

Then

$$\mathbb{E} \left| \frac{1}{(n-\ell+1)} \sum_{j=0}^{n-\ell} \frac{\overline{X}_\ell^2}{\ell} - \frac{2\overline{X}_\ell S_j(\ell)}{\ell} \right| \leq \underbrace{\frac{\text{Var}(\overline{X}_\ell)}{\ell}}_{W_1} + \underbrace{\frac{2}{\ell(n-\ell+1)^2} \sum_{j=0}^{n-\ell} \sum_{k=0}^{n-\ell} |\text{Cov}(S_j(\ell), S_k(\ell))|}_{W_2},$$

We first focus on  $W_1$ . We have

$$\text{Var}(\overline{X}_\ell) = \text{Var} \left( \frac{1}{(n-\ell+1)} \sum_{j=0}^{n-\ell} S_j(\ell) \right),$$

then by Lemma 2.4.2 we obtain

$$\frac{\text{Var}(\overline{X}_\ell)}{\ell} = O\left(\frac{n * \ell^2 * \ell * C(\ell, \ell)}{\ell(n-\ell+1)^2}\right).$$

Under our condition made on  $\ell$  we have  $C(\ell, \ell) \leq K^{2x} n^{\frac{2}{5}}$ . Then

$$O\left(\frac{n * \ell^2 * \ell * C(\ell, \ell)}{\ell(n-\ell+1)^2}\right) = O\left(\frac{\ell^2}{n^{\frac{3}{5}}}\right).$$

As  $n$  goes to infity we obtaion

$$\frac{\ell^2}{n^{\frac{3}{5}}} \sqrt{\frac{n}{\ell}} = \frac{\ell^{3/2}}{n^{\frac{1}{10}}} \rightarrow 0.$$

hence  $W_1 \rightarrow 0$  as  $n \rightarrow \infty$ . Now we focus on  $W_2$  where

$$W_2 = \frac{2}{(n-\ell+1)^2 \ell} \sum_{j=0}^{n-\ell} \text{Var}(S_j(\ell)) + \frac{2}{(n-\ell+1)^2 \ell} \sum_{j=0}^{n-\ell} \sum_{k \neq j}^{n-\ell} \text{Cov}(S_j(\ell) S_k(\ell)).$$

We have

$$\begin{aligned} \sum_{j=0}^{n-\ell} \sum_{k \neq j}^{n-\ell} \text{Cov}(S_j(\ell) S_k(\ell)) &= 2 \sum_{j=0}^{n-\ell-1} \sum_{k=j+1}^{n-\ell} \text{Cov}(S_j(\ell), S_k(\ell)) \\ &\leq 2 \sum_{r=1}^{n-\ell-1} (n-\ell-r) |\text{Cov}(S_0(\ell), S_r(\ell))| \\ &\leq 2(n-\ell) \sum_{r=1}^{n-\ell-1} \sum_{k=1}^{\ell} \sum_{s=1+r}^{\ell+r} |\text{Cov}(X_k, X_s)|. \end{aligned}$$

Where

$$\begin{aligned}
 \sum_{r=1}^{n-\ell-1} \sum_{k=1}^{\ell} \sum_{s=1+r}^{\ell+r} |\text{Cov}(X_k, X_s)| &= \sum_{r=1}^{2\ell-1} \sum_{k=1}^{\ell} \sum_{s=1+r}^{\ell+r} |\text{Cov}(X_k, X_s)| + \sum_{r=2\ell}^{n-\ell-1} \sum_{k=1}^{\ell} \sum_{s=1+r}^{\ell+r} |\text{Cov}(X_k, X_s)| \\
 &\leq 2(2\ell-1)\ell^2 \sup_i \|X_i\|_2^2 + \sum_{r=2\ell}^{n-\ell-1} \sum_{j=r+1-\ell}^r |\text{Cov}(X_0, X_j)| (j-r+\ell) + \sum_{r=2\ell}^{n-\ell-1} \sum_{j=r+1}^{r+\ell-1} |\text{Cov}(X_0, X_j)| (\ell+r-j) \\
 &\leq 2(2\ell-1)\ell^2 \sup_i \|X_i\|_2^2 + \ell \sum_{r=2\ell}^{n-\ell-1} \sum_{j=r+1-\ell}^{r+\ell-1} |\text{Cov}(X_0, X_j)|,
 \end{aligned}$$

such that

$$\sum_{r=2\ell}^{n-\ell-1} \sum_{j=r+1-\ell}^{r+\ell-1} |\text{Cov}(X_0, X_j)| \leq 2K_2(\ell-1)(C(1,1) + 8) \sup_i \|X_i\|_{2+\delta}^2 \sum_{r=2\ell}^{n-\ell-1} \alpha^*(r+1-\ell)^{\frac{\delta}{2+\delta}}.$$

Also we have

$$\begin{aligned}
 \text{Var}(S_j(\ell)) &\leq \ell \text{Var}(X_1) + 2 \sum_{r=1}^{\ell-1} (\ell-r) |\text{Cov}(X_1, X_{1+r})| \\
 &\leq 2 \sum_{r=0}^{\ell-1} (\ell-r) |\text{Cov}(X_1, X_{1+r})| \\
 &\leq 4\ell(C(1,1) + 8) \sup_i \|X_i\|_{2+\delta}^2 \sum_{r=0}^{\ell-1} \alpha^*(r)^{\frac{\delta}{\delta+2}}.
 \end{aligned}$$

Then

$$\begin{aligned}
 &\frac{2}{(n-\ell+1)^2 \ell} \sum_{j=0}^{n-\ell} \sum_{k \neq j}^{n-\ell} |\text{Cov}(S_j(\ell), S_k(\ell))| \\
 &\leq \frac{2}{(n-\ell+1)^2 \ell} 2(n-\ell) \\
 &\times \left\{ 2(2\ell-1)\ell^2 \sup_i \|X_i\|_2^2 + 2K_2(\ell-1)\ell \sum_{r=2\ell}^{n-\ell} (C(1,1) + 8) \sup_i \|X_i\|_{2+\delta}^2 \alpha^*(r+1-\ell)^{\frac{\delta}{2+\delta}} \right\} \\
 &= O\left(\frac{\ell^2}{n}\right).
 \end{aligned} \tag{2.88}$$

Then we obtain,

$$\sqrt{\frac{n}{\ell}} \times W_2 = O\left(\frac{1}{\sqrt{n}\sqrt{\ell}} + \frac{\ell^{3/2}}{\sqrt{n}}\right). \tag{2.89}$$

which goes to zero as n goes to infity. This concludes the proof.  $\square$

## 2.5 Asymptotic stationarity and conditionally independent processes

Our results can be used for sequences of r.v's converging toward a stationary state. Assume that  $(X_n)_{(n \in \mathbb{N})}$  is a sequence of asymptotically stationary r.v's. Roughly speaking

there is a stationary sequence  $(Y_n)_{(n \in \mathbb{N})}$  such that for all bounded function  $g : \mathbb{R}^k \rightarrow \mathbb{R}$  and  $n \in \mathbb{N}$  we have

$$\mathbb{E}(g(X_n, \dots, X_{n+k})) \xrightarrow{n \rightarrow \infty} \mathbb{E}(g(Y_0, \dots, Y_k)). \quad (2.90)$$

In applications we are interested by sequences of conditionally independent processes given the entire memory of the factor. Consider that r.v's  $X_n$  are conditionally independent given the unbounded memory of the factor denoted by  $\underline{V}_n$ , where  $\underline{V}_n = (V_1, \dots, V_n)$ . We have

$$\begin{aligned} \mathbb{E}(g(X_n, \dots, X_{n+k})) &= \mathbb{E}(\mathbb{E}(g(X_n, \dots, X_{n+k}) | \underline{V}_{n+k})) \\ &= \int_{\underline{V}_{n+k}} \mathbb{E}(g(X_n, \dots, X_{n+k}) | \underline{V}_{n+k} = \underline{v}_{n+k}) dF_{\underline{V}_{n+k}}(\underline{v}_{n+k}) \end{aligned}$$

where

$$\mathbb{E}(g(X_n, \dots, X_{n+k}) | \underline{V}_{n+k} = \underline{v}_{n+k}) = \int_{x_n, \dots, x_{n+k}} g(x_n, \dots, x_{n+k}) \prod_{j=n}^{n+k} N^j(dx_j, \underline{v}_j) \quad (2.91)$$

where  $\prod_{j=n}^{n+k} N^j(dx_j, \underline{V}_j)$  denotes the transition kernel of  $(X_n, \dots, X_{n+k})$  given  $\underline{V}_{n+k}$ . Thus, (2.90) may be rewritten as it follows by (2.92)

$$\begin{aligned} &\int_{x_n, \dots, x_{n+k}} g(x_n, \dots, x_{n+k}) \left[ \prod_{j=0}^k \int_{\underline{V}_{n+j}} N^{n+j}(\cdot, \underline{v}_{n+j}) dF_{\underline{V}_{n+k}}(\underline{v}_{n+k}) \right] dx_n, \dots, dx_{n+k} \\ &\xrightarrow{n \rightarrow \infty} \int_{y_0, \dots, y_k} g(y_0, \dots, y_k) \left[ \prod_{j=0}^k \int_{\underline{V}_\infty} N(\cdot, \underline{v}_\infty) dF_{\underline{V}_\infty}(\underline{v}_\infty) \right] dy_0, \dots, dy_k. \end{aligned} \quad (2.92)$$

In our context,  $(N^{n+j}(\cdot, \underline{V}_{n+j}))$  (resp.  $N^j(\cdot, \underline{V}_\infty)$ ), corresponds to the conditional law of  $X_n$  (resp.  $Y_j$ ) given  $\underline{V}_{n+j}$  (resp.  $\underline{V}_\infty$ ).

Let  $\mathcal{N}_{n+j}(\cdot) = \int_{\underline{V}_{n+j}} N^{n+j}(\cdot, \underline{v}_{n+j}) dF_{\underline{V}_{n+j}}(\underline{v}_{n+j}) = \mathbb{E}_{\underline{V}_{n+j}}(N^{n+j}(\cdot, \underline{V}_{n+j}))$  be a new measure of probability and recall that  $g : \mathbb{R}^k \rightarrow \mathbb{R}$  is a bounded function. In this case if (2.92) holds, then we say that  $\mathcal{N}_{n+j}(\cdot)$  converges weakly to the measure of probability  $\mathcal{N}_\infty(\cdot)$  where  $\mathcal{N}_\infty(\cdot) = \int_{\underline{V}_\infty} N(\cdot, \underline{V}_\infty) dF_{\underline{V}_\infty}(\underline{v}_\infty) = \mathbb{E}_{\underline{V}_\infty}(N(\cdot, \underline{V}_\infty))$ .

Thus, it follows that to insure (2.90), it is sufficient that the kernel transition  $(N^n(\cdot, \underline{V}_n))$ , goes to the measure  $N(\cdot, \underline{V}_\infty)$  as  $n \rightarrow \infty$ .

## 2.6 Applications

First of all, note that processes that we consider in the following examples are asymptotically stationary. We shall use results of Section 2.3 by considering that the stationary regime has been reached. A sufficient condition in our examples that insures the asymptotic stationarity of the conditional law and then of the non conditional law is to choose a stationary factor.

### 2.6.1 Example in a discrete case

Consider the sequence of r.v's  $(I_i)_{i \in \mathbb{N}}$  such that for all  $i \in \mathbb{N}$ ,  $I_i$  are conditionally independent with respect to the vector of the factor  $\underline{V}_i$  where  $\underline{V}_i = (V_1, \dots, V_i)$ . We consider that

the sequence  $(V_i)_{i \in \mathbb{N}}$  is a sequence of independent and identically distributed Bernoulli r.v's with parameter  $q$ .

We shall assume that the conditional law of  $I_i$  has the following structure:

$$\theta_i = \mathbb{P}(I_i = 1 | \underline{V}_i = \underline{v}_i) = K \sum_{j=1}^i \frac{1 + v_j}{2^{i-j}}, \quad (2.93)$$

where  $K$  is a constant. In this case we shall say that the conditioning is with respect to the unbounded memory of the factor.

For this example we have proved in Chapter 1 that the sequence  $(I_i)_{i \in \mathbb{N}}$  satisfies the mixing condition (2.1).

As we know the law of a Bernoulli r.v  $I_i$  is defined by its parameter  $p_i = \mathbb{P}(I_i = 1)$ . Hence to prove the asymptotic stationarity of the sequences  $(I_i)_{i \geq 1}$ , it is sufficient to prove the asymptotic stationarity of the sequence  $(\mathbb{E}[\Theta_i])_{i \geq 1}$ , where  $\Theta_i = \mathbb{P}(I_i = 1 | \underline{V}_i) = K \sum_{j=1}^i \frac{1+V_j}{2^{i-j}}$ . In fact,  $\mathbb{P}(I_i = 1) = \mathbb{E}_{\underline{V}_i}[\mathbb{P}(I_i = 1 | \underline{V}_i)] = \mathbb{E}[\Theta_i]$ . The sequence of the factor is stationary then  $\forall i \geq 1$  we have  $\mathbb{E}(V_i) = \mathbb{E}(V_1)$ . Hence we have

$$\mathbb{P}(I_i = 1) = K(1 + \mathbb{E}(V_1)) \sum_{j=1}^i \frac{1}{2^{i-j}} = K(1 + \mathbb{E}(V_1)) \frac{2^i - 1}{2^{i-1}}.$$

When  $i$  goes to infinity, it is easy to see that

$$\mathbb{P}(I_i = 1) \rightarrow 2K(1 + \mathbb{E}(V_1)).$$

In order to give some comparisons between our model and the classical one i.e. with model for which the conditioning is done with respect to the last state of the factor, we shall consider the following conditional law

$$\mathbb{P}(I_i = 1 | V_i = v_i) = K(1 + v_i), \quad (2.94)$$

and we shall say that the conditioning is of order one.

When the conditioning is with respect to the unbounded memory of the factor we obtain the following result for the asymptotic value of the variance of  $S_n$  where  $S_n = \sum_{i=1}^n I_i$ .

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(S_n)}{n} = \sigma^2 = \frac{8}{3} K^2 q(1 - q) + 2K(1 + q - 2K - 4Kq - 2Kq^2). \quad (2.95)$$

As  $\mathbb{E}(S_n) = \sum_{i=1}^n \mathbb{E}(I_i) = \sum_{i=1}^n \mathbb{E}(I_i) = 2K(1 + q)(n - 1 + \frac{1}{2^n})$  we have

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(S_n)}{n} = \mu = 2K(1 + q). \quad (2.96)$$

In the case where the conditioning is of order one i.e. model given by (2.94) we obtain

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(S_n)}{n} = \sigma^2 = K(1 + q)(1 - K(1 + q)), \quad (2.97)$$

and

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(S_n)}{n} = \mu = K(1+q). \quad (2.98)$$

(for the proofs of (2.97) and (2.95) see Appendix A.5). Note that, in models (2.94) and (2.93), r.v.'s  $I_i$  are positively correlated ( $\text{Cov}(I_i, I_j) > 0$  for all  $i \neq j$ ). We mention that for fixed  $i$ , the covariance function  $\text{Cov}(I_i, I_j)_{i \neq j}$  is a decreasing function of  $j$ .

Also, we note that the covariance function is an increasing function of the variability of the factor. In fact, in our case the variance of the factor is equal to  $q(1-q)$  and  $\frac{\partial \text{Cov}(I_i, I_j)}{\partial q(1-q)} > 0$  (see Appendix B.2). This is in adequation with the result of Frostig and Denuit (2008)[11] in which the authors mention that in model with one factor, the more the factor is variable the more the  $I_i$  are positively dependent.

- **Comparison between the limit standard deviations: Exact evaluation**

In order to give some comparison between the value of the limit variance derived for models (2.94) and (2.93), we must have equality between their asymptotic stationary laws. From (2.96) and (2.98), it is easy to see that this condition is true if  $K'(1+q) = 2K(1+q)$  where  $K'$  is a constant. This implies that  $K' = 2K$ .

	<i>Model with <math>(I_i V_i)</math></i>	<i>Model with <math>(I_i \underline{V}_i)</math></i>
Standard deviation: $\sigma$	0.4221	0.4637

Table 2.1: Exact values of  $\sigma$  in the Bernoulli discrete model,  $K = 0.24$ ,  $q = 0.6$ .

We mention that by the notation  $(I_i|V_i)$  (resp.  $(I_i|\underline{V}_i)$ ) we mean that the r.v.  $I_i$  is conditioned with respect to  $V_i$  (resp.  $\underline{V}_i$ ).

From Table 2.1 we see that the standard deviation for models with unbounded memory is larger than the one in model with a memory of the factor of length one.

- **Estimation of the variance and asymptotic normality of the normalized sum**

Consider the model in which the conditionnal law is defined by (2.93). To illustrate the convergence of the law of the normalized sum to a normal law we made some Monte Carlo simulations. We denote by  $n$  the number of periods and by  $m$  the number of Monte Carlo simulations. In Figure 2.1 we present results of simulations that confirm the asymptotic normality.

Note that the test of normality *Jarque-Bera* confirms the asymptotic normality of the normalized sum. Denote by *AMSE* the absolute mean squared error and by *RMSE* the relative mean squared error. Put  $\bar{B}_{n,2} = \frac{1}{m} \sum_{i=1}^m B_{n,2}^i$ . For  $n = 200000$  we choose  $x = 445$  and for  $n = 50000$  we choose  $x = 220$ .

The results of Monte Carlo simulations presented in Table 2.2 show the convergence of the estimated standard deviation derived by our estimator  $B_{n,2}$  to the real value of  $\sigma$  where  $\sigma$  is the limit standard deviation.

- **Comparison between the Value at Risk (VaR) and the Conditional Tail Expectation (CTE)**

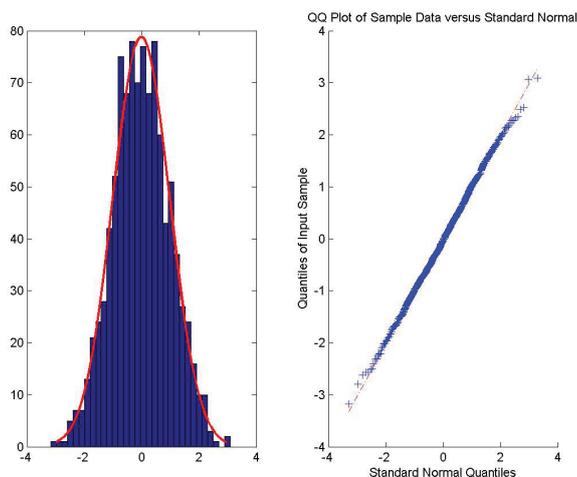


Figure 2.1: Normfit and QQplot of the normalized sum  $S_n$ : the case of Bernoulli sample data with  $m = 1000$ ,  $n = 20000$ ,  $q = 0.6$ ,  $x = 60$ .

n	m	$\bar{B}_{n,2}$	AMSE	RMSE
200000	20	0,4816	0,000263	0,001223
	50	0,4607	0,000148	0,000689
	100	0,4621	<b>0,000135</b>	<b>0,000627</b>
50000	200	0,4627	0,000262	0,001217
	300	0,4615	0,000261	0,001213
	400	0,4624	0,000258	0,001198

Table 2.2: Standard deviations estimated in the Bernoulli discrete model,  $K = 0.24$ ,  $q = 0.6$ .

In this paragraph we consider two risk measures: the VaR and CTE. These measures are very used in risk theory. Note that for a r.v.  $X$ ,  $\text{VaR}_\alpha(X)$  is the value at Risk of  $X$  at level  $\alpha$  where  $0 \leq \alpha < 1$ . By definition  $\text{VaR}_\alpha(X) = F_X^{-1}(\alpha)$  where if  $F_X$  is the distribution function of  $X$  then  $F_X^{-1}$  is its quantile function. At a level  $0 \leq \alpha < 1$ , the CTE is defined by  $\mathbb{E}(X|X > \text{VaR}_\alpha(X))$ .

Next we present three methods to evaluate the VaR and CTE. The first one is based on a normal approximation method using the exact values of  $\sigma$  and  $\mu$  (i.e. (2.98), (2.97), (2.96) and (2.95)). The second one is based on a normal approximation method using the estimation of  $\sigma$  derived by our estimator. The third one is based on a direct evaluation from the sampled data.

- Normal approximation method using the exact value of  $\sigma$  and  $\mu$

In order to evaluate the VaR and CTE in models (2.94) and (2.93) we use an approximation method. In fact as  $n \rightarrow \infty$ , we have  $F(S_n) \approx N(n\mu, n\sigma^2)$ . where  $\mathbb{E}(S_n) = n\mu$  and  $\text{Var}(S_n) = n\sigma^2$ . Let  $Y$  be a r.v. with a standard normal distribution. Recall that if a r.v.  $X$  has a normal distribution function with parameter  $\mu$  and  $\sigma^2$  (i.e.  $X = \sigma Y + \mu$ ) then, at a level  $0 \leq \alpha < 1$ ,  $\text{VaR}_\alpha(X) = \sigma \text{VaR}_\alpha(Y) + \mu$  and  $\text{CTE}_\alpha(X) = \sigma \text{CTE}_\alpha(Y) + \mu$ .

Take  $n = 50000$ ,  $q = 0.6$  and  $K = 0.24$ . In this case for both models we have

$$E(S_n) = 50000 * 2 * 0.24 * (1 + 0.6) = 38400.$$

If we consider model with  $(I_i|V_i)$  then

$$\text{Var}(S_n) = 50000 * 0.4637^2 = 10750,884.$$

If we consider model with  $(I_i|V_i)$  then

$$\text{Var}(S_n) = 50000 * 0.4221^2 = 8908,420.$$

In Table 2.3 we indicate the values of VaR and CTE for different levels of  $\alpha$ .

	<i>Model with <math>(I_i V_i)</math></i>		<i>Model with <math>(I_i \underline{V}_i)</math></i>	
$\alpha$	$\text{VaR}_\alpha(S_n)$	$\text{CTE}_\alpha(S_n)$	$\text{VaR}_\alpha(S_n)$	$\text{CTE}_\alpha(S_n)$
0.75	38463,616	38519,965	38469,885	38531,785
0.95	38555,266	38594,719	38570,564	38613,905
0.975	38584,998	38620,676	38603,225	38642,419
0.99	38619,543	38651,540	38641,175	38676,324
0.995	38643,134	38672,959	38667,096	38699,861

Table 2.3: VaR and CTE in the Bernoulli discrete model: normal approximation method using the exact value of  $\sigma$  and  $\mu$ ,  $n = 50000$ ,  $K = 0.24$ ,  $q = 0.6$ .

From results of Table 2.3, we remark that, for all level  $\alpha \geq 0.75$ , the VaR and CTE in model with unbounded memory (i.e. (2.93)) have higher values than those derived in the classical model (i.e.(2.94)). In risk theory we are interested by quantile of higher orders. We conclude that in this example the classical model subestimate the risk and it is interesting to consider all the memory of the factor.

#### – Normal approximation using the estimation of $\sigma$

We consider that  $F(S_n) \approx N(\mathbb{E}(S_n), n\bar{B}_{n,2}^2)$  where  $\mathbb{E}(S_n)$  is approximated by  $2K(1+q)n$  and  $\bar{B}_{n,2}^2$  is the variance of  $S_n$  estimated by our estimator. See Table (2.4).

	Model with $(I_i V_i)$	
	VaR $_{\alpha}(S_n)$	CTE $_{\alpha}(S_n)$
0.75	38468, 785	38530, 392
0.95	38568, 987	38640, 501
0.975	38601, 493	38674, 245
0.99	38639, 262	38697, 671
0.995	38665, 061	38697, 670

Table 2.4: VaR and CTE in the Bernoulli discrete model: normal approximation using the estimation of  $\sigma$  and the exact value of  $\mu$ ,  $n = 50000$ ,  $m = 300$ ,  $K = 0.24$ ,  $q = 0.6$ ,  $x = 220$ .

– **Simulation method**

We consider the simulated data  $(S_n^j)_{1 \leq j \leq m}$  where  $m$  is the number of Monte Carlo simulations and we estimate the VaR and CTE directly from the simulated data. See Table (2.5).

	Model with $(I_i V_i)$	
$\alpha$	VaR $_{\alpha}(S_n)$	CTE $_{\alpha}(S_n)$
0.75	38466	38523, 474
0.95	38554	38609, 875
0.975	38603	38651, 000
0.99	38624	38692, 500
0.995	38641	38715.33

Table 2.5: VaR and CTE in the Bernoulli discrete model: simulation method,  $n = 50000$ ,  $m = 300$ ,  $K = 0.24$ ,  $q = 0.6$ ,  $x = 220$ .

### 2.6.2 Examples in an absolutely continuous case

Consider the sequence of r.v's  $(X_i)_{i \in \mathbb{N}}$  such that for all  $i \in \mathbb{N}$ ,  $X_i = B_i$  are conditionally independent with respect to the vector of the factor  $V_i$  where  $V_i = (V_1, \dots, V_i)$ . In this subsection we consider that  $(X_i)_{i \in \mathbb{N}}$  are Pareto r.v's and that the sequence  $(V_i)_{i \in \mathbb{N}}$  is a sequence of independent and identically distributed Bernoulli r.v's with parameter  $q$ . Next we provide some results derived from a Monte Carlo simulations.

In this case we consider the same example given in Chapter 1, Subsection 1.4.2.1, on which the r.v.  $X_i$  given  $V_i$  has a Pareto distribution function with a shape parameter  $\beta$  and a scale parameter  $\theta_i$  where

$$\theta_i^\beta = \sum_{j=1}^i \frac{h(V_j)}{2^{i-j}}.$$

In this case we have proved that the sequence  $(X_i)_{i \in \mathbb{N}}$  is mixing and satisfies (2.1). As in the discrete case we shall choose  $h(V_j) = 1 + V_j$ . In this case

$$\theta_i^\beta = K \left( \sum_{j=1}^i \frac{1 + V_j}{2^{i-j}} \right).$$

We also consider the classical model in which  $X_i$  given  $V_i$  has a Pareto distribution function with scale parameter  $\theta_i$  where

$$\theta_i^\beta = h(V_j) = K(1 + V_j).$$

We recall that in order to give some comparisons between the classical model and our model, we assume that we have equality between the stationary laws of these two models. This may be obtained if we consider that

$$K \mathbb{E} \left( \sum_{j=1}^i \frac{1 + V_j}{2^{i-j}} \right) = K' \mathbb{E}(1 + V_j),$$

which means that  $K' = 2K$ . In the following we choose  $K = 1$ . Recall that to use our estimator r.v.'s  $X_i$  must respect some moment conditions. In fact, we need a moment of order  $4 + 3\delta$ . We choose  $\beta = 5$  and in a limit case we choose  $\beta = 4$ . In the case of our model, we note that the test of normality *Jarque-Bera* confirms the asymptotic normality of the normalized sum when  $\beta = 4$  (see Figure 2.2) and  $\beta = 5$  (see Figure 2.3).

In Table 2.6 we present values of the estimated standard deviations. In particular when

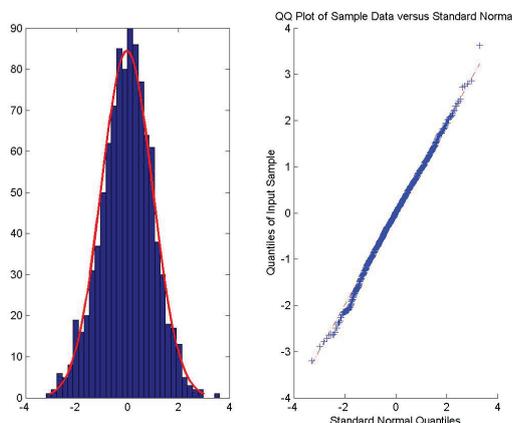


Figure 2.2: Normfit and QQplot of the normalized sum  $S_n$  in the Pareto model,  $m = 1000$ ,  $n = 10000$ ,  $q = 0.5$ ,  $x = 50$ ,  $\beta = 4$ .

the model is with an unbounded memory, we consider the particular case on which r.v.'s  $\theta_i^\alpha$ ,  $(i \in \mathbb{N})$  are independent. This allows us to derive the impact of the correlation between r.v.'s  $\theta_i^\alpha$ ,  $(i \in \mathbb{N})$  on the standard deviation of r.v.  $X_i$ . This can be obtained by regenerating in each time  $i$  all the vector  $\underline{V}_i$ . More clearly, in time  $i = 1$  we generate r.v.  $V_1$  and in time  $i = 2$  we not only generate r.v.  $V_2$  as in the dependent case but we also regenerate r.v.  $V_1$  and so on. Next we shall call the case on which  $\theta_i^\alpha$  are independent by the *Independent*

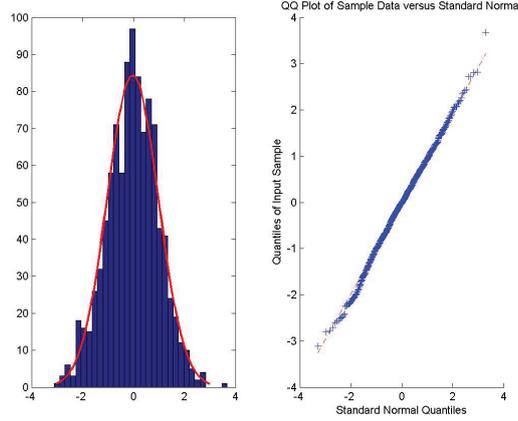


Figure 2.3: Normfit and QQplot of the normalized sum  $S_n$  in the Pareto model,  $m = 1000$ ,  $n = 10000$ ,  $q = 0.5$ ,  $x = 50$ ,  $\beta = 5$ .

case and the case on which  $\theta_i^\alpha$  are dependent by the *Dependent case*.

From Table 2.6, we remark that the standard deviation in model with unbounded memory of the factor, is higher than the one obtained in the classical model. From the two first columns, we see the effect of the correlation introduced by our long memory of the factor. We present in Table 2.7 the values of VaR and CTE evaluated by a normal approxima-

	<i>Model with <math>(X_i V_i)</math></i>		<i>Model with <math>(X_i V_i)</math></i>
	<i>Dependent case</i>	<i>Independent case</i>	
$\beta = 4$	0.6333	0.6226	0.6320
$\beta = 5$	0.4134	0.4039	0.4121

Table 2.6: Standard deviations in the Pareto model,  $m = 1000$ ,  $n = 10000$ ,  $q = 0.5$ ,  $x = 50$ .

tion method such that  $F(S_n) \approx N(\mathbb{E}(S_n), n\bar{B}_{n,2}^2)$ . These results confirm that, for a level  $\alpha \geq 0.75$ , a conditioning of order one subestimates the value of the risk in comparison with a conditioning with respect to the unbounded memory of the factor.

## Appendix

**Lemma 2.6.1.** *[[10], Lemma 4.6]*

Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of centered r.v.'s. For positive integer  $r$ , define the coefficients of weak dependence as the non-decreasing sequence  $(C_{r,q})_{(q \geq 2)}$  such that

$$C_{r,q} := \sup |Cov(X_{t_1} \times \dots \times X_{t_m}, X_{t_{m+1}} \times \dots \times X_{t_q})|,$$

	Model with $(X_i V_i)$		Model with $(X_i V_i)$	
$\alpha$	$\text{VaR}_\alpha(S_n)$	$\text{CTE}_\alpha(S_n)$	$\text{VaR}_\alpha(S_n)$	$\text{CTE}_\alpha(S_n)$
0.75	15453, 088	15478, 205	15550, 736	15575, 802
0.95	15493, 578	15512, 820	15591, 259	15610, 527
0.975	15511, 333	15525, 352	15607, 931	15622, 487
0.99	15525, 728	15538, 922	15621, 318	15636, 245
0.995	15534, 896	15547, 655	15633, 908	15645, 204

Table 2.7: VaR and CTE in the Pareto model,  $n = 10000$ ,  $m = 1000$ ,  $K = 0.24$ ,  $q = 0.5$ ,  $x = 50$ ,  $\beta = 5$ .

where the supremum is taken over the multi-indices  $1 \leq t_1 \leq \dots \leq t_q$  with  $t_{m+1} - t_m = r$ .

Define  $S_n = \sum_{i=1}^n X_i$ . Then

$$\mathbb{E}(S_n^2) \leq 2n \sum_{r=0}^{n-1} C_{r,2} \quad \mathbb{E}(S_n^4) \leq 4! \left[ n^2 \left( \sum_{r=0}^{n-1} C_{r,2} \right)^2 + n \sum_{r=0}^{n-1} (r+1)^2 C_{r,4} \right]. \quad (2.99)$$

**Theorem 2.6.1.** [[35], Theorem 4.1]

Let  $\{X_{ni}, 1 \leq i \leq k_n\}$  be a triangular array of r.v.'s such that the following conditions hold.

1. There is a constant  $C_1 > 0$  such that for every  $0 \leq a \leq b \leq k_n$ ,

$$\text{Var}\left(\sum_{j=a}^b X_{nj}\right) \leq C_1 \sum_{j=a}^b \text{Var}(X_{nj}), \quad (2.100)$$

- 2.

$$\liminf_{n \rightarrow \infty} \frac{\text{Var}\left(\sum_{j=1}^{k_n} X_{nj}\right)}{\sum_{j=1}^{k_n} \text{Var}(X_{nj})} > 0, \quad (2.101)$$

- 3.

$$\left| \text{Cov}\left(\exp(it \sum_{j=a}^b X_{nj}), \exp(it \sum_{j=b+u}^c X_{nj})\right) \right| \leq h_t(u) \sum_{j=a}^c \text{Var}(X_{nj}), \quad (2.102)$$

for every  $1 \leq a \leq b \leq c \leq k_n$  where  $h_t(u) \geq 0$ ,  $\sum h_t(2^i) < \infty$  and  $u$  is of the form  $u = [(c-a)^{1-\epsilon}]$  for  $0 < \epsilon < 1$ ;

- 4.

$$\frac{1}{\sigma_n^2} \sum_{i=1}^{k_n} \mathbb{E} X_{ni}^2 I(|X_{nj}| > \epsilon \sigma_n) \rightarrow 0, \quad (2.103)$$

as  $n \rightarrow \infty$  for every  $\epsilon > 0$ , where  $\sigma_n^2$  denotes  $\text{Var}(\sum_{i=1}^{k_n} X_{ni})$ . Then  $\frac{S_n}{\sigma_n} \xrightarrow{d} N(0, 1)$  as  $n \rightarrow \infty$ .

A careful analysis of the proof of Theorem 1.2 in Peligrad and Shao (1995) [33] gives the following statement.

**Lemma 2.6.2.** *Let  $(X_n)_{n \geq 1}$  be a stationary  $\alpha_{(u,v)}^*$ -mixing sequences for all  $(u, v) \in \mathbb{N}^*$  with  $\mathbb{E}(X_1) = 0$  and  $\mathbb{E}|X_i|^{4+3\delta} < \infty$ . Assume that  $D_{\infty,2,2} < \infty$  holds. If in addition (2.20) holds and if*

$$\left\{ \left| \frac{S_j(\ell)}{\sqrt{\ell}} \right|^{4+2\delta}, \ell \geq 1 \right\}, \quad (2.104)$$

is uniformly integrable, then

$$\begin{aligned} \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \sum_{i=1}^{\ell-1} \text{Cov} \left( \left( \frac{S(\ell)}{\sqrt{\ell}} \right)^2, \left( \frac{S_i(\ell)}{\sqrt{\ell}} \right)^2 \right) &= \int_0^1 \text{Cov} (\sigma^2 W(1)^2, \sigma^2 (W(1+t) - W(t))^2) dt \\ &= \frac{2\sigma^4}{3}, \end{aligned} \quad (2.105)$$

where  $(W(t), 0 \leq t \leq 1)$  is a standard wiener process and where  $\frac{\text{Var}(S(\ell))}{\sqrt{\ell}} \rightarrow \sigma^2$  as  $\ell \rightarrow \infty$ .

## Bibliography

- [1] Bernstein, S.N.: Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.* **97**, (1927), 1–59.
- [2] Blum, J.R., Hanson D.L. and Koopmans, L.H.: On the strong law of large number for a class of stochastic processes. *Z. Wahrsch. Verw. Gebiete*, no. **2**, (1963), 1–11.
- [3] Bradley R.C.: Basic properties of strong mixing conditions. *A survey and some open questions. Probability Surveys.* **2**, (2005), 107–144.
- [4] Bradley, R.C.: Basic Properties of strong mixing conditions. *In: E.Eberlein and M.S. Taqqu, eds., Dependence in Probability and Statistics. Progress Probab. Statist.* **11**, (1986), 165–192.
- [5] Chuanrong, L. and Zhengyan, L.: Limit theory for mixing dependent random variables. *Mathematics and its Applications.* **378**, (1996), xii+426.
- [6] Dalibor, V.: Approximation of stationary processes and the central limit problem. *Probability Theory and Mathematical Statistics, Lecture Notes in Mathematics.* **1299/1988**,(1988), 532–540.
- [7] Dalibor, V.: Approximating martingales and the central limit theorem for strictly stationary processes. *Stochastic Processes and their Applications.* **44**, (1993), 41–74.
- [8] Davydov, YU.A.: Convergence of distributions generated by stationary stochastic processe. *Theory of probability and its applications.* **13**, (1968), 691–696.
- [9] Davydov, YU.A.: The invariance principle for stationary processes. *Theory of probability and its applications.* **3**, (1970), 487–498.

- 
- [10] Dedecker, J., Doukhan, P., Lang, G., León R, J.R., Louhichi, S. and Prieur, C.: Weak dependence: with examples and applications. *Lecture Notes in Statistics, Springer, New York*. **190**, (2007).
- [11] Denuit, M. and Frostig, E.: Comparaison of dependence in factor models with application to credit risk portfolio. *Probability in the Engineering and informational sciences*. **22**, (2008), 151-160.
- [12] Doukhan, P.: Mixing properties and examples. *Lecture note in statistics*. **85**, (1994).
- [13] Doukhan, P. and Louhichi, S.: A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*. **84**, (1999), 313–342.
- [14] Doukhan, P., León R, J.R. and Jakubowicz, J.: Variance estimation with applications. *Dependence in probability, analysis and number theory Kendrick Press Heber City UT*. **1299**, (2010), 203–231.
- [15] Doukhan, P. and Witenberger, O.: An invariance principle for weakly dependent stationary general models. *Probability and Mathematical Statistics*. (2007), 45–73.
- [16] Gordin, M.I.: On the central limit theorem for stationary random processes. *Dokl. Akad. Nauk. SSSR*. **188**, (1969), 739-741.
- [17] Gordin, M.I.: A remark on the martingale method for proving the central limit theorem for stationary sequences. *Journal of Mathematical Sciences*. **133**, (2004).
- [18] Hall, P. and Heyde, C.C.: Martingale limit theory and its application. *New York: academic press*. (1980).
- [19] Hao, Yu. and Shao, Q.M.: Weak convergence for weighted empirical processes of dependent sequences. *The annal of probability*. **24**, (1996), 2098–2127.
- [20] Herrndorf, N.: A functional central limit theorem for strongly mixing sequences of random variables. *Gibiete*. **69**, (1985), 541–550.
- [21] Ibragimov, I.A.: Some limit theorems for stationary processes. *Theory of probability and its applications*. **VII**, (1962), 349–382.
- [22] Ibragimov I.A.: A note on the central limit theorem for dependent random variables. *Theory Of Probability and its applications*. **XX**, (1975), 135–141.
- [23] Ibragimov, I.A. and Linnik, Y.U.: Independent and stationary sequences of random variables. *Wolters-Noordhoff Publishing, Groningen*. (1971).
- [24] Kacem, M., Loisel, S., and Maume-Deschamps, V.: Some mixing properties of conditionally independent processes. *Submitted to communication in statistics: theory and method*. (2013).
- [25] Kolmogorov, A. and Rozanov, Y.A.: On a strong mixing condition for stationary gaussian processes. *Theory of probability and its applications*. **V**, (1960), 204–208.
- [26] Kim, T.Y.: Moment bounds for non-stationary dependent sequences. *Journal of applied probability*. **31**, (1994), 731–742.

- [27] Merlevède, F. and Peligrad, M. Utev, S.: Recent advances in invariance principles for stationary sequences. *Probability Surveys*. **3**, (2006), 1–36.
- [28] Mcleish, D.L.: Invariance principle for dependent variables. *Z. Wahrsch. Verw. Gebiete*. **32**, (1975), 165–178.
- [29] Obrein, G.L.: Limit theorems for sums of chain-dependent processes. *Journal of applied probability*. **11**, (1974), 582–587.
- [30] Peligrad, M.: Invariance Principle for mixing sequences of random variables. *The Annal of Probability*. **10**, (1982), 968–981.
- [31] Peligrad, M.: Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. In: *E.Eberlein and M.S. Taqqu, eds., Dependence in Probability and Statistics*.
- [32] Peligrad, M. and Shao, Q.M.: On the invariance principle for  $\rho$ - mixing sequences of random variables. *Chinese Ann. Math.Ser.B*. **10**, (1989), 427–433.
- [33] Peligrad, M. and Shao, Q.M.: Estimation of the variance of Partial Sums for  $\rho$ -mixing random variables. *Journal of Multivariate Analysis*. **52**, (1995), 140–157.
- [34] Peligrad, M. and Suresh, R. : Estimation of variance of partial sums of an associated sequence of random variables. *Stochastic processes and their applications*. (1995), 307–319.
- [35] Peligrad, M. and Utev, S.: Central limit theorem for linear processes. *The annals of probability*. **25**, (1997), 443–456.
- [36] Philipp, W. and Stout, W.F.: Almost sure invariance principles for partial sums of weakly dependent random variables. *Memoirs of the American Mathematical Society*. **2**, (1975), iv+140.
- [37] Revuz, D.: Markov chains. *Second edition. North-Holland Mathematical Library, 11. North-Holland Publishing Co., Amsterdam*, (1984).
- [38] Rio, E.: Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29**, (1993), 587–597.
- [39] Rio, E.: Inégalités de moments pour les suites stationnaires et fortement mélangeantes. *C.R. Acad. Sci. Paris, Série*. **10**, (1994), 355–360.
- [40] Rosenblatt, M.: A central limit theorem and a strong mixing condition. *Proceeding national academy of sciences*. **42**, (1956), 43–47.
- [41] Strongway, Shi.: Estimation of the variance for strongly mixing sequences. *Applied. Math.J.Chinese Univ.Ser.B*, (2000), **15**, 45–54.
- [42] Utev, S.: Central limit theorem for deendent random variables. *Probability theory and mathematical statistics*. **2**, (1990), 519–528.
- [43] Yokoyama, R.: Moment Bound for stationary mixing sequences. *Wahrscheinlichkeitstheorie* **52**, (1980), 45–57.

- [44] Young, W.C.: On classes of summable functions and their Fourier series. *Proc. Royal Soc. Lond.* **87**, (1912), 225–229.



## Part II

Extrema convexes pour des  
distributions discrètes décroissantes:  
effets de la contrainte de convexité



## Chapitre 3

# Extrema convexes pour des distributions discrètes décroissantes

## Résumé

Dans ce chapitre on considère la classe des distributions discrètes avec des fonctions de masse de probabilité qui sont décroissantes globalement sur  $\mathbb{N}$ . On considère que le support de ces distributions est inclus dans  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Les extrema convexes dans cette classe sont connus (voir Denuit *et al* (1999)). En assurance et en finance on fait face parfois à des situations où peu d'informations sont connues sur les distributions de certaines variables des modèles. On peut par exemple avoir à notre disposition uniquement des informations concernant les moments ou la forme de la distribution. Notre but est de montrer comment les contraintes de forme supplémentaires de type convexité modifient ces extrema. Deux cas sont considérés : la f.m.p. est globalement convexe dans  $\mathbb{N}$  ou elle est convexe uniquement à partir d'un point  $m$  fixé. Les extrema convexes correspondants sont calculés en utilisant une simple propriété de croisement entre les f.m.p. Plusieurs illustrations en théorie de la ruine sont présentées. Ces résultats apportent un complément à deux œuvres récentes de Lefèvre et Loisel (2010),(2012). Ce chapitre est un article en cours intitulé « Convex extrema for nonincreasing discrete distributions : effects of convexity constraints » co-écrit avec Claude Lefèvre et Stéphane Loisel.

### 3.1 Introduction

The convex order is a classical stochastic order that compares any two distributions with equal means. More precisely, a random variable (r.v.)  $X$  is smaller in the convex sense than a random variable  $Y$ , which is denoted by  $X \leq_{cx} Y$ , if

$$E[f(X)] \leq E[f(Y)] \text{ for all convex functions } f. \quad (3.1)$$

This implies that  $E(X) = E(Y)$  as desired, and  $Var(X) \leq Var(Y)$ , hence the convex order is sometimes named the variability order. Properties and applications of the convex order can be found e.g. in the books by Ross (1996)[15], Rolski (1976)[14], Müller and Stoyan (2002)[13], Goovaerts *et al.* (1990,1998)([7; 8]) and Shaked and Shanthikumar (1994,2007)([17; 16]).

The construction of convex extrema is an important theoretical problem with many practical implications. This problem has been investigated for discrete random variables by e.g. Lefèvre and Utev (1996)[12], Denuit and Lefèvre (1997)[2], Courtois *et al.* (2006)[1], Denuit *et al.* (1999a,1999c)[4; 6], Lefèvre and Picard (1993)[11] and for continuous random variables by e.g. Denuit *et al.* (1998,1999b)([3; 5]); see also the recent works by Lefèvre and Loisel (2010,2012)[9; 10].

The present paper is concerned with r.v.'s that are valued in a finite set  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ , for some fixed  $n \in \mathbb{N}^*$ . Let us recall that for such variables, a condition of convex ordering equivalent to (3.1) is that

$$E(X) = E(Y) \quad \text{and} \quad \sum_{i=k}^n F_X(i) \geq \sum_{i=k}^n F_Y(i), \quad 0 \leq k \leq n, \quad (3.2)$$

where  $F_X$  denotes the distribution function (d.f.) of  $X$ .

More precisely, our starting point is the special class of r.v.'s whose probability mass function (p.m.f.) is nonincreasing on a support  $\mathcal{D}_n \equiv \{0, 1, \dots, n\}$ . Distributions of that type are met in various probability models proposed in economics and biosciences. This is especially true in insurance for which the claim number distributions are often observed to be nonincreasing.

**Proposition 3.1.1** (Denuit *et al.* (1999c)[6]). *Let  $N$  be an arbitrary r.v. with a nonincreasing p.m.f. on  $\mathcal{D}_n$ . Denote  $\nu = E(N)$ . In this case  $N_{min} \leq_{cx} N \leq_{cx} N_{max}$  where*

$$N_{max} = \begin{cases} 0 & \text{with probab. } 1 - 2\nu/(n+1), \\ 1, \dots, n & \text{with equal probab. } 2\nu/n(n+1), \end{cases} \quad (3.3)$$

and, defining  $\xi$  as the integer on  $[0, n-1]$  satisfying  $\xi < 2\nu \leq \xi + 1$ ,

$$N_{min} = \begin{cases} 0, \dots, \xi & \text{with equal probab. } 2(\xi + 1 - \nu)/(\xi + 1)(\xi + 2), \\ \xi + 1 & \text{with probab. } (2\nu - \xi)/(\xi + 2). \end{cases} \quad (3.4)$$

These two extrema convex bounds are illustrated with a numerical example in Figure 3.1

Now, for certain situations, it may be possible to obtain further information on the shape of the p.m.f. In this paper, we make the assumption that the p.m.f. is not only nonincreasing on  $\mathcal{D}_n$ , but it is also convex, globally on  $\mathbb{N}$  or at least on some parts of  $\mathbb{N}$ .

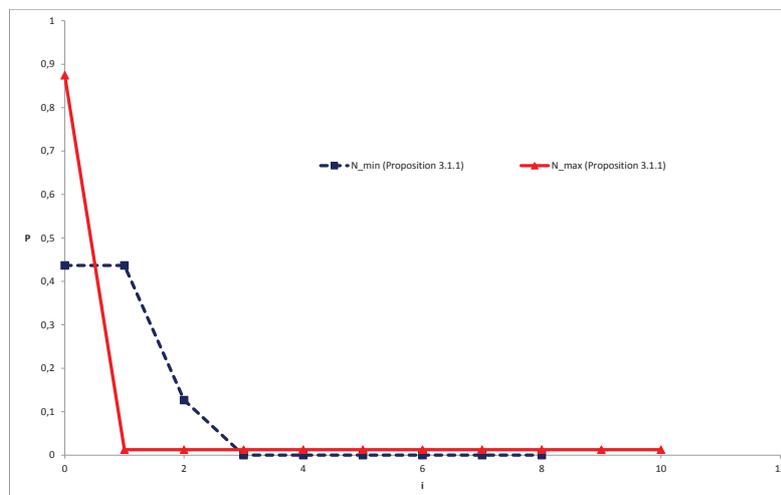


Figure 3.1: Convex extrema in the class of discrete r.v.'s with decreasing p.m.f. on a support  $\mathcal{D}_n$ ,  $\nu = 0.69$  and  $n = 10$ .

Our purpose is to point out how the convex extrema (3.3), (3.4) are modified under an additional convexity constraint.

The method of proof will exploit some simple crossing properties between two distributions that are briefly presented in Section 3.2. Three different cases are then discussed. Section 3.3 deals with the case the p.m.f. is assumed to be convex on  $\mathbb{N}$ . This situation has been studied in Lefèvre and Loisel (2012)[10], but in a more complex frameworks and using another approach that is based on the concept of multiple monotonicity. In Sections 3.4 and 3.5, convexity is assumed to hold only from a fixed point  $m = 1$  and  $m > 1$ , respectively. Finally, in Section 3.6, the influence of the choice of  $n$  and  $m$  on the extrema is discussed numerically, and several applications to some ruin problems are presented for illustration.

To simplify the presentation, positive (resp. negative) a function means below nonnegative (resp. nonpositive), and increasing (resp. decreasing) means nondecreasing (resp. nonincreasing).

## 3.2 Crossing properties and convex ordering

This section recalls some simple crossing results that will be used in the sequel. Although known, these results are presented with a short proof for the ease of presentation. Let  $S^-$  be the operator which, applied to a function  $f$ , counts the number of sign changes of  $f$  over its domain, zero terms being discarded. Two functions  $f$  and  $g$  cross each other  $k$  times,  $k \geq 0$ , if  $S^-(f - g) = k$ .

Let  $X$  and  $Y$  be two r.v.'s valued on  $\mathcal{D}_n$ , with (distinct) p.m.f.  $P_X, P_Y$  and distribution functions (d.f.)  $F_X, F_Y$ .

**Proposition 3.2.1.**

$$\text{If } E(X) = E(Y), \text{ then } S^-(P_Y - P_X) \geq 2. \quad (3.1)$$

**Proof.** As  $\sum_{i=0}^n [P_X(i) - P_Y(i)] = 0$ , the two p.m.f. have at least one crossing point. Suppose that there is a single crossing. Then, one of the two r.v.'s has necessarily a larger mean, which is in contradiction with the assumption made.  $\diamond$

**Lemma 3.2.2.**

$$\text{If } S^-(P_Y - P_X) = 2, \text{ then } S^-(F_Y - F_X) = 1. \quad (3.2)$$

**Proof.**  $S^-(P_Y - P_X) = 2$  means that the function  $P_Y - P_X$  has opposite signs on three consecutive intervals  $I_1, I_2, I_3$  of  $\mathcal{D}_n$ . Suppose that the sequence of signs is  $+, -, +$ , for instance. Then, the function  $F_Y - F_X$  is positive increasing on  $I_1$ , decreasing on  $I_2$ , and negative increasing on  $I_3$  since  $F_Y(n) = F_X(n) = 1$ . Thus,  $F_Y - F_X$  has one sign change on  $I_2$ . This is illustrated in Figure 3.2.  $\diamond$

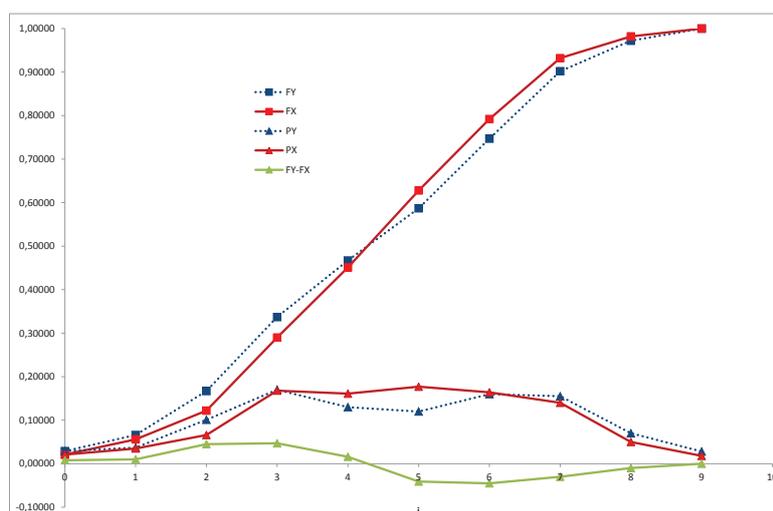


Figure 3.2: An example where  $S^-(P_Y - P_X) = 2$  and  $S^-(F_Y - F_X) = 1$ .

We are now ready to derive a crossing type condition that implies a convex ordering. This result is not new (see e.g. Denuit and Lefèvre (1997)).

**Proposition 3.2.3.**

$$\text{If } E(X) = E(Y) \text{ and } S^-(P_Y - P_X) = 2 \text{ with } P_Y \geq P_X \text{ near } n, \text{ then } X \leq_{cx} Y. \quad (3.3)$$

**Proof.** The two means being equal,  $P_Y - P_X$  has at least two sign changes by Proposition 3.2.1. Thus, the assumption that there are exactly two sign changes is admissible. By Lemma 3.2.2,  $F_Y - F_X$  has then one sign change. Moreover, as  $P_Y \geq P_X$  near  $n$ , one has  $F_Y \leq F_X$  near  $n$ , so that the consecutive signs of  $F_Y - F_X$  are  $+, -$ .

Now,  $E(X) = E(Y)$  is equivalent to

$$\sum_{i=0}^n F_Y(i) = \sum_{i=0}^n F_X(i).$$

For any integer  $k \in [0, n - 1]$ , this can be rewritten as

$$\sum_{i=k+1}^n [F_Y(i) - F_X(i)] = - \sum_{i=0}^k [F_Y(i) - F_X(i)]. \quad (3.4)$$

If  $k$  is before the sign change of  $F_Y - F_X$ , then  $F_Y(i) - F_X(i) \geq 0$  for  $i = 0, \dots, k$ , so that the left hand side of (3.4) is negative i.e.

$$\sum_{i=k+1}^n F_X(i) \geq \sum_{i=k+1}^n F_Y(i). \quad (3.5)$$

If  $k$  is after the sign change of  $F_Y - F_X$ , then  $F_Y(i) - F_X(i) \leq 0$  for  $i = k + 1, \dots, n$ , so that (3.5) holds too. In other words, the condition (3.2) is satisfied, hence  $X \leq_{cx} Y$ .  $\diamond$   
 Roughly speaking, Proposition 3.2.3 states that  $X$  is smaller than  $Y$  in the convex order if the p.m.f. of  $Y$  is heavier on the extremes (near 0 and  $n$ ).

### 3.3 For decreasing p.m.f. that are convex on $\mathbb{N}$

Consider the class of r.v.'s which have discrete p.m.f. Our aim is to derive explicit expressions of convex extrema in this class of r.v.'s under constraint of global decreasingness and convexity on  $\mathbb{N}$ . We start by introducing the following lemma which will be useful in some following proofs.

**Lemma 3.3.1.** *Consider two discrete r.v.'s  $X$  and  $Y$  valued on  $\mathcal{D}_n$  with globally decreasing and convex p.m.f. on  $[s, \infty)$  where  $s \in [1, n]$ . In particular we consider that the p.m.f. of  $X$  in  $[s, n + 1]$  is linear. If we have  $j \in [s, n]$  such that  $P(Y = j) \leq P(X = j)$ , then by the decreasingness and the convexity we have*

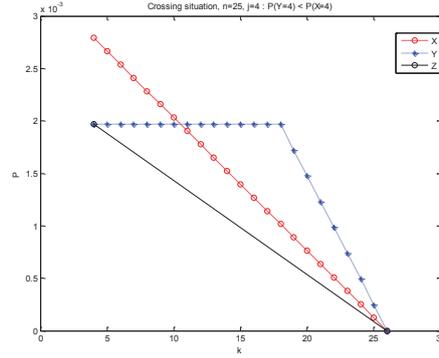
$$P(Y = k) \leq P(X = k), \quad \forall k > j.$$

**Proof.** Consider two r.v.'s as defined in Lemma 3.3.1 and recall that by global decreasingness  $P(Y = n + 1) = P(X = n + 1) = 0$ . Let  $j$  and  $k$  be two positive integers such that  $j \in [s, n]$  and  $k > j$ . In addition assume that  $P(Y = j) \leq P(X = j)$  and  $P(Y = k) \geq P(X = k)$ . It is clear that, on  $[j, n]$ , the p.m.f. of  $Y$  is above the segment of line  $Z$  (see Figure 3.3). Hence we conclude that  $Y$  is not convex which is absurd.  $\diamond$

Denote by  $\mathbb{J}_n$  the set of all r.v.'s with discrete distribution function valued on  $\mathcal{D}_n$  (i.e.  $P(X = j) = 0 \quad \forall j \geq n + 1$ ), with decreasing p.m.f. and fixed mean  $\nu$ .

#### 3.3.1 The upper bound

Let  $M$  and  $\tilde{N}_{\max}$  be two arbitrary r.v.'s in  $\mathbb{J}_n$  that have globally convex p.m.f. on  $\mathbb{N}$  such that  $Var(M) \neq Var(\tilde{N}_{\max})$ . Using a reasoning by contradiction, we prove that the convex


 Figure 3.3: Crossing situation,  $j = s = 4$  and  $n = 25$ .

maximum bound is attained for the r.v.  $\tilde{N}_{\max}$  such that

$$\begin{cases} P(\tilde{N}_{\max} = 0) = a, \\ P(\tilde{N}_{\max} = i) = b(n+1-i), \forall 1 \leq i \leq n. \end{cases} \quad (3.6)$$

where  $a$  and  $b$  are defined such that (3.6) is a true p.m.f., with fixed mean  $\nu$  i.e.

$$\begin{cases} a + \sum_{i=1}^n b(n+1-i) = 1, \\ \sum_{i=1}^n b(n+1-i)i = \nu, \end{cases} \quad (3.7)$$

hence,

$$\begin{cases} b = (6\nu) / (n(n+1)(n+2)), \\ a = 1 - bn(n+1)/2. \end{cases} \quad (3.8)$$

In this case it is possible to have explicit expression for  $a$  and  $b$  in function of  $\nu$  and  $n$ . The p.m.f. of  $\tilde{N}_{\max}$  is then as defined in the following proposition:

**Proposition 3.3.2.** *The convex maximum in the class of r.v.'s with discrete distribution that are with globally decreasing and convex p.m.f. on  $\mathbb{N}$  is attained for the r.v.  $\tilde{N}_{\max}$  where for  $0 < \nu \leq \frac{n}{3}$*

$$\begin{cases} P(\tilde{N}_{\max} = 0) = 1 - 3\nu / (n+2), \\ P(\tilde{N}_{\max} = i) = (6\nu(n+1-i)) / (n(n+1)(n+2)), \forall 1 \leq i \leq n. \end{cases} \quad (3.9)$$

To compare graphically bounds (3.9) and (3.3) see Figure 3.4. Note that the condition made on  $\nu$  comes from the constraints of global decreasingness and convexity on  $\mathbb{N}$ . The bound (3.9) is not in line with the one derived in Lefèvre and Loisel (2010)[9] for r.v.'s with decreasing and convex p.m.f. in a support  $\{0, \dots, n\}$ . At the same time this bound is in line with bound (3.15) in Lefèvre and Loisel (2012)[10] available for r.v.'s with decreasing and convex p.m.f. on  $\mathbb{N}$ .

**Proof.** The proof is based on Propositions 3.2.1 and 3.2.3. In fact, according to Proposition 3.2.1, we have at least two crossing points between p.m.f. of r.v.'s  $M$  and  $\tilde{N}_{\max}$ . In addition, if there exist exactly two crossing points such that  $P(\tilde{N}_{\max}) \geq P(M)$  near  $n$ , then under Proposition 3.2.3, we conclude that  $M \leq_{cx} \tilde{N}_{\max}$ .

We denote by  $P(M=0) = a_1$ ,  $P(\tilde{N}_{\max}=0) = a$ ,  $P(M=1) = c_1$  and  $P(\tilde{N}_{\max}=1) = c$ .

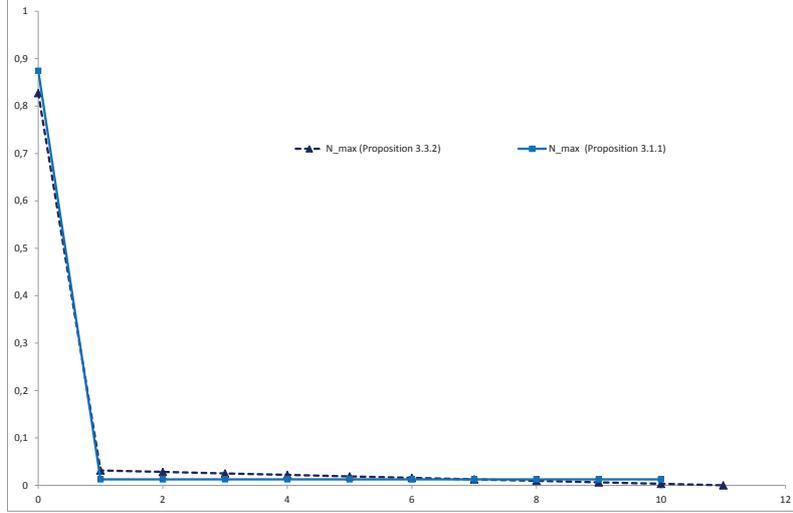


Figure 3.4: The upper bound (3.9) vs the upper bound (3.3),  $\nu = 0.69$  and  $n = 10$ .

1.  $a_1 \geq a$

Assume that  $c_1 \leq c$ , in this case we observe, at most, one crossing point in  $]0, 1]$ . Furthermore, by Proposition 3.2.1, it is necessary to have, at least, another crossing point in  $]1, n]$ . However, if this is the case then under Lemma 3.3.1 the constraint of convexity is violated. Thus this is absurd. Now consider that  $c_1 > c$ , then there is at most one crossing point over  $\mathcal{D}_n$ . This is absurd by Proposition 3.2.1.

2.  $a_1 < a$

Assume that  $c_1 > c$ , then a crossing point is observed in  $]0, 1]$ . By Lemma 3.3.1, we cannot get more than one crossing point in  $]1, n]$ . If not, the convexity constraint is not fulfilled. Hence, we have a situation with at most two crossing points over  $\mathcal{D}_n$  such that  $P(M) \leq P(\tilde{N}_{\max})$  near  $n$ . In this case under Proposition 3.2.3,  $M \leq_{cx} \tilde{N}_{\max}$ . The case where  $c_1 \leq c$  is absurd according to Lemma 3.3.1 and Proposition 3.2.1.

◇

Consequently, the maximum variance denoted by  $Var(\tilde{N}_{\max})$  is obtained for

$$\begin{aligned} Var(\tilde{N}_{\max}) &= \sum_{i=1}^n b(n+1-i)i^2 - \nu^2 \\ &= \frac{6\nu}{n(n+1)(n+2)} \frac{n(n+2)(n+1)^2}{12} - \nu^2, \end{aligned}$$

hence,

$$\boxed{Var(\tilde{N}_{\max}) = \frac{\nu(n+1)}{2} - \nu^2.} \tag{3.10}$$

### 3.3.2 The lower bound

Consider the class of r.v.'s with discrete p.m.f. that are globally convex and decreasing on  $\mathbb{N}$ . In this case the convex minimum order is given by the Proposition (3.3.3).

**Proposition 3.3.3.** *Let  $\xi$  be an integer in  $[0, n-1]$  such that  $\xi \leq 3\nu \leq \xi+1$  where  $n \geq \xi+1$ . The convex minimum  $\tilde{N}_{\min}$  in the class of r.v.'s which have a globally decreasing and convex distribution is as follows*

$$P\left(\tilde{N}_{\min} = j\right) = (\xi + 1 - j)\pi_1 + (\xi + 2 - j)\pi_2, \quad \forall j \in \{0, \dots, \xi + 1\}, \quad (3.11)$$

where  $\pi_1 = 2(\xi + 1 - 3\nu)/((\xi + 2)(\xi + 1))$ ,  $\pi_2 = 2(3\nu - \xi)/((\xi + 2)(\xi + 3))$ ,  $\pi_1 \geq 0$  and  $\pi_2 \geq 0$ .

For a comparison between bounds (3.11) and (3.4) see Figure 3.5.

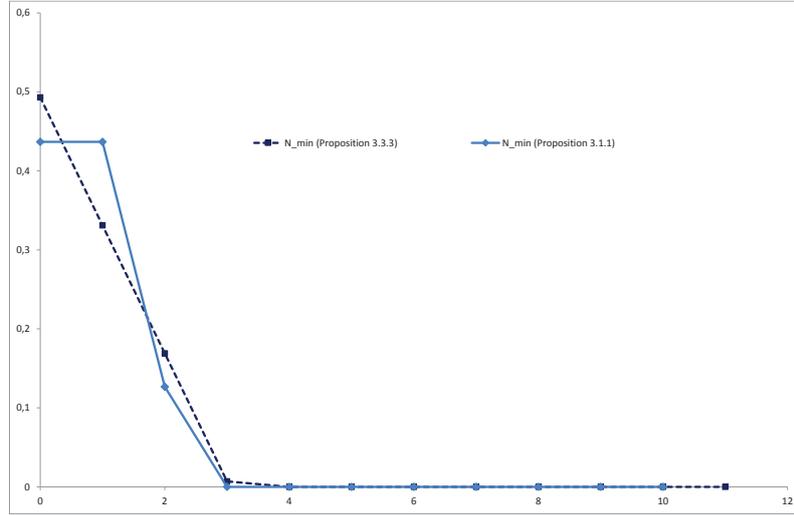


Figure 3.5: The lower bound (3.11) vs the lower bound (3.4),  $\nu = 0.69$  and  $n = 10$ .

**Proof.** This bound is obtained as follows: we consider that the distribution of  $\tilde{N}_{\min}$  is defined by

$$\begin{cases} P\left(\tilde{N}_{\min} = i\right) = a - bi, \quad \forall i \in \{0, \dots, \xi + 1\}, \\ \text{where } 0 < b < a < 1, \end{cases} \quad (3.12)$$

thus, we have  $P(\tilde{N}_{\min} = \xi + 2) = 0$ . Consider  $\pi_2 = P(\tilde{N}_{\min} = \xi + 1)$  where under the convexity assumption  $P(\tilde{N}_{\min} = \xi) \geq 2\pi_2$ . Consider that  $P(\tilde{N}_{\min} = \xi) = \pi_1 + 2\pi_2$ , with  $\pi_1 > 0$ , then immediately we have  $b = \pi_1 + \pi_2$  and  $a = (1 + \xi)\pi_1 + (2 + \xi)\pi_2$ , so (3.11) is true, where  $\pi_1$  and  $\pi_2$  are obtained under the following constraints

$$\begin{cases} \sum_{i=0}^{\xi+1} (\xi + 1 - i)\pi_1 + (\xi + 2 - i)\pi_2 = 1, \\ \sum_{i=1}^n i((\xi + 1 - i)\pi_1 + (\xi + 2 - i)\pi_2) = \nu. \end{cases} \quad (3.13)$$

Let  $M$  and  $\tilde{N}_{\min}$  be two arbitrary random variables in  $\mathbb{J}_n$  with p.m.f. globally convex on  $\mathbb{N}$  such that  $V(M) \neq V(\tilde{N}_{\min})$ .

1.  $a_1 > a$

Since Lemma 3.3.1 it is easy to see that it is impossible to get more than two crossing points over  $\mathcal{D}_n$ . This is due to the assumptions of decreasingness and convexity. Also, by the same Lemma, a situation with exactly two crossing points is observed if and only if there is  $j \in \{0, \dots, \xi + 1\}$  such that  $P(M = j) < P(\tilde{N}_{\min} = j)$  and  $P(M = \xi + 2) > 0$ . In this case it is clear that  $P_M > P_{\tilde{N}_{\min}}$  near  $n$ . Hence, by Proposition 3.2.3,  $\tilde{N}_{\min} \leq_{cx} M$ .

2.  $a_1 \leq a$

By convexity and decreasingness we have at most one crossing point which is absurd by Proposition 3.2.1.

◇

Consequently, we have

$$\begin{aligned} \text{Var}(\tilde{N}_{\min}) &= \sum_{i=1}^{\xi+1} ((\xi + 1 - i)\pi_1 + (\xi + 2 - i)\pi_2)i^2 - \nu^2 \\ &= (\xi + 1)\left(\nu - \frac{\xi}{6}\right) - \nu^2. \end{aligned} \tag{3.14}$$

Note that the lower bound (3.11) is in agreement with Corollary 5.4 in Lefèvre and Loisel (2010)[9] hold for discrete r.v.'s with p.m.f. that are decreasing and convex on  $\mathcal{D}_n$ . It is also in agreement with bound (3.14) in Lefèvre and Loisel (2012)[10] hold for discrete r.v.'s with p.m.f. that are decreasing and convex globally on  $\mathbb{N}$ .

The bounds (3.11) and (3.9) are illustrated in Figure 3.6 for illustration.

### 3.4 For decreasing p.m.f. that are convex in $[1, +\infty)$

Note that if a random variable  $X$  has a convex p.m.f. in  $[1, +\infty)$ , then it is either convex (see Section 3.3) or concave until 2. This section is devoted to extrema convex for r.v.'s which have decreasing p.m.f. that are convex in  $[1, +\infty)$  and concave until 2. In application, a typical example is to consider a Poisson distribution with parameter ( $\lambda$ ):  $(2 - \sqrt{2}) < \lambda \leq 1$ . Let  $M$  be a r.v. in  $\mathbb{J}_n$ . In comparison with the previous section we impose the following conditions: the p.m.f. of the r.v.  $M$  is concave in  $[0, 2]$  and convex in  $[1, +\infty)$ .

#### 3.4.1 The upper bound

**Proposition 3.4.1.** *Consider  $\frac{1}{3} < \nu < \frac{n}{3}$ . Consider the class of r.v.'s with discrete decreasing p.m.f. that are concave until  $[0, 2]$  and convex in  $[1, +\infty)$ . The convex maximum*

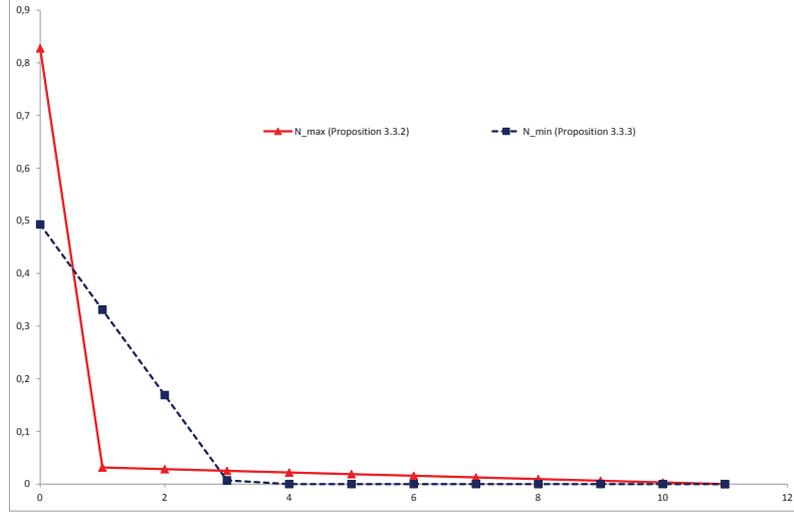


Figure 3.6: Convex extrema in the class of decreasing p.m.f. that are convex on  $\mathbb{N}$  with  $\nu = 0,69$  et  $n = 10$ .

in this class of r.v.'s is obtained for the r.v.  $\tilde{N}_{\max}$  defined by (3.15)

$$\begin{cases} P\left(\tilde{N}_{\max} = 0\right) = (2n(n+4) + 6 - 6\nu(n+1)) / (3(n+2)(n+1)), \\ P\left(\tilde{N}_{\max} = 1\right) = (n(n+4) - 3\nu(n-2)) / (3(n+2)(n+1)), \\ P\left(\tilde{N}_{\max} = i\right) = (2(3\nu - 1)(n+1-i)) / ((n-1)(n+2)(n+1)), \quad \forall 2 \leq i \leq n+1. \end{cases} \quad (3.15)$$

We refer the proof of Proposition 3.4.1 in Appendix C.1. For a comparison between (3.15) and (3.9) see Figure 3.7.

Consequently, we have

$$\begin{aligned} \text{Var}(\tilde{N}_{\max}) &= c + \sum_{i=2}^n b(n+1-i)i^2 - \nu^2 \\ &= c + b \frac{1}{12} n(n-1)(n^2 + 5n + 10) - \nu^2 \\ &= \left(\frac{1}{2}n + 1\right)\nu - \frac{1}{6}n - \nu^2. \end{aligned} \quad (3.16)$$

**Corollary 3.4.1.** *Let  $X$  be a r.v. in  $\mathbb{J}_n$  valued on  $\mathcal{D}_n$ . In addition, consider that the p.m.f. of  $X$  is decreasing on  $\mathbb{N}$  and convex in  $[1, \infty)$ . We know that on the set  $[0, 2]$  the distribution of  $X$  is either convex or concave (see Section 3.3). Then the upper bound of  $X$  is the supremum of the two bounds (3.15) and (3.9). From bounds (3.16) and (3.10) the supremum is attained when the random variable is convex on the set  $[0, 2]$  i.e for (3.10).*

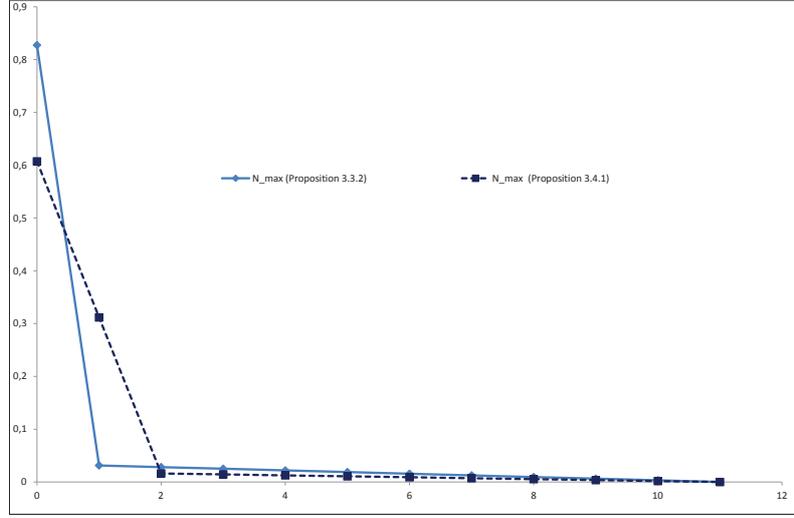


Figure 3.7: The upper bound (3.15) vs the upper bound (3.9) with  $\nu = 0.69$ ,  $n = 10$ .

### 3.4.2 The lower bound

#### 3.4.2.1 Case: $\frac{1}{3} \leq \nu \leq \frac{1}{2}$

**Proposition 3.4.2.** Consider the class of discrete r.v.'s which have a decreasing p.m.f. that are concave until  $[0, 2]$  and convex in  $[1, +\infty)$ . The convex minimum is attained for the r.v.  $\tilde{N}_{\min}$  given by (3.17)

$$\begin{cases} P(\tilde{N}_{\min} = 0) = 1 - \nu, \\ P(\tilde{N}_{\min} = 1) = \nu, \\ P(\tilde{N}_{\min} = i) = 0, \forall i > 1. \end{cases} \quad (3.17)$$

**Proof.** It is easy to see that it cannot exist more than two crossing points between p.m.f. of  $M$  and  $\tilde{N}_{\min}$ . In addition if a two crossing situations holds, then necessarily we must have  $P(M = 2) > P(\tilde{N}_{\min} = 2)$  hence  $\tilde{N}_{\min} \leq_{cx} M$ .  $\diamond$  Consequently we have,

$$\boxed{\text{Var}(\tilde{N}_{\min}) = \nu(1 - \nu).} \quad (3.18)$$

#### 3.4.2.2 Case: $\xi < 3\nu \leq \xi + 1 + \frac{2}{(\xi+4)}$ where $3\nu - \xi > \frac{2}{\xi+3}$ for all $\xi \geq 1$ and $n \geq \xi + 1$

**Proposition 3.4.3.** Consider  $\xi < \xi + \frac{2}{(\xi+3)} < 3\nu \leq \xi + 1 + \frac{2}{(\xi+4)}$  where  $3\nu - \xi > \frac{2}{\xi+3}$ . Taking the class of discrete r.v.'s with decreasing p.m.f. that are concave until  $[0, 2]$  and

convex in  $[1, +\infty)$ , the convex minimum is attained for the r.v.  $\tilde{N}_{\min}$  given by

$$\begin{cases} P(\tilde{N}_{\min} = 0) = \pi_2(1 + \xi) + \pi_1\xi, \\ P(\tilde{N}_{\min} = i) = \pi_2(2 + \xi - i) + \pi_1(1 + \xi - i), \quad \forall \quad 1 \leq i \leq \xi + 1, \\ P(\tilde{N}_{\min} = i) = 0, \quad \forall \quad i > \xi + 1, \end{cases} \quad (3.19)$$

where

$$\pi_1 = \frac{2(\xi^2 - 3\nu\xi + 5\xi - 12\nu + 6)}{(\xi + 2)(\xi + 5)\xi},$$

and

$$\pi_2 = \frac{2(-3\xi - 2 + 9\nu - \xi^2 + 3\nu\xi)}{(\xi + 5)(\xi + 1)(\xi + 2)}.$$

where  $\pi_1 \geq 0$  and  $\pi_2 \geq 0$ .

For a comparison between bounds (3.19) and (3.11) see Figure 3.8. We refer the proof of Proposition 3.4.3 in Appendix C.2.

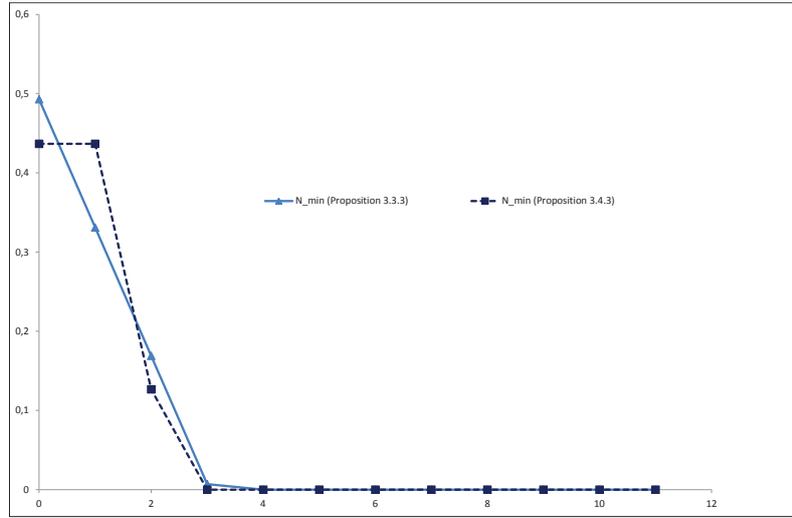


Figure 3.8: The lower bound (3.19) vs the lower bound (3.11),  $\nu = 0.69$ ,  $n = 10$ .

Consequently, we have

$$\boxed{Var(\tilde{N}_{\min}) = \frac{1}{12} (2 + \xi) (1 + \xi) (\pi_1\xi^2 + \pi_2\xi^2 + \pi_1\xi + 5\pi_2\xi + 6\pi_2) - \nu^2.} \quad (3.20)$$

In Figure 3.9 we present the upper bound (3.15) and the lower bound (3.19) for illustration.

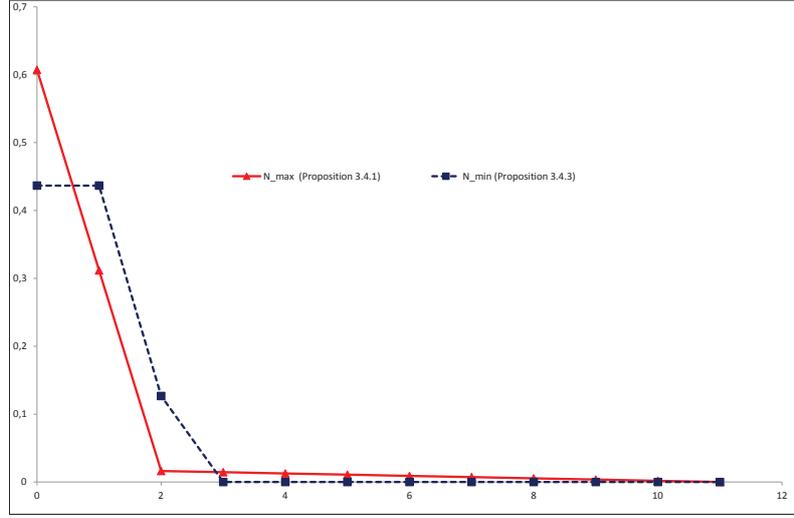


Figure 3.9: The lower bound (3.19) vs the upper bound (3.15),  $\nu = 0.69$ ,  $n = 10$ .

### 3.5 For decreasing p.m.f. that are convex in $[m, +\infty)$ for fixed $m > 1$

In this section we consider discrete r.v.'s with decreasing p.m.f. that are convex from a fixed point  $m > 1$ . Our aim is to find the convex extrema bounds in this family. We first forecast in the upper bound.

#### 3.5.1 The upper bound

**Proposition 3.5.1.** *Let  $n$  and  $m$  be integers and let  $\nu$  be a positive real number such that  $0 < 3\nu \leq \frac{(m^2 + nm + m + n^2 + 2n)}{(m+n+2)}$ . The convex maximum order in the class of r.v.'s with p.m.f. that are globally decreasing on  $\mathbb{N}$  and convex from a fixed point  $m > 1$  is attained for the r.v.  $\tilde{N}_{\max}$  given by*

$$\begin{cases} P\left(\tilde{N}_{\max} = 0\right) = 1 - (3\nu(n+m)) / (m^2 + nm + m + n^2 + 2n) , \\ P\left(\tilde{N}_{\max} = i\right) = (6\nu(n+1-m)) / (m + n^3 + 3n^2 + 2n - m^3) \quad \forall 1 \leq i \leq m-1, \\ P\left(\tilde{N}_{\max} = i\right) = (6\nu(n+1-i)) / (m + n^3 + 3n^2 + 2n - m^3) \quad \forall m \leq i \leq n. \end{cases} \quad (3.21)$$

We refer the proof of Proposition 3.5.1 in Appendix C.3. Consequently, the supremum of the variance is attained for the predefined r.v.  $\tilde{N}_{\max}$  where for all  $1 < m \leq n$ , we have

$$\boxed{\text{Var}(\tilde{N}_{\max}) = \frac{6\nu((m^2/12) - (m^4/12) + (n/6) + (n^4/12) + (n^3/3) + (5n^2/12))}{(m + n^3 + 3n^2 + 2n - m^3)} - \nu^2.} \quad (3.22)$$

As illustration see Figure 3.10.

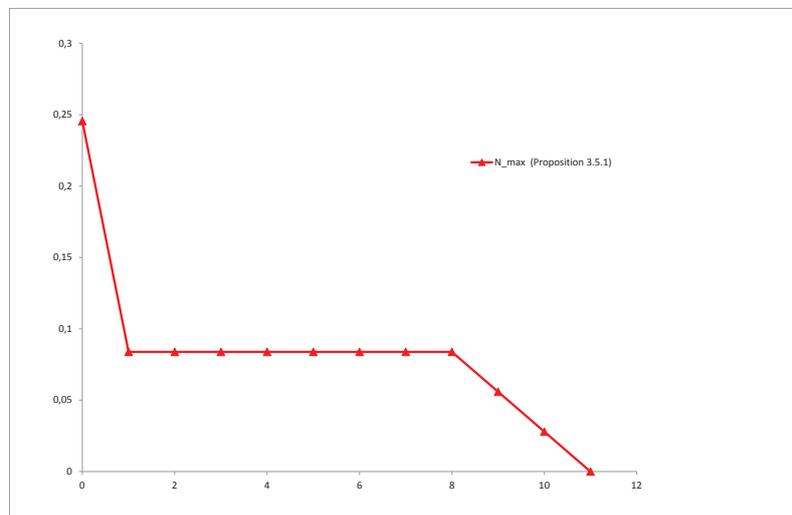


Figure 3.10: The upper bound (3.21),  $m = 8$ ,  $\nu = 3.8$ ,  $n = 10$ .

**Note 3.5.2.** If  $m = 1$  or  $m = 0$ , then this bound corresponds to the bound (3.10) obtained for discrete r.v.'s with globally decreasing and convex p.m.f.

**Note 3.5.3.** Let  $\mathcal{F}_m$  be a set of decreasing r.v.'s valued on  $\mathcal{D}_n$  with equal mean  $\nu$  and with convex p.m.f. for a fixed point  $m$ :  $1 < m \leq n$ . Let us denote by  $\tilde{N}_{\max, m}$  the convex maximum bound in the set  $\mathcal{F}_m$  and define the more general set  $\mathcal{G} = (\mathcal{F}_m)_{(m \in \{1, \dots, n-1\})}$ . As  $\mathcal{F}_p \subset \mathcal{F}_m$  for all  $p \leq m$ , we deduce that the convex maximum bound in the set  $\mathcal{G}$  is  $\tilde{N}_{\max, n}$ .

### 3.5.2 The lower bound

#### 3.5.2.1 Case: $0 \leq \xi < 2\nu \leq (\xi + 1) \leq m$

**Proposition 3.5.4.** Let  $n$  and  $m$  be two integers and  $\nu$  be a nonnegative real number such that  $0 \leq \xi < 2\nu \leq \xi + 1 \leq m$ . The convex minimum in the class of r.v.'s with p.m.f. that are globally decreasing on  $\mathbb{N}$  and convex from a fixed point  $m > 1$  is attained for the r.v.  $\tilde{N}_{\min}$  where

$$\begin{cases} P\left(\tilde{N}_{\min} = i\right) = 2(\xi + 1 - \nu)/(\xi + 1)(\xi + 2), & \forall 0 \leq i \leq \xi, \\ P\left(\tilde{N}_{\min} = \xi + 1\right) = (2\nu - \xi)/(\xi + 2), \\ P\left(\tilde{N}_{\min} = i\right) = 0, \forall i > \xi + 1. \end{cases} \quad (3.23)$$

Note that this bound is equal to the bound (3.12) in Lefèvre and Loisel (2012)[10] obtained for discrete r.v.'s with decreasing p.m.f. globally on  $\mathbb{N}$ . It is also equal to the

bound (5.3) derived in Lefèvre and Loisel (2010)[9] for discrete r.v.'s with decreasing p.m.f. on  $\mathcal{D}_n$ . Consequently,

$$\boxed{\text{Var}(\tilde{N}_{\min}) = -\frac{\xi}{3} (\xi - 4\nu + 1) + \nu - \nu^2.} \quad (3.24)$$

We refer the proof of Proposition 3.5.4 in Appendix C.4.

**3.5.2.2 Case:**  $m < \xi \leq 3\nu \leq \xi + 1$ ,  $\frac{m(m+1)}{m+\xi+2} \leq 3\nu - \xi$  and  $n \geq \xi + 1$

**Proposition 3.5.5.** *Let  $m, \xi$  and  $n$  be three integers where  $1 < m < n$  and  $\nu$  be nonnegative real number such that  $m < \xi \leq \xi + \frac{m(m+1)}{(m+2+\xi)} \leq 3\nu \leq \xi + 1 + \frac{m(m+1)}{(m+3+\xi)}$  and  $\frac{m(m+1)}{m+\xi+2} \leq 3\nu - \xi$ . The convex minimum in the class of r.v.'s with p.m.f. that are globally decreasing on  $\mathbb{N}$  and convex from a fixed point  $m > 1$  is attained for the r.v.  $\tilde{N}_{\min}$  such that*

$$\begin{cases} P\left(\tilde{N}_{\min} = i\right) = \pi_2(2 + \xi - m) + \pi_1(1 + \xi - m), \quad \forall \quad 0 \leq i \leq m, \\ P\left(\tilde{N}_{\min} = i\right) = \pi_2(2 + \xi - i) + \pi_1(1 + \xi - i), \quad \forall \quad m + 1 \leq i \leq \xi + 1, \\ P\left(\tilde{N}_{\min} = i\right) = 0, \quad \forall \quad i > \xi + 1, \end{cases} \quad (3.25)$$

where

$$\pi_1 = 2 \frac{m^2 - 3\nu m + m\xi + 2m + 3 - 9\nu - 3\nu\xi + \xi^2 + 4\xi}{(\xi + 2)(2m + \xi + 3)(\xi + 1 - m)},$$

and

$$\pi_2 = -\frac{2(m^2 - 3\nu m + m\xi + m - 6\nu - 3\nu\xi + \xi^2 + 2\xi)}{(2m + \xi + 3)(\xi + 2 - m)(\xi + 2)},$$

where  $\pi_1 \geq 0$  and  $\pi_2 \geq 0$ .

Consequently we have,

$$\begin{aligned} \text{Var}(\tilde{N}_{\min}) = & (1 + \xi - m) \left( \frac{1}{6} \pi_1 m(m+1)(2m+1) + \frac{1}{6} (\pi_1(1+\xi) + \pi_2(2+\xi)) \right. \\ & (2m^2 + 5m + 2\xi m + 6 + 7\xi + 2\xi^2) \\ & \left. - (\pi_1 + \pi_2) \left( \frac{1}{4} m + \frac{1}{2} + \frac{1}{4} \xi \right) (m^2 + m + 2 + 3\xi + \xi^2) \right) + \\ & \frac{1}{6} \pi_2 (2 + \xi - m) m(m+1)(2m+1) - \nu^2. \end{aligned}$$

Note that if we take  $m = 1$  and  $\xi = 1$ , then in this case this bound coincides with the bound obtained in Proposition 3.4.2.

See Figure 3.11 for an illustration of the lower bounds obtained from (3.25) and (3.23). We refer the proof of Proposition 3.5.4 in Appendix C.5.

## 3.6 Some numerical illustration

Since 2001, the European Commission has begun to establish the new regulation ‘Solvency II’ common to all countries members of the European Union. This regulation will be applied from October 2014. In particular, it sets the harmonization of methods of valuation of

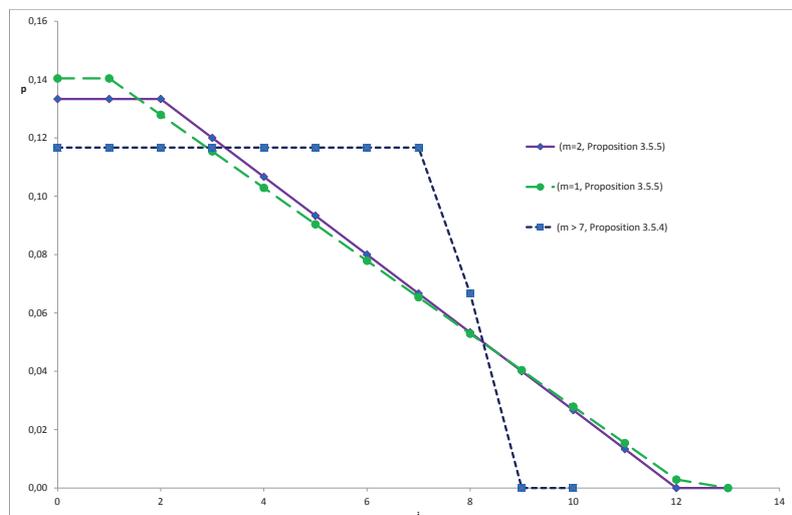


Figure 3.11: The lower bounds for decreasing p.m.f. that are convex on  $[m, +\infty)$ ,  $\nu = 3.8$ ,  $n = 10$  for  $m = 2$  and  $m = 1$ ,  $m > 7$ .

liabilities with a risk margin and the implementation of new rules to estimate the solvency capital requirement (SCR). This last is the main monitoring tool of the control authorities and is calculated so that all risks to which the entity is exposed are taken into consideration. The Basic SCR is estimated firstly for each business line, then all individual SCR are aggregated with respect to a correlation matrix. Insurance companies have the choice between two options: they can either adopt a standard approach or an internal model.

In the standard model of ‘Solvency II’ we distinguish two approaches: a scenario based approach and a factor based approach. In the scenario based approach, we measure the sensitivity to a shock of each individual risk and in the factor based approach we apply fixed factors to approximate the risk. In this last case the SCR is defined by  $SCR = q\sigma$  where  $\sigma$  is the standard error of the random loss and  $q > 0$  is called a quantile factor. So,  $q = 3$  is usually chosen for claim amount with a moderate tail distribution. For heavy tailed risks, a more relevant value is  $q = 5$  or  $6$ . In the sequel we use a factor based approach to approximate the SCR.

Traditionally, computing or approximating the distribution function of the aggregate claim amount has been one of the central points in insurance mathematics. In order to investigate the distribution of the aggregate claim amount we consider individual model or collective model. Note that a collective risk model is often adopted to describe the occurrence of large claim. The total number  $N$  of claim occurring in a given period is random, typically it has a Poisson, binomial or negative binomial distribution. Further r.v.’s claim size  $W$  are strictly positive and are assumed to be independent and identically distributed and independent of  $N$ .

Consider the r.v.  $S = \sum_{i=1}^N W_i$  describing the aggregate claim amount. In this case we say that  $S$  has a compound distribution where  $E(S) = E(N)E(W)$  and  $Var(S) = E(N)Var(W) + [E(W)]^2Var(N)$ . In practice, we have limited information about the distribution of  $N$ . In fact, in general only the mean  $E(N)$  is available and the distribution of  $N$  is assumed to be Poisson with parameter  $\lambda$ . For example with respect to some French data, one could use Poisson distribution with parameter  $\lambda(C_{27}) = 0.37$  for business line  $C_{27}$  (drought and earthquake) and  $\lambda(C_{35}) = 0.69$  for business line  $C_{35}$  (construction-damages to building). Recall that we are in a situation where a large amount of claim occurs without being in a catastrophic situation. In reality practitioners consider that observing more than a certain number  $n$  of claims correspond to a catastrophe. For this reason we consider that  $n$  is fixed.

Instead of Poisson distribution we consider distributions that are decreasing on  $\mathbb{N}$  which are globally or partially convex. Recall that the probability mass function of a Poisson distribution is decreasing convex if  $\lambda \leq 2 - \sqrt{2}$  and decreasing but not convex if  $2 - \sqrt{2} < \lambda \leq 1$ . The bounds for  $SCR = q\sigma$  are simply given by

$$SCR(\tilde{N}_{min}) = q\sqrt{Var(W)E(N) + E^2(W)Var(\tilde{N}_{min})}$$

and

$$SCR(\tilde{N}_{max}) = q\sqrt{Var(W)E(N) + E^2(W)Var(\tilde{N}_{max})}$$

In the numerical illustrations we choose,  $E[W(C_{27})] = 1000$ ,  $Var[W(C_{27})] = 2500^2$  and  $E[W(C_{35})] = 2000$ ,  $Var[W(C_{35})] = 7000^2$  (in thousand of euros).

### 3.6.1 Convex extrema for decreasing convex distribution globally on $\mathbb{N}$

This gives  $Var(\tilde{N}_{max}) = 1.8981$  and  $SCR(\tilde{N}_{max}) = 12311.85$  for business line  $C_{27}$  where  $n(C_{27}) = 10$ . These bounds are of course sharper than those obtained for distribution that are decreasing and convex not globally on  $\mathbb{N}$  studied by Lefèvre and Loisel (2010)[9] which are recalled here:  $Var(\tilde{N}_{max}) = 2.453$  and  $SCR(\tilde{N}_{max}) = 13098.2$ .

(For maximum bounds see Table 3.1 and for minimum bounds see Table 3.2).

	$C_{27}$		$C_{35}$	
$n$	$Var(\tilde{N}_{max})$	$SCR(\tilde{N}_{max})$	$Var(\tilde{N}_{max})$	$SCR(\tilde{N}_{max})$
5	0.9731	10875.73	1.5939	11858.77
10	1.8981	12311.85	3.3189	41171.37
20	3.7481	14770.97	6.7689	46817.53
30	5.5981	16875.47	10.2189	51852.50
40	7.4481	18745.18	13.6689	56440.07

Table 3.1: Maximum bounds for decreasing convex distribution globally on  $\mathbb{N}$ .

	$C_{27}$	$C_{35}$
$\text{Var}(\tilde{N}_{min})$	0.2698	0.5939
$\text{SCR}(\tilde{N}_{min})$	9641.7	36092.7

Table 3.2: Minimum bounds for decreasing convex distribution globally on  $\mathbb{N}$ .

### 3.6.2 Convex extrema for decreasing distributions that are concave until 2 and convex on $[1, \infty)$

For business line  $C_{27}$ ,  $\text{Var}(\tilde{N}_{max}) = 0.416$  and  $\text{SCR}(\tilde{N}_{max}) = 9911.680$  where  $n(C_{27}) = 10$ . For business line  $C_{35}$  (construction-damages to building),  $\text{Var}(\tilde{N}_{max}) = 3.781$  and  $\text{SCR}(\tilde{N}_{max}) = 41971.70$  for  $n(C_{35}) = 20$ . It is clear that these bounds are sharper than those derived in Section 3.6.1. (For maximum bounds see Table 3.3 and for minimum bounds see Table 3.4.)

	$C_{27}$		$C_{35}$	
$n$	$\text{Var}(\tilde{N}_{max})$	$\text{SCR}(\tilde{N}_{max})$	$\text{Var}(\tilde{N}_{max})$	$\text{SCR}(\tilde{N}_{max})$
5	0.3238	9742.012	1.105	37099.218
10	0.416	9911.680	1.997	38790.82
20	0.600	10239.220	3.781	41971.70
30	0.7832	10556.595	5.564	44926.18
40	0.966	10864.672	7.347	47698.30

Table 3.3: Maximum bounds for decreasing distributions that are concave until 2 and convex on  $[1, \infty)$ .

	$C_{27}$	$C_{35}$
$\text{Var}(\tilde{N}_{min})$	0.2331	0.4672
$\text{SCR}(\tilde{N}_{min})$	9572.96	35838.24

Table 3.4: Minimum bounds for decreasing distributions that are concave until 2 and convex on  $[1, \infty)$ .

### 3.6.3 Decreasing distribution globally on $\mathbb{N}$ and convex on $[m, +\infty)$ for fixed $m \geq 1$

For different fixed values of  $n$  and  $m$  we evaluate the variance and the SCR. (For maximum bounds see Tables 3.5, 3.6, 3.7 and 3.8, and for minimum bounds see Table 3.9.)

We give in the following table minimum bounds: for business line  $C_{27}$  we use Proposition 3.5.4 for any fixed  $m \geq 1$  in fact for this business line  $0 < 2 \times 0.37 \leq 1$ ; for

$n(C_{27}) \setminus m$	1	2	8	18	25	28	35	38
5	0.9731	0.9949	—	—	—	—	—	—
10	1.8981	1.9057	2.2409	—	—	—	—	—
20	3.7481	3.7503	3.8869	4.6899	—	—	—	—
30	5.5981	5.5992	5.6714	6.1820	6.8210	7.1509	—	—
40	7,4481	7.4487	7.4931	7.8404	8.3148	8.5710	9.2742	9.6147

Table 3.5: Maximum variances for decreasing distribution globally on  $\mathbb{N}$  and convex on  $[m, +\infty)$  for fixed  $m \geq 1$  ( business line  $C_{27}$ ).

$n(C_{27}) \setminus m$	1	2	8	18	25	28	35	38
5	10875.73	10911.75	—	—	—	—	—	—
10	12311.85	12322.95	12803.22	—	—	—	—	—
20	14770.98	14773.65	14939.16	15877.228	—	—	—	—
30	16875.49	16876.65	16953.48	17487.19	18133.01	18457.58	—	—
40	18745.18	18745.75	18788.34	19118.17	19559.73	19794.09	20423.55	20721.47

Table 3.6: Maximum SCR for decreasing distribution globally on  $\mathbb{N}$  and convex on  $[m, +\infty)$  for fixed  $m \geq 1$  (line of business  $C_{27}$ ).

$n(C_{35}) \setminus m$	1	2	8	18	25	28	35	38
20	6.7689	6.7731	7.0276	8.5253	—	—	—	—
30	10.2189	10.2209	10.3556	11.3078	12.4994	13.1147	—	—
40	13.6689	13.6701	13.7528	14.4004	15.2851	15.7630	17.0744	17.7094

Table 3.7: Maximum variances for decreasing distribution globally on  $\mathbb{N}$  and convex on  $[m, +\infty)$  for fixed  $m \geq 1$  (line of business  $C_{35}$ ).

$n(C_{35}) \setminus m$	1	2	8	18	25	28	35	38
20	46817.53	46823.99	47213.71	49444.95	—	—	—	—
30	51852.50	51855.28	52041.97	53343.07	54927.89	55728.60	—	—
40	56440.07	56441.60	56546.99	57365.65	58465.49	59051.10	60628.98	17.7095

Table 3.8: Maximum SCR for decreasing distribution globally on  $\mathbb{N}$  and convex on  $[m, +\infty)$  for fixed  $m \geq 1$  (line of business  $C_{35}$ ).

business line  $C_{35}$  we use Proposition 3.5.5 if  $m = 1$  and Proposition 3.5.4 for all  $m > 1$  as  $1 < 2 \times 0.69 \leq 2$ . From Tables 3.5, 3.6, 3.7 and 3.8 we remark that for fixed  $n$  the choice

	$C_{27}$	$C_{35}$
$\text{Var}(\tilde{N}_{min})$	0.2331	0.4672
$\text{SCR}(\tilde{N}_{min})$	9572.96	35838.24

Table 3.9: Minimum bounds for decreasing distribution globally on  $\mathbb{N}$  and convex on  $[m, +\infty)$  where  $m \geq 1$ .

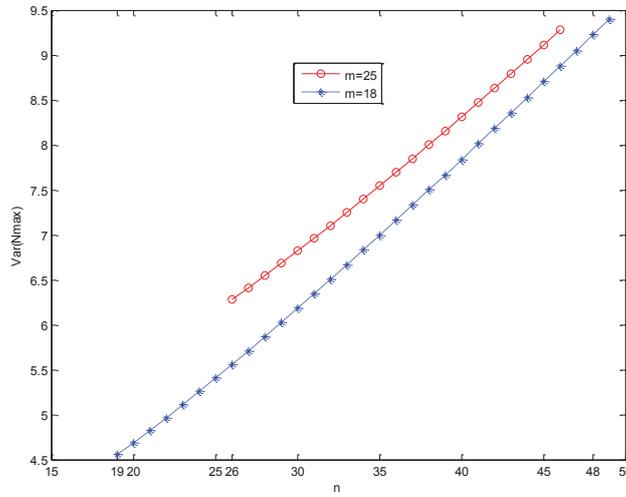


Figure 3.12:  $(\tilde{N}_{max})$  for different  $m$  and  $n$ ,  $\nu = 0.37$

of the maximal value  $m$  influences considerably the value of the variance and of the SCR. Also, for fixed  $m$  we observe that the choice of the maximal value  $n$  influences considerably the value of the variance and of the SCR (see Figure 3.12).

Let us prove algebraically that  $a$  is an increasing function of  $n$  and that  $b$  is a decreasing function of  $n$ . From (3.21) for all fixed  $m \geq 1$  we denote by

$$a(n) = 1 - \frac{3\nu(n+m)}{m^2 + nm + m + n^2 + 2n} = 1 - \frac{3\nu}{(m+1) + \frac{n(n+1)}{m+n}}.$$

Hence the behavior of  $a(n)$  follows the behavior of  $\beta(n) = \frac{n(n+1)}{m+n}$ , where

$$\frac{\partial\beta(n)}{\partial n} = \frac{(n+m)(2n+1) - n(n+1)}{(n+m)^2}.$$

Note that the sign of this derivative always follows the sign of the numerator. We note that  $(n+m)(2n+1) - n(n+1) = 2n^2 + 2nm + m = 0$  has a general form of a second degree equation where the value of the discriminant denoted by  $\Delta$  is equal to  $4m(m-1)$ . Two cases arise: the first one is when  $m$  is equal to 1. In this case  $\Delta = 0$  and we have a double solution equal to  $-1$ . As  $n$  is a strictly positive integer the sign of the partial derivative

is always positive in the first case. The second case is when  $m > 1$ . In this case  $\Delta > 0$  and the equation has two distinct solutions denoted by  $n_1 = -m - \sqrt{m(m-1)}$  and by  $n_2 = -m + \sqrt{m(m-1)}$  which are both negatives. As  $n > 0$  the sign of this polynomial is always positive on  $[1, \infty)$ . Hence we deduce that  $a(n)$  is a strictly increasing function of  $n$  for all fixed  $m \geq 1$ . By the same manner denote by

$$b(n) = \frac{6\nu}{m - m^3 + n(n^2 + 3n + 2)}.$$

We note that the behavior of  $b(n)$  when  $n$  varies follows the inverse behavior of  $\alpha(n) = (n(n^2 + 3n + 2))^{-1}$ . Hence for all fixed  $m \geq 1$ , it is very easy to see that  $b(n)$  is a decreasing function of  $n$ . By the same argument we can prove that for fixed  $n$

$$a(m) = 1 - \frac{3\nu(n+m)}{m(m+n+1) + (n^2 + 2n)}$$

is a strictly nonincreasing function of  $m$  and

$$b(m) = \frac{6\nu}{m(1-m^2) + n(n^2 + 3n + 2)}$$

is a strictly increasing function of  $m$ .

---

## Bibliography

- [1] Courtois, C., Denuit, M. and Van Belleghem, S.: Discrete  $s$ -convex extremal distributions: theory and applications. *Applied Mathematics Letters*. **19**, (2006), 1367-177.
- [2] Denuit, M. and Lefèvre, C.: Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance: Mathematics and Economics*. **20**, (1997), 197-213.
- [3] Denuit, M., Lefèvre, C. and Shaked, M.: The  $s$ -convex orders among real random variables, with applications. *Mathematical Inequalities and Applications*. **1**, (1998), 585-613.
- [4] Denuit, M., Lefèvre, C. and Utev, S.: Stochastic orderings of convex/concave type on an arbitrary grid. *Mathematics of Operations Research*. **24**, (1999a), 835-846.
- [5] Denuit, M., De Vylder, F.E. and Lefèvre, C.: Extremal generators and extremal distributions for the continuous  $s$ -convex stochastic orderings. *Insurance: Mathematics and Economics*. **24**, (1999b), 201-217.
- [6] Denuit, M., Lefèvre, C. and Mesfioui, M.: On  $s$ -convex stochastic extrema for arithmetic risks. *Insurance: Mathematics and Economics* **25**. (1999c), 143-155.
- [7] Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E. and Bauwelinckx, T.: *Effective Actuarial Methods*. North-Holland, Amsterdam. (1990).
- [8] Goovaerts, M.J., Kaas, R. and Van Heerwaarden A.E.: Ordering of Actuarial Risks. *Caire Education Series 1, Caire, Brussels*. (1998).
- [9] Lefèvre, C. and Loisel, S.: Stationary-excess operator and convex stochastic orders. *Insurance: Mathematics and Economics*. **47**, (2010), 64-75.
- [10] Lefèvre, C. and Loisel, S.: On multiply monotone risk distributions. *Accepted, to appear in Journal of Applied Probability*. (2012).
- [11] Lefèvre, C. and Picard, P.: An unusual stochastic order relation with some applications in sampling and epidemic theory. *Advances in Applied Probability*. **25**, (1993), 63-81.
- [12] Lefèvre, C. and Utev, S.: Comparing sums of exchangeable Bernoulli random variables. *Journal of Applied Probability*. **33**. (1996), 285-310.
- [13] Müller, A. and Stoyan, D.: Comparison Methods for Stochastic Models and Risks. *Wiley, New York*. (2002).
- [14] Rolski, T.: Order relations in the set of distribution functions and their applications in queueing theory. *Dissertations Math. CXXXII, Polish Academy of Sciences, Warsaw*. (1976).
- [15] Ross, S.M.: Stochastic Processes. *Wiley, New York*. (1996).
- [16] Shaked, M. and Shanthikumar, J.G.: Stochastic Orders and their Applications. *Springer, New York*. (2007).

- [17] Shaked, M. and Shanthikumar, J.G.: Stochastic Orders and their Applications. *Academic press, New York.* (1994).

# Conclusion et perspectives

Cette thèse porte sur deux parties principales qui émanent de la théorie du risque. La première est liée à l'étude des modèles à facteur commun pour la modélisation de la dépendance. La deuxième est liée à l'étude des extrema convexes sous des contraintes de convexité dans un cadre discret. Nous rappelons pour chacune des deux parties les problématiques ainsi que nos principaux résultats.

Dans la première partie nous avons considéré des modèles de dépendance dans lesquels les risques dépendent d'un facteur commun. En théorie du risque ces modèles sont connus sous le nom des modèles avec mélange commun. Ces derniers ont la propriété suivante : les variables sont conditionnellement indépendantes par rapport à un facteur commun. Rappelons que dans les modèles avec mélange commun classiques, le conditionnement se fait par rapport à une mémoire du facteur de longueur  $\ell$  finie ( $\ell$  généralement égale à un).

Dans notre étude nous proposons une généralisation de ces modèles en considérant que le conditionnement se fait par rapport à une mémoire non bornée du facteur. En d'autres termes, nous considérons que l'intégralité de la mémoire du facteur influe sur le comportement des variables de risque. Puisque le vecteur aléatoire du facteur de conditionnement est de longueur variable au cours du temps, le type de dépendance temporelle étudié dans cette thèse peut être qualifié de dynamique. En actuariat cette modélisation permet aux organismes assureurs de modéliser les risques dans un cadre général qui tient compte de l'évolution de l'environnement dans lequel évolue leur activité. En effet, les crises passées et d'autres facteurs importants affectent la fréquence et les montants des sinistres. De tels événements modifient les prix et peuvent avoir des conséquences à long terme sur les caractéristiques des contrats offerts sur le marché de l'assurance. Par conséquent, il n'est pas toujours suffisant de compter sur une mémoire finie pour modéliser les risques.

Dans ce cadre, la principale problématique abordée dans le premier chapitre de la thèse est d'étudier les propriétés des processus de risque avec conditionnement d'ordre non borné. Vu la particularité du conditionnement qu'on étudie, nous considérons que nos processus satisfont une condition de mélange conditionnel. On rappelle que le mélange en tant qu'outil en statistique est un concept de dépendance faible. De plus, il est connu que les processus mélangeants ont des propriétés asymptotiques intéressantes.

L'étude des propriétés des processus en question est liée à de nombreuses pistes de recherche comme l'analyse des processus agrégés en théorie du risque. Cette dernière constitue la problématique que nous avons étudiée dans le deuxième chapitre de la thèse. En effet, l'agrégation des risques permet d'avoir une vision globale sur le risque engendré par une

activité. Particulièrement, on s'intéresse au comportement asymptotique de la loi agrégée. La question que nous avons posée est : sous quelles conditions la somme des variables aléatoires conditionnellement indépendantes par rapport à une mémoire non bornée, peut-elle converger vers une loi normale ?

Dans cette perspective nous avons donné dans le premier chapitre un résultat dans lequel nous avons proposé quelques propriétés de mélange pour les processus de risque. Dans le deuxième chapitre nous avons donné un théorème central limite auto normalisé pour des suites fortement mélangeantes. L'estimateur de la variance que nous avons proposé est adapté au type de mélange de nos processus de risque. Au fil du deuxième chapitre, nous avons fortement insisté sur le fait que nos processus ne satisfont pas les propriétés de mélange classiques. Pour cela nous avons dû adapter la plupart des outils et résultats de la littérature existante avant de les utiliser.

Nous avons donné des réponses à nos problématiques posées dans cette première partie. Pour autant, les résultats que nous avons donné ne constituent qu'un travail qui ouvre la voie pour plusieurs autres perspectives. On peut regarder l'impact d'un conditionnement de mémoire non bornée sur les risques encourus par une activité d'assurance et ceci dans un cadre pluriannuel et non pas asymptotique. Dans ce cas, il est intéressant d'étudier la loi de la somme agrégée non asymptotique ainsi que la probabilité de ruine et de voir l'impact de la mémoire non bornée du conditionnement sur les mesures de risques VaR et TVaR. Il est aussi intéressant de généraliser notre étude au cas multivarié. En effet, on peut considérer différentes branches d'activité et voir l'impact de la dépendance introduite via une mémoire non bornée d'un facteur sur les risques d'une même branche d'activité et sur le risque inter-branches. Aussi il est important d'étudier le cas où la loi du facteur est différente de la loi de Bernoulli et où la loi de comptage est une loi de Poisson. En effet, cette dernière est très courante en théorie de risque. Une autre piste est d'approfondir notre étude dans le cas d'une modélisation Markovienne dans laquelle le facteur est une chaîne de Markov et de voir l'impact d'une telle modélisation sur la corrélation temporelle entre les risques.

Dans la première partie de la thèse on suppose qu'on dispose des informations statistiques complètes pour les distributions de pertes. Nous pouvons parfois faire face à des situations dans lesquelles on dispose uniquement d'une connaissance statistique incomplète sur certaines composantes des modèles. Dans ce contexte, dans la deuxième partie de cette thèse nous nous sommes concentrés sur les ordres convexes des distributions discrètes et décroissantes sous des contraintes de convexité. Les extrema convexes correspondants sont calculés en utilisant de simples propriétés de croisement entre deux distributions. Plusieurs illustrations en théorie de la ruine sont présentées. Cette deuxième partie ouvre la voie pour étudier d'autres cas tels que le cas où la fonction de masse de probabilité est décroissante globalement sur  $\mathbb{N}$  et convexe jusqu'à un point  $m$  du support.

## Appendix A

# Detailed proofs of some results of the first part

## A.1 Proof of equations 31 and 32 of the introduction

We have

$$\begin{aligned}\Psi_{S'_{n_1}|\underline{V}_n}(t) &= \mathbb{E}(e^{itS'_{n_1}}|\underline{V}_n) = \mathbb{E}\left\{e^{\frac{it}{\sqrt{n}}\sum_{j=1}^n X_j - \mathbb{E}(X_j|\underline{V}_j)}|\underline{V}_n\right\} \\ &= \mathbb{E}\left\{\prod_{j=1}^n e^{\frac{it}{\sqrt{n}}\sum_{j=1}^n X_j - \mathbb{E}(X_j|\underline{V}_j)}|\underline{V}_n\right\} \\ &= \prod_{j=1}^n \mathbb{E}\left\{e^{\frac{it}{\sqrt{n}}\sum_{j=1}^n X_j - \mathbb{E}(X_j|\underline{V}_j)}|\underline{V}_n\right\},\end{aligned}$$

by the Taylor development of the exponential function we write

$$\begin{aligned}\mathbb{E}(e^{\frac{it}{\sqrt{n}}(X_j - \mathbb{E}(X_j|\underline{V}_j))}|\underline{V}_n) &= \mathbb{E}\left\{1 + it\frac{X_j - \mathbb{E}(X_j|\underline{V}_j)}{\sqrt{n}} - \frac{t^2}{2n}(X_j - \mathbb{E}(X_j|\underline{V}_j))^2 + o\left(\frac{t^2}{n}\right)|\underline{V}_j\right\} \\ &= 1 - \frac{t^2}{2n}(\mathbb{E}(X_j^2|\underline{V}_j) - \mathbb{E}^2(X_j|\underline{V}_j)) + o\left(\frac{t^2}{n}\right) \\ &= 1 - \frac{t^2}{2n}\text{Var}(X_j|\underline{V}_j) + o\left(\frac{t^2}{n}\right),\end{aligned}$$

in fact,

$$\text{Var}(X_i|\underline{V}_i) = \mathbb{E}(\{X_i - \mathbb{E}(X_i|\underline{V}_i)\}^2|\underline{V}_i) = \mathbb{E}(X_i^2|\underline{V}_i) - \mathbb{E}^2(X_i|\underline{V}_i).$$

Then

$$\begin{aligned}\Psi_{S'_{n_1}}(t) &= \mathbb{E}_{\underline{V}_n}(\Psi_{S'_{n_1}|\underline{V}_n}(t)) = \mathbb{E}_{\underline{V}_n}\left\{\prod_{j=1}^n \left(1 - \frac{t^2}{2n}\text{Var}(X_j|\underline{V}_j) + o\left(\frac{t^2}{n}\right)\right)\right\} \\ &= \mathbb{E}_{\underline{V}_n}\left\{e^{\sum_{j=1}^n \ln\left[1 - \frac{t^2}{2n}\text{Var}(X_j|\underline{V}_j) + o\left(\frac{t^2}{n}\right)\right]}\right\}.\end{aligned}$$

The Taylor development of  $\ln(1 - x)$  allows us to write

$$\begin{aligned}\Psi_{S'_{n_1}}(t) &= \mathbb{E}_{\underline{V}_n}\left\{e^{\sum_{j=1}^n \left[-\frac{t^2}{2n}\text{Var}(X_j|\underline{V}_j) + o\left(\frac{t^2}{n}\right)\right]}\right\} \\ &= \mathbb{E}_{\underline{V}_n}\left\{e^{-\frac{t^2}{2n}\sum_{j=1}^n \text{Var}(X_j|\underline{V}_j)}\right\} \times e^{no\left(\frac{t^2}{n}\right)}.\end{aligned}$$

We also derive easily that

$$\Psi_{S'_{n_2}}(t) = \mathbb{E}(e^{itS'_{n_2}}) = E_{\underline{V}_n}\left(e^{itS'_{n_2}}\right) = e^{-\frac{t^2}{2n}E_{\underline{V}_n}\left\{\sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)]^2 + o\left(\frac{t^2}{n}\right)\right\}}. \quad (\text{A.1})$$

In fact,

$$\begin{aligned}
& \mathbb{E}_{\underline{V}_n} \left( e^{itS'_{n2}} \right) \\
&= \mathbb{E}_{\underline{V}_n} \left\{ e^{\frac{it}{\sqrt{n}} \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)]} \right\} \\
&= \mathbb{E}_{\underline{V}_n} \left\{ 1 + \frac{it}{\sqrt{n}} \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)] - \frac{t^2}{2n} \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)]^2 + o\left(\frac{t^2}{n}\right) \right\} \\
&= 1 + \mathbb{E}_{\underline{V}_n} \left\{ \sum_{j=1}^n \frac{it}{\sqrt{n}} [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)] - \frac{t^2}{2n} \left( \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)] \right)^2 + o\left(\frac{t^2}{n}\right) \right\} \\
&= 1 + \mathbb{E}_{\underline{V}_n} \left\{ -\frac{t^2}{2n} \left( \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)] \right)^2 + o\left(\frac{t^2}{n}\right) \right\} \\
&= e^{-\frac{t^2}{2n} \mathbb{E}_{\underline{V}_n} \left\{ \sum_{j=1}^n [\mathbb{E}(X_j|\underline{V}_j) - \mathbb{E}(X_j)] \right\}^2 + o\left(\frac{t^2}{n}\right)}.
\end{aligned}$$

## A.2 Model of application in a continuous case

**Example A.2.1.** *In this example we consider that  $(X_i)_{i \in \mathbb{N}}$  is a sequence of dependent not identically distributed Exponential r.v's which models the claim amount. The conditional kernel transition is denoted by*

$$N^i(dx_i, \underline{V}_i) = f_{\underline{V}_i}^i(\cdot) dx_i.$$

Consider that the conditional density of  $X_i$  given  $\underline{V}_i \in \mathbb{N}$ , has the form

$$f_{\underline{V}_i}^i(\theta_i; x_i) = \frac{1}{\theta_i} \exp\left(-\frac{x_i}{\theta_i}\right) \text{ for } x_i \geq 0,$$

where the parameter  $\theta_i$  is depending on  $\underline{V}_i$  with the following structure:

$$\theta_i = \sum_{j=1}^i \frac{h(V_j)}{2^{i-j}},$$

where the function  $0 < h(v) \leq \tau_2$ . This model have the following representation, which will be interesting in section : let  $Z_i$  be an exponential random variable with parameter 1. Consider that

$$X_i|\underline{V}_i = Z_i * \theta_i,$$

in this case  $X_i|\underline{V}_i$  is exponential with parameter  $(1/\theta_i)$ . As presented in the Pareto example we want to prove that the sequence  $(X_i)_{i \in \mathbb{N}}$  is conditionally mixing when conditioning is done with respect to the vector of the factor  $\underline{V}_i$ .

The conditional density of  $(X_{i1}, \dots, X_{iu})|\underline{V}_{iu}$  is denoted by

$$H_{x_{j1}, \dots, x_{jv}}(V_1, \dots, V_{jv}) = \left( \prod_{k=1}^v \theta_{i_k} \right)^{-1} \exp\left(-\sum_{k=1}^v \frac{x_{i_k}}{\theta_{i_k}}\right).$$

For this aim we evaluate

$$\left| \text{Cov}(H_{x_{i1}, \dots, x_{iu}}(\underline{V}_{iu}), H_{x_{j1}, \dots, x_{jv}}(\underline{V}_{jv})) \right|.$$

**Proposition A.2.1.** *If  $0 < \tau_1 \leq h(v) \leq \tau_2$  for any  $v$  then the process of the Example A.2.1 satisfies 1.12 and 1.10 of Assumption 1.2.1.*

*Proof.* Firstly we prove that condition 1.12 is satisfied. We have  $\frac{1}{2\tau_2} \leq \frac{1}{\theta_i} \leq \frac{1}{\tau_1}$  and  $-\frac{1}{\tau_1} \leq -\frac{1}{\theta_i} \leq -\frac{1}{2\tau_2}$ , hence

$$\sup_{\underline{V}} \theta_{i_k}^{-1} \exp\left(-\frac{x_{i_k}}{\theta_{i_k}}\right) \leq \frac{2\tau_2}{\tau_1} \frac{1}{2\tau_2} \exp\left(-\frac{x_{i_k}}{2\tau_2}\right),$$

and

$$\int \mathcal{N}^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v})_{\infty}(dx_{j_1}, \dots, dx_{j_v}) \leq \left(\frac{2\tau_2}{\tau_1}\right)^{-v} \underbrace{\int_{x_{j_1}, \dots, x_{j_v}} \left(\prod_{k=1}^v 2\tau_2\right)^{-1} \exp\left(-\sum_{k=1}^v \frac{x_{i_k}}{2\tau_2}\right)}_{=1} < \infty.$$

So (1.12) is satisfied. Secondly we prove that (1.10) is satisfied. We have shown that

$$\begin{aligned} & \left| \text{Cov} \left( H_{x_{i_1}, \dots, x_{i_u}}(\underline{V}_{i_u}), H_{x_{j_1}, \dots, x_{j_v}}(\underline{V}_{j_v}) \right) \right| \\ & \leq \|H_{x_{i_1}, \dots, x_{i_u}}(V_1, \dots, V_{i_u})\|_a \cdot \sup_{\underline{V}} |H_{x_{j_1}, \dots, x_{j_v}}(V_1, \dots, V_{j_v}) - H_{x_{j_1}, \dots, x_{j_v}}(0, \dots, 0, V_{i_u+\ell+1}, \dots, V_{j_v})| \\ & + \eta_{(u,v)}^*V(\ell) \|H_{x_{i_1}, \dots, x_{i_u}}(V_1, \dots, V_{i_u})\|_a \cdot \|H_{x_{j_1}, \dots, x_{j_v}}(V_1, \dots, V_{j_v})\|_b. \end{aligned}$$

Denote by

$$\tilde{\theta}_i = \sum_{j=1}^i \frac{h(\tilde{V}_j)}{2^{i-j}}.$$

We have

$$\begin{aligned} & \left| H_{x_{j_1}, \dots, x_{j_v}}(V_1, \dots, V_{j_v}) - H_{x_{j_1}, \dots, x_{j_v}}(0, \dots, 0, V_{i_u+\ell+1}, \dots, V_{j_v}) \right| \\ & = \left| \frac{1}{\prod_{k=1}^v \theta_{j_k}} \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) - \frac{1}{\prod_{k=1}^v \tilde{\theta}_{j_k}} \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\tilde{\theta}_{j_k}}\right) \right| \\ & = \left| \frac{\prod_{k=1}^v \tilde{\theta}_{j_k}}{\prod_{k=1}^v \theta_{j_k} \tilde{\theta}_{j_k}} \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) - \frac{\prod_{k=1}^v \theta_{j_k}}{\prod_{k=1}^v \theta_{j_k} \tilde{\theta}_{j_k}} \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\tilde{\theta}_{j_k}}\right) \right| \\ & \leq \frac{1}{\prod_{k=1}^v \theta_{j_k} \tilde{\theta}_{j_k}} \left[ \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) \left| \prod_{k=1}^v \tilde{\theta}_{j_k} - \prod_{k=1}^v \theta_{j_k} \right| \right. \\ & \quad \left. + \prod_{k=1}^v \theta_{j_k} \left| \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) - \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\tilde{\theta}_{j_k}}\right) \right| \right] \\ & \leq \frac{1}{\prod_{k=1}^v \theta_{j_k} \tilde{\theta}_{j_k}} \left[ \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) \left| \prod_{k=1}^v \tilde{\theta}_{j_k} - \prod_{k=1}^v \theta_{j_k} \right| \right] \\ & + \frac{1}{\prod_{k=1}^v \tilde{\theta}_{j_k}} \left| \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\theta_{j_k}}\right) - \exp\left(-\sum_{k=1}^v \frac{x_{j_k}}{\tilde{\theta}_{j_k}}\right) \right|, \end{aligned}$$

In one hand

$$\left| \prod_{k=1}^v \tilde{\theta}_{j_k} - \prod_{k=1}^v \theta_{j_k} \right| \leq \frac{(2\tau_1)^v}{2^p} \sum_{i=1}^v \frac{\binom{v}{i}}{2^i} \leq \frac{(3\tau_1)^v}{2^p},$$

where  $p = j_1 - i_u - \ell$  and where

$$\frac{1}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \leq \frac{1}{\tau_1^v}.$$

In fact,

$$\sum_{i=1}^v \frac{\binom{v}{i}}{2^i} < \sum_{i=0}^v \frac{\binom{v}{i}}{2^i} < \frac{1}{2^v} \sum_{i=0}^v \binom{v}{i} 2^{v-i} 1^v < \left(\frac{3}{2}\right)^v,$$

hence

$$\left| \prod_{k=1}^v \tilde{\theta}_{j_k} - \prod_{k=1}^v \theta_{j_k} \right| \leq \frac{(2\tau_1)^v 3^v}{2^p 2^v} \leq \frac{(3\tau_1)^v}{2^p},$$

and then

$$\frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\tilde{\theta}_k})}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k \tilde{\theta}_k} \left| \prod_{k=j_1}^{j_v} \tilde{\theta}_k - \prod_{k=j_1}^{j_v} \theta_k \right| \leq \frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\theta_k}) 3^v}{\prod_{k=j_1}^{j_v} \theta_k 2^p}.$$

In the other hand

$$\begin{aligned} \frac{1}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \left| \exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\theta_k}) - \exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\tilde{\theta}_k}) \right| &= \frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\tilde{\theta}_k})}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \left| \exp(\sum_{k=j_1}^{j_v} x_k (\frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k})) - 1 \right| \\ &\leq \frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\tilde{\theta}_k})}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \exp(u) \left| \sum_{k=j_1}^{j_v} x_k (\frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k}) \right| \\ &\leq \frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\tilde{\theta}_k} + u)}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \sum_{k=j_1}^{j_v} x_k \left| \frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k} \right| \\ &\leq \frac{\exp(-\sum_{k=j_1}^{j_v} x_k (\frac{1}{\theta_k} - \frac{2\tau_2}{2^p \tau_1^2}))}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \sum_{k=j_1}^{j_v} x_k \left| \frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k} \right|, \end{aligned}$$

where  $u \in \left[0, \left| \sum_{k=j_1}^{j_v} x_k (\frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k}) \right| \right]$ . We have

$$\left| \tilde{\theta}_k - \theta_k \right| \leq \frac{2\tau_2}{2^p}.$$

Hence

$$\left| \frac{1}{\theta_k} - \frac{1}{\tilde{\theta}_k} \right| = \left| \frac{\tilde{\theta}_k - \theta_k}{\theta_k \tilde{\theta}_k} \right| \leq \frac{2\tau_2}{2^p \tau_1^2}.$$

Then we conclude that

$$\begin{aligned} &\int_{\underline{V} x_{j_1}, \dots, x_{j_v}} \left| D(\underline{V}_{j_v}, \tilde{V}_{i_u, j_v, \ell}) \right| \\ &\leq \underbrace{\frac{3^v}{2^p} \int_{\underline{V} x_{j_1}, \dots, x_{j_v}} \left| \frac{\exp(-\sum_{k=j_1}^{j_v} \frac{x_k}{\theta_k})}{\prod_{k=j_1}^{j_v} \theta_k} \right|}_{1} + \frac{2\tau_2}{2^p \tau_1^2} \int_{\underline{V} x_{j_1}, \dots, x_{j_v}} \left| \frac{\exp(-\sum_{k=j_1}^{j_v} x_k (\frac{1}{\theta_k} - \frac{2\tau_2}{2^p \tau_1^2}))}{\prod_{k=j_1}^{j_v} \tilde{\theta}_k} \sum_{k=j_1}^{j_v} x_k \right| \\ &\leq \frac{3^v}{2^p} + \frac{2\tau_2}{2^p \tau_1^{2+v}} \int_{\underline{V} x_{j_1}, \dots, x_{j_v}} \left| \exp(-\sum_{k=j_1}^{j_v} x_k (\frac{1}{\tau_1} - \frac{2\tau_2}{2^p \tau_1^2})) \sum_{k=j_1}^{j_v} x_k \right| < \infty. \end{aligned}$$

### A.3 Complement to the proof of Theorem 1.2.1

**Assumption A.3.1.**  $K_v$  and  $K'_v$  denote positive constants depending on  $v$ ,  $(\kappa(\ell))_{\ell \in \mathbb{N}}$  denotes a decreasing to zero sequence and  $\tilde{N}^{j_1, \dots, j_v}$  is a positive finite measure on  $\mathbb{R}^v$  having total mass  $K_v$ . Let  $F_{\underline{V}}$  be the distribution function of the vector of r.v's  $\underline{V}$ . Consider  $k = j_1 - i_u - \ell$  with  $0 < \ell \leq j_1 - i_u$ . For multi indices  $(i_1, \dots, i_u)$  and  $(j_1, \dots, j_v)$  satisfying  $(\star)$ , we consider the following assumptions.

The positive kernel

$$|D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| \text{ is } \|\cdot\|_{\infty}\text{-dominated by } K_v \|N^{j_1, \dots, j_v}\|_1 \kappa(k). \quad (\text{A.1})$$

$$|D(\cdot, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| \text{ is } \|\cdot\|_{\infty}\text{-dominated by } \tilde{N}^{j_1, \dots, j_v} \kappa(k). \quad (\text{A.2})$$

$$\int |D(dx_{j_1} \dots dx_{j_v}, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| dF_{\underline{V}_{j_v}}(\underline{v}_{j_v}) \leq \kappa(k) K_v. \quad (\text{A.3})$$

$$N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v}) \text{ is } \|\cdot\|_{\infty}\text{-dominated by } K'_v \|N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v})\|_1. \quad (\text{A.4})$$

$$N^{j_1, \dots, j_v}(\cdot, \underline{V}_{j_v}) \text{ is } \|\cdot\|_{\infty}\text{-dominated by a finite measure } \mathcal{N}_{\infty}^{j_1, \dots, j_v} \quad (\text{A.5})$$

and  $\mathcal{N}_{\infty}^{j_1, \dots, j_v}$  has total mass  $K'_v$ .

Recall that when (A.1) or (A.2) of Assumption A.3.1 are satisfied, then we use Lemma 1.3.1 and when (A.4) of assumption A.3.1 is satisfied, then we use Lemma 1.3.2. Let  $f$  and  $g$  be two real valued bounded functions. Note that by Hölder inequality we have  $|\text{Cov}(f, g)| \leq 2 \|f\|_p \|g\|_q$  where  $\frac{1}{p} + \frac{1}{q} = 1$ . For a r.v.  $Y$  we denote by  $\eta_{(u,v)}^Y(r) = \eta^Y(r) C^Y(u, v)$ . Denote by  $(\star\star)$  the following quantity

$$|\text{Cov}(\mathbb{E}(f(X_{i_1}, \dots, X_{i_u}) | \underline{V}_{i_u}), \mathbb{E}(g(X_{j_1}, \dots, X_{j_v}) | \underline{V}_{j_v}))|.$$

1.  $(V_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing + (A.2) of Assumption A.3.1  $\Rightarrow (X_i)_{i \in \mathbb{N}}$  is  $\Phi_{(u,v)}^*$ -mixing.

*Proof.* In this case

$$\begin{aligned} (\star\star) &\leq 2 \|f\|_1 \kappa(r - \ell) \int |g| \cdot \tilde{N}^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}) + \\ &\quad \Psi_{(u,v)}^{*V}(\ell) \int |fg| \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \cdot \|N^{j_1, \dots, j_v}\|_1(dx_{j_1} \dots dx_{j_v}) \\ &\leq 2K_v \|f\|_1 \|g\|_{\infty} \kappa(r - \ell) + \Psi_{(u,v)}^{*V}(\ell) \|f\|_1 \|g\|_{\infty}. \end{aligned}$$

□

2.  $(V_i)_{i \in \mathbb{N}}$  is  $\Phi_{(u,v)}^*$ -mixing + (A.2) and (A.5) of Assumption A.3.1  $\Rightarrow (X_i)_{i \in \mathbb{N}}$  is  $\Phi_{(u,v)}^*$ -mixing.

*Proof.* In this case

$$\begin{aligned}
 (\star\star) &\leq 2\|f\|_1\kappa(r-\ell) \int |g| \cdot \tilde{N}^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}) + \\
 &\quad \Phi_{(u,v)}^{*V}(\ell) \int |fg| \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \cdot \mathcal{N}_\infty^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}) \\
 &\leq 2K_v\kappa(r-\ell) \|f\|_1 \|g\|_\infty + K'_v \Phi_{(u,v)}^{*V}(\ell) \|f\|_1 \|g\|_\infty.
 \end{aligned}$$

□

3.  $(V_i)_{i \in \mathbb{N}}$  is  $\alpha_{(u,v)}^*$ -mixing + (A.1) and (A.4) of Assumption A.3.1  $\Rightarrow (X_i)_{i \in \mathbb{N}}$  is  $\Psi_{(u,v)}^*$ -mixing.

*Proof.* In this case

$$\begin{aligned}
 (\star\star) &\leq 2K_v\kappa(r-\ell)\|f\|_1 \int |g| \|N^{j_1, \dots, j_v}\|_1(dx_{j_1} \dots dx_{j_v}) + \alpha_{(u,v)}^{*V}(\ell) \times \\
 &\quad \int |f| |g| K'_u K'_v \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \|N^{j_1, \dots, j_v}\|_1(dx_{j_1} \dots dx_{j_v}) \\
 &\leq 2K_v\kappa(r-\ell) \|f\|_1 \|g\|_1 + K'_u K'_v \alpha_{(u,v)}^{*V}(\ell) \|f\|_1 \|g\|_1.
 \end{aligned}$$

□

4.  $(V_i)_{i \in \mathbb{N}}$  is  $\alpha_{(u,v)}^*$ -mixing + (A.3) and (A.4) of Assumption A.3.1  $\Rightarrow (X_i)_{i \in \mathbb{N}}$  is  $\Phi_{(u,v)}^*$ -mixing.

*Proof.* In this case

$$\begin{aligned}
 (\star\star) &\leq 2K'_v \|g\|_\infty \int |f| \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \times \\
 &\quad \int |D(dx_{j_1} \dots dx_{j_v}, \underline{v}_{j_v}, \tilde{v}_{i_u, j_v, \ell})| dF_{\underline{V}_{j_v}} + \\
 &\quad \alpha_{(u,v)}^{*V}(\ell) \int |fg| \|N^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \cdot \mathcal{N}_\infty^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}) \\
 &\leq 2K_v K'_u \kappa(r-\ell) \|f\|_1 \|g\|_\infty + \alpha_{(u,v)}^{*V}(\ell) K'_u K'_v \|f\|_1 \|g\|_\infty.
 \end{aligned}$$

□

5.  $(V_i)_{i \in \mathbb{N}}$  is  $\alpha_{(u,v)}^*$ -mixing + (A.3) and (A.5) of Assumption A.3.1  $\Rightarrow (X_i)_{i \in \mathbb{N}}$  is  $\alpha_{(u,v)}^*$ -mixing.

*Proof.* In this case

$$\begin{aligned}
 (\star\star) &\leq 2K_v\kappa(r-\ell)\|f\|_\infty \|g\|_\infty \int |f| \|\mathcal{N}_\infty^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) + \\
 &\quad \alpha_{(u,v)}^{*V}(\ell) \int |fg| \|\mathcal{N}_\infty^{i_1, \dots, i_u}\|_1(dx_{i_1} \dots dx_{i_u}) \cdot \mathcal{N}_\infty^{j_1, \dots, j_v}(dx_{j_1} \dots dx_{j_v}) \\
 &\leq 2K_v K'_v \kappa(r-\ell) \|f\|_\infty \|g\|_\infty + K'_u K'_v \alpha_{(u,v)}^{*V}(\ell) \|f\|_\infty \|g\|_\infty.
 \end{aligned}$$

□

## A.4 Complement to the proof of Proposition 1.4.2

The fact that (A.5) is satisfied is straightforward from the expression of  $F_{x_{i_1}, \dots, x_{i_u}}(\underline{V}_{i_u})$  and the fact that  $\tau_1 \leq h(v) \leq \tau_2$  implies:  $\tau_1 \leq (\theta_i)^\alpha \leq 2\tau_2$ .

On one other hand,

$$\begin{aligned} & \left| F_{x_{j_1}, \dots, x_{j_v}}(v_1, \dots, v_{j_v}) - F_{x_{j_1}, \dots, x_{j_v}}(0, \dots, 0, v_{i_u+\ell+1}, \dots, v_{j_v}) \right| \\ &= \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} \left| \prod_{k=1}^v \theta_{j_k}^\alpha \mathbf{1}_{\{x_{j_k} \geq \theta_{j_k}\}} - \prod_{k=1}^v \tilde{\theta}_{j_k}^\alpha \mathbf{1}_{\{x_{j_k} \geq \tilde{\theta}_{j_k}\}} \right|, \end{aligned}$$

where

$$\tilde{\theta}_i = \sum_{j=1}^i \frac{h(\tilde{V}_j)}{2^{i-j}},$$

with  $\tilde{V}_j = 0$  for  $j = 1, \dots, i_u + \ell$  and  $\tilde{V}_j = V_j$  for  $j \geq i_u + \ell + 1$ . Proceeding as in Subsection 4.2.1, we have that

$$\left| \prod_{p=1}^v \theta_{i_p}^\alpha - \prod_{p=1}^v \tilde{\theta}_{i_p}^\alpha \right| \leq \frac{(4\tau_2)^v}{2^p},$$

with  $p = j_1 - i_u - \ell$ . Also, consider a derivable function  $f$  such that  $f : x \rightarrow x^{\frac{1}{\alpha}}$  where  $f' = \frac{1}{\alpha} x^{\frac{1}{\alpha}-1} = \frac{1}{\alpha} x^{\frac{1-\alpha}{\alpha}} = \frac{1}{\alpha x^{\frac{\alpha-1}{\alpha}}}$ . Consider that  $x = \theta^\alpha$ . In addition we have  $\tau_1 < \theta^\alpha < 2\tau_2$  thus  $f' = \frac{1}{\alpha(\theta^\alpha)^{\frac{\alpha-1}{\alpha}}} < \frac{1}{\alpha(\tau_1)^{\frac{\alpha-1}{\alpha}}}$ . By applying Taylor's equality we obtain

$$|f(\theta_i^\alpha) - f(\tilde{\theta}_i^\alpha)| \leq |\theta_i^\alpha - \tilde{\theta}_i^\alpha| \sup |f'|,$$

hence

$$|\theta_i - \tilde{\theta}_i| \leq |\theta_i^\alpha - \tilde{\theta}_i^\alpha| \sup |f'| \leq \frac{2\tau_2}{2^p} \frac{1}{\alpha \tau_1^{\frac{\alpha-1}{\alpha}}}, \quad (\text{A.6})$$

Let us denote by  $\Theta_i$  the interval  $[\min(\theta_i, \tilde{\theta}_i), \max(\theta_i, \tilde{\theta}_i)]$ . We get:

$$\begin{aligned} D(\underline{V}_{j_v}, \tilde{V}_{i_u, j_v, \ell}) &\leq \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} \frac{(4\tau_2)^v}{2^p} \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \geq \tau_1^{\frac{1}{\alpha}}\}} \\ &\quad + \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} (2\tau_2)^v \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \in \Theta_{j_k}\}}. \end{aligned}$$

We have

$$\begin{aligned}
 & \int_{\underline{V}} \int_{x_{j_1}, \dots, x_{j_v}} \left| D(\underline{V}_{j_v}, \tilde{V}_{i_u, j_v, \ell}) \right| \\
 & \leq \int_{\underline{V}} \int_{x_{j_1}, \dots, x_{j_v}} \left| \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} \frac{(4\tau_2)^v}{2^p} \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \geq \tau_1^{\frac{1}{\alpha}}\}} \right| + \int_{\underline{V}} \int_{x_{j_1}, \dots, x_{j_v}} \left| \frac{\alpha^v}{\prod_{k=1}^v x_{j_k}^{\alpha+1}} (2\tau_2)^v \prod_{k=1}^v \mathbf{1}_{\{x_{j_k} \in \Theta_{j_k}\}} \right| \\
 & \leq \underbrace{\frac{(4\tau_2)^v}{2^p} \prod_{k=1}^v \int_{x_{j_k}} \int_{\underline{V}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \mathbf{1}_{\{x_{j_k} \geq \tau_1^{\frac{1}{\alpha}}\}} \right|}_{< \infty} + (2\tau_2)^v \prod_{k=1}^v \int_{x_{j_k}} \int_{\underline{V}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \mathbf{1}_{\{x_{j_k} \in \Theta_{j_k}\}} \right|.
 \end{aligned}$$

Using Fubini's Theorem, we have

$$\int_{x_{j_k}} \int_{\underline{V}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \mathbf{1}_{\{x_{j_k} \in \Theta_{j_k}\}} \right| = \int_{x_{j_k} \in \Theta_{j_k}} \int_{\underline{V}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \right| = \int_{\underline{V}} \int_{x_{j_k} \in \Theta_{j_k}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \right|.$$

Using the Lebesgue measure of  $\Theta_i$  is  $|\theta_i - \tilde{\theta}_i|$  we then obtain,

$$\int_{\underline{V}} \int_{x_{j_k} \in \Theta_{j_k}} \left| \frac{\alpha}{x_{j_k}^{\alpha+1}} \right| \leq \int_{\underline{V}} \frac{|\theta_{j_k} - \tilde{\theta}_{j_k}|}{|\theta_{j_k}^\alpha * \tilde{\theta}_{j_k}^\alpha|}.$$

Finally by (A.6) we conclude that (A.3) is fulfilled.  $\square$

## A.5 Proof of equations (2.97) and (2.95)

*Proof of (2.95).* We have  $S_n = \sum_{i=1}^n I_i$ . In this case

$$\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(I_i) + 2 \sum_{i=1}^{n-1} \sum_{j>i}^n \text{Cov}(I_i, I_j),$$

where

$$\text{Var}(I_i) = \mathbb{E}(\text{Var}(I_i = 1 | \underline{V}_i)) = \mathbb{E}(P_i(1 - P_i)),$$

where  $P_i = P(I_i = 1 | \underline{V}_i)$  hence

$$\begin{aligned} \text{Var}(I_i) &= \mathbb{E}\left(K \sum_{j=1}^i \frac{1 + V_j}{2^{i-j}} \left(1 - K \sum_{j=1}^i \frac{1 + V_j}{2^{i-j}}\right)\right) \\ &= K(1 + q) \frac{2^i - 1}{2^{i-1}} \left(1 - K(1 + q) \frac{2^i - 1}{2^{i-1}}\right), \end{aligned}$$

hence

$$\begin{aligned} \sum_{i=1}^n \text{Var}(I_i) &= \frac{K}{3} (20K(1 + q^2) - 12nK(1 + q^2) + 40Kq + \frac{K}{4^{n-1}}(1 + q^2 + 2qK) + \\ &\quad \frac{3}{2^{n-1}}(1 + q - 4K - 8kq - 4kq^2) - 6(1 + q - n + 4Knq - nq)). \end{aligned}$$

In addition for all  $j > i$

$$\begin{aligned} \text{Cov}(I_i, I_j) &= \mathbb{E}\left(K^2 \sum_{\ell=1}^i \sum_{k=1}^j \frac{(1 + V_\ell)(1 + V_k)}{2^{i-\ell} 2^{j-k}}\right) - K^2 \mathbb{E}\left(\sum_{\ell=1}^i \frac{(1 + V_\ell)}{2^{i-\ell}}\right) \mathbb{E}\left(\sum_{k=1}^j \frac{(1 + V_k)}{2^{j-k}}\right) \\ &= \sum_{\ell=1}^i \sum_{k=1}^j K^2 \frac{\mathbb{E}(V_k V_\ell) - \mathbb{E}(V_k) \mathbb{E}(V_\ell)}{2^{(i+j)-(\ell+k)}} \\ &= \frac{4K^2 q(1 - q)}{3} \frac{(4^i - 1)}{2^{(i+j)}}. \end{aligned}$$

Hence,

$$\sum_{i=1}^{n-1} \sum_{j>i}^n \text{Cov}(I_i, I_j) = K^2 q(1 - q) \left(\frac{4n}{3} - \frac{8}{9} \frac{1}{4^n} - \frac{28}{9} + \frac{4}{2^n}\right).$$

Then

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(S_n)}{n} = \frac{8}{3} K^2 q(1 - q) + 2K(1 + q - 2K - 4Kq - 2Kq^2).$$

□

*Proof of (2.97).* In the case where the conditioning is of order one, i.e when  $P(I_i|V_i) = P(I_i|V_i) = K(1 + V_i)$  we have for all  $i \neq j$  the  $\text{Cov}(I_i, I_j) = 0$ . In fact, as the r.v's  $I_i, i \leq n$  take values in  $\{0, 1\}$  we have

$$\mathbb{E}(I_i) = \mathbb{E}(I_i^2) = \mathbb{E}(\mathbb{E}(I_i = 1|V_i)) = K(1 + \mathbb{E}(V_i)),$$

hence

$$\text{Var}(I_i) = \mathbb{E}(I_i^2) - \mathbb{E}^2(I_i) = K(1 + q)(1 - K(1 + q)),$$

$$\text{Cov}(I_i, I_j) = \mathbb{E}(I_i I_j) - \mathbb{E}(I_i)\mathbb{E}(I_j),$$

where by conditional independence we have

$$\begin{aligned} \mathbb{E}(I_i I_j) &= \mathbb{E}(\mathbb{E}(I_i I_j|V_i, V_j)) = \mathbb{E}(P(I_i = 1, I_j = 1|V_i, V_j)) = \mathbb{E}(P(I_i = 1|V_i)P(I_j = 1|V_j)) \\ &= K^2 \mathbb{E}((1 + V_i)(1 + V_j)) \\ &= K^2(1 + 2\mathbb{E}(V_1) + \mathbb{E}(V_i V_j)), \end{aligned}$$

hence by assumption of independence of r.v's  $V_i, i \leq n$  we obtain for all  $i \neq j$

$$\begin{aligned} \text{Cov}(I_i, I_j) &= \mathbb{E}(I_i I_j) - \mathbb{E}^2(I_i)\mathbb{E}^2(I_j) \\ &= K^2(1 + \mathbb{E}(V_i) + \mathbb{E}(V_j) + \mathbb{E}(V_i V_j)) - K^2(1 + \mathbb{E}(V_i))(1 + \mathbb{E}(V_j)) \\ &= K^2(\mathbb{E}(V_i V_j) - \mathbb{E}(V_i)\mathbb{E}(V_j)) = 0 \end{aligned}$$

In this case

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(S_n)}{n} = \text{Var}(I_1) = K(1 + q)(1 - K(1 + q)).$$

□



## Appendix B

### Simulations and some examples related to the first part

## B.1 Choice of $x$ in Theorems 2.3.2 and 2.3.3

Note that Theorems 2.3.2 and 2.3.3 hold for small value of  $x$  where we recall that in these Theorems  $\ell = \lceil \ln(n)/5K \rceil + x$ . The real number  $x$  ensures to have reasonable number of terms in the blocks  $S_j(\ell)$  (See Chapter 2, Section 2.3). Here we consider the example given in Chapter 2, Section 2.6.1. We estimate the variance of the sum  $S_k = \sum_{i=1}^k X_i$ , for all  $k = 1, \dots, n$ . The data that we consider are simulated with the following parameters:  $K = 0.24$ , and  $q = 0.6$ ,  $n = 1000$ ,  $m = 20$  where  $n$  is the number of period and  $m = 20$  is the number of Monte Carlo simulations.

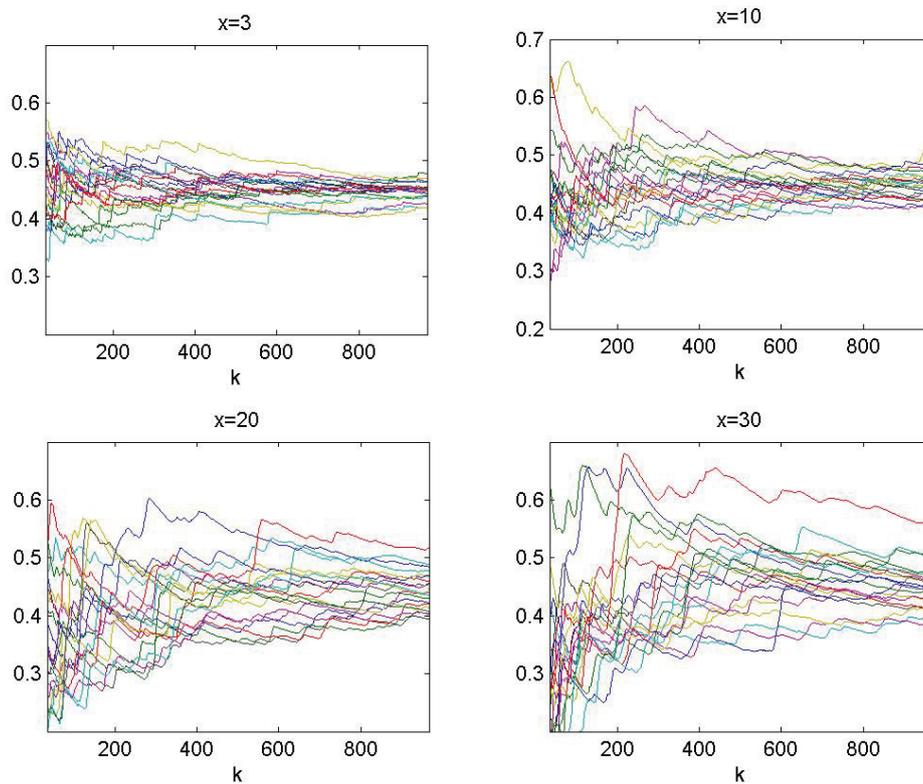


Figure B.1: Estimation of  $\sigma$  of the sum  $S_k$  over time  $k$  where  $k = 35, \dots, 1000$  for different values of  $x$ .

Graphically we can deduce that it is not interesting to have very large value for  $x$ .

## B.2 Discrete case: the factor $V$ is a Bernoulli r.v. with parameter $q$

### Behaviour of the variance in function of time

Consider the example given on Chapter 2, section 2.6.1. As we have seen in Chapter 2, when the factor is a sequence of independent Bernoulli r.v.'s with parameter  $q$ , the exact value of the variance of the sum is easy to compute. In Figure B.2 we present, for different choice of  $q$ , the time series  $(Var(I_i))_{i \in \mathbb{N}}$ .

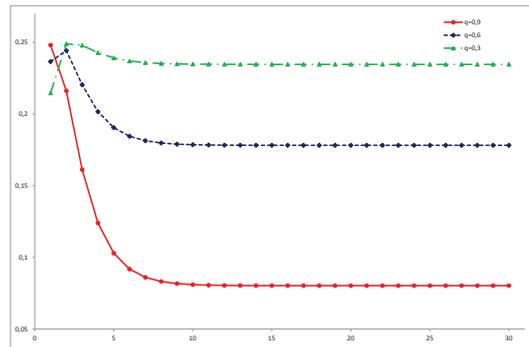


Figure B.2: Behaviour of  $Var(I_i)$ ,  $i = 1 \dots, 30$

This Figure shows how the individual variance  $Var(I_i)$  converges to a positive finite value as  $i \rightarrow \infty$ .

### Behaviour of the covariance processes in function of time

Consider the process  $(Cov(I_i, I_j))_{j > i}$ . For fixed  $i$ , we illustrate in Figure B.4 the behaviour of the covariance processes. It is clear from this figure that the covariance function decrease to zero as  $|i - j|$  is large enough. Figure B.4 illustrates how the covariance function increase in function of the variability of the factor. We recall that if  $q = 0.9$  then  $\sigma_V^2 = 0,09$ , if  $q = 0.6$  then  $\sigma_V^2 = 0.24$  and if  $q = 0.3$  then  $\sigma_V^2 = 0.21$ .

### Behaviour of the covariance processes in function of the variability of the factor

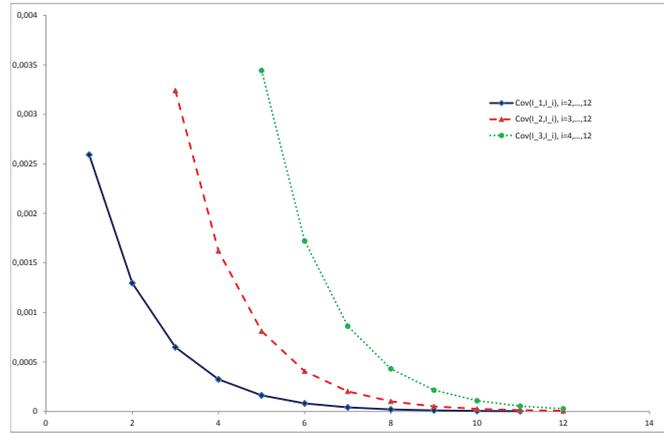


Figure B.3: Behaviour of  $Cov(I_i, I_j)$ ,  $i$  is fixed,  $j > i$ ,  $q = 0.9$

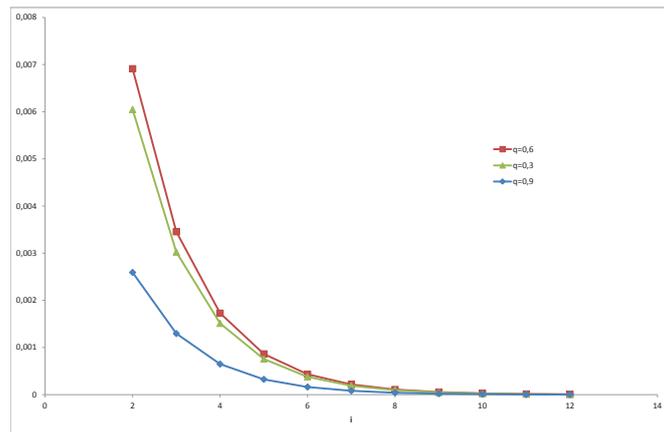


Figure B.4: Behaviour of  $Cov(I_1, I_i)$ ,  $i = 2, \dots, 12$ ,  $q = 0.9$ ,  $q = 0.6$ ,  $q = 0.3$

### B.3 Application when the factor $V$ is a Markov chain

Let  $(V_i)_{i \in \mathbb{N}}$  be a finite ergodic Markov chain with states  $\alpha = (1, \dots, s)$  and stationary probabilities  $\pi = \{\pi_1, \dots, \pi_s\}$  which admits an asymptotic distribution  $\pi^* = \lim_{n \rightarrow \infty} \pi^n$ . Consider an individual risk model in insurance such that the claim size is a Bernoulli random variable  $I$ . Consider that in each period  $1 \leq i \leq n$  the parameter of the Bernoulli random variable is given by  $P(I_i = 1 | \underline{V}_i) = \sum_{j=1}^i \frac{K(1+V_j)}{2^{i-j}}$  where  $\underline{V}_i = (V_1, \dots, V_i)$ . We have proved in Chapter 1 that with this choice of the conditional probability, the sequence  $(I_i)_{i \in \mathbb{N}}$  is mixing.

We want to prove by some simulation the asymptotic normality of  $S_n = \sum_{i=1}^n I_i$ . Consider the following numerical example in which the factor  $V_i$  is a Markov chain with two states  $s = \{1, 0\}$  and with the following transition matrix where 0.7 corresponds to the probability to be in state 0 and to stay in state 0

$$\begin{pmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{pmatrix}$$

The stationary distribution  $\pi$  is such that  $\pi_0 = 0.6667$  and  $\pi_1 = 0.3333$ .

We do some simulation for  $n = 1000$  and  $m = 1000$  where  $n$  is the number of period and  $m$  is the number of Monte Carlo simulations. From Figure B.5, we deduce graphically that the asymptotic distribution of the normalised sum is normal (the *JarqueBera* test confirms this graphical result).

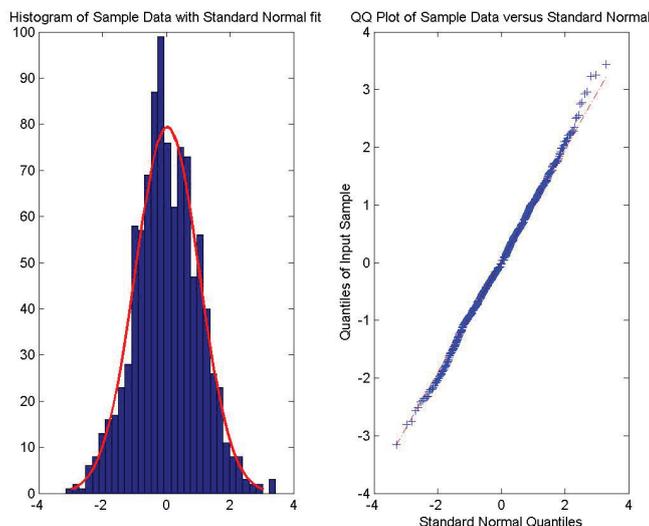


Figure B.5: Normfit and QQplot of the normalized sum  $S_n$  of the Bernoulli sample data,  $m = 1000$ ,  $n = 1000$ , (the factor is a Markov chain)

$n$	$(I_i V_i)$	$(I_i \underline{V}_i)$
100	0.1984	0.2322
300	0.2058	0.2521
1000	0.2156	0.2611

Table B.1: Estimation of the variance when the factor is a Markov chain,  $m = 1000$ ,  $n = 1000$ .

## Appendix C

# Proofs of some results of the second part

## C.1 Proof of Proposition 3.4.1

This bound is obtained as follows: we assume that

$$\begin{cases} P(\tilde{N}_{\max} = 0) = a, \\ P(\tilde{N}_{\max} = 1) = c, \\ P(\tilde{N}_{\max} = i) = b(n+1-i), \forall 2 \leq i \leq n+1, \end{cases} \quad (\text{C.1})$$

where  $a$ ,  $b$  and  $c$  are such that (C.1) is a true p.m.f. with fixed mean  $\nu$ . Moreover, note that if  $N$  is a r.v. with p.m.f. that is concave in  $[0, 2]$  and convex in  $[1, +\infty)$ , then by the constraint of convexity we must have

$$P(N = 2) - P(N = 1) \leq \frac{P(N = 3) - P(N = 1)}{2}$$

and by concavity constraint we have

$$P(N = 0) - P(N = 1) \leq \frac{P(N = 0) - P(N = 2)}{2}.$$

These constraints lead to

$$\begin{cases} bn \leq c, \\ a - c \leq c - b(n-1). \end{cases} \quad (\text{C.2})$$

Following Proposition 3.2.3 we have  $X \leq_{cx} Y$  if the p.m.f. of the r.v.  $Y$  has heavier values on extremes than those taken by the r.v.  $X$ . For this we shall consider that  $P(N = 0) - P(N = 1) = (P(N = 0) - P(N = 2))/2$ . Hence  $a$ ,  $b$  and  $c$  are determined by solving the following system

$$\begin{cases} a - c = c - b(n-1), \\ a + c + \sum_{i=2}^n b(n+1-i) = 1, \\ c + \sum_{i=2}^n b(n+1-i)i = \nu, \\ \text{where } 0 < bn < c < a < 1. \end{cases} \quad (\text{C.3})$$

Consider two r.v.'s  $M$  and  $\tilde{N}_{\max}$  in  $\mathbb{J}_n$  with p.m.f. concave in  $[0, 2]$  and convex in  $[1, +\infty)$ . Denote by  $P(M = 0) = a_1$ ,  $P(M = 1) = c_1$ ,  $P(M = 2) = d_1$  and  $P(\tilde{N}_{\max} = 2) = d$ .

1.  $a_1 < a$  and  $c_1 > c$

In this case we observe a crossing point in  $]0, 1[$ . By Proposition 3.2.1 there must exist, at least, another crossing point in  $]1, n]$ . Two cases arise:

- 1.1 If  $d_1 \leq d$ , then we have a crossing point in  $]1, 2[$ . In addition, by global convexity in  $[1, \infty)$  it cannot exist another crossing point in  $]2, n]$  (see Figure C.1 for illustration).
- 1.2 If  $d_1 > d$ , then by Lemma 3.3.1 we have a single crossing point in  $]2, n]$  (see Figure C.2 for illustration).

Note that in the previous two cases the p.m.f. of  $M$  and  $\tilde{N}_{\max}$  have exactly two crossing points in the set  $\mathcal{D}_n$ . In addition, near  $n$ ,  $P(M) \leq P(\tilde{N}_{\max})$ . Thus, by Proposition 3.2.3 we have  $M \leq_{cx} \tilde{N}_{\max}$ .

2.  $a_1 < a$  and  $c_1 = c$

Two cases arise:

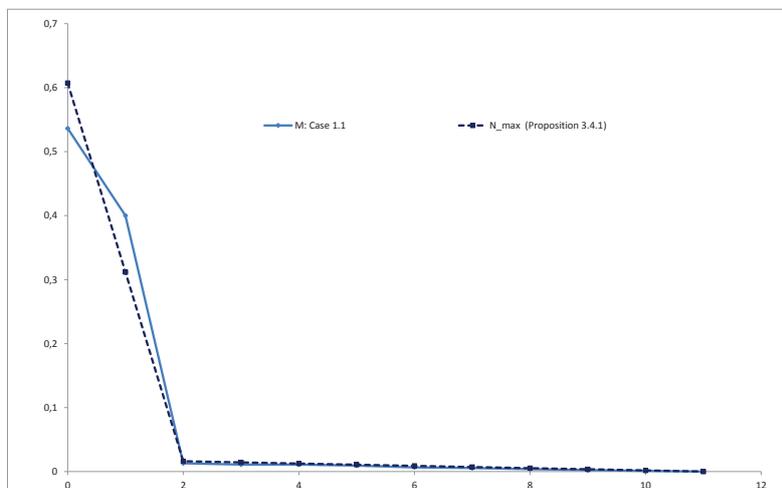


Figure C.1: The upper bound (3.15) vs the case 1.1,  $\nu = 0.69$ ,  $n = 10$ .

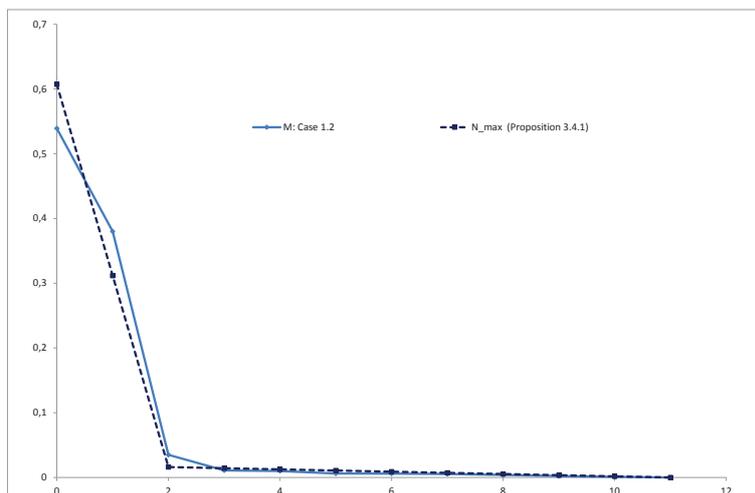


Figure C.2: The upper bound (3.15) vs the case 1.2,  $\nu = 0.69$ ,  $n = 10$ .

2.1 if  $d_1 > d$  then in this case we observe a first crossing point in  $]0, 2[$ . In addition, by Proposition 3.3.1 we have exactly one crossing point in  $]2, n]$  where near  $n$  we have  $P(M) \leq P(\tilde{N}_{\max})$  (see Figure C.3 for illustration).

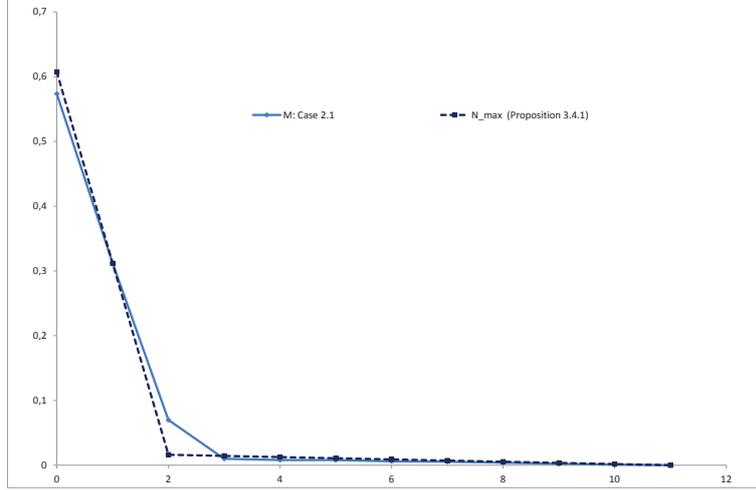


Figure C.3: The upper bound (3.15) vs the case 2.1,  $\nu = 0.69$ ,  $n = 10$ .

2.2 if  $d_1 \leq d$ , then the two crossing points must be in  $]2, n]$  which is not possible by Lemma 3.3.1. Hence, this case is absurd.

3.  $a_1 < a$  and  $c_1 < c$   
Two cases arise:

3.1 If  $d_1 > d$ , then the constraint of concavity is not fulfilled.

3.2 If  $d_1 \leq d$ , then it is clear that there is no crossing point until the point 2. In addition by Lemma 3.3.1, in order to ensure the constraint of convexity, there is no crossing between p.m.f. of  $M$  and  $\tilde{N}_{\max}$  in  $]2, n]$ . So, the constraint of equal mean is violated.

4.  $a_1 \geq a$  and  $c_1 \geq c$

In this case by Lemma 3.3.1 we have at most one crossing point over  $\mathcal{D}_n$ .

5.  $a_1 \geq a$  and  $c_1 \leq c$

In this case, to ensure the assumption of concavity, we must have  $d_1 < d$ . Hence we have a crossing point in  $]0, 2[$ . So, if a second crossing point exists, then it is necessarily in  $]2, n]$ . If it is the case, then by Lemma 3.3.1, the constraint of convexity is violated. So, this case is absurd.

## C.2 Proof of Proposition 3.4.3

The lower bound (3.19) is derived such that the distribution of  $\tilde{N}_{\min}$  has the following p.m.f. where  $a$  and  $b$  are two positive integers :

$$\begin{cases} P\left(\tilde{N}_{\min} = 0\right) = a - b, \\ P\left(\tilde{N}_{\min} = i\right) = a - bi, \forall 1 \leq i \leq \xi + 1, \\ P\left(\tilde{N}_{\min} = i\right) = 0, \forall i > \xi + 1, \\ \text{where } 0 < b < a < 1. \end{cases} \quad (\text{C.4})$$

Therefore,  $a$  and  $b$  are derived such that (C.4) is a true p.m.f. and such that the mean of the r.v.  $\tilde{N}_{\min}$  is equal to  $\nu$ .

$$\begin{cases} (a - b) + \sum_{i=1}^{\xi+1} (a - bi) = 1, \\ (a - b) + \sum_{i=2}^{\xi+1} (a - bi)i = \nu. \end{cases} \quad (\text{C.5})$$

Hence it follows that

$$a = \frac{2((2\xi + 3)(\xi + 2)(\xi + 1) - 3\nu(4 + \xi^2 + 3\xi))}{\xi(\xi + 5)(\xi + 2)(\xi + 1)}, \quad (\text{C.6})$$

and

$$b = \frac{6(\xi + 1 - 2\nu)}{\xi(\xi + 5)(\xi + 1)}, \quad (\text{C.7})$$

where  $0 < b < a < 1$  and  $\xi < 2\nu < \xi + 1$ .

Note that from (C.4) we have  $P(\tilde{N}_{\min} = 0) = P(\tilde{N}_{\min} = 1) = a - b$  and  $P(\tilde{N}_{\min} = 2) = a - 2b$ . So, the assumption of concavity in  $[0, 2]$  is valid.

Assume that  $P(\tilde{N}_{\min} = \xi + 1) = \pi_2$ . Moreover, we have  $P(\tilde{N}_{\min} = \xi + 2) = 0$ , then by constraint of convexity we must have  $P(\tilde{N}_{\min} = \xi) = \pi_1 + 2\pi_2$ . Hence, it follows that  $b = (\pi_1 + \pi_2)$  and  $a = \pi_2(2 + \xi) + \pi_1(\xi + 1)$ .

Denote by  $P(\tilde{N}_{\min} = 0) = P(\tilde{N}_{\min} = 1) = a^*$ ,  $P(M = 0) = a_1^*$ ,  $P(M = 1) = c_1$ ,  $P(M = 2) = d_1$ , and  $P(\tilde{N}_{\min} = 2) = d^*$ .

1.  $a_1^* > a^*$  and  $c_1 \geq a^*$

In this case we remark that in  $]0, 1[$  we have no crossing point. Now, assume that a first crossing point is observed in  $[1, \infty[$ . This means that there is  $j \in [1, \xi + 1[$  such that  $P(M = j) < P(\tilde{N}_{\min} = j)$ . Further, consider that a second crossing point exists then, this means that there is  $k \in [j + 1, \xi + 1[$  where  $j < k \leq \xi$  such that  $P(M = k) > P(\tilde{N}_{\min} = k)$ . If a third crossing point exists, then the constraint of convexity is violated. Hence at most, we have two crossing points where near  $n$ ,  $P(M) > P(\tilde{N}_{\min})$ . Then  $\tilde{N}_{\min} \leq_{cx} M$  (see Figure C.4 for illustration).

2.  $a_1^* > a^*$  and  $c_1 < a^*$

In this case we have a first crossing point in  $]0, 1[$ . By the constraint of convexity we have at most one crossing point in  $]1, \xi + 2[$ . Hence we are in a situation with exactly two crossing points. As  $a_1^* > a^*$  it follows immediately that near  $n$   $P(M) > P(\tilde{N}_{\min})$ . So, by Proposition 3.2.3,  $\tilde{N}_{\min} \leq_{cx} M$  (see Figure C.5 for illustration).

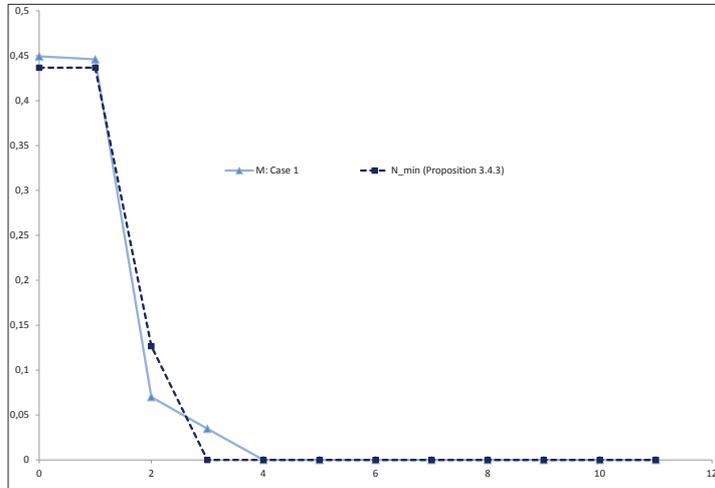


Figure C.4: The lower bound (3.19) vs the case 1,  $\nu = 0.69$ ,  $n = 10$ .

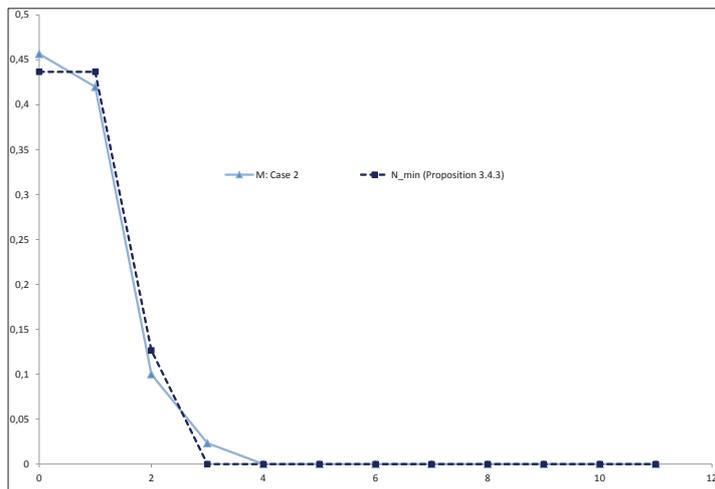


Figure C.5: The lower bound (3.19) vs the case 2,  $\nu = 0.69$ ,  $n = 10$ .

3.  $a_1^* \leq a^*$

In this case by the assumption of decreasingness we must have  $c_1 \leq a^*$ .

Three cases arise:

- 3.1 If  $P(M = \xi + 1) > P(\tilde{N}_{\min} = \xi + 1)$ , then, by the assumption of convexity we have at most one crossing point in the set  $\mathcal{D}_n$ . This is absurd under Proposition 3.2.1.
- 3.2 If  $P(M = \xi + 1) \leq P(\tilde{N}_{\min} = \xi + 1)$  and  $P(M = \xi + 2) > 0$ , then we observe a crossing point in  $]\xi + 1, \xi + 2[$ . In addition, by the assumption of convexity and according to Lemma 3.3.1 we have no crossing point in  $[0, \xi + 1]$ . So, this case is absurd because of the assumption of equal means which requires at least two crossing points.
- 3.3 If  $P(M = \xi + 1) \leq P(\tilde{N}_{\min} = \xi + 1)$  and  $P(M = \xi + 2) = 0$ , then under Lemma 3.3.1 we have no crossing point over  $\mathcal{D}_n$ . This is absurd.

### C.3 Proof of Proposition 3.5.1

This bound is obtained as follows:

$$\begin{cases} P(\tilde{N}_{\max} = 0) = a, \\ P(\tilde{N}_{\max} = i) = b(n + 1 - m), \forall 1 \leq i \leq m - 1, \\ P(\tilde{N}_{\max} = i) = b(n + 1 - i), \forall m \leq i \leq n + 1, \\ \text{where } 0 < b(n + 1 - m) < a < 1, \end{cases} \quad (\text{C.8})$$

such that

$$\begin{cases} a + (m - 1)b(n + 1 - m) + \sum_{i=m}^n b(n + 1 - i) = 1, \\ b(n + 1 - m) \sum_{i=1}^{m-1} i + \sum_{i=m}^n b(n + 1 - i)i = \nu, \end{cases} \quad (\text{C.9})$$

Next we present a reasoning by contradiction to prove Proposition 3.5.1. Let  $M$  be a random variable in  $\mathbb{J}_n$  with a p.m.f. that is decreasing globally on  $\mathbb{N}$  and convex from a fixed point  $m > 1$ , where  $0 < 3\nu \leq \frac{(m^2 + nm + m + n^2 + 2n)}{(m + n + 2)}$ . Denote by  $P(M = 0) = a_1$ ,  $P(M = 1) = c_1$ ,  $P(\tilde{N}_{\max} = 0) = a$  and  $P(\tilde{N}_{\max} = 1) = c$ . For this, we enumerate all possible p.m.f. that may be taken by the r.v.  $M$  and we prove that in all cases  $M \leq_{cx} \tilde{N}_{\max}$ .

1.  $a_1 < a$  and  $c_1 \leq c$

In this case by Lemma 3.3.1 there cannot get a crossing point on  $\mathcal{D}_n$  which is absurd.

2. If  $a_1 < a$  and  $c_1 > c$ , then we have a first crossing point in  $]0, 1[$ .

- 2.1 In addition, if  $P(M = m - 1) \leq P(\tilde{N}_{\max} = m - 1)$ , then by nonincreasingness we have necessarily, a crossing point in  $]1, m - 1[$ . In addition, under the Lemma 3.3.1 there is not any crossing point in  $]m, \infty)$  (for illustration see Figure C.6).
- 2.2 If  $P(M = m - 1) > P(\tilde{N}_{\max} = m - 1)$ , then by Lemma 3.3.1 there is at most one crossing point on  $]m, n[$ . Hence in both cases we have exactly two crossing points over  $\mathcal{D}_n$  where near  $n$  we have  $M \leq_{cx} \tilde{N}_{\max}$ .

3.  $a_1 \geq a$  and  $c_1 > c$

In this case by the same reasoning as in the previous case there cannot exist more than one crossing point over  $\mathcal{D}_n$  which is absurd. In fact, if  $P(M = m - 1) < P(\tilde{N}_{\max} = m - 1)$ , then we observe only one crossing point in  $]0, m - 1[$  and by Lemma 3.3.1 there is not another crossing point on  $]m - 1, n[$ . If  $P(M = m - 1) \geq P(\tilde{N}_{\max} = m - 1)$ , then by Lemma 3.3.1 there cannot exist more than one crossing over  $\mathcal{D}_n$ .

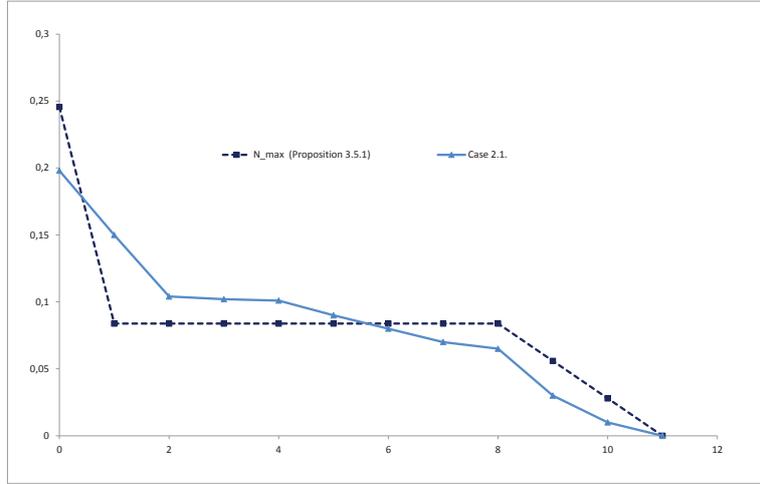


Figure C.6: The upper bound (3.21) vs the case 2.1,  $\nu = 0.69$ ,  $n = 10$ .

4.  $a_1 \geq a$  and  $c_1 \leq c$

In this case we observe a first crossing point in  $]0, 1[$  but, by the assumption of decreasingness and according to lemma 3.3.1 there is no other crossing point over  $\mathcal{D}_n$  which is absurd.  $\diamond$

## C.4 Proof of Proposition 3.5.4

This bound is obtained as follows: we assume that

$$\begin{cases} P(\tilde{N}_{\min} = i) = a, & \forall 0 \leq i \leq \xi, \\ P(\tilde{N}_{\min} = \xi + 1) = b, \\ P(\tilde{N}_{\min} = i) = 0, & \forall i \geq \xi + 2, \end{cases} \quad (\text{C.10})$$

where a and b are such that this is a true p.m.f., with fixed mean  $\nu$  i.e.

$$\begin{cases} (\xi + 1) * a + b = 1, \\ \xi(\xi + 1) * a/2 + (\xi + 1) * b = \nu, \end{cases} \quad (\text{C.11})$$

Mention that we don't have any constraint related with convexity below point  $\xi + 1$  as  $m \geq \xi + 1 > 1$ . Denote by  $P(M = 0) = a_1$  and by  $P(\tilde{N}_{\min} = 0) = a$ .

1.  $a_1 > a$  and  $P(M = \xi) < P(\tilde{N}_{\min} = \xi)$

In this case, by assumption of nonincreasingness, we have exactly one crossing point on  $]0, \dots, \xi[$ . Two cases arise:

1.1  $P(M = \xi + 1) > P(\tilde{N}_{\min} = \xi + 1)$  (see Figure C.7 for illustration).

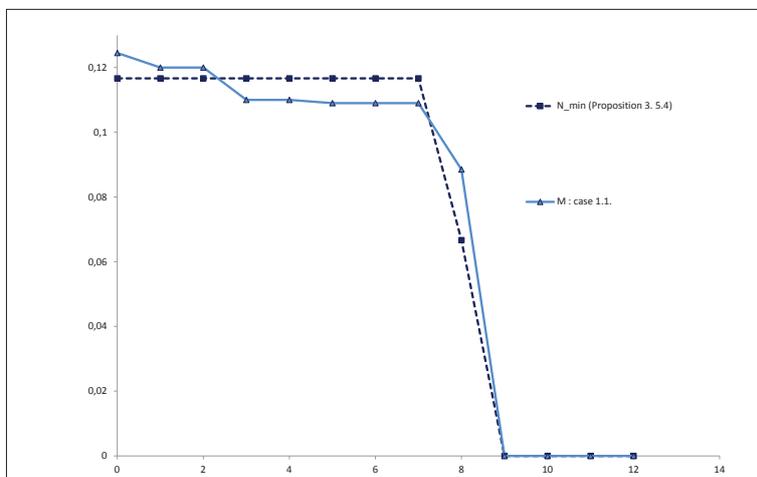


Figure C.7: The lower bound (3.23) vs the case 1.1 ,  $m = 8$ ,  $\nu = 3.8$ ,  $n = 10$ .

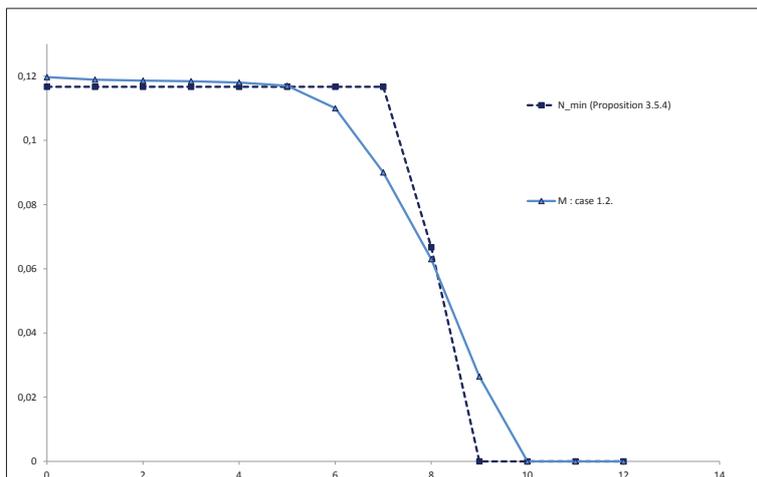


Figure C.8: The lower bound (3.23) vs the case 1.2. ,  $m = 8$ ,  $\nu = 3.8$ ,  $n = 10$ .

1.2  $P(M = \xi + 1) < P(\tilde{N}_{\min} = \xi + 1)$  (see Figure C.8 for illustration).

Note that in these two cases we have at most two crossing points. In addition in the presence of two crossing, we have  $P(M) \geq P(\tilde{N}_{\min})$  near  $n$ . Hence,  $\tilde{N}_{\min} \leq_{cx} M$ .

2.  $a_1 > a$  and  $P(M = \xi) > P(\tilde{N}_{\min} = \xi)$

In this case we have at most two crossing points over  $\mathcal{D}_n$ . More precisely this case arises if and only if we have in addition  $P(M = \xi + 1) < P(\tilde{N}_{\min} = \xi + 1)$  and  $P(M = \xi + 2) > 0$ , (see Figure C.9 for illustration). Thus in this case  $P_{\tilde{N}_{\min}} \leq P_M$  near  $n$ .

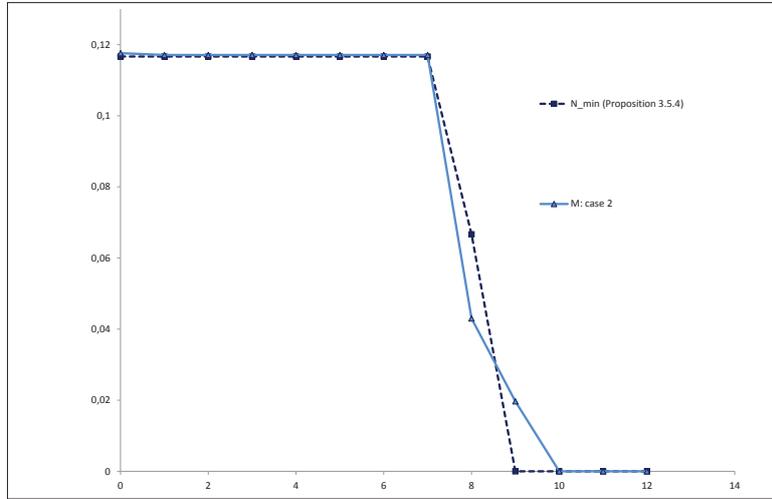


Figure C.9: The lower bound (3.23) vs the case 2.,  $m = 8$ ,  $\nu = 3.8$ ,  $n = 10$ .

So, by Proposition 3.2.3 we conclude that  $\tilde{N}_{\min} \leq_{cx} M$ .

It is clear that if  $P(M = \xi) = P(\tilde{N}_{\min} = \xi)$ , then we have at most one crossing which is absurd under Proposition 3.2.1.

3.  $a_1 \leq a$

We note that we have at most one crossing point over  $\mathcal{D}_n$  which is absurd following Proposition 3.2.1.

## C.5 Proof of Proposition 3.5.5

The distribution of  $\tilde{N}_{\min}$  is obtained as explained in the following optimisation program:

$$\begin{cases} P(\tilde{N}_{\min} = i) = a - bm, & \forall 0 \leq i \leq m, \\ P(\tilde{N}_{\min} = i) = a - bi, & \forall m + 1 \leq i \leq \xi + 1, \\ P(\tilde{N}_{\min} = i) = 0, & \forall i > \xi + 1, \\ \text{where } 0 < b < a < 1, \end{cases} \quad (\text{C.12})$$

where

$$\begin{cases} (m+1)(a-bm) + \sum_{i=m+1}^{\xi+1} (a-bi) = 1, \\ m(m+1)(a-bm)/2 + \sum_{i=m+1}^{\xi+1} (a-bi)i = \nu. \end{cases} \quad (\text{C.13})$$

It follows that

$$a = \frac{2(2\xi^3 + 9\xi^2 + 13\xi + 6 + m^3 - m - 3\nu m^2 - 3\nu m - 6\nu - 3\nu\xi^2 - 9\nu\xi)}{(\xi+2)(\xi^3 + 6\xi^2 - 3m^2\xi - 3m\xi + 11\xi + 6 - 3m^2 - 5m + 2m^3)}$$

and

$$b = \frac{6(\xi+1-2\nu)}{\xi^3 + 6\xi^2 - 3m^2\xi - 3m\xi + 11\xi + 6 - 3m^2 - 5m + 2m^3}.$$

Since (C.12), we have  $P(\tilde{N}_{\min} = \xi + 2) = 0$  and  $P(\tilde{N}_{\min} = \xi + 1) = \pi_2$ . In addition, to ensure the constraint of convexity, we consider that  $P(\tilde{N}_{\min} = \xi) = \pi_1 + 2\pi_2$  where  $\pi_1 \geq 0$ . Hence, we can deduce that  $b = (\pi_1 + \pi_2)$  and  $a = \pi_2(2 + \xi) + \pi_1(\xi + 1)$ .

Denote by  $P(M = m) = a_1$  and by  $P(\tilde{N}_{\min} = m) = a$ .

1.  $a_1 > a$

In this case there are, at most, two crossing points near  $n$   $P_{\tilde{N}_{\min}} \leq P_M$ . Thus, by Proposition 3.2.3,  $\tilde{N}_{\min} \leq_{cx} M$ . See Figure C.10 for illustration.

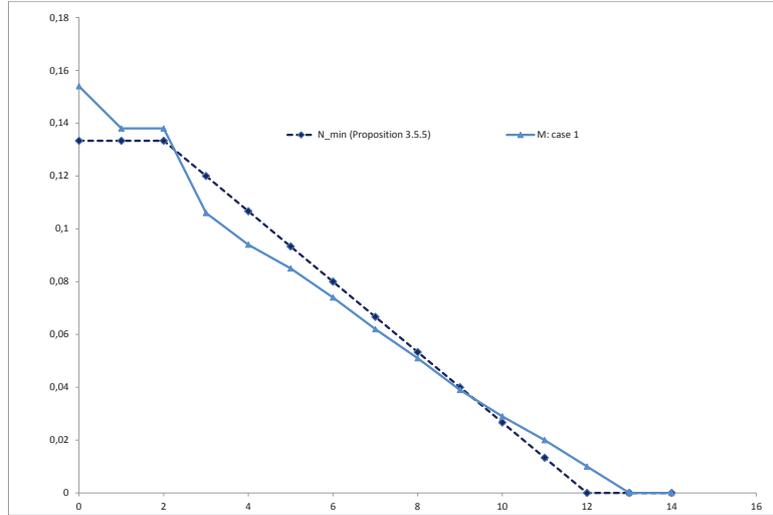


Figure C.10: The lower bound (3.25) vs the case 1,  $m = 2$ ,  $\nu = 3.8$ ,  $n = 14$ .

2.  $a_1 < a$

According to the assumptions of nonincreasingness and convexity, there exists, at most, two crossing over  $\mathcal{D}_n$  such that near  $n$   $P_{\tilde{N}_{\min}} \leq P_M$ . Thus, by Proposition 3.2.3,  $\tilde{N}_{\min} \leq_{cx} M$ . See Figure C.11 for illustration.

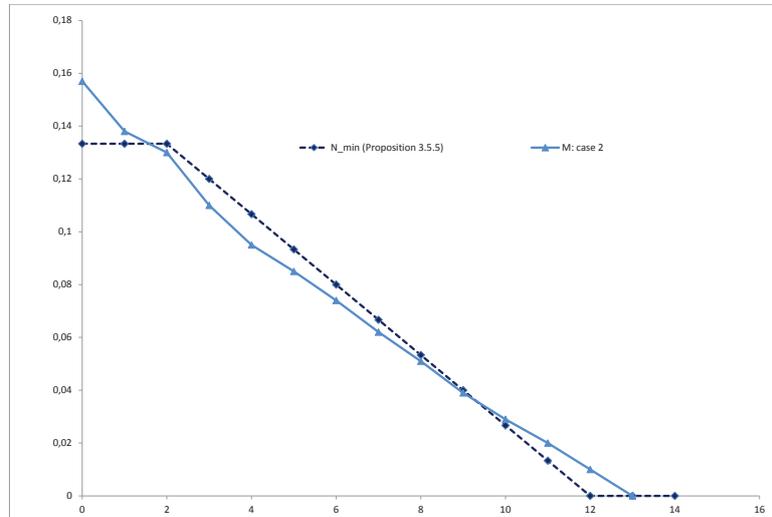


Figure C.11: The lower bound (3.25) vs case 2,  $m = 2$ ,  $\nu = 3.8$ ,  $n = 14$ .

3. If  $a_1 = a$ , then by convexity and nonincreasingness it is clear that there exists, at most, one crossing which is absurd under Proposition 3.2.1.

# Bibliographie

- [1] AHMED, A., LEO, R. et PROSCHAN, F. (1981). Generalized association, with application in multivariate statistics. *Annals of Statistics*, 9:168–176.
- [2] ALBRECHER, H. et BOXMA, O. J. (2004). A ruin model with dependence between claim sizes and claim intervals. *Insurance Math. Econom.*, 35(2):245–254.
- [3] ALBRECHER, H., CONSTANTINESCU, C. et LOISEL, S. (2011). Explicit ruin formulas for models with dependence among risks. *Insurance Math. Econom.*, 48(2):265–270.
- [4] ALBRECHER, H. et TEUGELS, J. L. (2006). Exponential behavior in the presence of dependence in risk theory. *J. Appl. Probab.*, 43(1):257–273.
- [5] AMBAGASPITIYA, R. S. (1999). On the distributions of two classes of correlated aggregate claims. *Insurance: Mathematics & Economics*, 24(3):301–308.
- [6] ASMUSSEN, S. (1989). Risk theory in a Markovian environment. *Scandinavian Actuarial Journal*, (2):69–100.
- [7] ASMUSSEN, S. (2000). *Ruin probabilities*, volume 2 de *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ.
- [8] ASMUSSEN, S. et ALBRECHER, H. (2010). *Ruin probabilities*. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second édition.
- [9] BAUERLE, N. et MULLER, A. (1998). Modeling and comparing dependences in multivariate risk portfolio. *Astin Bulletin.*, pages 59–76.
- [10] BERNSTEIN, S. (1927). Sur l’extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.*, 97(1):1–59.
- [11] BIARD, R., LEFÈVRE, C., LOISEL, S. et NAGARAJA, H. (2011). Asymptotic finite-time ruin probabilities for a class of path-dependent heavy-tailed claim amounts using poisson spacings. *J. Math. Anal. Appl.*, 27(5):503–518.
- [12] BILLINGSLEY, P. (1968). *Convergence of probability measures*. New York.
- [13] BIRKEL, T. (1988). Moment bounds for associated sequences. *Ann. Probab.*, 16(3): 1184–1193.

- [14] BLUM, J. R., HANSON, D. L. et KOOPMANS, L. H. (1963). On the strong law of large numbers for a class of stochastic processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:1–11.
- [15] BOKUCHAVA, I. V. (1978). The central limit theorem for conditionally independent sequences. *Soobshch. Akad. Nauk Gruzin. SSR*, 91(2):305–308.
- [16] BOKUCHAVA, I. V. (1982). *The central limit theorem for conditionally independent sequences*. “Metsniereba”, Tbilisi.
- [17] BOKUCHAVA, I. V., KVATADZE, Z. A. et SHERVASHIDZE, T. L. (1987). *On limit theorems for random vectors controlled by a Markov chain*. VNU Sci. Press, Utrecht.
- [18] BOKUCHAVA, I. V. et SHERVASHIDZE, T. L. (1984). *Limit theorems for conditionally independent random vectors*. “Metsniereba”, Tbilisi.
- [19] BOUDREAULT, M., COSSETTE, H., LANDRIAULT, D. et MARCEAU, E. (2006). On a risk model with dependence between interclaim arrivals and claim sizes. *Scand. Actuar. J.*, (5):265–285.
- [20] BOWERS, N., GERBER, H., HICKMAN, J., JONES, D. et NESBITT, C. (1997). Actuarial mathematics. *Society of Actuaries, Schaumburg, IL*.
- [21] BRADLEY, R. C. (1986). *Basic properties of strong mixing conditions*, volume 11 de *Progr. Probab. Statist.* Birkhäuser Boston, Boston, MA.
- [22] BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144. Update of, and a supplement to, the 1986 original.
- [23] BRADLEY, R. C. (2007). *Introduction to Strong Mixing Conditions*. Kendrick Press, Heber City, Utah.
- [24] BRADLEY, R. C. et BRYC, W. (1985). Multilinear forms and measures of dependence between random variables. *J. Multivariate Anal.*, 16(3):335–367.
- [25] BRILLINGER, D. R. (1962). A note on the rate of convergence of a mean. *Biometrika*, 49:574–576.
- [26] BRYC, W. et SMOLEŃSKI, W. (1993). Moment conditions for almost sure convergence of weakly correlated random variables. *Proc. Amer. Math. Soc.*, 119(2):629–635.
- [27] BULINSKIĬ, A. V. (1987). Limit theorems under weak dependence conditions. *In Probability theory and mathematical statistics, Vol. I (Vilnius, 1985)*, pages 307–326. VNU Sci. Press, Utrecht.
- [28] CAI, J. et DICKSON, D. C. M. (2004). Ruin probabilities with a Markov chain interest model. *Insurance Math. Econom.*, 35(3):513–525.
- [29] CARLSTEIN, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.*, 14(3):1171–1179.
- [30] CHARPENTIER, A. (2011). La loi des grands nombres et le théorème central limite comme base de l’assurabilité? *Risques*, 86:136–141.

- 
- [31] COGBURN, R. (1960). Asymptotic properties of stationary sequences. *Univ. California Publ. Statist.*, 3:99–146.
- [32] CONSTANTINESCU, C., HASHORVA, E. et JI, L. (2011). Archimedean copulas in finite and infinite dimensions - with application to ruin problems. *Insurance Math. Econom.*, 49:487–495.
- [33] COSSETTE, H., GAILLARDETZ, P. et MARCEAU, E. (2002a). Common mixture in the individual risk model. *Schweiz. Aktuarver. Mitt.*, (2):131–157.
- [34] COSSETTE, H., GAILLARDETZ, P., MARCEAU, E. et RIOUX, J. (2002b). On two dependent individual risk models. *Insurance Math. Econom.*, 30(2):153–166.
- [35] COSSETTE, H., LANDRIAULT, D. et MARCEAU, E. (2004). Compound binomial risk model in a Markovian environment. *Insurance Math. Econom.*, 35(2):425–443.
- [36] COSSETTE, H. et MARCEAU, E. (2000). The discrete-time risk model with correlated classes of business. *Insurance: Mathematics & Economics*, 26(2-3):133–149.
- [37] COULON-PRIEUR, C. et DOUKHAN, P. (2000). A triangular central limit theorem under a new weak dependence condition. *Statist. Probab. Lett.*, 47(1):61–68.
- [38] COURTOIS, C., DENUIT, M. et VAN BELLEGEM, S. (2006). Discrete  $s$ -convex extremal distributions: theory and applications. *Applied Mathematics Letters. An International Journal of Rapid Publication*, 19(12):1367–1377.
- [39] COUSIN, A. (2008). Analyse du risque et couverture des tranches de CDO synthétique. *Thèse de l'université  $\frac{1}{2}$  Claude Bernard Lyon 1*.
- [40] CRAMÈR, H. (1930). *On the Mathematical Theory of Risk*. Skandia Jubilee Volume, Stockholm.
- [41] DAVYDOV, J. A. (1968). The convergence of distributions generated by stationary stochastic processes. *Theory of Probability and Its Application*, 13:691–696.
- [42] DAVYDOV, J. A. (1969). The strong mixing property for Markov chains with a countable number of states. *Doklady Akademii Nauk SSSR*, 187:252–254.
- [43] DAVYDOV, J. A. (1970). The invariance principle for stationary processes. *Theory of Probability and Its Application*, 3:487–498.
- [44] DE VYLDER, E. (1996). *Advanced risk theory a self-contained introduction Actuariat*. Éd. de l'Université de Bruxelles.
- [45] DEDECKER, J. et DOUKHAN, P. (2003). A new covariance inequality and applications. *Stochastic Process. Appl.*, 106(1):63–80.
- [46] DEDECKER, J., DOUKHAN, P., LANG, G., LEÓN R., J. R., LOUHICHI, S. et PRIEUR, C. (2007). *Weak dependence: with examples and applications*, volume 190 de *Lecture Notes in Statistics*. Springer, New York.
- [47] DEDECKER, J. et PRIEUR, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields*, 132(2):203–236.

- [48] DEHLING, H., DENKER, M. et PHILIPP, W. (1986). Central limit theorems for mixing sequences of random variables under minimal conditions. *Ann. Probab.*, 14(4):1359–1370.
- [49] DENKER, M. (1986). *Uniform integrability and the central limit theorem for strongly mixing processes*, volume 11 de *Progr. Probab. Statist.* Birkhäuser Boston, Boston, MA.
- [50] DENUIT, M. et CHARPENTIER, A. (2004). *Mathématiques de l'assurance non vie, volume 1.* Economica, Paris.
- [51] DENUIT, M. et CHARPENTIER, A. (2005). *Mathématiques de l'assurance non vie, volume 2.* Economica, Paris.
- [52] DENUIT, M., DE VYLDER, E. et LEFÈVRE, C. (1999a). Extremal generators and extremal distributions for the continuous  $s$ -convex stochastic orderings. *Insurance: Mathematics & Economics*, 24(3):201–217.
- [53] DENUIT, M., DHAENE, J., GOOVAERTS, M. J. et KAAS, R. (2005). Actuarial Theory for Dependent Risks : Measures, Orders and Models.
- [54] DENUIT, M. et FROSTIG, E. (2008). Comparison of dependence in factor models with application to credit risk portfolios. 22(1):151–160.
- [55] DENUIT, M. et FROSTIG, E. (2009). Dependence in failure times due to environmental factors. *Statistics & Probability Letters*, 79(4):487–495.
- [56] DENUIT, M. et LEFÈVRE, C. (1997). Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance Math. Econom.*, 20(3):197–213.
- [57] DENUIT, M., LEFÈVRE, C. et MESFIOUI, M. (1999b). On  $s$ -convex stochastic extrema for arithmetic risks. *Insurance: Mathematics & Economics*, 25(2):143–155.
- [58] DENUIT, M., LEFÈVRE, C. et SHAKED, M. (1998). The  $s$ -convex orders among real random variables, with applications. *Mathematical Inequalities & Applications*, 1(4):585–613.
- [59] DENUIT, M., LEFÈVRE, C. et SHAKED, M. (2000). Stochastic convexity of the Poisson mixture model. *Methodology and Computing in Applied Probability*, 2(3):231–254.
- [60] DENUIT, M., LEFÈVRE, C. et UTEV, S. (1999c). Stochastic orderings of convex/concave-type on an arbitrary grid. *Mathematics of Operations Research*, 24(4): 835–846.
- [61] DENUIT, M., LEFÈVRE, C. et UTEV, S. (2002). Measuring the impact of dependence between claims occurrences. *Insurance: Mathematics and Economics*, 30:1–19.
- [62] DICKSON, D. C. M. (2005). *Insurance risk and ruin.* International Series on Actuarial Science. Cambridge University Press, Cambridge.
- [63] DOOB, J. L. (1990). *Stochastic processes.* John Wiley & Sons Inc., New York.
- [64] DOUKHAN, P. (1994). *Mixing properties and examples*, volume 85 de *Lecture Notes in Statistics.* Springer-Verlag, New York. Properties and examples.

- 
- [65] DOUKHAN, P., FERMANIAN, J.-D. et LANG, G. (2009). An empirical central limit theorem with applications to copulas under weak dependence. *12(1):65–87*.
- [66] DOUKHAN, P., JAKUBOWICZ, J. et LEÓN, J. R. (2010). Variance estimation with applications. *In Dependence in probability, analysis and number theory*, pages 203–231. Kendrick Press, Heber City, UT.
- [67] DOUKHAN, P. et LOUHICHI, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications.*, 84(12):313–342.
- [68] DOUKHAN, P., MASSART, P. et RIO, E. (1994). The functional central limit theorem for strongly mixing processes. *Annales de l’Institut Henri Poincaré. Probabilités et Statistiques*, 30(1):63–82.
- [69] DOUKHAN, P. et PORTAL, F. (1983). Moments de variables aléatoires mélangeantes. *C. R. Acad. Sci. Paris Sér. I Math.*, 297(2):129–132.
- [70] DOUKHAN, P. et WINTENBERGER, O. (2007). An invariance principle for weakly dependent stationary general models. *Probability and Mathematical Statistics*, 27(1):45–73.
- [71] DROUET MARI, D. et KOTZ, S. (2001). *Correlation and dependence*. Imperial College Press, London.
- [72] DZIUBDZIELA, W. (1986). Limit laws for  $k$ th order statistics from conditionally mixing arrays of random variables. *J. Appl. Probab.*, 23(3):679–687.
- [73] EFRON, B. (1982). *The jackknife, the bootstrap and other resampling plans*, volume 38. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa.
- [74] EMBRECHTS, P., MCNEIL, A. J. et FREY, R. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press, Princeton.
- [75] ESARY, J. D., PROSCHAN, F. et WALKUP, D. W. (1967). Association of random variables, with applications. 38:1466–1474.
- [76] FIGER, C. C. (2001). The one factor creditmetrics model in the new basel capital accord. *RiskMetrics Journal*, 2(1):9–18.
- [77] FISHBURN, P. C. et LAVALLE, I. H. (1995). Stochastic dominance on unidimensional grids. *Mathematics of Operations Research*, 20(3):513–525.
- [78] FREES, E. W. et VALDEZ, E. A. (1998). Understanding relationships using copulas. *N. Am. Actuar. J.*, 2(1):1–25.
- [79] FREY, R. et MCNEIL, A. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6:59–92.
- [80] FROSTIG, E. (2003). Ordering ruin probabilities for dependent claim streams. *Insurance: Mathematics & Economics*, 32(1):93–114.
- [81] GERBER, H. (1979). *An introduction to mathematical risk theory*. Foundation Monograph Series. University of Pennsylvania Wharton School S.S. Huebner.

- [82] GERBER, H. U. (1988). Mathematical fun with ruin theory. *Insurance Math. Econom.*, 7(1):15–23.
- [83] GOOVAERTS, M. J. et DHAENE, J. (1996). The compound Poisson approximation for a portfolio of dependent risks. *Insurance Math. Econom.*, 18(1):81–85.
- [84] GOOVAERTS, M. J., KAAS, R. et van HEERWAARDEN, A. E. (1998). *Ordering of actuarial risks*. Caire, Brussels.
- [85] GOOVAERTS, M. J., KAAS, R., van HEERWAARDEN, A. E. et BAUWELINCKX, T. (1990). Effective actuarial methods. 3:xiv+316.
- [86] GORDIN, M. I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR*, 188:739–741.
- [87] GORDIN, M. I. (2004a). A remark on the martingale method for proving the central limit theorem for stationary sequences. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 311.
- [88] GORDY, M. B. (2003). A risk factor model foundation for ratings based bank capital rules. *Journal of Financial Intermediation*, 12(3):199–232.
- [89] GRIGORESCU, S. et OPRISAN, G. (1976). Limit theorems for  $J - X$  processes with a general state space. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(1):65–73.
- [90] HALL, P. et HEYDE, C. C. (1980). *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.
- [91] HEILMANN, W.-R. (1988). *Fundamentals of risk theory*. Verlag Versicherungswirtschaft e.V., Karlsruhe.
- [92] HERRNDORF, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153.
- [93] HERRNDORF, N. (1985). A functional central limit theorem for strongly mixing sequences of random variables. 69(4):541–550.
- [94] HIRSCHFELD, O. (1935). A connection between correlation and contingency. 31:520–524.
- [95] HOEFFDING, W. (1955). The extrema of the expected value of a function of independent random variables. *Annals of Mathematical Statistics*, 26:268–275.
- [96] HURLIMANN, W. (1999). Extremal moment methods and stochastic orders. *Application in actuarial sciences, Monograph manuscript*.
- [97] IBRAGIMOV, I. A. (1959). Some limit theorems for stochastic processes stationary in the strict sense. *Dokl. Akad. Nauk SSSR*, 125:711–714.
- [98] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Teor. Veroyatnost. i Primenen.*, 7:361–392.

- 
- [99] IBRAGIMOV, I. A. (1963). A central limit theorem for a class of dependent random variables. *Teor. Verojatnost. i Primenen.*, 8:89–94.
- [100] IBRAGIMOV, I. A. (1975). A note on the central limit theorem for dependent random variables. *Theory of probability and its applications*, XX(1):135–141.
- [101] IBRAGIMOV, I. A. et LINNIK, Y. V. (1971). *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.
- [102] IBRAGIMOV, I. A. et ROZANOV, Y. A. (1978). *Gaussian random processes*, volume 9 de *Applications of Mathematics*. Springer-Verlag, New York. Translated from the Russian by A. B. Aries.
- [103] IOSIFESCU, M. (1980). *Recent advances in mixing sequences of random variables*. Bcdprime lgar. Akad. Nauk, Sofia.
- [104] IOSIFESCU, M. et THEODORESCU, R. (1969). *Random processes and learning*. Springer-Verlag, New York. Die Grundlehren der mathematischen Wissenschaften, Band 150.
- [105] JANSEN, K., HAEZENDONCK, J. et GOOVAERTS, M. J. (1986). Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance: Mathematics & Economics*, 5(4):315–334.
- [106] JIANG, X. et HAHN, M. (2008). A self-normalized central limit theorem for  $\rho$ -mixing stationary sequences. *Statist. Probab. Lett.*, 78(12):1541–1547.
- [107] JOE, H. (1997). *Multivariate models and dependence concepts*, volume 73. Chapman & Hall.
- [108] KABANOV, Y. M. (2009). Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics.*, 2:237–248.
- [109] KACEM, M., LEFÈVRE, C. et LOISEL, S. (2013a). Convex extrema for nonincreasing discrete distributions: effects of convexity constraints. *Working paper*.
- [110] KACEM, M., LOISEL, S. et MAUME DESCHAMPS, V. (2013b). Some mixing properties of conditionally independent processes. *Submitted to communication in statistics: theory and method*.
- [111] KACEM, M. et MAUME DESCHAMPS, V. (2013). Self normalised central limit theorem for some mixing processes. *Working paper*.
- [112] KARLIN, S. et STUDDEN, W. J. (1966). *Tchebycheff systems: With applications in analysis and statistics*. Interscience Publishers John Wiley & Sons, New York-London-Sydney.
- [113] KEMPERMAN, J. H. B. (1968). The general moment problem, a geometric approach. *Ann. Math. Statist*, 39:93–122.

- [114] KIM, T. Y. (1994). Moment bounds for non-stationary dependent sequences. *J. Appl. Probab.*, 31(3):731–742.
- [115] KLUGMAN, S. A., PANJER, H. H. et WILLMOT, G. E. (2008). *Loss models*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Hoboken, NJ, third édition. From data to decisions.
- [116] KOCHERLAKOTA, S. et KOCHERLAKOTA, K. (1992). *Bivariate discrete distributions*, volume 132 de *Statistics: Textbooks and Monographs*. Marcel Dekker Inc., New York.
- [117] KOLMOGOROV, A. N. et ROZANOV, J. A. (1960). On a strong mixing condition for stationary Gaussian processes. *Teor. Veroyatnost. i Primenen.*, 5:222–227.
- [118] KVATADZE, Z. A. et SHERVASHIDZE, T. L. (1988). On limit theorems for conditionally independent random variables controlled by a finite Markov chain. 1299:250–258.
- [119] LEFÈVRE, C. et LOISEL, S. (2010). Stationary-excess operator and convex stochastic orders. *Insurance: Mathematics & Economics*, 47(1):64–75.
- [120] LEFÈVRE, C. et LOISEL, S. (2012). On multiply monotone risk distribution. *Accepted, to appear in Journal of Applied Probability*.
- [121] LEFÈVRE, C. et PICARD, P. (1993). An unusual stochastic order relation with some applications in sampling and epidemic theory. *Advances in Applied Probability*, 25(1):63–81.
- [122] LEFÈVRE, C. et UTEV, S. (1995). Poisson approximation for the final state of a generalized epidemic process. *The Annals of Probability*, 23(3):1139–1162.
- [123] LEFÈVRE, C. et UTEV, S. (1996). Comparing sums of exchangeable bernoulli random variables. 33:285–310.
- [124] LEHTONEN, T. et NYRHINEN, H. (1992). On asymptotically efficient simulation of ruin probabilities in a Markovian environment. *Scand. Actuar. J.*, (1):60–75.
- [125] LIN, Z. et LU, C. (1996). *Limit theory for mixing dependent random variables*, volume 378 de *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht.
- [126] LOUHICHI, S. (2002). Moment inequalities for sums of certain dependent random variables. *Teor. Veroyatnost. i Primenen.*, 47(4):747–763.
- [127] LUNDBERG, F. (1903). *I. Approximerad framställning av sannolikhetsfunktioner: II. Aterforsakring av kollektivrisker*. Almqvist & Wiksell, Uppsala.
- [128] LUNDBERG, F. (1926). *Forsakringsteknisk Riskutjamning*. F. Englund's Boktryckeri AB, Stockholm.
- [129] MARCEAU, E. (2013). *Modélisation et évaluation quantitative des risques en actuariat : Modèle sur une période*. Springer .
- [130] MARKOWITZ, H. (1952). Portfolio selection. *Journal of Finance*, (VII):77–91.
- [131] MARSHALL, A. W. et OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, 62:30–44.

- 
- [132] MARSHALL, A. W. et OLKIN, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association*, 83(403):834–841.
- [133] MAXWELL, M. et WOODROOFE, M. (1997). A local limit theorem for hidden Markov chains. *Statist. Probab. Lett.*, 32(2):125–131.
- [134] MCLEISH, D. L. (1975). Invariance principles for dependent variables. *Probability Theory and Related Fields*, 32:165–178.
- [135] MCNEIL, A. J., FREY, R. et EMBRECHTS, P. (2005). *Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ. Concepts, techniques and tools.
- [136] MENG, Q., ZHANG, X. et GUO, J. (2008). On a risk model with dependence between claim sizes and claim intervals. 78(13):1727–1734.
- [137] MERLEVÈDE, F., PELIGRAD, M. et UTEV, S. (2006). Recent advances in invariance principles for stationary sequences. *Probab. Surv.*, 3:1–36.
- [138] MIKOSCH, T. (2009). *Non-life insurance mathematics*. Universitext. Springer-Verlag, Berlin, second édition.
- [139] MÜLLER, A. et STOYAN, D. (2002). *Comparison methods for stochastic models and risks*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.
- [140] NELSEN, R. B. (1999). *An introduction to copulas*, volume 139 de *Lecture Notes in Statistics*. Springer-Verlag, New York.
- [141] NORBERG, R. (1989). Actuarial analysis of dependent lives. *Schweizerische Vereinigung der Versicherungsmathematiker. Mitteilungen*, (2):243–254.
- [142] O'BRIEN, G. L. (1974). Limit theorems for sums of chain-dependent processes. *J. Appl. Probability*, 11:582–587.
- [143] ODAIRA, H. et YOSHIHARA, K.-i. (1972). Functional central limit theorems for strictly stationary processes satisfying the strong mixing condition. *Kōdai Math. Sem. Rep.*, 24:259–269.
- [144] PANJER, H. H. et WILLMOT, G. E. (1992). *Insurance risk models*. Society of Actuaries, Schaumburg, IL.
- [145] PELIGARD, M. et SURESH, R. (1995). Estimation of variance of partial sums of an associated sequence of random variables. *Stochastic Process. Appl.*, 56(2):307–319.
- [146] PELIGRAD, M. (1986). *Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (a survey)*, volume 11 de *Progr. Probab. Statist.* Birkhäuser Boston, Boston, MA.
- [147] PELIGRAD, M. (1987). The convergence of moments in the central limit theorem for  $\rho$ -mixing sequences of random variables. *Proc. Amer. Math. Soc.*, 101(1):142–148.
- [148] PELIGRAD, M. et SHAO, Q. M. (1994). Self-normalized central limit theorem for sums of weakly dependent random variables. *J. Theoret. Probab.*, 7(2):309–338.

- [149] PELIGRAD, M. et SHAO, Q. M. (1995). Estimation of the variance of partial sums for  $\rho$ -mixing random variables. *J. Multivariate Anal.*, 52(1):140–157.
- [150] PELIGRAD, M. et SHAO, Q.-M. (1996). A note on estimation of variance for  $\rho$ -mixing sequences. *Statist. Probab. Lett.*, 26(2):141–145.
- [151] PETROV, V. V. (1995). *Limit theorems of probability theory*, volume 4 de *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York. Sequences of independent random variables, Oxford Science Publications.
- [152] PHILIPP, W. (1986). *Invariance principles for independent and weakly dependent random variables*, volume 11 de *Progr. Probab. Statist.*
- [153] PHILIPP, W. et STOUT, W. (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Memoirs of the American Mathematical Society*, (issue 2, 161):iv+140.
- [154] PRAKASA, R. (2009). Conditional independence, conditional mixing and conditional association. *Ann. Inst. Statist. Math.*, 61(2):441–460.
- [155] PRÉKOPA, A. (1990). The discrete moment problem and linear programming. *Discrete Applied Mathematics. The Journal of Combinatorial Algorithms, Informatics and Computational Sciences*, 27(3):235–254.
- [156] REVUZ, D. (1984). *Markov chains.*, volume 11. Second edition. North-Holland Mathematical Library,.
- [157] RIO, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 29(4):587–597.
- [158] RIO, E. (1994). Inégalités de moments pour les suites stationnaires et fortement mélangeantes. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(4):355–360.
- [159] RIO, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31. Springer-Verlag.
- [160] ROLSKI, T., SCHMIDLI, H., SCHMIDT, V. et TEUGELS, J. (1999). *Stochastic processes for insurance and finance*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.
- [161] ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, 42:43–47.
- [162] ROSS, S. M. (1996). *Stochastic processes*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second édition.
- [163] ROTAR, V. (2006). Actuarial model the mathematics of insurance. *Chapman & Hall*, 38.
- [164] ROUSSAS, G. G. et IOANNIDES, D. (1987). Moment inequalities for mixing sequences of random variables. *Stochastic Anal. Appl.*, 5(1):61–120.
- [165] SHAKED, M. et J. G. SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and their applications*. Academic press, New York.

- 
- [166] SHAKED, M. et J. G. SHANTHIKUMAR, J. G. (2007). *Stochastic Orders and their applications*. Springer, New York.
- [167] SHAKED, M. et SPIZZICHINO, F. (1998). Positive dependent properties of conditionally independent random lifetimes. *Mathematics of Operations Research*, 23:944–959.
- [168] SHAO, Q. et YU, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *The Annals of Probability*, 24(4):2098–2127.
- [169] SHERVASHIDZE, T. (1999). Local limit theorems for conditionally independent random variables controlled by a finite Markov chain. *Rossiiskaya Akademiya Nauk. Teoriya Veroyatnostei i ee Primeneniya*, 44(1):143–148.
- [170] SHI, S. (2000). Estimation of the variance for strongly mixing sequences. *Appl. Math. J. Chinese Univ. Ser. B*, 15(1):45–54.
- [171] STOUT, W. F. (1974). *Almost sure convergence*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London. Probability and Mathematical Statistics, Vol. 24.
- [172] STRAUB, E. (1988). *Non-life insurance mathematics*. Springer-Verlag, Berlin.
- [173] SUNDT, B. (1993). An introduction to non-life insurance mathematics. 3rd ed., *Verlag Versicherungswirtschaft e. V., Karlsruhe*.
- [174] UTEV, S. A. (1985). *Extremal problems in moment inequalities*, volume 5. Novosibirsk.
- [175] UTEV, S. A. (1990). Central limit theorem for dependent random variables. pages 519–528.
- [176] VEIJANEN, A. (1990). On conditionally mixing processes. *Statist. Probab. Lett.*, 9(1):59–65.
- [177] VOLKONSKIĬ, V. A. et ROZANOV, Y. A. (1959). Some limit theorems for random functions. I. *Theor. Probability Appl.*, 4:178–197.
- [178] VOLNÝ, D. (1987). Martingale decompositions of stationary processes. *Yokohama Math. J.*, 35(1-2):113–121.
- [179] VOLNÝ, D. (1988). *Approximation of stationary processes and the central limit problem*, volume 1299 de *Lecture Notes in Math*. Springer, Berlin.
- [180] VOLNÝ, D. (1993). Approximating martingales and the central limit theorem for strictly stationary processes. *Stochastic Processes and their Applications*, 44(1):41–74.
- [181] WANG, S. (1998). *Aggregation of correlated risk portfolio: model and algorithms*. *Proceedings of the Casualty actuarial society*.
- [182] WANG, S. et DHAENE, J. (1998). Comonotonicity, correlation order and premium principles. *Insurance: Mathematics & Economics*, 22(3):235–242.
- [183] WILSON, T. (1997). Portfolio credit risk i. *RISK*.

- [184] WITHER, C. (1983). Central limit theorems for dependent variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 57:509–534.
- [185] YOKOYAMA, R. (1980). Moment bounds for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete*, 52(1):45–57.
- [186] YOUNG, W. (1912). On classes of summable functions and their fourier series. *Proc. Royal Soc. Lond.*, 87.
- [187] YUEN, K.-C. et GUO, J. (2006). Some results on the compound Markov binomial model. *Scand. Actuar. J.*, (3):129–140.
- [0]