



HAL
open science

Statistique des processus stables et des processus à longue mémoire

Caroline Robet

► **To cite this version:**

Caroline Robet. Statistique des processus stables et des processus à longue mémoire. Mathématiques [math]. Université Bretagne Loire, 2019. Français. NNT: . tel-02357610

HAL Id: tel-02357610

<https://hal.science/tel-02357610v1>

Submitted on 10 Nov 2019

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.

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ DE NANTES
COMUE UNIVERSITÉ BRETAGNE LOIRE

Ecole Doctorale N° 601

Mathématiques et Sciences et Technologies de l'Information et de la Communication

Spécialité : *Mathématiques et leurs Interactions*

Par

Caroline ROBET

**Statistique des processus stables et des processus
à longue mémoire**

Thèse présentée et soutenue à l'UNIVERSITÉ DE NANTES, le 20 septembre 2019

Unité de recherche : Laboratoire de Mathématiques Jean Leray (LMJL)

Jury :

Président : **M. Jean-Marc BARDET**, Professeur des universités, Université Paris 1
Rapporteurs : **M. Jean-François COEURJOLLY**, Professeur des universités, Université du Québec à Montréal
M. Clément DOMBRY, Professeur des universités, Université de Franche-Comté
Examineur : **M. Mathieu RIBATET**, Professeur des universités, École Centrale Nantes
Directeurs de thèse : **Mme Anne PHILIPPE**, Professeur des universités, Université de Nantes
M. Jacques LÉVY VÉHEL, Directeur de recherche INRIA, en disponibilité

Remerciements

Cette thèse marque la fin de mes études et je tiens à remercier tous ceux qui m'ont accompagnée pendant ce long périple.

Mes premiers remerciements vont à Anne qui m'a encadrée pendant mon stage de L3. Mon choix s'est tout naturellement porté sur toi lorsqu'il a fallu que je cherche un directeur de thèse. Merci de m'avoir accompagnée avec bienveillance pendant ces trois années et de m'avoir consacré du temps lorsque j'étais bloquée dans mon travail. Je tiens également à remercier Jacques, mon co-directeur. Merci pour les nombreuses idées que tu m'as suggérées pendant cette thèse.

Je remercie chaleureusement Jean-François COEURJOLLY et Clément DOMBRY d'avoir accepté de rapporter ma thèse. Les remarques, questions et corrections faites m'ont permis d'améliorer ce travail. Enfin, je remercie Jean-Marc BARDET et Mathieu RIBATET d'avoir accepté de faire partie de mon jury.

Merci à Marie-Claude VIANO et Mohamedou OULD HAYE qui ont contribué au travail que je présente aujourd'hui à travers des articles co-écrits. En particulier, merci Mohamedou pour ces nombreux échanges par mail qui ont largement permis d'améliorer notre travail.

Comme chacun le sait, une thèse repose également sur beaucoup de démarches administratives. Je remercie le secrétariat composé d'Anaïs, Annick, Brigitte, Stéphanie pour leur travail remarquable et leur patience pour gérer tous les papiers. Merci à Claude et Anh de s'occuper parfaitement du CRDM et de nous laisser emprunter des livres qu'ils ne sont pas sûrs de revoir avant longtemps. Merci Éric et Saïd pour leur aide en cas de soucis informatique.

Je remercie les doctorants/postdocs/ATER sans qui je n'aurais jamais pu passer d'aussi bons moments pendant ces trois années (et un peu plus) au LMJL, du fin fond du couloir du bas (où il faut aller chercher les gens pour manger) jusqu'aux bureaux perdus dans les couloirs d'informaticiens.

Tout d'abord, je commence par le bureau le plus éloigné (mais le plus près pour

aller en pause). Merci Fabien pour ta gentillesse inégalée. Ces parties d'Hanabi (magnifiquement remportées) avec toi chez Solène sur fond de Beyoncé resteront mythiques : *manette tu jettes!* Merci Alexandre pour m'avoir appris à jouer à la princesse, pour ton humour, pour tes nouveaux talents de photoshopeur. Merci Amiel pour ta bonne humeur, pour la relecture attentive de certains chapitres de cette thèse et pour tes talents de joueur hors-pair. Je te souhaite de toujours rester *heureux comme un pigeon*.

Merci au trio Pierre, Damien et Valentin qui ont participé à mon initiation à la taroinche, signe d'une bonne intégration parmi la communauté des doctorants. Vous avez su me mettre à l'aise dès mon stage de Master 2. Je te remercie Damien pour ton pot de thèse le jour de ma soutenance de Master qui m'a donné l'impression d'une grande fête en mon honneur.

Merci Vytauté, demie-sœur de thèse, pour ta bonne humeur lors de tes séjours en France. Bonne continuation au Danemark. Merci Radek, tu es revenu soutenir (et surtout tu as fait ton pot) en France pour nous revoir, tu as fait le bon choix ! Merci Guillaume, tu m'auras bien faire rire. J'espère que tu continues à tester le plus de crêperies possible. Merci Guillem. Merci Rhiannon, j'ai beaucoup appris sur l'Écosse grâce à toi. Merci Jean-Michel, mon demi *petit* frère de thèse, je te souhaite une bonne soutenance à venir, que je sais, sera excellente. Merci Claire B., digne représentante des Mauges, j'espère que tu retrouveras un jour les sièges de ta voiture. Merci Arthur, tu as mis du temps pour qu'on te connaisse ... ce n'était que pour mieux te révéler sur la piste de danse de la SMAI.

Merci Moudhaffar. J'ai eu l'honneur de te croiser malgré tes horaires décalés. Merci Florian et Victor M.D ! Vous avez su m'aider en informatique comme tout bon numéricien qui se doit. Florian, alias Blachèèèère le râleur invétéré, je n'oublierai pas ton combat contre l'école doctorale. Merci Victor d'avoir relu mes premiers écrits en anglais. J'ai été ravie de tes visites régulières dans le bureau où tu as pu partager ta passion pour la raclette.

Enfin, je passe à mon bureau que j'ai adoré. Merci Olivier, cher co-bureau, ta stratégie de non-réponse face à mes (fréquentes) discussions n'a pas toujours été efficace. Merci pour la bonne humeur du bureau 23 et pour ton aide sur la moindre démarche administrative à effectuer (tu es incollable !). Merci Caro pour toutes ces discussions, ces nombreuses rigolades, ces orgies de Schoko Bons dans le bureau. Ton départ m'a laissée en Caro solitaire. Merci Germain pour ces deux ans qui ne furent pas de tout repos pour toi (la reine d'Angleterre aime la tranquillité). Tu es

arrivé entouré de Caro, ce n'était pas facile. Fais attention, ton amour des frites finira par te mener en Belgique. Je te souhaite une bonne poursuite comme enseignant. Merci Matilde pour ta joie, ton accent italien, pour ta force de persuasion pour nous convaincre de sortir boire un verre.

Passons au bureau d'à côté, plus connu sous le nom de "bureau de la glandouille" et célèbre pour son canapé rouge qui a traversé les générations. J'ai fréquemment arpenté ce bureau en quête d'inspiration. Merci Victor VdR. Cette année en ta compagnie a été bien trop courte. Merci pour ton humour, pour tes reprises de chansons. Ton nom fait désormais partie intégrante de la taroinche : "il y a anguille sous Da Roche", "La Da Rochatte". Tu m'avais presque fait oublier comment je me prénommiais avec tes Robeeeeeeet. Merci Johan d'avoir été là malgré ta non-appartenance au LMJL. Merci de nous avoir accueilli pour l'after soutenance de Thomas W et d'avoir organisé des rencontres doctorales très sympas à Angers. Merci Thomas B. (fervent disciple de Mélenchon) pour ces presque 2 ans en tant que voisin de bureau. Merci à toi d'avoir fait une exception sur ton régime pour goûter ma galette des rois. Ton exil à Bressuire ne fut pas facile mais tu es maintenant de retour parmi nous. Merci Wallez, merci pour ton amour de Bertrand de la cafétéria, pour tes numéros de magie, pour ta pédagogie et ta passion à partager ce que tu connais. Merci Matthieu *AvecDeuxT* pour ton humour incomparable, pour ton don de beatbox, pour tes blagues aux cartes "120 sans" et "- 120 cœur. - Oh non, il n'y aura encore pas de gagnants!" en face tu as eu le droit à notre belge préféré "Jean-Claude J'entame". Je te souhaite beaucoup de manifestations pacifistes à venir. Merci Côme pour ta présence, qui, contrairement à ta matinalité, se passe désormais de commentaires, et pour tes vidéos de chats toujours trop mignonnes ! Tu fais désormais partie de la *Cômunity*. Merci Mohamad, le célèbre *Magnifaïk, ma chérie !* n'a plus la même signification depuis que je te connais. Tu nous as apporté des *cerfs* libanais délicieux.

Passons à un bureau plus sérieux et amateur de thés et tisanes en tout genre. Merci Thomas G pour tes bons gâteaux que tu nous as régulièrement ramenés, pour avoir animé un loup-garou sur fond de musique vietnamienne. Tu fais désormais partie des murs du laboratoire. Tu auras su nous surprendre chaque année avec tes réapparitions après les vacances. Merci Noémie pour ta joie (ton rire est très communicatif!), j'espère que tu continues toujours à chanter comme lorsque tu étais dans le chœur universitaire. Merci Niccolo, rare collègue en aléatoire. Merci Simon, grâce à toi Matt Pokora peut désormais affirmer qu'il est docteur

agrégé en mathématiques. Merci Zeinab pour ta gentillesse et ton humour quotidiens. Tu auras su m'amadouer avec les *maamoul* aux pistaches que tu as ramenés pendant tes vacances. Mes connaissances en arabe n'ont que peu progressé pendant cette thèse malgré ton sérieux en tant que professeur. Je peux néanmoins commander des *kusa*, même si j'aime ça couci-couça. Merci Maha pour la balade qu'on a fait ensemble au bord de l'Erdre, pour m'avoir dit "c'est la première fois que je te vois sérieuse" après mon exposé des rencontres doctorales. Merci Anh pour m'avoir bien fait rire en me demandant si je bouclais mes cheveux tous les matins, pour bien t'occuper de Trung (courage, encore deux ans), pour m'avoir appris l'essentiel du vietnamien *Côme không o' đâý*.

Merci Claire C., avec ton accent chantant, tu nous as transporté dans le sud. Tu as été un temps accompagnée par Matthieu et Côme dans ton bureau mais ce ne fut qu'éphémère. Merci Andrès pour avoir mis l'ambiance à la soirée du labo et pour ton pré-repas de Noël cosmopolite.

Passons désormais aux bureaux exilés du côté des informaticiens. Merci Emmanuel. Tu as été pendant un an l'homme invisible de ce labo mais je ne t'ai pas davantage vu après l'obtention d'un bureau. On a passé un bon moment en conférence à Oléron ensemble. Merci Cécile pour ton pot de départ chocolaté à Centrale, pour nous avoir aussi rejoint dans le club des sportifs occasionnels, pour la super ambiance à Guidel. Merci Trung pour ton insouciance, pour ton français approximatif (qui s'améliore à grand pas) qui m'a beaucoup fait rire et qui m'a forcée à te parler en français. Je garde en souvenir ton saut dans la fontaine à Brest ainsi que ton port de tongs en hiver comme preuve de ta capacité à faire abstraction du froid. Merci Meissa.

Merci Solène, alias Bubulle, Bulbizarre, Qulbutoke,... fidèle complice dans toute ma thèse. Merci de t'être laissé embarquer au step. On pourrait désormais être des championnes ensemble grâce à Ben (Tap up, Reverse, celle-là c'est ma chanson préférée, Tcha-tcha-tcha, Superman, Beyoncé, et on casse le chocolat ...). Merci de m'avoir laissé un coin de ton canapé lorsque je voulais rester sur Nantes. Merci aussi à toi Benjamin, bien que n'étant pas doctorant, tu auras passé beaucoup de temps au labo et aux soirées ! Merci Hélène. Tu t'es fait discrète en début de thèse... c'était pour pouvoir mieux te révéler ensuite. Merci pour les soirées jeux chez toi qui ont vraiment été drôlissimes, en particulier pour la course de chameaux la plus lente de l'histoire, pour ton Koh-Lanta meurtrier où Benjamin finit seul. Ta collection est impressionnante. Merci aussi à Laura. Ma participation à ton

enquête policière m'a bien plu. Je te souhaite de continuer à avoir cette passion pour l'enseignement. Merci Hala pour ta cuisine libanaise délicieuse. En particulier, merci pour ton anniversaire à Brest où on s'est fait un régal. Merci d'avoir essayé quelques séances de step avec Solène et moi, c'était très sympa. Merci Nicolas pour ta langue de vipère qui m'aura bien fait rire. Merci Beatrice pour ce très bon séjour à la SMAI avec toi. La soirée de gala a été l'occasion d'en apprendre un peu plus sur toi : *one bidet by day*. Je te souhaite bon courage pour ta thèse en Italie. Merci Maël.

Enfin, merci aux autres doctorants qu'on peut parfois croiser. Merci El Housaine et Fakhrielddine pour votre rencontre à Brest. Merci Fatima. Merci Karzan pour ton goûter.

Merci à tous les chercheurs du labo. Merci Marianne, Philippe (pour tes fréquents passages dans le couloir du bas, pour ton soutien au comité de suivi de thèse), Frédéric L., Hélène, Nicolas P. (pour ton encadrement dans mon stage de M1), Éric (pour ce super cours de stat en collaboration), Gilles, Gabriel, Marie-Anne (pour nos nombreuses discussions), Colette, Fred, Anaïs, Erwan, Jo, Marco, François N., Baptiste, Paul, Aymeric, Bertrand, Christophe B., Sébastien, Friedrich, Samuel, Laurent, Vincent, Bernard, Gueorgui, Benoît, Mehdi... En particulier, je remercie les membres de l'équipe SPAN. Je remercie également tous les doctorants, post-doctorants, ATER et titulaires que j'ai pu croiser au laboratoire ou lors des conférences ainsi que mes amis rennais qui ont continué à me soutenir pendant la thèse : Ninon, Adrien, Kévin, Joackim, Maylis, Maud ... et surtout, toi, Laura merci infiniment pour ton aide. Notre fil de discussion sur Messenger est témoin de notre soutien mutuel pendant ces trois années. Merci Valentine, ce Master 2 n'aurait pas été le même sans nos soirées du mercredi.

Merci à ma famille et à mes amis qui sont contents pour moi sans jamais être parvenus à comprendre ce que je faisais de mes journées. En particulier, merci Maman de m'avoir demandé si *l'école* s'était bien passé dès qu'on se voyait et pour tous les petits plats que tu m'as préparés (crêpes, soupes....), merci Guillaume pour avoir financé ma thèse avec tes impôts :p, merci Aurélie pour les nombreuses journées Just Dance. Merci Julien, Marie, Steven, Fanny, Florent, Toska ... et tous les autres !

Enfin, merci à toi Mathieu, le seul et unique. Malgré ton aversion pour les mathématiques, tu es parvenu jusqu'au bout de cette thèse à mes côtés.

Table des matières

1	Introduction	15
1.1	Processus stables	16
1.1.1	Quelques rappels sur les lois et processus stables	16
1.1.2	Représentations des processus stables	19
1.2	Processus stationnaires du second ordre	24
1.2.1	Processus stationnaires du second ordre : fonction d'autocovariance, densité spectrale et mémoire	24
1.2.2	Théorème central limite et convergence du processus des sommes partielles	28
1.3	Organisation de la thèse	31
1.3.1	Partie I : Estimations et simulations de processus stables	31
1.3.2	Partie II : Échantillonnage de processus stationnaires du second ordre définis en temps continu	35
I	Estimation et simulation de processus stables	41
2	Estimateurs explicites et combinés pour les paramètres de lois stables	43
2.1	Introduction	44
2.2	Méthodes d'estimation	47
2.2.1	Log-moments dans le cas symétrique	48
2.2.2	Performance numérique de chaque estimateur	50
2.2.3	Estimateur combiné	56
2.3	Lois stables asymétriques	61
2.3.1	Adaptation des estimateurs pour le cas asymétrique	61
2.3.2	Test de symétrie	63
2.4	Cas de variables stables non-identiquement distribuées	66

2.4.1	Perturbations déterministes	66
2.4.2	Perturbations aléatoires	67
2.5	Quelques applications de l'estimateur combiné	68
2.5.1	Résultats numériques sur des données synthétisées : mouve- ment de Lévy multistable	68
2.5.2	Application en finance	73
2.6	Annexe	77
3	Simulation de processus multistables	83
3.1	Processus localisables	84
3.2	Processus multistables définis par un champ stochastique	85
3.2.1	Cas des espaces de mesure finie	85
3.2.2	Cas des espaces de mesure σ -finie	87
3.3	Grille multistable	90
3.4	Exemples de simulation	103
3.4.1	Mouvement de Lévy multistable	103
3.4.2	Mouvement linéaire multifractionnaire multistable	105
3.5	Lemmes techniques	107
4	Estimation de la fonction d'intensité de processus multistables	115
4.1	Introduction	115
4.2	Convergence des moments empiriques de processus multistables autour d'un point	117
4.3	Estimation de la fonction d'intensité du processus multistable	122
4.3.1	Exemple d'estimation avec des mouvements de Lévy multi- stables	125
4.3.2	Exemple d'estimation avec des mouvements linéaires multi- fractionnaires multistables	127
II Échantillonnage de processus stationnaires du second ordre définis en temps continu		131
5	Échantillonnage aléatoire de processus stationnaires à temps continu	133
5.1	Introduction	134
5.2	Résultats généraux dans le domaine temporel	136
5.2.1	Loi du processus échantillonné	136

5.2.2	Mémoire du processus échantillonné	139
5.3	Processus à longue mémoire	142
5.3.1	Préservation du paramètre du mémoire lorsque $\mathbb{E}[T_1] < \infty$	142
5.3.2	Perte de mémoire lorsque $\mathbb{E}[T_1] = \infty$	152
5.4	Comparaison entre le sous échantillonnage et l'agrégation temporelle	154
5.5	Annexe	160
6	Étude spectrale de l'échantillonnage de processus stationnaires	165
6.1	Introduction	166
6.2	Existence d'une densité spectrale pour le processus échantillonné .	167
6.3	Asymptotique du périodogramme	171
6.4	Inférence pour le paramètre de longue mémoire	176
6.4.1	Convergence de l'estimateur local Whittle	177
6.4.2	Variance asymptotique	183
6.5	Annexe	188
6.5.1	Preuve du Lemme 6.2.1	188
6.5.2	Preuve de la Proposition 6.2.2	189
6.5.3	Preuve de la Proposition 6.2.3	198
6.5.4	Preuve de la Proposition 6.4.1	200
	Bibliographie	209

Table des figures

2.1	Log-rapport pour l'erreur quadratique des estimateurs de Koutrouvelis et des log-moments du paramètre α pour différentes tailles d'échantillons n	54
2.2	Densités empiriques des estimateurs de Koutrouvelis et des log-moments pour les paramètres α et σ avec un échantillon de taille $n = 500$	55
2.3	Poids moyen de l'estimateur des log-moments $\hat{\alpha}_{LOG}$ dans l'estimateur combiné en fonction de α	58
2.4	Densités empiriques des estimateurs des log-moments, de Koutrouvelis et combiné pour le paramètre α	60
2.5	Fonctions de répartitions empiriques de la p -valeur du test de symétrie pour différentes valeurs de α , β et n	65
2.6	Trajectoires de mouvements de Lévy multistables	70
2.7	Boîtes à moustaches pour l'estimateur combiné appliqué à un mouvement de Lévy multistable	71
2.8	Biais, erreur standard et erreur quadratique moyenne en fonction du temps pour l'estimateur combiné appliqué à un mouvement de Lévy multistable	72
2.9	Erreur quadratique moyenne pour l'estimateur combiné appliqué à un mouvement de Lévy multistable par rapport à la valeur de α	73
2.10	Évolution de l'indice financier <i>S&P 500</i> et de ses rendements au cours du temps entre 1996 et 2017	74
2.11	Fonction de répartition empirique de la p -valeur du test de symétrie pour les rendements de <i>S&P 500</i>	75
2.12	Valeurs de l'estimateur combiné $\hat{\alpha}_{COMB}$ de l'indice <i>S&P 500</i> pour une fenêtre glissante de taille n	76

TABLE DES FIGURES

2.13	Estimateurs combinés symétriques et asymétriques appliqués à l'indice financier S&P 500	77
3.1	Champ de mouvements de Lévy stables et mouvement de Lévy multistable	104
3.2	Mouvement de Lévy multistable et son processus approché	105
3.3	Simulation d'un mouvement linéaire multifractionnaire multistable	107
4.1	Convergence des moments empiriques autour d'un point du mouvement de Lévy multistable	118
4.2	Convergence des moments empiriques autour d'un point du mouvement linéaire multifractionnaire multistable	119
4.3	Estimation de la fonction d'intensité de mouvements de Lévy multistables grâce à un rapport de moments empiriques	126
4.4	Estimation de la fonction d'intensité de mouvements linéaires multifractionnaires multistables grâce à un rapport de moments empiriques	128
4.5	Estimation de la fonction d'intensité de mouvements linéaires multifractionnaires multistables grâce à un rapport de moments empiriques	129
5.1	Perte du caractère gaussien après échantillonnage aléatoire	138
5.2	Réalisations de trajectoires de la statistique R/S pour différentes valeurs de paramètre de mémoire	151
6.1	Réalisations des valeurs du périodogrammes aux fréquences de Fourier pour différentes valeurs de paramètre de mémoire	183

Notations

$\mathbb{1}_A$	fonction indicatrice de l'ensemble A
Λ_α	mouvement log-fractionnaire stable
$[x]$	partie entière du réel x
$\llbracket p, q \rrbracket$	ensemble des entiers n tels que $p \leq n \leq q$ avec $p, q \in \mathbb{N}$ et $p \leq q$
\mathbb{R}	ensemble des réels
$\mathcal{C}[0, 1]$	ensemble des fonctions continues de $[0, 1]$ dans \mathbb{R} .
$\mathcal{D}[0, 1]$	espace de Skorokhod : ensemble des fonctions càdlàg (continues à droite et ayant une limite à gauche en tout point) de $[0, 1]$ dans \mathbb{R} .
$\mathcal{E}(\lambda)$	loi exponentielle de moyenne $1/\lambda$
$\mathcal{F}_\alpha(E, \mathcal{E}, m)$	ensemble des fonctions g mesurables sur l'espace (E, \mathcal{E}, m) de mesure σ -finie, telles que $\ g\ _\alpha := (\int_E g(x) ^\alpha m(dx))^{1/\alpha}$ soit fini
$\mathcal{N}(\mu, \sigma^2)$	loi normale de moyenne μ et de variance σ^2
\mathcal{O}_p	$X_n = \mathcal{O}_p(a_n)$ si $\forall \varepsilon > 0, \exists M > 0, N \in \mathbb{N}$ tels que $\forall n > N, \mathbb{P}(X_n/a_n > M) < \varepsilon$
Leb	mesure de Lebesgue sur \mathbb{R}
\mathbb{N}	ensemble des entiers naturels
$\stackrel{Loi}{\equiv}$	égalité en loi
$\xrightarrow{Loi}, \xrightarrow{proba}, \xrightarrow{p.s.}$	convergence en loi, convergence en probabilité, convergence presque sûre
L_α	mouvement de Lévy α -stable

NOTATIONS

$L_{\alpha,H}$	mouvement linéaire fractionnaire α -stable bien équilibré, d'indice de Hurst H
o_p	$X_n = o_p(a_n)$ si $\forall \varepsilon > 0, \mathbb{P}(X_n/a_n > \varepsilon) \xrightarrow[n \rightarrow \infty]{} 0$, i.e $X_n/a_n \xrightarrow[n \rightarrow \infty]{proba} 0$
$S\alpha S$	loi α -stable symétrique
$S_\alpha(\sigma, \beta, \mu)$	loi stable d'indice α , de paramètre d'échelle σ , d'asymétrie β et de paramètre de position μ
i.i.d.	indépendants et de même loi
p.s.	presque sûrement



1

Introduction

Afin de modéliser simplement des phénomènes qui évoluent au cours du temps, il est souvent pertinent de supposer que le fait étudié évolue aléatoirement et de le représenter par un processus stochastique. La classe des processus stochastiques est très large et permet de modéliser de nombreuses caractéristiques observées sur le phénomène que l'on cherche à étudier. Le choix du processus stochastique pour le représenter permet ainsi de prendre en compte les éventuels dépendances au passé, sauts, invariants du phénomène, régularité de la trajectoire...

Dans de nombreux cas, les accroissements du processus sont supposés stationnaires. Cela permet de caractériser des données dont les accroissements semblent suivre la même loi au cours du temps. Si, pour de telles données, le phénomène étudié présente des queues de distribution lourdes (à savoir que la probabilité qu'un accroissement soit supérieure à M est de l'ordre de $M^{-\alpha}$ avec $0 < \alpha < 2$), on représente alors ce phénomène à l'aide d'un processus α -stable. Le paramètre α contrôle alors l'intensité des sauts dans la trajectoire. En particulier, dans le domaine de la finance, on peut constater que de nombreux indices présentent des queues lourdes. Cependant, ce type de processus décrit mal l'évolution du marché financier pour lequel le comportement va être différent suivant la période étudiée. En effet, les indices financiers n'auront pas le même aspect en période de forte croissance par rapport à une période de crise économique. Pour cela, on consi-

dère des processus multistables qui sont des processus dont l'intensité α évolue au cours du temps et qui, localement, se comportent comme des processus stables. L'étude locale des processus stochastiques et différentes méthodes de construction de processus multistables ont été traitées dans [Fal02 ; Fal03 ; Fal+09 ; LGLV12]. La première partie de cette thèse consiste en l'estimation des paramètres pour une loi stable puis à la simulation de processus multistables et l'estimation de la fonction d'intensité de tels processus.

Dans un deuxième temps, on se focalise sur des données en temps continu présentant une faible stationnarité au niveau de leur loi. On va alors supposer que le processus construit pour modéliser ces données a une espérance constante au cours du temps et une covariance invariante par translation. De tels processus sont utilisés dans des domaines variés (hydrologie, économie, télécommunications...).

En temps discret, ce type de processus a été largement étudié mais pour des phénomènes en temps continu l'étude est plus compliquée puisqu'il n'est généralement plus possible d'avoir accès à l'intégralité d'une trajectoire du processus. Pour de tels processus, on n'observe habituellement leurs valeurs qu'à des instants discrets correspondant aux instants d'échantillonnage à partir de mesures ponctuelles.

Des travaux ont été menés dans le cadre d'échantillonnage déterministe régulier ([Cha98], [Hwa00]) ou irrégulier ([TC05a]). Cependant, pour des domaines comme la géologie ou encore l'archéologie pour lesquels des données peuvent provenir par exemple d'une éruption volcanique, il est souhaitable de penser que les observations proviennent d'un échantillonnage aléatoire. De même, pour des données issues de machines, on peut représenter l'erreur de précision à l'aide d'un aléa sur les temps d'observations. Une étude de certains échantillonnage aléatoire est menée dans [SS60 ; Phi+19]. La deuxième partie de cette thèse est ainsi axée sur l'étude par échantillonnage aléatoire de processus faiblement stationnaires à temps continu.

1.1 Processus stables

1.1.1 Quelques rappels sur les lois et processus stables

Dans [ST94], un travail très approfondi sur les processus stables est mené. Nous allons uniquement rappeler dans cette section les principales définitions

et propriétés qui vont nous servir dans la première partie de cette thèse. On commence par rappeler la définition d'une variable aléatoire stable utilisée dans [ST94].

Definition 1.1.1

Une variable aléatoire X suit une loi stable s'il existe des paramètres $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\sigma \geq 0$ et $\mu \in \mathbb{R}$ tels que sa fonction caractéristique soit de la forme :

$$\mathbb{E} \left[e^{itX} \right] = \begin{cases} \exp \left(-\sigma^\alpha |t|^\alpha \left(1 - i\beta \text{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right) \right) + i\mu t \right), & \text{si } \alpha \neq 1, \\ \exp \left(-\sigma |t| \left(1 + i\beta \text{sign}(t) \frac{2 \log |t|}{\pi} \right) + i\mu t \right), & \text{si } \alpha = 1 \end{cases} \quad (1.1.1)$$

où

$$\text{sign}(t) = \begin{cases} 1 & \text{si } t > 0, \\ 0 & \text{si } t = 0, \\ -1 & \text{si } t < 0 \end{cases}$$

Une loi stable est ainsi caractérisée par quatre paramètres. On écrit

$$X \sim S_\alpha(\sigma, \beta, \mu)$$

pour indiquer que X suit une loi stable d'indice de stabilité $\alpha \in]0, 2]$, de paramètre d'échelle $\sigma \in \mathbb{R}^+$, d'asymétrie $\beta \in [-1, 1]$ et de paramètre de position $\mu \in \mathbb{R}$. On écrira également $X \sim S_\alpha S$ si X est symétrique α -stable, i.e. si $\beta = \mu = 0$. Les densités des variables aléatoires α -stables existent et sont continues mais n'ont pas de formes explicites à part dans de rares cas (décrits ci-dessous).

Exemple 1.1.1. 1. La loi normale $S_2(\sigma, 0, \mu) = \mathcal{N}(\mu, 2\sigma^2)$ admet pour densité

$$\frac{1}{\sqrt{4\pi\sigma^2}} \exp \left(-\frac{(x - \mu)^2}{4\sigma^2} \right).$$

2. La loi de Cauchy $S_1(\sigma, 0, \mu)$ admet pour densité

$$\frac{\sigma}{\pi((x - \mu)^2 + \sigma^2)}.$$

3. La loi de Lévy $S_{1/2}(\sigma, 1, \mu)$ admet pour densité

$$\sqrt{\frac{\sigma}{2\pi}} \frac{1}{(x - \mu)^{3/2}} \exp\left(-\frac{\sigma}{2(x - \mu)}\right) \mathbb{1}_{x > \mu}.$$

On rappelle ensuite des propriétés sur le comportement des queues et sur les moments de lois stables dans le cas non-gaussien.

Proposition 1.1.1

Soit $X \sim S_\alpha(\sigma, \beta, \mu)$ avec $0 < \alpha < 2$. Alors

$$\begin{cases} \lim_{\lambda \rightarrow +\infty} \lambda^\alpha P(X > \lambda) &= C_\alpha \frac{1 + \beta}{2} \sigma^\alpha, \\ \lim_{\lambda \rightarrow +\infty} \lambda^\alpha P(X < -\lambda) &= C_\alpha \frac{1 - \beta}{2} \sigma^\alpha, \end{cases} \quad (1.1.2)$$

où

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin(x) dx \right)^{-1} = \begin{cases} \frac{1 - \alpha}{\Gamma(2 - \alpha) \cos(\pi\alpha/2)} & \text{si } \alpha \neq 1, \\ 2/\pi & \text{si } \alpha = 1. \end{cases} \quad (1.1.3)$$

Le comportement des queues de distribution sert ensuite à décrire les moments des lois stables.

Proposition 1.1.2 ([ST94], Prop 1.2.16)

Soit $X \sim S_\alpha(\sigma, \beta, \mu)$ avec $0 < \alpha < 2$. Alors pour $p > 0$, on a

$$\mathbb{E}[|X|^p] < \infty \quad \text{si et seulement si } p < \alpha.$$

Proposition 1.1.3

Soit $X \sim S_\alpha(\sigma, \beta, 0)$ avec $0 < \alpha < 2$ (si $\alpha = 1$, on suppose de plus que β est égal à 0). Alors pour tout $p \in]0, \alpha[$, il existe une constante $c_{\alpha, \beta}(p)$ telle que

$$\mathbb{E}[|X|^p]^{1/p} = c_{\alpha, \beta}(p) \sigma. \quad (1.1.4)$$

L'expression de la constante est donnée dans [HJ84] par :

$$c_{\alpha, \beta}(p)^p = \frac{2^{p-1} \Gamma(1 - \frac{p}{\alpha})}{p \int_0^\infty \frac{\sin^2(u)}{u^{p+1}} du} \left(1 + \beta^2 \tan^2 \frac{\alpha\pi}{2} \right)^{p/2\alpha} \cos \left(\frac{p}{\alpha} \arctan \left(\beta \tan \frac{\alpha\pi}{2} \right) \right). \quad (1.1.5)$$

On rappelle à présent la définition d'un processus α -stable.

Définition 1.1.4

Soit $0 < \alpha \leq 2$. Un processus stochastique $\{X_t, t \in T\}$ où $T \subset \mathbb{R}$ est appelé α -stable si toutes ses lois fini-dimensionnelles sont α -stables.

En particulier, les processus 2-stables sont les processus gaussiens.

Définition 1.1.5

Un processus stochastique $\{X(t), t \in T\}$ où $T \subset \mathbb{R}$ est H -auto-similaire si pour tout $a > 0$, $\{X(at), t \in T\}$ et $\{a^H X(t), t \in T\}$ ont les mêmes lois fini-dimensionnelles.

Nous allons voir dans la suite que les processus auto-similaires interviennent dans de nombreux exemples de processus stables. Cette propriété associée à la stationnarité des accroissements va nous permettre de caractériser le comportement local des processus étudiés.

1.1.2 Représentations des processus stables

Nous rappelons à présent deux représentations des processus α -stables qui nous serviront par la suite dans les chapitres 3 et 4. Nous allons supposer dans cette partie que $\beta = \mu = 0$.

Soit (E, \mathcal{E}, m) un espace de mesure σ -finie et $\mathcal{E}_0 = \{A \in \mathcal{E} : m(A) < \infty\}$. On commence par définir une mesure aléatoire α -stable M . Soit (Ω, \mathcal{F}, P) l'espace probabilisé sous-jacent et $L^0(\Omega)$ l'ensemble des variables aléatoires réelles sur cet espace.

Définition 1.1.6

Une fonction d'ensemble $M : \mathcal{E}_0 \rightarrow L^0(\Omega)$ telle que :

- $\forall A \in \mathcal{E}_0, M(A) \sim S_\alpha(m(A)^{1/\alpha}, 0, 0)$,
- $\forall A_1, \dots, A_k \in \mathcal{E}_0$ disjoints, les $M(A_i)$ sont indépendants,
- $\forall (A_j)_{j \geq 1} \in \mathcal{E}_0^{\mathbb{N}}$ disjoints tels que $\bigcup_{j \geq 1} A_j \in \mathcal{E}_0, M(\bigcup_{j \geq 1} A_j) = \sum_{j \geq 1} M(A_j)$ p.s.

est appelée une mesure aléatoire symétrique α -stable sur (E, \mathcal{E}) de mesure de contrôle m .

On souhaite ensuite définir l'intégrale d'une fonction contre la mesure aléatoire symétrique α -stable M . Pour une fonction f étagée de la forme $f(x) = \sum_{j=1}^q c_j \mathbb{1}_{A_j}(x)$

(où les A_j sont des ensembles disjoints inclus dans \mathcal{E}_0), on définit

$$I(f) := \int_E f(x)M(dx) = \sum_{j=1}^q c_j M(A_j).$$

On souhaite, par extension, construire cette intégrale stochastique sur un espace plus large de fonctions. On définit

$$\mathcal{F}_\alpha(E, \mathcal{E}, m) = \{g : g \text{ est mesurable et } \|g\|_\alpha < \infty\}$$

où $\|\cdot\|_\alpha$ est la quasi-norme donnée par

$$\|g\|_\alpha = \left(\int_E |g(x)|^\alpha m(dx) \right)^{1/\alpha}.$$

On peut alors définir l'intégrale α -stable de mesure de contrôle m d'une fonction $f \in \mathcal{F}_\alpha(E, \mathcal{E}, m)$ notée

$$I(f) := \int_E f(x)M(dx) \tag{1.1.6}$$

et qui vérifie

$$I(f) \sim S_\alpha(\|f\|_\alpha, 0, 0),$$

(voir le chapitre 3 de [ST94] pour le passage de la construction pour les fonctions étagées à la construction sur l'espace $\mathcal{F}_\alpha(E, \mathcal{E}, m)$). On donne quelques exemples de processus qui peuvent s'écrire comme intégrale d'une fonction contre une mesure aléatoire symétrique α -stable M .

Exemple 1.1.2. Le **mouvement de Lévy α -stable** est défini par :

$$L_\alpha(t) := \int_0^\infty \mathbb{1}_{x \leq t}(x)M(dx) = \int_0^t M(dx), \quad t \geq 0$$

où M est une mesure aléatoire symétrique α -stable sur \mathbb{R}_+ de mesure de contrôle Leb. Il a les propriétés suivantes :

1. $L_\alpha(0) = 0$ p.s.,
2. L_α est à accroissements indépendants,
3. $L_\alpha(t) - L_\alpha(s) \sim S_\alpha((t-s)^{1/\alpha}, 0, 0)$ pour tout $0 \leq s \leq t$.

Le processus L_α est un processus de Lévy. Il est $1/\alpha$ -auto-similaire et à accroissements stationnaires. C'est d'ailleurs le seul processus α -stable qui vérifie ces deux propriétés si $0 < \alpha < 1$. Ce n'est plus le cas si $\alpha \geq 1$. Par exemple, le processus $X(t) = tY$

avec $Y \sim S_1(1, 0, 0)$ est 1-auto-similaire à accroissements stationnaires. Le cas $\alpha = 2$ correspond au cas particulier du **mouvement Brownien**.

Exemple 1.1.3. Processus des moyennes mobiles symétriques α -stable

Soit $0 < \alpha \leq 2$ et f une fonction réelle mesurable telle que $\int_{\mathbb{R}} |f(x)|^\alpha dx < \infty$. On définit un processus de moyennes mobiles X à partir de f et d'une mesure aléatoire M symétrique α -stable sur \mathbb{R} de mesure de contrôle Leb par

$$X(t) = \int_{\mathbb{R}} f(t-x)M(dx), \quad t \in \mathbb{R}.$$

C'est un processus stationnaire.

Exemple 1.1.4. Le processus d'Ornstein-Uhlenbeck stable est défini par :

$$X(t) := \int_{-\infty}^t e^{-\lambda(t-x)}M(dx), \quad t \in \mathbb{R}$$

où M est une mesure aléatoire symétrique α -stable sur \mathbb{R} de mesure de contrôle Leb et λ est un paramètre strictement positif. C'est un processus des moyennes mobiles symétrique α -stable. Il est donc stationnaire.

Exemple 1.1.5. Le processus d'Ornstein-Uhlenbeck Rétrograde stable est défini par :

$$X(t) := \int_t^{\infty} e^{-\lambda(x-t)}M(dx), \quad t \in \mathbb{R}$$

où M est une mesure aléatoire symétrique α -stable sur \mathbb{R} de mesure de contrôle Leb et λ est un paramètre strictement positif.

Exemple 1.1.6. Le Mouvement linéaire fractionnaire α -stable est défini par :

$$L_{\alpha,H}(a, b; t) := \int_{\mathbb{R}} f_{\alpha,H}(a, b; t, x)M(dx)$$

où

$$f_{\alpha,H}(a, b; t, x) = a \left[(t-x)_+^{H-1/\alpha} - (-x)_+^{H-1/\alpha} \right] + b \left[(t-x)_-^{H-1/\alpha} - (-x)_-^{H-1/\alpha} \right],$$

a et b sont des réels avec $|a| + |b| > 0$, $0 < \alpha \leq 2$ et $0 < H < 1$ tels que $H \neq 1/\alpha$, M mesure aléatoire symétrique α -stable sur \mathbb{R} de mesure de contrôle Leb. $L_{\alpha,H}(a, b; \cdot)$ est un processus H -auto-similaire à accroissements stationnaires. De plus, lorsque $a = b = 1$, il est appelé **mouvement linéaire fractionnaire α -stable bien équilibré** et est noté $L_{\alpha,H}$.

On a alors

$$L_{\alpha,H}(t) := \int_{\mathbb{R}} [|t-x|^{H-1/\alpha} - |x|^{H-1/\alpha}] M(dx).$$

Le cas $\alpha = 2$ pour $L_{\alpha,H}$ correspond au **mouvement Brownien fractionnaire**.

Exemple 1.1.7. Le **mouvement log-fractionnaire stable** est défini par :

$$\Lambda_{\alpha}(t) := \int_{\mathbb{R}} [\ln |t-x| - \ln |x|] M(dx), \quad t \in \mathbb{R}$$

où M est une mesure aléatoire symétrique α -stable sur \mathbb{R} de mesure de contrôle Leb. Ce processus est bien défini seulement pour $\alpha \in]1, 2]$. Il est $1/\alpha$ -auto-similaire à accroissements stationnaires.

On donne ensuite une proposition qui montre le résultat d'une intégrale stochastique après changement de mesure de contrôle.

Proposition 1.1.7 (Changement de mesure de contrôle)

Soit M_m et M_{ν} deux mesures aléatoires symétriques α -stable sur un espace mesurable (E, \mathcal{E}) de mesures de contrôle σ -finies respectives m et ν , vérifiant

$$\nu(dx) = \frac{1}{r(x)} m(dx)$$

où $r : E \rightarrow \mathbb{R}^+$. Alors pour tout $f \in \mathcal{F}_{\alpha}(E, \mathcal{E}, m)$,

$$\int_E f(x) M_m(dx) \stackrel{\text{Loi}}{=} \int_E f(x) r(x)^{1/\alpha} M_{\nu}(dx).$$

Cette proposition va être très utile par la suite pour la construction des processus multistables dans le cas de mesure σ -finie.

En effet, si m est une mesure σ -finie sur (E, \mathcal{E}) , on peut trouver une suite de parties $(E_n)_{n \in \mathbb{N}} \in \mathcal{E}^{\mathbb{N}}$ disjointes deux à deux telles que $E = \bigcup_{n \in \mathbb{N}} E_n$ et $m(E_n) < \infty$ pour tout n . Si on pose $r : E \rightarrow \mathbb{R}^+$ la fonction qui à $x \in E_n$ associe $m(E_n)$ et $\widehat{m}(dx)$ la mesure sur (E, \mathcal{E}) telle que $\widehat{m}(dx) = \frac{1}{r(x)} m(dx)$, on se ramène au cas d'une construction pour une mesure finie puisque \widehat{m} est une mesure de probabilité. Dans le cas de mesure σ -finie, on peut donc toujours trouver une fonction r à valeurs dans \mathbb{R}^+ telle que $\widehat{m}(dx) = \frac{1}{r(x)} m(dx)$ soit une mesure de probabilité.

On pourra regarder dans [ST94] (chap3) pour plus de détails sur la construction de cette intégrale stable.

Nous donnons désormais une autre représentation d'une intégrale α -stable qui va nous servir pour l'étude des processus multistables. Cette représentation de Ferguson-Klass-LePage repose sur une somme convergente de temps d'arrivée d'un processus de Poisson. On trouvera plus de détails sur cette représentation dans [Ben+01], [FK72], [Ros90] et [ST94].

On suppose désormais que (E, \mathcal{E}, m) est un espace de mesure finie et $0 < \alpha < 2$. On considère $(\Gamma_i)_{i \geq 1}$ une suite de temps d'arrivée d'un processus de Poisson avec un taux d'arrivée unitaire : $\Gamma_i = \sum_{j=1}^i \varepsilon_j$ où les (ε_j) sont i.i.d. $\mathcal{E}(1)$. Soit $(V_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $\widehat{m} = m/m(E)$ sur E et $(\gamma_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$. On suppose finalement que les trois suites $(\Gamma_i)_{i \geq 1}$, $(V_i)_{i \geq 1}$ et $(\gamma_i)_{i \geq 1}$ sont indépendantes.

On donne alors le théorème de représentation en série de Ferguson-Klass-LePage de l'intégrale stochastique α -stable (voir la version de [ST94] Théorème 3.10.1) :

Théorème 1.1.8 (Représentation de Ferguson-Klass-LePage, cas symétrique)

Soit M une mesure aléatoire symétrique α -stable avec $0 < \alpha < 2$ de mesure de contrôle m et f une fonction appartenant à $\mathcal{F}_\alpha(E, \mathcal{E}, m)$. On définit la série

$$S(f) := (C_\alpha m(E))^{1/\alpha} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha} f(V_i), \quad (1.1.7)$$

où C_α est la constante définie par (1.1.3).

La série est presque sûrement convergente et on a

$$I(f) \stackrel{\text{Loi}}{=} S(f)$$

où $I(f)$ définie par (1.1.6) est l'intégrale stochastique de la fonction f contre une mesure aléatoire M symétrique α -stable de mesure de contrôle m .

Corollaire 1.1.9

Soient $f_1, \dots, f_k \in \mathcal{F}_\alpha(E, \mathcal{E}, m)$, alors

$$(I(f_1), \dots, I(f_k)) \stackrel{\text{Loi}}{=} (S(f_1), \dots, S(f_k)).$$

On rappelle que deux processus sont dit égaux en loi s'ils ont les mêmes lois finidimensionnelles.

Exemple 1.1.8. Soit $(L_\alpha(t))_{t \in [0,1]}$ le processus de Lévy α -stable sur $[0, 1]$. On a vu que

$$(L_\alpha(t))_{t \in [0,1]} \stackrel{\text{Loi}}{=} \left(\int_0^1 \mathbb{1}_{[0,t]}(x) M(dx) \right)_{t \in [0,1]}$$

où M est une mesure aléatoire symétrique α -stable sur $([0, 1], \mathcal{B}([0, 1]))$ de mesure de contrôle Leb. D'après le théorème de représentation de Ferguson-Klass-LePage et le corollaire précédent, on a

$$(L_\alpha(t))_{t \in [0,1]} \stackrel{\text{Loi}}{=} \left(C_\alpha^{1/\alpha} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha} \mathbb{1}_{[0,t]}(V_i) \right)_{t \in [0,1]}$$

où les Γ_i , γ_i et V_i sont décrits précédemment. Ici les $(V_i)_i$ forment une suite de variables aléatoires i.i.d. de loi uniforme sur $[0, 1]$.

1.2 Processus stationnaires du second ordre

1.2.1 Processus stationnaires du second ordre : fonction d'autocovariance, densité spectrale et mémoire

Dans la suite, on s'intéresse à des processus stochastiques définis sur $I \subset \mathbb{R}$ où I correspond, suivant le processus étudié, à \mathbb{N} (temps discret) ou à \mathbb{R}^+ (temps continu). On rappelle les définitions et propriétés des processus stationnaires du second ordre (voir [Gir+12]).

Définition 1.2.1

Un processus stochastique à valeurs réelles $X = (X_t)_{t \in I}$ est dit *strictement stationnaire* si pour tout $n \geq 1$, pour tout $t_1, \dots, t_n \in I$, la loi jointe du vecteur aléatoire $(X_{t_1+s}, \dots, X_{t_n+s})$ ne dépend pas de $s \in I$.

Définition 1.2.2

Un processus stochastique à valeurs réelles $X = (X_t)_{t \in I}$ est dit *stationnaire du second ordre* si

- $\forall t \in I, X_t \in L^2$,
- $\forall t \in I, E[X_t] = \mu$ (l'espérance est constante au cours du temps),
- $\forall s, t \in I, \text{Cov}(X_s, X_{t+s}) = \text{Cov}(X_0, X_t)$. La fonction de covariance est invariante par translation.

On note alors σ_X sa fonction d'autocovariance définie par

$$\sigma_X(t) = \text{Cov}(X_0, X_t).$$

Les processus stationnaires du second ordre sont aussi appelés dans la littérature processus à covariance stationnaire ou processus faiblement stationnaires.

Dans le cas de processus gaussiens, la loi du processus est entièrement déterminée par la moyenne et par la fonction de covariance. Ainsi, tandis que les processus strictement stationnaires et L^2 sont des processus stationnaires du second ordre, la réciproque est vraie pour les processus gaussiens.

Les principales propriétés satisfaites par une fonction d'autocovariance d'un processus stationnaire du second ordre sont redonnées dans la proposition suivante.

Proposition 1.2.3

Soit X un processus stationnaire du second ordre sur I de fonction d'autocovariance σ_X . Alors pour tout $t \in I$,

$$\sigma_X(0) = \text{Var}(X_t) \geq 0, \quad |\sigma_X(t)| \leq \sigma_X(0) \quad \text{et} \quad \sigma_X(t) = \sigma_X(-t).$$

On donne maintenant une représentation spectrale de la fonction d'autocovariance.

Proposition 1.2.4

Toute fonction d'autocovariance σ_X associée à un processus stationnaire du second ordre peut s'écrire :

- dans le cas discret, $\sigma_X(k) = \int_{-\pi}^{\pi} e^{ivk} dF(v)$, $k \in \mathbb{N}$,
- dans le cas continu, $\sigma_X(t) = \int_{\mathbb{R}} e^{ivt} dF(v)$, $t \in \mathbb{R}^+$,

où $F/\sigma_X(0)$ est une fonction de répartition correspondant à une variable aléatoire à valeurs p.s. dans $[-\pi, \pi]$ (resp. \mathbb{R}) dans le cas discret (resp. dans le cas continu). La fonction F est appelée la fonction de répartition spectrale du processus X .

Si la mesure F est absolument continue par rapport à la mesure de Lebesgue, alors la fonction f définie par $f(v) = \frac{dF(v)}{dv}$ est appelée la densité spectrale du processus X .

Définition 1.2.5

- Dans le cas discret, une fonction réelle f sur $[-\pi, \pi]$ est appelée une densité spectrale du processus stationnaire du second ordre $(X_k)_{k \in \mathbb{N}}$ si

$$f \geq 0 \text{ sur } [-\pi, \pi] \text{ et } \sigma_X(k) = \int_{-\pi}^{\pi} e^{ivk} f(v) dv, \forall k \in \mathbb{Z}.$$

- Dans le cas continu, une fonction réelle f sur \mathbb{R} est appelée une densité spectrale du processus stationnaire du second ordre $(X_t)_{t \in \mathbb{R}^+}$ si

$$f \geq 0 \text{ sur } \mathbb{R} \text{ et } \sigma_X(t) = \int_{\mathbb{R}} e^{ivt} f(v) dv, \forall t \in \mathbb{R}.$$

En particulier, une densité spectrale est une fonction paire et intégrable sur $[-\pi, \pi]$ (resp. sur \mathbb{R}) dans le cas discret (resp. dans le cas continu). Les processus stationnaires du second ordre n'admettent pas forcément de densité spectrale. Cependant, sous certaines hypothèses sur la fonction d'autocovariance, on a des résultats d'existence :

Proposition 1.2.6

- Dans le cas discret, si $\sum_{k \geq 0} |\sigma_X(k)| < \infty$, alors X admet une densité spectrale définie pour $v \in [-\pi, \pi]$ par

$$f(v) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{-ikv} \sigma_X(k).$$

Cette somme converge uniformément pour $v \in [-\pi, \pi]$ et f est continue.

- Dans le cas continu, si $\int_{\mathbb{R}} |\sigma_X(t)| dt < \infty$, alors X admet une densité spectrale définie pour $v \in \mathbb{R}$ par

$$f(v) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itv} \sigma_X(t) dt.$$

Cette intégrale converge uniformément pour $v \in \mathbb{R}$, i.e.

$$\sup_{v \in \mathbb{R}} \left| f(v) - \frac{1}{2\pi} \int_{-n}^n e^{-itv} \sigma_X(t) dt \right| \xrightarrow{n \rightarrow \infty} 0$$

et f est continue.

Un critère plus général d'existence est donné dans la proposition suivante.

Proposition 1.2.7

• Dans le cas discret, si $\sum_{k \geq 0} \sigma_X(k)^2 < \infty$, alors X admet une densité spectrale $f \in L^2([-\pi, \pi])$ définie pour $v \in [-\pi, \pi]$ par

$$f(v) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{-ikv} \sigma_X(k)$$

et

$$\int_{-\pi}^{\pi} \left| f(v) - \frac{1}{2\pi} \sum_{k=-n}^n e^{-ikv} \sigma_X(k) \right|^2 dv \xrightarrow{n \rightarrow \infty} 0.$$

• Dans le cas continu, si $\int_{\mathbb{R}} \sigma_X(t)^2 dt < \infty$, alors X admet une densité spectrale $f \in L^2(\mathbb{R})$ définie pour $v \in \mathbb{R}$ par

$$f(v) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itv} \sigma_X(t) dt$$

et on a de même

$$\int_{\mathbb{R}} \left| f(v) - \frac{1}{2\pi} \int_{-n}^n e^{-itv} \sigma_X(t) dt \right|^2 dv \xrightarrow{n \rightarrow \infty} 0.$$

Définition 1.2.8

Soit $X = (X_k)_{k \in \mathbb{N}}$ un processus stationnaire du second ordre. On définit le périodogramme au point $u \in [-\pi, \pi]$ par

$$I_n(u) := \frac{1}{2\pi n} \left| \sum_{k=1}^n e^{iku} X_k \right|^2.$$

Le périodogramme est une version empirique de la densité spectrale pour un processus X stationnaire du second ordre à temps discret. Bien que ce ne soit pas un estimateur consistant de la densité spectrale $f(u)$, il vérifie que $\mathbb{E}[I_n(u)] \xrightarrow{n \rightarrow \infty} f(u)$ en tout point de continuité de la fonction f .

On souhaite désormais décrire et mesurer la dépendance d'un processus stationnaire du second ordre dans le domaine temporel à partir de la fonction d'autocovariance et dans le domaine des fréquences grâce à la densité spectrale.

Définition 1.2.9

Soit X un processus stationnaire du second ordre.

- Dans le cas discret, X est dit à courte mémoire (resp. longue mémoire) si $\sum_{k \in \mathbb{Z}} |\sigma_X(k)| < \infty$ (resp. $\sum_{k \in \mathbb{Z}} |\sigma_X(k)| = \infty$).
- Dans le cas continu, X est dit à courte mémoire (resp. longue mémoire) si $\int_{\mathbb{R}} |\sigma_X(t)| dt < \infty$ (resp. $\int_{\mathbb{R}} |\sigma_X(t)| dt = \infty$).

La caractérisation dans le domaine temporel de la courte et longue mémoire est ainsi basée sur le comportement asymptotique de la fonction d'autocovariance. Dans le cas de la courte mémoire, la condition de sommabilité ou intégrabilité de la fonction d'autocovariance implique par la Proposition 1.2.6 l'existence d'une densité spectrale f bornée. Sous l'hypothèse d'une longue mémoire, la densité spectrale, si elle existe, est généralement non bornée en zéro.

1.2.2 Théorème central limite et convergence du processus des sommes partielles

Enfin, on donne des résultats de type théorème central limite dans le cas de processus stationnaires en temps discret. Soit $X = (X_k)_{k \in \mathbb{N}}$, un processus stationnaire du second ordre. On définit alors $S_n = \sum_{j=1}^n X_j$.

Si $(X_j)_{j \in \mathbb{N}}$ est un processus gaussien centré, on a $\frac{S_n}{\sqrt{\text{Var}(S_n)}} \sim \mathcal{N}(0, 1)$ et ainsi la question de la loi limite de S_n revient à trouver le comportement asymptotique de $\text{Var}(S_n)$. Dans le cas de processus linéaires, on a le résultat suivant :

Théorème 1.2.10 (Théorème 16.6.5 de [IL71])

Soit $X = (X_k)_{k \in \mathbb{N}}$ un processus linéaire stationnaire du second ordre de la forme $X_k = \sum_{j=0}^{\infty} a_j \zeta_{k-j}$ où les innovations $(\zeta_k)_{k \in \mathbb{N}}$ sont i.i.d, centrées et L^2 . Si $\text{Var}(S_n) \xrightarrow{n \rightarrow \infty} \infty$ alors $\frac{S_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, 1)$

On a également des résultats plus forts : on peut montrer la convergence faible (on définit plus précisément cette convergence par la suite) du processus des sommes partielles $S_n(\cdot)$ défini pour $\tau \in [0, 1]$ par $S_n(\tau) = \sum_{j=1}^{\lfloor n\tau \rfloor} X_j$. À n fixé, $S_n(\cdot)$ est une fonction étagée appartenant à $\mathcal{D}[0, 1]$.

On rappelle les notions de convergence faible (voir [Bil68]) pour une suite de processus stochastiques.

Définition 1.2.11

On dit que la suite de processus stochastiques $(Y_n(\cdot))_n$ dans $\mathcal{D}[0, 1]$ converge faiblement vers le processus stochastique $Y(\cdot) \in \mathcal{D}[0, 1]$ si pour toute fonction $f : \mathcal{D}[0, 1] \rightarrow \mathbb{R}$ bornée et continue (par rapport à la J_1 -topologie de Skorokhod dans $\mathcal{D}[0, 1]$), on a

$$\mathbb{E}[f(Y_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[f(Y)].$$

En particulier, cette convergence faible entraîne la convergence en loi de fonctionnelles continues du processus. Par exemple, on obtient

$$\sup_{\tau \in [0, 1]} |Y_n(\tau)| \xrightarrow[n \rightarrow \infty]{\text{Loi}} \sup_{\tau \in [0, 1]} |Y(\tau)| \quad \text{et} \quad \int_0^1 Y_n^2(\tau) d\tau \xrightarrow[n \rightarrow \infty]{\text{Loi}} \int_0^1 Y^2(\tau) d\tau.$$

Dans la suite, on va s'intéresser uniquement à des situations pour lesquelles le processus limite est continu : $\mathbb{P}(Y(\cdot) \in \mathcal{C}[0, 1]) = 1$. Dans ce cas, il est naturel de discuter de la convergence faible de $(Y_n(\cdot))_n$ dans l'espace $\mathcal{D}[0, 1]$ muni de la topologie uniforme (donnée par la norme uniforme et la σ -algèbre $\mathcal{B}_0[0, 1]$ des sous-ensembles de $\mathcal{D}[0, 1]$ générés par les boules fermées pour la norme uniforme). Cette convergence est appelée la convergence faible dans $\mathcal{D}[0, 1]$ muni de la norme uniforme. On donne une définition équivalente :

Définition 1.2.12

On dit que la suite de processus stochastique $(Y_n(\cdot))_n$ dans $\mathcal{D}[0, 1]$ converge faiblement dans $\mathcal{D}[0, 1]$ muni de la norme uniforme vers le processus stochastique $Y(\cdot)$ continu si pour toute fonction $f : \mathcal{D}[0, 1] \rightarrow \mathbb{R}$ bornée, mesurable (par rapport à $\mathcal{B}_0[0, 1]$) et continue (par rapport à la norme uniforme), on a

$$\mathbb{E}[f(Y_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[f(Y)].$$

Définition 1.2.13

On dit que la suite de processus stochastique $(Y_n(\cdot))_n$ dans $\mathcal{D}[0, 1]$ est tendue dans $\mathcal{D}[0, 1]$ muni de la norme uniforme si $\forall \varepsilon > 0$, il existe un ensemble compact $K \subset \mathcal{D}[0, 1]$ avec $K \in \mathcal{B}_0[0, 1]$ tel que pour tout $n \geq 1$,

$$\mathbb{P}(Y_n(\cdot) \in K) > 1 - \varepsilon.$$

La topologie uniforme est plus forte que la J_1 -topologie. La convergence faible dans $\mathcal{D}[0, 1]$ muni de la norme uniforme nous permet d'étendre la classe de fonctionnelles f pour lesquelles on a la convergence $\mathbb{E}[f(Y_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[f(Y)]$.

La méthode classique pour montrer la convergence faible d'une suite de processus dans $\mathcal{D}[0, 1]$ muni de la norme uniforme s'effectue en deux étapes :

1. on prouve la convergence en loi fini-dimensionnelles,
2. on vérifie que la suite de processus est tendue dans $\mathcal{D}[0, 1]$ muni de la norme uniforme.

On donne un critère de tension pour une suite de processus :

Théorème 1.2.14 (Théorème 15.5 de [Bil68])

Soit $(Y_n(\cdot))$ suite de processus dans $\mathcal{D}[0, 1]$ telle que $|Y_n(0)| = \mathcal{O}_p(1)$. On suppose qu'il existe une suite de fonctions croissantes et continues à droite $(F_n)_n$ sur $[0, 1]$ qui sont uniformément bornées et qui vérifient

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{|t-s| \leq \delta} |F_n(t) - F_n(s)| = 0.$$

De plus, on suppose qu'il existe $C > 0$, $\gamma > 0$ et $\alpha > 1/2$ tels que

$$\mathbb{E}[|Y_n(u) - Y_n(s)|^{2\gamma}] \leq C|F_n(u) - F_n(s)|^{2\alpha}, \quad 0 \leq s < u \leq 1$$

alors la suite $(Y_n(\cdot))_n$ est tendue dans $\mathcal{D}[0, 1]$ muni de la norme uniforme. Et si la suite $(Y_n(\cdot))_n$ converge faiblement dans $\mathcal{D}[0, 1]$ muni de la norme uniforme alors le processus limite est continu p.s.

Revenons à la convergence du processus des sommes partielles $S_n(\cdot)$. À l'aide du critère de tension ci-dessus, on peut montrer la tension de $\left(\frac{S_n(\cdot)}{\sqrt{\text{Var}(S_n)}}\right)_n$:

Proposition 1.2.15 (Prop 4.4.2 dans [Gir+12])

Soit $X = (X_k)_{k \in \mathbb{N}}$ un processus stationnaire du second ordre tel que $\text{Var}(S_n) = n^{1+2d}L(n)$ avec $0 < d < 1/2$ et L à variation lente. Alors la suite des processus des sommes partielles $\left(\frac{S_n(\cdot)}{\sqrt{\text{Var}(S_n)}}\right)_n$ est tendue par rapport à la norme uniforme.

Dans le cas de processus linéaires, on a le résultat de convergence en loi fini-dimensionnelles suivant :

Proposition 1.2.16 (Thm 2 de [Dav70])

Soit $X = (X_k)_{k \in \mathbb{N}}$ un processus linéaire stationnaire du second ordre de la forme $X_k = \sum_{j=0}^{\infty} a_k \zeta_{k-j}$ où les innovations $(\zeta_k)_{k \in \mathbb{N}}$ sont i.i.d centrées et L^2 . Si $\text{Var}(S_n) = n^{1+2d}L(n)$ avec $-1/2 < d < 1/2$ et L à variation lente alors

$$\frac{S_n(\cdot)}{\sqrt{\text{Var}(S_n)}} \xrightarrow[n \rightarrow \infty]{} B_{d+1/2}$$

où la convergence est en loi fini-dimensionnelles et $B_{d+1/2}$ est le mouvement Brownien fractionnaire de paramètre $d + 1/2$.

Dans le chapitre 5, on va montrer un résultat de convergence de la suite des processus des sommes partielles dans $\mathcal{D}[0, 1]$ muni de la norme uniforme pour le processus étudié.

1.3 Organisation de la thèse

1.3.1 Partie I : Estimations et simulations de processus stables

Nous avons rappelé à présent les propriétés sur les lois et processus stables qui nous serviront par la suite. Dans la première partie de cette thèse, on va s'intéresser à l'estimation de lois et processus (localement) stables et à la simulation de ces processus.

Chapitre 2

Le chapitre 2 correspond à la version française de l'article [LV+19] soumis en collaboration avec J. Lévy Véhel et A. Philippe. Dans ce chapitre, on s'intéresse à des méthodes d'estimation pour les lois stables.

En particulier, on estime l'indice de stabilité α ainsi que le paramètre d'échelle σ pour une loi stable $S_\alpha(\sigma, 0, 0)$ en utilisant les expressions des log-moments :

$$\mathbb{E}[\log |W|] = \left(\frac{1}{\alpha} - 1\right) \gamma + \log \sigma \quad \text{et} \quad \text{Var}(\log |W|) = \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12},$$

où $W \sim S_\alpha(\sigma, 0, 0)$ et γ est la constante d'Euler. Ces estimateurs des log-moments sont fortement consistants et admettent un théorème central limite. De plus, ces estimateurs sont robustes dans le cas de variables non identiquement distribuées

sous certaines conditions sur la non-stationnarité. Par exemple, nous montrons la consistance et un théorème central limite pour l'estimateur des log-moments dans le cas d'un échantillon (X_1, \dots, X_n) de variables indépendantes tels que $X_i \sim S_{\alpha_i}(\sigma, 0, 0)$ où les $(\alpha_i)_i$ sont déterministes ou aléatoires et suffisamment proches de α (voir Section 2.4).

Un certain nombre d'estimateurs existent déjà pour les paramètres des lois stables comme ceux proposés par Fama-Roll [FR71] et Mc Culloch [McC86] qui reposent sur les quantiles empiriques ou celui donné par Koutrouvelis [Kou80; Kou81] qui est basé sur la fonction caractéristique. Une étude des estimateurs existants pour les lois stables est menée [Wer95] qui conclut que la méthode de régression de Koutrouvelis est un bon estimateur.

Après comparaison des performances numériques des estimateurs de Koutrouvelis et des log-moments pour les paramètres α et σ , il apparaît que les performances dépendent de la valeur de α avec un estimateur des log-moments meilleur pour α petit et inversement. Nous donnons ensuite un estimateur combiné qui permet d'améliorer les performances de notre estimateur pour le cas de petits échantillons.

Enfin, nous étendons nos estimateurs au cas de variables aléatoires asymétriques et nous construisons un test sur l'asymétrie des données.

Ces estimateurs sont ensuite appliqués dans le cas de processus multistables (dont nous continuons l'étude aux chapitres suivants) ainsi que sur l'estimation d'un indice financier.

Chapitres 3 et 4

Les chapitres 3 et 4 s'inscrivent dans le prolongement des travaux sur les processus multistables définis à partir de la représentation de Ferguson-Klass-LePage effectué dans [LGLV12]. Le chapitre 3 donne une méthode de simulation rapide de tels processus tandis que le chapitre 4 est consacré à l'estimation de la fonction qui paramétrise ces processus.

Dans le chapitre 3, on commence par rappeler la définition des processus localisables (voir [Fal02], [Fal03]) et des processus multistables définis à partir d'un champ stochastique dans le cas d'espace de mesure finie ou σ -finie (voir [LGLV12]). En particulier, à partir d'une fonction d'intensité α , le processus multistable Y est défini comme le processus diagonal du champ de processus stables

$(X(t, u))_{(t, u) \in [0, 1]^2}$, c'est à dire $Y(t) = X(t, t)$, où le champ est de la forme

$$X(t, u) = C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} f(t, u, V_i), \quad (t, u) \in [0, 1]^2,$$

(voir Théorème 3.2.1 pour les notations).

On va simuler le processus multistable grâce au champ. Si on veut simuler le processus multistable aux instants $(k/N)_{k \in \{1, \dots, N\}}$, il faut connaître les valeurs du champ en position $(\frac{k}{N}, \frac{k}{N})$. Comme, à u fixé, $X(\cdot, u)$ est un processus $\alpha(u)$ -stable, il faut simuler les N processus $X(\cdot, \frac{k}{N})$ aux instants $(\frac{k}{N})_{k \in \{1, \dots, N\}}$. Cette méthode de calcul nécessite beaucoup de temps de calcul lorsque N est grand puisqu'il faut alors calculer N points de N processus stables.

Nous donnons dans ce chapitre, une simulation approchée de ces processus à l'aide d'une grille fixée de pas $1/K$ avec $K \in \mathbb{N}^*$. Cette méthode consiste à discrétiser de manière régulière l'intervalle $]0, 2[$ parcouru par la fonction α en $(2K - 1)$ valeurs possibles. Au lieu de simuler le processus diagonal $Y = (X(t, t))_{t \in [0, 1]}$, on simule alors le processus approché $(\tilde{X}_K(t, t))_{t \in [0, 1]}$ défini à partir du champ :

$$\tilde{X}_K(t, u) = C_{\alpha_K(u)}^{1/\alpha_K(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha_K(u)} f(t, u, V_i), \quad (1.3.1)$$

$$\text{où } \alpha_K(u) = \max \left(\frac{\lfloor K\alpha(u) \rfloor}{K}, \frac{1}{K} \right).$$

Cette construction approchée permet une simulation beaucoup plus rapide des processus multistables puisqu'elle nécessite au maximum la simulation de N points de $(2K - 1)$ processus stables.

Nous montrons le résultat suivant sur l'erreur entre le vrai processus et sa construction approchée avec une grille :

Théorème A (Théorème 3.3.1)

Soit $t > 0$ fixé. Sous de bonnes conditions, on a

$$X(t, t) - \widetilde{X}_K(t, t) = (\alpha(t) - \alpha_K(t))Z_K(t)$$

avec $Z_K(t) \xrightarrow[K \rightarrow \infty]{Loi} Z_\infty(t)$ où la variable aléatoire $Z_\infty(t)$ est non-dégénérée. En conséquence,

$$|X(t, t) - \widetilde{X}_K(t, t)| = \mathcal{O}_{\mathbb{P}}\left(\frac{1}{K}\right).$$

Nous illustrons cette méthode de simulation pour la réalisation d'un mouvement de Lévy multistable et pour un mouvement linéaire multifractionnaire multistable. Par exemple, la simulation par le champ d'un mouvement de Lévy multistable est environ 1800 fois plus lente que son approximation à l'aide d'une grille de pas $1/K$ où $K = 2000$ lorsque $N = 2 \cdot 10^5$.

Après avoir donné une méthode de simulation rapide dans le chapitre 3, le chapitre 4 est consacré à l'estimation de la fonction d'intensité des processus multistables. Au point $t_0 \in]0, 1[$, l'estimateur de $\alpha(t_0)$ est basé sur les moments empiriques d'ordre $p \in]0, 2]$ du processus multistable Y définis par :

$$S_N(p) = \left(\frac{1}{n(N)} \sum_{k \in I_N(t_0)} \left| Y\left(\frac{k+1}{N}\right) - Y\left(\frac{k}{N}\right) \right|^p \right)^{1/p}$$

sur une fenêtre de taille $\frac{I_N(t_0)}{N}$ centrée autour du point t_0 où $N \geq 1$, $n(N)$ est une suite d'entiers pairs et

$$I_N(t_0) = \left\{ \lfloor Nt_0 \rfloor - \frac{n(N)}{2}, \dots, \lfloor Nt_0 \rfloor + \frac{n(N)}{2} - 1 \right\}.$$

Dans [LG13], Le Guével a lui-aussi étudié un estimateur consistant de la fonction d'intensité basé sur les moments empiriques (voir Section 4.3 pour plus de détails).

Suivant la valeur de p , on a le résultat suivant de convergence pour $S_N(p)$:

Théorème B (Théorème 5.4 de [LG13] et Théorème 4.2.2)

Soit $n(N)$ tel que $n(N) \xrightarrow{N \rightarrow \infty} \infty$ et $\frac{n(N)}{N} \xrightarrow{N \rightarrow \infty} 0$. Alors, sous de bonnes conditions, on a

$$\forall p \in]0, \alpha(t_0)[, \quad N^{H(t_0)} S_N(p) \xrightarrow[N \rightarrow \infty]{\text{proba}} \mathbb{E} [|X(1, t_0)|^p]^{1/p},$$

et

$$\forall p \in]\alpha(t_0), 2[, \quad N^{H(t_0)} n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} S_N(p) \xrightarrow[N \rightarrow \infty]{\text{Loi}} W^{1/p},$$

où $W \sim S_{\alpha(t_0)/p}(\sigma_{p, \alpha(t_0)}, 1, 0)$.

On utilise ensuite l'expression de $\mathbb{E} [|X(1, t_0)|^p]^{1/p}$ pour $0 < p < \alpha(t_0)$ afin de construire un estimateur de $\alpha(t_0)$ dont on montre la convergence en probabilité vers $\alpha(t_0)$ si $0 < p < \alpha(t_0)$ et vers p sinon.

Théorème C (Théorème 4.3.1)

Pour $0 < \theta < 1$, soit $\Phi_\theta : [0, 1[\rightarrow]-\infty, 1]$ la fonction strictement décroissante définie par $\Phi_\theta(x) = \frac{\Gamma(1-\theta x)^{1/\theta}}{\Gamma(1-x)}$ et de bijection réciproque Φ_θ^{-1} . On peut alors définir l'estimateur

$$\hat{\alpha}_{N,p,\theta}(t_0) = \frac{p}{\Phi_\theta^{-1} \left(\min \left[1, \left(\frac{S_N(\theta p)}{S_N(p)} \right)^p \frac{1-p}{\Gamma(2-p) \cos(\frac{\pi p}{2})} \left(\frac{\Gamma(2-\theta p) \cos(\frac{\pi \theta p}{2})}{1-\theta p} \right)^{1/\theta} \right] \right)}.$$

On se place sous les conditions du théorème précédent. On a alors,

$$\hat{\alpha}_{N,p,\theta}(t_0) \xrightarrow[N \rightarrow \infty]{\text{proba}} \begin{cases} \alpha(t_0), & \text{si } p \in]0, \alpha(t_0)[, \\ p, & \text{si } p \in]\alpha(t_0), 2[. \end{cases}$$

1.3.2 Partie II : Échantillonnage de processus stationnaires du second ordre définis en temps continu

L'échantillonnage de processus a toujours été une question intéressante puisqu'il permet de modéliser à la fois des données en temps continu dont nous n'avons accès qu'à un certain nombre d'observations (grâce à un capteur par exemple) soit de modéliser un processus dont il nous manque des données.

Une première approche d'échantillonnage proposée consiste à regarder de manière régulière le processus aux instants $(nh)_{n \in \mathbb{Z}}$, avec h un pas de temps fixé ([BT58]). Cependant, cette méthode ne permet pas de retrouver précisément la densité spectrale f du processus initial (et donc aussi sa fonction de covariance) car cette méthode produit des alias de f , c'est à dire d'autres densités spectrales compatibles avec les valeurs du processus aux instants $(nh)_{n \in \mathbb{Z}}$.

Ainsi, [SS60] proposent d'étudier certains échantillonnages irréguliers afin d'avoir des méthodes sans alias. Une première étude aux instants aléatoires $(nh + \gamma_n)_{n \in \mathbb{Z}}$ où h est un pas de temps fixé et les γ_n sont des variables i.i.d. $\mathcal{N}(0, \sigma^2)$ avec $\sigma \ll h$ mène à un échantillonnage qui admet lui-aussi des alias. Une deuxième étude pour un échantillonnage aléatoire *additif* aux instants $(t_n)_{n \in \mathbb{Z}}$ tels que $t_n = t_{n-1} + \gamma_n$ avec des $(\gamma_n)_n$ i.i.d. de densité p sur \mathbb{R}^+ et supposée dans L^2 mène à un échantillonnage sans alias (sous certaines conditions sur la loi commune des γ_n) [SS60; GR57]. On peut trouver d'autres échantillonnages dans la littérature. Par exemple, [BB10] proposent d'étudier des processus gaussiens stationnaires aux instants $(t_0^{(n)}, \dots, t_n^{(n)})$ tels que $t_{k+1}^{(n)} - t_k^{(n)} = \delta_n L_k$ avec $(\delta_n)_n$ suite déterministe positive tendant vers 0 et $(L_k)_k$ suite de variables aléatoires positives tandis que [LM94] s'intéressent à un échantillonnage à l'aide d'un processus ponctuel stationnaire.

Dans une deuxième partie, nous nous sommes ainsi intéressés aux processus stationnaires du second ordre définis en temps continu $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+}$ de fonction d'autocovariance σ_X . On cherche ainsi à décrire le comportement du processus aux instants d'observations $(T_n)_n$ en supposant que les inter-observations $(T_n - T_{n-1})_n$ sont i.i.d. à densité sur \mathbb{R}^+ (échantillonnage aléatoire additif de [SS60]). On définit alors le processus échantillonné $(Y_n)_n$ par

$$Y_n = X_{T_n} \quad \text{pour } n \in \mathbb{N}.$$

De nombreuses études ont été menées pour des processus stationnaires du second ordre gaussiens, linéaires avec des innovations i.i.d. ou encore fonctions d'un processus linéaire ou gaussien (voir [Gir+12]). Ces résultats classiques ne sont pas applicables puisque même sous l'hypothèse d'une loi connue pour le processus initial, on ne peut généralement rien préciser sur la loi du processus échantillonné.

Chapitre 5

Le chapitre 5 correspond à la version française de l'article [Phi+19] écrit en collaboration avec A. Philippe et M-C. Viano dans lequel on s'intéresse au comportement de la mémoire et de la loi du processus après échantillonnage aléatoire. Dans [PV10], Philippe et Viano ont traité le cas de processus stationnaires du second ordre en temps discrets qui sont échantillonnés aléatoirement. Bien que nos résultats obtenus sont similaires à ceux de [PV10], les méthodes d'étude ne sont pas les mêmes. En effet, dans le cas discret, la loi des inter-arrivées est une variable aléatoire discrète à valeurs dans \mathbb{N}^* et le processus initial et son échantillonné sont tous les deux à temps discrets ce qui facilite l'étude.

On aborde en particulier dans ce chapitre des questions naturelles que l'on se pose sur les processus X et Y : *Les caractéristiques du processus initial sont-elles conservées après échantillonnage aléatoire ? Peut-on retrouver les propriétés du processus initial X à partir du processus échantillonné Y ?*

Dans ce chapitre, nous traitons uniquement des propriétés dans le domaine temporel de ces processus. Nous montrons que Y est lui un processus stationnaire du second ordre dont la fonction d'autocovariance σ_Y vérifie

$$\begin{cases} \sigma_Y(0) = \sigma_X(0), \\ \sigma_Y(h) = \mathbb{E}[\sigma_X(T_h)], \quad h \geq 1. \end{cases}$$

De plus, on montre que dans le cas d'un processus initial X gaussien, les marginales de Y sont gaussiennes. Cependant, si la variable aléatoire $\sigma_X(T_1)$ n'est pas constante p.s. alors Y n'est pas un processus gaussien (voir Proposition 5.2.2).

Sous des hypothèses peu restrictives, la courte mémoire est conservée après échantillonnage alors que la conservation de la longue mémoire nécessite des inter-arrivées intégrables. Cette condition sur le moment d'ordre 1 des inter-observations avaient déjà été soulignée par [PV10].

Pour la longue mémoire, le cas d'une fonction d'autocovariance à variation régulière de la forme

$$(\star) \quad \sigma_X(t) = t^{-1+2d}L(t)$$

(avec $0 < d < 1/2$ et L à variation lente et décroissante à partir d'un certain rang) est étudié plus précisément afin de donner davantage d'informations sur le comportement de la mémoire après échantillonnage. En particulier, on montre que le paramètre de mémoire d est conservé après échantillonnage lorsque $\mathbb{E}[T_1] < \infty$.

Ce résultat est similaire à ceux obtenus par [PV10] (échantillonnage aléatoire) et par [Cha98] et [Hwa00] qui étudient l'échantillonnage déterministe et l'agrégation temporelle de processus FARIMA(0, d , 0) dont le paramètre de mémoire d est conservé après l'échantillonnage.

On a vu que le caractère gaussien du processus initial était perdu après échantillonnage. On peut cependant, sous certaines hypothèses, retrouver de manière asymptotique un comportement gaussien pour le processus des sommes partielles après normalisation. C'est l'objet du théorème ci-dessous qui étend les résultats classiques pour le processus des sommes partielles concernant les processus linéaires à innovations i.i.d (voir Chapitre 4 de [Gir+12]).

Théorème D (Théorème 5.3.4)

Soit X un processus gaussien de covariance vérifiant (\star) . Alors, si $\mathbb{E}[T_1] < \infty$, on obtient

$$\gamma_d^{-1/2} L(n)^{-1/2} n^{-1/2-d} S_n(\cdot) \xrightarrow[n \rightarrow \infty]{} B_{\frac{1}{2}+d}(\cdot), \quad (1.3.2)$$

où la convergence a lieu dans $\mathcal{D}[0, 1]$ muni de la norme uniforme, $B_{\frac{1}{2}+d}$ est le mouvement brownien fractionnaire de paramètre $\frac{1}{2} + d$ et $\gamma_d = \frac{\mathbb{E}[T_1]^{-1+2d}}{d(1+2d)}$.

Cela nous conduit à étudier la statistique R/S introduite par Hurst [Hur51 ; Hur56] puis utilisée dans [Rip57 ; MW69b ; MT79 ; Ber94 ; Li+11] dont nous montrons la convergence en loi (Proposition 5.3.1). Cette statistique nous donne une manière heuristique pour estimer le paramètre de longue mémoire d .

Dans le cas d'inter-arrivées non-intégrables, on montre un encadrement sur la mémoire de Y dépendant du paramètre de mémoire d et du paramètre

$$\beta := \sup\{\gamma : \mathbb{E}[T_1^\gamma] < \infty\} \in]0, 1[$$

dépendant de la loi de T_1 . En particulier, échantillonner un processus à longue mémoire avec des inter-arrivées non intégrables peut conduire à un processus à courte mémoire.

Enfin, nous étudions le processus agrégé :

$$Z_n = \int_{T_n}^{T_{n+1}} X_t dt,$$

qui, lui aussi, est stationnaire du second ordre. Dans le cas d'une fonction σ_X décroissante sur \mathbb{R}^+ , nous montrons que le processus agrégé hérite des propriétés

du processus discret Y défini après échantillonnage.

Chapitre 6

Le chapitre 6 correspond à un travail effectué en collaboration avec M. Ould Haye et A. Philippe. Ce chapitre est consacré à l'étude spectrale après échantillonnage du processus X stationnaire du second ordre à temps continu. Dans le cas d'un processus X à temps discret, il n'est pas possible de reconstruire la densité spectrale de f_X à partir de la connaissance d'une version échantillonnée obtenue par un échantillonnage déterministe dû à un phénomène d'alias. Cet aliasing peut être supprimé en prenant un échantillonnage aléatoire (différents schémas d'échantillonnage peuvent être trouvés dans [SS60], [Beu70] et [Mas71]).

On commence par montrer des résultats d'existence des densités spectrales de X et de son échantillonné Y ainsi que le lien entre ces deux densités spectrales. Plus particulièrement, la Proposition 6.2.1 montre l'existence dans le cas d'une fonction σ_X bornée par une fonction décroissante et dans L^p avec $1 \leq p \leq 2$ alors que la Proposition 6.2.2 montre que sous l'hypothèse de l'existence d'une densité spectrale avec une certaine forme ou sur des conditions sur la loi de T_1 , on a existence d'une densité spectrale pour Y avec son expression.

Le reste du chapitre se consacre à l'étude spectrale dans le cas particulier d'un processus initial gaussien à longue mémoire de fonction de covariance de la forme (\star) admettant une densité spectrale f_X avec un échantillonnage poissonnien. En particulier, dans ce contexte, on a existence d'une densité spectrale f_Y pour Y .

Pour ces processus, en définissant pour $j \in \{1, \dots, \lfloor n/2 \rfloor\}$, le périodogramme de (Y_1, \dots, Y_n) à la fréquence de Fourier $\lambda_j = \frac{2\pi j}{n}$: $I_n(\lambda_j) = \frac{1}{2\pi n} \left| \sum_{k=1}^n Y_k e^{ik\lambda_j} \right|^2$, nous montrons alors la convergence de ce périodogramme après normalisation :

Théorème E (Théorème 6.3.1)

Sous les hypothèses décrites ci-dessus, on a alors pour tout nombre de fréquences de Fourier $s \geq 1$ fixé, et pour tous $j_1, \dots, j_s \in \{1, \dots, \lfloor n/2 \rfloor\}$ entiers tous distincts,

$$\left(\frac{I_n(\lambda_{j_1})}{f_Y(\lambda_{j_1})}, \dots, \frac{I_n(\lambda_{j_s})}{f_Y(\lambda_{j_s})} \right) \xrightarrow[n \rightarrow +\infty]{\text{Loi}} \left(L_{j_1}(d)[U^2(j_1) + V^2(j_1)], \dots, L_{j_s}(d)[U^2(j_s) + V^2(j_s)] \right),$$

où $(U(1), \dots, U(\lfloor n/2 \rfloor), V(1), \dots, V(\lfloor n/2 \rfloor))$ est un vecteur gaussien centré de

matrice de covariance connue s'exprimant en fonction de (j_1, \dots, j_s) et d (voir 6.3.2) et $L_j(d)$ est une constante dépendant de j et d .

Enfin, à l'aide d'une condition sur les cumulants d'ordre 4 vérifiée par le processus Y (Proposition 6.4.1), on montre à la fois la convergence de l'estimateur local Whittle pour le paramètre d à l'aide des travaux de cet estimateur pour les processus non-linéaires de [Dal+06] et la convergence de l'estimateur HAC [Aba+09] de la variance asymptotique nécessaire pour inférer la moyenne inconnue. La convergence de l'estimateur HAC nous permet également de tester la courte mémoire du processus à l'aide de la statistique V/S [MW69a; Kwi+92; LS96; Gir+03; Gir+06].



Estimation et simulation de processus stables



2

Estimateurs explicites et combinés pour les paramètres de lois stables

Sommaire

2.1	Introduction	44
2.2	Méthodes d'estimation	47
2.2.1	Log-moments dans le cas symétrique	48
2.2.2	Performance numérique de chaque estimateur	50
2.2.3	Estimateur combiné	56
2.3	Lois stables asymétriques	61
2.3.1	Adaptation des estimateurs pour le cas asymétrique	61
2.3.2	Test de symétrie	63
2.4	Cas de variables stables non-identiquement distribuées	66
2.4.1	Perturbations déterministes	66
2.4.2	Perturbations aléatoires	67
2.5	Quelques applications de l'estimateur combiné	68
2.5.1	Résultats numériques sur des données synthétisées : mouvement de Lévy multistable	68
2.5.2	Application en finance	73

2.6 Annexe 77

Dans ce chapitre, on s'intéresse à l'estimation de l'indice de stabilité et du paramètre d'échelle de variables aléatoires stables. Bien que la littérature sur ce sujet soit conséquente, aucun résultat théorique précis ne semble disponible. Dans la suite, nous étudions un estimateur basé sur les log-moments qui sont toujours bien définis pour ces variables aléatoires. Cet estimateur a l'avantage de pouvoir s'écrire sous une expression explicite simple ce qui nous permet d'en déduire des résultats de convergence presque sûre et un théorème central limite. Nous montrons ensuite comment améliorer la précision de cet estimateur en le combinant avec ceux déjà définis. La forme explicite nous permet également de considérer le cas de données non identiquement distribuées, et nous montrons que nos résultats restent vrais, à condition que les écarts par rapport à la stationnarité soient "petits". À l'aide d'un centrage et d'une symétrie, nous élargissons les estimateurs précédents aux variables stables asymétriques et nous construisons un test pour vérifier l'asymétrie des données. En applications, nous montrons numériquement que l'indice de stabilité du mouvement de Lévy multistable peut être estimé avec précision et nous considérons un indice financier, à savoir le S&P 500, où nous trouvons que l'indice de stabilité évolue dans le temps en fonction des événements financiers majeurs.

2.1 Introduction

La famille de loi α -stable est omniprésente en théorie des probabilités puisque ces lois apparaissent comme les limites de sommes renormalisées de variables aléatoires indépendantes et identiquement distribuées. La densité d'une loi α -stable existe et est continue mais elle n'a pas de forme explicite sauf dans le cas de variables gaussiennes, de Cauchy, de Lévy ou dans le cas des variables constantes. Les lois stables non-gaussiennes sont un modèle de représentation de phénomènes réels comportant des sauts. De plus, lorsque $\alpha < 2$, leur densité admet un comportement de loi à queue lourde, résultant d'une décroissance lente de la probabilité de voir apparaître des événements extrêmes.

Les lois stables ont été largement utilisées ces dernières années pour la modélisation dans de nombreux domaines tels que la biomédecine [SG+13], la géophysique [Yan+09], l'économie [Man97], la finance [Man97], le trafic sur In-

ternet [Dim+11] et bien d'autres. Les lois stables sont caractérisées par quatre paramètres : un paramètre d'échelle généralement noté σ , (ce paramètre est proportionnel à la variance dans le cas gaussien), un paramètre de stabilité noté α , qui détermine le poids de la queue de distribution et qui varie dans l'intervalle $]0, 2]$, un paramètre de localisation μ comparable à la moyenne dans le cas de lois gaussiennes, et un paramètre d'asymétrie β appartenant à l'intervalle $[-1, 1]$. Afin de pouvoir utiliser des lois stables pour la modélisation des phénomènes du monde réel, il faut être capable d'estimer ces paramètres ci-dessus. Un certain nombre d'estimateurs sont couramment utilisés comme ceux proposés par Fama-Roll [FR71], Mc Culloch [McC86] et Koutrouvelis [Kou80 ; Kou81]. Ces estimateurs présentent un inconvénient : ils ne possèdent pas d'expression explicite simple. En conséquence, à notre connaissance, aucun résultat théorique n'est connu à leurs sujets, tel que des théorèmes de convergence presque sûre ou des théorèmes central limite. Leurs lois asymptotiques ainsi que leurs variances asymptotiques ne sont donc accessibles que par des simulations numériques.

De plus, l'absence de formes explicites simples rend difficile l'évaluation théorique de leur performance dans des situations qui s'écartent légèrement des hypothèses classiques des échantillons aléatoires indépendants et identiquement distribués. Or cela est souhaitable lorsque l'on veut traiter des données du monde réel qui, souvent, ne vérifient pas ces hypothèses idéales. Il est possible de trouver d'autres estimateurs pour les lois stables (moments négatifs dans [DI17], maximum de vraisemblance à partir d'une approximation de transformée de Fourier [Nol98]...). Ces estimateurs sont complexes à étudier.

Notre principal objectif est d'étudier les propriétés théoriques d'un estimateur des moments. Puisque les variables aléatoires stables non gaussiennes ne possèdent pas de variance finie et, même dans certains cas une moyenne bien définie, nous allons considérer un estimateur basé sur les moments logarithmiques. Cette idée n'est pas nouvelle, car il a été remarqué depuis longtemps que les log-moments existent toujours pour les variables aléatoires stables et qu'ils constituent des outils de travail pratiques. Par exemple, [MN95] considère le même estimateur que celui que nous étudions et l'applique à l'identification des canaux aveugles tandis que [Wan+15] l'utilise pour l'estimation d'un bruit α -stable pour l'erreur aléatoire d'un gyroscope laser. La forme de cet estimateur est connue mais pas son comportement. Grâce à son expression simple, nous sommes en mesure de prouver une convergence presque sûre et un théorème central limite à la fois dans un

cadre identiquement distribué et indépendant et dans le cas d'un léger écart de stationnarité. Nous comparons la performance de notre estimateur avec la méthode de régression de Koutrouvelis [Kou80 ; Kou81]. Les résultats dépendent de la valeur de α et de la taille de l'échantillon. Nous combinons ensuite ces deux estimateurs à l'aide d'une technique récemment mise au point dans [LR16] pour améliorer leur performance, en particulier dans le cas des petits échantillons.

Comme applications, nous montrons des résultats numériques à la fois sur des données synthétisées (mouvement symétrique de Lévy multistable) et sur un indice financier (S&P 500), qui confirment nos résultats théoriques, à savoir que l'estimateur est capable de suivre des variations assez lisses de l'indice de stabilité dans le temps.

Une loi stable est caractérisée par quatre paramètres. Nous écrivons $X \sim S_\alpha(\sigma, \beta, \mu)$ pour indiquer que X suit une loi stable d'indice de stabilité $\alpha \in]0, 2]$, de paramètre d'échelle $\sigma \in \mathbb{R}^+$, d'asymétrie $\beta \in [-1, 1]$ et paramètre de position $\mu \in \mathbb{R}$ (voir [ST94]).

Il existe plusieurs paramétrages des lois stables, chacun ayant ses avantages et ses inconvénients. Le suivant, de fonction caractéristique ϕ définie ci-après, est probablement le plus populaire :

$$\phi(t) = \begin{cases} \exp\left(-\sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu t\right) & \text{si } \alpha \neq 1, \\ \exp\left(-\sigma |t| \left(1 + i\beta \operatorname{sign}(t) \frac{2 \log |t|}{\pi}\right) + i\mu t\right) & \text{si } \alpha = 1 \end{cases} \quad (2.1.1)$$

où

$$\operatorname{sign}(t) = \begin{cases} 1 & \text{si } t > 0, \\ 0 & \text{si } t = 0, \\ -1 & \text{si } t < 0. \end{cases}$$

Pour une loi stable, la densité n'est pas connue de manière explicite sauf pour quelques exceptions (loi gaussienne, loi de Cauchy, ...). Il est donc difficile d'utiliser les méthodes d'estimation classiques basées sur la densité, comme le maximum de vraisemblance ou les statistiques bayésiennes. Bien que la densité puisse être estimée numériquement, par exemple par inversion de Fourier, cette procédure entraîne des erreurs d'approximation qui ne peuvent être facilement évaluées.

Sauf dans le cas gaussien, une autre difficulté vient du fait que les variables

aléatoires stables ont des moments infinis dès que l'ordre est plus grand que l'indice de stabilité. Plus précisément, lorsque $X \sim S_\alpha(\sigma, \beta, \mu)$ avec $0 < \alpha < 2$, $\mathbb{E}[|X|^p] < \infty$ si et seulement si $0 < p < \alpha$ (voir Proposition 1.1.2). Cela implique qu'il n'est pas possible d'utiliser la méthode des moments afin d'estimer les paramètres. Nous étudions dans la section suivante un estimateur simple basé sur les moments logarithmiques qui permet d'effectuer des calculs explicites.

Dans la section 2.2, nous étudions des estimateurs de α et σ pour des variables aléatoires stables symétriques (c'est-à-dire lorsque $\mu = \beta = 0$) : des estimateurs des log-moments et des estimateurs combinés construits avec l'estimateur de Koutrouvelis.

Dans la section 2.3, nous étendons les estimateurs des log-moments, de Koutrouvelis et combiné au cas asymétrique en étudiant deux façons d'adapter l'estimateur log-moments. Les propriétés des estimateurs des log-moments nous permettent également de proposer une méthode pour tester l'asymétrie des données. Dans la section 2.4, nous étudions le cas de variables non identiquement distribuées où nous prouvons la robustesse des estimateurs des log-moments sous certaines conditions pour les perturbations. Dans la section 2.5, nous effectuons des simulations numériques pour tester notre estimateur pour un mouvement de Lévy multistable et également pour des données réelles avec l'étude d'un indice financier.

2.2 Méthodes d'estimation

Nous construisons un estimateur pour les paramètres α et σ en utilisant un estimateur combiné dont la procédure générale de construction est décrite dans [LR16]. Dans notre cas particulier, $\theta = {}^t(\alpha, \sigma)$ sont les paramètres à estimer et nous avons accès à p estimateurs pour α et q estimateurs pour σ .

Nous considérons des estimateurs moyennés de θ de la forme

$$\hat{\theta}_\lambda = {}^t\lambda \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_p \\ \hat{\sigma}_1 \\ \vdots \\ \hat{\sigma}_q \end{pmatrix}, \quad \lambda \in \Lambda, \quad (2.2.1)$$

où ${}^t\lambda$ désigne la transposée de λ et Λ est un sous-ensemble donné de $\mathcal{M}_{(p+q)\times 2}(\mathbb{R})$.

On mesure la performance de $\hat{\theta}_\lambda$ pour $\lambda \in \Lambda$ à l'aide de l'erreur quadratique moyenne (MSE). On sélectionne λ^* qui satisfait la condition :

$$\lambda^* = \operatorname{argmin}_{\lambda \in \Lambda} \mathbb{E} [\| \hat{\theta}_\lambda - \theta \|^2].$$

Dans ce qui suit, nous présentons un estimateur classique : celui de Koutrouvelis ([Kou80], [Kou81]). Le choix de cet estimateur particulier est motivé par les résultats rapportés dans [Wer95] qui montrent qu'il fonctionne généralement mieux que d'autres méthodes telles que celles de Fama-Roll et Mc Culloch. Nous voulons combiner l'estimateur de Koutrouvelis décrit dans la section 2.2.2 avec l'estimateur des log-moments (Section 2.2.1) pour obtenir de meilleures performances surtout quand $\alpha < 1$.

2.2.1 Log-moments dans le cas symétrique

Pour une loi stable symétrique, les log-moments absolus peuvent être explicités [LG13], cela nous permet d'en déduire des expressions reliant les log-moments aux paramètres α et σ et d'obtenir un moyen de les estimer. Tout d'abord, on remarque la propriété suivante :

Proposition 2.2.1

Soit $Z \sim S_\alpha(1, 0, 0)$, alors

$$\mathbb{E}[|\log |Z||^p] < \infty \quad \text{pour tout } p > 0. \quad (2.2.2)$$

Démonstration : Voir l'annexe 2.6. ■

Ces espérances peuvent être calculées explicitement en remarquant que :

$$\mathbb{E}[(\log |Z|)^p] = \left. \frac{d^p \mathbb{E}[|Z|^t]}{dt^p} \right|_{t=0}, \quad (2.2.3)$$

et en utilisant le résultat suivant :

Proposition 2.2.2

Soit $Z \sim S_\alpha(1, 0, 0)$. Alors, pour tout $0 < t < \min(\alpha, 1)$,

$$\mathbb{E}[|Z|^t] = \frac{\Gamma(1 - t/\alpha)}{\Gamma(1 - t) \cos(\pi t/2)}. \quad (2.2.4)$$

Démonstration : Voir l'annexe 2.6. ■

On en déduit l'espérance et la variance d'une variable $\log |Z|$ où Z suit une loi $S_\alpha(1, 0, 0)$: $\mathbb{E}[\log |Z|] = \left(\frac{1}{\alpha} - 1\right) \gamma$ et $\text{Var}(\log |Z|) = \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12}$, où γ est la constante d'Euler ($\gamma = 0.57721 \dots$).

Théorème 2.2.3

Soit (X_1, \dots, X_n) une suite de variables aléatoires i.i.d stables symétriques réduites $S_\alpha(1, 0, 0)$. On définit

$$\hat{\alpha}_n(X_1, \dots, X_n) = \frac{\gamma}{\gamma + \frac{1}{n} \sum_{i=1}^n \log |X_i|}.$$

Alors, $\hat{\alpha}_n \xrightarrow[n \rightarrow +\infty]{p.s.} \alpha$. De plus, avec $f(x) = \frac{\pi^2}{6x^2} + \frac{\pi^2}{12}$,

$$\frac{\sqrt{n}(\hat{\alpha}_n - \alpha)\gamma}{\hat{\alpha}_n^2 \sqrt{f(\hat{\alpha}_n)}} \xrightarrow[n \rightarrow +\infty]{Loi} \mathcal{N}(0, 1). \quad (2.2.5)$$

Démonstration : La preuve de ce résultat se trouve en annexe 2.6. ■

Généralement, σ est inconnu et on doit estimer conjointement les deux paramètres d'intérêt. Soit W suivant une loi $S_\alpha(\sigma, 0, 0)$, alors on a la réduction $\frac{W}{\sigma} \sim S_\alpha(1, 0, 0)$ et on déduit les log-moments de W à partir de ceux de Z . On obtient

$$\mathbb{E}[\log |W|] = \left(\frac{1}{\alpha} - 1\right) \gamma + \log \sigma \quad \text{et} \quad \text{Var}(\log |W|) = \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12}.$$

Théorème 2.2.4

Soit (X_1, \dots, X_n) une suite de variables aléatoires i.i.d. de loi stable symétrique $S_\alpha(\sigma, 0, 0)$. On définit les estimateurs $\hat{\alpha}_{LOG}^{(n)}$ et $\hat{\sigma}_{LOG}^{(n)}$ (en fonction de

(X_1, \dots, X_n) par

$$\begin{cases} \hat{\alpha}_{LOG}^{(n)} = \left(\max \left(\frac{6}{\pi^2(n-1)} \sum_{i=1}^n \left[\log |X_i| - \frac{1}{n} \sum_{k=1}^n \log |X_k| \right]^2 - \frac{1}{2}, \frac{1}{4} \right) \right)^{-1/2}, \\ \hat{\sigma}_{LOG}^{(n)} = \exp \left(\frac{1}{n} \sum_{i=1}^n \log |X_i| - \left(\frac{1}{\hat{\alpha}_{LOG}^{(n)}} - 1 \right) \gamma \right) \end{cases} \quad (2.2.6)$$

Alors, on a

$$(\hat{\alpha}_{LOG}^{(n)}, \hat{\sigma}_{LOG}^{(n)}) \xrightarrow[n \rightarrow +\infty]{p.s.} (\alpha, \sigma). \quad (2.2.7)$$

et

$$\sqrt{n} \left(\begin{pmatrix} \hat{\alpha}_{LOG}^{(n)} \\ \hat{\sigma}_{LOG}^{(n)} \end{pmatrix} - \begin{pmatrix} \alpha \\ \sigma \end{pmatrix} \right) \xrightarrow[n \rightarrow +\infty]{Loi} \mathcal{N}(0, A_{\alpha, \sigma}), \quad (2.2.8)$$

où la matrice $A_{\alpha, \sigma}$ est explicitée dans la remarque ci-dessous.

Remarque 2.2.1. La fonction max dans l'expression de $\hat{\alpha}_{LOG}^{(n)}$ assure que l'estimateur est toujours calculable et qu'il est à valeurs dans $]0, 2]$ presque sûrement. La matrice $A_{\alpha, \sigma}$ s'exprime en fonction des moments de $\log |X_1|$, de α et σ par la relation :

$$A_{\alpha, \sigma} = F_{\alpha, \sigma} G_{\alpha, \sigma} \Sigma_{\alpha, \sigma} {}^t G_{\alpha, \sigma} {}^t F_{\alpha, \sigma}$$

où

$$F_{\alpha, \sigma} = \begin{pmatrix} \sigma & \gamma\sigma/\alpha^2 \\ 0 & 1 \end{pmatrix}, \quad (2.2.9)$$

$$G_{\alpha, \sigma} = \begin{pmatrix} 1 & 0 \\ 6 \left(\left(\frac{1}{\alpha} - 1 \right) \gamma + \log \sigma \right) \alpha^3 / \pi^2 & -3\alpha^3 / \pi^2 \end{pmatrix}, \quad (2.2.10)$$

et

$$\Sigma_{\alpha, \sigma} = \begin{pmatrix} \text{Var}(\log |X_1|) & \text{Cov}(\log |X_1|, (\log |X_1|)^2) \\ \text{Cov}(\log |X_1|, (\log |X_1|)^2) & \text{Var}((\log |X_1|)^2) \end{pmatrix}. \quad (2.2.11)$$

Démonstration : La preuve de ce résultat se trouve en annexe 2.6. ■

2.2.2 Performance numérique de chaque estimateur

L'estimateur de Koutrouvelis (voir [Kou80] et [Kou81]) est basé sur l'exploitation de l'expression explicite du logarithme itéré de la fonction caractéristique ϕ .

Dans le cas symétrique, il prend la forme particulièrement simple suivante :

$$\log(\log(|\phi(t)|^2)) = \log(2\sigma^\alpha) + \alpha \log |t|. \quad (2.2.12)$$

La fonction caractéristique empirique donnée par $\hat{\phi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j}$ basée sur les observations i.i.d (X_j) est un estimateur consistant de ϕ . Nous estimons ces paramètres en régressant $y = \log(\log(|\hat{\phi}_n(t)|^2))$ sur $w = \log |t|$ dans le modèle

$$y_k = m + \alpha w_k + \epsilon_k \quad (2.2.13)$$

où $m = \log(2\sigma^\alpha)$, $t_k = \frac{\pi k}{25}$ pour $k \in \{1, \dots, K\}$ avec K dépendant de la valeur de α et de la taille de l'échantillon, et ϵ_k désigne un terme d'erreurs. Dans nos simulations, nous utilisons une version plus facile de la méthode de régression de Koutrouvelis qui est plus adaptée au cas symétrique (voir [Wer95]). Nous décrivons l'algorithme :

- On définit une erreur admissible ($tol = 0.05$ dans la simulation) et le nombre maximal d'itérations ($iter_{max} = 10$).
- Une regression linéaire appliquée à la méthode des quantiles de Mc Culloch [McC86] nous fournit des estimations initiales $\hat{\alpha}^{(n)}$ et $\hat{\sigma}^{(n)}$ pour les paramètres, où n désigne la taille de l'échantillon observé.
- Tant que le nombre d'itérations est inférieur à $iter_{max}$ et que $|\hat{s} - 1| > tol$ (cette condition démarre à la deuxième itération) :
 - on trouve le nombre de points K dont on va se servir pour la régression, dépendant de $\hat{\alpha}^{(n)}$ comme dans la regression classique de Koutrouvelis,
 - on définit $w = (w_k)_{k \in \{1 \dots K\}}$ et $y = (y_k)_{k \in \{1 \dots K\}}$ par $w_k = \log |\pi k / 25|$,

$$y_k = \log\left(-\log\left(\left|\hat{\phi}_n\left(\frac{\pi k}{25\hat{\sigma}^{(n)}}\right)\right|^2\right)\right),$$

- on calcule la nouvelle valeur de $\hat{\alpha}^{(n)}$ donnée par

$$\hat{\alpha}^{(n)} = \min\left(\frac{\sum_{k=1}^K (w_k - \text{mean}(w))(y_k - \text{mean}(y))}{\sum_{k=1}^K (w_k - \text{mean}(w))^2}, 2\right)$$

- on met à jour $\hat{s} = \exp\left(\frac{\text{mean}(y - \hat{\alpha}^{(n)} w) - \log(2)}{\hat{\alpha}^{(n)}}\right)$
- on met à jour $\hat{\sigma}^{(n)} = \hat{\sigma}^{(n)} \hat{s}$.

Cette version modifiée de Koutrouvelis donne des performances (en termes

d'erreurs quadratiques moyennes) similaires à l'original. Cependant, elle est beaucoup plus rapide car cette version ne nécessite pas l'estimation des paramètres β et μ (qui nécessite l'inversion numérique de matrices de taille $n \times n$).

En général, la méthode de régression de Koutrouvelis a une erreur quadratique moyenne plus petite que les autres méthodes d'estimation classiques basées sur des quantiles empiriques (par exemple la méthode Fama-Roll [FR71] ou la méthode de Mc Culloch [McC86]). Une comparaison des performances de ces méthodes classiques est effectuée dans [Wer95]. Pour chaque paire de valeurs pour (α, σ) , r échantillons indépendants de taille n de variables aléatoires i.i.d. stables symétriques ont été générés en utilisant la méthode de simulation donnée par Chambers *et al* [Cha+76]. L'erreur quadratique moyenne empirique des lois d'échantillonnage de α et σ est donnée par

$$MSE_{\alpha} = \frac{1}{r} \sum_{i=1}^r (\hat{\alpha}_i - \alpha)^2, \quad MSE_{\sigma} = \frac{1}{r} \sum_{i=1}^r (\hat{\sigma}_i - \sigma)^2,$$

où $\hat{\alpha}$ (resp. $\hat{\sigma}$) est un estimateur de α (resp. σ).

Dans la suite, nous utilisons les abréviations "KOUT" et "LOG" pour désigner respectivement l'estimateur de Koutrouvelis ou des log-moments. Pour chaque α , les comportements de $\hat{\alpha}_{KOUT}$ et $\hat{\alpha}_{LOG}$ sont similaires pour toute valeur de σ (Table 2.1) alors que, pour chaque valeur de σ , $\hat{\sigma}_{KOUT}$ et $\hat{\sigma}_{LOG}$ s'améliorent lorsque α augmente (Tableau 2.2). En outre, lorsque α est fixe, $\hat{\sigma}_{KOUT}$ et $\hat{\sigma}_{LOG}$ ont le même comportement pour tout σ . Pour cette raison, seul le cas où σ est égal à 1 est représenté dans la Table 2.2.

		$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 1$	$\alpha = 1.4$	$\alpha = 1.8$
$\sigma = 10$	LOG	$9.06 \cdot 10^{-5}$	$9.06 \cdot 10^{-4}$	$4.67 \cdot 10^{-3}$	$1.97 \cdot 10^{-2}$	$3.10 \cdot 10^{-2}$
	KOUT	$4.70 \cdot 10^{-4}$	$2.35 \cdot 10^{-3}$	$3.75 \cdot 10^{-3}$	$8.20 \cdot 10^{-3}$	$4.17 \cdot 10^{-3}$
$\sigma = 1$	LOG	$8.07 \cdot 10^{-5}$	$1.06 \cdot 10^{-3}$	$4.47 \cdot 10^{-3}$	$2.20 \cdot 10^{-2}$	$3.19 \cdot 10^{-2}$
	KOUT	$4.27 \cdot 10^{-4}$	$2.11 \cdot 10^{-3}$	$3.91 \cdot 10^{-3}$	$7.58 \cdot 10^{-3}$	$4.28 \cdot 10^{-3}$
$\sigma = 0.1$	LOG	$8.93 \cdot 10^{-5}$	$9.59 \cdot 10^{-4}$	$4.43 \cdot 10^{-3}$	$2.20 \cdot 10^{-2}$	$2.97 \cdot 10^{-2}$
	KOUT	$4.66 \cdot 10^{-4}$	$1.96 \cdot 10^{-3}$	$4.22 \cdot 10^{-3}$	$7.58 \cdot 10^{-3}$	$4.29 \cdot 10^{-3}$

TABLE 2.1 – Erreur quadratique moyenne pour $\hat{\alpha}_{LOG}$ et $\hat{\alpha}_{KOUT}$ ($r = 500$ et $n = 500$).

		$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 1$	$\alpha = 1.4$	$\alpha = 1.8$
$\sigma = 10$	LOG	7.10	$9.86 \cdot 10^{-1}$	$6.08 \cdot 10^{-1}$	$6.07 \cdot 10^{-1}$	$5.43 \cdot 10^{-1}$
	KOUT	11.5	$9.92 \cdot 10^{-1}$	$4.82 \cdot 10^{-1}$	$3.56 \cdot 10^{-1}$	$1.72 \cdot 10^{-1}$
$\sigma = 1$	LOG	$8.20 \cdot 10^{-2}$	$9.45 \cdot 10^{-3}$	$6.52 \cdot 10^{-3}$	$6.98 \cdot 10^{-3}$	$5.77 \cdot 10^{-3}$
	KOUT	$1.18 \cdot 10^{-1}$	$9.82 \cdot 10^{-3}$	$4.44 \cdot 10^{-3}$	$3.83 \cdot 10^{-3}$	$1.56 \cdot 10^{-3}$
$\sigma = 0.1$	LOG	$6.73 \cdot 10^{-4}$	$8.77 \cdot 10^{-5}$	$6.71 \cdot 10^{-5}$	$6.54 \cdot 10^{-5}$	$5.89 \cdot 10^{-5}$
	KOUT	$8.63 \cdot 10^{-4}$	$8.60 \cdot 10^{-5}$	$4.75 \cdot 10^{-5}$	$3.35 \cdot 10^{-5}$	$1.82 \cdot 10^{-5}$

TABLE 2.2 – Erreur quadratique moyenne pour $\hat{\sigma}_{LOG}$ et $\hat{\sigma}_{KOUT}$ ($r = 500$ et $n = 500$).

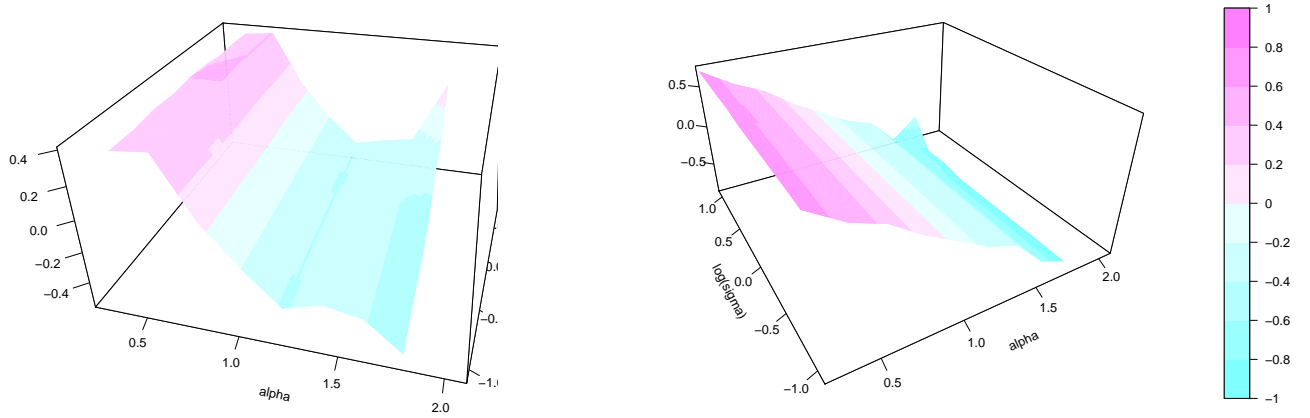
Avec une étude par simulation de variables stables, nous comparons les erreurs quadratiques moyennes empiriques de α et σ pour les méthodes présentées plus haut. Nous représentons le rapport logarithmique des erreurs quadratiques moyennes pour α et σ pour les deux méthodes d'estimation à partir de plusieurs échantillons de taille $r \times n$. Pour chaque échantillon de $r \times n$ observations, α et σ sont fixes

$$\alpha \in \{0.2; 0.4; 0.6; 0.8; 1; 1.2; 1.4; 1.6; 1.8\}$$

et

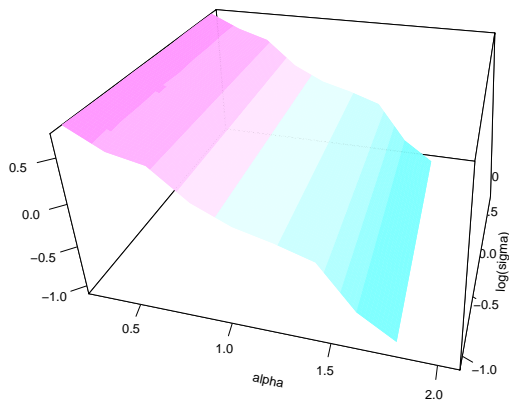
$$\log(\sigma) \in \{-1; -0.75; -0.5; -0.25; 0; 0.25; 0.5; 0.75; 1\}$$

(voir Figure 2.1). Les Tables 2.1-2.2 et les Figures 2.1-2.2 montre que l'estimateur des log-moments a de meilleures performances que celui de Koutrouvelis lorsque $\alpha < 1$, alors que c'est le contraire pour $\alpha > 1$. Les différences de performance accroissent pour des valeurs extrêmes de α .



(a) $n=100$

(b) $n=500$



(c) $n=2000$

FIGURE 2.1 – Log du rapport entre l'erreur quadratique moyenne pour le paramètre α pour l'estimateur de Koutrouvelis et celui des log-moments avec $r = 1000$ et $n = 100$ dans la Figure (a), $n = 500$ dans la Figure (b), $n = 2000$ dans la Figure (c).

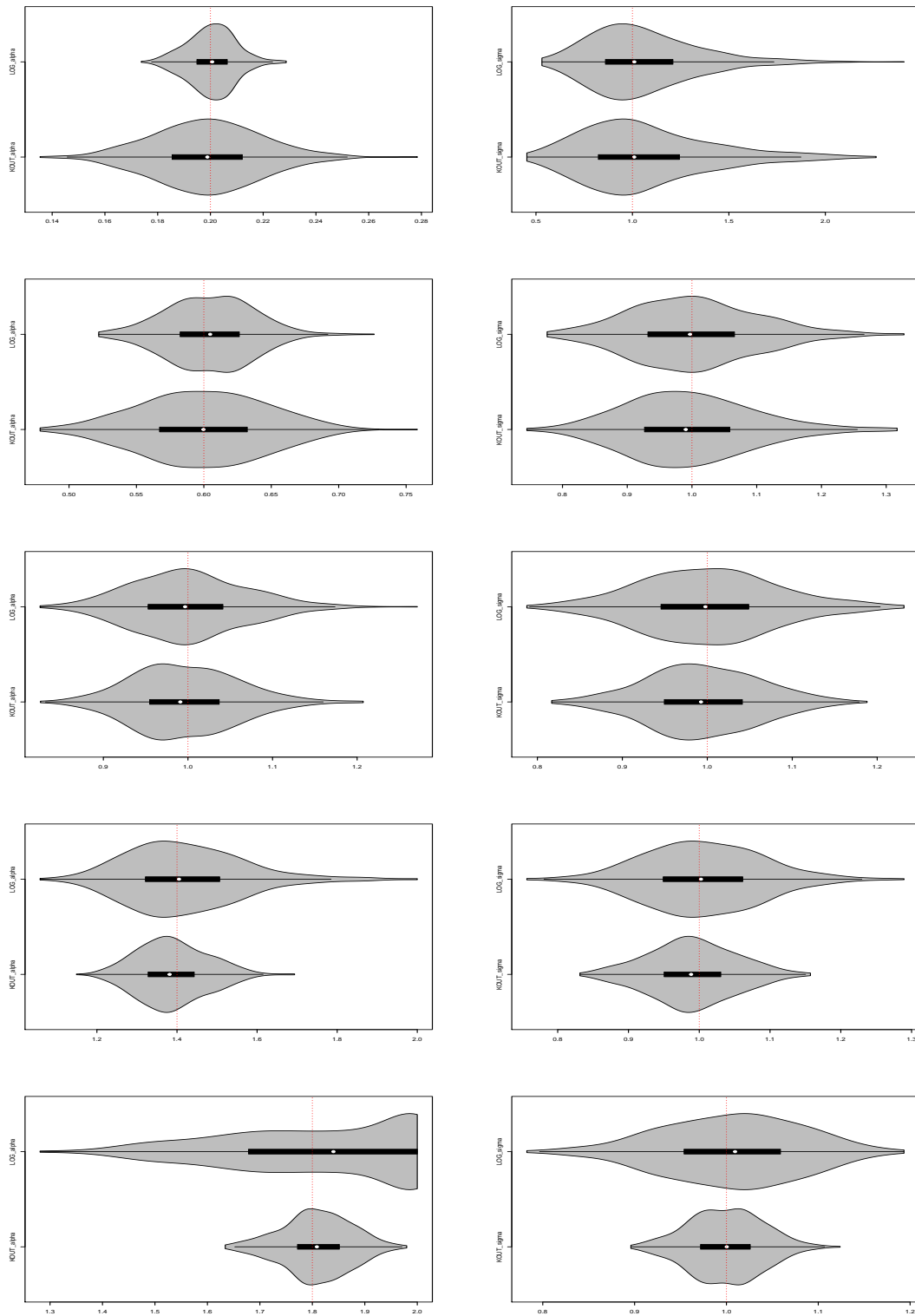


FIGURE 2.2 – Densités empiriques des estimateurs des log-moments (LOG) et de Koutrouvelis (KOUT) pour α (1ère colonne) et σ (2ème colonne) avec $r = 500$ et un échantillon de taille $n = 500$. Ici, on fixe $\sigma = 1$ et pour chaque ligne α change en prenant ses valeurs dans $\{0.2; 0.6; 1; 1.4; 1.8\}$.

2.2.3 Estimateur combiné

Comme nous l'avons vu précédemment, l'estimateur de régression de Koutrouvelis est préféré lorsque $\alpha > 1$ tandis que le log-estimateur devient significativement meilleur que celui de Koutrouvelis lorsque $\alpha (\leq 1)$ diminue. Nous voulons construire un estimateur de α qui sera au moins aussi bon que le meilleur estimateur, pour chaque α , pour de petits échantillons ($n \approx 100$). Pour ce faire, nous utilisons un estimateur combiné dont la procédure générale de construction est décrite dans [LR16]. Dans notre cas particulier, $\theta = {}^t(\alpha, \sigma)$ sont les paramètres à estimer et nous avons accès aux estimateurs $\hat{\alpha}_{KOUT}$, $\hat{\alpha}_{LOG}$ et $\hat{\sigma}_{KOUT}$ ¹.

On considère des estimateurs moyennés de θ de la forme :

$$\hat{\theta}_\lambda = {}^t\lambda \begin{pmatrix} \hat{\alpha}_{KOUT} \\ \hat{\alpha}_{LOG} \\ \hat{\sigma}_{KOUT} \end{pmatrix}, \quad \lambda \in \Lambda, \quad (2.2.14)$$

où ${}^t\lambda$ désigne le vecteur transposé de λ et $\Lambda \subset \mathcal{M}_{3 \times 2}(\mathbb{R})$.

Un moyen pratique de mesurer la performance de $\hat{\theta}_\lambda$ consiste à la comparer à celle de $\hat{\theta}^*$, qui est défini comme la meilleure combinaison linéaire $\hat{\theta}_\lambda$ obtenue pour un vecteur non aléatoire $\lambda \in \Lambda$. Plus précisément, $\hat{\theta}^*$ est la combinaison

linéaire ${}^t\lambda^* \begin{pmatrix} \hat{\alpha}_{KOUT} \\ \hat{\alpha}_{LOG} \\ \hat{\sigma}_{KOUT} \end{pmatrix}$ minimisant l'erreur quadratique moyenne (MSE), i.e.

$$\lambda^* = \operatorname{argmin}_{\lambda \in \Lambda} \mathbb{E}[\|\hat{\theta}_\lambda - \theta\|^2].$$

De toute évidence, plus l'ensemble Λ est grand et mieux ce sera. Cependant, choisir l'ensemble de l'espace $\Lambda = \mathcal{M}_{3 \times 2}(\mathbb{R})$ n'est en général pas exploitable. Nous devons imposer certaines conditions sur l'ensemble Λ afin d'avoir un formule explicite pour λ^* .

On définit

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

et on considère l'espace de contraintes

$$\Lambda_{\max} = \{\lambda \in \mathcal{M}_{3 \times 2}(\mathbb{R}) / {}^t\lambda J = I_2\} = \left\{ \begin{pmatrix} a & b \\ 1-a & -b \\ 0 & 1 \end{pmatrix} / (a, b) \in \mathbb{R}^2 \right\}$$

1. Comme $\hat{\sigma}_{KOUT}$ et $\hat{\sigma}_{LOG}$ ont des performances similaires, nous n'utilisons qu'un seul de ces estimateurs dans la combinaison.

où I_2 est la matrice identité. L'erreur quadratique moyenne $\mathbb{E}[\|\hat{\theta}_\lambda - \theta\|^2]$ est minimisée sur l'espace Λ_{\max} pour une unique valeur :

$$\lambda^* = \Sigma^{-1} J ({}^t J \Sigma^{-1} J)^{-1}, \quad (2.2.15)$$

où Σ est la matrice de Gram :

$$\Sigma = \mathbb{E} \left(\begin{pmatrix} \hat{\alpha}_{KOUT-\alpha} \\ \hat{\alpha}_{LOG-\alpha} \\ \hat{\sigma}_{KOUT-\sigma} \end{pmatrix} \begin{pmatrix} \hat{\alpha}_{KOUT-\alpha} \\ \hat{\alpha}_{LOG-\alpha} \\ {}^t \hat{\sigma}_{KOUT-\sigma} \end{pmatrix} \right). \quad (2.2.16)$$

Comme la matrice Σ est inconnu, l'estimateur moyenné $\hat{\theta}_{\max}$ est obtenu en remplaçant Σ par une estimation $\hat{\Sigma}$:

$$\begin{cases} \hat{\lambda}_{\max} = \hat{\Sigma}^{-1} J ({}^t J \hat{\Sigma}^{-1} J)^{-1}, \\ \hat{\theta}_{\max} = \hat{\lambda}_{\max}^\top \begin{pmatrix} \hat{\alpha}_{KOUT} \\ \hat{\alpha}_{LOG} \\ \hat{\sigma}_{KOUT} \end{pmatrix}. \end{cases} \quad (2.2.17)$$

Un moyen d'estimer Σ est d'utiliser une méthode de Monte Carlo. On commence par calculer une première estimation des paramètres :

$$\begin{cases} \hat{\alpha}_0 = \frac{\hat{\alpha}_{KOUT} + \hat{\alpha}_{LOG}}{2}, \\ \hat{\sigma}_0 = \hat{\sigma}_{KOUT}. \end{cases} \quad (2.2.18)$$

Puis, on simule B échantillons de taille n d'une loi stable symétrique de paramètres $\hat{\alpha}_0$ et $\hat{\sigma}_0$. Les trois estimateurs sont ensuite calculés pour chacun des B échantillons ce qui nous donne $\hat{\alpha}_{KOUT}^{(b)}$, $\hat{\alpha}_{LOG}^{(b)}$ et $\hat{\sigma}_{KOUT}^{(b)}$ pour $b = 1, \dots, B$, et chaque coefficient de la matrice est estimée par sa valeur empirique.

Les expériences montrent que les erreurs entraînées par l'estimation de Σ sont négligeables par rapport à l'avantage d'avoir plusieurs estimateurs pour de petits échantillons. On peut remarquer que des estimateurs similaires pourraient être construits en combinant plus de 2 estimateurs pour α . Par exemple, il serait possible d'ajouter l'estimateur des quantiles de Mc Culloch. Cela augmenterait la taille de la matrice de covariance dont l'estimation serait moins bonne et entraînerait un risque de construire un estimateur combiné qui ne serait jamais meilleur que chacun des estimateurs pris individuellement.

Le poids moyen du log-estimateur dans la combinaison est représenté dans la

Figure 2.3. Dans la Table 2.3, nous donnons les erreurs quadratiques moyennes pour différentes valeurs de α . Pour chaque valeur, on peut remarquer que l'estimateur combiné à partir des estimateurs de Koutrouvelis et des log-moments est toujours meilleur que chaque estimateur pris séparément. Ceci est confirmé par les graphiques de la Figure 2.4 comparant les lois empiriques de chaque estimateur.

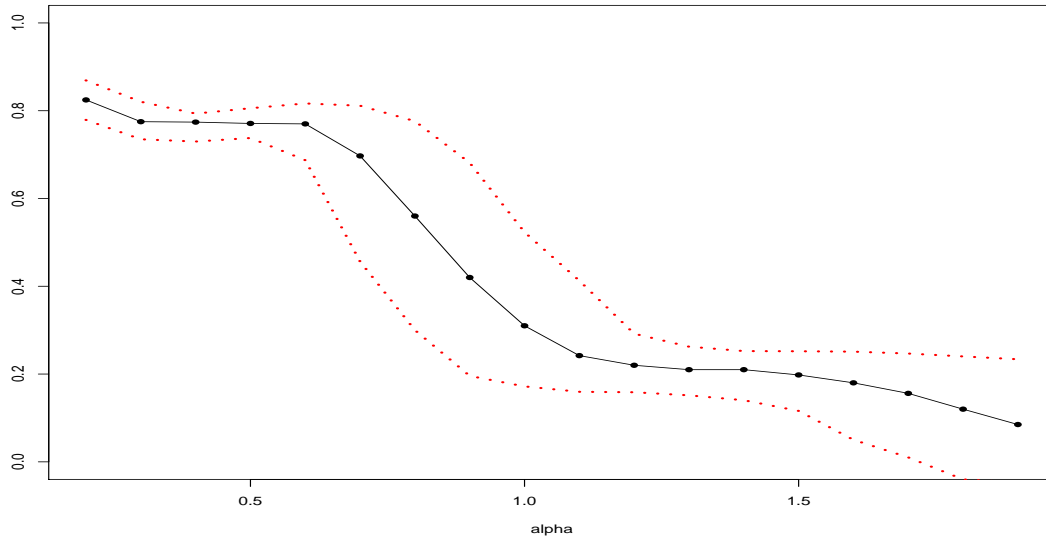


FIGURE 2.3 – Poids moyen de l'estimateur des log-moments $\hat{\alpha}_{LOG}$ dans l'estimateur combiné en fonction de α . La borne supérieure (resp. inférieure) de l'intervalle correspond au quantile à 95% (resp. 5%) pour $r = 500$ itérations de la combinaison avec $n = 100$ et $B = 1000$.

α	0.2	0.3	0.4	0.5	0.6	0.7
KOUT	$2.4 \cdot 10^{-3}$	$3.1 \cdot 10^{-3}$	$5.0 \cdot 10^{-3}$	$8.3 \cdot 10^{-3}$	$1.3 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$
LOG	$5.0 \cdot 10^{-4}$	$9.1 \cdot 10^{-4}$	$1.9 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$5.6 \cdot 10^{-3}$	$7.6 \cdot 10^{-3}$
COMB	$4.0 \cdot 10^{-6}$	$7.2 \cdot 10^{-4}$	$1.5 \cdot 10^{-3}$	$2.9 \cdot 10^{-3}$	$5.2 \cdot 10^{-3}$	$6.8 \cdot 10^{-3}$
α	0.8	0.9	1	1.1	1.2	1.3
KOUT	$1.4 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$2.8 \cdot 10^{-2}$	$3.4 \cdot 10^{-2}$
LOG	$1.2 \cdot 10^{-3}$	$2.0 \cdot 10^{-2}$	$3.3 \cdot 10^{-2}$	$4.6 \cdot 10^{-2}$	$5.9 \cdot 10^{-2}$	$7.8 \cdot 10^{-2}$
COMB	$8.3 \cdot 10^{-3}$	$1.1 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$	$3.1 \cdot 10^{-2}$
α	1.4	1.5	1.6	1.7	1.8	1.9
KOUT	$4.0 \cdot 10^{-2}$	$4.1 \cdot 10^{-2}$	$4.1 \cdot 10^{-2}$	$2.8 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$
LOG	$8.5 \cdot 10^{-2}$	$8.8 \cdot 10^{-2}$	$9.9 \cdot 10^{-2}$	$8.7 \cdot 10^{-2}$	$7.4 \cdot 10^{-2}$	$7.6 \cdot 10^{-2}$
COMB	$3.3 \cdot 10^{-2}$	$3.4 \cdot 10^{-2}$	$3.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$

TABLE 2.3 – Erreurs quadratiques moyennes pour les estimateurs de Koutrouvelis (KOUT), des Log-moments (LOG) et combiné (COMB) pour α avec $r = 500$, $n = 100$, $B = 1000$ et $\sigma = 1$.

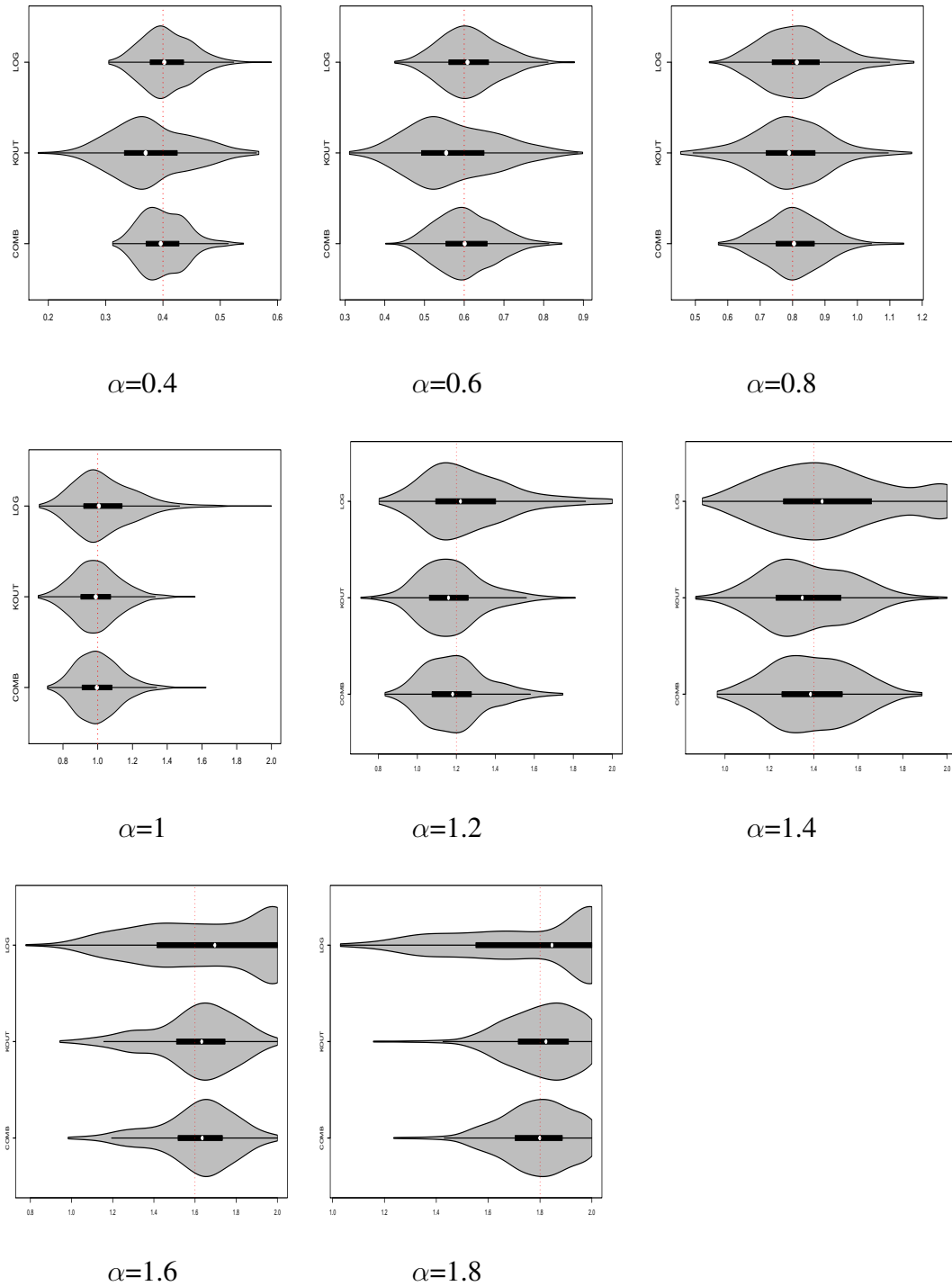


FIGURE 2.4 – Densités empiriques des estimateurs des log-moments (LOG), de Koutrouvelis (KOUT) et combiné (COMB) pour le paramètre α avec $r = 500$, $n = 100$, $B = 1000$ et $\sigma = 1$.

2.3 Lois stables asymétriques

2.3.1 Adaptation des estimateurs pour le cas asymétrique

On se place désormais dans le cas où X suit une loi $S_\alpha(\sigma, \beta, 0)$, on a alors

$$\begin{aligned}\mathbb{E}[\log |X|] &= \left(\frac{1}{\alpha} - 1\right) \gamma + \log \sigma - \frac{\log |\cos \theta|}{\alpha}, \\ \mathbb{E}[(\log |X| - \mathbb{E}[\log |X|])^2] &= \text{Var}(\log |X|) = \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12} - \frac{\theta^2}{\alpha^2},\end{aligned}$$

où γ est la constante d'Euler et $\theta = \arctan\left(\beta \tan \frac{\alpha\pi}{2}\right)$. (voir [Kur01] Proposition 4).

Soit (X_1, \dots, X_{2n}) une suite de $2n$ variables aléatoires i.i.d. de loi stable $S_\alpha(\sigma, \beta, 0)$. Nous utilisons la centro-symétrisation introduite dans [Kur01] aux données observées pour obtenir les n variables aléatoires i.i.d. symétriques de lois stables $S_\alpha(2\sigma, 0, 0)$:

$$(X_{2k} - X_{2k-1})_{k \in \{1, \dots, n\}}.$$

Nous estimons ensuite α à partir de $\hat{\alpha}_{LOG}^{(n)}(X_2 - X_1, \dots, X_{2n} - X_{2n-1})$, où $\hat{\alpha}_{LOG}^{(n)}$ est introduit dans le 2.2.4. On peut remarquer qu'il est possible d'estimer β en utilisant l'expression de $\text{Var}[\log |X|]$.

Une autre façon d'estimer α est d'utiliser les $(2n - 1)$ variables aléatoires $(X_k - X_{k-1})_{k \in \{2, \dots, 2n\}}$. Dans ce cas, nous conservons la même taille d'échantillon mais nous perdons l'indépendance des variables.

Les erreurs quadratiques moyennes de la première méthode ne dépendent pas de β puisque les variables aléatoires $(X_{2m} - X_{2m-1})$ sont symétriques et indépendantes. Numériquement, nous observons que les propriétés de la deuxième estimation ne dépendent pas non plus de β . De plus, les résultats numériques donnés dans la Table 2.4 indiquent que les deux estimations ont une performance similaire pour $\alpha \leq 1$.

$2n$		$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1$
100	n i.i.d	$1.01 \cdot 10^{-3}$	$4.44 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$	$3.30 \cdot 10^{-2}$	$10.7 \cdot 10^{-2}$
	$2n - 1$ dep.	$0.970 \cdot 10^{-3}$	$4.14 \cdot 10^{-3}$	$1.08 \cdot 10^{-2}$	$2.44 \cdot 10^{-2}$	$6.08 \cdot 10^{-2}$
500	n i.i.d	$1.80 \cdot 10^{-4}$	$7.90 \cdot 10^{-4}$	$2.10 \cdot 10^{-3}$	$4.68 \cdot 10^{-3}$	$1.01 \cdot 10^{-2}$
	$2n - 1$ dep.	$1.75 \cdot 10^{-4}$	$7.58 \cdot 10^{-4}$	$1.86 \cdot 10^{-3}$	$3.81 \cdot 10^{-3}$	$7.33 \cdot 10^{-3}$
1000	n i.i.d	$9.12 \cdot 10^{-5}$	$3.86 \cdot 10^{-4}$	$9.90 \cdot 10^{-4}$	$2.23 \cdot 10^{-3}$	$4.79 \cdot 10^{-3}$
	$2n - 1$ dep.	$8.94 \cdot 10^{-5}$	$3.69 \cdot 10^{-4}$	$8.88 \cdot 10^{-4}$	$1.84 \cdot 10^{-3}$	$3.48 \cdot 10^{-3}$

 TABLE 2.4 – Erreurs quadratiques moyennes pour l'estimation de α avec les log-moments en partant de $2n$ variables aléatoires i.i.d. $S_\alpha(1, \beta, 0)$.

La méthode de Koutrouvelis ne varie pas avec des lois asymétriques car le module de la fonction caractéristique ne dépend que de α et σ (voir performances dans les Tables 2.1 et 2.3). Nous combinons l'estimateur des log-moments après symétrie (nommé par la suite LOG sym.) avec l'estimateur de Koutrouvelis de la même manière qu'avec les variables symétriques (voir Section 2.2.3) pour obtenir un nouvel estimateur dont les performances numériques sont rapportées dans la Table 2.5.

α	0.2	0.4	0.6	0.8	1
COMB	$2.39 \cdot 10^{-3}$	$2.21 \cdot 10^{-3}$	$5.88 \cdot 10^{-3}$	$9.08 \cdot 10^{-3}$	$1.67 \cdot 10^{-2}$
KOUT	$6.14 \cdot 10^{-3}$	$4.92 \cdot 10^{-3}$	$1.10 \cdot 10^{-2}$	$1.36 \cdot 10^{-2}$	$1.72 \cdot 10^{-2}$
LOG sym.	$1.90 \cdot 10^{-3}$	$4.36 \cdot 10^{-3}$	$1.34 \cdot 10^{-2}$	$2.67 \cdot 10^{-2}$	$7.29 \cdot 10^{-2}$
α	1.2	1.4	1.6	1.8	
COMB	$2.70 \cdot 10^{-2}$	$3.53 \cdot 10^{-2}$	$3.54 \cdot 10^{-2}$	$2.20 \cdot 10^{-2}$	
KOUT	$2.71 \cdot 10^{-2}$	$4.15 \cdot 10^{-2}$	$4.16 \cdot 10^{-2}$	$2.41 \cdot 10^{-2}$	
LOG sym.	$1.24 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	

 TABLE 2.5 – Erreurs quadratiques moyennes pour l'estimation de α avec l'estimateur combiné (COMB), l'estimateur de Koutrouvelis (KOUT) et des log-moments après symétrisation (LOG sym.). Ici, on a choisi $r = 500$, $n = 100$, $B = 1000$, $\beta = 0.6$ et $\sigma = 1$.

Pour les données asymétriques (voir le tableau 2.5), les estimateurs combinés ont encore de bonnes performances, mais nous perdons en termes d'erreurs quadratiques moyennes comparé au cas symétrique.

2.3.2 Test de symétrie

Sous l'hypothèse d'observer des lois stables i.i.d, nous proposons un test pour vérifier l'asymétrie des données. Nous voulons tester $H_0 : \beta = 0$ contre $H_1 : \beta \neq 0$ grâce aux propriétés des estimateurs étudiés dans la section précédente.

Soit (X_1, \dots, X_{2n}) une suite de $2n$ variables aléatoires i.i.d. de loi stable $S_\alpha(\sigma, \beta, 0)$. Sous l'hypothèse nulle H_0 , les estimateurs $\widehat{\alpha}_{LOG}((X_{2k})_{k \in [1, n]})$ et $\widehat{\alpha}_{LOG}((X_{2k} - X_{2k-1})_{k \in [1, n]})$ sont tous les deux consistants, et donc la différence entre ces estimateurs tend vers zéro. Pour des variables asymétriques, cette convergence ne se produit pas puisque $\widehat{\alpha}_{LOG}((X_{2k})_{k \in [1, n]})$ ne converge plus vers α . Ces faits suggèrent de construire un test basé sur la différence de ces estimateurs. Posons

$$\begin{aligned} L_1 &:= E[\log |Z|] = \left(\frac{1}{\alpha} - 1\right) \gamma + \log \sigma, \\ L_2 &:= E[(\log |Z| - E[\log |Z|])^2] = \frac{\pi^2}{6\alpha^2} + \frac{\pi^2}{12}, \\ L_3 &:= E[(\log |Z| - E[\log |Z|])^3] = 2\zeta(3) \left(\frac{1}{\alpha^3} - 1\right), \\ L_4 &:= E[(\log |Z| - E[\log |Z|])^4] = \pi^4 \left(\frac{3}{20\alpha^4} + \frac{1}{12\alpha^2} + \frac{19}{240}\right), \\ C &:= Cov((\log |X_2| - E[\log |X_2|])^2, (\log |X_2 - X_1| - E[\log |X_2 - X_1|])^2) \\ &= E[(\log |X_2| - E[\log |X_2|])^2 (\log |X_2 - X_1| - E[\log |X_2 - X_1|])^2] - L_2^2. \end{aligned}$$

où Z est une variable aléatoire de loi $S_\alpha(\sigma, 0, 0)$ et ζ désigne la fonction zêta de Riemann $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ et $\zeta(3) = 1.2020569 \dots$

Proposition 2.3.1

Pour $w \in (0, 1)$, on définit la région critique

$$R_w = \left\{ n \frac{\left(\widehat{\alpha}_{LOG}(\{X_{2k}\}_{k \in [1, n]}) - \widehat{\alpha}_{LOG}(\{X_{2k} - X_{2k-1}\}_{k \in [1, n]})\right)^2}{\frac{18\widehat{\alpha}_{LOG}^6}{\pi^4} (\widehat{L}_4 - \widehat{L}_2^2 - \widehat{C})} > t_w \right\}$$

où t_w est le quantile d'ordre $1 - w$ d'une loi du χ^2 à 1 degré de liberté, et où \widehat{L}_4 , \widehat{L}_2 et \widehat{C} sont respectivement les moments empiriques de L_4 , L_2 et C .

On décide de rejeter l'hypothèse nulle si $(X_1, \dots, X_{2n}) \in R_w$. Ce test de région critique R_w est asymptotiquement de niveau w et est asymptotiquement consistant sous H_1 .

Démonstration : Sous l'hypothèse nulle, avec $Y_k = \log |X_{2k}|$ et $Z_k = \log |X_{2k} - X_{2k-1}|$

$$\sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n \begin{pmatrix} (Y_k - \bar{Y}_n)^2 \\ (Z_k - \bar{Z}_n)^2 \end{pmatrix} - \begin{pmatrix} L_2 \\ L_2 \end{pmatrix} \right) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} L_4 - L_2^2 & C \\ C & L_4 - L_2^2 \end{pmatrix} \right).$$

Et donc par la delta-méthode dans le cas multidimensionnel, on obtient

$$\sqrt{n} \left(\begin{pmatrix} \widehat{\alpha}_{LOG}((X_{2k})_{k \in [1, n]}) \\ \widehat{\alpha}_{LOG}((X_{2k} - X_{2k-1})_{k \in [1, n]}) \end{pmatrix} - \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \right) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{9\alpha^6}{\pi^4} \begin{pmatrix} L_4 - L_2^2 & C \\ C & L_4 - L_2^2 \end{pmatrix} \right),$$

et

$$\frac{\sqrt{n} \widehat{\alpha}_{LOG}((X_{2k})_{k \in [1, n]}) - \widehat{\alpha}_{LOG}((X_{2k} - X_{2k-1})_{k \in [1, n]})}{\sqrt{\frac{18\alpha^6}{\pi^4} (L_4 - L_2^2 - C)}} \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}(0, 1).$$

Finalement, en appliquant le lemme de Slutsky et en utilisant la consistance de \widehat{L}_4 , \widehat{L}_2 et \widehat{C} , on obtient que le niveau asymptotique vaut w .

Sous H_1 , on a :

$$\widehat{\alpha}_{LOG}((X_{2k})_{k \in [1, n]}) - \widehat{\alpha}_{LOG}((X_{2k} - X_{2k-1})_{k \in [1, n]}) \xrightarrow[n \rightarrow \infty]{p.s.} \frac{\pi\alpha}{\sqrt{\pi^2 - 6\theta^2}} - \alpha \neq 0,$$

avec $\theta = \arctan(\beta \tan \frac{\alpha\pi}{2})$. Alors, sous l'hypothèse alternative $\beta \neq 0$, on a un test consistant puisque $\mathbb{P}_{H_1}(R_w) \xrightarrow[n \rightarrow \infty]{} 1$. ■

Dans la Table 2.6, on utilise la méthode de Monte Carlo pour évaluer la probabilité de rejeter l'hypothèse nulle H_0 . Pour les petites valeurs de α , le niveau empirique converge lentement vers w : cela est dû à la forme de la densité qui est concentrée autour de l'origine et la faible qualité pour l'estimation du coefficient $L_4 - L_2^2 - C$ dans ce cas. L'estimation converge assez lentement vers ce coefficient (dont la valeur augmente lorsque α diminue).

	$\alpha = 0.2$	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8
$2n = 200$	0.065	0.074	0.084	0.087	0.088	0.097	0.088	0.06	0.042
$2n = 1000$	0.0953	0.1019	0.0971	0.89	0.074	0.065	0.067	0.061	0.035
$2n = 10^4$	0.18	0.16	0.10	0.071	0.051	0.050	0.048	0.051	0.053

TABLE 2.6 – Probabilités de rejeter l'hypothèse nulle sous H_0 pour différentes tailles d'échantillons et différentes valeurs de α . Le niveau asymptotique du test est fixé à $w = 5\%$.

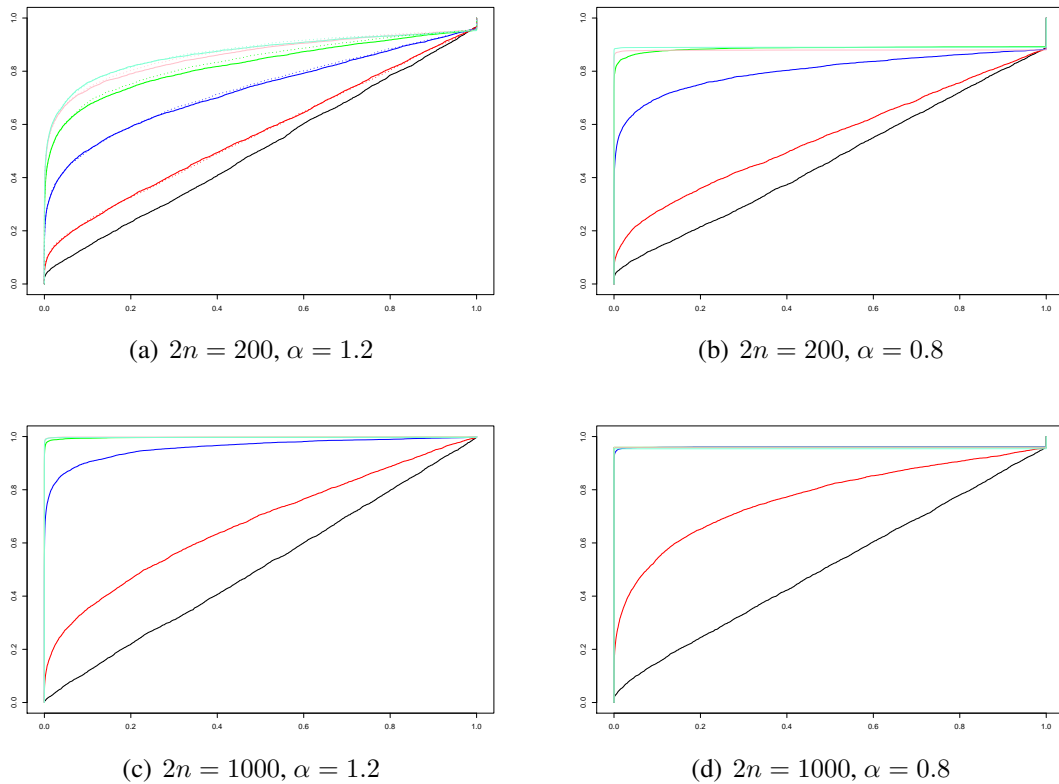


FIGURE 2.5 – Représentation des fonctions de répartition empiriques de la p -valeur du test de symétrie pour différentes valeurs de α , β et n : $\beta = 0$ (noir), $\beta = 0.2$ (rouge), $\beta = 0.4$ (bleu foncé), $\beta = 0.6$ (vert), $\beta = 0.8$ (rose) et $\beta = 1$ (bleu clair). Dans la Figure (a), on a ajouté en pointillés les fonctions de répartition empiriques pour $\beta \in \{-1; -0.8; -0.6; -0.4; -0.2\}$ qui correspondent exactement à celles pour les valeurs positives.

Dans la Figure 2.5, on peut voir que la puissance augmente lorsque β s'éloigne de 0 ou lorsque la taille de l'échantillon augmente. Cette convergence sous l'hypothèse alternative dépend des valeurs de α et β . On observe que la p -valeur prend la valeur 1 avec probabilité non-nulle pour les échantillons de petites tailles. Ce saut dû à la troncature dans l'estimateur des log-moments défini dans le Théorème 2.2.4 disparaît asymptotiquement.

2.4 Cas de variables stables non-identiquement distribuées

Dans les applications, il peut être nécessaire de savoir analyser des phénomènes non stationnaires. Par exemple, il semble plausible que les rendements financiers qui affichent des sauts verront l'intensité de ces sauts dépendre d'événements externes, comme les crises financières (voir la section suivante pour une illustration sur le S&P 500). Parfois, les variations pour α sont faibles, et il est intéressant d'examiner le comportement de notre estimateur dans les situations où les données disponibles s'écartent légèrement de l'hypothèse i.i.d. Dans la suite, nous examinons deux cas : des petites perturbations déterministes et aléatoires de α , conduisant à des variables aléatoires qui ne sont pas identiquement distribuées. Nous ne renonçons pas ici à l'hypothèse d'indépendance, bien qu'il s'agisse d'une extension souhaitable. Cette généralisation fera l'objet d'un travail futur.

2.4.1 Perturbations déterministes

Soit $(X_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires indépendantes et X une variable aléatoire indépendante des (X_i) telles que

$$X_i \sim S_{\alpha_i}(\sigma_i, 0, 0) \text{ et } X \sim S_{\alpha}(\sigma, 0, 0).$$

On pose $Y_i = \log |X_i|$ et $Y = \log |X|$. On suppose que, pour tout entier i , $\alpha_i = \alpha + \varepsilon_i$ et $\sigma_i = \sigma + \eta_i$ avec ε_i et η_i déterministes tels que

$$\frac{|\varepsilon_i|}{\alpha} \leq c_{\alpha} < 1 \text{ et } \frac{|\eta_i|}{\sigma} \leq c_{\sigma} < 1.$$

Proposition 2.4.1

Sous les conditions $\frac{1}{n} \sum_{i=1}^n |\varepsilon_i| \xrightarrow[n \rightarrow \infty]{} 0$ et $\frac{1}{n} \sum_{i=1}^n |\eta_i| \xrightarrow[n \rightarrow \infty]{} 0$, on a

$$\hat{\alpha}_{LOG}^{(n)} \xrightarrow[n \rightarrow \infty]{p.s.} \alpha,$$

où $\hat{\alpha}_{LOG}^{(n)}$ est défini dans le Théorème 2.2.4.

Démonstration : Voir l'annexe 2.6. ■

Proposition 2.4.2

Soit $\Sigma_{\alpha,\sigma}$ la matrice de covariance entre Y et Y^2 :

$$\Sigma_{\alpha,\sigma} = \begin{pmatrix} \text{Var}(Y) & \text{Cov}(Y, Y^2) \\ \text{Cov}(Y, Y^2) & \text{Var}(Y^2) \end{pmatrix}, \quad (2.4.1)$$

et posons

$$H_{\alpha,\sigma} = \begin{pmatrix} 6\left(\left(\frac{1}{\alpha}-1\right)\gamma+\log\sigma\right)\alpha^3/\pi^2 \\ -3\alpha^3/\pi^2 \end{pmatrix}. \quad (2.4.2)$$

Sous les conditions $\frac{1}{\sqrt{n}} \sum_{i=1}^n |\varepsilon_i| \xrightarrow[n \rightarrow \infty]{} 0$ et $\frac{1}{\sqrt{n}} \sum_{i=1}^n |\eta_i| \xrightarrow[n \rightarrow \infty]{} 0$, le théorème central limite suivant s'applique à $\hat{\alpha}_{LOG}^{(n)}$:

$$\sqrt{n} \left(\hat{\alpha}_{LOG}^{(n)} - \alpha \right) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}(0, {}^t H_{\alpha,\sigma} \Sigma_{\alpha,\sigma} H_{\alpha,\sigma}). \quad (2.4.3)$$

Démonstration : Voir l'annexe 2.6. ■

2.4.2 Perturbations aléatoires

Soit X une variable aléatoire de loi $S_\alpha(\sigma, 0, 0)$. Pour tout entier i , on pose $\alpha_i = \alpha + \varepsilon_i$ où ε_i est une variable aléatoire. On suppose que les perturbations ε_i vérifient la condition suivante : il existe une constante $c_\alpha \in]0, 1[$ telle que :

$$\mathbb{P} \left(\frac{\varepsilon_i}{\alpha} \in \left[-c_\alpha, \min\left(c_\alpha, \frac{2}{\alpha} - 1\right) \right] \right) = 1.$$

Cette condition assure que les perturbations sont petites relativement à la valeur de α et que pour tout i , α_i soit bien dans l'intervalle $]0, 2]$. Soit (X_i) une suite de variables indépendantes entre elles et de X telles que $X_i \sim S_{\alpha_i}(\sigma, 0, 0)$ (sachant α_i). On pose alors $Y = \log |X|$ et $Y_i = \log |X_i|$ pour $i \in \mathbb{N}$.

Proposition 2.4.3

Sous la condition $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[|\varepsilon_i|] \xrightarrow{n \rightarrow \infty} 0$, on a $\hat{\alpha}_{LOG}^{(n)} \xrightarrow[n \rightarrow \infty]{p.s.} \alpha$.

Si, de plus, $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E}[|\varepsilon_i|] \xrightarrow{n \rightarrow \infty} 0$, alors on a le théorème central limite suivant :

$$\sqrt{n} \left(\hat{\alpha}_{LOG}^{(n)} - \alpha \right) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}(0, H_{\alpha, \sigma} \Sigma_{\alpha, \sigma} {}^t H_{\alpha, \sigma})$$

où $\Sigma_{\alpha, \sigma}$ et $H_{\alpha, \sigma}$ sont définis dans (2.4.1) et (2.4.2).

Démonstration : Voir l'annexe 2.6. ■

2.5 Quelques applications de l'estimateur combiné

2.5.1 Résultats numériques sur des données synthétisées : mouvement de Lévy multistable

Nous utilisons maintenant nos estimateurs log-moments et combinés dans le cas du mouvement de Lévy multistable défini dans [FLV09]. (voir aussi [LGLV12] pour d'autres propriétés de ce processus). L'idée de base est de permettre à l'indice de stabilité d'évoluer avec le temps, de sorte que l'intensité des sauts, qui est régie par α , varie selon une trajectoire. Une telle caractéristique est couramment rencontrée dans les séries temporelles observées dans des domaines tels que la finance ou la biomédecine. Rappelons brièvement la définition de ces processus.

Soit $\alpha : [0, 1] \rightarrow]0, 2[$ une fonction de classe \mathcal{C}^1 . On note $r^{<s>} = \text{sign}(r)|r|^s$ pour $r \in \mathbb{R}$ et $s \in \mathbb{R}$. Le mouvement de Lévy multistable est défini par

$$M_\alpha(t) = C_{\alpha(t)} \sum_{(X,Y) \in \Pi} \mathbb{1}_{]0,t]}(X) Y^{<-1/\alpha(t)>} \tag{2.5.1}$$

où $C_\theta = \left(\int_0^\infty u^{-\theta} \sin(u) du \right)^{-1/\theta}$ et Π est un processus de Poisson sur $\mathbb{R}^+ \times \mathbb{R}$ ayant pour mesure moyenne le mesure de Lebesgue dans le plan \mathcal{L}^2 . Ce processus est

simulé en utilisant le champ

$$X(t, u) = C_{\alpha(u)} \sum_{(X, Y) \in \Pi} \mathbb{1}_{(0, t]}(X) Y^{<-1/\alpha(u)>}.$$

Pour chaque $u \in]0, 1[$, $X(\cdot, u)$ est un processus $\alpha(u)$ -stable à accroissements indépendants qui peut être simulé à l'aide du programme RSTAB disponible dans [ST04] ou dans [ST94]. L'intervalle $[0, 1]$ est discrétisé en N sous-intervalles de même longueur et $X(\cdot, u)$ est calculé aux points $\left(\frac{k}{N}\right)_{k \in \{1, \dots, N\}}$ à partir de la somme cumulée de N variables stables indépendantes d'exposant caractéristique $\alpha(u)$.

Dans la Figure 2.6, nous affichons des exemples de trajectoires de processus multistables pour différentes fonctions α . Ensuite, nous estimons ces fonctions en tout point t_0 grâce à l'estimateur combiné appliqué sur une fenêtre de n observations autour de t_0 .

Dans la Figure 2.7, on itère 100 fois la simulation et l'estimation pour un processus multistable ayant pour fonction d'intensité $\alpha(t) = 1.5 - 0.48 \sin(2\pi(t + 1/4))$. Pour chaque point où la fonction α est estimée, on obtient la loi empirique de l'estimateur combiné. Cette procédure est répétée pour plusieurs tailles de fenêtre (100, 200, 1000 et 2000). Nous observons que l'erreur type qui correspond à l'écart-type de l'estimateur combiné diminue lorsque la taille de la fenêtre n augmente alors que le biais augmente pour n grand. Enfin, l'erreur quadratique moyenne diminue lorsque n augmente jusqu'à $n = 1000$ et augmente ensuite pour une valeur supérieure. La figure 2.8 représente le biais, l'erreur-type et l'erreur quadratique moyenne en fonction de t pour diverses valeurs de la taille de la fenêtre.

L'erreur quadratique moyenne en fonction de α est représentée dans la Figure 2.9. L'erreur quadratique moyenne ne varie pas beaucoup selon la valeur de $\alpha \in [1, 2]$ lorsque n est fixé.

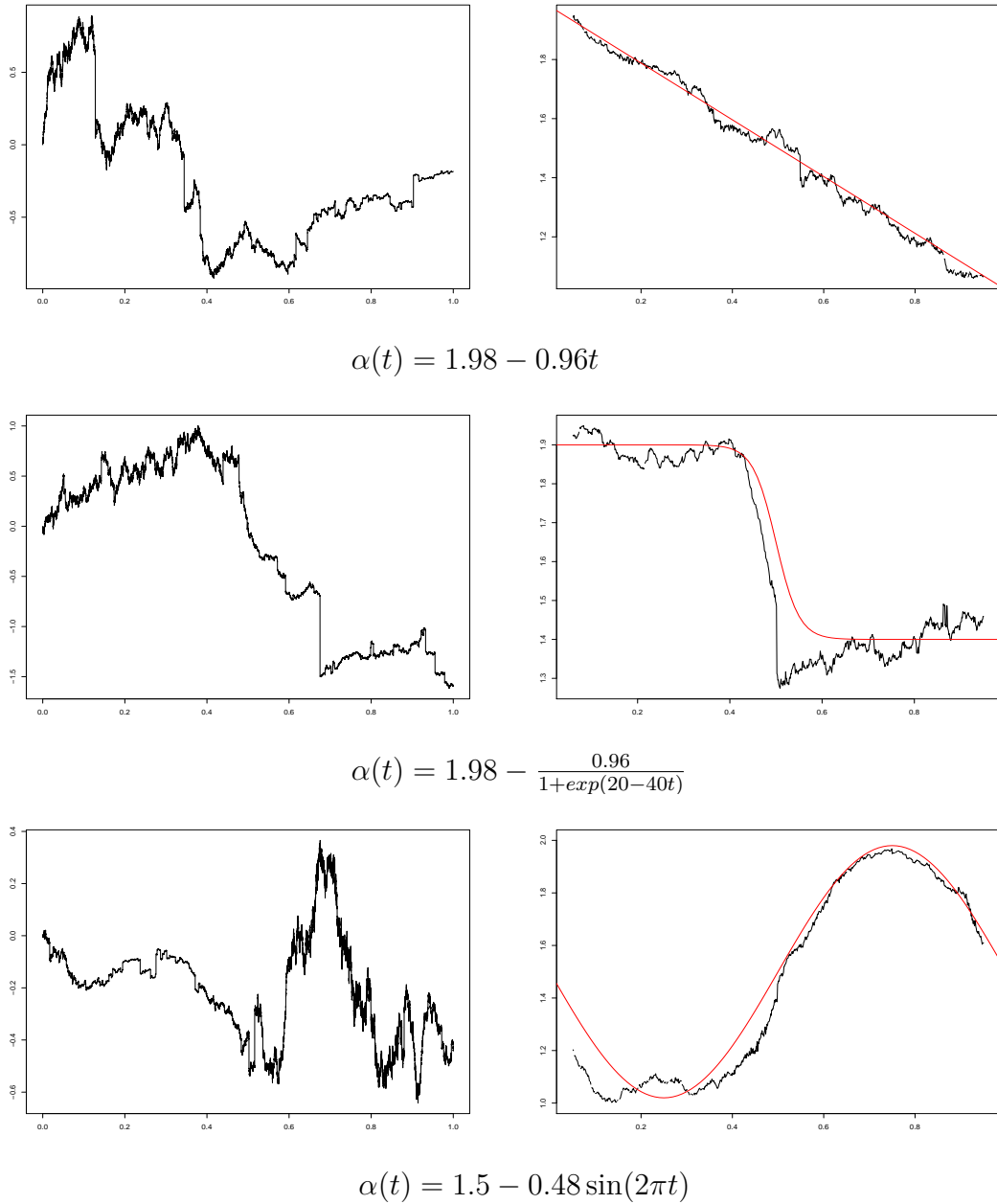
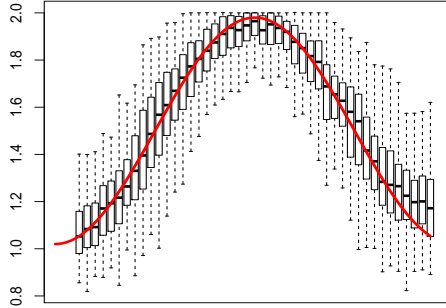
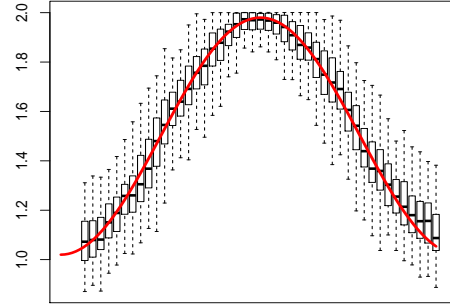


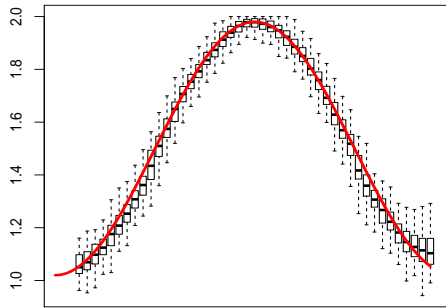
FIGURE 2.6 – Trajectoires de mouvements de Lévy multistables sur $(0,1)$ avec $N = 20000$ points dans la première colonne. Les fonctions $\alpha(t)$ (rouge) et $\hat{\alpha}_{COMB}(t)$ (noir) sont représentés dans la deuxième colonne avec $n = 2000$.



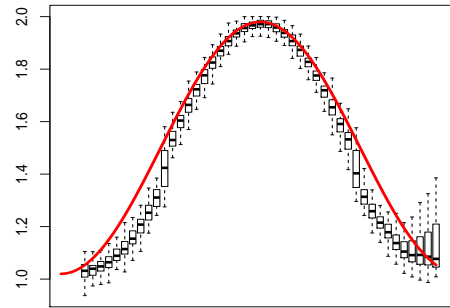
(a) $n = 100$



(b) $n = 200$

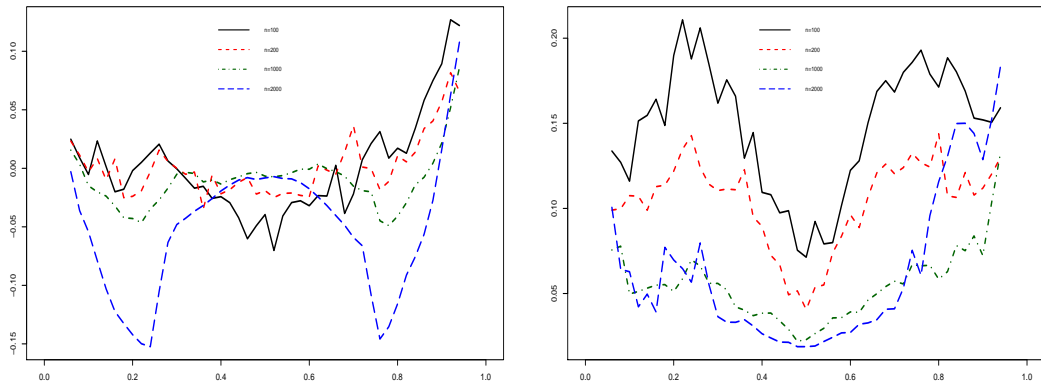


(c) $n = 1000$



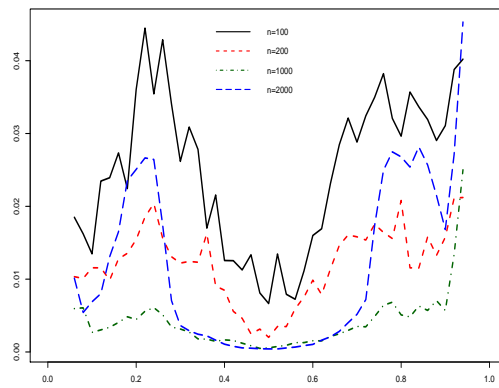
(d) $n = 2000$

FIGURE 2.7 – Boîtes à moustaches pour l'estimateur combiné $\hat{\alpha}_{COMB}$ appliqué à 100 itérations d'un mouvement de Lévy multistable d'exposant caractéristique $\alpha(t) = 1.5 - 0.48 \sin(2\pi(t + 1/4))$ (fonction tracée en rouge). Les boîtes à moustaches représentent le comportement de l'estimateur pour différentes tailles n de fenêtres.



(a)

(b)



(c)

FIGURE 2.8 – Représentation du biais (a), de l’erreur standard (b) et de l’erreur quadratique moyenne (c) en fonction du temps pour l’estimateur combiné $\hat{\alpha}_{COMB}$ pour $n = 100$ (ligne noire), $n = 200$ (ligne discontinue rouge), $n = 1000$ (mélange pointillé et trait vert) et $n = 2000$ (ligne discontinue bleue). Les statistiques sont évaluées sur les mêmes trajectoires que dans la Figure 2.7.

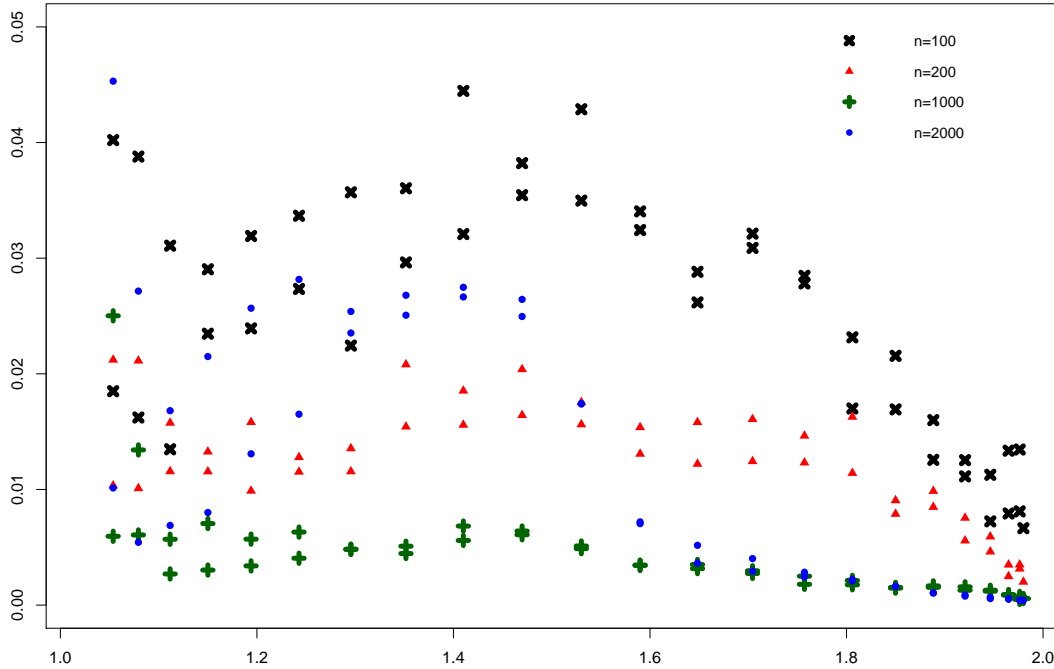


FIGURE 2.9 – Erreur quadratique moyenne pour l'estimateur combiné $\hat{\alpha}_{COMB}$ appliqué à un mouvement de Lévy multistable avec $\alpha(t) = 1.5 - 0.48 \sin(2\pi(t + 1/4))$ par rapport à la valeur de α . Les statistiques sont évaluées sur les mêmes trajectoires que dans la Figure 2.7.

Comme le montrent ces figures, des estimations raisonnables sont obtenues sur ces données simulées du fait que les variations de α sont plutôt lentes par rapport à la fréquence d'échantillonnage : cette caractéristique assure que centrer une fenêtre autour d'un t_0 donné et traiter tous les points dans cette fenêtre comme ayant la même valeur α est une approximation acceptable en ce qui concerne les estimations.

2.5.2 Application en finance

Cette dernière section traite des données réelles. Nous voulons appliquer l'estimateur combiné pour estimer l'exposant caractéristique de l'indice financier Standard & Poor's 500 (abrégé en *S&P 500*, voir Figure 2.10). Il s'agit d'un indice boursier américain basé sur les 500 plus grosses sociétés cotées en bourse. Les

rendements boursiers de l'indice financier *S&P 500*, qui correspondent au taux de croissance renormalisé $((Y_{t+1} - Y_t)/Y_t)_t$, sont supposés être des variables aléatoires indépendantes stables.

Dans la figure 2.11, nous testons d'abord la symétrie des données pour une fenêtre glissante de taille 1000 en utilisant le test défini dans la proposition 2.3.1. Nous représentons la fonction de distribution empirique de la valeur p des rendements *S&P 500* depuis 1996. La fonction de distribution cumulative est très proche de la fonction uniforme. Ainsi, on décide de ne pas rejeter l'hypothèse nulle de données symétriques pour les rendements du *S&P 500* depuis 1996. En conséquence, le paramètre d'intensité des sauts est estimé en appliquant l'estimateur défini dans le théorème 2.2.4 à l'aide d'une fenêtre glissante de différentes tailles pendant la période 1996 – 2017 (voir Figure 2.12).

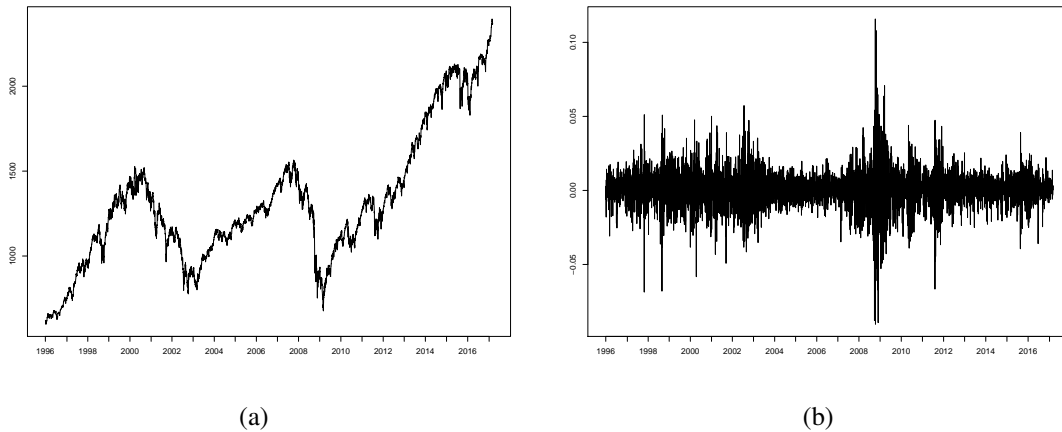


FIGURE 2.10 – Évolution de l'indice financier *S&P 500* (a) et de ses rendements (b) au cours du temps entre 1996 et 2017.

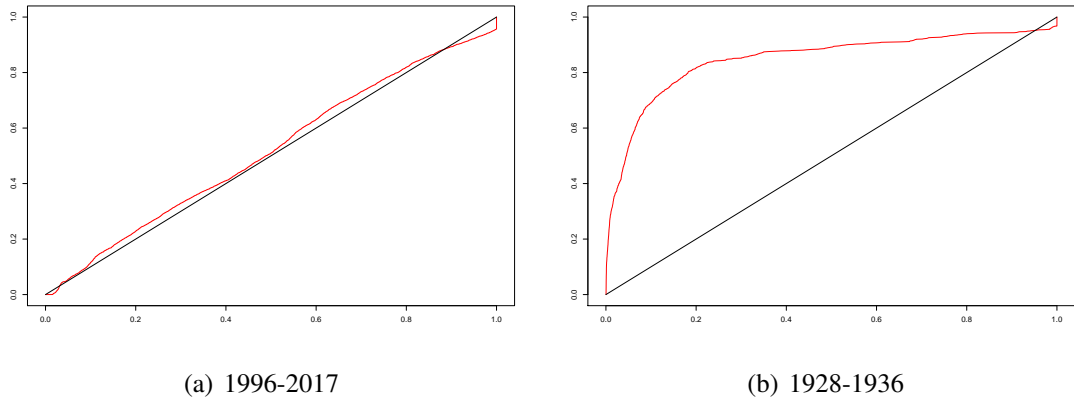
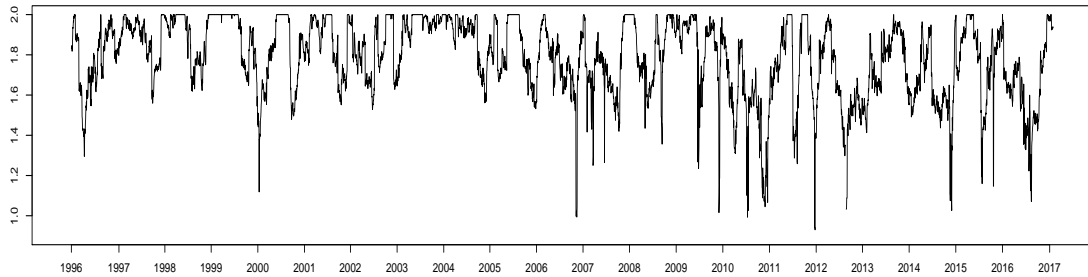
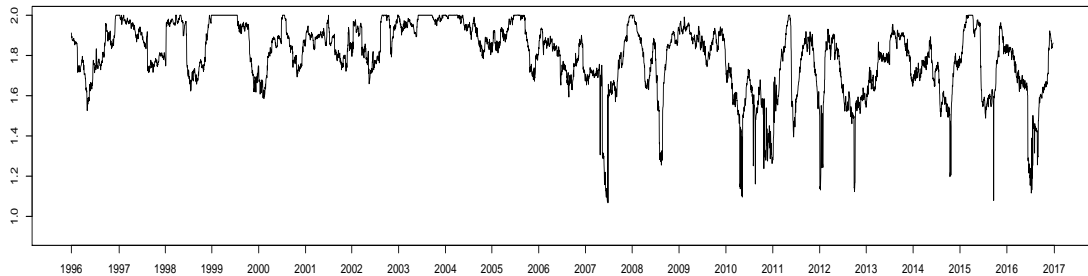


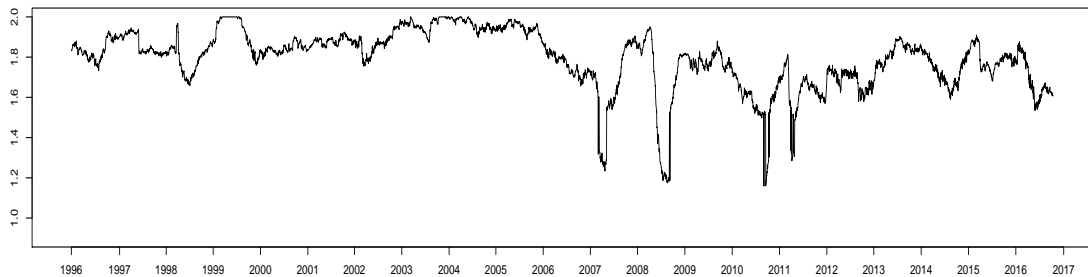
FIGURE 2.11 – Représentation de la fonction de répartition empirique de la p -valeur du test de symétrie pour les rendements de *S&P 500* entre 1996 et 2017 (Figure (a)) et entre 1928 et 1936 (Figure (b)) à l'aide d'une fenêtre glissante de taille 1000.



(a) $n=50$



(b) $n=100$



(c) $n=200$

FIGURE 2.12 – Valeurs de l'estimateur combiné $\hat{\alpha}_{COMB}$ de l'indice *S&P 500* pour une fenêtre glissante de taille n pour les jours ouvrés entre 1996 et 2017.

Nous rejetons l'hypothèse de symétrie entre 1929 et 1936 (voir Figure 2.11). Dans la Figure 2.13, l'estimation de l'exposant caractéristique se fait entre 1929 et 1936 en utilisant l'estimateur combiné asymétrique défini dans la section 2.3. Une

chute brutale est observée à la fin de l'année 1929. Ce changement correspond au krach boursier de Wall Street. L'estimation des données symétriques (en rouge) est ajoutée dans la figure pour voir la différence entre les deux estimations, en particulier pendant la crise.

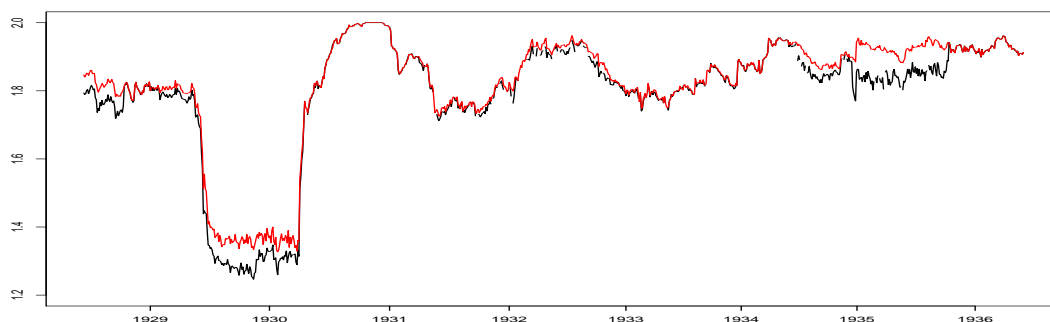


FIGURE 2.13 – Valeurs des estimateurs combinés symétriques (en rouge) et asymétriques (en noir) appliqués à l'indice financier S&P 500 autour du krach boursier de Wall Street (fenêtre glissante de taille $n = 200$ observations).

2.6 Annexe

Démonstration de la Proposition 2.2.1 : Soit Z une variable aléatoire stable de loi

$S_\alpha(1, 0, 0)$. On sait que :

- Z a une densité bornée,
- $\lim_{\lambda \rightarrow +\infty} \lambda^\alpha \mathbb{P}(Z > \lambda) = C_\alpha$,
- $\lim_{\lambda \rightarrow +\infty} \lambda^\alpha \mathbb{P}(Z < -\lambda) = C_\alpha$,

(Voir [ST94], Propriété 1.2.15). Donc pour tout $p > 0$,

$$\mathbb{E}[|\log |Z||^p] = \int_0^\infty \mathbb{P}(|\log |Z||^p > x) dx < \infty. \quad \blacksquare$$

Démonstration de la Proposition 2.2.2 : On utilise les formules des Propriétés 1.2.17 et 1.2.15 de [ST94] pour calculer $\mathbb{E}[|Z|^t]$ pour $0 < t < \alpha$:

$$\mathbb{E}[|Z|^t] = \frac{2^{t-1} \Gamma(1-t/\alpha)}{t \int_0^{+\infty} \frac{\sin^2(u)}{u^{t+1}} du} = \frac{2^{t-1} \Gamma(1-t/\alpha)}{\int_0^{+\infty} \frac{\sin(2u)}{u^t} du} = \frac{(1-t) \Gamma(1-t/\alpha)}{\Gamma(2-t) \cos(\pi t/2)}.$$

De plus, si $0 < t < 1$, on a :

$$\mathbb{E}[|Z|^t] = \frac{\Gamma(1 - t/\alpha)}{\Gamma(1 - t) \cos(\pi t/2)}. \quad \blacksquare$$

Démonstration du Théorème 2.2.3 : D'après la loi forte des grands nombres pour les variables aléatoires $(\log |X_i|)_i$ qui sont dans L^1 (Proposition 2.2.1), et par le *continuous mapping theorem* avec $g(x) = \frac{\gamma}{\gamma+x}$, on obtient $\hat{\alpha}_n \xrightarrow[n \rightarrow \infty]{p.s.} \alpha$. On applique ensuite le théorème central limite et la delta méthode avec la fonction $f(x) = \frac{\pi^2}{6x^2} + \frac{\pi^2}{12}$ pour obtenir

$$\frac{\sqrt{n}(\hat{\alpha}_n - \alpha)\gamma}{\hat{\alpha}_n^2 \sqrt{f(\hat{\alpha}_n)}} \xrightarrow{Loi} \mathcal{N}(0, 1). \quad \blacksquare$$

Démonstration du Théorème 2.2.4 : On prouve ce théorème à l'aide de la loi forte des grands nombres, du *continuous mapping theorem*, du théorème central limite multidimensionnel et de la delta méthode. On a besoin des log-moments d'ordre 3 et 4 d'une loi stable $Z \sim S_\alpha(1, 0, 0)$ pour le calcul de la matrice de covariance. On a

$$\mathbb{E}[(\log |Z| - \mathbb{E}[\log |Z|])^3] = 2\zeta(3) \left(\frac{1}{\alpha^3} - 1 \right),$$

et donc :

$$\mathbb{E}[(\log |Z|)^3] = \frac{4\gamma^3 + 2\gamma\pi^2 + 8\zeta(3)}{4\alpha^3} + \frac{-12\gamma^3 - 2\gamma\pi^2}{4\alpha^2} + \frac{12\gamma^3 + \gamma\pi^2}{4\alpha} + \frac{-4\gamma^3 - \gamma\pi^2 - 8\zeta(3)}{4},$$

et

$$\begin{aligned} \mathbb{E}[(\log |Z| - \mathbb{E}[\log |Z|])^4] &= \pi^4 \left(\frac{3}{20\alpha^4} + \frac{1}{12\alpha^2} + \frac{19}{240} \right), \\ \mathbb{E}[(\log |Z|)^4] &= \frac{240\gamma^4 + 240\gamma^2\pi^2 + 36\pi^4 + 1920\zeta(3)\gamma}{240\alpha^4} + \frac{-960\gamma^4 - 480\gamma^2\pi^2 - 1920\zeta(3)\gamma}{240\alpha^3} \\ &+ \frac{1440\gamma^4 + 360\gamma^2\pi^2 + 20\pi^4}{240\alpha^2} + \frac{-960\gamma^4 - 240\gamma^2\pi^2 - 1920\zeta(3)\gamma}{240\alpha} \\ &+ \frac{240\gamma^4 + 120\gamma^2\pi^2 + 19\pi^4 + 1920\zeta(3)\gamma}{240} \end{aligned}$$

où ζ est la fonction zêta de Riemann $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ et $\zeta(3) = 1.2020569 \dots$ \blacksquare

Démonstration de la Proposition 2.4.1 : Par la loi forte des grands nombres de Kolmogorov pour des variables aléatoires non-identiquement distribuées (Théorème

2.3.10 de [SS94]), on obtient

$$\frac{1}{n} \sum_{i=1}^n Y_i^2 - \left(\frac{1}{n} \sum_{i=1}^n Y_i \right)^2 - \left[\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] \right)^2 \right] \xrightarrow{p.s.} 0.$$

Or

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] - \mathbb{E}[Y] &= \frac{\gamma}{n} \sum_{i=1}^n \left(\frac{1}{\alpha_i} - \frac{1}{\alpha} \right) + \frac{1}{n} \sum_{i=1}^n (\log \sigma_i - \log \sigma) \\ &= \frac{\gamma}{n} \sum_{i=1}^n \frac{-\varepsilon_i}{(\alpha_i^*)^2} + \frac{1}{n} \sum_{i=1}^n \frac{\eta_i}{\sigma_i^*}, \end{aligned}$$

et

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \mathbb{E}[Y^2] \\ &= \frac{\gamma^2 + \pi^2/6}{n} \sum_{i=1}^n \left(\frac{1}{\alpha_i^2} - \frac{1}{\alpha^2} \right) - \frac{2\gamma^2}{n} \sum_{i=1}^n \left(\frac{1}{\alpha_i} - \frac{1}{\alpha} \right) \\ &\quad + \frac{2\gamma}{n} \sum_{i=1}^n \left(\left(\frac{1}{\alpha_i} - 1 \right) \log \sigma_i - \left(\frac{1}{\alpha} - 1 \right) \log \sigma \right) + \frac{1}{n} \sum_{i=1}^n \left((\log \sigma_i)^2 - (\log \sigma)^2 \right) \\ &= \frac{\gamma^2 + \pi^2/6}{n} \sum_{i=1}^n \frac{-2\varepsilon_i}{\tilde{\alpha}_i^3} - \frac{2\gamma^2}{n} \sum_{i=1}^n \frac{-\varepsilon_i}{(\alpha_i^*)^2} + \frac{2\gamma}{n} \sum_{i=1}^n \left(\frac{1}{\check{\sigma}_i} \left(\frac{1}{\check{\alpha}_i} - 1 \right) \eta_i - \frac{\log(\check{\sigma}_i)}{(\check{\alpha}_i)^2} \varepsilon_i \right) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{2 \log \tilde{\sigma}_i}{\tilde{\sigma}_i} \eta_i \end{aligned}$$

où $\tilde{\alpha}_i, \alpha_i^*, \check{\alpha}_i$ (respectivement $\tilde{\sigma}_i, \sigma_i^*, \check{\sigma}_i$) sont compris entre α et α_i (respectivement σ et σ_i).

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] - \mathbb{E}[Y] \right| \leq \frac{\gamma}{\alpha^2(1-c_\alpha)^2} \frac{1}{n} \sum_{i=1}^n |\varepsilon_i| + \frac{1}{\sigma(1-c_\sigma)} \frac{1}{n} \sum_{i=1}^n |\eta_i|$$

et donc

$$\begin{aligned} &\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \mathbb{E}[Y^2] \right| \\ &\leq \left(\frac{2\gamma^2 + \pi^2/3}{\alpha^3(1-c_\alpha)^3} + \frac{2\gamma^2}{\alpha^2(1-c_\alpha)^2} + \frac{2\gamma \max(|\log(\sigma - \sigma c_\sigma)|, |\log(\sigma + \sigma c_\sigma)|)}{\alpha^2(1-c_\alpha)^2} \right) \frac{1}{n} \sum_{i=1}^n |\varepsilon_i| \\ &\quad + \left(\frac{2\gamma}{\sigma(1-c_\sigma)} \left(\frac{1}{\alpha(1-c_\alpha)} + 1 \right) + \frac{2 \max(|\log(\sigma - \sigma c_\sigma)|, |\log(\sigma + \sigma c_\sigma)|)}{\sigma(1-c_\sigma)} \right) \frac{1}{n} \sum_{i=1}^n |\eta_i|. \end{aligned}$$

Sous les conditions :

$$1. \frac{1}{n} \sum_{i=1}^n |\varepsilon_i| \xrightarrow[n \rightarrow \infty]{} 0,$$

$$2. \frac{1}{n} \sum_{i=1}^n |\eta_i| \xrightarrow[n \rightarrow \infty]{} 0,$$

on obtient

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[Y^2] \quad \text{et} \quad \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[Y].$$

D'après le *continuous mapping theorem* avec $g(x, y) = \frac{\pi}{\sqrt{\max(6(y-x^2) - \frac{\pi^2}{2}, \frac{\pi^2}{4})}}$, on obtient alors :

$$\hat{\alpha}_n - \alpha \xrightarrow[p.s.]{} 0. \quad \blacksquare$$

Preuve de la Proposition 2.4.2

On définit la matrice de covariance $\Sigma_{\alpha, \sigma}^{(n)}$:

$$\Sigma_{\alpha, \sigma}^{(n)} := \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \text{Var}(Y_i) & \frac{1}{n} \sum_{i=1}^n \text{Cov}(Y_i, Y_i^2) \\ \frac{1}{n} \sum_{i=1}^n \text{Cov}(Y_i, Y_i^2) & \frac{1}{n} \sum_{i=1}^n \text{Var}(Y_i^2) \end{pmatrix}.$$

Avec les conditions de la Proposition 2.4.1, on a

$$\Sigma_{\alpha, \sigma}^{(n)} \xrightarrow[n \rightarrow \infty]{} \Sigma_{\alpha, \sigma} := \begin{pmatrix} \text{Var}(Y) & \text{Cov}(Y, Y^2) \\ \text{Cov}(Y, Y^2) & \text{Var}(Y^2) \end{pmatrix}.$$

Par le théorème central limite pour des variables aléatoires non identiquement distribuées (Théorème 3.3.9 in [SS94]), comme $\sup_k \mathbb{E} \left[\left| \begin{pmatrix} Y_k \\ (Y_k)^2 \end{pmatrix} \right|^4 \right] < \infty$, on a,

$$\sqrt{n} \left(\begin{pmatrix} \frac{1}{n} \sum_{i=1}^n Y_i \\ \frac{1}{n} \sum_{i=1}^n Y_i^2 \end{pmatrix} - \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] \\ \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] \end{pmatrix} \right) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}(0, \Sigma_{\alpha, \sigma}).$$

Sous les conditions de la Proposition 2.4.2, on obtient

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \mathbb{E}[Y^2] \right) \xrightarrow[n \rightarrow \infty]{} 0,$$

et

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] - \mathbb{E}[Y] \right) \xrightarrow[n \rightarrow \infty]{} 0,$$

ainsi

$$\sqrt{n} \left(\begin{pmatrix} \frac{1}{n} \sum_{i=1}^n Y_i \\ \frac{1}{n} \sum_{i=1}^n Y_i^2 \end{pmatrix} - \begin{pmatrix} \mathbb{E}[Y] \\ \mathbb{E}[Y^2] \end{pmatrix} \right) \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, \Sigma_{\alpha, \sigma}).$$

En appliquant la delta méthode avec $g(x, y) = \frac{\pi}{\sqrt{\max(6(y-x^2) - \frac{\pi^2}{2}, \frac{\pi^2}{4})}}$, on obtient le résultat.

Démonstration de la Proposition 2.4.3 : Par la loi forte des grands nombres pour des variables aléatoires non identiquement distribuées (Théorème 2.3.10 de [SS94]), on a

$$\frac{1}{n} \sum_{i=1}^n Y_i^2 - \left(\frac{1}{n} \sum_{i=1}^n Y_i \right)^2 - \left[\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] \right)^2 \right] \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

Avec le même calcul, on obtient

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \mathbb{E}[Y^2] &= \frac{\gamma^2 + \pi^2/6}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{1}{\alpha_i^2} - \frac{1}{\alpha^2} \right] + \frac{-2\gamma^2 + 2 \log(\sigma)}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{1}{\alpha_i} - \frac{1}{\alpha} \right] \\ &= \frac{\gamma^2 + \pi^2/6}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{-2\varepsilon_i}{(\tilde{\alpha}_i)^3} \right] + \frac{-2\gamma^2 + 2 \log(\sigma)}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{-\varepsilon_i}{(\alpha_i^*)^2} \right] \end{aligned}$$

et

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] - \mathbb{E}[Y] = \frac{\gamma}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{1}{\alpha_i} - \frac{1}{\alpha} \right] = \frac{\gamma}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{-\varepsilon_i}{(\alpha_i^*)^2} \right]$$

avec α_i^* et $\tilde{\alpha}_i \in (\min(\alpha, \alpha_i), \max(\alpha, \alpha_i))$. Comme α_i^* et $\tilde{\alpha}_i$ sont presque sûrement inclus dans l'intervalle $[\alpha(1 - c_\alpha), \alpha(1 + c_\alpha)]$, on a alors

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] - \mathbb{E}[Y^2] \right| \leq \left(\frac{\gamma^2 + \pi^2/6}{(\alpha(1 - c_\alpha))^3} + \frac{|-2\gamma^2 + 2 \log(\sigma)|}{(\alpha(1 - c_\alpha))^2} \right) \frac{1}{n} \sum_{i=1}^n \mathbb{E}[|\varepsilon_i|],$$

et

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] - \mathbb{E}[Y] \right| \leq \frac{\gamma}{(\alpha(1 - c_\alpha))^2} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[|\varepsilon_i|].$$

Si on suppose $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[|\varepsilon_i|] \rightarrow 0$, on obtient

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i^2] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[Y^2] \quad \text{et} \quad \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Y_i] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[Y].$$

Le théorème central limite se prouve comme dans la Proposition 2.4.2. ■

Simulation de processus multistables

Sommaire

3.1 Processus localisables	84
3.2 Processus multistables définis par un champ stochastique	85
3.2.1 Cas des espaces de mesure finie	85
3.2.2 Cas des espaces de mesure σ -finie	87
3.3 Grille multistable	90
3.4 Exemples de simulation	103
3.4.1 Mouvement de Lévy multistable	103
3.4.2 Mouvement linéaire multifractionnaire multistable	105
3.5 Lemmes techniques	107

Le but de ce chapitre est de simuler de manière approchée des processus multistables définis grâce à un champ dont nous rappelons la construction issue de [LGLV12] dans la Section 3.2. Ces processus sont définis à partir d'une fonction d'intensité $\alpha : \mathbb{R} \rightarrow]0, 2[$ qui décrit le comportement local du processus. Plus précisément, cela signifie qu'autour du point t_0 , un processus multistable se comporte comme un processus $\alpha(t_0)$ -stable.

Dans la Section 3.1, on commence par redéfinir les processus localisables dont on donne des exemples. Puis, nous rappelons dans la Section 3.2 la construction de processus multistables à partir d'un champ dans des espaces de mesure finie ou σ -finie. Nous construisons alors une méthode de simulation approchée à l'aide d'une grille sur le champ dans la Section 3.3 pour laquelle nous donnons l'erreur en probabilité entre le processus exact et sa simulation approchée. Enfin, la Section 3.4 est consacrée à l'étude numérique des performances de cette méthode de simulation sur les exemples du mouvement de Lévy multistable et du mouvement linéaire multifractionnaire multistable.

3.1 Processus localisables

Pour commencer, on rappelle la définition d'un processus localisable (voir [Fal02], [Fal03]).

Définition 3.1.1

Le processus $Y = (Y(t))_{t \in \mathbb{R}}$ est dit h -localisable au point $u \in \mathbb{R}$ s'il existe un processus non trivial $Y'_u = (Y'_u(t))_{t \in \mathbb{R}}$ tel que

$$\lim_{r \rightarrow 0} \frac{Y(u + rt) - Y(u)}{r^h} = Y'_u(t),$$

où la convergence est au sens des lois fini-dimensionnelles. Lorsque cette limite existe, Y'_u est appelé forme locale ou processus tangent de Y au point u .

Lorsqu'elle existe, la h -forme locale Y'_u est h -auto-similaire, c'est à dire pour tout $c > 0$ et $t \in \mathbb{R}$, $Y'_u(ct) \stackrel{Loi}{=} c^h Y'_u(t)$. Sous des conditions supplémentaires (voir [Fal02], [Fal03]), il est possible de montrer que Y'_u doit également avoir des accroissements stationnaires pour presque tout u .

Réciproquement, les processus auto-similaires à accroissements stationnaires admettent le résultat de localisation suivant :

Proposition 3.1.2 (voir [Fal02])

Soit $Y = (Y(t))_{t \in \mathbb{R}}$ un processus h -auto-similaire à accroissements stationnaires. Alors Y est h -localisable pour tout $u \in \mathbb{R}$ de forme locale $Y'_u = Y$.

La classe des processus localisables comprend ainsi plusieurs processus définis dans le chapitre d'introduction :

- le mouvement de Lévy α -stable L_α (voir Exemple 1.1.2) qui est, pour tout u , $1/\alpha$ -localisable de forme locale L_α ,
- le mouvement linéaire fractionnaire α -stable $L_{\alpha,H}(a, b; \cdot)$ (voir Exemple 1.1.6) qui est, pour tout u , H -localisable de forme locale $L_{\alpha,H}(a, b; \cdot)$,
- le mouvement log-fractionnaire stable (Exemple 1.1.7) qui est, pour tout u , $1/\alpha$ -localisable de forme locale lui-même.

Le processus d'Ornstein-Uhlenbeck rétrograde stable est lui-aussi localisable (cela est prouvé dans [Fal+09]). Au point u , il est $1/\alpha$ -localisable de forme locale L_α .

Pour construire des processus localisables de forme locale donnée, [LGLV12] se sert de champs stochastiques $\{X(t, v)\}_{(t,v) \in \mathbb{R}^2}$ où le processus $X(\cdot, v)$ est localisable pour tout v . À partir de ce champ, il est possible de décrire la forme locale du processus diagonal $Y = \{X(t, t)\}_{t \in \mathbb{R}}$. En particulier, on construit ainsi des processus dont la forme locale évolue au cours du temps.

3.2 Processus multistables définis par un champ stochastique

La notion de processus multistable est redéfinie dans cette section et une méthode de simulation de tels processus est donnée. À partir de la représentation de Ferguson-Klass-LePage dans le cas symétrique d'intégrale stochastique α -stable (voir Section 1.1.2), on rappelle la construction de processus où α varie au cours du temps. Cette construction est décrite dans [LGLV12]. D'autres approches pour construire des processus multistables existent. En particulier, dans [FLV09], des processus multifractionnaires multistables ont été construits à partir d'un processus de Poisson.

3.2.1 Cas des espaces de mesure finie

Soit (E, \mathcal{E}, m) un espace de mesure finie avec $m \neq 0$. Soit α une fonction de classe C^1 définie sur \mathbb{R} et à valeurs dans $]0, 2[$, b une fonction de classe C^1 définie sur \mathbb{R} et $f(t, u, \cdot)$ une famille de fonctions telles que, pour tout $(t, u) \in \mathbb{R}^2$, $f(t, u, \cdot) \in \mathcal{F}_{\alpha(u)}(E, \mathcal{E}, m)$.

On considère $(\Gamma_i)_{i \geq 1}$ une suite de temps d'arrivée d'un processus de Poisson

avec un taux d'arrivée unitaire : $\Gamma_i = \sum_{j=1}^i \varepsilon_j$ où les (ε_j) sont i.i.d. de loi $\mathcal{E}(1)$. Soit $(V_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $\widehat{m} = m/m(E)$ sur E et $(\gamma_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$. On suppose finalement que les trois suites $(\Gamma_i)_{i \geq 1}$, $(V_i)_{i \geq 1}$ et $(\gamma_i)_{i \geq 1}$ sont indépendantes.

On rappelle alors la définition des processus multistables basée sur la représentation de Ferguson-Klass-LePage des intégrales stables.

Théorème 3.2.1 ([LGLV12], Théorème 3.2)

On considère le champ stochastique défini par :

$$X(t, u) = b(u) \left(C_{\alpha(u)} m(E) \right)^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} f(t, u, V_i), \quad (3.2.1)$$

où C_α est la constante définie par (1.1.3). Soit $u > 0$ fixé. On suppose que le processus $X(\cdot, u)$ est localisable en u d'exposant $h \in]0, 1[$ et de forme locale $X'_u(\cdot, u)$. On suppose de plus qu'il existe $\varepsilon_0 > 0$ tel que :

(C1) $v \rightarrow f(t, v, x)$ est dérivable sur $B(u, \varepsilon_0)$ pour tout $t \in B(u, \varepsilon_0)$ et pour presque tout $x \in E$ (on notera cette dérivée f'_2),

(C2)

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f(t, w, x) \log |f(t, w, x)||^{\alpha(w)} \right) \widehat{m}(dx) < \infty,$$

(C3)

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f'_2(t, w, x)|^{\alpha(w)} \right) \widehat{m}(dx) < \infty.$$

Alors le processus diagonal Y défini par

$$Y(t) = X(t, t) \quad (3.2.2)$$

est localisable en u d'exposant h et de forme locale Y'_u telle que $Y'_u(t) = X'_u(t, u)$.

On peut remarquer que pour tout $t \in B(u, \varepsilon_0)$, $w \in B(u, \varepsilon_0)$ et $x \in E$,

$$\begin{aligned} |f(t, w, x)| &= |f(t, w, x)| \mathbb{1}_{|f(t, w, x)| < 1/e} + |f(t, w, x)| \mathbb{1}_{|f(t, w, x)| > e} \\ &\quad + |f(t, w, x)| \mathbb{1}_{1/e \leq |f(t, w, x)| \leq e}, \\ &\leq 2 |f(t, w, x) \log |f(t, w, x)|| + e. \end{aligned}$$

On obtient alors la majoration suivante :

$$|f(t, w, x)|^{\alpha(w)} \leq \max(1, 2^{\alpha(w)-1}) \left(|2f(t, w, x) \log |f(t, w, x)||^{\alpha(w)} + e^{\alpha(w)} \right).$$

La condition **(C2)** implique ainsi la condition suivante :

(C4)

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f(t, w, x)|^{\alpha(w)} \right) \widehat{m}(dx) < \infty.$$

Exemple 3.2.1. Mouvement de Lévy multistable, cas compact

Soit $T > 0$ fixé, $\alpha : [0, T] \rightarrow]1, 2[$ et $b : [0, T] \rightarrow]0, \infty[$ deux fonctions C^1 , et $\widehat{m}(dx)$ la loi uniforme sur $[0, T]$. Pour $t \in [0, T]$, posons

$$Y(t) = b(t) \left(C_{\alpha(t)} T \right)^{1/\alpha(t)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} \mathbf{1}_{[0, t]}(V_i).$$

Alors, Y est $1/\alpha(u)$ -localisable au point u de forme locale $Y'_u = b(u)L_{\alpha(u)}$.

3.2.2 Cas des espaces de mesure σ -finie

Lorsque l'espace E est de mesure infinie, il n'est pas possible d'utiliser la représentation des processus multistables définie dans le Théorème 3.2.1. En particulier, aucun processus défini sur l'espace $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$ ne pourra s'écrire sous cette forme. Comme décrit dans l'introduction dans la Proposition 1.1.7, il est possible de se ramener au cas d'une mesure finie grâce à un changement de mesure.

Soit (E, \mathcal{E}, m) un espace de mesure σ -finie et $r : E \rightarrow \mathbb{R}^+$ telle que la mesure $\widehat{m}(dx) = \frac{1}{r(x)} m(dx)$ soit une mesure de probabilité. On se donne α une fonction de classe C^1 définie sur \mathbb{R} et à valeurs dans $]0, 2[$, b une fonction de classe C^1 définie sur \mathbb{R} et $f(t, u, \cdot)$ une famille de fonctions telles que, pour tout $(t, u) \in \mathbb{R}^2$, $f(t, u, \cdot) \in \mathcal{F}_{\alpha(u)}(E, \mathcal{E}, m)$.

On considère $(\Gamma_i)_{i \geq 1}$ une suite de temps d'arrivée d'un processus de Poisson avec un taux d'arrivée unitaire : $\Gamma_i = \sum_{j=1}^i \varepsilon_j$ où les (ε_j) sont i.i.d. $\mathcal{E}(1)$. Soit $(V_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi \widehat{m} sur E et $(\gamma_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$. On suppose finalement que les trois suites $(\Gamma_i)_{i \geq 1}$, $(V_i)_{i \geq 1}$ et $(\gamma_i)_{i \geq 1}$ sont indépendantes.

Théorème 3.2.2 ([LGLV12], Théorème 4.1)

On considère le champ stochastique défini par :

$$X(t, u) = b(u)C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} r(V_i)^{1/\alpha(u)} f(t, u, V_i), \quad (3.2.3)$$

où C_α est la constante définie par (1.1.3). Soit $u > 0$ fixé. On suppose que le processus $X(\cdot, u)$ est localisable en u d'exposant $h \in]0, 1[$ et de forme locale $X'_u(\cdot, u)$. On suppose de plus qu'il existe $\varepsilon_0 > 0$ tel que :

(C1) $v \rightarrow f(t, v, x)$ est dérivable sur $B(u, \varepsilon_0)$ pour tout $t \in B(u, \varepsilon_0)$ et pour presque tout $x \in E$ (on notera cette dérivée f'_2),

(C2')

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f(t, w, x) \log |f(t, w, x)||^{\alpha(w)} \right) m(dx) < \infty,$$

(C3')

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f'_2(t, w, x)|^{\alpha(w)} \right) m(dx) < \infty,$$

(C4')

$$\sup_{t \in B(u, \varepsilon_0)} \int_E \sup_{w \in B(u, \varepsilon_0)} \left(|f(t, w, x) \log r(x)|^{\alpha(w)} \right) m(dx) < \infty.$$

Alors le processus diagonal Y défini par

$$Y(t) = X(t, t) \quad (3.2.4)$$

est localisable en u d'exposant h et de forme locale Y'_u telle que $Y'_u(t) = X'_u(t, u)$.

On donne désormais les versions multistables de certains processus stables définis dans la Section 1.1.2 dans le cas d'un espace non-compact. Dans [LGLV12], les hypothèses nécessaires afin d'appliquer le Théorème 3.2.2 sont vérifiées pour ces processus.

Exemple 3.2.2. Mouvement de Lévy multistable, cas non-compact

On se place sur l'espace de mesure σ -finie $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+), \text{Leb})$. Soient $\alpha : \mathbb{R} \rightarrow]1, 2[$ et $b : \mathbb{R} \rightarrow]0, \infty[$ deux fonctions \mathcal{C}^1 .

On prend $r(x) = \sum_{j \geq 1} 2^j \mathbb{1}_{x \in [j-1, j[}$ de sorte que $\widehat{m}(dx) = \sum_{j \geq 1} \frac{1}{2^j} \mathbb{1}_{x \in [j-1, j[} dx$ soit une

mesure de probabilité sur \mathbb{R}^+ . Pour $t \in \mathbb{R}^+$, posons

$$Y(t) = b(t)C_{\alpha(t)}^{1/\alpha(t)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} \sum_{j \geq 1} 2^{j/\alpha(t)} \mathbb{1}_{[0,t] \cap [j-1,j[}(V_i).$$

Alors, Y est $1/\alpha(u)$ -localisable au point u de forme locale $Y'_u = b(u)L_{\alpha(u)}$.

Exemple 3.2.3. Processus d'Ornstein-Uhlenbeck Rétrograde multistable

On se place sur l'espace de mesure σ -finie $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$. Soient $\lambda > 0$, $\alpha \in \mathcal{C}^1(\mathbb{R},]1, 2[)$ et $b \in \mathcal{C}^1(\mathbb{R},]0, \infty[)$.

On prend $r(x) = \sum_{j \geq 1} 2^{j+1} \mathbb{1}_{|x| \in [j-1, j[}$ de sorte que $\widehat{m}(dx) = \sum_{j \geq 1} \frac{1}{2^{j+1}} \mathbb{1}_{|x| \in [j-1, j[} dx$ soit une mesure de probabilité sur \mathbb{R} . Pour $t \in \mathbb{R}$, posons

$$Y(t) = b(t)C_{\alpha(t)}^{1/\alpha(t)} \sum_{i,j=1}^{\infty} \gamma_i \Gamma_i^{-\frac{1}{\alpha(t)}} 2^{(j+1)/\alpha(t)} e^{-\lambda(V_i-t)} \mathbb{1}_{V_i \geq t} \mathbb{1}_{|V_i| \in [j-1, j[}.$$

Alors, Y est $1/\alpha(u)$ -localisable au point u de forme locale $Y'_u = b(u)L_{\alpha(u)}$.

Exemple 3.2.4. Mouvement linéaire multifractionnaire multistable

On se place sur l'espace de mesure σ -finie $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$. Soient $\alpha : \mathbb{R} \rightarrow]0, 2[$, $b : \mathbb{R} \rightarrow]0, \infty[$ et $h : \mathbb{R} \rightarrow]0, 1[$ trois fonctions \mathcal{C}^1 . On prend $r(x) = \frac{\pi^2}{3} \sum_{j \geq 1} j^2 \mathbb{1}_{|x| \in [j-1, j[}$ de sorte que $\widehat{m}(dx) = \frac{3}{\pi^2} \sum_{j \geq 1} \frac{1}{j^2} \mathbb{1}_{|x| \in [j-1, j[} dx$ soit une mesure de probabilité sur \mathbb{R} . Pour $t \in \mathbb{R}$, posons

$$Y(t) = b(t)C_{\alpha(t)}^{1/\alpha(t)} \sum_{i,j=1}^{\infty} \gamma_i \Gamma_i^{-\frac{1}{\alpha(t)}} \left(\frac{\pi^2 j^2}{3} \right)^{\frac{1}{\alpha(t)}} \mathbb{1}_{|V_i| \in [j-1, j[} \left[|t - V_i|^{h(t) - \frac{1}{\alpha(t)}} - |V_i|^{h(t) - \frac{1}{\alpha(t)}} \right].$$

Alors, Y est $h(u)$ -localisable au point u de forme locale $Y'_u = b(u)L_{\alpha(u), h(u)}$.

Exemple 3.2.5. Mouvement log-fractionnaire multistable

On se place sur l'espace de mesure σ -finie $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$. Soient $\alpha : \mathbb{R} \rightarrow]1, 2[$ et $b : \mathbb{R} \rightarrow]0, \infty[$ deux fonctions \mathcal{C}^1 . On prend $r(x) = \frac{\pi^2}{3} \sum_{j \geq 1} j^2 \mathbb{1}_{|x| \in [j-1, j[}$ de sorte que

$\widehat{m}(dx) = \frac{3}{\pi^2} \sum_{j \geq 1} \frac{1}{j^2} \mathbb{1}_{|x| \in [j-1, j[} dx$ soit une mesure de probabilité sur \mathbb{R} . Pour $t \in \mathbb{R}$, posons

$$Y(t) = b(t) C_{\alpha(t)}^{1/\alpha(t)} \sum_{i,j=1}^{\infty} \gamma_i \Gamma_i^{-\frac{1}{\alpha(t)}} \left(\frac{\pi^2 j^2}{3} \right)^{\frac{1}{\alpha(t)}} \mathbb{1}_{|V_i| \in [j-1, j[} (\ln |t - V_i| - \ln |V_i|).$$

Alors, Y est $1/\alpha(u)$ -localisable au point u de forme locale $Y'_u = b(u) \Lambda_{\alpha(u)}$.

3.3 Grille multistable

Il est souhaitable de savoir simuler les processus multistables définis précédemment sur un intervalle de temps de la forme $[0, T]$ pour $T > 0$ fixé. Pour alléger les notations, dans la suite, T sera fixé à 1. On utilise dans cette partie la construction des processus multistables sur l'espace $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$ issus de champs aléatoires définis dans (3.2.1) de la forme :

$$X(t, u) = C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} f(t, u, V_i), \quad (t, u) \in [0, 1]^2.$$

Le processus multistable correspond alors au processus diagonal $Y(t) = X(t, t)$. Une autre manière de construire ce processus à partir d'un champ consiste à regarder le champ

$$\widehat{X}(t, \alpha) = C_{\alpha}^{1/\alpha} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha} \widehat{f}(t, \alpha, V_i), \quad t \in [0, 1], \alpha \in]0, 2],$$

avec $f(t, t, V_i) = \widehat{f}(t, \alpha(t), V_i)$ puis d'intersecter ce champ avec la courbe $(t, \alpha(t))$ ce qui donne

$$Y(t) = \widehat{X}(t, \alpha(t)).$$

Lorsque l'on change de fonction d'intensité α , on doit simuler un nouveau champ dans le cas $\{X(t, u)\}_{t \in [0, 1], u \in]0, 1[}$ puisque ce champ est défini pour une fonction α fixée. L'avantage Le champ $\{\widehat{X}(t, \alpha)\}_{t \in [0, 1], \alpha \in]0, 2[}$ a l'avantage de pouvoir donner des réalisations de processus multistables avec différentes fonctions d'intensité à partir d'une réalisation du champ.

Une première méthode pour simuler ces processus aux points $\frac{1}{N}, \frac{2}{N}, \dots, 1$ consiste à créer tout d'abord N processus stables $X(\cdot, \frac{k}{N})$ pour k allant de $\{1, \dots, N\}$, en prenant une même valeur de graine pour la simulation de chacun des processus

stables.

Le processus diagonal Y est alors obtenu aux points $\left(\frac{k}{N}\right)_{k \in \{1, \dots, N\}}$ par

$$Y\left(\frac{k}{N}\right) = X\left(\frac{k}{N}, \frac{k}{N}\right).$$

À condition de savoir simuler de manière exacte les processus stables qui composent le champ, cette première méthode est une méthode de simulation exacte. Cependant, elle est très coûteuse en temps de calcul car elle nécessite de calculer N points pour N processus stables.

Une nouvelle méthode consiste à discrétiser de manière régulière l'intervalle $]0, 2[$, qui est l'axe correspondant à l'indice de stabilité du processus, en $(2K - 1)$ valeurs possibles. On ne simule plus le processus diagonal $Y = (X(t, t))_{t \in [0, 1]}$ mais un processus approché $(\tilde{X}_K(t, t))_{t \in [0, 1]}$ défini à partir du champ :

$$\tilde{X}_K(t, u) = C_{\alpha_K(u)}^{1/\alpha_K(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha_K(u)} f(t, u, V_i), \quad (3.3.1)$$

$$\text{où } \alpha_K(u) = \max\left(\frac{\lfloor K\alpha(u) \rfloor}{K}, \frac{1}{K}\right).$$

Pour u fixé, $t \rightarrow \tilde{X}_K(t, u)$ est un processus $\alpha_K(u)$ -stable.

De plus, comme $\alpha_K(u) \in \left\{\frac{1}{K}, \frac{2}{K}, \dots, \frac{2K-1}{K}\right\}$, il faut calculer au maximum N points de $(2K - 1)$ processus stables afin d'obtenir les valeurs approchées du processus diagonal aux points $\left(\frac{k}{N}\right)_{k \in \{1, \dots, N\}}$.

Le résultat suivant donne l'erreur en probabilité entre le processus multistable défini par (3.2.1) et son approximation (3.3.1).

Théorème 3.3.1

Soit $t > 0$ fixé. On suppose que la condition **(C2)** est vérifiée au point t . Soit $K \in \mathbb{N}^*$ la constante de discrétisation de l'axe de l'intensité des sauts dans le champ stochastique tel que le pas vérifie $\frac{1}{K} < \varepsilon_0$, où ε_0 apparaît dans **(C2)**. Alors, on a

$$X(t, t) - \widetilde{X}_K(t, t) = (\alpha(t) - \alpha_K(t))Z_K(t)$$

avec

$$Z_K(t) \xrightarrow[K \rightarrow \infty]{Loi} b(t) \sum_{i=1}^{\infty} \gamma_i \left(g'(\alpha(t)) + g(\alpha(t)) \frac{\log \Gamma_i}{\alpha(t)^2} \right) \Gamma_i^{-1/\alpha(t)} f(t, t, V_i)$$

où la fonction g est définie sur $]0, 2[$ par $g(w) = (m(E)C_w)^{1/w}$. La limite est une série presque sûrement convergente.

En conséquence,

$$|X(t, t) - \widetilde{X}_K(t, t)| = \mathcal{O}_{\mathbb{P}} \left(\frac{1}{K} \right).$$

Démonstration : Soit $K \in \mathbb{N}^*$ tel que $1/K < \varepsilon_0$, on a

$$X(t, t) - \widetilde{X}_K(t, t) = b(t) \sum_{i=1}^{\infty} \gamma_i \left(g(\alpha(t)) \Gamma_i^{-1/\alpha(t)} - g(\alpha_K(t)) \Gamma_i^{-1/\alpha_K(t)} \right) f(t, t, V_i).$$

Comme les variables aléatoires Γ_i ne sont pas indépendantes, la somme est décomposée en deux afin de faciliter l'étude :

$$\begin{aligned} X(t, t) - \widetilde{X}_K(t, t) &= b(t) \sum_{i=1}^{\infty} \gamma_i (\Phi_i(\alpha(t)) - \Phi_i(\alpha_K(t))) f(t, t, V_i) \\ &\quad + b(t) \sum_{i=1}^{\infty} \gamma_i (\Psi_i(\alpha(t)) - \Psi_i(\alpha_K(t))) f(t, t, V_i), \end{aligned}$$

où

$$\begin{aligned} \Phi_i(w) &= g(w) i^{-1/w}, \\ \Psi_i(w) &= g(w) \left(\Gamma_i^{-1/w} - i^{-1/w} \right). \end{aligned}$$

Φ_i est dérivable sur $]0, 2[$ et Ψ_i est presque sûrement dérivable sur $]0, 2[$:

$$\begin{aligned}\Phi'_i(w) &= g'(w)i^{-1/w} + g(w)\frac{\log i}{w^2}i^{-1/w}, \\ \Psi'_i(w) &= g'(w)\left(\Gamma_i^{-1/w} - i^{-1/w}\right) + g(w)\left(\frac{\log \Gamma_i}{w^2}\Gamma_i^{-1/w} - \frac{\log i}{w^2}i^{-1/w}\right).\end{aligned}$$

Il existe une suite déterministe $w_{i,K} \in [\alpha_K(t), \alpha(t)]$ et une suite de variables aléatoires $x_{i,K} \in [\alpha_K(t), \alpha(t)]$ tels que :

$$X(t, t) - \tilde{X}_K(t, t) = (\alpha(t) - \alpha_K(t))b(t) \sum_{i=1}^{\infty} (M_{i,K} + N_{i,K} + P_{i,K} + Q_{i,K}),$$

où $M_{i,K}$, $N_{i,K}$, $P_{i,K}$ et $Q_{i,K}$ sont définis par :

$$\begin{aligned}M_{i,K} &= \gamma_i g'(w_{i,K}) i^{-1/w_{i,K}} f(t, t, V_i), \\ N_{i,K} &= \gamma_i g(w_{i,K}) \frac{\log i}{w_{i,K}^2} i^{-1/w_{i,K}} f(t, t, V_i), \\ P_{i,K} &= \gamma_i g'(x_{i,K}) \left(\Gamma_i^{-1/x_{i,K}} - i^{-1/x_{i,K}} \right) f(t, t, V_i), \\ Q_{i,K} &= \gamma_i g(x_{i,K}) \left(\frac{\log \Gamma_i}{x_{i,K}^2} \Gamma_i^{-1/x_{i,K}} - \frac{\log i}{x_{i,K}^2} i^{-1/x_{i,K}} \right) f(t, t, V_i).\end{aligned}\tag{3.3.2}$$

Afin d'alléger les formules, posons, pour $i \geq 1$

$$W_i = f(t, t, V_i).\tag{3.3.3}$$

On souhaite désormais appliquer le théorème de Billingsley¹ avec

$$Y_{n,K} = \sum_{i=1}^n (M_{i,K} + N_{i,K} + P_{i,K} + Q_{i,K}),$$

et

$$Z_K = \sum_{i=1}^{\infty} (M_{i,K} + N_{i,K} + P_{i,K} + Q_{i,K}).$$

Fixons $n \in \mathbb{N}^*$. Pour $i \in \{1, \dots, n\}$, $w_{i,K} \in [\alpha_K(t), \alpha(t)]$ et $x_{i,K}$ appartient presque sûrement à l'intervalle $[\alpha_K(t), \alpha(t)]$. Alors, on a

$$w_{i,K} \xrightarrow{K \rightarrow \infty} \alpha(t) \text{ et } x_{i,K} \xrightarrow{K \rightarrow \infty} \alpha(t).$$

Par le *continuous mapping theorem*, on obtient

$$\begin{aligned} M_{i,K} &\xrightarrow{K \rightarrow \infty} M_{i,\infty} := \gamma_i g'(\alpha(t)) i^{-1/\alpha(t)} W_i, \\ N_{i,K} &\xrightarrow{K \rightarrow \infty} N_{i,\infty} := \gamma_i g(\alpha(t)) \frac{\log i}{\alpha(t)^2} i^{-1/\alpha(t)} W_i, \\ P_{i,K} &\xrightarrow{K \rightarrow \infty} P_{i,\infty} := \gamma_i g'(\alpha(t)) \left(\Gamma_i^{-1/\alpha(t)} - i^{-1/\alpha(t)} \right) W_i, \\ Q_{i,K} &\xrightarrow{K \rightarrow \infty} Q_{i,\infty} := \gamma_i g(t) \frac{1}{\alpha(t)^2} \left(\log(\Gamma_i) \Gamma_i^{-1/\alpha(t)} - \log(i) i^{-1/\alpha(t)} \right) W_i \end{aligned}$$

Alors,

$$Y_{n,K} \xrightarrow{K \rightarrow \infty} Y_n := \sum_{i=1}^n (M_{i,\infty} + N_{i,\infty} + P_{i,\infty} + Q_{i,\infty}).$$

De plus, $\sum_{i=1}^n M_{i,\infty}$, $\sum_{i=1}^n N_{i,\infty}$, $\sum_{i=1}^n P_{i,\infty}$ et $\sum_{i=1}^n Q_{i,\infty}$ sont des séries presque sûre-

1.

Théorème 3.3.2 (Thm 4.2 de [Bil68])

On suppose que les trois conditions suivantes sont vérifiées :

1. pour tout $n \geq 1$, $Y_{n,K} \xrightarrow{K \rightarrow \infty} Y_n$,
2. $Y_n \xrightarrow{n \rightarrow \infty} Y$,
3. pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P}(|Z_K - Y_{n,K}| > \varepsilon) = 0.$$

Alors, $Z_K \xrightarrow{K \rightarrow \infty} Y$.

ment convergente (voir [LGLV12]). On a ainsi

$$Y_n \xrightarrow[n \rightarrow \infty]{p.s.} \sum_{i=1}^{\infty} (M_{i,\infty} + N_{i,\infty} + P_{i,\infty} + Q_{i,\infty}).$$

Soit $\varepsilon > 0$, on veut montrer que

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P}(|Z_K - Y_{n,K}| > \varepsilon) = 0, \quad (3.3.4)$$

c'est à dire que

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} (M_{i,K} + N_{i,K} + P_{i,K} + Q_{i,K}) \right| > \varepsilon \right) = 0.$$

Si on montre que,

$$\begin{aligned} \lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} \right| > \varepsilon \right) &= \lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} \right| > \varepsilon \right) \\ &= \lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon \right) = \lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} Q_{i,K} \right| > \varepsilon \right) = 0, \end{aligned}$$

alors on obtiendra (3.3.4).

- Vérification de l'hypothèse pour $M_{i,K} = \gamma_i a'(w_{i,K}) i^{-1/w_{i,K}} W_i$.

On a

$$\mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} \right| > \varepsilon \right) \leq \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| \leq 2} \right| > \frac{\varepsilon}{2} \right) + \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \frac{\varepsilon}{2} \right).$$

D'après l'inégalité de Tchebychev pour les variables aléatoires indépendantes et centrées $(M_{i,K})_i$,

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| \leq 2} \right| > \varepsilon/2 \right) &\leq \frac{4}{\varepsilon^2} \sum_{i=n+1}^{\infty} \mathbb{E} \left(M_{i,K}^2 1_{|M_{i,K}| \leq 2} \right) \\ &\leq \frac{16}{\varepsilon^2} \sum_{i=n+1}^{\infty} \mathbb{E} \left(\left(\frac{M_{i,K}}{2} \right)^2 1_{|M_{i,K}|/2 \leq 1} \right) \\ &\leq \frac{16}{\varepsilon^2} \sum_{i=n+1}^{\infty} \int_0^1 \mathbb{P} \left(\left(\frac{M_{i,K}}{2} \right)^2 > x \right) dx \\ &\leq \frac{16}{\varepsilon^2} \int_0^1 \sum_{i=n+1}^{\infty} \mathbb{P} \left(|M_{i,K}| > 2x^{1/2} \right) dx. \end{aligned}$$

Comme $2x^{1/2} \in]0, 2[$ pour $x \in]0, 1[$, on obtient d'après le Lemme 3.5.1 que

$$\sum_{i=n+1}^{\infty} \mathbb{P} \left(|M_{i,K}| > 2x^{1/2} \right) \leq \frac{M}{2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2}} \sum_{j=\lfloor \frac{n+1}{M} 2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2} \rfloor}^{\infty} \mathbb{P} (|W_1|^{w_{i,K}} > j).$$

D'où,

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| \leq 2} \right| > \varepsilon/2 \right) \\ & \leq \frac{16}{\varepsilon^2} \int_0^1 \frac{M}{2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2}} \sum_{j=\lfloor \frac{n+1}{M} 2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2} \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) dx. \end{aligned}$$

Comme la constante M ne dépend pas de K, on obtient

$$\begin{aligned} & \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| \leq 2} \right| > \varepsilon/2 \right) \\ & \leq \frac{16M}{2^{\alpha(t)-\varepsilon_0} \varepsilon^2} \int_0^1 \frac{1}{x^{\alpha(t)/2}} \sum_{j=\lfloor \frac{n+1}{M} 2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2} \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) dx. \end{aligned}$$

Posons

$$m_n(x) = \frac{1}{x^{\alpha(t)/2}} \sum_{j=\lfloor \frac{n+1}{M} 2^{\alpha(t)-\varepsilon_0} x^{\alpha(t)/2} \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right).$$

Par (C4), on a

$$\begin{aligned} & \sum_{j=1}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) \\ & = \mathbb{E} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) \right) \\ & = \int_R \max(|f(t, t, x)|^{\alpha(t)}, |f(t, t, x)|^{\alpha(t)-\varepsilon_0}) \widehat{m}(x) dx \\ & \leq 1 + \int_R |f(t, t, x)|^{\alpha(t)} 1_{|f(t, t, x)| > 1} \widehat{m}(x) dx \\ & \leq 1 + \mathbb{E} \left(\sup_{w \in B(t, \varepsilon_0)} (|f(t, w, V_1)|^{\alpha(w)}) \right) < \infty \end{aligned}$$

et alors pour $x \in]0, 1[$,

$$m_n(x) \xrightarrow[n \rightarrow \infty]{} 0,$$

$$|m_n(x)| \leq \frac{1}{x^{\alpha(t)/2}} \left(1 + \mathbb{E} \left[\sup_{w \in B(t, \varepsilon_0)} (|f(t, w, V_1)|^{\alpha(w)}) \right] \right).$$

On a donc domination de la fonction $m_n(\cdot)$ par une fonction intégrable. En utilisant le théorème de convergence dominée, on conclut que :

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| \leq 2} \right| > \varepsilon/2 \right) = 0.$$

On s'intéresse désormais au second terme $\mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2 \right)$.

Soit $\widetilde{m}_K = \inf (n \geq 1 : \forall i \geq n, |M_{i,K}| \leq 2)$.

Comme $\sum_{i=1}^{\infty} \mathbb{P}(|M_{i,K}| > 2) < \infty$ (on peut le voir à partir du Lemme 3.5.1), le lemme de Borel-Cantelli permet de déduire que $\mathbb{P}(\widetilde{m}_K < \infty) = 1$. Alors,

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2 \right) \\ &= \sum_{m=1}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2, \widetilde{m}_K = m \right) \\ &= \sum_{m=n+1}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^m M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2, \widetilde{m}_K = m \right) \\ &\leq \sum_{m=n+1}^{\infty} \mathbb{P}(\widetilde{m}_K = m) \\ &\leq \sum_{m=n+1}^{\infty} \mathbb{P}(|M_{m-1,K}| > 2). \end{aligned}$$

En utilisant le Lemme 3.5.1, on a

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2 \right) \\ &\leq \frac{M}{2^{\alpha(t)-\varepsilon_0}} \sum_{j=\lfloor \frac{n+1}{M} 2^{\alpha(t)-\varepsilon_0} \rfloor}^{\infty} \mathbb{P} \left(\max (|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right). \end{aligned}$$

On en déduit :

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} 1_{|M_{i,K}| > 2} \right| > \varepsilon/2 \right) = 0.$$

Au final,

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} M_{i,K} \right| > \varepsilon \right) = 0.$$

- Vérification de l'hypothèse pour $N_{i,K} = \gamma_i g(w_{i,K}) \frac{\log i}{w_{i,K}^2} i^{-1/w_{i,K}} W_i$:

Soit $\varepsilon > 0$ et K tel que $1/K < \varepsilon_0$ où ε_0 apparaît dans **(C2)**. On a

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} \right| > \varepsilon \right) &\leq \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}| \leq 2} \right| > \varepsilon/2 \right) \\ &\quad + \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}| > 2} \right| > \varepsilon/2 \right). \end{aligned}$$

Par l'inégalité de Tchebychev pour les variables aléatoires indépendantes et centrées $(N_{i,K})_i$, on a

$$\mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}| \leq 2} \right| > \varepsilon/2 \right) \leq \frac{16}{\varepsilon^2} \int_0^1 \sum_{i=n+1}^{\infty} \mathbb{P} \left(|N_{i,K}| > 2x^{1/2} \right) dx.$$

Comme $2x^{1/2} \in]0, 2[$ pour $x \in]0, 1[$, on obtient par le Lemme 3.5.2 que pour tout $n \geq i_\alpha - 1$,

$$\begin{aligned} &\sum_{i=n+1}^{\infty} \mathbb{P} \left(|N_{i,K}| > 2x^{1/2} \right) \\ &\leq \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1 \log |W_1||^{\alpha(t)}, |W_1 \log |W_1||^{\alpha(t)-\varepsilon_0}) > j \right) \\ &+ \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) \\ &+ \frac{M}{\eta(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \eta(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right), \end{aligned}$$

où $\tau(x,t) = 2^{\alpha(t)-\varepsilon_0} x^{\frac{\alpha(t)}{2}}$ et $\eta(x,t) = \min \left(1, \frac{(2\sqrt{x})^{\alpha(t)-\varepsilon_0}}{|\log 2\sqrt{x}|^{\alpha(t)-\varepsilon_0}}, \frac{(2\sqrt{x})^{\alpha(t)}}{|\log 2\sqrt{x}|^{\alpha(t)}} \right)$. Comme M ne dépend pas de K , on obtient alors

$$\begin{aligned}
 & \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}| \leq 2} \right| > \varepsilon/2 \right) \\
 & \leq \frac{16}{\varepsilon^2} \int_0^1 \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1 \log |W_1||^{\alpha(t)}, |W_1 \log |W_1||^{\alpha(t)-\varepsilon_0}) > j \right) \\
 & + \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) \\
 & + \frac{M}{\eta(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \eta(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) dx.
 \end{aligned}$$

Posons

$$\begin{aligned}
 c_n(x) &= \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1 \log |W_1||^{\alpha(t)}, |W_1 \log |W_1||^{\alpha(t)-\varepsilon_0}) > j \right) \\
 & + \frac{M}{\tau(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \tau(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right) \\
 & + \frac{M}{\eta(x,t)} \sum_{j=\lfloor \frac{n+1}{M} \eta(x,t) \rfloor}^{\infty} \mathbb{P} \left(\max(|W_1|^{\alpha(t)}, |W_1|^{\alpha(t)-\varepsilon_0}) > j \right).
 \end{aligned}$$

Par **(C2)** et **(C4)**, on a pour $x \in]0, 1[$,

$$c_n(x) \xrightarrow{n \rightarrow \infty} 0,$$

$$\begin{aligned}
 |c_n(x)| &\leq \frac{M}{2^{\alpha(t)-\varepsilon_0} x^{\frac{\alpha(t)}{2}}} \left(\exp(1) + \mathbb{E} \left(|W_1 \log |W_1||^{\alpha(t)} \right) \right) \\
 &+ \frac{M}{2^{\alpha(t)-\varepsilon_0} x^{\frac{\alpha(t)}{2}}} \left(1 + \mathbb{E} \left(|W_1|^{\alpha(t)} \right) \right) \\
 &+ \frac{M}{\min \left(1, \frac{(2\sqrt{x})^{\alpha(t)-\varepsilon_0}}{|\log 2\sqrt{x}|^{\alpha(t)-\varepsilon_0}}, \frac{(2\sqrt{x})^{\alpha(t)}}{|\log 2\sqrt{x}|^{\alpha(t)}} \right)} \left(1 + \mathbb{E} \left(|W_1|^{\alpha(t)} \right) \right)
 \end{aligned}$$

Le terme de droite est une fonction intégrable, on peut donc utiliser le théorème de convergence dominée pour conclure que

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}| \leq 2} \right| > \varepsilon/2 \right) = 0.$$

On considère maintenant le deuxième terme $\mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}|>2} \right| > \varepsilon/2 \right)$.

Soit $\tilde{n}_K = \inf \{ n \geq 1 : \forall i \geq n, |N_{i,K}| \leq 2 \}$. Comme $\sum_{i=1}^{\infty} \mathbb{P}(|N_{i,K}| > 2) < \infty$ (on le voit à partir du Lemme 3.5.2), on en déduit par le lemme de Borel-Cantelli que $\mathbb{P}(\tilde{n}_K < \infty) = 1$. On a ainsi,

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}|>2} \right| > \varepsilon/2 \right) &= \sum_{m=n+1}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^m N_{i,K} 1_{|N_{i,K}|>2} \right| > \varepsilon/2, \tilde{n}_K = m \right) \\ &\leq \sum_{m=n+1}^{\infty} \mathbb{P}(\tilde{n}_K = m) \\ &\leq \sum_{m=n+1}^{\infty} \mathbb{P}(|N_{m-1,K}| > 2). \end{aligned}$$

Comme pour les $M_{i,K}$, on a par le Lemme 3.5.2,

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} 1_{|N_{i,K}|>2} \right| > \varepsilon/2 \right) = 0.$$

On en déduit

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} N_{i,K} \right| > \varepsilon \right) = 0.$$

- Vérification de l'hypothèse pour $P_{i,K} = \gamma_i g'(x_{i,K}) \left(\Gamma_i^{-1/x_{i,K}} - i^{-1/x_{i,K}} \right) W_i$.

Soit $\Omega_i = \{ \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \}$ et $\bar{\Omega}_i$ son complémentaire. On a d'après [LGLV12]

$$\sum_{i=1}^{\infty} \mathbb{P}(\bar{\Omega}_i \cup \{|P_{i,K}| > 1\}) < \infty.$$

Soit $\tilde{p}_K = \inf \left\{ n \geq 1 : \forall i \geq n, |P_{i,K}| < 1 \text{ et } \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right\}$. On utilise de nouveau le lemme de Borel-Cantelli pour obtenir que

$$\mathbb{P}(\tilde{p}_K < \infty) = 1.$$

On peut alors décomposer suivant la valeur de \widetilde{p}_K :

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon \right) \\ &= \sum_{m=1}^{n+1} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right) + \sum_{m=n+2}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right). \end{aligned}$$

Tout d'abord, on a

$$\begin{aligned} & \sum_{m=n+2}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right) \leq \sum_{m=n+2}^{\infty} \mathbb{P}(\widetilde{p}_K = m) \\ & \leq \sum_{m=n+2}^{\infty} \mathbb{P}(\{|P_{m-1,K}| > 1\} \cup \overline{\Omega_{m-1}}) \\ & \leq \sum_{m=n+2}^{\infty} \mathbb{P}(\{|P_{m-1,K}| > 1\} \cap \Omega_{m-1}) + \mathbb{P}(\overline{\Omega_{m-1}}) \\ & \leq \sum_{m=n+1}^{\infty} \mathbb{P} \left(|P_{m,K}| > 1, \frac{1}{2} \leq \frac{\Gamma_m}{m} \leq 2 \right) + \mathbb{P} \left(\frac{\Gamma_m}{m} < \frac{1}{2} \right) + \mathbb{P} \left(\frac{\Gamma_m}{m} > 2 \right). \end{aligned} \tag{3.3.5}$$

Ensuite, en utilisant le Lemme 3.5.3, on trouve que

$$\mathbb{P} \left(|P_{m,K}| > 1, \frac{1}{2} \leq \frac{\Gamma_m}{m} \leq 2 \right) \leq \begin{cases} \frac{M}{m^{\frac{1}{\zeta} + \frac{\alpha(t) - \varepsilon_0}{2\zeta}}}, \text{ avec } \alpha(t) < \zeta < 1 + \frac{\alpha(t) - \varepsilon_0}{2} & , \text{ si } \alpha(t) \geq 1, \\ \frac{M}{m^{1 + \frac{\alpha(t) - \varepsilon_0}{2}}} & , \text{ si } \alpha(t) < 1. \end{cases} \tag{3.3.6}$$

Finalement, on obtient

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \sum_{m=n+2}^{\infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right) = 0.$$

Pour $1 \leq m \leq n + 1$, on pose $g_{m,n,K} = \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right)$. Alors on a

$$\begin{aligned}
 g_{m,n,K} &\leq P \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} 1_{|P_{i,K}| \leq 1, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2} \right| > \varepsilon \right) \\
 &\leq \varepsilon \mathbb{E} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} 1_{|P_{i,K}| \leq 1, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2} \right| \right) \\
 &\leq \varepsilon \sum_{i=n+1}^{\infty} \mathbb{E} \left(|P_{i,K}| 1_{|P_{i,K}| \leq 1, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2} \right) \\
 &\leq \varepsilon \sum_{i=n+1}^{\infty} \int_0^1 \mathbb{P} \left(|P_{i,K}| > x, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) dx \\
 &\leq \varepsilon \int_0^1 \sum_{i=n+1}^{\infty} \mathbb{P} \left(|P_{i,K}| > x, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) dx.
 \end{aligned}$$

En utilisant le Lemme 3.5.3, on a alors pour $x \in]0, 1[$,

$$\sum_{i=n+1}^{\infty} \mathbb{P} \left(|P_{i,K}| > x, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \xrightarrow{n \rightarrow \infty} 0,$$

et

$$\sum_{i=n+1}^{\infty} \mathbb{P} \left(|P_{i,K}| > x, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \leq \begin{cases} \frac{\widetilde{M}}{x^{\alpha(t)/\zeta}}, & \text{avec } \zeta > \alpha(t) & \text{si } \alpha(t) \geq 1, \\ \frac{\widetilde{M}}{x^{\alpha(t)}} & & \text{si } \alpha(t) < 1, \end{cases}$$

qui sont des fonctions intégrables sur $]0, 1[$. En appliquant le théorème de convergence dominée, on obtient

$$\limsup_K |g_{m,n,K}| \xrightarrow{n \rightarrow \infty} 0.$$

Ensuite, par les équations (3.3.5) et (3.3.6), on a $|g_{m,n,K}| \leq \mathbb{P}(\widetilde{p}_K = m) \leq g_m$ avec

$$\sum_{m=1}^{\infty} g_m < \infty \text{ et donc}$$

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \sum_{m=1}^{n+1} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon, \widetilde{p}_K = m \right) \leq \lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \sum_{m=1}^{\infty} g_{m,n,K} = 0.$$

On en conclut que

$$\lim_{n \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left(\left| \sum_{i=n+1}^{\infty} P_{i,K} \right| > \varepsilon \right) = 0.$$

- Vérification de l'hypothèse pour $Q_{i,K} = \gamma_i g(x_{i,K}) \left(\frac{\log \Gamma_i}{x_{i,K}^2} \Gamma_i^{-1/x_{i,K}} - \frac{\log i}{x_{i,K}^2} i^{-1/x_{i,K}} \right) W_i$
En utilisant le Lemme 3.5.1, la preuve est la même que pour $P_{i,K}$. ■

3.4 Exemples de simulation

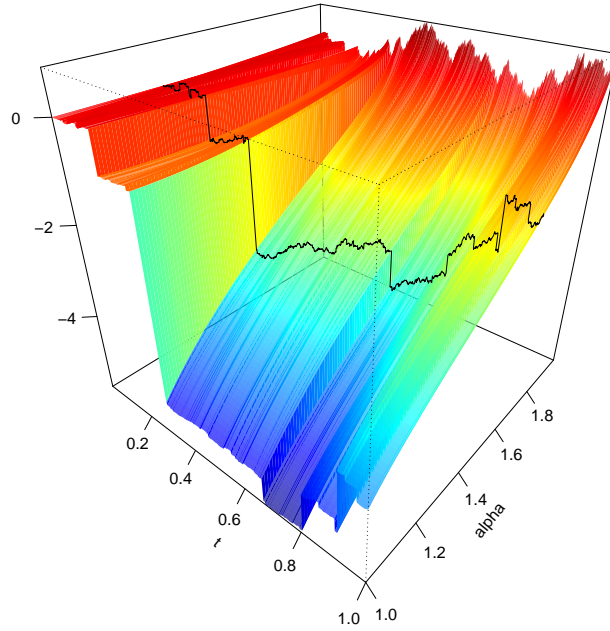
3.4.1 Mouvement de Lévy multistable

On considère $\alpha : [0, 1] \rightarrow]1, 2[$ une fonction \mathcal{C}^1 . Pour $t \in [0, 1]$, le mouvement de Lévy multistable est défini comme le processus diagonal du champ :

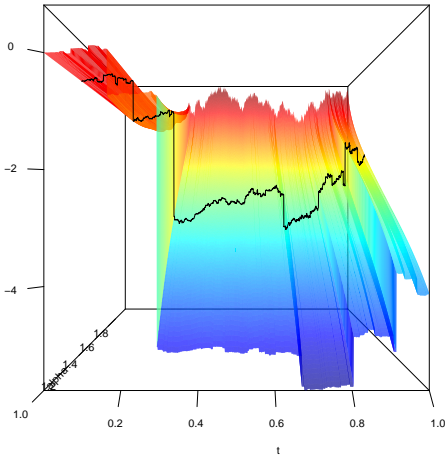
$$X(t, u) = C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} \mathbb{1}_{[0,t]}(V_i).$$

où $(\Gamma_i)_{i \geq 1}$ est une suite de temps d'arrivée d'un processus de Poisson avec un taux d'arrivée unitaire : $\Gamma_i = \sum_{j=1}^i \varepsilon_j$ avec (ε_j) i.i.d. de loi $\mathcal{E}(1)$, $(V_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0, 1]$ et $(\gamma_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$ telles que les trois suites $(\Gamma_i)_{i \geq 1}$, $(V_i)_{i \geq 1}$ et $(\gamma_i)_{i \geq 1}$ sont indépendantes.

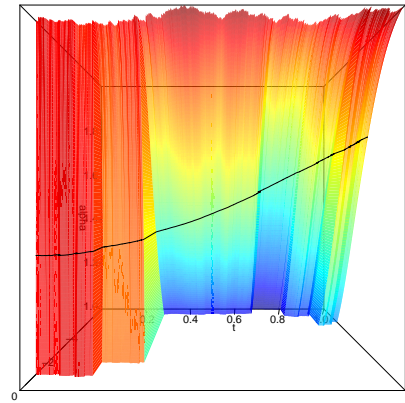
À u fixé, $X(\cdot, u)$ est un mouvement de Lévy $\alpha(u)$ -stable qui est un processus $\alpha(u)$ -stable à accroissements stationnaires et indépendants. On simule ces accroissements en utilisant le programme RSTAB (voir par exemple [ST94]).



(a)



(b)



(c)

FIGURE 3.1 – Les trois figures représentent le même champ stochastique $(\widehat{X}(t, \alpha))_{t \in [0,1], \alpha \in [1,2]}$ issu de mouvements de Lévy stables sous différents points de vue. Le mouvement de Lévy multistable est représenté sur le champ par la ligne noire $(t, \alpha(t), Y(t))$ où $\alpha(t) = \frac{2}{1+0.5 \exp(-t^2)}$.

Dans la Figure 3.2, le mouvement de Lévy multistable (en noir) est obtenu en utilisant le champ de mouvements de Lévy stables de la Figure 3.1, avec $N = 2000$. En rouge, le processus approché par une grille pour laquelle $K = 50$. À partir de $K = 100$, il n'est plus possible de discerner à cette échelle la différence entre les processus sur cet exemple.

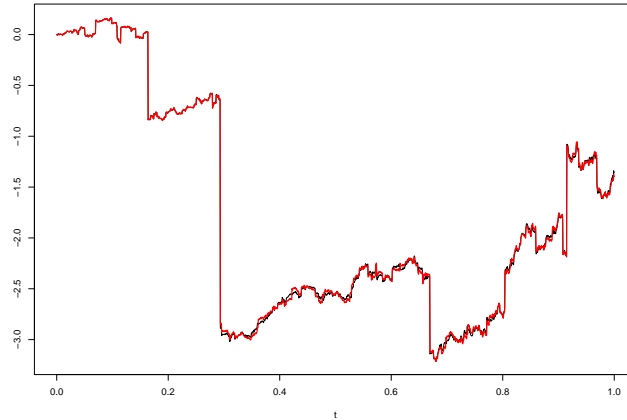


FIGURE 3.2 – Mouvement de Lévy multistable exact (noir) et approché par une grille de pas $1/50$ (rouge) avec $N = 2000$. La simulation correspond au même champ que dans la Figure 3.1.

Pour $N = 2.10^5$, la simulation grâce au champ d'un mouvement de Lévy multistable nécessite environ 6 heures pour l'ordinateur utilisé alors que son approximation à l'aide d'une grille de pas $1/K$ où $K = 2000$ est obtenue en seulement deux minutes avec une erreur en probabilité de l'ordre de $1/K$, soit une vitesse de simulation de l'ordre de 1000 fois plus rapide. Le gain de temps est donc conséquent.

3.4.2 Mouvement linéaire multifractionnaire multistable

On considère $\alpha : [0, 1] \rightarrow]0, 2[$ et $h : [0, 1] \rightarrow]0, 1[$ deux fonctions \mathcal{C}^1 . Pour $t \in [0, 1]$, le mouvement linéaire multifractionnaire multistable est défini comme le processus diagonal du champ :

$$X(t, u) = C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} \left[|t - V_i|^{h(u) - \frac{1}{\alpha(u)}} - |V_i|^{h(u) - \frac{1}{\alpha(u)}} \right],$$

où $(\Gamma_i)_{i \geq 1}$ est une suite de temps d'arrivée d'un processus de Poisson avec un taux d'arrivée unitaire : $\Gamma_i = \sum_{j=1}^i \varepsilon_j$ avec (ε_j) i.i.d. de loi $\mathcal{E}(1)$, $(V_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0, 1]$ et $(\gamma_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$ telles que les trois suites $(\Gamma_i)_{i \geq 1}$, $(V_i)_{i \geq 1}$ et $(\gamma_i)_{i \geq 1}$ sont indépendantes. À u fixé, $X(\cdot, u)$ est un mouvement linéaire $h(u)$ -fractionnaire $\alpha(u)$ -stable qui est simulé en utilisant le programme LFSN de Stoev et Taqqu [ST04].

Pour ce processus, deux fonctions apparaissent : α et h . Une méthode par le champ consiste à représenter le champ $(X(t, u))_{(t, u) \in [0, 1]^2}$ puis à prendre la diagonale. Ici pour la grille, la fonction d'intensité α va être restreinte, après approximation par la fonction α_K , aux $2K - 1$ valeurs : $\{\frac{1}{K}, \dots, \frac{2K-1}{K}\}$ (comme décrit dans la construction de la grille multistable) et la fonction de localisation h va, elle aussi, être restreinte aux valeurs $\{\frac{1}{K}, \dots, \frac{K-1}{K}\}$.

Cette double grille donne aussi des erreurs en probabilité de l'ordre de $1/K$. La présence des deux fonctions contraint à simuler davantage de processus stables. En effet, il faut dans ce cas simuler de l'ordre de $N \times K^2$ processus stables. En réalité, cet ordre maximum n'est rarement atteint puisque généralement la fonction α ne parcourt pas l'intégralité de l'intervalle $]0, 2]$ et donc il n'est pas nécessaire de simuler les $2K - 1$ possibilité pour α (de même pour h qui ne parcourt généralement pas $]0, 1[$ tout entier). Dans la Figure 3.3, on donne un exemple de simulation d'un mouvement linéaire multifractionnaire multistable. Pour simuler ce processus aux N points $(\frac{k}{N})_{k \in \{1, \dots, N\}}$ avec $N = 3496$, il a fallu 2,4 heures à l'aide du champ contre 10,3 minutes à l'aide de la double grille de pas $1/200$ (sur l'ordinateur utilisé), soit une vitesse 14 fois plus rapide.

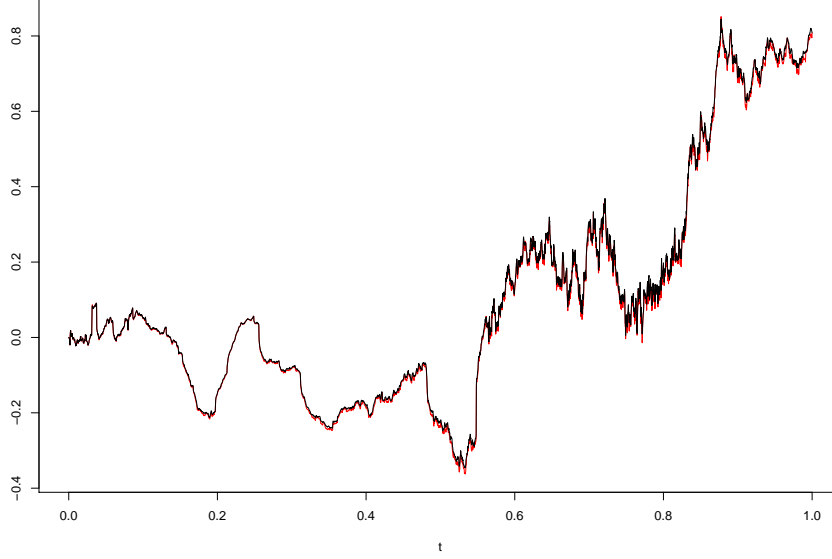


FIGURE 3.3 – Mouvement linéaire multifractionnaire multistable de fonction d'intensité $\alpha(t) = 1.41 + 0.57t$ et de fonction de localisation $h(t) = 0.725 + 0.175 \sin(2\pi t)$ par le champ (en noir) et par la double grille de pas $1/200$ (en rouge) avec $N = 3496$.

3.5 Lemmes techniques

Lemme 3.5.1

Il existe une constante $M \geq 2$ telle que pour tout $\lambda \in]0, 2]$ et $n \geq 1$,

$$\sum_{i=n+1}^{\infty} \mathbb{P}(|M_{i,K}| > \lambda) \leq \frac{M}{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})} \sum_{j=\lfloor \frac{n+1}{M} \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)}) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j)$$

où les $M_{i,K}$ sont définis en (3.3.2) et W_1 en (3.3.3).

Démonstration : Fixons $0 < \lambda \leq 2$. Alors on a :

$$\begin{aligned} \mathbb{P}(|M_{i,K}| > \lambda) &= \mathbb{P}(|g'(w_{i,K})| i^{-1/w_{i,K}} |W_i| > \lambda) \\ &\leq \mathbb{P} \left(|W_i|^{w_{i,K}} \sup_{w \in B(\alpha(t), \varepsilon_0)} (|g'(w)|^w) > i \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)}) \right). \end{aligned}$$

Comme g' est borné sur les compacts, on obtient que

$$M := \max(2, \sup_{w \in B(\alpha(t), \varepsilon_0)} (|g'(w)|^w)) < \infty.$$

Alors, $\mathbb{P}(|M_{i,K}| > \lambda) \leq \mathbb{P}\left(|W_1|^{w_{i,K}} > i \frac{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})}{M}\right)$. D'où,

$$\begin{aligned} \sum_{i=n+1}^{\infty} \mathbb{P}(|M_{i,K}| > \lambda) &\leq \sum_{i=n+1}^{\infty} \mathbb{P}\left(|W_1|^{w_{i,K}} > i \frac{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})}{M}\right) \\ &\leq \sum_{i=n+1}^{\infty} \mathbb{P}\left(|W_1|^{w_{i,K}} > \lfloor i \frac{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})}{M} \rfloor\right) \\ &\leq \sum_{j=\lfloor \frac{n+1}{M} \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)}) \rfloor}^{\infty} \lfloor \frac{M}{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})} \rfloor \mathbb{P}(|W_1|^{w_{i,K}} > j) \\ &\leq \frac{M}{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})} \sum_{j=\lfloor \frac{n+1}{M} \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)}) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j). \quad \blacksquare \end{aligned}$$

Lemme 3.5.2

Il existe une constante $M \geq 2$ telle que pour tout $\lambda \in]0, 2]$ et n suffisamment grand,

$$\begin{aligned} \sum_{i=n+1}^{\infty} \mathbb{P}(|N_{i,K}| > \lambda) &\leq \frac{M}{\ell(\lambda, t)} \sum_{j=\lfloor \frac{n+1}{M} \ell(\lambda, t) \rfloor}^{\infty} \mathbb{P}(|W_1 \log |W_1||^{w_{i,K}} > j) \\ &\quad + \frac{M}{\ell(\lambda, t)} \sum_{j=\lfloor \frac{n+1}{M} \ell(\lambda, t) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j) \\ &\quad + \frac{M}{m(\lambda, t)} \sum_{j=\lfloor \frac{n+1}{M} m(\lambda, t) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j), \end{aligned}$$

où $\ell(\lambda, t) = \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})$ et $m(\lambda, t) = \min\left(1, \frac{\lambda^{\alpha(t)-\varepsilon_0}}{|\log \lambda|^{\alpha(t)-\varepsilon_0}}, \frac{\lambda^{\alpha(t)}}{|\log \lambda|^{\alpha(t)}}\right)$. Les $N_{i,K}$ sont définis en (3.3.2) et W_1 en (3.3.3).

Démonstration : Fixons $0 < \lambda \leq 2$. Alors :

$$\begin{aligned} \mathbb{P}(|N_{i,K}| > \lambda) &= \mathbb{P}\left(|g(w_{i,K})| \frac{\log i}{w_{i,K}^2} i^{-1/w_{i,K}} |f(t, t, V_i)| > \lambda\right) \\ &\leq \mathbb{P}\left(|W_i|^{w_{i,K}} \sup_{w \in B(\alpha(t), \varepsilon_0)} \left(\left(\frac{|g(w)|}{w^2}\right)^w\right) > \frac{i}{(\log i)^{w_{i,K}}} \min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})\right). \end{aligned}$$

Posons

$$L := \max\left(2, \sup_{w \in B(\alpha(t), \varepsilon_0)} \left(\left(\frac{|g(w)|}{w^2}\right)^w\right)\right) < \infty.$$

On a ainsi

$$\mathbb{P}(|N_{i,K}| > \lambda) \leq \mathbb{P}\left(|W_i|^{w_{i,K}} \frac{L}{\min(\lambda^{\alpha(t)-\varepsilon_0}, \lambda^{\alpha(t)})} > \frac{i}{(\log i)^{w_{i,K}}}\right).$$

Soit $g_{i,K}(x) = \frac{x}{(\log x)^{w_{i,K}}}$ pour $x > 1$, $i \geq 1$ et $K \in \mathbb{N}^*$ tel que $1/K < \varepsilon_0$. Pour tout i et K , $g_{i,K}$ est strictement croissante sur l'intervalle $[8, \infty)$ et $\lim_{x \rightarrow \infty} g_{i,K}(x) = \infty$. Il existe $A \geq 8$ indépendant de i et K tel que pour tout $z \geq A$,

$$g_{i,K}(2z(\log z)^{w_{i,K}}) = 2z \left(\frac{\log(z)}{\log(z) + w_{i,K} \log \log(z)}\right)^{w_{i,K}} \geq z.$$

On a alors pour tout $z \geq A$, $i \in \mathbb{N}^*$ et $K \in \mathbb{N}^*$,

$$2z(\log z)^{w_{i,K}} \geq g_{i,K}^{-1}(z).$$

Soit $i_\alpha \in \mathbb{N}$ entier ne dépendant que de α tel que $\forall i \geq i_\alpha$, $\frac{i}{(\log i)^{\alpha(t)}} \geq A$.

Alors, pour $i \geq i_\alpha$,

$$\mathbb{P}(|N_{i,K}| > \lambda) \leq \mathbb{P}\left(|W_i|^{w_{i,K}} \frac{L}{\ell(\lambda, t)} > \frac{i}{(\log i)^{w_{i,K}}}\right)$$

et $g_{i,K}(i) = \frac{i}{(\log i)^{w_{i,K}}} \geq \frac{i}{(\log i)^{\alpha(t)}} \geq A$.

Donc,

$$\begin{aligned}
 & \mathbb{P}(|N_{i,K}| > \lambda) \\
 & \leq \mathbb{P}\left(g_{i,K}^{-1}\left(|W_i|^{w_{i,K}} \frac{L}{\ell(\lambda,t)}\right) > i\right) \\
 & \leq \mathbb{P}\left(2|W_i|^{w_{i,K}} \frac{L}{v} \left|\log\left(|W_i|^{w_{i,K}} \frac{L}{\ell(\lambda,t)}\right)\right|^{w_{i,K}} > i\right) \\
 & \leq \mathbb{P}\left(\frac{2L|W_i|^{w_{i,K}}}{\ell(\lambda,t)} |\log(|W_i|^{w_{i,K}}) + \log L - \log(\ell(\lambda,t))|^{w_{i,K}} > i\right) \\
 & \leq \mathbb{P}\left(\frac{6L \max(1, 3^{w_{i,K}-1})}{\ell(\lambda,t)} |W_i|^{w_{i,K}} |\log(|W_i|^{w_{i,K}})|^{w_{i,K}} > i\right) \\
 & \quad + \mathbb{P}\left(\frac{6L \max(1, 3^{w_{i,K}-1})}{\ell(\lambda,t)} |W_i|^{w_{i,K}} (\log L)^{w_{i,K}} > i\right) \\
 & \quad + \mathbb{P}\left(\frac{6L \max(1, 3^{w_{i,K}-1})}{\ell(\lambda,t)} |W_i|^{w_{i,K}} |\log(\ell(\lambda,t))|^{w_{i,K}} > i\right) \\
 & \leq \mathbb{P}\left(|W_1 \log |W_1||^{w_{i,K}} > \frac{i}{K_1} \ell(\lambda,t)\right) \\
 & \quad + \mathbb{P}\left(|W_1|^{w_{i,K}} > \frac{i}{K_2} \ell(\lambda,t)\right) \\
 & \quad + \mathbb{P}\left(|W_1|^{w_{i,K}} > \frac{i}{K_3} \min\left(1, \left(\frac{\lambda}{|\log \lambda|}\right)^{\alpha(t)-\varepsilon_0}, \left(\frac{\lambda}{|\log \lambda|}\right)^{\alpha(t)}\right)\right),
 \end{aligned}$$

où $K_1 \geq 2$, $K_2 \geq 2$ et $K_3 \geq 2$ sont des constantes qui ne dépendent pas de i ni de K . Soient $M = \max(K_1, K_2, K_3) \geq 2$ et $n \geq i_\alpha - 1$, on a alors

$$\begin{aligned}
 & \sum_{i=n+1}^{\infty} \mathbb{P}(|N_{i,K}| > \lambda) \\
 & \leq \frac{M}{\ell(\lambda,t)} \sum_{j=\lfloor \frac{n+1}{M} \ell(\lambda,t) \rfloor}^{\infty} \mathbb{P}(|W_1 \log |W_1||^{w_{i,K}} > j) \\
 & \quad + \frac{M}{\ell(\lambda,t)} \sum_{j=\lfloor \frac{n+1}{M} \ell(\lambda,t) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j) \\
 & \quad + \frac{M}{m(\lambda,t)} \sum_{j=\lfloor \frac{n+1}{M} m(\lambda,t) \rfloor}^{\infty} \mathbb{P}(|W_1|^{w_{i,K}} > j).
 \end{aligned}$$

■

Lemme 3.5.3

Il existe une constante $M > 0$ telle que pour tout $\lambda > 0$,

$$\mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \leq \begin{cases} \frac{M}{i^{\frac{1}{\zeta} + \frac{\alpha(t) - \varepsilon_0}{2\zeta}}} s_\zeta(\lambda, t) & \text{si } \alpha(t) \geq 1, \\ \frac{M}{i^{1 + \frac{\alpha(t) - \varepsilon_0}{2}}} s_1(\lambda, t) & \text{si } \alpha(t) < 1, \end{cases}$$

où ζ un réel tel que $\alpha(t) < \zeta < 1 + \frac{\alpha(t) - \varepsilon_0}{2}$ et $s_\zeta(\lambda, t) = \min(\lambda^{\frac{\alpha(t) - \varepsilon_0}{\zeta}}, \lambda^{\frac{\alpha(t)}{\zeta}})$. Les $P_{i,K}$ sont définis en (3.3.2).

Démonstration : On a

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ &= \mathbb{P} \left(|g'(x_{i,K})| i^{-1/x_{i,K}} \left(\frac{\Gamma_i^{-1/x_{i,K}}}{i^{-1/x_{i,K}}} - 1 \right) |W_i| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right). \end{aligned}$$

On applique alors la formule des accroissements finis à la fonction $y \mapsto y^{-1/x_{i,K}}$ sur $[\frac{1}{2}, 2]$,

$$\left| \left(\frac{\Gamma_i}{i} \right)^{-1/x_{i,K}} - 1 \right| \leq \left| \frac{\Gamma_i}{i} - 1 \right| \frac{1}{\alpha(t) - \varepsilon_0} \left(\frac{1}{2} \right)^{-1-1/(\alpha(t) - \varepsilon_0)}.$$

Soit $K_\alpha = \max \left(2, \sup_{w \in B(\alpha(t), \varepsilon_0)} \left(\left| \frac{g'(w) 2^{1+\alpha(t) - \varepsilon_0}}{\alpha(t) - \varepsilon_0} \right|^w \right) \right)$ et alors

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \mathbb{P} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^{\alpha(t) - \varepsilon_0} |W_i|^{x_{i,K}} > \frac{i}{K_\alpha} \min(\lambda^{\alpha(t) - \varepsilon_0}, \lambda^{\alpha(t)}), \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \mathbb{P} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^{\alpha(t) - \varepsilon_0} \max(|W_i|^{\alpha(t)}, |W_i|^{\alpha(t) - \varepsilon_0}) > \frac{i}{K_\alpha} \min(\lambda^{\alpha(t) - \varepsilon_0}, \lambda^{\alpha(t)}) \right). \end{aligned}$$

• Si $\alpha(t) \geq 1$:

on peut supposer ε_0 suffisamment petit pour avoir $\alpha(t) < 1 + (\alpha(t) - \varepsilon_0)/2$. Fixons alors

$$\zeta \in (\alpha(t), 1 + (\alpha(t) - \varepsilon_0)/2).$$

On a

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \mathbb{P} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^{\frac{\alpha(t)-\varepsilon_0}{\zeta}} \left[\max(|W_i|^{\alpha(t)}, |W_i|^{\alpha(t)-\varepsilon_0}) \right]^{\frac{1}{\zeta}} > \frac{i^{1/\zeta}}{K_\alpha^{1/\zeta}} s_\zeta(\lambda, t) \right). \end{aligned}$$

Par l'inégalité de Markov et l'indépendance de V_1 et Γ_i :

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \frac{K_\alpha^{1/\zeta}}{i^{1/\zeta} s_\zeta(\lambda, t)} \mathbb{E} \left(\left[\max(|W_i|^{\alpha(t)}, |W_i|^{\alpha(t)-\varepsilon_0}) \right]^{1/\zeta} \right) \mathbb{E} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^{(\alpha(t)-\varepsilon_0)/\zeta} \right). \end{aligned}$$

Puis par l'inégalité d'Hölder, on obtient

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \frac{K_\alpha^{1/\zeta}}{i^{1/\zeta} s_\zeta(\lambda, t)} \mathbb{E} \left(\max(|W_i|^{\alpha(t)}, |W_i|^{\alpha(t)-\varepsilon_0}) \right)^{1/\zeta} \mathbb{E} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^2 \right)^{(\alpha(t)-\varepsilon_0)/2\zeta} \\ & = \frac{\widetilde{K}}{i^{1/\zeta} s_\zeta(\lambda, t)} \frac{1}{i^{(\alpha(t)-\varepsilon_0)/\zeta}} \mathbb{E} \left(|\Gamma_i - i|^2 \right)^{(\alpha(t)-\varepsilon_0)/2\zeta} \\ & = \frac{\widetilde{K}}{i^{1/\zeta + (\alpha(t)-\varepsilon_0)/2\zeta} s_\zeta(\lambda, t)} \end{aligned}$$

puisque $\mathbb{E}(\Gamma_i) = i$ et $\text{Var}(\Gamma_i) = i$.

• Si $\alpha(t) < 1$, on a :

$$\begin{aligned} & \mathbb{P} \left(|P_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2 \right) \\ & \leq \frac{K_\alpha}{i s_1(\lambda, t)} \mathbb{E} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^{\alpha(t)-\varepsilon_0} \right) \mathbb{E} \left(\max(|W_1|^{\alpha(t)-\varepsilon_0}, |W_1|^{\alpha(t)}) \right) \\ & \leq \frac{\hat{K}}{i s_1(\lambda, t)} \mathbb{E} \left(\left| \frac{\Gamma_i}{i} - 1 \right|^2 \right)^{\frac{\alpha(t)-\varepsilon_0}{2}} \\ & \leq \frac{\hat{K}}{i^{1 + \frac{\alpha(t)-\varepsilon_0}{2}} s_1(\lambda, t)}. \end{aligned}$$

■

Lemma 3.5.1

Il existe une constante $M > 0$ telle que pour tout $\lambda > 0$,

$$\mathbb{P}\left(|Q_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2\right) \leq \begin{cases} \frac{M}{s_\zeta(\lambda, t)} \frac{(\log i)^{\alpha(t)/\zeta}}{i^{1/\zeta + (\alpha(t) - \varepsilon_0)/2\zeta}} & \text{si } \alpha(t) \geq 1, \\ \frac{M}{s_1(\lambda, t)} \frac{(\log i)^{\alpha(t)}}{i^{1 + (\alpha(t) - \varepsilon_0)/2}} & \text{si } \alpha(t) < 1, \end{cases}$$

où ζ est un réel tel que $\alpha(t) < \zeta < 1 + \frac{\alpha(t) - \varepsilon_0}{2}$, $s_\zeta(\lambda, t) = \min(\lambda^{\frac{\alpha(t) - \varepsilon_0}{\zeta}}, \lambda^{\frac{\alpha(t)}{\zeta}})$ et $Q_{i,K}$ est défini en (3.3.2).

Démonstration : On a

$$\begin{aligned} & \mathbb{P}\left(|Q_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2\right) \\ &= \mathbb{P}\left(\left|\frac{g(x_{i,K})}{x_{i,K}^2}\right| i^{-1/x_{i,K}} \log i \left|\frac{\log \Gamma_i}{\log i} \left(\frac{\Gamma_i}{i}\right)^{-1/x_{i,K}} - 1\right| |W_i| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2\right). \end{aligned}$$

On a alors par la formule des accroissements finis appliquée aux fonctions $y \mapsto \log(y)y^{-1/x_{i,K}}$ et $y \mapsto y^{-1/x_{i,K}}$ sur $[\frac{1}{2}, 2]$,

$$\begin{aligned} & \left|\frac{\log \Gamma_i}{\log i} \left(\frac{\Gamma_i}{i}\right)^{-1/\alpha(x_{i,K})} - 1\right| \\ & \leq \left|\frac{\log \frac{\Gamma_i}{i}}{\log i} \left(\frac{\Gamma_i}{i}\right)^{-1/x_{i,K}}\right| + \left|\left(\frac{\Gamma_i}{i}\right)^{-1/x_{i,K}} - 1\right| \\ & \leq \frac{1}{\log i} \left(1 + \frac{\log 2}{\alpha(t) - \varepsilon_0}\right) \left(\frac{1}{2}\right)^{-1 - \frac{1}{\alpha(t) - \varepsilon_0}} \left|\frac{\Gamma_i}{i} - 1\right| + \frac{1}{\alpha(t) - \varepsilon_0} \left(\frac{1}{2}\right)^{-1 - \frac{1}{\alpha(t) - \varepsilon_0}} \left|\frac{\Gamma_i}{i} - 1\right| \\ & \leq L \left|\frac{\Gamma_i}{i} - 1\right|. \end{aligned}$$

Soit $K_\alpha := \max\left(2, \sup_{w \in B(\alpha(t), \varepsilon_0)} \left(\left|\frac{Lg(w)}{w^2}\right|^w\right)\right)$ et alors :

$$\begin{aligned} & \mathbb{P}\left(|Q_{i,K}| > \lambda, \frac{1}{2} \leq \frac{\Gamma_i}{i} \leq 2\right) \\ & \leq \mathbb{P}\left(\left|\frac{\Gamma_i}{i} - 1\right|^{\alpha(t) - \varepsilon_0} \max(|W_1|^{\alpha(t) - \varepsilon_0}, |W_1|^{\alpha(t)}) > \frac{i}{K_\alpha (\log i)^{\alpha(t)}} s_1(\lambda, t)\right). \end{aligned}$$

La fin de la preuve est la même que pour le Lemme 3.5.3. ■



4

Estimation de la fonction d'intensité de processus multistables

Sommaire

4.1	Introduction	115
4.2	Convergence des moments empiriques de processus multistables autour d'un point	117
4.3	Estimation de la fonction d'intensité du processus multistable	122
4.3.1	Exemple d'estimation avec des mouvements de Lévy multistables	125
4.3.2	Exemple d'estimation avec des mouvements linéaires multi-fractionnaires multistables	127

4.1 Introduction

Dans ce chapitre, nous considérons les processus multistables étudiés dans le chapitre 3 qui ont été définis dans [LGLV12]. Après avoir donné une méthode de simulation rapide de tels processus, nous souhaitons désormais construire un estimateur de la fonction d'intensité.

Nous rappelons les notations utilisées dans le cadre de processus multistables. Soit l'espace de mesure σ -finie $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$ et $r : \mathbb{R} \rightarrow \mathbb{R}^+$ telle que la mesure \widehat{m} définie par $\widehat{m}(dx) = \frac{1}{r(x)}dx$ soit une mesure de probabilité. On se donne α une fonction de classe C^1 définie sur \mathbb{R} et à valeurs dans $]0, 2[$ qui est la fonction d'intensité du processus et $f(t, u, \cdot)$ une famille de fonctions telles que, pour tout $(t, u) \in \mathbb{R}^2$, $f(t, u, \cdot) \in \mathcal{F}_{\alpha(u)}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \text{Leb})$.

On considère le champ stochastique $(X(t, u))_{(t,u) \in \mathbb{R}^2}$ construit à partir de la fonction d'intensité $\alpha \in C^1(\mathbb{R},]0, 2[)$ et de la famille de fonctions $f(t, u, \cdot)$ et qui est défini dans (3.2.3) par

$$X(t, u) = C_{\alpha(u)}^{1/\alpha(u)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(u)} r(V_i)^{1/\alpha(u)} f(t, u, V_i).$$

Le processus multistable Y est défini dans (3.2.4) comme processus diagonal de ce champ stochastique, c'est à dire $Y(t) = X(t, t)$.

Soit U un intervalle ouvert de \mathbb{R} et $t_0 \in U$ fixé. On souhaite estimer $\alpha(t_0)$ qui correspond à la valeur de la fonction d'intensité au point t_0 . Pour cela, on considère $N \geq 1$ et $n(N)$ une suite d'entiers pairs et on pose alors l'intervalle d'entiers

$$I_N(t_0) = \left\{ \lfloor Nt_0 \rfloor - \frac{n(N)}{2}, \dots, \lfloor Nt_0 \rfloor + \frac{n(N)}{2} - 1 \right\}.$$

Après renormalisation (par N^{-1}), cet intervalle va correspondre à une fenêtre d'observation autour du point t_0 . On définit les moments empiriques $(S_N(p))_{p>0}$ du processus Y autour du point t_0 par

$$S_N(p) = \left(\frac{1}{n(N)} \sum_{k \in I_N(t_0)} \left| Y \left(\frac{k+1}{N} \right) - Y \left(\frac{k}{N} \right) \right|^p \right)^{1/p}. \quad (4.1.1)$$

Dans la Section 4.2, nous montrons (après normalisation) la convergence en probabilité des moments d'ordre p lorsque $0 < p < \alpha(t_0)$ et la convergence en loi de ces mêmes moments vers une loi stable lorsque $p > \alpha(t_0)$. La Section 4.3 est consacrée à l'étude d'un estimateur de la fonction d'intensité basé sur les moments empiriques du processus dont nous montrons les performances sur des exemples de mouvements de Lévy multistables et de mouvements linéaires multifractionnaires multistables.

4.2 Convergence des moments empiriques de processus multistables autour d'un point

Théorème 4.2.1 ([LG13] Théorème 5.4)

Soit $c = \min_{u \in U} \alpha(u)$ et $d = \max_{u \in U} \alpha(u)$. On suppose que :

- **(R1)** $v \rightarrow f(u, v, x)$ est dérivable sur U pour tout $u \in U$ et pour presque tout $x \in \mathbb{R}$ (cette dérivée sera notée f'_2),

- **(M1)** il existe $\delta > \frac{d}{c} - 1$ tel que :

$$\sup_{u \in U} \int_{\mathbb{R}} \left[\sup_{w \in U} (|f(u, w, x)|^{\alpha(w)}) \right]^{1+\delta} r(x)^\delta dx < \infty,$$

- **(M2)** il existe $\delta > \frac{d}{c} - 1$ tel que :

$$\sup_{u \in U} \int_{\mathbb{R}} \left[\sup_{w \in U} (|f'_2(u, w, x)|^{\alpha(w)}) \right]^{1+\delta} r(x)^\delta dx < \infty,$$

- **(M3)** il existe $\delta > \frac{d}{c} - 1$ tel que :

$$\sup_{u \in U} \int_{\mathbb{R}} \left[\sup_{w \in U} (|f(u, w, x) \log |r(x)||^{\alpha(w)}) \right]^{1+\delta} r(x)^\delta dx < \infty,$$

- le processus $X(\cdot, t_0)$ est $H(t_0)$ -auto-similaire avec $H(t_0) < 1$ et à accroissements stationnaires,

- en posant $h_{j,t_0}(x) = f(j+1, t_0, x) - f(j, t_0, x)$, on a

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}} |h_{0,t_0}(x) h_{j,t_0}(x)|^{\frac{\alpha(t_0)}{2}} dx = 0,$$

- $n(N) \xrightarrow{N \rightarrow \infty} \infty$ et $\frac{n(N)}{N} \xrightarrow{N \rightarrow \infty} 0$,

Alors, pour tout $p \in]0, \alpha(t_0)[$,

$$N^{H(t_0)} S_N(p) \xrightarrow[N \rightarrow \infty]{\text{proba}} \mathbb{E} [|X(1, t_0)|^p]^{1/p}.$$

À t_0 fixé, par la construction du champ, le processus $X(\cdot, t_0)$ est un processus $\alpha(t_0)$ -stable. En particulier, la variable $X(1, t_0)$ suit la loi $S_{\alpha(t_0)}(\|f(1, t_0, \cdot)\|_{\alpha(t_0)}, 0, 0)$.

En utilisant les propriétés des lois stables (voir (1.1.5)), on a

$$\mathbb{E} [|X(1, t_0)|^p]^{1/p} = \left(\frac{(1-p)\Gamma(1 - \frac{p}{\alpha(t_0)})}{\Gamma(2-p) \cos(\frac{\pi p}{2})} \right)^{1/p} \|f(1, t_0, \cdot)\|_{\alpha(t_0)}. \quad (4.2.1)$$

On illustre ce théorème sur les exemples du mouvement de Lévy multistable et du mouvement linéaire multifractionnaire multistable dans les Figures 4.1 et 4.2. On remarque numériquement que la convergence en probabilité de $N^{H(t_0)}S_N(p)$ est plus rapide lorsque $p \in]0, \alpha(t_0)[$ s'éloigne de la valeur $\alpha(t_0)$.

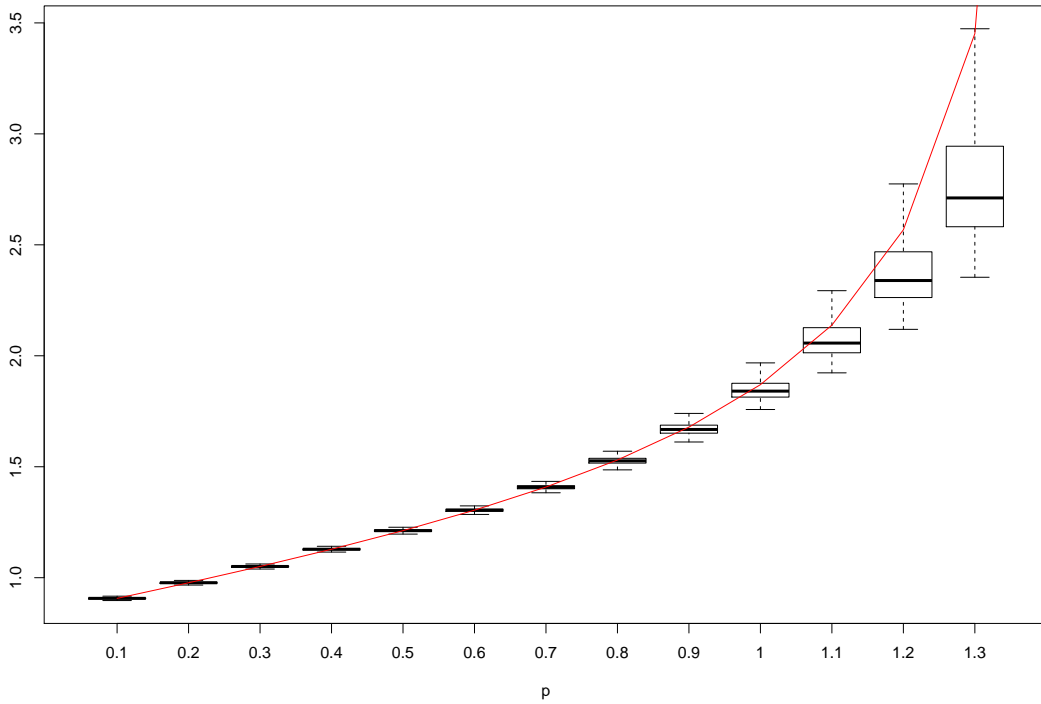


FIGURE 4.1 – Convergence des moments empiriques (en noir, boîtes à moustaches pour 1000 itérations avec $N = 10^5$ points) du mouvement de Lévy multistable de fonction d'intensité $\alpha(t) = \frac{2}{1+0.5 \exp(-t^2)}$ autour du point $t_0 = 0.5$ ($\alpha(t_0) \simeq 1.44$) vers la courbe théorique $p \rightarrow \mathbb{E} [|X(1, t_0)|^p]^{1/p}$ (en rouge).

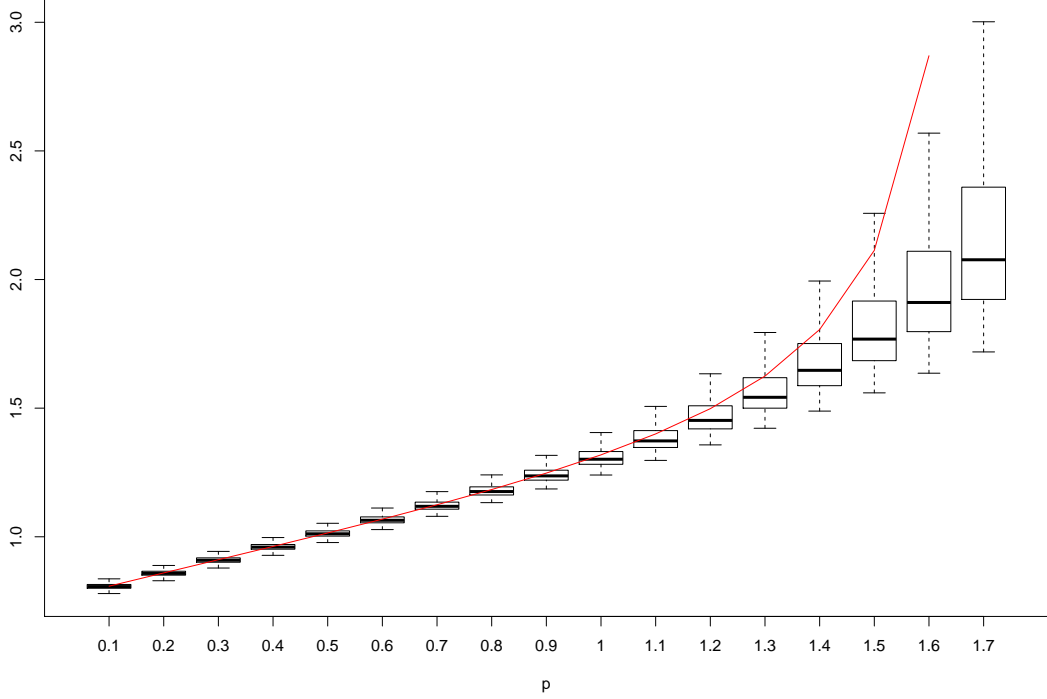


FIGURE 4.2 – Convergence des moments empiriques (en noir, boîtes à moustaches pour 1000 itérations avec $N = 15784$ points) du mouvement linéaire multifractionnaire multistable de fonction d'intensité $\alpha(t) = 1.41 + 0.57t$ et de fonction de localisation $h(t) = 0.725 + 0.175 \sin(2\pi t)$ autour du point $t_0 = 0.5$ ($\alpha(t_0) = 1.695$) vers la courbe théorique $p \rightarrow \mathbb{E}[|X(1, t_0)|^p]^{1/p}$ (en rouge).

On définit ensuite les moments empiriques autour de t_0 du processus $X(\cdot, t_0)$ qui est $\alpha(t_0)$ -stable,

$$\widetilde{S}_N(p) = \left(\frac{1}{n(N)} \sum_{k \in I_N(t_0)} \left| X\left(\frac{k+1}{N}, t_0\right) - X\left(\frac{k}{N}, t_0\right) \right|^p \right)^{1/p}.$$

Lorsqu'on est proche de t_0 , les processus $X(\cdot, t_0)$ et Y vont avoir des propriétés semblables. Le but est donc de décrire les propriétés du processus diagonal Y à partir du processus $X(\cdot, t_0)$ dont l'étude est plus facile.

Théorème 4.2.2

Soit $c = \min_{u \in U} \alpha(u)$ et $d = \max_{u \in U} \alpha(u)$. On suppose que :

- les conditions **(R1)**, **(M1)**, **(M2)** et **(M3)** définies dans le Théorème 4.2.1

sont vérifiées,

- $X(\cdot, t_0)$ est $H(t_0)$ -auto-similaire avec $H(t_0) < 1$ et à accroissements stationnaires,
- $\lim_{j \rightarrow \infty} \int_{\mathbb{R}} |h_{0,t_0}(x)h_{j,t_0}(x)|^{\frac{\alpha(t_0)}{2}} m(dx) = 0$, où $h_{j,t_0}(x) = f(j+1, t_0, x) - f(j, t_0, x)$,
- $n(N) = N^\delta$ avec $0 < \delta < c \frac{1-H(t_0)}{c+1}$,
- pour tout $p \in]\alpha(t_0), 2[$,

$$N^{H(t_0)} n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} \widetilde{S}_N(p) \xrightarrow[N \rightarrow \infty]{Loi} W^{1/p},$$

avec $W \sim S_{\alpha(t_0)/p}(\sigma_{p,\alpha(t_0)}, 1, 0)$.

Alors, pour tout $p \in]\alpha(t_0), 2[$,

$$N^{H(t_0)} n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} S_N(p) \xrightarrow[N \rightarrow \infty]{Loi} W^{1/p}. \quad (4.2.2)$$

Démonstration : Soit $p \in]\alpha(t_0), 2[$, on cherche à montrer que

$$N^{H(t_0)} n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} (S_N(p) - \widetilde{S}_N(p)) \xrightarrow[N \rightarrow \infty]{proba} 0 \quad (4.2.3)$$

afin de conclure par le lemme de Slutsky qu'on a bien la convergence (4.2.2) pour $S_N(p)$. Pour obtenir (4.2.3), il suffit de montrer que $A_N(p) \xrightarrow[N \rightarrow \infty]{proba} 0$ et

$B_N(p) \xrightarrow[N \rightarrow \infty]{proba} 0$ où

$$A_N(p) = N^{pH(t_0)} n(N)^{-p/\alpha(t_0)} \sum_{k \in I_N(t_0)} \left| X\left(\frac{k+1}{N}, \frac{k+1}{N}\right) - X\left(\frac{k+1}{N}, t_0\right) \right|^p,$$

$$B_N(p) = N^{pH(t_0)} n(N)^{-p/\alpha(t_0)} \sum_{k \in I_N(t_0)} \left| X\left(\frac{k}{N}, \frac{k}{N}\right) - X\left(\frac{k}{N}, t_0\right) \right|^p.$$

Soit $x > 0$,

$$\begin{aligned} \mathbb{P}(A_N(p) > x) &= \mathbb{P}\left(N^{pH(t_0)-p\delta/\alpha(t_0)} \sum_{k \in I_N(t_0)} \left| X\left(\frac{k+1}{N}, \frac{k+1}{N}\right) - X\left(\frac{k+1}{N}, t_0\right) \right|^p > x\right) \\ &\leq \sum_{k \in I_N(t_0)} \mathbb{P}\left(\left| X\left(\frac{k+1}{N}, \frac{k+1}{N}\right) - X\left(\frac{k+1}{N}, t_0\right) \right| > \frac{x^{1/p}}{N^{\delta/p+H(t_0)-\delta/\alpha(t_0)}}\right). \end{aligned}$$

À l'aide des conditions **(R1)**, **(M1)**, **(M2)** et **(M3)**, on peut appliquer la Proposition 4.9 de [LGLV13] : il existe $K_U > 0$ tel que pour N suffisamment grand et

pour tout $k \in I_N(t_0)$,

$$\begin{aligned} \mathbb{P}(A_N(p) > x) &\leq N^\delta K_U \frac{N^{d[\delta-1+H(t_0)-\frac{\delta}{\alpha(t_0)}+\frac{\delta}{p}]} }{x^{d/p}} \left(1 + \left| \log \frac{N^{\delta-1+H(t_0)-\frac{\delta}{\alpha(t_0)}+\frac{\delta}{p}}}{x^{1/p}} \right|^d \right) \\ &+ N^\delta K_U \frac{N^{c[\delta-1+H(t_0)-\frac{\delta}{\alpha(t_0)}+\frac{\delta}{p}]} }{x^{c/p}} \left(1 + \left| \log \frac{N^{\delta-1+H(t_0)-\frac{\delta}{\alpha(t_0)}+\frac{\delta}{p}}}{x^{1/p}} \right|^c \right). \end{aligned}$$

Ainsi, $\mathbb{P}(A_N(p) > x) \xrightarrow[N \rightarrow \infty]{} 0$ lorsque $\alpha(t_0) < p$ et $0 < \delta < c \frac{1-H(t_0)}{c+1}$. La preuve est la même pour $B_N(p)$. ■

Sous l'hypothèse que le processus $X(\cdot, t_0)$ est $H(t_0)$ -auto-similaire et à accroissements stationnaires, la dernière hypothèse du théorème est un résultat de convergence en loi d'une somme normalisée de variables aléatoires stables de même loi. En effet, il faut montrer que

$$n(N)^{\frac{-1}{\alpha(t_0)/p}} \sum_{k \in I_N(t_0)} \left| N^{H(t_0)} \left(X \left(\frac{k+1}{N}, t_0 \right) - X \left(\frac{k}{N}, t_0 \right) \right) \right|^p \xrightarrow[N \rightarrow \infty]{Loi} W$$

où pour tout $k \in I_N(t_0)$, on a l'égalité en loi suivante :

$$N^{H(t_0)} \left(X \left(\frac{k+1}{N}, t_0 \right) - X \left(\frac{k}{N}, t_0 \right) \right) \stackrel{Loi}{=} X(1, t_0).$$

Pour le processus de Lévy multistable, les variables aléatoires

$$\left(N^{H(t_0)} \left(X \left(\frac{k+1}{N}, t_0 \right) - X \left(\frac{k}{N}, t_0 \right) \right) \right)_{k \in I_N(t_0)}$$

sont i.i.d $\alpha(t_0)$ -stable symétriques avec un paramètre d'échelle unitaire (car associé à la fonction $f(1, t_0, x) = \mathbb{1}_{[0,1]}(x)$ qui vérifie $\|f(1, t_0, \cdot)\|_{\alpha(t_0)} = 1$).

Alors,

$$\begin{cases} \lambda^{\frac{\alpha(t_0)}{p}} \mathbb{P}(|X(1, t_0)|^p > \lambda) \xrightarrow[\lambda \rightarrow \infty]{} C_{\alpha(t_0)}, \\ \lambda^{\frac{\alpha(t_0)}{p}} \mathbb{P}(|X(1, t_0)|^p < -\lambda) \xrightarrow[\lambda \rightarrow \infty]{} 0. \end{cases}$$

Le cas d'une convergence de la somme de variables i.i.d. vers une loi stable a été étudiée notamment dans les travaux de [GK68], [Fel71] et [Pet95]. Une condition de *tail balance* pour la loi des variables est nécessaire et suffisante afin d'obtenir

la convergence vers une loi stable. Plus précisément, en utilisant le Théorème 5 du chapitre 7 de [GK68], la loi de $N^{H(t_0)} \left(X \left(\frac{k+1}{N}, t_0 \right) - X \left(\frac{k}{N}, t_0 \right) \right)$ appartient au domaine d'attraction d'une loi stable $S_{\frac{\alpha(t_0)}{p}}(\tilde{\sigma}, \tilde{\beta}, 0)$ et on a les relations :

$$\begin{cases} C_{\alpha(t_0)} = \frac{1+\tilde{\beta}}{2} C_{\alpha(t_0)/p} \tilde{\sigma}^{\alpha(t_0)/p}, \\ 0 = \frac{1-\tilde{\beta}}{2} C_{\alpha(t_0)} \tilde{\sigma}^{\alpha(t_0)/p}. \end{cases}$$

D'où, $\tilde{\beta} = 1$ et $\tilde{\sigma} = \left(\frac{C_{\alpha(t_0)}}{C_{\alpha(t_0)/p}} \right)^{p/\alpha(t_0)} \|f(1, t_0, \cdot)\|_{\alpha(t_0)}^p$. La dernière hypothèse du théorème est donc vérifiée pour les mouvements de Lévy multistables avec pour paramètre d'échelle $\sigma_{p, \alpha(t_0)} = \left(\frac{C_{\alpha(t_0)}}{C_{\alpha(t_0)/p}} \right)^{p/\alpha(t_0)}$. Les autres hypothèses du théorème ont été vérifiées dans le cadre de mouvement de Lévy multistables lorsque $\alpha \in \mathcal{C}^1([0, 1],]1, 2[)$ dans [LG13].

Dans le cas de variables aléatoires dépendantes, cette hypothèse est plus compliquée à vérifier puisque les conditions pour obtenir une convergence d'une somme de variables aléatoires dépendantes vers une loi stable sont complexes. Il nous faut dans le cas dépendant des conditions sur la convergence des lois fini-dimensionnelles vers une mesure de probabilité à variation régulière d'indice $\alpha \in]0, 2[$, des conditions de mélange et des conditions anti-groupage (voir [DH95], [Bar+11] ou [LeP+81]). Il faudrait pouvoir exprimer ces hypothèses à l'aide de conditions que doit vérifier la famille de fonctions $f(\cdot, \cdot, \cdot)$ qui engendre le processus multistable.

4.3 Estimation de la fonction d'intensité du processus multistable

Dans [LG13], un premier estimateur de la fonction d'intensité d'un processus multistable est proposé. Pour cet estimateur, Le Guével propose de considérer les moments empiriques $S_N(p)$ d'ordre p autour du point d'étude t_0 tels que définis dans (4.1.1). À partir de deux paramètres $p_0 \in]0, \min_{u \in U} \alpha(u)[$ et $\gamma \in]0, 1[$ qu'il faut fixer, il définit pour $p \in [p_0, 2]$ les rapports

$$\tilde{R}_N(p) = \frac{S_N(p_0)}{S_N(p)}, \quad \text{et} \quad R_\alpha(p) = \frac{\mathbb{E}[|Z|^{p_0}]^{1/p_0}}{\mathbb{E}[|Z|^p]^{1/p}} \mathbb{1}_{p < \alpha}, \quad \text{où } Z \sim S_\alpha(1, 0, 0).$$

Comme la fonction $\alpha \rightarrow \left(\int_{p_0}^2 |\tilde{R}_N(p) - R_\alpha(p)|^\gamma dp \right)^{1/\gamma}$ est une fonction continue, l'ensemble $\operatorname{argmin}_{\alpha \in [0,2]} \left(\int_{p_0}^2 |\tilde{R}_N(p) - R_\alpha(p)|^\gamma dp \right)^{1/\gamma}$ est un fermé non vide. Il définit alors l'estimateur de $\alpha(t_0)$ par :

$$\tilde{\alpha}_N(t_0) = \min \left(\operatorname{argmin}_{\alpha \in [0,2]} \left(\int_{p_0}^2 |\tilde{R}_N(p) - R_\alpha(p)|^\gamma dp \right)^{1/\gamma} \right),$$

et montre la consistance de cet estimateur (l'estimateur converge même dans L^r pour tout $r > 0$).

Cet estimateur nécessite le choix de deux paramètres p_0 et γ , l'approximation d'une intégrale ainsi que la minimisation de celle-ci. Numériquement, on peut constater sur les exemples du mouvement de Lévy multistable et du mouvement linéaire multifractionnaire multistable (voir Figure 4 – 9 de [LG13]) que cet estimateur n'est pas toujours très performant pour $N = 20000$ et $n(N) = 3000$.

Dans cette section, nous proposons un nouvel estimateur basé lui-aussi sur les moments d'ordre p du processus multistable.

Théorème 4.3.1

Soit $0 < \theta < 1$, on définit la fonction $\Phi_\theta(x) = \frac{\Gamma(1-\theta x)^{1/\theta}}{\Gamma(1-x)}$. Cette fonction est strictement décroissante de $[0, 1[$ à valeurs dans $]-\infty, 1]$ de bijection réciproque Φ_θ^{-1} . On peut alors définir l'estimateur

$$\hat{\alpha}_{N,p,\theta}(t_0) = \frac{p}{\Phi_\theta^{-1} \left(\min \left[1, \left(\frac{S_N(\theta p)}{S_N(p)} \right)^p \frac{1-p}{\Gamma(2-p) \cos(\frac{\pi p}{2})} \left(\frac{\Gamma(2-\theta p) \cos(\frac{\pi \theta p}{2})}{1-\theta p} \right)^{1/\theta} \right] \right)}. \quad (4.3.1)$$

On se place sous les conditions du Théorème 4.2.2. On a alors,

$$\hat{\alpha}_{N,p,\theta}(t_0) \xrightarrow[N \rightarrow \infty]{\text{proba}} \begin{cases} \alpha(t_0), & \text{si } p \in]0, \alpha(t_0)[, \\ p, & \text{si } p \in]\alpha(t_0), 2[. \end{cases} \quad (4.3.2)$$

Remarque 4.3.1. Cet estimateur dépend lui aussi de 2 paramètres : p et θ mais il présente l'avantage d'être très facilement calculable.

Démonstration : Soit $0 < \theta < 1$ fixé et $x \in [0, 1[$. alors,

$$\begin{aligned}\Phi'_\theta(x) &= \frac{-\theta\Gamma'(1-\theta x)1/\theta\Gamma(1-\theta x)^{1/\theta-1}\Gamma(1-x) + \Gamma'(1-x)\Gamma(1-\theta x)^{1/\theta}}{\Gamma(1-x)^2} \\ &= \frac{\Gamma(1-\theta x)^{1/\theta}}{\Gamma(1-x)} \left(\frac{\Gamma'(1-x)}{\Gamma(1-x)} - \frac{\Gamma'(1-\theta x)}{\Gamma(1-\theta x)} \right) \\ &= \frac{\Gamma(1-\theta x)^{1/\theta}}{\Gamma(1-x)} (\Psi(1-x) - \Psi(1-\theta x))\end{aligned}$$

où Ψ désigne la fonction digamma qui est strictement croissante sur $[0, 1[$. Ainsi, Φ_θ est une fonction strictement décroissante sur $[0, 1[$ de fonction réciproque $\Phi_\theta^{-1} :]-\infty, 1] \mapsto [0, 1[$. En utilisant le Théorème 4.2.1 et l'expression de $\mathbb{E}[|X(1, t_0)|^p]$ rappelée dans (4.2.1), on obtient que pour $p \in]0, \alpha(t_0)[$,

$$\left(\frac{S_N(\theta p)}{S_N(p)} \right)^p \frac{1-p}{\Gamma(2-p)\cos(\frac{\pi p}{2})} \left(\frac{\Gamma(2-\theta p)\cos(\frac{\pi\theta p}{2})}{1-\theta p} \right)^{1/\theta} \xrightarrow[N \rightarrow \infty]{proba} \frac{\Gamma(1-\theta p/\alpha(t_0))^{1/\theta}}{\Gamma(1-p/\alpha(t_0))}$$

et donc

$$\hat{\alpha}_{N,p,\theta}(t_0) \xrightarrow[N \rightarrow \infty]{proba} \alpha(t_0).$$

Lorsque $p \in]\alpha(t_0), 2[$, on distingue deux cas.

Si $\alpha(t_0) < \theta p < p$, on a

$$\frac{S_N(\theta p)}{S_N(p)} = \frac{N^{H(t_0)}n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{\theta p}} S_N(\theta p)}{N^{H(t_0)}n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} S_N(p)} n(N)^{-\frac{1-\theta}{\theta p}},$$

et alors on obtient à partir du Théorème 4.2.2,

$$N^{H(t_0)}n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} S_N(p) \xrightarrow[N \rightarrow \infty]{Loi} W^{1/p},$$

et

$$N^{H(t_0)}n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{\theta p}} S_N(\theta p) \xrightarrow[N \rightarrow \infty]{Loi} \tilde{W}^{1/\theta p}.$$

Par le lemme de Slutsky, on obtient que

$$\frac{S_N(\theta p)}{S_N(p)} \xrightarrow[N \rightarrow \infty]{proba} 0.$$

Si $\theta p < \alpha(t_0) < p$,

$$\frac{S_N(\theta p)}{S_N(p)} = \frac{N^{H(t_0)}S_N(\theta p)}{N^{H(t_0)}n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}} S_N(p)} n(N)^{\frac{-1}{\alpha(t_0)} + \frac{1}{p}},$$

en utilisant les Théorèmes 4.2.1, 4.2.2 et le lemme de Slutsky, on a alors

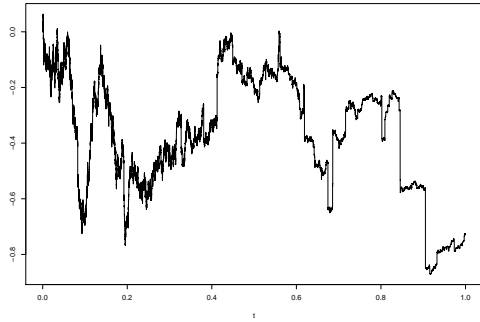
$$\frac{S_N(\theta p)}{S_N(p)} \xrightarrow[N \rightarrow \infty]{\text{proba}} 0.$$

Comme Φ_θ^{-1} est continue et que $\Phi_\theta^{-1}(0) = 1$, on obtient alors le résultat. ■

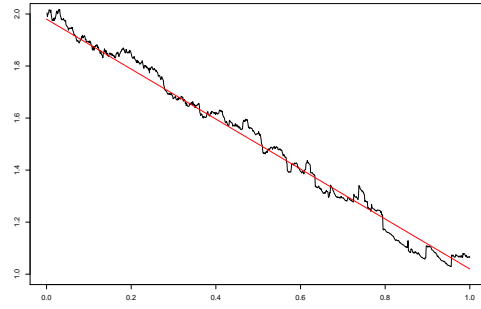
4.3.1 Exemple d'estimation avec des mouvements de Lévy multistables

Lorsque $\alpha \in \mathcal{C}^1([0, 1],]1, 2[)$, comme décrit après le Théorème 4.2.2, les mouvements de Lévy multistables vérifient toutes les conditions nécessaires à l'application du théorème de convergence des moments d'ordre p . En particulier, pour $0 < \theta < 1$, on a la convergence en probabilité de l'estimateur $\hat{\alpha}_{N,p,\theta}(t_0)$ défini en (4.3.1) vers $\alpha(t_0)$ pour $0 < p < \alpha(t_0)$ et vers p pour $\alpha(t_0) < p < 2$.

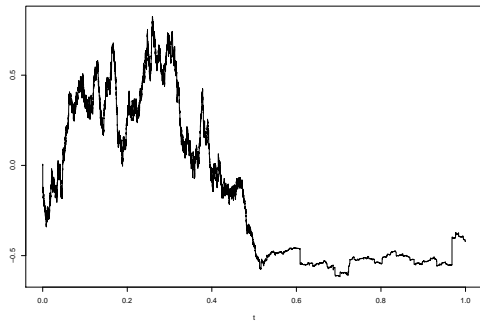
Pour les simulations, on va fixer le paramètre $p < 1$ et $\theta = 1/2$. Dans la Figure 4.3, on va s'intéresser à l'estimation de la fonction d'intensité des mouvements de Lévy multistables qu'on va considérer. On peut remarquer que l'estimateur arrive bien à estimer la fonction d'intensité même en cas de changements rapides.



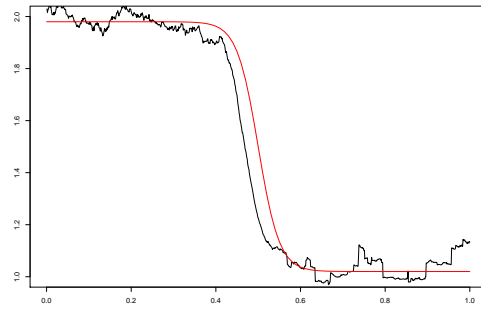
(a) mouvement de Lévy multistable de fonction d'intensité définie dans (b)



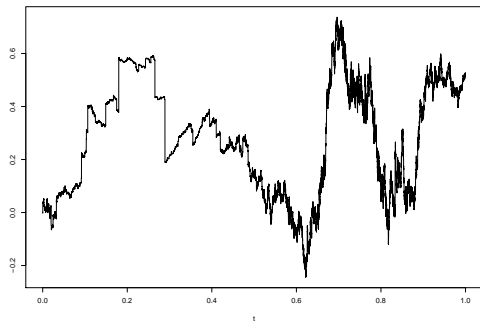
(b) $\alpha(t) = 1.98 - 0.96t$



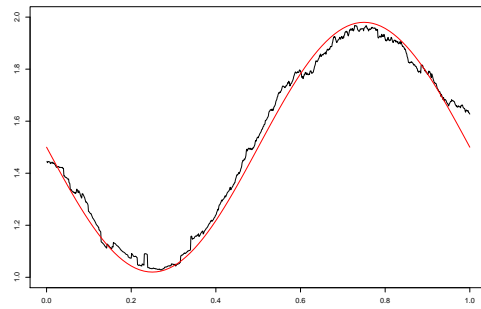
(c) mouvement de Lévy multistable de fonction d'intensité définie dans (d)



(d) $\alpha(t) = 1.98 - \frac{0.96}{1 + \exp(20 - 40t)}$



(e) mouvement de Lévy multistable de fonction d'intensité définie dans (f)



(f) $\alpha(t) = 1.5 - 0.48 \sin(2\pi t)$

FIGURE 4.3 – Trajectoires de mouvements de Lévy multistables sur $[0, 1]$ avec $N = 20000$ points et $n(N) = 2042$ pour l'estimation de α . La première colonne représente des réalisations de mouvements de Lévy multistables avec la fonction d'intensité α (en rouge) correspondante et son estimation $\hat{\alpha}_N$ (en noir) dans la deuxième colonne.

4.3.2 Exemple d'estimation avec des mouvements linéaires multifractionnaires multistables

Bien que la convergence de l'estimateur pour toute valeur de p n'a pas été vérifiée pour les mouvements linéaires multifractionnaires multistables, ces processus vérifient les conditions du Théorème 4.2.1 lorsque $h - \frac{1}{\alpha}$ est une fonction positive (voir Section 6.1 de [LG13]) où α est la fonction d'intensité et h est la fonction de localisation de ce processus. Pour de tels processus, on a ainsi, pour $0 < p < \alpha(t_0)$,

$$\hat{\alpha}_{N,p,\theta}(t_0) \xrightarrow[N \rightarrow \infty]{\text{proba}} \alpha(t_0).$$

Comme h est à valeurs dans $]0, 1[$, pour ces processus α à valeurs dans $]1, 2[$. On prend donc $p < 1$ pour l'estimation de la fonction d'intensité des mouvements linéaires multifractionnaires multistables considérés.

Les mouvements linéaires multifractionnaires multistables sont des processus difficiles à étudier car ils sont paramétrés par deux fonctions. Cependant, on constate dans les Figures 4.4 et 4.5 que la fonction d'intensité est bien estimée et ceci pour différentes formes pour α et h .

Nous n'avons pas étudié la fonction de localisation dans nos travaux. Cependant, il est possible de l'estimer à partir des log-moments du processus sur une fenêtre d'observation (voir [LG13]).

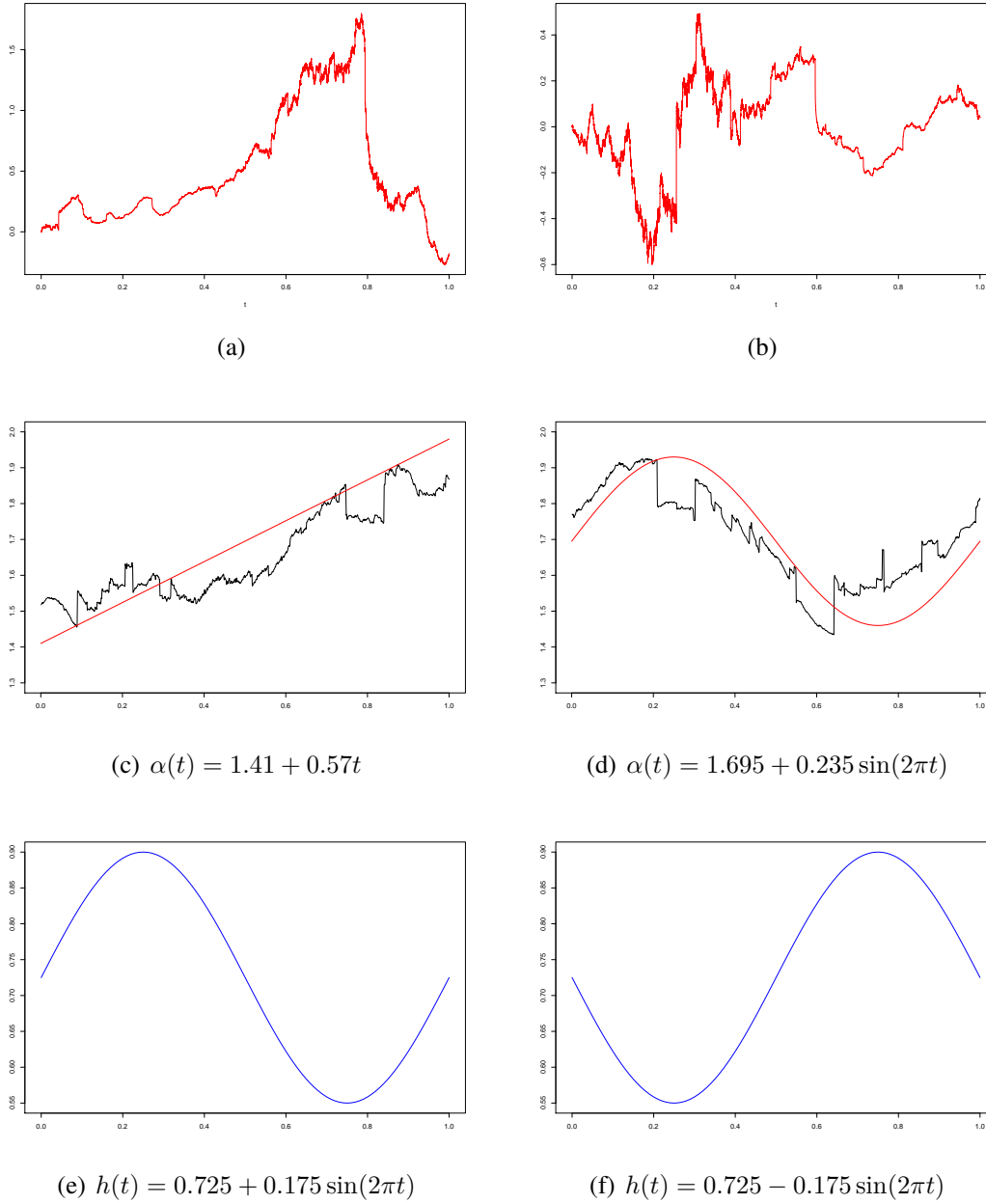


FIGURE 4.4 – Trajectoires de mouvements linéaires multifractionnaires multistables sur $[0, 1]$ avec $N = 32168$ points et $n(N) = 3000$ pour l'estimation de α . La première ligne représente des réalisations de mouvements linéaires multifractionnaires multistables avec la fonction d'intensité α (en rouge) correspondante et son estimation $\hat{\alpha}_N$ (en noir) dans la deuxième ligne, et la fonction de localisation h associée dans la troisième ligne.

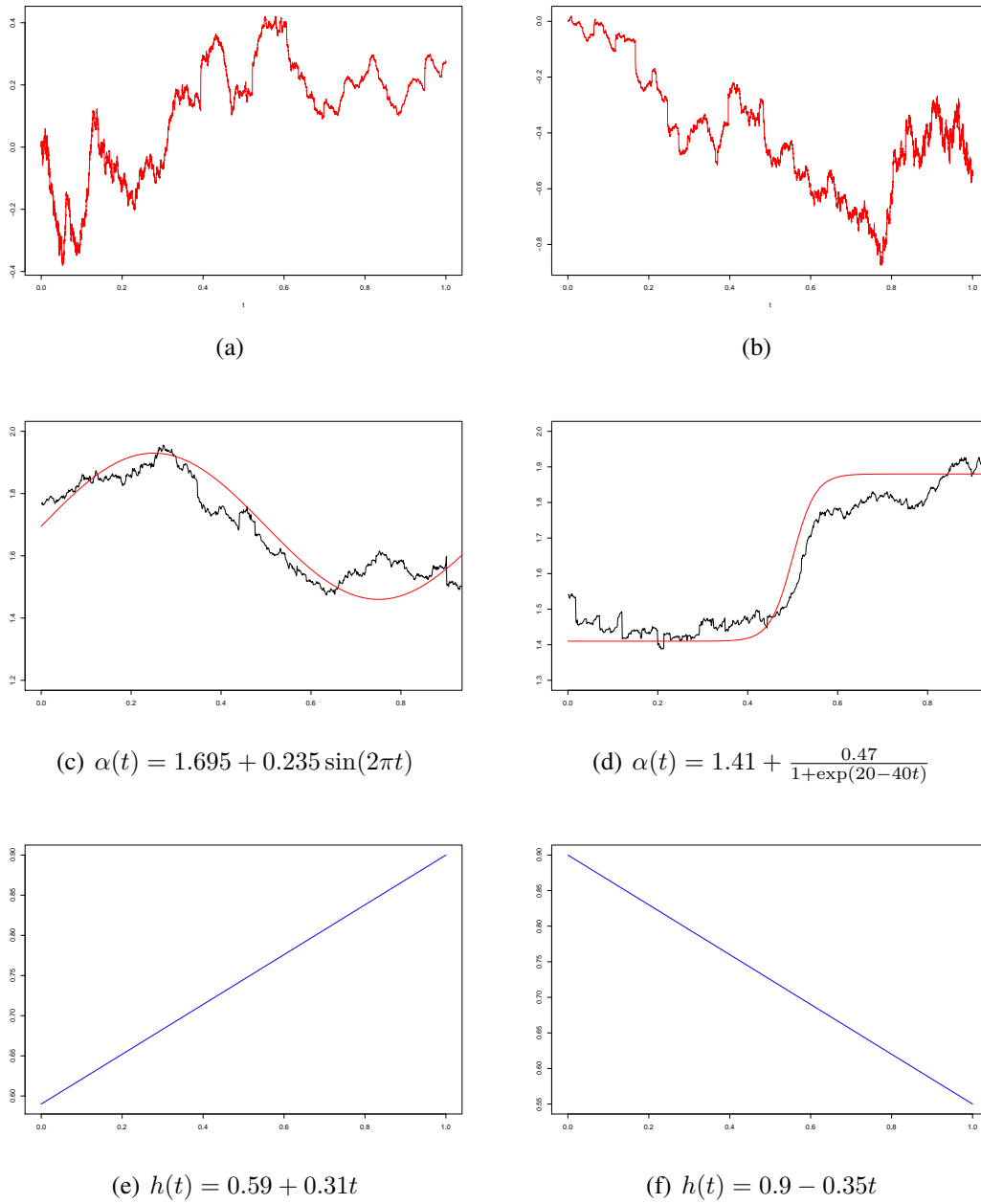


FIGURE 4.5 – Trajectoires de mouvements linéaires multifractionnaires multistables sur $[0, 1]$ avec $N = 32168$ points et $n(N) = 3000$ pour l'estimation de α . La première ligne représente des réalisations de mouvements linéaires multifractionnaires multistables avec la fonction d'intensité α (en rouge) correspondante et son estimation $\hat{\alpha}_N$ (en noir) dans la deuxième ligne, et la fonction de localisation h associée dans la troisième ligne.



**Échantillonnage de processus
stationnaires du second ordre
définis en temps continu**

Échantillonnage aléatoire de processus stationnaires à temps continu

Sommaire

5.1	Introduction	134
5.2	Résultats généraux dans le domaine temporel	136
5.2.1	Loi du processus échantillonné	136
5.2.2	Mémoire du processus échantillonné	139
5.3	Processus à longue mémoire	142
5.3.1	Préservation du paramètre du mémoire lorsque $\mathbb{E}[T_1] < \infty$	142
	Normalité asymptotique du processus des sommes partielles	145
	Résultats statistiques dans le cas de la longue mémoire	148
5.3.2	Perte de mémoire lorsque $\mathbb{E}[T_1] = \infty$	152
5.4	Comparaison entre le sous échantillonnage et l'agrégation temporelle	154
5.5	Annexe	160

5.1 Introduction

Les séries temporelles à longue dépendance ont diverses applications dans de nombreux domaines, notamment l'hydrologie, l'économie et les télécommunications (voir [Ber+13] chap. 2). La plupart des articles sur ce sujet considèrent les processus à temps discret. Cependant, certains modèles et méthodes d'estimation ont été étendus aux processus en temps continu (voir [TC05a], [Via+94], [CR96], [Com96]).

Dans [TC05a], Tsai et Chan ont introduit le modèle des processus à temps continu CARFIMA(p, H, q) (autorégressifs, moyenne mobile, fractionnairement intégrés). Sous la condition de forte dépendance $H \in]1/2, 1[$, ils donnent la fonction d'autocovariance du processus CARFIMA qui est stationnaire. Ces propriétés sont étendues au cas $H \in]0, 1[$ dans [Tsa09]. Dans [Via+94], des processus ARMA fractionnaire et à temps continu sont construits. Sous certaines conditions, ces processus sont gaussiens stationnaires et centrés. De plus, Viano *et al* donnent des résultats sur la fonction d'autocovariance et précisent la dépendance asymptotique de ces processus. Dans [CR96], Comte et Renault présentent une famille de modèles à longue mémoire : le processus fractionnaire à moyenne mobile en temps continu.

L'inférence statistique pour les processus en temps continu est généralement construite à partir d'un processus échantillonné. Dans [TC05a], la méthode d'estimation est basée sur l'estimation par maximum de vraisemblance pour des données de séries temporelles à espacement déterministe irrégulier.

Nous nous intéressons aux données irrégulièrement espacées lorsque les intervalles d'échantillonnage sont des variables aléatoires positives indépendantes et de même loi. À la vue des résultats en temps discret, l'échantillonnage aléatoire a un effet sur la structure de dépendance du processus. En effet, [PV10] montrent que l'intensité de la longue mémoire est préservée lorsque la loi des intervalles d'échantillonnage est L^1 , mais ils montrent aussi des situations menant à une réduction de la mémoire.

Dans ce chapitre, $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+}$ est un processus stationnaire du second ordre à temps continu de fonction d'autocovariance $\sigma_X(\cdot)$ et $(T_n)_{n \geq 0}$ est une marche aléatoire indépendante de \mathbf{X} . Les intervalles d'échantillonnage $T_{j+1} - T_j = \Delta_j$ sont

supposés i.i.d de densité commune s à support sur \mathbb{R}^+ . L'origine des observations est fixée à l'origine des temps : $T_0 = 0$. Sans perte de généralités, \mathbf{X} est supposé être un processus centré.

Nous adoptons la définition la plus courante de la longue mémoire. A savoir, pour un processus stationnaire \mathbf{U} ayant une fonction de covariance σ_U

$$\begin{aligned} \int_{\mathbb{R}^+} |\sigma_U(x)| dx = \infty & \quad \text{dans le cas à temps continu,} \\ \sum_{h \geq 0} |\sigma_U(h)| = \infty & \quad \text{dans le cas à temps discret.} \end{aligned}$$

Deux méthodes d'échantillonnage, toutes deux basées sur la marche aléatoire $(T_n)_{n \geq 0}$, sont examinées ci-après.

Les sections 5.2 et 5.3 sont consacrées à l'étude des processus échantillonnés à des moments aléatoires. Plus précisément, nous considérons le processus en temps discret \mathbf{Y} défini par

$$Y_n = X_{T_n}, \quad n = 0, 1, \dots \quad (5.1.1)$$

Dans la section 5.2, nous étudions le comportement du processus échantillonné dans le cas général. Nous prouvons que dans le cas d'un processus initial gaussien, l'échantillonné \mathbf{Y} n'est pas gaussien, et que sous des conditions assez faibles sur la covariance σ_X , échantillonner un processus à courte mémoire \mathbf{X} produit toujours une courte mémoire pour \mathbf{Y} , et que lorsque la loi d'échantillonnage satisfait $\mathbb{E}[T_1] < \infty$ le processus échantillonné \mathbf{Y} possède une longue mémoire si c'est le cas pour \mathbf{X} .

Dans la section 5.3, nous présentons la situation plus spécifique d'une covariance à variation régulière pour laquelle la préservation ou la non-préservation de la mémoire peut être quantifiée. En particulier, nous prouvons qu'une loi d'échantillonnage à queue lourde peut transformer un processus à longue mémoire \mathbf{X} en un processus à courte mémoire \mathbf{Y} .

La section 5.4 se tourne vers un schéma d'agrégation temporelle. Le processus échantillonné en temps discret \mathbf{Z} est obtenu en agrégeant le processus \mathbf{X} sur les intervalles définis par la marche aléatoire

$$Z_h = \int_{T_h}^{T_{h+1}} X_t dt.$$

Nous montrons que du point de vue de la transmission de la mémoire, les résultats

sont similaires à ceux obtenus dans les sections précédentes.

5.2 Résultats généraux dans le domaine temporel

Soit Y le processus défini à partir de X par

$$Y_n = X_{T_n} \quad n \in \mathbb{N} \quad (5.2.1)$$

où $(T_{j+1} - T_j)_j$ sont les intervalles d'échantillonnage supposés i.i.d. Le processus à temps discret Y est lui-aussi stationnaire du second ordre de fonction d'autocovariance

$$\begin{cases} \sigma_Y(0) = \sigma_X(0), \\ \sigma_Y(h) = \mathbb{E}[\sigma_X(T_h)], \quad h \geq 1. \end{cases} \quad (5.2.2)$$

5.2.1 Loi du processus échantillonné

On regarde la loi du processus Y afin de déterminer si les propriétés classiques que le processus initial peut avoir sont conservées après échantillonnage.

Proposition 5.2.1

Soit X un processus strictement stationnaire. Alors, le processus échantillonné Y est un processus à temps discret strictement stationnaire, ie la loi jointe de $(Y_{k_1+p}, \dots, Y_{k_n+p})$ ne dépend pas de p pour tout $k_1, \dots, k_n \in \mathbb{N}$ et $n \geq 1$.

Démonstration : Soit $n \geq 1, p \in \mathbb{N}$ et $0 \leq k_1 < \dots < k_n$. Pour $(y_1, \dots, y_n) \in \mathbb{R}^n$,

$$\begin{aligned} & P(Y_{k_1+p} \leq y_1, \dots, Y_{k_n+p} \leq y_n) \\ &= P(X_{T_{k_1+p}} \leq y_1, \dots, X_{T_{k_n+p}} \leq y_n) \\ &= \mathbb{E} \left[P \left(X_{\Delta_0 + \dots + \Delta_{k_1+p-1}} \leq y_1, \dots, X_{\Delta_0 + \dots + \Delta_{k_n+p-1}} \leq y_n \mid \Delta_0, \dots, \Delta_{k_n+p-1} \right) \right] \\ & \quad \text{par la stricte stationnarité de } X, \text{ on obtient} \\ &= \mathbb{E} \left[P \left(X_{\Delta_p + \dots + \Delta_{k_1+p-1}} \leq y_1, \dots, X_{\Delta_p + \dots + \Delta_{k_n+p-1}} \leq y_n \mid \Delta_0, \dots, \Delta_{k_n+p-1} \right) \right] \\ &= P(X_{U_0 + \dots + U_{k_1-1}} \leq y_1, \dots, X_{U_0 + \dots + U_{k_n-1}} \leq y_n) \\ & \quad \text{où } U_i = \Delta_{i+p} \text{ sont i.i.d de densité } s \\ &= P(Y_{k_1} \leq y_1, \dots, Y_{k_n} \leq y_n) \end{aligned}$$

■

La proposition suivante donne le premier résultat d'une perte d'information après échantillonnage : la normalité du processus n'est pas préservée par échantillonnage aléatoire.

Proposition 5.2.2

Si X est un processus gaussien alors les marginales du processus échantillonné Y sont gaussiennes. De plus, pour σ_X non constante presque partout sur le support de s , Y n'est pas un processus gaussien.

Démonstration : Soit Φ_U la fonction caractéristique de la variable aléatoire U .

On a

$$\Phi_{Y_k}(t) = \mathbb{E} \left[\mathbb{E}[e^{itX_{T_k}} | T_k] \right].$$

Sachant T_k , la loi conditionnelle de X_{T_k} est une loi normale centrée de variance $\sigma_X(0)$. Ainsi,

$$\Phi_{Y_k}(t) = e^{-\sigma_X(0)t^2/2}$$

et $Y_k \sim \mathcal{N}(0, \sigma_X(0))$.

Supposons que Y est un processus gaussien, alors $Y_1 + Y_2$ est une variable aléatoire gaussienne,

$$\Phi_{Y_1+Y_2}(t) = e^{-\text{Var}(Y_1+Y_2)t^2/2} = e^{-\sigma(0)t^2} e^{-t^2\mathbb{E}[\sigma_X(T_2-T_1)]}.$$

Comme X est un processus gaussien, $X_{T_1} + X_{T_2}$ est une variable gaussienne sachant (T_1, T_2) . Cela nous donne une autre manière de calculer $\Phi_{Y_1+Y_2}$:

$$\begin{aligned} \Phi_{Y_1+Y_2}(t) &= \Phi_{X_{T_1}+X_{T_2}}(t) \\ &= \mathbb{E} \left[\exp \left\{ -\frac{t^2}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \begin{pmatrix} \sigma_X(0) & \sigma_X(T_2-T_1) \\ \sigma_X(T_2-T_1) & \sigma_X(0) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \right] \\ &= e^{-\sigma(0)t^2} \mathbb{E} \left[e^{-t^2\sigma_X(T_2-T_1)} \right]. \end{aligned}$$

Ainsi, pour tout t , $e^{-t^2\mathbb{E}[\sigma_X(T_2-T_1)]} = \mathbb{E} \left[e^{-t^2\sigma_X(T_2-T_1)} \right]$. C'est le cas d'égalité dans l'inégalité de Jensen, donc $\sigma_X(T_2 - T_1)$ est constant presque sûrement. ■

Exemple 5.2.1. Dans la Figure 5.1, nous illustrons l'absence de normalité du processus échantillonné.

Pour X processus gaussien de fonction d'autocovariance $\sigma_X(t) = (1 + t^{0.9})^{-1}$, nous simulons en Figure (a) la densité jointe du couple de variables aléatoires centrées (Y_1, Y_2)

pour des intervalles Δ_j de loi exponentielle de moyenne 1. Pour cela, on commence par simuler l'intervalle de temps selon une loi exponentielle $\mathcal{E}(1)$. Puis, pour chaque valeur simulée de $T_2 - T_1$, le couple (X_{T_1}, X_{T_2}) est gaussien et est simulé à l'aide de la décomposition de Cholesky de la matrice de covariance (voir [Rip87]). Par cette méthode, $p = 200000$ réalisations de (Y_1, Y_2) sont simulées et la densité bivariable de (Y_1, Y_2) est approchée par une méthode de noyau.

Dans la Figure (b), la densité d'un couple gaussien centré (W_1, W_2) ayant la même matrice de covariance que (Y_1, Y_2) : $\Sigma_{Y_1, Y_2} = \begin{pmatrix} \sigma_X(0) & \mathbb{E}[\sigma_X(T_1)] \\ \mathbb{E}[\sigma_X(T_1)] & \sigma_X(0) \end{pmatrix}$ est représentée afin de comparer le comportement du processus échantillonné avec le processus gaussien correspondant. Dans ce cas, la densité bivariable a une expression explicite

$$f_{(W_1, W_2)}(w_1, w_2) = \left(2\pi\sqrt{\det(\Sigma_{Y_1, Y_2})}\right)^{-1} \exp\left(-\frac{1}{2}(w_1 \ w_2)\Sigma_{Y_1, Y_2}^{-1}(w_1 \ w_2)^T\right)$$

où Σ_{Y_1, Y_2} est la matrice de covariance commune. La forme de la loi pour le processus échantillonné est très différente de la loi gaussienne.

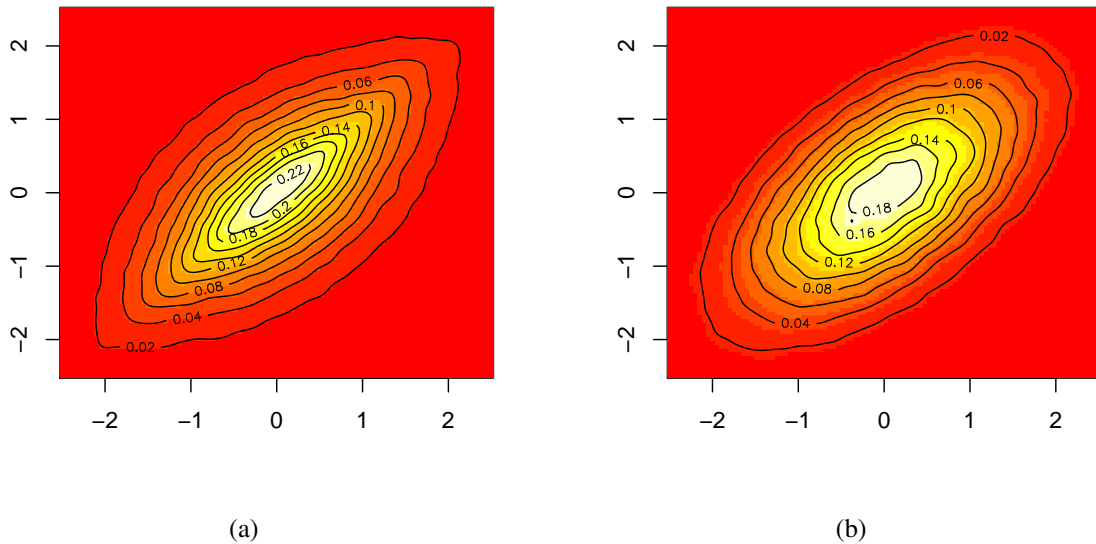


FIGURE 5.1 – Dans la Figure (a), la densité bivariable du couple centré (Y_1, Y_2) est représentée pour des intervalles Δ_j ayant une loi exponentielle de moyenne 1 et un processus initial gaussien de fonction d'autocovariance $\sigma_X(t) = (1 + t^{0.9})^{-1}$. Dans (b), on représente la densité bivariable du vecteur gaussien centré (W_1, W_2) de matrice de covariance Σ_{Y_1, Y_2} comme (Y_1, Y_2) .

5.2.2 Mémoire du processus échantillonné

On cherche désormais à décrire la mémoire du processus après échantillonnage.

Proposition 5.2.3

Soit $p \geq 1$. S'il existe une fonction $\sigma_*(\cdot)$ bornée strictement positive et décroissante sur \mathbb{R}^+ , telle que

1. $|\sigma_X(t)| \leq \sigma_*(t), \quad \forall t \in \mathbb{R}^+$
2. $\int_{\mathbb{R}^+} \sigma_*^p(t) dt < \infty$

alors, la fonction d'autocovariance du processus échantillonné Y est dans ℓ^p , i.e $\sum_{h \geq 0} |\sigma_Y(h)|^p < \infty$.

Remarque 5.2.1. Cette proposition confirme une intuition : l'échantillonnage aléatoire ne peut pas produire une longue mémoire à partir d'une courte mémoire. Le cas particulier $p = 1$ implique que si X a une courte mémoire alors, le processus échantillonné Y a aussi une courte mémoire.

Démonstration : Il suffit clairement de prouver que

$$\sum_{h \geq 1} \mathbb{E}[\sigma_*^p(T_h)] < \infty. \quad (5.2.3)$$

On utilise l'inégalité

$$\Delta_h \sigma_*^p(T_h + \Delta_h) = (T_{h+1} - T_h) \sigma_*^p(T_{h+1}) \leq \int_{T_h}^{T_{h+1}} \sigma_*^p(t) dt, \quad \forall h \geq 0. \quad (5.2.4)$$

En prenant l'espérance au membre de droite de l'inégalité et en remarquant que Δ_h et T_h sont indépendants, on obtient pour tout $a > 0$,

$$\begin{aligned} \mathbb{E}[\Delta_h \sigma_*^p(T_h + \Delta_h)] &= \int_{\mathbb{R}^+} u \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) \\ &= \int_0^a u \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) + \int_a^{+\infty} u \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) \\ &\geq \int_0^a u \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) + a \int_a^{+\infty} \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) \\ &= \int_0^a u \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) + a \left(\int_{\mathbb{R}^+} \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) - \int_0^a \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) \right) \\ &= \int_0^a (u - a) \mathbb{E}[\sigma_*^p(T_h + u)] dS(u) + a \mathbb{E}[\sigma_*^p(T_{h+1})] \end{aligned}$$

De plus, $\sigma_*^p(T_h + u) \leq \sigma_*^p(T_h)$ et $u - a \leq 0$ pour $u \in [0, a]$, ce qui conduit à

$$\mathbb{E} [\Delta_h \sigma_*^p(T_h + \Delta_h)] \geq \left(\int_{[0, a[} (u - a) dS(u) \right) \mathbb{E} [\sigma_*^p(T_h)] + a \mathbb{E} [\sigma_*^p(T_{h+1})]. \quad (5.2.5)$$

Il est possible de choisir a tel que $S([0, a]) < 1$. Pour un tel choix, on obtient

$$0 \leq - \int_{[0, a[} (u - a) dS(u) =: \ell(a) \leq aS([0, a]) < a.$$

La somme des inégalités dans (5.2.5) donne, pour tout $K \geq 0$

$$\begin{aligned} \mathbb{E} \left[\sum_{h=1}^{\infty} \Delta_h \sigma_*^p(T_{h+1}) \right] &\geq \sum_{h=1}^K [-\ell(a) \mathbb{E} [\sigma_*^p(T_h)] + a \mathbb{E} [\sigma_*^p(T_{h+1})]] \\ &= a (\mathbb{E} [\sigma_*^p(T_{K+1})] - \mathbb{E} [\sigma_*^p(T_1)]) + (a - \ell(a)) \sum_{h=1}^K \mathbb{E} [\sigma_*^p(T_h)] \\ &\geq -a \sigma_*^p(0) + (a - \ell(a)) \sum_{h=1}^K \mathbb{E} [\sigma_*^p(T_h)], \end{aligned}$$

ce qui implique

$$\mathbb{E} \left[\sum_{h=1}^{\infty} \Delta_h \sigma_*^p(T_{h+1}) \right] \geq -a \sigma_*^p(0) + (a - \ell(a)) \sum_{h \geq 1} \mathbb{E} [\sigma_*^p(T_h)]$$

Finalement, en utilisant (5.2.4)

$$\mathbb{E} \left[\sum_{h \geq 1} \Delta_h \sigma_*^p(T_{h+1}) \right] \leq \mathbb{E} \left[\sum_{h \geq 1} \int_{T_h}^{T_{h+1}} \sigma_*^p(t) dt \right] \leq \int_{\mathbb{R}^+} \sigma_*^p(t) dt < \infty$$

et par conséquent, comme $a - \ell(a) > 0$

$$\sum_{h=1}^{\infty} \mathbb{E} [\sigma_*^p(T_h)] < \infty. \quad (5.2.6) \quad \blacksquare$$

Proposition 5.2.4

Supposons que $\sigma_X(\cdot)$ est une fonction strictement positive et décroissante sur \mathbb{R}^+ à partir d'un certain rang, i.e il existe $t_0 \geq 0$ tel que $\sigma_X(\cdot)$ est strictement positive et décroissante sur $[t_0, \infty)$. Si $\mathbb{E}[T_1] < \infty$ et \mathbf{X} est à longue mémoire alors \mathbf{Y} l'est aussi.

Démonstration : Soit h_0 le premier indice (aléatoire) tel que $T_{h_0} \geq t_0$. Pour tout

$h \geq h_0$,

$$\int_{T_h}^{T_{h+1}} \sigma_X(t) dt \leq (T_{h+1} - T_h) \sigma_X(T_h). \quad (5.2.7)$$

Après avoir fait la somme, on obtient

$$\sum_{h \geq 1} \mathbb{1}_{h \geq h_0} \int_{T_h}^{T_{h+1}} \sigma_X(t) dt \leq \sum_{h \geq 1} \mathbb{1}_{h \geq h_0} \Delta_h \sigma_X(T_h).$$

Maintenant, en prenant l'espérance et en remarquant par la loi forte des grands nombre que $T_h \xrightarrow{p.s.} \infty$ (puisque $\mathbb{E}[T_1] = \mathbb{E}[\Delta_1] > 0$), et en particulier $h_0 < \infty$ p.s., d'où

$$\mathbb{E} \left[\int_{T_{h_0}}^{\infty} \sigma_X(t) dt \right] \leq \mathbb{E} \left[\sum_{h=1}^{\infty} \Delta_h \sigma_X(T_h) \mathbb{1}_{h_0 \leq h} \right].$$

Dans l'inégalité, le membre de gauche est $+\infty$, et, comme Δ_h est indépendant de $\sigma_X(T_h) \mathbb{1}_{h_0 \leq h}$, le membre de droite est $\mathbb{E}[T_1] \sum_{h \geq 1} \mathbb{E}[\sigma_X(T_h) \mathbb{1}_{h_0 \leq h}]$. En conséquence, comme $\mathbb{E}[T_1] < \infty$,

$$\sum_{h \geq 1} \mathbb{E}[\sigma_X(T_h) \mathbb{1}_{h_0 \leq h}] = \infty. \quad (5.2.8)$$

Il reste à remarquer que $\mathbb{E}[h_0] < \infty$ (voir par exemple [Fel71] p.185), ce qui implique

$$\sum_{h \geq 1} \mathbb{E}[\sigma_X(T_h) \mathbb{1}_{h_0 > h}] \leq \sigma_X(0) \sum_{h \geq 1} P(h_0 \geq h) \leq \sigma_X(0) \mathbb{E}[h_0] < \infty,$$

ce qui entraîne, par (5.2.8) que $\sum_{h \geq 1} |\mathbb{E}[\sigma_X(T_h)]| = \infty$. ■

Remarque 5.2.2. Dans cette proposition, les hypothèses sur la positivité et la décroissance de la fonction d'autocovariance sont raisonnables dans la plupart des modèles. À moins d'étudier un phénomène dont on peut choisir l'échantillonnage (par exemple, pour des données financières), la condition d'intégrabilité des intervalles Δ_j est la plus difficile à vérifier lorsqu'on veut traiter avec des données réelles (par exemple, lorsque les données sont issues de la géologie) puisque le processus sous-jacent n'est alors pas observé.

5.3 Processus à longue mémoire

Dans cette section, nous considérons les processus de longue mémoire avec une forme spécifique de fonction de covariance. L'hypothèse est satisfaite, par exemple pour les modèles CARFIMA.

Définition 5.3.1

— Une fonction L sur $[0, \infty[$ est dite à variation lente en l'infini si L est strictement positive sur $[t_0, \infty[$ pour un certain $t_0 > 0$ et

$$\lim_{x \rightarrow +\infty} \frac{L(ax)}{L(x)} = 1, \quad \forall a > 0.$$

— Une fonction f sur $[0, \infty[$ est dit à variation régulière en l'infini d'indice $\delta \in \mathbb{R}$, si f est strictement positive sur $[t_0, \infty[$ pour un certain $t_0 > 0$ et

$$\lim_{x \rightarrow +\infty} \frac{f(ax)}{f(x)} = a^\delta, \quad \forall a > 0.$$

À partir de maintenant, la fonction d'autocovariance σ_X est une fonction à variation régulière en l'infini de la forme

$$\sigma_X(t) = t^{-1+2d}L(t), \quad \forall t \geq 1 \tag{5.3.1}$$

où $0 < d < 1/2$ et L est à variation lente en l'infini et décroissante à partir d'un certain rang. Cela signifie que le paramètre d résume la mémoire de \mathbf{X} et pour être plus précis par rapport aux résultats de la section 5.2, nous décrivons ce qui peut arriver à d après échantillonnage.

5.3.1 Préservation du paramètre du mémoire lorsque $\mathbb{E}[T_1] < \infty$

Théorème 5.3.2

Sous l'hypothèse (5.3.1), si $0 < \mathbb{E}[T_1] < \infty$, le processus à temps discret \mathbf{Y} a une longue mémoire et sa covariance se comporte en l'infini comme

$$\sigma_Y(h) \sim (h\mathbb{E}[T_1])^{-1+2d}L(h), \quad h \rightarrow \infty.$$

Remarque 5.3.1. On peut écrire

$$\sigma_Y(h) = h^{-1+2d}\tilde{L}(h)$$

où \tilde{L} est à variation lente en l'infini et $\tilde{L}(h) \sim \mathbb{E}[T_1]^{-1+2d}L(h)$ lorsque $h \rightarrow \infty$. En particulier, \mathbf{X} et \mathbf{Y} ont le même paramètre de mémoire d .

Démonstration :

- On commence par prouver que

$$\liminf_{h \rightarrow \infty} \frac{\sigma_Y(h)}{(h\mathbb{E}[T_1])^{-1+2d}L(h)} \geq 1.$$

Soit $0 < c < \mathbb{E}[T_1]$, et $h \in \mathbb{N}$ tel que $ch \geq 1$,

$$\sigma_Y(h) \geq \mathbb{E}[\sigma_X(T_h) \mathbf{1}_{T_h > ch}] \geq \inf_{t > ch} \{L(t)t^{2d}\} \mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h}\right]$$

Grâce à l'inégalité d'Hölder, on a

$$(P(T_h > ch))^2 \leq \mathbb{E}[T_h] \mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h}\right],$$

et donc

$$\mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h}\right] \geq \frac{(P(T_h > ch))^2}{h\mathbb{E}[T_1]}.$$

Pour résumer, on a ainsi

$$\sigma_Y(h) \geq \inf_{t > ch} \{L(t)t^{2d}\} \frac{(P(T_h > ch))^2}{h\mathbb{E}[T_1]}$$

$$\frac{\sigma_Y(h)}{(h\mathbb{E}[T_1])^{-1+2d}L(h)} \geq \inf_{t > ch} \{L(t)t^{2d}\} \frac{(P(T_h > ch))^2}{(h\mathbb{E}[T_1])^{2d}L(h)} \quad (5.3.2)$$

En utilisant [Bin+89] (Th 1.5.3, p23), on obtient, comme $d > 0$

$$\inf_{t \geq ch} \{L(t)t^{2d}\} \sim L(ch)(ch)^{2d} \quad \text{lorsque } h \rightarrow \infty. \quad (5.3.3)$$

La loi forte des grands nombres implique que $T_h/h \xrightarrow{p.s.} \mathbb{E}[T_1]$. Comme $c < \mathbb{E}[T_1]$, on a $P(T_h > ch) \rightarrow 1$ et le membre de droite de (5.3.2) tend vers $(c/\mathbb{E}[T_1])^{2d}$

comme $h \rightarrow \infty$. Finalement, pour tout $c < \mathbb{E}[T_1]$,

$$\liminf_{h \rightarrow \infty} \frac{\sigma_Y(h)}{(h\mathbb{E}[T_1])^{-1+2d}L(h)} \geq \left(\frac{c}{\mathbb{E}[T_1]}\right)^{2d}$$

En prenant la limite lorsque $c \rightarrow \mathbb{E}[T_1]$, on obtient la borne inférieure.

• Nous allons maintenant prouver que

$$\limsup_{h \rightarrow \infty} \frac{\sigma_Y(h)}{(h\mathbb{E}[T_1])^{-1+2d}L(h)} \leq 1.$$

Pour cela, nous utilisons une preuve similaire à celle présentée dans [Shi+10] (Thm 1). On pose pour $h \geq 1$ et $0 < s < 1$,

$$\begin{aligned} \mu_h &= \mathbb{E}[T_h] = h\mathbb{E}[T_1] \\ T_{h,s} &= \sum_{j=0}^{h-1} \Delta_j \mathbf{1}_{\Delta_j \leq \mu_h^s / \sqrt{h}} \\ \mu_{h,s} &= \mathbb{E}[T_{h,s}] = h\mathbb{E}\left[\Delta_0 \mathbf{1}_{\Delta_0 \leq \mu_h^s / \sqrt{h}}\right] \end{aligned}$$

Comme $\mathbb{E}[T_1] < \infty$, on a pour $\frac{1}{2} < s < 1$, $\mu_{h,s} \sim \mu_h$ lorsque $h \rightarrow \infty$.

Soit $\frac{1}{2} < s < \tau < 1$, t_0 tel que $L(\cdot)$ est décroissante sur $[t_0, \infty[$ et h tel que $\mu_{h,s} - \mu_{h,s}^\tau \geq t_0$,

$$\begin{aligned} \sigma_Y(h) &= \mathbb{E}\left[T_h^{-1+2d}L(T_h) \mathbf{1}_{T_{h,s} \geq \mu_{h,s} - \mu_{h,s}^\tau}\right] + \mathbb{E}\left[T_h^{-1+2d}L(T_h) \mathbf{1}_{T_{h,s} < \mu_{h,s} - \mu_{h,s}^\tau}\right] \\ &= M_1 + M_2 \end{aligned}$$

$$\begin{aligned} M_1 &\leq \mathbb{E}\left[T_{h,s}^{-1+2d}L(T_{h,s}) \mathbf{1}_{T_{h,s} \geq \mu_{h,s} - \mu_{h,s}^\tau}\right] \leq (\mu_{h,s} - \mu_{h,s}^\tau)^{-1+2d} L(\mu_{h,s} - \mu_{h,s}^\tau) \\ &= (h\mathbb{E}[T_1])^{-1+2d}L(h) \left(\frac{\mu_{h,s} - \mu_{h,s}^\tau}{h\mathbb{E}[T_1]}\right)^{-1+2d} \frac{L(\mu_{h,s} - \mu_{h,s}^\tau)}{L(h)} \end{aligned} \quad (5.3.4)$$

Comme $\tau < 1$ et $1/2 < s < 1$, $\left(\frac{\mu_{h,s} - \mu_{h,s}^\tau}{h\mathbb{E}[T_1]}\right)^{-1+2d} \rightarrow 1$ lorsque $h \rightarrow \infty$. Alors,

$$\frac{L(\mu_{h,s} - \mu_{h,s}^\tau)}{L(h)} = \frac{L\left(\frac{h\mathbb{E}[T_1] \mu_{h,s} - \mu_{h,s}^\tau}{h\mathbb{E}[T_1]}\right)}{L(h\mathbb{E}[T_1])} \frac{L(h\mathbb{E}[T_1])}{L(h)}$$

Comme on a la convergence uniforme de $\frac{L(h\mathbb{E}[T_1] \cdot)}{L(h\mathbb{E}[T_1])}$ vers 1 sur chaque intervalle

$[a, b]$ et comme $\frac{\mu_{h,s} - \mu_{h,s}^\tau}{h\mathbb{E}[T_1]} \rightarrow 1$, on obtient

$$\frac{L(\mu_{h,s} - \mu_{h,s}^\tau)}{L(h)} \rightarrow 1$$

lorsque $h \rightarrow \infty$. On obtient

$$M_1 \leq (\mu_{h,s} - \mu_{h,s}^\tau)^{-1+2d} L(\mu_{h,s} - \mu_{h,s}^\tau) \sim (h\mathbb{E}[T_1])^{-1+2d} L(h) \quad (5.3.5)$$

Comme $\sup_{t \in \mathbb{R}^+} |\sigma_X(t)| = \sigma_X(0) < \infty$, on a

$$M_2 \leq \sigma_X(0) P(T_{h,s} < \mu_{h,s} - \mu_{h,s}^\tau) = \sigma_X(0) P(-T_{h,s} + \mathbb{E}[T_{h,s}] > \mu_{h,s}^\tau).$$

On applique l'inégalité d'Hoeffding aux variables $Z_j = -\Delta_j \mathbf{1}_{\Delta_j \leq \mu_h^s / \sqrt{h}}$ qui sont p.s. dans $[-\frac{\mu_h^s}{\sqrt{h}}, 0]$ pour obtenir,

$$M_2 \leq \sigma_X(0) \exp\left(-2 \left(\frac{\mu_{h,s}^\tau}{\mu_h^s}\right)^2\right)$$

et $\left(\frac{\mu_{h,s}^\tau}{\mu_h^s}\right)^2 \sim (h\mathbb{E}[T_1])^{2(\tau-s)}$. Finalement,

$$M_2 = o((h\mathbb{E}[T_1])^{-1+2d} L(h)). \quad (5.3.6)$$

Avec (5.3.5) et (5.3.6), on obtient la borne supérieure. ■

Normalité asymptotique du processus des sommes partielles

Nous avons vu dans la section précédente que le caractère gaussien du processus initial n'était pas maintenu après échantillonnage (voir Proposition 5.2.2). Cependant, nous allons montrer dans cette partie que nous pouvons trouver asymptotiquement un processus gaussien.

On définit

$$S_n(\tau) = \sum_{j=1}^{[n\tau]} Y_j, \quad 0 \leq \tau \leq 1 \quad (5.3.7)$$

le processus des sommes partielles de \mathbf{Y} .

Dans le théorème 5.3.4, on va pouvoir montrer que sous l'hypothèse gaus-

sienne pour X le processus des sommes partielles normalisé converge vers un processus gaussien. Pour prouver la convergence du processus des sommes partielles normalisé, nous devons d'abord obtenir un résultat sur la convergence en probabilité de variance conditionnelle de S_n .

Lemme 5.3.3

Si X est un processus gaussien avec une covariance à variation régulière $\sigma_X(t) = L(t)t^{-1+2d}$, avec $0 < d < 1/2$ et L est à variation lente en l'infini et décroissante à partir d'un certain rang (hypothèse 5.3.1).

Alors, si $\mathbb{E}[T_1] < \infty$,

$$L(n)^{-1}n^{-1-2d}\text{Var}(X_{T_1} + \dots + X_{T_n}|T_1, \dots, T_n) \xrightarrow[n \rightarrow \infty]{\text{proba}} \gamma_d \quad (5.3.8)$$

où

$$\gamma_d := \frac{\mathbb{E}[T_1]^{-1+2d}}{d(1+2d)}. \quad (5.3.9)$$

Démonstration : Voir l'annexe. ■

Théorème 5.3.4

Si X est un processus gaussien à covariance à variation régulière $\sigma_X(t) = L(t)t^{-1+2d}$, avec $0 < d < 1/2$ et L à variation lente en l'infini et décroissante à partir d'un certain rang (hypothèse 5.3.1). Alors, si $\mathbb{E}[T_1] < \infty$, on obtient

$$\gamma_d^{-1/2}L(n)^{-1/2}n^{-1/2-d}S_n(\cdot) \xrightarrow[n \rightarrow \infty]{} B_{\frac{1}{2}+d}(\cdot), \quad (5.3.10)$$

où la convergence a lieu dans $\mathcal{D}[0, 1]$ muni de la norme uniforme, $B_{\frac{1}{2}+d}$ est le mouvement brownien fractionnaire de paramètre $\frac{1}{2} + d$ et γ_d est défini dans (5.3.9).

Démonstration : On commence par prouver la convergence des lois fini dimensionnelles de

$$\gamma_d^{-1/2}L(n)^{-1/2}n^{-1/2-d}S_n(\cdot)$$

vers celles du mouvement brownien fractionnaire $B_{\frac{1}{2}+d}(\cdot)$.

Il suffit de montrer que pour tout $k \geq 1$, $b_1, \dots, b_k \in \mathbb{R}$, $0 \leq t_1, \dots, t_k \leq 1$,

$$A_n := \gamma_d^{-1/2}L(n)^{-1/2}n^{-1/2-d} \sum_{i=1}^k b_i S_n(t_i)$$

vérifie

$$A_n \xrightarrow{Loi} \sum_{i=1}^k b_i B_{\frac{1}{2}+d}(t_i).$$

Si $t_1 = \dots = t_k = 0$, alors

$$\gamma_d^{-1/2} L(n)^{-1/2} n^{-1/2-d} \sum_{i=1}^k b_i S_n(t_i) = \sum_{i=1}^k b_i B_{\frac{1}{2}+d}(t_i) = 0.$$

Sinon, on peut choisir n suffisamment grand pour avoir $[n \max_i(t_i)] \geq 1$ et on pose $T^{(n)} = (T_1, \dots, T_{[n \max_i(t_i)]})$ afin de calculer la fonction caractéristique de A_n

$$\Phi_{A_n}(t) = \mathbb{E}[e^{itA_n}] = \mathbb{E}[e^{-\frac{t^2}{2} \text{Var}(A_n|T^{(n)})}]$$

et

$$\begin{aligned} & \text{Var}(A_n|T^{(n)}) \\ &= \sum_{i,j=1}^k b_i b_j \gamma_d^{-1} L(n)^{-1} n^{-1-2d} \mathbb{E}[S_n(t_i) S_n(t_j)|T^{(n)}] \\ &= \sum_{i,j=1}^k \frac{b_i b_j \gamma_d^{-1} L(n)^{-1} n^{-1-2d}}{2} \left[\text{Var}(S_n(t_i)|T^{(n)}) + \text{Var}(S_n(t_j)|T^{(n)}) - \text{Var}(S_n(t_i) - S_n(t_j)|T^{(n)}) \right] \end{aligned}$$

Par le Lemme 5.3.3,

$$L(n)^{-1} n^{-1-2d} \text{Var}(Y_1 + \dots + Y_n | T_1, \dots, T_n) \xrightarrow[n \rightarrow \infty]{proba} \gamma_d.$$

donc

$$\gamma_d^{-1} L(n)^{-1} n^{-1-2d} \text{Var}(S_n(t_i)|T^{(n)}) \xrightarrow[n \rightarrow \infty]{proba} t_i^{1+2d}$$

pour $t_i > t_j$

$$\begin{aligned} & \gamma_d^{-1} L(n)^{-1} n^{-1-2d} \text{Var}(S_n(t_i) - S_n(t_j)|T^{(n)}) \\ &= \gamma_d^{-1} L(n)^{-1} n^{-1-2d} \text{Var}(Y_{[nt_i]+1} + \dots + Y_{[nt_j]})|T^{(n)}) \\ & \xrightarrow[n \rightarrow \infty]{proba} (t_i - t_j)^{1+2d}. \end{aligned}$$

Finalement,

$$\text{Var}(A_n|T^{(n)}) \xrightarrow[n \rightarrow \infty]{proba} \sum_{i,j=1}^k b_i b_j r_{\frac{1}{2}+d}(t_i, t_j)$$

où $r_{\frac{1}{2}+d}$ est la fonction de covariance du mouvement brownien fractionnaire .

Par le *continuous mapping theorem*, pour $t \in \mathbb{R}$ fixé,

$$\exp\left(-\frac{t^2}{2}\text{Var}(A_n|T^{(n)})\right) \xrightarrow[n \rightarrow \infty]{\text{proba}} \exp\left(-\frac{t^2}{2} \sum_{i,j=1}^k b_i b_j r_{\frac{1}{2}+d}(t_i, t_j)\right)$$

et

$$P\left(\left|\exp\left(-\frac{t^2}{2}\text{Var}(A_n|T^{(n)})\right)\right| \leq 1\right) = 1,$$

ainsi par le théorème de convergence dominée, on obtient

$$\exp\left(-\frac{t^2}{2}\text{Var}(A_n|T^{(n)})\right) \xrightarrow[n \rightarrow \infty]{L^1} \exp\left(-\frac{t^2}{2} \sum_{i,j=1}^k b_i b_j r_{\frac{1}{2}+d}(t_i, t_j)\right).$$

En particulier,

$$\Phi_{A_n}(t) \xrightarrow[n \rightarrow \infty]{} \exp\left(-\frac{t^2}{2} \sum_{i,j=1}^k b_i b_j r_{\frac{1}{2}+d}(t_i, t_j)\right) = \Phi_{\sum_{i=1}^k b_i B_{\frac{1}{2}+d}(t_i)}(t).$$

La suite des processus des sommes partielles $\gamma_d^{-1/2} L(n)^{-1/2} n^{-1/2-d} S_n(\cdot)$ est tendue dans $\mathcal{D}(0, 1)$ muni de la norme uniforme (par application directe de la Proposition 1.2.15) et on obtient ainsi la convergence dans $\mathcal{D}[0, 1]$ muni de la norme uniforme. ■

Résultats statistiques dans le cas de la longue mémoire

En conséquence de ce théorème limite, nous utilisons la méthode d'estimation non paramétrique basée sur la statistique du *re-scaled range* (R/S) pour estimer d à partir du processus échantillonné Y . La statistique R/S est définie comme le quotient entre R_n et S_n où

$$R_n := \max_{1 \leq k \leq n} \sum_{j=1}^k (Y_j - \bar{Y}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (Y_j - \bar{Y}_n) \quad (5.3.11)$$

et

$$S_n := \left(\frac{1}{n} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2 \right)^{1/2}. \quad (5.3.12)$$

Proposition 5.3.1

Sous les mêmes hypothèses que le Théorème 5.3.4, on a

$$\frac{1}{L(n)^{1/2}n^{1/2+d}} \frac{R_n}{S_n} \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{R}(1) := \sqrt{\frac{\gamma_d}{\sigma_X(0)}} \left(\max_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) - \min_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) \right) \quad (5.3.13)$$

où $B_{\frac{1}{2}+d}^0(t) = B_{\frac{1}{2}+d}(t) - tB_{\frac{1}{2}+d}(1)$ est le pont brownien fractionnaire et γ_d est la constante définie dans (5.3.9).

Démonstration : On utilise l'égalité

$$\sum_{j=1}^k (Y_j - \bar{Y}_n) = \sum_{j=1}^k Y_j - \frac{k}{n} \sum_{j=1}^n Y_j = S_n \left(\frac{k}{n} \right) - \frac{k}{n} S_n(1)$$

et la convergence du processus des sommes partielles du Théorème 5.3.4, pour avoir

$$\frac{R_n}{L(n)^{1/2}n^{1/2+d}} \xrightarrow[n \rightarrow \infty]{Loi} \sqrt{\gamma_d} \left(\max_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) - \min_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) \right).$$

Ensuite, nous étudions la convergence en probabilité de

$$S_n^2 = \frac{1}{n} \sum_{j=1}^n Y_j^2 - \left(\frac{1}{n} \sum_{j=1}^n Y_j \right)^2.$$

Comme $\text{Var} \left(\sum_{j=1}^n Y_j \right) \sim Cn^{1+2d}$, on a pour $\varepsilon > 0$

$$P \left(\left| \frac{1}{n} \sum_{j=1}^n Y_j \right| > \varepsilon \right) \leq \frac{1}{n^2 \varepsilon^2} \text{Var} \left(\sum_{j=1}^n Y_j \right) \xrightarrow[n \rightarrow \infty]{} 0$$

et

$$\begin{aligned} P \left(\left| \frac{1}{n} \sum_{j=1}^n Y_j^2 - \sigma_X(0) \right| > \varepsilon \right) &\leq \frac{1}{n^2 \varepsilon^2} \text{Var} \left(\sum_{j=1}^n Y_j^2 \right) \\ &= \frac{1}{n^2 \varepsilon^2} \sum_{j=1}^n \sum_{k=1}^n \text{Cov} \left(Y_j^2, Y_k^2 \right) \\ &= \frac{1}{n^2 \varepsilon^2} \sum_{j=1}^n \sum_{k=1}^n \left(\mathbb{E} \left[\mathbb{E} [X_{T_j}^2 X_{T_k}^2 | T_j, T_k] \right] - \sigma_X(0)^2 \right) \end{aligned}$$

Pour $(s, t) \in (\mathbb{R}^+)^2$, on décompose X_s^2 et X_t^2 dans le système orthogonal complet

des polynômes d'Hermite $(H_k)_{k \geq 0}$, $\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right)^2 = H_0\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right) + H_2\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right)$,
alors

$$\begin{aligned} \frac{\mathbb{E}[X_s^2 X_t^2]}{\sigma_X(0)^2} &= \mathbb{E}\left[H_0\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right) H_0\left(\frac{X_t}{\sqrt{\sigma_X(0)}}\right)\right] + \mathbb{E}\left[H_2\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right) H_0\left(\frac{X_t}{\sqrt{\sigma_X(0)}}\right)\right] \\ &+ \mathbb{E}\left[H_0\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right) H_2\left(\frac{X_t}{\sqrt{\sigma_X(0)}}\right)\right] + \mathbb{E}\left[H_2\left(\frac{X_s}{\sqrt{\sigma_X(0)}}\right) H_2\left(\frac{X_t}{\sqrt{\sigma_X(0)}}\right)\right] \end{aligned}$$

En utilisant la propriété d'orthogonalité des polynômes d'Hermite pour un couple gaussien à variances unitaires (voir par exemple Prop 2.4.1 de [Gir+12]), on obtient

$$\begin{aligned} \mathbb{E}[X_s^2 X_t^2] &= \sigma_X^2(0) \left[1 + 2\text{Cov}^2\left(\frac{X_s}{\sqrt{\sigma_X(0)}}, \frac{X_t}{\sqrt{\sigma_X(0)}}\right)\right] \\ &= \sigma_X^2(0) + 2\sigma_X^2(t-s) \end{aligned}$$

Finalement,

$$\begin{aligned} P\left(\left|\frac{1}{n} \sum_{j=1}^n Y_j^2 - \sigma_X(0)\right| > \varepsilon\right) &\leq \frac{2}{n^2 \varepsilon^2} \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}\left[\sigma_X^2(T_j - T_k)\right] \\ &= \frac{4}{n^2 \varepsilon^2} \sum_{j=0}^{n-1} (n-j) \mathbb{E}\left[\sigma_X^2(T_j)\right] \end{aligned}$$

Si $0 \leq d \leq 1/4$, $\frac{1}{n^2} \sum_{j=0}^{n-1} (n-j) \mathbb{E}\left[\sigma_X^2(T_j)\right] \leq \frac{1}{n} \sum_{j=0}^{\infty} \mathbb{E}\left[\sigma_X^2(T_j)\right] \xrightarrow{n \rightarrow \infty} 0$ en appliquant la Proposition 5.2.3 avec $p = 1$ à la fonction σ_X^2 . Si $1/4 < d < 1/2$, on utilise le Théorème 5.3.2 qui donne

$$\frac{1}{n^2} \sum_{j=0}^{n-1} (n-j) \mathbb{E}\left[\sigma_X^2(T_j)\right] \sim C n^{-2+4d}$$

et on obtient dans les deux cas que

$$P\left(\left|\frac{1}{n} \sum_{j=1}^n Y_j^2 - \sigma_X(0)\right| > \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

On conclut que $S_n \xrightarrow[n \rightarrow \infty]{\text{proba}} \sqrt{\sigma_X(0)}$ et

$$\frac{1}{L(n)^{1/2} n^{1/2+d}} \frac{R_n}{S_n} \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{R}(1) := \sqrt{\frac{\gamma d}{\sigma_X(0)}} \left(\max_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) - \min_{0 \leq t \leq 1} B_{\frac{1}{2}+d}^0(t) \right) \quad \blacksquare$$

Dans le cas où $L(t) = c > 0$ pour tout $t > t_0$, en prenant le logarithme dans l'expression (5.3.13), on obtient l'identité heuristique suivante :

$$\log\left(\frac{R_n}{S_n}\right) \sim (1/2 + d) \log(n) + \log(\sqrt{c}\mathcal{R}(1))$$

On estime alors la pente de la droite de régression à partir des points $(\log(n), \log(R_n/S_n))$ qui fournit une estimation R/S de d . Remarquons que dans le cas plus général où L est une fonction à variation lente décroissante à partir d'un certain rang, on a $\log\left(\frac{R_n}{S_n}\right) \sim (1/2 + d) \log(n) + \log(L(n))/2 + \log(\mathcal{R}(1))$ et $\log(L(n))$ est négligeable comparé à $\log(n)$.

Dans la Figure 5.2, nous avons représenté plusieurs réalisations de la statistique R/S d'un processus échantillonné à partir d'inter-arrivées i.i.d. $\mathcal{E}(1)$ et d'un processus initial gaussien de covariance $\sigma_X(t) = (1 + |t|)^{-1+2d}$ où on fait varier la valeur de d .

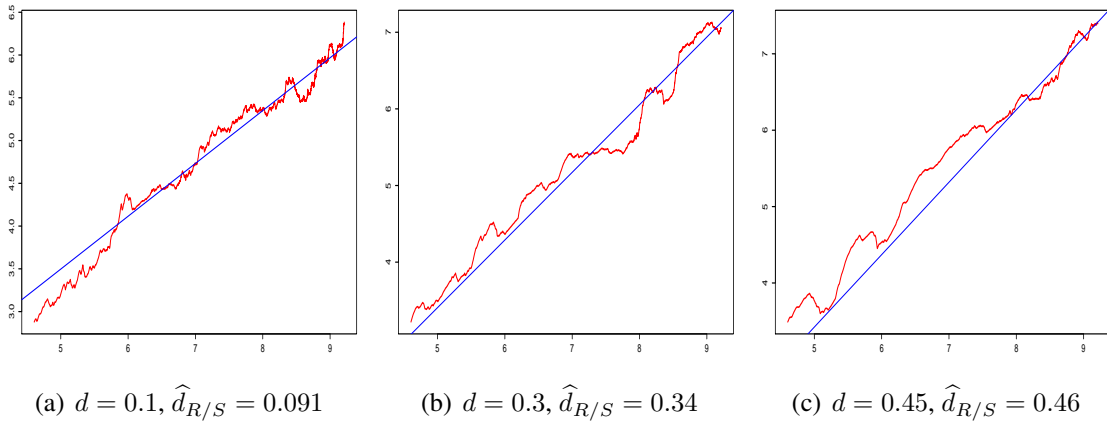


FIGURE 5.2 – Réalisations de trajectoires de la statistique $n \rightarrow \log(R/S(n))$ (en rouge) pour différentes valeurs de paramètre de mémoire avec des intervalles Δ_j ayant une loi exponentielle de moyenne 1 et un processus initial gaussien de fonction d'autocovariance $\sigma_X(t) = (1 + |t|)^{-1+2d}$. En bleu, on représente la droite de régression dont la pente correspond à $\hat{d}_{R/S} + 1/2$. Les axes des abscisses sont en échelle logarithmique.

5.3.2 Perte de mémoire lorsque $\mathbb{E}[T_1] = \infty$

Le phénomène est le même que dans le cas discret (voir [PV10]) : à partir d'un processus à longue mémoire, une loi d'échantillonnage à queue lourde peut conduire à un processus à courte mémoire.

Proposition 5.3.2

Supposons que la fonction d'autocovariance de \mathbf{X} vérifie

$$|\sigma_X(t)| \leq c \min(1, t^{-1+2d}) \quad \forall t \in \mathbb{R}^+ \quad (5.3.14)$$

où $0 < d < 1/2$. Si pour un certain $\beta \in]0, 1[$

$$\liminf_{x \rightarrow \infty} (x^\beta P(T_1 > x)) > 0 \quad (5.3.15)$$

(ce qui entraîne $\mathbb{E}[T_1^\beta] = \infty$) alors

$$|\sigma_Y(h)| \leq Ch^{\frac{-1+2d}{\beta}}. \quad (5.3.16)$$

Démonstration : De l'hypothèse (5.3.14), on a

$$|\sigma_Y(h)| \leq \mathbb{E}[|\sigma_X(T_h)|] \leq c\mathbb{E}[\min\{1, T_h^{-1+2d}\}]$$

Ainsi si on pose S^{*h} la fonction de répartition de T_h , on a par intégration par parties

$$\begin{aligned} \mathbb{E}[\min\{1, T_h^{-1+2d}\}] &= \int_0^1 dS^{*h}(x) + \int_1^\infty x^{-1+2d} dS^{*h}(x) \\ &= S^{*h}(1) + (1-2d) \int_1^\infty x^{-2+2d} S^{*h}(x) dx - S^{*h}(1) \\ &= (1-2d) \int_1^\infty x^{-2+2d} S^{*h}(x) dx. \end{aligned} \quad (5.3.17)$$

De l'hypothèse (5.3.15) sur la queue de la loi d'échantillonnage, il existe $C > 0$ et $x_0 \geq 1$ tels que

$$\forall x \geq x_0, \quad P(T_1 > x) \geq Cx^{-\beta}.$$

De plus, pour $x \in [1, x_0]$,

$$x^\beta P(T_1 > x) \geq P(T_1 > x_0) \geq Cx_0^{-\beta}.$$

On obtient $\forall x \geq 1$, $P(T_1 > x) \geq \tilde{C}x^{-\beta}$ avec $\tilde{C} = Cx_0^{-\beta}$ et donc

$$\begin{aligned} S^{*h}(x) &= P(T_h \leq x) \leq P\left(\max_{0 \leq l \leq h-1} \Delta_l \leq x\right) = P(T_1 \leq x)^h \\ &\leq \left(1 - \tilde{C}x^{-\beta}\right)^h \leq e^{-\frac{\tilde{C}h}{x^\beta}}. \end{aligned} \quad (5.3.18)$$

En rassemblant les équations (5.3.17) et (5.3.18), on aboutit à

$$\begin{aligned} \mathbb{E}[\min\{1, T_h^{-1+2d}\}] &\leq (1 - 2d) \int_1^\infty x^{-2+2d} e^{-\frac{\tilde{C}h}{x^\beta}} dx \\ &= \frac{1 - 2d}{\beta} h^{-(1-2d)/\beta} \int_0^h u^{(1-2d)/\beta-1} e^{-\tilde{C}u} du \end{aligned}$$

et le résultat en découle puisque

$$\int_0^h u^{(1-2d)/\beta-1} e^{-\tilde{C}u} du \xrightarrow{h \rightarrow \infty} \int_0^\infty u^{(1-2d)/\beta-1} e^{-\tilde{C}u} du. \quad \blacksquare$$

La proposition suivante stipule que la borne pour le paramètre de mémoire du processus échantillonné trouvée dans la Proposition 5.3.2 est optimale sous certaines hypothèses supplémentaires.

Proposition 5.3.3

Supposons que

$$\sigma_X(t) = t^{-1+2d}L(t)$$

où $0 < d < 1/2$ et L est à variation lente en l'infini et monotone à partir d'un certain rang. Si

$$\beta := \sup\{\gamma : \mathbb{E}[T_1^\gamma] < \infty\} \in]0, 1[\quad (5.3.19)$$

alors, pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ tel que

$$\sigma_Y(h) \geq C_\varepsilon h^{-\frac{1-2d}{\beta}-\varepsilon}, \quad \forall h \geq 1. \quad (5.3.20)$$

Démonstration : Soit $\varepsilon > 0$. On a

$$\frac{\sigma_X(T_h)}{h^{-\frac{1-2d}{\beta}-\varepsilon}} = \frac{T_h^{-1+2d}}{h^{-\frac{1-2d}{\beta}-\varepsilon}} L(T_h) = \frac{T_h^{-1+2d-\frac{\beta\varepsilon}{2}}}{h^{-\frac{1-2d}{\beta}-\varepsilon}} T_h^{\frac{\beta\varepsilon}{2}} L(T_h) = \left(\frac{T_h}{h^\delta}\right)^{-1+2d-\frac{\beta\varepsilon}{2}} T_h^{\frac{\beta\varepsilon}{2}} L(T_h)$$

où

$$\delta = \frac{(1-2d)/\beta + \varepsilon}{1-2d + \frac{\beta\varepsilon}{2}} = \frac{1}{\beta} \left(\frac{1-2d + \beta\varepsilon}{1-2d + \beta\varepsilon/2} \right).$$

En utilisant la Proposition 1.3.6 de [Bin+89], on a

$$T_h^{\frac{\beta\varepsilon}{2}} L(T_h) \xrightarrow[h \rightarrow \infty]{p.s.} +\infty$$

De plus, $\delta > \frac{1}{\beta}$. Cela implique en utilisant (5.3.19) que $\mathbb{E}[T_1^{1/\delta}] < \infty$. Ensuite, la loi forte des grands nombres de Marcinkiewicz-Zygmund (voir [Sto74] Théorème 3.2.3) conduit à :

$$\frac{T_h}{h^\delta} \xrightarrow[h \rightarrow \infty]{p.s.} 0. \quad (5.3.21)$$

Enfin en appliquant le Lemme de Fatou :

$$\frac{\sigma_Y(h)}{h^{-\frac{1-2d}{\beta}-\varepsilon}} \xrightarrow[h \rightarrow \infty]{} \infty. \quad \blacksquare$$

5.4 Comparaison entre le sous échantillonnage et l'agrégation temporelle

Les mêmes notations que dans les sections précédentes sont conservées : on part de $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+}$, un processus centré stationnaire du second ordre à temps continu de fonction d'autocovariance $\sigma_X(\cdot)$ et d'une marche aléatoire $(T_n)_{n \geq 0}$ indépendante de \mathbf{X} . Les intervalles d'échantillonnage $T_{j+1} - T_j = \Delta_j$ sont i.i.d. de densité de probabilité commune s à support sur \mathbb{R}^+ . De plus, $T_0 = 0$.

Nous voulons considérer le processus agrégé $\int_{T_h}^{T_{h+1}} X_t dt$ comme une variable aléatoire. Pour cela, nous avons besoin d'hypothèses supplémentaires sur l'intégrabilité de \mathbf{X} pour garantir que $\int_a^b |X_t| dt < \infty$ pour $0 \leq a < b < \infty$ et sur la mesurabilité de \mathbf{X} . Si nous supposons que \mathbf{X} a des trajectoires continues ([MS70; Kub70; Kaw81]) ou que $\sup_{t \in [a,b]} E(|X_t|) < \infty$ pour tout $0 \leq a < b < \infty$ et si le processus $X : [0, \infty[\times \Omega \rightarrow \mathbb{R}$ est conjointement mesurable, l'intégrale stochastique $\int_a^b X_t dt$ est bien définie et est une variable aléatoire.

Dans cette section, nous supposons que ces conditions sont remplies et nous définissons le processus agrégé comme suit

$$Z_h = \int_{T_h}^{T_{h+1}} X_t dt, \quad h \in \mathbb{N}. \quad (5.4.1)$$

Ce processus agrégé est bien défini puisque $P(0 \leq T_h < T_{h+1} < \infty) = 1$. Dans [TC05c], le processus est agrégé sur des intervalles de la forme $[h\Delta, (h+1)\Delta]$ où Δ est un pas déterministe. La structure de corrélation de la suite r -différenciée $\left\{ \nabla^r \left(\int_{h\Delta}^{(h+1)\Delta} L(u) du \right) \right\}_{h \in \mathbb{N}}$ est étudiée en supposant que le processus de la r -ième dérivé de L est un modèle CARFIMA(p, H, q). En particulier, ils donnent des résultats sur le comportement de la structure de corrélation lorsque $\Delta \rightarrow \infty$.

Dans cette partie, nous comparons les fonctions d'autocovariance des processus agrégés et échantillonnés.

Lemme 5.4.1

Le processus agrégé défini dans (5.4.1) est un processus stationnaire du second ordre et sa fonction d'autocovariance est donnée par

$$\sigma_Z(0) = 2 \iint_{0 \leq s < t} \sigma_X(t-s) P(t < T_1) dt ds$$

et pour $h \geq 1$,

$$\sigma_Z(h) = \iint_{0 \leq s < t} \sigma_X(t-s) P(0 \leq s < T_1, T_h \leq t < T_{h+1}) dt ds.$$

Démonstration : On a

$$\begin{aligned} \mathbb{E}[Z_h] &= \mathbb{E} \left[\int_{T_h}^{T_{h+1}} X_t dt \right] = \int \mathbb{E}[X_t \mathbf{1}_{T_h \leq t < T_{h+1}}] dt \\ &= \int \mathbb{E}[X_t] \mathbb{E}[\mathbf{1}_{T_h \leq t < T_{h+1}}] dt = 0 \end{aligned}$$

car X et T sont indépendants.

Soit $p \in \mathbb{N}$ et $h \geq 1$,

$$\mathbb{E}[Z_{h+p}Z_p] \tag{5.4.2}$$

$$\begin{aligned} &= \mathbb{E} \left[\int_{T_p}^{T_{p+1}} X_s \, ds \int_{T_{h+p}}^{T_{h+p+1}} X_t \, dt \right] \\ &= \iint \mathbb{E}[X_s X_t \mathbf{1}_{T_p \leq s < T_{p+1}, T_{h+p} \leq t < T_{h+p+1}}] \, dt \, ds \\ &= \iint_{s < t} \sigma_X(t-s) P(T_p \leq s < T_{p+1}, T_{h+p} \leq t < T_{h+p+1}) \, dt \, ds \\ &= \mathbb{E} \left[\mathbb{E} \left[\iint_{s < t} \sigma_X(t-s) \mathbf{1}_{0 \leq s - T_p < \Delta_p, \Delta_p + \dots + \Delta_{p+h-1} \leq t - T_p < \Delta_p + \dots + \Delta_{p+h}} \, du \, dv \middle| T_p \right] \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\iint_{u < v} \sigma_X(v-u) \mathbf{1}_{0 \leq u < \Delta_p, \Delta_p + \dots + \Delta_{p+h-1} \leq v < \Delta_p + \dots + \Delta_{p+h}} \, dt \, ds \middle| T_p \right] \right] \\ &= \iint_{u < v} \sigma_X(v-u) P(0 \leq u < \Delta_p, \Delta_p + \dots + \Delta_{p+h-1} \leq v < \Delta_p + \dots + \Delta_{p+h}) \, du \, dv \\ &= \iint_{u < v} \sigma_X(v-u) P(0 \leq u < T_1, T_h \leq v < T_{h+1}) \, du \, dv \tag{5.4.3} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[Z_p^2] &= \mathbb{E} \left[\int_{T_p}^{T_{p+1}} X_s \, ds \int_{T_p}^{T_{p+1}} X_t \, dt \right] \\ &= 2 \iint_{s < t} \mathbb{E}[X_s X_t \mathbf{1}_{T_p \leq s < t < T_{p+1}}] \, dt \, ds \\ &= 2 \iint_{0 \leq u < v} \sigma_X(v-u) P(v < T_1) \, du \, dv \tag{5.4.4} \end{aligned}$$

Comme (5.4.3) et (5.4.4) ne dépendent pas de p , la stationnarité du processus est prouvée. ■

Exemple 5.4.1. Échantillonnage par un processus de Poisson

Dans le cas d'un échantillonnage de Poisson, les intervalles entre deux observations $T_{j+1} - T_j = \Delta_j$ sont i.i.d. de loi commune une loi exponentielle. Soit N le processus de comptage associé à la marche aléatoire $(T_n)_{n \geq 0}$:

$$N(t) = \sum_{j \geq 1} \mathbf{1}_{T_j \leq t}.$$

Le processus N est à accroissements stationnaires et indépendants et, pour tout t fixé, $N(t)$ suit une loi de Poisson de paramètre $t\ell$ où $\ell = \mathbb{E}[T_1]^{-1}$. On pose $p(t, k) = P(N(t) = k)$.

Pour $h \geq 1$,

$$\begin{aligned}\sigma_Z(h) &= \iint_{0 \leq s < t} \sigma_X(t-s) P(N(s)=0, N(t)-N(s)=h) dt ds \\ &= \iint_{0 \leq s < t} \sigma_X(t-s) p(s,0) p(t-s, h) dt ds \\ &= \iint_{\mathbb{R}_+ \times \mathbb{R}_+} \sigma_X(x) p(y,0) p(x, h) dx dy \\ &= \int_{\mathbb{R}_+} \sigma_X(x) p(x, h) dx \int_{\mathbb{R}_+} p(y,0) dy.\end{aligned}$$

De plus

$$\int_{\mathbb{R}_+} p(y,0) dy = \int_{\mathbb{R}_+} P(T_1 > y) dy = \mathbb{E}[T_1] = \ell^{-1}$$

Ainsi, on a

$$\sigma_Z(h) = \ell^{-1} \int_{\mathbb{R}_+} \sigma_X(x) e^{-\ell x} \ell^h x^h \frac{1}{h!} dx$$

Comme T_{h+1} suit une loi gamma distribution de paramètre $(h+1, \lambda)$, on a

$$\sigma_Z(h) = \ell^{-2} \mathbb{E}[\sigma_X(T_{h+1})]$$

conduisant à une forme explicite de la fonction d'autocovariance pour le processus agrégé avec un échantillonnage poissonnien

$$\sigma_Z(h) = \mathbb{E}[T_1]^2 \mathbb{E}[\sigma_X(T_{h+1})] = \mathbb{E}[T_1]^2 \sigma_Y(h+1), \quad \text{pour } h \geq 1 \quad (5.4.5)$$

et avec le même calcul

$$\sigma_Z(0) = 2\mathbb{E}[T_1]^2 \mathbb{E}[\sigma_X(T_1)] = 2\mathbb{E}[T_1]^2 \sigma_Y(1). \quad (5.4.6)$$

Proposition 5.4.1

Si \mathbf{X} est un processus gaussien alors les marginales du processus agrégé \mathbf{Z} sont identiquement distribuées mais non gaussiennes. En particulier, \mathbf{Z} n'est pas un processus gaussien.

Démonstration : Commençons par montrer que $Z_0 = \int_0^{T_1} X_t dt$ n'est pas gaussien.

Sachant T_1 , la loi conditionnelle de Z_0 est une loi normale centrée de variance $\int_0^{T_1} \int_0^{T_1} \sigma_X(t-s) dt ds$. On peut alors en déduire la densité bivariée de (Z_0, T_1) :

$$f_{(Z_0, T_1)}(x, y) = \frac{\exp\left(-\frac{x^2}{2 \int_0^y \int_0^y \sigma_X(t-s) dt ds}\right)}{\sqrt{2\pi \int_0^y \int_0^y \sigma_X(t-s) dt ds}} s(y),$$

où s est la densité de la variable aléatoire T_1 . Posons $V = \int_0^{T_1} \int_0^{T_1} \sigma_X(t-s) dt ds$, on a

$$f_{Z_0}(x) = \mathbb{E} \left[\frac{\exp\left(-\frac{x^2}{2V}\right)}{\sqrt{2\pi V}} \right].$$

On a $\text{Var}(Z_0) = \mathbb{E}[V]$. Si $\mathbb{E}[V] = \infty$, Z_0 n'est pas gaussien puisque les lois normales admettent une variance finie.

On suppose désormais que $\mathbb{E}[V] < \infty$. Soit W une variable aléatoire gaussienne centrée de variance $\mathbb{E}[V]$. La densité de W est alors

$$f_W(x) = \frac{\exp\left(-\frac{x^2}{2\mathbb{E}[V]}\right)}{\sqrt{2\pi\mathbb{E}[V]}}.$$

La variable aléatoire Z_0 est ainsi gaussienne si et seulement si pour presque tout $x \in \mathbb{R}$,

$$\mathbb{E} \left[\frac{\exp\left(-\frac{x^2}{2V}\right)}{\sqrt{V}} \right] = \frac{\exp\left(-\frac{x^2}{2\mathbb{E}[V]}\right)}{\sqrt{\mathbb{E}[V]}}.$$

La variable aléatoire V est positive presque sûrement. Pour $x \in \mathbb{R}$ et $t > 0$, on définit $\Psi_x(t) = \frac{e^{-\frac{x^2}{2t}}}{\sqrt{t}}$, Z_0 est gaussienne si et seulement si pour presque tout $x \in \mathbb{R}^*$, $\Psi_x(\mathbb{E}[V]) = \mathbb{E}[\Psi_x(V)]$. On peut remarquer que $\mathbb{E}[\Psi_x(V)]$ est bien définie et finie pour $x \neq 0$ puisque $0 < \Psi_x(V) < \frac{e^{-1/2}}{x}$ p.s.

- Si $\mathbb{E} \left[\frac{1}{\sqrt{V}} \right] < \infty$, posons $h(x) = \mathbb{E}[\Psi_x(V)] - \Psi_x(\mathbb{E}[V])$. En utilisant, pour $y > 0$ fixé, la continuité de $x \rightarrow \Psi_x(y)$ et la relation $|\Psi_x(y)| \leq \frac{1}{\sqrt{y}}$, on en déduit du théorème de continuité des intégrales à paramètres, la continuité de h sur \mathbb{R} . De plus $h(0) = \mathbb{E} \left[\frac{1}{\sqrt{V}} \right] - \frac{1}{\sqrt{\mathbb{E}[V]}} > 0$ par l'inégalité de Jensen puisque V n'est pas constante p.s. Ainsi, il existe $\varepsilon > 0$ tel que pour tout $x \in [-\varepsilon, \varepsilon]$, $\mathbb{E}[\Psi_x(V)] > \Psi_x(\mathbb{E}[V])$.
- Si $\mathbb{E} \left[\frac{1}{\sqrt{V}} \right] = \infty$. Alors, on peut trouver $M > 0$ tel que $\mathbb{E} \left[\frac{1}{\sqrt{V}} \mathbf{1}_{V \geq M} \right] > \frac{e^{1/2}}{\sqrt{\mathbb{E}[V]}}$.
On a alors pour $x \in]0, \sqrt{M}]$,

$$\mathbb{E}[\Psi_x(V)] \geq \mathbb{E}[\Psi_x(V) \mathbf{1}_{V \geq M}] \geq e^{-\frac{x^2}{2M}} \mathbb{E} \left[\frac{1}{\sqrt{V}} \mathbf{1}_{V \geq M} \right] > \frac{1}{\sqrt{\mathbb{E}[V]}} > \Psi_x(\mathbb{E}[V]).$$

On en déduit que Z_0 n'est pas une variable gaussienne.

Finalement, pour $h \geq 1$, $Z_h = \int_{T_h}^{T_h + \Delta_h} X_t dt$. Par les mêmes arguments que pour Z_0 , on a

$$\begin{aligned} f_{(Z_h, T_h, \Delta_h)}(x, y, z) &= \frac{\exp\left(-\frac{x^2}{2 \int_y^{y+z} \int_y^{y+z} \sigma_X(t-s) dt ds}\right)}{\sqrt{2\pi \int_y^{y+z} \int_y^{y+z} \sigma_X(t-s) dt ds}} f_{T_h}(y) s(z) \\ &= \frac{\exp\left(-\frac{x^2}{2 \int_0^z \int_0^z \sigma_X(t-s) dt ds}\right)}{\sqrt{2\pi \int_0^z \int_0^z \sigma_X(t-s) dt ds}} f_{T_h}(y) s(z) \end{aligned}$$

On en déduit $f_{Z_h}(x) = \mathbb{E}[\Psi_x(V)] = f_{Z_0}(x)$. Finalement, $\forall h \geq 0$, les variables aléatoires Z_h sont identiquement distribuées avec une loi non-gaussienne. ■

Lemme 5.4.2

Si σ_X est décroissante sur \mathbb{R}^+ , alors pour $h \geq 1$

$$\sigma_Z(h) \leq \mathbb{E}[T_1]^2 \sigma_Y(h-1) \quad (5.4.7)$$

Démonstration : Dans la partie droite de la relation

$$\sigma_Z(h) = \mathbb{E}\left[\int_{s < t} \sigma_X(t-s) \mathbf{1}_{0 \leq s < T_1 \leq T_h \leq t < T_{h+1}} dt ds\right],$$

on a $T_h - T_1 \leq t - s \leq T_{h+1}$ ce qui conduit à

$$\sigma_Z(h) \leq \mathbb{E}[T_1(T_{h+1} - T_h) \sigma_X(T_h - T_1)].$$

Finalement, $T_{h+1} - T_h$ est indépendant de (T_1, T_h) et $T_h - T_1$ est indépendant de T_1 . Le résultat en découle. ■

Grâce au Lemme 5.4.2, certaines propriétés du processus échantillonné peuvent être directement transposées au cas du processus agrégé. La proposition suivante les réunit.

Proposition 5.4.2

1. Si σ_X est bornée par une fonction strictement positive, décroissante et intégrable et si $\mathbb{E}[T_1] < \infty$ alors $\sum_{h \geq 0} |\sigma_Z(h)| < \infty$.
2. Si $\mathbb{E}[T_1] < \infty$ et $\sigma_X(t) = t^{-1+2d} L(t)$ où L est une fonction à variation

lente en l'infini décroissante et $0 < d < 1/2$ alors,

$$\sigma_Z(h) \leq \sigma_1(h) \sim \mathbb{E}[T_1]^{1+2d} h^{-1+2d} L(h) \quad h \rightarrow \infty$$

3. Supposons $\liminf_{x \rightarrow \infty} [x^\beta P(T_1 > x)] > 0$ pour un certain $\beta \in (0, 1)$ (ce qui implique $\mathbb{E}[T_1^\beta] = \infty$). Si σ_X est décroissante sur \mathbb{R}^+ et $|\sigma_X(t)| \leq c \min(1, t^{-1+2d})$ avec $0 < d < 1/2$, alors

$$|\sigma_Z(h)| \leq C h^{-(1-2d)/\beta}.$$

Démonstration : Ces résultats sont des conséquences immédiates du Lemme 5.4.2 et de, respectivement, la Proposition 5.2.3 avec $p = 1$, le théorème 5.3.2 et la Proposition 5.3.2. ■

5.5 Annexe

Preuve du Lemme 5.3.3.

Pour la preuve du Lemme 5.3.3, nous avons besoin du résultat intermédiaire suivant :

Lemme 5.5.1

Si $\mathbb{E}[T_1] < \infty$ et si \mathbf{X} a une fonction d'autocovariance à variation régulière

$$\sigma_X(t) = L(t)t^{-1+2d}$$

avec $0 < d < 1/2$ et L une fonction à variation lente à l'infini et décroissante à partir d'un certain rang. Alors,

$$\text{Var}(\sigma_X(T_h)) = o(L(h)^2 h^{-2+4d}) \quad \text{lorsque } h \rightarrow +\infty \quad (5.5.1)$$

Démonstration : Par le Théorème 5.3.2, on a $\mathbb{E}[\sigma_X(T_h)] \underset{h \rightarrow \infty}{\sim} L(h)(h\mathbb{E}[T_1])^{-1+2d}$.

Pour obtenir le résultat, il suffit donc de prouver que

$$\mathbb{E}[\sigma_X(T_h)^2] \underset{h \rightarrow \infty}{\sim} L(h)^2 (h\mathbb{E}[T_1])^{-2+4d}.$$

Afin de montrer le comportement asymptotique de $\mathbb{E}[\sigma_X(T_h)^2]$, nous allons suivre la même preuve que pour le Théorème 5.3.2 :

- Soit $0 < c < \mathbb{E}[T_1]$, et $h \in \mathbb{N}$ tel que $ch \geq 1$,

$$\mathbb{E}[\sigma_X(T_h)^2] \geq \mathbb{E}[\sigma_X(T_h)^2 \mathbf{1}_{T_h > ch}] \geq \mathbb{E}[L(T_h)^2 T_h^{-2+4d} \mathbf{1}_{T_h > ch}] \geq \inf_{t > ch} \{L(t)^2 t^{4d}\} \mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h^2}\right]$$

Grâce aux inégalités de Jensen et d'Hölder,

$$\mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h^2}\right] \geq \mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h}\right]^2 \text{ et } P(T_h > ch)^2 \leq \mathbb{E}[T_h] \mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h}\right],$$

ce qui donne

$$\mathbb{E}\left[\frac{\mathbf{1}_{T_h > ch}}{T_h^2}\right] \geq \frac{P(T_h > ch)^4}{\mathbb{E}[T_h]^2}.$$

d'où

$$\frac{\mathbb{E}[\sigma_X(T_h)^2]}{L(h)^2 (h\mathbb{E}[T_1])^{-2+4d}} \geq \frac{\inf_{t > ch} \{L(t)^2 t^{4d}\}}{L(h)^2 h^{4d} \mathbb{E}[T_1]^{4d}} P(T_h > ch)^4. \quad (5.5.2)$$

Ensuite, pour $c < \mathbb{E}[T_1]$, on a $P(T_h > ch) \rightarrow 1$ et $\inf_{t > ch} \{L(t)^2 t^{4d}\} \sim L(ch)^2 (ch)^{4d}$.
Finalement, pour tout $c < \mathbb{E}[T_1]$,

$$\liminf_{h \rightarrow \infty} \frac{\mathbb{E}[\sigma_X(T_h)^2]}{L(h)^2 (h\mathbb{E}[T_1])^{-2+4d}} \geq \left(\frac{c}{\mathbb{E}[T_1]}\right)^{4d}$$

En prenant la limite lorsque $c \rightarrow \mathbb{E}[T_1]$, on obtient

$$\liminf_{h \rightarrow \infty} \frac{\mathbb{E}[\sigma_X(T_h)^2]}{L(h)^2 (h\mathbb{E}[T_1])^{-2+4d}} \geq 1$$

- Soit $\frac{1}{2} < s < \tau < 1$, t_0 tel que $L(\cdot)$ soit décroissante et positive sur $[t_0, \infty)$ et h tel que $\mu_{h,s} - \mu_{h,s}^\tau \geq t_0$ (les notations correspondent à celles utilisées dans le Théorème 5.3.2),

$$\begin{aligned} \mathbb{E}[\sigma_X(T_h)^2] &= \mathbb{E}\left[L(T_h)^2 T_h^{-2+4d} \mathbf{1}_{T_{h,s} \geq \mu_{h,s} - \mu_{h,s}^\tau}\right] + \mathbb{E}\left[\sigma(T_h)^2 \mathbf{1}_{T_{h,s} < \mu_{h,s} - \mu_{h,s}^\tau}\right] \\ &\leq L(\mu_{h,s} - \mu_{h,s}^\tau)^2 (\mu_{h,s} - \mu_{h,s}^\tau)^{-2+4d} + \sigma_X(0)^2 P(T_{h,s} < \mu_{h,s} - \mu_{h,s}^\tau) \end{aligned}$$

$$\begin{aligned} \frac{\mathbb{E}[\sigma_X(T_h)^2]}{L(h)^2(h\mathbb{E}[T_1])^{-2+4d}} &\leq \left(\frac{L(\mu_{h,s} - \mu_{h,s}^\tau)}{L(h)} \right)^2 \left(\frac{\mu_{h,s} - \mu_{h,s}^\tau}{h\mathbb{E}[T_1]} \right)^{-2+4d} \\ &\quad + \sigma_X(0)^2 \frac{P(T_{h,s} < \mu_{h,s} - \mu_{h,s}^\tau)}{L(h)^2(h\mathbb{E}[T_1])^{-2+4d}} \end{aligned}$$

On a donc

$$\limsup_{h \rightarrow \infty} \frac{\mathbb{E}[\sigma_X(T_h)^2]}{L(h)^2(h\mathbb{E}[T_1])^{-2+4d}} \leq 1. \quad \blacksquare$$

On retourne à la preuve du Lemme 5.3.3. Posons

$$W_n = L(n)^{-1} n^{-1-2d} \sum_{i=1}^n \sum_{j=1}^n \sigma_X(T_j - T_i) = L(n)^{-1} n^{-1-2d} \text{Var}(X_{T_1} + \dots + X_{T_n} | T_1, \dots, T_n).$$

On cherche à prouver que W_n converge en probabilité vers γ_d . Pour cela, on montre que $\mathbb{E}[W_n] \xrightarrow[n \rightarrow \infty]{} \gamma_d$ et $\text{Var}(W_n) \xrightarrow[n \rightarrow \infty]{} 0$.

• Comme \mathbf{X} est un processus centré $E[W_n] = L(n)^{-1} n^{-1-2d} \text{Var}(Y_1 + \dots + Y_n)$. Or par le Théorème 5.3.2,

$$\sigma_Y(h) \sim L(h)(h\mathbb{E}[T_1])^{-1+2d} \quad h \rightarrow \infty,$$

d'où

$$L(n)^{-1} n^{-1-2d} \text{Var}(Y_1 + \dots + Y_n) \xrightarrow[n \rightarrow \infty]{} \gamma_d \quad (5.5.3)$$

(voir par exemple [Gir+12] Prop 3.3.1 p.43)

et on obtient

$$E[W_n] \xrightarrow[n \rightarrow \infty]{} \gamma_d.$$

• De plus,

$$\begin{aligned} \text{Var}(W_n) &= L(n)^{-2} n^{-2-4d} \text{Var} \left(\sum_{i=1}^n \sum_{j=1}^n \sigma_X(T_j - T_i) \right) \\ &\leq L(n)^{-2} n^{-2-4d} \left(\sum_{i=1}^n \sum_{j=1}^n \sqrt{\text{Var}(\sigma_X(T_j - T_i))} \right)^2 \\ &= \left(2n^{-1-2d} L(n)^{-1} \sum_{h=1}^n (n-h) \sqrt{\text{Var}(\sigma_X(T_h))} \right)^2 \end{aligned}$$

Par le Lemme 5.5.1, $\sqrt{\text{Var}(\sigma_X(T_h))} = o(L(h)h^{-1+2d})$ et $2 \sum_{h=1}^n (n-h)L(h)h^{-1+2d} \sim \frac{L(n)n^{1+2d}}{d(1+2d)}$. On a alors

$$2 \sum_{h=1}^n (n-h)\sqrt{\text{Var}(\sigma_X(T_h))} = o(L(n)n^{1+2d})$$

Au final, $\text{Var}(W_n) = o(1)$ ce qui signifie que $\text{Var}(W_n) \xrightarrow[n \rightarrow \infty]{} 0$. On obtient

$$W_n \xrightarrow[n \rightarrow \infty]{L^2, p} \gamma_d.$$



6

Étude spectrale de l'échantillonnage de processus stationnaires

Sommaire

6.1	Introduction	166
6.2	Existence d'une densité spectrale pour le processus échantillonné	167
6.3	Asymptotique du périodogramme	171
6.4	Inférence pour le paramètre de longue mémoire	176
6.4.1	Convergence de l'estimateur local Whittle	177
6.4.2	Variance asymptotique	183
6.5	Annexe	188
6.5.1	Preuve du Lemme 6.2.1	188
6.5.2	Preuve de la Proposition 6.2.2	189
	Sous la condition (C1)	192
	Sous la condition (C2)	193
	Sous la condition (C3)	195
6.5.3	Preuve de la Proposition 6.2.3	198
6.5.4	Preuve de la Proposition 6.4.1	200

6.1 Introduction

Dans ce chapitre, nous prouvons que, sous des conditions assez générales, l'existence de la densité spectrale est préservée par échantillonnage aléatoire. Certains processus à temps continu ont déjà été étudiés auparavant du point de vue de l'étude spectrale. Notamment, dans [TC05a] et [Tsa09], Tsai et Chan ont étudié le modèle des processus à temps continu CARFIMA(p,H,q) (autorégressifs, moyenne mobile, fractionnairement intégrés) pour lesquels ils donnent l'expression de la densité spectrale (voir [TC05b]). Également, dans [Via+94], des processus ARMA fractionnaire et à temps continu sont construits. Sous certaines conditions, ces processus sont gaussiens stationnaires et centrés. De plus, Viano *et al* donnent des résultats sur la densité spectrale de ces processus.

L'inférence statistique pour les processus en temps continu est généralement construite à partir d'un processus échantillonné (voir [TC05a], [TC05b], [Cha96], [Com96]). Différents types d'échantillonnage peuvent être envisagés. Sous l'hypothèse d'identifiabilité du modèle, Chambers [Cha96] considère l'estimation du paramètre de longue mémoire d'un processus ARMA fractionnaire à temps continu avec des données à temps discrets en utilisant le comportement de la densité spectrale pour les basses fréquences. Comte [Com96] a étudié deux méthodes pour l'estimation avec des données régulièrement espacées : la méthode de vraisemblance de Whittle (introduite dans [Whi62]) et l'approche semi-paramétrique de Geweke et Porter-Hudak [GP83]. Pour un échantillonnage aléatoire particulier, Bardet et Bertrand [BB10] proposent d'étudier un estimateur de la densité spectrale basé sur l'analyse par ondelettes des processus gaussiens stationnaires qu'ils considèrent. [LM94] proposent un estimateur pour la densité spectrale d'un processus stationnaire à courte mémoire à temps continu échantillonné à l'aide d'un processus ponctuel stationnaire. Dans le cas de processus gaussiens stationnaires, l'estimation de la densité spectrale est souvent basée sur une méthode utilisant le périodogramme (voir [SS60], [Mas78a], [Mas78b] ou [LM94]).

Les notations de ce chapitre sont les mêmes que dans le chapitre précédent : $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+}$ est un processus stationnaire du second ordre à temps continu de fonction d'autocovariance $\sigma_X(\cdot)$, $(T_n)_{n \geq 0}$ est une marche aléatoire indépendante de \mathbf{X} , les intervalles d'échantillonnage $T_{j+1} - T_j = \Delta_j$ sont i.i.d de densité commune s à support sur \mathbb{R}^+ . L'origine des observations est fixée à l'origine des temps : $T_0 = 0$. Sans perte de généralités, \mathbf{X} est supposé être un processus centré.

La section 6.2 présente quelques cas de préservation de l'existence d'une den-

sité spectrale lorsque le spectre de X est absolument continu. La section 6.3 est consacrée à l'étude du comportement asymptotique du périodogramme normalisé de Y sous l'hypothèse d'un processus initialement gaussien avec une covariance à variation régulière. La loi asymptotique correspond alors à une statistique de type χ^2 pondéré. Enfin, la section 6.4 présente des résultats d'inférence pour le paramètre d de longue mémoire sous les conditions d'un processus initial gaussien à longue mémoire. Une condition vérifiée par les cumulants d'ordre 4 permet de montrer la convergence de l'estimateur local Whittle et de l'estimateur HAC de la variance asymptotique nécessaire à l'étude de la moyenne du processus et permettant de tester la mémoire du processus.

6.2 Existence d'une densité spectrale pour le processus échantillonné

De nombreux estimateurs pour les processus à longue mémoire sont basés sur le comportement de la densité spectrale (voir [Gir+12] pour un récapitulatif et une analyse de ses estimateurs et [Gro+18]). Dans cette section, nous étudions l'existence de la densité spectrale du processus échantillonné Y . Dans la proposition suivante, nous montrons l'existence de densité spectrale pour X et Y sous une condition L^p pour la fonction d'autocovariance et nous établissons la relation entre les densités spectrales de X et Y .

Proposition 6.2.1

Soit $1 \leq p \leq 2$. On suppose qu'il existe une fonction positive bornée $\sigma_*(\cdot)$, décroissante sur \mathbb{R}^+ , telle que

1. $|\sigma_X(t)| \leq \sigma_*(t), \quad \forall t \in \mathbb{R}^+,$
2. $\int_{\mathbb{R}^+} \sigma_*^p(t) dt < \infty.$

Alors les processus \mathbf{X} et \mathbf{Y} admettent tous les deux une densité spectrale et on a les relations :

$$\sigma_Y(k) = \int_{\mathbb{R}} \hat{S}(x)^k f_X(x) dx, \quad k \in \mathbb{N}, \quad (6.2.1)$$

$$f_Y(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} (e^{-i\lambda} \hat{S}(x))^k f_X(x) dx, \quad \lambda \in [-\pi, \pi], \quad (6.2.2)$$

où

$$\hat{S}(x) = \mathbb{E}(e^{ixT_1}) = \rho_x e^{i\tau_x}, \quad (6.2.3)$$

est la fonction caractéristique de la loi des inter-arrivées. Les fonctions ρ et τ sont définis comme le module et l'argument de la fonction caractéristique.

Remarque 6.2.1. Les hypothèses de cette proposition correspondent à celles de la Proposition 5.2.3 lorsque $p \in [1, 2]$.

Démonstration : On a dans ce cas,

$$\int_{\mathbb{R}^+} |\sigma_X(t)|^2 dt \leq \sigma_X(0)^{2-p} \int_{\mathbb{R}^+} \sigma_*^p(t) dt < \infty.$$

De plus, par la Proposition 5.2.3, on a aussi $\sum_{h \geq 0} |\sigma_Y(h)|^p < \infty$ et ainsi on a

$$\sum_{h \geq 0} |\sigma_Y(h)|^2 \leq \sigma_X(0)^{2-p} \sum_{h \geq 0} |\sigma_Y(h)|^p < \infty.$$

\mathbf{X} et \mathbf{Y} admettent ainsi des densités spectrales $f_X \in L^2(\mathbb{R})$ et $f_Y \in L^2([-\pi, \pi])$.

Pour $\lambda \in [-\pi, \pi]$,

$$f_Y(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{-ik\lambda} \mathbb{E}(\sigma_X(T_k)) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{-ik\lambda} \mathbb{E} \left(\int_{\mathbb{R}} e^{ixT_k} f_X(x) dx \right).$$

Comme $\mathbb{E} \left(\int_{\mathbb{R}} |e^{ixT_k} f_X(x)| dx \right) = \int_{\mathbb{R}} f_X(x) dx = E[X_0^2] < \infty$, on en déduit par le

théorème de Fubini que

$$\begin{aligned} f_Y(\lambda) &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} e^{-ik\lambda} \mathbb{E} \left(e^{ixT_k} \right) f_X(x) dx \\ &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} e^{-ik\lambda} \widehat{S}(x)^k f_X(x) dx \end{aligned}$$

où la somme converge dans $L^2([-\pi, \pi])$. ■

Dans cette proposition, lorsque $p \neq 1$, la fonction d'auto-covariance est carré sommable sans forcément être sommable. On est donc dans un cas à longue mémoire. En particulier, cette proposition s'applique pour le modèle de processus à longue mémoire avec une fonction d'autocovariance à variation régulière (modèle défini dans (5.3.1)) lorsque le paramètre de mémoire d appartient à $]0, \frac{1}{4}[$. Cette proposition s'applique aussi dans le cas où $-1/2 < d \leq 0$.

On se place désormais dans le cadre de l'existence d'une densité spectrale f_X pour \mathbf{X} . On commence par montrer le lien entre σ_Y , la fonction d'autocovariance de \mathbf{Y} , et f_X à travers la loi des inter arrivées.

Lemme 6.2.1

Supposons que \mathbf{X} admet une densité spectrale f_X . Alors pour tout $k \in \mathbb{Z}$

$$\sigma_Y(k) = \lim_{r \rightarrow 1^-} \int_{-\pi}^{\pi} e^{ikx} g(r, x) dx, \quad (6.2.4)$$

où

$$g(r, x) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - r^2 |\widehat{S}(\lambda)|^2}{|1 - r e^{ix} \widehat{S}(\lambda)|^2} f_X(\lambda) d\lambda, \quad (6.2.5)$$

et \widehat{S} est la fonction caractéristique de T_1 .

Démonstration : Voir l'annexe 6.5.1. ■

Proposition 6.2.2

Supposons que \mathbf{X} admet une densité spectrale notée f_X de la forme

$$f_X(\lambda) = |\lambda|^{-2d} \phi(\lambda) \quad (6.2.6)$$

où ϕ est une fonction positive, intégrable et bornée au voisinage de zéro, et où $0 \leq d < 1/2$. Si l'une de ces conditions est vérifiée :

- **(C1)** $d = 0$, i.e f_X est bornée au voisinage de zéro,
- **(C2)** $\mathbb{E}(T_1) < \infty$,
- **(C3)** la densité de T_1 vérifie

$$s(x) \sim cx^{-\gamma} \quad \text{lorsque } x \rightarrow \infty, \quad (6.2.7)$$

avec $1 < \gamma < 2$ et $c > 0$ (en particulier $\mathbb{E}(T_1) = \infty$),
alors le processus échantillonné Y admet une densité spectrale f_Y définie sur $[-\pi, \pi]$ (sauf éventuellement au point $x = 0$) par

$$f_Y(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - |\widehat{S}(\lambda)|^2}{|1 - e^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda) d\lambda. \quad (6.2.8)$$

Remarque 6.2.2. Si le processus X a une courte mémoire dans le sens où X admet une densité spectrale f_X continue bornée, alors le résultat de la Proposition 6.2.2 reste valide.

Démonstration : Voir l'annexe 6.5.2 pour les preuves sous les différentes conditions. ■

La proposition suivante donne une forme plus explicite de la densité spectrale de Y sous des conditions peu restrictives lorsque les arrivées suivent une loi exponentielle.

Proposition 6.2.3

On suppose que T_1 suit une loi exponentielle (de paramètre 1) et que X admet une densité spectrale de la forme

$$f_X(\lambda) = |\lambda|^{-2d} \phi(\lambda), \quad (6.2.9)$$

avec $\phi(0) \neq 0$ et ϕ dérivable au voisinage de zéro. Alors, pour $x \in]0, \pi]$,

$$f_Y(x) = x^{-2d} f_Y^*(x) \quad (6.2.10)$$

où f_Y^* est une fonction positive et continue sur $[-\pi, \pi]$ et

$$f_Y^*(x) = \frac{\phi(0)}{2} + \frac{\sigma_X(0)}{2\pi} x^{2d} + o(x^{2d}), \quad \text{lorsque } x \rightarrow 0. \quad (6.2.11)$$

Démonstration : La preuve se trouve en annexe 6.5.3 ■

6.3 Asymptotique du périodogramme

Dans cette section, on se place dans la cadre plus spécifique d'un processus initial \mathbf{X} gaussien stationnaire centré à longue mémoire de fonction de covariance

$$\sigma_X(t) = t^{-1+2d}L(t), \text{ avec } 0 < d < 1/2,$$

où on suppose que L est une fonction à variation lente à l'infini ayant une limite finie : $\lim_{n \rightarrow \infty} L(n) = \ell > 0$. On suppose de plus que \mathbf{X} admet une densité spectrale f_X dont on suppose qu'elle vérifie les hypothèses de la Proposition 6.2.3.

Nous échantillons aléatoirement ce processus aux instants d'échantillonnage $(T_i)_i$ tiré suivant un processus de Poisson. En particulier, on se place dans un cadre où $\mathbb{E}[T_1] < \infty$ pour lequel l'existence d'une densité spectrale pour \mathbf{X} en induit une pour \mathbf{Y} (voir Proposition 6.2.2 sous l'hypothèse **(C2)**).

Dans la suite, nous étendons certains faits bien connus sur les propriétés des périodogrammes au processus échantillonné \mathbf{Y} . En particulier, un des nos principaux résultats consiste à établir que le périodogramme normalisé de \mathbf{Y} converge asymptotiquement vers une statistique de type χ^2 pondéré.

Soit f_Y la densité spectrale de \mathbf{Y} . Pour $j \in \{1, \dots, \lfloor n/2 \rfloor\}$, on définit

$$I_n(\lambda_j) = \frac{1}{2\pi n} \left| \sum_{k=1}^n Y_k e^{ik\lambda_j} \right|^2$$

le périodogramme de (Y_1, \dots, Y_n) à la fréquence de Fourier $\lambda_j = \frac{2\pi j}{n}$.

Dans [HB93], Hurvich et Beltrao se sont intéressés à des processus en temps discret à longue mémoire de périodogramme I_n et de densité spectrale f . Pour ces processus assez généraux, ils étudient le comportement du rapport $\frac{I_n(\lambda_j)}{f(\lambda_j)}$ où $n \rightarrow \infty$ et j est fixé.

Alors que pour une fréquence ω quelconque fixée, on a $\mathbb{E} \left[\frac{I_n(\omega)}{f(\omega)} \right] \xrightarrow{n \rightarrow \infty} 1$ et pour M fréquences fixées $\left(\frac{I_n(\omega_1)}{f(\omega_1)}, \dots, \frac{I_n(\omega_M)}{f(\omega_M)} \right)$ tend asymptotiquement vers M variables i.i.d exponentielles ([Han73]), ce n'est plus le cas lorsqu'on étudie les basses fréquences λ_j (dépendant de n). Ce sont ces fréquences qui comptent lorsqu'on

cherche à estimer le paramètre de mémoire d à partir d'une méthode basée sur le périodogramme. En particulier, si le processus à temps discret étudié est gaussien, les variables $\left(\frac{I_n(\lambda_1)}{f(\lambda_1)}, \dots, \frac{I_n(\lambda_M)}{f(\lambda_M)}\right)$ sont asymptotiquement corrélées et on peut expliciter les variances et covariances. Comme la loi limite de $\frac{I_n(\lambda_j)}{f(\lambda_j)}$ dépend de j et de d , on a besoin de considérer un nombre de fréquences de Fourier tendant vers l'infini pour les méthodes de régression du paramètre de mémoire d afin d'avoir un estimateur asymptotiquement non biaisé.

Théorème 6.3.1

On se place sous les hypothèses décrites ci-dessus sur \mathbf{X} , \mathbf{Y} et sur les inter-arrivées. Alors pour tout nombre de fréquences de Fourier $s \geq 1$ fixé, et pour tous $j_1, \dots, j_s \in \{1, \dots, \lfloor n/2 \rfloor\}$ entiers tous distincts, on a

$$\left(\frac{I_n(\lambda_{j_1})}{f_Y(\lambda_{j_1})}, \dots, \frac{I_n(\lambda_{j_s})}{f_Y(\lambda_{j_s})} \right) \xrightarrow[n \rightarrow +\infty]{Loi} \left(L_{j_1}(d)[U^2(j_1) + V^2(j_1)], \dots, L_{j_s}(d)[U^2(j_s) + V^2(j_s)] \right), \quad (6.3.1)$$

où le vecteur $(U(1), \dots, U(\lfloor n/2 \rfloor), V(1), \dots, V(\lfloor n/2 \rfloor))$ est un vecteur gaussien centré. De plus, pour $\forall j, k \in \{1, \dots, \lfloor n/2 \rfloor\}$, on a $U(j), V(k)$ sont indépendants et

$$\begin{cases} \text{Var}(U(j)) = \frac{1}{2} - \frac{R_j(d)}{L_j(d)}, \\ \text{Var}(V(j)) = \frac{1}{2} + \frac{R_j(d)}{L_j(d)}, \\ \text{Cov}(U(j), U(k)) = \frac{L_{j,k}(d) - R_{j,k}(d)}{\sqrt{L_j(d)L_k(d)}}, \quad \text{pour } j \neq k, \\ \text{Cov}(V(j), V(k)) = \frac{L_{j,k}(d) + R_{j,k}(d)}{\sqrt{L_j(d)L_k(d)}}, \quad \text{pour } j \neq k. \end{cases} \quad (6.3.2)$$

Les expressions de $L_j(d)$, $R_j(d)$, $L_{j,k}(d)$ et $R_{j,k}(d)$ sont données dans la remarque ci-dessous.

Remarque 6.3.1. Pour $j, k \in \{1, \dots, \lfloor n/2 \rfloor\}$, les quantités $L_j(d)$, $R_j(d)$, $L_{j,k}(d)$ et $R_{j,k}(d)$ apparaissant dans le Théorème 6.3.1 sont définies par :

$$L_j(d) = \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2} \left| \frac{\lambda}{2\pi j} \right|^{-2d} d\lambda, \quad (6.3.3)$$

$$R_j(d) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)(2\pi j + \lambda)} \left| \frac{\lambda}{2\pi j} \right|^{-2d} d\lambda, \quad (6.3.4)$$

$$L_{j,k}(d) = \frac{(jk)^d}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi k - \lambda)(2\pi j - \lambda)} \left| \frac{\lambda}{2\pi} \right|^{-2d} d\lambda, \quad (6.3.5)$$

et

$$R_{j,k}(d) = \frac{(jk)^d}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi k + \lambda)(2\pi j - \lambda)} \left| \frac{\lambda}{2\pi} \right|^{-2d} d\lambda. \quad (6.3.6)$$

Démonstration : À l'aide de la décomposition

$$\frac{I_n(\lambda_j)}{f_Y(\lambda_j)} = \left| \frac{1}{\sqrt{2\pi n f_Y(\lambda_j)}} \sum_{r=1}^n Y_r \cos(r\lambda_j) + i \frac{1}{\sqrt{2\pi n f_Y(\lambda_j)}} \sum_{r=1}^n Y_r \sin(r\lambda_j) \right|^2,$$

on va montrer le résultat plus général

$$\begin{aligned} & \left(\frac{1}{\sqrt{2\pi n f_Y(\lambda_{j_1})}} \sum_{r=1}^n X_{T_r} \cos(r\lambda_{j_1}), \frac{1}{\sqrt{2\pi n f_Y(\lambda_{j_1})}} \sum_{r=1}^n X_{T_r} \sin(r\lambda_{j_1}), \dots, \right. \\ & \quad \left. \frac{1}{\sqrt{2\pi n f_Y(\lambda_{j_s})}} \sum_{r=1}^n X_{T_r} \cos(r\lambda_{j_s}), \frac{1}{\sqrt{2\pi n f_Y(\lambda_{j_s})}} \sum_{r=1}^n X_{T_r} \sin(r\lambda_{j_s}) \right) \\ & \xrightarrow[n \rightarrow +\infty]{Loi} \left(\sqrt{L_{j_1}(d)}[U(j_1), V(j_1)], \dots, \sqrt{L_{j_s}(d)}[U(j_s), V(j_s)] \right). \end{aligned} \quad (6.3.7)$$

Conditionnellement à (T_1, \dots, T_n) , le vecteur (Y_1, \dots, Y_n) est gaussien. Donc le vecteur défini dans (6.3.7) a pour fonction caractéristique

$$\mathbb{E} \left[\exp \left(-\frac{1}{2} u' \Sigma_T^{-1} u \right) \right], \quad \text{pour } u \in \mathbb{R}^{2s}$$

où Σ_T est la matrice de covariance, conditionnellement à (T_1, \dots, T_n) , du vecteur dans (6.3.7). La fonction caractéristique étant bornée, pour montrer la convergence (6.3.7), il suffit de montrer que

$$\Sigma_T \xrightarrow[n \rightarrow +\infty]{proba} \Sigma, \quad (6.3.8)$$

où Σ est la matrice de covariance du vecteur limite

$$\left(\sqrt{L_{j_1}(d)}(U(j_1), V(j_1)), \dots, \sqrt{L_{j_s}(d)}(U(j_s), V(j_s)) \right).$$

Dans ce qui suit, la constante c peut changer d'une ligne à une autre. Conditionnellement à (T_1, \dots, T_n) , l'entrée de la matrice Σ_T en position (i, k) est de la forme

$$\frac{1}{2\pi n \sqrt{f_Y(\lambda_{j_i}) f_Y(\lambda_{j_k})}} \sum_{r=1}^n \sum_{s=1}^n \sigma_X(T_r - T_s) h_{i,k}(r, s)$$

où

$$h_{i,k}(r, s) = \cos(r\lambda_{j_i}) \cos(s\lambda_{j_k}), \text{ ou } \cos(r\lambda_{j_i}) \sin(s\lambda_{j_k}), \text{ ou } \sin(r\lambda_{j_i}) \sin(s\lambda_{j_k}).$$

(lorsque i et k sont fixés, la forme de $h_{i,k}(r, s)$ est la même pour tout r et s). Ainsi, les entrées de $\mathbb{E}(\Sigma_T)$ vont être de la forme

$$\begin{aligned} & \frac{1}{2\pi n \sqrt{f_Y(\lambda_{j_i}) f_Y(\lambda_{j_k})}} \sum_{r=1}^n \sum_{s=1}^n \mathbb{E}(\sigma_X(T_r - T_s)) h_{i,k}(r, s) \\ &= \frac{1}{2\pi n \sqrt{f_Y(\lambda_{j_i}) f_Y(\lambda_{j_k})}} \sum_{r=1}^n \sum_{s=1}^n \sigma_Y(r - s) h_{i,k}(r, s). \end{aligned}$$

On a alors $\mathbb{E}(\Sigma_T) \xrightarrow[n \rightarrow +\infty]{} \Sigma$ par application du Théorème 5 de [HB93] dont la seule condition requise est la stationnarité du second ordre pour le processus Y et le comportement (6.2.10) de sa densité spectrale. Pour finir la preuve, il suffit de prouver que

$$\text{Var}(\Sigma_T) \xrightarrow[n \rightarrow +\infty]{} 0, \quad (6.3.9)$$

i.e. que les variances des entrées de Σ_T convergent vers zéro. Par Cauchy-Schwarz, il est suffisant de regarder ce qui se passe pour les coefficients diagonaux de la matrice. On traite ces coefficients dans le cas de cosinus (ceux avec des sinus se traitent de la même manière). À j fixé, en utilisant le fait que $f(\lambda_j) \sim bn^{2d}$, on peut écrire à l'aide de l'inégalité de Cauchy-Schwarz

$$\begin{aligned} & \text{Var} \left(\frac{1}{2\pi n f_Y(\lambda_j)} \sum_{r=1}^n \sum_{s=1}^n \sigma_X(T_r - T_s) \cos(r\lambda_j) \cos(s\lambda_j) \right) \\ & \sim \frac{c}{n^{2+4d}} \sum_{r,s,r',s'=1}^n \text{Cov}(\sigma_X(T_r - T_s) \cos(r\lambda_j) \cos(s\lambda_j), \sigma_X(T_{r'} - T_{s'}) \cos(r'\lambda_j) \cos(s'\lambda_j)) \\ & \leq \frac{c}{n^{2+4d}} \left(\sum_{r=1}^n \sum_{s=1}^n \sqrt{\text{Var}(\sigma_X(T_r - T_s))} \right)^2 \leq \frac{c}{n^{4d}} \left(\sum_{h=1}^n \sqrt{\text{Var}(\sigma_X(T_h))} \right)^2. \quad (6.3.10) \end{aligned}$$

Comme σ_X est bornée et que $T_h \xrightarrow[h \rightarrow +\infty]{p.s.} \infty$, on a $\text{Var}(\sigma_X(T_h)) \sim \ell^2 \text{Var}(T_h^{-1+2d})$ lorsque $h \rightarrow \infty$. Or, en utilisant que $\frac{\Gamma(h-a)}{\Gamma(h)} = h^{-a} \left(1 + \frac{a(a+1)}{2h} + O\left(\frac{1}{h^2}\right) \right)$ lorsque $h \rightarrow \infty$, on a

$$\begin{aligned} \text{Var}(T_h^{2d-1}) &= \mathbb{E}(T_h^{4d-2}) - \left(\mathbb{E}(T_h^{2d-1})\right)^2 = \frac{\Gamma(h - (2 - 4d))}{\Gamma(h)} - \left(\frac{\Gamma(h - (1 - 2d))}{\Gamma(h)}\right)^2 \\ &= h^{4d-2} \left[1 + \frac{(1-2d)(3-4d)}{h} - \left(1 + \frac{(1-2d)(2-2d)}{h} \right) + O\left(\frac{1}{h^2}\right) \right] \\ &= h^{4d-2} \left[\frac{(1-2d)^2}{h} + O\left(\frac{1}{h^2}\right) \right] \sim ch^{4d-3}. \end{aligned}$$

Ainsi la variance dans (6.3.10) est bornée par $c(n^{-1} + n^{-4d})$ et donc converge vers

zéro. ■

6.4 Inférence pour le paramètre de longue mémoire

Nous supposons toujours dans cette partie que \mathbf{X} est un processus gaussien stationnaire centré à longue mémoire de fonction de covariance

$$\sigma_X(t) = t^{-1+2d}L(t),$$

où on suppose que L est une fonction à variation lente à l'infini ayant une limite finie et $0 < d < 1/2$. On suppose de plus que \mathbf{X} admet une densité spectrale f_X .

Afin de simplifier les notations, on va supposer que la fonction à variation lente L est de la forme $L(t) = c(\frac{1+t}{t})^{-1+2d}$ de sorte que σ_X vérifie l'hypothèse suivante :

Hypothèse (*) : σ_X est une fonction à variation régulière de la forme

$$\sigma_X(t) = c(1+t)^{-1+2d}, \text{ pour } t \geq 0,$$

avec $0 < d < 1/2$ et $c > 0$.

La proposition suivante montre que, bien qu'il ne s'agisse pas d'un processus linéaire, \mathbf{Y} satisfait une condition sur les cumulants d'ordre 4.

Proposition 6.4.1

On suppose que les inter-arrivées sont i.i.d $\mathcal{E}(\lambda)$. Si \mathbf{X} est un processus gaussien stationnaire vérifiant la condition () alors \mathbf{Y} est un processus stationnaire du 4^{ième} ordre et il existe $C > 0$ tel que,*

$$\sup_{h \in \mathbb{N}} \sum_{r,s=0}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{2d},$$

et

$$\sum_{h,r,s=0}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{4d}.$$

Démonstration : La preuve se trouve en annexe. ■

Cette condition va nous permettre à la fois de montrer la convergence d'un estimateur du paramètre de mémoire d ainsi que l'estimation de la variance asymptotique, nécessaire par exemple pour inférer la moyenne (inconnue) du processus original X .

6.4.1 Convergence de l'estimateur local Whittle

Dans de nombreux travaux, la question de l'estimation du paramètre de longue mémoire d a été traitée sous différentes conditions sur le processus. Dans le cas où le comportement de la densité spectrale est connue autour de l'origine et de la forme $f(u) \sim b|u|^{-2d}$, on va utiliser des estimateurs semi-paramétriques.

Geweke et Porter-Hudak [GP83] ont introduit l'estimateur du log-périodogramme \hat{d}_{GPH} basé sur la minimisation de la fonction objectif

$$\frac{1}{m_n} \sum_{j=1}^{m_n} \left(\log(b\lambda_j^{-2d}) - \log(I_n(\lambda_j)) \right)^2,$$

pour $b > 0$ et $d \in [-1/2, 1/2]$ avec $m_n \rightarrow \infty$ et $m_n = o(n)$. Sous l'hypothèse d'un processus gaussien avec un paramètre m_n bien choisi et de conditions techniques sur la régularité de la densité spectrale, Robinson [Rob95b] montre la loi asymptotique de cet estimateur

$$\sqrt{m_n}(\hat{d}_{GPH} - d) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}\left(0, \frac{\pi^2}{24}\right).$$

Sous de bonnes hypothèses, [Hur+98] montre que la taille de la bande m_n optimale au sens de l'erreur quadratique moyenne est de l'ordre de $\mathcal{O}(n^{4/5})$.

L'estimateur local Whittle permet aussi d'estimer d à partir des basses fréquences [Rob95a] et est défini par :

$$\hat{d}_{LW} = \underset{\alpha \in [-1/2, 1/2]}{\operatorname{argmin}} U_n(\alpha)$$

où la fonction objectif local U_n est définie par

$$U_n(\alpha) = \log \left(\frac{1}{m_n} \sum_{j=1}^{m_n} \lambda_j^{2\alpha} I_n(\lambda_j) \right) - \frac{2\alpha}{m_n} \sum_{j=1}^{m_n} \log \lambda_j.$$

avec $I_n(\lambda_j)$ correspondant au périodogramme de (Y_1, \dots, Y_n) pris à la fréquence

de Fourier $\lambda_j = \frac{2\pi j}{n}$. Ici la taille de la bande de fréquences $m = m_n$ vérifie $m_n \rightarrow \infty$ et $m_n = o(n)$. Dans [Rob95a], Robinson montre que dans le cas d'un processus linéaire avec des conditions sur les innovations, sur la régularité de la densité spectrale et sur m_n , on a la consistance de \hat{d}_{LW} . Avec quelques hypothèses supplémentaires, on a aussi

$$\sqrt{m_n}(\hat{d}_{LW} - d) \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}\left(0, \frac{1}{4}\right),$$

ce qui fait de l'estimateur local Whittle un meilleur estimateur que l'estimateur du log-périodogramme puisqu'il a une plus petite variance asymptotique.

On a vu que dans les conditions de cette section que $f_Y(\lambda) \sim b|\lambda|^{-2d}$ lorsque $\lambda \rightarrow 0$ avec $0 < b < \infty$ et $0 < d < 1/2$. On va s'intéresser à la convergence de l'estimateur local Whittle pour le processus échantillonné qui n'est ni un processus linéaire à innovations i.i.d. ni un processus gaussien ce qui empêche d'appliquer directement les résultats connus sur cet estimateur. On va s'appuyer sur les travaux de [Dal+06] sur la convergence de l'estimateur local Whittle pour des processus non linéaires.

D'après le Théorème 1 de [Dal+06], si les périodogrammes normalisés

$$\eta_j^* := \frac{I_n(\lambda_j)}{b\lambda_j^{-2d}}$$

vérifient la loi faible des grands nombres :

$$\frac{1}{m_n} \sum_{j=1}^{m_n} \eta_j^* \xrightarrow[n \rightarrow \infty]{proba} 1, \quad (6.4.1)$$

alors on a $\hat{d}_{LW} \xrightarrow[n \rightarrow \infty]{proba} d$. En particulier, il faut trouver une condition suffisante afin d'avoir cette loi faible des grands nombres. Pour cela, [Dal+06] proposent d'étudier la quantité

$$\Delta_{m_n} := \max_{1 \leq k \leq m_n} \mathbb{E} \left| \sum_{j=1}^k \eta_j^* - \mathbb{E}[\eta_j^*] \right|$$

pour laquelle ils prouvent que si $\Delta_{m_n} = o(m_n)$ alors la loi faible (6.4.1) pour les η_j^* est vérifiée. Si, de plus, on suppose davantage de régularité pour f_Y , on peut montrer une vitesse de convergence pour l'estimateur.

Proposition 6.4.2

| On suppose que X est un processus gaussien stationnaire vérifiant la condition

(*) et que les inter-arrivées sont i.i.d $\mathcal{E}(\lambda)$.

Alors pour toute taille de bande de fréquences $m = m_n$ vérifiant

$$m_n \rightarrow \infty \text{ et } m_n = o(n),$$

on a la convergence de l'estimateur local Whittle pour le paramètre de longue mémoire :

$$\widehat{d}_{LW} \xrightarrow[n \rightarrow \infty]{\text{proba}} d.$$

Si, de plus, \mathbf{X} satisfait les conditions de la Proposition 6.2.3 et que $m_n = n^a$ avec $0 < a < 1$, alors on a

$$\log(n)(\widehat{d}_{LW} - d) = o_p(1)$$

Démonstration : D'après ce qui a été énoncé précédemment, il suffit de montrer que $\Delta_{m_n} = o(m_n)$ pour toute suite (m_n) telle que $m_n \rightarrow \infty$ et $m_n = o(n)$. Pour cela, on va suivre l'idée de la preuve du Lemme 1 de [Dal+06].

On rappelle que $f_Y(\lambda) \sim b|\lambda|^{-2d}$ lorsque $\lambda \rightarrow 0$ avec $0 < b < \infty$ et $0 < d < 1/2$. On pose alors

$$v(\lambda_j) = \frac{1}{\sqrt{2\pi n b \lambda_j^{-2d}}} \sum_{t=1}^n Y_t e^{it\lambda_j},$$

de telle sorte que $\eta_j^* = |v(\lambda_j)|^2$. On peut alors écrire

$$\begin{aligned} \text{Cov}(\eta_j^*, \eta_p^*) &= \text{Cov}(v(\lambda_j), v(\lambda_p))\text{Cov}(\bar{v}(\lambda_j), \bar{v}(\lambda_p)) \\ &\quad + \text{Cov}(v(\lambda_j), \bar{v}(\lambda_p))\text{Cov}(\bar{v}(\lambda_j), v(\lambda_p)) \\ &\quad + \text{cum}(v(\lambda_j), v(\lambda_p), \bar{v}(\lambda_j), \bar{v}(\lambda_p)). \end{aligned}$$

Ainsi, pour $1 \leq k \leq m_n$, on a

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{j=1}^k \eta_j^* - \mathbb{E}[\eta_j^*] \right)^2 \right] &= \sum_{j,p=1}^k \text{Cov}(\eta_j^*, \eta_p^*) \\ &\leq \sum_{j,p=1}^k \left(|\text{Cov}(v(\lambda_j), v(\lambda_p))|^2 + |\text{Cov}(v(\lambda_j), \bar{v}(\lambda_p))|^2 \right) \\ &\quad + \left| \sum_{j,p=1}^k \text{cum}(v(\lambda_j), v(\lambda_p), \bar{v}(\lambda_j), \bar{v}(\lambda_p)) \right| =: i_{n,1}(k) + i_{n,2}(k). \end{aligned}$$

Ainsi,

$$\Delta_{m_n} \leq \max_{1 \leq k \leq m_n} (i_{n,1}(k) + i_{n,2}(k))^{1/2} \leq \max_{1 \leq k \leq m_n} (i_{n,1}(k)^{1/2} + i_{n,2}(k)^{1/2}).$$

À l'aide du Lemme 4 de [Dal+06], on a

$$i_{n,1}(k) \leq C \sum_{1 \leq j \leq p \leq m_n} (j^{-2d} p^{-2+2d} \log^2 m_n + \mathbb{1}_{j=p}) \leq C(\log^3 m_n + m_n) \leq C m_n.$$

Il reste maintenant à majorer $i_{n,2}(k)$ uniformément pour $1 \leq k \leq m_n$:

$$\begin{aligned} i_{n,2}(k) &\leq \sum_{j,p=1}^k \frac{\lambda_j^{2d} \lambda_p^{2d}}{b^2 (2\pi n)^2} \left| \sum_{t_1, t_2, t_3, t_4=1}^n e^{i(t_1-t_2)\lambda_j} e^{i(t_3-t_4)\lambda_p} \text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4}) \right| \\ &\leq C n^{-2} \left(\sum_{j=1}^{m_n} \left(\frac{j}{n} \right)^{2d} \right)^2 \sum_{t_1, t_2, t_3, t_4=1}^n |\text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})| \\ &\leq C n^{-2} \left(\frac{m_n^{2d+1}}{n^{2d}} \right)^2 \sum_{t_1, t_2, t_3, t_4=1}^n |\text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})|. \end{aligned}$$

Comme l'expression de $\text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})$ est invariante par n'importe quelle permutation de $\{t_1, t_2, t_3, t_4\}$ et invariante par translation (grâce à la stationnarité d'ordre 4 du processus), on a alors :

$$\begin{aligned} i_{n,2}(k) &\leq C n^{-2} \left(\frac{m_n}{n} \right)^{4d} m_n^2 4 \sum_{\substack{t_1, t_2, t_3, t_4=1 \\ t_1 \leq \min(t_2, t_3, t_4)}}^n |\text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})| \\ &\leq C n^{-2} \left(\frac{m_n}{n} \right)^{4d} m_n^2 \sum_{\substack{t_1, t_2, t_3, t_4=1 \\ t_1 \leq \min(t_2, t_3, t_4)}}^n |\text{cum}(Y_0, Y_{t_2-t_1}, Y_{t_3-t_1}, Y_{t_4-t_1})| \\ &\leq C n^{-2} \left(\frac{m_n}{n} \right)^{4d} m_n^2 \sum_{t_1=1}^n \sum_{h,s,r=0}^n |\text{cum}(Y_0, Y_h, Y_s, Y_r)| \end{aligned}$$

En utilisant la Proposition 6.4.1, on a alors la majoration,

$$i_{n,2}(k) \leq C n^{-2} \left(\frac{m_n}{n} \right)^{4d} m_n^2 n n^{4d}$$

On a alors

$$\frac{\Delta_{m_n}}{m_n} \leq C \left(m_n^{-1/2} + \frac{m_n^{2d}}{n^{1/2}} \right). \quad (6.4.2)$$

Donc, si $m_n = O(n^\gamma)$ avec $\gamma < \frac{1}{4d}$, on obtient la convergence du quotient $\frac{\Delta_{m_n}}{m_n}$

vers 0. Comme $0 < d < 1/2$, on a la convergence pour toute suite (m_n) telle que $m_n = O(n^\gamma)$ avec $0 < \gamma \leq 1/2$.

On va enfin majorer plus finement la quantité $i_{n,2}(k)$ afin d'avoir un résultat de convergence lorsque $m_n = O(n^\gamma)$ avec $\gamma > 1/2$. Pour cela, on va écrire $i_{n,2}(k)$ sous une autre forme :

$$i_{n,2}(k) = \left| \sum_{t_1, t_2, t_3, t_4=1}^n B_k(t_1 - t_2) B_k(t_3 - t_4) \text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4}) \right|,$$

où $B_k(t) = \frac{1}{2\pi n} \sum_{j=1}^k b^{-1} \lambda_j^{2d} e^{it\lambda_j}$. D'où,

$$i_{n,2}(k) \leq \sum_{t_1, t_2, t_3, t_4=1}^n |B_k(t_1 - t_2)| |B_k(t_3 - t_4)| |\text{cum}(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})|.$$

Comme pour tout $u \in \mathbb{Z}$, $|B_k(-u)| = |\bar{B}_k(u)| = |B_k(u)|$, l'expression est symétrique en t_1, t_1, t_3, t_4 donc

$$i_{n,2}(k) \leq C \sum_{\substack{t_1, t_2, t_3, t_4=1 \\ t_1 \leq \min(t_2, t_3, t_4)}}^n |B_k(t_2 - t_1)| |B_k(t_3 - t_4)| |\text{cum}(Y_0, Y_{t_2-t_1}, Y_{t_3-t_1}, Y_{t_4-t_1})|$$

De plus, pour $t_1 \leq \min(t_2, t_3, t_4)$, l'expression dans la somme ci-dessus est symétrique en t_3, t_4 , il suffit donc de faire le calcul pour $t_3 \leq t_4$,

$$\begin{aligned} i_{n,2}(k) &\leq C \sum_{\substack{t_1, t_2, t_3, t_4=1 \\ t_1 \leq \min(t_2, t_3, t_4) \text{ et } t_3 \leq t_4}}^n |B_k(t_2 - t_1)| |B_k(t_4 - t_3)| |\text{cum}(Y_0, Y_{t_2-t_1}, Y_{t_3-t_1}, Y_{t_4-t_1})| \\ &\leq C n \sum_{u_1, u_2, u_3=0}^n |B_k(u_1)| |B_k(u_2)| |\text{cum}(Y_0, Y_{u_1}, Y_{u_3}, Y_{u_2+u_3})| \\ &\leq C n \left(\sum_{u=0}^n |B_k(u)| \right)^2 \sup_{0 \leq u_1, u_2 \leq n} \sum_{u_3=0}^n |\text{cum}(Y_0, Y_{u_1}, Y_{u_3}, Y_{u_2+u_3})| \\ &\leq C n \left(\sum_{u=0}^n |B_k(u)| \right)^2 \sup_{0 \leq u_1 \leq n} \sum_{r=0}^{2n} \sum_{s=0}^{2n} |\text{cum}(Y_0, Y_{u_1}, Y_s, Y_r)| \end{aligned}$$

Il reste donc à majorer $B_k(t)$ pour $t \in \mathbb{N}$.

Pour $t = 0$,

$$|B_k(0)| = \left| \frac{1}{2\pi n} \sum_{j=1}^k b^{-1} \lambda_j^{2d} \right| \leq C n^{-1} \sum_{j=1}^{m_n} \frac{j^{2d}}{n^{2d}} \leq C \left(\frac{m_n}{n} \right)^{2d}.$$

À $t \geq 1$ fixé, si on pose $s_p = \sum_{j=1}^p e^{it\lambda_j}$, on a par sommation par parties

$$B_k(t) = Cn^{-1-2d} \left(k^{2d} s_k + \sum_{j=1}^{k-1} (j^{2d} - (j+1)^{2d}) s_j \right).$$

Comme

$$|s_p| = \left| e^{i\lambda t} \frac{1 - e^{ip\lambda t}}{1 - e^{i\lambda t}} \right| \leq \frac{2}{|1 - e^{i\lambda t}|} \leq \frac{Cn}{t},$$

on a alors

$$|B_k(t)| \leq \frac{Cn^{-2d}}{t} \left(k^{2d} + \sum_{j=1}^{k-1} |j^{2d} - (j+1)^{2d}| \right) \leq \frac{Cn^{-2d} m_n^{2d}}{t}.$$

On en déduit alors que

$$i_{n,2}(k) \leq C n \left(\frac{m_n}{n} \right)^{4d} \log^2 n n^{2d},$$

et donc

$$\frac{\Delta_{m_n}}{m_n} \leq C \left(m_n^{-1/2} + \frac{(\sqrt{n})^{1-2d}}{m_n^{1-2d}} \log(n) \right). \quad (6.4.3)$$

Si $m_n = O(n^\gamma)$ avec $\gamma > 1/2$, on a alors $\frac{\Delta_{m_n}}{m_n} \xrightarrow[n \rightarrow \infty]{} 0$. On a ainsi la loi faible des grand nombres pour η_j^* et donc la convergence en probabilité de l'estimateur local Whittle.

Si on ajoute les hypothèses de la Proposition 6.2.3, afin de montrer la vitesse de convergence, il est suffisant de montrer que

$$\Delta_{m_n} = o(m_n^{1-\delta}), \quad \text{pour un certain } \delta > 0.$$

En utilisant les majorations (6.4.2) et (6.4.3), on arrive au résultat. ■

Dans la Figure 6.1, nous avons représenté plusieurs réalisations du périodogramme associé à (Y_1, \dots, Y_n) aux fréquences de Fourier à partir d'inter-arrivées i.i.d. $\mathcal{E}(1)$ et d'un processus initial gaussien de covariance $\sigma_X(t) = (1 + |t|)^{-1+2d}$ où on fait varier la valeur de d . Pour un bon choix de taille de la bande m_n , les estimateurs du log-périodogramme (GPH) et local Whittle (LW) donnent des résultats satisfaisant pour l'estimation du paramètre de longue mémoire après échantillonnage aléatoire.

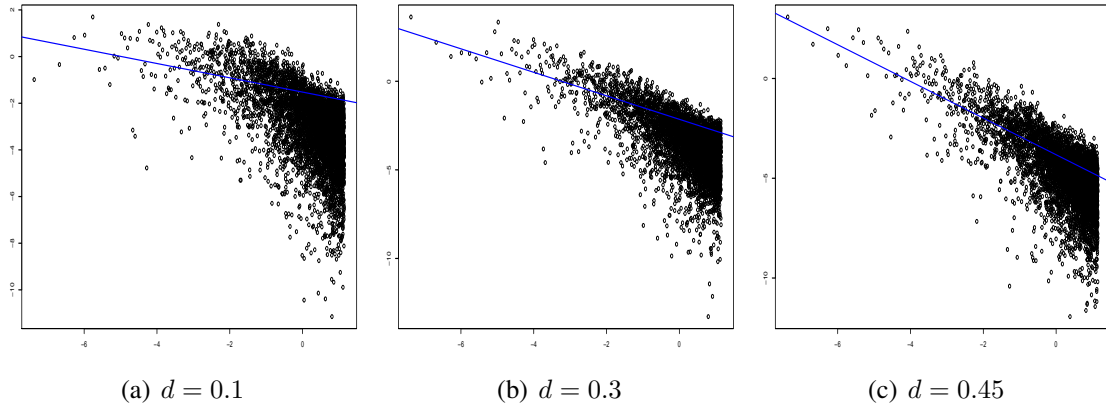


FIGURE 6.1 – Réalisations des valeurs du périodogrammes aux fréquences de Fourier $(I_n(\lambda_j))_{j \in \{1, \dots, \lfloor n/2 \rfloor\}}$ ($n = 10000$) pour différentes valeurs de paramètre de mémoire avec des intervalles Δ_j ayant une loi exponentielle de moyenne 1 et un processus initial gaussien de fonction d'autocovariance $\sigma_X(t) = (1+t)^{-1+2d}$. En bleu, on représente la droite de régression dont la pente correspond à $-2\hat{d}_{GPH}$. Les estimateurs du log-périodogramme (GPH) et local Whittle (LW) sont calculés avec la bande de fréquence $m_n = 300$: Figure (a) $d = 0.1$, $\hat{d}_{GPH} = 0.15$, $\hat{d}_{LW} = 0.15$; Figure (b) $d = 0.3$, $\hat{d}_{GPH} = 0.33$, $\hat{d}_{LW} = 0.37$; Figure (c) $d = 0.45$, $\hat{d}_{GPH} = 0.46$, $\hat{d}_{LW} = 0.47$. Les deux axes sont en échelle logarithmique.

6.4.2 Variance asymptotique

On se place sous les mêmes hypothèses que dans la partie précédente, à savoir : \mathbf{X} est un processus gaussien stationnaire centré à longue mémoire admettant une densité spectrale f_X et vérifiant l'hypothèse (\star) :

Hypothèse (\star) : σ_X est une fonction à variation régulière de la forme

$$\sigma_X(t) = c(1+t)^{-1+2d}, \text{ pour } t \geq 0,$$

avec $0 < d < 1/2$ et $c > 0$.

Pour permettre l'estimation de la moyenne μ commune des processus \mathbf{X} et \mathbf{Y} , on définit la moyenne empirique

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$$

qui est l'estimateur naturel de la moyenne. D'après le Théorème 5.3.4, on a la convergence

$$n^{1/2-d} \frac{\bar{Y}_n - \mu}{S_Y} \xrightarrow[n \rightarrow \infty]{Loi} \mathcal{N}(0, 1), \quad (6.4.4)$$

où

$$S_Y^2 := \frac{c\mathbb{E}[T_1]^{-1+2d}}{d(1+2d)} \quad (6.4.5)$$

correspond à la variance asymptotique de $n^{1/2-d}\bar{Y}_n$.

Il est possible de donner une autre expression de cette variance asymptotique à l'aide de la densité spectrale. En effet, la densité spectrale de Y au voisinage de l'origine est de la forme

$$f_Y(\lambda) \sim b|\lambda|^{-2d} \quad (6.4.6)$$

lorsque $\lambda \rightarrow 0$. On a alors l'égalité suivante :

$$\text{Var}(n^{1/2-d}\bar{Y}_n) = n^{-1-2d} \int_{-\pi}^{\pi} \left(\frac{\sin(n\lambda/2)}{\sin(\lambda/2)} \right)^2 f_Y(\lambda) d\lambda.$$

En utilisant l'hypothèse (6.4.6) sur le comportement de f_Y en 0 et par un changement de variables, on a alors

$$\begin{aligned} \text{Var}(n^{1/2-d}\bar{Y}_n) &\xrightarrow[n \rightarrow \infty]{} 4b \int_{-\infty}^{\infty} \left(\frac{\sin(\lambda/2)}{\lambda} \right)^2 |\lambda|^{-2d} d\lambda \\ &= 2b \frac{\Gamma(1-2d) \sin(\pi d)}{d(1+2d)}. \end{aligned}$$

Ainsi, on a deux expressions de la variance asymptotique S_Y^2 suivant qu'on regarde la fonction d'autocovariance ou la densité spectrale du processus :

$$S_Y^2 = 2b \frac{\Gamma(1-2d) \sin(\pi d)}{d(1+2d)} = \frac{c\mathbb{E}[T_1]^{-1+2d}}{d(1+2d)}.$$

Dans les deux cas, la variance asymptotique s'exprime en fonction de paramètres inconnus. Il faut donc estimer S_Y^2 . Le problème de l'estimation de la variance à long terme S_Y^2 revient ainsi à l'estimation du paramètre de mémoire d et de l'estimation du paramètre b de la densité spectrale en l'origine ou du paramètre $c\mathbb{E}[T_1]^{-1+2d}$ qui décrit le comportement asymptotique de la fonction de covariance.

Supposons qu'on ait un estimateur \hat{S}_n^2 de S_Y^2 tel que $\hat{S}_n^2 \xrightarrow[n \rightarrow \infty]{proba} S_Y^2$ et un estima-

teur \hat{d}_n de d tel que $\log(n)(\hat{d}_n - d) \xrightarrow[n \rightarrow \infty]{\text{proba}} 0$. Pour $\gamma \in]0, 1[$, on peut alors avoir un intervalle de confiance asymptotique de niveau $1 - \gamma$ pour μ :

$$\left[\bar{Y}_n - \frac{\hat{S}_n}{n^{1/2-\hat{d}_n}} z_{1-\gamma/2}, \bar{Y}_n + \frac{\hat{S}_n}{n^{1/2-\hat{d}_n}} z_{1-\gamma/2} \right],$$

où $z_{1-\gamma/2}$ est le quantile d'ordre $1 - \gamma/2$ de la loi $\mathcal{N}(0, 1)$. On peut remarquer que la longueur de cet intervalle est une fonction croissante de \hat{d}_n . En particulier, plus la mémoire est longue (d proche de $1/2$) et plus l'intervalle de confiance va être grand.

Dans la suite, on cherche donc un estimateur consistant de S_Y^2 . On définit pour $h \in \{1, \dots, (n-1)\}$

$$\hat{\sigma}(h) = \frac{1}{n} \sum_{j=1}^{n-h} (Y_j - \bar{Y}_n)(Y_{j+h} - \bar{Y}_n)$$

la fonction de covariance empirique de \mathbf{Y} . Pour q_n une taille de fenêtre telle que $q_n \xrightarrow[n \rightarrow \infty]{} +\infty$ avec $q_n = O(n^{1-\varepsilon})$ pour un certain $\varepsilon > 0$, on définit l'estimateur HAC (consistant en cas d'hétéroscédasticité et d'auto-corrélation) avec noyau de Bartlett [Aba+09] $\hat{S}_{q_n, Y}^2(d)$ de la variance à long terme S_Y^2 par

$$\hat{S}_{q_n, Y}^2(d) = q_n^{-2d} \left(\hat{\sigma}_Y(0) + 2 \sum_{h=1}^{q_n} \left(1 - \frac{h}{q_n}\right) \hat{\sigma}_Y(h) \right).$$

D'autres estimateurs HAC sont possibles avec des noyaux différents ([Phi+07]). Cet estimateur dépend du paramètre de mémoire d . Dans le théorème suivant, on montre qu'avec un estimateur \hat{d}_n tendant vers d suffisamment rapidement, on peut remplacer d par \hat{d}_n dans l'expression de $\hat{S}_{q_n, Y}^2(d)$ afin d'avoir un estimateur consistant de S_Y^2 .

Théorème 6.4.1

Soit \hat{d}_n un estimateur consistant du paramètre de mémoire d tel que

$$\log(n)(\hat{d}_n - d) = o_P(1).$$

Soit q_n une taille de fenêtre telle que $q_n \xrightarrow{n \rightarrow \infty} +\infty$ et $q_n = O(\sqrt{n})$. Alors, on a

$$\hat{S}_{q_n, Y}^2(\hat{d}_n) \xrightarrow[n \rightarrow \infty]{\text{proba}} S_Y^2.$$

Remarque 6.4.1. En particulier, si on se place sous les hypothèses de la Proposition 6.2.3, on a vu dans la Proposition 6.4.2 que l'estimateur local Whittle vérifie bien la condition $\log(n)(\hat{d}_n - d) = o_P(1)$. On a ainsi convergence de l'estimateur HAC basé sur l'estimateur local Whittle de d .

Démonstration : La preuve consiste à l'application du Théorème 2.2 de [Aba+09] où la condition sur les cumulants d'ordre 4 :

$$\sup_h \sum_{r,s=1}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq cn^{2d}$$

est vérifiée dans la Proposition 6.4.1. ■

En particulier, si on se place sous les hypothèses du Théorème 6.4.1, on peut construire un test sur la mémoire du processus. En effet, on peut considérer la statistique $V/S(n)$ [MW69a ; Kwi+92 ; LS96 ; Gir+03 ; Gir+06] définie par

$$V/S(n) = \frac{1}{\hat{S}_{q_n, Y}^2(0) n^2} \sum_{k=1}^n (S'_k - \bar{S}'_n)^2,$$

où $S'_k = \sum_{j=1}^k (Y_j - \bar{Y}_n)$ et $\bar{S}'_n = \frac{1}{n} \sum_{k=1}^n S'_k$.

On veut alors tester $H_0 : d = 0$ contre $H_1 : d > 0$. En utilisant la Proposition 6.4.1 sur les cumulants d'ordre 4, on a alors d'après [Gir+03] (Prop. 3.5 et 3.6) que

$$\begin{cases} V/S(n) \xrightarrow[n \rightarrow \infty]{\text{Loi}} U_{V/S} & \text{sous } H_0, \\ \left(\frac{q_n}{n}\right)^{2d} V/S(n) \xrightarrow[n \rightarrow \infty]{\text{Loi}} Z_{V/S} & \text{sous } H_1, \end{cases}$$

où

$$U_{V/S} = \int_0^1 \left(B^0(u) - \int_0^1 B^0(v) dv \right)^2 du \text{ et } Z_{V/S} = \int_0^1 \left(B_{1/2+d}^0(u) - \int_0^1 B_{1/2+d}^0(v) dv \right)^2 du$$

(B^0 désigne le pont Brownien et $B_{1/2+d}^0$ désigne le pont Brownien fractionnaire de paramètre de Hurst $1/2 + d$).

Soit $c_{1-\alpha}$ le quantile d'ordre $1 - \alpha$ de $U_{V/S}$, le test qui rejette l'hypothèse H_0 au profit de l'hypothèse H_1 lorsque $V/S(n) > c_{1-\alpha}$ est asymptotiquement de niveau α et est consistant sous H_1 , c'est à dire,

$$\begin{cases} \mathbb{P}(V/S(n) > c_{1-\alpha}) \xrightarrow[n \rightarrow \infty]{} \alpha & \text{sous } H_0, \\ \mathbb{P}(V/S(n) > c_{1-\alpha}) \xrightarrow[n \rightarrow \infty]{} 1 & \text{sous } H_1. \end{cases}$$

À partir de l'estimateur de longue mémoire $\hat{S}_{q_n, Y}^2$, d'autres tests existants pour la longue mémoire sont également des tests asymptotiquement de niveau α et consistants sous H_1 pour le processus échantillonné :

— le test utilisant la statistique $R/S(n)$ modifiée [Lo91] définie par

$$R/S(n) = n^{-1/2} \frac{\max_{j \in \llbracket 1, n \rrbracket} S_j - \min_{j \in \llbracket 1, n \rrbracket} S_j}{\hat{S}_{q_n, Y}^2(0)},$$

— le test utilisant la statistique KPSS [Kwi+92] définie par

$$KPSS(n) = \frac{1}{n^2 \hat{S}_{q_n, Y}^2(0)} \sum_{k=1}^n S'_k{}^2.$$

L'étude de [KT02] permet de comparer ces différents tests (voir également le chapitre 9 de [Gir+12] et [Ber+13] pour une description plus complète de toutes ces statistiques). En particulier, les résultats numériques de [Gir+03] montrent que la statistique V/S génère un test de meilleure puissance que les autres tests cités ci-dessus. Il existe d'autres tests pour la longue mémoire dont nous n'avons pas étudié la convergence dans le cas du processus échantillonné (voir par exemple [Gro+18]).

6.5 Annexe

6.5.1 Preuve du Lemme 6.2.1

Pour $0 \leq r < 1$ et $x \in [-\pi, \pi]$, on pose

$$g(r, x) := \frac{1}{2\pi} \left(\sum_{j=0}^{\infty} \sigma_Y(j) r^j [e^{ijx} + e^{-ijx}] - \sigma_Y(0) \right).$$

Pour $0 \leq r < 1$ et $x \in [-\pi, \pi]$, $g(r, x)$ est bien défini puisque la suite $(\sigma_Y(j))_{j \geq 0}$ est bornée.

On a de plus,

$$\int_{-\pi}^{\pi} \sum_{j=0}^{\infty} |\sigma_Y(j) r^j [e^{ijx} + e^{-ijx}]| dx \leq 2\pi \sum_{j=0}^{\infty} \sigma_Y(0) 2r^j < \infty,$$

on en déduit alors par Fubini que pour $k \in \mathbb{Z}$, on a

$$\int_{-\pi}^{\pi} e^{ikx} g(r, x) dx = \frac{1}{2\pi} \sum_{j=0}^{\infty} \sigma_Y(j) r^j \int_{-\pi}^{\pi} [e^{i(k+j)x} + e^{i(k-j)x}] dx - \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} \sigma_Y(0) dx.$$

On a ainsi

$$\int_{-\pi}^{\pi} e^{ikx} g(r, x) dx = \begin{cases} \sigma_Y(k) r^k & \text{si } k > 0, \\ \sigma_Y(-k) r^{-k} & \text{si } k < 0, \\ \sigma_Y(0) & \text{si } k = 0. \end{cases}$$

D'où pour tout $k \in \mathbb{Z}$, on a

$$\lim_{r \rightarrow 1^-} \int_{-\pi}^{\pi} e^{ikx} g(r, x) dx = \sigma_Y(k).$$

Il reste alors à montrer que $g(r, x)$ s'exprime sous la forme énoncée dans le Lemme 6.2.1. Comme $\mathbb{E} \left(\int_{\mathbb{R}} |e^{ixT_j} f_X(x)| dx \right) = \int_{\mathbb{R}} f_X(x) dx = E[X_0^2] < \infty$, on en déduit par le théorème de Fubini que

$$\begin{aligned} \sigma_Y(j) &= \mathbb{E}(\sigma_X(T_j)) \\ &= \mathbb{E} \left(\int_{\mathbb{R}} e^{i\lambda T_j} f_X(\lambda) d\lambda \right) \\ &= \int_{\mathbb{R}} \hat{S}(\lambda)^j f_X(\lambda) d\lambda \end{aligned} \tag{6.5.1}$$

On a alors

$$\begin{aligned} g(r, x) &= \frac{1}{2\pi} \left(\sum_{j=0}^{\infty} \int_{\mathbb{R}} \widehat{S}(\lambda)^j r^j [e^{ijx} + e^{-ijx}] f_X(\lambda) d\lambda - \int_{\mathbb{R}} f_X(\lambda) d\lambda \right) \\ &= \frac{1}{2\pi} \left(\int_{\mathbb{R}} \sum_{j=0}^{\infty} \widehat{S}(\lambda)^j r^j [e^{ijx} + e^{-ijx}] f_X(\lambda) d\lambda - \int_{\mathbb{R}} f_X(\lambda) d\lambda \right), \end{aligned}$$

d'après Fubini puisque

$$\sum_{j=0}^{\infty} \int_{\mathbb{R}} |\widehat{S}(\lambda)^j r^j [e^{ijx} + e^{-ijx}] f_X(\lambda)| d\lambda \leq \sum_{j=0}^{\infty} 2r^j \int_{\mathbb{R}} f_X(\lambda) d\lambda < \infty.$$

Ainsi, en utilisant la parité de f_X , on obtient

$$\begin{aligned} g(r, x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{1}{1 - re^{ix}\widehat{S}(\lambda)} + \frac{1}{1 - re^{-ix}\widehat{S}(\lambda)} - 1 \right) f_X(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{1}{1 - re^{ix}\widehat{S}(\lambda)} + \frac{1}{1 - re^{-ix}\widehat{S}(-\lambda)} - 1 \right) f_X(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - r^2 |\widehat{S}(\lambda)|^2}{|1 - re^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda) d\lambda, \end{aligned}$$

ce qui conclut la preuve.

6.5.2 Preuve de la Proposition 6.2.2

Tout d'abord, si $x \neq 0$, $g(r, x)$ admet une limite lorsque $r \rightarrow 1^-$ et cette limite est égale au membre de droite dans (6.2.8). En effet, en utilisant le théorème de Lebesgue, il suffit de montrer que l'intégrande $\frac{1-r^2|\widehat{S}(\lambda)|^2}{|1-re^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda)$ dans l'expression (6.5.6) de $g(r, x)$ est uniformément bornée en r par une fonction intégrable.

Comme \widehat{S} est une fonction caractéristique, \widehat{S} est une fonction continue.

Soit $x \in]0, \pi/2]$. Soit $\delta > 0$ tel que

$$|1 - \widehat{S}(\lambda)| < \frac{\sin x}{\sqrt{2}}, \text{ pour tout } |\lambda| \leq \delta. \quad (6.5.2)$$

Ainsi, pour $|\lambda| \leq \delta$, $\widehat{S}(\lambda) = 1 + h(\lambda)$ avec $|h(\lambda)| < \sin x/\sqrt{2}$. On utilise alors l'inégalité entre deux complexes a, b , $|a - b| \geq ||a| - |b||$ et l'inégalité entre deux

réels $\alpha, \beta, \alpha^2 + \beta^2 \geq \frac{(\alpha+\beta)^2}{2}$, pour en déduire que pour tout $|\lambda| \leq \delta$

$$\begin{aligned} |1 - re^{ix}\widehat{S}(\lambda)| &= |1 - re^{ix} - re^{ix}h(\lambda)| \\ &\geq \frac{1 - r \cos x}{\sqrt{2}} + \frac{r \sin x}{\sqrt{2}} - r|h(\lambda)| \\ &\geq \frac{1 - \cos x}{\sqrt{2}}. \end{aligned} \quad (6.5.3)$$

Puis, comme T_1 est une variable à densité, il existe $\varepsilon \in]0, 1[$ tel que $|\widehat{S}(\lambda)| < \varepsilon$ pour tout $|\lambda| > \delta$. Ainsi, pour tout $|\lambda| > \delta$,

$$|1 - re^{ix}\widehat{S}(\lambda)| \geq 1 - r|\widehat{S}(\lambda)| \geq 1 - r\varepsilon > 1 - \varepsilon \quad (6.5.4)$$

En rassemblant, (6.5.3) et (6.5.4), on aboutit à :

$$\forall x \in]0, \pi/2], \forall \lambda \in \mathbb{R}, \quad \left| \frac{1 - r^2 |\widehat{S}(\lambda)|^2}{|1 - re^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda) \right| \leq \left(\frac{1}{(1 - \cos(x))^2} + \frac{1}{(1 - \varepsilon)^2} \right) f_X(\lambda).$$

Pour $x \in]\pi/2, \pi]$, on choisit $\delta > 0$ tel que

$$|1 - \widehat{S}(\lambda)| < -\frac{\cos x}{\sqrt{2}}, \text{ pour tout } |\lambda| \leq \delta$$

et on a alors

$$|1 - re^{ix}\widehat{S}(\lambda)| = |1 - re^{ix} - re^{ix}h(\lambda)| \geq \frac{1}{\sqrt{2}}.$$

Ainsi, avec un $\tilde{\varepsilon} \in]0, 1[$, on a

$$\forall x \in]\pi/2, \pi], \forall \lambda \in \mathbb{R}, \quad \left| \frac{1 - r^2 |\widehat{S}(\lambda)|^2}{|1 - re^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda) \right| \leq \left(2 + \frac{1}{(1 - \tilde{\varepsilon})^2} \right) f_X(\lambda).$$

On a donc pour tout $x \neq 0$ de l'intervalle $[-\pi, \pi]$,

$$\lim_{r \rightarrow 1^-} g(r, x) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - |\widehat{S}(\lambda)|^2}{|1 - e^{ix}\widehat{S}(\lambda)|^2} f_X(\lambda) d\lambda. \quad (6.5.5)$$

De plus, pour tout $r \in [0, 1[$, $g(r, \cdot)$ est une fonction positive et paire puisque

$$\begin{aligned} g(r, -x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - r^2 |\widehat{S}(\lambda)|^2}{|1 - re^{-ix}\widehat{S}(\lambda)|^2} f_X(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - r^2 |\widehat{S}(-\lambda)|^2}{|1 - re^{-ix}\widehat{S}(-\lambda)|^2} f_X(-\lambda) d\lambda \\ &= g(r, x) \end{aligned}$$

car f_X est paire et $\widehat{S}(-\lambda) = \overline{\widehat{S}(\lambda)}$.

Par conséquence, comme on a obtenu dans (6.2.4) que

$$\sigma_Y(k) = \lim_{r \rightarrow 1^-} \int_{-\pi}^{\pi} e^{ikx} g(r, x) dx,$$

la preuve de la proposition consiste simplement à échanger la limite et l'intégration afin de montrer que

$$\sigma_Y(k) = \int_{-\pi}^{\pi} e^{ikx} \lim_{r \rightarrow 1^-} g(r, x) dx.$$

car cela implique que

$$f_Y(x) = \lim_{r \rightarrow 1^-} g(r, x).$$

Nous allons maintenant prouver que les conditions pour appliquer le théorème de Lebesgue sont bien vérifiées dans (6.2.4).

On commence par écrire $g(r, x)$ à partir de la fonction caractéristique de T_1 qui s'écrit sous la forme exponentielle $\widehat{S}(\lambda) = \rho_\lambda e^{i\tau\lambda}$ et du noyau de Poisson défini pour $s \in [0, 1[$ et $t \in \mathbb{R}$ par

$$P_s(t) = \frac{1}{2\pi} \left(\frac{1 - s^2}{1 - 2s \cos(t) + s^2} \right).$$

On a

$$\begin{aligned}
 g(r, x) &= g(r, -x) \\
 &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - (r\rho_\lambda)^2}{|1 - re^{-ix}\hat{S}(\lambda)|^2} f_X(\lambda) d\lambda \\
 &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1 - (r\rho_\lambda)^2}{1 - 2r\rho_\lambda \cos(\tau_\lambda - x) + (r\rho_\lambda)^2} f_X(\lambda) d\lambda \\
 &= \int_{\mathbb{R}} P_{r\rho_\lambda}(\tau_\lambda - x) f_X(\lambda) d\lambda
 \end{aligned} \tag{6.5.6}$$

On rappelle quelques propriétés du noyau de Poisson qui sont utilisées dans les preuves de cette annexe. Il admet entre autres les propriétés suivantes (voir par exemple [PV10] pour les preuves) :

$$\bullet \forall 0 \leq s < 1 - \eta < 1, \forall t \in \mathbb{R}, \quad 0 \leq 2\pi P_s(t) \leq 2/\eta \tag{6.5.7}$$

$$\bullet \forall x \in]-\pi, \pi[, r \in [0, 1[, \int_{-\pi}^{\pi} P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda = 1 \tag{6.5.8}$$

$$\bullet \text{ si } 0 < \delta < |t| \leq \pi, \text{ alors } P_s(t) < P_s(\delta) \tag{6.5.9}$$

$$\bullet 2\pi \sup_{0 < s < 1} P_s(t) = \begin{cases} 1/|\sin(t)| & \text{si } t \in (-\pi/2, \pi/2) \\ 1 & \text{si } t \in (-\pi, -\pi/2] \cup [\pi/2, \pi) \end{cases} \tag{6.5.10}$$

où ρ_λ et τ_λ sont définis dans (6.2.3) par $\hat{S}(\lambda) = E(e^{i\lambda T_1}) = \rho_\lambda e^{i\tau_\lambda}$.

Sous la condition (C1)

On a

$$|\hat{S}(\lambda)| < 1, \quad \forall \lambda > 0 \quad \text{et} \quad |\hat{S}(\lambda)| \rightarrow 0 \quad \text{lorsque } |\lambda| \rightarrow \infty$$

(voir [Fel71] p.501 et p.514). Ainsi, grâce à la continuité de $|\hat{S}(\lambda)|$, on a

$$\sup_{|\lambda| > \varepsilon} |\hat{S}(\lambda)| < 1, \quad \forall \varepsilon > 0.$$

L'intégrale (6.5.6) est séparée en deux : choisissons $\varepsilon \in [0, \pi]$ tel que f_X soit bornée sur $[-\varepsilon, \varepsilon]$.

$$\int_{-\varepsilon}^{\varepsilon} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda \leq \sup_{y \in [-\varepsilon, \varepsilon]} [f_X(y)] \int_{-\varepsilon}^{\varepsilon} P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda$$

En utilisant la propriété (6.5.8), on aboutit alors à :

$$\begin{aligned} \int_{-\varepsilon}^{\varepsilon} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda &\leq \sup_{y \in [-\varepsilon, \varepsilon]} [f_X(y)] \int_{-\pi}^{\pi} P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda \\ &= \sup_{y \in [-\varepsilon, \varepsilon]} [f_X(y)]. \end{aligned} \quad (6.5.11)$$

Puis, à l'aide de la propriété (6.5.7) du noyau de Poisson avec

$$s = r\rho_\lambda < \rho_\lambda, \quad t = \tau_\lambda \pm x, \quad \text{et} \quad \eta_\varepsilon = 1 - \sup_{|\lambda| > \varepsilon} |\hat{S}(\lambda)|$$

on a

$$\begin{aligned} \int_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda &\leq \frac{1}{\pi\eta_\varepsilon} \int_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]} f_X(\lambda) d\lambda \\ &\leq \frac{1}{\pi\eta_\varepsilon} \int_{\mathbb{R}} f_X(\lambda) d\lambda. \end{aligned} \quad (6.5.12)$$

En rassemblant les équations (6.5.11) et (6.5.12), on aboutit à : pour tout $r \in [0, 1[$ et pour tout $x \in]-\pi, \pi[$,

$$|e^{ikx} g(r, x)| \leq 4\pi \sup_{y \in [-\varepsilon, \varepsilon]} [f_X(y)] + \frac{1}{\pi\eta_\varepsilon} \int_{\mathbb{R}} f_X(\lambda) d\lambda,$$

et la preuve est finie en utilisant le théorème de Lebesgue.

Sous la condition (C2)

La preuve consiste à trouver une fonction g intégrable sur $[-\pi, \pi]$ telle que

$$|g(r, x)| \leq g(x), \quad \forall r \in [0, 1[, x \in]-\pi, \pi[. \quad (6.5.13)$$

Pour cela, nous avons besoin de l'estimation suivante de $\hat{S}(\lambda)$ autour de zéro :

$$\begin{aligned}
 |1 - \hat{S}(\lambda)| &= \left| (1 - e^{i\lambda}) \int_{\mathbb{R}_+} \frac{1 - e^{ix\lambda}}{1 - e^{i\lambda}} dS(x) \right| \\
 &= |1 - e^{i\lambda}| \left| \int_{\mathbb{R}_+} e^{\frac{i(x-1)\lambda}{2}} \frac{\sin(x\lambda/2)}{\sin(\lambda/2)} dS(x) \right| \\
 &\leq |1 - e^{i\lambda}| \int_{\mathbb{R}_+} (1+x) dS(x) \\
 &= |2i \sin(\lambda/2) e^{i\lambda/2}| (1 + \mathbb{E}(T_1)) \\
 &= 2 |\sin(\lambda/2)| (1 + \mathbb{E}(T_1)) \\
 &\leq |\lambda| (1 + \mathbb{E}(T_1)) \\
 &\leq C|\lambda|.
 \end{aligned} \tag{6.5.14}$$

Maintenant on utilise que pour $|u| \leq u_0 \leq 1$, on a

$$|1 - u| \geq 1 - u_0, \quad |\sin(\arg(1 - u))| \leq u_0 \quad \text{et} \quad |\arg(1 - u)| \leq \pi u_0/2.$$

De cela et de l'inégalité (6.5.14), on obtient :

$$\text{si } |\lambda| \leq \frac{1}{C} \text{ alors } |\tau_\lambda| \leq \frac{\pi C |\lambda|}{2}. \tag{6.5.15}$$

On rappelle l'expression de $g(r, x)$ défini dans (6.2.5) :

$$g(r, x) = \int_{\mathbb{R}} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda.$$

Soit λ_0 tel que ϕ est bornée sur $[-\lambda_0, \lambda_0]$. Pour $x \in (0, \pi)$ fixé, posons

$$b(x) = \min \left\{ \lambda_0, \frac{1}{C}, \frac{x}{\pi C}, \frac{2(\pi - x)}{\pi C} \right\},$$

on sépare \mathbb{R} en quatre intervalles :

$$]-\infty, -b(x)], \quad]-b(x), 0], \quad]0, b(x)], \quad]b(x), +\infty[.$$

On va étudier les deux derniers intervalles (la preuve pour les deux autres est similaire) :

$$I_1(x) = \int_0^{b(x)} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda \quad \text{et} \quad I_2(x) = \int_{b(x)}^{+\infty} f_X(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda.$$

• **Borne pour I_1 :**

De (6.5.15), comme $|\lambda| \leq \frac{1}{C}$ pour λ dans $]0, b(x)[$, on a

$$|\tau_\lambda| \leq \frac{\pi C |\lambda|}{2} \leq \frac{\pi C b(x)}{2} \leq x/2$$

ce qui implique

$$\frac{x}{2} \leq |\tau_\lambda - x|$$

et $|\tau_\lambda - x| \leq |\tau_\lambda| + x \leq \frac{\pi C b(x)}{2} + x \leq \pi$.

À l'aide des propriétés (6.5.9) et (6.5.10), on a ainsi

$$P_{r\rho_\lambda}(\tau_\lambda - x) \leq P_{r\rho_\lambda}(x/2) \leq \frac{1}{2\pi} \frac{1}{\sin(x/2)} \leq \frac{1}{2x}.$$

En conséquence,

$$I_1(x) \leq \frac{\sup_{[0, b(x)]} \phi(\cdot)}{2x} \int_0^{b(x)} \lambda^{-2d} d\lambda = \frac{\sup_{[0, b(x)]} \phi(\cdot)}{2x} \frac{b(x)^{-2d+1}}{-2d+1} \leq C_1 x^{-2d}$$

comme $b(x) \leq x/(\pi C)$ et $-2d+1 > 0$.

• **Borne pour I_2 :**

Lorsque $\lambda > b(x)$, on a $\lambda^{-2d} \leq C_2 \max\{x^{-2d}, 1\}$ pour une certaine constante C_2 .

Ainsi

$$I_2(x) \leq C_2 \max\{x^{-2d}, 1\} \int_{\mathbb{R}} \phi(\lambda) P_{r\rho_\lambda}(\tau_\lambda - x) d\lambda. \quad (6.5.16)$$

Comme ϕ est bornée sur un voisinage de zéro, les arguments utilisés dans la preuve de la Proposition 6.2.2 avec la condition **(C1)** nous permettent de montrer que l'intégrale (6.5.16) est bornée par une constante.

Finalement $I_1 + I_2$ est bornée par la fonction $g(x) = C_1 x^{-2d} + C_3 \max\{x^{-2d}, 1\}$ intégrable sur $[-\pi, \pi]$ et la proposition est prouvée.

Sous la condition (C3)

• On commence par donner le comportement local de $\hat{S}(\lambda)$ autour de zéro sous l'hypothèse (6.2.7).

$$1 - \hat{S}(\lambda) = \int_{\mathbb{R}^+} (1 - \cos(\lambda x)) s(x) dx + i \int_{\mathbb{R}^+} \sin(\lambda x) s(x) dx$$

D'après l'hypothèse sur la densité s ,

$$\begin{aligned} \int_{\mathbb{R}^+} (1 - \cos(\lambda x))s(x)dx &\underset{\lambda \rightarrow 0}{\sim} c \int_{\mathbb{R}^+} (1 - \cos(\lambda x))x^{-\gamma} dx \\ &= c|\lambda|^{\gamma-1} \int_{\mathbb{R}^+} (1 - \cos(y))y^{-\gamma} dy \\ &=: c|\lambda|^{\gamma-1}c_\gamma \end{aligned}$$

et

$$\begin{aligned} \int_{\mathbb{R}^+} \sin(\lambda x)s(x)dx &\underset{\lambda \rightarrow 0}{\sim} c \int_{\mathbb{R}^+} \sin(\lambda x)x^{-\gamma} dx \\ &= c|\lambda|^{\gamma-1} \int_{\mathbb{R}^+} \sin(y)y^{-\gamma} dy \\ &=: c|\lambda|^{\gamma-1}d_\gamma. \end{aligned}$$

Il est clair que $c_\gamma > 0$, et $d_\gamma > 0$. Ensuite,

$$|\lambda|^{1-\gamma}(1 - \hat{S}(\lambda)) \underset{\lambda \rightarrow 0}{\longrightarrow} Z, \text{ où } \Re(Z) > 0 \text{ et } \Im(Z) > 0. \quad (6.5.17)$$

• Dans la suite, on prend $\lambda > 0$. Si λ est suffisamment petit (ie $0 \leq \lambda \leq \lambda_0$),

$$c_3\lambda^{\gamma-1} \leq \tau_\lambda \leq c'_3\lambda^{\gamma-1} \quad (6.5.18)$$

où c_3 et c'_3 sont des constantes strictement négatives, et

$$1 - c_4\lambda^{\gamma-1} \leq \rho_\lambda \leq 1 - c'_4\lambda^{\gamma-1} \quad (6.5.19)$$

où c_4 et c'_4 sont des constantes strictement positives. Soit λ_1 tel que ϕ soit borné sur $[0, \lambda_1]$. Pour $x \in (0, \pi)$ fixé, on définit

$$c(x) = \min \left\{ \lambda_0, \lambda_1, \left(\frac{\pi - x}{-c_3} \right)^{1/(\gamma-1)}, \left(\frac{x}{-c'_3} \right)^{1/(\gamma-1)} \right\}.$$

on déduit de (6.5.18) que pour $\lambda \in (0, c(x))$

$$0 < x - c'_3\lambda^{\gamma-1} \leq x - \tau_\lambda \leq x - c_3\lambda^{\gamma-1} \leq \pi \quad (6.5.20)$$

Ensuite, on sépare \mathbb{R} en quatre intervalles

$$(-\infty, -c(x)[, [-c(x), 0[,]0, c(x)],]c(x), +\infty).$$

On va uniquement s'intéresser aux intégrales sur les deux derniers intervalles.

• Lorsque $\lambda \in]0, c(x)]$, l'inégalité (6.5.20) et les propriétés (6.5.9) et (6.5.10) du noyau de Poisson nous amène au résultat suivant :

$$P_{r\rho\lambda}(\tau\lambda - x) \leq P_{r\rho\lambda}(x - c'_3\lambda^{\gamma-1})$$

avec

$$P_{r\rho\lambda}(x - c'_3\lambda^{\gamma-1}) \leq \begin{cases} \frac{1}{2\pi \sin(x - c'_3\lambda^{\gamma-1})} \leq \frac{1}{4(x - c'_3\lambda^{\gamma-1})} & \text{si } x - c'_3\lambda^{\gamma-1} \in [0, \pi/2] \\ \frac{1}{2\pi} \leq \frac{1}{2(x - c'_3\lambda^{\gamma-1})} & \text{si } x - c'_3\lambda^{\gamma-1} \in [\pi/2, \pi] \end{cases} \quad (6.5.21)$$

$$\begin{aligned} I_1(x) &:= \int_0^{c(x)} f_X(\lambda) P_{r\rho\lambda}(\tau\lambda - x) d\lambda \leq \sup_{[0, \lambda_1]} \phi(\cdot) \int_0^{c(x)} \frac{C_1 \lambda^{-2d}}{x - c'_3\lambda^{\gamma-1}} d\lambda \\ &\leq C'_1 x^{-1} \int_0^{c(x)} \frac{\lambda^{-2d}}{1 - c'_3 x^{-1} \lambda^{\gamma-1}} d\lambda \\ &\leq C'_2 x^{\frac{-2d+1}{\gamma-1} - 1} \int_0^1 \frac{u^{\frac{-2d+1}{\gamma-1} - 1}}{1 + u} du. \end{aligned}$$

Comme $\frac{-2d+1}{\gamma-1} - 1 > -1$, l'intégrale ci-dessus est finie, ce qui implique

$$I_1(x) \leq C'_3 x^{\frac{-2d+1}{\gamma-1} - 1},$$

qui est une fonction intégrable sur $[-\pi, \pi]$.

• Finalement,

$$I_2(x) = \int_{c(x)}^{+\infty} f_X(\lambda) P_{r\rho\lambda}(\tau\lambda - x) d\lambda \leq c(x)^{-2d} \int_{\mathbb{R}} \phi(\lambda) P_{r\rho\lambda}(\tau\lambda - x) d\lambda$$

Cette dernière intégrale a déjà été étudiée dans la preuve de la Proposition 6.2.2 avec la condition (C1) puisque ϕ est bornée au voisinage de zéro :

$$\int_{\mathbb{R}} \phi(\lambda) P_{r\rho\lambda}(\tau\lambda - x) d\lambda \leq C_4$$

et $x \rightarrow c(x)^{-2d}$ est une fonction intégrable car $\frac{-2d}{\gamma-1} > -1$.

La preuve est terminée en rassemblant les résultats obtenus pour I_1 et I_2 .

6.5.3 Preuve de la Proposition 6.2.3

Comme f_Y est paire, on va considérer $x \in]0, \pi]$. On commence par utiliser la forme donnée par la Proposition 6.2.2 en utilisant que $\hat{S}(\lambda) = (1 - i\lambda)^{-1}$,

$$f_Y(x) = \frac{1}{2\pi} \int_0^\infty \frac{\lambda^2}{(\lambda - \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda \quad (6.5.22)$$

$$+ \frac{1}{2\pi} \int_0^\infty \frac{\lambda^2}{(\lambda + \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda. \quad (6.5.23)$$

On étudie alors les deux intégrales dans (6.5.22) au voisinage de $x = 0$.

$$\int_0^\infty \frac{\lambda^2}{(\lambda + \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda \xrightarrow{x \rightarrow 0} \int_0^\infty f_X(\lambda) d\lambda = \frac{\sigma_X(0)}{2} \quad (6.5.24)$$

puisque l'intégrande croît vers $f_X(\lambda)$ lorsque $x \rightarrow 0$. On traite maintenant le cas de la première intégrale dans (6.5.22).

$$\begin{aligned} \int_0^\infty \frac{\lambda^2}{(\lambda - \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda &= \int_0^{\sin x} \frac{\lambda^2}{(\lambda - \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda \\ &+ \int_{\sin x}^{2\sin x} \frac{\lambda^2}{(\lambda - \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda \\ &+ \int_{2\sin x}^\infty \frac{\lambda^2}{(\lambda - \sin x)^2 + (1 - \cos x)^2} f_X(\lambda) d\lambda. \end{aligned}$$

Pour simplifier les notations, et comme au voisinage de zéro, $\sin x \sim x$, on va prendre $x = \sin x$ et $1 - \cos x = x^2/2$. On fait le changement de variable $t = \lambda/x$, et on utilise que $f_X(\lambda) = \lambda^{-2d}\phi(\lambda)$, on obtient pour la première intégrale ci-dessus le résultat suivant avec $u(t) \in]0, 1[$ et $v(t) \in]0, 1[$

$$\begin{aligned}
& \int_0^x \frac{\lambda^2}{(\lambda-x)^2 + x^4/4} f_X(\lambda) d\lambda \\
&= x^{-2d} \int_0^1 \frac{t^{2-2d} x}{(1-t)^2 + x^2/4} \phi(xt) dt \\
&= x^{-2d} \int_0^1 \frac{(1-t)^{2-2d} x}{t^2 + x^2/4} \phi(x(1-t)) dt \\
&= x^{-2d} \int_0^1 (1 - (2-2d)(1-u(t))^{1-2d} t) \frac{x}{t^2 + x^2/4} (\phi(0) + \phi'(v(t))x(1-t)) dt \\
&= x^{-2d} \left[\int_0^1 \frac{x\phi(0)}{t^2 + x^2/4} dt + O\left(x \int_0^1 \frac{t}{t^2 + x^2/4} dt + x \int_0^1 \frac{x}{t^2 + x^2/4} dt\right) \right] \\
&= x^{-2d} \left[\phi(0) \arctan(2/x) + O\left(x(\ln(1+x^2/4) - \ln(x/2)) + x \arctan(2/x)\right) \right] \\
&= x^{-2d} \left(\frac{\pi}{2} \phi(0) + O(x \ln x) \right). \tag{6.5.25}
\end{aligned}$$

De même, on a

$$\begin{aligned}
& \int_x^{2x} \frac{\lambda^2}{(\lambda-x)^2 + x^4/4} f_X(\lambda) d\lambda \\
&= x^{-2d} \int_1^2 \frac{t^{2-2d} x}{(1-t)^2 + x^2/4} \phi(xt) dt \\
&= x^{-2d} \int_0^1 \frac{(1+t)^{2-2d} x}{t^2 + x^2/4} \phi(x(1+t)) dt \\
&= x^{-2d} \int_0^1 (1 + (2-2d)(1+u(t))^{1-2d} t) \frac{x}{t^2 + x^2/4} (\phi(0) + \phi'(v(t))x(1+t)) dt \\
&= x^{-2d} \left(\frac{\pi}{2} \phi(0) + O(x \ln x) \right). \tag{6.5.26}
\end{aligned}$$

Enfin, on a lorsque $x \rightarrow 0$

$$\int_{2x}^{\infty} \frac{\lambda^2}{(\lambda-x)^2 + x^4/4} f_X(\lambda) d\lambda \rightarrow \int_0^{\infty} f_X(\lambda) d\lambda = \frac{\sigma_X(0)}{2}, \tag{6.5.27}$$

comme l'intégrande est bornée par $4f_X(\lambda)$, on peut appliquer le théorème de Lebesgue. En combinant (6.5.24), (6.5.25), (6.5.26) et (6.5.27), on aboutit à

$$f_Y(x) = x^{-2d} f_Y^*(x), \quad \text{avec } f_Y^*(x) = \frac{\phi(0)}{2} + \frac{\gamma_X(0)}{2\pi} x^{2d} + o(x^{2d}) \quad \text{lorsque } x \rightarrow 0,$$

et f_Y^* est continue et positive sur $[-\pi, \pi]$ ce qui achève la preuve.

6.5.4 Preuve de la Proposition 6.4.1

Sans perte de généralité, on peut supposer pour la preuve que $\lambda = 1$ et $c = 1$. Afin de simplifier l'expression des calculs, on pose $\alpha := 1 - 2d \in]0, 1[$. Dans cette preuve, la constante C peut changer d'une ligne à une autre.

Le processus Y est un processus stationnaire du 4^{ième} ordre car pour tout $(k, h, r, s) \in \mathbb{N}^4$, la loi de $(Y_k, Y_{k+h}, Y_{k+r}, Y_{k+s})$ sachant $(T_1, \dots, T_{k+\max(h,r,s)})$ est la loi normale multivariée centrée de matrice de covariance $M(T_k, T_{k+h}, T_{k+r}, T_{k+s})$ définie par

$$M(T_k, T_{k+h}, T_{k+r}, T_{k+s}) := \begin{pmatrix} \sigma_X(0) & \sigma_X(T_{k+h}-T_k) & \sigma_X(T_{k+r}-T_k) & \sigma_X(T_{k+s}-T_k) \\ \sigma_X(T_{k+h}-T_k) & \sigma_X(0) & \sigma_X(T_{k+r}-T_{k+h}) & \sigma_X(T_{k+s}-T_{k+h}) \\ \sigma_X(T_{k+r}-T_k) & \sigma_X(T_{k+r}-T_{k+h}) & \sigma_X(0) & \sigma_X(T_{k+s}-T_{k+r}) \\ \sigma_X(T_{k+s}-T_k) & \sigma_X(T_{k+s}-T_{k+h}) & \sigma_X(T_{k+s}-T_{k+r}) & \sigma_X(0) \end{pmatrix}$$

La loi de cette matrice est indépendante de k d'où la stationnarité du 4^{ième} ordre de Y . On doit ainsi montrer que

$$\sup_{h \in \mathbb{N}} \sum_{r,s=0}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{2d}. \quad (6.5.28)$$

Par symétrie de l'expression, il suffit de montrer que

$$\sup_{h \in \mathbb{N}} \sum_{\substack{r,s=0 \\ r=s}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{2d} \quad (6.5.29)$$

et

$$\sup_{h \in \mathbb{N}} \sum_{\substack{r,s=0 \\ r>s}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{2d}. \quad (6.5.30)$$

On utilise ensuite l'expression reliant le cumulante avec le cumulante conditionnel afin d'avoir une expression plus simple à calculer :

$$\text{cum}(Y_0, Y_h, Y_r, Y_s) = \sum_{\substack{\Pi \in \{\text{partitions} \\ \text{de } \{0, h, r, s\}\}}}^n \text{cum}(\text{cum}(Y_{\Pi_1} | T_h, T_r, T_s), \dots, \text{cum}(Y_{\Pi_b} | T_h, T_r, T_s))$$

tel que pour toute partition Π de $\{0, h, r, s\}$, Π est séparé par les blocs de partitions Π_1, \dots, Π_b avec $b \leq 4$. Si un des blocs est de taille 1 (on peut supposer que c'est Π_1), alors $\text{cum}(Y_{\Pi_1} | T_h, T_r, T_s) = \mathbb{E}[Y_{\Pi_1} | T_h, T_r, T_s] = 0$. Si un des blocs est de taille 3 ou 4 (on peut supposer que c'est Π_1), alors $\text{cum}(Y_{\Pi_1} | T_h, T_r, T_s) = 0$ car sachant

(T_h, T_r, T_s) , Y_{Π_1} est gaussien. Il ne reste donc plus que les produits de deux blocs de taille 2 qui rentrent en compte dans le calcul du cumulatif. On a donc

$$\begin{aligned} \text{cum}(Y_0, Y_h, Y_r, Y_s) &= \text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s)) + \text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s)) \\ &\quad + \text{Cov}(\sigma_X(T_s), \sigma_X(T_r - T_h)) \end{aligned}$$

• On commence par montrer l'inégalité (6.5.30) pour laquelle on suppose $r > s$. Pour $h \in \{1, \dots, n\}$ fixé, on décompose la somme de l'inégalité sur les 5 ensembles $\{h < s < r\}$, $\{h = s < r\}$, $\{s < h < r\}$, $\{s < r = h\}$ et $\{s < r < h\}$.

— Sur l'ensemble $\{h < s < r\}$,

$$\begin{aligned} \sum_{\substack{r,s=0 \\ h < s < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| &\leq \sum_{\substack{r,s=0 \\ h < s < r}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_s - T_h))| \\ &\quad + |\text{Cov}(\sigma_X(T_s), \sigma_X(T_r - T_h))| \end{aligned}$$

On décompose alors suivant la forme donnée par le théorème central limite en posant :

$$A_r := \sqrt{r} \left(\frac{T_r}{r} - 1 \right) \quad \text{et} \quad B_{s-h} := \sqrt{s-h} \left(\frac{T_s - T_h}{s-h} - 1 \right)$$

pour avoir

$$\begin{aligned} &\text{Cov}(\sigma_X(T_r), \sigma_X(T_s - T_h)) \\ &= r^{-\alpha} (s-h)^{-\alpha} \text{Cov} \left(\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}} \right)^{-\alpha}, \left(1 + \frac{1}{s-h} + \frac{B_{s-h}}{\sqrt{s-h}} \right)^{-\alpha} \right). \end{aligned}$$

Or

$$\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}} \right)^{-\alpha} = \left(1 + \frac{1}{r} \right)^{-\alpha} - \alpha \left(1 + \frac{1}{r} \right)^{-\alpha-1} \frac{A_r}{\sqrt{r}} + \frac{\alpha(\alpha+1)}{2} \left(1 + \frac{1}{r} + \xi_r \right)^{-\alpha-2} \frac{A_r^2}{r}$$

et

$$\begin{aligned} \left(1 + \frac{1}{s-h} + \frac{B_{s-h}}{\sqrt{s-h}}\right)^{-\alpha} &= \left(1 + \frac{1}{s-h}\right)^{-\alpha} - \alpha \left(1 + \frac{1}{s-h}\right)^{-\alpha-1} \frac{B_{s-h}}{\sqrt{s-h}} \\ &\quad + \frac{\alpha(\alpha+1)}{2} \left(1 + \frac{1}{s-h} + \tilde{\xi}_{s-h}\right)^{-\alpha-2} \frac{B_{s-h}^2}{s-h} \end{aligned}$$

où ξ_r (resp. $\tilde{\xi}_{s-h}$) est compris strictement entre 0 et $\frac{A_r}{\sqrt{r}}$ (resp. entre 0 et $\frac{B_{s-h}}{\sqrt{s-h}}$). On a ainsi

$$\text{Cov}(\sigma_X(T_r), \sigma_X(T_s - T_h)) \quad (6.5.31)$$

$$\begin{aligned} &= r^{-\alpha}(s-h)^{-\alpha} \left[\alpha^2 \left(1 + \frac{1}{r}\right)^{-\alpha} \left(1 + \frac{1}{s-h}\right)^{-\alpha} \text{Cov}\left(\frac{A_r}{\sqrt{r}}, \frac{B_{s-h}}{\sqrt{s-h}}\right) \right. \\ &\quad - \frac{\alpha^2(\alpha+1)}{2} \left(1 + \frac{1}{s-h}\right)^{-\alpha-1} \text{Cov}\left(\frac{B_{s-h}}{\sqrt{s-h}}, \left(1 + \frac{1}{r} + \xi_r\right)^{-\alpha-2} \frac{A_r^2}{r}\right) \\ &\quad - \frac{\alpha^2(\alpha+1)}{2} \left(1 + \frac{1}{r}\right)^{-\alpha-1} \text{Cov}\left(\frac{A_r}{\sqrt{r}}, \left(1 + \frac{1}{s-h} + \tilde{\xi}_{s-h}\right)^{-\alpha-2} \frac{B_{s-h}^2}{s-h}\right) \\ &\quad \left. + \frac{\alpha^2(\alpha+1)^2}{4} \text{Cov}\left(\left(1 + \frac{1}{r} + \xi_r\right)^{-\alpha-2} \frac{A_r^2}{r}, \left(1 + \frac{1}{s-h} + \tilde{\xi}_{s-h}\right)^{-\alpha-2} \frac{B_{s-h}^2}{s-h}\right) \right]. \end{aligned}$$

De plus,

$$\begin{aligned} \left| \alpha^2 \left(1 + \frac{1}{r}\right)^{-\alpha} \left(1 + \frac{1}{s-h}\right)^{-\alpha} \text{Cov}\left(\frac{A_r}{\sqrt{r}}, \frac{B_{s-h}}{\sqrt{s-h}}\right) \right| &\leq \alpha^2 \text{Cov}\left(\frac{A_r}{\sqrt{r}}, \frac{B_{s-h}}{\sqrt{s-h}}\right) \\ &= \frac{\alpha^2}{r} \\ &\leq C \frac{1}{\sqrt{s-h}} \frac{1}{\sqrt{r}} \end{aligned}$$

et

$$\begin{aligned}
& \left| \text{Cov} \left(\frac{B_{s-h}}{\sqrt{s-h}}, \left(1 + \frac{1}{r} + \xi_r\right)^{-\alpha-2} \frac{A_r^2}{r} \right) \right| \\
& \leq \text{Var} \left(\frac{B_{s-h}}{\sqrt{s-h}} \right)^{1/2} \text{Var} \left(\left(1 + \frac{1}{r} + \xi_r\right)^{-\alpha-2} \frac{A_r^2}{r} \right)^{1/2} \\
& \leq \frac{1}{\sqrt{s-h}} \mathbb{E} \left[\left(1 + \frac{1}{r} + \xi_r\right)^{-2\alpha-4} \frac{A_r^4}{r^2} \right]^{1/2} \\
& \leq \frac{1}{\sqrt{r}\sqrt{s-h}} \left(\frac{\mathbb{E}[A_r^4]}{r} + \frac{1}{r} \mathbb{E} \left[\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}}\right)^{-2\alpha-4} A_r^4 \right] \right)^{1/2} \\
& \leq \frac{1}{\sqrt{r}\sqrt{s-h}} \left(\frac{3}{r} + \frac{6}{r^2} + \frac{1}{r} \mathbb{E} \left[\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}}\right)^{-2\alpha-4} A_r^4 \right] \right)^{1/2}
\end{aligned}$$

or

$$\begin{aligned}
& \mathbb{E} \left[\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}}\right)^{-2\alpha-4} A_r^4 \right] \\
& = r^{2\alpha+2} \mathbb{E} \left[(T_r + 1)^{-2\alpha-4} (T_r - r)^4 \right] \\
& \leq r^{2\alpha+2} \mathbb{E} \left[T_r^{-2\alpha-4} (T_r - r)^4 \right] \\
& = r^{2\alpha+2} \int_0^\infty x^{-2\alpha-4} (x - r)^4 \frac{x^{r-1}}{\Gamma(r)} e^{-x} dx \\
& = \frac{r^{2\alpha+2}}{\Gamma(r)} (\Gamma(r - 2\alpha) - 4r\Gamma(r - 2\alpha - 1) + 6r^2\Gamma(r - 2\alpha - 2) - 4r^3\Gamma(r - 2\alpha - 3) + r^4\Gamma(r - 2\alpha - 4))
\end{aligned}$$

On utilise alors le développement asymptotique de $\frac{\Gamma(r-a)}{\Gamma(r)}$:

$$\frac{\Gamma(r-a)}{\Gamma(r)} \sim r^{-a} \left(1 + \frac{a(a+1)}{2r} + o\left(\frac{1}{r}\right) \right), \quad \text{lorsque } r \rightarrow \infty$$

pour conclure que $\mathbb{E} \left[\left(1 + \frac{1}{r} + \frac{A_r}{\sqrt{r}}\right)^{-2\alpha-4} A_r^4 \right] = o(r)$

Finalement,

$$\left| \text{Cov} \left(\frac{B_{s-h}}{\sqrt{s-h}}, \left(1 + \frac{1}{r} + \xi_r\right)^{-\alpha-2} \frac{A_r^2}{r} \right) \right| = \frac{1}{\sqrt{s-h}} o\left(\frac{1}{\sqrt{r}}\right).$$

Les deux autres termes intervenant dans l'expression de l'équation (6.5.31)

se traitent de la même façon et on a alors

$$\text{Cov}(\sigma_X(T_r), \sigma_X(T_s - T_h)) \leq Cr^{-\alpha-1/2}(s-h)^{-\alpha-1/2}$$

donc

$$\sum_{\substack{r,s=0 \\ h < s < r}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_s - T_h))| \leq C \sum_{s=h+1}^n \sum_{r=s+1}^n r^{-\alpha-1/2}(s-h)^{-\alpha-1/2} \leq Cn^{1-\alpha}$$

On fait de même pour avoir

$$\text{Cov}(\sigma_X(T_s), \sigma_X(T_r - T_h)) \leq s^{-\alpha-1/2}(r-h)^{-\alpha-1/2}$$

d'où

$$\sum_{\substack{r,s=0 \\ h < s < r}}^n |\text{Cov}(\sigma_X(T_s), \sigma_X(T_r - T_h))| \leq C \sum_{s=h+1}^n s^{-\alpha-1/2} \sum_{r=s+1}^n (r-h)^{-\alpha-1/2} \leq Cn^{1-2\alpha}$$

Sur l'ensemble $\{h < s < r\}$, on a donc bien

$$\sum_{\substack{r,s=0 \\ h < s < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{1-\alpha}$$

— Sur l'ensemble $\{h = s < r\}$,

$$\sum_{\substack{r,s=0 \\ h=s < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| = 0 \leq Cn^{1-\alpha}$$

— Sur l'ensemble $\{s < h < r\}$,

$$\begin{aligned} \sum_{\substack{r,s=0 \\ s < h < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| &\leq \sum_{\substack{r,s=0 \\ s < h < r}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s))| \\ &\quad + |\text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s))| \end{aligned}$$

Par rapport au cas $\{h < s < r\}$, il suffit de permuter les rôles de h et s pour

obtenir :

$$\sum_{\substack{r,s=0 \\ s < h < r}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s))| \leq C \sum_{s=0}^{h-1} \sum_{r=s+1}^n r^{-\alpha-1/2} (h-s)^{-\alpha-1/2} \leq Cn^{1-\alpha}.$$

De même,

$$\text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s)) \leq h^{-\alpha-1/2} (r-s)^{-\alpha-1/2}$$

$$\begin{aligned} \sum_{\substack{r,s=0 \\ s < h < r}}^n |\text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s))| &\leq C \sum_{s=0}^{h-1} h^{-\alpha-1/2} \sum_{r=s+1}^n (r-s)^{-\alpha-1/2} \\ &\leq C \sum_{s=0}^{h-1} s^{-\alpha-1/2} \sum_{r=1}^n r^{-\alpha-1/2} \\ &\leq Ch^{-\alpha+1/2} n^{-\alpha+1/2} \\ &\leq Cn^{-2\alpha+1} \end{aligned}$$

Sur l'ensemble $\{s < h < r\}$, on a donc bien

$$\sum_{\substack{r,s=0 \\ s < h < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{1-\alpha}$$

— Sur l'ensemble $\{s < r = h\}$,

$$\sum_{\substack{r,s=0 \\ s < r=h}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| = 2 \sum_{\substack{r,s=0 \\ s < r=h}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_s))|$$

On utilise la même preuve que pour le cas $h < r = s$, afin de montrer que

$$\sum_{\substack{r,s=0 \\ s < r=h}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_s))| \leq C \sum_{s=0}^{r-1} r^{-\alpha} (r-s)^{-\alpha} \leq Cn^{1-2\alpha}$$

— Sur l'ensemble $\{s < r < h\}$,

$$\begin{aligned} \sum_{\substack{r,s=0 \\ s < r < h}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| &\leq \sum_{\substack{r,s=0 \\ s < r < h}}^n |\text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s))| \\ &\quad + |\text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s))| \end{aligned}$$

Par rapport au cas $\{s < h < r\}$, il suffit de permuter les rôles de h et r pour obtenir :

$$\begin{aligned} \sum_{\substack{r,s=0 \\ s < r < h}}^n |\text{Cov}(\sigma_X(T_h), \sigma_X(T_r - T_s))| &\leq Ch^{-\alpha-1/2} \sum_{r=1}^{h-1} \sum_{s=0}^{r-1} (r-s)^{-\alpha-1/2} \\ &\leq Ch^{-1/2-\alpha} \sum_{r=1}^{h-1} r^{-\alpha+1/2} \\ &\leq Ch^{1-2\alpha} \\ &\leq Cn^{1-\alpha} \end{aligned}$$

De même,

$$\begin{aligned} |\text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s))| &\leq Cr^{-\alpha-1/2} (h-s)^{-\alpha-1/2} \\ \sum_{\substack{r,s=0 \\ s < r < h}}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_h - T_s))| &\leq C \sum_{r=1}^{h-1} \sum_{s=0}^{r-1} r^{-\alpha-1/2} (h-s)^{-\alpha-1/2} \\ &\leq Cn^{-2\alpha+1} \end{aligned}$$

Sur l'ensemble $\{s < r < h\}$, on a donc bien

$$\sum_{\substack{r,s=0 \\ s < r < h}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{1-\alpha}$$

On a donc montré que

$$\sup_{h \in \{1, \dots, n\}} \sum_{\substack{r,s=0 \\ s < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| \leq Cn^{2d}$$

Le cas $h > n$ se traite de la même manière que le cas $s < r < h \leq n$.

• Il reste à montrer l'inégalité (6.5.29) pour laquelle on suppose $r = s$, on a ainsi

$$\text{cum}(Y_0, Y_h, Y_r, Y_s) = 2\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_h))$$

- Si $h = r$, $\text{cum}(Y_0, Y_h, Y_r, Y_s) = 0$
- Si $h < r$, on écrit de nouveau T_r et $T_r - T_h$ sous les formes données par le

théorème central limite :

$$\sqrt{r} \left(\frac{T_r}{r} - 1 \right) = A_r \text{ et } \sqrt{r-h} \left(\frac{T_r - T_h}{r-h} - 1 \right) = B_{r-h}$$

pour avoir

$$|\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_h))| \leq Cr^{-\alpha}(r-h)^{-\alpha}$$

— Si $h > r$, $\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_h)) = 0$ par indépendance des variables aléatoires T_r et $T_r - T_h$.

Ainsi,

$$\begin{aligned} \sup_{h \in \mathbb{N}} \sum_{\substack{r,s=0 \\ r=s}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)| &= 2 \sup_{h \in \{1, \dots, (n-1)\}} \sum_{r=h+1}^n |\text{Cov}(\sigma_X(T_r), \sigma_X(T_r - T_h))| \\ &\leq C \sup_{h \in \{1, \dots, (n-1)\}} \sum_{r=h+1}^n r^{-\alpha}(r-h)^{-\alpha} \\ &\leq C \sup_{h \in \{1, \dots, (n-1)\}} \sum_{r=h+1}^n (r-h)^{-2\alpha} \\ &\leq C \sum_{r=1}^n r^{-2\alpha} \\ &\leq Cn^{1-\alpha} \end{aligned}$$

Ce qui conclut la preuve.

Enfin, le cas de la somme triple est plus facile à traiter grâce à la symétrie de l'expression :

$$\sum_{h,r,s=0}^n |\text{cum}(Y_k, Y_{k+h}, Y_{k+r}, Y_{k+s})| \leq C \sum_{\substack{h,r,s=0 \\ h < s \leq r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)|.$$

On peut remarquer par les calculs fait précédemment qu'on peut négliger dans la somme les termes où 2 indices sont égaux. On a alors

$$\sum_{h,r,s=0}^n |\text{cum}(Y_k, Y_{k+h}, Y_{k+r}, Y_{k+s})| \leq C \sum_{\substack{h,r,s=0 \\ h < s < r}}^n |\text{cum}(Y_0, Y_h, Y_r, Y_s)|.$$

Puis on utilise le cas $h < s < r$ fait précédemment, on a alors

$$\sum_{\substack{h,r,s=0 \\ h < s < r}}^n |\text{cum}(Y_k, Y_{k+h}, Y_{k+r}, Y_{k+s})| \leq C \sum_{\substack{h,r,s=0 \\ h < s < r}}^n r^{-\alpha-1/2}(s-h)^{-\alpha-1/2} + s^{-\alpha-1/2}(r-h)^{-\alpha-1/2}$$

or

$$\begin{aligned} \sum_{\substack{h,r,s=0 \\ h < s < r}}^n r^{-\alpha-1/2}(s-h)^{-\alpha-1/2} &\leq C \sum_{r=1}^n \sum_{s=1}^{r-1} \sum_{h=0}^s r^{-\alpha-1/2}(s-h)^{-\alpha-1/2} \\ &\leq C \sum_{r=1}^n r^{-\alpha-1/2} \sum_{s=1}^{r-1} s^{-\alpha+1/2} \\ &\leq C \sum_{r=1}^n r^{-\alpha-1/2} r^{-\alpha+3/2} \\ &\leq C n^{2-2\alpha} = C n^{4d}. \end{aligned}$$

Puis,

$$\begin{aligned} \sum_{\substack{h,r,s=0 \\ h < s < r}}^n s^{-\alpha-1/2}(r-h)^{-\alpha-1/2} &\leq C \sum_{r=1}^n \sum_{s=1}^{r-1} s^{-\alpha-1/2} \sum_{h=0}^s (r-h)^{-\alpha-1/2} \\ &\leq C \sum_{r=1}^n \sum_{s=1}^{r-1} s^{-\alpha-1/2} s(r-s)^{-\alpha-1/2} \\ &\leq C \sum_{s=1}^n s^{-\alpha+1/2} \sum_{r=s+1}^n (r-s)^{-\alpha-1/2} \\ &\leq C n^{-\alpha+3/2} n^{-\alpha+1/2} = C n^{4d}. \end{aligned}$$

Bibliographie

- [Aba+09] K. M. ABADIR, W. DISTASO et L. GIRAITIS, « Two estimators of the long-run variance : Beyond short memory », in : *Econometrics* 150.1 (2009), p. 56.
- [BB10] J.-M. BARDET et P. BERTRAND, « A Non-Parametric Estimator of the Spectral Density of a Continuous-Time Gaussian Process Observed at Random Times », in : *Scandinavian Journal of Statistics* 37 (2010), p. 458-476.
- [Bar+11] K. BARTKIEWICZ, A. JAKUBOWSKI, T. MIKOSCH et O. WINTENBERGER, « Stable limits for sums of dependent infinite variance random variables », in : *Probability Theory and Related Fields* 150.3 (2011), p. 337-372.
- [Ben+01] V. BENTKUS, A. JUOZULYNAS et V. PAULAUSKAS, « Lévy–LePage Series Representation of Stable Vectors : Convergence in Variation », in : *Journal of Theoretical Probability* 14.4 (2001), p. 949-978.
- [Ber94] J. BERAN, *Statistics for Long-Memory Processes*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis, 1994.
- [Ber+13] J. BERAN, Y. FENG, S. GHOSH et R. KULIK, *Long-Memory Processes : Probabilistic properties and statistical methods*. Springer, 2013.
- [Beu70] F. J. BEUTLER, « Alias-free randomly timed sampling of stochastic processes », in : *IEEE Transactions on Information Theory* 16.2 (1970), p. 147-152.
- [Bil68] P. BILLINGSLEY, *Convergence of probability measures*, John Wiley & Sons, Inc., New York-London-Sydney, 1968, p. xii+253.
- [Bin+89] N. BINGHAM, C. GOLDIE et J. TEUGELS, *Regular Variation*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1989.
- [BT58] R. B. BLACKMAN et J. W. TUKEY, « The measurement of power spectra from the point of view of communications engineering — Part I », in : *The Bell System Technical Journal* 37.1 (1958), p. 185-282.

-
- [Cha+76] J. M. CHAMBERS, C. L. MALLOWS et B. W. STUCK, « A Method for Simulating Stable Random Variables », in : *Journal of the American Statistical Association* 71.354 (1976), p. 340-344.
- [Cha96] M. J. CHAMBERS, « The Estimation of Continuous Parameter Long-Memory Time Series Models », in : *Econometric Theory* 12 (1996), 374–390.
- [Cha98] M. J. CHAMBERS, « Long-memory and aggregation in macroeconomic time-series », in : *International Economic Review* 39.4 (1998), p. 1053-1072.
- [Com96] F. COMTE, « Simulation and estimation of long memory continuous time models », in : *Journal of Time Series Analysis* 17.1 (1996), p. 19-36.
- [CR96] F. COMTE et E. RENAULT, « Long memory continuous time models », in : *Journal of Econometrics* 73.1 (1996), p. 101 -149.
- [Dal+06] V. DALLA, L. GIRAITIS et J. HIDALGO, « Consistent estimation of the memory parameter for nonlinear time series », in : *Journal of Time Series Analysis* 27.2 (2006), p. 211-251.
- [DI17] T. T. N. DANG et J. ISTAS, « Estimation of the Hurst and the stability indices of a H -self-similar stable process », in : *Electronic Journal of Statistics* 11.2 (2017), p. 4103-4150.
- [DH95] R. A. DAVIS et T. HSING, « Point process and partial sum convergence for weakly dependent random variables with infinite variance », in : *Ann. Probab.* 23.2 (1995), p. 879-917.
- [Dav70] Y. DAVYDOV, « Invariance principle for stationary processes. », in : *Theory of Probability and Its Applications* 15 (1970), p. 498-509.
- [Dim+11] I. A. DIMITRIADIS, C. ALBEROLA-LÓPEZ, M. MARTÍN-FERNÁNDEZ, P. C. de-la HIGUERA, F. SIMMROSS-WATTENBERG et J. I. ASENSIO-PÉREZ, « Anomaly Detection in Network Traffic Based on Statistical Inference and α -Stable Modeling », in : *IEEE Transactions on Dependable and Secure Computing* 8 (2011), p. 494-509.
- [FLV09] K. J. FALCONER et J. LÉVY VÉHEL, « Multifractional, Multistable, and Other Processes with Prescribed Local Form », in : *Journal of Theoretical Probability* 22.2 (2009), p. 375-401.
- [Fal02] K. J. FALCONER, « Tangent Fields and the Local Structure of Random Fields », in : *Journal of Theoretical Probability* 15.3 (2002), p. 731-750.

-
- [Fal03] K. J. FALCONER, « The Local Structure of Random Processes », in : *Journal of the London Mathematical Society* 67.3 (2003), p. 657-672.
- [Fal+09] K. J. FALCONER, R. LE GUÉVEL et J. LÉVY VÉHEL, « Localizable moving average symmetric stable and multistable processes », in : *Stoch. Models* 25.4 (2009), p. 648-672, ISSN : 1532-6349.
- [FR71] E. F. FAMA et R. ROLL, « Parameter estimates for symmetric stable distributions », in : *Journal of the American Statistical Association* 66 (1971), p. 331-338.
- [Fel71] W. FELLER, *An introduction to probability theory and its applications*, t. 2, John Wiley & Sons, 1971.
- [FK72] T. S. FERGUSON et M. J. KLASS, « A Representation of Independent Increment Processes without Gaussian Components », in : *Ann. Math. Statist.* 43.5 (1972), p. 1634-1643.
- [GP83] J. GEWEKE et S. PORTER-HUDAK, « The Estimation and Application of Long Memory Time Series Model », in : *Journal of Time Series Analysis* 4 (1983), p. 221 -238.
- [Gir+03] L. GIRAITIS, P. KOKOSZKA, R. LEIPUS et G. TEYSSIERE, « Rescaled variance and related tests for long memory in volatility and levels », in : *Journal of Econometrics* 112.2 (2003), p. 265-294.
- [Gir+06] L. GIRAITIS, R. LEIPUS et A. PHILIPPE, « A Test for Stationarity versus Trends and Unit Roots for a Wide Class of Dependent Errors », in : *Econometric Theory* 22.6 (2006), p. 989-1029.
- [Gir+12] L. GIRAITIS, H. L. KOUL et D. SURGAILIS, *Large Sample Inference for Long Memory Processes*, Imperial College Press, 2012.
- [GK68] B. V. GNEDENKO et A. N. KOLMOGOROV, *Limit distributions for sums of independent random variables*, Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition, Addison-Wesley Publishing Co., 1968, p. ix+293.
- [GR57] U. GRENANDER et M. ROSENBLATT, *Statistical Analysis of Stationary Time Series*, Wiley publication in mathematical statistics, Wiley, 1957.
- [Gro+18] G. GROMYKOV, M. OULD HAYE et A. PHILIPPE, « A frequency-domain test for long range dependence », in : *Statistical Inference for Stochastic Processes* 21.3 (2018), p. 513-526.

-
- [Han73] E. HANNAN, « The Central Limit Theorem for Time Series Regression », in : *Stochastic Processes and their Applications* 9 (1973), p. 281-289.
- [HJ84] C. D. HARDIN JR., *Skewed stable variables and processes*, Technical Report 79, Center for Stochastic Processes at the University of North Carolina, Chapel Hill, 1984.
- [Hur51] H. E. HURST, « Long-Term Storage Capacity of Reservoirs », in : *Transactions of the American Society of Engineers* 116 (1951), p. 770-799.
- [Hur56] H. E. HURST, « Methods of using long-term storage in reservoirs », in : *Proceedings of the Institution of Civil Engineers* 5.5 (1956), p. 519-543.
- [HB93] C. M. HURVICH et K. I. BELTRAO, « Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series », in : *Journal of Time Series Analysis* 14.5 (1993), p. 455-472.
- [Hur+98] M. C. HURVICH, R. DEO et J. BRODSKY, « The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series », in : *Journal of Time Series Analysis* 19 (1998), p. 19-46.
- [Hwa00] S. HWANG, « The Effects of Systemetic Sampling and Temporal Aggregation on Discrete Time Long Memory Processes and their Finite Sample Properties », in : *Econometric Theory* 16.3 (2000), 347-372.
- [IL71] I. IBRAGIMOV et Y. LINNIK, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen, 1971.
- [Kaw81] T. KAWATA, « Uniform Convergence of Random Trigonometric Series and Sample Continuity of Weakly Stationary Processes », in : (1981).
- [KT02] A. KIRMAN et G. TEYSSIÈRE, « Microeconomic Models for Long Memory in the Volatility of Financial Time Series », in : *Studies in Nonlinear Dynamics & Econometrics* 5.4 (2002), p. 1-23.
- [Kou80] I. A. KOUTROUVELIS, « Regression-type estimation of the parameters of stable laws », in : *Journal of the American Statistical Association* 75.372 (1980), p. 918-928.
- [Kou81] I. A. KOUTROUVELIS, « An iterative procedure for the estimation of the parameters of stable laws », in : *Communications in Statistics. B. Simulation and Computation* 10.1 (1981), p. 17-28.

-
- [Kub70] I. KUBO, « On a necessary condition for the sample path continuity of weakly stationary processes », in : *Nagoya Math. J.* 38 (1970), p. 103-111.
- [Kur01] E. E. KURUOGLU, « Density parameter estimation of skewed alpha-stable distributions », in : *IEEE Transactions on Signal Processing* 49.10 (2001), p. 2192-2201.
- [Kwi+92] D. KWIATKOWSKI, P. C. PHILLIPS, P. SCHMIDT et Y. SHIN, « Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root? », in : *Journal of Econometrics* 54.1 (1992), p. 159 -178.
- [LR16] F. LAVANCIER et P. ROCHET, « A general procedure to combine estimators », in : *Computational Statistics & Data Analysis* 94 (2016), p. 175-192.
- [LG13] R. LE GUÉVEL, « An estimation of the stability and the localisability functions of multistable processes », in : *Electronic Journal of Statistics* 7 (2013), p. 1129-1166.
- [LGLV12] R. LE GUÉVEL et J. LÉVY VÉHEL, « A Ferguson-Klass-LePage series representation of multistable multifractional motions and related processes », in : *Bernoulli* 18.4 (2012), p. 1099-1127.
- [LGLV13] R. LE GUÉVEL et J. LÉVY VÉHEL, « Incremental moments and Hölder exponents of multifractional multistable processes », in : *ESAIM. Probability and Statistics* 17 (2013), p. 135-178.
- [LS96] D. LEE et P. SCHMIDT, « On the power of the KPSS test of stationarity against fractionally-integrated alternatives », in : *Journal of Econometrics* 73.1 (1996), p. 285-302.
- [LeP+81] R. LEPAGE, M. WOODROOFE et J. ZINN, « Convergence to a Stable Distribution Via Order Statistics », in : *The Annals of Probability* 9.4 (1981), p. 624-632.
- [LV+19] J. LÉVY VÉHEL, A. PHILIPPE et C. ROBET, « Explicit and combined estimators for stable distributions parameters », in : *preprint* (2019).
- [Li+11] W. LI, C. YU, A. CARRIQUIRY et W. KLIEMANN, « The asymptotic behavior of the R/S statistic for fractional Brownian motion », in : *Statistics & Probability Letters* 81.1 (2011), p. 83 -91.

-
- [LM94] K.-S. LII et E. MASRY, « Spectral estimation of continuous-time stationary processes from random sampling », in : *Stochastic Processes and their Applications* 52.1 (1994), p. 39 -64.
- [Lo91] A. W. LO, « Long-Term Memory in Stock Market Prices », in : *Econometrica* 59.5 (1991), p. 1279-1313.
- [MN95] X. MA et C. L. NIKIAS, « Parameter estimation and blind channel identification in impulsive signal environments », in : *IEEE transactions on signal processing* 43.12 (1995), p. 2884-2897.
- [MT79] B. MANDELBROT et M. TAQQU, « Robust R/S of long run serial correlation. In Proc. of the 42nd Session of the I.S.I », in : *Bullet. of the I.S.I* 48 (1979), p. 69-104.
- [Man97] B. B. MANDELBROT, *Fractals and Scaling In Finance : Discontinuity, Concentration, Risk*, 1st, Springer Publishing Company, Incorporated, 1997.
- [MW69a] B. B. MANDELBROT et J. R. WALLIS, « Computer Experiments with Fractional Gaussian Noises : Part 3, Mathematical Appendix », in : *Water Resources Research* 5.1 (1969), p. 260-267.
- [MW69b] B. B. MANDELBROT et J. R. WALLIS, « Robustness of The Rescaled Range R/S in The Measurement of Noncyclic Long-Run Statistical Dependence », in : *Water Resources Research* 5 (1969), p. 967-988.
- [MS70] M. MARCUS et L. SHEPP, « Continuity of Gaussian Processes », in : *Transactions of the American Mathematical Society* 151 (1970).
- [Mas71] E. MASRY, « Random sampling and reconstruction of spectra », in : *Information and Control* 19.4 (1971), p. 275 -288.
- [Mas78a] E. MASRY, « Alias-free sampling : An alternative conceptualization and its applications », in : *IEEE Trans. Information Theory* 24 (1978), p. 317-324.
- [Mas78b] E. MASRY, « Poisson sampling and spectral estimation of continuous-time processes », in : *IEEE Transactions on Information Theory* 24.2 (1978), p. 173-183.
- [McC86] J. H. MCCULLOCH, « Simple consistent estimators of stable distribution parameters », in : *Communications in Statistics. B. Simulation and Computation* 15.4 (1986), p. 1109-1136.
- [Nol98] J. P. NOLAN, « Parameterizations and modes of stable distributions », in : *Statistics & Probability Letters* 38.2 (1998), p. 187-195.

-
- [Pet95] V. V. PETROV, *Limit theorems of probability theory*, t. 4, Oxford Studies in Probability, Sequences of independent random variables, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995, p. xii+292, ISBN : 0-19-853499-X.
- [PV10] A. PHILIPPE et M.-C. VIANO, « Random sampling of long-memory stationary processes », in : *Journal of Statistical Planning and Inference* 140 (2010), p. 1110-1124.
- [Phi+19] A. PHILIPPE, C. ROBET et M.-C. VIANO, « Random discretization of stationary continuous time processes », in : *preprint* (2019).
- [Phi+07] P. C. B. PHILLIPS, Y. SUN et S. JIN, « Long run variance estimation and robust regression testing using sharp origin kernels with no truncation », in : *J. Statist. Plann. Inference* 137.3 (2007), p. 985-1023.
- [Rip87] B. RIPLEY, *Stochastic simulation*, Wiley Series in Probability and Statistics, J. Wiley, 1987.
- [Rip57] W RIPPLE, « The capacity of storage reservoirs for water supply », in : *Minutes of the Proceedings of the Institution of Civil Engineers* 71.1883 (1957), p. 270-278.
- [Rob95a] P. M. ROBINSON, « Gaussian Semiparametric Estimation of Long Range Dependence », in : *Ann. Statist.* 23.5 (1995), p. 1630-1661.
- [Rob95b] P. M. ROBINSON, « Log-Periodogram Regression of Time Series with Long Range Dependence », in : *Ann. Statist.* 23.3 (1995), p. 1048-1072.
- [Ros90] J. ROSINSKI, « On Series Representations of Infinitely Divisible Random Vectors », in : *Ann. Probab.* 18.1 (1990), p. 405-430.
- [SG+13] D. SALAS-GONZALEZ, J. M. GÓRRIZ, J. RAMÍREZ, I. A. ILLÁN et E. W. LANG, « Linear intensity normalization of FP-CIT SPECT brain images using the α -stable distribution », in : *NeuroImage* 65 (2013), p. 449 -455.
- [ST94] G. SAMORODNITSKY et M. S. TAQQU, *Stable non-Gaussian random processes*, Stochastic Modeling, Stochastic models with infinite variance, Chapman & Hall, New York, 1994.
- [SS94] P. SEN et J. SINGER, *Large Sample Methods in Statistics : An Introduction with Applications*, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis, 1994.

-
- [SS60] H. S. SHAPIRO et H. S. SILVERMAN, « Alias-Free Sampling of Random Noise », in : *Journal of the Society for Industrial and Applied Mathematics* 8.2 (1960), p. 225-248.
- [Shi+10] X. SHI, Y. WU et Y. LIU, « A note on asymptotic approximations of inverse moments of nonnegative random variables », in : *Statistics & Probability Letters* 80.15 (2010), p. 1260 -1264.
- [ST04] S. STOEV et M. TAQQU, « Simulation methods for linear fractional stable motion and farima using the fast fourier transform », in : *Fractals* 12.1 (2004), p. 95-121.
- [Sto74] W. F. STOUT, *Almost sure convergence*, Academic Press New York, 1974.
- [Tsa09] H. TSAI, « On continuous-time autoregressive fractionally integrated moving average processes », in : *Bernoulli* 15 (2009).
- [TC05a] H. TSAI et K. S. CHAN, « Maximum likelihood estimation of linear continuous time long memory processes with discrete time data », in : *Journal of the Royal Statistical Society Series B* 67 (2005), p. 703-716.
- [TC05b] H. TSAI et K. S. CHAN, « Quasi-Maximum Likelihood Estimation for a Class of Continuous-time Long-memory Processes », in : *Journal of Time Series Analysis* 26 (2005), p. 691-713.
- [TC05c] H. TSAI et K. S. CHAN, « Temporal Aggregation of Stationary And Nonstationary Discrete-Time Processes », in : *Journal of Time Series Analysis* 26 (2005), p. 613-624.
- [Via+94] M. VIANO, C. DENIAU et G. OPPENHEIM, « Continuous-time fractional ARMA processes », in : *Statistics & Probability Letters* 21.4 (1994), p. 323 -336.
- [Wan+15] X. WANG, K. LI, P. GAO et S. MENG, « Research on Parameter Estimation Methods for Alpha Stable Noise in a Laser Gyroscope's Random Error », in : *Sensors* 15.8 (2015), p. 18550-18564.
- [Wer95] R. WERON, *Performance of the estimators of stable law parameters*, HSC Research Reports, Hugo Steinhaus Center, Wroclaw University of Technology, 1995.
- [Whi62] P. WHITTLE, « Gaussian estimation in stationary time series », in : *Bulletin of the International Statistical Institute* 39 (1962), p. 105-129.

- [Yan+09] C. YANG, K. HSU et K. CHEN, « The use of the Levy-stable distribution for geophysical data analysis », in : *Hydrogeology Journal* 17.5 (2009), p. 1265-1273.

Titre : Statistique des processus stables et des processus à longue mémoire

Mot clés : processus (multi)stables, log-moments, stationnarité du second ordre, échantillonnage aléatoire, longue mémoire, approches spectrale et temporelle.

Resumé : Ce manuscrit, séparé en deux parties, débute par l'étude des lois et processus α -stables et des processus multistables. Après avoir construit et étudié un estimateur basé sur les log-moments de lois stables, on améliore ses performances en le combinant avec l'estimateur de Koutrouvelis. Puis, nous donnons une méthode approchée afin de simuler rapidement un processus multistable et nous construisons un estimateur de la fonction d'intensité de ce processus à l'aide du rapport de moments empiriques. La deuxième partie est consacrée à l'étude des processus stationnaires du se-

cond ordre à longue mémoire en temps continu. Ce processus est échantillonné à des instants d'observations aléatoires tels que les inter-arrivées soient i.i.d. Le comportement du processus échantillonné est alors étudié dans les domaines temporel et fréquentiel. Une étude plus précise dans le cas d'une fonction d'autocovariance à variation régulière permet de montrer l'évolution de la mémoire après échantillonnage. De plus, pour un processus initialement gaussien, on étudie le périodogramme, les sommes partielles et la convergence de l'estimateur local Whittle pour le paramètre de mémoire.

Title : Statistics of stable processes and long memory processes

Keywords : (multi)stable processes, log-moments, second-order stationarity, random sampling, long memory, spectral and temporal approaches.

Abstract : This manuscript is divided into two parts. The first one is devoted to the study of α -stable distributions and processes and multistable processes. After having built and studied an estimator based on log-moments of the stable distribution, an improvement is obtained by combining it with the Koutrouvelis estimator. Then, we give a non-exact method to simulate efficiently a multistable process, and we construct an estimator of its intensity function using an empirical moments ratio. The second part is devoted to the study of continuous

time second order stationary processes with long memory. This process is sampled at random observation times such that inter-arrivals are i.i.d. The behaviour of the sampled process is then studied in time and frequency domains. For autocovariance functions with regular variation, we study the evolution of the memory after sampling. In addition, for an initially Gaussian process, the periodogram, partial sums and convergence of the local Whittle estimator for the memory parameter are studied.