



HAL
open science

Modélisation de séries financières à l'aide de processus invariants d'échelle. Application à la prédiction du risque.

Alexey Kozhemyak

► **To cite this version:**

Alexey Kozhemyak. Modélisation de séries financières à l'aide de processus invariants d'échelle. Application à la prédiction du risque.. Théorie des représentations [math.RT]. Ecole Polytechnique X, 2006. Français. NNT: . pastel-00002224

HAL Id: pastel-00002224

<https://pastel.hal.science/pastel-00002224v1>

Submitted on 29 Jul 2010

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 pour l'obtention du titre de
DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : Mathématiques Appliquées

présentée et soutenue par

Alexey KOZHEMYAK

**Modélisation de séries financières à l'aide
de processus invariants d'échelle.
Application à la prédiction du risque.**

Jury

Nicole	EL KAROUI	(Présidente du jury)
Jean-Philippe	BOUCHAUD	(Rapporteur)
Marc	HOFFMANN	(Rapporteur)
Emmanuel	BACRY	(Directeur de thèse)
Dominique	GUÉGAN	(Examineur)
Jean-François	MUZY	(Examineur)

ÉCOLE POLYTECHNIQUE – CENTRE DE MATHÉMATIQUES APPLIQUÉES

le 7 décembre 2006

Remerciements

Je tiens en premier lieu à remercier très chaleureusement Emmanuel Bacry, mon directeur de thèse, pour sa grande disponibilité de tous les instants et son soutien au cours de ces trois années. Son exigence et son ouverture scientifique, sa rigueur, son enthousiasme très communicatif, ses critiques et ses encouragements m'ont beaucoup aidé.

Je tiens également à remercier Jean-François Muzy, pour les nombreuses discussions que nous avons eues, pour toutes ses précieuses remarques ainsi que pour les bons moments passés lors du séjour en Corse.

Je n'aurais jamais pensé à travailler dans le domaine de la finance mathématique si je n'avais pas suivi les excellents cours de Nicole El Karoui, d'abord à l'École Polytechnique et ensuite au DEA de Probabilités et Applications de l'Université Paris VI. Je la remercie vivement pour cela mais aussi pour les nombreuses interactions que nous avons eues au cours de ma thèse.

Je remercie aussi tous les membres du jury du temps qu'ils m'ont consacré. Mes remerciements vont tout particulièrement à Jean-Philippe Bouchaud et à Marc Hoffmann qui ont accepté la tâche de rapporteur. Ma reconnaissance va également à Dominique Guégan qui me fait l'honneur d'être membre de mon jury.

Je voudrais exprimer ma reconnaissance à Wolfgang Härdle pour m'avoir invité au laboratoire d'économétrie à l'Institut de Statistique et Econométrie de l'Université Humboldt de Berlin et à Kai Detlefsen pour son aide lors de ce séjour.

Toute ma gratitude va aux différents membres du CMAP, à son ancien directeur Vincent Giovangigli, à la secrétaire Jeanne Bailleul, au groupe de travail de finance, à mes camarades Charles et Benjamin et à Sylvain Ferrand pour son aide immense en informatique. J'ai eu énormément de chance de pouvoir travailler au sein d'une équipe aussi agréable et sympathique.

J'ai également une pensée pour mes amis Sergueï, Alexandre, Peter et Alexei.

Enfin, je dédie ce mémoire à toute ma famille pour leurs encouragements précieux, à Anastasia et à Ivan pour leur patience.

Table des matières

Introduction	17
I Modélisation des cours boursiers	23
1 Faits stylisés et modèles classiques	25
1.1 Description des données réelles	25
1.1.1 Données journalières	26
1.1.2 Données à haute fréquence	28
1.2 Faits stylisés	28
1.2.1 Notations	29
1.2.2 Stationnarité et ergodicité des rendements logarithmiques	30
1.2.3 Distribution des rendements	31
1.2.4 Autocorrélation des rendements logarithmiques	36
1.2.5 Mémoire longue des marchés financiers	38
1.2.6 Volatilité réalisée	39
1.2.7 Accumulation des volatilités	40
1.2.8 Corrélation entre le rendement et la volatilité instantanée	40
1.2.9 Conclusion	42
1.3 Modèles classiques	42
1.3.1 Modèles stochastiques de Bachelier et de Samuelson	42
1.3.2 Modélisation de cours financiers par des processus de Lévy	44
1.3.3 Modèles de type ARCH	45
1.3.4 Modèles à volatilité stochastique	51
1.4 Conclusion	53
2 Multifractalité	55
2.1 Processus autosimilaires	57
2.1.1 Lois infiniment divisibles	57
2.1.2 Lois stables	59
2.1.3 Processus stables	61
2.1.4 Processus autosimilaires	62
2.1.5 Processus autosimilaires à accroissements stationnaires	63
2.1.6 Analyse des étendues normalisées	64
2.1.7 Exemples	66
2.1.8 Conclusion	68
2.2 Analyse multifractale	68
2.2.1 Invariance d'échelle	69
2.2.2 Autosimilarité au sens stochastique	72

2.2.3	Régularité locale, exposant de Hölder	76
2.2.4	Spectre de singularités	76
2.2.5	Formalisme multifractal	77
2.3	Faits stylisés multifractals	81
2.3.1	Hétérogénéité du marché	81
2.3.2	Invariance d'échelle des rendements	81
2.3.3	Autocorrélation des magnitudes	83
2.4	Modèles multifractals	83
2.4.1	Modèles en cascade discrète	84
2.4.2	Modèle MMAR de Mandelbrot, Fisher et Calvet	90
2.4.3	Modèle MSM de Calvet et Fisher	92
2.4.4	Modèles en cascade continue	94
2.4.5	Modèle MPCP de Barral et Mandelbrot	97
2.4.6	Modèle MRW de Bacry et Muzy	97

II Modèle MRW 99

3 Modèle MRW log-infiniment divisible 101

3.1	Mesure multifractale MRM	102
3.1.1	Construction et non dégénérescence de la mesure MRM	102
3.1.2	Propriétés de la mesure MRM	106
3.1.3	Moments d'ordre négatif de la mesure MRM	110
3.1.4	Exemples	116
3.1.5	Construction pour une invariance d'échelle asymptotique	119
3.2	Construction du processus multifractal MRW	119
3.2.1	Construction d'un mouvement brownien subordonné	119
3.2.2	Propriété du processus MRW	121
3.3	Invariance par changement du temps intégral	121
3.4	Modèle MRW log-normal	123
3.4.1	Construction alternative de la mesure MRM log-normale	123
3.4.2	Propriétés du processus MRW log-normale	124
3.4.3	Simulation numérique	127
3.4.4	Faits stylisés	128
3.4.5	Résumé de la section	135
3.5	Conclusion	137

4 Événements extrêmes 139

4.1	Exposants des queues des distributions	140
4.1.1	Estimation des exposants des queues	140
4.1.2	Loi de distribution de maximum	141
4.2	Statistique des valeurs extrêmes multifractales	142
4.2.1	Résultats dans le cas des processus stationnaires	142
4.2.2	Limite asymptotique	143

4.2.3	Comportement des estimateurs classiques	143
4.2.4	Expériences numériques	146
4.3	Application aux séries financières	147
4.4	Conclusion	148
5	Approximation limite de petite intermittence	151
5.1	Moments d'ordre généralisés de la mesure MRM log-normale	152
5.1.1	Représentation intégrale des moments généralisés	152
5.1.2	Développement en série de Taylor des moments généralisés centrés	154
5.2	Moments généralisés du logarithme de la mesure MRM log-normale	157
5.2.1	Existence des moments	157
5.2.2	Développement en série de Taylor	158
5.3	Magnitude renormalisée	161
5.3.1	Définition	162
5.3.2	Existence de la magnitude renormalisée	162
5.3.3	Autocovariance des accroissements de la magnitude renormalisée	164
5.4	Approximation log-normale de la mesure MRM	165
5.4.1	Convergence en loi vers la magnitude renormalisée	165
5.4.2	Approximation des moments généralisés logarithmiques	165
5.4.3	Approximation des moments généralisés	167
5.5	Approximation du processus MRW	168
5.5.1	Approximation des moments généralisés	168
5.5.2	Approximation des moments généralisés logarithmiques	169
5.6	Simulations numériques	170
6	Estimation des paramètres	175
6.1	Régression de la fonction d'autocovariance des magnitudes	176
6.1.1	Fonction d'autocovariance empirique	176
6.1.2	Régime asymptotique "historique infini"	177
6.2	Estimation des paramètres par la méthode GMM	178
6.2.1	Méthode des moments généralisés	179
6.2.2	Estimation des paramètres du modèle MRW log-normal	180
6.2.3	Application aux données journalières	181
6.3	Estimation des paramètres lorsque le temps intégral est grand	182
6.3.1	Estimation du temps intégral	184
6.3.2	Estimation de la variance	185
6.3.3	Estimation du coefficient d'intermittence	186
6.3.4	Modèle MRW log-normal n'a que deux paramètres effectifs	186
6.4	Estimation du coefficient d'intermittence	186
6.4.1	Représentation matricielle de la fonction d'autocovariance empirique	188
6.4.2	Propriétés de la différence des fonctions d'autocovariance empiriques	189
6.4.3	Application au logarithme de la mesure MRM log-normale	190
6.4.4	Application au logarithme du processus MRW log-normal	192
6.4.5	Application aux données journalières	192

III	Prédiction des risques	195
7	Prédiction de volatilité	197
7.1	Position du problème de prédiction de volatilité	198
7.2	Prédiction de volatilité dans le cadre du modèle GARCH	199
7.2.1	Prédiction	199
7.2.2	Paramètres utilisés	199
7.3	Prédiction de volatilité dans le cadre du modèle MRW	200
7.3.1	Prédiction de volatilité linéaire et de volatilité quadratique	201
7.3.2	Prédiction de volatilité logarithmique	202
7.3.3	Paramètres utilisés	202
7.4	Evaluation des méthodes de prédiction de volatilité	203
7.4.1	Mesures MAE, MSE, R^2	203
7.4.2	Régression de Mincer-Zarnowitz	204
7.5	Applications aux données journalières	204
7.6	Conclusion	206
8	Prédiction de Valeur à Risque	211
8.1	Définitions	212
8.2	Prédiction de VaR	212
8.2.1	Différentes méthodes de prédiction de VaR	212
8.2.2	Calcul implicite de la VaR	213
8.3	Prédiction de VaR dans le cadre du modèle GARCH	214
8.4	Prédiction de VaR dans le cadre du modèle MRW	215
8.5	Evaluation des différentes méthodes de prédiction de VaR	215
8.5.1	Test statistique de Kupiec	216
8.5.2	Test statistique de Christoffersen	217
8.5.3	Quantiles empiriques	217
8.6	Applications aux données journalières	218
8.7	Conclusion	218
IV	Vers une modélisation de séries financières à haute fréquence	225
9	Données à haute fréquence	227
9.1	Présentation des séries intra-journalières	228
9.2	Effets saisonniers	228
9.2.1	Mise en évidence des effets saisonniers	229
9.2.2	Etude de l'activité intra-journalière	230
9.2.3	"Durées empiriques" des nuits, des jours fériés et des week-ends	232
9.3	Deux modèles d'évolution des prix intra-journaliers	232
9.3.1	Volatilité saisonnière multiplicative	233
9.3.2	Volatilité saisonnière subordonnée	233
9.3.3	Choix du modèle	233
9.4	Estimation des paramètres	235

9.4.1	Volatilité saisonnière multiplicative	235
9.4.2	Volatilité saisonnière subordonnée	235
9.5	Prédiction de volatilité	237
9.6	Prédiction de VaR conditionnelle	239
9.7	Conclusion	241
Conclusion		243
A	Eléments de l'économétrie linéaire	245
B	Dimension de Hausdorff	249
C	Simulation de variables gaussiennes corrélées	251
D	Inégalités	255
E	Fonction d'autocovariance de la magnitude renormalisée	261
F	Fonction d'autocovariance de magnitude	263
G	Lemme auxiliaire	267
H	Article	269
Bibliographie		303

Table des figures

1.1	Taux de change GBP/USD	27
1.2	Rendements logarithmiques du taux de change GBP/USD	29
1.3	Déformation de la distribution des rendements du taux de change GBP/USD	33
1.4	Quantiles empiriques des rendements du taux de change GBP/USD	36
1.5	Autocorrélation des rendements du taux de change GBP/USD	37
1.6	Autocorrélation des rendements absolus du taux de change GBP/USD	39
1.7	Volatilité des rendements du taux de change GBP/USD	41
2.1	Chou romanesco et suite qui tend vers la flocon de Koch	56
2.2	Analyse R/S des rendements journaliers du taux de change GBP/USD	66
2.3	Exemple de l'exposant multifractal $\zeta(q)$	79
2.4	Exemple du spectre de singularités $D(\alpha)$	80
2.5	Moments des rendements du taux de change GBP/USD	82
2.6	Autocovariance des magnitudes du taux de change GBP/USD	84
2.7	Construction de la mesure binomiale	85
2.8	Construction de la cascade multiplicative	87
2.9	Représentation 2D de la cascade multiplicative de Mandelbrot	95
2.10	Représentation 2D de la cascade dans le modèle MSM de Calvet et Fisher	96
3.1	Cônes $\mathcal{A}_{l,T}(t_1)$, $\mathcal{A}_{l,T}(t_2)$ et $\mathcal{A}_{l,T}(t_1, t_2)$ intervenant dans le modèle MRW	104
3.2	Cône $\mathcal{B}_{l,T}(t)$ et domaine $\mathcal{C}_T(t)$	110
3.3	Réalisation du processus MRW log-normal	129
3.4	Accroissements du processus MRW log-normal	130
3.5	Déformation de la distribution des accroissements du processus MRW	131
3.6	Quantiles empiriques des accroissements du processus MRW	132
3.7	Autocorrélation des accroissements du processus MRW log-normal	133
3.8	Autocorrélation des accroissements absolus du processus MRW	134
3.9	Carrés des accroissements du processus MRW log-normal	135
3.10	Moments des accroissements du processus MRW log-normal	136
3.11	Autocovariance des magnitudes du processus MRW log-normal	137
4.1	Comportement de l'estimateur de Zipf appliqué à la mesure MRW	146
4.2	Diagramme rang-fréquence pour le cours du taux de change GBP/USD	147
5.1	Distributions du logarithme de la mesure MRM et l'approximation normale	170
5.2	Distributions de la mesure MRM et l'approximation log-normale	171
5.3	Distributions du logarithme de la mesure MRM et l'approximation normale	172
5.4	Distributions de la mesure MRM et l'approximation log-normale	173
6.1	Autocovariance des magnitudes du processus MRW log-normal	178

6.2	Autocovariance des magnitudes du processus MRW log-normal	187
8.1	Quantile empirique de la prédiction de VaR du taux de change GBP/USD	222
8.2	Quantile empirique de la prédiction de VaR du taux de change GBP/USD	223
9.1	Prix du contrat à terme sur le taux de change GBP/USD	229
9.2	Autocorrélation des rendements carrés du contrat à terme GBP/USD . .	230
9.3	Transformée de Fourier des rendements carrés du contrat à terme GBP/USD	231
9.4	Structure de volatilité saisonnière intra-journalière des taux de change . .	231
9.5	Choix du modèle de volatilité saisonnière	234
9.6	Autocovariance des magnitudes désaisonnalisées	236
9.7	Autocovariance des logarithmes de la volatilité agrégée	237
9.8	Autocovariance moyenne des magnitudes	238

Liste des tableaux

1.1	Forex. Statistiques	34
1.2	CAC 40. Statistiques	35
1.3	Forex. Estimation de paramètres du modèle GARCH(1,1)	49
1.4	CAC 40. Estimation de paramètres du modèle GARCH(1,1)	50
3.1	Statistiques d'une réalisation du processus MRW log-normal	129
6.1	Forex. Estimations GMM des paramètres du modèle MRW log-normal	182
6.2	CAC 40. Estimations des paramètres du modèle MRW log-normal	183
7.1	Prédiction de volatilité du taux de change GBP/USD	205
7.2	Forex. Evaluation de la prédiction de volatilité	207
7.3	Forex. Résumé de la prédiction de volatilité	208
7.4	CAC 40. Résumé de la prédiction de volatilité	209
8.1	Prédiction de VaR du taux de change GBP/USD	219
8.2	Forex. Résumé de la prédiction de VaR	220
8.3	CAC 40. Résumé de la prédiction de VaR	221
9.1	Activité des contrats à terme sur taux de change	232
9.2	Prédiction de volatilité en utilisant les données à haute fréquence	239
9.3	Prédiction de VaR en utilisant les données à haute fréquence	240

‘Contrariwise,’ continued Tweedledee, ‘if it was so, it might be; and if it were so, it would be; but as it isn’t, it ain’t. That’s logic.’

‘Through the Looking Glass’ by Lewis Carroll

‘In fact, the more head-downwards I am, the more I keep inventing new things.’

‘Through the Looking Glass’ by Lewis Carroll

Table des notations

$\lfloor x \rfloor$	partie entière du nombre réel x
$\Re[x], \Im[x]$	parties réelle et imaginaire du nombre complexe x
$\mathbb{I}d_n$	matrice identité d'ordre n
X^T, M^T	vecteur X transposé, matrice M transposée
$f(t), X(t)$	fonction, processus à temps continu
$X[k] = X(kl)$	processus à temps discret ou échantillonnage du processus à temps continu
$\delta_l X(t) = X(t) - X(t-l)$	accroissement du processus à l'instant t et à l'échelle l
$\delta_m X[k] = \delta_{ml} X(kl)$	accroissement du processus à temps discret ou de l'échantillonnage du processus à temps continue
$\mathbb{P}[X], \mathbb{E}[X], \text{Var}[X]$	probabilité, espérance et variance de la variable X
$\mathbb{P}[X Y], \mathbb{E}[X Y], \text{Var}[X Y]$	probabilité, espérance et variance conditionnelles de la variable X sachant Y
$\text{Cov}[X, Y], \text{Corr}[X, Y]$	covariance, corrélation des variables X et Y
$X \stackrel{\mathcal{L}}{=} Y$	égalité en loi des variables aléatoires X et Y
$X = Y$ p.s.	égalité presque sûr des variables aléatoires X et Y
$\{X(t)\}_t \stackrel{\mathcal{L}}{=} \{Y(t)\}_t$	égalité en loi des processus aléatoires X et Y
$\text{VaR}_p(l)$	Valeur à Risque de niveau p et de l'échelle l
$\text{VaR}_p(t, m, h)$	Valeur à Risque conditionnelle de niveau p à l'instant t de l'échelle m et de l'horizon h
$\zeta(q)$	exposant multifractal
$X(t)$	processus MRW
$M(t) = M([0, t])$	mesure MRM de l'intervalle $[0, t]$

Introduction

Modélisation de cours boursiers

La théorie de la finance concerne l'étude des prix des produits financiers et plus précisément de leurs évolutions temporelles. Un modèle financier repose le plus souvent sur une représentation des prix d'actifs financiers (ou des niveaux de taux d'intérêt). Une telle représentation peut être recherchée dans une perspective de meilleure compréhension des marchés financiers, ou encore dans une perspective d'élaboration d'outils permettant d'améliorer la gestion financière : gestion des risques, allocation d'actifs, création de nouveaux produits financiers.

Une modélisation financière réalise toujours un compromis entre adéquation aux observations dans les marchés financiers et commodité d'utilisation. Une modélisation qui cherche à reproduire toutes les propriétés statistiques des observations conduit généralement à un modèle compliqué qui se révèle souvent difficile à utiliser, les calculs théoriques étant généralement difficile à mener. A l'inverse, une simplification excessive du modèle permet, certes, de mener à bout de nombreux calculs, par exemple, pricing des options et gestion du portefeuille, mais risque de ne pas être cohérent avec la réalité financière. Notre travail se situe entre les deux approches, bien que plus proche de la première. Nous analysons les principales propriétés statistiques des séries financières et nous proposons une modélisation cohérente de la plupart de ces propriétés. Cependant, nous ne nous contentons pas d'un modèle purement descriptif : la majeure partie de ce travail est motivée par l'utilisation du modèle ainsi construit pour la prédiction de risque.

Une caractéristique des prix des produits financiers est que leur évolution ne peut s'expliquer par des modèles déterministes simples. La plupart des modélisations financières reposent sur une représentation probabiliste. Ainsi, le problème de la détermination du prix à une date future revient à déterminer la loi de probabilité du prix de l'actif, ce dernier étant vu comme une variable aléatoire.

En statistique, une des lois les plus utilisées est la loi gaussienne. Ses différentes propriétés, comme la stabilité, le fait que deux paramètres (moyenne et variance) suffisent à la

caractériser ou bien encore le théorème central limite font que cette loi s'adapte bien dans de nombreuses études, et notamment dans le domaine des séries chronologiques financières.

La première modélisation des cours boursiers, proposée en 1900 par Bachelier [Bac00], a consisté à modéliser les cours financiers par des mouvements browniens (cf. section 1.3.1). Il faut attendre plus de 60 ans pour qu'un autre modèle voit le jour et s'installe. En 1960, Samuelson [Sam65] propose de retenir cette modélisation pour les rendements, plutôt que pour les cours eux-mêmes. Il est important de souligner que cette modélisation de Samuelson est à la base de la théorie très populaire de Black et Scholes [BS73]. Il faut noter que les propriétés très spécifiques de la modélisation gaussienne permettent de faciliter beaucoup de calculs théoriques dans la théorie financière.

Dans les années 60, Mandelbrot étudie des fluctuations boursières, pour lesquelles il était tout à fait clair que le modèle gaussien ne convenait pas. En effet, l'analyse des queues de distribution des rendements des actifs financiers, *a priori* indispensable pour une mesure fine des risques, montre que les valeurs des exposants de queues sont beaucoup trop grandes pour correspondre à des distributions normales. Mandelbrot utilise alors les lois de Pareto pour mettre en évidence un nouveau modèle de variation des prix basé sur les "lois α -stables" (cf. section 2.1.2) dont la variance est infinie. Dans l'article [Man63] Mandelbrot affirme que son modèle décrit de façon réaliste la variation des prix observés sur certaines bourses des valeurs. Deux ans plus tard, Fama [Fam65] valide ce modèle pour le marché des actions.

Avec l'importance croissante des marchés et de leur volatilité, en particulier sur le marché des changes depuis la disparition des accords de Bretton-Woods en 1971, l'évaluation dynamique de portefeuilles a fait objet de très nombreuses recherches. A la fin des années 80, un grand nombre de travaux semblent rejeter la modélisation en finance par des variables aléatoires indépendantes identiquement distribuées quelle que soit leur loi, on remet en question l'hypothèse d'indépendance des rendements. On s'intéresse alors au processus de volatilité qui a toujours été un paramètre essentiel pour la gestion de portefeuille. La formalisation par Engle d'une corrélation à courte portée (exponentielle) des volatilités a conduit à la modélisation ARCH [Eng82] (cf. section 1.3.3) et plus tard aux modèles à volatilité stochastique [Tay86, Tay94, GHR95, Hes93, BNS01] (cf. section 1.3.4).

Avec la mise en évidence de la mémoire longue de la volatilité (cf. section 1.2.5) (décroissance lente de la fonction d'autocovariance), une autre classe de processus a suscité depuis peu un intérêt croissant en finance, la classe des processus invariants d'échelle. Ces processus sont essentiellement des processus dont les trajectoires présentent des irrégularités à toutes échelles.

Dans le domaine des processus stochastiques, les idées d'invariance d'échelle ont donné naissance à la vaste classe des processus autosimilaires (cf. section 2.1) et, plus généralement, aux processus multifractals.

La notion de "fractale" a été introduite la première fois par Mandelbrot pour décrire le phénomène de turbulence et elle est devenue très populaire dans ce domaine notamment à la suite des travaux de Frisch et Parisi [FP85, PF85, Fri95]. L'analyse multifractale

(cf. section 2.2) permet de caractériser finement la structure singulière d'un objet fractal. L'importance de l'invariance d'échelle en finance a été mise en évidence assez récemment.

Historiquement, les premiers objets multifractals proposés dans la littérature ont été les mesures multifractales construites à partir de cascades qui ont été développées par l'école russe [Yag66] pour la modélisation en turbulence. Après les premiers travaux de Mandelbrot [Man74a, Man74b, Man89, Man03], de nombreuses études ont été consacrées aux cascades aléatoires multiplicatives discrètes (cf. section 2.4). Cependant, les processus obtenus à l'aide de ces cascades ne sont pas satisfaisant pour une modélisation financière, leur défaut majeur concerne la non stationnarité des accroissements et l'introduction d'un rapport d'échelle privilégié arbitraire.

Les notions relatives aux cascades continues infiniment divisibles ont été abondamment utilisées, dans de nombreux domaines, avant même que l'existence de tels objets ait été prouvée [CGH90, Nov94, Cas96, AMR97]. Au cours des années 90, l'urgence de construire de tels processus se faisait de plus en plus pressante. Ce n'est qu'en 2001, que l'utilisation de cônes dans le demi-plan supérieur, en vue de construire des cascades continues infiniment divisibles, apparaît pour la première fois dans [SM01]. Cependant, dans ce travail, aucune construction de processus ou de mesure stochastique n'est encore réellement proposée. Peu de temps après, une construction effective unidimensionnelle (court-circuitant donc l'approche standard bi-dimensionnelle des cascades multiplicatives) de la mesure MRM log-normale (cf. chapitre 3) est décrite dans [BDM01]. Aucune preuve d'existence n'est donnée, mais plusieurs calculs formels sont effectués et, notamment, l'invariance d'échelle exacte est démontrée sous condition d'existence. Nous étudions cette construction dans la section 3.4.1.

Parallèlement, Barral et Mandelbrot [BM02] ont explicité la construction de mesures MRM dans le cas log-poisson composé. Ces auteurs ne se sont cependant pas intéressés aux propriétés d'invariance d'échelle de ces mesures et se sont attachés à l'étude de leurs spectres de singularités. Les mesures MRM ainsi construites n'ont d'ailleurs pas d'invariance d'échelle exacte, elles possèdent une invariance d'échelle asymptotique.

En 2003, Bacry et Muzy ont formulé [BM03] une description commune à tout ces modèles, permettant une généralisation au cas log-infiniment divisible ainsi que le choix du type d'invariance d'échelle souhaitée (exacte ou asymptotique).

Notons que Chainais, Riedi et Abry ont publié un travail [CRA03] effectué indépendamment du travail de Bacry et Muzy, dans lequel ils construisent les mesures MRM log-infiniment divisibles dans un cadre moins générale que [BM03] (notamment convergence L^2 et pas d'invariance d'échelle exacte).

Lorsque nous avons commencé notre travail, plusieurs travaux avaient déjà été effectués dans le domaine de modélisation financière avec des processus multifractals, par exemple, [MDB00, CF01, Bou01, BMP01, CF02, Poc03, CF04].

Plan de la thèse

Dans la première partie de ce travail nous commençons par rappeler (chapitre 1) les propriétés statistiques, ou **faits stylisés**, généralement observées dans les marchés financiers. Nous les illustrons, principalement, sur des données issues de marchés des changes (le taux de change GBP/USD) et nous décrivons les modèles classiques d'évolutions des cours financiers. Dans le chapitre 2, après avoir effectué une brève revue des processus autosimilaires, nous introduisons des méthodes de base de l'analyse multifractale en nous appuyant sur la notion d'invariance d'échelle et, dans la suite, nous étudions les faits stylisés fractals observés dans les marchés financiers et nous décrivons brièvement des modèles multifractals.

Dans la seconde partie, plus mathématique, nous étudions en détail les propriétés du modèle MRW log-normal qui permet de reproduire la plupart des faits stylisés. Nous effectuons cette étude en plusieurs étapes. Dans un premier temps, dans le chapitre 3, nous introduisons et étudions les propriétés du modèle plus général MRW log-infinitement divisible, et notamment ses propriétés d'invariance d'échelle. Dans ce chapitre, notre contribution originale se résume aux résultats présentés dans les sections 3.1.3 et 3.3 qui concernent les propriétés d'existence des moments d'ordre négatifs de la mesure MRM et d'invariance d'échelle par changement du temps intégral.

Le grand avantage du modèle MRW log-normal est qu'il dépend de trois paramètres seulement : le coefficient d'intermittence λ^2 , qui quantifie à la fois la multifractalité du processus (l'écart au mouvement brownien) et la force des corrélations de la volatilité, le temps intégral T , qui caractérise l'échelle au delà de laquelle le processus cesse de posséder l'invariance d'échelle, et la variance σ^2 , qui joue le rôle d'un facteur multiplicatif.

Tous les résultats des chapitres suivant le chapitre 3 sont entièrement originaux.

Dans le chapitre 4 nous nous intéressons au problème de l'estimation de l'exposant de queues des distributions du processus MRW log-normal. En finance, de nombreux travaux semblent prouver que les moments des rendements logarithmiques d'ordre 3 ou 4 explosent. Le modèle MRW log-normal quant à lui prédit une explosion à l'ordre 30! Les résultats de ce chapitre montrent que les estimateurs les plus couramment utilisés sont extrêmement instables et ont besoin d'une quantité faramineuse de données pour atteindre le régime asymptotique. Quantité qui n'est pas accessible en pratique. Les estimateurs ainsi obtenus sont *a priori* très fortement biaisés et permettent d'expliquer les écarts observés (cf. section 4.3).

Les études du comportement de la mesure log-normale et du processus MRW log-normal, effectuées dans le chapitre 5, nous conduisent à énoncer, dans ce chapitre, une série de théorèmes qui montre la validité d'une approximation log-normale de la volatilité (stochastique) du processus MRW log-normale, sous condition de faible intermittence ($\lambda^2 \ll 1$), ce qui est le cas des séries financières, comme nous le verrons dans le chapitre 6. Les résultats théoriques obtenus dans ce chapitre sont essentiels au reste de ce travail, puisque tous les chapitres suivants les utilisent.

Dans un premier temps, ils permettent d'expliquer la structure de corrélation de la volatilité entièrement régie par le coefficient d'intermittence λ^2 et le temps intégral T . La connaissance d'une telle structure de corrélation nous permet, dans le chapitre 6, de bien poser le problème d'estimation des paramètres du modèle MRW log-normal, et notamment de montrer que, en pratique, le régime asymptotique "historique infini" n'est pas atteint, et que le paramètre T n'est pas réellement un paramètre, il peut être choisi arbitrairement. Les résultats théoriques, à la fin de ce chapitre 6, permettent de montrer que le coefficient d'intermittence peut être estimé dans une limite asymptotique "haute fréquence", ce qui, en pratique explique pourquoi les estimations, obtenues par la régression linéaire, sont souvent bonnes même si la limite asymptotique "historique infini" est loin d'être atteinte. Un des résultats principaux de ce chapitre montre que dans le cas des séries financières il ne reste que deux paramètres effectifs du modèle MRW log-normal : le coefficient d'intermittence λ^2 et la variance "locale".

Dans le troisième partie, nous utilisons les résultats des précédents chapitres afin de construire des prédictors de risque : prédictors de volatilité, dans le chapitre 7, et de Valeur à Risque (VaR), dans le chapitre 8. Nous effectuons les comparaisons entre ces différentes méthodes de prédiction de volatilité et de VaR avec les prédictors, classiquement utilisés en économétrie, basés sur les modèles GARCH(1,1). Afin de favoriser les prédictions obtenues dans le cadre des modèles GARCH(1,1), nous avons choisi d'effectuer les prédictions pour ces modèles en *in-sample* et celles pour le modèle MRW log-normal en *out-of-sample*. De plus, seul le paramètre de variance σ^2 est réellement estimé, le temps intégral T est fixé arbitrairement à l'échelle d'observation et le coefficient d'intermittence λ^2 est fixé pour tous les cours. Notons que les résultats de prédiction sont peu dépendants de la valeur choisie de λ^2 . Les résultats obtenus à l'aide des méthodes développées dans le cadre du modèle MRW log-normal (cf. sections 7.5 et 8.6) sont systématiquement plus performants que ceux obtenus dans le cadre des modèles GARCH(1,1).

La quatrième partie de ce travail est constitué d'un chapitre de prospective (chapitre 9) qui montre qu'on peut construire un modèle basé sur le modèle MRW log-normal qui permet de reproduire les structures de dépendance saisonnière des données financières intra-journalières. La saisonnalité est modélisée en utilisant une subordination du processus MRW log-normal par un temps physique qui porte l'intégralité de la structure saisonnière journalière.

Dans le cadre de ce modèle nous proposons (cf. section 9.5) une méthode de prédiction de volatilité qui utilise à la fois les données journalières (prix de clôture) et les données intra-journalières. Cette méthode permet d'améliorer les résultats obtenus dans le chapitre 7.

Première partie

Modélisation des cours boursiers

Faits stylisés et modèles classiques

La section 1.1 de ce chapitre présente les données réelles que nous utilisons tout au long de ce travail. Les données journalières sont issues du marché des actions de Paris et du marché des changes Forex. Les données à haute fréquence sont issues du marché des contrats à terme sur les taux de changes.

Dans la section 1.2, nous présentons les **faits stylisés**, c'est-à-dire les propriétés statistiques observées dans la plupart des marchés financiers. Nous les illustrons, en général, sur les données journalières du taux de change GBP/USD et, dans certains cas, sur l'ensemble des données journalières. Les principaux faits stylisés sont la non normalité ou les queues épaisses de la distribution des rendements, la lente décroissance de la fonction d'autocorrélation des rendements logarithmiques absolus (des valeurs absolues des logarithmes des rendements) et le phénomène d'accumulation des volatilités. Les faits stylisés observés indiquent le caractère non gaussien des rendements logarithmiques et la présence de dépendances très particulières entre eux.

Dans la deuxième partie de ce chapitre, dans la section 1.3, nous effectuons une brève revue des modèles d'évolution des prix des actifs financiers utilisés traditionnellement en finance : les modèles de Bachelier et de Samuelson, le modèle exponentielle-Lévy, les modèles de type ARCH, les modèles à volatilité stochastique. Pour chacun de ces modèles nous indiquerons les faits stylisés qui sont reproduits et ceux qui ne le sont pas.

1.1 Description des données réelles

Deux types de données seront utilisées tout au long de ce travail : des cours journaliers et des cours intra-journaliers (*intraday*). Dans cette section nous les décrivons en détails.

1.1.1 Données journalières

Composantes de l'indice français CAC 40

Le **CAC 40** est le principal indice boursier sur la place de Paris. Créé par la Compagnie des Agents de Change, il est déterminé à partir des cours de 40 actions cotées en continu sur le Premier Marché parmi les 100 sociétés les plus capitalisées sur Euronext Paris.

Ces valeurs, représentatives des différentes branches d'activités, reflètent en principe la tendance globale de l'économie des grandes entreprises françaises et leur liste est revue régulièrement pour maintenir cette représentativité.

L'indice CAC 40 est mis à jour toutes les 30 secondes pendant la journée de 9h00 à 17h30. CAC, qui signifiait Compagnie des Agents de Change, signifie aujourd'hui Cotation Assistée en Continu : l'indice donne donc, en continu, une idée de l'évolution du marché.

Chacune des 40 sociétés pondère l'indice en fonction de la quantité de titres disponibles sur le marché. Les pondérations varient d'une société à l'autre en fonction de l'importance de sa capitalisation et des échanges survenus sur la valeur.

Quand une société n'est plus cotée, elle est remplacée par une des valeurs du CAC Next 20 répondant aux exigences financières de cotation dans l'indice CAC 40 (liquidité du titre, capitalisation boursière suffisante, échange de titres quotidiens important...)

Il ne faut pas confondre l'évolution à long terme du CAC 40 et celle à long terme de l'ensemble des valeurs cotées sur cette bourse, puisque les valeurs qui se comportent mal finissent par être remplacées dans sa liste par d'autres ayant pris l'avantage sur elles. Ainsi un portefeuille comprenant les 40 valeurs du CAC 40 du jour se dégradera progressivement et son évolution ne sera plus celle du CAC 40 trois mois plus tard et ainsi de suite au fil du temps.

Dans le cadre de cette thèse nous utilisons les prix de clôture d'un ensemble des cours de 29 actions qui entraient dans la composition de l'indice CAC 40 au 11 avril 2005 et qui avaient le plus long historique de cotation. Ces cours étaient fournis par la base de donnée de DataStream. Tous les cours utilisés comprennent environ 3770 points de 1990 à 2005 et ils sont ajustés afin de tenir compte des dividendes et des divisions¹ (*splits*).

Taux de change Forex

Issu de la contraction des termes anglais Foreign Exchange, **Forex** est le surnom universellement donné au marché des changes. Ce marché mondial, qui est essentiellement interbancaire, est le deuxième marché financier de la planète en terme de volume global, derrière celui des taux d'intérêt. C'est néanmoins le plus concentré et le premier pour la liquidité des produits les plus traités, comme la parité euro/dollar.

¹Une société peut décider de diviser la valeur nominale de ses actions. Dans ce cas, le nombre d'actions composant le capital se multiplie dans la même proportion.

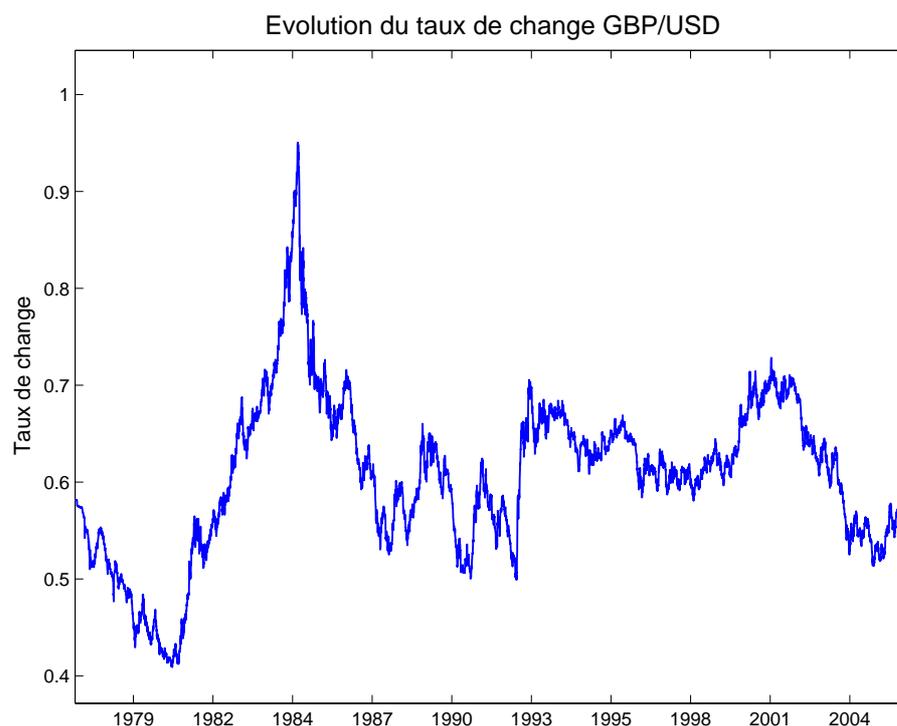


FIG. 1.1 – Valeurs du clôturé du taux de change (GBP/USD) de la livre britannique par rapport au dollar américain, du 01 juillet 1977 au 20 mars 2006, 7208 points. Son comportement présente un caractère erratique.

Le marché des changes existe sous sa forme actuelle, dite régime des changes flottants, depuis mars 1973 et l'abandon de la fixité des taux de change des diverses monnaies par rapport à l'étalon dollar issue des accords de Bretton Woods en 1944. Malgré le fort développement de l'euro, le dollar reste le pivot dominant.

Le marché Forex est le plus grand marché financier du monde, avec un chiffre d'affaire quotidien de plus de 2 trillions de dollars. C'est un marché de gré à gré, ce qui signifie qu'il n'y a aucun lieu physique, pas de bourse et d'horaires où les ordres sont appariés. Au contraire, ce marché est ouvert vingt-quatre heures par jour grâce à un réseau électronique de banques, d'entreprises et d'individus qui négocient une monnaie contre une autre.

Pour nos études nous utilisons les valeurs du clôturé des taux de change journaliers sur la période de 1977 à 2006 des différentes devises par rapport au dollar américain (USD) : du dollar canadien (CAD), du yen japonais (JPY), du franc suisse (CHF) et de la livre britannique (GBP), présenté sur la figure 1.1. Toutes ces séries ont été téléchargées de la page Web de *Federal Reserve Bank of St. Louis* <http://research.stlouisfed.org/>.

Dans nos études basées sur des données journalières nous négligeons les effets saisonniers provoqués par les week-ends et les jours fériés. Les séries financières journalières sont donc

considérées comme étant uniformément échantillonnées.

Dans la section 9.2 nous étudions en détails les effets saisonniers.

1.1.2 Données à haute fréquence

La disponibilité récente de données financières à haute fréquence, c'est-à-dire ayant des fréquences correspondantes aux échelles inférieures à un jour, permet de réaliser l'étude des propriétés statistiques des marchés de façon très fine. Il est actuellement possible de disposer de bases de données très détaillées pour certaines grandes devises du marché des changes, où toutes les transactions (prix et volume) et les cotations (cours acheteur *bid* et cours vendeurs *ask*²) sont enregistrés.

Contrats à terme sur taux de change

Un *future* ou **contrat à terme** est un contrat standardisé négocié sur un marché organisé permettant de s'assurer ou de s'engager sur un prix pour une quantité déterminée d'un produit donné (le sous-jacent) à une date future. La particularité des contrats à terme est d'aboutir à la livraison physique de la matière première à l'échéance. Les gérants indiciaires ont donc pour mission de repositionner leur portefeuille avant l'échéance pour éviter la livraison physique. Cette procédure appelée *roll-over*, consiste à vendre le contrat que le gestionnaire détient en portefeuille et qui est en général le contrat d'échéance la plus proche pour acheter le contrat d'échéance suivante. Le *roll-over* s'effectue le dernier jour ou la dernière semaine du mois précédant le début de la période de livraison sur un critère de liquidité.

Afin de tenir compte de la procédure de *roll-over*, il est courant d'utiliser les prix des contrats à terme ajustés : au moment du changement de contrat on calcule le rapport entre le prix de l'ancien contrat à terme et celui du nouveau, et on utilise ce coefficient pour pondérer le cours du nouveau contrat à terme.

Dans ce travail, nous utilisons les prix intra-journaliers ajustés échantillonnés à 10 minutes de 7h20 à 13h50 sur la période de 1999 à 2005 des contrats à termes sur le taux de change des différentes devises par rapport au dollar américain (USD) : de la livre britannique (GBP), du yen japonais (JPY) et du franc suisse (CHF).

1.2 Faits stylisés

Nous passons en revue les principales propriétés des séries financières de prix et de rendement. Ce qui nous permettra au passage d'introduire un certain nombre de définitions essentielles.

²Il existe toujours l'écart entre le cours acheteur et le cours vendeur, appelé la *fourchette bid-ask*.

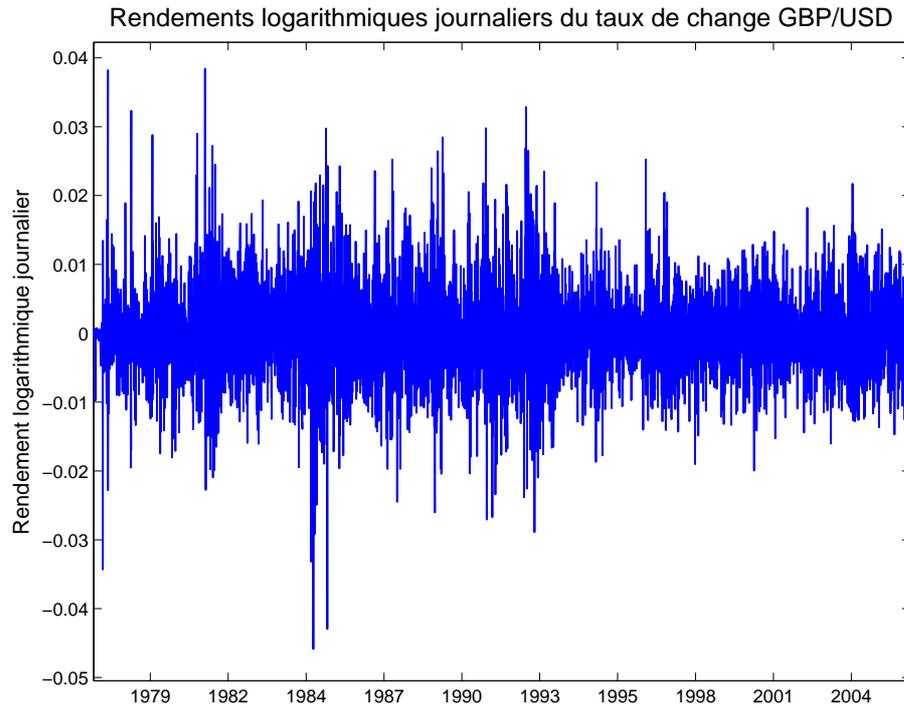


FIG. 1.2 – Rendements logarithmiques journaliers du taux de change GBP/USD. On peut noter les pics aigus et des cycles semblent apparaître.

1.2.1 Notations

Nous noterons $S(t)$ le cours en fonction du temps t d'un actif financier qui pourra être une action, un indice financier, un taux de change ou un contrat à terme. Nous définissons son logarithme $X(t) = \ln(S(t))$ et de même, nous définissons le rendement logarithmique du cours pour une échelle de temps τ donnée

$$\delta_\tau X(t) = X(t) - X(t - \tau) = \ln \left(\frac{S(t)}{S(t - \tau)} \right). \quad (1.1)$$

Notons qu'en pratique les modèles considèrent que les rendements logarithmiques des cours peuvent prendre un continuum des valeurs dans l'ensemble \mathbb{R} que nous supposons muni d'une probabilité \mathbb{P} , la probabilité historique, sous laquelle nous observons les cours. Ce n'est pas le cas dans les marchés financiers à cause de l'existence d'une valeur minimale de modification des prix. En effet, les prix auxquels les opérations sont réalisées (et les prix cotés) sont modifiés par une grandeur minimale, qui dépend du marché : le *tick*. Cependant, nous ne cherchons pas à modéliser cette discontinuité.

La figure 1.2 montre les rendements journaliers du taux de change GBP/USD.

1.2.2 Stationnarité et ergodicité des rendements logarithmiques

Nous devons souligner que la plupart des études statistiques de séries chronologiques sont basées sur les hypothèses de stationnarité et d'ergodicité. En termes simples, le problème est de savoir si les propriétés probabilistes des séries financières que nous étudions restent stables dans le temps et si chaque réalisation contient toute l'information statistique. Si ce n'est pas le cas, il est clair que les estimations statistiques par des moyennes temporelles ne peuvent pas être réellement utilisées.

Définition 1.1

Un processus stochastique $\{X(t)\}_t$ est dit **strictement stationnaire** si pour tout n , tous $t_1 < \dots < t_n$ et tout h on a l'égalité en loi suivante :

$$(X(t_1), \dots, X(t_n)) \stackrel{L}{=} (X(t_1 + h), \dots, X(t_n + h)) \quad (1.2)$$

Définition 1.2

Un processus stochastique $\{X(t)\}_t$ stationnaire est dit **strictement ergodique** si tous les moments généralisés empiriques calculés sur une réalisation convergent vers les moments théoriques, ainsi pour tout nombre entier n et tous décalages h_1, \dots, h_n , nous avons

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X(t + h_1) \cdots X(t + h_n) = \mathbb{E}[X(h_1) \cdots X(h_n)]. \quad (1.3)$$

Les notions de stationnarité et d'ergodicité au sens strict sont rarement nécessaires en pratique, il suffit généralement à requérir la **stationnarité** et l'**ergodicité au sens faible** du processus étudié.

Définition 1.3

Un processus stochastique $\{X(t)\}_t$ est dit **faiblement stationnaire** si :

- l'espérance $\mathbb{E}[X(t)]$ ne dépend pas de t ,
- les moments du second ordre sont finis et la fonction d'autocovariance $\text{Cov}[X(t), X(s)]$ ne dépend que de la différence $t - s$.

Cette dernière propriété implique, en particulier, que la variance de $X(t)$ est constante en fonction du temps.

Définition 1.4

Soit $\{X(t)\}_t$ un processus aléatoire stationnaire au sens faible. Si la moyenne temporelle converge en probabilité vers l'espérance

$$\lim_{T \rightarrow \infty} \bar{X}_T = \mathbb{E}[X], \quad \text{où } \bar{X}_T = \frac{1}{T} \sum_{t=1}^T X(t) \quad (1.4)$$

et la covariance temporelle converge en probabilité vers la covariance

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T (X(t) - \bar{X}_T)(X(t+h) - \bar{X}_T) = \text{Cov}[X(t), X(t+h)] \quad (1.5)$$

lorsque T tend vers l'infini, alors le processus $\{X(t)\}_t$ est **ergodique au sens faible**.

En fait, nous n'avons pas de raison de supposer que les propriétés statistiques des séries financières restent constantes au cours du temps. Néanmoins la plupart des tests statistiques effectués pour vérifier la stationnarité des rendements logarithmiques ne rejettent pas cette hypothèse (voir [Pag96] dans le cadre d'un modèle ARMA).

De plus, dans le cadre de ce travail nous nous intéressons à des séries financières sur des périodes de l'ordre 10-20 ans. Sur ces séries, plusieurs travaux [Lo91, CKW95, TTW99] mettent en évidence l'existence de corrélations temporelles à longue portée, corrélations qui ne permettent pas d'effectuer de façon efficace des tests statistiques de stationnarité et d'ergodicité. Nous étudions l'existence de corrélation temporelle à longue portée dans les sections 1.2.4 et 1.2.5.

Nous admettons les hypothèses de stationnarité et d'ergodicité, alors que la taille des séries financières est parfois trop petite pour pouvoir assurer une convergence des moments empiriques. Notons que, la consistance des résultats de prédiction des risques, qui sont présentés dans la partie III, justifie *a posteriori* l'hypothèse de stationnarité. En conséquence nous pouvons raisonnablement faire des statistiques par des moyennes temporelles sur nos données financières.

1.2.3 Distribution des rendements

Loi de probabilité et exposants de queues

La première modélisation des cours boursiers par un mouvement brownien, décrite dans la section 1.3.1, repose sur l'hypothèse que les rendements (logarithmiques) suivent une loi normale³. Cette modélisation reste encore aujourd'hui au cœur de la théorie financière.

En pratique, on observe des distributions de rendements logarithmiques ayant des queues plus épaisses que des distributions normales. Ce fait a été remarqué pour la première fois par Mandelbrot dans [Man63] pour les prix du coton et a été depuis lors observé pour différents marchés. Ce sera encore plus flagrant si l'on considère les rendements sur de petits intervalles de temps, ce qui n'est possible que grâce aux données à haute fréquence.

La décroissance d'une loi de distribution peut être décrite grâce à son exposant de queue.

Définition 1.5

On définit l'**exposant de queue** de la loi de distribution X par le nombre réel q_c , qui peut prendre la valeur infinie, tel que

$$q_c = \sup \left\{ q; \limsup_x \mathbb{P}[X > x] x^q < +\infty \right\}. \quad (1.6)$$

Le résultat classique suivant permet de relier l'existence des moments et l'exposant de queue de distribution.

³Egalement appelée loi gaussienne.

Proposition 1.1

Soit X une variable aléatoire positive. L'exposant de queue q_c vérifie la relation suivante

$$q_c = \sup_q \{q; \mathbb{E}[X^q] < +\infty\}. \quad (1.7)$$

On peut classer les distributions en trois types :

- les distributions à **queue légère**, $q_c = \infty$, où tous les moments de la distribution existent, et la fonction de répartition décroît exponentiellement dans les queues,
- les distributions à **queue lourde (ou épaisse)**, $q_c < \infty$, caractérisées par une fonction de répartition qui décroît comme une certaine loi de puissance,
- les distributions bornées, prenant des valeurs contenues dans un ensemble borné, et ne possédant donc pas de queue.

De récentes études (voir par exemple [GMAS98]) semblent indiquer que l'exposant de queue caractérisant le comportement en loi de puissance des queues de la distribution des rendements serait "universellement" dans l'intervalle $[3, 4]$ (cf. chapitre 4).

On a également observé le fait que les queues de la distribution des rendements deviennent de moins en moins épaisses lorsque l'échelle de temps considéré (pour le calcul des rendements) augmente, jusqu'à se rapprocher d'une loi normale pour de grands intervalles de temps.

La figure 1.3 représente les densités empiriques obtenues pour différents rendements logarithmiques, de haut en bas : journaliers, à 5 jours, à 10 jours et à 20 jours. En guise de comparaison, nous avons tracé sur le même graphique la courbe correspondant à la densité gaussienne standard. La déformation de la distribution empirique des rendements a été étudiée dans [GBP⁺96].

De façon générale, il existe beaucoup de modèles différents pour les distributions des rendements. Certains supposent que les actifs financiers ont des rendements assez proches de distributions stables, d'autres, de distributions de Student, qui ne font pas partie de la classe des distributions stables (cf. section 2.1.2). Il existe également la classe ARCH des modèles conditionnels hétéroscédastiques [Eng82, Bol86, BCK92] (cf. section 1.3.3), et beaucoup d'auteurs présument que ces modèles décrivent mieux les données que des modèles non conditionnels. Tout ces modèles s'accordent sur le caractère à queue épaisse des distributions des rendements. On est donc loin d'une modélisation par un processus gaussien [Man73a, Man73c, Man73b, Orl99, Wal03].

Skewness et kurtosis

Une manière de quantifier plus précisément les déviations par rapport à une loi normale est de calculer la *skewness* et la *kurtosis* des rendements logarithmiques.

Définition 1.6

Soit X une variable aléatoire. La *skewness*, définie comme le troisième moment centré

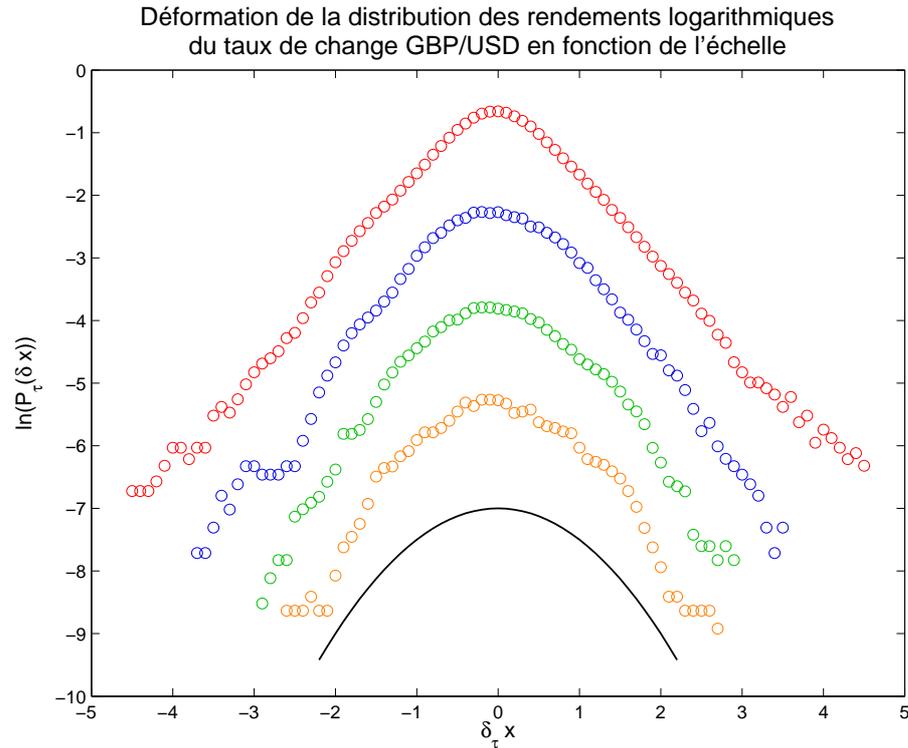


FIG. 1.3 – Déformation de la distribution empirique des rendements logarithmiques normalisés par leur écart-type du taux de change GBP/USD en fonction de l'échelle, de haut en bas : $\tau = 1$ jour, 5 jours, 10 jours et 20 jours. L'échelle est logarithmique et les courbes sont décalées arbitrairement pour une meilleure lisibilité. Le tracé montre clairement (voir, par exemple, le changement de convexité des queues) que les distributions qui correspondent aux petites échelles (en haut) possèdent un caractère leptokurtique, alors que celles qui correspondent aux grandes échelles (en bas) sont proches de la distribution gaussienne standard, représentée par le trait plein.

normalisé par le cube de l'écart-type,

$$\text{Skew}(X) = \frac{\mathbb{E}[X - \mathbb{E}[X]]^3}{\left(\mathbb{E}[X - \mathbb{E}[X]]^2\right)^{3/2}}, \quad (1.8)$$

permet de mesurer l'asymétrie de la distribution. La **kurtosis**, ou le quatrième moment centré normalisé par le carré de la variance,

$$\text{Kurt}(X) = \frac{\mathbb{E}[X - \mathbb{E}[X]]^4}{\left(\mathbb{E}[X - \mathbb{E}[X]]^2\right)^2}, \quad (1.9)$$

mesure la **leptokurticité**, ou non gaussianité, de la distribution.

Pour une variable gaussienne la skewness est nulle et la kurtosis égale à 3. Une variable aléatoire ayant un excès de kurtosis a une distribution plus large que la gaussienne et elle

	Moyenne	Variance	Skewness	Kurtosis	Test de JB
CAD/USD	0.0013	0.1054	-0.0065	5.7243	0.0000
JPY/USD	-0.0115	0.4624	-0.4802	6.9016	0.0000
CHF/USD	-0.0090	0.5680	-0.0192	5.6205	0.0000
GBP/USD	-0.0003	0.3856	0.1011	6.2693	0.0000

TAB. 1.1 – Forex. Moyenne, variance, skewness, kurtosis et p -valeur du test statistique de Jarque et Bera des rendements logarithmiques journaliers. Le test de normalité de Jarque et Bera est rejeté pour chacun des cours puisque la p -valeur est presque nulle.

est susceptible de donner naissance à des événements de plus grande amplitude.

Le test de Jarque et Bera [JB87] est un test de normalité fondé sur les propriétés de symétrie et d'aplatissement de la loi normale. La statistique du test est donnée par

$$JB(X) = \frac{n}{6} \left(\text{Skew}(X)^2 + \frac{(\text{Kurt}(X) - 3)^2}{4} \right), \quad (1.10)$$

où n est la taille d'échantillon. Sous l'hypothèse de normalité, la statistique de Jarque et Bera (1.10) est asymptotiquement distribuée selon la loi du $\chi^2(2)$. Notons que, la présence de mémoire longue dans les séries (cf. section 1.2.5), peut fausser ses résultats.

Dans les tableaux 1.1 et 1.2 nous donnons les statistiques : moyenne, variance, skewness, kurtosis et p -valeur du test statistique de Jarque et Bera, pour les rendements logarithmiques journaliers des données de Forex et CAC 40. Certaines composantes du CAC 40, notamment "Alcatel", "Saint Gobain" et "Vivendi Universal", présentent un très fort excès de kurtosis. Cependant, il faut noter que dans le chapitre 4, sur les événements extrêmes, nous mettons en évidence que la petite taille des séries financières ne permet pas d'assurer la convergence de moments empiriques d'ordre élevé. L'asymétrie des rendements est aussi visible.

Quantiles empiriques

Une autre façon de tester, visuellement, l'adéquation ou non des rendements avec une distribution donnée, par exemple gaussienne, est de faire une analyse des quantiles, en traçant les quantiles empiriques contre les quantiles théoriques d'une loi donnée. Si les distributions sont identiques, le tracé doit coïncider approximativement avec la droite $y = x$. Si au contraire elles sont différentes, on doit observer des déviations.

Dans la figure 1.4, nous avons soumis à cette analyse les rendements logarithmiques à plusieurs échelles. La déviation de la normalité, qui se traduit par un tracé non linéaire, est plus forte aux petites échelles. Les déviations diminuent lorsque l'échelle croît mais elles le font lentement et restent considérables même au bout de vingt jours.

	Moyenne	Variance	Skewness	Kurtosis	Test de JB
Accor	0.0164	4.1302	-0.1946	6.5537	0.0000
Air Liquide	0.0351	2.9136	0.0965	5.3988	0.0000
Alcatel	-0.0140	9.2631	-0.8909	28.2650	0.0000
Axa	0.0311	5.0735	0.0287	7.5109	0.0000
Bouygues	0.0426	5.2142	0.1934	7.2622	0.0000
Capgemini	-0.0173	8.7527	-0.0704	8.3784	0.0000
Carrefour	0.0504	3.4422	-0.0804	6.1774	0.0000
Casino Guichard	0.0356	3.5825	0.2289	5.4663	0.0000
Danone	0.0314	2.3257	-0.0291	6.6562	0.0000
Essilor International	0.0580	3.9690	0.0948	7.9286	0.0000
L'Oréal	0.0585	3.8911	-0.0182	5.0729	0.0000
Lafarge	0.0209	4.0134	-0.0058	4.8109	0.0000
Lagardere	0.0278	6.2287	0.1995	7.2883	0.0000
LVMH	0.0314	4.1532	0.2594	6.4667	0.0000
Michelin	0.0280	4.4332	0.0279	5.9599	0.0000
Pernod Ricard	0.0429	3.6197	0.0807	6.0166	0.0000
PSA Peugeot Citroën	0.0263	3.6814	-0.0151	6.5678	0.0000
Pinault Printemps	0.0489	4.5730	0.2025	8.4324	0.0000
Publicis	0.0398	5.9546	0.1334	6.4795	0.0000
Saint Gobain	0.0280	4.2504	-0.4348	12.0470	0.0000
Sanofi-Aventis	0.0583	4.1524	-0.0258	5.4426	0.0000
Schneider Electric	0.0176	5.1456	-0.2732	8.7748	0.0000
Société Générale	0.0445	4.5186	0.0363	6.8456	0.0000
Suez	0.0102	4.2901	-0.0742	9.9460	0.0000
TF1	0.0495	6.2216	0.1541	8.2373	0.0000
Thales	0.0197	5.7341	0.0075	5.5864	0.0000
Total	0.0622	3.2843	-0.1098	4.9802	0.0000
Vinci	0.0375	4.4338	0.2606	5.7971	0.0000
Vivendi Universal	0.0009	5.8363	-1.2667	26.1106	0.0000

TAB. 1.2 – CAC 40. Moyenne, variance, skewness, kurtosis et p -valeur du test statistique de Jarque et Bera des rendements logarithmiques journaliers. Le test de normalité de Jarque et Bera est toujours rejeté puisque la p -valeur est presque nulle.

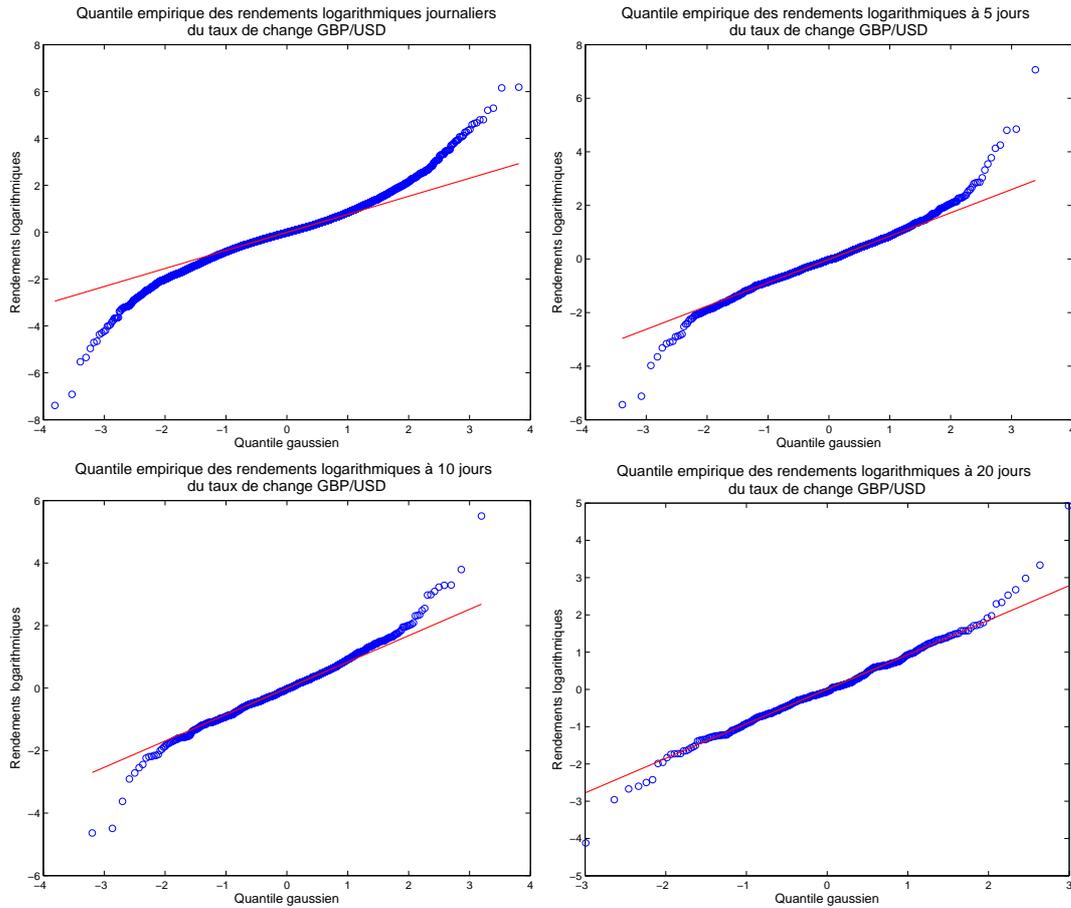


FIG. 1.4 – Quantiles empiriques des rendements logarithmiques normalisés par leur écart-type du taux de change GBP/USD, aux échelles journalière (en haut à gauche), à 5 jours (en haut à droite), à 10 jours (en bas à gauche) et à 20 jours (en bas à droite). Les tracés montrent clairement la convergence de la distribution vers la gaussienne, présentée par le trait plein, lorsque l'échelle croît.

1.2.4 Autocorrélation des rendements logarithmiques

Il est naturel de s'intéresser à la fonction d'autocorrélation de la série des rendements logarithmiques

$$\gamma(h) = \text{Corr}[\delta_\tau X(t), \delta_\tau X(t+h)] \quad (1.11)$$

Nous avons représenté la fonction d'autocorrélation empirique des rendements logarithmiques journaliers du taux de change GBP/USD sur la figure 1.5. Sur ce tracé, la fonction d'autocorrélation décroît extrêmement rapidement avec h et peut être significativement considérée comme nulle dès que h est différent de 0.

La structure institutionnelle du marché, dite **microstructure**, notamment les principes d'organisation du marché, tels que l'existence des ticks (cf. section 1.2.1), les modes de

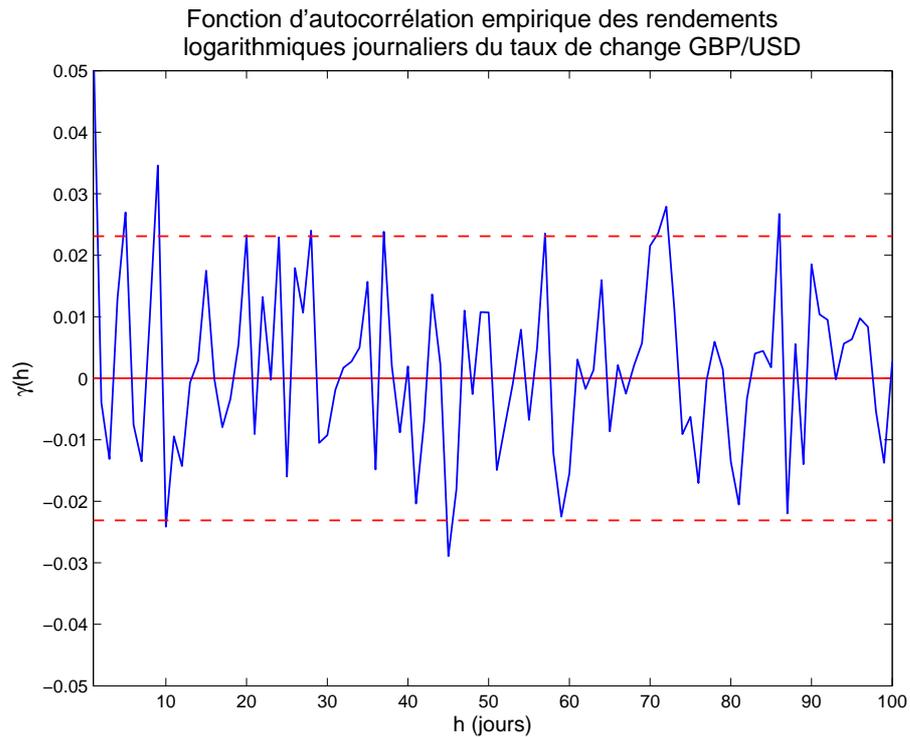


FIG. 1.5 – Fonction d'autocorrélation empirique des rendements logarithmiques journaliers du taux de change GBP/USD. Les traits tiretés représentent l'intervalle de confiance à 95%.

cotation, de transmission et d'exécution des ordres, peut avoir un impact sur les propriétés statistiques des actifs financiers.

Néanmoins, les études de la fonction d'autocorrélation sur les données à haute fréquence [Goo89, GF91, Eve95b, GO95], montrent qu'elle décroît exponentiellement vite avec le temps et peut être considérée comme nulle dès que h est supérieur à 5 minutes, temps au delà duquel les effets de microstructure [Lyo01] deviennent négligeables.

La justification la plus fréquente de cette absence de corrélation invoque le principe d'absence d'arbitrage des marchés liquides. Si de telles corrélations, faciles à détecter, existaient, il serait aisé pour un intervenant de concevoir une stratégie d'arbitrage qui en moyenne lui permettrait de gagner de l'argent. Or la mise en œuvre de ces stratégies devrait éliminer mécaniquement ces corrélations, au moins pour des temps supérieurs au temps de réaction du marché, qui typiquement est de l'ordre de quelques minutes, pour revenir à l'équilibre. Néanmoins, ce type d'arguments mérite un examen statistique plus approfondi car des subtilités existent [WB03].

1.2.5 Mémoire longue des marchés financiers

Une autre observation importante est celle de **longue mémoire**. Cette notion est apparue dans les années 1950 à partir des travaux en hydrologie de Hurst, voir la section 2.1.6, et depuis, cette notion a été appliquée dans divers autres domaines tels que le domaine de l'économie, de la météorologie, de la géophysique, de la finance. A titre d'exemple, la présence d'une composante de mémoire longue a été détectée dans des séries du taux de change [Che93, BO99, Vel99], dans des séries de prix d'actions en Bourse [CL95, CDFN95].

Si le processus des rendements logarithmiques semble effectivement non autocorrélé, ce qui est mis en évidence dans la section 1.2.4, il n'en est pas de même pour celui des rendements logarithmiques absolus (c-à-d. la valeur absolue des rendements logarithmiques) ou carrés dont la fonction d'autocorrélation semble faiblement décroître dans le temps.

Les études empiriques [MDD⁺95, CPB97, PS00] effectuées sur des données financières différentes conduit aux fonctions d'autocorrélation $\gamma_a(h)$ des rendements logarithmiques absolus significativement non nulle avec une décroissance en loi de puissance, dite **décroissance hyperbolique**,

$$\gamma_a(h) \underset{h \rightarrow \infty}{\sim} \frac{c}{h^\alpha}, \quad (1.12)$$

avec l'exposant α compris entre 0.2 et 0.4.

Une telle fonction a la propriété de ne pas présenter d'échelle caractéristique : les corrélations ne possèdent pas de durée typique au delà de laquelle elles décroîtraient rapidement. Une autre façon de le voir est qu'il est impossible de décrire correctement ces corrélations par une fonction exponentielle.

Nous aborderons le sujet de l'absence d'échelle caractéristique et de la propriété d'invariance d'échelle dans la section 2.2.1.

La faible valeur de l'exposant α se traduit par une décroissance des autocorrélations qui ne sont pas absolument sommable

$$\sum_{h=1}^{\infty} |\gamma_a(h)| = +\infty. \quad (1.13)$$

On parle dans ce cas de **mémoire longue** ou de **corrélations à longue portée** [Ber94]. Leur existence conduit à des problèmes délicats de statistique et la plupart des résultats classiques ne peuvent plus s'appliquer simplement. En particulier, le comportement d'une somme de variables ainsi corrélées n'obéit plus en général au théorème de la limite centrale classique [Ros61, DM79, Taq79, GS85, MBK06]. Dans le chapitre 4 nous étudions en détails le comportement des estimateurs classiques de l'exposant de queue de distributions.

Contrairement à la fonction d'autocorrélation des rendements logarithmiques, tracée sur la figure 1.5, la fonction d'autocorrélation des rendements logarithmiques absolus reste très longtemps significativement positive et indique donc l'existence d'une dépendance entre les différents rendements logarithmiques. Une régression "log-log" de cette fonction

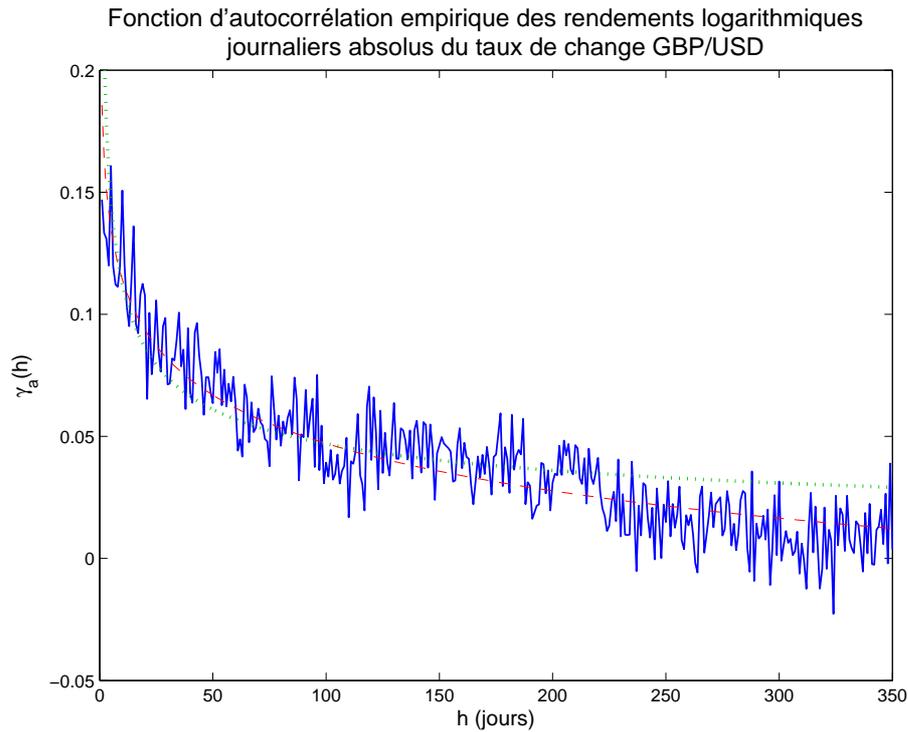


FIG. 1.6 – Fonction d'autocorrélation empirique des rendements logarithmiques journaliers absolus du taux de change GBP/USD. Le trait pointillé (vert) correspond à la courbe d'ajustement par une loi de puissance $\gamma_a(h) = \frac{0.27}{h^{0.38}}$ qui est en accord avec (1.12). Le trait tireté (rouge) correspond à la courbe d'ajustement $\gamma_a(h) = \left(\frac{550}{h}\right)^{0.027} - 1$ dont la forme provient du modèle MRW log-normal (cf. équation 3.139).

d'autocorrélation contre le temps indique qu'elle peut être bien approximée par une loi de puissance comme le montre le tracé sur la figure 1.6.

1.2.6 Volatilité réalisée

Une mesure de l'agitation du marché au cours du temps peut-être obtenue grâce à la **volatilité réalisée** définie comme suit.

Définition 1.7

Supposons donnés des rendements logarithmiques uniformément échantillonnés $\{X[t]\}_t = \{X(t\tau)\}_t$. La **volatilité réalisée** ou **volatilité historique** au temps t et à l'échelle τ est définie par

$$v[t] = \left[\sum_{j=1}^n |\delta_1 X[t - n + j]|^2 \right]^{1/2}, \quad (1.14)$$

où n est le nombre d'observations que l'on fait intervenir. En d'autres termes, $n\tau$ est la fenêtre de temps sur laquelle on se place.

Si les rendements ont une moyenne fort différente de zéro, c'est-à-dire en présence d'une tendance, on peut également centrer la volatilité réalisée en la définissant par

$$v_c[t] = \left[\frac{1}{n-1} \sum_{j=1}^n \left| \delta_1 X[t-n+j] - \frac{1}{n} \delta_n X[t] \right|^2 \right]^{1/2}, \text{ où } n > 1. \quad (1.15)$$

Dans un cas particulier, où n est égal à 1, les volatilités réalisées sont données par les rendements logarithmiques absolus.

En pratique, les séries des rendements logarithmiques ne sont pas uniformément échantillonnées, et le carré de la volatilité réalisée $v^2[t, t']$ correspondant à l'intervalle de temps $[t, t']$ est calculée en sommant les carrés des rendements logarithmiques

$$v^2[t, t'] = \sum_{k=1}^m (X[s_k] - X[s_{k-1}])^2, \text{ avec } t = s_0 < \dots < s_m = t', \quad (1.16)$$

où s_0, \dots, s_m est la grille d'échantillonnage.

Par exemple, dans le cas de données à haute fréquence la volatilité réalisée journalière est souvent calculée par une somme de rendements logarithmiques intra-journaliers au carré.

1.2.7 Accumulation des volatilités

Une propriété importante observée universellement sur les séries financières est le phénomène d'**accumulation des volatilités** (*volatility clustering*) : le fait qu'il semble y avoir des périodes de forte volatilité et d'autres de volatilité plus faible. Ce regroupement des volatilités par paquets est entièrement dû aux corrélations des séries financières. À cause de celles-ci, un grand mouvement, correspondant à une forte volatilité, a des chances d'être suivi par un mouvement de même ampleur (mais pas nécessairement dans le même sens), de même pour un mouvement de faible amplitude. Par ailleurs, à cause de l'absence d'échelle caractéristique dans les corrélations, ces périodes de forte ou faible volatilité n'ont pas non plus de longueur typique.

Sur la figure 1.7 nous avons représenté la volatilité réalisée des rendements logarithmiques journaliers du taux de change GBP/USD. Le comportement intermittent des rendements logarithmiques est très visible, on voit clairement que les fortes variations se rassemblent en paquets.

1.2.8 Corrélation entre le rendement et la volatilité instantanée

Les asymétries dans la dynamique de la volatilité sont souvent référencées dans la littérature sous le vocable **effet de levier** depuis que Black [Bla76] avait noté que les rendements sont négativement corrélés avec les variations de leur volatilité, au sens où la volatilité a tendance à augmenter en réponse à des "mauvaises nouvelles" (rendements nets moins

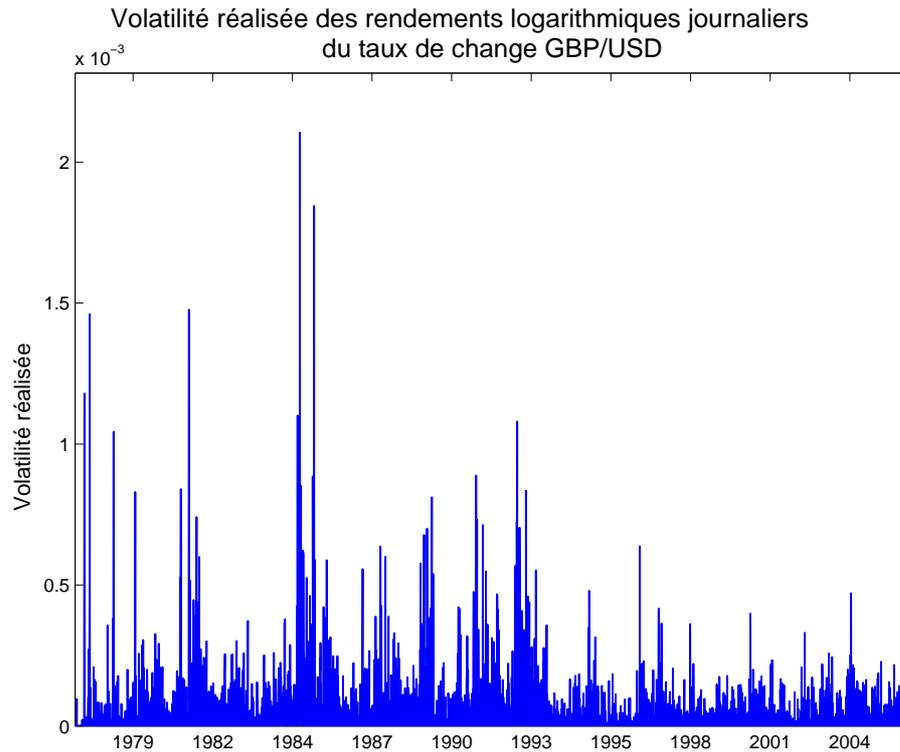


FIG. 1.7 – Volatilité réalisée obtenue par les carrés des rendements logarithmiques journaliers du taux de change GBP/USD. On observe le regroupement des volatilités par paquets.

élevés que prévu) et à chuter en réponse à des bonnes nouvelles (rendements nets plus élevés que prévu). Le levier financier peut en effet expliquer cette corrélation négative entre rendement d'aujourd'hui et volatilité de demain mais quel que soit le signe de cette corrélation, celle-ci introduira un certain type d'asymétrie dans la distribution observée. L'existence d'un effet de levier introduit souvent un effet de skewness dans la distribution des rendements.

Par ailleurs, autant les corrélations de la volatilité sont largement documentées, autant les études empiriques qui se sont intéressées aux corrélations entre le rendement et la volatilité sont peu nombreuses. Récemment, Bouchaud et al. [BMP01] se sont penchés sur ce problème. Ils ont étudié la **fonction de levier**

$$\mathcal{L}(t, s) = \frac{\mathbb{E}[\delta_\tau X(t)(\delta_\tau X(s))^2]}{\mathbb{E}[(\delta_\tau X(t))^2]} \quad (1.17)$$

pour plusieurs actions et indices.

Empiriquement, cette fonction, pour les actions, est nettement négative, dans le cas $t < s$, et peut être bien reproduite par une fonction exponentielle. Tandis que la même fonction calculée entre volatilités passées et rendements futurs, ce qui correspond au cas $t > s$, est

à peu près nulle. Les auteurs présentent ensuite un modèle qui permettrait d'expliquer le phénomène. Selon eux, cette asymétrie n'aurait pas d'origine économique mais apparaîtrait mécaniquement, par le simple retard des agents à prendre en compte les niveaux de prix.

Pour les indices, la situation est différente. L'effet de levier semble à la fois plus important mais décroît aussi beaucoup plus vite. Dans l'article [BMP01], les auteurs pensent que des effets de panique sur les marchés peuvent toucher les indices financiers, à propos desquels l'information est la plus vite transmise, et être à l'origine de cette asymétrie plus forte.

1.2.9 Conclusion

Toutes ces observations ont conduit à des modèles d'évolution des cours plus généraux que ceux comportant un mouvement brownien ou une semi-martingale, par exemple, les modèles de type ARCH ou ceux contenant des processus multifractals. Ce genre de modèle semble alors mieux tenir compte des propriétés statistiques précitées, mais nombreux problèmes deviennent trop difficile à résoudre théoriquement ou deviennent mal posé. Par exemple, jusqu'à aujourd'hui, on ne sait pas définir une intégrale stochastique satisfaisante par rapport à ce type de processus ou évaluer le prix de l'option européenne standard. Il existe donc de ce côté grand nombre de questions ouvertes en probabilité, mais nous ne les aborderons pas dans ce travail.

1.3 Modèles classiques

Dans cette section nous effectuons une brève revue des modèles d'évolution des prix des actifs financiers. Pour chaque modèle nous présentons les faits stylisés, étudiés dans la section 1.2, qui sont reproduits et nous expliquons ses principales faiblesses.

1.3.1 Modèles stochastiques de Bachelier et de Samuelson

Mouvement brownien

Historiquement, le **mouvement brownien** est associé à l'analyse de mouvements qui évoluent au cours du temps de manière si désordonnée qu'il semble difficile de prévoir leur évolution, même dans un intervalle de temps très court. Il joue un rôle central dans la théorie des processus aléatoires, parce que dans de nombreux problèmes théoriques ou appliqués, le mouvement brownien ou les diffusions que l'on en déduit fournissent des modèles limites simples sur lesquels de nombreux calculs peuvent être faits.

Le mouvement brownien a été nommé en hommage au botaniste écossais Robert Brown. Au cours de ses études en 1827 [Bro28], il avait observé au microscope des grains de pollen en suspension dans de l'eau. Ces grains de pollen étaient animés de mouvements erratiques incessants.

En 1900, Louis Bachelier [Bac00] introduit le mouvement brownien pour modéliser la dynamique des prix des actions à la Bourse. En 1905, Albert Einstein [Ein05, Ein56] construit un modèle probabiliste de mouvement brownien.

Définition 1.8

Un **mouvement brownien (standard)** est un processus aléatoire réel $\{B(t); t \in \mathbb{R}^+\}$ à trajectoires continues, tel que

- $B(0) = 0$,
- Tout accroissement $B(t) - B(s)$ suit une loi gaussienne centrée de variance $|t - s|$,
- Pour tous $t_1 < \dots < t_n$, les accroissements $\{B(t_{i+1}) - B(t_i)\}$ sont indépendants.

Notons que cette définition suffit à garantir que les accroissements du mouvement brownien sont stationnaires. Le mouvement brownien possède la propriété d'invariance suivante.

Proposition 1.2

Si $\{B(t); t \in \mathbb{R}^+\}$ est un mouvement brownien, pour tout $c > 0$, il en est de même du processus $\{B^c(t); t \in \mathbb{R}^+\}$ défini comme $B^c(t) = c^{-1}B(c^2t)$.

Preuve - 1.2 -

Par construction, B^c est un processus à trajectoires continues dont les accroissements $\{B^c(t) - B^c(s)\}$ sont indépendants de variance $|t - s|$, ce qui conduit au résultat. ■

Modélisation de cours boursiers par le mouvement brownien

Afin de prendre en compte le comportement très erratique des cours des actifs financiers, Bachelier les modélise à l'aide d'un mouvement brownien avec tendance

$$S(t) = S(0) + \mu t + \sigma B(t), \quad (1.18)$$

où μ est la tendance et σ est la volatilité de l'actif. Dans une telle modélisation la variable $S(t)$ peut prendre des valeurs négatives, ce qui n'a pas de sens financier.

En 1960, Samuelson [Sam65] propose de retenir cette modélisation pour les rendements, plutôt que pour les cours eux-mêmes

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dB(t) \quad (1.19)$$

Il est important de souligner que cette modélisation de Samuelson est à la base de la théorie de Black et Scholes [BS73]. A l'aide du lemme d'Itô on a la solution de cette équation différentielle stochastique

$$S(t) = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right) \quad (1.20)$$

Il est évident que dans ce cas les cours des actifs financiers restent toujours positifs.

Faits stylisés

Les modèles de Bachelier et de Samuelson permettent de reproduire l'absence de corrélation temporelle des rendements logarithmiques (cf. section 1.2.4), mais ils ne permettent pas de reproduire les queues épaisses et la déformation de la distribution des rendements logarithmiques en fonction de l'échelle (cf. section 1.2.3), la mémoire longue (cf. section 1.2.5), l'effet de levier (cf. section 1.2.8) et, en conséquence, le phénomène d'accumulation des volatilités (cf. section 1.2.7).

Grâce à leur simplicité, les modèles de Bachelier et de Samuelson sont devenus des outils incontournables en mathématiques financières, en particulier pour l'évaluation des options.

1.3.2 Modélisation de cours financiers par des processus de Lévy

Processus de Lévy

Mandelbrot dans ses travaux a montré que les rendements des actifs financiers à un jour, ou une semaine ne sont clairement pas statistiquement gaussiens, en particulier que la probabilité de grands mouvements de ces rendements est plus grande que celle que le monde gaussien quantifie (cf. section 1.2.3). En fait, la loi des variations des rendements des cours boursiers se rapproche plutôt des distributions dites de **Lévy α -stable** qui allie une forte densité de probabilité autour de zéro à la prise en compte des valeurs extrêmes, les queues de distribution très étirées [LVW02].

Définition 1.9

*Un processus en temps continu $\{X(t); t \geq 0\}$ est un **processus de Lévy** si ses incréments sont stationnaires et indépendants.*

Autrement dit, si $\{X(t); t \geq 0\}$ est un processus de Lévy, les accroissements $X(t_i) - X(s_i)$ sur des intervalles disjoints $(s_i, t_i]$ sont des variables aléatoires indépendants et leur loi ne dépend que de $t_i - s_i$.

Malgré cette simplicité apparente, les processus de Lévy ont des nombreuses propriétés intéressantes et constituent un domaine d'étude en plein développement : plusieurs ouvrages ont été publiés récemment [Ber96, Sat99] et une série de conférences internationales dédiées aux processus de Lévy et leurs applications a rencontré un grand succès [BNMR01]. Nous étudions plus en détail le mouvement de Lévy α -stable dans la Chapitre 2.1.

Modèle exponentielle-Lévy

On obtient le modèle exponentielle-Lévy de cours des actifs financiers en remplaçant dans le modèle de Samuelson (1.19) le mouvement brownien par un processus de Lévy.

Mandelbrot [Man63] a proposé un modèle pour les prix du coton selon un processus α -stable à variance infinie. Ce modèle a été fort critiqué à cause de la non existence des moments du second ordre du processus (cf. proposition 2.6). Voir également [Wal90, Wal95].

La tentative de résolution du problème d'observation de rendements non gaussiens, étudié dans la section 1.2.3, par la conservation de l'hypothèse d'indépendance des rendements et l'utilisation de distributions α -stables, ne permet pas de reproduire tous les faits stylisés. En particulier ces lois décroissent selon une loi de puissance avec un indice compris entre 0 et 2, tandis que des distributions empiriques des rendements décroissent selon une loi de puissance d'exposant plus élevé.

Plusieurs modèles, plus en accord avec ces observations, ont été proposés faisant intervenir distribution de Student [BG74], distributions hyperboliques [EKP98, Rai00, CGMY03], distribution normal inverse gaussien [BN98] ou distributions hyperboliques généralisées [Rai00, CGMY03]. Ces lois parviennent relativement bien à reproduire la distribution des rendements à une échelle de temps donnée mais elles sont en général incapables de les représenter simultanément pour chaque échelle de temps. De plus elles ne permettent pas de reproduire la présence de corrélation temporelle de volatilité.

Faits stylisés

Le modèle exponentielle-Lévy permet de reproduire l'absence de corrélation temporelle (cf. section 1.2.4) et les queues épaisses de la distribution des rendements logarithmiques (cf. section 1.2.3), mais il ne permet pas de reproduire la déformation de la distribution des rendements logarithmiques en fonction de l'échelle (cf. section 1.2.3), la mémoire longue (cf. section 1.2.5), l'effet de levier (cf. section 1.2.8) et le phénomène d'accumulation des volatilités (cf. section 1.2.7).

1.3.3 Modèles de type ARCH

Dans cette section nous considérons les modèles à temps discret.

Dans les années 70 s'est développée la classe des **modèles ARMA**, (*AutoRegressive Moving Average*). Ces modèles sont dits linéaires car sont fondés sur la représentation de la valeur présente de la série temporelle comme fonction linéaire de ses valeurs passées et de la valeur présente d'un certain bruit blanc.

Les **modèles ARCH** (*AutoRegressive Conditional Heteroskedasticity*) ont été introduits par Engle en 1982 [Eng82]. Dans ces modèles, ce sont les seconds moments conditionnels qui varient en fonction du temps suivant un certain modèle de régression. Cela correspond au fait que l'on s'est rendu compte que les moments du second ordre de séries temporelles issues des marchés spéculatifs dépendent du temps, et les chercheurs en finance ont donc commencé à modéliser cette dépendance [BCK92].

Puisque la variance conditionnelle d'un processus du type ARCH varie dans le temps, ces processus sont appelés **conditionnellement hétéroscédastiques**. Néanmoins, inconditionnellement, ces processus sont homoscedastiques. En d'autres termes, la variance non conditionnelle est identique à chaque instant du processus.

Pour un bon aperçu des modèles de type ARCH, on peut consulter le livre [Gou92]. Nous rappelons les notions d'homoscédasticité et hétéroscédasticité en appendice A.

Modèle ARCH

Dans la définition qui suit, nous supposons que la variance conditionnelle $\sigma[t]$ du processus $\{r[t]\}_t$ peut prendre des valeurs dans l'ensemble \mathbb{R}^+ que nous supposerons muni d'une probabilité \mathbb{P} et d'une filtration \mathcal{F}_t . Typiquement, cet ensemble d'information contient tous les éléments indicés par $t - 1$ et par les périodes précédentes.

L'absence de corrélation des rendements logarithmiques et la présence de structures de dépendance non linéaires conduit naturellement à une classe de modèles où les rendements logarithmiques $\{r[t]\}_t$ s'écrivent comme le produit de deux variables aléatoires

$$r[t] = \sigma[t]\epsilon[t]. \quad (1.21)$$

Le processus $\{\epsilon[t]\}_t$ est un bruit blanc, non obligatoirement gaussien, qui assure la décorrélation des rendements

$$\mathbb{E}[\epsilon[t]] = 0, \quad \text{Var}[\epsilon[t]] = 1, \quad (1.22)$$

Le processus de la variance conditionnelle $\{\sigma[t]\}_t$, indépendant de $\{\epsilon[t]\}_t$, contrôle la volatilité et, à travers lui, toute la complexité de la structure de dépendance. Le processus des rendements $\{r[t]\}_t$ est univarié dans la plupart des applications financières.

Définition 1.10

Le modèle linéaire **ARCH**(q), introduit par Engle [Eng82], suppose la paramétrisation suivante pour $\sigma^2[t]$,

$$\sigma^2[t] = \mathbb{E}[r^2[t]|\mathcal{F}_t] = \omega + \sum_{k=1}^q \alpha_k r^2[t-k], \quad (1.23)$$

où les paramètres satisfont $\omega > 0$ et $\alpha_j \geq 0$ pour tout j .

La version la plus simple de (1.23) est le processus **ARCH**(1),

$$\sigma^2[t] = \mathbb{E}[r^2[t]|\mathcal{F}_t] = \omega + \alpha_1 r^2[t-1]. \quad (1.24)$$

Il est simple de déduire la condition sur les paramètres ω et α_1 pour qu'un modèle ARCH(1) soit stationnaire.

Proposition 1.3

Si $\alpha_1 < 1$ alors le processus ARCH(1) est asymptotiquement stationnaire.

Par conséquent, la limite σ^2 de la variance non conditionnelle $\text{Var}[r[t]]$ lorsque t tend vers l'infini dépend des paramètres du processus ARCH(1) et, en général, elle est différente de la variance conditionnelle $\sigma^2[t]$.

Le moment du quatrième ordre est donné par

$$\lim_{t \rightarrow \infty} \mathbb{E}[r[t]^4] = \frac{3\omega^2(1 + \alpha_1)}{(1 - 3\alpha_1^2)(1 - \alpha_1)}. \quad (1.25)$$

La condition d'existence de ce moment est $3\alpha_1^2 < 1$. On en déduit l'expression de la kurtosis

$$\lim_{t \rightarrow \infty} \text{Kurt}(r[t]) = \lim_{t \rightarrow \infty} \frac{\mathbb{E}[r[t]^4]}{\mathbb{E}[r[t]^2]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3. \quad (1.26)$$

Elle est toujours supérieure à celle de la loi normale. Les queues de la distribution d'un processus ARCH(1) sont donc plus épaisses que pour un processus gaussien.

Modèle GARCH

Le modèle ARCH a parfois du mal à se calibrer sur les données réelles et le nombre de paramètres α_j nécessaires est souvent trop important. Plusieurs variantes du modèle ARCH ont été proposées. Une variante particulièrement populaire est le modèle ARCH généralisée, ou **GARCH** (*Generalized AutoRegressive Conditional Heteroskedasticity*), suggéré par Bollerslev [Bol86] et Taylor [Tay86] indépendamment.

Définition 1.11

Le modèle **GARCH**(p, q) peut s'écrire

$$\sigma^2[t] = \mathbb{E}[r^2[t] | \mathcal{F}_t] = \omega + \sum_{k=1}^p \alpha_k r^2[t - k] + \sum_{k=1}^q \beta_k \sigma^2[t - k], \quad (1.27)$$

Dans le modèle GARCH, la variance conditionnelle $\sigma^2[t]$ dépend aussi bien de ses propres valeurs passées que des valeurs retardées de $r^2[t]$. Cela signifie que $\sigma^2[t]$ dépend effectivement de toutes les valeurs passées de $r^2[t]$.

Le modèle GARCH avec très peu de paramètres permet de rendre compte des principaux faits stylisés souvent aussi bien qu'un modèle ARCH ayant de nombreux paramètres. En particulier, un modèle simple souvent utilisé est le modèle GARCH(1,1)

$$\sigma^2[t] = \mathbb{E}[r^2[t] | \mathcal{F}_t] = \omega + \alpha_1 r^2[t - 1] + \beta_1 \sigma^2[t - 1]. \quad (1.28)$$

En pratique, nous pouvons éliminer les termes de variance conditionnelle dans la partie droite de l'expression (1.28), par une procédure récursive, pour pouvoir estimer les paramètres du modèle GARCH(1,1). Le problème ressemble essentiellement à l'estimation d'un modèle à moyenne mobile ou d'un modèle ARMA avec une composante moyenne

mobile. Le modèle GARCH(1,1) peut être représentée récursivement, comme le modèle ARCH(∞), de la même manière qu'un processus AR(1) peut être vu comme un MA(∞), pour donner

$$\sigma^2[t] = \frac{\omega}{1 - \beta_1} + \alpha_1 \sum_{k=0}^{\infty} \beta_1^k r^2[t - 1 - k] \quad (1.29)$$

Il est possible de démontrer [Bol86] que sous la condition $\alpha_1 + \beta_1 < 1$ le processus GARCH(1,1) est asymptotiquement stationnaire au sens faible. De plus, la variance non conditionnelle est donnée par

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}. \quad (1.30)$$

Dans le cas d'un processus GARCH(1,1), il est possible de montrer que la kurtosis est de la forme

$$\text{Kurt}(r[t]) = \frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - \alpha_1^2(\mu_4 - 1)} \text{Kurt}(\epsilon[t]), \text{ où } \mu_4 = \mathbb{E}[\epsilon[t]^4]. \quad (1.31)$$

Modèle HARCH

Le modèle HARCH a été introduit dans [MDD⁺97].

Définition 1.12

Le modèle **HARCH**(n) peut s'écrire

$$\sigma^2[t] = \gamma_0 + \sum_{j=1}^n \gamma_j \left(\sum_{k=1}^j r[t - k] \right)^2, \quad (1.32)$$

où $\gamma_0 > 0$, $\gamma_n > 0$ et $\gamma_j \leq 0$ pour tout $j = 1, \dots, n - 1$.

La somme $\sum_{k=1}^j r[t - k]$ n'est rien d'autre que le rendement sur un intervalle de longueur j . On fait donc intervenir le carré des rendements sur des intervalles de longueurs différentes. Ce qui différencie ce modèle des modèles ARCH et GARCH.

Les auteurs de [ESDM96] étudient les conditions sous lesquelles ce processus est stationnaire et admet des moments finis.

Notons que ce modèle peut rendre compte de l'hypothèse d'hétérogénéité des marchés, décrite dans la section 2.3.1.

Estimation des paramètres du modèle GARCH(1,1)

Les modèles ARCH/GARCH sont classiquement estimés par la méthode de quasi-maximum de vraisemblance. Différentes hypothèses peuvent être retenues concernant des *alea* antérieurs à ceux de l'échantillon. La plus simple est de supposer qu'ils sont nuls, mais il est plus

	GARCH(1,1) gaussien				GARCH(1,1) Student				
	ω	α_1	β_1	$\alpha_1 + \beta_1$	ν	ω	α_1	β_1	$\alpha_1 + \beta_1$
CAD/USD	0.0020	0.135	0.845	0.980	6.686	0.0015	0.116	0.868	0.984
JPY/USD	0.0215	0.091	0.863	0.954	3.877	0.0114	0.104	0.888	0.992
CHF/USD	0.0160	0.090	0.888	0.978	5.805	0.0116	0.077	0.909	0.986
GBP/USD	0.0134	0.066	0.905	0.971	4.915	0.0021	0.072	0.928	1.000

TAB. 1.3 – Forex. Résultats de l'estimation de paramètres des modèles GARCH(1,1) gaussien (à gauche) et Student (à droite) par la méthode de quasi-maximum de vraisemblance pour les séries des rendements logarithmiques journaliers. On observe que la somme $\alpha_1 + \beta_1$ est toujours assez proche 1, qui correspond à la limite de stabilité.

réaliste de supposer qu'ils sont égaux à leur espérance non conditionnelle. La convergence des estimateurs de quasi-maximum de vraisemblance est démontrée dans [LH94].

En supposant que le bruit blanc $\{\epsilon[t]\}_t$ est gaussien la fonction de vraisemblance s'écrit

$$L(\omega, \alpha_1, \beta_1) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma[t]^2}} \exp\left(-\frac{y[t]^2}{2\sigma[t]^2}\right), \quad (1.33)$$

où la variance conditionnelle $\sigma[t]$ est donnée par la dynamique (1.28). Si le bruit blanc $\{\epsilon[t]\}_t$ suit une loi de Student à $\nu > 2$ degrés de liberté la fonction de vraisemblance prend la forme suivante

$$L(\nu, \omega, \alpha_1, \beta_1) = \prod_{t=1}^T \Gamma\left(\frac{\nu+1}{2}\right) \Gamma\left(\frac{\nu}{2}\right)^{-1} \frac{1}{\sqrt{(\nu-2)\sigma[t]^2}} \left(1 + \frac{y[t]^2}{(\nu-2)\sigma[t]^2}\right)^{-\frac{\nu+1}{2}}. \quad (1.34)$$

Notons que lorsque ν tend vers l'infini la distribution de Student (1.34) converge vers la distribution gaussienne (1.33).

Résultats de l'estimation des paramètres pour les données journalières

Dans les tableaux 1.3 et 1.4 nous représentons les résultats de l'estimation de paramètres des modèles GARCH(1,1) avec un bruit blanc (1.21) gaussien et t -Student par la méthode de quasi-maximum de vraisemblance pour les séries des rendements logarithmiques journaliers. Pour réaliser ces estimation nous avons utilisé toutes les données disponibles. Pour assurer la stabilité de la routine de minimisation les rendements journaliers ont été multipliés par le facteur 100. Par ailleurs, nous n'avons pas mis les intervalles de confiance pour ne pas charger les tableaux.

	GARCH(1,1) gaussien				ν	GARCH(1,1) Student			
	ω	α_1	β_1	$\alpha_1 + \beta_1$		ω	α_1	β_1	$\alpha_1 + \beta_1$
Accor	0.127	0.085	0.883	0.967	6.527	0.066	0.086	0.901	0.987
Air Liquide	0.036	0.068	0.920	0.988	8.721	0.030	0.070	0.921	0.991
Alcatel	0.128	0.150	0.850	1.000	5.161	0.067	0.114	0.886	1.000
Axa	0.080	0.097	0.886	0.984	7.795	0.057	0.095	0.895	0.990
Bouygues	0.053	0.072	0.919	0.991	5.491	0.048	0.080	0.915	0.995
Capgemini	0.153	0.086	0.899	0.985	5.235	0.173	0.121	0.870	0.991
Carrefour	0.043	0.061	0.926	0.987	7.372	0.037	0.061	0.929	0.990
Casino Guichard	0.070	0.050	0.930	0.981	5.159	0.093	0.079	0.900	0.979
Danone	0.050	0.072	0.907	0.979	6.344	0.037	0.074	0.912	0.986
Essilor	0.311	0.110	0.812	0.922	4.122	0.152	0.098	0.873	0.971
L'Oréal	0.072	0.092	0.891	0.983	13.689	0.061	0.086	0.900	0.986
Lafarge	0.043	0.049	0.940	0.989	8.215	0.035	0.051	0.941	0.992
Lagardere	0.166	0.092	0.885	0.977	5.063	0.067	0.112	0.888	1.000
LVMH	0.050	0.070	0.918	0.988	7.028	0.033	0.079	0.917	0.995
Michelin	0.061	0.060	0.927	0.987	6.589	0.057	0.059	0.929	0.988
Pernod Ricard	0.108	0.087	0.885	0.971	5.925	0.100	0.089	0.886	0.976
Peugeot Citroën	0.122	0.098	0.870	0.968	6.448	0.082	0.086	0.894	0.981
Pinault Printemps	0.071	0.057	0.927	0.985	4.152	0.034	0.075	0.925	1.000
Publicis	0.150	0.103	0.874	0.977	4.377	0.160	0.135	0.855	0.990
Saint Gobain	0.070	0.099	0.890	0.989	6.989	0.062	0.078	0.908	0.986
Sanofi-Aventis	0.083	0.085	0.897	0.982	6.716	0.075	0.087	0.899	0.986
Schneider Electric	0.101	0.098	0.887	0.984	6.469	0.131	0.111	0.869	0.980
Société Générale	0.106	0.109	0.870	0.979	5.903	0.077	0.107	0.881	0.989
Suez	0.189	0.089	0.857	0.946	5.792	0.108	0.093	0.879	0.972
TF1	0.103	0.092	0.891	0.984	4.613	0.093	0.116	0.881	0.997
Thales	0.069	0.067	0.922	0.989	6.871	0.062	0.074	0.918	0.993
Total	0.019	0.054	0.941	0.995	9.633	0.015	0.052	0.945	0.997
Vinci	0.005	0.018	0.981	0.999	4.719	0.016	0.046	0.954	1.000
Vivendi Universal	0.063	0.079	0.907	0.986	6.152	0.065	0.089	0.899	0.989

TAB. 1.4 – CAC 40. Résultats de l'estimation de paramètres des modèles GARCH(1,1) gaussien (à gauche) et Student (à droite) par la méthode de quasi-maximum de vraisemblance pour les séries des rendements logarithmiques journaliers. On observe que la somme $\alpha_1 + \beta_1$ est toujours assez proche de 1, qui correspond à la limite de stabilité.

Fait stylisés

Les modèles ARCH, GARCH et HARCH ont joué (et jouent encore) un rôle important dans la modélisation des séries financières. Ces modèles peuvent reproduire l'absence de corrélation temporelle (cf. section 1.2.4). Il est possible de montrer en utilisant des résultats des [Kes73, Gol91] que les queues de distributions suivent une loi de Pareto ce qui permet de reproduire la non gaussianité observée des rendements logarithmiques (cf. section 1.2.3). Ces modèles peuvent reproduire le phénomène d'accumulation des volatilités des données financières, décrit dans la section 1.2.7.

Cependant, les processus de type ARCH sont mal adaptés à la modélisation de longues séries. En particulier, ils ne permettent de reproduire que partiellement la longue dépendance observée dans les séries financières. En effet, ils présentent des corrélations à courte portée, qui décroissent exponentiellement avec le temps et sont donc incapables de reproduire le phénomène de mémoire longue (cf. section 1.2.5). Par ailleurs, ces modèles ne permettent pas de reproduire l'effet de levier (cf. section 1.2.8).

On peut voir dans les tableaux 1.3 et 1.4 que la somme des paramètres estimés $\alpha_1 + \beta_1$ est toujours assez proche de 1. Cela se traduit par une grande sensibilité des prédictions de volatilité et de Valeur à Risque (cf. partie III) à une variation des paramètres.

Lorsque l'on utilise des séries chronologiques sur de longues périodes les chocs de volatilité semblent être fortement persistants [FSS87, Akg89], et la somme $\alpha_1 + \beta_1$ du modèle est proche de 1. Si cette somme est égale à 1, alors le modèle GARCH est appelé IGARCH [EB86], qui est à la frontière de la non stationnarité. En effet, un choc sur la variance conditionnelle a un effet qui ne s'élimine pas asymptotiquement [Bol86], de même la variance non conditionnelle n'existe pas.

Par ailleurs, sur de courtes périodes les paramètres du modèle GARCH semblent être sensiblement plus faible [LL90].

De même, les modèle de type ARCH ne rendent pas compte des comportements des queues de distribution en fonction de l'échelle (cf. section 1.2.3). Pour tenter de résoudre ces imperfections, de nombreuses généralisations ont été proposées : FIGARCH [BBM96], EGARCH [Nel91], LM-ARCH [DG96]. L'inconvénient de la majorité de ces extensions est que dans la plupart des cas les processus construits ne sont pas stationnaires, ce qui traduit inévitablement, en pratique, par des difficultés numériques.

Enfin, rappelons que ces modèles sont à temps discret et ne sont pas stables par agrégation temporelle [DN93]. Cela limite l'utilisation de ces modèles en pratique.

1.3.4 Modèles à volatilité stochastique

Dans le modèle de Samuelson-Black-Scholes, les rendements logarithmiques vérifient

$$r[t] = \mu + \sigma \epsilon[t], \quad (1.35)$$

avec μ et σ constants et $\epsilon[t]$ indépendantes et identiquement distribuées de loi normale standard. On peut généraliser ceci en supposant que σ n'est plus constante mais est lui-même un processus stochastique. On parle alors de modèle à **volatilité stochastique**.

Comme dans le cas des modèles de type ARCH nous considérons la représentation du processus des rendements par le produit de bruit blanc $\epsilon[t]$ et de processus de volatilité $\sigma[t]$. Le plus souvent la volatilité s'écrit sous une forme exponentielle

$$\sigma^2[t] = \sigma^2 \exp(h[t]) \quad (1.36)$$

Le modèle le plus simple [Tay86, Tay94, GHR95] suppose une dynamique autorégressive pour le processus $h[t]$

$$h[t] = \alpha h[t-1] + \lambda \eta[t], \quad (1.37)$$

avec $\alpha \in [0, 1)$ afin d'assurer la stationnarité. Les bruits $\epsilon[t]$ et $\eta[t]$ sont, dans la plus simple formulation du modèle, des bruits blanc gaussiens, indépendants l'un de l'autre. La supposition que $\epsilon[t]$ suit une loi non gaussienne permet d'augmenter la leptokurticité des rendements. La relaxation de l'hypothèse d'indépendance des bruits blancs $\epsilon[t]$ et $\eta[t]$ permet introduire l'asymétrie dans le processus $r[t]$.

Même dans le cas le plus simple où les deux innovations sont gaussiennes et indépendantes, la loi de distribution des rendements possède des queues épaisses. Par exemple, la kurtosis peut être calculée explicitement

$$\text{Kurt}(r[t]) = 3 \exp\left(\frac{\lambda^2}{1 - \alpha^2}\right), \quad (1.38)$$

elle est toujours supérieur à 3.

Contrairement aux modèles du type ARCH, dans un modèle à volatilité stochastique il existe deux "source aléatoires" $\epsilon[t]$ et $\eta[t]$. En conséquence, ce modèle est largement plus difficile à estimer que les modèles ARCH, ce qui restreint, en pratique leur utilisation. Récemment de nouvelles méthodes d'estimation sont apparues qui les rendent plus compétitifs [FH98, SK98].

Les modèles à volatilité stochastique peuvent être construits à temps continu

$$\frac{dS_t}{S_t} = \mu dt + \sigma_t dB_t \quad (1.39)$$

$$d\sigma_t = \alpha(\sigma_t) dt + \beta(\sigma_t) d\tilde{B}_t \quad (1.40)$$

Plusieurs dynamiques ont été proposées pour la volatilité. Ainsi, le modèle le plus connu a été proposé par Heston [Hes93], il correspond à la limite du modèle GARCH (1.27),

$$d\sigma_t^2 = \kappa(\nu - \sigma_t^2) dt + \theta \sigma_t d\tilde{B}_t \quad (1.41)$$

Les mouvements browniens B_t et \tilde{B}_t peuvent être corrélés pour introduire de l'asymétrie. Comme leur analogue en temps discret, ces processus possèdent des queues épaisses, qui peuvent encore être alourdies si le processus B_t est remplacé par un processus de Lévy [BNS01].

Faits stylisés

Les modèles à volatilité stochastique permettent de reproduire des lois plus réalistes pour les rendements logarithmiques, les queues de ces distributions sont plus épaisses que celles de la loi normale, de même ils permettent d'introduire l'asymétrie dans les distributions (cf. section 1.2.3).

Ces modèles permettent de reproduire l'absence de corrélation temporelle (cf. section 1.2.4) et le phénomène d'accumulation des volatilités des données financières (cf. section 1.2.7). En relaxant l'hypothèse de l'indépendance des "sources aléatoires", il est ainsi possible de reproduire l'effet de levier, étudié dans la section 1.2.8.

La longue mémoire qui présente dans les cours financiers, (cf. section 1.2.5) reste le principal élément que cette modélisation ne permet pas d'écrire.

1.4 Conclusion

D'un point de vue statistique les séries financières sont donc des objets complexes, qui possèdent nombre de propriétés aujourd'hui bien connues et résumées sous le nom de "faits stylisés". Les rendements logarithmiques sont décorrélés (cf. section 1.2.4), sans être pour autant indépendants. Leur volatilité exhibe en effet des dépendances à longue portée (cf. section 1.2.5) qui rendent non triviales les propriétés d'agrégation temporelle des rendements logarithmiques. En particulier, ils sont distribués selon une loi asymétrique avec des queues épaisses qui persiste sur une large échelle de temps (cf. section 1.2.3).

Assez étrangement ces propriétés semblent vérifiées sur la plupart des marchés suffisamment liquides (actions, indices, change), ce qui laisse penser que des mécanismes sous-jacents communs pourraient régir leur fonctionnement. A l'interface de l'économie, des mathématiques se développent actuellement des recherches sur des modèles simples de marchés [Lux98, Hor01, KT02, GB03] visant à reproduire les caractéristiques observées empiriquement, avec le but ultime de mieux maîtriser le comportement d'un marché financier.

Les modélisations classiques : processus ARCH (cf. section 1.3.3), à volatilité stochastique (cf. section 1.3.4), de Lévy (cf. section 1.3.2) ne rendent compte que partiellement des propriétés que nous venons de décrire. C'est pourquoi nous présentons une autre classe de processus qui permettrait de combler ces lacunes, les processus multifractals.

Multifractalité

Le terme de **fractale** a été proposé pour la première fois par Benoit Mandelbrot en 1974, sur la racine latine *fractus* qui signifie brisé, pour décrire la géométrie des objets pouvant être caractérisés par une dimension non entière [Man77, Man82]. Une des propriétés fondamentales de ces objets fractals est la présence d'une autosimilarité qui se résume par le fait qu'il existe des transformations impliquant des dilatations qui laissent l'objet invariant (éventuellement dans un sens statistique).

Les objets fractals ont l'originalité de présenter des détails à toutes les échelles de longueur, ce qui explique les difficultés rencontrées pour les caractériser. Un des grands mérites de Mandelbrot est d'avoir su reconnaître que de nombreuses structures ou dynamiques complexes multi-échelles observées dans de nombreux domaines des sciences fondamentales et appliquées, possédaient des propriétés d'invariance d'échelle.

Aujourd'hui, les idées résultant de cette notion de fractale sont utilisées dans bien des domaines scientifiques et techniques, au nombre desquelles nous pouvons citer : la physique, la chimie, la biologie, la médecine, la finance, la géographie, l'informatique, la synthèse d'images (graphisme informatique), ...

Des exemples mathématiques de fractales sont très nombreux : les ensembles de Julia et de Mandelbrot, l'ensemble de Cantor, le triangle et le tapis de Sierpinski ou le flocon de Koch. Les fractales peuvent être des fractales déterministes ou stochastiques. Elles apparaissent souvent dans l'étude des systèmes chaotiques.

Le succès des différents ouvrages de Mandelbrot a donné un élan expérimental à de nombreuses branches de la physique, de la chimie, de la biologie, de l'astronomie, de la science des matériaux, de la géologie [Man89], de la météorologie [MSL96], de l'étude de trafic Internet [LTWW94, RCRB99] ainsi qu'en mathématiques financières [Man88, Eve95a, MDB00]. Les riches propriétés des processus multifractals en font des candidats idéaux à la modélisation de phénomènes complexes.

On dit qu'un objet ou un phénomène présente une invariance d'échelle lorsqu'aucune échelle (de temps ou d'espace) n'est privilégiée. Un objet **monofractal** est un objet

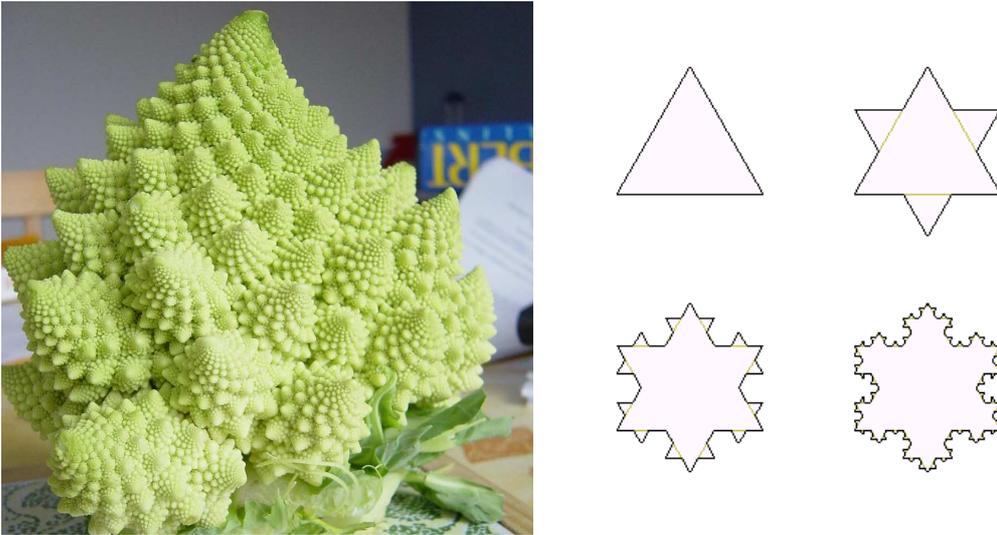


FIG. 2.1 – Chou romanesco (à gauche), sa géométrie autosimilaire fait qu'il est souvent cité comme un exemple de fractale naturel. Suite qui tend vers la flocon de Koch (à droite), c'est un exemple de fractale numérique.

invariant par des transformations géométriques de dilatation : sans plus d'information sur l'objet, on est incapable de distinguer l'objet lui-même de l'un de ses détails, convenablement dilaté, comme des flocons de neige, naturels ou mathématiques, ou les trajectoires du mouvement brownien (dans un sens statistique). Un objet **multifractal** est plus complexe dans le sens où il est toujours invariant par dilatation, mais le facteur de dilatation nécessaire pour ne pouvoir distinguer le détail de l'objet entier dépend du détail observé.

En pratique, la propriété d'invariance d'échelle ne s'étend pas à l'infini, le plus souvent elle est vérifiée sur un domaine d'échelles fini mais suffisamment grand pour être observée. L'invariance d'échelle a acquis une place centrale en physique statistique dans l'étude des phénomènes critiques.

L'importance de l'invariance d'échelle en finance a été mise en évidence assez récemment.

Ce chapitre commence par une brève revue des processus autosimilaires (cf. section 2.1). En nous appuyant sur la notion d'invariance d'échelle (cf. section 2.2.1), nous exposons l'approche de l'analyse multifractal (cf. section 2.2) et nous présentons de manière simple et relativement intuitive les processus multifractals, qui peuvent être vue comme une généralisation des processus autosimilaires (cf. section 2.2.2). Nous présentons, enfin (cf. section 2.3), les **faits stylisés fractals**, c'est-à-dire les propriétés statistiques fractales observés dans la plupart des marchés financiers et décrivons brièvement, à la fin de ce chapitre (cf. section 2.4), des modèles multifractals d'évolution des cours financiers.

2.1 Processus autosimilaires

Dans cette section nous rappelons les propriétés des lois infiniment divisibles et, en particulier, des lois stables. Puis nous abordons les processus autosimilaires et, un cas particulier, les processus autosimilaires à accroissements stationnaires.

2.1.1 Lois infiniment divisibles

Lorsque nous faisons du traitement des séries chronologiques, il nous arrive fréquemment d'utiliser la **moyenne** et la **variance empiriques**.

Définition 2.1

Soit $\{X(t)\}_t$ un processus tel que la variance $\text{Var}[X(t)] = \sigma^2$ est indépendante du temps t . Considérons la série d'observations $\{X[t]\}_{t=1,\dots,T}$ du processus $\{X(t)\}_t$, alors la **variance empirique** est donnée par

$$\widehat{\sigma^2} = \frac{1}{T-1} \sum_{t=1}^T (X[t] - \bar{X}_T)^2, \quad (2.1)$$

où \bar{X}_T est la **moyenne empirique** définie par

$$\bar{X}_T = \frac{1}{T} \sum_{t=1}^T X[t]. \quad (2.2)$$

Comme le nombre d'observations T est fini, la variance empirique $\widehat{\sigma^2}$ est évidemment finie. Dans certains cas cependant, elle est très grande. La question se pose alors de savoir si la variance théorique σ^2 est finie ou infinie.

Dans le cadre des distributions à variance infinie sont souvent utilisées les lois α -stables, dont le moment d'ordre 2 est infini dès que α est strictement inférieur à 2, le cas de α égal à 2 correspond à la loi normale. Ces lois sont utilisées dans de nombreux domaines tels que la finance [BEK98], la télécommunication [BCT98], ...

Avant de définir les lois α -stables, nous allons introduire une famille de lois plus générale : les lois infiniment divisibles. C'est à partir de ces lois que sera précisée la forme de la fonction caractéristique des lois stables. De même, les lois infiniment divisibles sont à la base de la construction des mesures multifractales log-infiniment divisibles (cf. section 3.1).

Définition 2.2

Une variable aléatoire réelle a une distribution **infiniment divisible** si et seulement si pour tout n ils existent des variables aléatoires X_1, \dots, X_n indépendantes et de même loi telles que

$$X \stackrel{\mathcal{L}}{=} X_1 + \dots + X_n. \quad (2.3)$$

Les variables aléatoires X_i n'ont pas la même loi que X . En revanche, comme nous verrons dans les exemples suivants, elles appartiennent à la même famille de loi.

L'intérêt principal de telles lois réside dans la solution du problème suivant : Comment détermine-t-on toutes les distributions qui s'expriment comme limite d'une somme de n variables aléatoires réelles indépendantes et identiquement distribuées ? En effet, on a le théorème suivant.

Théorème 2.1

Une variable aléatoire réelle X est la limite d'une somme de n variables aléatoires réelles indépendantes et identiquement distribuées si et seulement si X est infiniment divisible.

Preuve - 2.1 -

La démonstration est détaillée dans [Shi84]. ■

Théorème 2.2 (Lévy-Khintchine)

Si X a une distribution infiniment divisible, alors sa fonction caractéristique s'écrit

$$\varphi_X(t) = \exp \left(i\mu t + \int_{-\infty}^{+\infty} \frac{e^{itx} - 1 - it \sin x}{x^2} \nu(dx) \right), \quad (2.4)$$

où μ est un réel et ν est une mesure qui attribue une masse finie à tout intervalle fini et telle que les deux intégrales suivantes

$$\int_x^{+\infty} \frac{\nu(dy)}{y^2} \quad \text{et} \quad \int_{-\infty}^{-x} \frac{\nu(dy)}{y^2} \quad (2.5)$$

sont convergentes pour tout $x > 0$

Preuve - 2.2 -

La démonstration est détaillée dans [Fel71]. ■

Proposition 2.3

Les accroissements d'un processus de Lévy, défini dans la section 1.3.2, sont infiniment divisibles, et réciproquement une loi infiniment divisible correspond à un processus de Lévy.

Preuve - 2.3 -

Soit $X(t) - X(s)$ un accroissement du processus de Lévy, alors pour tout n entier

$$X(t) - X(s) = (X(t) - X(t - \tau)) + \dots + (X(s + \tau) - X(s)), \quad \text{où } \tau = (t - s)/n. \quad (2.6)$$

Les accroissements $(X(t - k\tau) - X(t - (k+1)\tau))$ sont indépendants identiquement distribués, d'après la définition du processus de Lévy 1.9, et donc ils peuvent jouer le rôle des variables X_1, \dots, X_n dans la définition 2.2.

La démonstration de la deuxième partie de l'affirmation est détaillée dans [Bre92]. Elle repose sur le théorème de Kolmogorov pour les familles cohérentes de lois de probabilités. ■

Une des caractérisations des lois infiniment divisibles est que leur fonction caractéristique peut s'écrire comme puissance $n^{\text{ième}}$ d'une autre fonction caractéristique.

Exemple 1

Beaucoup de lois connues sont infiniment divisibles. Si une variable aléatoire X suit :

- la loi normale $\mathcal{N}(m, \sigma^2)$: sa fonction caractéristique s'écrit

$$\varphi_X(t) = \exp\left(imt - \frac{t^2\sigma^2}{2}\right) = \left[\exp\left(i\frac{m}{n}t - \frac{t^2\frac{\sigma^2}{n}}{2}\right)\right]^n \quad (2.7)$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi normale $\mathcal{N}\left(\frac{m}{n}, \frac{\sigma^2}{n}\right)$,

- la loi de Cauchy $\mathcal{C}(c)$: sa fonction caractéristique s'écrit

$$\varphi_X(t) = \exp(-c|t|) = \left[\exp\left(-\frac{c}{n}|t|\right)\right]^n \quad (2.8)$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi de Cauchy $\mathcal{C}\left(\frac{c}{n}\right)$,

- la loi de Poisson $\mathcal{P}(\lambda)$: sa fonction caractéristique s'écrit

$$\varphi_X(t) = \exp(\lambda(e^{it} - 1)) = \left[\exp\left(\frac{\lambda}{n}(e^{it} - 1)\right)\right]^n \quad (2.9)$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi de Poisson $\mathcal{P}\left(\frac{\lambda}{n}\right)$,

- la loi Gamma $\Gamma(r, \lambda)$: sa fonction caractéristique s'écrit

$$\varphi_X(t) = \left(1 - \frac{it}{\lambda}\right)^{-r} = \left[\left(1 - \frac{it}{\lambda}\right)^{-\frac{r}{n}}\right]^n \quad (2.10)$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi Gamma $\Gamma\left(\frac{r}{n}, \lambda\right)$,

- Il en va de même pour la loi exponentielle, égale à $\Gamma(1, \lambda)$, et la loi du $\chi^2(m)$, égale à $\Gamma\left(\frac{m}{2}, \frac{1}{2}\right)$,
- une loi à support borné n'est pas infiniment divisible,
- un mélange fini¹ de lois normales n'est pas infiniment divisible.

2.1.2 Lois stables

Nous allons introduire une famille de **lois stables** [Lév37, TS94, Tsa97].

¹Une loi dont la distribution est donnée par la somme des densités des composantes du mélange.

Définition 2.3

Une variable aléatoire réelle a une distribution **stable** si et seulement si pour tout k et toute famille X_1, \dots, X_k indépendantes et identiquement distribuées de même loi que X , il existe $a_k > 0$ et b_k , deux réels, tels que

$$X_1 + \dots + X_k \stackrel{\mathcal{L}}{=} a_k X + b_k. \quad (2.11)$$

Lorsque $b_k = 0$, on parle de distribution **strictement stable**.

On peut montrer qu'il existe une constante α , $0 < \alpha \leq 2$, telle que $a_k = k^{1/\alpha}$ pour $k \in \mathbb{N}$. La démonstration est détaillée dans [Fel71].

Proposition 2.4

Si X est stable, X est infiniment divisible.

Preuve - 2.4 -

Il suffit de prendre des variables aléatoires $Y_j = \frac{X_j - \frac{b_k}{k}}{a_k}$, pour $j = 1, \dots, k$. Comme les X_j sont indépendantes, les Y_j sont aussi indépendantes et

$$Y_1 + \dots + Y_k = \frac{X_1 + \dots + X_k}{a_k} - \frac{b_k}{a_k}. \quad (2.12)$$

Or $X_1 + \dots + X_k \stackrel{\mathcal{L}}{=} a_k X + b_k$, d'où $Y_1 + \dots + Y_k \stackrel{\mathcal{L}}{=} X$. ■

La réciproque est fautive, voir le contre exemple, ci-dessous, la loi de Poisson.

Théorème 2.5 (Lévy-Khintchine)

Si X a une distribution stable, alors sa fonction caractéristique s'écrit :

$$\varphi_X(t) = \begin{cases} \exp \{ i\mu t - \gamma |t| [1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \ln |t|] \}, & \text{si } \alpha = 1, \\ \exp \{ i\mu t - \gamma |t|^\alpha [1 - i\beta \operatorname{sign}(t) \tan \frac{\alpha\pi}{2}] \}, & \text{si } \alpha \neq 1 \end{cases} \quad (2.13)$$

Preuve - 2.5 -

La démonstration est détaillée dans [GK68]. ■

Une loi stable est définie par quatre paramètres :

- α : paramètre principal, $0 < \alpha \leq 2$. Il caractérise les queues de distribution. Plus α diminue, plus les queues sont épaisses. C'est pourquoi on parle de lois α -stables. Le cas de α égal à 2 correspond à la loi normale,
- μ : paramètre de position. Il caractérise la moyenne de loi (lorsque $\alpha > 1$),
- γ : paramètre de dispersion. Dans le cas gaussien c'est la moitié de la variance.
- β : paramètre de symétrie, $-1 \leq \beta \leq 1$. Attention : il ne faut pas confondre symétrie et symétrique par rapport à 0. Si $\beta = 0$, la loi est symétrique par rapport au paramètre μ . Si de plus, μ est égal à 0, la loi est dite symétrique α -stable, de fonction caractéristique $\varphi_X(t) = \exp(-\gamma|t|^\alpha)$.

Par convention, $S_\alpha(\mu, \beta, \gamma)$ représentera une loi stable de paramètres α , μ , β et γ . Il est assez courant dans la littérature de noter $\gamma = \sigma^\alpha$.

Exemple 2

Là aussi, certaines des lois connues appartiennent à cette classe

- la loi normale $\mathcal{N}(m, \sigma^2)$ est la loi $S_2\left(m, \beta, \frac{\sigma^2}{2}\right)$ et réciproquement la loi $S_2(\mu, \beta, \gamma)$ est la loi normale $\mathcal{N}(\mu, 2\gamma)$,
- la loi de Cauchy généralisée de densité $f(x) = \frac{1}{\pi} \frac{\sigma}{\sigma^2 + (x-m)^2}$ est la loi $S_1(m, 0, \sigma)$,

La loi de Poisson n'est pas stable. En effet, soient X_1 et X_2 deux variables aléatoires suivant une loi de Poisson. Supposons que X_1 et X_2 soient stables, alors il existe $\alpha > 0$ et b , tels que

$$X_1 + X_2 \stackrel{\mathcal{L}}{=} aX_1 + b \quad (2.14)$$

Par égalité des moyennes et des variances, nous pouvons voir que

$$\begin{cases} 2\lambda = a\lambda + b \\ 2\lambda = a^2\lambda \end{cases} \implies \begin{cases} b = (2 - \sqrt{2})\lambda \\ a = \sqrt{2} \end{cases}, \quad (2.15)$$

ce qui entraîne une contradiction car $X_1 + X_2$ a ses valeurs uniquement dans \mathbb{N} alors que $\sqrt{2}X_1 + (2 - \sqrt{2})\lambda$ n'a pas des valeurs dans \mathbb{N} .

Proposition 2.6

Si X suit une loi $S_\alpha(\mu, \beta, \gamma)$, alors

1. Si $\alpha = 2$, X suit une loi normale et donc pour tout p , $\mathbb{E}[|X|^p] < +\infty$,
2. Si $0 < \alpha < 2$, $\begin{cases} \mathbb{E}[|X|^p] < +\infty, & \text{pour } 0 \leq p < \alpha, \\ \mathbb{E}[|X|^p] = +\infty, & \text{pour } p \geq \alpha. \end{cases}$

Dès que α est strictement inférieur à 2, la variance d'une loi α -stable est infinie. Dès que α est strictement inférieur à 1, c'est la moyenne qui devient infinie. D'après la proposition 1.1 l'exposant de queue q_c est égal à α .

2.1.3 Processus stables

Définition 2.4

Un vecteur aléatoire $\underline{X} = (X_1, \dots, X_n)$ est α -**stable** si pour tout k et pour toute famille $\underline{X}^{(1)}, \dots, \underline{X}^{(k)}$ de même loi que \underline{X} , il existe $a_k > 0$ et $B^{(k)}$, tels que

$$\underline{X}^{(1)} + \dots + \underline{X}^{(k)} \stackrel{\mathcal{L}}{=} a_k \underline{X} + B^{(k)} \quad (2.16)$$

Lorsque $B^{(k)}$ est le vecteur nul, on parle de loi **strictement stable**.

Une conséquence de la proposition suivante est que le paramètre a_k joue le même rôle que dans le cas univarié. Il existe une constante α , $0 < \alpha \leq 2$, telle que $a_k = k^{1/\alpha}$.

Proposition 2.7

Pour \underline{X} un vecteur α -stable, toute combinaison linéaire des composantes de \underline{X} est une variable aléatoire réelle α -stable.

Définition 2.5

Un processus $\{X(t); t \in T \subset \mathbb{R}\}$ est α -stable si pour tout $n \in \mathbb{N}$ et pour tout $t_1, \dots, t_n \in T$ le vecteur aléatoire $(X(t_1), \dots, X(t_n))$ est α -stable.

Les processus stables sont une généralisation des processus gaussiens. Ils ont joué un rôle important en modélisation en finance, voir la section 1.3.2.

2.1.4 Processus autosimilaires

Définition et exemples

Les processus **autosimilaires** sont des processus de distribution invariante par certains changements d'échelle de temps et d'espace. Ils ont été introduits et étudiés par Kolmogorov à l'occasion de ses études sur les fluides turbulents. Le terme "autosimilaire" est dû à Mandelbrot, et est devenu standard.

Définition 2.6

Un processus $\{X(t); t \in \mathbb{R}\}$ est un processus **autosimilaire** d'indice de Hurst² $H > 0$ si pour tout $a > 0$, on a

$$\{X(at)\}_t \stackrel{\mathcal{L}}{=} \{a^H X(t)\}_t. \quad (2.17)$$

Un théorème de Lamperti [Lam62] assure qu'un tel indice H , lorsqu'il existe, est unique. Une conséquence immédiate de la définition 2.6 est qu'un processus autosimilaire vérifie nécessairement la condition

$$X(0) = 0 \text{ p.s.} \quad (2.18)$$

Proposition 2.8

Un processus autosimilaire non dégénéré ne peut pas être stationnaire.

Preuve - 2.8 -

En effet, si le processus $\{X(t)\}_t$ était stationnaire, pour tout t_1, \dots, t_n et pour tout h nous aurions la suite d'égalités en loi

$$X(t) = X(at + (1-a)t) \stackrel{\mathcal{L}}{=} X(at) \stackrel{\mathcal{L}}{=} a^H X(t). \quad (2.19)$$

Nous arrivons à une contradiction lorsque a tend vers l'infini. ■

²Cette notion a été introduite par Hurst dans ses travaux en hydrologie [Hur65].

Exemple 3

- Un mouvement brownien standard est autosimilaire d'indice $H = 1/2$,
- Un mouvement de Lévy α -stable est autosimilaire d'indice $H = 1/\alpha$.

En conséquence, on peut dire que l'exposant H mesure la déviation d'un processus autosimilaire par rapport au mouvement brownien, qui correspond au cas de base d'un exposant égal à $1/2$.

Dans la définition de processus autosimilaire, on demande que les vecteurs aléatoires, de dimension finie, $(X(at_1), \dots, X(at_n))$ soient invariants en loi par changement d'échelle. Contrairement à l'autosimilarité déterministe, on ne demande pas que les réalisations, les trajectoires du processus se répètent lorsqu'on agrandit certaines parties. C'est "l'allure" générale des trajectoires qui sera invariante par changement d'échelle (c-à-d. leurs propriétés statistiques).

2.1.5 Processus autosimilaires à accroissements stationnaires

Parmi les processus autosimilaires, la classe la plus étudiée est celle des processus à accroissements stationnaires³. Pourtant, il n'existe pas de représentation générale pour cette classe de processus. Par conséquent, beaucoup de travaux portent sur des exemples particuliers comme le mouvement de Lévy α -stable et le mouvement brownien fractionnaire introduits dans la section 2.1.7 qui interviennent dans de nombreux domaines d'application.

Nous allons nous intéresser à l'existence des moments de ce genre de processus. Le résultat suivant montre que si l'on désire l'existence de certains moments, cela limite les valeurs admissibles pour H .

Proposition 2.9

Soit $\{X(t); t \in \mathbb{R}\}$ un processus autosimilaire à accroissements stationnaires, supposons que $\mathbb{P}[X(1) \neq 0] > 0$, alors la relation

$$\mathbb{E}[|X(1)|^\gamma] < \infty \tag{2.20}$$

implique

$$\begin{cases} H \in (0, 1/\gamma), & \text{si } 0 < \gamma < 1, \\ H \in (0, 1], & \text{si } \gamma \geq 1. \end{cases} \tag{2.21}$$

Preuve - 2.9 -

La preuve de cette proposition 2.9 est très technique et détaillée dans [TS94]. ■

³Rappelons que la définition des processus stationnaires est donnée dans la section 1.2.2.

Proposition 2.10

Soit $\{X(t); t \in \mathbb{R}\}$ un processus autosimilaire à accroissements stationnaires d'indice de Hurst H , et soit $t_1 < t_2 < t_3 < t_4$, alors la covariance $\text{Cov}[X(t_4) - X(t_3), X(t_2) - X(t_1)]$ est négative si $0 < H < \frac{1}{2}$, dans ce cas on parle de **persistance**, est nulle si $H = \frac{1}{2}$ et est positive si $\frac{1}{2} < H < 1$, dans ce cas on parle de **anti-persistance**.

Preuve - 2.10 -

On peut se ramener facilement au cas du lemme 2.11. ■

Lemme 2.11

Soit $\{X(t); t \in \mathbb{R}\}$ un processus autosimilaire à accroissements stationnaires d'indice de Hurst H , et soit $t_1 < t_2 < t_3$, alors la covariance $\text{Cov}[X(t_3) - X(t_2), X(t_2) - X(t_1)]$ est négative si $0 < H < \frac{1}{2}$, est nulle si $H = \frac{1}{2}$, est positive si $\frac{1}{2} < H < 1$.

Preuve - 2.11 -

En utilisant l'identité

$$\begin{aligned} 2\text{Cov}[X(t_3) - X(t_2), X(t_2) - X(t_1)] \\ = \text{Var}[X(t_3) - X(t_1)] - \text{Var}[X(t_3) - X(t_2)] - \text{Var}[X(t_2) - X(t_1)] \end{aligned} \quad (2.22)$$

et la définition du processus autosimilaire à accroissement stationnaires, nous avons

$$\begin{aligned} 2\text{Cov}[X(t_3) - X(t_2), X(t_2) - X(t_1)] \\ = ((t_3 - t_1)^{2H} - (t_3 - t_2)^{2H} - (t_2 - t_1)^{2H}) \text{Var}[X(1)] \\ = (t_3 - t_1)^{2H} \left(1 - \left(\frac{t_2 - t_1}{t_3 - t_1} \right)^{2H} - \left(1 - \frac{t_3 - t_2}{t_3 - t_1} \right)^{2H} \right) \text{Var}[X(1)]. \end{aligned} \quad (2.23)$$

Ce qui, d'après la propriété de convexité de la fonction $f(x) = 1 - x^{2H} - (1 - x)^{2H}$, où $0 < x < 1$, conduit au résultat annoncé. ■

2.1.6 Analyse des étendues normalisées

Plusieurs méthodes statistiques existent pour estimer l'indice de Hurst (2.17). En 1965, Hurst [Hur65] a proposé une méthode appelée **analyse des étendues normalisées**, ou *Rescaled Range Analysis (R/S Analysis)*. Notons que l'objectif initial de l'auteur était de modéliser la série temporelle de la hauteur des crues du Nil, de l'antiquité à nos jours.

La méthode consiste à estimer, pour un intervalle de taille donné, le rapport entre l'étendue R de la série centrée et intégrée, et l'écart-type S de la série originale. En d'autres termes, la série est considérée comme la "cause" générant le déplacement d'une particule (dont rend compte la série centrée et intégrée). On s'intéresse à la différence entre les limites

minimale et maximale du déplacement, dans un certain intervalle temporel, normalisé par l'écart-type de la cause ayant généré le déplacement.

La méthode préconisée par Hurst est construite sur la statistique $Q(n) = \frac{R(n)}{S(n)}$, dite **étendue normalisée**, où n désigne la taille de l'échantillon, elle permet de détecter des cycles non périodique. On obtient ainsi l'estimation de l'indice de Hurst H avec

$$R(n) = \max_{1 \leq k \leq n} \sum_{j=1}^k (X[j] - \bar{X}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (X[j] - \bar{X}_n), \text{ l'étendue,} \quad (2.24)$$

$$S^2(n) = n^{-1} \sum_{j=1}^n (X[j] - \bar{X}_n)^2, \text{ la variance empirique,} \quad (2.25)$$

$$\bar{X}_n = n^{-1} \sum_{j=1}^n X[j], \text{ la moyenne empirique.} \quad (2.26)$$

Hurst a montré que l'étendue normalisée $Q(n)$ est liée à la taille n de l'intervalle considéré par la relation suivante

$$Q(n) = an^H, \quad (2.27)$$

où H est le paramètre de Hurst, équivalent à l'indice d'autosimilarité (2.17), et a est une constante.

En pratique, cette méthode se fait en plusieurs étapes :

- Tout d'abord, on détermine une suite des entiers $(k_j)_{j=1, \dots, m}$ de longueur m , telle que $1 < k_m < \dots < k_1 = n$. En pratique, on choisit la longueur m telle que k_m est de l'ordre de 10. Souvent on utilise la suite définie par Davies et Harte [DH87] telle que

$$k_j = \begin{cases} \lfloor \frac{n}{j} \rfloor, & \text{pour } 1 \leq j \leq 6, \\ \lfloor \frac{k_{j-1}}{1.15^j} \rfloor, & \text{pour } 6 < j \leq m, \end{cases} \quad (2.28)$$

- Ensuite, pour chaque k_j , on considère une série de subdivisions indépendantes de taille k_j , et pour chaque subdivision, on détermine la statistique $Q(k_j)$,
- Puis, on procède à la pondération moyenne $\bar{Q}(k_j)$ des statistiques $Q(k_j)$ par nombre de subdivisions,
- Enfin, l'exposant H peut être estimé par la méthode des moindres carrés, ou par l'estimation de la pente de la régression de $\ln(\bar{Q}(k_j))$ sur $\ln(k_j)$.

L'analyse de Hurst est délicate pour les séries brèves. Des fluctuations importantes peuvent être obtenues avec des séries de moins de 500 données. Ceci explique pourquoi les subdivisions les plus faibles utilisées ne sont pas inférieures à 10. Enfin l'exposant de Hurst est sensible à la taille de l'échantillon [IFMAV98].

Le seul avantage de cette méthode c'est qu'elle permet d'obtenir un estimateur qui possède des bonnes propriétés de robustesse [MT79].

Sur la figure 2.2 nous avons tracé la dépendance de l'étendue normalisée $\bar{Q}(k)$ en fonction de k , aux échelles logarithmiques, pour la série des rendements logarithmiques journaliers du taux de change GBP/USD. Nous avons tracé également la droite d'ajustement qui permet d'estimer l'indice de Hurst pour cette série.

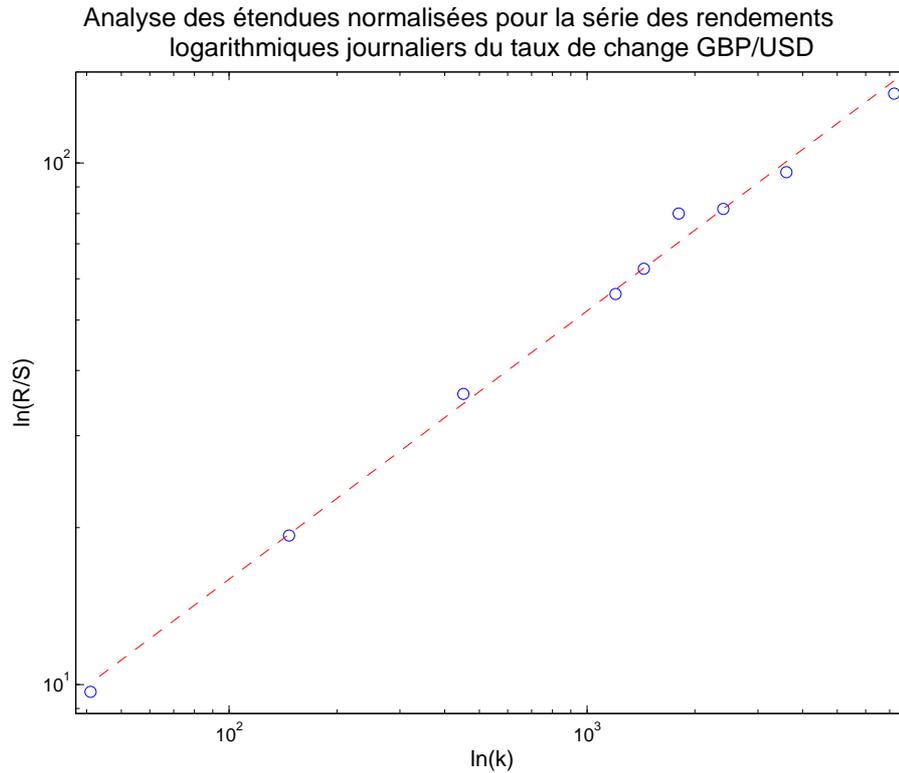


FIG. 2.2 – Analyse des étendues normalisées des rendements logarithmiques journaliers du taux de change GBP/USD. Le trait tireté correspond à l'ajustement $\bar{Q}(k) = 0.4k^{0.5142}$. Puisque l'exposant est proche de 0.5 les rendements logarithmiques sont décorrélés (voir la proposition 2.10).

2.1.7 Exemples

Mouvement de Lévy α -stable

Nous allons présenter l'exemple d'un processus stable autosimilaire d'indice de Hurst $H = 1/\alpha$ à accroissements stationnaires indépendants.

Définition 2.7

Un processus stochastique $\{X(t); t \geq 0\}$ est un **mouvement de Lévy α -stable standard** s'il vérifie les propriétés suivantes

- $X(0) = 0$,
- Tout accroissement $X(t) - X(s)$, suit une loi α -stable $S_\alpha(0, \beta, |t - s|)$, où $0 < \alpha \leq 2$ et $-1 \leq \beta \leq 1$,
- Pour tous $0 = t_0 < t_1 < \dots < t_n$, les accroissements $\{X(t_{i+1}) - X(t_i); 0 \leq i < n\}$ sont indépendants.

Pour $\alpha = 2$ le mouvement de Lévy α -stable correspond au mouvement brownien standard.

Mouvement brownien fractionnaire

L'exemple d'un processus autosimilaire gaussien à accroissements stationnaires, mais non nécessairement indépendants est un **mouvement brownien fractionnaire**. Il s'agit d'un type de processus très important au niveau des applications, surtout à cause de son caractère à longue mémoire.

Le mouvement brownien fractionnaire a été introduit par Kolmogorov pour l'étude de la turbulence [Kol40]. Cependant, l'étude et l'utilisation des processus gaussiens autosimilaires a réellement commencé avec l'article fondateur de Mandelbrot et Van Ness [MN68] sur les mouvements browniens fractionnaires. Ces processus sont utilisés en sciences de la terre (hydrologie, géophysique), en physique (turbulence [Kol41, Shi99]) mais aussi en économie dans différents modèles.

Définition 2.8

Un **mouvement brownien fractionnaire (standard)** de l'indice de Hurst H , où H est un nombre réel appartient à l'intervalle $(0, 1)$, est un processus $\{B^H(t); t \in \mathbb{R}^+\}$ gaussien centré dont la covariance est donnée par

$$\text{Cov}[B^H(s), B^H(t)] = \frac{1}{2} \left(s^{2H} + t^{2H} - |t - s|^{2H} \right). \quad (2.29)$$

Le résultat suivant [TS94] implique que tout processus gaussien autosimilaire d'indice $H < 1$ à accroissements stationnaires est un mouvement brownien fractionnaire, grâce à la caractérisation des processus gaussiens par leurs fonctions de moyenne et de covariance.

Proposition 2.12

Supposons que $\{B^H(t); t \in \mathbb{R}\}$ soit un processus autosimilaire d'indice H à accroissements stationnaires non dégénéré et de variance $\text{Var}[B^H(t)]$ finie, alors l'indice $H \in (0, 1]$, $B^H(0) = 0$ p.s. et

$$\text{Cov}[B^H(s), B^H(t)] = \frac{1}{2} \left(|s|^{2H} + |t|^{2H} - |t - s|^{2H} \right) \text{Var}[B^H(1)]. \quad (2.30)$$

De plus, pour tout $t \in \mathbb{R}$

$$\begin{cases} \mathbb{E}[B^H(t)] = 0, & \text{si } H < 1, \\ B^H(t) = tB^H(1) \text{ p.s.}, & \text{si } H = 1. \end{cases} \quad (2.31)$$

Le mouvement brownien fractionnaire présente des accroissements dépendants. Lorsque $1/2 < H < 1$ les corrélations sont positives et à longue portée, tandis que la condition $0 < H < 1/2$ conduit à des corrélations négatives. Pour $H = 1/2$, les accroissements sont indépendants. Dans le cas $1/2 < H < 1$, en raison des corrélations positives qui ont pour effet de favoriser une tendance haussière ou baissière (le processus est dit **persistant**).

Les processus gaussiens autosimilaires ou localement autosimilaires introduits simultanément dans [BJR97, PLV96] constituent des généralisations naturelles des mouvements

browniens fractionnaires. En particulier, on a introduit les processus aléatoires avec facteur d'autosimilarité variable au cours du temps, cette variabilité pouvant être au choix très régulière ou au contraire extrêmement brutale. Des estimateurs statistiques des paramètres de ces processus aléatoires sont à chaque fois construits et étudiés [BBCI98, BBCI99].

Lors de l'application du modèle du mouvement brownien fractionnaire à un cas concret, il est nécessaire d'estimer aussi précisément que possible le paramètre de Hurst H . Il existe de nombreuses méthodes qui permettent de réaliser cette tâche et il est difficile d'en choisir une *a priori*. Théoriquement, même si des résultats d'efficacité asymptotique sont démontrés pour certains estimateurs [Dah89, Pog95, JHJ01], rien ne permet de connaître leur comportement dans toutes les circonstances (signaux de longueur réduite, bruit, ...)

Parmi les processus à longue mémoire, les mouvements browniens fractionnaires ont beaucoup de succès à cause de leur autosimilarité qui en facilite l'analyse. Cependant, il a été souligné [Rog97, Che03] que le processus $\{B^H(t)\}_t$ n'est ni une semi-martingale ni un processus de Markov⁴, et qu'il existe des opportunités d'arbitrages, ce qui signifie que le mouvement brownien fractionnaire ne peut pas être un modèle financièrement cohérent pour les rendements logarithmiques.

2.1.8 Conclusion

Nous avons présenté deux types de processus autosimilaires : les mouvements de Lévy α -stable et les mouvements browniens fractionnaires, le mouvement brownien standard appartient simultanément aux deux classes. D'autres processus autosimilaires existent [PT02a, PT02b], par exemple les processus stables qui généralisent les mouvements de Lévy α -stables en autorisant des incréments dépendants. L'autosimilarité est une propriété mathématique très agréable car la connaissance du processus à une échelle suffit pour en déduire son comportement à toutes les échelles. En revanche, cette propriété entraîne des contraintes qui la rendent mal adaptée à la modélisation de phénomènes réels et complexes où les propriétés à grande échelle diffèrent notablement de celles à petite échelle. Les processus multifractals ont pour but de restituer cette richesse de comportement.

2.2 Analyse multifractale

L'analyse multifractale a été introduite par des physiciens [FP85, Fri95, HJK⁺86] au milieu des années 80. Le but était de proposer une interprétation de certains phénomènes observés dans les écoulements de fluides en régime turbulent. L'utilisation de processus multifractals s'est largement développée et s'est imposée dans de nombreux domaines : géophysique [Man89], climatologie [MSL96], traitement d'images [TP00], étude de trafic Internet [LTWW94, AV98, RCRB99], turbulence [CAP99, CLAB00], finance [Eve95a, Man88, MDB00]. Les riches propriétés de ces processus en font des candidats idéaux à

⁴Sauf dans le cas $H = 1/2$ où $B^H(t)$ se réduit au mouvement brownien standard.

la modélisation de multiples phénomènes complexes. Sur le plan plus théorique, ils ont également intéressé les mathématiciens et leur étude rigoureuse est un sujet particulièrement riche [Jaf97a, Jaf97b].

Mathématiquement, l'analyse multifractale étudie des processus et des mesures singulières à travers leurs comportements singuliers.

Dans cette section, nous présentons de manière simple et relativement intuitive les processus multifractals en nous appuyant sur la notion d'invariance d'échelle. La propriété d'autosimilarité au sens stochastique, vérifiée par les processus multifractals, nous permet, dans la suite, d'étudier les densités de probabilité de leurs accroissements en fonction de l'échelle. Puis, nous aborderons les principaux outils de l'analyse multifractale, tels que l'exposant de Hölder et le spectre de singularités, et le lien avec des processus multifractals, qui s'exprime par le moyen du formalisme multifractal.

2.2.1 Invariance d'échelle

Lois de puissance

On dit qu'un phénomène présente une **invariance d'échelle** lorsqu'aucune échelle n'est privilégiée, ce qui se traduit par l'**absence d'échelle caractéristique**. En d'autres termes, on ne peut pas identifier dans le système ou le signal étudié des échelles jouant un rôle spécifique.

Le paradigme de l'absence d'échelle caractéristique est la loi de puissance. En effet, si une quantité $f(\tau)$ qui décrit à l'échelle τ un phénomène présentant une invariance d'échelle, varie en loi de puissance

$$f(\tau) = C\tau^\zeta, \quad (2.32)$$

alors, le rapport des valeurs de $f(\tau_1)$ et $f(\tau_2)$ associées à deux échelles τ_1 et τ_2 ne dépend que du rapport de ces deux échelle τ_1/τ_2 :

$$\frac{f(\tau_1)}{f(\tau_2)} = \left(\frac{\tau_1}{\tau_2}\right)^\zeta. \quad (2.33)$$

En conséquence aucune échelle ne peut être privilégiée, comme, par exemple, dans le cas des corrélations à longue portée, étudié dans la section 1.2.5.

Fonction de structure, moments et processus monofractals et multifractals

Le plus souvent la quantité $f(\tau)$ correspond à un moment empirique d'ordre q à l'échelle τ . Dans le contexte de l'étude de la turbulence pleinement développée, en particulier de l'étude de la nature singulière du champ de vitesse d'un écoulement turbulent, Parisi et Frisch [PF85] ont proposé une description multifractale basée sur l'étude du comportement en lois de puissance des **fonctions de structure** [Fri95, ASM75].

Définition 2.9

Pour une série d'observations $\{x[t]; t = 0, 1, \dots\}$ à l'échelle τ et de taille d'observation L la **fonction de structure** est donnée par

$$S(\tau, q) = \frac{1}{\lfloor L/\tau \rfloor} \sum_{k=1}^{\lfloor L/\tau \rfloor} |x[k\tau] - x[k\tau - \tau]|^q. \quad (2.34)$$

Cette fonction évoquera les fonctions de partition rencontrées en physique statistique. Elle peut être aussi interprétée comme un moment empirique absolu d'ordre q des accroissements à l'échelle τ . Dans ce cadre, dire qu'un processus $\{X(t)\}_t$ à accroissements stationnaires est invariant d'échelle revient à dire que les moments absolus de $\delta_\tau X(t)$ suivent une loi de puissance

$$\mathbb{E}[|\delta_\tau X(t)|^q] \sim C(q)\tau^{\zeta(q)}. \quad (2.35)$$

Définition 2.10

Nous allons distinguer deux classes de processus selon la linéarité de l'exposant $\zeta(q)$, appelé **exposant multifractal** :

- si $\zeta(q)$ est une fonction linéaire, alors on parle de processus **monofractal**,
- si $\zeta(q)$ est une fonction non linéaire, alors on parle de processus **multifractal**.

Les processus autosimilaires à accroissements stationnaires, décrits dans la section 2.1.5, caractérisés par un indice de Hurst H , sont monofractals. En effet, en utilisant la stationnarité des accroissements et la définition de l'autosimilarité, nous trouvons que

$$\mathbb{E}[|\delta_\tau X(t)|^q] = \mathbb{E}[|\delta_\tau X(\tau)|^q] = \mathbb{E}[|\delta_1 X(1)|^q] \tau^{qH}, \text{ pour tout } \tau \text{ positif}, \quad (2.36)$$

et donc $C(q) = \mathbb{E}[|\delta_1 X(1)|^q]$ et $\zeta(q) = qH$ correspond alors à une fonction linéaire de q . En conséquence, le processus $\{X(t)\}_t$ est monofractal.

Les faits stylisés, que nous étudierons dans la section 2.3, mettent en évidence le caractère multifractal des séries financières. Il est donc naturel de chercher à construire des processus multifractals et d'étudier précisément leurs propriétés.

Propriété de l'exposant multifractal $\zeta(q)$

Proposition 2.13

Si la loi d'échelle (2.35) est valide lorsque τ tend vers zéro, c-à-d. si

$$\mathbb{E}[|\delta_\tau X(t)|^q] \underset{\tau \rightarrow 0}{\sim} C(q)\tau^{\zeta(q)}, \quad (2.37)$$

alors la fonction $\zeta(q)$ est concave.

Preuve - 2.13 -

L'application de l'inégalité de Hölder permet de mettre en évidence la concavité de l'exposant multifractal $\zeta(q)$. En effet, supposons que les moments d'ordre q_1 et q_2 sont finis, alors pour tout $\alpha \in [0, 1]$, nous avons

$$\mathbb{E} \left[|\delta_\tau X(t)|^{\alpha q_1 + (1-\alpha)q_2} \right] \leq \mathbb{E} [|\delta_\tau X(t)|^{q_1}]^\alpha \mathbb{E} [|\delta_\tau X(t)|^{q_2}]^{1-\alpha} \quad (2.38)$$

en prenant le logarithme et en notant $q = \alpha q_1 + (1 - \alpha)q_2$, nous avons

$$\ln(C(q)) + \zeta(q) \ln(\tau) \leq \alpha (\ln(C(q_1)) + \zeta(q_1) \ln(\tau)) + (1-\alpha) (\ln(C(q_2)) + \zeta(q_2) \ln(\tau)). \quad (2.39)$$

Ensuite, en divisant par $\ln(\tau)$ et en faisant tendre τ vers 0, ce qui est légitime puisque par l'hypothèse la loi d'échelle est valide lorsque τ tend vers zéro, nous trouvons que

$$\zeta(q) \geq \alpha \zeta(q_1) + (1 - \alpha) \zeta(q_2), \quad (2.40)$$

ce qui prouve que $\zeta(q)$ est une fonction concave. ■

La relation (2.39) permet également de montrer que le caractère multifractal d'un processus ne peut pas s'étendre à tout l'ensemble des valeurs positives \mathbb{R}^+ de l'échelle.

Proposition 2.14

Si l'exposant $\zeta(q)$ est non linéaire, alors la loi d'échelle (2.35) ne peut pas être valide au voisinage de zéro et l'infini simultanément, c'est-à-dire qu'on ne peut pas avoir à la fois

$$\mathbb{E} [|\delta_\tau X(t)|^q] \underset{\tau \rightarrow 0}{\sim} C(q) \tau^{\zeta(q)}, \quad (2.41)$$

$$\mathbb{E} [|\delta_\tau X(t)|^q] \underset{\tau \rightarrow \infty}{\sim} C(q) \tau^{\zeta(q)}, \quad (2.42)$$

Preuve - 2.14 -

En effet, si l'on suppose que l'invariance d'échelle (2.11) est vérifiée lorsque τ tend vers $+\infty$, en divisant la relation (2.39) par $\ln(\tau)$ et en faisant tendre, cette fois-ci, τ vers $+\infty$, on trouvera que

$$\zeta(q) \leq \alpha \zeta(q_1) + (1 - \alpha) \zeta(q_2), \quad (2.43)$$

ce qui est en contradiction avec la concavité et la non linéarité de la fonction $\zeta(q)$. ■

En conséquence, si nous supposons que le comportement des moments absolus est décrit par une loi de puissance (2.35) lorsque l'échelle τ appartient à un intervalle I qui contient zéro et si, de plus, l'exposant $\zeta(q)$ est non linéaire, alors l'intervalle I doit être forcément borné.

Invariances d'échelle

Pour conclure cette section nous allons définir rigoureusement les différentes propriétés d'**invariances d'échelle** que nous serons amenés à utiliser.

Définition 2.11

Soit $\{X(t); t \geq 0\}$ un processus à accroissements stationnaires. On distingue trois types d'invariance d'échelle.

- Le processus $\{X(t)\}_t$ possède une **invariance d'échelle exacte** caractérisée par l'exposant multifractal $\zeta(q)$ si les moments absolus des accroissements $\delta_\tau X(t)$ à l'échelle τ inférieure à certaine échelle T , dit **temps intégral**, varie en loi de puissance

$$\mathbb{E}[|\delta_\tau X(t)|^q] = C(q)\tau^{\zeta(q)}, \quad \text{pour } 0 \leq \tau \leq T, \quad (2.44)$$

où le facteur $C(q)$ égale au moment absolu d'ordre q à l'échelle $\tau = 1$.

- Le processus $\{X(t)\}_t$ possède une **invariance d'échelle asymptotique** si les moments absolus des accroissements $\delta_\tau X(t)$ à l'échelle τ sont caractérisés par un comportement en loi de puissance au voisinage de zéro

$$\mathbb{E}[|\delta_\tau X(t)|^q] \sim C(q)\tau^{\zeta(q)}, \quad \text{pour } \tau \rightarrow 0^+. \quad (2.45)$$

- Le processus $\{X(t)\}_t$ possède une **invariance d'échelle discrète** si les moments absolus des accroissements $\delta_\tau X(t)$ à l'échelle τ sont caractérisés par un comportement en loi de puissance pour un sous-ensemble discret $\{\tau_n\}_n$ d'échelles

$$\mathbb{E}[|\delta_{\tau_n} X(t)|^q] \sim C(q)\tau_n^{\zeta(q)}, \quad \text{pour } \tau_n \rightarrow 0^+. \quad (2.46)$$

Notons que dans le cas d'invariance d'échelle exacte au delà du temps intégral T le processus cesse de présenter un caractère multifractal.

2.2.2 Autosimilarité au sens stochastique

Nous avons déjà fait remarquer dans la section 2.1.8 que les processus autosimilaires à accroissements stationnaires, étudiés dans la section 2.1.5, sont mal adaptés à la modélisation de phénomènes réels observés sur les marchés financiers. Il existe de nombreuses raisons à cela. Le caractère multifractal (que nous étudierons précisément dans la section 1.2.3) des séries financières en est une. Le fait que les distributions des rendements ont des profils différents aux différentes échelles (queues épaisses à petites échelles, loi presque gaussienne à grandes échelles), alors que, comme le montre la proposition suivante, dans le cas de processus autosimilaires, ces distributions ne varient que par un facteur de dilatation, en est une autre.

Proposition 2.15

Un processus autosimilaire à accroissements stationnaires est un processus monofractal, il possède une invariance d'échelle exacte pour toute échelle positive, et la densité de probabilité de ses accroissements dépend de l'échelle de la façon suivante

$$P_{\delta_l X}(x) = (L/l)^H P_{\delta_L X}((L/l)^H x), \quad (2.47)$$

pour toutes échelles l et L .

Preuve - 2.15 -

Nous avons déjà mis en évidence le caractère monofractal des processus autosimilaires à accroissements stationnaires à l'aide de la relation (2.36). La même relation montre qu'ils possèdent une invariance d'échelle exacte (2.44). Il reste à étudier le comportement de la distribution de leurs accroissements en fonction de l'échelle.

La définition d'un processus autosimilaire à accroissements stationnaires implique

$$\delta_l X \stackrel{\mathcal{L}}{=} (l/L)^H \delta_L X \quad (2.48)$$

ce qui entraîne directement le résultat annoncé (2.47). \blacksquare

La démonstration précédente est basée sur le fait que la variable $(l/L)^H$ est déterministe. Une façon naturelle de rendre compte de la variation du profil de la densité de probabilité des rendements logarithmiques en fonction de l'échelle (cf. section 1.2.3) ainsi que de la multifractalité de ces rendements (étudié dans la section 2.3 décrivant les faits stylisés multifractals) consiste, comme nous allons le voir, à supposer que dans la définition 2.6 a^H n'est plus une constante mais une variable aléatoire. Cela nous amène à la définition suivante.

Définition 2.12

Un processus $\{X(t); t \in \mathbb{R}\}$ est un processus **autosimilaire au sens stochastique** si pour tout $0 < a \leq 1$, il existe une variable aléatoire positive W_a , indépendante du processus $\{X(t)\}_t$, dont la loi dépend seulement de a , et une échelle intégrale $T > 0$, telles que

$$\{X(at)\}_{0 \leq t \leq T} \stackrel{\mathcal{L}}{=} \{W_a X(t)\}_{0 \leq t \leq T}. \quad (2.49)$$

Proposition 2.16

La densité de probabilité des accroissements $P_{\delta_l X}(x)$ d'un processus autosimilaire au sens stochastique à accroissements stationnaires, vérifie la relation suivante

$$P_{\delta_l X}(x) = \int_{-\infty}^{\infty} G_{l/L}(u) P_{\delta_L X}(e^{-u}x) e^{-u} du, \quad (2.50)$$

pour toutes échelles $l \leq L \leq T$, où le **noyau autosimilaire** $G_{l/L}(u)$ est la densité de probabilité de la variable aléatoire $\ln(W_{l/L})$.

Preuve - 2.16 -

La définition 2.12, un processus autosimilaire à accroissements stationnaires implique

$$\delta_l X \stackrel{\mathcal{L}}{=} W_{l/L} \delta_L X, \quad (2.51)$$

qui est une généralisation de la relation (2.48). On en déduit que

$$\begin{aligned} \mathbb{P}[\delta_l X \in [x, x + dx]] &= \mathbb{P}[W_{l/L} \delta_L X \in [x, x + dx]] \\ &= \mathbb{E} \left[\mathbb{P}[\delta_L X \in [e^{-u}x, e^{-u}x + e^{-u}dx] \mid \ln(W_{l/L}) = u] \right], \end{aligned} \quad (2.52)$$

ce qui entraîne directement le résultat annoncé (2.50). \blacksquare

L'équation (2.50) est une généralisation directe de l'équation (2.47). Elle a été introduite dans le cadre de la théorie de la turbulence développée par Castaing et al. [CGH90]. On peut l'interpréter de la façon suivante : la densité de probabilité $P_{\delta_l X}(x)$ correspondant à la petite échelle l s'écrit comme la pondération par le noyau autosimilaire $G_{l/L}(u)$ de la densité de probabilité $P_{\delta_L X}(e^{-u}x)e^{-u}$ correspondant à la grande échelle L .

Théorème 2.17

Si le processus $\{X(t)\}_t$, qui possède une invariance d'échelle exacte (2.44), est multifractal et autosimilaire au sens stochastique, alors le noyau autosimilaire $G_{l/L}(u)$, introduit par l'équation (2.50), est relié à l'exposant multifractal $\zeta(q)$ par la relation

$$\zeta(q) = \ln \left(\widehat{G}_{l/L}(-iq) \right) / \ln(l/L), \quad (2.53)$$

où $\widehat{G}_{l/L}(k)$ est la fonction caractéristique de $\ln(W_{l/L})$ obtenue par la transformée de Fourier du noyau autosimilaire $G_{l/L}(u)$

$$\widehat{G}_{l/L}(k) = \mathbb{E} \left[e^{ik \ln(W_{l/L})} \right] = \int_{-\infty}^{\infty} G_{l/L}(u) e^{iku} du. \quad (2.54)$$

Preuve - 2.17 -

En utilisant un simple raisonnement de semi-groupe, nous pouvons démontrer que la transformée de Fourier $\widehat{G}_{l/L}(k)$ est de la forme suivante

$$\widehat{G}_{l/L}(k) = (l/L)^{F(k)}, \quad (2.55)$$

ce qui implique que la variable $\ln(W_{l/L})$ est infiniment divisible, voir la section 2.1.1.

En faisant le calcul du moment d'ordre q de l'accroissement absolu $\delta_l X(t)$, nous pouvons déduire que

$$\mathbb{E}[|\delta_l X(t)|^q] = \widehat{G}_{l/L}(-iq) \mathbb{E}[|\delta_L X(t)|^q] = \mathbb{E}[|\delta_L X(t)|^q] \left(\frac{l}{L} \right)^{F(-iq)}. \quad (2.56)$$

La comparaison avec (2.44) montre que le préfacteur $C(q)$ est donné par

$$C(q) = \mathbb{E}[|\delta_L X(t)|^q] / L^{F(-iq)} \quad (2.57)$$

et que l'exposant multifractal est relié au noyau autosimilaire par la relation

$$\zeta(q) = F(-iq). \quad (2.58)$$

Finalement, en utilisant l'équation (2.55) nous pouvons déduire le résultat souhaité (2.53). ■

Remarque sur les processus monofractals et multifractals

Dans le cas d'un processus monofractal, la variable aléatoire $W_{l/L}$, qui intervient dans la définition 2.12 de l'autosimilarité au sens stochastique, est dégénérée puisque toute la masse de probabilité est concentrée en un seul point $(l/L)^H$. En conséquence, la densité de probabilité de $\ln(W_{l/L})$, donnée par le noyau autosimilaire $G_{l/L}(u)$, est simplement une δ -fonction de Dirac

$$G_{l/L}(u) = \delta(u - H \ln(l/L)). \quad (2.59)$$

En utilisant la relation (2.50), on déduit que les densités de probabilité $P_{\delta_l X}$ et $P_{\delta_L X}$ des accroissements $\delta_l X$ et $\delta_L X$ correspondant aux échelles l et L ont le même profil (cf. équation (2.47)).

Dans le cas d'un processus multifractal, le profil de distribution varie en fonction de l'échelle et peut être, par exemple, gaussien pour les grandes échelles, ce qui correspond aux fait stylisés décrits dans la section 1.2.3. D'après le théorème 2.53, la non linéarité de l'exposant multifractal $\zeta(q)$ entraîne que le noyau $G_{l/L}(u)$ est différent d'une δ -fonction de Dirac. En conséquence, la variable aléatoire $\ln(W_{l/L})$ est non dégénérée.

Mesure d'intermittence

La plus simple mesure du caractère aléatoire de la variable $\ln(W_{l/L})$ est donnée par sa variance (quand celle-ci est finie). Un calcul simple basé sur la relation (2.53) donne

$$\text{Var}[\ln(W_{l/L})] = -\zeta''(0) \ln(l/L). \quad (2.60)$$

Définition 2.13

Le *coefficient d'intermittence*⁵ λ^2 est défini par

$$\lambda^2 = -\zeta''(0), \quad (2.61)$$

ce qui correspond à la courbure de l'exposant multifractal $\zeta(q)$ autour de zéro.

⁵Cette notion a été introduite dans le cadre de la théorie de turbulence développée.

Notons que les processus monofractals sont caractérisés par λ^2 égal à zéro, et les processus multifractals sont caractérisés par λ^2 positif. Ainsi le coefficient d'intermittence λ^2 "mesure" en quelque sorte l'écart au cas monofractal.

Dans le cas le plus simple du modèle log-normal, décrit dans la section 3.4, l'exposant multifractal $\zeta(q)$ correspond à une parabole entièrement définie par le coefficient d'intermittence λ^2 . En conséquence $G_{l/L}$ est une loi gaussienne et donc la variable aléatoire $W_{l/L}$ suit une loi log-normale.

2.2.3 Régularité locale, exposant de Hölder

Commençons par définir les espaces de Hölder homogènes.

Définition 2.14

Soit $\alpha > 0$. Une fonction $f : \mathbb{R} \mapsto \mathbb{R}$ appartient à la classe $C^\alpha(\mathbb{R})$ s'il existe une constante $C > 0$ et un polynôme P de degré inférieur à $\lfloor \alpha \rfloor$ tels que pour tout $t, t' \in \mathbb{R}$,

$$|f(t) - P(t - t')| \leq C|t - t'|^\alpha. \quad (2.62)$$

On va par ailleurs définir la régularité ponctuelle à partir de cette notion. La régularité ponctuelle d'une fonction autour d'un point t_0 se mesure à l'aide de l'exposant de Hölder ponctuel.

Définition 2.15

Soit $t_0 \in \mathbb{R}$ et soit $\alpha > 0$. Une fonction localement bornée $f : \mathbb{R} \mapsto \mathbb{R}$ appartient à la classe $C^\alpha(t_0)$ s'il existe une constante $C > 0$ et un polynôme P de degré inférieur à $\lfloor \alpha \rfloor$ tel que sur un voisinage de t_0 ,

$$|f(t) - P(t)| \leq C|t - t_0|^\alpha. \quad (2.63)$$

L'exposant de Hölder de f en t_0 est

$$h_f(t_0) = \sup\{\alpha | f \in C^\alpha(t_0)\}. \quad (2.64)$$

Si l'exposant de Hölder $h_f(t_0)$ est strictement supérieur à un nombre entier m alors la fonction $f(t)$ est m fois dérivable en t_0 , tandis qu'une discontinuité de la fonction $f(t)$ en t_0 implique $h_f(t_0) = 0$. L'exposant de Hölder peut également prendre des valeurs non entières, ainsi les trajectoires du mouvement brownien sont caractérisées par un exposant $h_f(t)$ égal presque partout à $1/2$.

2.2.4 Spectre de singularités

L'exposant de Hölder est un outil théorique puissant pour la caractérisation du comportement singulier d'un objet. Cependant, en pratique, il reste difficile à utiliser. Son estimation est souvent très instable, voire n'a pas de sens. En effet, rien n'oblige cet exposant à varier

lentement d'un point à l'autre. On peut construire, par exemple, des fonctions qui sont continues et dont l'exposant de Hölder est partout discontinu [Jaf97a, Jaf97b].

Dans le cas de processus autosimilaires au sens stochastique, l'exposant de Hölder varie généralement sur un continuum de valeurs $[\alpha_*, \alpha^*]$ et cela au voisinage de tout point (α_* et α^* ne dépendent pas du point considéré), voir la section 2.2.5 sur le formalisme multifractal. Dans ce cas, essayer d'associer un exposant de Hölder à chaque point d'une réalisation échantillonnée n'a aucun sens.

Il est donc naturel, pour décrire le comportement singulier d'un objet, d'introduire une nouvelle notion moins locale que l'exposant de Hölder mais dont on peut espérer que son estimation sur une réalisation échantillonnée ait un sens et soit robuste.

Définition 2.16

Soit $f : \mathbb{R} \mapsto \mathbb{R}$. On note $E_f(\alpha)$ l'ensemble de niveau α , qui peut prendre la valeur $+\infty$,

$$E_f(\alpha) = \{t | h_f(t) = \alpha\}, \quad (2.65)$$

l'ensemble des points où l'exposant de Hölder vaut α .

En général les ensembles de niveau α présentent une structure extrêmement complexe, impossible à décrire simplement par une forme euclidienne : ils sont des objets typiques de la géométrie fractale [HJK⁺86, Fal90].

Définition 2.17

Soit $f : \mathbb{R} \mapsto \mathbb{R}$. Le **spectre de singularités** de f , noté $D(\alpha)$, est l'application qui à α associe la dimension de Hausdorff de l'ensemble de niveau α .

Pour que cette définition soit complète, nous rappelons la notion de dimension de Hausdorff [Fal90] en appendice B.

Le spectre de singularités $D(\alpha)$ représente en quelque sorte la fréquence avec laquelle une singularité d'ordre α apparaît dans la fonction. Une interprétation non rigoureuse mais intuitive consiste à décomposer l'intervalle $[0, T]$, où T est le temps intégral, en intervalles de taille τ . Le nombre d'intervalles $N(\tau, \alpha)$, où l'accroissement $\delta_\tau f(t)$ se comporte comme τ^α est [HJK⁺86]

$$N(\tau, \alpha) \sim \tau^{-D(\alpha)}. \quad (2.66)$$

Donc, le spectre de singularités décrit comment "l'histogramme" $N(\tau, \alpha)$ varie lorsque τ tend vers zéro.

2.2.5 Formalisme multifractal

La définition du spectre de singularités reste difficile à utiliser en pratique : elle nécessite en effet la détermination de la mesure de Hausdorff, tâche difficile, d'ensembles présentant eux-mêmes une structure très complexe. Beaucoup d'auteurs [MBA91, BMA93, MBA94, Jaf97a, Jaf97b], et tout particulièrement en physique [FP85], ont donc cherché une méthode

pratique pour déterminer ce spectre. Ces travaux ont aboutit à l'analyse multifractale que nous allons présenter dans cette section.

Dans la suite nous présentons un calcul non rigoureux mais intuitif qui relie le spectre de singularités et l'exposant multifractal.

Si la loi d'échelle (2.35) est vérifiée au voisinage de zéro, il est possible d'exprimer le moment $\mathbb{E}[|\delta_\tau X|^q]$ en faisant intervenir le spectre de singularités, lorsque τ tend vers zéro. Au voisinage d'une singularité d'ordre α on a, dans l'intervalle de taille τ

$$|X(t) - X(t - \tau)|^q \sim \tau^{\alpha q}. \quad (2.67)$$

Si la dimension de cette singularité est $D(\alpha)$, il existe environ $\tau^{-D(\alpha)}$ de tels intervalles, chacun de taille τ . La contribution totale de cette singularité à l'espérance est donc environ

$$\tau^{\alpha q + 1 - D(\alpha)}. \quad (2.68)$$

D'après la méthode du col, l'ordre de grandeur de l'espérance est donné par la plus grande contribution qui correspond à l'exposant le plus faible, lorsque τ tend vers 0, d'où la relation

$$\zeta(q) = \inf_{\alpha} (q\alpha + 1 - D(\alpha)). \quad (2.69)$$

L'exposant multifractal $\zeta(q)$ est la transformée de Legendre de $D(\alpha) - 1$. Sous certaines conditions, essentiellement sous condition de concavité, le spectre de singularité $D(\alpha)$ peut alors être déduite de $\zeta(q)$ par la transformation de Legendre inverse.

Définition 2.18

On dit que le **formalisme multifractal** est vérifié si et seulement si l'exposant multifractal $\zeta(q)$ est relié au spectre de singularités $D(\alpha)$ par une transformation de Legendre.

Plus précisément le spectre de singularités $D(\alpha)$ est donné par

$$D(\alpha) = \begin{cases} D^*(\alpha), & \text{pour tout } \alpha \text{ tel que } \alpha_* \leq \alpha \leq \alpha^*, \\ -\infty, & \text{sinon,} \end{cases} \quad (2.70)$$

où $D^*(\alpha)$ est la transformée de Legendre de $\zeta(q)$

$$D^*(\alpha) = 1 + \min_q (q\alpha - \zeta(q)), \quad (2.71)$$

$$\zeta(q) = 1 + \min_{\alpha} (q\alpha - D^*(\alpha)), \quad (2.72)$$

et α_* et α^* définis par

$$\alpha_* = \inf\{\alpha | D(\alpha) = 0\}, \quad (2.73)$$

$$\alpha^* = \sup\{\alpha | D(\alpha) = 0\}. \quad (2.74)$$

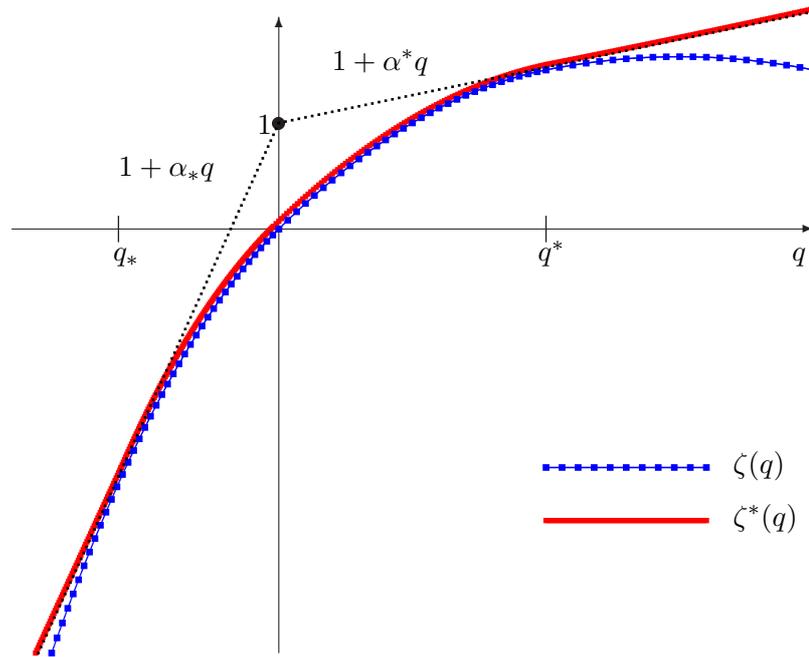


FIG. 2.3 – Exemple de l'exposant multifractal $\zeta(q)$ (trait pointillé) et la transformée de Legendre $\zeta^*(q)$ du spectre de singularités (trait solid).

Le formalisme multifractal a été démontré sur plusieurs classes de fonctions déterministes [Jaf97a, Jaf97b] ou stochastiques [BM02]. Dans ce dernier cas, le formalisme multifractal a été démontré au sens presque sûr, c-à-d. qu'il est valide sur presque toute réalisation.

Grâce à la concavité de l'exposant multifractal $\zeta(q)$, il est possible d'interpréter l'exposant de Hölder α comme la valeur de la dérivée de $\zeta(q)$ et réciproquement q comme la valeur de la dérivée de $D^*(\alpha)$. En effet, étant donné la valeur fixée de $q = q_0$, on a

$$D^*(\alpha_0) = 1 + q_0\alpha_0 - \zeta(q_0), \quad (2.75)$$

$$\alpha_0 = \frac{d\zeta}{dq}(q_0), \quad (2.76)$$

$$q_0 = \frac{dD^*}{d\alpha}(\alpha_0). \quad (2.77)$$

Ce que nous présentons sur la figure 2.4.

Notons de même que le spectre de singularités $D(\alpha)$ peut être vu comme la transformée de Legendre de la fonction $\zeta^*(q)$, définie comme

$$\zeta^*(q) = \begin{cases} 1 + \alpha^*q, & \text{pour } q < q^*, \\ \zeta(q), & \text{pour } q^* \leq q \leq q_*, \\ 1 + \alpha_*q, & \text{pour } q_* < q, \end{cases} \quad (2.78)$$

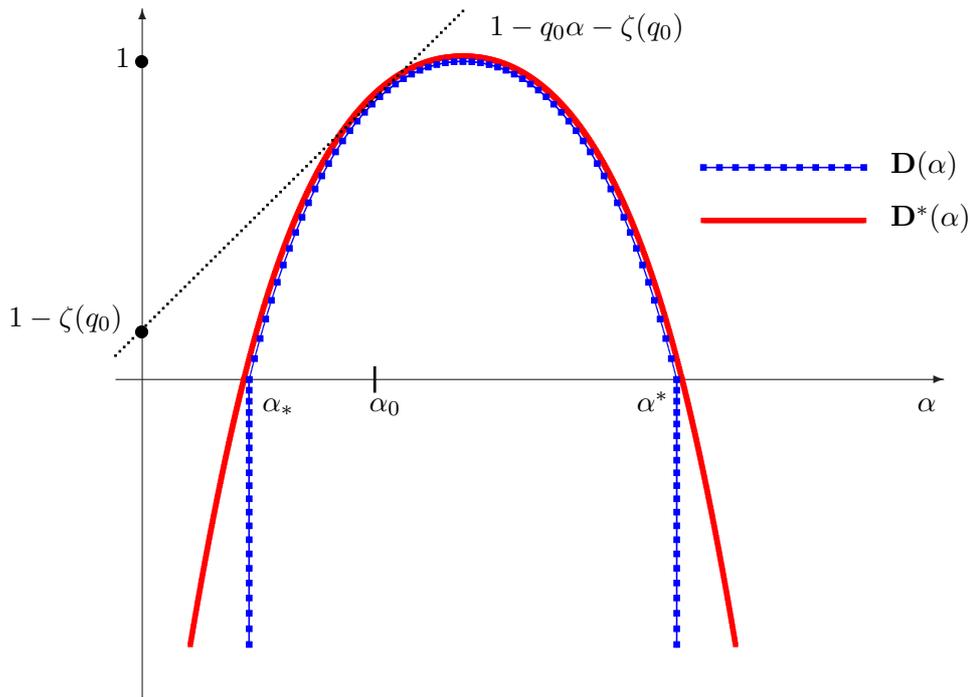


FIG. 2.4 – Exemple du spectre de singularités $D(\alpha)$ (trait pointillé) et la transformée de Legendre $D^*(\alpha)$ de l'exposant multifractal (trait solid).

où q_* et q^* donnés par

$$q_* = \frac{dD}{d\alpha}(\alpha_*), \quad (2.79)$$

$$q^* = \frac{dD}{d\alpha}(\alpha^*). \quad (2.80)$$

Il est important de noter qu'en pratique, sous les conditions usuelles, on ne peut estimer que la fonction $\zeta^*(q)$ et non pas l'exposant multifractal $\zeta(q)$ [Mol96, LAC04, MBK06].

Une interprétation probabiliste des **dimensions fractales négatives** en termes de grandes déviations a été proposée par Mandelbrot [Man90, MR97].

Cas monofractal et multifractal

Lorsque l'exposant multifractal $\zeta(q) = qH$ est une fonction linéaire, il est facile de voir que le formalisme multifractal conduit à un spectre réduit au point H , et $D(H) = 1$ ($D(\alpha) = -\infty$ pour $\alpha \neq H$). En revanche lorsque l'exposant multifractal $\zeta(q)$ est une fonction non linéaire, le spectre de singularités n'est plus dégénéré presque sûrement.

2.3 Faits stylisés multifractals

L'intérêt pour les processus multifractals en finance est relativement récent, la littérature dans ce domaine est encore peu abondante mais plusieurs études [VA98, PS99, PS00, SSL00, MDB00, CF01, CF02] ont déjà mis en évidence les faits stylisés multifractals dans plusieurs séries financières à l'aide de différentes méthodes : ondelettes [AMS98] ou fonction de structure [FCM97, CF01, CF02].

2.3.1 Hétérogénéité du marché

Une hypothèse classique en économie est celle d'un **marché homogène**, où tous les investisseurs se comportent de la même façon, en investisseurs rationnels maximisant leur fonction d'utilité. Les investisseurs réagissent donc tous de la même façon aux événements.

L'hypothèse de **marché hétérogène** est une nouvelle idée en modélisation. Cette hypothèse suppose que les agents ont des perceptions du marché, des degrés d'information et des règles de comportement différents, ce qui peut expliquer l'absence de l'échelle caractéristique.

Pour le marché des changes, il apparaît que les agents ont des profils de risques, des localisations géographiques et des contraintes institutionnelles différents. Il y a eu un certain nombre d'articles dans les années 1990 rejetant l'hypothèse traditionnelle d'agents rationnels interprétant toutes les informations de la même manière et n'ayant aucune raison d'avoir des horizons de temps différents, et ce, pas uniquement pour le marché des devises.

Dans les travaux [MDD⁺95, MDD⁺97], il est montré que c'est l'horizon de temps sur lequel les investisseurs influencent le marché qui est l'aspect le plus important des différences de comportement et de perception du marché. Pour eux, les différences entre agents vont essentiellement se traduire par des horizons de temps différents. Ils tentent d'observer cela pour le marché des devises. L'idée est que les agents agissant à court terme évaluent le marché à de plus hautes fréquences et ont une plus courte mémoire que ceux travaillant à plus long terme. Une augmentation rapide de 0.5% suivie par une baisse rapide de 0.5% est un événement important pour un agent à court terme, mais pas pour une banque centrale ou un autre investisseur à long terme. Les agents à long terme s'intéressent, quant à eux, plus à la tendance générale sur de longs intervalles. Ils regardent le marché (et sa volatilité) sur un maillage plus grossier. Les agents à court terme et à long terme n'auront donc pas les mêmes ensembles d'opportunités, ils réagiront différemment à une même information.

2.3.2 Invariance d'échelle des rendements

Comme nous avons noté dans la section 1.2.3, la distribution des rendements change avec l'échelle de temps en se rapprochant d'une loi normale pour de grands intervalles de temps.

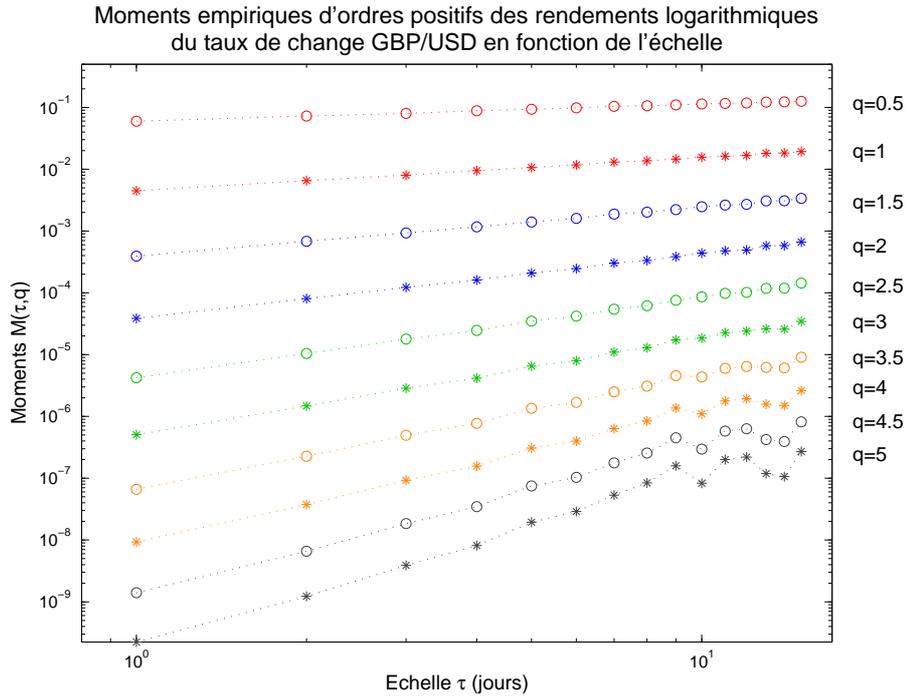


FIG. 2.5 – Fonction de partition (2.81) des rendements logarithmiques du taux de change GBP/USD, pour l'ordre q variant de 0.5 à 5 et l'échelle τ variant de 1 jour à 15 jours. Le comportement linéaire des moments aux échelles logarithmiques confirme que les rendements logarithmiques possède une invariance d'échelle.

Pour mettre en évidence le caractère multifractal d'une série financière il est naturel d'évaluer les différents moments des rendements logarithmiques $\mathbb{E}[|\delta_\tau X(t)|^q]$ par leurs moyennes empiriques

$$M(\tau, q) = \frac{1}{\lfloor N/\tau \rfloor} \sum_{k=1}^{\lfloor N/\tau \rfloor} |X(k\tau) - X(k\tau - \tau)|^q, \quad (2.81)$$

où N est la longueur de la série. La fonction $M(\tau, q)$ devrait se comporter en fonction de τ comme une loi de puissance si la série est multifractale. Pour le vérifier, en pratique, on trace les logarithmes des moyennes empiriques $\ln(M(\tau, q))$ en fonction de $\ln(\tau)$ pour détecter le comportement linéaire.

Pour des moments supérieurs à 5, les fluctuations statistiques sont assez élevées, ce qui introduit une grande incertitude de la signification statistique des résultats.

Sur la figure 2.5 nous avons évalué les différents moments absolus des rendements logarithmiques du taux de change GBP/USD par leur moyenne empirique (2.81). L'hypothèse de l'invariance d'échelle doit se traduire par un tracé linéaire aux échelles logarithmiques, ce qui est confirmé au moins pour les premières valeurs de q . Pour les ordres q supérieurs à 5 et pour les échelles supérieures à 15 jours, les fluctuations statistiques sont amplifiées,

c'est-à-dire que seulement quelques termes dominant la fonction de partition (2.81), et, si les tracés conservent une allure qualitativement linéaire, l'incertitude est grande et la signification statistique des résultats obtenus est sujette à caution.

2.3.3 Autocorrélation des magnitudes

Dans la section 1.2.5 nous avons déjà mis en évidence la présence de mémoire longue dans les séries financières en calculant des autocovariances des rendements absolus ou quadratiques. Une autre observation importante est celle de l'autocovariance des logarithmes des rendements (logarithmiques) absolus, appelés **magnitudes**. Une étude effectuée sur des données réelles montre que cette fonction d'autocovariance se comporte comme

$$\gamma_m(h) = \text{Cov}[\ln |\delta_\tau X(t)|, \ln |\delta_\tau X(t+h)|] \sim \begin{cases} -\lambda^2 \ln \frac{T}{h}, & \text{pour } l \ll h \leq T, \\ 0, & \text{pour } h > T, \end{cases} \quad (2.82)$$

où le facteur λ^2 est de l'ordre de 0.02 et le coefficient T peut varier entre quelques mois et quelques années. Une propriété remarquable de cette fonction d'autocovariance est qu'elle ne dépend pas du choix de l'échelle τ .

La figure 2.6 représente la fonction d'autocovariance empirique des magnitudes journaliers du taux de change GBP/USD. Elle reste très longtemps significativement positive et indique donc l'existence d'une autocorrélation des magnitudes. Une régression effectuant de cette fonction d'autocovariance contre le logarithme du temps indique qu'elle peut être bien approximée par une fonction décroissante logarithmiquement jusqu'à zéro, comme le montre le tracé tireté sur la figure.

2.4 Modèles multifractals

Les premiers objets multifractals proposés dans la littérature ont été les mesures multifractales construites à partir de cascades qui ont été développées par l'école russe [Yag66] pour la modélisation en turbulence. Après les premiers travaux de Mandelbrot [Man74a, Man74b, Man89, Man03], de nombreuses études ont été consacrées aux cascades aléatoires multiplicatives discrètes. Leurs convergence et propriétés de régularité ont été extensivement étudiés notamment dans les travaux suivants [KP76, Gui87, Mol96, Mol97]. A notre connaissance, les constructions précises des processus multifractals à temps continus à accroissements stationnaires sont apparues initialement dans les travaux de Bacry et Muzy [MDB00, BDM01, MB02, BM03] et de Barral et Mandelbrot [BM02]. Dans ces travaux les processus multifractals sont obtenus comme des limites de processus stochastiques.

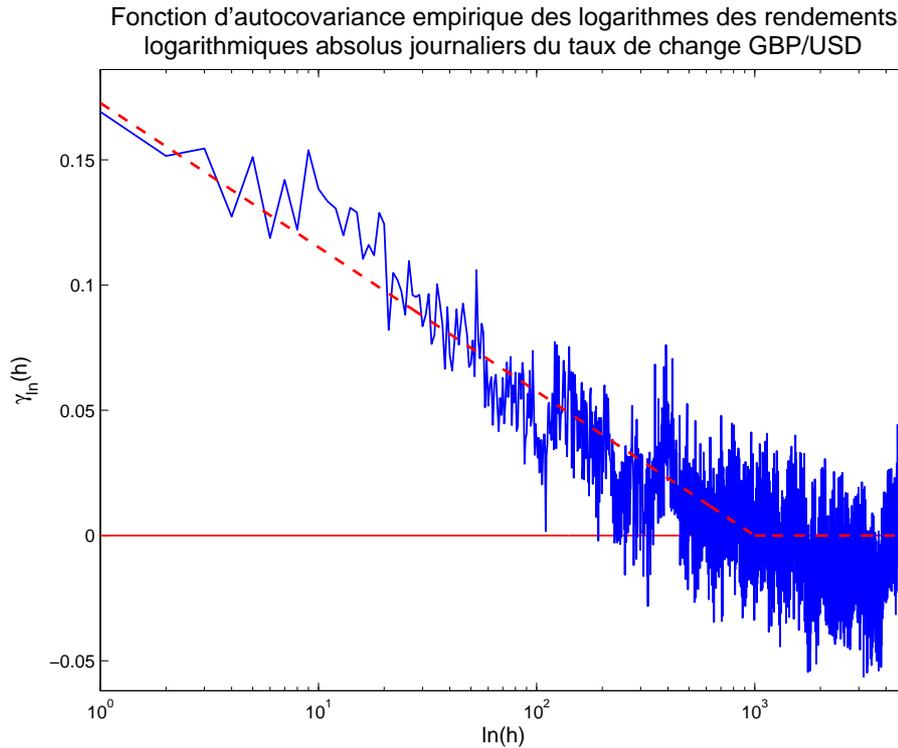


FIG. 2.6 – Fonction d'autocovariance empirique des magnitudes journalières du taux de change GBP/USD. Le trait tireté correspond à la régression par une fonction décroissante logarithmique $0.025 \ln(1000/h) \mathbf{1}_{\{h \leq 1000\}}$.

2.4.1 Modèles en cascade discrète

Mesure binomiale

L'exemple de cascade le plus simple, non aléatoire, est la mesure binomiale, ou mesure de Bernoulli, sur l'intervalle $[0, 1]$, introduite par Mandelbrot [Man74a, Man89].

La mesure binomiale μ sur l'intervalle $[0, 1]$ est définie comme la limite d'une suite de mesures μ_k . Considérons la mesure de probabilité uniforme μ_0 sur l'intervalle $[0, 1]$ et deux réels positifs m_0 et m_1 , tels que $m_0 + m_1 = 1$.

A la première étape, on divise l'intervalle $[0, 1]$ en deux intervalles $[0, 1/2]$ et $[1/2, 1]$. La mesure μ_1 est construite en répartissant uniformément la masse m_0 sur l'intervalle $[0, 1/2]$ et la masse m_1 sur l'intervalle $[1/2, 1]$:

$$\mu_1([0, 1/2]) = m_0, \quad \mu_1([1/2, 1]) = m_1.$$

A la seconde étape, l'intervalle $[0, 1/2]$ est à nouveau divisé en deux intervalles $[0, 1/4]$ et

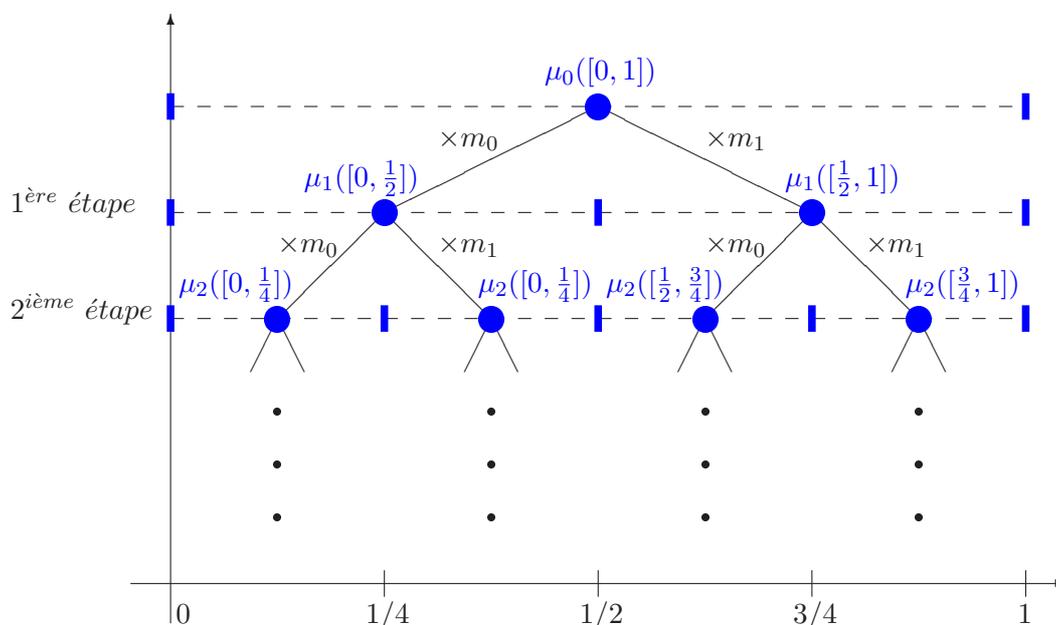


FIG. 2.7 – Construction de la mesure binomiale

$[1/4, 1/2]$ qui reçoivent respectivement une fraction m_0 et m_1 de la mesure $\mu_1([0, 1/2])$:

$$\mu_2([0, 1/4]) = \mu_1([0, 1/2])m_0 = m_0m_0, \quad \mu_2([1/4, 1/2]) = \mu_1([1/2, 1])m_1 = m_0m_1.$$

La même procédure est appliquée à l'intervalle $[1/2, 1]$:

$$\mu_2([1/2, 3/4]) = \mu_1([1/2, 1])m_0 = m_1m_0, \quad \mu_2([3/4, 1]) = \mu_1([1/2, 1])m_1 = m_1m_1,$$

ce qui conduit à la construction d'une mesure μ_2 .

On continue dans la suite par itération : à l'étape $k + 1$, étant donnée la mesure μ_k et l'intervalle $[t, t + 2^{-k}]$, où t est un nombre dyadique, la masse $\mu_k([t, t + 2^{-k}])$ est répartie selon respectivement les fractions m_0 et m_1 sur deux intervalles $[t, t + 2^{-k-1}]$ et $[t + 2^{-k-1}, t + 2^{-k}]$.

L'application de cette procédure à tous les intervalles $[t, t + 2^{-k}]$ permet construire la mesure μ_{k+1} . Comme nous avons noté ci-dessus la mesure binomiale μ est définie comme

la limite (au sens de la topologie faible) des mesures μ_k :

$$\mu = \lim_{k \rightarrow \infty} \mu_k. \quad (2.83)$$

Il est assez naturel d'étendre la construction binomiale. Ainsi, on peut considérer la mesure multinomiale, où les intervalles sont divisés en $n > 2$ sous-intervalles à chaque étape. Chacun d'entre eux reçoit une fraction m_j (pour $j = 1, \dots, n$) de la masse totale. Pour assurer la conservation de la mesure totale, il est nécessaire d'imposer la condition de normalisation $\sum_{j=1}^n m_j = 1$.

Cascades multiplicatives

Il est possible de généraliser les constructions précédentes en considérant la classe plus vaste des cascades multiplicatives qui ne possèdent plus de caractère déterministe. L'idée consiste à considérer comme générateur de la cascade une variable aléatoire W , positive, d'espérance 1, plutôt que des valeurs discrètes déterministes.

On part de la distribution uniforme de l'intervalle $[0, 1]$. En utilisant la même division diadique comme dans le cas de la cascade binomiale, à la $n^{\text{ième}}$ étape de construction, si l'on introduit des adresses diadiques $k_1 \dots k_n$, où $k_j = 0, 1$, on multiplie les densités des sous-intervalles $I_{k_1 \dots k_n}$ par des variables aléatoires positives $W_{k_1 \dots k_n}$ indépendantes et distribuées de même loi que W . En conséquence, la masse de l'intervalle diadique $I_{k_1 \dots k_n}$ est simplement égale à

$$M_n(I_{k_1 \dots k_n}) = 2^{-n} \prod_{j=1}^n W_{k_1 \dots k_n}. \quad (2.84)$$

Théorème 2.18 (Kahane-Peyrière)

Sous la condition

$$\mathbb{E}[W \ln(W)] < 1, \quad (2.85)$$

*la construction d'une cascade multiplicative converge presque sûrement vers un processus non dégénéré M , qu'on appelle la **mesure multiplicative**.*

Preuve - 2.18 -

La démonstration est détaillée dans [KP76]. ■

Par construction, la mesure multiplicative vérifie la propriété suivante

$$M[I_{k_1 \dots k_n}] \stackrel{\mathcal{L}}{=} \frac{1}{2} W M[I_{k_1 \dots k_{n-1}}]. \quad (2.86)$$

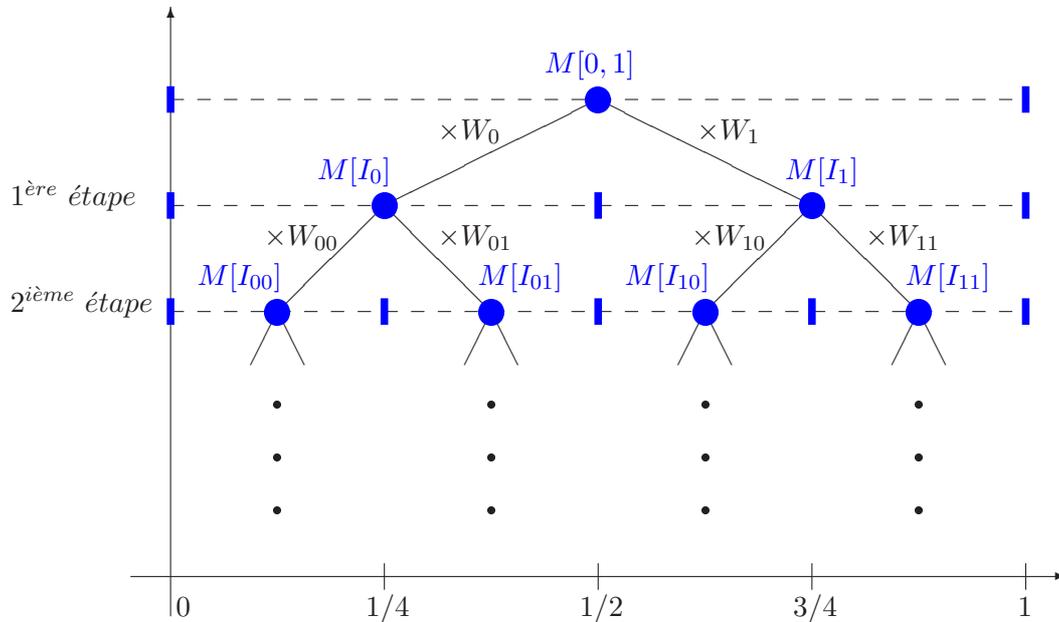


FIG. 2.8 – Construction de la cascade multiplicative

Propriétés de la mesure multiplicative

Proposition 2.19

La mesure multiplicative M vérifie l'identité en loi

$$M \stackrel{\mathcal{L}}{=} \frac{W_1}{2} M_1 + \frac{W_2}{2} M_2, \tag{2.87}$$

où M_1 et M_2 sont des copies de M , W_1 et W_2 sont des copies de W , et M_1, M_2, W_1 et W_2 sont toutes indépendantes.

Cette identité appelée *star equation* de Mandelbrot [Man01], peut être vue comme une généralisation de l'équation définissant les lois stables, étudiées dans la section 2.1.2. Elle a fait l'objet de multiple travaux en probabilité [KP76, DL83, Gui90, Liu02].

Le caractère multifractal de la cascade multiplicative est démontré, par exemple, dans le travail de Molchan [Mol96].

Proposition 2.20

La mesure multiplicative possède l'invariance d'échelle discrète (2.46), dont l'exposant multifractal $\zeta_M(q)$ est donnée par

$$\zeta_M(q) = q - \log_2(\mathbb{E}[W^q]). \quad (2.88)$$

Preuve - 2.20 -

Par itération de l'équation (2.86), il est possible de calculer le moment d'ordre q

$$\mathbb{E}[(M[I_{k_1 \dots k_n}])^q] = 2^{-nq}(\mathbb{E}[W^q])^n \mathbb{E}[(M[0, 1])^q] = \mathbb{E}[(M[0, 1])^q | I_{k_1 \dots k_n}]^{\zeta_M(q)}, \quad (2.89)$$

où $|I_{k_1 \dots k_n}| = 2^{-n}$. La fonction $\zeta_M(q)$ est alors donnée par l'équation (2.88). Et donc le comportement des moments présente le caractère multifractal. ■

Un résultat important concerne les queues des lois de probabilités solutions de la *star equation* de Mandelbrot :

Théorème 2.21 (Guivarc'h)

Si $\mathbb{E}[W \ln(W)] < 1$, alors la distribution de probabilité de M solution de la *star equation* de Mandelbrot est unique, à une homothétie près, et il existe une constante K telle que :

$$\mathbb{P}[M > x] \underset{x \rightarrow +\infty}{\sim} K x^{-q_c}, \quad (2.90)$$

où q_c est l'unique solution de $\zeta_M(q) = 1$ telle que $q_c > 1$.

Preuve - 2.21 -

La démonstration est détaillée dans [Gui87, Gui90]. ■

Ce théorème implique en particulier que les moments de M divergent pour un ordre q plus grand que q_c .

Une autre quantité intéressante est la structure de covariance de la mesure multiplicative. La proposition suivante montre la présence de la structure compliquée de corrélations dans les cascades multiplicatives.

Proposition 2.22

Considérons deux intervalles disjoints $I_{k_1 \dots k_{n_1}}$ et $I_{l_1 \dots l_{n_2}}$, et supposons que les premiers m indices des adresses diadiques soient identiques $k_1 \dots k_m = l_1 \dots l_m$, alors nous avons

$$\mathbb{E}[M[I_{k_1 \dots k_{n_1}}]^{q_1} M[I_{l_1 \dots l_{n_2}}]^{q_2}] = \frac{\mathbb{E}[(M[0, 1])^{q_1 + q_2}]}{2^{m\zeta_M(q_1 + q_2) - m\zeta_M(q_1) - m\zeta_M(q_2) + n_1\zeta_M(q_1) + n_2\zeta_M(q_2)}}. \quad (2.91)$$

Preuve - 2.22 -

En utilisant la relation (2.86) nous pouvons déduire que

$$\begin{aligned} & \mathbb{E} \left[\left(M[I_{k_1 \dots k_{n_1}}] \right)^{q_1} \left(M[I_{l_1 \dots l_{n_2}}] \right)^{q_2} \right] \\ &= \frac{1}{2^{n_1 q_1 + n_2 q_2}} \mathbb{E} \left[(M[0, 1])^{q_1 + q_2} \right] (\mathbb{E} [W^{q_1 + q_2}])^m (\mathbb{E} [W^{q_1}])^{(n_1 - m)} (\mathbb{E} [W^{q_2}])^{(n_2 - m)} \end{aligned} \quad (2.92)$$

En exprimant les moments de W à l'aide de l'équation (2.88), on déduit l'expression (2.91). ■

Exemple d'une cascade multiplicative, cascade log-normale

Dans le cas le plus simple, W est distribuée selon une loi log-normale

$$\ln(W) \sim \mathcal{N}(-2\lambda^2 \ln(2), 4\lambda^2 \ln(2)), \quad (2.93)$$

où la moyenne est définie par la condition $\mathbb{E}[W] = 1$. Un calcul gaussien montre que

$$\mathbb{E}[W^q] = 2^{2q(q-1)\lambda^2}, \quad (2.94)$$

et alors l'exposant multifractal est donné par

$$\zeta_M^{\ln}(q) = q(1 + 2\lambda^2) - 2\lambda^2 q^2. \quad (2.95)$$

L'expression (2.95) de l'exposant multifractal permet d'interpréter le paramètre $4\lambda^2$ comme le coefficient d'intermittence défini par (2.61).

Par ailleurs, d'après la théorème 2.21, les moments divergent au delà du seuil $q_c = \frac{1}{2\lambda^2} > 1$ tel que $\zeta_M^{\ln}(q_c) = 1$.

Qualités et défauts des cascades discrètes

Notons que l'extension des cascades binomiales et multiplicatives à un intervalle quelconque $[0, T]$ est immédiate en suivant les mêmes étapes décrites ci-dessus.

La construction de la mesure binomiale peut être généralisée. On obtient alors une vaste classe de mesures multifractales construites par cascades multiplicatives qui conduisent à des mesures aléatoires dont les moments divergent à partir d'un certain ordre.

Cependant, il est facile de montrer à partir de l'expression (2.91) que ces mesures multiplicatives ne sont pas stationnaires ce qui peut limiter leur application. La construction utilise une discrétisation diadique de l'intervalle dont la rigidité est le principal défaut de la mesure multiplicative. En effet, l'invariance d'échelle n'est pas strictement continue

mais restreinte aux sous-intervalles diadiques, ce qui correspond à une invariance d'échelle discrète (2.46). Nous verrons plus loin comment il est possible de résoudre ce problème.

Une autre limitation provient de la construction non causale, "de haut en bas", de la cascade construite sur un intervalle $[0, t]$ il sera nécessaire de recommencer entièrement la construction pour l'avoir sur l'intervalle $[0, t + \tau]$.

2.4.2 Modèle MMAR de Mandelbrot, Fisher et Calvet

Définition

Pour construire une classe de processus multifractals à partir d'une mesure qui possède des caractéristiques multifractales, Mandelbrot, Fisher et Calvet [MFC97, CFM97, FCM97] proposent d'utiliser la subordination (voir [MT67, Fel68]) d'un mouvement brownien fractionnaire. Leur modèle MMAR (Multifractal Model of Asset Return) est destiné à la modélisation de séries financières. Dans ce modèle, on suppose que le processus de prix logarithmiques $X(t) = \ln(P(t)/P(0))$ est donné par

$$X(t) = B^H(\theta(t)) \quad (2.96)$$

où B^H est un mouvement brownien fractionnaire et θ la fonction de répartition d'une mesure aléatoire multifractale sur $[0, T]$, appelée le **temps physique** ou **trading time**, indépendante de B^H . Les auteurs sous-entendent que la mesure multifractale est une mesure multiplicative décrite précédemment.

Remarque sur la notion de "trading time"

Notons qu'en finance l'utilisation d'un processus subordinateur est une vieille suggestion qui remonte à Clark [Cla73]. Plus récemment, pour mieux rendre compte du fonctionnement du marché qui connaît des périodes de hausse ou de baisse d'activité ("lunch effect", "end-of-the-day effect"), certains proposent d'utiliser un temps déformé, calibré sur une mesure de cette activité, à la place du temps physique [MDD⁺95, AG00].

Moments

Il est clair que les propriétés du processus $X(t)$ en grande partie dépendent de celles de la fonction $\theta(t)$. Ainsi, si l'on note $\{\mathcal{F}_t^\theta; t \in [0, T]\}$ la filtration engendrée par le processus $\theta(t)$, le moment d'ordre q de processus X s'écrira comme :

$$\begin{aligned} \mathbb{E}[|X(t)|^q] &= \mathbb{E}[|B^H(\theta(t))|^q] = \mathbb{E}\left[\mathbb{E}\left[|B^H(\theta(t))|^q \mid \mathcal{F}_t^\theta\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[|\theta(t)^H B^H(1)|^q \mid \mathcal{F}_t^\theta\right]\right] = \mathbb{E}[\theta(t)^{qH}] \mathbb{E}[|B^H(1)|^q], \quad (2.97) \end{aligned}$$

où nous avons utilisé l'indépendance entre B^H et θ . D'après ce calcul, la finitude du moment d'ordre q du processus X dépend de l'existence du moment $\mathbb{E}[\theta(t)^{qH}]$. Comme les moments de $\theta(t)$ divergent au delà d'un ordre q_c alors le moment $\mathbb{E}[|X(t)|^q]$ diverge pour $q > q_c/H$. Notons que la condition de divergence ne dépend pas de t .

Invariance d'échelle

En utilisant la relation (2.97) on peut relier les exposants multifractals des processus $\theta(t)$ et $X(t)$ par la formule

$$\zeta_X(q) = \zeta_\theta(qH) \quad (2.98)$$

En effet, en utilisant la définition de l'exposant multifractal (2.11) et la relation (2.97), on a

$$C^X(q)t^{\zeta_X(q)} \sim \mathbb{E}[|X(t)|^q] = \mathbb{E}[\theta(t)^{qH}] \mathbb{E}[|B^H(1)|^q] \sim \mathbb{E}[|B^H(1)|^q] C_\theta(q)t^{\zeta_\theta(qH)} \quad (2.99)$$

Par exemple, dans le cas, où la fonction θ est la fonction de répartition d'une mesure multiplicative log-normale (2.93), l'exposant s'écrit

$$\zeta_X^{\ln}(q) = \zeta_\theta^{\ln}(qH) = qH(1 + 2\lambda^2) - 2\lambda^2 q^2 H^2. \quad (2.100)$$

En particulier, lorsque $H = 1/2$, l'expression devient

$$\zeta_X^{\ln}(q) = \frac{q}{2}(1 + 2\lambda^2) - \frac{\lambda^2 q^2}{2}. \quad (2.101)$$

Corrélations

Lorsque $H = 1/2$ (on compose avec un mouvement brownien standard B), les accroissements du processus MMAR sont décorrélés mais ils ne sont pas indépendants :

$$\mathbb{E}[(\delta_\tau X(t))^{2q} (\delta_\tau X(t+\tau))^{2q}] = \mathbb{E}[(\delta_\tau \theta(t))^q (\delta_\tau \theta(t+\tau))^q]^2 (\mathbb{E}[(B(1))^{2q}])^2. \quad (2.102)$$

La structure de corrélation de la mesure multifractale permet au processus MMAR de tenir compte du phénomène de longue mémoire observé sur les marchés financiers.

Choisir H différent de $1/2$ permet d'avoir des accroissements corrélés, positivement si $H > 1/2$ et négativement si $H < 1/2$.

Faits stylisés

La construction du processus MMAR repose sur la subordination d'un mouvement brownien fractionnaire par une mesure multifractale. C'est à travers ce changement de temps que le processus acquiert la dépendance des incréments et les propriétés de multifractalité.

Le modèle MMAR permet de modéliser la plupart des faits stylisés classiques décrits dans la section 1.2 à l'exception de la corrélation entre le rendement et la volatilité instantanée ainsi qu'il permet de reproduire les faits stylisés multifractals grâce au caractère multifractal de la mesure.

Cependant, le modèle MMAR possède au moins un défaut majeur. Il dépend de la mesure multifractale construite d'avance, autrement dit ce modèle n'est pas causal ce qui limite son utilisation pour modéliser des comportements des actifs financiers.

De même, nous devons noter que sans préciser la mesure multifractale le modèle MMAR reste incomplet. Dans les articles [MFC97, CFM97, FCM97] les auteurs sous-entendent implicitement l'utilisation de la mesure multiplicative, étudiée dans la section 2.4.1. Dans ce cas, le modèle MMAR hérite tous les défauts de la mesure multiplicative, et, en particulier, l'invariance d'échelle restreinte aux intervalles diadiques.

2.4.3 Modèle MSM de Calvet et Fisher

Dans [CF01], Calvet et Fisher présentent une extension très intéressante du modèle MMAR. Leur but est de s'affranchir des contraintes de la mesure multiplicative dans la construction initiale du processus MMAR.

Au lieu de diviser à chaque étape un intervalle en sous-intervalles de manière déterministe, ils proposent d'utiliser des instants aléatoires, où les multiplicateurs changent. Ces instants suivent une loi de Poisson dont l'intensité progresse à chaque étape. Le processus PMM (Poisson Multifractal Model) obtenu possède des propriétés comparables à celles du processus MMAR mais possède en plus un caractère stationnaire et causal. Par ailleurs, grâce au choix du processus de Poisson, la construction de Calvet et Fisher hérite d'une structure markovienne. Ceci leur permet de bâtir un processus discret qui converge vers le processus PMM précédent mais dont la construction est causale et peut se faire lorsque le temps progresse.

Nous considérons une série $\{X[t]; t \in \mathbb{N}\}$ définie en temps discret, qui sert à modéliser des prix logarithmiques espacés régulièrement dans le temps, avec \bar{k} composants $M_1[t], \dots, M_{\bar{k}}[t]$ caractérisés par les fréquences $\gamma_1, \dots, \gamma_{\bar{k}}$. Les rendements logarithmiques journaliers s'écrivent comme

$$\delta_1 X[t] = X[t] - X[t-1] = \sigma (M_1[t] \cdots M_{\bar{k}}[t])^{1/2} \epsilon[t], \quad (2.103)$$

où le paramètre σ est une constante positive et $\epsilon[t]$ est un bruit blanc gaussien. Les **multiplicateurs** ou les **composantes de volatilité** $M_k[t]$ sont stationnaires, non négatifs d'espérance 1. De plus, nous supposons qu'ils sont indépendants à tout instant t . En conséquence, σ^2 correspond à la variance non conditionnel des rendements logarithmiques journaliers $\delta_1 X[t]$.

L'équation (2.103) définit le modèle à volatilité stochastique $\delta_1 X[t] = \sigma[t] \epsilon[t]$ avec la structure multiplicative de volatilité $\sigma[t] = \sigma (M_1[t] \cdots M_{\bar{k}}[t])^{1/2}$. Les propriétés de volatilité

sont gouvernées par la dynamique stochastique des multiplicateurs $M_k[t]$, pour lesquels nous admettons la construction récursive. Si nous supposons qu'ils sont connus à l'instant $t - 1$ alors pour tout $k \in \{1, \dots, \bar{k}\}$ le multiplicateur $M_k[t]$ correspondant à la période suivante est tiré indépendamment selon une loi fixée à l'avance M avec la probabilité γ_k et reste égal à sa valeur présente $M_k[t - 1]$ avec la probabilité $1 - \gamma_k$.

Les probabilités de transition $\gamma_1, \dots, \gamma_{\bar{k}}$ sont spécifiées dans [CF01] de telle façon que

$$\gamma_k = 1 - (1 - \gamma_1)^{(b^{k-1})} = 1 - (1 - \gamma_{\bar{k}})^{(b^{\bar{k}-k})}, \quad (2.104)$$

où $\gamma_1 \in (0, 1)$ et $b \in (1, \infty)$. Pour les petites valeurs de k les probabilités de transition satisfont les approximations suivantes

$$\gamma_k \simeq \gamma_1 b^{k-1} \simeq \gamma_{\bar{k}} b^{\bar{k}-k}. \quad (2.105)$$

La construction décrite impose seulement deux restrictions sur la loi fixée M : $M \geq 0$ et $\mathbb{E}[M] = 1$. Calvet et Fisher proposent deux types de modèles les plus simples. Dans le premier cas M est distribuée selon une loi binomiale qui prend deux valeurs équiprobables m_0 et $2 - m_0$, dans le deuxième M suite une loi log-normale $\ln(M) \sim \mathcal{N}(m_0, -m_0^2/2)$.

Le processus $X[t]$ construit de cette façon, dit le processus **MSM**(\bar{k}), (*Markov-Switching Multifrequency*), dépend de cinq paramètres : m_0 , σ , b , $\gamma_{\bar{k}}$ et \bar{k} . Calvet et Fisher proposent d'utiliser la méthode de maximum de vraisemblance pour estimer ces paramètres. L'utilisation de cette méthode est critiquée dans [Lux04], la matrice de transition de taille $2^k \times 2^k$, où la valeur de \bar{k} peut monter jusqu'à 10, entraîne des calculs d'une grande complexité.

De même, il est impossible d'appliquer la méthode de maximum de vraisemblance dans le cas de la loi log-normale. Pour éviter ces inconvénients Lux dans [Lux03] suggère d'utiliser la méthode GMM (*Generalized Method of Moments*).

Faits stylisés

Le modèle MSM de Calvet et Fisher à travers la structure multiplicative de la volatilité permet de reproduire la plupart des faits stylisés classiques et multifractals : l'absence de corrélation temporelle (cf. section 1.2.4), les queues de distribution des accroissements sont plus épaisses que celles de la loi normale, le phénomène d'accumulation des volatilités (cf. section 1.2.7), la présence de la mémoire longue (cf. section 1.2.5), la variation de distribution des accroissements en fonction de l'échelle 1.2.3 et la variation des moments absolus en loi de puissance (cf. section 2.3.2).

La construction du processus MSM est explicitement symétrique, en conséquence, ce modèle ne permet pas de reproduire l'asymétrie observée dans les données financières (cf. section 1.2.8).

Bien qu'à accroissements stationnaires, le modèle MSM de Calvet et Fisher est formellement proche d'une cascade discrète avec un nombre fini d'étapes de construction : b joue

le rôle du facteur d'échelle, comme le facteur 2 dans le cadre des cascades binomiales, et \bar{k} correspond au nombre d'étapes de construction. En d'autres termes, le processus MSM ne possède, au mieux (dans la limite où le paramètre \bar{k} tend vers l'infini), que l'invariance d'échelle discrète (2.46), puisque le facteur b introduit un rapport d'échelles privilégié. Le fait de choisir le paramètre \bar{k} fini se traduit par une absence théorique d'invariance d'échelle.

Outre l'absence théorique d'invariance d'échelle due à la finitude du paramètre \bar{k} , l'inconvénient principal du modèle MSM de Calvet et Fisher réside dans l'impossibilité d'estimer ce paramètre. En d'autres termes, comme l'expliquent Calvet et Fisher eux-mêmes, il existe autant de modèles que de choix de \bar{k} . A noter que Calvet et Fisher proposent d'estimer les autres paramètres par une méthode de maximum de vraisemblance. Cependant, pour ce faire, ils sont obligés de simplifier à l'extrême la loi des multiplicateurs $\{M_k[t]\}_k$ de façon à ce qu'ils aient un nombre fini d'états.

2.4.4 Modèles en cascade continue

L'idée générale de cette section est de trouver une manière de rendre continue les constructions "classiques" par cascades discrètes. L'outil principal, qui préserve une certaine intuition géométrique, consiste à intégrer une mesure aléatoire bi-dimensionnelle sur un domaine conique.

Représentation bi-dimensionnelle des cascades discrètes

Dans cette section nous allons montrer que la cascade multiplicative et la cascade dans le modèle MSM de Calvet et Fisher peuvent être obtenues par des constructions bi-dimensionnelles sur un demi-plan (t, s) , où la variable t représente le temps et la variable s représente l'échelle.

Dans le cas d'une mesure multiplicative, nous considérons un ensemble de masses aléatoires $W_{k_1 \dots k_n}$, placées aux points $(t_{k_1 \dots k_n}, s_{k_1 \dots k_n})$ avec

$$t_{k_1 \dots k_n} = \frac{1}{2^{n+1}} + \sum_{j=1}^n \frac{k_j}{2^j}, \quad (2.106)$$

$$s_{k_1 \dots k_n} = \frac{1}{2^n}. \quad (2.107)$$

où $k_1 \dots k_n$ sont des adresses diadiques. Ce qui est représenté sur la figure 2.9.

Si nous définissons le cône $\mathcal{C}(t)$ par

$$\mathcal{C}(t) = \left\{ (t', s'); 0 \leq s', |t' - t| \leq \frac{1}{2} s' \right\}, \quad (2.108)$$

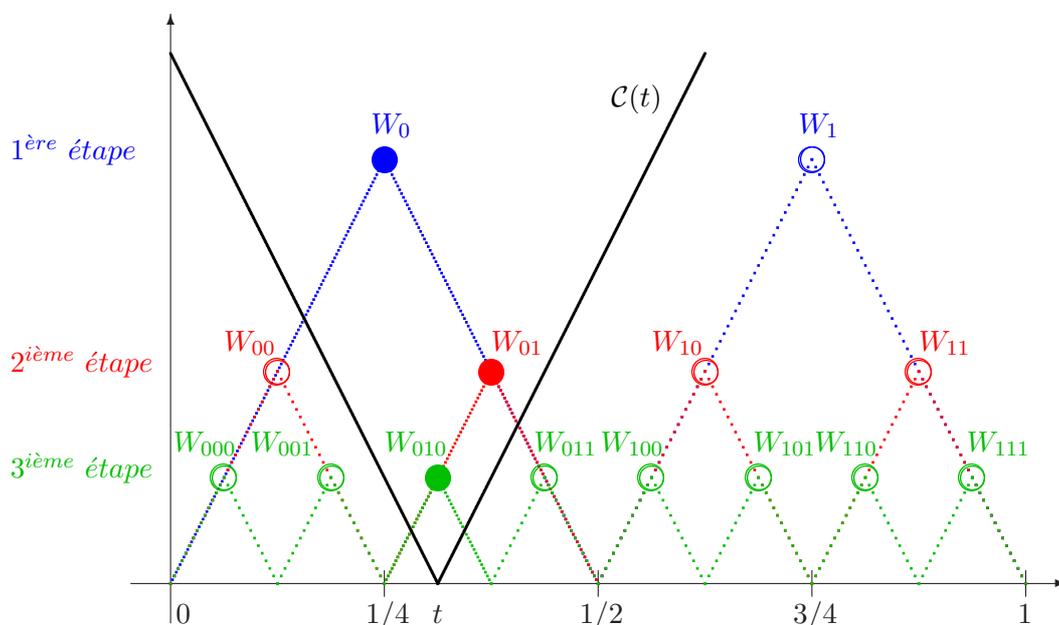


FIG. 2.9 – Représentation 2D de la cascade multiplicative de Mandelbrot

alors la valeur de la mesure multiplicative est simplement égale à

$$dM(x) = \prod_{(t_j, s_j) \in \mathcal{C}(x)} W_j = e^{\sum_{(t_j, s_j) \in \mathcal{C}(x)} \ln(W_j)}. \quad (2.109)$$

Dans le cas du modèle de Calvet et Fisher, nous considérons tout d'abord les instants de transitions $\{t_k^j\}_j$ des multiplicateurs M_k avec $1 \leq k \leq \bar{k}$. Ensuite nous considérons un ensemble de masses aléatoires $M_k[t_k^j]$, placées aux points $(t_k^j/2, t_k^j)$. La figure 2.10 représente cette construction.

En utilisant la même définition (2.108) du cône $\mathcal{C}(t)$, il est simple de voir que la volatilité stochastique du modèle de Calvet et Fisher est égale à

$$\sigma[t]^2 = \sigma^2 \prod_{(t_j, s_j) \in \mathcal{C}(t)} M_j = \sigma^2 e^{\sum_{(t_j, s_j) \in \mathcal{C}(t)} \ln(M_j)}. \quad (2.110)$$

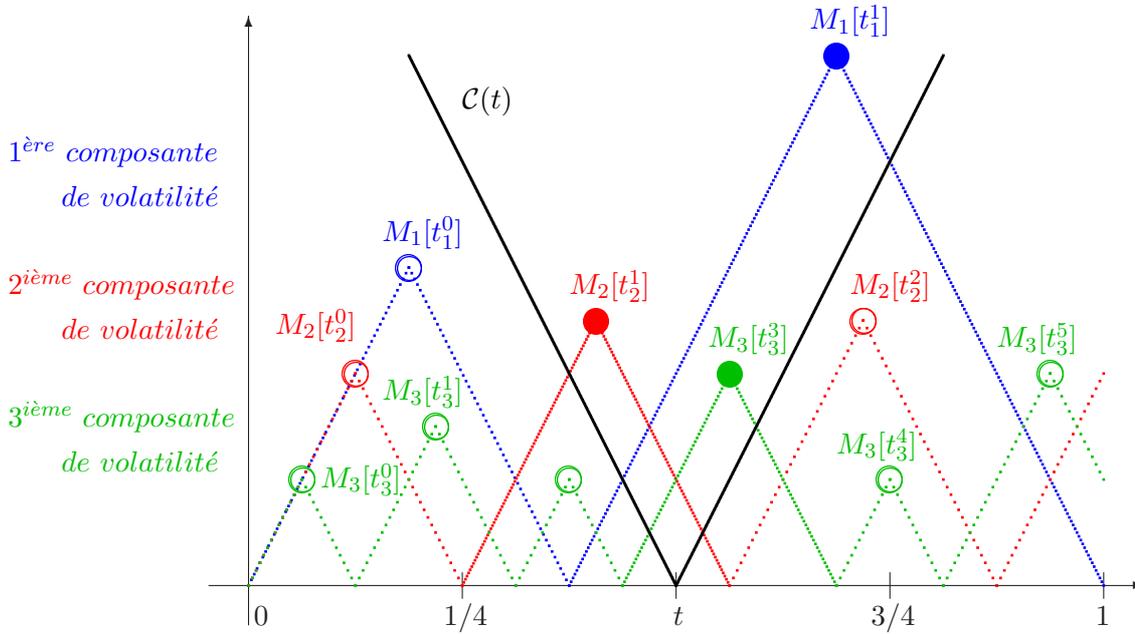


FIG. 2.10 – Représentation 2D de la cascade dans le modèle MSM de Calvet et Fisher

Cascades continues

La généralisation naturelle des représentations (2.109) et (2.110) consiste à remplacer les masses ponctuelles par une distribution continue de masse aléatoire sur un demi-plan (t, s) , ce qui conduit à une représentation intégrale

$$dM(t) = e^{\int_{C(t)} d\omega(t', s')}. \tag{2.111}$$

Cette représentation, introduite initialement par Schmitt et Marsan [SM01] utilise la notion d'un champ aléatoire dispersé indépendamment $d\omega(t, s)$ (*independently scattered random measure*) introduite par Rajput et Rosinski [RR89]. Elle est à la base de toutes les constructions de processus multifractals autosimilaires au sens stochastique connus à ce jour.

2.4.5 Modèle MPCP de Barral et Mandelbrot

La première construction d'une cascade continue a été réalisée par Barral et Mandelbrot dans [BM02]. Dans leur modèle **MPCP** (*Multifractal Product of Cylindrical Pulses*) la mesure aléatoire $d\omega(t, s)$ à tout niveau de l'échelle s est donnée par un processus de Poisson composé d'intensité $r(s) = s^{-1}$.

Il faut noter que Barral et Mandelbrot n'ont pas étudié les propriétés d'invariance d'échelle de leur construction. Ils ont focalisé leur travail sur les propriétés de régularité [Bar03]. Ils ont démontré la validité du formalisme multifractal, introduit dans la section 2.2.5, dans le cadre du modèle MPCP.

2.4.6 Modèle MRW de Bacry et Muzy

Modèle MRW log-gaussien

Les spécialistes de la modélisation financière sont plus habitués à raisonner en termes de volatilité qu'en termes de temps déformé, ce qui est le cas dans les modèles MMAR et MSM, étudiés dans les sections 2.4.2 et 2.4.3.

Dans l'article [BDM01] Bacry, Delour et Muzy ont proposé une construction du processus multifractal **MRW log-gaussien** qui n'est pas basée sur un modèle de cascade. La multifractalité y est introduite par la dynamique de la volatilité, ce qui rend le modèle plus proche, dans l'esprit, des processus ARCH (cf. section 1.3.3) ou des modèles à volatilité stochastique (cf. section 1.3.4).

Modèle MRW log-infiniment divisible

Bacry et Muzy s'inspirent [BM03] des travaux de Barral et Mandelbrot [BM02] sur les produits de "pulses" multifractals, qui font appel à un processus de Poisson, et des travaux de Schmitt et Marsan [SM01] sur les cascades multiplicative continues, pour tenter de présenter une version unifiée de la construction des différents processus multifractals. Leur modèle **MRW log-infiniment divisible** gagne en généralité puisqu'il englobe notamment le modèle MPCP de Barral et Mandelbrot [BM02] et MRW log-gaussien de Bacry, Delour et Muzy [BDM01]. De même il permet d'étudier l'existence de processus limites considérés.

Notons que Chainais, Riedi et Abry ont publié un travail [CRA03] effectué indépendamment du travail de Bacry et Muzy, dans lequel ils construisent les mesures MRM log-infiniment divisibles dans un cadre moins générale que [BM03] (notamment convergence L^2 et pas d'invariance d'échelle exacte).

Notons que les processus multifractals log-infiniment divisible font l'objet de plusieurs études en turbulence [CAP99, CLAB00, Cha06].

Dans la partie suivante, nous étudions en détail le modèle MRW log-infinitement divisible, et le cas particulier correspondant au modèle MRW log-normale.

Deuxième partie

Modèle MRW

Modèle MRW log-infiniment divisible

Ce chapitre traite des processus originalement définis dans [BM03]. Ces processus correspondent à des processus stochastiques stationnaires, multifractals qui possèdent une invariance d'échelle exacte (2.44) et qui permettent de reproduire fidèlement les faits stylisés étudiés dans les sections 1.2 et 2.3.

La construction s'effectue en deux temps. Dans un premier temps (cf. définition 3.3), la mesure MRM log-infiniment divisible est définie au moyen d'un champs log-infiniment divisible défini sur le demi-plan supérieur de \mathbb{R}^2 et d'une famille de cônes sur ce demi-plan. Cette construction peut être vue comme une généralisation continue des cascades discrètes présentées dans la section 2.4.1 du chapitre précédent. Cette mesure existe (cf. théorème 3.1), n'est pas dégénérée (cf. théorème 3.3), est multifractale et possède une invariance d'échelle exacte (cf. proposition 3.7).

Dans un second temps, la subordination du mouvement brownien par la mesure MRM log-infiniment divisible, comme dans le cadre du modèle MMAR introduit dans la section 2.4.2, permet de définir le processus MRW log-infiniment divisible (cf. définition 3.4). C'est à travers la subordination que le processus MRW acquiert ses propriétés de multifractalité.

Après un exposé relativement détaillé de la construction de la mesure MRM et du processus MRW log-infiniment divisibles, nous abordons ses propriétés principales. Notre contribution originale se résume aux résultats présentés dans les sections 3.1.3 et 3.3 qui concernent les propriétés d'existence des moments d'ordre négatifs de la mesure MRM et d'invariance d'échelle par changement du temps intégral.

Dans la section 3.4.1, le cas particulier du modèle MRW log-normale est introduit en utilisant la construction effective unidimensionnelle décrite dans [BDM01]. C'est à l'aide de cette construction que nous pouvons calculer des moments d'ordres entiers des accroissements du processus MRW log-normal (cf. équations (3.130) et (3.131)).

A la fin de ce chapitre, nous étudierons en détails (cf. section 3.4.4) les faits stylisés que le modèle MRW log-normal permet de reproduire.

3.1 Mesure multifractale MRM

3.1.1 Construction et non dégénérescence de la mesure MRM

Champ aléatoire infiniment divisible

Le champ aléatoire infiniment divisible étudié dans cette section a été introduit par Rajput et Rosinski, nous suivons donc essentiellement leur article [RR89].

Nous considérons l'espace mesuré $(\mathcal{S}^+, \mu(dt, dl))$, où \mathcal{S}^+ est le demi-plan

$$\mathcal{S}^+ = \{(t, l); t \in \mathbb{R}, l \in \mathbb{R}^+\} \quad (3.1)$$

muni de la mesure, dite mesure de Haar à gauche (invariante par le groupe translations-dilatations),

$$\mu(dt, dl) = l^{-2} dt dl. \quad (3.2)$$

Nous noterons $\mathcal{B}(\mathcal{S}^+, \mu)$, l'ensemble des parties \mathcal{A} du demi-plan \mathcal{S}^+ qui sont μ -mesurables.

Définition 3.1

Le **champ aléatoire infiniment divisible** \mathcal{P} défini sur le demi-plan $(\mathcal{S}^+, \mu(dt, dl))$ et associé à la mesure de Lévy $\nu(dx)$ est défini comme l'application de $\mathcal{B}(\mathcal{S}^+, \mu)$ dans \mathbb{R} qui vérifie les propriétés suivantes

– pour tout élément \mathcal{A} de $\mathcal{B}(\mathcal{S}^+, \mu)$, la variable $\mathcal{P}(\mathcal{A})$ est une variable aléatoire infiniment divisible dont la fonction caractéristique est donnée par

$$\mathbb{E} \left[e^{iq\mathcal{P}(\mathcal{A})} \right] = e^{\varphi(q)\mu(\mathcal{A})}, \quad (3.3)$$

où la fonction $\varphi(q)$ est égale au logarithme de la fonction caractéristique associée à la mesure de Lévy $\nu(dx)$, par la formule de Lévy-Khintchine (2.4)

$$\varphi(q) = imq + \int_{-\infty}^{+\infty} \frac{e^{iqx} - 1 - iq \sin x}{x^2} \nu(dx), \quad (3.4)$$

– pour toute famille $\{\mathcal{A}_n\}_n$ d'éléments disjoints de $\mathcal{B}(\mathcal{S}^+, \mu)$, les variables $\{\mathcal{P}(\mathcal{A}_n)\}_n$ sont des variables aléatoires indépendantes qui vérifient

$$\mathcal{P} \left(\bigcup_{n=1}^{\infty} \mathcal{A}_n \right) = \sum_{n=1}^{\infty} \mathcal{P}(\mathcal{A}_n) \text{ p.s.} \quad (3.5)$$

Il est possible d'introduire la fonction convexe $\psi(q)$ telle que pour tout q réel positif et pour tout élément \mathcal{A} de $\mathcal{B}(\mathcal{S}^+, \mu)$, nous avons

$$\begin{cases} \psi(q) = +\infty, & \text{si } \mathbb{E} [e^{q\mathcal{P}(\mathcal{A})}] = +\infty, \\ \mathbb{E} [e^{q\mathcal{P}(\mathcal{A})}] = e^{\mu(\mathcal{A})\psi(q)}, & \text{sinon.} \end{cases} \quad (3.6)$$

De même, si l'on définit \tilde{q} par

$$\tilde{q} = \sup_q \{q \geq 0; \psi(q) < +\infty\}, \quad (3.7)$$

il est possible de réaliser un prolongement dans le plan complexe de la fonction $\varphi(q)$, défini sur l'axe réel par (3.3), de telle façon que

$$\psi(q) = \varphi(-iq), \quad (3.8)$$

pour tout complexe q tel que $0 \leq \Re[q] < \tilde{q}$.

Pour conclure nous définissons la filtration \mathcal{F}_l de l'espace de probabilité Ω sur lequel le champ aléatoire \mathcal{P} est défini

$$\mathcal{F}_l = \sigma\{\mathcal{P}(dt, dl'); l' \geq l\}. \quad (3.9)$$

Cônes et mesure MRM

Définition 3.2

Pour tout l et T réels positifs tels que $0 < l < T$, nous définissons le cône $\mathcal{A}_{l,T}(t)$ par

$$\mathcal{A}_{l,T}(t) = \left\{ (t', l'); l \leq l', |t' - t| \leq \frac{1}{2} \min(l', T) \right\}. \quad (3.10)$$

C'est à l'aide de ces cônes que nous définissons la mesure MRM (*Multifractal Random Measure*) log-infinitement divisible.

Définition 3.3

Soit T réel positif. Le processus $\{\omega_{l,T}(t); t \in \mathbb{R}\}$ est défini par

$$\omega_{l,T}(t) = \frac{1}{2} \mathcal{P}(\mathcal{A}_{l,T}(t)), \quad (3.11)$$

où $\{\mathcal{A}_{l,T}(t)\}_t$ la famille de cônes donnée par (3.10) et \mathcal{P} le champ aléatoire infinitement divisible défini sur le demi-plan \mathcal{S}^+ (cf. définition 3.1).

La mesure MRM $M_T(dt)$ est définie comme la limite des mesures $M_{l,T}(dt)$ lorsque l tend vers 0, si elle existe

$$M_T(dt) = \lim_{l \rightarrow 0^+} M_{l,T}(dt), \quad (3.12)$$

où la mesure $M_{l,T}(dt)$ est donnée par

$$M_{l,T}(dt) = e^{2\omega_{l,T}(t)} dt, \text{ pour } l > 0. \quad (3.13)$$

Ainsi, pour tout ensemble I Lebesgue-mesurable, on a

$$M_{l,T}(I) = \int_I e^{2\omega_{l,T}(t)} dt. \quad (3.14)$$

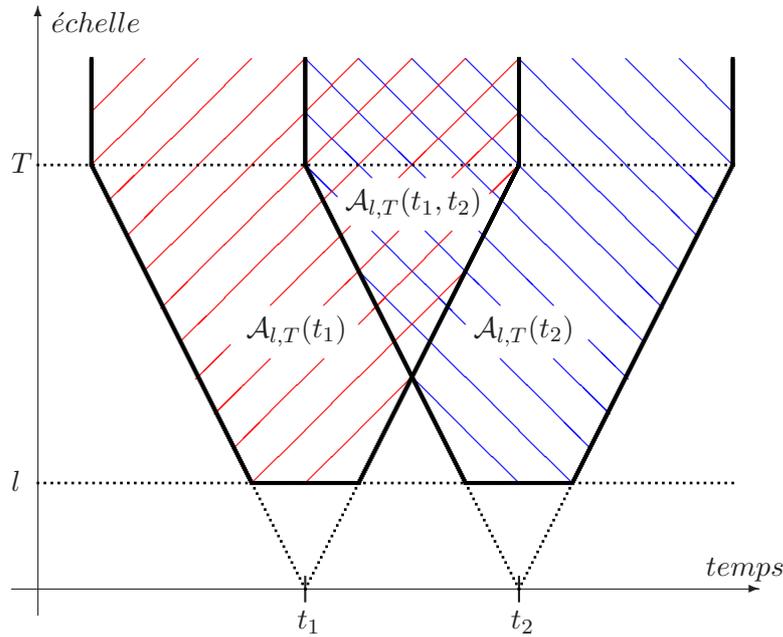


FIG. 3.1 – Cônes $\mathcal{A}_{l,T}(t_1)$, $\mathcal{A}_{l,T}(t_2)$ et $\mathcal{A}_{l,T}(t_1, t_2)$ intervenant dans la construction du modèle MRM log-infiniment divisible (cf. définition 3.3). Le paramètre T joue le rôle du temps intégral. La mesure MRM est obtenue à limite lorsque l tend vers zéro.

On peut remarquer que l’addition au champ \mathcal{P} d’une densité uniforme mène à la même mesure MRM à un facteur près. Ainsi, sans perte de généralité, on peut supposer que la fonction $\psi(q)$, définie par (3.6), satisfait

$$\psi(1) = 0. \tag{3.15}$$

Ce qui nous mène à la représentation suivante de la fonction $\psi(q)$ à l’aide de la formule (3.4)

$$\psi(q) = \int_{-\infty}^{+\infty} \frac{e^{qx} - 1 - q(e^x - 1)}{x^2} \nu(dx) \tag{3.16}$$

Le théorème suivant démontre l’existence de la mesure MRM comme limite des mesures (3.13).

Théorème 3.1 (Bacry-Muzy)

- la suite des mesures $M_{l,T}(dt)$ converge faiblement vers la mesure $M_T(dt)$ lorsque l tend vers zéro avec la probabilité 1,
- pour tout t réel, $M_T(\{t\}) = 0$,
- pour tout ensemble K borné dans \mathbb{R} , $M_T(K) < +\infty$ et $\mathbb{E}[M_T(K)] \leq |K|$.

Preuve - 3.1 -

De l'identité $\psi(1) = 0$ on déduit que $\mathbb{E}[e^{2\omega_{l,T}(t)}] = 1$. Il est facile de montrer que pour tout ensemble I mesurable par rapport à une mesure de Lebesgue la suite $\{M_{l,T}(I)\}_l$ est une martingale, par rapport à la filtration \mathcal{F}_l définie par (3.9), positive et continue à gauche. En utilisant la théorie générale de [Kah87], on démontre le théorème 3.1. ■

Non dégénérescence de la mesure MRM

Proposition 3.2 (Bacry-Muzy)

Soit $q > 0$ tel que $\zeta_M(q) > 1$, alors $\sup_l \mathbb{E} [M_{l,T}([0, t])^q] < +\infty$. La mesure $M_{l,T}$ est alors uniformément intégrable.

Preuve - 3.2 -

La démonstration est détaillée dans [BM03], inspirée des travaux [BM02, KP76]. ■

D'après ce résultat, la suite $\{M_{l,T}([0, t])^q\}_l$ est uniformément intégrable s'il existe $\epsilon > 0$ tel que $\zeta_M(q + \epsilon) > 1$. En effet, pour $x > 0$, on a

$$\begin{aligned} x^\epsilon \mathbb{E} \left[M_{l,T}([0, t])^q \mathbf{1}_{\{M_{l,T}([0, t]) > x\}} \right] &\leq \mathbb{E} \left[M_{l,T}([0, t])^{q+\epsilon} \mathbf{1}_{\{M_{l,T}([0, t]) > x\}} \right] \\ &\leq \mathbb{E} [M_{l,T}([0, t])^{q+\epsilon}] \leq \sup_l \mathbb{E} [M_{l,T}([0, t])^{q+\epsilon}] < +\infty \end{aligned} \quad (3.17)$$

et donc

$$\lim_{x \rightarrow \infty} \mathbb{E} \left[M_{l,T}([0, t])^q \mathbf{1}_{\{M_{l,T}([0, t]) > x\}} \right] \leq \lim_{x \rightarrow \infty} \frac{1}{x^\epsilon} \sup_l \mathbb{E} [M_{l,T}([0, t])^{q+\epsilon}] = 0. \quad (3.18)$$

Théorème 3.3 (Bacry-Muzy)

Soient $q > 0$ et $\epsilon > 0$ tels que $\zeta_M(q + \epsilon) > 1$, alors

$$\lim_{l \rightarrow 0^+} \mathbb{E}[M_{l,T}([0, t])^q] = \mathbb{E}[M_T([0, t])^q]. \quad (3.19)$$

Preuve - 3.3 -

La suite $\{M_{l,T}([0, t])^q\}_l$ converge en loi vers $M_T([0, t])^q$ et la condition $\zeta_M(q + \epsilon) > 1$ entraîne l'intégralité uniforme de $\{M_{l,T}([0, t])^q\}_l$. On peut donc passer à la limite sous le signe espérance. La démonstration est détaillée dans [Bil68]. ■

Dans le cas particulier du théorème 3.3, où l'ordre du moment q est égal à 1, nous avons

$$\lim_{l \rightarrow 0^+} \mathbb{E}[M_{l,T}([0, t])] = \mathbb{E}[M_T([0, t])] = t, \quad (3.20)$$

ce qui signifie que la mesure MRM log-infinitement divisible n'est pas dégénérée.

3.1.2 Propriétés de la mesure MRM

Dans cette section nous étudions les propriétés d'invariance d'échelle de la mesure MRM et les conditions d'existence de ses moments d'ordre q . Pour cela, nous devons d'abord étudier plus en profondeur le processus $\{\omega_{l,T}(t)\}_t$.

Fonction caractéristique du processus $\{\omega_{l,T}(t)\}_t$

Rappelons que le processus $\{\omega_{l,T}(t)\}_t$ est défini par la relation (3.11).

La relation (3.3) permet d'obtenir l'expression analytique des quantités de la forme

$$\mathbb{E} \left[e^{\sum_{m=1}^n p_m \mathcal{P}(\mathcal{A}_m)} \right],$$

où les $\{\mathcal{A}_m\}_m$ sont des sous-ensembles disjoints de \mathcal{S}^+ et $\{p_m\}_m$ sont des nombres complexes arbitraires. Dans le cas général, où les sous-ensembles $\{\mathcal{A}_m\}_m$ ne sont pas disjoints, il est alors naturel d'introduire des intersections

$$\mathcal{A}_{l,T}(t, t') = \mathcal{A}_{l,T}(t) \cap \mathcal{A}_{l,T}(t') \quad (3.21)$$

qui permettent de faire une décomposition de $\{\mathcal{A}_m\}_m$ en des sous-ensembles disjoints.

Proposition 3.4 (Bacry-Muzy)

Pour tout $n \in \mathbb{N}$, pour tous points $t_1 \leq \dots \leq t_n$ et toutes valeurs p_1, \dots, p_n la fonction caractéristique du vecteur $\{\omega_{l,T}(t_m)\}_{1 \leq m \leq n}$ s'écrit comme

$$\mathbb{E} \left[e^{\sum_{m=1}^n i p_m \omega_{l,T}(t_m)} \right] = e^{\sum_{k=1}^n \sum_{j=1}^k \alpha(k,j) \rho_{l,T}(t_j - t_k)}, \quad (3.22)$$

où la fonction $\rho_{l,T}(t)$ est donnée par

$$\rho_{l,T}(t) = \mu(\mathcal{A}_{l,T}(0, t)) \quad (3.23)$$

et les coefficients $\alpha(k, j)$ sont définis par

$$\alpha(k, j) = \varphi(r_{j,k}) + \varphi(r_{j+1,k+1}) - \varphi(r_{j,k-1}) - \varphi(r_{j+1,k}), \quad (3.24)$$

avec

$$r_{j,k} = \begin{cases} \frac{1}{2} \sum_{m=j}^k p_m, & \text{pour } j \leq k, \\ 0, & \text{pour } j > k. \end{cases} \quad (3.25)$$

De plus, on a l'identité suivante

$$\sum_{k=1}^n \sum_{j=1}^k \alpha(k, j) = \varphi\left(\frac{1}{2} \sum_{k=1}^n p_k\right). \quad (3.26)$$

Preuve - 3.4 -

La preuve de cette proposition 3.4 est très calculatoire et est détaillée dans [BM03]. ■

Propriétés d'invariance d'échelle de la fonction $\rho_{l,T}(t)$

Les propriétés d'invariance d'échelle de la mesure MRM reposent sur le comportement de la fonction $\rho_{l,T}(t)$, définie par l'équation (3.23), lorsque l'on varie l'échelle l et le paramètre T .

Le calcul direct de la fonction $\rho_{l,T}(t)$ donne

$$\rho_{l,T}(t) = \begin{cases} \ln\left(\frac{T}{l}\right) + 1 - \frac{t}{l}, & \text{si } 0 < t \leq l, \\ \ln\left(\frac{T}{t}\right), & \text{si } l < t \leq T, \\ 0, & \text{si } T < t. \end{cases} \quad (3.27)$$

Il est simple de voir que $\rho_{l,T}(t)$ satisfait les relations suivantes :

$$\rho_{\lambda l,T}(\lambda t) = \rho_{l,T}(t) - \ln \lambda, \quad \text{pour } t \leq T \text{ et } \lambda \in (0, 1), \quad (3.28)$$

$$\rho_{l,T}(t) = \rho_{l,\lambda T}(t) - \ln \lambda, \quad \text{pour } t \leq \lambda T \text{ et } \lambda \in (0, 1). \quad (3.29)$$

$$\rho_{\lambda l,\lambda T}(\lambda t) = \rho_{l,T}(t), \quad \text{pour tout } \lambda \text{ positif}. \quad (3.30)$$

Ces relations permettent d'établir les diverses propriétés d'invariance d'échelle de la mesure $M_{l,T}(t)$ puis celles de la mesure MRM limite.

Propriété d'autosimilarité stochastique de la mesure $M_{l,T}(t)$

Proposition 3.5 (Bacry-Muzy)

La mesure $M_{l,T}(dt)$ vérifie la propriété d'autosimilarité stochastique, pour $\lambda \in (0, 1)$

$$\{M_{\lambda l,T}([0, \lambda t])\}_t \stackrel{\mathcal{L}}{=} W_\lambda \{M_{l,T}([0, t])\}_t, \quad \text{pour } t < T, \quad (3.31)$$

avec $W_\lambda = \lambda e^{2\Omega_\lambda}$, où Ω_λ est une variable aléatoire infiniment divisible dont la fonction caractéristique est donnée par

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.32)$$

Preuve - 3.5 -

En utilisant la proposition 3.4 et la relation (3.28), on a

$$\mathbb{E} \left[e^{\sum_k ip_k \omega_{\lambda l,T}(\lambda t_k)} \right] = \lambda^{-\varphi\left(\frac{1}{2} \sum_k p_k\right)} \mathbb{E} \left[e^{\sum_k ip_k \omega_{l,T}(t_k)} \right] = \mathbb{E} \left[e^{\sum_k ip_k (\Omega_\lambda + \omega_{l,T}(t_k))} \right]. \quad (3.33)$$

où la variable aléatoire Ω_λ est indépendante du processus $\{\omega_l(t)\}_t$ et sa fonction caractéristique s'écrit comme

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.34)$$

Finalement, il reste à calculer l'intégrale

$$\begin{aligned} \{M_{\lambda l, T}([0, \lambda t])\}_t &= \left\{ \int_0^{\lambda t} e^{2\omega_{\lambda l, T}(u)} du \right\}_t = \lambda \left\{ \int_0^t e^{2\omega_{\lambda l, T}(\lambda u)} du \right\}_t \\ &\stackrel{\mathcal{L}}{=} \lambda e^{2\Omega_\lambda} \left\{ \int_0^t e^{2\omega_{l, T}(u)} du \right\}_t = W_\lambda \{M_{l, T}([0, t])\}_t \end{aligned} \quad (3.35)$$

où les variables W_λ et Ω_λ sont liées par la relation $W_\lambda = \lambda e^{2\Omega_\lambda}$. ■

Invariance d'échelle exacte, multifractalité et autosimilarité stochastique de $M_T(dt)$

Avant d'énoncer le théorème général, nous démontrons une propriété d'invariance d'échelle de la mesure MRM dont nous aurons besoin par la suite.

Proposition 3.6

La mesure $M_T(dt)$ vérifie la propriété d'invariance suivante

$$\{M_{\lambda T}([0, \lambda t])\}_t \stackrel{\mathcal{L}}{=} \lambda \{M_T([0, t])\}_t, \quad (3.36)$$

pour tout λ positif.

Preuve - 3.6 -

En utilisant la propriété (3.30), vérifiée par la fonction $\rho_{l, T}(t)$, et la proposition 3.4, on a

$$\mathbb{E} \left[e^{\sum_k i p_k \omega_{\lambda l, \lambda T}(\lambda t_k)} \right] = \mathbb{E} \left[e^{\sum_k i p_k \omega_{l, T}(t_m)} \right]. \quad (3.37)$$

Le calcul de l'intégrale (3.14) conduit à

$$\begin{aligned} \{M_{\lambda l, \lambda T}([0, \lambda t])\}_t &= \left\{ \int_0^{\lambda t} e^{2\omega_{\lambda l, \lambda T}(u)} du \right\}_t = \lambda \left\{ \int_0^t e^{2\omega_{\lambda l, \lambda T}(\lambda u)} du \right\}_t \\ &\stackrel{\mathcal{L}}{=} \lambda \left\{ \int_0^t e^{2\omega_{l, T}(u)} du \right\}_t = \lambda \{M_{l, T}([0, t])\}_t. \end{aligned} \quad (3.38)$$

Pour démontrer la relation (3.36) il suffit de passer à la limite $l \rightarrow 0$. ■

Théorème 3.7 (Bacry-Muzy)

La mesure MRM vérifie la propriété d'autosimilarité stochastique (2.49), pour $\lambda \in (0, 1)$

$$\{M_T([0, \lambda t])\}_t \stackrel{\mathcal{L}}{=} W_\lambda \{M_T([0, t])\}_t, \text{ pour } t \leq T, \quad (3.39)$$

avec $W_\lambda = \lambda e^{2\Omega_\lambda}$, où Ω_λ est une variable aléatoire infiniment divisible dont la fonction caractéristique est donnée par

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.40)$$

De plus, pour tout $t \leq T$, s'il existe $\epsilon > 0$ tel que $\zeta_M(q + \epsilon) > 1$, en d'autres termes, si le moment d'ordre q existe, alors la mesure MRM possède une invariance d'échelle exacte

$$\mathbb{E} [M_T([0, t])^q] = \left(\frac{t}{T}\right)^{\zeta_M(q)} \mathbb{E} [M_T([0, T])^q], \text{ pour } q > 0 \quad (3.41)$$

s'il existe $\epsilon > 0$ tel que $\zeta_M(q + \epsilon) > 1$. De plus, si la mesure de Lévy $\nu(dx)$ n'est pas identiquement nulle alors $\zeta_M(q)$ n'est pas linéaire et la mesure M_T est multifractale.

Preuve - 3.7 -

Le résultat est une conséquence directe de la proposition 3.5. Il suffit de calculer l'espérance

$$\mathbb{E} [M_{\lambda, T}([0, t])^q] = \mathbb{E} [W_\lambda^q] \mathbb{E} [M_{l, T}([0, T])^q] = \lambda^{q-\psi(q)} \mathbb{E} [M_{l, T}([0, T])^q], \quad (3.42)$$

avec $\lambda = t/T$ et passer à la limite sous le signe espérance d'après le théorème 3.3. ■

Finitude des moments de $M_T(dt)$. Exposant de queue

Théorème 3.8 (Bacry-Muzy)

1. *Soit $q > 0$. Si $\zeta_M(q) > 1$ alors $\mathbb{E} [M_T([0, t])^q] < +\infty$.*
2. *Soit $q > 1$. Si $M_T \neq 0$ alors $\mathbb{E} [M_T([0, t])^q] < +\infty$ entraîne $\zeta_M(q) \geq 1$.*

Preuve - 3.8 -

La preuve de la première affirmation est relativement technique et se trouve dans [BM03]. C'est essentiellement une conséquence de la proposition 3.2 et du fait que $M_{l, T}([0, t])^q$ est une sous-martingale par rapport à la filtration \mathcal{F}_l définie par (3.9).

Pour démontrer la deuxième affirmation notons que si la mesure MRM est non dégénérée $M \neq 0$ et si le moment d'ordre q positif est fini $\mathbb{E} [M_T([0, t])^q] < +\infty$ alors en utilisant le théorème 3.7, il est simple de voir que $\psi(q) \neq +\infty$ ou $\zeta_M(q) \neq -\infty$. En divisant l'intervalle $[0, t]$ en deux sous-intervalles, on a

$$\mathbb{E} [M_T([0, t])^q] \geq \mathbb{E} [M_T([0, t/2])^q] + \mathbb{E} [M_T([t/2, t])^q] = 2\mathbb{E} [M_T([0, t/2])^q]. \quad (3.43)$$

L'application de l'invariance d'échelle exacte (3.41) mène à

$$\mathbb{E} [M_T([0, t])^q] \geq 2^{1-\zeta_M(q)} \mathbb{E} [M_T([0, t])^q], \quad (3.44)$$

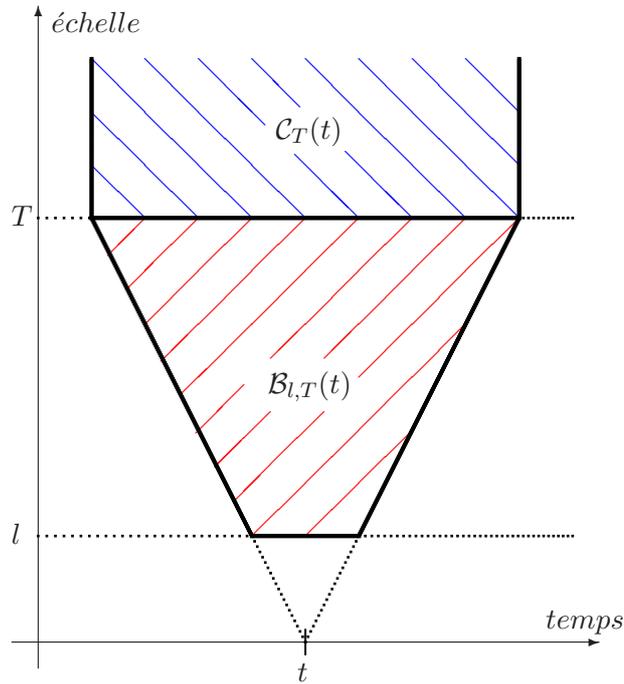


FIG. 3.2 – Cône $\mathcal{B}_{l,T}(t)$ et domaine $\mathcal{C}_T(t)$ qui interviennent dans la démonstration de l'existence des moments d'ordres négatifs de la mesure MRM log-infiniment divisible.

et en conséquence $\zeta_M(q) \geq 1$. ■

D'après la proposition 1.1 l'exposant de queue q_c de la mesure $M_T([0, t])$ est donné par la solution de l'équation $\zeta_M(q_c) = 1$ telle que $q_c > 1$.

3.1.3 Moments d'ordre négatif de la mesure MRM

Les résultats de cette section sont originaux.

Dans cette section, nous allons montrer que si le coefficient d'intermittence λ^2 est assez petit, certains moments d'ordre négatif de la mesure MRM log-infiniment divisible sont finis.

Notations

On décompose (cf. figure 3.2) le cône $\mathcal{A}_{l,T}(t)$ en deux parties : $\mathcal{A}_{l,T}(t) = \mathcal{B}_{l,T}(t) \cup \mathcal{C}_T(t)$ où $\mathcal{B}_{l,T}(t)$ désigne le cône centré en t , qui s'arrête en l aux petites échelles et en T aux

grandes échelles

$$\mathcal{B}_{l,T}(t) = \left\{ (t', l'); l \leq l' \leq T, |t' - t| \leq \frac{l'}{2} \right\} \quad (3.45)$$

et où $\mathcal{C}_T(t)$ désigne le complémentaire :

$$\mathcal{C}_T(t) = \left\{ (t', l'); T \leq l', |t' - t| \leq \frac{T}{2} \right\}. \quad (3.46)$$

On pose

$$\eta_{l,T}(t) = \frac{1}{2} \mathcal{P}(\mathcal{B}_{l,T}(t)). \quad (3.47)$$

et

$$\delta_T(t) = \frac{1}{2} \mathcal{P}(\mathcal{C}_T(t)). \quad (3.48)$$

On a donc $\omega_{l,T}(t) = \eta_{l,T}(t) + \delta_T(t)$. Nous définissons également la mesure $N_T(dt)$ comme la limite des mesures $N_{l,T}(dt)$

$$N_T(dt) = \lim_{l \rightarrow 0^+} N_{l,T}(dt), \quad (3.49)$$

où la mesure $N_{l,T}(dt)$ est donnée par

$$N_{l,T}(dt) = e^{2\eta_{l,T}(t)} dt, \text{ pour } l > 0 \quad (3.50)$$

ou au sens de Lebesgue, pour tout ensemble mesurable I , on a

$$N_{l,T}(I) = \int_I e^{2\eta_{l,T}(t)} dt. \quad (3.51)$$

Existence des moments d'ordres négatifs

La proposition suivante relie l'existence d'un moment de la mesure $M_T(dt)$ à celui de la mesure $N_T(dt)$.

Proposition 3.9

Pour tout $q \in \mathbb{R}$ tel que $\psi(q) \neq +\infty$, l'existence du moment d'ordre q de la mesure $N_T([0, t])$ entraîne l'existence de celui de la mesure $M_T([0, t])$, où t est fixé,

$$\mathbb{E}[N_T([0, t])^q] < \infty \implies \mathbb{E}[M_T([0, t])^q] < \infty. \quad (3.52)$$

Preuve - 3.9 -

Dans le lemme 4 de [BM03], ce résultat est démontré pour $q \geq 0$. Démontrons le pour q négatif. Nous avons, directement d'après les définitions ci-dessus,

$$M_T([0, t]) \geq N_T([0, t]) \inf_{u \in [0, t]} e^{\delta_T(u)} \quad (3.53)$$

Remarquons, comme dans la démonstration du lemme 5 de [BM03], que pour $u \in [0, t]$, nous pouvons décomposer $\delta_T(u)$ en trois termes

$$\delta_T(u) = \delta_T^0(t) + \delta_T^g(u, t) + \delta_T^d(u, t), \quad (3.54)$$

où $\delta_T^0(t)$, $\delta_T^g(u, t)$ et $\delta_T^d(u, t)$ sont des variables aléatoires indépendantes définies par

$$\delta_T^0(t) = \frac{1}{2} \mathcal{P}(\mathcal{C}_T^0(t)), \quad \text{où } \mathcal{C}_T^0(t) = \bigcap_{u \in [0, t]} \mathcal{C}_T(u), \quad (3.55)$$

$$\delta_T^g(u, t) = \frac{1}{2} \mathcal{P}(\mathcal{C}_T^g(u, t)), \quad \text{où } \mathcal{C}_T^g(u, t) = (\mathcal{C}_T(u) \cap \mathcal{C}_T(0)) \setminus \mathcal{C}_T^0(t), \quad (3.56)$$

$$\delta_T^d(u, t) = \frac{1}{2} \mathcal{P}(\mathcal{C}_T^d(u, t)), \quad \text{où } \mathcal{C}_T^d(u, t) = (\mathcal{C}_T(u) \cap \mathcal{C}_T(t)) \setminus \mathcal{C}_T^0(t). \quad (3.57)$$

Ce qui permet de déduire

$$M_T([0, t]) \geq N_T([0, t]) e^{2\delta_T^0(t)} \inf_{u \in [0, t]} e^{2\delta_T^g(u, t)} \inf_{u \in [0, t]} e^{2\delta_T^d(u, t)} \quad (3.58)$$

et donc pour $q < 0$

$$M_T([0, t])^q \leq N_T([0, t])^q e^{2q\delta_T^0(t)} \sup_{u \in [0, t]} e^{2q\delta_T^g(u, t)} \sup_{u \in [0, t]} e^{2q\delta_T^d(u, t)}. \quad (3.59)$$

Les variables $e^{2\delta_T^g(u, t)}$ et $e^{2\delta_T^d(t-u, t)}$ sont des martingales par rapport à u . Comme la fonction $f(x) = 1/x$ est convexe $e^{-2\delta_T^g(u, t)}$ et $e^{-2\delta_T^d(t-u, t)}$ sont des sous-martingales. D'après le théorème de Doob sur les moments du supremum de sous-martingale [Doo01]

$$\mathbb{E} [M_T([0, t])^q] \leq \mathbb{E} [N_T([0, t])^q] \mathbb{E} [e^{2q\delta_T^0(t)}] \mathbb{E} [e^{2q\delta_T^g(t, t)}] \mathbb{E} [e^{2q\delta_T^d(0, t)}], \quad (3.60)$$

ce qui implique directement la proposition. ■

Pour montrer la finitude des moments négatifs de $M_T([0, t])$ il suffit donc de montrer la finitude des moments négatifs de $N_T([0, t])$. Avant d'énoncer le théorème principal 3.11 nous avons besoin de la proposition suivante

Proposition 3.10

Pour tout $t < T$, on a presque sûrement la minoration suivante

$$N_T([0, t]) \geq Y_1 N_1 + Y_2 N_2, \quad (3.61)$$

où les variables Y_1 et Y_2 sont de mêmes loi que Y , indépendantes de N_1 et N_2 et vérifient

$$\psi(-q) < +\infty \text{ implique } \mathbb{E} [Y^{-q}] < +\infty, \text{ pour } q > 0, \quad (3.62)$$

et où N_1 et N_2 sont indépendantes et de même loi que $N_T([0, t])$.

Preuve - 3.10 -

Remarquons tout d'abord que

$$\eta_{sl,sT}(st) \stackrel{\mathcal{L}}{=} \eta_{l,T}(t), \text{ pour tout } s, \quad (3.63)$$

ce qui peut se voir par un simple changement de variable dans l'intégrale qui définit η . Nous en déduisons que

$$N_{sl,sT}([0, st]) \stackrel{\mathcal{L}}{=} sN_{l,T}([0, t]) \quad (3.64)$$

et donc en passant à la limite $l \rightarrow 0^+$

$$N_{sT}([0, st]) \stackrel{\mathcal{L}}{=} sN_T([0, t]). \quad (3.65)$$

Pour s positif inférieur à $t/2$ la mesure $N_{l,T}$ vérifie l'inégalité suivante

$$N_{l,T}([0, t]) \geq N_{l,T}([0, s]) + N_{l,T}([t-s, t]). \quad (3.66)$$

Les frontières des domaines coniques $\mathcal{B}_{l,T}(s)$ et $\mathcal{B}_{l,T}(t-s)$ s'intersectent au point $(t/2, t-2s)$. Nous décomposons alors la variable $\eta_{l,T}(u)$ en deux termes

$$\eta_{l,T}(u) = \eta_{l,t-2s}(u) + \eta_{t-2s,T}(u), \quad (3.67)$$

où $\eta_{l,t-2s}(u)$ et $\eta_{t-2s,T}(u)$ sont indépendants. De plus, $\eta_{l,t-2s}(u)$ pour $u \in [0, s]$ est indépendant de $\eta_{l,t-2s}(v)$ pour $v \in [t-s, t]$. En conséquence,

$$N_{l,T}([0, t]) \geq \inf_{u \in [0, s]} e^{\eta_{t-2s,T}(u)} N_{l,t-2s}([0, s]) + \inf_{u \in [t-s, t]} e^{\eta_{t-2s,T}(u)} N_{l,t-2s}([t-s, t]). \quad (3.68)$$

et en passant à la limite $l \rightarrow 0^+$, nous avons

$$N_T([0, t]) \geq \inf_{u \in [0, s]} e^{\eta_{t-2s,T}(u)} N_{t-2s}([0, s]) + \inf_{u \in [t-s, t]} e^{\eta_{t-2s,T}(u)} N_{t-2s}([t-s, t]). \quad (3.69)$$

En prenant $s = \frac{t^2}{T+2t}$, nous avons la minoration souhaitée (3.61) avec

$$N_1 = \frac{T+2t}{t} N_{t-2s}([0, s]), \quad (3.70)$$

$$N_2 = \frac{T+2t}{t} N_{t-2s}([t-s, t]), \quad (3.71)$$

$$Y_1 = \frac{t}{T+2t} \inf_{u \in [0, s]} e^{\eta_{t-2s,T}(u)}, \quad (3.72)$$

$$Y_2 = \frac{t}{T+2t} \inf_{u \in [t-s, t]} e^{\eta_{t-2s,T}(u)}, \quad (3.73)$$

grâce à l'identité $\frac{s}{t} = \frac{t-2s}{T}$ et l'égalité en loi (3.65). Par un argument de sous-martingale comme dans la proposition 3.9, nous avons de plus l'implication (3.62). ■

En appliquant la minoration (3.61) une deuxième fois à N_1 et N_2 , nous avons

$$N_T([0, t]) \geq Y_1 N_1 + Y_2 N_2 + Y_3 N_3 + Y_4 N_4, \quad (3.74)$$

où les variables $\{N_j\}_{j=1}^4$ sont indépendantes entre elles et de même loi que $N_T([0, t])$ et les variables $\{Y_j\}_{j=1}^4$ sont de même loi et indépendantes de $\{N_j\}_{j=1}^4$.

Nous sommes maintenant prêts à énoncer le théorème principal de cette section qui va donner une condition d'existence des moments d'ordre négatif.

Théorème 3.11

Supposons qu'il existe λ_0^2 et q_0 positifs tels que le moment d'ordre négatif $\mathbb{E}[Y^{-q_0}]$ est borné uniformément en le coefficient d'intermittence $\lambda^2 < \lambda_0^2$. De plus, supposons que la mesure $N_T(dt)$ converge vers la mesure de Lebesgue uniformément lorsque le coefficient d'intermittence λ^2 tend vers 0. Il existe donc $h > 0$ tel que pour tout intervalle I , on a

$$\mathbb{E} \left[N_T(I)^{-h} \right] < +\infty, \text{ pour tout } \lambda^2 < \lambda_0^2. \quad (3.75)$$

Preuve - 3.11 -

Nous suivons de très près la démonstration de Molchan [Mol96].

Soit $\varphi(s) = \mathbb{E} [e^{-sN_T(I)}]$, alors nous avons

$$\int_0^\infty s^{h-1} \varphi(s) ds = \Gamma(h) \mathbb{E} \left[N_T(I)^{-h} \right] \quad (3.76)$$

ce qui signifie qu'il suffit de démontrer l'existence de h tel que l'intégrale de gauche (3.76) est finie uniformément en λ^2 .

En utilisant l'inégalité (3.74) et l'inégalité de Jensen, nous avons

$$\varphi(s) \leq \mathbb{E} \left[\prod_{j=1}^4 e^{-sY_j N_j} \right] = \mathbb{E} \left[\prod_{j=1}^4 \varphi(sY_j) \right] \leq \mathbb{E} [\varphi^4(sY)] = \int_0^\infty \varphi^4(sx) dF(x), \quad (3.77)$$

où $F(x)$ est une loi de Y .

Ensuite, en fixant le point $x_0 = s^{-1/2}$ il est possible d'effectuer la majoration suivante

$$\varphi(sx) \leq \begin{cases} 1, & \text{si } x \leq x_0, \\ \varphi(s^{1/2}), & \text{si } x > x_0, \end{cases} \quad (3.78)$$

alors, nous avons

$$\varphi(s) \leq \int_0^{x_0} \varphi^4(sx) dF(x) + \int_{x_0}^\infty \varphi^4(sx) dF(x) \leq F(s^{-1/2}) + \varphi^4(s^{1/2}). \quad (3.79)$$

A l'aide de l'inégalité de Bienaymé-Tchebychev, nous pouvons majorer le premier terme

$$\varphi(s) \leq F(s^{-1/2}) + \varphi^4(s^{1/2}) \leq \mathbb{E} [Y^{-q_0}] s^{-q_0/2} + \varphi^4(s^{1/2}). \quad (3.80)$$

ce qui implique l'inégalité

$$\varphi(s) \leq (\mathbb{E} [Y^{-q_0}] s^{2q-q_0/2} + \varphi^2(s^{1/2})) (s^{-2q} + \varphi^2(s^{1/2})), \quad (3.81)$$

pour tout $q < q_0/4$.

La fonction $\varphi(s)$ tend vers zéro quand $\lambda \rightarrow 0^+$ et $s \rightarrow +\infty$, en effet, nous avons

$$\varphi(s) = \int_0^{|I|/2} e^{-sx} d\mathbb{P}_{N_T(I)}(x) + \int_{|I|/2}^{\infty} e^{-sx} d\mathbb{P}_{N_T(I)}(x) \leq \mathbb{P} \left[N_T(I) \leq \frac{|I|}{2} \right] + e^{-s|I|/2}. \quad (3.82)$$

La probabilité $\mathbb{P} \left[N_T(I) \leq \frac{|I|}{2} \right]$ converge vers zéro uniformément lorsque $\lambda \rightarrow 0^+$, puisque par hypothèse la mesure $N_T(I)$ converge vers $|I|$ uniformément en λ^2 .

En conséquence, on peut fixer λ_0^2 et $s_0 > 1$ tels que

$$\varphi(s) \leq \frac{1}{\sqrt{2}} (s^{-2q} + \varphi^2(s^{1/2})), \text{ pour } s > s_0 \text{ et } \lambda^2 < \lambda_0^2. \quad (3.83)$$

Le lemme 3.12 montre que la fonction caractéristique peut être minorée par

$$\varphi(x) \leq \frac{1}{\sqrt{2}} (x^{-\alpha} + x^{-\beta}), \text{ pour tout } x > x_0, \quad (3.84)$$

où α et β sont des réels positifs et x_0 est supérieur à s_0 .

Pour terminer la démonstration du théorème 3.11, il suffit de choisir h inférieur à α et β , qui sont uniformes en $\lambda^2 < \lambda_0^2$, pour assurer la finitude de l'intégrale (3.76). ■

Lemme 3.12

Supposons qu'il existe q et λ_0^2 positifs et s_0 supérieur à 1, tels que

$$\varphi(s) \leq \frac{1}{\sqrt{2}} (s^{-2q} + \varphi^2(s^{1/2})), \text{ pour } s > s_0 \text{ et } \lambda^2 < \lambda_0^2, \quad (3.85)$$

alors ils existent $\alpha > 0$, $\beta > 0$ et $x_0 > s_0$ tels que

$$\varphi(x) \leq \frac{1}{\sqrt{2}} (x^{-\alpha} + x^{-\beta}), \text{ pour tout } x > x_0. \quad (3.86)$$

Preuve - 3.12 -

Par récurrence, nous trouvons

$$\varphi(s^{2^n}) \leq \frac{1}{\sqrt{2}} \left(a_{n+1} s^{-2^{n+1}q} + \varphi^{2^{n+1}}(s^{1/2}) \right), \text{ où } a_1 = 1, a_{n+1} = 1 + a_n^2. \quad (3.87)$$

Si l'on prend $Q \geq \frac{1+\sqrt{5}}{2}$ alors on a $Q^2 - Q - 1 \geq 0$, et nous avons donc la majoration $Q^{2^n} \geq Q + a_n$. En effet, pour $n = 1$ nous avons

$$Q^2 \geq Q + 1 = Q + a_1, \quad (3.88)$$

en supposant que l'inégalité est valide pour n nous allons la démontrer pour $n + 1$

$$Q^{2^{n+1}} = (Q^{2^n})^2 \geq Q^2 + a_n^2 = Q + a_{n+1} + (Q^2 - Q - 1) \geq Q + a_{n+1}. \quad (3.89)$$

Si l'on fixe $s = s_0 + Q^{1/q}$, alors pour tout x supérieur à s il existe un entier n tel que

$$s^{2^n} \leq x < s^{2^{n+1}}, \text{ ou } 2^n \leq \frac{\ln(x)}{\ln(s)} < 2^{n+1}, \quad (3.90)$$

nous avons donc

$$\varphi(x) \leq \varphi(s^{2^n}) \leq \frac{1}{\sqrt{2}} \left(\left(\frac{Q}{s^q} \right)^{2^{n+1}} + \varphi^{2^{n+1}}(s^{1/2}) \right) \leq \frac{1}{\sqrt{2}} \left(x^{-\alpha} + x^{-\beta} \right), \quad (3.91)$$

avec $\alpha = q - \frac{\ln(Q)}{\ln(s)} > 0$ et $\beta = -\frac{\ln(\varphi(s^{1/2}))}{\ln(s)} > 0$. ■

3.1.4 Exemples

Dans cette section, nous allons illustrer la construction de la mesure MRM pour différentes mesures de Lévy $\nu(dx)$.

Mesure de Lebesgue

Dans le cas le plus simple, où la mesure de Lévy $\nu(dx)$ est égale à zéro. La mesure MRM associée est réduite alors à la mesure de Lebesgue, c'est-à-dire que $M(dt) = dt$ et $\zeta_M(q) = q$. Dans ce cas, le coefficient d'intermittence λ^2 (2.61) est égal à zéro.

MRM log-normale

Si la mesure de Lévy $\nu(dx)$ représentée par une masse à l'origine

$$\nu(dx) = 4\lambda^2\delta(x)dx, \text{ avec } \lambda^2 > 0, \quad (3.92)$$

alors il est facile de montrer, à partir de (3.16) que $\psi(q)$ est la fonction génératrice des cumulants de la loi normale $\psi(q) = -2\lambda^2q + 2\lambda^2q^2$. Ainsi, l'exposant \tilde{q} défini par (3.7) est $\tilde{q} = +\infty$ et l'exposant multifractal est une parabole

$$\zeta_M^{\text{ln}}(q) = q(1 + 2\lambda^2) - 2\lambda^2q^2. \quad (3.93)$$

L'exposant de queue est, d'après le théorème 3.8, par $q_c = \frac{1}{2\lambda^2}$.

Ce modèle correspond au celui proposé par Bacry, Delour et Muzy [BDM01]. Il est particulièrement intéressant car il n'utilise que deux paramètres : le temps intégral T et le coefficient d'intermittence λ^2 . Une section lui est consacrée (cf. section 3.4) car il sera utilisé dans toute la suite de ce travail.

MRM log-Poisson

Nous considérons la mesure de Lévy

$$\nu(dx) = 4\lambda^2\delta(x - c)dx \quad (3.94)$$

qui représente une masse au point c . Cette distribution correspond à la loi de Poisson de paramètre d'échelle $4\lambda^2/c^2$ et d'intensité c . Dans ce cas l'exposant multifractal est donné par

$$\zeta_M^{\text{lp}}(q) = q \left(1 + \frac{\lambda^2}{c^2} (e^c - 1) \right) - \frac{\lambda^2}{c^2} (e^{qc} - 1). \quad (3.95)$$

et on a de nouveau $\tilde{q} = +\infty$. Ce modèle correspond au cadre proposé par She et Lévêque des champs de turbulence [SL94]. Notons, cependant, que ce cadre de travail ne donnait aucune construction explicite de mesure multifractale, l'existence même de ce type de processus n'avait pas encore été démontrée. On peut noter que l'exposant $\zeta_M^{\text{lp}}(q)$ converge vers $\zeta_M^{\text{ln}}(q)$ quand c tend vers zéro, ainsi que l'exposant de queue est infini lorsque c est négatif et assez petit.

MRM log-Poisson composée

Il est possible de généraliser le cas précédent. Considérons la mesure canonique de Lévy $\nu(dx)$ qui vérifie

$$\int_{-\infty}^{\infty} \frac{\nu(dx)}{x^2} = C < +\infty, \quad (3.96)$$

c'est-à-dire que la distribution de $\nu(dx)$ n'est pas concentrée à l'origine. Il est simple de voir que $F(dx) = \frac{\nu(dx)}{Cx^2}$ est une mesure de probabilité. La loi infiniment divisible dans ce cas est une loi de Poisson d'échelle C composée avec la distribution F [Fel71]. En faisant le calcul de l'exposant multifractal, on a

$$\zeta_M^{\text{pc}}(q) = q(1 + C(\mathbb{E}[W] - 1)) - C(\mathbb{E}[W^q] - 1), \quad (3.97)$$

où la variable W telle que $\ln(W)$ est distribuée selon la loi $F(dx)$. La mesure obtenue par cette construction correspond (cf. section 3.1.5), après troncature à l'échelle T des cônes $\mathcal{A}_{l,T}$, et donc, perte de l'invariance d'échelle exacte, au produit des pulsations cylindriques de Barral et Mandelbrot [BM02]. Dans ce travail, les auteurs se sont attachés à démontrer la validité du formalisme multifractal (cf. section 2.2.5).

Dans cet exemple, $\tilde{q} = \sup_q \{\mathbb{E}[W^q] < +\infty\} > 1$.

MRM Log- α -stable

Nous considérons le cas où la mesure de Lévy $\nu(dx)$ correspond à la densité α -stable

$$\nu(dx) = \begin{cases} C|x|^{1-\alpha}, & \text{si } x \leq 0, \\ 0, & \text{si } x > 0, \end{cases} \quad (3.98)$$

où $C > 0$ et $0 < \alpha < 2$. Le calcul direct mène à $\tilde{q} = +\infty$ et

$$\zeta_M^{\text{ls}}(q) = q(1 + \sigma^\alpha) - \sigma^\alpha |q|^\alpha. \quad (3.99)$$

De telles distributions ont été utilisées dans le contexte de la turbulence et de la géophysique [SLSL92].

MRM Log-gamma

La famille des distributions Gamma correspond à la mesure de Lévy de la forme

$$\nu(dx) = C\gamma^2 x e^{-\gamma x} dx, \text{ pour } x \geq 0. \quad (3.100)$$

Pour cette classe de distributions, on a $\tilde{q} = \gamma$, d'où l'on déduit la condition nécessaire pour construire la mesure MRM $\gamma > 1$. Dans ce cas l'exposant multifractal est donné par

$$\zeta_M^{\text{lg}}(q) = q \left(1 + C\gamma^2 \ln \left(\frac{\gamma}{\gamma - 1} \right) \right) - C\gamma^2 \ln \left(\frac{\gamma}{\gamma - q} \right). \quad (3.101)$$

Notons que si $C = 4\lambda^2$ l'exposant $\zeta_M^{\text{lg}}(q)$ tend vers $\zeta_M^{\text{ln}}(q)$ quand γ tend vers $+\infty$. L'exposant de queue q_c est donné par la solution de l'équation $W(q)e^{W(q)} = q$, où $W(q)$ est la fonction de Lambert.

3.1.5 Construction pour une invariance d'échelle asymptotique

La construction de la mesure MRM peut être généralisée en modifiant arbitrairement la forme des cônes à grande échelle. Dans ce cas, on peut montrer que la mesure construite possède l'invariance d'échelle asymptotique (2.45) et non plus exacte (2.44).

Plus de détails sur la construction MRM dans le cas général, peuvent être trouvés dans l'article de Bacry et Muzy [BM03].

3.2 Construction du processus multifractal MRW

3.2.1 Construction d'un mouvement brownien subordonné

Les premières tentatives de construction d'une marche¹ aléatoire multifractale **MRW** (*Multifractal Random Walk*) en utilisant la subordination d'un mouvement brownien fractionnaire, étudié dans la section 2.1.7, apparaissent dans [BDM01, MB02]. La multifractalité y est introduite par la dynamique de la volatilité ce qui rend le modèle plus proche, dans l'esprit, des processus à volatilité stochastique, voir la section 1.3.4.

Dans la suite, Bacry et Muzy ont porté leur attention sur le cas de la subordination d'un mouvement brownien standard.

L'approche la plus simple consiste à subordonner un mouvement brownien $B(t)$ à l'aide d'un processus croissant $M([0, t])$. La subordination du mouvement brownien à un processus croissant a été introduite par Mandelbrot et Taylor [MT67] et ce sujet a été intensivement étudié dans la littérature mathématique. La subordination multifractale a été considérée par Mandelbrot et al. [Man99] ainsi que dans le modèle MMAR (cf. section 2.4.2) et est largement utilisée pour construire des processus multifractals.

Définition 3.4

Soient $\{B(t); t > 0\}$ un mouvement brownien de variance σ^2 et $M_T(dt)$ une mesure MRM non dégénérée indépendante de $B(t)$. Le processus MRW subordonné est défini par

$$X_T^{(s)}(t) = B(M_T([0, t])). \quad (3.102)$$

Une construction alternative consiste à intégrer la mesure $M_{l,T}$ par rapport au processus de Wiener et ensuite passer à la limite $l \rightarrow 0$.

Définition 3.5

Soient $dB(u)$ un processus de Wiener de variance σ^2 et $\omega_l(u)$ un processus introduit par (3.11) indépendant de $dB(u)$. Le processus MRW est défini comme la limite

$$X_T(t) = \lim_{l \rightarrow 0^+} X_{l,T}(t), \quad (3.103)$$

¹Attention, le mot "marche" peut prêter à confusion ici. Les processus MRW correspondent, comme nous le verrons, à des processus à temps continu.

où $X_{l,T}(t)$ est donné par

$$X_{l,T}(t) = \int_0^t e^{\omega_{l,T}(u)} dB(u). \quad (3.104)$$

Notons que le processus $X_{l,T}(t)$ est bien défini, car

$$\int_0^t \mathbb{E} \left[\left(e^{\omega_{l,T}(u)} \right)^2 \right] du = \int_0^t \mathbb{E} \left[e^{2\omega_{l,T}(u)} \right] du = t. \quad (3.105)$$

L'équivalence des deux dernières définition est énoncée dans le théorème suivant :

Théorème 3.13 (Bacry-Muzy)

S'il existe $\epsilon > 0$ tel que $\zeta_M(1 + \epsilon) > 1$, alors

$$X_T(t) = \lim_{l \rightarrow 0^+} X_{l,T}(t) \stackrel{\mathcal{L}}{=} \lim_{l \rightarrow 0^+} B(M_{l,T}([0, t])) = B(M_T([0, t])) = X_T^{(s)}(t), \quad (3.106)$$

où $B(t)$ et $dB(u)$ sont un mouvement brownien et un processus de Wiener correspondant de variance σ^2 indépendants de la probabilité \mathcal{P} . De même,

$$\mathbb{E}[X_T(t)^2] = \sigma^2 t. \quad (3.107)$$

Preuve - 3.13 -

La preuve de théorème 3.13 est détaillée dans [BM03]. ■

En pratique, pour modéliser des séries temporelles réelles, nous avons besoin d'une représentation à temps discret (échantillonnage uniforme) du processus MRW.

Proposition 3.14

L'échantillonnage uniforme $\{X_T(n\tau)\}_n$ à l'échelle τ du processus MRW peut être représenté comme

$$\{\delta_\tau X_T(n\tau)\}_n \stackrel{\mathcal{L}}{=} \{\epsilon[n] \sqrt{\delta_\tau M_t([0, n\tau])}\}_n, \quad (3.108)$$

où $\{\epsilon[n]\}_n$ est un bruit blanc gaussien centré de variance σ^2 .

Preuve - 3.14 -

La représentation (3.108) est une conséquence de l'équation (3.102). ■

3.2.2 Propriété du processus MRW

Théorème 3.15 (Bacry-Muzy)

Le processus $X_T(t)$ vérifie les propriétés suivantes :

1. Autosimilarité stochastique, étudiée dans la section 2.2.2, pour $\lambda \in (0, 1)$

$$\{X_T(\lambda t)\}_t \stackrel{\mathcal{L}}{=} W_\lambda \{X_T(t)\}_t, \text{ pour } t \leq T, \quad (3.109)$$

avec $W_\lambda = \lambda^{1/2} e^{\Omega_\lambda}$, où Ω_λ est une variable aléatoire infiniment divisible indépendante de $\{X_T(t)\}_t$ dont la fonction caractéristique est donnée par

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}, \quad (3.110)$$

2. Invariance d'échelle exacte (2.44) et multifractalité. On pose $\zeta_X(q) = \zeta_M(\frac{q}{2})$. S'il existe $\epsilon > 0$ tel que $\zeta_X(q + \epsilon) > 1$, en d'autres termes, si le moment d'ordre $q/2$ de la mesure M_T existe, alors pour tout $t \leq T$,

$$\mathbb{E}[|X_T(t)|^q] = \left(\frac{t}{T}\right)^{\zeta_X(q)} \mathbb{E}[|X_T(T)|^q], \text{ pour } q > 0. \quad (3.111)$$

De plus, si la mesure de Lévy $\nu(dx)$ n'est pas identiquement nulle alors $\zeta_M(q)$ et donc $\zeta_X(q)$ n'est pas linéaire et la processus X_T est multifractal.

3. Existence de moments d'ordres positifs de $X_T(t)$:

(a) Si $\zeta_X(q) > 1$ alors $\mathbb{E}[|X_T(t)|^q] < +\infty$,

(b) Soit $q > 1$. Si $\mathbb{E}[|X_T(t)|^q] < +\infty$ alors $\zeta_X(q) \geq 1$.

Preuve - 3.15 -

Ce théorème est une conséquence directe des théorèmes de la section 3.1.2. ■

3.3 Invariance par changement du temps intégral

Les résultats de cette section sont originaux.

Dans les sections 3.1.2 et 3.2.2 nous avons fait remarquer que la mesure MRM et le processus MRW possèdent les propriétés d'autosimilarité stochastique, données par les équations (3.39) et (3.109). Ces propriétés reposent sur le comportement (3.28) de la fonction $\rho_{l,T}(t)$ lorsque t varie.

Il est donc naturel d'utiliser la relation (3.29), qui exprime le comportement de la fonction $\rho_{l,T}(t)$ lorsque le temps intégral T varie, pour étudier les propriétés d'invariance par changement du temps intégral. Comme nous allons le voir, celles-ci ont la même forme que les propriétés d'invariance d'échelle (cf. théorème 3.7) de la mesure MRM ou du processus MRW log-infiniment divisibles.

Théorème 3.16

La mesure MRM vérifie la propriété suivante pour $\lambda \in (0, 1)$

$$\{M_T([0, t])\}_t \stackrel{\mathcal{L}}{=} W_\lambda \{M_{\lambda T}([0, t])\}_t, \text{ pour } t \leq \lambda T \quad (3.112)$$

avec $W_\lambda = e^{2\Omega_\lambda}$, où Ω_λ est une variable aléatoire infiniment divisible dont la fonction caractéristique est donnée par

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.113)$$

De plus, pour tout $t \leq T_1 \leq T_2$, s'il existe $\epsilon > 0$ tel que $\zeta_M(q + \epsilon) > 1$, en d'autres termes, si le moment d'ordre q existe, alors la mesure MRM vérifie la propriété suivante

$$\mathbb{E} [M_{T_2}([0, t])^q] = \left(\frac{T_1}{T_2}\right)^{\zeta_M(q)-q} \mathbb{E} [M_{T_1}([0, t])^q], \text{ pour } q > 0. \quad (3.114)$$

Preuve - 3.16 -

En utilisant la proposition 3.4 et la relation (3.29), on obtient

$$\mathbb{E} \left[e^{\sum_k ip_k \omega_{l,T}(t_k)} \right] = \lambda^{-\varphi(\frac{1}{2} \sum_k p_k)} \mathbb{E} \left[e^{\sum_k ip_k \omega_{l,\lambda T}(t_m)} \right] = \mathbb{E} \left[e^{\sum_k ip_k (\Omega_\lambda + \omega_{l,\lambda T}(t_k))} \right], \quad (3.115)$$

où la variable aléatoire Ω_λ est indépendante du processus $\{\omega_l(t)\}_t$ et sa fonction caractéristique s'écrit comme

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.116)$$

Il suffit de calculer l'espérance

$$\mathbb{E} [M_{\lambda l,T}([0, t])^q] = \mathbb{E} [W_\lambda^q] \mathbb{E} [M_{l,T}([0, T])^q] = \lambda^{q-\psi(q)} \mathbb{E} [M_{l,T}([0, T])^q], \quad (3.117)$$

avec $\lambda = t/T$ et passer à la limite sous le signe espérance d'après le théorème 3.3. ■

Théorème 3.17

Le processus $X(t)$ vérifie la propriété suivante pour $\lambda \in (0, 1)$

$$\{X_T(t)\}_t \stackrel{\mathcal{L}}{=} W_\lambda \{X_{\lambda T}(t)\}_t, \text{ pour } t \leq \lambda T, \quad (3.118)$$

avec $W_\lambda = e^{\Omega_\lambda}$, où Ω_λ est une variable aléatoire infiniment divisible indépendante de $\{X_T(t)\}_t$ dont la fonction caractéristique est donnée par

$$\mathbb{E} [e^{iq\Omega_\lambda}] = \lambda^{-\varphi(\frac{q}{2})}. \quad (3.119)$$

De même, pour tout $t \leq T_1 \leq T_2$, s'il existe $\epsilon > 0$ tel que $\zeta_X(q + \epsilon) > 1$, en d'autres termes, si le moment d'ordre q existe, alors le processus MRW vérifie la propriété suivante

$$\mathbb{E} [|X_{T_2}([0, t])|^q] = \left(\frac{T_1}{T_2}\right)^{\zeta_X(q)-q} \mathbb{E} [|X_{T_1}([0, t])|^q], \text{ pour } q > 0. \quad (3.120)$$

Preuve - 3.17 -

Ce résultat est une conséquence directe du théorème 3.16 ■

3.4 Modèle MRW log-normal

Dans les sections précédentes, nous avons étudiés les propriétés communes à toutes les mesures MRM ou tous les processus MRW, quelle que soit la loi infiniment divisible $\nu(dx)$ utilisée. Le cas log-normal va jouer un rôle central dans notre travail, nous étudions donc dans cette section les propriétés qui lui sont propres.

Dans la suite nous allons travailler avec un temps intégral T fixé. Afin d'alléger le texte, nous nous permettrons parfois d'omettre l'indice T dans les notations $\mathcal{A}_{l,T}(t)$, $\omega_{l,T}(t)$, $M_{l,T}(dt)$ et $M_T(dt)$.

3.4.1 Construction alternative de la mesure MRM log-normale

Dans le cas log-normal, le processus $\{\omega_l(t)\}_t$ étant normal, il peut être directement défini par sa moyenne et sa fonction d'autocovariance. Il n'est pas nécessaire de passer par la représentation bi-dimensionnelle introduite dans le cas général infiniment divisible (cf. définition 3.3). La construction de la mesure MRM log-normale s'en trouve simplifiée. Notons que cette représentation unidimensionnelle avait été découverte [BDM01] avant la construction générale bi-dimensionnelle. Certaines propriétés de cette mesure limite, et notamment son existence, n'avait pas alors été démontrée. L'existence peut, en fait, être prouvée à partir de la théorie du chaos gaussien de Kahane [Kah87] ou encore découle de la construction générale bi-dimensionnelle des sections précédentes.

Le processus gaussien $\{\omega_l(t)\}_t$ est entièrement déterminé par sa moyenne

$$\mathbb{E}[\omega_l(t)] = -\lambda^2 \left(\ln \left(\frac{T}{l} \right) + 1 \right) \quad (3.121)$$

et par sa fonction d'autocovariance

$$\rho_l(\tau) = \text{Cov}[\omega_l(t), \omega_l(t + \tau)] = \begin{cases} \lambda^2 \left(\ln \left(\frac{T}{l} \right) + 1 - \frac{\tau}{l} \right), & \text{si } 0 \leq \tau < l, \\ \lambda^2 \ln \left(\frac{T}{\tau} \right), & \text{si } l \leq \tau < T, \\ 0, & \text{si } T \leq \tau < +\infty, \end{cases} \quad (3.122)$$

la construction de la mesure MRM donnée par la définition 3.3 reste alors valide. La mesure MRM $M_T(dt)$ est définie comme la limite des mesures $M_{l,T}(dt)$

$$M_T(dt) = \lim_{l \rightarrow 0^+} M_{l,T}(dt), \quad (3.123)$$

où la mesure $M_{l,T}(dt)$ est donnée par

$$M_{l,T}(dt) = e^{2\omega_{l,T}(t)} dt, \text{ pour } l > 0 \quad (3.124)$$

ou au sens de Lebesgue, pour tout ensemble mesurable I , on a

$$M_{l,T}(I) = \int_I e^{2\omega_{l,T}(t)} dt. \quad (3.125)$$

3.4.2 Propriétés du processus MRW log-normale

Paramètres du modèle

Le modèle MRW log-normal fait intervenir trois paramètres : σ , T et λ^2 .

- Le premier, σ , correspond simplement à la volatilité moyenne (variance non conditionnelle) du modèle.
- Le deuxième, T , correspond au temps de décorrélation du processus $\{\omega_l(t)\}_t$. Ainsi, les variables $\omega_l(t)$ et $\omega_l(t')$ éloignées de plus de T sont indépendantes. En conséquence, les accroissements du processus MRW éloignés de plus de T sont aussi indépendants. Ce paramètre T correspond aussi au temps intégral du processus au delà duquel le comportement multifractal cesse d'exister (cf. proposition 2.14)
- Le troisième, λ^2 , est sans doute le plus important. Il gouverne la force des corrélations des variables $\omega_l(t)$: plus λ^2 est élevé, plus celles-ci sont importantes. Il correspond aussi au coefficient d'intermittence (cf. définition 2.13). Lorsque λ^2 est égal à zéro le processus $\{\omega_l(t)\}_t$ est dégénéré et devient nul, en effet $\mathbb{E}[\omega_l(t)^2] = 0$. Dans ce cas le processus MRW se réduit au mouvement brownien de volatilité σ^2 . Le paramètre λ^2 mesure donc "l'écart" au mouvement brownien ou encore la "quantité" de multifractalité.

Choix de la fonction d'autocovariance et stationnarité

Le choix (3.122) d'une fonction de covariance décroissant logarithmiquement avec la distance n'est pas anodin. Elle est, en effet, de la même forme que celle observée dans certains modèles de cascade multifractale [ABM98, BDM01]. Seul ce choix permet de garantir l'invariance d'échelle exacte (2.44) d'une mesure MRM.

En physique statistique, plusieurs travaux se sont récemment intéressés à des modèles de désordre présentant la même structure de corrélation. Ils mettent en évidence le caractère marginal de ces corrélations notamment en ce qui concerne l'étude des valeurs extrêmes [CD01].

Plutôt que de définir le processus $\{\omega_l(t)\}_t$ par leur structure de corrélation, il est possible de le donner sous une forme intégrale, à la manière de la représentation de Mandelbrot et Van Ness du mouvement brownien fractionnaire

$$\omega_l(t) = -\mathbb{E}[\omega_l(t)] + \int_{-\infty}^t K_l(t-u)dB(u), \quad (3.126)$$

avec le noyau K_l causal, vérifiant l'équation de convolution

$$K_l * K_l(t) = \text{Cov}[\omega_l(0), \omega_l(t)]. \quad (3.127)$$

Une transformée de Fourier de l'équation précédente montre que dans le régime $l \ll t \ll T$ le noyau K_l doit se comporter comme

$$K_l(t) \sim \frac{K_0}{\sqrt{t}}. \quad (3.128)$$

Une telle représentation met en évidence le caractère très particulier de la mesure MRM log-normale. La lente décroissance du noyau en $1/\sqrt{t}$ apparaît comme un cas marginal puisque une décroissance du noyau plus lente conduirait à un processus non stationnaire. Formellement, le processus $\{\omega_l(t)\}_t$ représente la limite d'un mouvement brownien fractionnaire lorsque le paramètre de Hurst H tend vers zéro.

Calculs des moments d'ordres entiers

Le théorème 3.15 montre que le processus MRW log-normal possède une invariance d'échelle exacte (2.44). Il est possible d'établir une formule explicite des moments d'ordre entiers des accroissements du processus MRW log-normal et d'obtenir ainsi le coefficient $C(q)$, pour q entier.

Le moment d'ordre entier q du processus $X(t)$ s'écrit grâce à l'indépendance des mouvement brownien $B(t)$ et mesure $M(dt)$

$$\mathbb{E}[X(t)^q] = \mathbb{E}[B(M([0, t]))^q] = \mathbb{E}[B(1)^q] \mathbb{E}[M([0, t])^{q/2}], \text{ pour tout } q \text{ positif,} \quad (3.129)$$

d'où l'on déduit que les moments impairs du processus MRW log-normal sont nuls.

Notons $q = 2p$. Le théorème de Wick [Wic50] conduit à la formule

$$\begin{aligned} \mathbb{E}[X(t)^{2p}] &= \mathbb{E}[B(1)^{2q}] \mathbb{E}[M([0, t])^q] = \frac{(2p)! \sigma^{2p}}{2^{2p} p!} \int_0^t du_1 \dots \int_0^t du_q \prod_{1 \leq i < j \leq q} \frac{T^{4\lambda^2}}{|u_i - u_j|^{4\lambda^2}} \\ &= \frac{(2p)! \sigma^{2p} t^p}{2^{2p} p!} \left(\frac{T}{t}\right)^{2p(p-1)\lambda^2} \int_0^1 du_1 \dots \int_0^1 du_q \prod_{1 \leq i < j \leq q} \frac{1}{|u_i - u_j|^{4\lambda^2}} = K_{2p} t^{\zeta_X^{\ln}(2p)}, \end{aligned} \quad (3.130)$$

où le préfacteur K_{2p} fait intervenir une intégrale, dite intégrale de Selberg [Sel44], dont le calcul explicite est donné dans [BDM01]

$$K_{2p} = \frac{(2p)! \sigma^{2p} T^{2p(p-1)\lambda^2}}{2^{2p} p!} \prod_{k=0}^{p-1} \frac{\Gamma(1 - 2(k+1)\lambda^2) \Gamma(1 - 2k\lambda^2)^2}{\Gamma(2 - 2(p+k-1)\lambda^2) \Gamma(1 - 2\lambda^2)}. \quad (3.131)$$

et l'exposant $\zeta_X^{\ln}(2p)$ est précisé ci-dessous (3.136).

Finitude des moments et queues de distribution

D'après le calcul précédent et le fait que la fonction Gamma est divergente pour tout réel négatif ou nul, pour que le moment d'ordre $q = 2p$ ne diverge pas il est nécessaire que

$$\begin{cases} \lambda^2 < \frac{1}{2}, \\ q < q_c = \frac{1}{\lambda^2}. \end{cases} \quad (3.132)$$

Notons que ce résultat s'étend aux moments d'ordre quelconque (non forcément entier). En effet, c'est une conséquence directe du théorème 3.15 qui montre que l'exposant critique q_c correspondant à l'explosion des moments est donné par l'équation $\zeta_X^{\ln}(q_c) = 1$. En insérant (3.136) dans cette dernière relation, on montre bien que $q_c = \frac{1}{\lambda^2}$.

En utilisant la proposition 1.1, on montre donc que pour tout q strictement supérieur à $q_c = \frac{1}{\lambda^2}$, il existe une constante C réelle positive, telle que pour tout x_0 réel positif, il existe x supérieur à x_0 , tel que

$$\mathbb{P}[M([0, t]) > x] > Cx^{-q/2}, \quad (3.133)$$

$$\mathbb{P}[X(t) > x] > Cx^{-q}. \quad (3.134)$$

Exposant multifractal

Nous avons vu dans le théorème 3.15 que, de façon générale, quelle que soit $\nu(dx)$, les exposants multifractals de la mesure MRM et du processus MRW vérifient

$$\zeta_X(q) = \zeta_M\left(\frac{q}{2}\right). \quad (3.135)$$

Dans le cas log-normal, en utilisant l'équation (3.93), nous obtenons :

$$\zeta_X^{\ln}(q) = \zeta_M^{\ln}\left(\frac{q}{2}\right) = \frac{q}{2}(1 + 2\lambda^2) - \frac{\lambda^2 q^2}{2}. \quad (3.136)$$

Nous pouvons préciser quelques valeurs remarquables de cette fonction. Notons tout d'abord que $\zeta_X^{\ln}(2) = 1$, puisque les accroissements sont décorrélés (cf. proposition 2.10). En d'autres termes le moment d'ordre 2 du MRW log-normal présente un exposant, typique du mouvement brownien standard.

Autocovariance de volatilité et mémoire longue

Dans le modèle MRW log-normal, les accroissements sont décorrélés mais non indépendants. Il est aisé de calculer les covariances entre les valeurs au carré des incréments. Dans [MDB00, BDM01] il est montré que pour des intervalles disjoints $[t - \tau, t]$ et $[t + h - \tau, t + h]$ éloignés l'un de l'autre de moins de T , on a pour $h \geq \tau$

$$\mathbb{E} [\delta_\tau X(t)^2 \delta_\tau X(t + h)^2] = \sigma^4 T^{4\lambda^2} \frac{(h + \tau)^{2-4\lambda^2} + (h - \tau)^{2-4\lambda^2} - 2h^{2-4\lambda^2}}{(1 - 4\lambda^2)(2 - 4\lambda^2)}. \quad (3.137)$$

Dans le cas particulier, où $\tau \ll h < T$, cette espérance peut être approximée par

$$\mathbb{E} [\delta_\tau X(t)^2 \delta_\tau X(t + h)^2] \sim \sigma^4 \tau^2 \left(\frac{T}{h}\right)^{4\lambda^2}, \quad (3.138)$$

c'est-à-dire une loi de puissance en h .

Il est possible de montrer, en utilisant l'approximation de petit intermittence (cf. chapitre 5) et le théorème 5.17, que la fonction d'autocovariance des accroissements absolus s'écrit

$$\text{Cov}[|\delta_\tau X(t)|, |\delta_\tau X(t+h)|] = \frac{2}{\pi} \left(\frac{T e^{3/2}}{\tau} \right)^{-\lambda^2} \left(\left(\frac{T}{h} \right)^{\lambda^2} - 1 \right), \text{ pour } \tau \ll h < T, \quad (3.139)$$

ce qui correspond au comportement observés dans les marchés (cf. section 1.2.4).

Stricto sensu, ce processus n'est pas à mémoire longue puisque sa mémoire est de taille finie, au delà d'une distance T , ses incréments sont indépendants. Cependant, la fonction d'autocovariance des carrés des accroissements décroît très lentement vers la valeur zéro. En pratique, si T est assez grand, par exemple, quelques années, le processus MRW log-normal permet de reproduire le phénomène de mémoire longue (cf. section 1.2.5) observée dans les séries financières. Nous renvoyons le lecteur à la section 3.4.4 qui porte sur l'étude de la reproduction des faits stylisés par le modèle MRW log-normal.

3.4.3 Simulation numérique

Toutes les définitions précédemment énoncées du processus MRW log-normal $X(t)$ consistent à le définir comme processus limite, lorsque l tend vers zéro, d'un processus à temps continu $X_l(t)$. A ce jour, il n'existe pas de construction directe de cet objet limite. La seule manière de réaliser des simulations numériques de ce processus est donc de fixer l "assez" petit et de simuler un échantillonnage du processus $X_l(t)$. En pratique, l représente une échelle en dessous de laquelle le processus $X_l(t)$ est très régulier. Pour que la simulation soit bonne, il suffira donc de choisir l assez petit devant le pas d'échantillonnage.

Soit τ le pas d'échantillonnage. Nous fixons l'échelle l , qu'on appelle **échelle de sur-échantillonnage**, telle que $\tau = nl$, avec $n \gg 1$ de façon à ce que $l \ll \tau$.

Pour simuler N points d'une trajectoire d'un MRW, il suffit de générer un processus gaussien $\{\omega_l[t]; t = l, \dots, Nnl\}$ de moyenne (3.121) et de fonction d'autocovariance (3.122). La proposition 3.14 suggère alors de simuler

$$\widetilde{M}_{l,\tau}([0, k\tau]) = l \sum_{j=1}^{kn} e^{2\omega_l[jl]}, \quad (3.140)$$

$$\widetilde{X}_{l,\tau}(k\tau) = \sqrt{\widetilde{M}_{l,\tau}([0, k\tau])} \epsilon[k], \quad (3.141)$$

où $\{\epsilon[k]; k = 1, \dots, N\}$ est un bruit blanc gaussien centré de variance σ^2 .

La simulation du processus $\{\omega_l[t]; t = l, \dots, Nnl\}$ n'est pas totalement triviale. La méthode que nous utilisons fait appel à sa représentation spectrale et utilise une transformation de Fourier rapide, en conséquence elle autorise la simulation de très longues séries. L'algorithme est expliqué en détail dans l'appendice C.

Il est possible de démontrer [BDM01, MB02, BM03] que les processus $\{\widetilde{M}_{l,\tau}([0, k\tau])\}_k$ et $\{\widetilde{X}_{l,\tau}([0, k\tau])\}_k$ convergent en moyenne quadratique vers $\{M([0, k\tau])\}_k$ et $\{X([0, k\tau])\}_k$ lorsque l'échelle de sur-échantillonnage l tend vers zéro.

Notons que l'expression (3.141) correspond exactement au processus MRW log-normal introduit dans [MDB00, BDM01].

3.4.4 Faits stylisés

Tout d'abord il faut noter que, par construction (3.4), le processus MRW est un processus à accroissements stationnaires. Les propriétés du processus MRW log-normal proviennent de la forme très particulière de la fonction de corrélation $\rho_l(t)$ (3.122). Il est possible de mener beaucoup de calculs explicitement avec ce modèle. En particulier, nous sommes capables de donner une formule explicite des moments d'ordre entiers positifs (cf. les équations (3.130) et (3.131)).

Le modèle MRW log-normal est d'un grand intérêt en finance puisque, comme nous allons le voir dans cette section, il reproduit fidèlement la majeure partie des faits stylisés classiques étudiés dans la section 1.2, ainsi que l'intégralité des faits stylisés multifractals étudiés dans la section 2.3.

Pour mener à bien cette étude, nous n'allons utiliser qu'une seule réalisation du processus MRW log-normal, présentée sur la figure 3.3, dont les paramètres sont les suivants

$$N = 2^{18}, \quad \text{nombre d'échantillon de la réalisation,} \quad (3.142)$$

$$\tau = 1, \quad \text{échelle d'échantillonnage,} \quad (3.143)$$

$$l = 2^{-4}, \quad \text{échelle de sur-échantillonnage,} \quad (3.144)$$

$$T = 2000, \quad \text{échelle intégrale,} \quad (3.145)$$

$$\lambda^2 = 0.02, \quad \text{coefficient d'intermittence,} \quad (3.146)$$

$$\sigma^2 = 1, \quad \text{variance.} \quad (3.147)$$

Notons que les simulations numériques plus approfondies montrent qu'il est inutile de prendre une échelle de sur-échantillonnage inférieure à $l = 2^{-4}$, les corrections apportées par l'utilisation d'une échelle plus fine sur les propriétés statistiques que nous étudions ici sont du second ordre.

Accroissements, skewness et kurtosis

Pour une meilleure lisibilité, la figure 3.4 ne représente que les 15000 premiers accroissements à l'échelle $\tau = 1$ de la réalisation de la figure 3.3. On observe l'allure qui est très différente de celle d'un bruit blanc gaussien : la variabilité des accroissements est plus importante et des cycles semblent apparaître, comme dans la figure 1.2 représentant les accroissements d'une série financière.

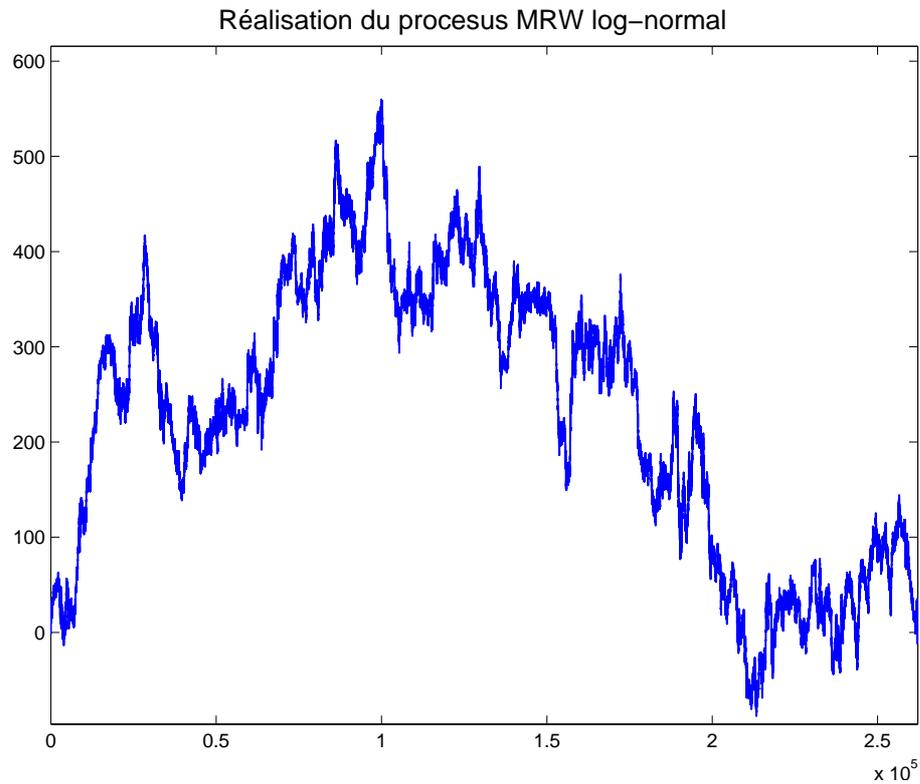


FIG. 3.3 – Réalisation comportant 2^{18} échantillons du processus MRW log-normal simulé en utilisant les paramètres donnés par les équations (3.142) à (3.147). Son comportement présente un caractère erratique, similaire à celui du taux de change GBP/USD, représenté dans la figure 1.1. Cette réalisation sera utilisée pour tous les calculs empiriques de cette section.

Echelle	Moyenne	Variance	Skewness	Kurtosis	Test de JB
1	-0.0000 (0)	1.0388 (1)	-0.0114 (0)	6.1597 (6.2394)	0.0000 (0.0000)
10	-0.0004 (0)	10.4027 (10)	-0.0427 (0)	5.1307 (5.1897)	0.0000 (0.0000)
50	-0.0005 (0)	52.3994 (50)	0.0481 (0)	4.8509 (4.5628)	0.0000 (0.0000)
100	-0.0010 (0)	105.3465 (100)	0.0487 (0)	4.2418 (4.3166)	0.0000 (0.0000)
500	0.0518 (0)	506.6979 (500)	0.0685 (0)	3.9481 (3.7951)	0.0000 (0.1001)
1000	0.1036 (0)	1010.0620 (1000)	0.1744 (0)	3.6811 (3.5904)	0.0409 (0.5302)

TAB. 3.1 – Moyenne, variance, skewness, kurtosis et p -valeur du test statistique de Jarque et Bera des accroissements de la réalisation de la figure 3.3 à différentes échelles. Pour des échelles inférieures à 500, la normalité est rejetée puisque la p -valeur est très proche de zéro. Les chiffres entre les parenthèses représentent les valeurs théoriques qui ont été calculées en utilisant les équations (3.130) et (3.131).

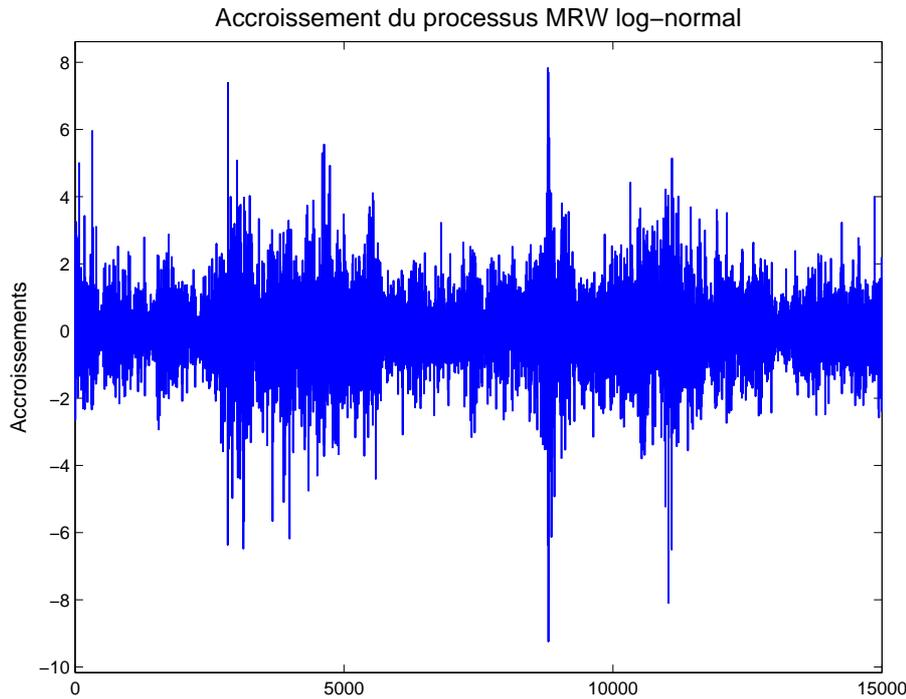


FIG. 3.4 – 15000 premiers accroissements de la réalisation de la figure 3.3. On observe l’allure très différente de celle d’un bruit blanc gaussien, alors qu’elle ressemble à celle des rendements logarithmiques journaliers du taux de change GBP/USD, voir la figure 1.2.

Dans le tableau 3.1 nous donnons les statistiques : moyenne, variance, skewness, kurtosis et p -valeur du test statistique de Jarque et Bera (1.10), pour les accroissements de la réalisation de la figure 3.3 à différentes échelles. On observe un très fort excès de kurtosis pour les petites échelles, la p -valeur du test statistique de Jarque et Bera est significativement non nulle pour l’échelle 1000 seulement.

Variation de la distribution des accroissements en fonction de l’échelle

Dans le théorème 3.17, nous avons vu que le processus MRW est autosimilaire au sens stochastique. La proposition 2.16 montre alors que le profil des distributions des accroissements change en fonction de l’échelle. La figure 3.5 représente les densités empiriques obtenues à partir de la réalisation de la figure 3.3, pour différentes échelles. On observe effectivement que les queues de la distribution des accroissements du processus MRW log-normal sont plus épaisses que celles de la distribution normales (trait plein). On observe également que les queues deviennent de moins en moins épaisses lorsque l’échelle de temps considéré (pour le calcul des accroissements) augmente, jusqu’à se rapprocher d’une loi normale pour de grands intervalles de temps.

Sur la figure 3.6, nous avons tracé les quantiles empiriques versus les quantiles théoriques

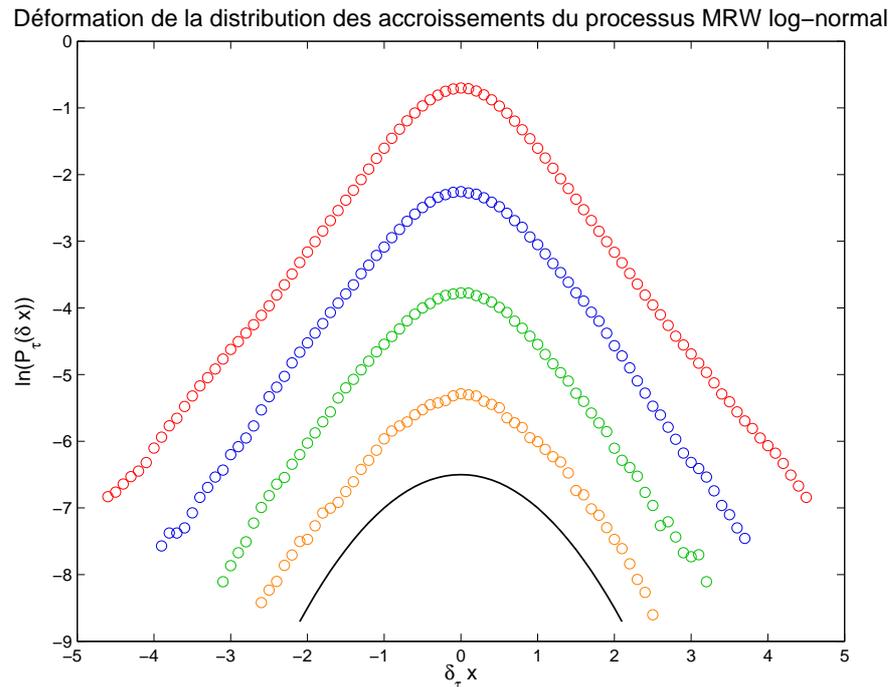


FIG. 3.5 – Déformation de la distribution empirique des accroissements normalisés par leur écart-type de la réalisation de la figure 3.3 en fonction de l'échelle. De haut en bas : $\tau = 1, 10, 50$ et 150 . L'échelle est logarithmique et les courbes sont décalées arbitrairement pour une meilleure lisibilité. Le tracé montre clairement (voir, par exemple, le changement de convexité des queues) que les distributions qui correspondent aux petites échelles (en haut) possèdent un caractère leptokurtique, alors que celles qui correspondent aux grandes échelles (en bas) sont proches de la distribution gaussienne standard, représentée par le trait plein. Un tel comportement est similaire à celui des rendements logarithmiques journaliers du taux de change GBP/USD, représenté dans la figure 1.3 de la section 1.2.3 sur l'étude de la distribution des rendements.

d'une loi normale, des accroissements de la réalisation de la figure 3.3 à différentes échelles. La déviation à la normalité, qui se traduit par un tracé non linéaire, est plus forte aux petites échelles. Les déviations diminuent avec l'échelle mais elles le font lentement et restent considérables même à l'échelle de vingt jours. Ce comportement ressemble beaucoup à celui des données du taux de change GBP/USD étudié dans la section 1.2.3.

Décorrélés des rendements

Sur la figure 3.7 nous avons représenté la fonction d'autocorrélation empirique des accroissements de la réalisation de la figure 3.3. Sur ce tracé, la fonction d'autocorrélation décroît extrêmement rapidement lorsque le décalage h croît et peut être considérée comme nulle dès que h est différent de zéro. Ce qui permet de reproduire l'absence de corrélation temporelle des séries financières étudiée dans la section 1.2.4.

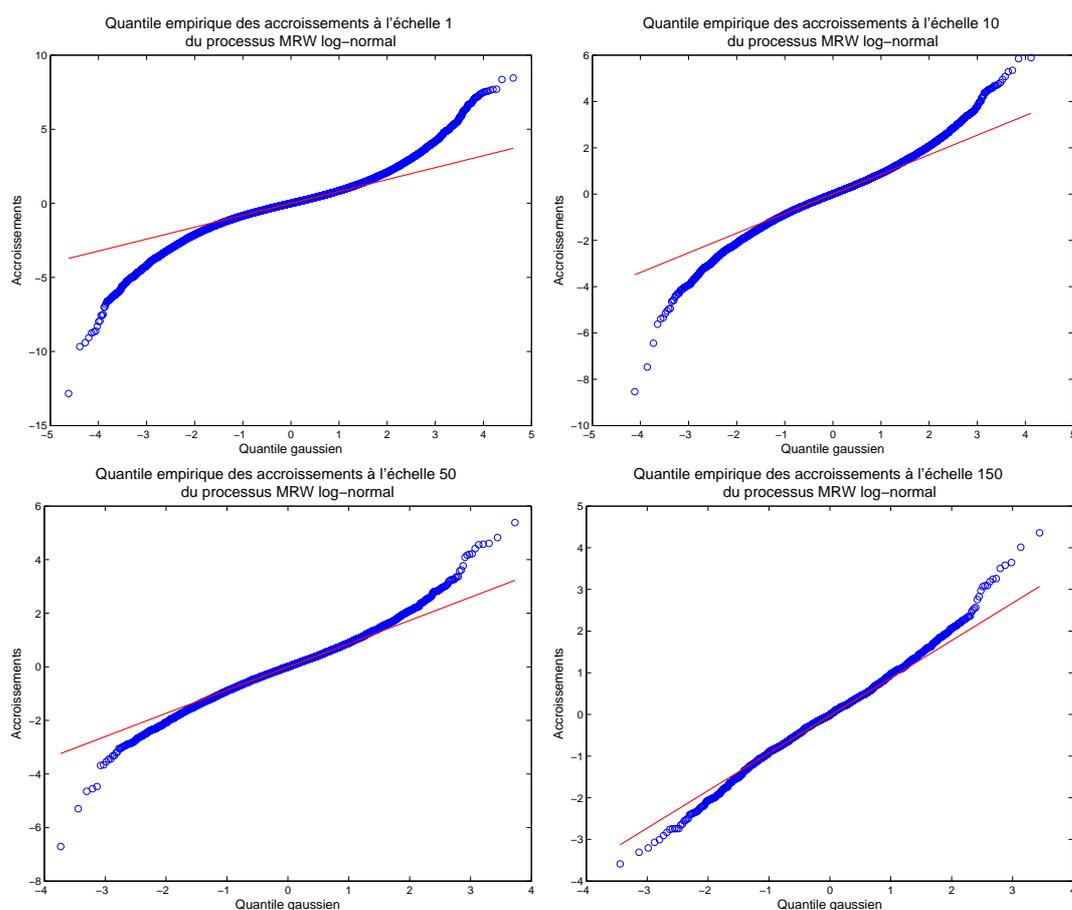


FIG. 3.6 – Quantiles empiriques des accroissements normalisés par leur écart-type de la réalisation de la figure 3.3, aux échelles 1 (en haut à gauche), 10 (en haut à droite), 50 (en bas à gauche) et 150 (en bas à droite). Les tracés montrent clairement les queues épaisses des distributions. Il est aussi possible d’observer la convergence de la distribution vers la gaussienne, présentée par le trait plein, lorsque l’échelle croît.

Dépendance non linéaire à décroissance lente

Nous avons montré (cf. équation (3.138)) que les carrés des accroissements d’un processus MRW log-normal ont des corrélations présentant une décroissance lente (en loi de puissance). Un calcul analogue montrerait que c’est aussi le cas des accroissements absolus. Comme le montre la figure 3.8, la fonction d’autocorrélation empiriques des accroissements absolus reste très longtemps significativement positive. Une régression ”log-log” de cette fonction d’autocorrélation en fonction du temps indique qu’elle peut être bien approximée par une loi de puissance. Ce comportement montre donc que le modèle MRW log-normal permet de reproduire le phénomène de mémoire longue, étudiée dans la section 1.2.5.

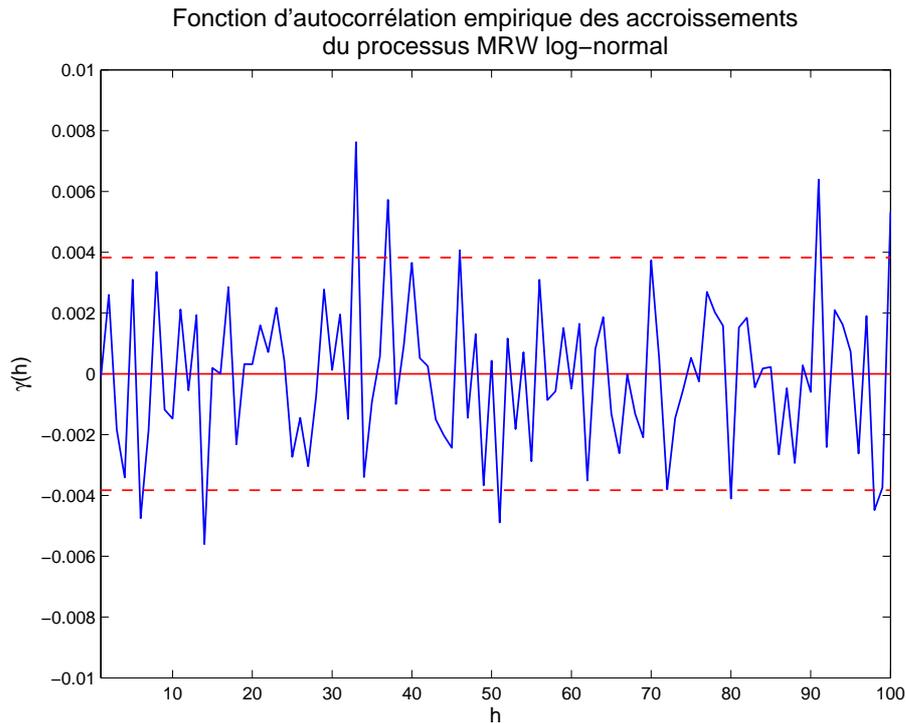


FIG. 3.7 – Fonction d'autocorrélation empirique normalisée des accroissements de la réalisation de la figure 3.3. Les traits tiretés représentent l'intervalle de confiance à 95%. Pour une meilleure lisibilité, la variance n'est pas représentée.

Accumulation des volatilités

Sur la figure 3.9 nous avons représenté les carrés des accroissements de la réalisation de la figure 3.3 dont le comportement intermittent ressemble celui de la volatilité réalisée obtenue par les carrés des rendements logarithmiques journaliers du taux de change GBP/USD, voir la figure 1.7. Le modèle MRW log-normal permet donc de reproduire le phénomène d'accumulation des volatilités, étudié dans la section 1.2.7.

Comportement des moments des accroissements

Le processus MRW log-normal possède une invariance d'échelle exacte et est multifractal (cf. théorème 3.17). Sur la figure 3.10 nous avons évalué les différents moments des accroissements absolus de la réalisation de la figure 3.3 par leur moyenne empirique (2.81). L'invariance d'échelle se traduit par un comportement linéaire aux échelles logarithmiques, ce qui est confirmé au moins pour les premières valeurs de q . Le fait que les pentes de ces comportements linéaires varient non linéairement en fonction de l'ordre du moment indique la multifractalité de la réalisation.

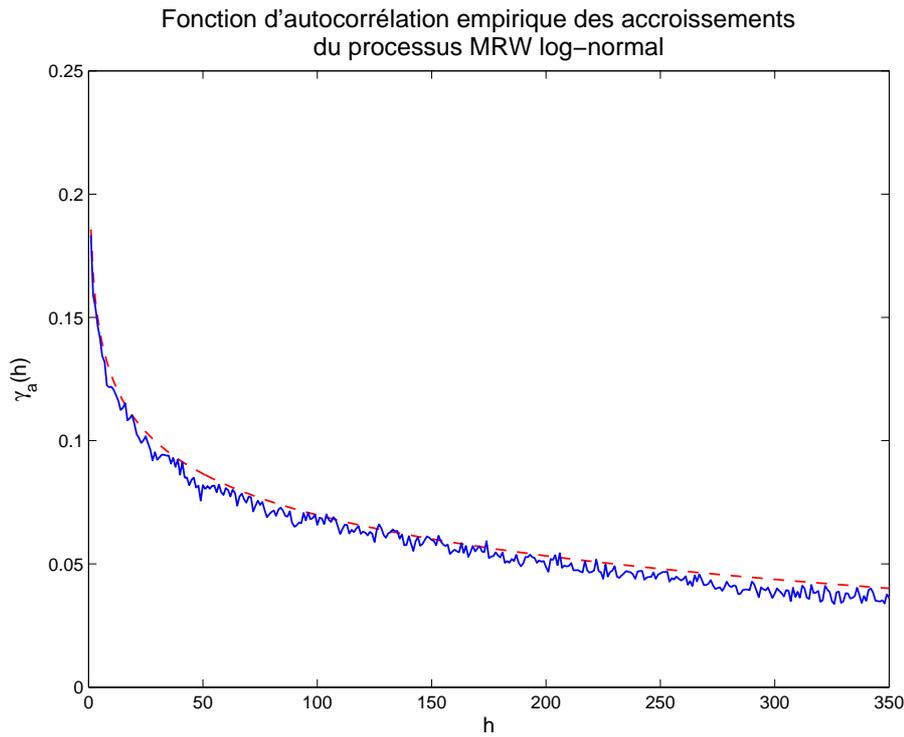


FIG. 3.8 – Fonction d'autocorrélation empirique normalisée des accroissements absolus de la réalisation de la figure 3.3. Le trait tireté correspond à la courbe régression par une loi de puissance $\gamma_a(h) = \left(\frac{T}{h}\right)^{0.02} - 1$.

Autocovariance des magnitudes

La figure 3.11 représente la fonction d'autocovariance empirique des magnitudes de la réalisation de la figure 3.3. Elle reste très longtemps significativement positive et indique donc l'existence d'une autocorrélation des magnitudes. Une régression de cette fonction d'autocovariance *versus* le logarithme du temps indique qu'elle peut être bien approximée par une fonction décroissante logarithmiquement jusqu'à zéro, comme le montre le tracé tireté sur la figure. Un tel comportement de la fonction d'autocovariance des magnitudes est justifié (cf. proposition 5.19) dans le cadre de l'étude de l'approximation limite de petit intermittence (cf. section 5.2). Le modèle MRW log-normal permet donc de reproduire le fait stylisé observé (cf. section 2.3.3).

Notons qu'une procédure semblable de régression permet d'estimer le coefficient d'intermittence λ^2 , (cf. section 6.4).

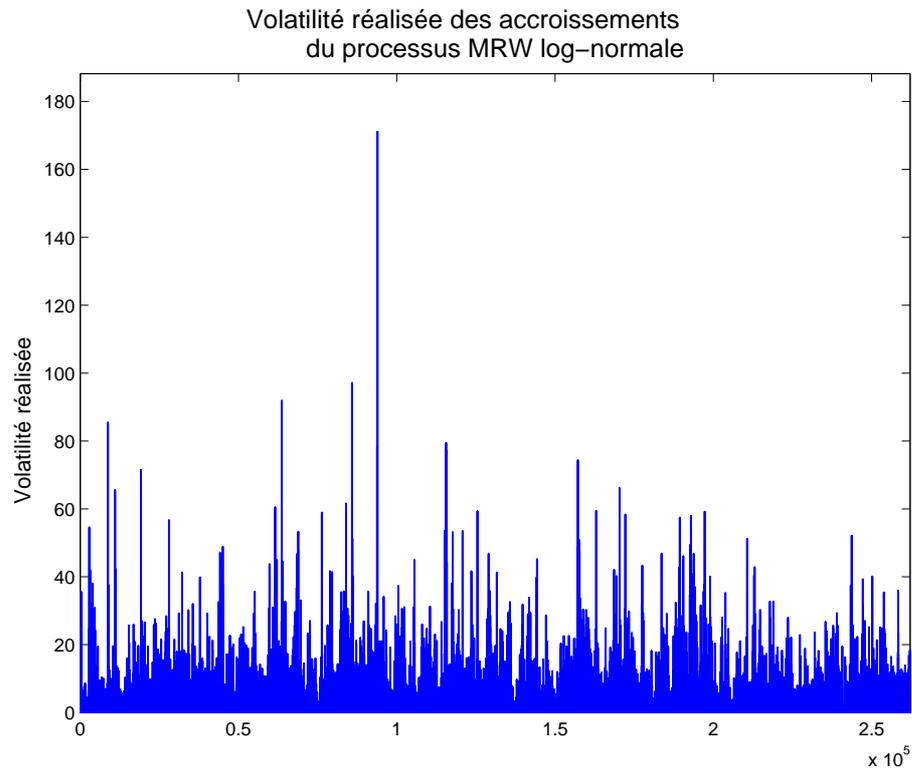


FIG. 3.9 – Carrés des accroissements de la réalisation de la figure 3.3. On observe le comportement intermittent qui ressemble à celui de la volatilité réalisée du taux de change GBP/USD, voir la figure 1.7.

3.4.5 Résumé de la section

Le modèle MRW log-normal permet de reproduire les queues épaisses (leptokurticité) des distributions des rendements logarithmiques observées dans les séries financières. Les queues varient en fonction de l'échelle et elles deviennent proche de gaussiennes pour les grandes échelles. Ce modèle permet donc de reproduire les comportements d'échelle (cf. section 1.2.3).

Ce modèle permet de reproduire l'absence de corrélation temporelle des rendements logarithmiques (cf. section 1.2.4) ainsi que les corrélations temporelles à longue portée. Il permet donc de reproduire le phénomène de mémoire longue (cf. section 1.2.5) et, en conséquence, le phénomène d'accumulation des volatilités (cf. section 1.2.7).

Grâce à son caractère multifractal, ce modèle permet de reproduire les faits stylisés multifractals observés dans les marchés, étudiés dans la section 2.3, tels que le comportement des moments absolus en fonction de l'échelle et de l'ordre (cf section 2.3.2) et la corrélation des magnitudes des rendements logarithmiques (cf. section 2.3.3).

Le modèle MRW log-normal permet de reproduire la plupart des faits stylisés observés dans

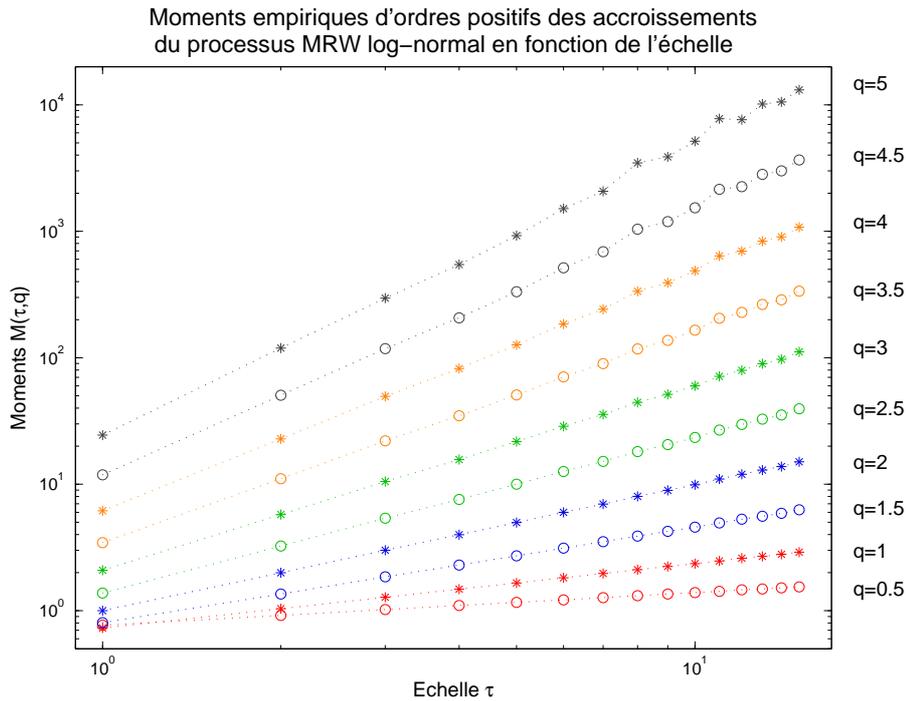


FIG. 3.10 – Fonction de partition (2.81) des accroissements de la réalisation de la figure 3.3, pour l'ordre q variant de 0.5 à 5 et l'échelle τ variant de 1 à 15. Le comportement linéaire des moments aux échelles logarithmiques confirme que les rendements logarithmiques possède une invariance d'échelle.

les marchés financiers. Cependant, il existe certains faits stylisés qui ne sont pas reproduits à l'aide de ce modèle. Ces faits stylisés caractérisent l'asymétrie dans les rendements observés.

Puisque le processus MRW log-normal est construit à l'aide de la subordination (3.102) d'un mouvement brownien la loi de distribution d'accroissements ne possède pas de skewness (1.8). De même, puisque le mouvement brownien est indépendant de la mesure MRM, le modèle MRW log-normal ne permet pas de reproduire l'effet de levier observé dans les marchés financiers (cf. section 1.2.8).

Bouchaud et Pochart dans [PB02] ont proposé de la construction d'une généralisation du mouvement MRW, le processus **SMRW** (*Skewed Multifractal Random Walk*), qui prend en compte l'effet de levier sans pour autant perdre ses propriétés multifractales. Néanmoins, le processus SMRW n'est pas un processus à temps continu, c'est un processus à temps discret. Le principal problème rencontré lorsque l'on fait tendre le pas de discrétisation vers 0 étant qu'à cette limite toutes les propriétés d'asymétrie (skewness et levier) disparaissent. La construction de processus à temps continu, autosimilaires au sens stochastique, à accroissements stationnaires et présentant skewness et effet de levier reste encore un problème ouvert aujourd'hui.

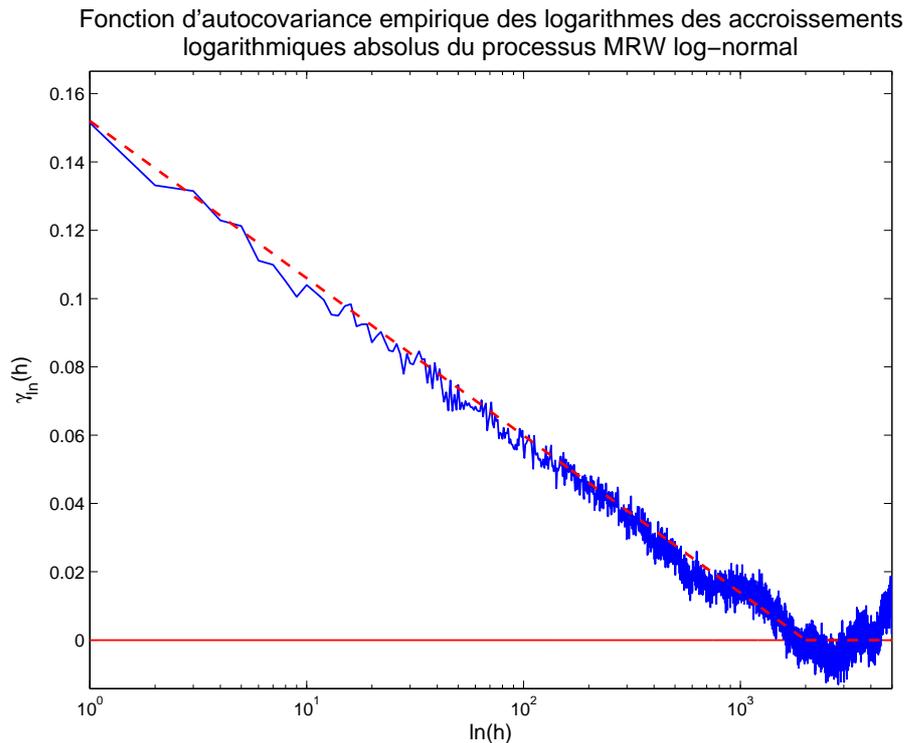


FIG. 3.11 – Fonction d'autocovariance empirique des magnitudes de la réalisation de la figure 3.3. Le trait tireté correspond à la courbe de régression par une fonction décroissant logarithmiquement $0.02 \ln(2000/h) \mathbf{1}_{\{h \leq 2000\}}$. On observe la décroissance de même type que celle de la fonction d'autocovariance empirique des magnitudes des rendements logarithmiques journaliers du taux de change GBP/USD, voir la figure 2.6.

3.5 Conclusion

Dans ce chapitre nous avons étudié la construction (cf. sections 3.1, 3.2 et 3.4) et les propriétés des processus MRW log-infiniment divisibles (cf. section 3.1.2, 3.2.2 et 3.4.2) qui forment à ce jour la seule classe de processus à accroissements stationnaires, invariants d'échelle et multifractals.

Il apparaît clairement que les processus MRW log-infiniment divisibles et, en particulier, les processus MRW log-normaux, ouvrent de très nombreux horizons tant à un niveau théorique (propriétés de régularité, caractérisation du spectre de singularités) qu'à un niveau applicatif (simulations numériques, estimation des paramètres).

Il est naturel d'utiliser le modèle MRW log-normal pour modéliser les variations des séries chronologiques financières, puisqu'il a de nombreux avantages par rapport aux autres modèles couramment utilisés en finance : il est à accroissements stationnaires, peut être vu comme un processus limite de marches aléatoires (d'où un caractère causal), rend compte parfaitement des propriétés multifractales des séries financières (sans faire

intervenir de rapport d'échelle privilégié) ainsi que de la structure complexe de corrélation des accroissements aux différentes échelles.

Ce processus permet avec trois paramètres uniquement de rendre compte de très nombreuses propriétés statistiques de séries financières (cf. section 3.4.4).

Événements extrêmes

La théorie des **valeurs extrêmes** classique cherche à décrire le comportement des occurrences maximales ou minimales d'un ensemble de variables aléatoires tel que, par exemple, les valeurs d'un processus temporel sur une période de temps fixe. Cette théorie a des applications pertinentes dans plusieurs domaines : en physique statistique dans l'étude des systèmes désordonnés à très basse température, en géologie dans l'étude des tremblements de Terre, en météorologie dans l'étude des averses, des inondations, des tempêtes et des canicules, en finance, dans l'étude des risques des marchés financiers [LLR83, EKM97, Sor03]. Les **événements extrêmes** sont particulièrement significatifs (car notablement probables) pour les phénomènes caractérisés par des densités de probabilité dont les queues décroissent en loi de puissance.

On se propose de considérer le problème important qui consiste à estimer et interpréter des exposants des queues de distributions dans le cas de processus de type cascades où, comme nous l'avons vu, les corrélations sont importantes. Cela nous permettra notamment d'identifier la présence de corrélations importantes sur un échantillon comme l'un des mécanismes possibles qui conduisent à l'observation de "queues épaisses". Ce type problème a déjà été abordé dans les travaux [Bou01, Sor03, New05, MBK06].

La théorie des valeurs extrêmes [Bor22, Mis23, FT28, Mis54] a été développée sous l'hypothèse que les variables aléatoires considérées sont indépendantes et identiquement distribuées. Une extension de cette théorie aux processus stationnaires (ou suites stationnaires) a été développée avec un affaiblissement de l'hypothèse d'indépendance, celle-ci étant remplacée par une hypothèse de dépendance "courte portée" [LLR83]. Dans le cas Gaussien, cela revient à supposer que la covariance du processus décroît suffisamment vite. Néanmoins, très peu de résultats théoriques sont connus si les corrélations ne décroissent pas de façon assez rapide ou si le régime asymptotique n'est pas atteint.

En finance, connaître (estimer) les exposants de queues des lois de probabilité des rendements est important en pratique pour évaluer le risque, mais aussi pour la modélisation en général, car beaucoup de modèles reposent sur l'hypothèse de l'existence de la variance des rendements [MF00]. La BIS (*Bank for International Settlements*) a par exemple établi des

règles à suivre par les banques pour contrôler leurs risques, mais la plupart des modèles utilisés dans ces règles sont encore basés sur l'hypothèse que les actifs financiers sont distribués suivant une distribution normale. Avoir plus d'information sur l'exposant de queue des séries financières est donc une question d'intérêt en soi.

Ce chapitre commence par une brève revue des propriétés des exposants des queues des distributions et notamment des méthodes classiques d'estimation (cf. section 4.1). Puis nous effectuons une étude statistique des valeurs extrêmes dans le cadre d'un modèle multifractal (cf. section 4.2).

Plusieurs études récentes [GMAS98, PGA⁺99, GPA⁺99, BP03] trouvent que les exposants de queues sont proches de 3, alors que le modèle MRW prédit un exposant de queue supérieur de l'ordre de 30. Dans les sections 4.2.3 et 4.3, nous montrons que les séries financières ne sont pas de taille suffisante pour atteindre les régimes asymptotiques des estimateurs de queues de distributions et que cela suffit pour expliquer un tel écart.

4.1 Exposants des queues des distributions

Nous rappelons que l'exposant de queue de distribution est introduit par la définition 1.5.

4.1.1 Estimation des exposants des queues

Estimer les exposants de queue (sans distribution connue *a priori*) est une tâche difficile et est un sujet actif de recherche en statistique. Pour déterminer l'exposant de queue d'une distribution, il faut énormément d'observations (puisque lié aux valeurs extrêmes de la distribution). Diverses méthodes issues de la théorie des valeurs extrêmes existent, et ce sujet est assez vaste.

Nous considérons un échantillon de N observations Z_1, \dots, Z_N indépendantes et identiquement distribuées de fonction de répartition F inconnue. Nous supposons que cette distribution est à queue lourde. Nous noterons les statistiques d'ordre décroissantes par $X_1 \geq \dots \geq X_N$. Décrivons brièvement différents estimateurs classiques de l'exposant de queue q_c associé à F . Soient j et k tels que $1 \leq j \leq k \ll N$,

- l'**estimateur de Zipf** ou la méthode de rang-fréquence [Zip29, Man97] est basé sur le comportement de la fonction de répartition

$$1 - F(X_j) \sim \frac{j}{N}, \text{ et alors } X_j \sim \left(\frac{j}{N}\right)^{-1/q_c} \quad (4.1)$$

d'où on déduit que l'exposant de queue peut être estimé comme la pente $\widehat{q}_c^Z(k, N)$ de la courbe rang-fréquence $[\ln(X_j), \ln(j)]_j$, pour $j = 1, \dots, k$.

Dans [New05], on démontre que l'estimateur de Zipf est biaisé, c'est pourquoi on utilise les estimateurs alternatifs comme ceux de Hill et Pickands [EKM97].

Soit k dépendant de N tel que $k(N) = o(N)$,

– l'estimateur de Hill est défini par

$$\hat{q}_c^H(k, N) = \frac{k-1}{\sum_{j=1}^{k-1} \ln(X_j/X_k)}, \quad (4.2)$$

– l'estimateur de Pickands est donné par

$$\hat{q}_c^P(k, N) = \frac{\ln(2)}{\ln\left(\frac{X_k - X_{2k}}{X_{2k} - X_{4k}}\right)}. \quad (4.3)$$

Concernant les estimateurs donnés ci-dessus et leurs propriétés, on peut aller voir les références suivantes [Hil75, Pic75].

4.1.2 Loi de distribution de maximum

L'étude mathématique des estimateurs, introduits dans la section précédente 4.1.1, telle que la consistance, le biais, la normalité asymptotique, repose sur la théorie des valeurs extrêmes [LLR83, EKM97]. Le résultat fondamental de cette théorie est le théorème de Fisher et Tippett [FT28] qui caractérise toutes les formes possibles de la distribution de la plus grande valeur Y_N d'un échantillon de taille N . A mesure que la dimension de l'échantillon augmente, la fonction de répartition $F_N(x) = \mathbb{P}[Y_N \leq x]$ de se rapproche asymptotiquement d'une distribution de Fréchet, Weibull ou Gumbel. Plus précisément, il existe deux suites, $a_N > 0$ et b_N telles que

$$\lim_{N \rightarrow +\infty} F_N\left(\frac{x - b_N}{a_N}\right) = F_Y(x) \quad (4.4)$$

où $F_Y(x)$ est une distribution de Fréchet,

$$F_Y(x) = \begin{cases} \exp(-x^{-q_c}), & \text{si } x > 0, \\ 0, & \text{sinon} \end{cases} \quad (4.5)$$

de Weibull

$$F_Y(x) = \begin{cases} \exp(-(-x)^{-1/q_c}), & \text{si } x < 0, \\ 1, & \text{sinon} \end{cases} \quad (4.6)$$

ou de Gumbel

$$F_Y(x) = \exp(-e^{-x}). \quad (4.7)$$

L'existence des suites a_N , et b_N , dans la définition (4.4) d'une loi des valeurs extrêmes repose en fait sur des propriétés vérifiées par la loi F . Des conditions nécessaires et suffisantes sur F ont été obtenues [Gal87] pour l'existence de la loi limite $F_Y(x)$.

Ces trois distributions peuvent être incorporées en une classe plus générale qui peut être représentée à la normalisation près sous la forme suivante, dite représentation de Jenkinson

$$F_Y^\xi(x) = \begin{cases} \exp(-(1 + \xi x)^{-1/\xi}), & \text{si } 1 + \xi x > 0, \\ 1, & \text{sinon.} \end{cases} \quad (4.8)$$

Le paramètre ξ est appelé **indice de queue**. Le cas $\xi > 0$ correspond à la loi de Fréchet, le cas $\xi < 0$ correspond à la loi de Weibull et le cas limite lorsque ξ tend vers zéro correspond à la loi de Gumbel.

Si la loi du maximum Y (normalisé comme ci-dessus) de N variables aléatoires indépendantes et identiquement distribuées de loi F converge vers F_Y^ξ alors on dit que le maximum est attiré par F_Y^ξ et, par extension, on dit que la loi F appartient au **domaine d'attraction** de F_Y^ξ [Mis54]. On dispose de caractérisations très précises du domaine d'attraction de chaque loi F en fonction de comportement de ces queues [BGT87]. Ces caractérisations sont souvent très techniques et difficilement vérifiables par le praticien, voir par exemple [Gal87, Res87], aussi nous n'entrerons pas dans les détails.

On peut, par exemple, montrer que la loi normale, la loi exponentielle et la loi log-normale appartiennent au domaine d'attraction de la loi de Gumbel. Cette propriété est due au fait que toutes ces lois possèdent des queues de distribution peu épaisses.

Les lois de Pareto, de Cauchy, de Student appartiennent au domaine d'attraction des lois de Fréchet. Ces lois se caractérisent par la présence de queues de distribution lourdes (non exponentielle) ayant tendance à générer de grandes valeurs.

La loi uniforme et les lois qui ont un support fini mais avec une asymptote en leur point terminal, par exemple les lois bêta, appartiennent au domaine d'attraction de la loi de Weibull.

4.2 Statistique des valeurs extrêmes multifractales

4.2.1 Résultats dans le cas des processus stationnaires

Les résultats de l'étude mathématique des estimateurs de l'exposant de queue, introduits dans le cas des variables indépendantes et identiquement distribuées, restent valides lorsque l'on utilise l'extension de la théorie des valeurs extrêmes aux processus aléatoires stationnaires autocorrélés dans les conditions qui assurent l'indépendance asymptotique des maximums [LLR83]. La différence principale est que l'on remplace le nombre des variables indépendantes N par le nombre "effectif" θN avec $0 < \theta < 1$, où la valeur θ , dit **l'indice extrême**, quantifie l'effet de dépendance. Dans le cas d'un processus Gaussien, les résultats théoriques restent valides lorsque la fonction d'autocovariance $\rho(x)$ décroît de façon suffisamment rapide

$$\rho(x) \ln(x) \xrightarrow{x \rightarrow +\infty} 0. \quad (4.9)$$

Une justification intuitive de ce résultat est proposée dans [CD01].

Le but principal de cette partie est d'étudier le comportement des estimateurs de queue appliqués aux échantillons provenant des processus multifractals. Nous suivons essentiellement l'article de Muzy, Bacry et Kozhemyak [MBK06].

4.2.2 Limite asymptotique

La fonction d'autocovariance $\rho(x)$ d'un processus multifractal décroît très lentement jusqu'à l'échelle intégrale T au delà de laquelle les valeurs du processus deviennent indépendantes. Ainsi, nous admettons que la condition (4.9) n'est pas vérifiée entre l'échelle de l'observation τ et l'échelle intégrale T . En revanche, elle est vérifiée au delà de l'échelle intégrale T . Pour être plus précis, la condition (4.9) se présente lorsque le nombre d'observation N tend vers l'infini. Cependant, il existe plusieurs possibilités d'atteindre la limite asymptotique $N \rightarrow +\infty$.

Soit L la taille de l'intervalle d'observation (dans une certaine unité de temps) et τ l'échelle d'échantillonnage (dans la même unité). Le nombre d'échantillons est alors égal à

$$N = \frac{L}{\tau}, \quad (4.10)$$

et le nombre d'échelles intégrales contenues dans la fenêtre d'observation N_T et le nombre d'échantillons contenus dans l'échelle intégrale N_τ sont donnés respectivement par

$$N_T = \frac{L}{T}, \quad N_\tau = \frac{T}{\tau}, \quad N = N_\tau N_T. \quad (4.11)$$

Introduisons l'exposant χ pour relier les valeurs N_T et N_τ

$$N_T \sim N_\tau^\chi. \quad (4.12)$$

Notons que cet exposant χ a été interprété par Mandelbrot comme "embedding dimension" dans [Man90].

Plus précisément, quand N tend vers l'infini

- le régime $\chi = 0$, "**haute fréquence**", correspond au cas limite $\tau \rightarrow 0$, tandis que L et T restent fixées. Dans ce cas, nous admettons que la condition (4.9) n'est pas vérifiée.
- le régime $\chi = +\infty$, "**historique infini**", correspond au cas limite $L \rightarrow +\infty$, tandis que τ et T restent fixées. Dans ce cas la condition (4.9) est satisfaite.

Comme nous allons voir dans la section suivante, le comportement des estimateurs de queues dépend fortement de la valeur de χ . Dans le premier cas ($\chi = 0$), rien ne garantit que les méthodes classiques puissent être appliquées pour estimer les exposants de queue, alors que dans le deuxième cas ($\chi = +\infty$), nous espérons que les résultats de la théorie des valeurs extrêmes restent valides et que les estimateurs soient consistants.

4.2.3 Comportement des estimateurs classiques

Le but de cette section est l'étude statistique des estimateurs classiques appliqués aux $k(N)$ valeurs extrêmes observées dans le régime asymptotique $N \rightarrow +\infty$ et $k(N) \rightarrow +\infty$. La limite $N \rightarrow +\infty$ est considérée conformément aux régimes décrits dans la section précédente par l'exposant χ . Une paramétrisation commode pour $k(N)$ est

$$k(N) \sim N^\nu, \quad (4.13)$$

avec $0 \leq \nu \leq 1$, la valeur $\nu = 0$ sera interprétée comme $k(N) \sim \ln(N)$. Nous notons $\widehat{q}_c(\nu, \chi)$ les valeurs de l'estimateur de l'exposant de queue en fonction de ν et χ .

Nous considérons l'exemple de l'estimateur de Pickands, les mêmes arguments pouvant être appliqués aux autres estimateurs tels que l'estimateur de Hill.

Théorème 4.1

Nous supposons que le formalisme multifractal, étudié dans la section 2.2.5, est vérifié. Définissons $\alpha_{\nu, \chi}$ tel que

$$k \sim N^\nu \sim N_T \left(\frac{\tau}{T} \right)^{-D^*(\alpha_{\nu, \chi})}, \quad (4.14)$$

où $D^(\alpha)$ est la transformée de Legendre de l'exposant multifractal. Dans ce cas, l'estimateur de Pickands $\widehat{q}_c^P(\nu, \chi)$ vérifie la relation*

$$\widehat{q}_c^P(\nu, \chi) \simeq q_{\nu, \chi}, \quad (4.15)$$

où $q_{\nu, \chi}$ est la valeur de la dérivée de $D^(\alpha)$ en $\alpha_{\nu, \chi}$*

$$q_{\nu, \chi} = \frac{dD^*}{d\alpha}(\alpha_{\nu, \chi}). \quad (4.16)$$

Preuve - 4.1 -

Nous donnons ici simplement une preuve basée sur des arguments "heuristiques". Nous renvoyons à la référence [MBK06], pour une approche mathématique plus rigoureuse fondée sur la résolution d'équations itératives non-linéaires en termes de fronts progressifs.

Puisque le formalisme multifractal est vérifié, la partie positive de la transformée de Legendre $D^*(\alpha)$ de l'exposant multifractal $\zeta(q)$ correspond au spectre de singularité $D(\alpha)$, voir la définition 2.18. En accord avec la définition (4.12) de χ , nous avons

$$N_T \sim N_\tau^\chi = \left(\frac{\tau}{T} \right)^{-\chi}, \quad (4.17)$$

$$N \sim N_\tau^{1+\chi} = \left(\frac{\tau}{T} \right)^{-(1+\chi)}, \quad (4.18)$$

et donc nous pouvons déduire que $\alpha_{\nu, \chi}$ vérifie la relation suivante

$$D^*(\alpha_{\nu, \chi}) \sim \nu(1 + \chi) - \chi. \quad (4.19)$$

Considérons ensuite $\alpha'_{\nu, \chi} = \alpha_{\nu, \chi} + \epsilon_1$, tel que

$$2k \sim N_T \left(\frac{\tau}{T} \right)^{-D^*(\alpha'_{\nu, \chi})} \simeq N_T \left(\frac{\tau}{T} \right)^{-D^*(\alpha_{\nu, \chi}) - q_{\nu, \chi} \epsilon_1}, \quad (4.20)$$

où $q_{\nu, \chi}$ est la valeur de la dérivée de $D^*(\alpha)$ en $\alpha_{\nu, \chi}$

$$q_{\nu, \chi} = \frac{dD^*}{d\alpha}(\alpha_{\nu, \chi}). \quad (4.21)$$

Dans le même genre d'idée, nous définissons $\alpha''_{\nu,\chi} = \alpha_{\nu,\chi} + \epsilon_2$, tel que

$$4k \sim N_T \left(\frac{\tau}{T} \right)^{-D^*(\alpha''_{\nu,\chi})} \simeq N_T \left(\frac{\tau}{T} \right)^{-D^*(\alpha_{\nu,\chi}) - q_{\nu,\chi} \epsilon_2}. \quad (4.22)$$

Grâce aux relations

$$X_k - X_{2k} \sim X_k \left(1 - \left(\frac{\tau}{T} \right)^{\epsilon_1} \right), \quad (4.23)$$

$$X_{2k} - X_{4k} \sim X_k \left(\left(\frac{\tau}{T} \right)^{\epsilon_1} - \left(\frac{\tau}{T} \right)^{\epsilon_2} \right), \quad (4.24)$$

l'estimateur de Pickands, défini par (4.3), vérifie la relation suivante

$$\widehat{q}_c^P(\nu, \chi) \simeq \ln(2) / \ln \left(\frac{1 - 2^{-1/q_{\nu,\chi}}}{2^{-1/q_{\nu,\chi}} - 4^{-1/q_{\nu,\chi}}} \right) = q_{\nu,\chi}. \quad (4.25)$$

■

Notons ce le même résultat peut être obtenu pour l'estimateur de Hill.

Exemple log-normal

Considérons le cas d'un processus MRW log-normal dont l'exposant $\zeta(q)$ est donné par

$$\zeta_X^{\ln}(q) = \frac{q}{2}(1 + 2\lambda^2) - \frac{q^2\lambda^2}{2}. \quad (4.26)$$

et la transformée de Legendre $D^*(\alpha)$ correspondante

$$D^*(\alpha) = -\frac{\alpha^2}{2\lambda^2} + \alpha \left(1 + \frac{1}{2\lambda^2} \right) - \frac{\lambda^2}{2} \left(1 - \frac{1}{2\lambda^2} \right)^2. \quad (4.27)$$

Pour calculer l'estimateur de queue $\widehat{q}_c(\nu, \chi)$, on doit résoudre les équations (4.19) et (4.16), ce qui amène au résultat suivant

$$\widehat{q}_c^{\ln}(\nu, \chi) = \sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}}. \quad (4.28)$$

Notons que la valeur estimée $\widehat{q}_c^{\ln}(\nu, \chi)$ de l'exposant de queue, qui est de l'ordre $\frac{1}{\lambda}$, est différente de la valeur non conditionnelle, qui est de l'ordre $\frac{1}{\lambda^2}$, d'après l'équation (3.134).

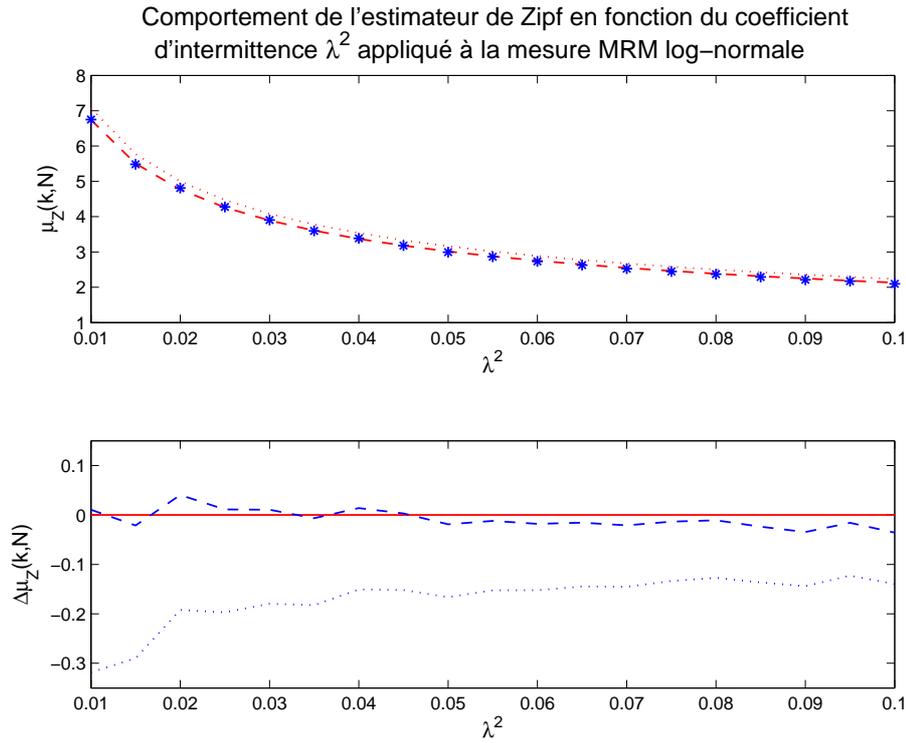


FIG. 4.1 – L'étude par la méthode de Monte Carlo du comportement de l'estimateur de Zipf appliqué à la mesure MRM log-normale en fonction du coefficient d'intermittence. Le nombre de tirages de Monte Carlo est égal à 10^4 , la taille d'échantillon N est égale à 2^{14} , le temps intégral T est égal à 2^{10} et le nombre de valeurs maximales utilisées k est égal à 2^4 , ce qui correspond aux coefficients ν égal à $2/7$ et χ égal à $2/5$. En haut, le trait bleu correspond aux résultats de l'estimation, le trait pointillé correspond aux valeurs théoriques (4.29) de l'estimateur de queue et le trait tireté correspond aux valeurs théoriques corrigées (voir dans le texte). En bas, les écarts entre les résultats de l'estimation et les valeurs théoriques sans correction (le trait pointillé) et avec la correction (trait tireté).

4.2.4 Expériences numériques

Dans la figure 4.1 nous avons représenté le comportement de l'estimateur de Zipf, introduit dans la section 4.1.1, appliqué à la mesure MRM log-normale en fonction du coefficient d'intermittence λ^2 . Le tracé a été obtenu par la méthode de Monte Carlo de 10^4 tirages. Sur la même figure nous avons représenté les valeurs théoriques de l'estimateur de queue

$$\hat{q}_c^{\ln}(\nu, \chi) = \sqrt{\frac{(1-\nu)(1+\chi)}{2\lambda^2}}. \quad (4.29)$$

Cette dernière formule est une conséquence directe de l'équation (4.28).

Les valeurs obtenue par les expériences numériques sont assez proches de celles théoriques. Néanmoins, nous avons observé un biais dépendant de la taille d'échantillon N et du temps

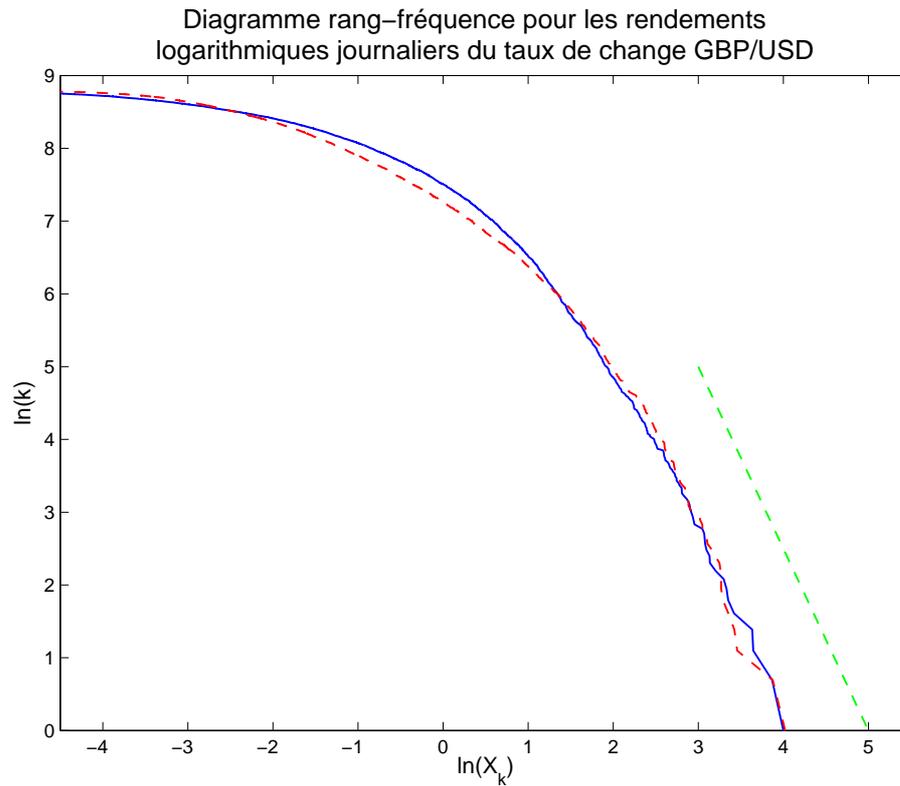


FIG. 4.2 – Les diagrammes rang–fréquence pour les carrés des rendements logarithmiques journaliers du taux de change GBP/USD (le trait plein) et pour une réalisation de la mesure MRM log-normale de paramètres $T = 2^{13}$ et $\lambda^2 = 0.05$ (le trait tireté rouge). On observe que les graphiques ont les comportements très similaires l'un à l'autre. La régression donne l'estimation de l'exposant de queue $\hat{q}_c^Z \simeq 2.5$ (le trait tireté vert).

intégral T . Nous expliquons la présence de ce biais par le fait que dans la démonstration du théorème 4.1 nous avons supposé que l'échantillon contient les observations indépendantes par paquets de taille T , ce qui n'est pas vérifié par la mesure MRM log-normale.

Pour corriger ce biais observé, il est nécessaire d'utiliser, dans le calcul du coefficient χ , une valeur effective du temps intégral qui est approximativement égale à $2T$, ce qui correspond au coefficient χ égal à $4/11$, dans le cas représenté dans la figure 4.1. Les traits tiretés visualisent ces corrections.

4.3 Application aux séries financières

Plusieurs études récentes [GMAS98, PGA⁺99, GPA⁺99, BP03] de données financières à haute fréquence trouvent que les queues de la distribution cumulée des rendements logarithmiques d'actions et d'indices sont représentées par une loi de puissance avec un

exposant de queue proche de 3. C'est la célèbre "loi cubique". Ces observations semblent contredire le modèle multifractal pour les rendements des actifs financiers.

Cependant, une simple application numérique du résultat obtenu précédemment montre que, dans le cadre du modèle MRW log-normal, la valeur de l'exposant de queue théorique et celle obtenue à l'aide d'un estimateur classique diffèrent d'un ordre de grandeur. En effet, Les valeurs typiques en finance sont $\tau \in [10^{-2}, 1]$ jour et $T \simeq L \simeq 10$ ans, ce qui correspond, grossièrement, à la valeur de $\chi \simeq 0$. En prenant $\nu \simeq 0.5$, l'estimateur classique de queue (4.28) est approximativement égal à

$$\hat{q}_c^{\ln}(\nu, \chi) \simeq \frac{1}{\sqrt{\lambda^2}} \simeq 5, \quad (4.30)$$

où nous avons utilisé la valeur typique 0.04 du coefficient d'intermittence λ^2 (voir les résultats d'estimations dans la section 6.2.3). Par ailleurs, la loi distribution des accroissements du processus MRW log-normale possède des queues épaisses (voir (3.134)) dont l'exposant est donnée par

$$q_c = \frac{1}{\lambda^2} \simeq 50. \quad (4.31)$$

Il y a donc bien une différence significative entre l'exposant de queue estimé typiquement sur un échantillon et la valeur théorique non conditionnelle.

Dans la figure 4.2 nous avons représenté les diagrammes rang-fréquence (voir la définition de l'estimateur de Zipf dans la section 4.1.1) pour les carrés des rendements logarithmiques journaliers du taux de change GBP/USD et pour une réalisation de la mesure MRM log-normale. Il est simple de voir que les graphiques ont les comportements très similaires l'un à l'autre. La mesure MRM log-normale est de paramètres $L = T = 2^{13}$, ce qui est proche de la taille de la série financière, et $\lambda^2 = 0.05$, ce qui correspond à l'exposant de queue $q_c = \frac{1}{\lambda^2} = 20$. Néanmoins, l'estimateur de Zipf \hat{q}_c^Z donne la valeur de l'exposant de queue égale à 2.5.

Notons que pour les petites valeurs de $\ln(X_k)$, nous pouvons observer l'écart entre les graphiques. En fait, l'approximation de la volatilité réalisée journalière par le carré du rendement logarithmique journalier devient très bruitée pour les petites valeurs des rendements.

4.4 Conclusion

Le but principal de ce chapitre était de mettre en évidence le fait que l'analyse d'exposants de queue des distributions des rendements logarithmiques d'actifs financiers est sensible à la présence de corrélation de la volatilité à longue portée. Les valeurs observées ne correspondent pas aux valeurs des exposants de queue des lois de distribution non conditionnelles. Autrement dit, on observe le comportement **non ergodique** des quantités dépendant des valeurs extrêmes, par exemple des moments d'ordre élevé des rendements.

L'étude de ce chapitre repose ici sur une approche phénoménologique. Elle a été menée de façon parallèle à partir de la résolution de l'équation d'itération de la cascade par une méthode de fronts progressifs [MBK06].

Approximation limite de petite intermittence

Les résultats de ce chapitre sont essentiels au reste de ce travail puisqu'ils sont la base de tous les chapitres qui suivent.

L'autosimilarité (déterministe ou stochastique) est une propriété très utile car la connaissance du processus à une échelle suffit pour en déduire son comportement à toutes échelles. Le théorème 3.15 met en évidence le caractère autosimilaire stochastique de la mesure MRM et du processus MRW. Ce qui signifie que, ces objets possèdent des propriétés de stabilité par changement d'échelle. Cependant, le théorème 3.15 n'a pas d'intérêt pratique, car il ne fournit pas de distribution à l'échelle l mais indique seulement comment elle peut varier avec l'échelle. En d'autres termes, il est très difficile de travailler avec le processus MRW ou avec la mesure MRM puisque on ne contrôle pas, en pratique, leurs distributions.

Comme nous le verrons dans le chapitre 6, les valeurs du coefficient d'intermittence λ^2 couramment observées pour les séries financières sont autour de 0.02. Rappelons que cette valeur quantifie la courbure de $\zeta(q)$ et "mesure" donc l'écart du processus MRW log-normal (resp. de la mesure MRM log-normale) au mouvement brownien (resp. mesure de Lebesgue). Il est donc naturel d'étudier les propriétés de ce processus sous condition de faible intermittence $\lambda^2 \ll 1$.

Dans ce chapitre, après une série de propositions techniques sur les moments de la mesure MRM log-normale (cf. section 5.1) et du logarithme de cette mesure (cf. section 5.2), nous montrons que, sous condition de faible intermittence $\lambda^2 \ll 1$, la mesure MRM log-normale peut être approximée par un processus log-normal, la **magnitude renormalisée** $\{\Omega(t)\}_t$ (cf. section 5.3), dont la fonction d'autocovariance admet des formules explicites. Le sens de cette approximation est multiple.

Le théorème 5.12 (et son corollaire 5.13) montre que le logarithme de la mesure MRM normalisée par λ converge en loi vers la magnitude renormalisée dans la limite où λ^2 tend

vers 0 :

$$\left\{ \frac{1}{2\lambda} \ln \left(\frac{\delta_\tau M(t)}{\tau} \right) \right\}_t \xrightarrow{\mathcal{L}} \left\{ \frac{\delta_\tau \Omega(t)}{\tau} \right\}_t. \quad (5.1)$$

Le théorème 5.14 compare les moments généralisés de ces deux quantités et montrent qu'ils sont proches lorsque λ^2 est petit. Plus précisément introduisons la notation suivante :

Définition 5.1

Soient $\{X_\lambda(t)\}_t$ et $\{Y_\lambda(t)\}_t$ des processus dépendant du paramètre λ^2 . Si les ordres zéro et les premiers ordres non triviaux qui suivent des développements en série de Taylor en λ^2 de leurs moment généralisés coïncident alors nous écrivons

$$X_\lambda(t) \stackrel{\lambda}{\simeq} Y_\lambda(t). \quad (5.2)$$

Le théorème 5.14 établit la relation

$$\ln \left(\frac{\delta_\tau M(t)}{\tau} \right) \stackrel{\lambda}{\simeq} 2\lambda \frac{\delta_\tau \Omega(t)}{\tau}. \quad (5.3)$$

Enfin le théorème 5.16 établit une relation similaire non plus sur les moments généralisés du logarithme de la mesure mais sur les moments généralisés de la mesure elle-même.

Dans la section 5.5, des résultats similaires sont établis sur les moments généralisés du processus MRW log-normal (théorème 5.17) et de son logarithme (théorème 5.18). Ce premier théorème est à la base des méthodes de prédiction linéaire et quadratique des chapitres 7 et 9. Quant au second théorème, il est à la base des méthodes d'estimation des paramètres présentées dans le chapitre 6 et est à la base de la méthode de prédiction logarithmique des chapitres 7, 8 et 9.

A la fin de ce chapitre, nous présentons des résultats d'expériences numériques qui permettent d'apprécier la précision de l'approximation log-normale au sens des distributions.

5.1 Moments d'ordre généralisés de la mesure MRM log-normale

5.1.1 Représentation intégrale des moments généralisés

Dans le cadre de l'étude des propriétés du processus MRW log-normale, nous avons effectué le calcul de moments d'ordres positifs (à un point) de leurs accroissements (3.130). Ces moments admettent une forme analytique obtenue à l'aide de la formule de Selberg [Sel44].

Dans cette section, nous considérons les moments généralisés d'ordres positifs (à n points) de la mesure MRM log-normale. Dans le cas général, à notre connaissance, ces moments n'ont pas de forme analytique. La proposition suivante fournit une représentation intégrale de ces moments généralisés.

Proposition 5.1

Soit n un nombre entier positif. Le moment généralisé de la mesure MRM log-normale correspondant aux intervalles I_1, \dots, I_n possède la représentation intégrale

$$\mathbb{E} \left[\frac{M(I_1)}{|I_1|} \dots \frac{M(I_n)}{|I_n|} \right] = \int_{I_1} \frac{du_1}{|I_1|} \dots \int_{I_n} \frac{du_n}{|I_n|} f_{\lambda^2}(u_1, \dots, u_n), \quad (5.4)$$

où la fonction $f_{\lambda^2}(u_1, \dots, u_n)$ est donnée par

$$f_{\lambda^2}(u_1, \dots, u_n) = \exp \left(4\lambda^2 \sum_{1 \leq i < j \leq n} \ln \left(\frac{T}{|u_i - u_j|} \right) \mathbf{1}_{\{|u_i - u_j| \leq T\}} \right). \quad (5.5)$$

Preuve - 5.1 -

Tout d'abord nous supposons que tous les intervalles ont la forme $I_j = [0, t_j]$.

Une simple généralisation du théorème 3.3, permet d'évaluer les moments généralisés comme la limite suivante

$$\mathbb{E} \left[\frac{M(I_1)}{|I_1|} \dots \frac{M(I_n)}{|I_n|} \right] = \lim_{l \rightarrow 0^+} \mathbb{E} \left[\frac{M_l(I_1)}{|I_1|} \dots \frac{M_l(I_n)}{|I_n|} \right]. \quad (5.6)$$

La représentation discrète (3.140) conduit à la formule

$$\mathbb{E} \left[\frac{M_l(I_1)}{|I_1|} \dots \frac{M_l(I_n)}{|I_n|} \right] = \frac{l^n}{|I_1| \dots |I_n|} \sum_{k_1=0}^{\lfloor |I_1|/l \rfloor - 1} \dots \sum_{k_n=0}^{\lfloor |I_n|/l \rfloor - 1} \mathbb{E} \left[e^{2\omega_l(k_1 l)} \dots e^{2\omega_l(k_n l)} \right]. \quad (5.7)$$

Par ailleurs, grâce au fait que $\text{Var}[\omega_l(t)] = -\mathbb{E}[\omega_l(t)]$, on peut facilement calculer

$$\mathbb{E} \left[e^{2\omega_l(k_1 l)} \dots e^{2\omega_l(k_n l)} \right] = \exp \left(\sum_{1 \leq i < j \leq n} \rho_l(k_i l - k_j l) \right). \quad (5.8)$$

Dans l'expression (5.7), on reconnaît une somme de Riemann (positives) qui converge, lorsque l tend vers zéro, vers l'intégrale multiple impropre (5.4) dans le cas où celle-ci est finie, c-à-d si le coefficient d'intermittence λ^2 est inférieur à $\frac{1}{2n}$.

En fait, on rencontre une petite difficulté pour montrer la convergence de la somme de Riemann vers l'intégrale impropre. En effet, lorsque l'on fait tendre l vers 0, des singularités apparaissent correspondant au cas où m indices parmi k_1, \dots, k_n coïncident. Ces singularités sont de l'ordre de $l^{-2m(m-1)\lambda^2}$. Le nombre de ces singularités est borné par Cl^{m-1-n} . Nous pouvons donc bien les négliger dans les sommes (5.7) pour λ^2 inférieur à $\frac{1}{2m}$, ce qui est vérifié si l'intégrale converge.

Dans le cas, où tous les intervalles I_1, \dots, I_n coïncident les uns avec les autres, l'intégrale multiple peut être évaluée à l'aide de la formule de Selberg [Sel44]. Néanmoins, dans le cas général cette intégrale n'a pas de forme analytique.

En utilisant la linéarité des intégrales, nous pouvons voir que la représentation intégrale reste valide même si l'on enlève la contrainte de la forme des intervalles I_1, \dots, I_n , qui a été faite au début de cette démonstration. ■

5.1.2 Développement en série de Taylor des moments généralisés centrés

Pour alléger les calculs de cette section nous introduisons les ensembles d'indices ordonnés

$$\mathcal{I}_m = \{i_1, \dots, i_m\} \subset \mathcal{I}_n = \{1, \dots, n\}. \quad (5.9)$$

Théorème 5.2

Soit n un nombre entier positif. Le moment généralisé centré de la mesure MRM log-normale correspondant aux intervalles I_1, \dots, I_n admet le développement en série de Taylor autour de la valeur 0 du coefficient d'intermittence λ^2

$$\mathbb{E} \left[\left(\frac{M(I_1)}{|I_1|} - 1 \right) \dots \left(\frac{M(I_n)}{|I_n|} - 1 \right) \right] = K(I_1, \dots, I_n) \lambda^n + o(\lambda^n), \quad (5.10)$$

où la constante $K(I_1, \dots, I_n)$ est donnée par

$$K(I_1, \dots, I_n) = \begin{cases} 2^n \sum_{P(\mathcal{I}_n)} \left[\prod_{k=1}^{n/2} \int_{I_{i_k}} \frac{du_{i_k}}{|I_{i_k}|} \int_{I_{j_k}} \frac{du_{j_k}}{|I_{j_k}|} \ln \left(\frac{T}{|u_{i_k} - u_{j_k}|} \right) \mathbf{1}_{\{|u_{i_k} - u_{j_k}| \leq T\}} \right], & \text{si } n \text{ pair,} \\ 0, & \text{sinon,} \end{cases} \quad (5.11)$$

où $P(\mathcal{I}_n)$ est l'ensemble de toutes les partitions à deux éléments (non ordonnés) de l'ensemble \mathcal{I}_n . Un élément de $P(\mathcal{I}_n)$ s'écrit donc de la forme $\{(i_k, j_k)\}_{k=1, \dots, n/2}$.

Preuve - 5.2 -

Pour alléger les calculs de cette démonstration nous introduisons les variables aléatoires

$$M_j = \frac{M(I_j)}{|I_j|}, \text{ pour } j = 1, \dots, n. \quad (5.12)$$

Nous considérons le moment centré généralisé correspondant aux intervalles I_1, \dots, I_n . Il est égal à une combinaison linéaire des moments généralisés

$$\mathbb{E}[(M_1 - 1) \dots (M_n - 1)] = \sum_{m=0}^n (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n} \mathbb{E}[M_{i_1} \dots M_{i_m}]. \quad (5.13)$$

En utilisant la représentation intégrale, donnée par la proposition 5.1, des moments généralisés, nous avons

$$\mathbb{E}[(M_1 - 1) \dots (M_n - 1)] = \sum_{m=0}^n (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n} \int_{I_{i_1}} \frac{du_{i_1}}{|I_{i_1}|} \dots \int_{I_{i_m}} \frac{du_{i_m}}{|I_{i_m}|} e^{\lambda^2 S(\mathcal{I}_m)}, \quad (5.14)$$

où $\mathcal{S}(\mathcal{I}_m)$ est la somme symétrique définie par

$$\mathcal{S}(\mathcal{I}_m) = \sum_{\substack{i_j, i_k \in \mathcal{I}_m \\ i_j < i_k}} X_{i_j i_k}, \quad (5.15)$$

avec

$$X_{i_j i_k} = 4 \ln \left(\frac{T}{|u_{i_j} - u_{i_k}|} \right) \mathbf{1}_{\{|u_{i_j} - u_{i_k}| \leq T\}}. \quad (5.16)$$

Il est simple de voir qu'il est possible d'effectuer l'intégration par rapport à toutes les variables u_1, \dots, u_n sous les signes intégrales dans l'expression (5.14)

$$\int_{I_{i_1}} \frac{du_{i_1}}{|I_{i_1}|} \cdots \int_{I_{i_m}} \frac{du_{i_m}}{|I_{i_m}|} e^{\lambda^2 \mathcal{S}(\mathcal{I}_m)} = \int_{I_1} \frac{du_1}{|I_1|} \cdots \int_{I_n} \frac{du_n}{|I_n|} e^{\lambda^2 \mathcal{S}(\mathcal{I}_m)}. \quad (5.17)$$

En utilisant la linéarité, nous pouvons changer l'ordre d'intégration et de sommation, et donc l'expression (5.14) devient

$$\mathbb{E}[(M_1 - 1) \cdots (M_n - 1)] = \int_{I_1} \frac{du_1}{|I_1|} \cdots \int_{I_n} \frac{du_n}{|I_n|} \sum_{m=0}^n (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n} e^{\lambda^2 \mathcal{S}(\mathcal{I}_m)}. \quad (5.18)$$

Nous supposons que les moments généralisés de la mesure MRM log-normale appartiennent à la classe des fonctions $n/2$ fois continuellement dérivables par rapport à λ^2 . Ce résultat découle du lemme 5.3 qui suit cette preuve¹. La k^i ème dérivée du moment généralisé centré par rapport à λ^2 en 0 est alors donnée par

$$\begin{aligned} & \left(\frac{\partial}{\partial \lambda^2} \right)^k \mathbb{E}[(M_1 - 1) \cdots (M_n - 1)] \Big|_{\lambda^2=0} \\ &= \int_{I_1} \frac{du_1}{|I_1|} \cdots \int_{I_n} \frac{du_n}{|I_n|} \sum_{m=0}^n (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n} (\mathcal{S}(\mathcal{I}_m))^k \end{aligned} \quad (5.19)$$

Fixons l'indice j arbitrairement, il est alors possible de regrouper les termes de la somme sous le signe intégrale dans l'expression (5.19)

$$\sum_{m=0}^n (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n} (\mathcal{S}(\mathcal{I}_m))^k = \sum_{m=0}^{n-1} (-1)^{n-m} \sum_{\mathcal{I}_m \subset \mathcal{I}_n \setminus \{j\}} (\mathcal{S}(\mathcal{I}_m))^k - (\mathcal{S}(\mathcal{I}_m \cup \{j\}))^k \quad (5.20)$$

En remarquant que la somme symétrique $\mathcal{S}(\mathcal{I}_m \cup \{j\})$ s'écrit

$$\mathcal{S}(\mathcal{I}_m \cup \{j\}) = \mathcal{S}(\mathcal{I}_m) + \mathcal{C}(j, \mathcal{I}_m), \quad (5.21)$$

¹Les conditions du lemme sont vérifiées puisque nous considérons le régime où λ^2 tend vers zéro.

où $\mathcal{C}(j, \mathcal{I}_m)$ est défini par

$$\mathcal{C}(j, \mathcal{I}_m) = \sum_{i_k \in \mathcal{I}_m} X_{j i_k}, \quad (5.22)$$

nous pouvons voir que

$$(\mathcal{S}(\mathcal{I}_m))^k - (\mathcal{S}(\mathcal{I}_m \cup \{j\}))^k = -\mathcal{C}(j, \mathcal{I}_m) \sum_{i=1}^k \frac{k!}{i!(k-i)!} (\mathcal{C}(j, \mathcal{I}_m))^{i-1} (\mathcal{S}(\mathcal{I}_m))^{k-i}. \quad (5.23)$$

Cette dernière relation signifie que chaque terme de la somme (5.20) contient au moins un facteur du type $X_{j i_m}$. Puisque l'indice j a été fixe arbitrairement, les ensembles d'indices des facteurs de chaque terme de la somme (5.20) doivent contenir tous les indices $1, \dots, n$. En donc, si l'ordre de la dérivée k est inférieur à $n/2$ la somme (5.20) est forcément égale à zéro.

En conséquence, le premier terme non trivial du développement en série de Taylor en λ^2 du moment généralisé centré est au moins de l'ordre $\frac{n+1}{2}$ si l'ordre du moment n est impair, et de l'ordre $\frac{n}{2}$ si l'ordre du moment n est pair. Dans ce dernier cas, les termes d'ordre minimal en λ^2 du développement en série de Taylor en λ^2 sont proportionnels à $\lambda^n X_{i_1 j_1} \cdots X_{i_{n/2} j_{n/2}}$, où l'ensemble d'indices $\{i_1, j_1, \dots, i_{n/2}, j_{n/2}\}$ contient toutes les valeurs $1, \dots, n$. Les termes de ce type peuvent être générés seulement par le développement de $e^{\lambda^2 \mathcal{S}(\mathcal{I}_n)}$, ce qui nous amène au résultat annoncé. ■

Lemme 5.3

Soit n un nombre entier positif. Supposons que $\lambda^2 < \frac{1}{2n} - 2\delta$, où δ est un nombre réel tel que $\frac{1}{2n} - 2\delta > 0$. Le moment généralisé de la mesure MRM log-normale correspondant aux intervalles I_1, \dots, I_n appartient donc à la classe des fonctions n fois continuellement dérivables par rapport à λ^2 . En d'autres termes la dérivation par rapport à λ^2 sous le signe intégrale dans (5.4) est licite

$$\begin{aligned} \left(\frac{\partial}{\partial \lambda^2} \right)^k \mathbb{E}[M(I_1) \cdots M(I_n)] &= \left(\frac{\partial}{\partial \lambda^2} \right)^k \int_{I_1} du_1 \cdots \int_{I_n} du_n f_{\lambda^2}(u_1, \dots, u_n) \\ &= \int_{I_1} du_1 \cdots \int_{I_n} du_n \left(\frac{\partial}{\partial \lambda^2} \right)^k f_{\lambda^2}(u_1, \dots, u_n), \end{aligned} \quad (5.24)$$

où la fonction $f_{\lambda^2}(u_1, \dots, u_n)$ est définie par l'équation (5.5).

Preuve - 5.3 -

La condition $\lambda^2 < \frac{1}{2n} - 2\delta$ permet d'assurer l'existence du moment généralisé (5.4), puisqu'il est majoré par le moment d'ordre n qui est fini sous cette condition, voir le théorème 3.8

$$\mathbb{E}[M(I_1) \cdots M(I_n)] \leq \mathbb{E}\left[M\left(\bigcup_{j=1}^n I_j\right)^n\right]. \quad (5.25)$$

La fonction $f_{\lambda^2}(u_1, \dots, u_n)$, définie par l'équation (5.5), et qui apparaît sous le signe d'intégrale (5.4) est positive et sa $k^{\text{ième}}$ dérivée par rapport à λ^2 s'écrit

$$\left(\frac{\partial}{\partial \lambda^2}\right)^k f_{\lambda^2}(u_1, \dots, u_n) = \left(4 \sum_{1 \leq i < j \leq n} \ln \left(\frac{T}{|u_i - u_j|}\right) \mathbf{1}_{\{|u_i - u_j| \leq T\}}\right)^k f_{\lambda^2}(u_1, \dots, u_n). \quad (5.26)$$

Il est simple de voir qu'il est possible de la majorer uniformément en λ^2 . En effet, en utilisant l'inégalité (D.20) de l'appendice D on déduit la borne supérieure

$$\left(4 \sum_{1 \leq i < j \leq n} \ln \left(\frac{T}{|u_i - u_j|}\right) \mathbf{1}_{\{|u_i - u_j| \leq T\}}\right)^k \leq \left(\frac{p}{e\epsilon}\right)^k f_{\epsilon}(u_1, \dots, u_n). \quad (5.27)$$

En conséquence, en prenant ϵ égal à δ on trouve la majoration uniforme en λ^2 de la $k^{\text{ième}}$ dérivée de la fonction $f_{\lambda^2}(u_1, \dots, u_n)$ par rapport à λ^2

$$\left(\frac{\partial}{\partial \lambda^2}\right)^k f_{\lambda^2}(u_1, \dots, u_n) \leq \left(\frac{k}{e\delta}\right)^k f_{\lambda^2+\delta}(u_1, \dots, u_n) \leq \left(\frac{k}{e\delta}\right)^k f_{\frac{1}{2n}-\delta}(u_1, \dots, u_n), \quad (5.28)$$

où nous avons utilisé le fait que la fonction $f_{\lambda^2}(u_1, \dots, u_n)$ est croissante en λ^2 .

La partie droite de l'inégalité (5.28) est intégrable, puisque l'intégrale est proportionnelle au moment d'ordre n , qui est fini, de la mesure MRM log-normale dont le coefficient d'intermittence est égale à $\frac{1}{2n} - \delta$. La majoration uniforme (5.28) implique donc l'intégrabilité de la $k^{\text{ième}}$ dérivée de la fonction $f_{\lambda^2}(u_1, \dots, u_n)$.

Finalement, la dérivation sous le signe intégrale (5.24) est licite, d'après le théorème de convergence dominée. \blacksquare

5.2 Moments généralisés du logarithme de la mesure MRM log-normale

5.2.1 Existence des moments

Proposition 5.4

Soit n un nombre entier positif. Il existe λ_0^2 tel que le moment généralisé du logarithme de la mesure MRM log-normale correspondant aux intervalles I_1, \dots, I_n est borné uniformément en le coefficient d'intermittence λ^2 , pour λ^2 inférieur à λ_0^2

$$\exists C > 0, \quad \mathbb{E} \left[\ln \left(\frac{M(I_1)}{|I_1|} \right) \cdots \ln \left(\frac{M(I_n)}{|I_n|} \right) \right] < C, \quad \text{pour tout } \lambda^2 < \lambda_0^2. \quad (5.29)$$

Preuve - 5.4 -

L'inégalité de Jensen nous conduit à la majoration suivante

$$\mathbb{E} \left[\ln \left(\frac{M(I_1)}{|I_1|} \right) \cdots \ln \left(\frac{M(I_n)}{|I_n|} \right) \right] \leq \frac{1}{n} \sum_{j=1}^n \mathbb{E} \left[\left| \ln \left(\frac{M(I_j)}{|I_j|} \right) \right|^n \right]. \quad (5.30)$$

La fonction $\psi(q)$, définie par (3.6), dans le cadre du modèle MRW log-normal, est égale à $\psi(q) = -2\lambda^2 q + 2\lambda^2 q^2$. Pour q_0 et λ_0^2 réels positifs, $\psi(-q_0)$ est finie pour tout $\lambda^2 < \lambda_0^2$. En utilisant la proposition 3.10 et le théorème de Doob [Doo01], on peut démontrer que l'espérance $\mathbb{E}[Y^{-q_0}]$ est bornée uniformément pour tout $\lambda^2 < \lambda_0^2$.

D'après la définition (3.49) de la mesure $N_T(dt)$, elle tend vers la mesure de Lebesgue, puisque son espérance est égale à la mesure de Lebesgue et sa variance tend vers zéro uniformément en le coefficient d'intermittence λ^2 .

D'après le théorème 3.11, dont les conditions sont vérifiées, il existe $h < 0$ et λ_0^2 tel que les moments d'ordre h de la mesure MRM log-normale sont finis et bornés uniformément en λ^2 , pour $\lambda^2 < \lambda_0^2$. En utilisant l'inégalité (D.22), pour tout x et h positifs et tout n entier

$$\left| \ln \left(\frac{M(I_j)}{|I_j|} \right) \right|^n \leq \left(\frac{n}{e} \right)^n \left[h^{-n} \left(\frac{M(I_j)}{|I_j|} \right)^{-h} + \frac{M(I_j)}{|I_j|} \right], \quad (5.31)$$

il est simple de déduire la majoration (5.29) uniforme en λ^2 du moment généralisé du logarithme de la mesure MRM log-normale. ■

5.2.2 Développement en série de Taylor

Théorème 5.5

Soit n un nombre entier positif. Le moment généralisé centré du logarithme de la mesure MRM correspondant aux intervalles I_1, \dots, I_n admet le développement en série de Taylor autour de la valeur 0 du coefficient d'intermittence λ^2

$$\mathbb{E} \left[\ln \left(\frac{M(I_1)}{|I_1|} \right) \cdots \ln \left(\frac{M(I_n)}{|I_n|} \right) \right] = K(I_1, \dots, I_n) \lambda^n + o(\lambda^n), \quad (5.32)$$

où l'expression de $K(I_1, \dots, I_n)$ est donné dans l'énoncé du théorème 5.2 (équation (5.11)). Rappelons que $K(I_1, \dots, I_n) = 0$ si n est impair.

Preuve - 5.5 -

Pour alléger les calculs nous introduisons les variables aléatoires

$$M_j = \frac{M(I_j)}{|I_j|}, \text{ pour } j = 1, \dots, n. \quad (5.33)$$

La difficulté principale de la démonstration vient du fait que le développement du logarithme $\ln(1+x)$ en série de Taylor en x autour de zéro converge seulement pour les valeurs de x comprises dans l'intervalle $(-1, 1)$.

Soit ϵ un réel positif inférieur à 1. Nous considérons le vecteur $M = (M_1, \dots, M_n)$ et l'ensemble compact B_ϵ dans \mathbb{R}^n définie par

$$B_\epsilon = \left\{ x = (x_1, \dots, x_n) \in \mathbb{R}^n; \max_{1 \leq k \leq n} (|x_k - 1|) \leq \epsilon \right\}. \quad (5.34)$$

Nous définissons les deux restes $R^{\ln}(\lambda)$ et $R^c(\lambda)$ par les deux relations suivantes

$$R^{\ln}(\lambda) = \mathbb{E}[\ln(M_1) \cdots \ln(M_n)] - \mathbb{E}[\ln(M_1) \cdots \ln(M_n) | M \in B_\epsilon], \quad (5.35)$$

$$R^c(\lambda) = \mathbb{E}[(M_1 - 1) \cdots (M_n - 1)] - \mathbb{E}[(M_1 - 1) \cdots (M_n - 1) | M \in B_\epsilon]. \quad (5.36)$$

En appliquant le lemme 5.7 aux restes $R^{\ln}(\lambda)$ et $R^c(\lambda)$, nous pouvons déduire que

$$|R^{\ln}(\lambda)| \leq \sqrt{\mathbb{E}[\ln(M_1)^2 \cdots \ln(M_n)^2]} \sqrt{\mathbb{P}[M \notin B_\epsilon]} + |\ln(1 - \epsilon)|^n \mathbb{P}[M \notin B_\epsilon], \quad (5.37)$$

$$|R^c(\lambda)| \leq \sqrt{\mathbb{E}[(M_1 - 1)^2 \cdots (M_n - 1)^2]} \sqrt{\mathbb{P}[M \notin B_\epsilon]} + \epsilon^n \mathbb{P}[M \notin B_\epsilon]. \quad (5.38)$$

Puisque nous nous intéressons au régime lorsque le coefficient d'intermittence λ^2 tend vers zéro, d'après les théorèmes 3.8 et 5.5, il existe λ_0^2 positif tel que les deux espérances dans les inégalités (5.37) et (5.38) sont bornés uniformément en λ^2 , pour λ^2 inférieur à λ_0^2 .

En appliquant le lemme 5.6 avec m égal à n , nous pouvons majorer les deux restes

$$|R^{\ln}(\lambda)| \leq C\lambda^{n+1}, \quad (5.39)$$

$$|R^c(\lambda)| \leq C\lambda^{n+1}, \quad (5.40)$$

où la constante C dépend de ϵ .

D'après le lemme 5.8, les définitions (5.35) et (5.36) et les majorations (5.39) et (5.40), nous avons

$$\mathbb{E}[\ln(M_1) \cdots \ln(M_n)] - \mathbb{E}[(M_1 - 1) \cdots (M_n - 1)] = o(\lambda^n). \quad (5.41)$$

Le résultat du théorème découle alors directement du théorème 5.2. \blacksquare

Lemme 5.6

Soient ϵ un réel positif inférieur à 1 et m un nombre entier positif, alors nous avons la majoration suivante

$$\mathbb{P}[M \notin B_\epsilon] \leq o(\lambda^{2m}), \quad (5.42)$$

où l'ensemble compact B_ϵ est défini par (5.34) et $o(\lambda^{2m})$ dépend de ϵ .

Preuve - 5.6 -

Puisque nous nous intéressons au régime où λ^2 tend vers zéro, alors sans perte de généralité on peut supposer que le moment centré d'ordre $2m + 2$ de la mesure MRM log-normale existe, grâce au théorème 3.8. En utilisant l'inégalité de Bienaymé-Tchebychev, nous avons

$$\mathbb{P}[M \notin B_\epsilon] \leq \sum_{j=1}^n \mathbb{P}[|M_j - 1| > \epsilon] \leq \sum_{j=1}^n \frac{1}{\epsilon^{2m+2}} \mathbb{E}[(M_j - 1)^{2m+2}] = o(\lambda^{2m}). \quad (5.43)$$

■

Lemme 5.7

Soit ϵ un réel positif inférieur à 1. Pour toute fonction $f(M)$ continue sur l'ensemble compact B_ϵ , défini par (5.34), nous avons la majoration suivante

$$\begin{aligned} \mathbb{E}[f(M)] - \mathbb{E}[f(M)|M \in B_\epsilon] \\ \leq \sqrt{\mathbb{E}[f(M)^2]} \sqrt{\mathbb{P}[M \notin B_\epsilon]} + \left(\sup_{M \in B_\epsilon} |f(M)| \right) \mathbb{P}[M \notin B_\epsilon]. \end{aligned} \quad (5.44)$$

Preuve - 5.7 -

En appliquant la loi des probabilités totales, nous avons

$$\begin{aligned} \mathbb{E}[f(M)] - \mathbb{E}[f(M)|M \in B_\epsilon] \\ = \mathbb{E}[f(M)|M \notin B_\epsilon] \mathbb{P}[M \notin B_\epsilon] + \mathbb{E}[f(M)|M \in B_\epsilon] (\mathbb{P}[M \in B_\epsilon] - 1) \\ = \mathbb{E}[f(M)|M \notin B_\epsilon] \mathbb{P}[M \notin B_\epsilon] - \mathbb{E}[f(M)|M \in B_\epsilon] \mathbb{P}[M \notin B_\epsilon], \end{aligned} \quad (5.45)$$

où le premier terme est majoré à l'aide de l'inégalité de Cauchy-Schwartz

$$\mathbb{E}[f(M)|M \notin B_\epsilon] \mathbb{P}[M \notin B_\epsilon] = \mathbb{E}[f(M) \mathbf{1}_{\{M \notin B_\epsilon\}}] \leq \sqrt{\mathbb{E}[f(M)^2]} \sqrt{\mathbb{P}[M \notin B_\epsilon]}, \quad (5.46)$$

tandis que le deuxième est majoré par le supremum sur l'ensemble compact B_ϵ

$$\mathbb{E}[f(M)|M \in B_\epsilon] \mathbb{P}[M \notin B_\epsilon] \leq \left(\sup_{M \in B_\epsilon} |f(M)| \right) \mathbb{P}[M \notin B_\epsilon]. \quad (5.47)$$

■

Lemme 5.8

Soit n un nombre entier positif. Sur l'ensemble compact B_ϵ , avec ϵ inférieur à 1, les deux développements en série de Taylor autour de la valeur 0 du coefficient d'intermittence λ^2

$$\mathbb{E} \left[\prod_{j=1}^n \ln(M_j) \middle| M \in B_\epsilon \right] = \mathbb{E} \left[\prod_{j=1}^n (M_j - 1) \middle| M \in B_\epsilon \right] + o(\lambda^n) \quad (5.48)$$

coïncident jusqu'à l'ordre n .

Preuve - 5.8 -

Grâce à l'identité $\ln(x) = x - 1 - (x - 1 - \ln(x))$, nous avons

$$\prod_{j=1}^n \ln(M_j) - \prod_{j=1}^n (M_j - 1) = \prod_{j=1}^n (M_j - 1 - (M_j - 1 - \ln(M_j))) - \prod_{j=1}^n (M_j - 1). \quad (5.49)$$

Le développement du premier produit de la partie droite nous amène à une combinaison linéaire de termes du type $\prod_{j=1}^k (M_{i_j} - 1 - \ln(M_{i_j})) \prod_{j=k+1}^n (M_{i_j} - 1)$ avec $1 \leq k \leq n$ dont l'espérance conditionnelle est majorée à l'aide de l'inégalité de Cauchy-Schwartz et de l'inégalité (D.28)

$$\begin{aligned} & \mathbb{E} \left[\prod_{j=1}^k (M_{i_j} - 1 - \ln(M_{i_j})) \prod_{j=k+1}^n (M_{i_j} - 1) \middle| M \in B_\epsilon \right] \\ & \leq \mathbb{E} \left[\prod_{j=1}^k (M_{i_j} - 1 - \ln(M_{i_j}))^2 \middle| M \in B_\epsilon \right]^{1/2} \mathbb{E} \left[\prod_{j=k+1}^n (M_{i_j} - 1)^2 \middle| M \in B_\epsilon \right]^{1/2} \\ & \leq C_\epsilon \mathbb{E} \left[\prod_{j=1}^k (M_{i_j} - 1)^4 \middle| M \in B_\epsilon \right]^{1/2} \mathbb{E} \left[\prod_{j=k+1}^n (M_{i_j} - 1)^2 \middle| M \in B_\epsilon \right]^{1/2}. \quad (5.50) \end{aligned}$$

En appliquant la majoration (5.40), avec n égal à 2 et à 4, il est possible d'enlever les conditionnements $M \in B_\epsilon$

$$\begin{aligned} & \mathbb{E} \left[\prod_{j=1}^k (M_{i_j} - 1 - \ln(M_{i_j})) \prod_{j=k+1}^n (M_{i_j} - 1) \middle| M \in B_\epsilon \right] \\ & = (O(\lambda^{4k}) + o(\lambda^{4n}))^{1/2} (O(\lambda^{2n-2k}) + o(\lambda^{2n}))^{1/2} = O(\lambda^{n+k}) = o(\lambda^n). \quad (5.51) \end{aligned}$$

■

5.3 Magnitude renormalisée

Nous avons établis deux théorèmes dans la limite de petite intermittence : le théorème 5.2 qui donne une approximation des moments de la mesure MRM log-normale et le théorème 5.5 qui donne une approximation des moments du logarithme de cette mesure. Ces deux résultats vont permettre d'établir, dans les sections 5.4 et 5.5, les approximations log-normales de la mesure MRM log-normale et du processus MRW log-normal.

Le processus gaussien à accroissements stationnaires qui intervient dans ces approximations log-normales est le processus de magnitude renormalisée que nous introduisons dans cette section.

5.3.1 Définition

Définition 5.2

Nous définissons le processus de **magnitude renormalisée** $\{\Omega(t)\}_t$ comme la limite, si elle existe,

$$\Omega(t) = \lim_{l \rightarrow 0^+} \Omega_l(t) \quad (5.52)$$

où le processus $\Omega_l(t)$ est donné par

$$\Omega_l(t) = \frac{1}{\lambda} \int_0^t (\omega_l(s) - \mathbb{E}[\omega_l]) ds, \quad (5.53)$$

et où le processus gaussien stationnaire $\{\omega_l(t)\}_t$ est défini dans la section 3.4.1.

Dans la suite, si $I = [t_0, t_1]$ est un intervalle, nous utiliserons la notation $\Omega(I)$ pour désigner l'accroissement de la magnitude renormalisée sur l'intervalle I

$$\Omega(I) = \Omega(t_1) - \Omega(t_0), \quad (5.54)$$

et la notation $\delta_\tau \Omega(t)$ pour désigner l'accroissement sur l'intervalle $[t - \tau, t]$

$$\delta_\tau \Omega(t) = \Omega(t) - \Omega(t - \tau). \quad (5.55)$$

5.3.2 Existence de la magnitude renormalisée

Les deux propositions suivantes démontrent respectivement la convergence des lois finies dimensionnelles du processus $\{\Omega_l(t)\}_t$ et la tension des processus $\{e^{2\Omega_l(t)}\}_t$. L'existence du processus limite $\{\Omega(t)\}_t$ en découle.

Proposition 5.9

Soient t_1, \dots, t_n , n réels, alors le vecteur gaussien $(\Omega_l(t_1), \dots, \Omega_l(t_n))$ converge, lorsque l tend vers zéro, vers un vecteur gaussien centré $(\Omega(t_1), \dots, \Omega(t_n))$, qui ne dépend pas de λ et dont la matrice de variance-covariance Σ est donnée par

$$(\Sigma)_{jk} = \text{Cov}[\Omega(t_j), \Omega(t_k)] = \frac{1}{\lambda^2} \int_0^{t_j} du \int_0^{t_k} dv \rho(u - v), \quad (5.56)$$

où la fonction $\rho(t)$ est définie par

$$\rho(t) = \begin{cases} \lambda^2 \ln \left(\frac{T}{|t|} \right), & \text{si } |t| < T, \\ 0, & \text{si } T \leq |t|. \end{cases} \quad (5.57)$$

Preuve - 5.9 -

La fonction $\rho_l(t)$ étant définie (cf. section 3.4.1) comme la fonction d'autocorrélation du processus $\omega_{l(t)}$, le vecteur $(\Omega_l(t_1), \dots, \Omega_l(t_n))$ est un vecteur gaussien centré dont la matrice de variance-covariance Σ_l est donnée par

$$(\Sigma_l)_{jk} = \text{Cov}[\Omega_l(t_j), \Omega_l(t_k)] = \frac{1}{\lambda^2} \int_0^{t_j} du \int_0^{t_k} dv \rho_l(|u - v|), \quad (5.58)$$

où la fonction $\rho_l(t)$ est définie par la relation (3.122).

La matrice Σ_l converge vers la matrice Σ (dont tous les coefficients sont finis) lorsque l tend vers zéro. Il suffit donc de démontrer que la matrice Σ est semi-définie positive. Dans l'appendice F, nous montrons que la fonction $\rho(t)$ peut s'écrire comme (F.7)

$$\rho(t) = \int_{-\infty}^{\infty} \tilde{\rho}(t-s)\tilde{\rho}(s)ds. \quad (5.59)$$

Si l'on définit le vecteur $V(s)$ par

$$V(s) = \left(\int_0^{t_1} \tilde{\rho}(u_1 - s)du_1, \dots, \int_0^{t_n} \tilde{\rho}(u_n - s)du_n \right)^T \quad (5.60)$$

alors la matrice Σ s'écrit comme

$$\Sigma = \int_{-\infty}^{\infty} V(s)V^T(s)ds, \quad (5.61)$$

grâce à l'identité $\tilde{\rho}(s) = \tilde{\rho}(-s)$. En conséquence, la matrice Σ est semi-définie positive. ■

Etant donné que les variables $\Omega_l(t)$ et $\Omega(t)$ sont gaussiennes, pour montrer la tension de la suite $\{e^{2\Omega_l(t)}\}_t$, il suffit de montrer que

Proposition 5.10

Il existe $\epsilon \in [0, 1[$ tel que

$$\mathbb{E} \left[(e^{2\Omega(t)} - e^{2\Omega(s)})^2 \right] = o((t-s)^{2-\epsilon}), \quad \forall t, s. \quad (5.62)$$

Preuve - 5.10 -

Un calcul simple nous amène à l'équation suivante

$$\begin{aligned} \mathbb{E} \left[(e^{2\Omega(t)} - e^{2\Omega(s)})^2 \right] &= e^{8\text{Var}[\Omega(t)]} + e^{8\text{Var}[\Omega(s)]} - 2e^{2\text{Var}[\Omega(t)+\Omega(s)]} \\ &= \left(e^{4\text{Var}[\Omega(t)]} - e^{4\text{Var}[\Omega(s)]} \right)^2 + 2e^{4\text{Var}[\Omega(t)]+4\text{Var}[\Omega(s)]} \left(1 - e^{-2\text{Var}[\Omega(t-s)]} \right). \end{aligned} \quad (5.63)$$

Le premier terme de (5.63) peut être estimé comme

$$\begin{aligned} & \left(e^{4\text{Var}[\Omega(t)]} - e^{4\text{Var}[\Omega(s)]} \right)^2 \\ &= 64t^2 e^{8t^2 \ln\left(\frac{Te^{3/2}}{t}\right)} (t-s)^2 + o((t-s)^2) = o((t-s)^{2-\epsilon}). \end{aligned} \quad (5.64)$$

Le deuxième terme de (5.63) peut être estimé comme

$$\begin{aligned} & 2e^{4\text{Var}[\Omega(t)]+4\text{Var}[\Omega(t)]} \left(1 - e^{-2\text{Var}[\Omega(|t-s|)]} \right) \\ &= 2e^{8t^2 \ln\left(\frac{Te^{3/2}}{t}\right)} \ln\left(\frac{Te^{3/2}}{|t-s|}\right) (t-s)^2 + o((t-s)^2) = o((t-s)^{2-\epsilon}). \end{aligned} \quad (5.65)$$

■

5.3.3 Autocovariance des accroissements de la magnitude renormalisée

A partir de l'expression (5.56), il est facile de montrer que la magnitude renormalisée est un processus gaussien à accroissements stationnaires. Comme nous allons le montrer dans les sections suivantes, ses accroissements vont permettre d'approximer les accroissements d'une mesure MRM log-normale ou d'un processus MRW log-normal. Afin de mieux les caractériser, donnons une expression explicite de leur fonction d'autocovariance ².

Proposition 5.11

Soient $\tau > 0$ et $h \geq \tau$. Pour tout t , on a

- si $h + \tau \leq T$,

$$\text{Cov}\left[\frac{\delta_\tau \Omega(t)}{\tau}, \frac{\delta_\tau \Omega(t+h)}{\tau}\right] = \ln\left(\frac{Te^{3/2}}{\tau}\right) + f\left(\frac{h}{\tau}\right), \quad (5.66)$$

où la fonction $f(u)$ est donnée par

$$f(u) = \begin{cases} -\ln(u) - \frac{(u+1)^2}{2} \ln\left(1 + \frac{1}{u}\right) - \frac{(u-1)^2}{2} \ln\left(1 - \frac{1}{u}\right), & \text{si } u \geq 2, \\ -2\ln(2), & \text{si } u = 1, \\ 0, & \text{si } u = 0, \end{cases} \quad (5.67)$$

- si $h \geq T + \tau$,

$$\text{Cov}[\delta_\tau \Omega(t), \delta_\tau \Omega(t+h)] = 0. \quad (5.68)$$

Preuve - 5.11 -

Ces résultats sont simples à obtenir à partir des relations (5.56) et (5.57). ■

²Nous n'explicitons pas la fonction d'autocovariance sur tout son domaine, les formules deviennent rapidement assez compliquées à écrire et sans grand intérêt. Nous nous contenterons de l'écrire lorsque les deux accroissements appartiennent à une même échelle intégrale ou lorsqu'ils sont distants de plus d'une échelle intégrale.

5.4 Approximation log-normale de la mesure MRM

5.4.1 Convergence en loi vers la magnitude renormalisée

Théorème 5.12

Soient I_1, \dots, I_n , n intervalles quelconques. Lorsque λ^2 tend vers zéro, on a la convergence en loi suivante

$$\left(\frac{1}{2\lambda} \ln \left(\frac{M(I_1)}{|I_1|} \right), \dots, \frac{1}{2\lambda} \ln \left(\frac{M(I_n)}{|I_n|} \right) \right) \xrightarrow{\mathcal{L}} \left(\frac{\Omega(I_1)}{|I_1|}, \dots, \frac{\Omega(I_n)}{|I_n|} \right). \quad (5.69)$$

Preuve - 5.12 -

Un calcul simple, basé sur le théorème de Wick [Wic50] et sur l'expression (5.56) de la matrice de covariance de $\Omega(t)$, permet de montrer que

$$\mathbb{E} \left[\frac{\Omega(I_1)}{|I_1|}, \dots, \frac{\Omega(I_n)}{|I_n|} \right] = \frac{1}{2^n} K(I_1, \dots, I_n), \quad (5.70)$$

où la constante $K(I_1, \dots, I_n)$ est donné par (5.11). En utilisant le théorème 5.5, on obtient immédiatement que, pour tout n ,

$$\lim_{\lambda \rightarrow 0} \mathbb{E} \left[\prod_{j=1}^n \frac{1}{2\lambda} \ln \left(\frac{M(I_j)}{|I_j|} \right) \right] = \mathbb{E} \left[\frac{\Omega(I_1)}{|I_1|}, \dots, \frac{\Omega(I_n)}{|I_n|} \right]. \quad (5.71)$$

Le lemme G.1, permet de conclure : il suffit juste de remarquer que la condition (G.1) est vérifiée pour une variable gaussienne X quelconque. ■

On peut énoncer un corollaire du théorème 5.12 portant sur le processus formé par les accroissements successifs de la mesure :

Corollaire 5.13

Si τ est un réel strictement positif, alors

$$\left\{ \frac{1}{2\lambda} \ln \left(\frac{\delta_\tau M(t)}{\tau} \right) \right\}_t \xrightarrow{\mathcal{L}} \left\{ \frac{\delta_\tau \Omega(t)}{\tau} \right\}_t. \quad (5.72)$$

A la fin de ce chapitre, dans la section 5.6, nous effectuons la comparaison de la distribution empiriques du logarithme de l'accroissement de la mesure MRM log-normale avec la distribution de l'approximation log-normale.

5.4.2 Approximation des moments généralisés logarithmiques

Théorème 5.14

Soit τ un réel strictement positif. Le processus $\{2\lambda\delta_\tau\Omega(t)/\tau\}_t$ permet de reproduire l'ordre zéro ainsi que le premier ordre non trivial qui suit du développement de Taylor du moment

généralisé du logarithme des accroissements de la mesure MRM log-normale

$$\ln\left(\frac{\delta_\tau M(t)}{\tau}\right) \stackrel{\lambda}{\simeq} 2\lambda \frac{\delta_\tau \Omega(t)}{\tau}. \quad (5.73)$$

Preuve - 5.14 -

Ce résultat est une conséquence directe du théorème 5.5 et du théorème de Wick [Wic50] appliqué à l'espérance du produit de variables gaussiennes $\{\delta_\tau \Omega(t_j)\}_j$ centrées dont la matrice de variance-covariance est donnée dans la proposition 5.9 par l'équation (5.56). ■

Notons qu'un résultat analogue mais plus général, portant non plus sur des accroissements de M disjoints et de même taille mais sur des accroissements non forcément disjoints et de tailles différentes peut être démontré tout aussi simplement.

Une application de ce théorème consiste à calculer une approximation de la fonction d'autocovariance des accroissements de la mesure MRM log-normale

Proposition 5.15

Pour tout $\tau > 0$ et $h \geq 0$, on a, pour tout t

$$\text{Cov}\left[\ln\left(\frac{\delta_\tau M(t)}{\tau}\right), \ln\left(\frac{\delta_\tau M(t+h)}{\tau}\right)\right] = 4\lambda^2 \text{Cov}\left[\frac{\delta_\tau \Omega(t)}{\tau}, \frac{\delta_\tau \Omega(t+h)}{\tau}\right] + o(\lambda^2), \quad (5.74)$$

où la fonction d'autocovariance des accroissements de la magnitude renormalisée est explicitée dans la proposition 5.11. De plus dès que $h+\tau \leq T$ alors le terme $o(\lambda^2)$ ne dépend pas de l'échelle intégrale T et ne dépend de τ qu'à travers le rapport h/τ .

Preuve - 5.15 -

La relation (5.74) est une conséquence directe du théorème précédent. Il reste juste à montrer que si $h + \tau \leq T$ alors $o(\lambda^2)$ ne dépend que de h/τ .

En utilisant la propriété d'invariance par changement du temps intégral, donnée par le théorème (3.16), puis la propriété d'invariance, donnée par la proposition 3.6, nous avons l'égalité en loi

$$\{\delta_\tau M_T(t)\}_{\tau \leq t \leq h} \stackrel{\mathcal{L}}{=} W_{h/T} \{\delta_\tau M_{h/\tau}(t/\tau)\}_{\tau \leq t \leq h}, \quad (5.75)$$

où $W_{h/T}$ est une variable log-normale qui vérifie

$$\mathbb{E}\left[e^{iq \ln(W_{h/T})}\right] = \left(\frac{h}{T}\right)^{-\varphi(q)}. \quad (5.76)$$

En utilisant (3.4) et (3.6), on obtient $\varphi(q) = -2i\lambda^2 q + 2\lambda^2 q^2$. Nous avons donc

$$\text{Var}\left[\ln(W_{h/T})\right] = 4\lambda^2 \ln\left(\frac{T}{h}\right). \quad (5.77)$$

Nous pouvons déduire de l'équation précédente (5.77) et de (5.75) que la différence $\text{Cov}[\ln(|\delta_\tau M_T(t)|), \ln(|\delta_\tau M_T(t+h)|)] - \text{Var}[\ln(W_{h/T})]$ ne dépend que de λ^2 et de h/τ . Or, en utilisant (5.56) et (5.57), on obtient, si $h + \tau \leq T$,

$$4\lambda^2 \text{Cov}\left[\frac{\delta_\tau \Omega(t)}{\tau}, \frac{\delta_\tau \Omega(t+h)}{\tau}\right] = 4\lambda^2 \int_{t-\tau}^t \frac{du}{\tau} \int_{t+h-\tau}^{t+h} \frac{dv}{\tau} \ln\left(\frac{h}{|u-v|}\right) + \text{Var}[\ln(W_{h/T})]. \quad (5.78)$$

En faisant le changement de variable $u' = u/\tau$ et $v' = v/\tau$ on montre que l'intégrale double ci-dessus ne dépend que de h/τ . En insérant cette expression dans l'équation (5.74), on montre donc que le $o(\lambda^2)$ qui apparaît dans cette équation ne dépend que de h/T . ■

Notons qu'un calcul analogue permet de calculer l'espérance

$$\mathbb{E}\left[\ln\left(\frac{\delta_\tau M(t)}{\tau}\right)\right] = -2\lambda^2 \ln\left(\frac{T e^{3/2}}{\tau}\right) + o(\lambda^2), \quad (5.79)$$

où le terme $o(\lambda^2)$ ne dépend ni de T et ni de τ .

5.4.3 Approximation des moments généralisés

La relation (5.73) précédemment démontrée pousserait à écrire que les moments généralisés des accroissements $\delta_\tau M(t)$ de la mesure MRM sont reproduits, à l'ordre zéro et au premier ordre non trivial, par ceux du processus $\tau e^{2\lambda \frac{\delta_\tau \Omega(t)}{\tau}}$. Il est facile de voir que cela est faux : les moyennes (c-à-d. l'ordre zéro) ne sont pas les mêmes. Cela est dû au fait que la moyenne de l'exponentielle d'une variable aléatoire n'est pas l'exponentielle de la moyenne de cette variable. Il est donc nécessaire de modifier légèrement le processus $\frac{\delta_\tau \Omega(t)}{\tau}$ en le décentrant.

Théorème 5.16

Soit τ un réel strictement positif. Le processus $\{\tau e^{2\lambda \delta_\tau \Omega(t)/\tau - 2\lambda^2 \text{Var}[\delta_\tau \Omega(t)/\tau]}\}_t$ permet de reproduire l'ordre zéro ainsi que le premier ordre non trivial qui suit du développement de Taylor du moment généralisé des accroissements de la mesure MRM log-normale

$$\delta_\tau M(t) \stackrel{\lambda}{\simeq} \tau e^{2\lambda \frac{\delta_\tau \Omega(t)}{\tau} - 2\lambda^2 \text{Var}\left[\frac{\delta_\tau \Omega(t)}{\tau}\right]}. \quad (5.80)$$

Preuve - 5.16 -

Le moment à n points du processus intervenant dans le terme de droite de (5.80) peut s'écrire :

$$m(t_1, \dots, t_n) = \tau^n e^{-2n\lambda^2 \text{Var}\left[\frac{\delta_\tau \Omega}{\tau}\right]} \mathbb{E}\left[e^{2\lambda \sum_{i=1}^n \frac{\delta_\tau \Omega(t_i)}{\tau}}\right] = \tau^n e^{2\lambda^2 \text{Var}\left[\sum_{i=1}^n \frac{\delta_\tau \Omega(t_i)}{\tau}\right]}. \quad (5.81)$$

En faisant un développement en série de Taylor de cette expression, et en remplaçant la variance de $\sum_{i=1}^n \delta_\tau \Omega(t_i)$ par son expression (en utilisant l'équation (5.56)), on obtient

$$m(t_1, \dots, t_n) = \tau^n + 4\tau^n \int_{t_n}^{t_n+\tau} du_1 \cdots \int_{t_n}^{t_n+\tau} du_n \sum_{1 \leq j \leq k \leq n} \rho(u_j - u_k) + o(\lambda^2), \quad (5.82)$$

où ρ est défini par (5.57). En rapprochant cette dernière équation des équations (5.4) et (5.5), on montre que

$$m(t_1, \dots, t_n) = \mathbb{E}[\delta_\tau M(t_1), \dots, \delta_\tau M(t_n)] + o(\lambda^2), \quad (5.83)$$

ce qui démontre le résultat énoncé. \blacksquare

Notons qu'un résultat analogue mais plus général, portant non plus sur des accroissements de M disjoints et de même taille mais sur des accroissements non forcément disjoints et de tailles différentes peut être démontré tout aussi simplement.

5.5 Approximation du processus MRW

Le processus MRW log-normal étant obtenu par subordination d'un mouvement brownien par la mesure MRM log-normale, il est clair que les moments généralisés des accroissements du MRW log-normal ou de leur logarithme (après avoir pris la valeur absolue) sont liés à ceux de la mesure MRM log-normale. Dans la section précédente, nous avons obtenu une approximation de ces moments dans le cas de la mesure MRM log-normale, nous pouvons donc déduire de ces approximations des approximations de ces moments dans le cas MRW log-normal.

5.5.1 Approximation des moments généralisés

Théorème 5.17

Soit τ un réel strictement positif. Soit $\{\epsilon[n]\}_n$ un bruit blanc gaussien de variance σ^2 . Le processus à temps discret $\{\tau^{1/2} \epsilon[n] e^{\delta_\tau \Omega(n\tau)/\tau - \lambda^2 \text{Var}[\delta_\tau \Omega/\tau]}\}_n$ permet de reproduire l'ordre zéro ainsi que le premier ordre non trivial qui suit du développement de Taylor du moment généralisé des accroissements du processus MRW log-normal $X(t)$

$$\delta_\tau X(n\tau) \stackrel{\lambda}{\simeq} \tau^{1/2} \epsilon[n] e^{\lambda \frac{\delta_\tau \Omega(n\tau)}{\tau} - \lambda^2 \text{Var}[\frac{\delta_\tau \Omega}{\tau}]}. \quad (5.84)$$

De plus, le premier ordre non trivial qui n'est pas l'ordre 0 est de l'ordre de λ^2 .

Preuve - 5.17 -

Ce résultat est une conséquence immédiate du fait que $X(t) \stackrel{\mathcal{L}}{=} B(M(t))$, où $B(t)$ est un mouvement brownien de variance σ^2 et du théorème 5.16. \blacksquare

5.5.2 Approximation des moments généralisés logarithmiques

Théorème 5.18

Soit τ un réel strictement positif. Soit $\{\epsilon[n]\}_n$ un bruit blanc gaussien de variance σ^2 . Le processus à temps discret $\{\ln(\tau^{1/2}) + \ln(|\epsilon[n]|) + \lambda\delta_\tau\Omega(n\tau)/\tau\}_n$ permet de reproduire l'ordre zéro ainsi que le premier ordre non trivial qui suit du développement de Taylor du moment généralisé du logarithme des accroissements du processus MRW log-normal $X(t)$

$$\ln |\delta_\tau X(n\tau)| \simeq \frac{\lambda}{2} \ln(\tau) + \ln(|\epsilon[n]|) + \lambda \frac{\delta_\tau \Omega(n\tau)}{\tau}. \quad (5.85)$$

De plus, le premier ordre non trivial et qui n'est pas l'ordre 0 est de l'ordre de λ^2 .

Preuve - 5.18 -

Ce résultat est une conséquence immédiate du fait que $X(t) \stackrel{\mathcal{L}}{=} B(M(t))$, où $B(t)$ est un mouvement brownien de variance σ^2 et du théorème 5.14. ■

On peut énoncer la propriété suivante sur la covariance

Proposition 5.19

Pour tout $\tau > 0$ et $h \geq 0$, on a, pour tout t

$$\begin{aligned} \text{Cov} [\ln(|\delta_\tau X(t)|), \ln(|\delta_\tau X(t+h)|)] \\ = \frac{\pi^2}{8} \delta(m) + \lambda^2 \text{Cov} \left[\frac{\delta_\tau \Omega(t)}{\tau}, \frac{\delta_\tau \Omega(t+h)}{\tau} \right] + o(\lambda^2), \end{aligned} \quad (5.86)$$

où la fonction d'autocovariance des accroissements de la magnitude renormalisée est explicitée dans la proposition 5.11. De plus dès que $h + \tau \leq T$ alors le terme $o(\lambda^2)$ ne dépend pas de l'échelle intégrale T et ne dépend de τ qu'à travers le rapport h/τ .

De plus, on a

$$\mathbb{E} [\ln(|\delta_\tau X(n\tau)|)] = -\frac{\gamma + \ln(2)}{2} - \lambda^2 \ln \left(\frac{T e^{3/2}}{\tau} \right) + o(\lambda^2), \quad (5.87)$$

où γ est une constante d'Euler et le terme $o(\lambda^2)$ ne dépend ni de T et ni de τ .

Preuve - 5.19 -

La relation (5.86) est une conséquence directe du théorème précédent. Pour démontrer que si $h + \tau \leq T$ alors $o(\lambda^2)$ ne dépend que de h/τ , on peut utiliser le même raisonnement que dans la preuve de la proposition 5.15. L'expression de l'espérance (5.87) est une conséquence directe de l'équation (5.79). ■

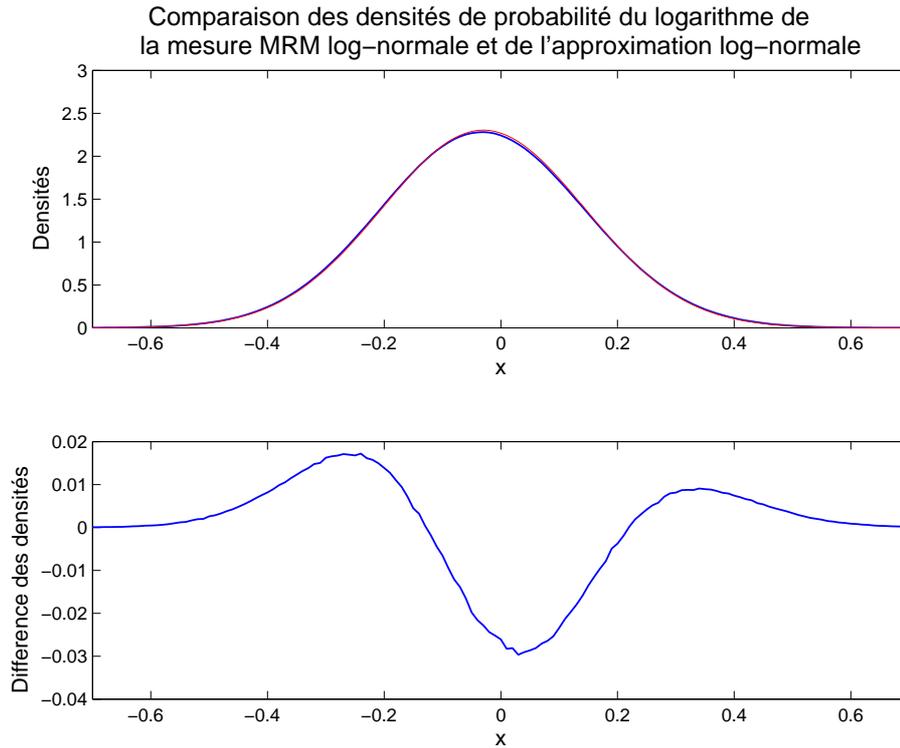


FIG. 5.1 – En haut, les densités de probabilité du logarithme de la mesure MRM log–normale. L’une est obtenue par la méthode Monte Carlo de $4 \cdot 10^9$ tirages de paramètres $T = 1$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$, et l’autre est donnée par l’approximation normale. En bas, l’écart entre les densités de probabilité.

Notons que si m est tel que $\tau \ll (m+1)\tau \leq T$ alors le terme dominant de l’équation (5.86) est

$$\text{Cov}[\ln(|\delta_\tau X(n\tau)|), \ln(|\delta_\tau X((n+m)\tau)|)] \simeq \lambda^2 \ln\left(\frac{T}{m\tau}\right), \quad (5.88)$$

ce qui permet d’expliquer le comportement (cf. figure 3.11) de la fonction d’autocovariance empirique des magnitudes, étudiée dans la section 3.4.4 sur les faits stylisés dans le cadre du modèle MRW log-normal.

5.6 Simulations numériques

Le théorème 5.12 démontre la convergence de l’approximation log–normale de la mesure MRM log–normale lorsque le coefficient d’intermittence λ^2 tend vers zéro. Dans cette section nous effectuons des comparaisons des distributions empiriques, obtenues par la méthode de Monte Carlo, de l’accroissement de la mesure MRM log–normale avec les distributions théoriques.

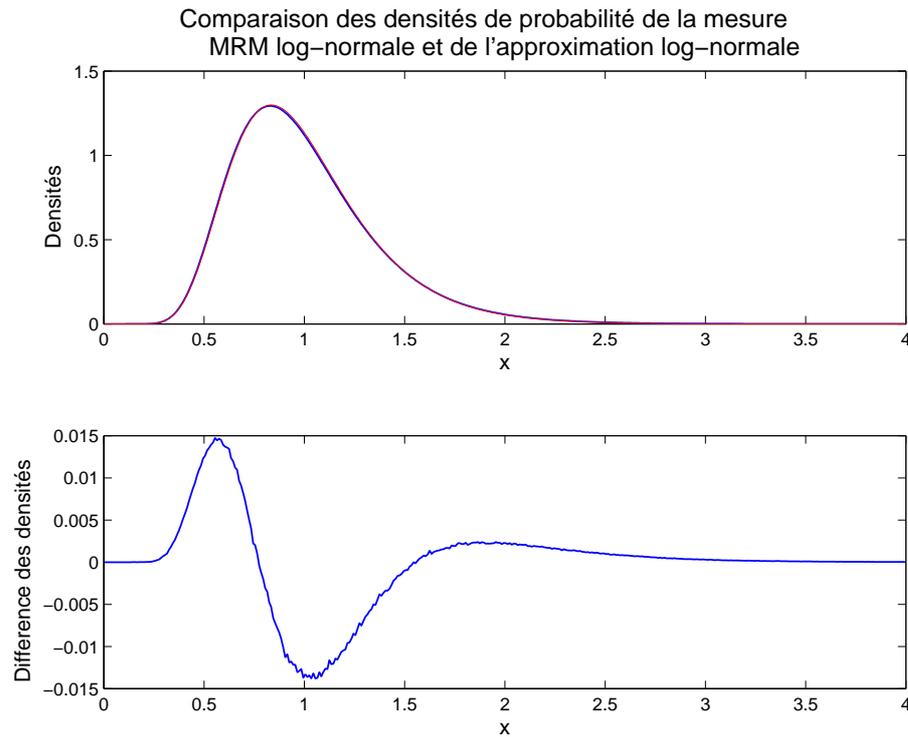


FIG. 5.2 – En haut, les densités de probabilité de la mesure MRM log-normale. L'une est obtenue par la méthode Monte Carlo de $4 \cdot 10^9$ tirages de paramètres $T = 1$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$, et l'autre est donnée par l'approximation log-normale. En bas, l'écart entre les densités de probabilité.

Les figures 5.1 et 5.2 représentent les résultats des simulations de Monte Carlo de la mesure MRM log-normale $M_T(1)$ et du logarithme $\ln(M_T(1))$. Nous avons effectué $4 \cdot 10^9$ tirages de l'accroissement à l'échelle 1 de la mesure MRM log-normale de temps intégral $T = 1$, de coefficient d'intermittence $\lambda^2 = 0.02$, de variance $\sigma^2 = 1$ et d'échelle de sur-échantillonnage $l = 2^{-4}$. Les figures 5.3 et 5.4 représentent les mêmes distributions lorsque le temps intégral $T = 250$.

Il est simple de voir que les approximations, log-normale et normale, permettent de reproduire fidèlement les vraies distributions, les écarts entre les densités de probabilité sont de l'ordre 10^{-2} , dans le cas où le temps intégral $T = 1$, et 10^{-3} dans le cas $T = 250$.

Notons que la distribution du logarithme de la mesure MRM log-normale du temps intégral $T = 250$, représentée dans la figure 5.3 théoriquement peut être obtenue à partir de celle du temps intégral $T = 1$, représentée dans la figure 5.1 par une convolution avec un noyau gaussien, d'espérance $-\lambda^2 \ln(T)$ et de variance $\lambda^2 \ln(T)$, grâce à la propriété d'invariance par changement du temps intégral, donnée par le théorème 3.16.

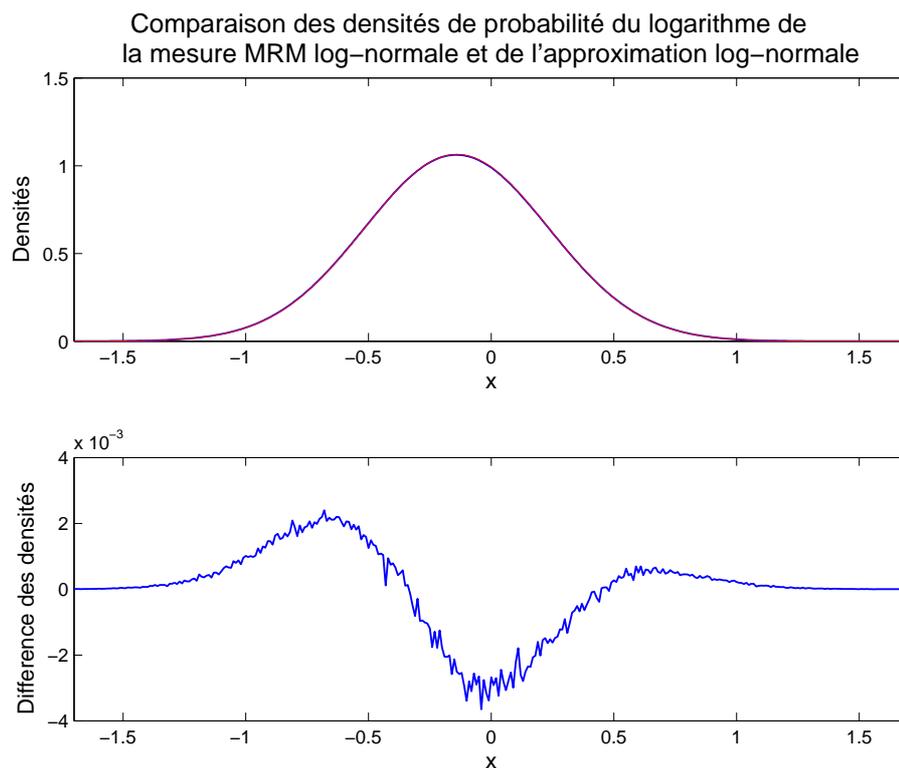


FIG. 5.3 – En haut, les densités de probabilité du logarithme de la mesure MRM log-normale. L'une est obtenue par la méthode Monte Carlo de 10^9 tirages de paramètres $T = 250$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$, et l'autre est donnée par l'approximation normale. En bas, l'écart entre les densités de probabilité.

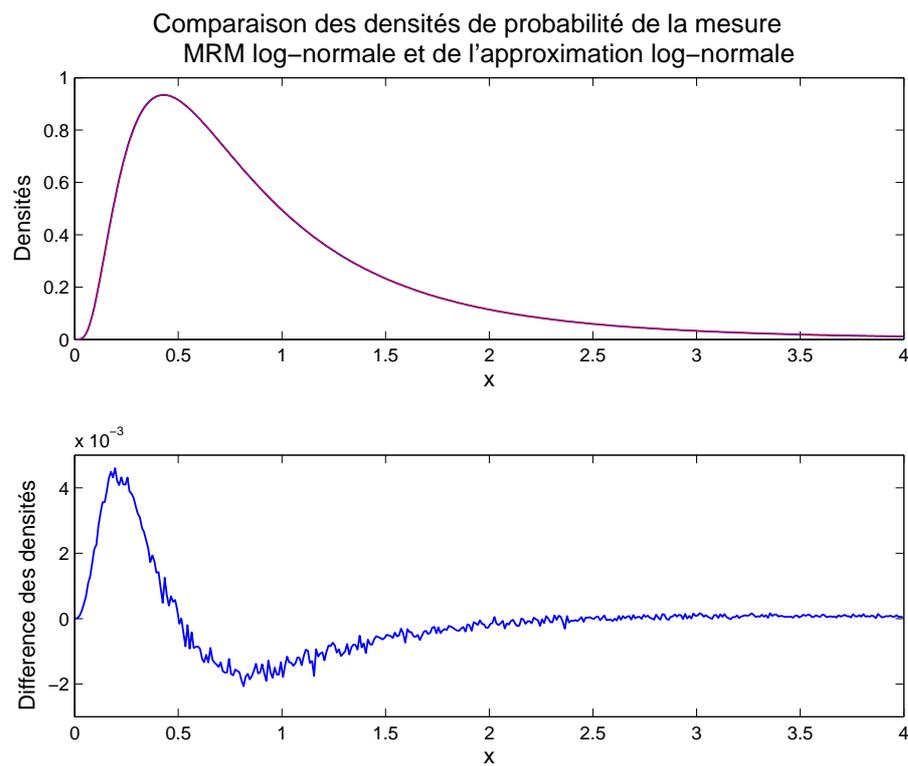


FIG. 5.4 – En haut, les densités de probabilité de la mesure MRM log-normale. L'une est obtenue par la méthode Monte Carlo de 10^9 tirages de paramètres $T = 250$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$, et l'autre est donnée par l'approximation log-normale. En bas, l'écart entre les densités de probabilité.

Estimation des paramètres

Malgré le grand nombre d'études consacrées aux modèles multifractals, très peu de travaux se sont intéressés à l'estimation de paramètres, voir néanmoins [OW00].

Nous rappelons que le modèle MRW log-normale, étudié en détails dans la section 3.4, est entièrement défini par l'ensemble de trois paramètres : la variance σ^2 , le temps intégral T et le coefficient d'intermittence λ^2 .

L'estimation de la variance σ^2 semble *a priori* ne pas poser de problèmes. Elle peut être estimée par la variance empirique des accroissements du processus MRW

$$\widehat{\sigma^2} = \frac{1}{N\tau} \sum_{k=1}^N (\delta_\tau X(k\tau))^2 \quad (6.1)$$

En revanche, l'estimation des deux autres paramètres n'est pas triviale.

Une méthode naïve d'estimation de λ^2 consiste à obtenir l'exposant multifractal $\zeta(q)$ (2.37) par régression du logarithme des moments empiriques des accroissements absolus contre le logarithme de l'échelle (cf. figure 2.5) puis à effectuer l'ajustement de l'exposant multifractal estimé $\widehat{\zeta}(q)$ par une fonction quadratique. Cette méthode manque de robustesse puisque il n'est pas possible de contrôler quantitativement les fluctuations statistiques des moments empiriques. Elle est néanmoins suffisante pour justifier l'approximation $\lambda^2 \ll 1$, [MDB00, BDM01].

Les résultats du chapitre 5 sur l'approximation limite de petite intermittence peuvent donc être appliqués. Notamment l'espérance et la fonction d'autocovariance du logarithme des accroissements du processus MRW log-normal à une échelle de temps quelconque sont données par la proposition 5.19. Il est donc naturel d'effectuer l'estimation des paramètres λ^2 et T du modèle MRW log-normal au moyen de la régression de la fonction d'autocovariance empirique correspondante.

Après avoir étudié les propriétés statistiques de ces estimateurs dans le cas du régime

asymptotique "historique infini"¹ (cf. section 6.1), c'est-à-dire dans le cas où l'échelle d'observation L (la longueur totale de la série) est telle que $L \gg T$, nous utilisons la méthode des moments généralisés (GMM) pour résoudre le problème d'estimation des paramètres du modèle log-normal (cf. section 6.2). L'étude des intervalles de confiance des paramètres par Monte Carlo (cf. section 6.2.3) montre que le paramètre T peut être considéré comme étant du même ordre de grandeur que l'échelle d'observation L . Le régime asymptotique $L \gg T$ "historique infini" n'est donc pas atteint. Les intervalles de confiance théoriques donnés par GMM ne sont donc pas utilisables. En fait, nous montrons (cf. section 6.3) que si T est de l'ordre de L , alors le paramètre T n'est pas réellement un paramètre et peut en fait être choisi arbitrairement. Il ne reste donc plus que deux paramètres à estimer : σ^2 et λ^2 . Les théorèmes 6.8 et 6.9 à la fin de ce chapitre montrent que le coefficient d'intermittence λ^2 peut être estimé dans une limite asymptotique "haute fréquence"², c'est-à-dire lorsque $\tau < T \simeq L$ (où τ est l'échelle d'échantillonnage) ce qui, en pratique, explique pourquoi les estimations GMM de λ^2 sont souvent bonnes même si la limite asymptotique "historique infini" est loin d'être atteinte.

6.1 Régression de la fonction d'autocovariance des magnitudes

6.1.1 Fonction d'autocovariance empirique

Définition 6.1

La *fonction d'autocovariance empirique* des observations $\{Y_\tau[k]; k = 1, \dots, N\}$, à l'échelle d'échantillonnage τ et à l'échelle d'observation $L = N\tau$, est donnée par

$$\hat{\gamma}_{\tau,N}[h] = \frac{1}{N} \sum_{k=1}^{N-h} (Y_\tau[k] - \hat{\mu}_{\tau,N})(Y_\tau[k+h] - \hat{\mu}_{\tau,N}), \quad (6.2)$$

où la *moyenne empirique* $\hat{\mu}_{\tau,N}$ est donnée par

$$\hat{\mu}_{\tau,N} = \frac{1}{N} \sum_{k=1}^N Y_\tau[k]. \quad (6.3)$$

Le lemme suivant permet de relier de façon statistique les estimateurs de la fonction d'autocovariance (6.2) et de l'espérance (6.3).

Lemme 6.1

La *fonction d'autocovariance empirique* et la *moyenne empirique*, définies par les équations (6.2) et (6.3), vérifient la relation suivante

$$\mathbb{E}[\hat{\gamma}_{\tau,N}[h]] - \gamma_\tau[h] = -\text{Var}[\hat{\mu}_{\tau,N}] - \frac{h}{N} \left(\gamma_\tau[h] + \text{Var}[\hat{\mu}_{\tau,N}] - 2\text{Cov}[\hat{\mu}_{\tau,h}, \hat{\mu}_{\tau,N}] \right). \quad (6.4)$$

¹Nous avons introduit la notion du régime asymptotique "historique infini" dans la section 4.2.2.

²Nous avons introduit la notion du régime asymptotique "haute fréquence" dans la section 4.2.2.

Preuve - 6.1 -

La fonction d'autocovariance empirique (6.2) peut s'écrire de la façon suivante

$$\hat{\gamma}_{\tau,N}[h] = \frac{1}{N} \sum_{k=1}^{N-k} ((Y_{\tau}[k] - \mu_{\tau}) - (\hat{\mu}_{\tau,N} - \mu_{\tau}))((Y_{\tau}[k+h] - \mu_{\tau}) - (\hat{\mu}_{\tau,N} - \mu_{\tau})), \quad (6.5)$$

Nous avons alors

$$\begin{aligned} \mathbb{E}[\hat{\gamma}_{\tau,N}[h]] &= \frac{N-h}{N} (\gamma_{\tau}[h] + \mathbb{V}\text{ar}[\hat{\mu}_{\tau,N}]) - 2\mathbb{E}\left[\frac{1}{N} \sum_{k=1}^{N-h} (Y_{\tau}[k] - \mu_{\tau})(\hat{\mu}_{\tau,N} - \mu_{\tau})\right] \\ &= \frac{N-h}{N} (\gamma_{\tau}[h] + \mathbb{V}\text{ar}[\hat{\mu}_{\tau,N}]) - 2\left(\mathbb{V}\text{ar}[\hat{\mu}_{\tau,N}] - \frac{h}{N} \mathbb{C}\text{ov}[\hat{\mu}_{\tau,h}, \hat{\mu}_{\tau,N}]\right). \end{aligned} \quad (6.6)$$

Ce qui nous conduit au résultat annoncé. \blacksquare

Puisque la variance et la covariance qui interviennent dans la relation (6.4) ne dépendent pas du décalage h , la fonction d'autocovariance empirique est *a priori* biaisée.

Dans la suite, nous allons étudier le comportement du biais (6.4) dans le cadre du modèle MRW log-normal en régime asymptotique "**historique infini**", introduit dans la section 4.2.2, qui correspond au cas limite où l'échelle d'observation $L = N\tau$ tend vers l'infini, alors que l'échelle d'échantillonnage τ reste fixée.

6.1.2 Régime asymptotique "historique infini"

Dans le régime asymptotique "historique infini", défini ci-dessus, nous avons le résultat classique suivant.

Proposition 6.2

Le biais (6.4) de la fonction d'autocovariance empirique tend vers zéro

$$\lim_{N \rightarrow \infty} \mathbb{E}[\hat{\gamma}_{\tau,N}[h]] = \gamma_{\tau}[h], \quad (6.7)$$

lorsque N tend vers l'infini et que τ et h sont fixés, si la limite suivante est égale à zéro

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{|k| < N} \left(1 - \frac{|k|}{N}\right) \gamma_{\tau}[k] = 0. \quad (6.8)$$

Dans le cadre du modèle MRW log-normal, la condition (6.8) est vérifiée puisque la fonction d'autocovariance $\gamma_{\tau}[h]$ s'annule pour les grandes valeurs du décalage h supérieures à T/τ en valeur absolue. La loi des grands nombres assure la convergence de la fonction d'autocovariance empirique vers celle théorique.

La figure 6.1 représente le comportement de la fonction d'autocovariance empirique dans le régime asymptotique "historique infini". Il est simple de voir que l'écart entre les traits empirique et théorique est assez faible, ce qui est en accord avec la proposition 6.2.

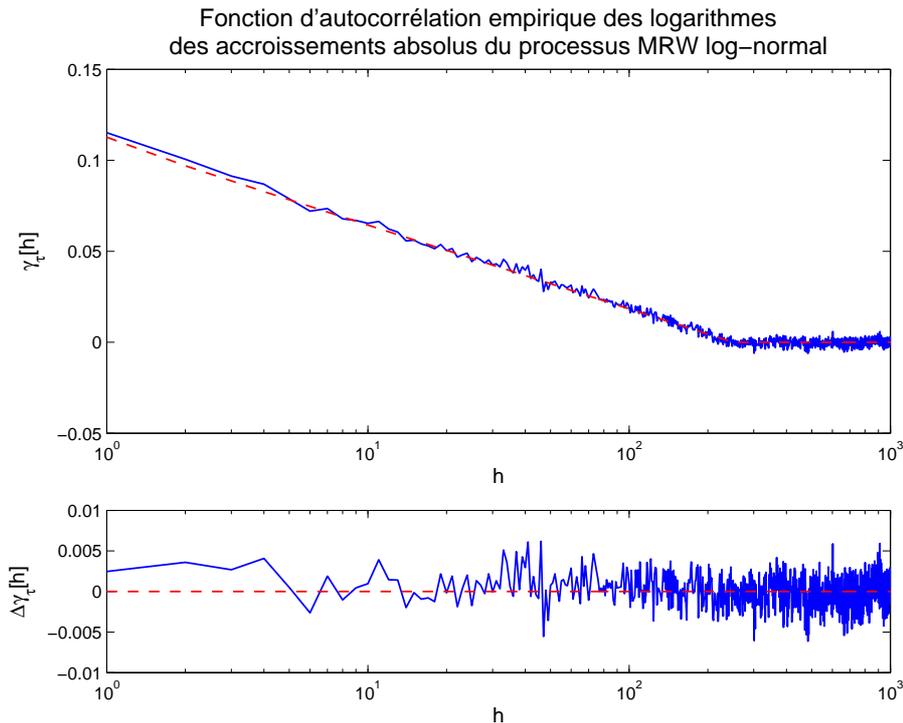


FIG. 6.1 – En haut, la fonction d'autocovariance empirique des magnitudes obtenue à partir d'une réalisation du processus MRW log-normal ($N = 2^{19}$ points, $\tau = 1$) de paramètres $T = 250$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$. Le trait tireté correspond aux valeurs théoriques. En bas, l'écart entre les fonctions d'autocovariance empirique et théorique est très faible, ce qui est en accord avec la proposition 6.2.

6.2 Estimation des paramètres par la méthode GMM

Il est simple de voir que tous les paramètres sont directement liés à certains moments des accroissements du processus MRW ou de leur logarithme. Il semble donc naturel d'utiliser la **méthode des moments généralisés** (GMM).

Les premières applications de la méthode GMM en finance ont été faites dans le cadre des modèles à volatilité stochastique (cf. section 1.3.4) [Tay86, MT90, JPR94, AS96, AS97, ACS99]. Dans le cadre des modèles multifractals, Calvet et Fisher ont utilisé [CF01, CF04] cette méthode pour estimer les paramètres de leur modèle MSM (cf. section 2.4.3), dans le cas où les composantes de volatilité suivent une loi binomiale. Dans le cadre du modèle MSM avec composantes de volatilité log-normale, Lux a effectué [Lux01, Lux03, Lux04] une comparaison de la méthodes GMM avec la méthode du maximum de vraisemblance.

Les résultat préliminaires de cette section ont été publiés dans [BKM06].

6.2.1 Méthode des moments généralisés

La **méthode des moments généralisés** ou **GMM** (de l'anglais *Generalized Method of Moment*) a été initialement proposé par Hansen dans [Han82], mais l'idée de base remonte à Sargan [Sar58].

Nous considérons le processus $\{Y_{\Theta}[k]\}_k$ caractérisé par le vecteur de paramètres Θ de dimension p . Etant donné une observation $\{Y_{\Theta_0}[k]\}_k$, notons $f(Y_{\Theta_0}[k], \Theta)$ la fonction de dimension r , supérieure à p , qui vérifie la **condition sur les moments**

$$\mathbb{E}[f(Y_{\Theta_0}[k], \Theta)] = 0, \text{ si et seulement si } \Theta = \Theta_0. \quad (6.9)$$

La condition sur les moments peut être approximée en utilisant la moyenne empirique

$$g_N(\Theta) = \frac{1}{N} \sum_{k=1}^N f(Y_{\Theta_0}[k], \Theta). \quad (6.10)$$

Définition 6.2

L'**estimateur GMM** est défini par

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} (g_N^T W_N g_N), \quad (6.11)$$

où W_N est une suite de matrices de pondération qui converge vers une matrice W_{∞} symétrique définie positive lorsque N tend vers $+\infty$.

Théorème 6.3 (Hansen)

Si les propriétés suivantes sont vérifiées

- le processus $\{Y[k]\}_k$ est ergodique,
- la série $\{f(Y_{\Theta_0}[k], \Theta)\}_k$ vérifie le théorème central limite, c'est-à-dire

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N f(Y_{\Theta_0}[k], \Theta) \rightarrow \mathcal{N}(0, V_{\Theta}), \quad (6.12)$$

où la matrice V_{Θ} est définie par

$$V_{\Theta} = \lim_{K \rightarrow +\infty} \sum_{k=-K}^K \mathbb{E}[f(Y_{\Theta_0}[k], \Theta) f(Y_{\Theta_0}[k], \Theta)^T]. \quad (6.13)$$

- la matrice $(r \times p)$ -dimensionnelle $Dg_N = \frac{\partial g_N}{\partial \Theta}$ a un rang p et elle converge vers

$$Df = \mathbb{E} \left[\frac{\partial f(Y_{\Theta_0}[k], \Theta)}{\partial \Theta} \right], \quad (6.14)$$

alors l'estimateur GMM $\hat{\Theta}$ est consistant et

$$\sqrt{N}(\hat{\Theta} - \Theta) \rightarrow \mathcal{N}(0, \Sigma), \quad (6.15)$$

où

$$\Sigma = (Df^T W_\infty Df)^{-1} Df^T W_\infty V_{\Theta_0} W_\infty Df (Df^T W_\infty Df)^{-1}. \quad (6.16)$$

Preuve - 6.3 -

La démonstration est détaillée dans [Han82]. ■

Proposition 6.4

L'estimateur GMM est optimal si la limite W_∞ des matrices de pondération W_N (6.11) est égale à $V_{\Theta_0}^{-1}$, définie par l'équation (6.13). Dans ce cas la matrice de variance-covariance de l'estimateur GMM est donnée par

$$\Sigma_{opt} = (Df^T V_{\Theta_0}^{-1} Df)^{-1}. \quad (6.17)$$

Preuve - 6.4 -

La démonstration est détaillée dans [Han82]. ■

En pratique, il est bien évidemment impossible d'utiliser la matrice de pondération optimale $V_{\Theta_0}^{-1}$, puisque le vecteur Θ_0 est inconnu. On suit donc généralement la procédure récursive suivante :

1. On choisit la matrice de pondération W_N quelconque, on peut prendre par exemple la matrice \mathbb{I}_N ou encore la matrice $V_{\hat{\Theta}}^{-1}$, où $\hat{\Theta}$ est un estimateur *a priori* de Θ_0 ,
2. On calcule (6.11) l'estimateur GMM $\hat{\Theta}$ en utilisant la matrice de pondération W_N ,
3. On fixe la matrice de pondération $W_N = V_{\hat{\Theta}}^{-1}$, où $\hat{\Theta}$ est l'estimateur GMM obtenu à l'étape 2,
4. On répète les étapes 2 et 3 jusqu'à ce que les valeurs successives de $\hat{\Theta}$ soient suffisamment proches l'une de l'autre.

Les intervalles de confiance de l'estimation $\hat{\Theta}$ peuvent être déterminés à l'aide de l'équation (6.15). Sous condition que le régime asymptotique est atteint. Si ce n'est pas le cas, seule une méthode de Monte Carlo permet d'obtenir ces intervalles de confiance.

6.2.2 Estimation des paramètres du modèle MRW log-normal

Soit $\{Z_\tau[k]\}_{k=1, \dots, N}$ la suite de logarithmes des accroissements absolus du processus MRW log-normale à l'échelle d'échantillonnage τ . Considérons le vecteur de paramètres

$$\Theta = \{\ln(\sigma), \lambda^2, \ln(T)\}. \quad (6.18)$$

Pour les conditions sur les moments (6.9), il paraît naturel de choisir la variance du processus $\{e^{2Z_\tau[k]}\}_k$, pour l'estimation de σ^2 , et la fonction d'autocovariance empirique du processus $\{Z_\tau[k]\}_k$ à différents décalages, pour l'estimation de λ^2 et de T (cf. section 6.1).

Ainsi nous considérons

$$g_N(\Theta) = \frac{1}{N} \sum_{k=1}^N \begin{pmatrix} \exp(2Z_\tau[k]) \\ (Z_\tau[k] - \mu_\tau)(Z_\tau[k - h_1] - \mu_\tau) \\ \vdots \\ (Z_\tau[k] - \mu_\tau)(Z_\tau[k - h_K] - \mu_\tau) \end{pmatrix} - \begin{pmatrix} \sigma^2\tau \\ C_{h_1} \\ \vdots \\ C_{h_K} \end{pmatrix}, \quad (6.19)$$

où

$$\mu_\tau = \mathbb{E}[Z_\tau[k]] = \frac{1}{2} \ln(\tau) + \mathbb{E}[\ln(|\epsilon[k]|)] + \lambda \mathbb{E}\left[\frac{\delta_\tau \Omega(k\tau)}{\tau}\right] \quad (6.20)$$

et

$$C_h = \text{Cov}[Z_\tau[k], Z_\tau[k - h]], \quad (6.21)$$

et h_1, \dots, h_K sont K différents décalages positifs et $\{\Omega(t)\}_t$ est la magnitude renormalisée (cf. section 5.3). Notons que les expressions analytiques de μ et C_h en fonction de σ , λ^2 et T sont données par la proposition 5.19.

6.2.3 Application aux données journalières

Dans cette section, nous présentons les résultats d'estimation par la méthode GMM des paramètres du modèle MRW log-normal pour les séries financières journalières.

Notons que pour unifier la procédure d'estimation des paramètres, les rendements journaliers ont été multipliés par un facteur 100, comme dans le cadre du modèle GARCH (cf. section 1.3.3).

L'application de la méthode GMM aux données financières pose le problème de traitement des accroissements nuls. Il arrive en effet que le prix de clôture du marché soit le même d'un jour sur l'autre. Il n'est pas alors possible de calculer le logarithme de tels accroissements. Pour éviter ce problème, nous utilisons la valeur du *tick* (cf. section 1.2.1) pour remplacer les accroissements du prix égaux à zéro. La valeur du *tick* peut être estimée par la valeur minimale d'accroissement absolu de prix observée.

Dans les tableaux 6.1 et 6.2 nous avons représenté les résultats de l'estimation des paramètres du modèle MRW log-normal par la méthode GMM. Pour l'estimation, nous avons utilisé les décalages $h_k = k$ jours, pour $1 \leq k \leq 100$. Nous pouvons remarquer que, très souvent, les estimations des temps intégrals T conduisent à des valeurs d'un ordre comparable à la taille de l'échantillon (29 ans, soit $\simeq 7200$ points, pour les taux de change et 15 ans, soit $\simeq 3770$ points, pour les composants de l'indice CAC 40). Le régime asymptotique "historique infini" n'est donc pas atteint. L'échelle d'observation des séries n'est pas suffisamment grande. Les intervalles théoriques prévus par GMM ne sont donc

	GMM				
	$\ln(\sigma)$	λ^2	$\ln(T)$	σ	T (jours)
CAD/USD	-1.121	0.0254	8.593	0.3260	5394
JPY/USD	-0.381	0.0224	5.579	0.6832	265
CHF/USD	-0.278	0.0186	5.074	0.7574	160
GBP/USD	-0.473	0.0193	6.975	0.6234	1070

TAB. 6.1 – Forex. Résultats de l'estimation des paramètres du modèle MRW log-normal. Les trois premières colonnes présentent les estimations par la méthode GMM. Les deux colonnes suivantes présentent les valeurs calculées à partir des trois premières. Pour l'estimation nous avons utilisé les décalages $h_k = k$ jours, pour $1 \leq k \leq 100$.

pas utilisables. Seul une méthode de Monte Carlo permet donc d'obtenir ces intervalles de confiance.

Une expérience de Monte Carlo avec 15000 tirages du processus MRW log-normal de λ^2 égal à 0.02 et de temps intégral égal à l'échelle d'observation des données journalières, montre que les intervalles de confiance pour l'estimateur du coefficient d'intermittence λ^2 sont de l'ordre $\pm 35\%$ et les intervalles de confiance pour l'estimateur du logarithme du temps intégral $\ln(T)$ sont de l'ordre $\pm 40\%$. En d'autres termes, l'estimation du coefficient d'intermittence par la méthode GMM vérifie la propriété suivante :

– si le coefficient d'intermittence $\lambda^2 = 0.02$ alors son estimateur $\widehat{\lambda^2} \in [0.013, 0.027]$,

et l'estimation du temps intégral par la méthode GMM vérifie les propriétés suivantes :

– si le temps intégral $T = 3770$ ($\simeq 15$ ans) alors son estimateur $\widehat{T} \in [140, 100000]$,

– si le temps intégral $T = 7200$ ($\simeq 29$ ans) alors son estimateur $\widehat{T} \in [200, 250000]$.

Ces résultats montrent que l'on peut considérer que, pour tous les taux de change et pour toutes les composantes du CAC 40, le temps intégral est de l'ordre de l'échelle d'observation.

Il est donc important de bien comprendre les problèmes d'estimation lorsque le temps intégral T est du même ordre que (voire supérieur à) l'échelle d'observation.

6.3 Estimation des paramètres lorsque le temps intégral est grand

Il faut noter que rien n'assure, en pratique, que l'échelle d'observation L est supérieure au temps intégral T . Ainsi, si le temps intégral T est supérieur à l'échelle d'observation L , d'après le théorème 3.17, le processus MRW log-normal $\{X_T(t)\}_{t \leq L}$ vérifie l'égalité en loi suivante

$$\{X_T(t)\}_{t \leq L} \stackrel{\mathcal{L}}{=} \{W_{L/T} X_L(t)\}_{t \leq L}, \quad (6.22)$$

où $W_{L/T}$ est une variable aléatoire log-normale dont la loi de probabilité est précisée dans le théorème 3.17 et qui est indépendante du processus MRW log-normal $\{X_L(t)\}_{t \leq L}$ de

	GMM			σ	T (jours)
	$\ln(\sigma)$	λ^2	$\ln(T)$		
Accor	0.711	0.0327	6.792	2.0363	891
Air Liquide	0.533	0.0157	8.183	1.7039	3580
Alcatel	1.122	0.0157	19.817	3.0718	404145504
Axa	0.812	0.0303	7.796	2.2529	2430
Bouygues	0.823	0.0228	8.341	2.2774	4193
Capgemini	1.093	0.0283	8.936	2.9834	7605
Carrefour	0.624	0.0183	8.220	1.8665	3716
Casino Guichard	0.626	0.0338	5.340	1.8707	209
Danone	0.420	0.0126	8.808	1.5218	6687
Essilor International	0.684	0.0266	6.384	1.9822	592
L'Oréal	0.675	0.0108	9.265	1.9641	10561
Lafarge	0.692	0.0156	7.343	1.9968	1546
Lagardere	0.908	0.0613	6.186	2.4800	486
LVMH	0.715	0.0275	7.602	2.0440	2003
Michelin	0.739	0.0174	6.579	2.0934	720
Pernod Ricard	0.639	0.0169	6.546	1.8945	697
PSA Peugeot Citroën	0.643	0.0251	5.737	1.9015	310
Pinault Printemps	0.761	0.0555	6.380	2.1411	590
Publicis	0.895	0.0473	8.317	2.4484	4092
Saint Gobain	0.720	0.0231	7.153	2.0554	1278
Sanofi-Aventis	0.709	0.0219	6.983	2.0326	1078
Schneider Electric	0.821	0.0216	6.874	2.2719	967
Société Générale	0.757	0.0219	7.993	2.1324	2959
Suez	0.719	0.0290	6.774	2.0518	875
TF1	0.920	0.0294	8.865	2.5094	7079
Thales	0.875	0.0272	6.250	2.3995	518
Total	0.594	0.0181	8.930	1.8116	7553
Vinci	0.727	0.0197	7.103	2.0688	1215
Vivendi Universal	0.888	0.0233	10.491	2.4294	35983

TAB. 6.2 – CAC 40. Résultats de l'estimation des paramètres du modèle MRW log-normal. Les trois premières colonnes présentent les estimations par la méthode GMM. Les deux colonnes suivantes présentent les valeurs calculées à partir des trois premières. Et la dernière colonne contient les résultats de la régression de la fonction d'autocovariance empirique non centrée. Pour les deux méthodes nous avons utilisé les décalages $h_k = k$ jours, pour $1 \leq k \leq 100$.

temps intégral égal à l'échelle d'observation L .

Etant donné une réalisation de taille L du processus MRW, la variable $W_{L/T}$ prend donc une valeur unique pour cette réalisation, et elle peut être considérée comme un simple facteur multiplicatif qui ne fait que modifier la variance du processus. En conséquence, il n'existe pas de possibilité d'estimer, dans ce cas, ni le temps intégral T , ni la variance σ^2 , sans information *a priori* sur ces valeurs.

Il est donc important, dans un premier temps, de comprendre à quelles estimations la méthode de régression de la fonction d'autocovariance empirique des magnitudes (cf. section 6.1) ou la méthode GMM (cf. section 6.2) conduisent.

6.3.1 Estimation du temps intégral

Afin d'alléger les calculs, nous considérons le régime asymptotique **"haute fréquence"**, introduit dans la section 4.2.2, qui correspond au cas limite où l'échelle d'échantillonnage τ tend vers zéro, alors que l'échelle d'observation $L = N\tau$ inférieure au temps intégral T , reste fixée.

Proposition 6.5

Fixons $L = N\tau$ l'échelle d'observation, où τ est l'échelle d'échantillonnage et N est le nombre d'observations. Supposons que $L < T$. Fixons un réel positif h et considérons le régime asymptotique "haute fréquence" (de telle façon que h/τ prend des valeurs entières), alors l'espérance de la fonction d'autocovariance empirique (6.2) est donnée par

$$\begin{aligned} \mathbb{E}[\widehat{\gamma}_\tau[h/\tau]] &= \lambda^2 \left[\ln \left(\frac{L}{he^{3/2}} \right) - \frac{h}{L} \ln \left(\frac{L}{he^{3/2}} \right) \right] \\ &\quad + \lambda^2 \left[\frac{h^2}{L^2} \ln \left(\frac{h}{L} \right) + \frac{(L-h)^2}{2L^2} \ln \left(1 - \frac{h}{L} \right) \right] - \left(1 - \frac{h}{L} \right) \frac{\pi^2}{8L} + o(\lambda^2), \end{aligned} \quad (6.23)$$

où $o(\lambda^2)$ ne dépend que de λ^2 .

Preuve - 6.5 -

Dans le régime asymptotique "haute fréquence", sous condition de grand temps intégral, $L \leq T$, la fonction d'autocovariance $\gamma_\tau[h/\tau]$ est donnée par

$$\gamma_\tau[h/\tau] = \lambda^2 \ln \left(\frac{T}{h} \right) + o(\lambda^2), \quad (6.24)$$

où $o(\lambda^2)$ ne dépend que de λ^2 (cf. proposition 5.15).

En utilisant la définition 6.3 de la moyenne empirique et la proposition 5.19, il est simple de voir que les approximations intégrales suivantes sont valides

$$\begin{aligned} \mathbb{V}\text{ar}[\hat{\mu}_{\tau,N}] &= \lambda^2 \int_0^L \frac{du}{L} \int_0^L \frac{dv}{L} \ln \left(\frac{T}{|u-v|} \right) + \frac{\pi^2}{8L} + o(\lambda^2) \\ &= \lambda^2 \ln \left(\frac{T e^{3/2}}{L} \right) + \frac{\pi^2}{8L} + o(\lambda^2), \end{aligned} \quad (6.25)$$

et

$$\begin{aligned} \mathbb{C}\text{ov}[\hat{\mu}_{\tau,h/\tau}, \hat{\mu}_{\tau,N}] &= \lambda^2 \int_0^h \frac{du}{h} \int_0^L \frac{dv}{L} \ln \left(\frac{T}{|u-v|} \right) + \frac{\pi^2}{8L} + o(\lambda^2) \\ &= \lambda^2 \ln \left(\frac{T e^{3/2}}{L} \right) - \lambda^2 \frac{h}{2L} \ln \left(\frac{h}{L} \right) + \lambda^2 \frac{(L-h)^2}{2Lh} \ln \left(1 - \frac{h}{L} \right) + \frac{\pi^2}{8L} + o(\lambda^2) \end{aligned} \quad (6.26)$$

Ce qui conduit au résultat annoncé (6.23), d'après le lemme 6.1. ■

L'espérance de la fonction d'autocovariance empirique des magnitudes (6.23) ne dépend pas du temps intégral. Le terme dominant de l'équation (6.23) est

$$\mathbb{E}[\hat{\gamma}_{\tau}[h]] \simeq \lambda^2 \ln \left(\frac{N}{h e^{3/2}} \right) = \lambda^2 \ln \left(\frac{L e^{-3/2}}{h} \right), \quad (6.27)$$

ce qui signifie que les valeurs de l'estimation du temps intégral obtenue par la régression de la fonction d'autocovariance empirique des magnitudes sont de l'ordre de $L e^{-3/2}$ indépendamment de la vraie valeur du temps intégral. En d'autres termes, nous pouvons faire confiance au résultat de l'estimation de T seulement si cette estimation est beaucoup plus petite que l'échelle d'observation. À l'inverse, les estimations qui sont du même ordre de grandeur³ que l'échelle d'observation peuvent être considérées comme l'indication d'un grand temps intégral, ce qui est le cas de beaucoup données financières (cf. tableaux 6.1 et 6.2).

6.3.2 Estimation de la variance

Comme nous l'avons vu (introduction de la section 6.3), il n'est pas possible d'estimer le paramètre de variance σ^2 , lorsque l'échelle d'observation L est inférieure au temps intégral T , à partir d'une seule réalisation. En effet, une seule réalisation du processus MRW log-normal de grand temps intégral peut être vue comme une réalisation du processus MRW

³Voir l'ordre de grandeur des intervalles de confiance dans la section 6.2.3.

log-normal de temps intégral égal à l'échelle d'observation multiplié par un facteur $W_{L/T}$ constant, que l'on ne peut pas estimer.

Si nous fixons arbitrairement le paramètre du temps intégral T égal à l'échelle d'observation L alors il est possible d'estimer le produit $\sigma^2 W_{L/T}$, qui joue le rôle d'une variance "locale" du processus MRW log-normal, par la variance empirique (6.1).

6.3.3 Estimation du coefficient d'intermittence

L'équation (6.23) montre que le comportement de la fonction d'autocovariance empirique des magnitudes en fonction du décalage h reste correcte, même si le temps intégral est supérieur à l'échelle d'observation. Ainsi, seul le coefficient d'intermittence λ^2 semble pouvoir être estimé de façon robuste à partir d'une seule réalisation du processus MRW log-normal.

Dans la section 6.4 nous démontrons que l'estimateur du coefficient d'intermittence λ^2 correspondant à la régression de la fonction d'autocovariance empirique des magnitudes est consistant.

6.3.4 Modèle MRW log-normal n'a que deux paramètres effectifs

Comme nous l'avons vu si l'on fixe arbitrairement le paramètre du temps intégral T à l'échelle d'observation L il ne reste que deux paramètres du modèle : la variance locale $\sigma^2 W_{L/T}$, que l'on peut estimer à partir de la variance empirique, et le coefficient d'intermittence λ^2 , que l'on peut estimer à partir de la fonction d'autocovariance empirique, comme nous le verrons dans la section suivante.

Pour effectuer les prédictions de volatilité et de la Valeur à Risque (cf. chapitres 7, 8 et 9) nous fixerons, en pratique, le temps intégral T à l'échelle d'observation L .

6.4 Estimation du coefficient d'intermittence

Tout au long de cette section nous supposons que l'échelle d'observation L est inférieure au temps intégral T . Dans la section précédente nous avons montré qu'il n'est pas possible d'estimer le temps intégral T et la variance σ^2 à partir d'une seule réalisation. Rien n'empêche par contre d'estimer le coefficient d'intermittence λ^2 .

La figure 6.2 représente le comportement de la fonction d'autocovariance empirique lorsque l'échelle d'observation est inférieure au temps intégral. Il est simple de voir que l'écart entre les traits empirique et théorique change peu en fonction du décalage h . Cette observation donne à penser qu'il est possible d'estimer le coefficient d'intermittence λ^2 par la régression de la fonction d'autocovariance empirique.

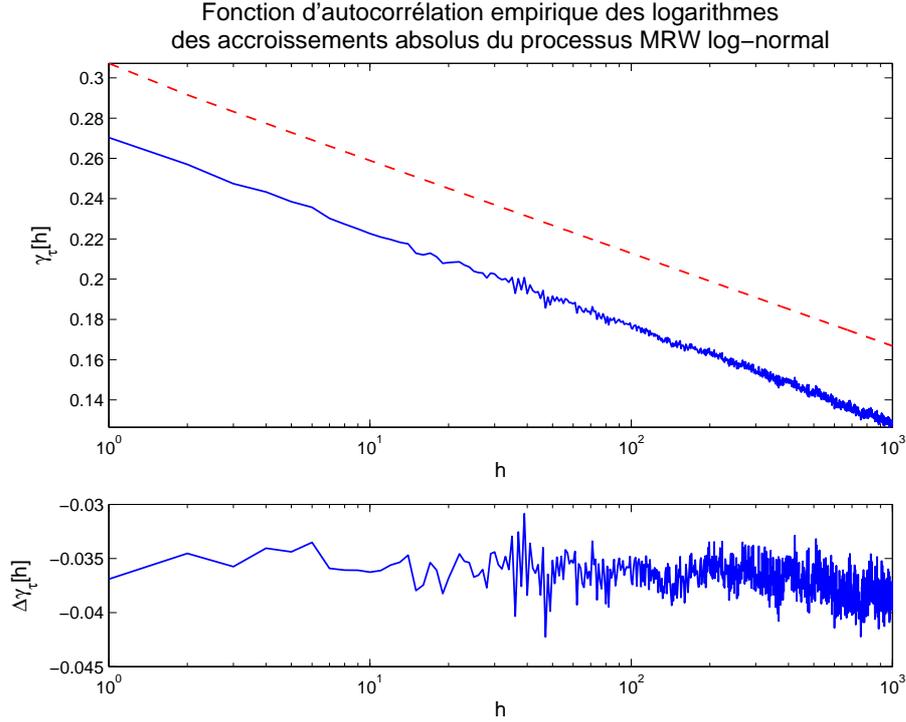


FIG. 6.2 – En haut, la fonction d'autocovariance empirique des magnitudes obtenue à partir d'une réalisation du processus MRW log-normal ($N = 2^{19}$ points et $\tau = 1$) de paramètres $T = 2^{22}$, $\lambda^2 = 0.02$ et $\sigma^2 = 1$. Le trait tireté correspond aux valeurs théoriques. En bas, l'écart entre les fonctions d'autocovariance empirique et théorique qui semble peu varier en fonction du décalage h .

Pour effectuer un calcul formel, nous fixons deux différents décalages h et h' tels que les intervalles de temps $h\tau$ et $h'\tau$, où τ est l'échelle d'échantillonnage, sont inférieurs à l'échelle d'observation L , et en conséquence, ils sont inférieurs au temps intégral T . Considérons la suite $\{Z_\tau[k]\}_{k=1,\dots,N}$, qui peut représenter, dans cette section, soit la série des logarithmes de la mesure MRM log-normale soit la série des logarithmes des accroissements absolus du processus MRW log-normal.

D'après les théorèmes 5.16 et 5.17, pour h entier différent de zéro, nous avons

$$\text{Cov}[Z_\tau[k], Z_\tau[k+h]] = \lambda^2 \left(\ln \left(\frac{T e^{3/2}}{\tau} \right) - f[h] \right) + o(\lambda^2), \quad (6.28)$$

où $o(\lambda^2)$ dépend du coefficient d'intermittence λ^2 et ne dépend ni du temps intégral T , ni de l'échelle d'échantillonnage τ , et la fonction $f[h]$ est donnée par la proposition 5.11. En conséquence, il est simple de voir que la différence entre les covariances (6.28) pour deux décalages h et h' ne dépend ni du temps intégral T , ni de l'échelle d'échantillonnage τ . Ce qui nous conduit naturellement à l'estimation du coefficient d'intermittence λ^2 par la régression.

Définition 6.3

L'estimateur du coefficient d'intermittence est donné par

$$\widehat{\lambda^2} = \frac{\widehat{\gamma}_{\tau,N}^0[h] - \widehat{\gamma}_{\tau,N}^0[h']}{f[h] - f[h']}, \quad (6.29)$$

où la **fonction d'autocovariance empirique non centrée** $\widehat{\gamma}_{\tau,N}^0[h]$ est donnée par

$$\widehat{\gamma}_{\tau,N}^0[h] = \frac{1}{N} \sum_{k=1}^{N-k} Y_{\tau}[k]Y_{\tau}[k+h]. \quad (6.30)$$

Le but de cette section est d'étudier les propriétés statistiques de l'estimateur (6.29) dans le régime asymptotique "données à haute fréquence".

6.4.1 Représentation matricielle de la fonction d'autocovariance empirique

Définition 6.4

Nous définissons la **matrice de décalage** de dimension $N \times N$

$$S = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad (6.31)$$

dont la $n^{\text{ième}}$ puissance est donnée par ses éléments

$$(S^n)_{jk} = \begin{cases} 1, & \text{si } k - j = n, \\ 0, & \text{sinon.} \end{cases} \quad (6.32)$$

Lemme 6.6

La fonction d'autocovariance empirique non centrée $\widehat{\gamma}_{\tau,N}^0[h]$, définie par l'équation (6.30), admet la représentation matricielle suivante

$$\widehat{\gamma}_{\tau,N}^0[h] = \frac{1}{N} \sum_{j,k=1}^N (S^h)_{jk} Y_{\tau}[j]Y_{\tau}[k]. \quad (6.33)$$

Preuve - 6.6 -

Notons que la multiplication du vecteur d'observations $(Y_{\tau}[1], \dots, Y_{\tau}[N])$ par la matrice

de décalage S effectue le décalage des indices

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} Y_\tau[1] \\ Y_\tau[2] \\ \vdots \\ Y_\tau[N-1] \\ Y_\tau[N] \end{pmatrix} = \begin{pmatrix} Y_\tau[2] \\ Y_\tau[3] \\ \vdots \\ Y_\tau[N] \\ 0 \end{pmatrix}. \quad (6.34)$$

Ce qui conduit directement à la représentation (6.33) de la fonction $\hat{\gamma}_{\tau,N}^0[h]$. ■

Définition 6.5

Nous définissons la différence des fonctions d'autocovariance empirique (6.30)

$$\hat{\gamma}_{\tau,N}^0[h, h'] = \hat{\gamma}_{\tau,N}^0[h] - \hat{\gamma}_{\tau,N}^0[h']. \quad (6.35)$$

Notons que, d'après le lemme 6.6, $\hat{\gamma}_{\tau,N}^0[h, h']$ s'écrit sous forme matricielle

$$\hat{\gamma}_{\tau,N}^0[h, h'] = \frac{1}{N} \sum_{j,k=1}^N (S^h - S^{h'})_{jk} Y_\tau[j] Y_\tau[k]. \quad (6.36)$$

6.4.2 Propriétés de la différence des fonctions d'autocovariance empiriques

Proposition 6.7

Si la suite $\{Y_\tau[k]\}_k$ est d'espérance μ , alors la différence des fonctions d'autocovariance empiriques non centrées (6.35) est asymptotiquement non biaisée

$$\mathbb{E}[\hat{\gamma}_{\tau,N}^0[h, h']] = \gamma_\tau[h] - \gamma_\tau[h'] + O\left(\frac{1}{N}\right), \quad (6.37)$$

et sa variance est donnée par la représentation matricielle suivante

$$\text{Var}[\hat{\gamma}_{\tau,N}^0[h, h']] = \frac{1}{N^2} \sum_{j,k,l,m=1}^N (DS^{hh'})_{jk} (DS^{hh'})_{lm} \text{Cov}[Y_\tau[j] Y_\tau[k], Y_\tau[l] Y_\tau[m]], \quad (6.38)$$

où $DS^{hh'}$ désigne la différence

$$DS^{hh'} = S^h - S^{h'}. \quad (6.39)$$

Preuve - 6.7 -

Il est simple de voir que l'espérance de $\hat{\gamma}_{\tau,N}^0[h, h']$ est donnée au premier ordre en $\frac{1}{N}$ par

$$\mathbb{E}[\hat{\gamma}_{\tau,N}^0[h]] = \frac{N-h}{N} (\gamma_\tau[h] + \mu^2) = \gamma_\tau[h] + \mu^2 + O\left(\frac{1}{N}\right). \quad (6.40)$$

En conséquence, $\widehat{\gamma}_{\tau,N}^0[h, h']$ est asymptotiquement non biaisée

$$\mathbb{E}[\widehat{\gamma}_{\tau,N}^0[h, h']] = \mathbb{E}[\widehat{\gamma}_{\tau,N}^0[h]] - \mathbb{E}[\widehat{\gamma}_{\tau,N}^0[h']] = \gamma_\tau[h] - \gamma_\tau[h'] + O\left(\frac{1}{N}\right). \quad (6.41)$$

La variance s'exprime en terme de covariance comme

$$\text{Var}[\widehat{\gamma}_{\tau,N}^0[h, h']] = \text{Cov}[\widehat{\gamma}_{\tau,N}^0[h, h'], \widehat{\gamma}_{\tau,N}^0[h, h']], \quad (6.42)$$

et la représentation (6.36) nous conduit au résultat souhaité. \blacksquare

6.4.3 Application au logarithme de la mesure MRM log-normale

Dans cette section, appliquons la proposition 6.7 au cas où

$$Z_\tau[k] = \frac{1}{2} \ln \left(\frac{M([(k-1)\tau, k\tau])}{\tau} \right). \quad (6.43)$$

Théorème 6.8

Dans le régime asymptotique "données à haute fréquence", l'estimateur (6.29) est un estimateur asymptotiquement non biaisé dont la variance est de l'ordre de

$$\text{Var}[\widehat{\lambda^2}] = \lambda^4 O\left(\frac{\ln(N)}{N}\right) + o(\lambda^4). \quad (6.44)$$

Preuve - 6.8 -

En utilisant la proposition 6.7, il est simple de voir que l'espérance de l'estimateur est asymptotiquement non biaisée

$$\mathbb{E}[\widehat{\lambda^2}] = \frac{\gamma_\tau[h] - \gamma_\tau[h']}{f[h] - f[h']} + O\left(\frac{1}{N}\right) = \lambda^2 + O\left(\frac{1}{N}\right) + o(\lambda^2). \quad (6.45)$$

Dans le régime asymptotique "données à haute fréquence" le terme $O\left(\frac{1}{N}\right)$ devient nulle.

Pour majorer la variance de l'estimateur, nous utilisons la représentation (6.38) et l'identité

$$Z_\tau[j]Z_\tau[k] = (Z_\tau[j] - \mu_\tau)(Z_\tau[k] - \mu_\tau) + \mu_\tau(Z_\tau[j] - \mu_\tau + Z_\tau[k] - \mu_\tau) + \mu_\tau^2. \quad (6.46)$$

Comme nous l'avons vu (cf. théorème 5.16) les moments généralisés des variables $\{Z_\tau[k]\}_{k=1,\dots,N}$ sont égaux au premier ordre non trivial en λ^2 aux moments généralisés de variables gaussiennes, d'après le théorème de Wick [Wic50], nous avons

$$\begin{aligned} \text{Cov}[(Z_\tau[j] - \mu_\tau)(Z_\tau[k] - \mu_\tau), (Z_\tau[l] - \mu_\tau)(Z_\tau[m] - \mu_\tau)] \\ = \gamma_\tau[j - l]\gamma_\tau[k - m] + \gamma_\tau[j - m]\gamma_\tau[k - l] + o(\lambda^4). \end{aligned} \quad (6.47)$$

En effectuant le contrôle des éléments $(DS^{hh'})_{jk}$ non nuls, d'après la proposition 5.15 nous avons

$$\begin{aligned} \mathbb{V}\text{ar}[\widehat{\gamma}_{\tau,N}^0[h, h']] &= \frac{1}{N^2} \sum_{j=1}^{N-h} \sum_{k=1}^{N-h} (\gamma_{\tau}[j-k]^2 + \gamma_{\tau}[j-k+h]\gamma_{\tau}[j-k-h]) \\ &\quad - \frac{2}{N^2} \sum_{j=1}^{N-h} \sum_{k=1}^{N-h'} (\gamma_{\tau}[j-k]\gamma_{\tau}[j-k+h-h'] + \gamma_{\tau}[j-k+h]\gamma_{\tau}[j-k-h']) \\ &\quad + \frac{2}{N^2} \sum_{j=1}^{N-h'} \sum_{k=1}^{N-h'} (\gamma_{\tau}[j-k]^2 + \gamma_{\tau}[j-k+h']\gamma_{\tau}[j-k-h']) + o(\lambda^4), \end{aligned} \quad (6.48)$$

où nous avons utilisé le fait que $\mu_{\tau} = O(\lambda^2)$.

En prenant les limites de sommation (6.48) de 0 à N , nous introduisons une erreur de l'ordre de $\frac{1}{N}$. Dans ce cas la variance de $\widehat{\gamma}_{\tau,N}^0[h, h']$ s'écrit

$$\begin{aligned} \mathbb{V}\text{ar}[\widehat{\gamma}_{\tau,N}^0[h, h']] &= \frac{1}{N^2} \sum_{|k|<N} (N-|k|)(\gamma_{\tau}[k+h] - \gamma_{\tau}[k+h'])(\gamma_{\tau}[k-h] - \gamma_{\tau}[k-h']) \\ &\quad + \frac{2}{N^2} \sum_{|k|<N} (N-|k|)\gamma_{\tau}[k](\gamma_{\tau}[k] - \gamma_{\tau}[k+h-h']) + O\left(\frac{1}{N}\right) + o(\lambda^4). \end{aligned} \quad (6.49)$$

Dans le régime asymptotique "données à haute fréquence" le premier terme de (6.49) se comporte comme

$$\frac{1}{N^2} \sum_{|k|<N} (N-|k|)(\gamma_{\tau}[k+h] - \gamma_{\tau}[k+h'])(\gamma_{\tau}[k-h] - \gamma_{\tau}[k-h']) \sim \lambda^4 K_{hh'}^{(1)} \frac{1}{N^3}, \quad (6.50)$$

où $K_{hh'}^{(1)}$ est une constante qui ne dépend que des décalages h et h' . Le deuxième terme de (6.49) se comporte comme

$$\frac{2}{N^2} \sum_{|k|<N} (N-|k|)\gamma_{\tau}[k](\gamma_{\tau}[k] - \gamma_{\tau}[k+h-h']) \sim \lambda^4 K_{hh'}^{(2)} \frac{\ln(N)}{N}. \quad (6.51)$$

où $K_{hh'}^{(2)}$ est une constante qui ne dépend que des décalages h et h' . En conséquence, le premier terme de l'équation (6.49) est d'un ordre plus petit que le deuxième, et nous pouvons donc le négliger.

Le comportement asymptotique de la variance $\mathbb{V}\text{ar}[\widehat{\gamma}_{\tau,N}^0[h, h']]$ est entièrement gouvernée par le deuxième terme de l'équation (6.49), ce qui nous conduit à (6.44). ■

6.4.4 Application au logarithme du processus MRW log-normal

Dans cette section nous supposons que les observations $\{Z_\tau[k]\}_{k=1,\dots,N}$ sont égales aux logarithmes des accroissements absolus du processus MRW log-normal

$$Z_\tau[k] = \ln(|\delta_\tau X(k\tau)|). \quad (6.52)$$

Théorème 6.9

Dans le régime asymptotique "données à haute fréquence", l'estimateur (6.29) est un estimateur asymptotiquement non biaisé dont la variance est majorée par

$$\mathbb{V}\text{ar}[\widehat{\lambda^2}] = \lambda^4 O\left(\frac{\ln(N)}{N}\right) + o(\lambda^4). \quad (6.53)$$

Preuve - 6.9 -

Sous condition de faible intermittence, $\lambda^2 \ll 1$, la suite $\{Z[k]\}_k$ peut être approximée par une somme d'une suite gaussienne et d'un bruit blanc non gaussien, grâce à l'approximation log-normale du processus MRW log-normal, donnée par l'équation (5.84).

En utilisant la proposition 6.7, il est simple de voir que l'espérance de l'estimateur est non biaisée au premier ordre en $\frac{1}{N}$ et λ^2

$$\mathbb{E}[\widehat{\lambda^2}] = \frac{\gamma_\tau[h] - \gamma_\tau[h']}{f[h] - f[h']} + O\left(\frac{1}{N}\right) = \lambda^2 + O\left(\frac{1}{N}\right) + o(\lambda^2). \quad (6.54)$$

Dans le régime asymptotique "données à haute fréquence" le terme $O(\frac{1}{N})$ devient nulle.

La variance de l'estimateur s'exprime par la somme de la variance (6.38), due à la mesure MRM log-normale, et la variance due au bruit qui n'est pas nulle que si deux ou plus indices parmi j, k, l et m coïncident. L'impact total dans la variance (6.38) de tels termes est de l'ordre de $\frac{1}{N}$. En conséquence, dans le régime "asymptotique" nous pouvons négliger le bruit blanc non gaussien et utiliser la démonstration du théorème 6.8. ■

6.4.5 Application aux données journalières

En effectuant une simple régression de la fonction d'autocovariance empirique (cf. section 6.1) ou en appliquant la méthode GMM (cf. section 6.2), nous nous sommes aperçus que, pour la plupart des séries financières, la valeur du temps intégral T peut être considérée comme étant de l'ordre de (voire supérieur à) l'échelle d'observation L . Ainsi, par exemple, la figure 2.6 représente le comportement de la fonction d'autocovariance empirique des rendements logarithmiques journaliers du taux de change GBP/USD. Dans ce cas, le temps intégral obtenu par la régression est égal à $\widehat{T} = 1000$ jours qui appartient à l'intervalle de confiance (à 95%) dans le cas où $T = L = 29$ ans (cf. section 6.2.3). Ce qui signifie que le régime asymptotique "historique infini" n'est pas atteint.

Il est simple de voir que les données journalières que nous disposons vérifient les conditions du régime asymptotique "haute fréquence" $\tau \ll T \simeq L$, alors le résultat théorique sur la convergence de l'estimateur du coefficient d'intermittence, donné par le théorème 6.9 est valide. En d'autres termes, les estimations du coefficient d'intermittence obtenues par la méthode GMM (cf. tableaux 6.1 et 6.2) sont correctes. Cela nous permet, d'un côté, de justifier *a posteriori* la validité de l'utilisation de l'approximation limite de petite intermittence pour l'étude des propriétés des estimateurs des paramètres du modèle MRW log-normal. De l'autre côté, ces résultats montrent qu'il est aussi possible d'utiliser cette approximation limite pour les prédictions de risques, que nous étudierons dans la partie III.

Troisième partie

Prédiction des risques

Prédiction de volatilité

La mesure du risque joue un rôle clé dans la recherche du prix des produits financiers. Une caractérisation précise et une prédiction des mouvements des prix des marchés financiers et des niveaux des taux d'intérêts permet de quantifier efficacement les risques de management. Le plus souvent, la notion de risque se résume à la variabilité des rendements. L'indicateur le plus courant est la volatilité.

De nombreuses modélisations de la volatilité des rendements ont été effectuées pour représenter sa dynamique et tenter de la prévoir. Parmi celles-ci, on trouve très fréquemment des modèles de type ARCH [BCK92] et, en particulier, le modèle le plus populaire GARCH introduit par Bollerslev en 1996 [Bol86] (cf. section 1.3.3).

Après la position du problème de prédiction de volatilité (cf. section 7.1), nous effectuons l'étude des différentes méthodes de prédiction de volatilité, qui sont toutes basées sur une prédiction linéaire. Nous utiliserons comme les modèles de référence les modèles GARCH(1,1) gaussien et Student dont les prédictions sont étudiées dans la section 7.2.

Dans le cadre du modèle MRW log-normal, nous proposons trois méthodes (cf. section 7.3), qui sont basées sur les résultats théoriques du chapitre 5 et notamment sur les approximations log-normales données par le théorème 5.17, pour les méthodes linéaire et quadratique, et le théorème 5.18 et la proposition 5.19, pour la méthode logarithmique.

Suivant les résultats du chapitre 6, nous allons choisir arbitrairement le temps intégral T égal à l'échelle d'observation L , qui est égale à la taille de la série pour les données journalières. Nous allons choisir pour le coefficient d'intermittence λ^2 une valeur typique 0.02 (cf. tableaux 6.1 et 6.2). Seule la variance σ^2 sera donc réellement estimée.

Afin de favoriser les prédictions obtenues dans le cadre des modèles GARCH(1,1), nous avons choisi d'effectuer les prédictions pour ces modèles en *in-sample* et celles pour le modèle MRW log-normal en *out-of-sample*. Comme nous le verrons dans la section 7.5, les méthodes de prédiction de volatilité, proposées dans le cadre du modèle MRW log-normal, sont systématiquement plus performantes que celles des modèles GARCH(1,1).

7.1 Position du problème de prédiction de volatilité

Néanmoins, il est important de souligner que la notion de volatilité dépend du modèle. Par exemple, dans le cadre du modèle GARCH, étudié dans la section 1.3.3, la volatilité conditionnelle au passé est déterministe (1.27) mais n'est pas directement observée, tandis que dans le cadre des modèles à volatilité stochastique, étudiés dans la section 1.3.4, ou dans le cadre du modèle MRW, étudié dans la section 3.4, la volatilité (non conditionnelle et conditionnelle au passé) est une variable aléatoire. En conséquence, le problème de prédiction de volatilité est *a priori* mal posé puisqu'il dépend du modèle.

Ainsi, afin de comparer les différents modèles nous devons tout d'abord poser un problème commun. La problématique naturelle consiste à prédire la volatilité réalisée, voir la section 1.2.6, sur la période considérée.

Pour simplifier le problème, nous supposons que les données financières sont espacées régulièrement dans le temps, ce qui nous permet d'introduire le processus des rendements logarithmiques d'un actif financier $\{\delta_1 X[t]; t = 1, 2, \dots\}$ à temps discret

$$\delta_1 X[k] = X(k\tau) - X((k-1)\tau), \text{ avec } k=1, 2, \dots, \quad (7.1)$$

où $X(t)$ est le logarithme du prix à l'instant t et τ est le pas d'échantillonnage. De même façon, nous notons la mesure MRM log-normale de l'intervalle $[t + h\tau - m\tau, t + h\tau]$

$$\delta_m M[t + h] = \delta_{m\tau} M(t + h\tau) \quad (7.2)$$

et l'accroissement de la magnitude renormalisée sur l'intervalle $[t + h\tau - m\tau, t + h\tau]$

$$\delta_m \Omega[t + h] = \delta_{m\tau} \Omega(t + h\tau). \quad (7.3)$$

En supposant l'absence de corrélation des rendements logarithmiques (cf. section 1.2.4), et la présence de structures de dépendance non linéaires (cf. section 1.2.5), nous considérons un modèle générique, dans lequel les rendements logarithmiques $\delta_1 X[t]$ s'écrivent comme le produit de deux variables aléatoires

$$\delta_1 X[t] = \sigma[t]\epsilon[t], \quad (7.4)$$

où le processus $\{\epsilon[t]\}_t$ est un bruit blanc indépendant du processus de volatilité $\{\sigma[t]\}_t$ et où la volatilité $\sigma[t]$ conditionnellement au passé peut être déterministe, comme dans le cas du modèle GARCH (cf. section 1.3.3), ou stochastique, comme dans le cadre du modèle MRW (cf. section 3.4).

L'indépendance entre le bruit blanc $\{\epsilon[t]\}_t$ et le processus de volatilité $\{\sigma[t]\}_t$ nous permet d'introduire la représentation du rendement logarithmique à l'échelle m et à l'horizon h

$$\delta_m X[t + h] = \sigma[t, m, h]\epsilon[t, m, h], \quad (7.5)$$

où l'innovation $\epsilon[t, m, h]$ est indépendante de la volatilité $\sigma[t, m, h]$.

Définition 7.1

Le problème de **prédiction de volatilité** consiste à déterminer (prévoir) à l'instant t la volatilité réalisée à l'échelle m et à l'horizon h

$$RV[t, m, h] = \sum_{j=1}^m (\delta_1 X[t + h - m + j])^2, \text{ pour } m \leq h, \quad (7.6)$$

sachant toute l'information disponible $\{\delta_1 X[t - j]\}_{j \geq 0}$ à l'instant t . On note $\widehat{\sigma}^2[t, m, h]$ le prédicteur correspondant.

7.2 Prédiction de volatilité dans le cadre du modèle GARCH

7.2.1 Prédiction

La prédiction de la volatilité correspondant à l'échelle 1 et à l'horizon 1 est donnée, grâce à la représentation (1.29) du modèle GARCH(1,1) comme le modèle ARCH(∞), par

$$\widehat{\sigma}^2[t, 1, 1] = \frac{\omega}{1 - \beta_1} + \alpha_1 \sum_{k=0}^{L_f} \beta_1^k (\delta_1 X[t - k])^2. \quad (7.7)$$

où L_f est une longueur du filtre linéaire.

En appliquant un raisonnement par récurrence, il est possible de déduire que la prédiction de volatilité correspondante à l'horizon h s'écrit comme [ABCD05]

$$\widehat{\sigma}^2[t, 1, h] = \sigma^2 + (\alpha_1 + \beta_1)^{h-1} (\widehat{\sigma}^2[t, 1, 1] - \sigma^2), \quad (7.8)$$

où σ^2 représente la variance non conditionnelle donnée par l'équation (1.30). De plus, grâce à la non corrélation d'accroissements, la variance conditionnelle correspondant à l'échelle m est égale à la somme des variances conditionnelles intermédiaires

$$\widehat{\sigma}^2[t, m, h] = m\sigma^2 + \sum_{k=1}^m (\alpha_1 + \beta_1)^{h-k} (\widehat{\sigma}^2[t, 1, 1] - \sigma^2). \quad (7.9)$$

7.2.2 Paramètres utilisés

Il est très important de noter que dans le cadre du modèle GARCH(1,1), nous effectuons la prédiction de volatilité en utilisant les paramètres obtenus par la méthode de quasi-maximum de vraisemblance, appliquée à toutes les données disponibles. En d'autres termes, la prédiction que nous effectuons correspond donc à une prédiction *in-sample*. Ce choix permet d'améliorer (par rapport à une prédiction *out-of-sample*) très notablement les prédictions et d'avoir ainsi un modèle de référence plus exigeant.

Les paramètres des modèles GARCH(1,1) gaussien et Student sont rapportés dans les tableaux 1.3 et 1.4.

7.3 Prédiction de volatilité dans le cadre du modèle MRW

Les résultats de cette section sont basés sur les propriétés théoriques du modèle MRW log-normale sous condition de faible intermittence $\lambda^2 \ll 1$, étudiées dans le chapitre 5.

Nous considérons l'approximation log-normale de la mesure MRM log-normale, (cf. équation (5.80)), ce qui permet de représenter la volatilité stochastique $\sigma[t, h, m]$ par une dynamique log-normale à l'aide de la magnitude renormalisée (cf. section 5.3)

$$\sigma^2[t, h, m] \stackrel{\lambda}{\simeq} m\tau e^{2\lambda \frac{\delta_m \Omega[t+h]}{m\tau} - 2\lambda^2 \text{Var} \left[\frac{\delta_m \Omega[t+h]}{m\tau} \right]}. \quad (7.10)$$

où $\Omega_m[t+h]$ est une variable gaussienne centrée dont la fonction d'autocovariance est donnée par les équations (5.56) et (5.57).

L'approximation log-normale du processus MRW, introduite par l'équation (5.84), permet de représenter les rendements logarithmiques $\delta_m X[t+h]$ de la façon suivante

$$\delta_m X[t+h] \stackrel{\lambda}{\simeq} \epsilon[t+h](m\tau)^{1/2} e^{\lambda \frac{\Omega_m[t+h]}{m\tau} - \lambda^2 \text{Var} \left[\frac{\delta_m \Omega[t+h]}{m\tau} \right]}, \quad (7.11)$$

où $\epsilon[t+h]$ est une variable gaussienne centrée de variance σ^2 et $\Omega_m[t+h]$ est définie comme ci-dessus.

Dans le cadre du modèle MRW log-normal nous présentons trois méthodes de prédiction de volatilité fondées sur la **régression linéaire** stationnaire¹, ce qui consiste à prédire une certaine fonction f de la volatilité $\sigma[t, m, h]$ par une combinaison linéaire des valeurs de la fonction f appliquée aux valeurs passées des rendements logarithmiques

$$\widehat{f(\sigma)}[t, m, h] = \mathbb{E}[f(\sigma[t, m, h])] + \sum_{k=0}^{L_f-1} \alpha_k \left(f(\delta_1 X[t-k]) - \mathbb{E}[f(\delta_1 X[t-k])] \right), \quad (7.12)$$

où L_f est la taille du filtre linéaire donné par les coefficients $\{\alpha_k\}_{k=0, \dots, L_f-1}$ qui peuvent être déterminés à l'aide de la fonction d'autocovariance du processus $\{f(\delta_1 X[t])\}_t$.

A priori dans le cas de processus à mémoire longue (cf. section 1.2.5) il est préférable d'utiliser un long historique pour réaliser une prédiction performante. Des expériences numériques montrent que, en pratique, pour le cas des processus MRW, l'utilisation de filtres de taille L_f égale à 100 jours est suffisante.

Chacune des deux premières méthodes de prédiction de volatilité, linéaire et quadratique, décrites dans la section suivante 7.3.1, fournit une valeur de la prédiction. La troisième méthode, décrite dans la section 7.3.2, fournit une prédiction de la loi conditionnelle de distribution de la volatilité $\sigma[t, m, h]$, ce que nous utiliserons, dans le chapitre 8, pour l'évaluation de la Valeur à Risque.

¹Dans l'appendice A nous faisons un rappel sur la régression linéaire.

7.3.1 Prédiction de volatilité linéaire et de volatilité quadratique

Prédiction de volatilité linéaire

La méthode de prédiction de volatilité linéaire $\widehat{\sigma}[t, m, h]$ consiste à utiliser la combinaison linéaire des rendements logarithmiques absolus $\{|\delta_1 X[t-k]|\}_{k=0, \dots, L_f-1}$, ce cas correspond au choix de la fonction $f(x) = |x|$ dans l'équation (7.12).

Pour déterminer les coefficients du filtre linéaire $\{\alpha_k\}_{k=0, \dots, L_f-1}$, on utilise :

– la variance de $|\delta_1 X[s]|$

$$\text{Var}\left[|\delta_1 X[s]|\right] \simeq \left(1 - \frac{2}{\pi} e^{-\lambda^2 \text{Var}\left[\frac{\delta_1 \Omega[s]}{\tau}\right]}\right) \sigma^2 \tau, \quad (7.13)$$

– la covariance entre $|\delta_1 X[s]|$ et $|\delta_1 X[s']|$, où $s \neq s'$

$$\text{Cov}\left[|\delta_1 X[s]|, |\delta_1 X[s']|\right] \simeq \frac{2}{\pi} e^{-\lambda^2 \text{Var}\left[\frac{\delta_1 \Omega[s]}{\tau}\right]} \left(e^{\lambda^2 \text{Cov}\left[\frac{\delta_1 \Omega[s]}{\tau}, \frac{\delta_1 \Omega[s']}{\tau}\right]} - 1\right) \sigma^2 \tau, \quad (7.14)$$

– la covariance entre $\sigma[t, m, h]$ et $|\delta_1 X[s]|$, où $s \leq t$

$$\begin{aligned} & \text{Cov}\left[\sigma[t, m, h], |\delta_1 X[s]|\right] \\ & \simeq \sqrt{\frac{2}{\pi}} e^{-\frac{\lambda^2}{2} \text{Var}\left[\frac{\delta_m \Omega[t+h]}{m\tau}\right] - \frac{\lambda^2}{2} \text{Var}\left[\frac{\delta_1 \Omega[s]}{\tau}\right]} \left(e^{\lambda^2 \text{Cov}\left[\frac{\delta_m \Omega[t+h]}{m\tau}, \frac{\delta_1 \Omega[s]}{\tau}\right]} - 1\right) \sigma m^{1/2} \tau. \end{aligned} \quad (7.15)$$

Prédiction de volatilité quadratique

La méthode de prédiction de volatilité quadratique $\widehat{\sigma^2}[t, m, h]$ consiste à utiliser la combinaison linéaire des rendements logarithmiques carrés $\{(\delta_1 X[t-k])^2\}_{k=0, \dots, L_f-1}$, ce cas correspond au choix de la fonction $f(x) = x^2$ dans l'équation (7.12).

Pour déterminer les coefficients du filtre linéaire $\{\alpha_k\}_{k=0, \dots, L_f-1}$, on utilise :

– la variance de $(\delta_1 X[s])^2$

$$\text{Var}\left[(\delta_1 X[s])^2\right] \simeq \left(3e^{4\lambda^2 \text{Var}\left[\frac{\delta_1 \Omega[s]}{\tau}\right]} - 1\right) \sigma^4 \tau^2, \quad (7.16)$$

– la covariance entre $(\delta_1 X[s])^2$ et $(\delta_1 X[s'])^2$, où $s \neq s'$

$$\text{Cov}\left[(\delta_1 X[s])^2, (\delta_1 X[s'])^2\right] \simeq \left(e^{4\lambda^2 \text{Cov}\left[\frac{\delta_1 \Omega[s]}{\tau}, \frac{\delta_1 \Omega[s']}{\tau}\right]} - 1\right) \sigma^4 \tau^2, \quad (7.17)$$

– la covariance entre $\sigma^2[t, m, h]$ et $(\delta_1 X[s])^2$, où $s \leq t$

$$\text{Cov}\left[\sigma^2[t, m, h], (\delta_1 X[s])^2\right] \simeq \left(e^{4\lambda^2 \text{Cov}\left[\frac{\delta_m \Omega[t+h]}{m\tau}, \frac{\delta_1 \Omega[s]}{\tau}\right]} - 1\right) \sigma^2 m \tau^2. \quad (7.18)$$

7.3.2 Prédiction de volatilité logarithmique

La dernière méthode de prédiction que nous étudions ici est celle de prédiction de volatilité logarithmique $\widehat{\ln(\sigma)}[t, m, h]$. Elle consiste à utiliser la combinaison linéaire des logarithmes des rendements logarithmiques absolus $\{\ln(|\delta_1 X[t-k]|)\}_{k=0, \dots, L_f-1}$, ce cas correspond au choix de la fonction $f(x) = \ln(|x|)$ dans l'équation (7.12).

Pour déterminer les coefficients du filtre linéaire $\{\alpha_k\}_{k=0, \dots, L_f-1}$, on utilise :

– la variance de $\ln(|\delta_1 X[s]|)$

$$\text{Var}\left[\ln(|\delta_1 X[s]|)\right] \simeq \frac{\lambda}{8} \frac{\pi^2}{8} + \lambda^2 \text{Var}\left[\frac{\delta_1 \Omega[s]}{\tau}\right], \quad (7.19)$$

– la covariance entre $\ln(|\delta_1 X[s]|)$ et $\ln(|\delta_1 X[s']|)$, où $s \neq s'$

$$\text{Cov}\left[\ln(|\delta_1 X[s]|), \ln(|\delta_1 X[s']|)\right] \simeq \lambda^2 \text{Cov}\left[\frac{\delta_1 \Omega[s]}{\tau}, \frac{\delta_1 \Omega[s']}{\tau}\right], \quad (7.20)$$

– la covariance entre $\ln(\sigma)[t, m, h]$ et $\ln(|\delta_1 X[s]|)$, où $s \leq t$

$$\begin{aligned} & \text{Cov}\left[\ln(\sigma)[t, m, h], \ln(|\delta_1 X[s]|)\right] \\ & \simeq \text{Cov}\left[\ln(m\tau) + \Omega_m[t+h], \ln(|\epsilon[s]|\tau^{1/2}) + \Omega_1[s]\right] = \text{Cov}[\Omega_m[t+h], \Omega_1[s]]. \end{aligned} \quad (7.21)$$

Puisque la variable $\Omega_m[t+h]$ est gaussienne sa loi de distribution est entièrement définie par sa moyenne, qui peut être approximée par la prédiction $\widehat{\sigma}[t, m, h]$, et par sa variance, qui peut être approximée par la variance des résidus de la régression linéaire

$$\begin{aligned} & \text{Var}\left[\ln(\sigma)[t, m, h] - \widehat{\ln(\sigma)}[t, m, h]\right] \\ & = \text{Var}\left[\Omega_m[t+h]\right] - \sum_{k=0}^{L_f-1} \alpha_k \text{Cov}\left[\Omega_m[t+h], \ln(|\epsilon[t-k]|\tau^{1/2}) + \Omega_1[t-k]\right] \end{aligned} \quad (7.22)$$

7.3.3 Paramètres utilisés

Notons que, en suivant les résultats du chapitre 6, nous effectuons la prédiction de volatilité en utilisant le temps intégral T fixé arbitrairement à l'échelle d'observation L (cf. section 6.3). Le coefficient d'intermittence λ^2 est fixé à 0.02 pour tous les cours, ce qui correspond à la valeur typique obtenue par estimation (cf. tableaux 6.1 et 6.2). Seule la variance σ^2 restent à estimer.

Les espérances $\mathbb{E}[f(\sigma[t, m, h])]$ et $\mathbb{E}[f(\delta_1 X[t-k])]$, qui interviennent dans la régression linéaire 7.12, dépendent de la variance σ^2 du processus MRW log-normal, qui est approximée à chaque instant à l'aide de l'estimation historique

$$\widehat{\sigma^2}[t] = \frac{1}{t\tau} \sum_{k=0}^{t-1} (\delta_1 X[k])^2. \quad (7.23)$$

La prédiction que nous effectuons correspond donc à une prédiction *out-of-sample*. Notons que les résultats de prédiction sont très robustes par rapport au choix de la valeur typique $\lambda^2 = 0.02$.

7.4 Evaluation des méthodes de prédiction de volatilité

Dans cette section nous allons décrire les critères d'évaluation de prédiction de volatilité au sens de la définition 7.1 qui fait intervenir la volatilité réalisée.

Il faut noter que pour les petites échelles de prédiction m la volatilité réalisée donne une approximation très bruitée de la volatilité théorique (GARCH ou MRW) qui est non observée. En conséquence, l'utilisation de la volatilité réalisée comme valeur de référence peut donner une vision erronée des performances de prédiction.

La majorité des évaluations de prédiction de volatilité qui sont proposées dans la littérature reposent sur un critère d'erreur quadratique moyenne mettant en jeu des rendements logarithmiques au carré, ou en valeur absolue, calculés sur l'horizon de prédiction.

7.4.1 Mesures MAE, MSE, R^2

Commençons par citer les critères classiques de pertinence de prévisions de volatilité :

– l'erreur en moyenne absolue MAE (de l'anglais *Mean Absolute Error*)

$$MAE[t, m, h] = \frac{1}{n} \sum_{t=1}^n \left| RV[t, m, h] - \widehat{\sigma}^2[t, m, h] \right|, \quad (7.24)$$

– l'erreur en moyenne quadratique MSE (de l'anglais *Mean Squared Error*)

$$MSE[t, m, h] = \frac{1}{n} \sum_{t=1}^n \left(RV[t, m, h] - \widehat{\sigma}^2[t, m, h] \right)^2, \quad (7.25)$$

– le coefficient de détermination $R^2[t, m, h]$

$$R^2[t, m, h] = 1 - \frac{MSE[t, m, h]}{TSS[t, m, h]}, \quad (7.26)$$

où $TSS[t, m, h]$ (de l'anglais *Total Sum of Squares*) est une variance empirique de la volatilité réalisée [CFA04]

$$TSS[t, m, h] = \frac{1}{n} \sum_{t=1}^n \left(RV[t, m, h] - \frac{1}{n} \sum_{s=1}^n RV[s, m, h] \right)^2. \quad (7.27)$$

Notons que l'erreur MSE est plus sensible aux événements extrêmes que l'erreur MAE.

7.4.2 Régression de Mincer-Zarnowitz

Une évaluation particulièrement populaire est obtenue par la régression de la série de volatilité réalisée sur la série de la volatilité prédite.

$$RV[t, m, h] = \gamma_0 + \gamma_1 \widehat{\sigma^2}[t, m, h] + \eta[t], \quad (7.28)$$

où $\{\eta[t]\}_t$ sont les résidus de la régression linéaire.

Cette équation de régression est l'analogie d'une procédure courante pour évaluer les prédictions d'une moyenne conditionnelle, appelée **régression de Mincer-Zarnowitz** à la suite de [MZ69].

L'évaluation typique d'efficacité des marchés consiste à vérifier que le coefficient γ_1 n'est pas significativement différent de 1 et la constante γ_0 de 0 pour démontrer l'absence de biais dans les anticipations. Puisque les erreurs de prévision doivent être nulles en espérance, c'est-à-dire orthogonales à n'importe quelle prévision formé rationnellement.

La méthode de la régression de Mincer-Zarnowitz est devenue couramment utilisée dans la littérature économétrique [PS90, WC95, ABM02]. Les intervalles de confiance des coefficients γ_0 et γ_1 sont étudiés dans [NW87, NW94, WM98]. Cependant, ces études ont été effectuées pour le régime asymptotique "historique infini" qui n'est pas atteint, en pratique (voir le chapitre 6 sur l'estimation des paramètres du modèle MRW log-normale). Par conséquent, nous n'utilisons cette régression pour évaluation de différentes méthodes de prédiction de volatilité qu'à titre indicatif.

7.5 Applications aux données journalières

Dans cette section nous représentons les résultats de prédiction de volatilité.

Les prédictions de volatilité pour des données des taux de change Forex ont été effectuées sur la période de 1990 à 2006. Les données de la période de 1977 à 1989 ont été utilisées pour initialiser les procédures de prédiction². Les prédictions de volatilité pour les données de l'indice CAC 40 ont été effectuées sur la période de 1992 à 2005. Les données de la période de 1990 à 1992 ont été utilisé pour initialiser les procédures de prédiction.

Le tableau 7.1 représente pour le taux de change GBP/USD les résultats d'évaluation des prédictions de volatilité aux différentes échelles et aux différents horizons obtenues en utilisant les modèles GARCH(1,1) gaussien et Student (cf. section 7.2) et les méthodes linéaire, quadratique et logarithmique (cf. section 7.3) dans le cadre du modèle MRW log-normal. Pour évaluer les différentes approches nous avons utilisé les mesures MAE (7.24) et MSE (7.25), le coefficient de détermination R^2 (7.26) et les coefficients γ_0 et γ_1 de la régression de Mincer-Zarnowitz (7.28).

²Ces périodes ont été utilisée dans [CF04].

Echelle	Horizon	nGARCH	tGARCH	MRW Lin	MRW Quad	MRW Log
MAE						
1 jour	1 jour	0.368	0.360	0.339	0.353	0.331
5 jours	5 jours	1.083	1.052	0.992	1.023	1.008
10 jours	10 jours	1.786	1.734	1.593	1.635	1.653
20 jours	20 jours	3.334	3.231	2.833	2.913	2.938
1 jour	6 jours	0.379	0.369	0.341	0.360	0.331
5 jours	10 jours	1.131	1.098	1.004	1.048	1.013
10 jours	20 jours	2.042	1.994	1.717	1.800	1.726
20 jours	40 jours	4.163	4.154	3.185	3.363	3.146
MSE, R^2						
1 jour	1 jour	0.416, 0.076	0.419, 0.071	0.417, 0.074	0.414, 0.081	0.428, 0.049
5 jours	5 jours	2.771, 0.212	2.854, 0.189	2.775, 0.211	2.715, 0.228	2.985, 0.151
10 jours	10 jours	6.944, 0.280	7.351, 0.238	6.853, 0.290	6.674, 0.308	7.577, 0.215
20 jours	20 jours	21.748, 0.266	24.350, 0.178	20.622, 0.304	20.011, 0.325	22.640, 0.236
1 jour	6 jours	0.426, 0.055	0.430, 0.045	0.425, 0.058	0.423, 0.061	0.432, 0.040
5 jours	10 jours	2.919, 0.171	3.039, 0.137	2.876, 0.183	2.842, 0.193	3.045, 0.135
10 jours	20 jours	8.397, 0.132	9.351, 0.033	7.894, 0.184	7.789, 0.194	8.269, 0.145
20 jours	40 jours	29.540, 0.007	36.367, -0.222	24.853, 0.165	24.989, 0.160	25.610, 0.139
γ_0, γ_1						
1 jour	1 jour	-0.054, 1.028	0.056, 0.761	0.032, 0.956	-0.002, 0.957	0.044, 1.013
5 jours	5 jours	-0.293, 1.026	0.336, 0.720	0.147, 0.935	-0.026, 0.958	0.223, 0.955
10 jours	10 jours	-0.798, 1.064	0.724, 0.694	0.254, 0.935	-0.092, 0.964	0.421, 0.941
20 jours	20 jours	-2.040, 1.090	1.827, 0.624	0.776, 0.881	0.051, 0.922	1.081, 0.881
1 jour	6 jours	-0.074, 1.045	0.079, 0.674	0.029, 0.967	-0.004, 0.948	0.043, 1.031
5 jours	10 jours	-0.518, 1.108	0.391, 0.669	0.122, 0.951	-0.053, 0.962	0.209, 0.972
10 jours	20 jours	-1.286, 1.126	1.117, 0.554	0.542, 0.844	0.174, 0.872	0.673, 0.865
20 jours	40 jours	-4.714, 1.318	2.868, 0.437	1.476, 0.771	0.845, 0.790	1.683, 0.791

TAB. 7.1 – Résultats d'évaluation des différentes prédictions de volatilité : *in-sample* GARCH(1,1) gaussien ("nGARCH"), Student ("tGARCH"), *out-of-sample* MRW log-normal linéaire ("MRW Lin"), quadratique ("MRW Quad") et logarithmique ("MRW Log"), du taux de change GBP/USD aux différentes échelles et aux différents horizons. Les paramètres des modèles sont précisés dans les sections 7.2.2 et 7.3.3. La table haute correspond à la mesure MAE, la table au milieu correspond à la mesure MSE et au coefficient R^2 et la table basse représente les coefficients γ_0 et γ_1 de la régression de Mincer-Zarnowitz. Lorsque les résultats des modèles MRW sont meilleurs que ceux des deux modèles GARCH, ceux-ci apparaissent en gras. On observe nettement que les modèles MRW sont plus performants que les modèles GARCH.

Rappelons que dans le cadre du modèle GARCH(1,1) nous avons effectué la prédiction *in-sample* dont les paramètres sont précisés dans la section 7.2.2, alors que dans le cadre du modèle MRW log-normale nous avons effectué la prédiction *out-of-sample* dont les paramètres sont précisés dans la section 7.3.3.

Notons que le modèle GARCH(1,1) gaussien fournit de meilleures prédictions que le modèle GARCH(1,1) Student pour la mesure MSE, alors que pour la mesure MAE la situation est inverse.

Il est simple de voir que les trois méthodes basées sur le modèle MRW log-normal sont plus performantes que les deux méthodes GARCH(1,1) pour la mesure MAE. Pour la mesure MSE, la meilleure méthode est la prédiction quadratique, ce qui n'est pas surprenant, puisque le calcul du filtre linéaire, dans ce cas, est basée sur la minimisation de la norme L^2 . Les deux autres méthodes donnent des résultats meilleurs que ceux des modèles GARCH aux grandes échelles et aux grands horizons.

On observe un comportement similaire pour toutes les séries de taux de change, comme le montre les tableaux tableau 7.2 et 7.3, et les séries des composants de l'indice CAC 40, comme le montre le tableau 7.4.

7.6 Conclusion

Dans ce chapitre nous avons étudié différentes méthodes de prédiction de volatilité dans le cadre du modèle MRW log-normal. Nous avons effectué la comparaison des prédictions *out-of-sample* obtenues par les méthodes linéaire, quadratique et logarithmique avec celles *in-sample* obtenues dans le cadre des modèles GARCH(1,1) gaussien et Student.

Les prédictions de volatilité données par le modèle MRW log-normal sont systématiquement plus performantes. Dans le prochain chapitre, afin d'affiner la comparaison, nous allons étudier le problème de prédiction de la Valeur à Risque.

	nGARCH	tGARCH	MRW Lin	MRW Quad	MRW Log
Echelle = 1 jour, Horizon = 1 jour					
CAD/USD	0.064	0.064	0.063	0.063	0.064
JPY/USD	1.299	1.320	1.296	1.288	1.327
CHF/USD	0.824	0.822	0.820	0.815	0.832
GBP/USD	0.416	0.419	0.417	0.414	0.428
Echelle = 5 jours, Horizon = 5 jours					
CAD/USD	0.380	0.370	0.356	0.352	0.380
JPY/USD	9.632	10.394	9.292	9.330	9.636
CHF/USD	5.114	5.081	4.931	4.876	5.136
GBP/USD	2.771	2.854	2.775	2.715	2.985
Echelle = 20 jours, Horizon = 20 jours					
CAD/USD	2.791	2.691	2.548	2.565	2.884
JPY/USD	59.638	77.832	54.955	56.065	59.363
CHF/USD	33.897	34.544	31.223	30.334	34.684
GBP/USD	21.748	24.350	20.622	20.011	22.640
Echelle = 10 jours, Horizon = 20 jours					
CAD/USD	1.028	1.006	0.995	0.996	1.093
JPY/USD	24.743	29.827	23.788	23.958	25.238
CHF/USD	13.161	13.477	12.390	12.240	13.150
GBP/USD	8.397	9.351	7.894	7.789	8.269
Echelle = 20 jours, Horizon = 40 jours					
CAD/USD	3.925	3.824	3.441	3.471	3.692
JPY/USD	70.788	109.035	67.282	66.691	71.121
CHF/USD	41.435	44.130	36.185	36.382	38.192
GBP/USD	29.540	36.367	24.853	24.989	25.610

TAB. 7.2 – Forex. Résultats d'évaluation en utilisant la mesure MSE des différentes prédictions de volatilité (voir la légende du tableau 7.1). Chaque table correspond aux différentes échelles et horizons. Les paramètres des modèles sont précisés dans les sections 7.2.2 et 7.3.3. Lorsque les résultats des modèles MRW sont meilleurs que ceux des deux modèles GARCH, ceux-ci apparaissent en gras. On observe que les prédictions *out-of-sample* linéaire ("MRW Lin") et quadratique ("MRW Quad") du modèle MRW log-normal sont plus performants que celles *in-sample* des modèles GARCH(1,1) gaussien ("nGARCH"), Student ("tGARCH"). La prédiction *out-of-sample* logarithmique ("MRW Log") est plus performant pour les grandes échelles et les grands horizons.

Echelle	Horizon	MRW Lin	MRW Quad	MRW Log
MAE				
1 jour	1 jour	4	4	4
5 jours	5 jours	4	4	4
10 jours	10 jours	4	4	4
20 jours	20 jours	4	4	4
10 jours	20 jours	4	4	4
20 jours	40 jours	4	4	4
MSE, R^2				
1 jour	1 jour	3	4	0
5 jours	5 jours	3	4	0
10 jours	10 jours	4	4	1
20 jours	20 jours	4	4	1
10 jours	20 jours	4	4	2
20 jours	40 jours	4	4	3

TAB. 7.3 – Forex. Nombre de cours, parmi les quatre taux de change Forex, pour lesquels les prédictions *out-of-sample* linéaire ("MRW Lin"), quadratique ("MRW Quad") et logarithmique ("MRW Log") du modèle MRW log-normal sont plus performants que celles *in-sample* des deux modèles GARCH(1,1) : gaussien ("nGARCH") et Student ("tStudent"). On observe que pour la mesure MAE tous les modèles MRW sont toujours meilleurs. Pour la mesure MSE les modèles MRW linéaire et quadratiques donnent toujours une meilleure prédiction que les modèles GARCH et le modèle MRW logarithmique est plus performant à grandes échelles et grands horizons.

Echelle	Horizon	MRW Lin	MRW Quad	MRW Log
MAE				
1 jour	1 jour	28	26	29
5 jours	5 jours	28	24	26
10 jours	10 jours	27	24	19
20 jours	20 jours	26	23	17
10 jours	20 jours	28	24	25
20 jours	40 jours	28	20	23
MSE, R^2				
1 jour	1 jour	13	22	2
5 jours	5 jours	17	19	2
10 jours	10 jours	16	19	4
20 jours	20 jours	21	19	8
10 jours	20 jours	22	22	8
20 jours	40 jours	17	16	8

TAB. 7.4 – CAC 40. Nombre de cours, parmi les vingt neuf composantes de l'indice CAC 40, pour lesquels les prédictions *out-of-sample* linéaire ("MRW Lin"), quadratique ("MRW Quad") et logarithmique ("MRW Log") du modèle MRW log-normal sont plus performants que celles *in-sample* des deux modèles GARCH(1,1) : gaussien ("nGARCH") et Student ("tStudent"). On observe que pour la mesure MAE tous les modèles MRW sont toujours meilleurs. Pour la mesure MSE les modèles MRW linéaire et quadratiques donnent toujours une meilleure prédiction que les modèles GARCH et le modèle MRW logarithmique est plus performant à grandes échelles et grands horizons.

Prédiction de Valeur à Risque

Les mesures du risque ont beaucoup évolué depuis que Markowitz a énoncé sa célèbre théorie de la diversification de portefeuille à la fin des années 1950, théorie qui devait révolutionner la gestion de portefeuille moderne. L'écart-type était alors la mesure du risque d'un portefeuille efficient. Mais pour un titre, cette mesure n'est pas appropriée. En effet, dans le cas d'un titre individuel, l'écart-type du rendement comprend les risques diversifiable et non diversifiable. Or, seul le risque non diversifiable, le risque présentée par la covariance entre le rendement du titre et les rendements des titres qui constituent un portefeuille hautement diversifié, est rémunéré par le marché. Une des difficultés fondamentales de la modélisation concerne l'estimation des facteurs du risque. Si l'on suppose que le nombre de titres qui compose le portefeuille est de N , le modèle de Markowitz exige l'estimation de N variances et $\frac{N(N-1)}{2}$ covariances. Quand N devient important, l'estimation de la matrice variance-covariance se présente comme un exercice très laborieux et les risques d'erreurs d'estimations sont loin d'être négligeables.

Durant les années 60, Sharp a proposé le modèle CAPM (*Capital Asset Pricing Model*) de l'évaluation es actifs financiers. Ce modèle est **monofactoriel** en ce sens qu'il ne distingue qu'un seul facteur explicatif du risque d'un titre, soit la corrélation entre le rendement de ce titre et celui du portefeuille du marché.

Au début des années 90, la VaR (de l'anglais *Value-at-Risk*), une nouvelle mesure du risque, a fait son entrée. Cette notion, originaire du secteur de l'assurance, a été importée sur les marchés financiers aux Etats-Unis par la banque *Bankers Trust* et popularisée par la banque *JP Morgan* en 1993 et son service (gratuit et public) *Riskmetrics*. La VaR sert à quantifier le risque de marché auquel sont soumis les portefeuilles bancaires. L'Accord de Bâle a imposé aux banques, en 1997, de détenir un montant de capital réglementaire pour palier les risques de marché. Or, ce capital est calculé à partir de la VaR. Cette mesure est ensuite devenue de plus en plus populaire pour évaluer le risque de portefeuilles institutionnels ou individuels. Elle permet entre autres d'évaluer les risques de type asymétrique, comme celui qui est associé aux options, l'écart-type et le ratio de Sharp ne permettent pas de prendre en compte ce risque de façon satisfaisante.

Par analogie avec la prédiction de volatilité (cf. chapitre 7), dans ce chapitre, nous utilisons les modèles GARCH(1,1) comme les modèles de référence. Nous étudions les méthodes de prédiction de VaR dans les sections 8.3, pour les modèles GARCH(1,1), et 8.4 pour le modèle MRW log-normal. Ces prédictions sont basées sur les prédicteurs de volatilité présentés dans la section précédente. Ainsi, pour le modèle GARCH(1,1), des résultats de prédictions *in-sample* sont présentés. Dans le cadre du modèle MRW, les prédictions sont *out-of-sample* et seul le paramètre σ^2 est estimé (T est fixé à L et $\lambda^2 = 0.02$ pour tous les cours). Après avoir effectué une brève revue des méthodes d'évaluation de prédiction de VaR (cf. section 8.5), nous présentons les résultats d'application de la prédiction de VaR aux données journalières (cf. section 8.6).

8.1 Définitions

La VaR correspond au montant de pertes maximum sur un horizon temporel donné et avec un niveau de confiance choisi. Elle dépend essentiellement de 4 paramètres :

- la distribution des résultats des portefeuilles,
- le niveau de confiance choisi, 95% ou 99% en général,
- l'horizon temporel choisi,
- l'échelle choisie.

Nous allons distinguer deux types de VaR : non conditionnelle et conditionnelle.

Définition 8.1

Soit $X(t)$ le logarithme du cours d'un actif financier. La **VaR non conditionnelle** du niveau p à l'échelle l est donnée par la relation

$$\mathbb{P}[\delta_l X(t+l) < -\text{VaR}_p(l)] = p, \quad \forall t. \quad (8.1)$$

Définition 8.2

Soit $X(t)$ le logarithme du cours d'un actif financier. La **VaR conditionnelle** à l'instant t du niveau p à l'échelle m et à l'horizon h est donnée par la relation

$$\mathbb{P}[\delta_m X(t+h) < -\text{VaR}_p(t, m, h) | X(s), s \leq t] = p. \quad (8.2)$$

8.2 Prédiction de VaR

8.2.1 Différentes méthodes de prédiction de VaR

Il existe essentiellement quatre méthodes pour prédire la VaR d'un actif : la méthode paramétrique, la méthode historique, la méthode de Monte Carlo et la méthode semi-paramétrique.

La méthode paramétrique repose souvent sur des hypothèses de normalité des rendements logarithmiques. Or, comme nous avons fait remarquer dans le chapitre 1, les distributions

des rentabilités historiques des variables de marché sont souvent très éloignées de la loi normale. On observe en particulier des distributions dissymétriques et présentant des queues épaisses. Il est évidemment possible de postuler une loi de distribution permettant de rendre compte de l'existence de queues de distribution épaisses (loi de Student, loi de Pareto, etc.), mais la nécessité d'imposer un choix *a priori* conduit à un fort risque de spécification appelé **risque de modèle**.

La méthode historique, au contraire, ne fixe pas *a priori* d'hypothèse sur la forme de la distribution des rendements logarithmiques et ne prend alors en compte que les données de l'échantillon pour prédire la VaR. Dans ce cas, la qualité des séries historiques est importante. En effet, il existe ainsi un risque non négligeable d'échantillon et il suffit de quelques données manquantes ou aberrantes pour rendre non pertinent le résultat obtenu. Comme pour la méthode paramétrique, cette méthode est rapide à mettre en œuvre.

La méthode de Monte Carlo reprend le principe de la VaR historique mais les données utilisées sont obtenues par simulation stochastique. Cette méthode est lourde à mettre en œuvre car elle nécessite de réaliser beaucoup de simulations pour obtenir une bonne précision. Elle est surtout utilisée dans le cadre de la prédiction de VaR sur des instruments complexes (options) proposant comme flux des fonctions non linéaires des actifs de base. Notons que cette méthode possède aussi le risque de modèle.

Il est possible de mettre en place une méthode de prédiction de la VaR basée sur une approche "intermédiaire" et qui réalise un bon compromis entre l'approche VaR historique (pas de risque de modèle) et l'approche paramétrique (limitation du risque d'échantillon), il s'agit d'une VaR semi-paramétrique, dans un environnement de queues de distribution épaisses. Un exemple de cette méthode de détermination de la perte potentielle maximale consiste d'abord à calculer une VaR en utilisant une formule de distribution normale puis en la corrigeant avec un développement dit de Cornish-Fisher [CF37] pour prendre en compte la skewness et la kurtosis (moments d'ordre 3 et d'ordre 4 de la distribution des rentabilités).

Les méthodes de prédiction que nous étudierons dans le cadre de ce travail sont toutes les méthodes semi-paramétrique.

8.2.2 Calcul implicite de la VaR

Pour simplifier le problème, comme dans le chapitre 7 sur la prédiction de volatilité, nous supposons que de données financières sont espacées régulièrement dans le temps, ce qui nous permet d'introduire le processus des rendements logarithmiques d'un actif financier $\{\delta_1 X[k]; k = 1, 2, \dots\}$ à temps discret

$$\delta_1 X[k] = X(k\tau) - X((k-1)\tau), \text{ avec } k=1,2,\dots, \quad (8.3)$$

où $X(t)$ est le logarithme du prix à l'instant t et τ est le pas d'échantillonnage.

L'indépendance entre le bruit blanc $\{\epsilon[t]\}_t$ et le processus de volatilité $\{\sigma[t]\}_t$ nous permet d'introduire la représentation du rendement logarithmique à l'échelle m et à l'horizon h

$$\delta_m X[t+h] = \sigma[t, m, h] \epsilon[t, m, h], \quad (8.4)$$

où l'innovation $\epsilon[t, m, h]$ est indépendante de la volatilité $\sigma[t, m, h]$.

Lemme 8.1

La VaR conditionnelle vérifie la relation suivante

$$\mathbb{E} \left[\Phi_\epsilon \left(\frac{-\text{VaR}_p[t, m, h]}{\sigma[t, m, h]} \right) \middle| \delta_1 X[s], s \leq t \right] = p, \quad (8.5)$$

où $\Phi_\epsilon(x)$ est la fonction de répartition de la variable aléatoire $\epsilon[t, m, h]$.

Preuve - 8.1 -

Il est simple de voir que la fonction de répartition du rendement logarithmique $\delta_m X[t+h]$ conditionnellement au passé est donnée par

$$\mathbb{P}[\delta_m X[t+h] < z | \delta_1 X[s], s \leq t] = \mathbb{P}[\sigma[t, m, h] \epsilon[t, m, h] < z | \delta_1 X[s], s \leq t]. \quad (8.6)$$

En utilisant la loi des espérances itérées et la définition de la VaR conditionnelle, on a

$$\begin{aligned} & \mathbb{P}[\delta_m X[t+h] < -\text{VaR}_p[t, m, h] | \delta_1 X[s], s \leq t] \\ &= \mathbb{E} \left[\mathbb{P}[\sigma[t, m, h] \epsilon[t, m, h] < -\text{VaR}_p[t, m, h] | \sigma[t, m, h]] \middle| \delta_1 X[s], s \leq t \right] \\ &= \mathbb{E} \left[\Phi_\epsilon \left(\frac{-\text{VaR}_p[t, m, h]}{\sigma[t, m, h]} \right) \middle| \delta_1 X[s], s \leq t \right] = p. \end{aligned} \quad (8.7)$$

■

Les méthodes de prédiction de VaR conditionnelle dans le cadre des modèles GARCH(1,1) et MRW log-normal sont basées sur l'inversion de la relation 8.5.

8.3 Prédiction de VaR dans le cadre du modèle GARCH

Dans le cadre du modèle GARCH(1,1), étudié dans la section 1.3.3, la volatilité $\sigma[t, m, h]$, sachant le passé, est déterministe. En appliquant, dans ce cas, le lemme 8.1, nous avons

$$\Phi_\epsilon \left(\frac{-\text{VaR}_p[t, m, h]}{\sigma[t, m, h]} \right) = p, \quad (8.8)$$

ce qui nous conduit à la relation suivante

$$\text{VaR}_p[t, m, h] = -\sigma[t, m, h] \Phi_\epsilon^{-1}(p), \quad (8.9)$$

où $\Phi_\epsilon^{-1}(p)$ est la quantile de la loi de probabilité de la variable $\epsilon[t, m, h]$.

La prédiction (7.9) de la volatilité $\sigma^2[t, m, h]$ nous amène à l'estimateur de la VaR conditionnelle suivante

$$\widehat{\text{VaR}}_p[t, m, h] = -\sqrt{\widehat{\sigma}^2[t, m, h]}\Phi_\epsilon^{-1}(p). \quad (8.10)$$

Pour la prédiction de VaR, dans le cadre du modèle GARCH(1,1), nous utilisons les mêmes paramètres que pour la prédiction de volatilité (cf. section 7.2.2). Il est très important de noter que cette prédiction correspond à une prédiction *in-sample*.

8.4 Prédiction de VaR dans le cadre du modèle MRW

Les résultats de cette section sont basés sur les propriétés théoriques du modèle MRW log-normale sous condition de faible intermittence $\lambda^2 \ll 1$ (cf. chapitre 5).

L'utilisation du lemme 8.1 pour le calcul de la VaR conditionnelle, dans le cadre du modèle MRW log-normal, nécessite la connaissance de la loi de la variable $\sigma[t, m, h]$ qui est égale à la loi conditionnelle de la mesure MRM log-normale conditionnellement aux accroissements passés du processus MRW log-normal.

Puisque la vraie loi conditionnelle n'a pas de forme analytique, nous utilisons, en pratique, son approximation log-normale obtenue par la méthode de la prédiction logarithmique de la volatilité (cf. section 7.3.2).

L'approximation de la VaR conditionnelle est donnée, dans le cas du modèle MRW log-normal, par l'inversion numérique de l'équation (8.5), où l'espérance dans la partie gauche peut être évaluée numériquement à l'aide de la méthode de quadrature de Gauss-Hermite [PTVF88].

Pour la prédiction de VaR, dans le cadre du modèle MRW log-normal, nous utilisons les mêmes paramètres que pour la prédiction de volatilité (cf. section 7.3.3). Notons que cette prédiction correspond à une prédiction *out-of-sample*.

8.5 Evaluation des différentes méthodes de prédiction de VaR

Le but de cette section est de présenter les différentes méthodes d'évaluation des modèles de prédiction de VaR, décrites ci-dessus dans les sections 8.3 et 8.4.

Il est bien connu qu'il existe beaucoup d'erreurs non systématiques de prédiction de VaR souvent liées au traitement de données financières, comme, par exemple, l'échantillonnage uniforme ou l'absence de certaines cotations. Ces erreurs non systématiques ont souvent un impact important sur les prédictors de VaR, qui deviennent biaisés.

Cependant, il est important de noter que la VaR conditionnelle théorique est inobservable. La seule évaluation envisageable consiste à vérifier que la suite de prédictions de VaR est consistante avec les rendements logarithmiques réalisés.

Nous présentons, par la suite, des tests statistiques de Kupiec (cf. section 8.5.1) et de Christoffersen (cf. section 8.5.2) basés sur l'étude de suite d'exception.

Définition 8.3

Nous définissons la **suite d'exceptions** $\{I_p[t]\}_t$ par la comparaison de rendements logarithmiques réalisés avec la prédiction de VaR conditionnelle

$$I_p[t] = \begin{cases} 1, & \text{si } \delta_m X[t+h] < -\text{VaR}_p[t, m, h], \\ 0, & \text{si } \delta_m X[t+h] \geq -\text{VaR}_p[t, m, h]. \end{cases} \quad (8.11)$$

8.5.1 Test statistique de Kupiec

Soit n le nombre d'exceptions dans un échantillon de taille N

$$n = \sum_{t=1}^N I_p[t], \quad (8.12)$$

le nombre d'exceptions voulant dire le nombre de fois que la perte réalisée excède la VaR estimée. Si le modèle de prédiction de VaR est correctement spécifié, n suit une distribution binomiale $B(N, p)$. L'hypothèse nulle du test de Kupiec [Kup95] (ou test statistique de Markov [Qua93a, Qua93b, BB03]) indique que les prévisions du modèle VaR sont justes. En d'autres termes, ce test est basé sur le fait que la probabilité de l'exception est égale à p

$$\mathbb{P}[I_p[t] = 1] = p. \quad (8.13)$$

Le test de Kupiec est le test du rapport de vraisemblance suivant

$$L_K = 2n \ln \left(\frac{n}{Np} \right) + 2(N-n) \ln \left(\frac{N-n}{N-Np} \right), \quad (8.14)$$

où p est le niveau de VaR. La statistique de Kupiec L_K , donnée par l'expression (8.14), est asymptotiquement distribuée selon la loi du $\chi^2(1)$. Cependant, pour des courts échantillons et des petite valeurs de probabilité p , le nombre d'exception n est petit (voire quelques unités) et la limite asymptotique n'est pas atteinte, et nous utilisons, en pratique, la vraie distribution de la statistique de Kupiec qui peut être déterminer à l'aide de la relation suivante

$$\mathbb{P} \left[\sum_{t=1}^N I_p[t] = n \right] = \frac{C_N^n}{p^N} = \frac{N!}{p^N n! (N-n)!}. \quad (8.15)$$

Ce test peut rejeter des modèles qui sous-estime ou sur-estime de façon importante la VaR, Kupiec a fait, cependant, remarquer dans [Kup95] que la puissance statistique de ce test est en général très faible.

8.5.2 Test statistique de Christoffersen

Le test statistique de Kupiec, étudié dans la section 8.5.1, ne permet pas de détecter si les observations sont indépendantes l'une de l'autre. Par exemple, la suite définie comme $I_p[t] = \mathbf{I}_{\{t \leq Np\}}$ est acceptée par le test de Kupiec.

Christoffersen dans l'article [Chr98] a proposé un test statistique plus avancé qui est capable de vérifier si la procédure de la prédiction de VaR est uniformément performante au cours du temps. Il permet donc de détecter si le modèle prend en compte le phénomène de l'accumulation des volatilités, étudiée dans la section 1.2.7.

Le test de Kupiec est basé seulement sur le fait (8.13) que la probabilité de l'exception est égale à p , alors que le test de Christoffersen est basé sur le fait que la probabilité de l'exception conditionnellement à la valeur précédente de la suite d'exceptions est égale à p

$$\mathbb{P}[I_p[t] = 1 | I_p[t-1] = 0] = \mathbb{P}[I_p[t] = 1 | I_p[t-1] = 1] = p. \quad (8.16)$$

Le test de Christoffersen est un test du rapport de vraisemblance dont la statistique est donnée par l'expression suivante

$$L_C = 2(n_{00} \ln(\hat{p}_{00}) + n_{01} \ln(\hat{p}_{01}) + n_{10} \ln(\hat{p}_{10}) + n_{11} \ln(\hat{p}_{11}) - n \ln(p) - (N - n) \ln(1 - p)), \quad (8.17)$$

où n est le nombre d'exceptions, donné par l'équation (8.12), n_{jk} est le nombre d'observations de valeur j suivies par la valeur k et les probabilités empiriques \hat{p}_{jk} sont définies comme

$$\hat{p}_{jk} = \frac{n_{jk}}{N}, \quad (8.18)$$

où j et k peuvent prendre les valeurs 0 et 1.

La statistique de Christoffersen L_C , définie par l'expression (8.17), est asymptotiquement distribuée selon la loi du $\chi^2(2)$. La puissance statistique du test de Christoffersen pour les courtes suites d'exceptions a été étudié dans [CP04].

8.5.3 Quantiles empiriques

Une façon d'évaluer visuellement les différents modèles de prédiction de VaR est de représenter les moyennes empiriques \hat{p} des suites d'exceptions correspondant aux différents niveaux de probabilité p versus les niveaux de probabilité p . Si le modèle de prédiction de VaR est correctement spécifié, le tracé doit coïncider approximativement avec la droite $y = x$. L'avantage de cette méthode est qu'elle permet d'obtenir visuellement des plages de valeurs de niveau de probabilité p pour lesquels le modèle considéré fournit des résultats satisfaisants.

Pour une meilleure lisibilité, nous représentons dans les figures 8.1 et 8.2 les excès de moyennes empiriques $\hat{p} - p$ versus les niveaux de probabilité p .

8.6 Applications aux données journalières

Dans cette section nous représentons les résultats de prédiction de VaR conditionnelle.

Comme dans le cas de prédiction de volatilité, les prédictions de VaR conditionnelle pour les données des taux de change Forex ont été effectuées sur la période de 1990 à 2006, les données de la période de 1977 à 1989 ont été utilisées pour initialiser les procédures de prédiction. Les prédictions de volatilité pour les données de l'indice CAC 40 ont été effectuées sur la période de 1992 à 2005, les données de la période de 1990 à 1992 ont été utilisées pour initialiser les procédures de prédiction.

Tout d'abord, nous effectuons la comparaison des résultats de prédiction de VaR conditionnelle du taux de change GBP/USD obtenus dans le cadre des différents modèles. Le tableau 8.1 représente les résultats d'évaluation des prédictions de volatilité à différentes échelles, horizons et niveaux de probabilité obtenues en utilisant les modèles GARCH(1,1) gaussien et Student (cf. section 7.2) et le modèle MRW log-normal. Pour évaluer les différentes approches nous avons utilisé le test statistique de Kupiec (cf. section 8.5.1).

Il est simple de voir que la prédiction de VaR conditionnelle *out-of-sample*, obtenue par le modèle MRW, est plus performante que celles *in-sample*, obtenues par les deux modèles GARCH(1,1) gaussien et Student. Notons, de plus, que les prédictions de VaR conditionnelle obtenues à l'aide du modèle GARCH(1,1) Student ne sont jamais acceptées par le test statistique de Kupiec.

Comme le montre les tableaux récapitulatifs 8.2 et 8.3, on observe un comportement similaire pour toutes les séries de taux de change et pour toutes les composantes de l'indice CAC 40.

Dans les figures 8.1 nous avons représenté les quantiles empiriques (cf. section 8.5.3) de la prédiction de VaR du taux de change GBP/USD à l'échelle 1 jour et à l'horizon 1 jour (à gauche) et à l'échelle 5 jours et à l'horizon 5 jours (à droite). Il est simple de voir que les prédictions de VaR conditionnelle obtenues à l'aide du modèle MRW log-normal sont plus performantes que celles obtenues par les deux modèles GARCH : gaussien et Student.

On observe un comportement similaire pour les quantiles empiriques moyennes des prédictions de VaR des composantes de l'indice CAC 40.

8.7 Conclusion

De nombreux travaux s'intéressent au problème de prédiction de risques et notamment de Valeur à Risque. De nombreux modèles existent déjà et sont couramment utilisés. Cependant des travaux numériques sur de très grandes quantités de données (réalisés par plusieurs équipes internationales) ont montré qu'aucune de ces modèles ne surpasse les autres de façon systématique.

	nGARCH	tGARCH	MRW
Echelle = 1 jour, Horizon = 1 jour			
0.5%	0.006 (0.321)	0.001 (0.000)	0.006 (0.231)
1%	0.010 (0.844)	0.002 (0.000)	0.013 (0.089)
5%	0.039 (0.000)	0.024 (0.000)	0.055 (0.151)
10%	0.077 (0.000)	0.061 (0.000)	0.106 (0.222)
20%	0.161 (0.000)	0.156 (0.000)	0.203 (0.606)
Echelle = 1 jour, Horizon = 6 jours			
0.5%	0.005 (0.762)	0.000 (0.000)	0.005 (0.939)
1%	0.008 (0.210)	0.002 (0.000)	0.010 (0.837)
5%	0.038 (0.000)	0.023 (0.000)	0.053 (0.452)
10%	0.072 (0.000)	0.058 (0.000)	0.103 (0.563)
20%	0.153 (0.000)	0.151 (0.000)	0.204 (0.527)
Echelle = 1 jour, Horizon = 11 jours			
0.5%	0.004 (0.447)	0.000 (0.000)	0.004 (0.599)
1%	0.007 (0.053)	0.002 (0.000)	0.010 (0.795)
5%	0.034 (0.000)	0.021 (0.000)	0.052 (0.529)
10%	0.071 (0.000)	0.056 (0.000)	0.103 (0.479)
20%	0.151 (0.000)	0.145 (0.000)	0.202 (0.719)
Echelle = 5 jours, Horizon = 5 jours			
0.5%	0.005 (0.972)	0.000 (0.004)	0.005 (0.972)
1%	0.010 (0.961)	0.001 (0.001)	0.012 (0.527)
5%	0.039 (0.147)	0.027 (0.001)	0.060 (0.195)
10%	0.079 (0.035)	0.066 (0.001)	0.117 (0.150)
20%	0.182 (0.161)	0.172 (0.033)	0.212 (0.423)
Echelle = 5 jours, Horizon = 10 jours			
0.5%	0.004 (0.579)	0.000 (0.004)	0.007 (0.369)
1%	0.007 (0.431)	0.000 (0.000)	0.011 (0.763)
5%	0.034 (0.031)	0.022 (0.000)	0.053 (0.708)
10%	0.074 (0.009)	0.063 (0.000)	0.114 (0.180)
20%	0.173 (0.044)	0.165 (0.008)	0.226 (0.078)

TAB. 8.1 – Résultats des prédictions de VaR conditionnelle à différentes échelles, horizons et niveaux de probabilité pour le taux de change GBP/USD en utilisant les modèles : GARCH(1,1) gaussien ("nGARCH"), GARCH(1,1) Student ("tGARCH") et MRW log-normal ("MRW"). Pour chaque prédiction on précise la probabilité empirique et la p -valeur du test statistique de Kupiec. Les résultats apparaissent en gras si le test de Kupiec est accepté pour le niveau de confiance de 95%. On observe que le modèle MRW est nettement plus performant que les deux modèles GARCH.

	nGARCH	tGARCH	MRW
Echelle = 1 jour, Horizon = 1 jour			
0.5%	1, 1	0, 0	3, 2
1%	1, 1	0, 0	3, 2
5%	3, 3	0, 0	2, 3
10%	2, 2	0, 0	4, 4
20%	1, 1	0, 0	4, 4
Echelle = 1 jour, Horizon = 6 jours			
0.5%	1, 1	0, 0	3, 2
1%	2, 1	0, 0	3, 2
5%	3, 2	0, 0	4, 2
10%	1, 0	0, 0	3, 3
20%	1, 0	0, 0	3, 3
Echelle = 1 jour, Horizon = 11 jours			
0.5%	1, 1	0, 0	3, 3
1%	2, 1	0, 0	3, 2
5%	3, 1	0, 0	4, 3
10%	1, 1	0, 0	3, 2
20%	1, 0	0, 0	3, 3
Echelle = 5 jours, Horizon = 5 jours			
0.5%	3, 3	3, 3	3, 4
1%	3, 3	2, 2	3, 4
5%	4, 4	0, 0	3, 3
10%	3, 4	1, 1	3, 3
20%	4, 4	1, 2	4, 4
Echelle = 5 jours, Horizon = 10 jours			
0.5%	3, 3	3, 3	3, 3
1%	3, 4	0, 1	3, 3
5%	3, 4	1, 1	3, 4
10%	2, 2	1, 1	4, 4
20%	3, 4	1, 1	4, 4

TAB. 8.2 – Forex. Nombre de cours, parmi les quatre taux de change Forex, pour lesquels les tests statistiques : de Kupiec (chiffre à gauche) et de Christoffersen (chiffre à droite) sont acceptés pour le niveau de confiance de 95%. Le chiffre apparaît en gras si le modèle MRW ("MRW") log-normal donne un meilleur résultat que les deux modèles GARCH(1,1) : gaussien ("nGARCH") et Student ("tGARCH").

	nGARCH	tGARCH	MRW
Echelle = 1 jour, Horizon = 1 jour			
0.5%	11, 15	2, 4	24, 23
1%	21, 21	0, 2	21, 20
5%	16, 13	0, 0	22, 13
10%	3, 5	0, 0	24, 20
20%	2, 2	0, 0	25, 24
Echelle = 1 jour, Horizon = 6 jours			
0.5%	7, 10	2, 7	26, 20
1%	22, 15	1, 1	22, 16
5%	11, 7	0, 0	20, 10
10%	2, 3	0, 0	21, 12
20%	0, 1	0, 0	25, 18
Echelle = 1 jour, Horizon = 11 jours			
0.5%	9, 10	5, 9	26, 19
1%	22, 15	1, 1	24, 14
5%	11, 6	0, 0	20, 8
10%	2, 3	0, 0	22, 9
20%	0, 0	0, 0	26, 19
Echelle = 5 jours, Horizon = 5 jours			
0.5%	22, 21	22, 22	23, 21
1%	23, 24	14, 19	26, 24
5%	23, 24	4, 4	21, 24
10%	19, 18	4, 4	23, 24
20%	14, 17	2, 7	21, 22
Echelle = 5 jours, Horizon = 10 jours			
0.5%	25, 20	23, 18	26, 18
1%	26, 25	14, 17	24, 23
5%	22, 19	3, 4	23, 24
10%	15, 18	1, 4	24, 24
20%	10, 14	0, 4	21, 22

TAB. 8.3 – CAC 40. Nombre de cours, parmi les vingt neuf composantes de l'indice CAC 40, pour lesquels les tests statistiques : de Kupiec (chiffre à gauche) et de Christoffersen (chiffre à droite) sont acceptés pour le niveau de confiance de 95%. Le chiffre apparaît en gras si le modèle MRW ("MRW") log-normal donne un meilleur résultat que les deux modèles GARCH(1,1) : gaussien ("nGARCH") et Student ("tGARCH").

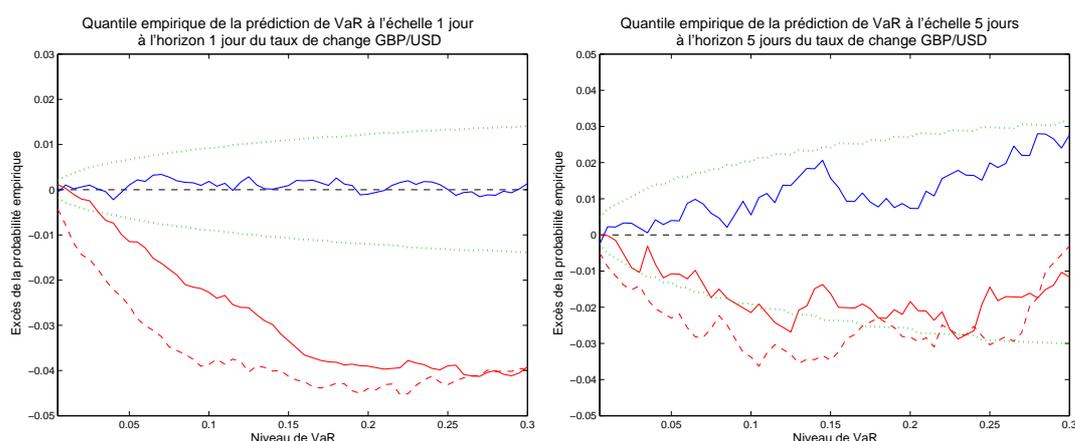


FIG. 8.1 – Quantile empirique de la prédiction de VaR conditionnelle du taux de change GBP/USD à l'échelle 1 jour et à l'horizon 1 jour (à gauche) et à l'échelle 5 jours et à l'horizon 5 jours (à droite). Les traits pointillés représentent l'intervalle de confiance à 95% du test statistique de Kupiec. Le trait plein bleu correspond au modèle MRW log-normal et les traits rouges correspondent aux modèles GARCH(1,1) gaussien (plein) et Student (tireté). Les paramètres des modèles GARCH et MRW sont précisés dans les sections 8.3 et 8.4. On observe que le modèle MRW log-normal est nettement plus performant que les deux modèles GARCH.

Les résultats des tests extensifs, qui ont été effectués sur les données journalières de Forex et du CAC 40, montrent que le modèle MRW log-normal semble surpasser assez nettement toutes les autres méthodes. Ceci est rendu possible, car le processus MRW log-normal permet une modélisation des cours multi-échelles. Contrairement à la plupart des autres modèles, le modèle MRW log-normal ne se cale pas sur une unique échelle de temps.

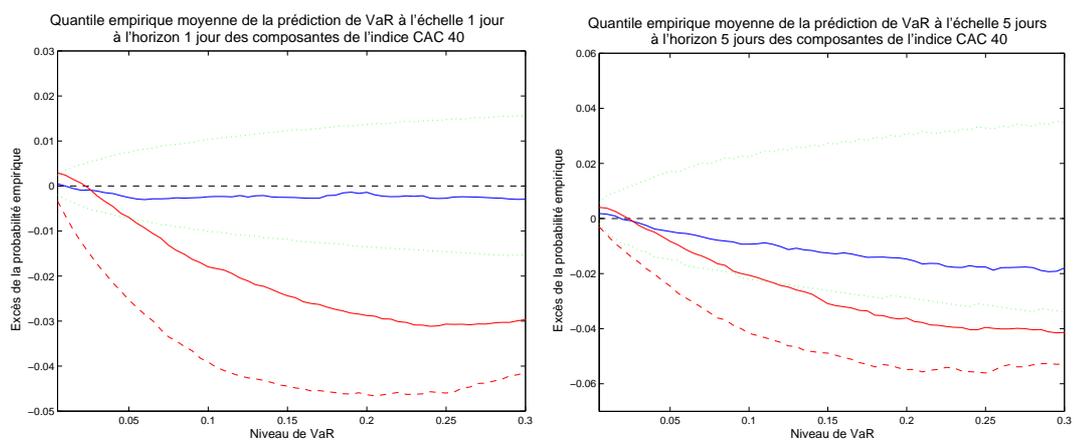


FIG. 8.2 – Quantile empirique moyenne des prédictions de VaR conditionnelle des composantes de l'indice CAC 40 à l'échelle 1 jour et à l'horizon 1 jour (à gauche) et à l'échelle 5 jours et à l'horizon 5 jours (à droite). Les traits pointillés représentent l'intervalle de confiance à 95% du test statistique de Kupiec. Le trait plein bleu correspond au modèle MRW log-normal et les traits rouges correspondent aux modèles GARCH(1,1) gaussien (plein) et Student (tireté). Les paramètres des modèles GARCH et MRW sont précisés dans les section 8.3 et 8.4. On observe que le modèle MRW log-normal est nettement plus performant que les deux modèles GARCH.

Quatrième partie

Vers une modélisation de séries financières à haute fréquence

Données à haute fréquence

Depuis le début des années 1990, on s'intéresse aux données intra-journalières, ce qui a permis de mettre en évidence des nouveaux comportements qui n'apparaissaient pas dans l'analyse des données journalières.

Le marché de changes produit des données intra-journalières ayant joué un rôle central en finance à haute fréquence. Contrairement aux autres marchés, ces données peuvent être disponibles sur de longues périodes et à de hautes fréquences, pour certaines devises on enregistre toutes les transactions. Ce marché est également très liquide.

Le marché des devises est caractérisé par une répartition géographique des différents marchés et différents types d'agents, avec des profils de risques différents et des contraintes institutionnelles différentes. Avant l'étude des données à haute fréquence, ces caractéristiques structurelles n'étaient pas apparentes dans les données, et peu considérées en modélisation.

On peut diviser ce marché en essentiellement deux parties : le marché au comptant (*spot market*) et le marché à terme (*forward market*)¹. On ne parlera que du marché à terme.

De façon générale, le traitement des données à haute fréquence nécessite la mise en oeuvre de toute une série de méthodes (filtrage, interpolation, désaisonnalisation, ...). Le récent livre [DGM⁺01] donne une bonne idée des méthodes couramment utilisées .

Ce chapitre de prospective montre que l'on peut construire un modèle basé sur le modèle MRW log-normal qui permet de reproduire les structures de dépendance saisonnière des données financières intra-journalières.

Nous illustrons les qualités de ce nouveau modèle sur les données intra-journalières du contrat à terme sur le taux de change GBP/USD échantillonnées à l'échelle 10 minutes. Notons que l'on peut montrer qu'à cette échelle, les effets de microstructure (cf. section 1.2.4) sont négligeables.

¹Il existe encore le marché des produits dérivés sur devises, mais c'est une plus petite partie du marché (en développement cependant)

Après avoir étudié les effets saisonniers (cf. section 9.2) nous introduisons deux modèles d'évolution des cours boursiers intra-journaliers (cf. section 9.3) : multiplicatif et subordonné. Ces deux modèles, construits à partir du modèle MRW log-normal, permettent de prendre en compte les effets saisonniers. Dans le premier modèle (cf. section 9.3.1), la saisonnalité journalière est traitée de façon "classique" : un facteur multiplicatif porte l'intégralité de la structure saisonnière journalière [TX95, CT96, DGM⁺01, BG01]. Dans le second (cf. section 9.3.2) la saisonnalité est modélisée en utilisant une subordination du processus MRW log-normal. Le temps physique (cf. section 2.4.2) porte alors l'intégralité de la structure saisonnière journalière.

Nous abordons le problème du choix du modèle qui permet de mieux reproduire les propriétés statistiques des séries à haute fréquence (cf. section 9.3.3). Dans la section 9.4, nous étudions brièvement l'estimation du coefficient d'intermittence λ^2 , dans le cadre des deux modèles introduits. A la fin du chapitre, nous présentons les résultats des prédictions de risque obtenus à l'aide de ces deux modèles.

Les résultats de ce chapitre montrent la supériorité sur bien des rapports du modèle subordonné par rapport au modèle multiplicatif.

9.1 Présentation des séries intra-journalières

Nous rappelons que nous disposons de séries de prix de contrats à terme sur les taux de change. Ces séries sont uniformément échantillonnées à l'échelle 10 minutes à l'aide d'une interpolation linéaire. De plus, elles sont ajustées pour tenir compte de la procédure de *roll-over* (cf. section 1.1.2).

Dans la figure 9.1 nous avons représenté les cours du contrat à terme sur le taux de change GBP/USD (données intra-journalières) et le taux de change lui-même (données journalières) pour la même période. Il est simple de voir que les dynamiques des deux séries se ressemblent beaucoup.

9.2 Effets saisonniers

De nombreuses études réalisées à partir de données de taux de change journalières et intra-journalières ont révélé la présence d'**effets saisonniers** dans la volatilité [MD⁺90, DMNP93, ER98, DGM⁺01, BG01]. Ainsi, par exemple, les rendements logarithmiques absolus moyens sont très élevés à l'ouverture des marchés où l'activité est grande. Au Japon, la période d'activité la plus faible a lieu autour du repas (vers midi), c'est à ce moment que l'on retrouve la plus faible valeur des rendements logarithmiques moyens absolus.

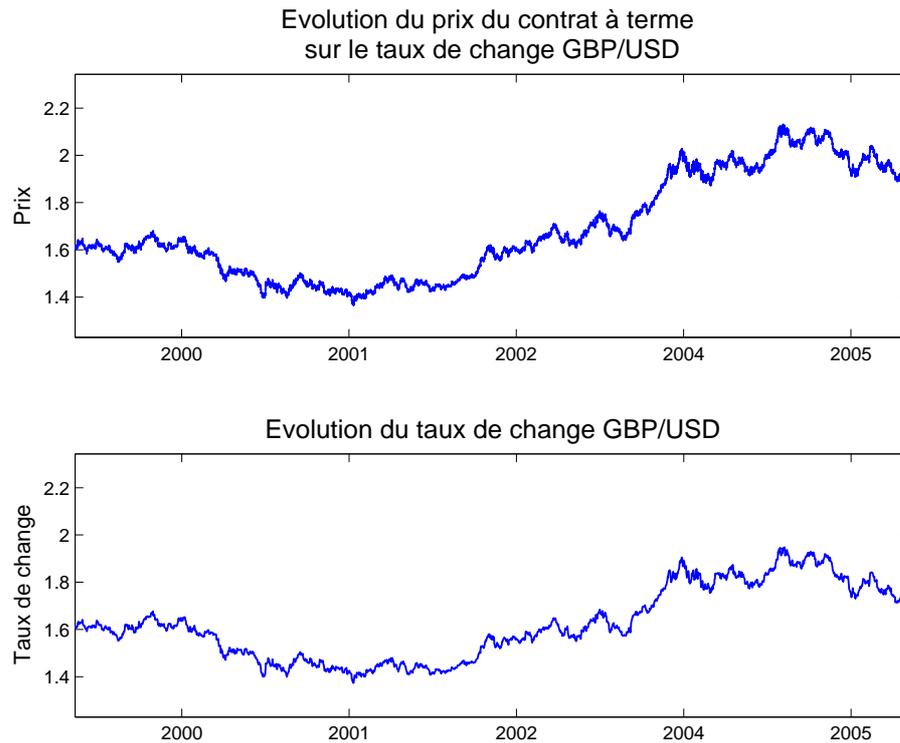


FIG. 9.1 – Prix intra-journaliers du contrat à terme sur le taux de change GBP/USD (en haut), valeurs de clôture du taux de change GBP/USD (en bas) pour la même période.

9.2.1 Mise en évidence des effets saisonniers

Dans la figure 9.2 nous avons représenté la fonction empirique d'autocorrélation des rendements logarithmiques intra-journaliers du contrat à terme sur le taux de change GBP/USD. Elle met en évidence un effet saisonnier avec un cycle de 24 heures.

Une autre méthode classique de détection des cycles saisonniers consiste à étudier l'amplitude de la transformée de Fourier de la série d'observations. Si les données présentent un effet saisonnier avec une période T la transformée de Fourier doit posséder les amplitudes plus élevées pour les fréquences multiple de $\frac{2\pi}{T}$.

Dans la figure 9.3 nous avons représenté l'amplitude de la transformée de Fourier de la série des rendements logarithmiques intra-journaliers du contrat à terme sur le taux de change GBP/USD². L'effet saisonnier avec un cycle de 24h précédemment observé est de nouveau visible.

Notons que l'on n'observe pas d'effet saisonnier correspondant à la période de 7 jours, en

²Pour ne pas introduire le changement de phase, en effectuant la transformation de Fourier, nous avons copié les rendements logarithmiques du jour précédent ouvert si le marché a été fermé (hors week-ends).

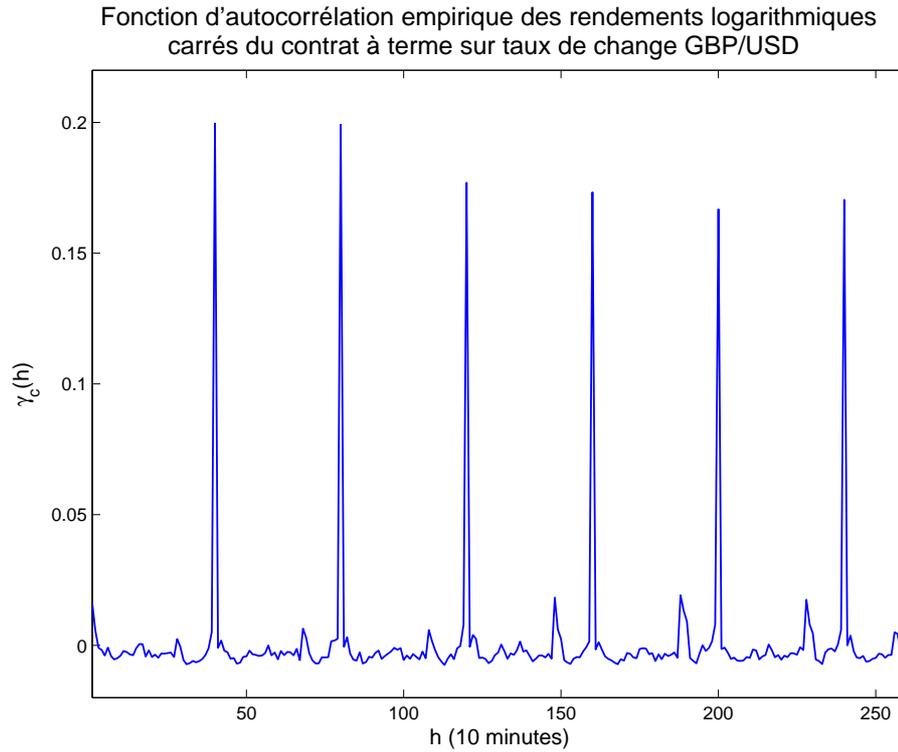


FIG. 9.2 – Fonction d'autocorrélation empirique des rendements logarithmiques intra-journaliers (à l'échelle 10 minutes) carrés du contrat à terme sur le taux de change GBP/USD. On observe de fortes corrélations pour les décalages proportionnel à 40, ce qui correspond à un jour ouvert (40×10 minutes = 6 heures 30 minutes).

d'autres termes c'est l'activité journalière qui est responsable de l'effet saisonnier observé.

9.2.2 Etude de l'activité intra-journalière

Afin d'étudier plus précisément les saisonnalités journalières, mises en évidence dans la section précédente, nous allons calculer les volatilités réalisées en fonction de l'heure dans la journée. On considère N_p périodes $\{\Delta_j\}_{j=1,\dots,N_p}$ dans la journée, par exemple, pour les données des contrats à terme sur les taux de change, on peut considérer $N_p = 39$ périodes à l'échelle 10 minutes de 7h20 à 13h50, et on calcule les variances empiriques $\{\widehat{\sigma_j^2}\}_{j=1,\dots,N_p}$ des rendements logarithmiques sur ces périodes.

Dans la figure 9.4 nous avons représenté les volatilités réalisées moyennes $\{\widehat{\sigma_j^2}\}$ en fonction de j pour les contrats à terme sur les taux de change GBP/USD et JPY/USD.

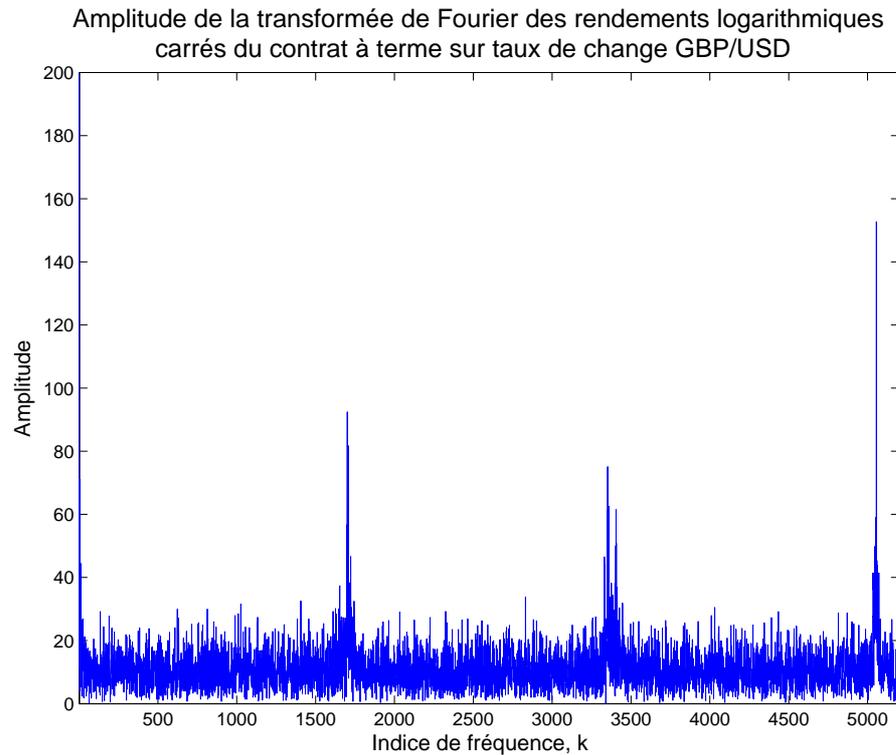


FIG. 9.3 – Amplitude de la transformée de Fourier de la série des rendements logarithmiques carrés intra-journaliers (à l'échelle 10 minutes) du contrat à terme sur le taux de change GBP/USD. On observe de fortes amplitudes pour les fréquences multiples de celle qui correspond à la période 40, ce qui correspond à un jour ouvert (40×10 minutes = 6 heures 30 minutes).

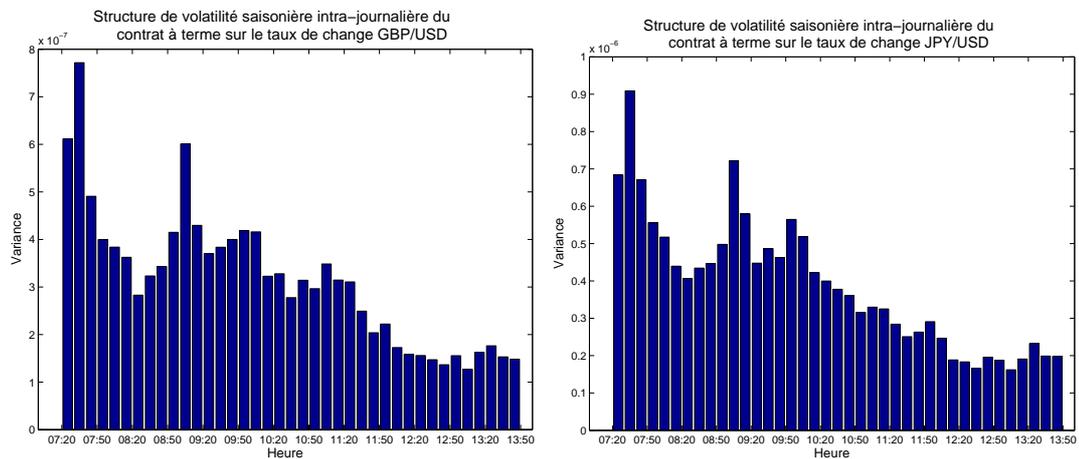


FIG. 9.4 – Structure de volatilité saisonnière intra-journalière des contrat à terme sur les taux de change GBP/USD et JPY/USD. On peut observer la haute activité à l'ouverture du marché, ainsi que la baisse de l'activité à midi.

	$\widehat{\sigma}_{\text{n}}^2/\widehat{\sigma}^2$	$\widehat{\sigma}_{\text{w}}^2/\widehat{\sigma}^2$	$\widehat{\sigma}_{\text{jf}}^2/\widehat{\sigma}^2$	$\widehat{\sigma}_{\text{lw}}^2/\widehat{\sigma}^2$
GBP/USD	0.517	0.520	0.629	1.495
JPY/USD	0.590	0.770	1.055	1.186
CHF/USD	0.448	0.542	0.654	1.272

TAB. 9.1 – "Durées empiriques" (en jours ouvrés) des différentes périodes : nuits ("n"), week-ends ("w"), jours fériés ("jf") et longs week-ends ("lw") (voir la note en bas de page 232).

9.2.3 "Durées empiriques" des nuits, des jours fériés et des week-ends

On observe des différences entre les prix d'ouverture et ceux de clôture, il est donc naturel de supposer que les cours continuent à évoluer, quand le marché est physiquement fermé, mais avec une "activité" plus faible que dans la journée.

Nous allons distinguer quatre type de périodes lorsque le marché est fermé : les nuits Δ_{n} , les week-ends Δ_{w} , les jours fériés Δ_{jf} et les longs week-ends Δ_{lw} ³. On note $\widehat{\sigma}_{\text{n}}^2$, $\widehat{\sigma}_{\text{w}}^2$, $\widehat{\sigma}_{\text{jf}}^2$ et $\widehat{\sigma}_{\text{lw}}^2$ les variances empiriques des rendements logarithmiques sur ces périodes.

Dans le tableau 9.1 nous avons représenté les "durées empiriques" des différentes périodes où le marché est fermé. Ces valeurs ont été obtenues par le calcul des variances empiriques des rendements logarithmiques normalisées par la variance empirique correspondant au jour complet

$$\widehat{\sigma}^2 = \sum_{j=1}^{N_p} \widehat{\sigma}_{\text{j}}^2 + \widehat{\sigma}_{\text{n}}^2. \quad (9.1)$$

Notons que l'on peut observer des différences entre les valeurs obtenues pour les contrats à terme sur les taux de changes correspondant aux différentes zones géographiques : européennes (la livre britannique (GBP) et le franc suisse (CHF)) et asiatique (le yen japonais (JPY)).

9.3 Deux modèles d'évolution des prix intra-journaliers

Le modèle MRW log-normal, tel qu'il a été introduit dans la section 3.4, n'intègre pas de saisonnalité journalière. Néanmoins, de simples modifications permettent de rendre le modèle MRW log-normal adapté à la représentation des données intra-journalières. Dans cette section nous étudierons deux modèles d'évolution des prix qui prennent en compte les effets saisonniers intra-journaliers.

³ En fait, il arrive que le marché soit fermé trois jours de suite le week-end et le lundi ou le vendredi.

9.3.1 Volatilité saisonnière multiplicative

Etant donné un modèle à temps continu de séries financières journalières, le problème qui consiste à comprendre comment modifier ce modèle afin d'en faire un modèle intra-journalier et donc de rendre compte des saisonnalités journalière observées précédemment est un problème classique qui est pratiquement toujours traité de la même façon [TX95, CT96, DGM⁺01, BG01]. La saisonnalité est modélisée par un facteur multiplicatif (périodique de période un jour) qu'il est facile d'estimer par des moyennes empiriques. Diviser par ce facteur multiplicatif revient à "désaisonnaliser" la série intra-journalière afin qu'elle puisse être bien modélisée par le modèle initiale.

Il est naturel d'appliquer cette méthode au modèle MRW log-normal. Le logarithme du prix s'écrit alors comme

$$\ln(P(t)) = B \left(\int_0^t \theta^2(s) dM(s) \right), \quad (9.2)$$

où $\theta(s)$ est une fonction à variation lente (voire constante par morceaux), qui peut être estimée par

$$\widehat{\theta}^2(s) = \sum_{j=1}^{N_p} \frac{\widehat{\sigma}_{\Delta_j}^2}{|\Delta_j|} \mathbf{1}_{\{s \in \Delta_j\}} + \frac{\widehat{\sigma}_{\Delta_n}^2}{|\Delta_n|} \mathbf{1}_{\{s \in \Delta_n\}} + \frac{\widehat{\sigma}_{\Delta_w}^2}{|\Delta_w|} \mathbf{1}_{\{s \in \Delta_w\}} + \frac{\widehat{\sigma}_{\Delta_{jf}}^2}{|\Delta_{jf}|} \mathbf{1}_{\{s \in \Delta_{jf}\}} + \frac{\widehat{\sigma}_{\Delta_{lw}}^2}{|\Delta_{lw}|} \mathbf{1}_{\{s \in \Delta_{lw}\}}. \quad (9.3)$$

9.3.2 Volatilité saisonnière subordonnée

Dans le cadre du modèle MMAR (cf. section 2.4.2), nous avons introduit la notion de temps physique (ou *trading time*), il est possible alors de prendre en compte les effets saisonniers à travers le changement de temps. Plus précisément, la volatilité du marché peut être modélisée par une représentation sous forme de processus stochastique subordonné, le logarithme du prix s'écrit alors comme

$$\ln(P(t)) = B \left(M \left(\int_0^t \theta^2(s) ds \right) \right), \quad (9.4)$$

où $\theta(s)$ est une fonction à variation lente (voire constante par morceaux), qui peut être estimée par (9.3).

9.3.3 Choix du modèle

Pour détecter quel modèle de volatilité saisonnière permet de mieux modéliser les données financières, nous étudions le comportement théorique des moments des accroissements absolus en fonction de l'ordre. Si nous supposons que l'intervalle $[t - \tau, t]$ appartient à un des intervalles $\{\Delta_j\}_{j=1, \dots, N_p}$, Δ_n , Δ_w , Δ_{jf} et Δ_{lw} , en d'autres termes si la fonction

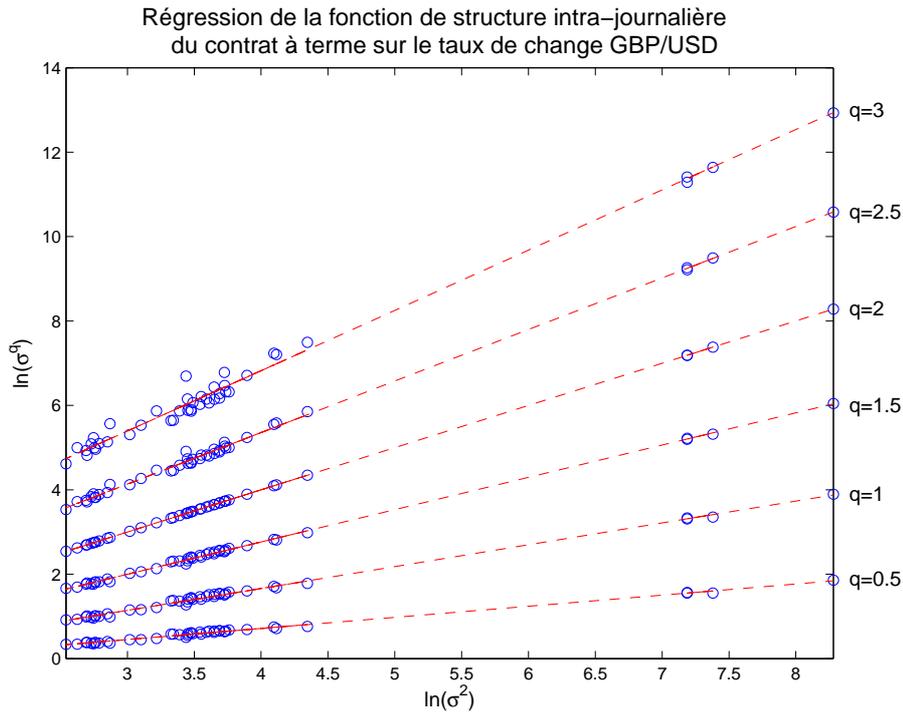


FIG. 9.5 – Résultats de l'étude de la dépendance de la fonction de structure intra-journalière du contrat à terme sur le taux de change GBP/USD. Les traits tiretés représentent les régressions linéaires des traits correspondant aux différentes valeurs de l'ordre q . On observe la dépendance linéaire, ce qui favorise le choix du modèle de la volatilité subordonnée.

$\theta(s)$ ne change pas de valeur sur cet intervalle, alors le moment d'ordre q du rendement logarithmique absolu est donné :

– dans le cadre de la volatilité multiplicatif, par

$$\sigma^q([t - \tau, t]) = \mathbb{E} \left[\left| \ln \left(\frac{P(t)}{P(t - \tau)} \right) \right|^q \right] = C(q) \theta^q(t) \tau^{\zeta_M(\frac{q}{2})}, \quad (9.5)$$

– dans le cadre de la volatilité subordonnée, par

$$\sigma^q([t - \tau, t]) = \mathbb{E} \left[\left| \ln \left(\frac{P(t)}{P(t - \tau)} \right) \right|^q \right] = C(q) \theta^{2\zeta_M(\frac{q}{2})}(t) \tau^{\zeta_M(\frac{q}{2})}. \quad (9.6)$$

Notons que les seconds moments théoriques sont égaux pour les deux modèles (9.5) et (9.6).

Nous introduisons la **fonction de structure intra-journalière**

$$\widehat{\sigma}^q = (\widehat{\sigma}^q(\Delta_1), \dots, \widehat{\sigma}^q(\Delta_{N_p}), \widehat{\sigma}^q(\Delta_n), \widehat{\sigma}^q(\Delta_w), \widehat{\sigma}^q(\Delta_{jf}), \widehat{\sigma}^q(\Delta_{lw})), \quad (9.7)$$

définie par les moments absolus empiriques de l'ordre q des rendements logarithmiques sur les intervalles $\{\Delta_j\}_{j=1, \dots, N_p}$, Δ_n , Δ_w , Δ_{jf} et Δ_{lw} .

Il est possible de comparer visuellement les deux modèles de volatilité saisonnière décrits ci-dessus en vérifiant si le logarithme de la fonction de structure intra-journalière de l'ordre q dépend linéairement du logarithme de la fonction de structure intra-journalière du second ordre. Si les données à haute fréquence suivent le modèle subordonné alors on doit observer la dépendance linéaire, sinon on doit observer celle non linéaire.

Dans la figure 9.5 nous avons représenté les résultats de la comparaison des deux modèles de volatilité saisonnière. Il est simple d'observer la dépendance linéaire, ce qui favorise le choix du modèle de la volatilité subordonnée.

Puisque nous n'avons pas effectué de tests statistiques qui peuvent réfuter un des deux modèles, dans la suite nous allons effectuer l'estimation des paramètres et la prédiction de volatilité et Valeur à Risque dans le cadre des deux modèles. Ces résultats nous permettront de confirmer *a posteriori* que le modèle de volatilité saisonnière subordonnée est mieux adapté à la représentation des données intra-journalières.

9.4 Estimation des paramètres

9.4.1 Volatilité saisonnière multiplicative

Dans le cadre du modèle de volatilité saisonnière multiplicative les accroissements d'un processus MRW log-normal uniformément échantillonné modélisent les rendements logarithmiques désaisonnalisés, c-à-d. les rendements logarithmiques normalisés par les écart-types empiriques $(\widehat{\sigma^2})^{1/2}$ correspondants. En conséquence, tous les estimateurs des paramètres du modèle MRW log-normal, introduits dans le chapitre 6, peuvent être appliqués à la série des rendements logarithmiques désaisonnalisés.

Dans la figure 9.6 nous avons représenté la fonction d'autocovariance empirique des magnitudes des rendements logarithmiques désaisonnalisés. Il est simple de voir qu'elle décroît plus lentement que la loi logarithmique, ce qui ne correspond pas au cas du modèle MRW log-normal (cf. figure 3.11). Strictement dit, une telle décroissance de la fonction d'autocovariance empirique indique un caractère non stationnaire de la série, ce qui met en doute la validité de l'approche classique de la désaisonnalisation (cf. section 9.3.1).

9.4.2 Volatilité saisonnière subordonnée

Régression de la fonction d'autocovariance du logarithme des volatilités agrégées

La première méthode d'estimation du coefficient d'intermittence, dans le cadre du modèle de volatilité saisonnière subordonné, consiste à étudier la fonction d'autocovariance du logarithmes des volatilités réalisées journalières (cf. section 1.2.6) obtenues par l'agrégation des rendements logarithmiques intra-journaliers au carré (cf. équation (1.16)). Dans cette

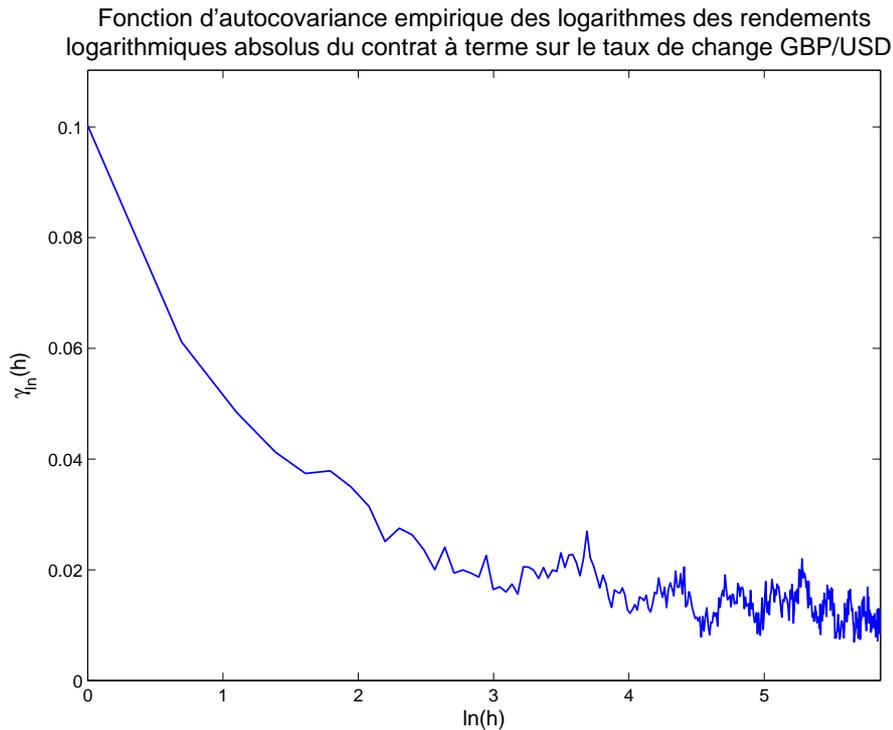


FIG. 9.6 – Fonction d'autocovariance empirique des magnitudes des rendements logarithmiques désaisonnalisés du contrat à terme sur le taux de change GBP/USD. On observe la décroissance plus lente que logarithmique, ce qui ne correspond pas au cas du modèle MRW log-normal.

approche, nous nous permettrons de négliger la présence des week-ends et des jours fériés, puisque nous ne cherchons ici que l'ordre de grandeur du coefficient d'intermittence.

La figure 9.7 représente les résultats de la méthode décrite appliquée aux prix du contrat à terme sur le taux de change GBP/USD. L'estimation du coefficient d'intermittence, obtenue par la régression linéaire, est égale à 0.0042. Il faut noter que même si l'estimation obtenue correspond à une intermittence assez faible les intervalles de confiance sont très larges, vue que la série est assez courte (1800 jours). Néanmoins, la décroissance logarithmique de la fonction d'autocovariance empirique donne un argument en faveur du modèle de volatilité saisonnière subordonnée.

Régression de la fonction d'autocovariance des magnitudes

La deuxième méthode d'estimation du coefficient d'intermittence, dans le cadre du même modèle, consiste à étudier le comportement de la fonction d'autocovariance obtenue par la moyenne des fonctions d'autocovariance empiriques des magnitudes correspondant aux différents intervalles dans la journée $\{\Delta_j\}_{j=1,\dots,N_p}$, ce qui permet de diminuer les fluctuations statistiques.

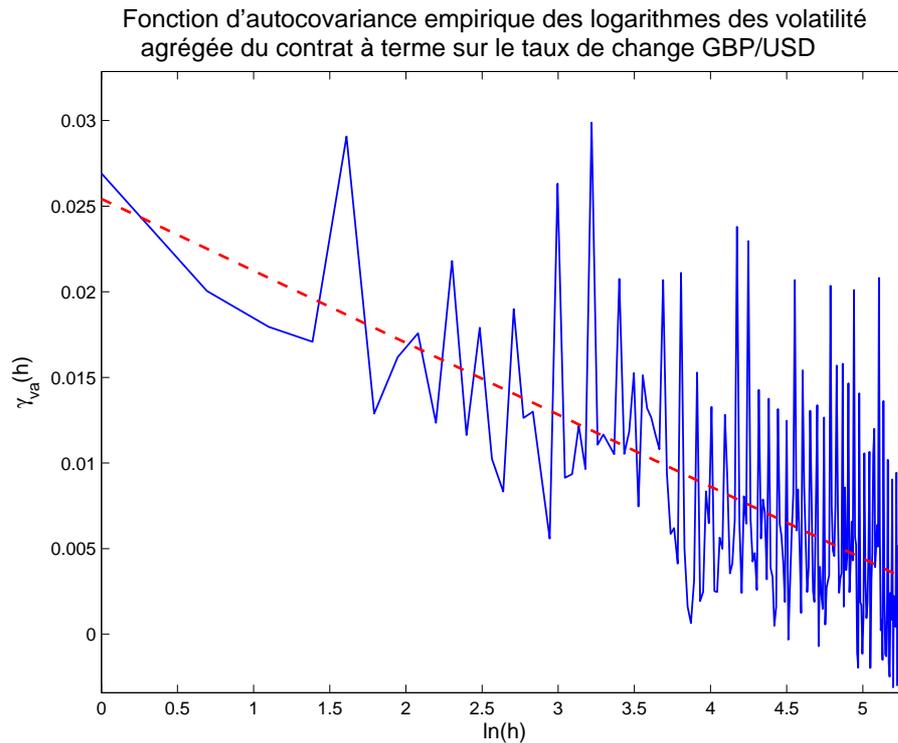


FIG. 9.7 – Fonction d'autocovariance empirique des logarithmes de la volatilité agrégée du contrat à terme sur le taux de change GBP/USD. Le trait tireté correspond à la courbe de régression par une fonction décroissante logarithmiquement $0.0042 \ln(425/h)$, ce qui correspond au cas du modèle MRW log-normal.

Dans la figure 9.8 nous avons représenté le résultat de la méthode décrite ci-dessus. Il est simple de voir que le tracé possède une décroissance logarithmique, ce qui est consistant avec le modèle MRW log-normal de volatilité saisonnière subordonnée. La faible valeur d'intermittence 0.0061 peut être expliquée par la petite taille de la série (1800 jours) ou par le fait de négliger les jours fériés qui introduisent les décalages dans le calcul des fonctions d'autocovariance empiriques que nous n'avons pas contrôlé.

9.5 Prédiction de volatilité

Les résultats obtenus dans les sections 9.3.3 et 9.4.2 nous permettent de choisir le modèle de volatilité saisonnière subordonnée pour reproduire des effets saisonniers. Dans cette section nous représentons la comparaison les résultats de la prédiction de volatilité obtenus en utilisant les modèles MRW log-normaux journalier et intra-journalier.

Dans le tableau 9.2 nous avons représenté les résultats d'évaluation des différentes prédictions de volatilité à l'aide des modèles MRW log-normal standards utilisant des données

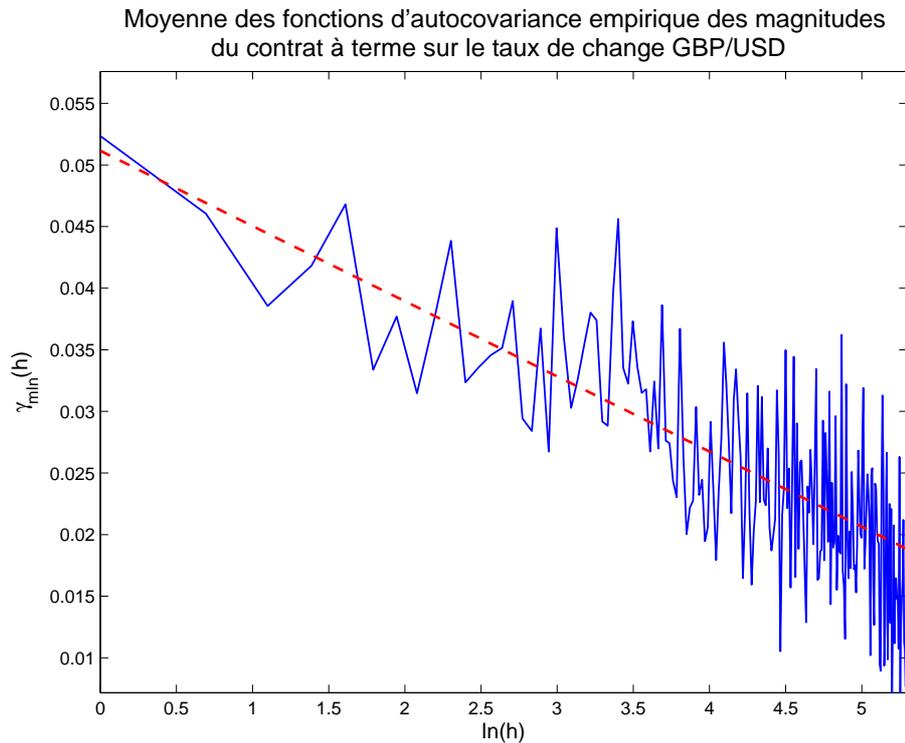


FIG. 9.8 – Fonctions d'autocovariance empirique obtenue par la moyenne des celles des magnitudes du contrat à terme sur le taux de change GBP/USD correspondant aux différents intervalles dans la journée $\{\Delta_j\}_{j=1,\dots,N_p}$. Le trait tireté correspond à la courbe de régression par une fonction décroissante logarithmiquement $0.0061 \ln(4685/h)$, ce qui correspond au cas du modèle MRW log-normal.

journalières et ceux de volatilité saisonnière subordonnée. La volatilité réalisée est calculée par les rendements logarithmiques intra-journaliers au carré.

Les méthodes basées sur le modèle standard utilisent des filtres linéaires de taille 100 jours calculés une fois au début de la procédure d'évaluation de la prédiction de volatilité. Alors que, les méthodes basées sur le modèle à haute fréquence utilisent des filtres linéaires de taille 4 jours recalculés à chaque étape de la procédure d'évaluation pour tenir compte de la grille du temps non uniforme.

Suivant les résultats du chapitre 6, nous allons choisir arbitrairement le temps intégral T égal à l'échelle d'observation $L = 1800$ jours. Nous allons choisir pour le coefficient d'intermittence λ^2 la valeur typique des séries financières égal à 0.02.

On observe que les deux méthodes ont des performances plus ou moins équivalentes, ce qui ne permet pas de choisir la meilleure méthode de prédiction de volatilité. Alors qu'il est naturel de construire les méthodes mixtes de prédiction de volatilité qui consistent à rendre la valeur moyenne des deux premières méthodes, en espérant de réduire la variance d'erreur de la prédiction. Il est simple de voir que toutes les nouvelles méthodes : linéaire,

Echelle	Horizon	Journalière			Intra-journalière			Mixte		
		Lin	Quad	Log	Lin	Quad	Log	Lin	Quad	Log
MAE										
1 jour	1 jour	0.167	0.169	0.165	0.160	0.165	0.163	0.157	0.165	0.156
5 jours	5 jours	0.497	0.493	0.508	0.499	0.526	0.532	0.451	0.492	0.461
10 jours	10 jours	0.852	0.851	0.885	0.852	0.928	0.928	0.760	0.861	0.787
20 jours	20 jours	1.598	1.578	1.662	1.537	1.645	1.704	1.385	1.571	1.418
1 jour	6 jours	0.165	0.170	0.161	0.154	0.166	0.156	0.155	0.165	0.152
5 jours	10 jours	0.504	0.513	0.509	0.487	0.531	0.516	0.463	0.508	0.468
10 jours	20 jours	0.944	0.958	0.931	0.904	0.939	0.976	0.859	0.932	0.857
20 jours	40 jours	1.826	1.835	1.788	1.598	1.680	1.713	1.594	1.731	1.589
MSE										
1 jour	1 jour	0.070	0.070	0.072	0.072	0.070	0.072	0.067	0.067	0.067
5 jours	5 jours	0.452	0.436	0.494	0.499	0.518	0.530	0.399	0.431	0.418
10 jours	10 jours	1.198	1.138	1.375	1.374	1.504	1.514	1.014	1.160	1.094
20 jours	20 jours	3.982	3.699	4.585	4.278	4.820	4.883	3.180	3.708	3.452
1 jour	6 jours	0.069	0.069	0.071	0.069	0.069	0.071	0.066	0.067	0.068
5 jours	10 jours	0.454	0.448	0.499	0.500	0.525	0.543	0.418	0.451	0.446
10 jours	20 jours	1.414	1.370	1.529	1.460	1.539	1.636	1.235	1.337	1.318
20 jours	40 jours	5.363	5.102	5.748	4.638	5.048	5.303	4.340	4.705	4.631

TAB. 9.2 – Résultats d'évaluation des prédictions de volatilité du contrat à terme sur le taux de change GBP/USD aux différentes échelles et aux différents horizons, en utilisant les différentes méthodes : linéaire ("Lin"), quadratique ("Quad") et logarithmique ("Log"), en appliquant les méthodes : journalière ("Journalière"), intra-journalière ("Intra Journalière") et mixte ("Mixte"), basées sur le modèle MRW log-normal. Les méthodes journalières utilisent le filtre linéaire de taille 100 jours, et celles intra-journalières utilisent le filtre linéaire de 4 jours. On observe les performances équivalentes des deux premières méthodes, alors que la méthode mixte est systématiquement plus performante.

quadratique et logarithmiques sont systématiquement les plus performantes.

9.6 Prédiction de VaR conditionnelle

Dans le tableau 9.3 nous avons représenté les résultats des prédictions VaR conditionnelle à différentes échelles, horizons et niveaux de probabilité obtenues à l'aide des modèles MRW

	MRW	MRW Intra
	Echelle = 1 jour,	Horizon = 1 jour
0.5%	0.003 (0.375)	0.010 (0.032)
1%	0.007 (0.351)	0.019 (0.005)
5%	0.050 (0.958)	0.076 (0.000)
10%	0.107 (0.436)	0.136 (0.000)
20%	0.196 (0.740)	0.213 (0.278)
	Echelle = 1 jour,	Horizon = 6 jours
0.5%	0.002 (0.172)	0.006 (0.696)
1%	0.007 (0.214)	0.012 (0.409)
5%	0.047 (0.579)	0.066 (0.012)
10%	0.111 (0.195)	0.126 (0.003)
20%	0.194 (0.582)	0.208 (0.509)
	Echelle = 1 jour,	Horizon = 11 jours
0.5%	0.004 (0.676)	0.003 (0.384)
1%	0.006 (0.117)	0.011 (0.772)
5%	0.053 (0.684)	0.058 (0.241)
10%	0.107 (0.434)	0.122 (0.015)
20%	0.195 (0.685)	0.207 (0.605)
	Echelle = 5 jours,	Horizon = 5 jours
0.5%	0.000 (0.121)	0.004 (0.847)
1%	0.000 (0.028)	0.004 (0.301)
5%	0.025 (0.050)	0.050 (0.988)
10%	0.075 (0.179)	0.120 (0.306)
20%	0.195 (0.746)	0.202 (0.974)
	Echelle = 5 jours,	Horizon = 10 jours
0.5%	0.000 (0.122)	0.004 (0.851)
1%	0.004 (0.307)	0.004 (0.304)
5%	0.029 (0.112)	0.037 (0.353)
10%	0.087 (0.388)	0.112 (0.526)
20%	0.204 (0.974)	0.203 (1.000)

TAB. 9.3 – Résultats des prédictions de VaR conditionnelle à différentes échelles, horizons et niveaux de probabilité pour le contrat à terme sur le taux de change GBP/USD en utilisant les modèles MRW log-normal journalier ("MRW") et à haute fréquence ("MRW Intra"). Pour chaque prédiction on précise la probabilité empirique et la p -valeur du test statistique de Kupiec. Les résultats apparaissent en gras si le test de Kupiec est accepté pour le niveau de confiance de 95%. On observe que les deux modèles ont les performances équivalentes.

log-normal : journalier et à haute fréquence.

Pour effectuer la prédiction de VaR, Nous avons utilisé les mêmes filtres linéaires que pour la prédiction de volatilité (cf. section 9.5).

On observe que les deux modèles ont les performances plus ou moins équivalentes, la plupart des tests statistiques de Kupiec (cf. section 8.5.1) sont acceptés. Pour choisir le meilleur modèle, il est donc nécessaire d'effectuer une étude plus fine de comparaison des modèles MRW log-normaux proposés.

9.7 Conclusion

D'après ce qui précède, le modèle MRW log-normal à volatilité saisonnière subordonnée semble assez bien reproduire les données intra-journalières. Notre étude laisse pourtant beaucoup de questions sans réponse comme autant de sujets de recherches futures. Sur un plan théorique, il serait intéressant de mieux comprendre la structure générale des données intra-journalières. Ainsi, dans nos études nous n'avons utilisé que les prix d'actifs financiers, alors qu'il est possible d'utiliser aussi l'information disponible sur les volumes de transactions intra-journalières.

Comme nous avons fait remarquer dans la section 9.5, l'application des méthodes mixtes de prédiction de volatilité qui utilisent à la fois les données de clôture et les données intra-journalières, permet d'améliorer les erreurs de prédiction. Par conséquent, il est très naturel d'étudier de près les méthodes mixtes de prédiction de volatilité, d'aborder le problème d'optimalité de telles méthodes. De même, il est envisagé de construire des prédicteurs mixtes de VaR.

Toujours en vue d'une utilisation pratique, il est nécessaire de développer des techniques robustes d'estimation du coefficient d'intermittence adaptées aux données financières intra-journalières. Le problème est extrêmement délicat en raison des effets saisonniers qui ne permettent pas de considérer les données comme étant uniformément échantillonnées, ce qui empêche d'appliquer directement le résultat du théorème 6.9.

Une autre direction de recherche porte sur la modélisation des données financières intra-journalières aux échelles inférieures à 5 minutes, c'est-à-dire les échelles pour lesquelles on ne peut pas négliger les effets de microstructure des marchés.

Conclusion

La modélisation des nombreuses propriétés statistiques des séries financières (décorrélation des rendements mais dépendances à longue portée des volatilités, accumulation des volatilités, distributions à queues épaisses à petites échelles et quasi-gaussiennes à grandes échelles, ...) nécessite l'utilisation de modèle relativement complexe.

Récemment a été proposée l'utilisation de la classe des processus multifractals et en particulier des processus MRW log-normaux. Ces processus sont bien adaptés pour la modélisation en finance : ils permettent de reproduire la plupart des faits stylisés classiques, ainsi que les faits stylisés fractals (invariance d'échelle exacte et la structure cohérente d'autocovariance des magnitudes).

Pour démontrer la plupart des propriétés du processus MRW log-normal, il nous a fallu utiliser sa généralisation log-infiniment divisible. De cette façon, nous avons obtenu les conditions d'existence des moments négatifs (et donc des moments du logarithme) de la mesure MRM et démontré la propriété d'invariance par changement du temps intégral.

Un des points souvent critiqués du modèle MRW log-normal concerne la valeur théorique des exposants de queues de distribution, valeur bien supérieure à celle estimée par de nombreux travaux économétriques. Une partie de notre travail, basée sur un calcul qualitatif, permet d'expliquer cette différence, qui provient essentiellement des dépendances à longue portée de la volatilité.

Une grande partie de notre travail est consacrée à l'approximation log-normale de la mesure MRM et du processus MRW sous condition de faible intermittence, $\lambda^2 \ll 1$. Les résultats théoriques, que nous avons obtenus, nous ont conduit à élaborer des méthodes d'estimation des paramètres du modèle et des méthodes de prédiction de risques.

Notre étude sur l'estimation des paramètres a permis de montrer que le nombre de paramètres effectifs du modèle se réduit à deux : le coefficient d'intermittence λ^2 , qui est en pratique fixé à la valeur typique 0.02, et la variance "locale", seul paramètre réellement estimé lors des applications. Nous avons également démontré que l'estimateur du coefficient d'intermittence $\widehat{\lambda^2}$, obtenu par la régression de la fonction d'autocovariance

des magnitudes, est asymptotiquement non biaisé et consistant.

L'utilisation du modèle MRW log-normal pour la prédiction des risques a été développée dans deux cadres : prédiction de volatilité et prédiction de Valeur à Risque. Les résultats empiriques obtenus montrent que le modèle MRW est systématiquement plus performant que les modèles du type GARCH(1,1), et ce, même si l'on favorise les modèles GARCH en estimant, pour chaque cours individuel, leur paramètres sur l'intégralité des données disponibles.

Il faut noter que les résultats théoriques, qui concernent l'approximation log-normale de la mesure MRM et du processus MRW, permettent de justifier rigoureusement les méthodes de prédiction de volatilité. Néanmoins, plusieurs quantités mériteraient d'être étudiées plus finement pour justifier rigoureusement la méthode de prédiction de Valeur à Risque conditionnelle qui repose sur l'approximation log-normale conditionnelle de la mesure MRM. Une approche possible consiste à étudier la validité du développement de Cornish-Fisher en le coefficient d'intermittence λ^2 de la distribution conditionnelle de la mesure MRM qui permet de justifier l'approximation que nous avons utilisée.

Dans le cadre de l'étude des données intra-journalières, nous avons montré que les méthodes mixtes de prédiction de volatilité, qui utilisent à la fois les filtres journaliers et intra-journaliers, permettent d'améliorer la performance de la prédiction. Il semble donc intéressant d'étudier l'optimisation de la méthode mixte pour la prédiction de volatilité, en effectuant l'analyse de filtres linéaires correspondant à différentes échelles d'échantillonnage de données. Il semble naturel d'envisager d'élaborer des méthodes mixtes de prédiction de Valeur à Risque.

L'intérêt applicatif de notre travail est multiple. Ainsi, la gestion des risques fait de plus en plus appel à des modèles perfectionnés, qui reposent sur des hypothèses réalistes sur les distributions des rendements de produits financiers de base, pour mieux contrôler ses positions. Nous espérons que les résultats, que nous avons obtenus dans le cadre du problème de prédiction de Valeur à Risque, et notamment la méthode performante de prédiction de la distribution de la volatilité conditionnelle, pourront être utilisés pour la meilleure gestion dynamique des risques de produits financiers "complexes", par exemple, produits dérivés.

Dans le cadre du problème de couverture d'une option, il est intéressant d'étudier si l'utilisation de la modélisation des cours des actifs sous-jacents par les processus MRW log-normaux permet d'améliorer la performance de la couverture. De cette façon, il est naturel d'effectuer l'étude de prédiction de Valeur à Risque de la volatilité implicite conditionnellement à la volatilité historique de l'option.

Éléments de l'économétrie linéaire

Modèle linéaire et estimateur des moindres carrés ordinaires

On considère une variable d'intérêt y appelée variable **dépendante** et un ensemble de K variables dites **explicatives** auquel on adjoint une constante. On dispose de N observations. On note $\underline{y} = (y_1, \dots, y_N)$ l'empilement des N observations de la variable dépendante. On définit de même les vecteurs $\underline{x}_1, \dots, \underline{x}_K$ et \underline{x} la matrice des variables explicatives à laquelle on adjoint le vecteur constant $e = (1, \dots, 1)^T$

$$\underline{x} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1K} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nK} \end{pmatrix}. \quad (\text{A.1})$$

Donc \underline{x} est une matrice de dimension $N \times (K + 1)$.

Définition A.1

Le **modèle linéaire** s'écrit sous forme matricielle

$$\underline{y} = \underline{x}b + \underline{u}, \quad (\text{A.2})$$

avec un vecteur des **résidus** ou **perturbation** $\underline{u} = (u_1, \dots, u_N)^T$.

Définition A.2

L'**estimateur des moindres carrés ordinaires (MCO)** est défini comme le vecteur $b = (b_0, b_1, \dots, b_K)^T$ de dimension $K + 1$, des coefficients de la combinaison linéaire de $e, \underline{x}_1, \dots, \underline{x}_K$ réalisant le minimum de la distance de \underline{y} à l'espace vectoriel de \mathbb{R}^N engendré par $e, \underline{x}_1, \dots, \underline{x}_K$, pour la norme euclidienne

$$\hat{b}_{mco} = \underset{b}{\operatorname{argmin}} \|\underline{y} - \underline{x}b\|^2. \quad (\text{A.3})$$

Proposition A.1

Sous l'hypothèse

H1 : les vecteurs $e, \underline{x}_1, \dots, \underline{x}_K$ sont indépendants,

l'estimateur des moindres carrés ordinaires existe, est unique et a pour expression

$$\hat{b}_{mco} = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{y}. \quad (\text{A.4})$$

Définition A.3

*On dit qu'un estimateur $\hat{b}(\underline{y}, \underline{x})$ est **sans biais** lorsque*

$$\mathbb{E}[\hat{b}(\underline{y}, \underline{x})] = b. \quad (\text{A.5})$$

Proposition A.2

Sous l'hypothèse

H2 : l'espérance des résidus conditionnellement aux variables explicatives est nulle,

$$\mathbb{E}[u_n | \underline{x}] = 0, \text{ pour tout } n. \quad (\text{A.6})$$

l'estimateur des moindres carrés ordinaires est sans biais.

Proposition A.3

Sous les hypothèses H1, H2,

H3 : les résidus sont conditionnellement aux variables explicatives des variances identiques

$$\text{Var}[u_n | \underline{x}] = \sigma^2, \text{ pour tout } n, \quad (\text{A.7})$$

H4 : les résidus sont non corrélées les unes avec les autres

$$\mathbb{E}[u_n u_m | \underline{x}] = 0, \text{ pour tout } n \neq m, \quad (\text{A.8})$$

la variance de l'estimateur des moindres carrés ordinaires conditionnellement aux variables explicatives est donnée par

$$\text{Var}[\hat{b}_{mco} | \underline{x}] = \sigma^2 (\underline{x}^T \underline{x})^{-1}. \quad (\text{A.9})$$

La variance non conditionnelle est donnée par

$$\text{Var}[\hat{b}_{mco}] = \sigma^2 \mathbb{E}[(\underline{x}^T \underline{x})^{-1}]. \quad (\text{A.10})$$

Théorème A.4 (Gauss-Markov)

Sous les hypothèses H1-H4 l'estimateur des moindres carrés ordinaires du modèle linéaire (A.2) est optimale dans la classe des estimateurs sans biais conditionnellement aux variables explicatives.

Homoscédasticité et hétéroscédasticité

Définition A.4

On appelle modèle linéaire **hétéroscédastique** le modèle dans lequel un vecteur de variables aléatoires \underline{y} dépend linéairement de $K + 1$ variables explicatives \underline{x}

$$\underline{y} = \underline{x}b + \underline{u}, \quad (\text{A.11})$$

avec les hypothèses

$$H1 : \mathbb{E}[\underline{u}|\underline{x}] = 0,$$

$$H2 : \mathbb{V}\text{ar}[\underline{u}|\underline{x}] = \Omega \text{ inversible},$$

$$H3 : \underline{x}^T \underline{x} \text{ est inversible}.$$

Le modèle est dit **homoscédastique** si l'hypothèse H2 est remplacée par l'hypothèse

$$H2 \text{ bis} : \mathbb{V}\text{ar}[\underline{u}|\underline{x}] = \sigma^2 \mathbb{I}d_N.$$

Proposition A.5

Sous les hypothèses H1-H3, l'estimateur des moindres carrés ordinaires est sans biais

$$\mathbb{E}[\hat{b}_{mco}|\underline{x}] = \mathbb{E}[(\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{y}|\underline{x}] = 0, \quad (\text{A.12})$$

et sa variance conditionnelle aux variables explicatives est donnée par

$$\mathbb{V}\text{ar}[\hat{b}_{mco}|\underline{x}] = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \Omega \underline{x} (\underline{x}^T \underline{x})^{-1}. \quad (\text{A.13})$$

Rappel sur la régression linéaire

La **régression** est une méthode de prédiction très utilisée en plusieurs domaines scientifiques : économie, statistique, traitement du signal, ...

Dans le cadre de la régression linéaire, nous cherchons à prédire, avec le plus de précision possible, la valeur $Y[t]$ prise par un processus stochastique faiblement stationnaire Y , dite variable **endogène**, à partir d'une série de variables **explicatives** $\{X[t - k]\}_{k=1, \dots, n}$ données par les valeurs prises par un processus faiblement stationnaire X .

$$Y[t] - \mathbb{E}[Y] = \sum_{k=1}^n \alpha_k (X[t - k] - \mathbb{E}[X]) + \eta[t], \quad (\text{A.14})$$

où $\eta[t]$ est une erreur de la régression qui exprime l'information manquante dans l'explication linéaire et $\{\alpha_k\}_{k=1, \dots, n}$ sont les coefficients du filtre linéaire à estimer.

Nous supposons que les variables explicatives $\{X[t - k]\}_{k=1, \dots, n}$ et l'erreur de la régression $\eta[t]$ sont décorrélées, ce qui nous conduit au système d'équations linéaires suivant

$$\begin{pmatrix} \rho_X[0] & \rho_X[1] & \cdots & \rho_X[n-1] \\ \rho_X[1] & \rho_X[0] & \cdots & \rho_X[n-2] \\ \vdots & \vdots & \ddots & \vdots \\ \rho_X[n-1] & \rho_X[n-2] & \cdots & \rho_X[0] \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \rho_{Y,X}[1] \\ \rho_{Y,X}[2] \\ \vdots \\ \rho_{Y,X}[n] \end{pmatrix}, \quad (\text{A.15})$$

où $\rho_X[\cdot]$ est la fonction d'autocovariance du processus X

$$\rho_X[k] = \text{Cov}[X[t], X[t - k]], \text{ pour tout entier } k, \quad (\text{A.16})$$

et $\rho_{x,Y}[\cdot]$ est la fonction de covariance croisée entre les processus Y et X

$$\rho_{Y,X}[k] = \text{Cov}[Y[t], X[t - k]], \text{ pour tout entier } k. \quad (\text{A.17})$$

Les coefficients $\{\alpha_k\}_{k=1,\dots,n}$ peuvent être calculés en inversant la matrice de Toeplitz, qui intervient dans le système (A.15), à l'aide de l'algorithme rapide de Levinson-Durbin [Lev47, Dur60, BD96] dont la complexité est de l'ordre de $O(n^2)$.

Dimension de Hausdorff

Définition B.1 (Mesure de Hausdorff s -dimensionnelle)

Soit E est un sous ensemble de \mathbb{R} et $s \geq 0$. Pour tout $\delta > 0$ on définit

$$\mathcal{H}_\delta^s(E) = \inf \left\{ \sum_{j=1}^{\infty} |U_j|^s \mid \{U_j\} \text{ est un recouvrement de } E \text{ tel que } |U_j| \leq \delta \right\} \quad (\text{B.1})$$

La **mesure de Hausdorff** s -dimensionnelle de E est une limite

$$\mathcal{H}^s(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(E). \quad (\text{B.2})$$

Cette mesure s -dimensionnelle est décroissante en s . De plus, elle vérifie les propriétés suivantes. Si $\mathcal{H}^s(E) > 0$ alors pour tout $s' < s$, $\mathcal{H}^{s'}(E) = \infty$ et si $\mathcal{H}^s(E) = 0$, pour tout $s' > s$, $\mathcal{H}^{s'}(E) = 0$. Il existe alors une valeur critique de s pour laquelle $\mathcal{H}^s(E)$ passe de $+\infty$ à 0, c'est la dimension de Hausdorff de E .

Définition B.2 (Mesure de Hausdorff)

Soit $E \subset \mathbb{R}$. La **dimension de Hausdorff** ou la **dimension fractale** de E est

$$\dim_H(E) = \sup \{s \mid \mathcal{H}^s(E) = \infty\} = \inf \{s \mid \mathcal{H}^s(E) = 0\}. \quad (\text{B.3})$$

Simulation de variables gaussiennes corrélées

Cet appendice est consacré à la présentation détaillée de la méthode de simulation utilisée pour générer les variables gaussiennes corrélées $\{\omega_l[k]\}_k$ dans la construction de MRW log-normal. Bien que la méthode soit déjà décrite dans certains ouvrages, il nous a semblé intéressant pour le lecteur de la rappeler brièvement ici. Nous supposons connue une méthode de génération de variables gaussiennes indépendantes. Si ce n'est pas le cas, nous renvoyons le lecteur à l'ouvrage [VTPF93]. Le problème est donc de trouver une méthode efficace pour générer des variables corrélées à partir des variables indépendantes.

Soit $\{X[k]\}_{k=0,\dots,n-1}$ un échantillon gaussien dont la matrice de variance covariance Σ , symétrique semi-définie positive, s'écrit au moyen de la fonction d'autocovariance γ

$$\Sigma = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \dots & \gamma(0) \end{pmatrix}. \quad (\text{C.1})$$

Une première méthode consiste à utiliser la décomposition de Cholesky de la matrice de covariance [LL97]. Toutefois, cette méthode est bien adaptée à la simulation de petits échantillons, car facile à mettre en œuvre. En revanche, elle l'est beaucoup moins pour de grandes séries, car très lente, sa complexité est de l'ordre de $O(n^2)$ [Kai86], et nécessite aussi un très grand espace mémoire.

Nous décrivons un algorithme, proposé par Wood et Chen [WC94], qui est très précis numériquement. Cet algorithme dont la complexité est de l'ordre de $O(n \log_2(n))$, est basé sur une transformation de Fourier rapide [CT65]. Il permet en toute généralité la simulation de longues séries de variables gaussiennes corrélées selon la structure souhaitée.

Nous considérons la matrice circulaire de Toeplitz C de dimension $m \times m$, où m est un nombre pair, tel que $m \geq 2(n-1)$,

$$C = \begin{pmatrix} c_0 & c_1 & \dots & c_{m-1} \\ c_{m-1} & c_0 & \dots & c_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \dots & c_0 \end{pmatrix}, \quad (\text{C.2})$$

où

$$c_j = \begin{cases} \gamma(j), & \text{si } 0 \leq j \leq \frac{m}{2}, \\ \gamma(m-j), & \text{si } \frac{m}{2} < j \leq m-1. \end{cases} \quad (\text{C.3})$$

Si la matrice C est semi-définie positive, c'est une matrice de variance-covariance d'un échantillon circulaire de taille m . Notons que

$$C = Q\Lambda Q^*, \quad (\text{C.4})$$

où Λ est une matrice diagonale des valeurs propres de la matrice C et Q est une matrice unitaire composée des vecteurs propres de la matrice C .

La matrice Λ est donnée par

$$\Lambda = \begin{pmatrix} \lambda_0 & 0 & \dots & 0 \\ 0 & \lambda_1 & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_{m-1} \end{pmatrix}, \quad (\text{C.5})$$

où

$$\lambda_j = \frac{1}{\sqrt{m}} \sum_{k=0}^{m/2-1} \gamma(k) \exp\left(i\frac{2\pi}{m}jk\right) + \frac{1}{\sqrt{m}} \sum_{k=m/2}^{m-1} \gamma(m-k) \exp\left(i\frac{2\pi}{m}jk\right), \quad (\text{C.6})$$

par construction les valeurs propres $\lambda_0, \dots, \lambda_{m-1}$ sont réelles. Et la matrice Q est indépendante de la matrice C et elle donnée par

$$Q = \frac{1}{\sqrt{m}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{i\frac{2\pi}{m}} & e^{i\frac{4\pi}{m}} & \dots & e^{i\frac{2(m-1)\pi}{m}} \\ 1 & e^{i\frac{4\pi}{m}} & e^{i\frac{8\pi}{m}} & \dots & e^{i\frac{4(m-1)\pi}{m}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{i\frac{2(m-1)\pi}{m}} & e^{i\frac{4(m-1)\pi}{m}} & \dots & e^{i\frac{2(m-1)^2\pi}{m}} \end{pmatrix}. \quad (\text{C.7})$$

Il est simple de voir que les valeurs propres $\lambda_0, \dots, \lambda_{m-1}$ est une transformée de Fourier discrète de la première ligne de la matrice C . De plus, pour tout vecteur v le produit Qv est une transformée de Fourier discrète du vecteur v . En conséquence le vecteur $Y = Q\Lambda^{1/2}Q^*Z$, où Z est un vecteur gaussien centré de matrice de variance-covariance identité d'ordre m , peut être calculé à l'aide de deux transformations de Fourier si la matrice C est semi-définie positive, (c-à-d. que sa valeur propre minimale est non négative).

Ensuite, nous présentons l'algorithme de simulation d'un échantillon gaussien $\{X[k]\}$:

1. On choisit la dimension¹ $m = 2^g \geq 2(n-1)$.
2. On définit les valeurs propres $\lambda_0, \dots, \lambda_{m-1}$ de la matrice C , en utilisant la transformation de Fourier rapide de la première ligne de la matrice C .
3. Si la valeur propre minimale est négative alors on augmente de 1 la valeur de g et revient à l'étape 1, sinon on calcule les racines carrés des valeurs propres.
4. On simule directement, voir ci-dessus, le vecteur gaussien centré Q^*Z , où Z est un vecteur de m variables gaussiennes indépendantes standards.
5. On définit le vecteur Y comme la transformée de Fourier de $\Lambda^{1/2}Q^*Z$.
6. On extrait n éléments successifs de vecteur Y pour former l'échantillon souhaité.

Pour que cet algorithme soit complet, il faut décrire la simulation du vecteur Q^*Z en étape 4. Il est simple de voir que les éléments du vecteur Q^*Z vérifient

$$\Im[(Q^*Z)_0] = \Im[(Q^*Z)_{m/2}] = 0, \quad (\text{C.8})$$

$$(Q^*Z)_j = (QZ)_{m-j}, \text{ pour } 1 \leq j < m/2. \quad (\text{C.9})$$

De plus, $(Q^*Z)_0, (Q^*Z)_{m/2}, \sqrt{2}\Re[(Q^*Z)_1], \dots, \sqrt{2}\Re[(Q^*Z)_{m/2-1}], \sqrt{2}\Im[(Q^*Z)_1], \dots, \sqrt{2}\Im[(Q^*Z)_{m/2-1}]$ sont les variables gaussiennes indépendantes standards. Pour simuler le vecteur Q^*Z , tirons alors deux séries U_0, \dots, U_{n-1} et V_1, \dots, V_{n-2} des variables gaussiennes indépendantes standards et construisons les variables complexes

$$(Q^*Z)_j = \begin{cases} U_j, & \text{si } j = 0, m/2, \\ \frac{U_j + iV_j}{\sqrt{2}}, & \text{si } j = 1, \dots, m/2 - 1, \\ \frac{U_{m-j} - iV_{m-j}}{\sqrt{2}}, & \text{si } j = m/2 + 1, \dots, m - 1. \end{cases} \quad (\text{C.10})$$

Pour conclure nous présentons deux conditions sur la forme de la fonction d'autocovariance $\gamma(h)$ pour que la matrice C de dimension minimale possible $m = 2(n-1)$ soit semi-définie positive :

- La fonction $\gamma(h)$ est non négative et convexe pour h positifs [DN97, Gne98].
- La fonction $\gamma(h)$ devient nulle pour $h \geq h_0$, où h_0 est quelconque, et de plus $n \geq h_0$ [DN97, CD99, Gne00].

¹La dimension égale à une puissance de 2 permet d'augmenter l'efficacité de l'algorithme grâce à l'utilisation de la transformation de Fourier rapide.

Inégalités

Proposition D.1

Pour tout réel x , on a

$$e^x \geq 1 + x. \quad (\text{D.1})$$

Preuve - D.1 -

La fonction $f(x) = e^x - 1 - x$ est strictement convexe et atteint son minimum au point $x_0 = 0$, où la dérivée s'annule

$$f'(x_0) = e^{x_0} - 1 = 0. \quad (\text{D.2})$$

Sa valeur minimale est égale à $f(x_0) = e^{x_0} - 1 - x_0 = 0$, donc

$$f(x) = e^x - 1 - x \geq f(x_0) = 0, \quad (\text{D.3})$$

d'où l'on déduit l'inégalité souhaitée (D.1). ■

Proposition D.2

Pour tout ϵ positif strictement inférieur à 1 et pour tout entier positif N , il existe $m_\epsilon(N)$ et $M_\epsilon(N)$ positifs tels que

$$m_\epsilon(N)|x|^{N+1} \leq \left| \ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k \right| \leq M_\epsilon(N)|x|^{N+1}, \quad \text{pour } x \in [-\epsilon, \epsilon]. \quad (\text{D.4})$$

De plus, pour N impair, on a

$$\ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k \leq 0, \quad \text{pour } x \in (-1, +\infty). \quad (\text{D.5})$$

Preuve - D.2 -

On considère la fonction $f(x)$ définie par

$$f_N(x) = (-1)^N \frac{\ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k}{x^{N+1}}. \quad (\text{D.6})$$

Lemme D.3

La fonction $f_N(x)$ est continue décroissante sur l'intervalle $(-1, +\infty)$.

Preuve - D.3 -

Elle vérifie la propriété limite

$$\lim_{x \rightarrow 0} f_N(x) = \frac{1}{N+1}, \quad (\text{D.7})$$

donc la fonction $f_N(x)$ est continue sur l'intervalle $(-1, +\infty)$. Sa dérivée s'écrit

$$f'_N(x) = (-1)^N \left[-\frac{(N+1)\ln(1+x)}{x^{N+2}} + \frac{1}{(1+x)x^{N+1}} - \sum_{k=1}^N \frac{(-1)^k(N-k+1)}{k} x^{k-N-2} \right], \quad (\text{D.8})$$

On considère la fonction $g_N(x)$ définie par

$$g_N(x) = x^{N+2} f'_N(x) = (-1)^N \left[-(N+1)\ln(1+x) + \frac{x}{1+x} - \sum_{k=1}^N \frac{(-1)^k(N-k+1)}{k} x^k \right]. \quad (\text{D.9})$$

Elle vérifie la propriété limite

$$\lim_{x \rightarrow 0} g(x) = 0, \quad (\text{D.10})$$

et donc la fonction $g_N(x)$ est continue sur l'intervalle $(-1, +\infty)$. Sa dérivée s'écrit

$$g'_N(x) = -\frac{x^{N+1}}{(1+x)^2}. \quad (\text{D.11})$$

En conséquence, la dérivée $f'_N(x)$ est négative

$$f'_N(x) \leq 0, \text{ pour } x \in (-1, +\infty), \quad (\text{D.12})$$

et donc, la fonction $f_N(x)$ est décroissante. ■

Lemme D.4

La fonction $f_N(x)$ est positive sur l'intervalle $(-1, +\infty)$.

Preuve - D.4 -

On donne la démonstration par récurrence.

La fonction $f_0(x) = \frac{\ln(1+x)}{x}$ est positive.

Supposons qu'on a déjà démontré que la fonction $f_N(x)$ est positive.

Pour x négatif la fonction $f_N(x)$ est positive puisque sa dérivée est négative.

Pour x positif la fonction $f_N(x)$ vérifie

$$f_N(x) \leq \frac{1}{1+N} \quad (\text{D.13})$$

en conséquence, on a

$$(-1)^N \left[\ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k \right] \leq \frac{x^{N+1}}{N+1}, \text{ pour } x > 0, \quad (\text{D.14})$$

et donc

$$(-1)^{N+1} \left[\ln(1+x) + \sum_{k=1}^{N+1} \frac{(-1)^k}{k} x^k \right] \geq 0, \quad (\text{D.15})$$

ce qui signifie que la fonction $f_{N+1}(x)$ est positive. ■

D'après les deux lemmes, on a

$$f_N(-\epsilon) \geq f_N(x) \geq f_N(\epsilon) > 0, \text{ pour } x \in [-\epsilon, \epsilon], \quad (\text{D.16})$$

Et donc

$$f_N(\epsilon)|x|^{N+1} \leq \left| \ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k \right| \leq f_N(-\epsilon)|x|^{N+1}, \text{ pour } x \in [-\epsilon, \epsilon]. \quad (\text{D.17})$$

De plus, si N est impair, on a

$$-x^{N+1} f_N(x) = \ln(1+x) + \sum_{k=1}^N \frac{(-1)^k}{k} x^k \leq 0. \quad (\text{D.18})$$

■

Proposition D.5

Pour tout réel x supérieur à -1 , on a

$$\ln(1+x) \leq x. \quad (\text{D.19})$$

Preuve - D.5 -

L'inégalité (D.19) est une conséquence directe de l'inégalité (D.5) pour N égal à 1. ■

Proposition D.6

Soit p un nombre entier, alors pour tout réel x et tout positif ϵ , on a

$$x^p \leq \left(\frac{p}{\epsilon} \right)^p e^{\epsilon x} \quad (\text{D.20})$$

Preuve - D.6 -

En appliquant l'inégalité (D.1) à $\frac{\epsilon x}{p} - 1$, on a

$$\frac{\epsilon x}{p} \leq e^{\frac{\epsilon x}{p} - 1}. \quad (\text{D.21})$$

En calculant la $p^{\text{ième}}$ puissance on en déduit l'inégalité souhaitée. ■

Proposition D.7

Soit n un nombre entier, alors pour tout positifs x et h , on a

$$|\ln(x)|^n \leq \left(\frac{n}{e}\right)^n (x^{-h} h^{-n} + x). \quad (\text{D.22})$$

Preuve - D.7 -

Il suffit de démontrer que

$$|\ln(x)|^n \leq \begin{cases} \left(\frac{n}{e}\right)^n x^{-h} h^{-n}, & \text{si } x \in (0, 1), \\ \left(\frac{n}{e}\right)^n x, & \text{si } x \geq 1. \end{cases} \quad (\text{D.23})$$

Pour démontrer la première affirmation, appliquons l'inégalité (D.19) à $\frac{1}{e} x^{-\frac{h}{n}} - 1$

$$\ln\left(\frac{1}{e} x^{-\frac{h}{n}} - 1 + 1\right) = -\frac{h}{n} \ln(x) - 1 \leq \frac{1}{e} x^{-\frac{h}{n}} - 1, \quad (\text{D.24})$$

ce qui nous amène, pour $x \in (0, 1)$

$$|\ln(x)|^n = (-\ln(x))^n \leq \left(\frac{n}{he} x^{-\frac{h}{n}}\right)^n = \left(\frac{n}{e}\right)^n x^{-h} h^{-n}. \quad (\text{D.25})$$

Tandis que pour la deuxième affirmation, appliquons l'inégalité (D.19) à $\frac{1}{e} x^{\frac{1}{n}} - 1$

$$\ln\left(\frac{1}{e} x^{\frac{1}{n}} - 1 + 1\right) = \frac{1}{n} \ln(x) - 1 \leq \frac{1}{e} x^{\frac{1}{n}} - 1, \quad (\text{D.26})$$

et en conséquence,

$$\ln(x) \leq \frac{n}{e} x^{\frac{1}{n}}, \quad (\text{D.27})$$

d'où nous pouvons déduire l'inégalité souhaitée, pour $x \geq 1$. ■

Proposition D.8

Pour tout ϵ positif strictement inférieur à 1, on a

$$\frac{\epsilon - \ln(1 + \epsilon)}{\epsilon^2} x^2 \leq x - \ln(1 + x) \leq -\frac{\epsilon + \ln(1 - \epsilon)}{\epsilon^2} x^2, \text{ pour } x \in [-\epsilon, \epsilon]. \quad (\text{D.28})$$

Preuve - D.8 -

L'inégalité (D.28) est une conséquence directe de l'inégalité (D.5) pour N égal à 1. ■

Proposition D.9

Pour tout ϵ positif strictement inférieur à 1, il existe M_ϵ tel que

$$\left| \ln(1 + x) - x + \frac{1}{2}x^2 \right| \leq M_\epsilon |x|^3, \text{ pour } x \in [-\epsilon, \epsilon]. \quad (\text{D.29})$$

Preuve - D.9 -

L'inégalité (D.29) est une conséquence directe de l'inégalité (D.4) pour N égal à 2. ■

Fonction d'autocovariance de la magnitude renormalisée

On note

$$\rho_T(t) = \lambda^2 \ln \left(\frac{T}{|t|} \right) \mathbf{1}_{\{|t| \leq T\}}, \quad (\text{E.1})$$

la limite des fonctions d'autocovariance (3.122), utilisée pour la construction de la mesure MRM log-normale, lorsque l tend vers zéro.

La fonction d'autocovariance de la magnitude renormalisée, introduite dans la section 5.3, est donnée par l'expression suivante

$$\rho_\Omega(\tau, \Delta_1, \Delta_2) = \text{Cov}[\Omega(\Delta_1) - \Omega(0), \Omega(\tau + \Delta_2) - \Omega(\tau)] = \int_0^{\Delta_1} du \int_\tau^{\tau + \Delta_2} dv \rho_T(u - v). \quad (\text{E.2})$$

Il est simple de voir que la fonction d'autocovariance (E.2) de la magnitude renormalisée est proportionnelle au coefficient d'intermittence λ^2 , alors pour alléger les calculs, sans perdre en généralité, on peut considérer λ^2 égal à 1.

Cas $\Delta_1 = \Delta_2 = \Delta \leq T$

$$\rho_\Omega(0, \Delta_1, \Delta_2) = \Delta^2 \ln \left(\frac{T e^{3/2}}{\Delta} \right). \quad (\text{E.3})$$

Cas $\Delta_1 \leq \Delta_2 \leq T$

$$\rho_{\Omega}(0, \Delta_1, \Delta_2) = \frac{\Delta_1^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_1} \right) + \frac{\Delta_2^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_2} \right) - \frac{(\Delta_2 - \Delta_1)^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_2 - \Delta_1} \right) \quad (\text{E.4})$$

Cas $\Delta_1 \leq T \leq \Delta_2 \leq T + \Delta_1$

$$\rho_{\Omega}(0, \Delta_1, \Delta_2) = \frac{H^2}{2} \ln \left(\frac{T e^{3/2}}{H} \right) - \frac{(\Delta_2 - T)^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_2 - T} \right) + T \Delta_2 - \frac{T^2}{4}. \quad (\text{E.5})$$

Cas $\Delta_1 \leq T \leq T + \Delta_1 \leq \Delta_2$

$$\rho_{\Omega}(0, \Delta_1, \Delta_2) = \frac{\Delta_1^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_1} \right) + T \Delta_1. \quad (\text{E.6})$$

Cas $T \leq \Delta_1 = \Delta_2 = \Delta$

$$\rho_{\Omega}(0, \Delta_1, \Delta_2) = 2T\Delta - \frac{T^2}{2}. \quad (\text{E.7})$$

Cas $T \leq \Delta_1 \leq \Delta_2$

$$\rho_{\Omega}(0, \Delta_1, \Delta_2) = T\Delta_1 + T\Delta_2 - \frac{T^2}{2} - \frac{(\Delta_2 - \Delta_1)^2}{2} \ln \left(\frac{T e^{3/2}}{\Delta_2 - \Delta_1} \right). \quad (\text{E.8})$$

Cas $\tau \neq 0$

$$\rho_{\Omega}(\tau, \Delta_1, \Delta_2) = \rho_{\Omega}(0, \Delta_1, \tau + \Delta_2) - \rho_{\Omega}(0, \Delta_1, \tau). \quad (\text{E.9})$$

Fonction d'autocovariance de magnitude

Nous rappelons la définition de la fonction $\rho(t)$

Définition F.1

La fonction $\rho(t)$ est définie comme une limite

$$\rho(t) = \lim_{l \rightarrow 0} \rho_l(t) = \begin{cases} \lambda^2 \ln\left(\frac{T}{t}\right), & \text{si } 0 \leq t < T, \\ 0, & \text{si } T \leq t, \end{cases} \quad (\text{F.1})$$

où la fonction $\rho_l(t)$ est définie par (3.122)

Proposition F.1

La transformée de Fourier $\hat{\rho}(\omega)$ de la fonction $\rho(t)$

$$\hat{\rho}(\omega) = \int_{-\infty}^{\infty} \rho(t) e^{i\omega t} dt \quad (\text{F.2})$$

est réelle et positive et elle est une fonction paire.

Preuve - F.1 -

La fonction $\rho(t)$ est paire $\rho(-t) = \rho(t)$ alors sa transformée de Fourier est réelle. Considérons la partie réelle de la transformée de Fourier de la fonction $\rho(t)$

$$\hat{\rho}(\omega) = \Re[\hat{\rho}(\omega)] = \int_{-\infty}^{\infty} \rho(t) \cos(\omega t) dt = 2 \int_0^T \ln\left(\frac{T}{t}\right) \cos(\omega t) dt \quad (\text{F.3})$$

en appliquant l'intégration par partie, nous avons

$$\hat{\rho}(\omega) = \Re[\hat{\rho}(\omega)] = \frac{2}{\omega} \int_0^T \frac{\sin(\omega t)}{t} dt = \frac{2}{\omega} \int_0^{\omega T} \frac{\sin(t)}{t} dt = \frac{2}{\omega} Si(\omega T) \geq 0, \quad (\text{F.4})$$

où nous avons utilisé le fait que la fonction $Si(x)$ s'annule seulement en zéro [AS65].

De plus, il est simple de voir que la transformée de Fourier vérifie

$$\hat{\rho}(-\omega) = \hat{\rho}(\omega), \quad (\text{F.5})$$

et donc elle est paire. ■

Proposition F.2

Soient t_1, \dots, t_n réels, alors la matrice Σ définie comme

$$(\Sigma)_{jk} = \rho(t_j - t_k) \quad (\text{F.6})$$

est symétrique semi-définie positive, c'est-à-dire que pour tout vecteur $a = (a_1, \dots, a_n)$, le produit $a^T \Sigma a$ est non négatif.

De plus, la fonction $\rho(t)$ peut être représentée comme une convolution

$$\rho(t) = \int_{-\infty}^{\infty} \tilde{\rho}(t-s)\tilde{\rho}(s)ds. \quad (\text{F.7})$$

Preuve - F.2 -

Nous pouvons définir la fonction $\tilde{\rho}(t)$ comme

$$\tilde{\rho}(t) = \int_{-\infty}^{\infty} \hat{\rho}^{1/2}(\omega) e^{-i\omega t} \frac{d\omega}{2\pi}. \quad (\text{F.8})$$

puisque la transformée de Fourier $\hat{\rho}(\omega)$ est positive. La valeur en $t_j - t_k$ de la fonction $\rho(t)$ peut être alors représentée sous forme de la convolution

$$\rho(t_j - t_k) = \int_{-\infty}^{\infty} \tilde{\rho}(t_j - s)\tilde{\rho}(s - t_k)ds, \quad (\text{F.9})$$

puisque la transformée de Fourier d'une convolution est égale au produit des transformées de Fourier

$$\widehat{f \circ g}(\omega) = \hat{f}(\omega)\hat{g}(\omega). \quad (\text{F.10})$$

La matrice Σ peut être donc représentée comme

$$\Sigma = \int_{-\infty}^{\infty} R(s)R^T(s)ds, \quad (\text{F.11})$$

où le vecteur $R(s)$ de dimension n est défini par

$$R(s) = \begin{pmatrix} \tilde{\rho}(t_1 - s) \\ \vdots \\ \tilde{\rho}(t_n - s) \end{pmatrix}. \quad (\text{F.12})$$

Pour tout vecteur $a = (a_1, \dots, a_n)$ le produit $a^T \Sigma a$ est donné par

$$a \Sigma a^T = \int_{-\infty}^{\infty} a^T R(s) R^T(s) a ds = \int_{-\infty}^{\infty} (a^T R(s))^2 ds \geq 0. \quad (\text{F.13})$$

■

Lemme auxiliaire

Lemme G.1

Soit X une variable aléatoire d -dimensionnelle telle que les moments $m^{(k)}(p)$ (k entier) des produits scalaires (p, X) sont finis et vérifient la condition limite

$$\lim_{k \rightarrow \infty} \frac{1}{k!} m^{(k)}(p) = 0, \text{ pour tout } p \in \mathbb{R}^d. \quad (\text{G.1})$$

Si $\{X_j\}_j$ une suite de variables aléatoires d -dimensionnelles telles que les moments $m_j^{(k)}(p)$ des produits scalaires (p, X_j) sont finis. Alors

$$\lim_{j \rightarrow \infty} m_j^{(k)}(p) = m^{(k)}(p) \implies X_j \xrightarrow{\mathcal{L}} X. \quad (\text{G.2})$$

Preuve - G.1 -

Cette démonstration est une généralisation multidimensionnelle du cas mono-dimensionnel traité dans [Chu74].

Pour p fixé, le développement en série de Taylor de $e^{i(p,x)}$ autour de zéro s'écrit

$$e^{i(p,x)} = \sum_{k=0}^{2n-1} \frac{i^k (p,x)^k}{k!} + \frac{i^{2n} (p,x)^{2n}}{(2n)!} e^{i(p,x)\xi}, \text{ avec } \xi \in [0, 1]. \quad (\text{G.3})$$

En conséquence, nous avons

$$\mathbb{E}\left[e^{i(p,X_j)}\right] = \sum_{k=0}^{2n-1} \frac{i^k}{k!} \mathbb{E}\left[(p, X_j)^k\right] + \frac{i^{2n}}{(2n)!} \mathbb{E}\left[(p, X_j)^{2n} e^{i(p,x)\xi}\right], \text{ avec } \xi \in [0, 1]. \quad (\text{G.4})$$

La différence des fonctions caractéristiques est donnée par

$$\left| \mathbb{E}\left[e^{i(p,X_j)}\right] - \mathbb{E}\left[e^{i(p,X)}\right] \right| \leq \sum_{k=0}^{2n-1} \frac{|m_j^{(k)}(p) - m^{(k)}(p)|}{k!} + \frac{m_j^{(2n)}(p) + m^{(2n)}(p)}{(2n)!}. \quad (\text{G.5})$$

Etant donné $\epsilon > 0$, il existe un nombre entier $n = n(\epsilon)$ tel que

$$\frac{2m^{(2n)}(p) + 1}{(2n)!} \leq \frac{\epsilon}{2}. \quad (\text{G.6})$$

Etant donné $n(\epsilon)$, il existe un nombre entier $j_0 = j_0(\epsilon)$ tel que, pour tout $j \geq j_0$

$$m_j^{(2n)}(p) \leq m^{(2n)}(p) + 1, \quad (\text{G.7})$$

et de même

$$\max_{0 \leq k \leq 2n-1} \left| m_j^{(k)}(p) - m^{(k)}(p) \right| \leq \frac{\epsilon}{2e}. \quad (\text{G.8})$$

Finalement, la partie gauche de l'équation (G.5) peut être majorée par

$$\left| \mathbb{E} \left[e^{i(p, X_j)} \right] - \mathbb{E} \left[e^{i(p, X)} \right] \right| \leq \frac{\epsilon}{2e} \sum_{k=0}^{2n-1} \frac{1}{k!} + \frac{(2m^{(2n)}(p) + 1)}{(2n)!} \leq \epsilon, \quad (\text{G.9})$$

et donc la suite $\{X_j\}_j$ converge en loi vers X . ■

Article

L'article [MBK06] qui suit porte sur le problème de l'estimation de la probabilité d'observation des événements extrêmes dans des réalisations de certains processus multifractals. Il est consacré également à l'étude du comportement des distributions de probabilité à queues épaisses associées à telles données. Le résultat crucial de cet article est que, dans le cadre des processus multifractals, l'approche classique des valeurs extrêmes n'est pas valide, à cause de corrélations à longue portée présentées dans ces processus. Le calcul quantitatif, qui fait appel à des outils de l'analyse multifractal, et notamment au spectre des singularités, justifié par les expériences numériques, permet de montrer que les valeurs des exposants de queues obtenues à l'aide des estimateurs classiquement utilisés (estimateur de Zipf, Hill ou Pickands) peuvent être différentes à celles théoriques. Les applications aux séries financières permettent d'expliquer l'écart observé entre les valeurs des exposants des queues indiquées par plusieurs travaux économétriques et les valeurs théoriques obtenues dans le cadre du modèle MRW log-normal.

Il s'agit du premier article publié dans le cadre de cette thèse. Plusieurs des résultats de cet article sont présentés au chapitre 4.

APS/123-QED

Extreme values and fat tails of multifractal fluctuations

J.F. Muzy*

*SPE UMR 6134, CNRS, Université de Corse, 20250 Corte, France*E. Bacry[†] and A. Kozhemyak[‡]*CMAP, Ecole Polytechnique, 91128 Palaiseau, France*

(Dated: June 23, 2006)

Abstract

In this paper we discuss the problem of the estimation of extreme event occurrence probability for data drawn from some multifractal process. We also study the heavy (power-law) tail behavior of probability density function associated with such data. We show that because of strong correlations, standard extreme value approach is not valid and classical tail exponent estimators should be interpreted cautiously. Extreme statistics associated with multifractal random processes turn out to be characterized by non self-averaging properties. Our considerations rely upon some analogy between random multiplicative cascades and the physics of disordered systems and also on recent mathematical results about the so-called multifractal formalism. Applied to financial time series, our findings allow us to propose an unified framemork that accounts for the observed multiscaling properties of return fluctuations, the volatility clustering phenomenon and the observed “inverse cubic law” of the return pdf tails.

PACS numbers: 0.250.-r, 05.40.-a, 05.45.-a, 89.65.Gh

*Electronic address: muzy@univ-corse.fr

[†]Electronic address: emmanuel.bacry@polytechnique.fr

[‡]Electronic address: alexey@cmplx.polytechnique.fr

I. INTRODUCTION

Statistics of extremes is an issue of prime importance in many situations where extreme events may appear to have disastrous effects or to govern the main observations. Such situations can be found in a wide range of fields from physics (e.g. disordered systems at low temperature), geology (earthquakes), meteorology (rainfalls, storms), insurance or finance [1–3]. Extreme events are particularly relevant for random phenomena involving a probability density function (pdf) which tails decrease very slowly and roughly follow a power-law. Such heavy tailed distributions have been observed in many natural phenomena. An important question concerns therefore the estimation and interpretation of pdf tail exponent as well as the identification of mechanisms leading to them. These problems are at the heart of an increasing number of works [3–5]. Probabilistic and statistical questions related to very high or very low values of random variables are addressed within the framework of extreme value theory. This theory has been originally developed for independent identically distributed (i.i.d.) random variables and more recently extended to stationary processes where independence condition has been relaxed [1]. However, when correlations are not weak enough very few results are known.

In this paper we aim at studying the statistics of extreme events and the (fat) tail exponent of fluctuations associated with multifractal random processes. Multifractal extreme fluctuations are interesting because they represent an example of strongly correlated random variables that do not satisfy standard mixing conditions of extreme value theory. But multifractals are also interesting because they are widely used to model of self-similar phenomena displaying multiscaling properties. Our purpose is to study probability of extreme event of multifractal processes and the associated power-law tail exponent. We will show that for such processes, the pdf tail exponent value observed (estimated) from experimental data may be different from the value associated with the unconditional theoretical pdf. We examine different experimental conditions depending upon the size of the observed sample L , the correlation length T and the observation scale τ . We emphasize that, under most usual conditions, the estimated tail exponent is smaller than the exponent one would expect without correlations. This result is intimately related to some non-self averaging property of usual tail exponent estimators and is the analog of the glassy behavior observed at low temperatures in disordered systems. We apply our phenomenological framework to multi-

fractal models of asset return fluctuations and show that the well known “inverse cubic law” of pdf tails can be naturally explained in terms of volatility correlations.

The paper is organized as follows: In section II, we briefly review the main results about extreme value theory and the commonly used tail exponent estimators. Multifractal processes, multiplicative cascades and their main mathematical properties are reviewed in section III. In sections IV we build an extreme value theory for multifractal cascades. The problem of the estimation of the power-law tail exponent associated with the cumulative probability distribution of multifractal fluctuations is addressed in section V. In section VI we illustrate our phenomenology by numerical examples of continuous as well as discrete cascades. Application to finance is considered in section VII while section VIII contains concluding remarks and questions for future research. Auxiliary computations or technical material are reported in Appendices.

II. FAT TAILS AND EXTREME VALUE STATISTICS

Let us briefly review the main estimators used to characterize the power-law tail behavior of some probability distribution. Let $F(x) = \mathbb{P}[Z \leq x]$ be the cumulative probability distribution (cdf) of some random variable Z . The variable Z is said to be of power-law type tail or Pareto type tail if, when $x \rightarrow +\infty$,

$$1 - F(x) \sim Cx^{-\mu} \quad (1)$$

where C is a positive normalization constant or a slowly varying function. The exponent μ is called the *tail exponent* of the distribution. The problem addressed in this section concerns the estimation of this exponent from empirical data.

A simple, widely used method relies on the so-called “Zipf” or “rank-frequency” plots (see e.g. [3, 5, 6]): Let $Z_1 \dots Z_N$ be N i.i.d. samples characterized by the same distribution function $F(x)$. Let us denote $X_1 \geq X_2 \geq \dots \geq X_N$ the rank ordered values of Z_i (sorted in descending order). Then, if the asymptotic behavior of $F(x)$ is Pareto like (as in Eq. (1)), one has, for $1 \leq j \leq k \ll N$:

$$1 - F(X_j) \sim \frac{j}{N} \implies X_j \sim \left(\frac{j}{N}\right)^{-1/\mu} \quad (2)$$

and therefore μ can be simply estimated as the slope of the *Zipf plot* $[\ln X_j, \ln(j)]$, $j = 1 \dots k$

(see [6] for exact results). In the following, we refer to this estimator as the *power-law fit estimator*. Since this estimator is biased [5], one should use alternative estimators.

Actually, there are many alternative tail exponent estimators [2]. The most commonly used are Hill or Pickands estimators that are defined as follows: Let $k(N) = o(N)$ be the maximum X rank used to estimate μ , the Hill estimator is simply

$$\mu_H(k, N) = \frac{k-1}{\sum_{i=1}^{k-1} \ln(X_i/X_k)} \quad (3)$$

while the Pickands estimator is:

$$\mu_P(k, N) = \frac{\ln(2)}{\ln\left(\frac{X_k - X_{2k}}{X_{2k} - X_{4k}}\right)} \quad (4)$$

The mathematical study of these estimators (consistency, bias, asymptotic normality) relies upon Extreme Value Theory [1, 2], i.e., the theory that deals with maxima and minima properties of random variables. According to this theory, the maximum value Y of N i.i.d. random variables (normalized properly), has a probability density function that asymptotically belongs to the Fisher-Tippett's Extreme Value distribution class. According to the shape of the pdf of Y , the pdf of the maximum can either be of the Frechet type ($F_Y(x) = e^{-x^{-\mu}}$), of the Gumbel type ($F_Y(x) = e^{-e^{-x}}$) or of the Weibull class [1].

Extreme value theory has also been extended to dependent (or correlated) stationary random processes [1]. Under mixing conditions ensuring asymptotic independence of maxima, it can be shown that the limit theorems established in the i.i.d. case still hold. The key difference is that the number N of independent variables is replaced by an "effective" number $N\theta$. The value $0 < \theta < 1$ that quantifies the effect of dependence is called the *extremal index* [1]. In the case of Gaussian processes, these theorems hold provided the covariance function $\rho(x)$ decreases sufficiently fast for large lags x , i.e.

$$\rho(x) \ln(x) \xrightarrow{x \rightarrow +\infty} 0 \quad (5)$$

A simple intuitive justification of $1/\ln(x)$ as the limiting case for the validity of standard extreme value theorems is provided in ref. [7].

The main purpose of this paper is to try to understand how these theoretical results apply to data sampled from a multifractal process.

III. MULTIFRACTAL PROCESSES

Multifractal processes are random functions that possess non trivial scaling properties. They are now widely used models in many areas of applied and fundamental fields. Well known examples are turbulence, internet traffic, rainfall distributions or finance. For the sake of simplicity we will consider only non-decreasing multifractal processes (often referred to as multifractal *measures* though their variations are not bounded) denoted hereafter $M(t)$. More general multifractal processes can be conveniently builded as a simple Brownian motion $B(t)$ compounded with the measure $M(t)$ considered as a stochastic time: $X(t) = B[M(t)]$. The statistical properties of $X(t)$ can be directly deduced from those of $M(t)$ (see e.g., [8, 9]).

A. Multiscaling

Multifractal processes are characterized by multiscaling properties of their variations. More precisely, if one defines the increments of $M(t)$ at scale τ , $M(t, \tau) = M(t + \tau) - M(t)$, multifractality can be loosely defined from the scaling behavior of the moments of $M(t, \tau)$:

$$\mathbb{E}[M(t, \tau)^q] \underset{\tau \rightarrow 0}{\sim} \tau^{\zeta(q)} \quad (6)$$

where $q \in \mathbb{R}$ is the order of the moment and the exponent $\zeta(q)$ is some nonlinear convex function often called the multifractal exponent spectrum. The simplest example of such function is the so-called *log-normal* spectrum for which $\zeta(q)$ is a simple parabola:

$$\zeta(q) = \left(1 + \frac{\lambda^2}{2}\right)q - \frac{\lambda^2}{2}q^2 \quad (7)$$

The coefficient λ^2 quantifies the curvature of $\zeta(q)$ (and hence the multifractality of the process) that is constant in the log-normal case. In the general case, one often calls $-\zeta''(0)$ the *intermittency coefficient*. Let us notice that, because of Hölder inequality for moments the scaling (6) with a non linear convex $\zeta(q)$ cannot hold at arbitrary large scales τ but is valid only in a domain bounded by some large scale T (actually the limit $\tau \rightarrow 0$ in Eq. (6) must be understood as $\tau/T \ll 1$). This scale T will be called the *integral scale*.

B. Singularity spectrum and multifractal formalism

Let us recall some classical results about the multifractal formalism. This formalism has been introduced in early eighties by Parisi and Frisch (see e.g. [10, 11]) in order to interpret

the above multiscaling properties of the moments in terms of pointwise regularity properties of the paths of the process $M(t)$. Let us define the local Hölder exponent $\alpha(t_0)$ at point (or time) t_0 as

$$M(t_0, \tau) \underset{\tau \rightarrow 0}{\sim} \tau^{\alpha(t_0)} \quad (8)$$

The limit $\tau \rightarrow 0$ means $\tau \ll T$ where T is the integral scale. The singularity spectrum $f^*(\alpha)$ can be introduced as the fractal (Hausdorff or packing) dimension of the iso-Hölder exponents sets:

$$f^*(\alpha) = \text{Dim}\{t, \alpha(t) = \alpha\} \quad (9)$$

Roughly speaking, this equation means that at scales $\tau \ll T$, the number of points where $M(t, \tau) \sim \tau^\alpha$ is

$$N(\tau, \alpha) \sim \tau^{-f^*(\alpha)}. \quad (10)$$

The *multifractal formalism* states that $f^*(\alpha)$ and $\zeta(q)$ as defined in Eq. (6) are basically Legendre transform one to each other. More precisely, if we define $f(\alpha)$ as the Legendre transform of $\zeta(q)$, i.e.,

$$\begin{aligned} f(\alpha) &= 1 + \min_q (q\alpha - \zeta(q)) \\ \zeta(q) &= 1 + \min_\alpha (q\alpha - f(\alpha)), \end{aligned}$$

then

$$f^*(\alpha) = f(\alpha), \quad \forall \alpha^* \geq \alpha \geq \alpha_*$$

where α_* and α^* are defined by

$$\begin{aligned} \alpha_* &= \inf\{\alpha, f^*(\alpha) = 0\} \\ \alpha^* &= \sup\{\alpha, f^*(\alpha) = 0\}. \end{aligned}$$

In the following sections, we will use the fact that q can be interpreted as a value of the derivative of $f(\alpha)$ and conversely α is a value of the derivative of $\zeta(q)$: for a given value of $q = q_0$ one has, from previous Legendre transform relationship and thanks to the convexity of $\zeta(q)$:

$$f(\alpha_0) = 1 + q_0\alpha_0 - \zeta(q_0) \quad (11)$$

$$\alpha_0 = \frac{d\zeta}{dq}(q_0) \quad (12)$$

$$q_0 = \frac{df}{d\alpha}(\alpha_0). \quad (13)$$

Let us note that, $f^*(\alpha)$ can be seen as the Legendre transform of the function $\zeta^*(q)$ simply defined as

$$\zeta^*(q) = \begin{cases} \zeta(q) & \text{for } q^* \leq q \leq q_* \\ \alpha_* q & \text{for } q > q_*, \\ \alpha^* q & \text{for } q < q^*, \end{cases}, \quad (14)$$

where

$$q_* = \frac{df}{d\alpha}(\alpha_*) \quad (15)$$

$$q^* = \frac{df}{d\alpha}(\alpha^*) \quad (16)$$

It is important to point out that, experimentally, under usual conditions, only $\zeta^*(q)$ (and not $\zeta(q)$) can be estimated (see e.g. refs. [12–14]).

C. Cascades

The paradigm of multifractal measures are multiplicative cascades originally introduced by the russian school for modelling the energy cascade in fully developed turbulence. After the early works of Mandelbrot [15–17], a lot of mathematical studies have been devoted to discrete random cascades [12, 18–21]. Very recently, continuous versions of these cascades have been defined: they share exact multifractal scaling with discrete cascades but they display continuous scaling and possess stationary increments [8, 9, 22, 23]. Let us summary the main properties of these constructions and set some notations.

The simplest discrete multifractal cascade can be constructed as follows: one starts with an interval of length T where the measure is uniformly spread, and split the interval in two equal parts: On each part, the density is multiplied by (positive) i.i.d. random factors W . Each of the two sub-intervals is again cut in two equal parts and the process is repeated infinitely. At construction step n , if one addresses a dyadic interval of length $T2^{-n}$ by a kneading sequence $k_1 \dots k_n$, with $k_i = 0, 1$, the “mass” of this interval (denoted as $I_{k_1 \dots k_n}$) is simply:

$$M_n(I_{k_1 \dots k_n}) = 2^{-n} \prod_{i=1}^n W_{k_1 \dots k_i} = 2^{-n} e^{\sum_{i=1}^n \omega_{k_1 \dots k_i}} \quad (17)$$

where all the $W_{k_1 \dots k_i} = e^{\omega_{k_1 \dots k_i}}$ are i.i.d such that $\mathbb{E}[W] = 1$. Peyrière and Kahane [18] proved that this construction converges almost surely towards a stochastic non decreasing

process M_∞ provided $\mathbb{E}[W \ln W] < 1$. The multifractality of M_∞ (hereafter simply denoted as M) and the validity of the previously described multifractal formalism have been studied by many authors (see e.g. [12]). An interesting additional property of cascades is that they are self-similar in the following stochastic sense:

$$M[I_{k_1 \dots k_n}] \underset{\text{law}}{=} 2^{-1} W M[I_{k_1 \dots k_{n-1}}] \quad (18)$$

and therefore the order q moments of $M(t, \tau)$ behave as a power-law:

$$\mathbb{E}[M[0, T2^{-n}]^q] = 2^{-nq} \mathbb{E}[W^q]^n \mathbb{E}[M[0, T]^q] \quad (19)$$

Comparison of Eqs. (19) and (6) with $\tau = T2^{-n}$ directly yields the expression of the spectrum $\zeta(q)$ in terms of cumulant generating function of $\omega = \ln W$:

$$\zeta(q) = q - \ln_2(\mathbb{E}[W^q]) \quad (20)$$

In that sense, $\zeta(q)$ is nothing but a *large deviation* spectrum. Let us mention that the validity of the multifractal formalism has been rigorously proved for cascades. In Appendix A we provide explicit expressions of $\zeta(q)$ for various laws of W .

Let us note that, as shown in [24, 25], the correlation function $\rho(x)$ of such cascades decreases slowly as

$$\rho(x) \simeq \ln(T/x) \quad \text{for lags } x \leq T. \quad (21)$$

If we consider that the data come from a sampling of successive independent cascades, the correlation function is 0 for lags above T , i.e.,

$$\rho(x) = 0, \quad x \geq T. \quad (22)$$

The integral scale T where cascading process “starts” can therefore be interpreted as a *correlation length* for the variations of the $M(t)$.

Because the previous construction involves dyadic intervals, and a ‘top-bottom’ construction, it is far from being stationary. In order to get rid of this drawback, as already mentioned, some continuous cascade constructions have been recently proposed and studied on a mathematical ground [8, 9, 22, 23]. Without entering into details, we just want to mention that such continuous cascades involve a family of infinitely divisible random processes $\omega_l(t)$ which correlation function basically follows Eqs (21) and (22). The process $e^{\omega_l(t)}$ is the analog of the density satisfying the self-similarity:

$$e^{\omega_{sl}(st)} \underset{\text{law}}{=} e^{\Omega_s} e^{\omega_l(t)}$$

Martingale theory allows one to prove the convergence of the continuous process $M[0, t] = \lim_{t \rightarrow 0} \int_0^t e^{\omega_l(u)} du$.

Let us mention that the validity of the multifractal formalism has been established by Molchan by for discrete cascades [12, 20] and generalized by Barral and Mandelbrot for continuous cascades [23].

In the sequel, we will be using indifferently the classical top-bottom model (i.e., we will consider that the data come from a sampling of a succession of independent realizations of the same cascade process) and the continuous cascade model (i.e., we will consider that the data simply come from a sampling of a continuous cascade). Most of the arguments will be done using the first model while numerical examples will be performed on data of the second model.

D. Cascades have fat tails

Let us first emphasize that the unconditional law of M can have a power law tail. Indeed, a simple argument involving the self-similarity of the limit measure allows one to obtain a simple bound: since the measure is additive, $M[0, 1] = M[0, 1/2] + M[1/2, 1]$, for $q > 1$ one gets

$$\mathbb{E}[M[0, 1]^q] \geq 2\mathbb{E}[M[0, 1/2]^q] = 2^{1-\zeta(q)}\mathbb{E}[M[0, 1]^q] \quad (23)$$

and finally $\mathbb{E}[M[0, 1]^q] < +\infty \Rightarrow \zeta(q) \geq 1$. The reverse implication is basically true. It is however trickier to obtain and we refer to [18] for a precise proof.

Since the power law tail exponent of a distribution is directly related to the maximum order finite moments, if one defines

$$\mu = 1 + \sup\{q, q > 1, \zeta(q) > 1\}, \quad (24)$$

then, cascades (discrete as well as continuous) have thick tails of exponent μ :

$$\mathbb{P}[M(t, \tau) \geq x] \underset{x \rightarrow +\infty}{\sim} x^{-\mu}. \quad (25)$$

E. Defining the asymptotic limit $N \rightarrow +\infty$

As we have seen (Eq. (21)), the covariance function $\rho(x)$ for multifractal processes has a very slow decay (slower than Eq. (5)) up to the *integral scale* T above which data are

independent. Thus we expect that mixing conditions are not valid in the range from the sampling scale τ to the integral scale T while they hold when “looking” above scale T . Let us try to be more precise: the mixing conditions are conditions in the limit when the total number of samples N goes to $+\infty$. However, there are several ways to reach this asymptotic limit. Let L be the length of the whole sequence and τ the sampling scale. The total number of samples is therefore,

$$N = \frac{L}{\tau},$$

while the number of integral scales N_T and the number of samples per integral scales N_τ are respectively:

$$N_T = \frac{L}{T} \quad , \quad N_\tau = \frac{T}{\tau}, \quad N = N_\tau N_T. \quad (26)$$

In order to control the relative values of N_T and N_τ , let us define the exponent χ as follows:

$$N_T \sim N_\tau^\chi \quad (27)$$

Let us note that this exponent χ has been already introduced by B.B. Mandelbrot as an “embedding dimension” [26] in order to discuss the concept of negative dimension (see below).

Thus, when $N \rightarrow +\infty$, if, for instance, $\chi = 0$, it means that we are in the case $\tau \rightarrow 0$ (while L and T are fixed), i.e., most of the data are lying between the lags τ and T . Consequently, we do not expect the mixing conditions to hold. On the contrary, if $\chi = +\infty$, it means that we are in the case $L \rightarrow +\infty$ (while τ and T are fixed), i.e, most of the data are lying between the lags T and L and consequently the mixing conditions are satisfied.

Thus, as it will be discussed in the next two sections, in the first case ($\chi = 0$), nothing guarantees that classical results of extreme value theory can be applied and that exponent tail estimators provide the expected values, whereas in the second case ($\chi = +\infty$), we expect the i.i.d. extreme value theorems to hold and the exponent tail estimators to be consistent. As we will see, one can go continuously from the first case to the second one. Actually we will show that both the extreme value distribution associated with cascades and the corresponding tail exponent estimator strongly depend on the value of χ .

IV. MULTIFRACTAL EXTREME VALUE STATISTICS

A. Cumulative probability distribution of the maximum

Let $\tau = T2^{-n}$. We call $X_1(N)$ the maximum value of $\ln(M(I_n)/\tau)$, where I_n is a short notation of the dyadic intervals $I_{k_1\dots k_n}$ of size τ . Let $P(x, N) = \mathbb{P}[X_1(N) < x]$ be the cumulative distribution function (cdf) of $X_1(N)$, i.e., the probability that $X_1(N)$ is smaller than x .

Let us recall that we consider that the data come from a sampling of successive independent realizations of the same cascade measure. Thus T (the integral scale) is fixed whereas L (the total length of the data) and τ (the sampling scale) are varying. We want to study the statistics $P(x, N)$ of the data in the limit $N \rightarrow +\infty$. As we explained in the previous section, these statistics will strongly depend on χ (Eq. (27)), i.e., on the way N_τ and N_T go to $+\infty$. We fix $\chi = r/p$ and we choose the following parametrization:

$$N_\tau \sim 2^{pm}, \quad N_T \sim 2^{rm}, \quad N = N_\tau N_T \sim 2^{(p+r)m},$$

when the integer parameter $m \rightarrow +\infty$.

Thus $P(x, m) \equiv P(x, N(m))$ is the cdf of $\ln M(I_{pm})/\tau$ where $\tau = T2^{-pm}$.

In Appendix B, we show that $P(x, m)$ can be simply expressed as:

$$P(x, m) = (P'(x, pm))^{2^m} \quad (28)$$

where $P'(x, n)$ is the cdf associated with the maximum of $\ln(M(I_n)/\tau)$ on a single integral scale instead of the whole data. It satisfies the renormalization equation:

$$P'(x, n+1) = [P'(x, n) \star g(x)]^2 \quad (29)$$

where \star stands for the convolution product and $g(x)$ is the probability density of $\omega = \ln(W)$.

Let us notice that the initial condition $P(x, 0)$ is precisely given by the law of $\ln(M)$ which exponential tail is described by Eq. (25):

$$P(x, 0) \underset{x \rightarrow +\infty}{\sim} 1 - Ce^{-\mu x} \quad (30)$$

B. Traveling front solutions

1. Case $\chi = +\infty$

When $r = 1$ and $p = 0$, Eq. (28) is simply the recurrence for the maximum cdf of i.i.d. random variables which solutions are Fisher-Tippett's fixed points reviewed in sec. II. More precisely, since the initial condition is exponentially decreasing (Eq. (30)), when $N \rightarrow +\infty$, $P(x, N)$ will have a Gumbel shape. Consequently, the law of the maximum value of M will belong to the Frechet class with a tail exponent μ as defined in Eq. (24).

2. Eq. (29) and the KPP equation

In order to solve the nonlinear problem (28) in the general case, let us first study Eq. (29) and its solutions. We will show that these solutions are traveling fronts so that, in some moving frame, Eq. (28) reduces to a standard (i.i.d.) extreme value problem.

Equation (29) is exactly satisfied by the cdf associated with the maximum value of random variables hierarchically correlated (generated additively along a Cayley tree). Such an equation has been studied in ref. [27, 28]. In ref. [7], the authors have considered a log-normal random process with log-correlated covariance, that can be considered as a continuous version of a random cascade. In this case, the obtained partial differential equation for the law of the maximum turns out to be the famous Kolmogorov-Petrovsky-Piscounov (KPP) equation. Thus Eq. (29), in the case where $g(x)$ is Gaussian, can be seen as a "discretized" version of the KPP equation.

It is well known that the KPP equation has traveling wavefront solutions connecting the homogeneous stable state to the unstable one. These solutions can be studied using linear stability analysis. As emphasized in [29], most of KPP features are somehow universal in the sense that the same analysis can be performed for a wide variety of "reaction-diffusion" problems involving non linear integro-differential or integro-difference equations. In particular the famous marginal stability criterium (see e.g. refs. [29–31] and below) for the selected front velocity can be generically applied for a large class of equations. Therefore, as explained in [28, 30, 31], one can apply the same techniques to get solutions of Eq. (29). Though these studies do not rely on a fully rigorous mathematical ground (as KPP), we reproduce a similar analysis in the following.

Let us first notice that both Eqs (28) and Eqs (29) admit two homogeneous solutions $P = 0$ and $P = 1$, the first one being (linearly) stable while the second one is unstable. A given cdf $P(x, N)$ (or $P'(x, n)$) will therefore connect the stable state to the unstable one. As in the above cited references, one can consider a traveling front solution of (29) $P'(x, n) = P'_{tw}(x - v'n)$ where v' is the front velocity. In order to compute this velocity, one performs a linear analysis in the vicinity of the unstable solution, i.e., in the limit $x \rightarrow +\infty$ where $P'_{tw}(x - v'n) \rightarrow 1$.

If we denote

$$Q'_{tw}(x) \equiv 1 - P'_{tw}(x) \ll 1$$

then, to the first order in $Q'_{tw}(x)$, Eq. (29) becomes :

$$Q'_{tw}(x - v') = 2Q'_{tw}(x) * g(x) + O(Q'^2_{tw})$$

If one seeks for exponential solutions:

$$Q'_{tw}(x) = Ce^{-qx} ;$$

then,

$$P'_{tw}(x, n) = 1 - Ce^{-q(x-v'(q)n)} , \quad (31)$$

where v' and q satisfy the “dispersion” relation:

$$v'(q) = q^{-1} \ln \left[2 \int g(x)e^{qx} dx \right]$$

Notice that $g(x)$ is the law of the logarithms of the weights of the cascade construction, and then, from Eq. (20)

$$\ln(2 \int g(x)e^{qx} dx) = \ln(2) (1 + q - \zeta(q))$$

which yields:

$$v'(q) = \frac{\ln(2) (1 + q - \zeta(q))}{q} \quad (32)$$

In the next section, one shall see which velocity (or which q value) is selected

3. Case $\chi \neq +\infty$

From Eqs. (28) and (31), one gets the following traveling front solution for P ,

$$P(x, m) = (1 - Ce^{-q(x-v'(q)pm)})^{2^m} ,$$

where $v'(q)$ is given by (32).

If $y = x - v'(q)pm - q^{-1}r \ln(2)m$, the cdf of y converges to the Gumbel shape:

$$P(y, m) = (1 - Ce^{-qy - r \ln(2)m})^{2^{rm}} \xrightarrow{m \rightarrow +\infty} e^{-Ce^{-qy}}$$

and thus $P(x, m)$ is itself a traveling Gumbel front:

$$P(x, m) = e^{-Ce^{-q(x - mv(q))}} \quad (33)$$

with the dispersion relation:

$$v(q) = \frac{\ln(2)(r + p(1 + q - \zeta(q)))}{q} \quad (34)$$

One can reproduce the same kind of analysis as in refs [28, 30–33]. In Appendix C, we provide a sketch of proof that one can use a standard Aronson-Weinberger stability criterium to compute the velocity and q value which are selected: Let q_{min} be the unique positive q value such that

$$v(q_{min}) = \min_{q>0} v(q) \quad (35)$$

and let

$$q_{*,\chi} = \min(\mu, q_{min}). \quad (36)$$

Then the selected velocity is simply $v(q_{*,\chi})$ and the shape of the traveling front is (up to a subdominant correction described in Appendix C)

$$P(x, m) = e^{-Ce^{-q_{*,\chi}(x - v(q_{*,\chi})m)}}$$

Let us recall (see section IV A) that $P(x, m)$ corresponds to the cdf of $\ln M(I_n)/\tau$ where $\tau = T2^{-n}$ and $n = pm$. Let us go back to $\ln M$.

4. From $\ln M(I_n)/\tau$ to $\ln M(I_n)$

The velocity $v(q)$ is a velocity related to the parameter m , i.e., $P(x, m) = P(x - v(q)m)$. It is convenient to compute the velocity $v_\tau(q)$ as respect to the “observable” scale $\ln(T/\tau)$, i.e., $P(x, m) = P(x - v_\tau(q) \ln(T/\tau))$. Clearly, $v_\tau(q) = v(q)m / \ln(T/\tau)$. Since $\tau = T2^{-n}$ and $n = pm$, one gets $v_\tau(q) = v(q)/p \ln(2)$. Now, if we switch from the cdf of $\ln M(I_n)/\tau$ to the cdf of $\ln M(I_n)$, one finally obtains the velocity:

$$v_{\ln M}(q) \equiv v(q)/p \ln(2) - 1 = -\frac{\zeta(q) - 1 - \chi}{q}$$

Let us note that the minimum of $v_{\ln M}(q)$ (i.e., the minimum of $v(q)$) is reached for q_{min} satisfying

$$-\chi = 1 + q \frac{d\zeta(q_{min})}{dq} - \zeta(q_{min}).$$

One can also notice that $1 + q \frac{d\zeta(q)}{dq} - \zeta(q)$ is the Legendre transform of $\zeta(q)$. We know from section III B that this Legendre transform is nothing but the spectrum $f(\alpha)$. According to the results of this section (and particularly Eq. (11)) it can be easily seen that this value of q_{min} corresponds to a singularity exponent α_{min} with

$$f(\alpha_{min}) = -\chi \quad (37)$$

$$q_{min} = \frac{df}{d\alpha}(\alpha_{min}) \quad (38)$$

$$\alpha_{min} = v_{\ln M}(q_{min}). \quad (39)$$

Moreover, since $f(\alpha_{min}) \leq 0$ then $df(\alpha_{min})/d\alpha > 0$ and consequently $q_{min} > 0$. Thus, this value of q_{min} does correspond to the one defined in Eq. (35).

We finally established that the law of the maximum of M at scale τ of a sample of length $\tau^{-\chi}$ is Fréchet when $\tau \rightarrow 0$ with a tail exponent $q_{*,\chi}$ that depends on χ :

$$\begin{aligned} q_{*,\chi} &= \min(q_{min}, \mu) \\ q_{min} &= \frac{df}{d\alpha}(\alpha_{min}) \\ f(\alpha_{min}) &= -\chi \end{aligned} \quad (40)$$

Let us remark that for $\chi = 0$, the value of $q_{*,\chi}$ is exactly the value q_* that is involved in the multifractal formalism as described at the end of section III B (Eq. (15)). For finite χ value, as already noticed more than a decade ago by Mandelbrot [26], the exponent χ that governs the size of the ‘‘supersample’’ allows one to explore negative dimensions (i.e., negative values of $f(\alpha)$).

Numerical evidence of the so-obtained results are reported in section VI A.

V. ORDER STATISTICS, TAIL EXPONENT ESTIMATION AND MULTIFRACTAL FORMALISM

A. Notations

The idea underlying tail exponent estimation is to study statistics of the $k(N)$ observed extreme values in the asymptotic regime $N \rightarrow +\infty$ and $k(N) \rightarrow +\infty$. As in the previous section, the limit $N \rightarrow +\infty$ is taken as explained in section III E, using the χ exponent (see Eqs (26) and (27)).

For the $k(N) \rightarrow +\infty$ limit, it is convenient to parametrize $k(N)$ as:

$$k(N) \sim N^\nu \quad (41)$$

where $0 \leq \nu \leq 1$, the value $\nu = 0$ being interpreted as $k(N) \sim \ln(N)$. We will denote $\hat{\mu}(\nu, \chi)$ the estimated tail exponent using one of the estimators reviewed in section II for some given χ and ν (χ is defined in the section III E). At scale $\tau = T2^{-n}$, as in section II, we denote $X_1 \dots X_N$ the rank ordered values of $M(I_n)$ over dyadic intervals.

B. Tail exponent estimators

Let us compute the expected value of the tail exponent estimator (Power-law fit, Pickands or Hill estimator) for fixed values of ν and χ . For the sake of simplicity we focus on Pickands estimator but the same argument equally applies to other estimators (e.g. Hill). Our heuristics will rely upon the multifractal formalism.

As recalled in section III B, the positive part of the Legendre transform of $\zeta(q)$ corresponds to the singularity spectrum, i.e., the Hausdorff dimension of sets of iso-regularity. Mandelbrot has proposed a long time ago, a probabilistic interpretation of *negative dimensions* (i.e., negative values of $f(\alpha)$) in terms of large deviation spectrum [26]. In this section, we will use this formalism in order to study the order statistics of multifractal fluctuations.

The Pickands estimator is simply defined as:

$$\hat{\mu}(\nu, \chi) = \ln(2) \left(\ln \frac{X_k - X_{2k}}{X_{2k} - X_{4k}} \right)^{-1} \quad (42)$$

Let us define $\alpha_{\nu, \chi}$ such that

$$k = N^\nu \sim N_T \left(\frac{\tau}{T} \right)^{-f(\alpha_{\nu, \chi})}$$

Using the definition of χ according to which $N_T \sim (\tau/T)^{-\chi}$, it follows that $\alpha_{\nu,\chi}$ satisfies

$$f(\alpha_{\nu,\chi}) = \nu - \chi(1 - \nu) \quad (43)$$

Let us now consider $\alpha'_{\nu,\chi} = \alpha_{\nu,\chi} + \epsilon_1$ such that

$$2k \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha'_{\nu,\chi})}$$

i.e., $2k \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha_{\nu,\chi}) + \epsilon_1 q_{\nu,\chi}}$, thus $2^{-1/q_{\nu,\chi}} \sim \left(\frac{\tau}{T}\right)^{\epsilon_1}$ where

$$q_{\nu,\chi} = \frac{df}{d\alpha}(\alpha_{\nu,\chi}) . \quad (44)$$

Along the same line if $\alpha''_{\nu,\chi} = \alpha_{\nu,\chi} + \epsilon_2$ such that

$$4k \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha''_{\nu,\chi})}$$

we have $4^{-1/q_{\nu,\chi}} \sim \left(\frac{\tau}{T}\right)^{\epsilon_2}$. Thanks to the fact that,

$$\begin{aligned} X_k - X_{2k} &\sim X_k (1 - (\tau/T)^{\epsilon_1}) \\ X_{2k} - X_{4k} &\sim X_k ((\tau/T)^{\epsilon_1} - (\tau/T)^{\epsilon_2}) \end{aligned}$$

one finally gets for expression (42):

$$\hat{\mu}(\nu, \chi) \simeq \ln(2) \left(\ln \frac{1 - 2^{-1/q_{\nu,\chi}}}{2^{-1/q_{\nu,\chi}} - 4^{-1/q_{\nu,\chi}}} \right)^{-1} = q_{\nu,\chi} \quad (45)$$

We then come to the conclusion that the Pickands tail estimator for a multifractal process strongly depends on the choice of the rank $k = N^\nu$ and the exponent χ . This is another difference with standard theory for i.i.d. random variables. The same kind of phenomenology can also be applied to the Power-law fit or Hill estimator. Let us notice that when $\nu = 0$ we consider only the 'extreme' of the 'extremes' for the tail exponent estimation and therefore we recover the tail of the law of the maximum value as discussed previously. Indeed, from Eqs. (43), (44) and (45), we get

$$\begin{aligned} \hat{\mu}(0, \chi) &= f'(\alpha_{0,\chi}) \\ f(\alpha_{0,\chi}) &= -\chi , \end{aligned}$$

that is, exactly Eqs. (40).

Numerical evidence illustrating these results are provided in the next section.

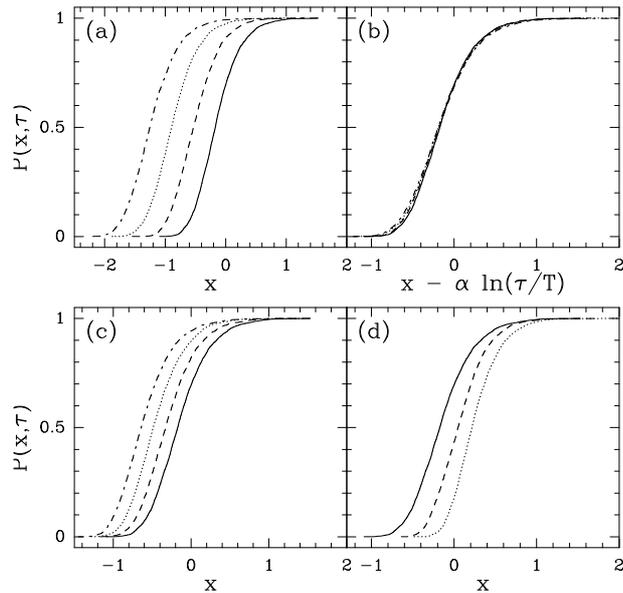


FIG. 1: Cumulative probability distribution of the maximum of $\ln(M)$ at scale $\tau = T2^{-n}$, $P(x, n)$ (also referred to as $P(x, \tau)$), as a function of x for various values of τ for a continuous log-normal cascade with $\lambda^2 = 0.2$. (a) $N_T = 1$ and $\ln_2(T/\tau) = 3$ (continuous line), 4, 5, 6 (dotted lines). According to the formalism of section IV, the cdf should be a front moving toward $x < 0$ at a “velocity” $\alpha_{*,\chi=0} = \alpha_* \simeq 0.47$. (b) All the cdf of Fig. (a) merge to a single curve when plotted in the “moving referential”. (c) Same plot as Fig. (a) but with $\chi = 1$. One expects a smaller velocity $\alpha_{*,\chi=1} \simeq 0.2$. (d) Same plot as figs. (a) and (c) for scales $\ln_2(T/\tau) = 3, 4, 5$ with $\chi = 3$. One observes, as expected, a negative velocity $\alpha_{*,\chi=3} \simeq -0.16$.

VI. NUMERICAL EXAMPLES

Most of our considerations rely on phenomenological scaling and asymptotic limit arguments. We therefore neglected prefactors and slowly varying corrections, i.e., finite size effects. In order to test our approach as well as to quantify the importance of finite-size corrections, it is therefore interesting to perform numerical simulations. In this section, we illustrate our purpose on specific examples such as those mentioned in Appendix A. As already explained, most of these simulations are performed on continuous cascades which construction algorithm is described in ref. [8].

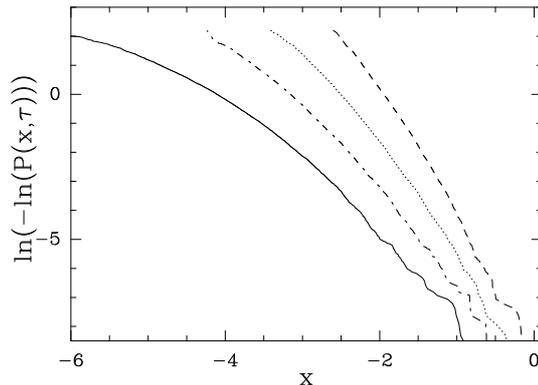


FIG. 2: “Gumbel plots” of the cumulative probability distribution of the maximum of $\ln(M)$ at scale $\tau = T2^{-n}$, $P(x, n)$ (also referred to as $P(x, \tau)$) for different values of χ . One sees that as χ increases from 0 to 1.5, $q_{*,\chi}$, the slope in the tail increases significantly.

A. Extreme value statistics

In order to check the results of section IV, we have generated $N = 5000$ independent realisations of log-normal continuous cascades of intermittency parameter $\lambda^2 = 0.2$ for various values of the parameters N_τ and N_T . From these independent samples, we have estimated the cdf of the maximum value of $\ln(M)$ at scale $\tau = T2^{-n}$. In Fig. 1(a) the cdf $P(x, n)$ (also referred to as $P(x, \tau)$) is plotted for $N_T = 1$ and $T/\tau = 8, 16, 32, 64$. According to Eqs. (33) and (40), we expect to observe, as the scale $\ln(\tau/T)$ decreases, a front traveling towards negative x at velocity $\alpha_{*,\chi=0} = \alpha_*$. This behavior is well verified and can be quantitatively checked in Fig. 1(b) where all the fronts merge when plotted versus $x - \alpha_* \ln(\tau/T)$ with $\alpha_* \simeq 0.47$ as given according to Eqs. (40) and (7):

$$\alpha_* = 1 + \frac{\lambda^2}{2} - \lambda \sqrt{2(1 + \chi)}.$$

In Fig. 1(c) the same analysis is performed for $N_T = T/\tau$ and therefore $\chi = 1$. One observes that the velocity decreases as expected (in that case $\alpha_{*,\chi=1} \simeq 0.2$). For χ large enough, $\alpha_{*,\chi}$ can become negative as illustrated in Fig. 1(c) where we have reproduced the plot of Fig. 1(a) for $\tau = 8, 16, 32$ and $\chi = 3$. According to previous formula, $\alpha_{*,\chi=3} \simeq -0.16$, a value compatible with observations where one sees the front moving towards positive x values as the scale decreases. The fact that $\alpha_{*,\chi}$ is negative for χ large enough can be easily understood as follows: because the measure is continuous, as τ decreases, the maximum is

expected to go to zero, but if in the same time, the number of independent integral scales is increased, the maximum is expected to increase. The exponent χ controls the balance between these two opposite effects. For χ large enough one explores negative dimensions that can be associated with negative α 's. Notice that in all plots one observes a slight change in the shape of the front as the scale goes to zero: this is not surprising because the asymptotic shape of the front depends on χ and is a priori different from the initial front at scale $T/\tau = 8$ considered in Fig. 1.

According to Eq. (33), the asymptotic shape of the fronts should be Gumbel, i.e.,

$$P(x) = e^{-Ce^{-q_{\star,\chi}x}}$$

with parameter $q_{\star,\chi}$ for the log-normal model:

$$q_{\star,\chi} = \sqrt{\frac{2(1+\chi)}{\lambda^2}}.$$

Notice that $q_{\star,\chi}$ is an increasing function of χ . In Fig. 2, we have plotted $\ln(-\ln(P(x)))$ versus x for various values of $T/\tau = N_\tau$ and N_T . We have chosen the values $(N_\tau = 512, N_T = 1)$, $(N_\tau = 512, N_T = 8)$, $(N_\tau = 512, N_T = 64)$ and $(N_\tau = 64, N_T = 512)$ that correspond respectively (if one neglects prefactors in our analysis) to $\chi = 0$, $\chi = 0.33$, $\chi = 0.67$, $\chi = 1.5$. For an exact Gumbel law, one would expect a straight line. One clearly observes strong deviations to the Gumbel shape because the asymptotic regime is not reached but the tails look straight and are very well estimated by the theoretical $q_{\star,\chi}$ values associated with the values of χ : one sees a systematic increase of $q_{\star,\chi}$ when one goes from $\chi = 0$ to $\chi = 1.5$.

B. Tail exponent estimators

Let us now check our results on the tail behavior of the estimated pdf from measure samples.

Let us first compute $\hat{\mu}_{\nu,\chi}$ as given by Eq. (45), i.e. solve Eqs. (43) and (44) for respectively log-Normal, log-Poisson and log-Gamma examples. After some simple algebra, we find in the log-normal case

$$\hat{\mu}_{\ln}(\nu, \chi) = \sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}} \quad (46)$$

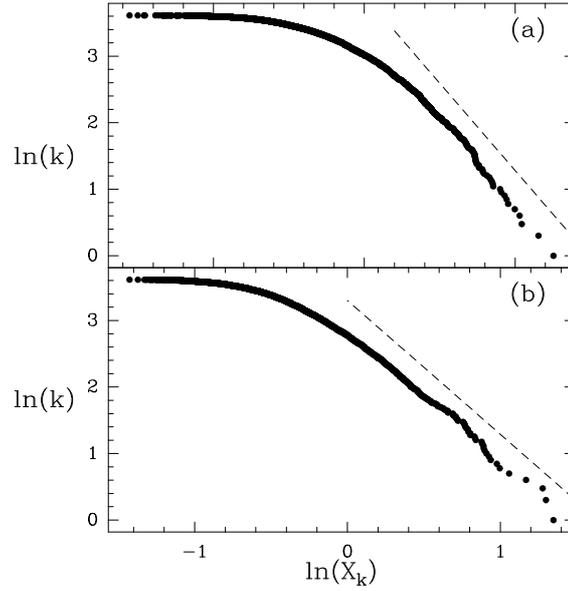


FIG. 3: Rank-Frequency (Zipf) plots in log-log representation for a log-Normal cascade (top) and a log-Gamma cascade (bottom). The slope of the right linear part provides an estimate of the tail exponent μ . Dashed lines indicate analytical expectations (see text).

Let us notice that for $\nu = 0$, one recovers previous tail exponent of the extreme values $\hat{\mu}(0, \chi) = q_{*, \chi}$. In the case of log-Poisson statistics (Eq. (A3)), the computation leads to an expression for $\hat{\mu}(\nu, \chi)$ that involves Lambert $W(x)$ function. A perturbation series in the limit $\delta \rightarrow 0$ of this expression gives:

$$\hat{\mu}_{lp}(\nu, \chi) = \sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}} + 2\delta \frac{(1+\chi)(\nu-1) - 1}{3\lambda^2} + \dots \quad (47)$$

Along the same line, the value of $\hat{\mu}(\nu, \chi)$ can be computed in the case of the log-gamma cascade (Eq. (A4)) which again involves the second branch of the Lambert function W_{-1} :

$$\hat{\mu}_{lg}(\nu, \chi) = \beta \left[1 - e^{1+W_{-1}\left(-e^{-1+\frac{(1+\chi)(\nu-1)}{\lambda^2\beta^2}}\right) + \frac{(1+\chi)(1-\nu)}{\lambda^2\beta^2}} \right] \quad (48)$$

A series expansion in the limit $\beta \rightarrow +\infty$ gives:

$$\hat{\mu}_{lg}(\nu, \chi) = \sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}} + 4 \frac{(\nu-1)(1+\chi)}{3\beta\lambda^2} + \dots \quad (49)$$

In Fig. 3, the rank ordering of the a log-normal and a log-Gamma (with $\beta = 4$) cascade processes with an intermittency coefficient $\lambda^2 = 0.2$ are plotted in doubly logarithmic scale

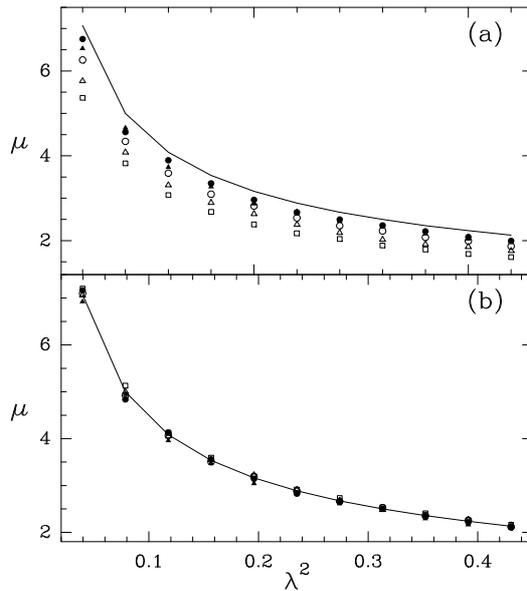


FIG. 4: Mean estimated tail exponent as a function of the intermittency coefficient λ^2 for continuous log-normal cascades. Each mean value is computed using 1000 cascade samples of length L . (a) Power-law fit estimator as a function of λ^2 for $T = 1024$, $L = 4096$, $k = 16$ (\bullet), $k = 32$ (\circ), $k = 64$ (\triangle), $k = 128$ (\square) and $T = 512$, $L = 8192$, $k = 32$ (\blacktriangle). The continuous line corresponds to the theoretical prediction $\mu_{ln} = \sqrt{2/\lambda^2}$ (Eq. (46)). (b) The same as in (a) but each curve has been rescaled by a factor $\sqrt{2(1-\nu)(1+\chi)}$ according to Eq. (46) (see text).

(such plots are often referred to as “rank-frequency” plots or “Zipf” plots [3, 5]). One clearly sees that the rightmost part of each distribution behaves as a power law. From Eqs. (46) and (48) (with $\chi = 0$ and $\nu = 0$), the slope of the plots should be respectively $\mu \simeq 3.1$ and $\mu \simeq 2.0$. One sees that these behaviors, reported on the figures as dashed lines, fit the data relatively well. One can observe that the the scaling range associated with the log-Gamma cascade is wider than for the log-Normal measure. Indeed, according to Appendix D (see Eq. (D1)), in the log-Normal case, this range should be around $0.2pq_* \ln(T/\tau) \simeq 0.7$ for $p = 0.1$ while its values for log-Gamma is expected to be $1.13pq_* \ln(T/\tau) \simeq 2$ for $p = 0.1$, i.e., more than two times wider than for the log-Normal case. This difference can be visually checked in Fig. 3.

In Fig. 4, we have plotted the mean value of the tail exponent estimator (power-law fit estimator) as a function of the intermittency parameter λ^2 in the case of a log-Normal

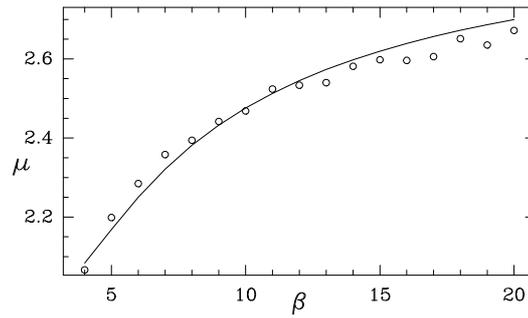


FIG. 5: Mean estimated (Power-law fit) tail exponent (\circ) as a function of β for discrete log-Gamma cascades with $\lambda^2 = 0.2$ and $T = 2048$. The Hill parameter is $k = 32$ and each mean value is computed from 1000 cascade samples. The continuous line corresponds to the analytical expression (48)

(continuous) cascade for various values of k , N_T and N_τ . The theoretical prediction (46) for $\nu = 0$ is shown in continuous line for comparison. The mean has been evaluated on 10^3 independent cascade samples. In Fig. 4(a), one can see that for all parameter values, the curves have the same decreasing behavior. All these curves collapse on the theoretical prediction $\sqrt{2/\lambda^2}$ as shown in Fig. 4(b), if one rescales each one by a factor $\sqrt{(1-\nu)(1+\chi)}$ where ν and χ are computed from expressions (41) and (27) by assuming that the prefactors are trivially 1. Empirically we find that the asymptotic phenomenology works quite well and prefactors or slowly varying behavior are negligible. As previously emphasized, the exact computation of prefactor values and finite size effects is beyond the scope of this paper and should involve more sophisticated mathematical tools (see appendix C).

In Fig. 5, we compare the expression (48) to the behavior of the tail exponent estimator as a function of β in the case of a discrete log-Gamma cascade: in that case, λ^2 is fixed ($\lambda^2 = 0.2$) while β varies. Once again, we can see that if we take into account finite k value through the value of ν and χ close their expected values, the analytical expression provides a very good fit of the data.

We have therefore illustrated, on two specific examples, that the phenomenology developed in previous section allows us to predict with a relative precision the tail behavior of multifractal measure samples. Let us now show how financial time series fluctuations can be described within this framework.

VII. APPLICATION TO FINANCIAL DATA

Multifractal models are, with many regards, well suited to account for return fluctuation of financial assets [6, 34]. Among the “stylized facts” characterizing the asset return fluctuations, the phenomenon of “volatility clustering” (called heteroskedasticity in econometrics) is the most important one. One of the key points raised in refs. [22, 35] is that these volatility correlations are found empirically very close to the “log-correlations” of continuous multifractal cascades [24]. Therefore the observed multiscaling properties of returns can be simply explained in terms of volatility persistence. In this section we want to stress that, for the same reason, namely the logarithmic shape of log-volatility correlations, the pdf of return fluctuations appear empirically as fat tailed with a rather small tail exponent.

In ref. [22, 35], we have shown that a parcimonious model of $X(t)$, some asset return value at time t , can be constructed as follows:

$$X(t) = B(M(t)) \tag{50}$$

where $B(t)$ is the standard Brownian motion and $M(t)$ is a multifractal continuous cascade as defined in refs. [8, 9]:

$$M(t) = \lim_{l \rightarrow 0} \int_0^t e^{\omega_l(t)} dt .$$

The process $\omega_l(t)$ plays exactly the same role as $\omega = \ln(W)$ in discrete cascades. It is easy to see that the variations of $M(t)$ in (50) can be interpreted as a stochastic variance. This quantity is called the “volatility” in finance [34].

If $M(t)$ has multiscaling (e.g. log-normal) properties then so do $X(t)$. Empirically, it has been determined using data from several markets, from various countries, that the intermittency coefficient characterizing the multifractal statistics of the volatility is close to $\lambda^2 = 0.2$ while the integral scale T is typically around 1 year [22]. On the other hand, many studies relying on high frequency data or on thousands of daily stock returns, have revealed that the financial return pdf have heavy tails with a tail exponent μ in the interval [3, 5]. This is the famous “inverse cubic” law for return fluctuations [34, 36]. This observation lead to one of the main objections raised against the previous multifractal model for asset returns [34]. Indeed, the *unconditional pdf* of the volatility associated with a log-normal multifractal cascades of coefficient $\lambda^2 = 0.15$ has a tail exponent $\mu \approx 13$ (Eq. (A2)). Within the “subordinated” model (50), this would mean that the tail of the volatility pdf is around

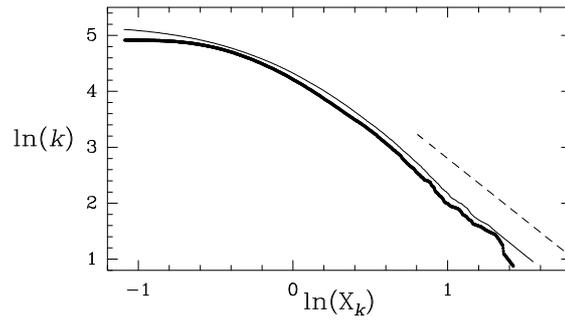


FIG. 6: Rank-Frequency plot of CAC40 daily volatility estimates (\bullet) as compared to similar plot for a log-Gamma continuous cascade with $T = 253$ days, $\lambda^2 = 0.05$ and $\beta = 4$ (thin line). The fit of the extreme tail provides an estimation $\mu \simeq 2$ (dashed line).

$2\mu \approx 26$, i.e., close to ten times the observed value ! However, the main message of this paper is that for multifractal fluctuations, the observed extreme events are far from being distributed as they were independent. In particular we have shown in section V that the estimators of the pdf tail exponent strongly depend on ν and χ and are “generically” smaller than μ . Typically, in finance $\tau \simeq 10^{-2} - 1$ day, $T \simeq 1 - 2$ years and $L \simeq 10$ years. Therefore, a rough estimate of χ and ν can be $\chi \simeq \nu \simeq 0.5$. If one uses these values in Eq. (46), one finds a typical value for the estimated tail within the log-normal model that is $2\hat{\mu} \simeq 6$. This is a value closer to the observations. In order to have a better fit of the tail behavior one could use a log-gamma model with the same intermittency coefficient and $\beta = 4$ (see Appendix A). In that case, Eq. (48) gives an estimator value $2\hat{\mu} \simeq 3.6$ that agrees with observations.

In Fig. 6 is reported a rank-frequency plot of estimated daily volatilities associated with the 40 stock values composing to the French CAC 40 index. The data are daily ‘open’ ‘high’ ‘low’ ‘close’ quotes over a mean period of 10 years. The daily volatilities are estimated using the widely used Garman-Klass method [37] and each volatility sample mean has been normalized to 1. For comparison the Zipf plot associated with a multifractal log-Gamma measure with $T = 1$ year, $\lambda^2 = 0.2$ and $\beta = 4$ has been also reported. One can see that both curves behave very similarly with a power-law tail exponent $\mu \simeq 2$. Let us note that for small volatilities values, the behavior of CAC40 volatility pdf is slightly different from the log-Gamma cascade probably because of the high frequency noise in the Garman-Klass volatility estimates. This figure illustrates very well our result: there is no discrepancy

between the value of the intermittency coefficient and the estimated pdf power-law behavior.

VIII. SUMMARY AND PROSPECTS

In this paper we have addressed the problem of extreme value statistics for multifractal processes. This problem is non trivial and possesses a rich phenomenology involving non ergodic behavior. In the case of multifractal processes, two important parameters govern the asymptotics: the overall sample length and the scale at which data are sampled. The exponent χ we have introduced, precisely quantifies how one defines the asymptotic limit as one changes these two parameters. The observed extreme value statistics result from a “competition” between an increase of the number of independent samples (which tends to increase typical extreme values) and a decrease of the sampling scale (which tends to decrease typical observed values). Consequently, the law of extremes continuously depends on χ , an exponent that turns out to be interpreted as a *negative dimension*.

This exponent naturally plays an important role when one wants to estimate the tail of the probability law associated with multifractal fluctuations. Using the phenomenology of the multifractal formalism we have shown that tail exponent estimators continuously depend on χ . They also depend on another exponent ν that quantifies “how many” extremes one uses for estimation. Under usual experimental conditions we have notably shown that the obtained exponent is usually smaller than the exponent expected from the unconditional law. Such non ergodic behavior are similar to those observed in the thermodynamics of disordered systems, at low temperature, under the freezing transition [7, 27].

The arguments and methods we used in this paper are mostly phenomenological and rely upon large deviation type arguments, multifractal formalism and traveling front solution of non linear iteration equations. Beyond the need for sitting it on rigorous bases, they are other appealing mathematical prospects raised by our approach, such as the possibility to address finite size effects in multifractal scaling laws or to define a precise statistical framework for studying them.

As far as applications are concerned, we have provided a direct use of our results in the field of econophysics where multifractal models for asset returns are popular. We have shown that the observed fat tails of return pdf are well reproduced by a multifractal model designed to account for the volatility clustering phenomenon. Other fields where multifractal

processes are involved can be potentially investigated along the same line. Conversely and perhaps more importantly, within this framework, multifractality appears as an alternative that can be invoked to explain the origin of fat tails as observed in many fields of applied science [3–5].

Acknowledgments

We thank J.P. Bouchaud for helpful discussions about the analogy of some multifractal problems addressed in this paper and similar problems arising in physics of disordered systems. This work has been supported by contract ACI *nouvelles interfaces des mathématiques*, Nb 54-03, from French ministère de l'éducation nationale.

APPENDIX A: LOG-NORMAL, LOG-POISSON AND LOG-GAMMA MULTIFRACTAL MEASURES

In this appendix we provide 3 examples of multifractal statistics to which we will refer all along the paper. We consider the 3 infinitely divisible laws of $\ln W$ (or equivalently ω): normal, Gamma and Poisson.

In the simplest case, $\omega = \ln W$ is normal of variance $\lambda^2/\ln(2)$ (and mean $-\lambda^2/2\ln(2)$ because $E(W) = 1$). Then $\mathbb{E}[W^q] = e^{-q\lambda^2/2 + \lambda^2 q^2/2}$ and we recover the expression of Eq. (7):

$$\zeta_{\ln}(q) = q(1 + \lambda^2/2) - \lambda^2 q^2/2 \quad (\text{A1})$$

The intermittency coefficient in this case is simply λ^2 . If one solves $\zeta(q) = 1$, $q > 1$ one gets

$$\mu = \frac{2}{\lambda^2}. \quad (\text{A2})$$

In the second example, n is a Poisson random variable of intensity $\gamma \ln(2)$ and $\omega = m_0 \ln(2) + n\delta$. Then $q - \ln_2 \mathbb{E}[e^{q\omega}] = q(1 - m_0) + \gamma(1 - e^{q\delta})$. If one sets $\zeta(1) = 1$ and $-\zeta''(0) = \lambda^2$, one obtains the spectrum:

$$\zeta_{\text{lp}}(q) = q \left(1 + \frac{\lambda^2}{\delta^2} (e^\delta - 1) \right) + \frac{\lambda^2}{\delta^2} (1 - e^{q\delta}) \quad (\text{A3})$$

Notice that when $\delta \rightarrow 0$, $\zeta_{\text{lp}} \rightarrow \zeta_{\ln}$. It is easy to show that, when δ is negative and small enough, for all $q > 1$, $\zeta_{\text{lp}}(q) > 1$ and therefore $\mu = +\infty$ in that case.

In the third example, ω is Gamma distributed: if x is a random variable of pdf $\beta^{\alpha \ln(2)} x^{\alpha \ln(2)-1} e^{-\beta x} / \Gamma(\alpha \ln(2))$ and $\omega = x + m_0 \ln(2)$, then $q - \ln_2 \mathbb{E}[e^{q\omega}] = q(1 - m_0) + \alpha \ln(1 - q/\beta)$. By setting $\zeta(1) = 1$ and $-\zeta''(0) = \lambda^2$, we have:

$$\zeta_{\text{lg}}(q) = q \left(1 - \lambda^2 \beta^2 \ln\left(\frac{\beta - 1}{\beta}\right) \right) + \lambda^2 \beta^2 \ln\left(\frac{\beta - q}{\beta}\right) \quad (\text{A4})$$

Notice that $\zeta_{\text{lg}} \rightarrow \zeta_{\text{ln}}$ when $\beta \rightarrow +\infty$. The solutions of $\zeta_{\text{lg}}(q) = 1$ can be obtained in terms of Lambert W function (satisfying $W(x)e^{W(x)} = x$) and therefore the value of μ can be exactly computed as a function of λ^2 and β .

APPENDIX B: PROOF OF EQ.28

In this appendix we prove Eq. (28). Let us first study how the law of the maximum of $\ln(M(I_n)/\tau)$ inside one integral scale varies when on changes the scale $\tau = T2^{-n+1}$ to $\tau = T2^{-n}$. Let $I_n(k)$, $k = 0 \dots 2^n - 1$ denote the dyadic intervals of size $T2^{-n}$ of the interval $[0, T]$. From the cascade construction, $\forall n$, the following stochastic equality can be easily proven:

$$\begin{aligned} M[I_n(k)] &\stackrel{fdd}{=} \frac{W_1}{2} M_1(I_{n-1}(k)), \quad k \in [0, 2^{n-1}) \\ M[I_n(k)] &\stackrel{fdd}{=} \frac{W_2}{2} M_2(I_{n-1}(k - 2^{n-1})), \quad k \in [2^{n-1}, 2^n] \end{aligned}$$

where W_1, W_2, M_1 and M_2 are independent copies of W and M respectively. The symbol *fdd* means an equality in law for all finite dimensional distributions. Therefore, if

$$X(n) = \max_k [\ln(M(I_n(k))/\tau)]$$

One has

$$X(n) \stackrel{law}{=} \max[\ln(W_1) + X_1(n-1), \ln(W_2) + X_2(n-1)]$$

If $g(x)$ denote the law of $\omega = \ln(W)$ and $P'(x, n) = \mathbb{P}[X(n) > x]$, the previous equality can be rewritten as

$$\begin{aligned} P'(x, n) &= \left[\int P'(z - x, n-1) g(z) dz \right]^2 \\ &= [g \star P'(x, n-1)]^2 \end{aligned}$$

Now if one has $N_T = 2^{r^m} > 1$ integral scales, the cdf of the maximum is given by (28).

APPENDIX C: ARONSON-WEINBERGER CRITERIUM AND FINITE SIZE EFFECTS ON FRONT SOLUTIONS OF EQ. (28)

In this Appendix we provide some additional technical details on the solutions of Eq. (28). We do not establish rigorous proofs but mostly recast some results from refs. [28–33] to our problem. If one linearizes Eq. (28) in the tail region $x \rightarrow +\infty$, one obtains the following recursion for $Q(x) = 1 - P(x, m)$:

$$Q(x, m + 1) = 2^{p+r} Q(x, m) * g^{(p)}(x) \quad (\text{C1})$$

where $g^{(p)}(x)$ is simply the product of p convolutions $g(x) * \dots * g(x)$. If one decomposes $Q(x, m)$ on Fourier modes:

$$\hat{Q}(k, m) = \int e^{-ikx} Q(x, m) dx$$

then (C1) becomes:

$$\hat{Q}(k, m + 1) = 2^{p+r} \hat{Q}(k, m) e^{pF(ik)}$$

where $F(ik) = \ln \mathbb{E} [e^{-ik\omega}]$ is the cumulant generating function of $\omega = \ln(W)$ the logarithm of cascade weights. The solution is therefore

$$\hat{Q}(k, m) = A(k) e^{m \ln(2) [r + p(1 + F(ik)/\ln(2))]}$$

where $A(k)$ is simply the Fourier transform of the initial condition. $Q(x, m)$ is obtained as the inverse Fourier transform:

$$Q(x, m) = (2\pi)^{-1} \int e^{ikx} A(k) e^{m \ln(2) [r + p(1 + F(ik)/\ln(2))]} dk$$

In a referential moving at velocity v , (i.e. $x_m = x_0 + vm$), the previous integral can be computed using a steepest descent method: One deforms the integral over the real axis to a contour in the complex plane of constant phase. In the limit $m \rightarrow +\infty$, the main contribution comes from saddle point of the function $F(ik)$ along this path:

$$Q(x_m, m) \sim A(k_*) e^{ik_* x_m + m \ln(2) [r + p(1 + F(ik_*)/\ln(2))]}$$

where k_* satisfies:

$$-ik_* v = \left. \frac{dF(ik)}{dk} \right|_{k=k_*} \quad (\text{C2})$$

Moreover, if the front is stationary in the moving frame, the selected velocity should be such that the real part of the exponent is zero:

$$\Re \left\{ ik_{\star}mv + m \ln(2) \left[r + p \left(1 + \frac{F(ik_{\star})}{\ln(2)} \right) \right] \right\} = 0 \quad (\text{C3})$$

and thus, if v is real, by setting $k = iq$, and thanks to the equality

$$F(-q) = \ln(2)(q - \zeta(q))$$

one can rewrite respectively Eqs. (C3) and (C2) as follows:

$$v(q) = \frac{\ln(2)[r+p(1+q-\zeta(q))]}{q}$$

$$\frac{\partial v(q)}{\partial q} \Big|_{q=q_{\star,\chi}} = 0$$

The first equation is the dispersion relationship (34) while the second is the standard Aronson-Weinberger criterium stating that the selected velocity is the minimum velocity, a velocity that corresponds to the marginally stable solution (Eq. (35)).

Notice that previous argument assumes that both $F(ik)$ and $A(k)$ are analytical functions. If the initial condition decreases exponentially, i.e., $A(k)$ has a complex pole which imaginary part is μ (as in Eq. (30)), then, as discussed in [29], $q_{\star,\chi}$ must be replaced by $\min(q_{\star,\chi}, \mu)$. This yields Eq. (36). As far as $F(ik)$ is concerned, it is easy to show that it is analytical in a strip around the real axis $\{z, \Im(z) \leq \mu\}$.

Let us finally remark that a general solution traveling at velocity v of Eq. (C1) can be written as:

$$Q(z_m) = A_1 e^{-q_1 z_m} + A_2 e^{-q_2 z_m}$$

where $z_m = z_0 + vm$ and q_1 and q_2 are the two complex conjugated roots of the dispersion relation (34). When $v = v(q_{\star,\chi})$, the selected minimum velocity, these two roots merge and generically the solution behaves as [7, 29, 30]:

$$Q(z_m) = (Az_m + B)e^{-q_{\star,\chi} z_m}$$

We see that the asymptotic shape of the front is not precisely Gumbel but has a subdominant correction factor.

Notice that the matching of the previous functional shape and the shape obtained in the previous saddle point analysis, when one accounts for the Gaussian corrections in the

integration around the saddle point, leads to famous Bramson logarithmic correction to the front velocity [7, 29, 30]:

$$x_m = mv(q_{*,\chi}) - \frac{3}{2q_{*,\chi}} \ln(m)$$

This “universal” logarithmic correction to the velocity is well known for solutions of KPP equation. A specific analysis of these finite-size corrections to scaling is beyond the scope of this paper and will be reported in a forthcoming study.

APPENDIX D: SCALING RANGE FOR TAIL ESTIMATION

One question that naturally arises from the analysis made in section V is the question of stability of $\hat{\mu}(\nu, \chi) = q_{\nu, \chi}$ as a function of ν (and thus of k) and χ . This question is linked to the question of the scaling range associated with the log-log representation of Eq. (2). This problem is important for practical purpose. Let $p = \Delta q/q$ be the precision above which one can detect a tail exponent variation (for the sake of simplicity we will consider $p \simeq 0.1$) Let $\ln(S)$ be the scaling range over which one observes the power-law. According to our description, this corresponds to a variation of the Hölder exponent, i.e.,

$$\ln(S) \simeq \Delta\alpha \ln(T/\tau)$$

This variation corresponds therefore to a variation of q that is

$$\Delta q \simeq \Delta\alpha \frac{\partial q}{\partial h(q)} = \frac{\Delta\alpha}{\zeta''(q)|_{q=q_{\nu, \chi}}}$$

and finally, if the scaling range is such that $\Delta q = pq$, one obtains

$$\ln(S) = pq_{\nu, \chi} \ln(T/\tau) \zeta''(q_{\nu, \chi}) \quad (\text{D1})$$

We see that this scaling range is related to the global scaling range rescaled by the factor that involves the intermittency coefficient evaluated at the value $q = q_{\nu, \chi}$.

-
- [1] M. R. Leadbetter, G. Lingren, and H. Rotzén, *Extremes and related properties of random sequences and processes* (Springer Verlag, Berlin, 1983).
- [2] P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling extremal events for insurance and finance* (Springer, Berlin, 1997).

-
- [3] D. Sornette, *Critical Phenomena in Natural Sciences* (Springer, Heidelberg, 2003).
- [4] J. P. Bouchaud, *Quantitative Finance* **1**, 105 (2001).
- [5] M. E. J. Newman, cond-mat/0412004.
- [6] B. B. Mandelbrot, *Fractals and scaling in finance. Discontinuity, concentration, risk* (Springer, New-York, 1997).
- [7] D. Carpentier and P. LeDoussal, *Phys. Rev. E* **63**, 026110 (2001).
- [8] J. F. Muzy and E. Bacry, *Phys. Rev. E* **66**, 056121 (2002).
- [9] E. Bacry and J. F. Muzy, *Comm. in Math. Phys.* **236**, 449 (2003).
- [10] U. Frisch, *Turbulence: The legacy of A.N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).
- [11] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. Shraiman, *Phys. Rev. A* **33**, 1141 (1986).
- [12] G. M. Molchan, *Comm. in Math. Phys.* **179**, 681 (1996).
- [13] B. Lahermes, P. Abry, and P. Chainais, *Int. J. of Wavelets, Multiresolution and Inf. Proc.* **2**, 497 (2004).
- [14] J. F. Muzy and E. Bacry, *Multifractal formalism, extreme values and negative dimensions: Applications to turbulence data* (2005), preprint.
- [15] B. B. Mandelbrot, *J. Fluid Mech.* **62**, 1057 (1974).
- [16] B. B. Mandelbrot, *C.R. Acad. Sci. Paris* **278**, 289 (1974).
- [17] B. B. Mandelbrot, *J. Stat. Phys.* **110**, 739 (2003).
- [18] J. P. Kahane and J. Peyrière, *Adv. in Mathematics* **22**, 131 (1976).
- [19] Y. Guivarc'h, *C.R. Acad. Sci. Paris* **305**, 139 (1987).
- [20] G. M. Molchan, *Phys. Fluids* **9**, 2387 (1997).
- [21] Q. S. Liu, *Asian J. of Math.* **6**, 145 (2002).
- [22] J. F. Muzy, J. Delour, and E. Bacry, *Eur. J. Phys. B* **17**, 537 (2000).
- [23] J. Barral and B. B. Mandelbrot, *Prob. Theory and Relat. Fields* **124**, 409 (2002).
- [24] A. Arneodo, J. F. Muzy, and D. Sornette, *Eur. Phys. J. B* **2**, 277 (1998).
- [25] A. Arneodo, E. Bacry, S. Manneville, and J. F. Muzy, *Phys. Rev. Lett.* **80**, 708 (1998).
- [26] B. B. Mandelbrot, *Physica A* **163**, 306 (1990).
- [27] B. Derrida and H. Spohn, *J. Stat. Phys.* **51**, 817 (1988).
- [28] D. S. Dean and S. N. Majumdar, *Phys. Rev. E* **64**, 045101 (2001).

- [29] W. Van-Saarloos, Phys. Rep. **386**, 29 (2003).
- [30] E. Brunet, Ph.D. thesis, Université de Paris VII (2000).
- [31] S. N. Majumdar and P. L. Krapivsky, Phys. Rev. E **65**, 036127 (2002).
- [32] E. Brunet and B. Derrida, Phys. Rev. E **62**, 2597 (1997).
- [33] S. N. Majumdar and P. L. Krapivsky, Phys. Rev. E **62**, 7735 (2000).
- [34] J. P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing* (Cambridge University Press, Cambridge, 2003).
- [35] E. Bacry, J. Delour, and J. F. Muzy, Phys. Rev. E **64**, 026103 (2001).
- [36] P. Gopikrishnam, M. Meyer, L. A. N. Amaral, and H. E. Stanley, Eur. Phys. J. B **3**, 139 (1998).
- [37] B. M. Garman and M. J. Klass, J. of Business **3**, 67 (1980).

Bibliographie

- [ABCD05] T. G. Andersen, T. Bollerslev, P. Christoffersen and F. X. Diebold, *Volatility and correlation forecasting*, In G. Elliott, C.W.J. Granger and Allan Timmermann (Eds.), *Handbook of Economic Forecasting*, Amsterdam : North-Holland, 2005.
- [ABM98] A. Arneodo, E. Bacry and J. F. Muzy, “Random cascades on wavelet dyadic trees”, *Journal of Mathematical Physics* **39** (1998), pp. 4124–4164.
- [ABM02] T. G. Andersen, T. Bollerslev and N. Meddahi, *Analytic evaluation of volatility forecasts*, CIRANO Working Paper 2002s-90, 2002.
- [ACS99] T. G. Andersen, H. J. Chung and B. E. Sørensen, “Efficient method of moments estimation of a stochastic volatility model : A monte carlo study”, *Journal of Econometrics* **91** (1999), pp. 61–87.
- [AG00] T. Ané and H. Geman, “Order flow, transaction clock and normality of asset returns”, *Journal of Finance* **9** (2000), pp. 2259–2284.
- [Akg89] V. Akgiray, “Conditional heteroskedasticity in time series of stock returns : evidence and forecasts”, *Journal of Business* **62** (1989), pp. 55–80.
- [AMR97] A. Arneodo, J. F. Muzy and S. Roux, “Experimental analysis of self-similarity and random cascade processes : application to fully developed turbulence data”, *Journal of Physique II France* **7** (1997), pp. 363–370.
- [AMS98] A. Arneodo, J. F. Muzy and D. Sornette, “”Direct” causal cascade in the stock market”, *European Physical Journal B* **2** (1998), pp. 277–282.
- [AS65] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, Dover Publications Inc., New York, 1965.
- [AS96] T. G. Andersen and B. E. Sørensen, “Gmm estimation of a stochastic volatility model : A monte carlo study”, *Journal of Business & Economic Statistics* **14** (1996), pp. 328–352.
- [AS97] ———, “Gmm and qml asymptotic standard deviations in stochastic volatility models : Comments on ruiz (1994)”, *Journal of Econometrics* **76** (1997), pp. 397–403.
- [ASM75] A. M. Yaglom A. S. Monin, *Statistical fluid mechanics*, MIT Press, Cambridge, 1975.
- [AV98] P. Abry and D. Veitch, “Wavelet analysis of long-range dependent traffic”, *IEEE Trans. on Info. Theory* **44** (1998), pp. 2–15.
- [Bac00] L. Bachelier, *Théorie de la spéculation*, Ph.D. thesis, Ecole Normale Supérieure, 1900.
- [Bar03] J. Barral, “Poissonian products of random weights : uniform convergence and related measures”, *Rev. Math. Ibero-Amer.* **19** (2003), pp. 1–44.

- [BB03] F. Bickenbach and E. Bode, “Evaluating the markov property in studies of economic convergence”, *International Regional Science Review* **26** (2003), pp. 363–392.
- [BBCI98] A. Benassi, P. Bertrand, S. Cohen and J. Istas, “Identification of the hurst exponent of a step multifractional brownian motion”, *Stat. Inf. Stoc. Proc.* **3** (1998), pp. 101–111.
- [BBCI99] ———, “Identification d’un processus gaussien multifractionnaire avec des ruptures sur la fonction d’échelle”, *Notes aux C.R.A.S.* **329** (1999), pp. 435–440.
- [BBM96] R. T. Baillie, T. Bollerslev and H. O. Mikkelsen, “Fractionally integrated generalized autoregressive conditional heteroskedasticity”, *Journal of Econometrics* **74** (1996), pp. 3–30.
- [BCK92] T. Bollerslev, R. Chou and K. Kroner, “Arch modelling in finance : A review of the theory and empirical evidence”, *Journal of Econometrics* **52** (1992), pp. 5–59.
- [BCT98] A. Bestavros, M. Crovella and M. Taqqu, *Heavy-tailed probability distributions in the world wide web*, Birkhäuser, Boston, 1998.
- [BD96] P. J. Brockwell and R. A. Davis, *Introduction to time series and forecasting*, Springer-Verlag, New York, 1996.
- [BDM01] E. Bacry, J. Delour and J. F. Muzy, “Multifractal random walks”, *Phys. Rev. E* **64** (2001), pp. 026103–026106.
- [BEK98] F. Bassi, P. Embrechts and M. Kafetzaki, *Risk management and quantile estimation*, Birkhäuser, Boston, 1998.
- [Ber94] J. Beran, *Statistics for long-memory processes*, Chapman & Hall, 1994.
- [Ber96] J. Bertoin, *Lévy processes*, Cambridge University Press, Cambridge, 1996.
- [BG74] R. Blattberg and N. Gonedes, “A comparison of stable and student distribution as statistical models for stock prices”, *Journal of Business* **47** (1974), pp. 244–280.
- [BG01] L. Bauwens and P. Giot, *Econometric modelling of stock market intraday activity*, Kluwer Academic Publisher, Boston, 2001.
- [BGT87] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation, encyclopedia of mathematics and its applications*, Cambridge Univ. Press, 1987.
- [Bil68] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons, Inc., 1968.
- [BJR97] A. Benassi, S. Jaffard and D. Roux, “Gaussian processes and pseudodifferential elliptic operators”, *Revista Mathematica Iberoamericana* **1** (1997), pp. 19–90.
- [BKM06] E. Bacry, A. Kozhemyak and J. F. Muzy, *Continuous cascade models for asset returns*, à paraître dans *Journal of Economic Dynamics and Control*, 2006.
- [Bla76] F. Black, “Studies of stock market volatility changes”, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* (1976), pp. 177–181.

- [BM02] J. Barral and B. B. Mandelbrot, “Multifractal products of cylindrical pulses”, *Prob. Theory and Relat. Fields* **124** (2002), pp. 409–430.
- [BM03] E. Bacry and J. F. Muzy, “Log-infinitely divisible multifractal process”, *Comm. in Math. Phys.* **236** (2003), pp. 449–475.
- [BMA93] E. Bacry, J. F. Muzy and A. Arneodo, “Singularity spectrum of fractal signals from wavelet analysis : exact results”, *J. Stat. Phys.* **70** (1993), pp. 635–674.
- [BMP01] J. P. Bouchaud, A. Matacz and M. Potters, “Leverage effect in financial market : the retarded volatility model”, *Physical Review Letters* **87** (2001), p. 228701.
- [BN98] O. E. Barndorff-Nielsen, “Processes of normal inverse gaussian type”, *Finance and Stochastics* **2** (1998), pp. 41–68.
- [BNMR01] O. E. Barndorff-Nielsen, T. Mikosch and S. Resnik, *Lévy processes – theory and applications*, Birkhäuser, Boston, 2001.
- [BNS01] O. E. Barndorff-Nielsen and N. Shephard, “Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics”, *Journal of Royal Statistical Society* **63** (2001), pp. 167–241.
- [BO99] J. Beran and D. Ocker, “Semifar forecasts, with applications to foreign exchanges rates”, *Journal of Statistical Planning and Inference* **80** (1999), pp. 137–153.
- [Bol86] T. Bollerslev, “A generalized autoregressive conditional heteroskedasticity”, *Journal of Econometrics* **31** (1986), pp. 307–327.
- [Bor22] L. Bortkiewicz, “Variations breite und mittlerer fehler”, *Sitzungsberichte der Berliner Mathematischen Gesellschaft* **21** (1922), pp. 3–11.
- [Bou01] J. P. Bouchaud, “Power laws in economics and finance : some ideas from physics”, *Quantitative Finance* **1** (2001), pp. 105–112.
- [BP03] J. P. Bouchaud and M. Potters, *Theory of financial risk and derivative pricing*, Cambridge University Press, Cambridge, 2003.
- [Bre92] L. Breiman, *Probability*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [Bro28] R. Brown, “A brief account of microscopical observations made in the months of june, july and august, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies”, *Philosophical Magazine* **4** (1928), pp. 161–173.
- [BS73] F. Black and M. Scholes, “The pricing of options and corporate liabilities”, *Journal of Political Economy* **3** (1973), pp. 637–654.
- [CAP99] P. Chainais, P. Abry and J. F. Pinton, “Intermittency and coherent structures in a turbulent flow : a wavelet analysis of joint pressure and velocity measurements”, *Phys. Fluids* **11** (1999), pp. 3524–3539.
- [Cas96] B. Castaing, “The temperature of turbulent flows”, *Journal de Physique II France* **6** (1996), pp. 105–114.

- [CD99] J. Chlès and P. Delfiner, *Geostatistics : Modeling spatial uncertainty*, Wiley, New York, 1999.
- [CD01] D. Carpentier and P. Le Doussal, “Glass transition of a particle in a random potential, front selection in non linear renormalization group, and entropic phenomena in liouville and sinh-gordon models”, *Physical Review E* **63** (2001), pp. 026110–026144.
- [CDFN95] K. V. Chow, K. C. Denning, S. Ferris and G. Noronha, “Long-term and short-term price memory in the stock market”, *Economics Letters* **49** (1995), pp. 287–293.
- [CF37] E. A. Cornish and R. A. Fisher, “Moments and cumulants in the specification of distributions”, *Review of the International Statistical Institute* **5** (1937), pp. 307–320.
- [CF01] L. Calvet and A. Fisher, “Forecasting multifractal volatility”, *Journal of Econometrics* **105** (2001), pp. 27–58.
- [CF02] ———, “Multifractality in asset returns : theory and evidence”, *Review of Economics and Statistics* **84** (2002), pp. 381–406.
- [CF04] ———, “How to forecast long-run volatility : Regime-switching and the estimation of multifractal processes”, *Journal of Financial Econometrics* **2** (2004), pp. 49–83.
- [CFA04] L. Calvet, A. Fisher and J. Adlai, “How to forecast long-run volatility : Regime switching and the estimation of multifractal processes”, *Journal of Financial Econometrics* **2** (2004), pp. 49–83.
- [CFM97] L. Calvet, A. Fisher and B. B. Mandelbrot, *Large deviation theory and the distribution of price changes*, Cowles Foundation Paper, 1165, 1997.
- [CGH90] B. Castaing, Y. Gagne and E. Hopfinger, “Velocity probability density functions of high reynolds number turbulence”, *Physica D* **46** (1990), pp. 177–200.
- [CGMY03] P. Carr, H. German, D. Madan and M. Yor, “Stochastic volatility for lévy processes”, *Mathematical Finance* **13** (2003), pp. 345–382.
- [Cha06] P. Chainais, “Multidimensional infinitely divisible cascades”, *The European Physical Journal B* **51** (2006), pp. 229–243.
- [Che93] Y. W. Cheung, “Long memory in foreign-exchanges rates”, *Journal of Business and Economic Statistics* **11** (1993), pp. 93–101.
- [Che03] P. Cheridito, “Arbitrage in fractional brownian motion models”, *Finance and Stochastics* **7** (2003), pp. 533–553.
- [Chr98] P. Christoffersen, “Evaluating interval forecasts”, *International Economic Review* **39** (1998), pp. 841–862.
- [Chu74] Kai Lai Chung, *A course in probability theory*, Academic Press Inc., London, second ed., 1974.

- [CKW95] N. J. Cutland, P. E. Kopp and W. Willinger, “From discrete to continuous stochastic calculus”, *Stochastics and Stochastics Reports* **52** (1995), pp. 173–192.
- [CL95] Y. W. Cheung and K. Lai, “A search of long memory in international stock market returns”, *Journal of International Money and Finance* **14** (1995), pp. 597–615.
- [Cla73] P. K. Clark, “A subordinated stochastic process model with finite variance for speculative prices”, *Econometrica* **41** (1973), pp. 135–156.
- [CLAB00] P. Chainais, E. Lévêque, P. Abry and Ch. Baudet, *Remarkable features of multiplier distributions in turbulence*, European Turbulence Conference VIII, 2000.
- [CP04] P. Christoffersen and D. Pelletier, “Backtesting value-at-risk : A duration-based approach”, *Journal of Financial Econometrics* **2** (2004), pp. 84–108.
- [CPB97] R. Cont, M. Potters and J.-P. Bouchaud, *Scaling in stock market data : stable laws and beyond*, Scale invariance and beyond, 1997.
- [CRA03] P. Chainais, R. Riedi and P. Abry, *Scale invariant infinitely divisible cascades*, in Int. Symp. on Physics in Signal and Image Processing, Grenoble, France, 2003.
- [CT65] J. W. Cooley and J. W. Tukey, “An algorithm for the machine calculation of complex fourier series”, *Math. Comput.* **19** (1965), pp. 297–301.
- [CT96] Y. Chang and S. J. Taylor, *Information arrivals and intraday exchange rate volatility*, Department of Accounting and Finance, The Management School, Lancaster, octobre, 1996.
- [Dah89] R. Dahlhaus, “Efficient parameter estimation for self-similar processes”, *The annals of Statistics* **17** (1989), pp. 1749–1766.
- [DG96] Z. Ding and C. W. J. Granger, “Modeling volatility persistence of speculative returns : a new approach”, *Journal of Econometrics* **73** (1996), pp. 185–215.
- [DGM⁺01] M. M. Dacorogna, R. Gençay, U. A. Müller, R. B. Olsen and O. V. Pictet, *An introduction to high frequency finance*, Academic Press, 2001.
- [DH87] R. B. Davies and D. S. Harte, “Tests for hurst effect”, *Biometrika* **74** (1987), pp. 95–101.
- [DL83] R. Durrett and T. M. Liggett, “Fixed points of the smoothing transformation”, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **45** (1983), pp. 281–285.
- [DM79] R. L. Dobrushin and P. Major, “Non-central limit theorems for non-linear functionals of gaussian fields”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **50** (1979), pp. 27–52.
- [DMNP93] M. M. Dacorogna, U. A. Müller, R. Nagler and O. V. Pictet, “A geographical model for the daily and weekly seasonal volatility in the foreign exchange market”, *Journal of International Money and Finance* **12** (1993), pp. 413–438.

- [DN93] F. C. Drost and T. E. Nijman, “Temporal aggregation of garch processes”, *Econometrica* **61** (1993), pp. 909–927.
- [DN97] C. R. Dietrich and G. N. Newsam, “Fast and exact simulation of stationary gaussian processes through circulant embedding of the covariance matrix”, *SIAM Journal on Scientific and Statistical Computing* **18** (1997), pp. 1088–1107.
- [Doo01] J. L. Doob, *Classical potential theory and its probabilistic counterpart*, Springer, Berlin, reprint of the 1984 ed., 2001.
- [Dur60] J. Durbin, “The fitting of time series models”, *Rev. Inst. Int. Stat* **28** (1960), pp. 233–243.
- [EB86] R. F. Engle and T. Bollerslev, “Modelling the persistence of conditional variances. with comments and a reply by the authors”, *Econometric Review* **5** (1986), pp. 1–87.
- [Ein05] A. Einstein, “Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen”, *Ann. Phys.* **17** (1905), pp. 549–560.
- [Ein56] ———, *Investigations on the theory of brownian movement*, Dover, New York, 1956.
- [EKM97] P. Embrechts, C. Kluppelberg and T. Mikosch, *Modelling extremal events for insurance and finance*, Springer, Berlin, 1997.
- [EKP98] E. Eberlein, U. Keller and K. Prause, “New insights into smile, mispricing and value at risk”, *Journal of Business* **71** (1998), pp. 371–406.
- [Eng82] R. F. Engle, “Autoregressive conditional heteroskedasticity with estimates of the variance of u.k. inflation”, *Econometrica* **50** (1982), pp. 987–1008.
- [ER98] R. F. Engle and J. R. Russell, “Autoregressive conditional duration : a new model for irregularly spaced transaction data”, *Econometrica* **66** (1998), pp. 1127–1162.
- [ESDM96] P. Embrechts, G. Samorodnitsky, M. M. Dacorogna and U. A. Müller, *How heavy are the tails of a stationary HARCH(k) process? a study of the moments*, preprint, 1996.
- [Eve95a] C. J. G. Evertsz, “Fractal geometry of finance time series”, *Fractals* **3** (1995), pp. 609–616.
- [Eve95b] ———, *Self-similarity of high-frequency usd/dem exchange rates*, Proceedings of the First International Conference on High Frequency Data in Finance, 1995.
- [Fal90] K. Falconer, *Fractal geometry*, Wiley, New-York, 1990.
- [Fam65] E. F. Fama, “The behavior of stock market price”, *Journal of Business* **38** (1965), pp. 34–195.
- [FCM97] A. Fisher, L. Calvet and B. B. Mandelbrot, *Multifractality of deutschmark/us dollar exchange rate*, Cowles Foundaton Paper, 1166, 1997.

- [Fel68] W. Feller, *An introduction to probability theory and its applications*, vol. I, John Wiley & Sons Inc., 2 ed., 1968.
- [Fel71] ———, *An introduction to probability theory and its applications*, vol. II, John Wiley & Sons Inc., 2 ed., 1971.
- [FH98] M. Fridman and L. Harris, “A maximum likelihood approach for non-gaussian stochastic volatility models”, *Journal of Business and Economic Statistics* **16** (1998), pp. 284–291.
- [FP85] U. Frisch and G. Parisi, *Fully developed turbulence and intermittency*, Proc. of Int. Summer school Phys. Enrico Fermi, 1985.
- [Fri95] U. Frisch, *Turbulence*, Cambridge Univ. Press, Cambridge, 1995.
- [FSS87] K. French, G. Schwert and R. Stambaugh, “Expected stock returns and volatility”, *Journal of Financial Economics* **19** (1987), pp. 3–29.
- [FT28] R. A. Fisher and L. H. C. Tippett, “Limiting forms of the frequency distributions of the largest or smallest member of a sample”, *Proc. Camb. Phil. Soc.* **24** (1928), pp. 180–190.
- [Gal87] J. Galambos, *The asymptotic theory of extreme order statistics*, Krieger, Melbourne, FL, 2nd ed., 1987.
- [GB03] I. Giardina and J. P. Bouchaud, “Bubbles, crashes and intermittency in agent based market models”, *European Physical Journal B* **31** (2003), pp. 421–437.
- [GBP⁺96] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner and Y. Dodge, “Turbulent cascades in foreign exchange markets”, *Nature* **386** (1996), pp. 767–770.
- [GF91] C. A. Goodhard and L. Figliuoli, “Every minut counts in financial markets”, *Journal of International Money and Finance* **10** (1991), pp. 23–52.
- [GHR95] E. Ghysels, A. C. Harvey and E. Renault, *Statistical methods in finance*, Elsevier Science, Amsterdam : North-Holland, 1995, volume 14 of Handbook of Statistics, chapter Stochastic Volatility, pages 128–198.
- [GK68] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley Publishing Co., Mass.-London-Don Mills, revised ed., 1968, Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu.
- [GMAS98] P. Gopikrishnan, M. Meyer, L. A. N. Amaral and H. E. Stanley, “Inverse cubic law for the distribution of stock price variations”, *Eur. Phys. J. B* **3** (1998), pp. 139–140.
- [Gne98] T. Gneiting, “Simple tests for the validity of correlation function models on the circle”, *Statistics and Probability Letters* **39** (1998), pp. 119–122.
- [Gne00] ———, *Compactly supported correlation functions*, Technical Report 45. National Research Center for Statistics and the Environment, University of Washington., 2000.
- [GO95] C. A. Goodhard and M. O’Hara, *High-frequency data in financial markets : issues and applications*, Proceedings of the First International Conference on High Frequency Data in Finance, 1995.

- [Gol91] C. W. Goldie, “Implicite renewal theory and tails of solutions of random equations”, *Annals of Applied Probability* **1** (1991), pp. 126–166.
- [Goo89] C. A. Goodhard, “News and the foreign exchange market”, *Proc. Manchester Statist. Soc.* (1989), pp. 1–79.
- [Gou92] C. Gouriéroux, *Modèles arch et applications financières*, Economica, Paris, 1992.
- [GPA⁺99] P. Gopikrishnan, V. Plerou, L. A. N. Amaral, M. Meyer, and H. E. Stanley, “Scaling of the distribution of fluctuations of financial market indices”, *Physical Review E* **60** (1999), pp. 5305–5316.
- [GS85] L. Giratis and D. Surgailis, “Clt and other limit theorems for functionals of gaussian processes”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **70** (1985), pp. 191–212.
- [Gui87] Y. Guivarc’h, “Remarques sur les solutions d’une équation fonctionnelle non linéaire de b. mandelbrot”, *C.R. Acad. Sci. Paris* **305** (1987), pp. 139–141.
- [Gui90] ———, “Sur une extension de la notion de loi semi-stable”, *Annales de l’IHP* **26** (1990), pp. 261–285.
- [Han82] L.P. Hansen, “Large sample properties of generalized method of moments estimators”, *Econometrica* **50** (1982), pp. 1029–1054.
- [Hes93] S. L. Heston, “A closed-form solution for options with stochastic volatility with applications to bond and currency options”, *The Review of Financial Studies* **6** (1993), pp. 327–343.
- [Hil75] B. M. Hill, “A simple general approach to inference about the tail of a distribution”, *Annals of Statistics* **13** (1975), pp. 331–341.
- [HJK⁺86] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia and B. I. Shraiman, “Fractal measures and their singularities : the characterization of strange sets”, *Physical Review A* **33** (1986), pp. 1141–1151.
- [Hor01] U. Horst, “Financial price fluctuation in a stock market model with many interacting agents”, *ISE Finance Award Series* **2** (2001), pp. 1–28.
- [Hur65] H. E. Hurst, *Long-term storage : An experimental study*, Constable, London, 1965.
- [Jaf97a] S. Jaffard, “Multifractal formalism for functions, part 1 : Results valid for all functions”, *SIAM J. Math. Anal.* **28** (1997), pp. 944–970.
- [Jaf97b] ———, “Multifractal formalism for functions, part 2 : Selfsimilar functions”, *SIAM J. Math. Anal.* **28** (1997), pp. 971–998.
- [JB87] C. M. Jarque and A. K. Bera, “A test for normality of observations and regression residuals”, *International Statistical Review* **55** (1987), pp. 163–172.
- [JHJ01] R. Jennane, R. Harba and G. Jacquet, “Analysis methods for fractional brownian motion : theory and comparative results”, *Traitement du Signal* **18** (2001), pp. 419–436.

- [JPR94] E. Jacquier, N. G. Polson and P. E. Rossi, “Bayesian analysis of stochastic volatility models”, *Journal of Business & Economic Statistics* **12** (1994), pp. 69–87.
- [Kah87] J. P. Kahane, “Positive martingales and random measures”, *Chi. Ann. of Math.* **8B** (1987), pp. 1–12.
- [Kai86] T. Kailath, “A theorem of I. Schur and its impact on modern signal processing” in *Schur methods in operator theory and signal processing*, I. Gohberg (ed.), pp. 9–30, Birkhauser-Verlag, Boston, 1986.
- [Kes73] H. Kesten, “Random difference equation and renewal theory for product of random matrices”, *Acta Mathematica* **131** (1973), pp. 207–248.
- [Kol40] A. N. Kolmogorov, “Wiensche spiralen und einige andere interessante kurven in hilbertschen raum”, *Comptes Rendus (Doklady) Acad. Sci. USSR (N.S.)* **26** (1940), pp. 115–118.
- [Kol41] ———, “The local structure of turbulence in an incompressible fluid at very high reynolds numbers”, *Dokl. Acad. Nauk USSR* **30** (1941), pp. 299–303.
- [KP76] J. P. Kahane and J. Peyrière, “Sur certaines martingales de benoît mandelbrot”, *Adv. in Mathematics* **22** (1976), pp. 131–145.
- [KT02] A. Kirman and G. Teyssière, “Microeconomic models for long memory in the volatility of financial time series”, *Studies in nonlinear dynamics and econometrics* **5** (2002), pp. 281–302.
- [Kup95] P. Kupiec, “Techniques for verifying the accuracy of risk measurement models”, *Journal of Derivatives* **2** (1995), pp. 73–84.
- [LAC04] B. Lahermes, P. Abry and P. Chainais, “New insights into the estimation of scaling exponents”, *International Journal of Wavelets, Multiresolution and Information Processing* **2** (2004), no. 4, pp. 497–523.
- [Lam62] J. W. Lamperti, “Semi-stable processes”, *Transaction of American mathematical society* **104** (1962), pp. 62–78.
- [Lév37] P. Lévy, *Théorie de l’addition des variables aléatoires*, Gauthier-Villars, Paris, second ed., 1937.
- [Lev47] N. Levinson, “The wiener rms error criterion in filter design and prediction”, *J. Math. Phys.* **25** (1947), pp. 261–278.
- [IFMAV98] I. M. De la Fuente, L. Martinez, J. M. Aguirregabiria and J. Veguillas, “ r/s analysis in strange attractors”, *Fractals* **6** (1998), pp. 95–100.
- [LH94] S. W. Lee and B. E. Hansen, “Asymptotic theory for the garch(1,1) quasi-maximum likelihood estimator”, *Econometric Theory* **10** (1994), pp. 29–52.
- [Liu02] Q. Liu, “An extension of a functional equation of poincaré and mandelbrot”, *Asian Journal of Mathematics* **6** (2002), pp. 145–168.
- [LL90] C. G. Lamoureux and W. D. Lastrapes, “Persistence in variance, structural change, and the garch model”, *Journal of Business & Economic Statistics* **8** (1990), pp. 225–234.

- [LL97] D. Lambertson and B. Lapeyre, *Introduction au calcul stochastique appliqués à la finance*, Ellipses, 1997.
- [LLR83] M. R. Leadbetter, G. Lingren and H. Rootzén, *Extremes and related properties of random sequences and processes*, Springer Verlag, Berlin, 1983.
- [Lo91] A. W. Lo, “Long-term memory in stock market prices”, *Econometrica* **59** (1991), pp. 1279–1313.
- [LTWW94] W. Leland, M. Taqqu, W. Willinger and D. Wilson, “On the self-similar nature of ethernet traffic”, *IEEE/ACM Transaction on Networking* (1994), pp. 1–15.
- [Lux98] T. Lux, “The socio-economic dynamics of speculative markets : interacting agents, chaos, and the fat tails of returns distributions”, *Journal of Economic Behavior and Organization* **33** (1998), pp. 143–165.
- [Lux01] ———, “Turbulence in financial markets : the surprising explanatory power of simple cascade models”, *Quantitative finance* **1** (2001), pp. 632–640.
- [Lux03] ———, *The multi-fractal model of asset returns : Its estimation via gmm estimation and its use for volatility forecasting*, University of Kiel, 2003.
- [Lux04] ———, *The markov-switching multi-fractal model of asset returns : Gmm estimation and linear forecasting of volatility*, University of Kiel, 2004.
- [LVW02] J. Lévy-Véhel and C. Walter, *Les marchés fractals*, Presses universitaires de France, Paris, 2002.
- [Lyo01] R. K. Lyons, *The microstructure approach to excnge rates*, MIT Press, Cambridge, 2001.
- [Man63] B. B. Mandelbrot, “The variation of certain speculative prices”, *Journal of Business* **36** (1963), pp. 394–419.
- [Man73a] ———, “Formes nouvelles du hasard dans les sciences”, *Economie appliquée* **26** (1973), pp. 307–319.
- [Man73b] ———, “Le problème de la réalité des cycles lents et le syndrome de joseph”, *Economie appliquée* **26** (1973), pp. 349–365.
- [Man73c] ———, “Le syndrome de la variance infinie et ses rapports avec la discontinuité des prix”, *Economie appliquée* **26** (1973), pp. 321–348.
- [Man74a] ———, “Intermittent turbulence in self-similar cascades : divergence of high moments and dimension of the carrier”, *Journal of Fluid Mechanics* **62** (1974), pp. 331–358.
- [Man74b] ———, “Multiplications aléatoires itérées et distributions invariantes par moyenne pondérée aléatoire”, *C.R. Acad. Sci. Paris* **278** (1974), pp. 289–292.
- [Man77] ———, *Fractals : Form, chance and dimensions*, Freeman, San Francisco, 1977.
- [Man82] ———, *The fractal geometry of nature*, Freeman, San Francisco, 1982.
- [Man88] ———, *Fractals and scaling in finance*, Springer, 1988.
- [Man89] ———, “Multifractal measures, especially for geophysicist”, *Pure and Applied Geophysics* **131** (1989), pp. 5–42.

- [Man90] ———, “Limit lognormal multifractal measures”, *Physica A* **163** (1990), pp. 306–315.
- [Man97] ———, *Fractals and scaling in finance. discontinuity, concentration, risk*, Springer, New-York, 1997.
- [Man99] ———, “A multifractal walk down street”, *Scientific American* **280** (1999), pp. 70–73.
- [Man01] ———, “Scaling in finance prices : Ii. multifractals and the star equation”, *Quantitative finance* **1** (2001), pp. 124–130.
- [Man03] ———, “Multifractal power law distributions : Negative and critical dimensions and other ”Anomalies”, explained by a simple example”, *Journal of Statistical Physics* **110** (2003), pp. 739–774.
- [MB02] J. F. Muzy and E. Bacry, “Multifractal stationary random measures and multifractal random walks with log-innitely divisible scaling laws”, *Phys. Rev. E* **66** (2002), pp. 056121–056150.
- [MBA91] J. F. Muzy, E. Bacry and A. Arneodo, “Wavelet and multifractal formalism for singular signals : applications to turbulence data”, *Phys. Rev. Lett.* **67** (1991), pp. 3515–3518.
- [MBA94] ———, “The multifractal formalism revisited with wavelets”, *Int J. of Bif. and Chaos* **4** (1994), pp. 245–302.
- [MBK06] J. F. Muzy, E. Bacry and A. Kozhemyak, “Extreme values and fat tails of multifractal fluctuations”, *Physical Review E* **73** (2006), p. 066114.
- [MD⁺90] U. A. Müller, M. M. Dacorogna, , R. B. Olsen, O. V. Pictet, M. Schwartz, and G. Morgenegg, “Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis”, *Journal of Banking and Finance* **14** (1990), pp. 1189–1208.
- [MDB00] J. F. Muzy, J. Delour and E. Bacry, “Modelling fluctuations of financial time series : from cascade process to stochastic volatility model”, *Eur. J. Phys. B* **17** (2000), pp. 537–548.
- [MDD⁺95] U. A. Müller, M. M. Dacorogna, R. D. Davé, O. V. Pictet, R. B. Olsen and J. R. Ward, *Fractals and intrinsic time, a challenge to econometricians*, preprint, O& A Research Group, 1995.
- [MDD⁺97] U. A. Müller, M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet and J.E. von Weizsächer, “Volatilities of different time resolutions - analysing the dynamics of market components”, *Journal of Empirical Finance* **4** (1997), pp. 213–239.
- [MF00] A. McNeil and R. Frey, “Estimation of tail-related risk measures for heteroecastic financial times series : an extreme value approach”, *Journal of Empirical Finance* **7** (2000), pp. 271–300.
- [MFC97] B. B. Mandelbrot, A. Fisher and L. Calvet, *A multifractal model of asset returns*, Cowles Foundation Discussion Paper, 1164, 1997.

- [Mis23] R. Mises, “Über die variationsbreite einer beobachtungsreihe”, Sitzungsberichte der Berliner Mathematischen Gesellschaft **22** (1923), pp. 3–8.
- [Mis54] ———, *La distribution de la plus grande de n valeur*, In Selected Papers II. Amer. Math. Soc., Providence, R.L, 1954.
- [MN68] B. Mandelbrot and J. W. Van Ness, “Fractional brownian motions, fractional gaussian noises and applications”, SIAM Review **10** (1968), pp. 422–437.
- [Mol96] G. M. Molchan, “Scaling exponents and multifractal dimensions for independent random cascades”, Comm. in Math. Phys. **179** (1996), pp. 681–702.
- [Mol97] ———, “Turbulent cascades : Limitations and a statistical test of the lognormal hypothesis”, Phys. Fluids **9** (1997), pp. 2387–2395.
- [MR97] B. B. Mandelbrot and R. H. Riedi, “Inverse measures, the inversion formula, and discontinuous multifractals”, Advances in Applied Mathematics **18** (1997), pp. 50–58.
- [MSL96] D. Marsan, D. Schertzer and S. Lovejoy, “Causal space-time multifractal process : predictability and forecasting of rain fields”, Journal of Geophysical Research **101** (1996), no. 26, pp. 333–346.
- [MT67] B. B. Mandelbrot and H. M. Taylor, “On the distribution of stock price differences”, Op. Research **15** (1967), pp. 1057–1062.
- [MT79] B. B. Mandelbrot and M. S. Taqqu, “Robust r/s analysis of long-run serial correlation”, Bulletin of the International Statistical Institute **48** (1979), pp. 69–99.
- [MT90] A. Melino and S. M. Turnbull, “Pricing foreign currency options with stochastic volatility”, Journal of Econometrics **45** (1990), pp. 239–265.
- [MZ69] J. Mincer and V. Zarnowitz, *The evaluation of economic forecasts*, Columbia University Press, New York, 1969, Economic Forecast and Expectations.
- [Nel91] D. B. Nelson, “Conditional heteroskedasticity in asset pricing : a new approach”, Econometrica **59** (1991), pp. 347–370.
- [New05] M. E. J. Newman, “Power laws, pareto distributions and zipf’s law”, Contemporary Physics **46** (2005), pp. 323–351.
- [Nov94] E. A. Novikov, “Infinitely divisible distributions in turbulence”, Phys. Rev. E **50** (1994), pp. 3303–3305.
- [NW87] W. Newey and K. West, “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix”, Econometrica **55** (1987), pp. 703–708.
- [NW94] ———, “Automatic lag selection in covariance matrix estimation”, Review of Economic Studies **61** (1994), pp. 631–654.
- [Orl99] A. Orléan, *Le pouvoir de la finance*, Odile Jacob, Paris, 1999.
- [OW00] M. Ossiander and E. C. Waymire, “Statistical estimation for multiplicative cascades”, The Ann. of Stat. **28** (2000), pp. 1533–1560.

- [Pag96] A. Pagan, “The econometrics of financial markets”, *Journal of Empirical Finance* **3** (1996), pp. 15–102.
- [PB02] B. Pochart and J. P. Bouchaud, “The skewed multifractal random walk with applications to option smiles”, *Quantitative finance* **2** (2002), pp. 303–314.
- [PF85] G. Parisi and U. Frisch, *Turbulence and predictability in geophysical fluid dynamics and climate dynamics*, Proc. of Int. School, 1985.
- [PGA⁺99] V. Plerou, P. Gopikrishnan, L. A. N. Amaral, M. Meyer, and H. E. Stanley, “Scaling of the distribution of fluctuations of financial market indices”, *Physical Review E* **60** (1999), pp. 6519–6529.
- [Pic75] J. Pickands, “Statistical inference using extreme order statistics”, *Annals of Statistics* **3** (1975), pp. 119–131.
- [PLV96] R. Peltier and J. Lévy-Vehel, *Multifractional brownian motion : definition and preliminary results*, Rapport de recherche de l’INRIA 2645 1995, 1996.
- [Poc03] B. Pochart, *Multifractal processes in finance and pricing of options by minimisation of extremal risks*, Ph.D. thesis, Ecole Polytechnique, Palaiseau, France, 2003.
- [Pog95] J. M. Poggi, *Statistiques quadratiques pour le brownien fractionnaire*, Univ. Paris-Sud, Orsay, preprint, 1995.
- [PS90] A. Pagan and W. Schwert, “Alternative models for conditional stock volatility”, *Journal of Econometrics* **45** (1990), pp. 267–290.
- [PS99] M. Pasquini and M. Serva, “Multiscaling and clustering of volatility”, *Physica A* **269** (1999), pp. 140–147.
- [PS00] ———, “Clustering of volatility as a multiscale phenomenon”, *The European Physical Journal B* **16** (2000), pp. 195–201.
- [PT02a] V. Pipiras and M. Taqqu, “Decomposition of self-similar stable mixed moving averages”, *Probability theory and related fields* **123** (2002), pp. 412–452.
- [PT02b] ———, “The structure of self-similar stable mixed moving averages”, *The annals of Probability* **30** (2002), pp. 898–932.
- [PTVF88] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical recipes in C*, Cambridge University Press, 1988.
- [Qua93a] D. T. Quah, “Empirical cross-section dynamics in economic growth”, *European Economic Review* **37** (1993), pp. 426–434.
- [Qua93b] ———, “Galton’s fallacy and tests of the convergence hypothesis”, *Scandinavian Journal of Economics* **95** (1993), pp. 427–443.
- [Rai00] S. Raible, *Lévy processes in finance : theory numerics and empirical facts*, Ph.D. thesis, Universität Freiburg, 2000.
- [RCRB99] R. H. Riedi, M. Crouse, V. Ribeiro and R. Baraniuk, “A multifractal wavelet model with application to tcp network traffic”, *IEEE Transaction on Information Theory* **45** (1999), pp. 992–1089.

- [Res87] S. I. Resnick, *Extreme values, regular variation and point processes*, Springer-Verlag, New York, 1987.
- [Rog97] L. C. G. Rogers, “Arbitrage with brownian motion”, *Math. Finance* **7** (1997), pp. 95–105.
- [Ros61] M. Rosenblatt, *Independence and dependence*, Proceedings of the 4th Berkley symposium of mathematical statistics and probability, 1961.
- [RR89] B. Rajput and J. Rosinski, “Spectral representation of infinitely divisible processes”, *Probab. Theory Relat. Fields* **82** (1989), pp. 451–487.
- [Sam65] P. Samuelson, “Rational theory of warrant pricing”, *Industrial Management Review* **6** (1965), pp. 13–32.
- [Sar58] J. D. Sargan, “The estimation of economic relationship using instrumental variables”, *Econometrica* **26** (1958), pp. 393–415.
- [Sat99] K. Sato, *Lévy processes and infinitely divisible distributions*, Cambridge University Press, Cambridge, 1999.
- [Sel44] A. Selberg, “Bemerkninger om et multipelt integral”, *Norsk Matematisk Tidsskrift* **26** (1944), pp. 71–78.
- [Shi84] A. N. Shiryaev, *Probability*, Springer-Verlag, 1984.
- [Shi99] ———, “Kolmogorov and the turbulence”, *MaPhySto Misc.* **12** (1999), pp. 1–24.
- [SK98] G. Sandmann and S. J. Koopman, “Estimation of stochastic volatility models via monte carlo maximum likelihood”, *Journal of Econometrics* **87** (1998), pp. 271–301.
- [SL94] Z. S. She and E. Lévêque, “Ununiversal scaling laws in fully developed turbulence”, *Phys. Rev. Lett.* **72** (1994), pp. 336–339.
- [SLSL92] F. Schmitt, D. Lavallé, D. Schertzer and S. Lovejoy, “Empirical determination of universal multifractal exponents in turbulent velocity fields”, *Phys. Rev. Lett.* **68** (1992), pp. 305–308.
- [SM01] F. Schmitt and D. Marsan, “Stochastique equation generating continuous multiplicative cascades”, *European Physical Journal B* **20** (2001), pp. 3–6.
- [Sor03] D. Sornette, *Critical phenomena in natural sciences*, Springer, Heidelberg, 2003.
- [SSL00] F. Schmitt, D. Schertzer and S. Lovejoy, “Multifractal fluctuations in finance”, *International Journal of Theoretical and Applied Finance* **3** (2000), pp. 361–364.
- [Taq79] M. S. Taqqu, “Convergence of intergrated processes of arbitrary hermite rank”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **50** (1979), pp. 53–83.
- [Tay86] S. J. Taylor, *Modelling financial time series*, Wiley, 1986.
- [Tay94] ———, “Modeling stochastic volatility”, *Mathematical Finance* **4** (1994), pp. 183–204.

- [TP00] A. Turiel and N. Parga, “Multifractal wavelet filter of natural images”, *Physical Review Letters* **85** (2000), pp. 3325–3328.
- [TS94] M. S. Taqqu and G. Samorodnitsky, *Stable non-gaussian random processes*, Chapman & Hall, 1994.
- [Tsa97] C. Tsallis, “Lévy distributions”, *Physics World* **10** (1997), pp. 42–45.
- [TTW99] V. Teverovsky, M. S. Taqqu and W. Willinger, “A critical look at lo’s modified r/s statistic”, *Journal of Statistical Planning and Inference* **80** (1999), pp. 211–227.
- [TX95] S. J. Taylor and X. Xu, “The incremental volatility information in one million foreign exchange quotations”, *Journal of Empirical Finance* **4** (1995), pp. 317–340.
- [VA98] N. Vandewalle and M. Ausloos, “Multi-affine analysis of typical currency exchange rates”, *The European Physical Journal B.* **4** (1998), pp. 257–261.
- [Vel99] C. Velasco, “Gaussian semiparametric estimation of non-stationary time series”, *Journal of Time Series Analysis* **20** (1999), pp. 87–127.
- [VTPF93] W. T. Vetterling, S. A. Teukolsky, B. P. Press and W. H. Flannery, *Numerical recipes in C : the art of scientific computing*, Cambridge University Press, 1993.
- [Wal90] C. Walter, “Levy-stables distributions and fractal structure on the paris market : an empirical examination”, *Proceedings of the 1st AFIR international colloquium* **3** (1990), pp. 241–259.
- [Wal95] ———, “Levy-stability-under-addition and fractal structure of markets : implications for the actuaries and emphasized axamination of matif national contract”, *Proceedings of the 5th AFIR international colloquium* (1995), pp. 1285–1330.
- [Wal03] ———, *1900-2000 : Un siècle de description statistique des fluctuations boursières ou les aléas du modèle de marche au hasard en finances*, Communication au colloque « Marché boursier ». Collège de France, 2003.
- [WB03] M. Wyart and J.-P. Bouchaud, *Self-referentiel behaviour, overreaction and conventions in financial markets*, cond-mat/0303584, 2003.
- [WC94] A. T. A. Wood and G. Chan, “Simulation of stationary gaussian processes in $[0, 1]^d$ ”, *Journal of Computational and Graphical Statistics* **3** (1994), pp. 409–432.
- [WC95] K. West and D. Cho, “The predictive ability of several models of exchange rate volatility”, *Journal of Econometrics* **69** (1995), pp. 367–391.
- [Wic50] G. C. Wick, “The evaluation of the collision matrix”, *Physical review* **80** (1950), pp. 268–272.
- [WM98] K. West and M. McCracken, “Regression-based tests of predictive ability”, *Iternational Economic Review* **39** (1998), pp. 817–840.

- [Yag66] A. M. Yaglom, “The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval”, *Soviet Physics Doklady* **11** (1966), pp. 26–29.
- [Zip29] G. K. Zipf, “Relative frequency as a determinant of phonetic change”, *Harvard Studies in Classical Philology* **15** (1929), pp. 1–95.