



Problèmes économétriques d'analyse des séries temporelles à mémoire longue

Béchir Dola

► To cite this version:

Béchir Dola. Problèmes économétriques d'analyse des séries temporelles à mémoire longue. Econométrie de la finance [q-fin.ST]. Université Panthéon-Sorbonne - Paris I, 2012. Français. NNT : 2012PA010052 . tel-00794676

HAL Id: tel-00794676

<https://theses.hal.science/tel-00794676>

Submitted on 26 Feb 2013

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.



École Doctorale De Sciences Économiques
Maison des Sciences Économiques de Paris

N° d'identification de la thèse

2	0	1	2	P	A	0	1	0	0	5	2
---	---	---	---	---	---	---	---	---	---	---	---

THÈSE DE DOCTORAT

pour obtenir le grade de :

DOCTEUR ÈS-SCIENCES

SPÉCIALITÉ : ÉCONOMÉTRIE DES SÉRIES TEMPORELLES

présentée et soutenue publiquement le : 13 décembre 2012 par

Béchir DOLA

sous le titre :

Problèmes économétriques d'analyse des séries temporelles à mémoire
longue

devant le jury composé de :

Président :	Dominique GUÉGAN	Professeur à l'Université Paris 1.
Directeur de Thèse :	Jean-Marc BARDET	Professeur à l'Université Paris 1.
Rapporteurs :	Pierre BERTRAND	Professeur à l'Université Blaise Pascal Clermont-Ferrand 2.
	Mohamed BOUTAHAR	Maître de Conférence (HDR) à Aix-Marseille Université.
Examineurs :	Paul DOUKHAN	Professeur à l'Université de Cergy-Pontoise, IUF.
	Valérie MIGNON	Professeur à l'Université Paris 10 Nanterre.

À ceux qui m'ont instruit. À ceux qui m'ont construit.

Remerciements

Mes remerciements sincères et toute ma gratitude vont au Professeur Jean-Marc Bardet, mon directeur de thèse. Je porte à son égard le plus grand estime pour ce qu'il a toujours été : ingénieux et rigoureux à toute épreuve, dans l'humilité spontanée et la bienveillance perspicace.

Je remercie Professeur Pierre Bertrand pour avoir accepté avec enthousiasme d'être le rapporteur de ma thèse. Son rapport minutieux et positif m'a permis d'actualiser les références interférées de certaines équations. Ses critiques relatives aussi bien aux modes de générations des données qu'aux interprétations des dates de ruptures des données financières m'ont permis de prendre conscience de ces insuffisances.

Je remercie le Directeur de Recherche Mr Mohamed Boutahar en lui présentant toute ma gratitude pour avoir accepté volontiers d'être, le rapporteur de ma thèse. Son rapport minutieux et positif pointe à raison certains aspects non élucidés dans certaines démonstrations ainsi que certaines erreurs et imprécisions glissées dans le texte originel que j'ai tâché de tout corriger dans cette nouvelle version.

Tout mon estime et toute ma considération sont adressés à mes examinateurs et membres du jury de ma thèse :

J'adresse tout d'abord mes chaleureux remerciements : au Professeur Paul Doukhan pour ses encouragements et son aiguillage permanent et bien aimable, toujours dans l'humour et la bonne humeur.

J'adresse toute ma considération au Professeur Dominique Guégan qui m'a toujours été d'un soutien inestimable à la fois par ses écrits et par son encouragement à toute épreuve.

Mes remerciements les meilleurs vont au Professeur Valérie Mignon qui m'a été le modèle de l'économiste accompli et qui m'honore par sa présence en tant que membre du jury de ma thèse.

Cette thèse n'aurait pas pu se réaliser comme elle l'est sans les conditions académiques humaines et logistiques du laboratoire SAMM. J'adresse ici tous mes remerciements à tous les membres du SAMM et à son équipe dirigeante en la personne de Pr Marie Cottrell relayée depuis septembre 2012 par Pr Jean-Marc Bardet qui m'ont soutenu et m'ont aidé à accomplir mes recherches dans les meilleures conditions. A dire vrai j'ai connu en eux deux directeurs d'exception, à tout point de vue.

Je garde de chacun de la quarantaine des membres du SAMM, les sentiments d'une amitié sincère et profonde et d'un soutien soutenu et indéfectible. J'adresse par la même à travers Dhaou et Fania les représentants des doctorants, à Anthikos le représentant de l'opposition, tous mes remerciements et toute ma gratitude à tous.

Un remerciement chaleureux et cordial va tout particulièrement au grand Érudit et Homme de Culture Professeur Xavier Greffe responsable des études doctorales en Économie de Paris 1- Panthéon-Sorbonne, pour ses efforts qui ont permis le bon déroulement de la soutenance de cette thèse.

L'accueil chaleureux et convivial dont j'ai toujours joui de la part de l'administration des études doctorales d'Économie, au bureau des thèses, à la MSE, en la personne de Mr Loïc Sorel et Mme Elda André et leur joie enthousiasmée à la nouvelle de l'imminence de la soutenance de ma thèse me rend redevable à leur égard. Je leur dis : merci Madame merci Monsieur.

Table des matières

1 Synthèse des travaux	9
Synthèse des travaux	9
1.1 Introduction	10
1.1.1 Quelques définitions	11
1.1.2 Motivation de notre travail	16
1.2 Épistémologie de la mémoire longue et son paradigme associé	18
1.2.1 Propension de la mémoire longue en économie	18
1.2.2 Carence de mémoire longue au sein des paradigmes économiques standards .	20
1.2.3 Le paradigme de la mémoire longue	28
1.2.4 La mémoire longue à l'oeuvre en Économie : une fusion intégrale est en perspective	29
1.3 Estimation et tests adaptatifs pour des processus LRD	30
1.3.1 Présentation de la statistique unidimensionnelle de Ratios d'Incréments : IR(m)	31
1.3.2 Estimateur issu de la Statistique Multidimensionnelle de Ratios d'Incréments	32
1.3.3 La version adaptative de l'estimateur	33
1.3.4 Application à des tests statistiques	34
1.4 Mise en œuvre empirique des différentes procédures d'estimation et de tests	36
1.4.1 Simulations	36
1.4.2 Applications aux séries financières	37
1.5 Conclusion	38
2 Épistémologie de la mémoire longue en économie et son “paradigme” associé	41
Épistémologie de la mémoire longue	41
2.1 L'épistémologie économique, ce labyrinthe des concepts problématiques	46
2.1.1 Problématiques et controverses autour du concept d'épistémologie	49
2.1.2 Controverses autour de la <i>scientificité</i> de l'économie	54

2.1.4	Paradigme de l'économétrie des séries temporelles en économie	60
2.1.5	Problématique de la cointégration des séries temporelles multivariées	62
2.1.6	Les origines diverses et multiples de la mémoire longue	63
2.1.7	Contexte général de l'apparition du phénomène de mémoire longue	66
2.1.8	Auto-similarité, accroissements indépendants, accroissements stationnaires	69
2.1.9	Méthodes, groupes et opérateurs de renormalisation	71
2.2	Economie et mémoire longue	73
2.2.1	Les théories économiques revisitées, eu égard au paradigme de la mémoire longue	73
2.2.2	Projet de Stiglitz de refonte intégrale de l'économie	79
2.3	Appréhension d'une réhabilitation LRD de l'utilitarisme de John Stuart Mill	83
3	An adaptive MIR estimator of the memory parameter and the goodness-of-fit test	87
	An adaptive MIR estimator	87
3.1	Introduction	88
3.2	The multidimensional increment ratio statistic and its statistical applications	89
3.3	Adaptive versions of the estimator and goodness-of-fit test	92
3.4	Simulations and Monte-Carlo experiments	95
3.4.1	Application of the IR estimator and tests applied to generated data	95
3.4.2	Comparison with other adaptive semiparametric estimator of the memory parameter	96
3.5	Proofs	99
4	New semiparametric stationarity tests based on adaptive MIR statistics	115
	Stationarity Adaptive tests	115
4.1	Introduction	116
4.2	The multidimensional increment ratio statistic	118
4.3	The adaptive version of the estimator	121
4.4	Stationarity and nonstationarity tests	123
4.4.1	A stationarity test	123
4.4.2	A new nonstationarity test	124
4.5	Results of simulations and application to Econometric and Financial data	125
4.5.1	Numerical procedure for computing the estimator and tests	125
4.5.2	Monte-Carlo experiments on several time series	125
4.5.3	Application to the the Stocks and the Exchange Rate Markets	127
4.6	Proofs	131

<i>TABLE DES MATIÈRES</i>	7
5 Conclusion et Perspectives	147
Conclusion et Perspectives	147
Bibliographie	149

Chapitre 1

Synthèse des travaux

Sommaire

1.1	Introduction	10
1.1.1	Quelques définitions	11
1.1.2	Motivation de notre travail	16
1.2	Épistémologie de la mémoire longue et son paradigme associé	18
1.2.1	Propension de la mémoire longue en économie	18
1.2.2	Carence de mémoire longue au sein des paradigmes économiques standards	20
1.2.3	Le paradigme de la mémoire longue	28
1.2.4	La mémoire longue à l'oeuvre en Économie : une fusion intégrale est en perspective	29
1.3	Estimation et tests adaptatifs pour des processus LRD	30
1.3.1	Présentation de la statistique unidimensionnelle de Ratios d'Incréments : $IR(m)$	31
1.3.2	Estimateur issu de la Statistique Multidimensionnelle de Ratios d'Incréments	32
1.3.3	La version adaptative de l'estimateur	33
1.3.4	Application à des tests statistiques	34
1.4	Mise en œuvre empirique des différentes procédures d'estimation et de tests	36
1.4.1	Simulations	36
1.4.2	Applications aux séries financières	37
1.5	Conclusion	38

1.1 Introduction

Ce chapitre de synthèse est consacré à la présentation des principaux résultats, inhérents aux travaux, détaillés aux chapitres 2, 3, 4 et dédiés aux approches et aux investigations, respectivement : épistémologique, statistique théorique et empirique économique, de la mémoire longue, générée par des processus, fortement corrélés (ou autocorrélés), stationnaires, ou à accroissements stationnaires. La mémoire longue ou la dépendance de longue portée “ *est un phénomène qui... se rapporte à la vitesse de décroissance de la dépendance statistique, (de type puissance)... plus lente qu’une décroissance exponentielle... (comme le sont) certains processus auto-similaires... mais tous les processus (de) dépendance de longue portée ne sont pas auto-similaires* ”¹. Il s’agit d’un “*phénomène ... largement observé dans la nature... caractérisé par des trajectoires affichant des trends apparents et des cycles, ... éphémères...* ”[51].

Les modèles de “*mémoire longue... à la fois dans le temps et dans l’espace* ”([48], p1) ont joué un rôle clé “*en sciences physiques au moins à partir de 1950, par des statisticiens dans divers champs comme l’hydrologie et la climatologie* ”([48], p1). Ils sont devenus incontournables en économétrie, “*à partir des années 1980 et en finance à partir de 1995* ”([48], p1) pour se populariser, et se généraliser, ainsi, à d’autres champs disciplinaires multiples et variés, “*comme l’agronomie, l’astronomie, la chimie, l’économie, l’ingénierie, les sciences environnementales, les géosciences, l’hydrologie, les mathématiques, la physique, et la statistique* ”([12], Preface). Plusieurs livres et revues lui ont été dédiés ([85], [4], [27], [83], [29],...) . Désormais, à l’image des neurosciences², des turbulences, des files d’attente, ou de la finance³, les théories écosystémiques (et géo-écologiques) des plus pertinentes adoptant les modèles de mémoire longue foisonnent et se multiplient.

En effet, rien que pour 2011-2012, qu’il s’agisse du développement durable⁴ ([30]), de l’énergie renouvelable [11], de l’écologie [69], du changement climatique [59], [19], [75], ou de la biodiversité [74], on détecte, au sein des données récoltées, dans chacune de ces disciplines respectives, de plus en plus, de phénomènes de “*mémoire longue... temporelle..ou... spatiale* ”([48]) avérés.

1. Wikipedia, avec adaptation

2. Long-Range Temporal Correlations in the EEG Bursts of Human Preterm Babies, Hartley C. and al, 2012.

3. “*j’ai été pris d’un goût extrême pour les bruits et turbulences (disait Mandelbrot). Grosso modo, il existe une ressemblance réelle entre ces deux domaines : le temps qu’il fait et les cours de la bourse sont également imprévisibles. Et j’allais trouver une traduction mathématique de cette idée banale en identifiant une notion sans précédent : l’idée de « self-similarité » géométrique. La self-similarité s’est avérée un concept fructueux, et si omniprésent que personne ne sait plus que c’est moi qui ai créé ce terme à Harvard en 1964. In La Recherche (n° 85, p. 5, janvier 1978 ; n° 171, p. 1334, novembre 1985)* ”

4. La «*conférence des Nations unies sur l’environnement humain* » de 1972 à Stockholm avait posé les fondations de «*l’éco-développement* ». Sa définition fut formalisée en 1987 dans le rapport Brundtland : «*Le développement durable est un mode de développement qui répond aux besoins du présent sans compromettre la capacité des générations futures de répondre aux leurs.* » in Le Figaro du 19 juin 2012. Sommet de Rio : le modeste bilan de la réunion de 1992.

Ces nouveaux domaines d'investigation de la mémoire longue, trouvent dans le modèle de régulation du flux du Nil de Hurst, présenté dans son fameux article “*long-term storage capacity of reservoirs*” [55], une modélisation statistique de queue, dans une file d'attente, réussie et facilement transposable. Il en est de même, pour la modélisation du trafic, sur réseaux Ethernet ([29], pp 373-407), qui s'en est fortement influencée, et qui devient, à son tour, une modélisation de référence, pour ces nouvelles disciplines écosystémiques, comme en témoignent, ces articles divers et multiples (cf. [57], [58], [72], [41], [68], [5]).

1.1.1 Quelques définitions

La mémoire longue, et son synonyme, la dépendance de longue portée, ou “persistance” ([83], p 1), est baptisée “*Effet Joseph*”, par Mandelbrot et ses disciples ([87], p 2), qui l'ont “*associée (délibérément, et ce depuis son origine) à la loi d'échelle et au comportement fractal (ou auto-similaire)*” ([87], p 2), et que Hurst appelait, lui “*mémoire de long-terme*”, expression inspirée, probablement, par la mémoire du lieu de sa découverte.

La mémoire longue “*occupe (aujourd'hui), un rôle clé, dans la modélisation des données macroéconomiques et financières*” ([83], p 1). La mémoire longue se présente sous des aspects variés, inhérents à la “*variété de points de vue*” ([87], Abstract) eu égard à un processus stochastique donné, et de ses manifestations multiples, dont traduisent ses “*onze définitions*” différentes relevées dans l'ouvrage de référence ([87], p1). Néanmoins, toutes “*ces définitions ont le désavantage de ne pas tenir compte, ni de la non stationnarité ni de la variance infinie*” ([50], p1). Ainsi “*Le point de vue le plus populaire, concernant la dépendance de longue portée, est de loin, (celui de) la décroissance lente des corrélations... restreintes aux processus stationnaires du second ordre*” ([87], p 29).

Remarque 1. *Depuis les théorèmes de décomposition de Wold et de Volterra, les notions de bruit blanc et de stationnarité jouent un rôle fondamental en statistique. Il existe principalement cinq types de bruits blancs et deux types de stationnarités : une stationnarité stricte (ou forte ou en distribution) et une stationnarité faible (ou du second ordre ou en covariance).*

Nous rappelons quelques définitions des notions qu'on désire utiliser à bon escient, dans la suite (cf. [28]).

Définition 1. (Bruit blanc)

Soit $(\xi_t)_{t \in \mathbb{Z}}$ un processus stochastique du second ordre ($\mathbb{E}(\xi_t^2) < \infty$). On appelle $(\xi_t)_{t \in \mathbb{Z}}$ un bruit blanc (avec la classification, dans la suite, allant du plus faible au plus fort)

1. faible si : $\mathbb{E}(\xi_t) = 0$, $V(\xi_t) = \sigma^2$, $cov(\xi_{t+k}, \xi_t) = 0$ pour tout t et tout $k \neq 0$
2. différence de martingale marginalement homoscedastique si : $\mathbb{E}(\xi_t | \xi_{t-1}) = 0$, $V(\xi_t) = \sigma^2$
3. différence de martingale conditionnellement homoscedastique si : $\mathbb{E}(\xi_t | \xi_{t-1}) = 0$, $V(\xi_t | \xi_{t-1}) = \sigma^2$

4. fort si les ξ_t sont indépendants, identiquement distribués et centrés. On les notera alors : $\xi_t \sim \text{IID}(0, \sigma^2)$
5. gaussien si les ξ_t sont gaussiens indépendants, identiquement distribués et centrés. On les notera alors : $\xi_t \sim \text{IIN}(0, \sigma^2)$

Remarque 2. On parlera de bruit blanc fort s'il est indépendant et identiquement distribué (IID) : la notion d'indépendance est plus forte que la nullité des autocorrélations, et le fait que le processus soit identiquement distribué est plus fort que la stabilité de ses deux premiers moments.

Définition 2. (Stationnarité forte)

Un processus $X = \{X_t, t \in \mathbb{Z}\}$ est dit stationnaire strict (ou au sens fort) si quelque soit $n, t_1, \dots, t_n$, et h on a : $(X_{t_1+h}, \dots, X_{t_n+h}) \stackrel{\mathcal{L}}{=} (X_{t_1}, \dots, X_{t_n})$ i.e. toutes les lois de dimension finie du processus translaté sont identiques.

Définition 3. (Stationnarité faible)

Un processus $X = \{X_t, t \in \mathbb{Z}\}$ est dit stationnaire du second ordre (ou au sens faible) si quelque soient les entiers relatifs t, s on a :

- $\mathbb{E}(X_t^2) < \infty$
- la moyenne du processus est constante et indépendante du temps : $\mathbb{E}(X_t) = m$
- les autocovariances ne dépendent que de l'écart (en valeur absolue) entre les dates des observations : $\text{cov}(X_t, X_s) = \gamma(|t - s|)$

Remarque 3. Les deux types de stationnarité faible et forte ne sont pas liés, en général. En effet on peut avoir la stationnarité forte comme pour les copies de processus SaS entre elles, sans avoir pour autant une quelconque stationnarité faible puisqu'elles n'ont même pas de variance. C'est le cas aussi des processus GARCH(p, q) définis par $X_t = \sigma_t \xi_t$, où ξ_t est IID(0, 1) et $\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$, ($\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$) $\forall i \in \{1, \dots, p\}, \forall j \in \{1, \dots, q\}$ et où la condition de stationnarité faible est définie par $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$, la condition de stationnarité forte est définie, elle, par $\mathbb{E} \log(\sum_{i=1}^p \alpha_i \xi_{t-i}^2 + \sum_{j=1}^q \beta_j) < 0$. A l'aide de l'inégalité de Jensen on déduit $\mathbb{E} \log(\sum_{i=1}^p \alpha_i \xi_{t-i}^2 + \sum_{j=1}^q \beta_j) < \log(\sum_{i=1}^p \alpha_i \mathbb{E}(\xi_{t-i}^2) + \sum_{j=1}^q \beta_j) = \log(\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j)$. C'est donc la stationnarité faible ($\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$) qui entraîne la stationnarité forte $\mathbb{E} \log(\sum_{i=1}^p \alpha_i \xi_{t-i}^2 + \sum_{j=1}^q \beta_j) < 0$ et non l'inverse.

Remarque 4. Exceptionnellement, les deux types de stationnarité sont liés, pour deux types de processus : les processus gaussiens et les processus stationnaires forts du second ordre. En effet :

1. Pour tout processus gaussien : les deux types de stationnarité coïncident.
2. Tout processus stationnaire fort du second ordre, est nécessairement stationnaire faible.

Théorème 1. (Théorème de Wold)

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire au second ordre. Alors (X_t) peut s'écrire sous la forme

$$X_t = c_0 + \sum_{i=0}^{\infty} \theta_i \xi_{t-i}$$

où la suite $(\theta_i)_{i \in \mathbb{N}}$ est convergente en L^2 et $(\xi_t)_{t \in \mathbb{Z}}$ est un bruit blanc au second ordre.

Théorème 2. (Théorème de Volterra)

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire au sens fort. Alors $(X_t)_{t \in \mathbb{Z}}$ peut s'écrire sous la forme

$$X_t = c_0 + \sum_{i=0}^{\infty} \theta_i \xi_{t-i} + \sum_{i,j=0}^{\infty} \theta_{ij} \xi_{t-i} \xi_{t-j} + \sum_{i,j,k=0}^{\infty} \theta_{ijk} \xi_{t-i} \xi_{t-j} \xi_{t-k} + \dots$$

où les suites $(\theta_i)_{i \in \mathbb{N}}$, $(\theta_{ij})_{i,j \in \mathbb{N}}$, $(\theta_{ijk})_{i,j,k \in \mathbb{N}}$, ... sont convergentes en L^2 et $(\xi_t)_{t \in \mathbb{Z}}$ est un bruit blanc fort gaussien.

Voici quelques exemples de processus stationnaires du second ordre, le bruit ξ_t étant un $IID(0, \sigma_t^2)$, avec $\sigma_t^2 < \infty$

- Les processus ARMA(p, q) et leur extension MA(∞), représentée par $X_t = \sum_{i \in \mathbb{N}} a_i \xi_{t-i}$, avec la condition $\sum_{i \in \mathbb{N}} |a_i|^2 < \infty$
- Les processus GARCH(p, q) et leur extension ARCH(∞) $\begin{cases} X_t = \sigma_t \xi_t, \\ \sigma_t^2 = b_0 + \sum_{j=1}^{\infty} b_j X_{t-j}^2 \end{cases}$, avec la condition $b_0 + \sum_{j=1}^{\infty} b_j^2 < 1$
- Les processus de Markov issus d'une mesure invariante.

Parmi les processus non stationnaires au second ordre les plus célèbres, notons :

- Les marches aléatoires, les processus linéaires (ARIMA, SARIMA, ...) etc...
- Les processus avec une tendance additive et/ou mutliplicative, les processus avec une saisonnalité, les processus avec un changement de régime...

Dans la suite nous donnons la définition principale d'un processus de longue mémoire ou processus de dépendance de longue portée (Long Range Dependence, en abrégé LRD)

Définition 4. Un processus stationnaire du second ordre (ou $\mathbb{L}^2$) est appelé processus de longue mémoire (ou à dépendance de longue portée (LRD)) si son autocovariance $\gamma(k)$ vérifie : $\lim_{n \rightarrow \infty} \sum_{k=0}^n |\gamma(k)| = \infty$. Il peut être défini par sa densité spectrale $f(\nu)$ vérifiant : $\lim_{\nu \rightarrow 0^+} f(\nu) = |\nu|^a L_3(|\nu|)$, avec $-1 < a < 0$ et $L_3(|\nu|)$ une fonction régulière à variations lentes.

Une dizaine d'autres définitions existent dans la littérature, pour ces LRD stationnaires ou à accroissements stationnaires, pour différents modèles : semi-paramétriques et non semi-paramétriques, gaussiens et non gaussiens.

Les cinq définitions suivantes (dont celle relative aux processus à accroissements stationnaires) sont les principales :

Définition 5. Un processus stationnaire (du second ordre) $(X_t, t \geq 0)$ (non nécessairement gaussien), d'autocovariance $\gamma_X(k)$, d'autocorrélation $\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)}$ et de densité spectrale $f_X(\omega)$, est un processus LRD ou à mémoire longue, de paramètre $-0.5 < d < 0.5$ ($d = H - \frac{1}{2}$) si l'une quelconque des propriétés suivantes est vérifiée ([29], p 14, 4.1-4.3, [12], p 42-43, 2.1-2.7) (extensibles avec adaptation aux processus à accroissements stationnaires à l'aide du filtre $(1 - L)$ (L étant l'opérateur retard $LX_{t+1} = X_t$))

1. Soient $\alpha \in (0, 1)$, $c_{\gamma_X} > 0$ et $L_1(k)$ une fonction à variation lente, telle que l'une des deux relations équivalentes suivantes, soit vérifiée :

$$\lim_{k \rightarrow \infty} \frac{\sum_{n=-k}^k \gamma_X(n)}{c_{\gamma_X} k^\alpha} = 1 \quad \Longleftrightarrow \quad \sum_{n=-k}^k \gamma_X(n) \underset{k \rightarrow \infty}{\sim} k^\alpha L_1(k)$$

2. Soient $\beta \in (0, 1)$, $c_{\rho_X} > 0$ et $L_2(k)$ une fonction à variation lente, telle que l'une des deux relations équivalentes suivantes, soit vérifiée :

$$\lim_{k \rightarrow \infty} \frac{\rho_X(k)}{c_{\rho_X} k^{-\beta}} = 1 \quad \Longleftrightarrow \quad \rho_X(k) \underset{k \rightarrow \infty}{\sim} k^{-\beta} L_2(k)$$

3. Soient $\delta \in (0, 1)$, $c_{f_X} > 0$ et $L_3(\omega)$ une fonction à variation lente, telle que l'une des deux relations équivalentes suivantes, soit vérifiée :

$$\lim_{\omega \rightarrow 0^+} \frac{f_X(\omega)}{c_{f_X} |\omega|^{-\delta}} = 1 \quad \Longleftrightarrow \quad f_X(\omega) \underset{\omega \rightarrow 0^+}{\sim} |\omega|^{-\delta} L_3(\omega)$$

$$\text{Idem pour } \omega \rightarrow \omega_0^+ \quad \lim_{\omega \rightarrow \omega_0^+} \frac{f_X(\omega)}{c_{f_X} |\omega - \omega_0|^{-\delta}} = 1 \quad \Longleftrightarrow \quad f_X(\omega) \underset{\omega \rightarrow \omega_0^+}{\sim} |\omega - \omega_0|^{-\delta} L_3(\omega)$$

4. Soient $\alpha \in (0, 1)$, $\beta \in (0, 1)$, $\delta \in (0, 1)$, $c_{f_{\Delta X}} > 0$, $c_{\gamma_{\Delta X}} > 0$, $c_{\rho_{\Delta X}} > 0$

$$\lim_{k \rightarrow \infty} \frac{\sum_{n=-k}^k \gamma_{\Delta X}(n)}{c_{\gamma_{\Delta X}} k^\alpha} = 1; \quad \lim_{k \rightarrow \infty} \frac{\rho_{\Delta X}(k)}{c_{\rho_{\Delta X}} k^{-\beta}} = 1; \quad \lim_{\omega \rightarrow 0^+} \frac{f_{\Delta X}(\omega)}{c_{f_{\Delta X}} |\omega|^{-\delta}} = 1;$$

$$\text{Idem pour } \omega \rightarrow \omega_0^+ \quad \lim_{\omega \rightarrow \omega_0^+} \frac{f_{\Delta X}(\omega)}{c_{f_{\Delta X}} |\omega - \omega_0|^{-\delta}} = 1$$

Si, en plus, la fonction d'autocovariance $\gamma_X(k)$ est ultimement monotone, ([29], p15, proposition 4.1) alors toutes ces définitions sont équivalentes, avec $\alpha = \delta = 1 - \beta = 2H - 1 = 2d$, et

$$L_3(x) = \frac{\alpha \Gamma(\alpha)}{2\pi} \sin(\frac{\pi\beta}{2}) L_1(\frac{1}{x}) = \frac{\Gamma(\alpha)}{\pi} \sin(\frac{\pi\beta}{2}) L_2(\frac{1}{x}), \text{ avec } c_{f_X} = \frac{\alpha \Gamma(\alpha)}{2\pi} \sin(\frac{\pi\beta}{2}) c_{\gamma_X} = \frac{\Gamma(\alpha)}{\pi} \sin(\frac{\pi\beta}{2}) c_{\rho_X}$$

La 4^{ème} définition de la mémoire longue est une extension au moyen du filtre Δ , aux processus fractionnaires non stationnaires ([87], Chapitre 6), de telle sorte qu'aujourd'hui “une série à mémoire longue signifie non seulement une série stationnaire, mais a fortiori, une série non stationnaire, dans laquelle la dépendance décroît, très lentement, au cours du temps” ([83], p 1).

Diebold et Inoue (2000) proposent une autre définition d'un processus LRD, capable de le distinguer d'un changement de régime, à partir de la vitesse de convergence de la variance de la somme partielle ainsi :

Définition 6. *L'ordre de convergence de la variance de la somme partielle $S_T = \sum_{t=1}^T X_t$ d'un processus LRD avec un paramètre de mémoire $-\frac{1}{2} < d < \frac{1}{2}$ est $\left(O(T^{2d+1})\right)$*

Dans la suite nous nous limiterons à l'ensemble des processus LRD stationnaires ou à accroissements stationnaires inhérents aux seuls modèles semi-paramétriques de paramètre de mémoire longue d tel que $d \in (-0.5, 1.5)$. Nous appelons cette classe, désormais, $SI(d)$. Il s'agit de l'ensemble des processus intégrés fractionnaires $X = (X_k)_{k \in \mathbb{Z}}$ pour $-0.5 < d < 1.5$ défini par :

Hypothèse $SI(d)$: $X = (X_t)_{t \in \mathbb{Z}}$ est une série chronologique s'il existe une fonction continue $f^* : [-\pi, \pi] \rightarrow [0, \infty[$ telle que :

1. pour $-0.5 < d < 0.5$, X est un processus stationnaire ayant une densité spectrale f satisfaisant

$$f(\lambda) = |\lambda|^{-2d} f^*(\lambda) \quad \text{pour tout } \lambda \in (-\pi, 0) \cup (0, \pi), \text{ avec } f^*(0) > 0. \quad (1.1)$$

2. pour $0.5 \leq d < 1.5$, $U = (U_t)_{t \in \mathbb{Z}} = X_t - X_{t-1}$ est un processus stationnaire ayant une densité spectrale f satisfaisant

$$f(\lambda) = |\lambda|^{2-2d} f^*(\lambda) \quad \text{pour tout } \lambda \in (-\pi, 0) \cup (0, \pi), \text{ avec } f^*(0) > 0. \quad (1.2)$$

Parmi toute la classe $SI(d)$, nous nous limiterons encore à la sous-classe des seuls processus gaussiens avec un comportement de la densité spectrale précisé plus haut. Nous appelons cette sous-classe $SIG(d, \beta)$: pour $-0.5 < d < 1.5$ et $\beta > 0$ définie par :

Hypothèse $SIG(d, \beta)$: $X = (X_t)_{t \in \mathbb{Z}}$ est un processus Gaussien tel qu'il existe $\epsilon > 0$, $c_0 > 0$, $c'_0 > 0$ and $c_1 \in \mathbb{R}$ satisfaisant :

1. pour $-0.5 < d < 0.5$, X est un processus stationnaire ayant une densité spectrale f satisfaisant pour tout $\lambda \in (-\pi, 0) \cup (0, \pi)$

$$f(\lambda) = c_0 |\lambda|^{-2d} + c_1 |\lambda|^{-2d+\beta} + O(|\lambda|^{-2d+\beta+\epsilon}) \quad \text{avec} \quad |f'(\lambda)| \leq c'_0 |\lambda|^{-2d-1}. \quad (1.3)$$

2. pour $0.5 \leq d < 1.5$, $U = (U_t)_{t \in \mathbb{Z}} = X_t - X_{t-1}$ est un processus stationnaire ayant une densité spectrale f satisfaisant, pour tout $\lambda \in (-\pi, 0) \cup (0, \pi)$

$$f(\lambda) = c_0 |\lambda|^{2-2d} + c_1 |\lambda|^{2-2d+\beta} + O(|\lambda|^{2-2d+\beta+\epsilon}) \quad \text{avec} \quad |f'(\lambda)| \leq c'_0 |\lambda|^{-2d+1}. \quad (1.4)$$

Les exemples suivants illustrent certaines propriétés relatives à ces processus intégrés fractionnaires gaussiens, stationnaires ou à accroissements stationnaires, les plus en vue, dans un cadre semi-paramétrique standard ([29], p 15-23) :

- Le mouvement brownien fractionnaire X_t (ou fBm) qui est un processus gaussien centré à accroissements stationnaires de variance $\gamma_X(t) = \sigma^2 |t|^{2d+1}$ avec $d \in (-0.5, 0.5)$ $\sigma^2 > 0$, $\Rightarrow X_t \in SIG(d+1, 2)$.
- Le bruit gaussien fractionnaire X_t (ou fGn) stationnaire centré de covariance $\text{Cov}(X_0, X_t) = \frac{1}{2}\sigma^2(|t+1|^{2d+1} + |t-1|^{2d+1} - 2|t|^{2d+1})$ avec $d \in (-0.5, 0.5)$, $\sigma^2 > 0 \Rightarrow X_t \in SIG(d, 2)$.
- Le processus FARIMA(p, d, q) gaussien avec $d \in (-0.5, 0.5)$ solution de l'équation $(I - L)^d \Phi(L) X_t = \Theta(L) \xi_t$ où $\Phi(L)$, $\Theta(L)$ sont deux polynômes de degrés respectivement p , q en L (L étant l'opérateur retard), de densité spectrale $f_X(\lambda) \sim \frac{\sigma^2}{2\pi} |\frac{\Theta(1)}{\Phi(1)}|^2 |\lambda|^{-2d}$, au voisinage de 0 $\Rightarrow X_t \in SIG(d, 2)$.
- Un processus gaussien de densité spectrale $f(\lambda) = \frac{1}{|\lambda|^{2d}} (1 + c_1 |\lambda|^\beta)$, $\beta \in (0, \infty) \Rightarrow X \in SIG(d, \beta)$.
Un tel processus peut être simulé par la procédure Paxson décrite dans la **section 1.4**

1.1.2 Motivation de notre travail

Nous avons développé une méthode adaptative d'estimation du paramètre d d'un processus LRD sous les hypothèses $SIG(d, \beta)$ générales qui ne préconisent que la forme asymptotique de la densité spectrale au voisinage de 0, inhérente à la mémoire longue.

Une conséquence intéressante, développée plus loin au chapitre 3, sera la mise en place d'un test de stationnarité adaptative dont les résultats numériques sont très satisfaisants.

Cette approche semi-paramétrique est mise en œuvre pour estimer le paramètre de mémoire longue, d tel que $(-0.5 < d < 1.25)$. Il s'agit de rendre la statistique univariée $IR_N(m)$ développée par Surgailis *et al*, (2008, [92]), multivariée $(IR_N(jm))$, pour pouvoir estimer, avec la meilleure qualité d'ajustement, la taille m de la fenêtre dans un modèle stationnaire semi-paramétrique. Une généralisation a pu être réalisée dans les mêmes conditions que précédemment, aux processus gaussiens à accroissements stationnaires.

Cette première partie théorique est assortie, ensuite, d'études empiriques comparatives avec trois autres estimateurs de la mémoire longue pour $(-0.5 < d < 1.25)$ que sont : l'estimateur $\tilde{d}^{MS}$ de Moulines et Soulier, dit aussi Fractionally Exponential $\tilde{d}^{FEXP}$ [70], l'estimateur $\tilde{d}^{ROB}$ dit aussi local de Whittle [81], l'estimateur $\tilde{d}^{WAV}$ des ondelettes du à Bardet et al (2008), [10]. L'estimateur $\tilde{d}^{IR}$ obtenu reste aussi bon si ce n'est pas meilleur que tous les autres estimateurs : $\tilde{d}^{MS}$, $\tilde{d}^{ROB}$, $\tilde{d}^{WAV}$ pour tous les échantillons, dont les résultats sont présentés plus loin, dans la **section 1.4**.

Nos résultats théoriques ont permis à la fois la construction d'un estimateur consistant mais aussi l'établissement d'un TCL pour tout processus à mémoire longue $I(d)$, aussi bien stationnaire que non stationnaire pour $(-0.5 < d < 1.25)$.

Deux articles détaillent les démonstrations de ces résultats dont le premier article [6] est publié au Journal JMVA(2012), le deuxième article étant soumis.

Les conséquences à la fois économétriques et économiques de ces résultats sont intéressantes et leurs répercussions sur l'épistémologie de la mémoire longue elle même sont immédiates. En effet : notre travail de thèse a permis de mettre en place :

- à la fois un estimateur adaptatif du paramètre de la mémoire d'un LRD et un intervalle de confiance qui le précise.
- des tests adaptatifs aussi bien pour la qualité d'ajustement que pour la stationnarité et la non stationnarité du processus.

Cela constitue de notre point de vue :

- une innovation statistique et économétrique : car à notre connaissance, il n'y a pas eu de méthode d'estimation semi-paramétrique robuste de la fenêtre m par optimisation d'un critère statistique construit à partir des données générées par un processus $I(d)$ pour $(-0.5 < d < 1.25)$ assorti de l'établissement d'un estimateur du paramètre de la mémoire longue mais aussi d'un TCL respectif.
- Une élaboration de deux nouveaux tests semi-paramétriques de stationnarité et de non stationnarité.
- une ouverture de nouveaux horizons économétriques prometteurs avec une généralisation probable aux processus linéaires. Un nouvel éclairage désormais paraît possible sur la cointégration fractionnaire, ...

Dès lors, il est désormais possible pour un praticien de :

- tester la stationnarité d'une variable économique à travers sa série chronologique.
- estimer le paramètre de la mémoire longue de cette série, si elle en a une, sans aucun prérequis paramétrique préalable.

De nouveaux regards sur les concepts, les fondements et les fondamentaux mêmes de l'économie, eu égard à ce phénomène de mémoire longue, sont désormais nécessaires. En effet :

- La mémoire longue établie pour les log-rendements en valeur absolue rend l'efficience (informationnelle) des marchés (lorsque $H = 0.5$), cette hypothèse centrale de l'analyse financière, obsolète (et ce pour tout $H > 0.5$). En effet une déconnexion durable, en cas de choc exogène, entre les cours de ces actifs financiers et leurs fondamentaux est désormais certaine. Ce schéma est généralisable, théoriquement, à tous les prix (selon la conception auto-similaire des prix des denrées de Mandelbrot évoquée plus loin).
- Toute l'épistémologie économétrique et économique, aussi bien de la racine unitaire, de la cointégration de deux variables et de leur relation d'équilibre de long-terme, se trouve porteuse, désormais, de l'empreinte du seul et unique phénomène de mémoire longue et de l'intégration fractionnaire $I(d)$ associée.
- La conception adaptative ([42], p 462) qui a déjà pris corps dans les anticipations des agents à la place des anticipations rationnelles antérieures, devraient plutôt se baser sur une optimisation d'un critère statistique, à l'instar de notre approche.

Dans cette synthèse nous commencerons par une introduction à l'épistémologie de la mémoire longue et à son paradigme associé, puis nous présenterons notre investigation des processus LRD semi-paramétriques sous les hypothèses $SIG(d, \beta)$ avec des procédures théoriques adaptatives à la fois d'estimation, mais aussi de test de la qualité d'ajustement, de stationnarité et de non stationnarité pour ces processus LRD. Enfin nous mettrons en œuvre ces différentes procédures théoriques en les appliquant à des séries chronologiques réelles et simulées avec les diagnostics respectifs qui s'imposent.

1.2 Épistémologie de la mémoire longue et son paradigme associé

1.2.1 Propension de la mémoire longue en économie

Il est désormais établi, à l'état actuel de la modélisation économétrique, qu'"en macroéconomie et en finance, la mémoire longue (appelée aussi dépendance de longue portée, ou dépendance forte ou persistance), occupe un rôle clé"([83], p1 contextualisé), aussi bien, en probabilité des processus stochastiques, qu'en statistique des séries temporelles.

Qu'en est-il, de la fondation économique, elle même, dans la pluralité de ses branches et la diversité de ses approches, vis à vis de ce phénomène de mémoire longue ?

Remarque 5. *Hormis la finance et assimilée, qui opèrent hors paradigmes via l'emploi d'outils purement techniques, neutres par vocation, il n'en est rien pour les autres branches de l'économie. En effet, chaque branche de l'économie est intimement liée, de fait à son propre paradigme, en quête de sens aux actions et aux politiques qui y sont menées. A de rares rares exceptions, comme pour les théories de la croissance endogène de l'école hétérodoxe, les paradigmes classiques et néoclassiques, en l'état, sont profondément statiques. Les paradigmes keynésiens et néokeynésiens, en l'état, bien que dynamiques, sont profondément de court terme. Il n'y a donc toujours pas de place pour une quelconque analyse économique, macroéconomique, ou économétrique, à mémoire longue, pour de très longs termes ou horizons, sans être en contradiction totale avec son propre paradigme.*

Remarque 6. *L'endogénéisation de la dynamique, par le passé et ce jusqu'en 1970, en économie, était purement déterministe. C'est à partir des travaux de Mandelbrot puis de Box Jenkins qu'on commence à prendre conscience de la nature probabiliste des processus et la spécificité de leurs dynamiques complexes. Ce type de dynamique probabiliste fractale fût-il de court terme n'existe dans aucun des dits paradigmes cités plus haut.*

Le témoignage de Mandelbrot, en la matière, en dit long : "En 1962, au moment où j'ai quitté IBM pour Harvard (disait Mandelbrot), je faisais alors de l'économie (à l'IBM) dans des conditions étranges, qui n'auraient jamais été soutenues financièrement ailleurs qu'à IBM. C'est comme

professeur d'économie qu'on m'avait invité à Harvard. Je m'occupais des choses très concrètes : j'avais introduit une idée qui semblait arbitraire, prise isolément, mais qui allait s'avérer être la base de la théorie des fractals. L'idée était que, dans l'étude des prix, il n'y avait aucune différence de nature entre les variations à court et à long-terme. On peut décrire, par exemple les changements du prix d'une denrée comme le coton sur quelques semaines ou sur plusieurs années comme deux phénomènes statistiquement identiques, sauf qu'ils se déroulent sur deux échelles différentes. Cela allait à l'encontre des idées reçues, qui voulaient que les variations quotidiennes soient dues à la spéculation, et les changements à long-terme aux lois fondamentales de l'économie"⁵.

Pourtant, de par la nature et les méthodes de collecte et de traitement des données économiques mêmes, l'occurrence d'opportunités d'une présence effective de mémoire longue, presque partout, en économie, est quasi certaine, avec son lot de jugement suspect et de conclusion douteuse. En effet :

1. Toute variable économique intégrée $I(d)$, en l'état, sans agrégation, avec $0 < d < 0.5$, tel un actif financier par exemple, intègre en son propre présent, tout son passé, de longue portée, ce qui pourrait causer, en cas de choc exogène, une déconnexion durable, avec sa valeur fondamentale et fausserait les calculs de l'équilibre microéconomique, avant d'atteindre l'étape même de l'agrégation macroéconomique.
2. La non stationnarité d'une variable au niveau microéconomique, très fréquente dans les séries économiques, qu'elle soit par l'effet de présence d'une tendance, d'une saisonnalité, ou d'un ou de plusieurs cycles, est stationnalisable par filtrages préalables successifs, en deux étapes, pour chacun, comme le veut la pratique conventionnelle, au moyen de polynômes de Gegenbauer([80], p 214). Néanmoins cette technique de stationnarisation préalable aurait des conséquences quasi-certaines sur la qualité d'ajustement statistique, dans un aller-retour, périlleux, plus que problématique pour la série filtrée : à la fois théorique et technique.
3. L'agrégation, sans changement adaptatif préalable d'échelles, a un effet intrusif de mémoire longue potentielle, en hautes ou en basses fréquences, et constitue un fait spécifique aux données macroéconomiques, avec des conséquences non négligeables, sur la consistance de leurs conclusions respectives, voire même sur la pertinence de leurs modèles macroéconomiques sous-jacents.
4. La cointégration, cette autre variabilité stochastique commune inter-variables (cette fois-ci et non intra-variable comme dans le cas précédent), de deux variables économiques intégrées non stationnaires individuellement, respectivement : $I(d_1)$ et $I(d_2)$ (avec $d_1 \geq 0.5$ et $d_2 \geq 0.5$) dont la combinaison linéaire, permet de déterminer, une relation stable de long-terme, exprimant, en dépit de leur non stationnarité individuelle, leur tendance stochastique commune.

Ainsi, l'appréhension des déséquilibres, suite à un choc exogène, que pourrait avoir lieu, une seule variable économique $I(d)$ avant son retour à son sentier d'équilibre d'origine, par rapport à son

5. in La Recherche (n° 85, p. 5, janvier 1978 ; n° 171, p. 1334, novembre 1985)

propre passé, voire deux variables cointégrées, $I(d)$ pour $d \geq 0.5$, avec une tendance stochastique commune, reste fondamentale en économie, dans la recherche de leur équilibre de long-terme. De tels déséquilibres risquent d'entraver inexorablement les échéances et causer des frictions structurelles voire des mises en cause de la véracité même, du modèle macroéconomique sous-jacent, proposé.

Or toutes les variables macroéconomiques sont des variables agrégées, en l'état (appelées agrégats), avec leurs lots de mémoires longues, effectives ou fallacieuses, selon le type de processus qui les génère comme certains processus AR de lois Béta (et autres). Ce phénomène comportemental est connu et reconnu, pour des données, récoltées différemment selon les modes de hautes ou de basses fréquences.

Dès lors, il reste à mesurer, et en premier abord, la distorsion des résultats d'analyse économique, menée à partir d'agrégats hétéroclites en terme de modes en hautes ou en basses fréquences.

En outre, la non stationnarité, cette fois, au niveau macroéconomique, relativement fréquente au sein des agrégats économiques, avec de fortes propensions de longue mémoire, en macroéconomie, peut provenir de la présence d'une tendance (comme le revenu permanent), d'une saisonnalité (comme les variations saisonnières) ou d'un ou plusieurs cycles économiques (comme le cycle des affaires), au sein de l'agrégat en question.

Dans de telles circonstances, et en premier abord, toute l'infrastructure microéconomique, relayant le calcul optimal, tel quel, au niveau macroéconomique, comme les modèles d'équilibre général dynamique DGE (Dynamic General Equilibrium), se trouve fortement remise en cause.

1.2.2 Carence de mémoire longue au sein des paradigmes économiques standards

L'épistémologie de l'économie, habitée par un esprit critique opiniâtre et persévérant, se réduit, de fait, à l'épistémologie des paradigmes et des communautés épistémiques économiques. En effet, combien même *“l'esprit critique est indispensable pour la recherche et la démarche scientifique. La réflexion épistémologique (réflexion sur les paradigmes et sur les groupes de la science normale socialement constitués autour de ces paradigmes) en est un élément incontournable... (elle) intervient... afin d'éclaircir les idées, d'orienter le débat et d'enrichir les connaissances ... l'épistémologie et la méthodologie (ces deux éléments de base) sont nécessaires pour mieux cerner le statut et le contenu de la connaissance économique([52], p 3).*

Mais en quoi la mémoire longue constitue t-elle un nouveau paradigme pour la macroéconomie et la finance ? Tout d'abord qu'est ce qu'un paradigme ?

Définition d'un paradigme

Une définition épistémologique, synthétique et opérationnelle d'un paradigme, est proposée par Hinti ([52], p 163) :

Définition 7. *Un paradigme est un ensemble de concepts, de croyances, de lois et de théories qui servent de modèle de référence à une communauté scientifique... , à un moment donné. Autrement dit un paradigme est un ensemble de principes généraux pour un cadre intellectuel ou idéologique à l'intérieur duquel se déroule un débat scientifique. C'est aussi une sorte de constellation complète de valeurs, techniques, croyances, communes aux membres d'une collectivité scientifique donnée et sur la base desquelles, théories et modèles peuvent se développer ([52]).*

Nous aurons, ainsi, autant de paradigmes que de *théories-cadres économiques*, basées chacune sur un “référentiel épistémique” et une “communauté épistémique” propres. En “adopt(ant) l'idée de matrice disciplinaire de Kühn (depuis l'Antiquité et ce jusqu'à nos jours ([52] p 166)), il est possible de dégager... , six étapes majeures, (en économie) qui peuvent ressembler au paradigme“([52], p 166). On a :

“1— *Le paradigme à connotation philosophique, allant de l'antiquité au moyen âge : l'économie est issue de la philosophie, de la morale... , la richesse n'est pas une finalité en soi.*

2— *Le paradigme des mercantilistes, du 16^e au 17^e siècles : les sources de la richesse sont la monnaie et les métaux précieux : or, diamant, etc... Ce paradigme fut dominant en Espagne, au Portugal, en France et en Angleterre.*

3— *Le paradigme des physiocrates, à partir du 18^e siècle : la terre agricole est la principale source de richesse. Quesnay, médecin, voit l'économie comme un organisme vivant dans lequel circule le surplus agricole, d'où sa conception organique du circuit économique“([52], p 166).*

Il n'en demeure pas moins qu’*“il est de coutumes de présenter la science économique comme comportant quatre courants (paradigmes) principaux : classiques, néoclassiques, keynésiens et contemporains.”([52], p 166).*

Paradigme de la théorie économique classique : 1790-1870

La théorie économique classique était fondée principalement par les économistes anglais : *Adam Smith (1723-1790), David Ricardo (1772-1823), Thomas Malthus (1766-1834), John Stuart Mill (1806-1873)* et les français : *Étienne Bonnot de Condillac (1715-1780), Anne Robert Jacques Turgot (1727-1781), Jean-Baptiste Say (1767-1832).* Karl Marx (1818-1883), adepte de Ricardo, se définit lui même, mais aussi par Shumpeter, comme un classique.

Dans “*le paradigme des classiques : l'économie politique est la recherche des lois de la reproduction et la consommation des richesses. Autrement dit, tout ce qui satisfait les besoins de la consommation et de la production. L'appropriation de la nature pour la transformer et l'adapter à nos besoins dépend du travail. Leur paradigme est composé de matrices telles que la recherche et l'harmonisation des intérêts individuels, le travail comme facteur de production (valeur-travail)... Les idées classiques*

étaient dominantes en Angleterre pendant longtemps, puis contestées et remplacées vers la fin du 19^e par le marginalisme (néoclassique). “⁶

Les principes fondateurs (ou postulats) du paradigme classique, dans sa première version, profondément “*optimiste(s) et progressiste(s) (quant) à l’intelligence humaine*”⁷ sont : l’invariance, l’exclusion de la causalité, la souveraineté du monde réel sur le monde monétaire, la souveraineté des choix conscients⁸, ou d’une manière plus explicite :

1. Il existe une nature humaine invariante.
2. La monnaie a un caractère totalement négligeable.
3. Tous les êtres humains participent ensemble à déterminer l’état du monde économique et social en l’absence de tout pouvoir.
4. La nature humaine, si elle est éclairée, peut parvenir à l’état optimal qui s’impose au pouvoir.

Paradigme de la théorie économique néoclassique (marginaliste) (1870-(1929 ?1970))

Les pères fondateurs du courant néoclassique (marginaliste) dont les œuvres vont révolutionner l’économie classique sont :

- Léon Walras(1834-1910), avec ses “*Eléments d’économie politique pure (Lausanne, 1874)*”.
- William Stanley Jevons(1835-1882), avec sa “ *Théorie de l’économie politique (The Theory of Political Economy, Manchester, 1871)*”.
- Carl Menger(1840-1921), avec ses “ *Principes d’économie (Grundsätze der Volkswirtschaftslehre, Vienne, 1871)*”.

Cette deuxième version (marginaliste) du paradigme classique, qui va compléter la première, s’intéresse plus au marché et “*ne s’intéresse pas à l’individu en tant que tel. Le marché (régulateur dynamique ([21] p 85) permettant de déterminer des situations d’équilibre (idem p 68)) est un éclairage donné en présentant d’emblée les résultats de choix collectifs*”([21], p 66) “*Ainsi, la seconde représentation classique par l’intermédiaire du marché ouvre un passage spontané à une analyse dynamique en termes d’élasticité de l’offre et de la demande* “([21] p 86) sur tous les marchés des biens et services concourant en cela à un équilibre général toujours accessible, constituant “*un point de référence ou d’attraction. (En effet) Même si les équilibres partiels ne s’obtiennent pas simultanément, l’équilibre général (lui) continuerait de s’imposer comme un appel permanent au fonctionnement (concurrentiel pur et parfait) d’une économie nationale* “([21], p 88). Ainsi : “*Dans l’acception la plus courante... (le paradigme de) l’économie néoclassique se fonde sur quatre postulats :*

1. *Les phénomènes économiques peuvent et doivent être étudiés à l’aide des mêmes méthodes que les phénomènes physiques ;*
2. *Les agents sont rationnels, leurs préférences peuvent être identifiées et quantifiées ;*

6. idem

7. Les quatre piliers de la science économique, Cotta A. et Calvet C., Ed Fayard 2005, p 42

8. idem, p 41

3. Les agents cherchent à maximiser l'utilité des biens consommés, tandis que les entreprises cherchent à maximiser leur profit ;
4. Les agents agissent chacun indépendamment, à partir d'une information complète et pertinente. (Ce postulat est récusé par l'école autrichienne, qui est fidèle au dualisme méthodologique et utilise une conception plus faible de la rationalité, et par certains autres courants qui peuvent parfois quand cette hypothèse seule est relâchée être considérés comme néo-classiques) "9.

Autrement dit : "(Dans) Le paradigme des néoclassiques : l'économie politique est la recherche des lois du calcul optimal des agents économiques pour une meilleure allocation des richesses. L'essentiel de leurs analyses a pour but la détermination des prix relatifs des biens, des services et des facteurs de production. Leur paradigme fait référence à des éléments tels que la valeur-utilité, la rémunération des facteurs de production à leur productivité marginale, l'optimisation... La révolution marginaliste est une vraie révolution (dans une optique) kühnienne parce qu'elle a provoqué des changements profonds dans la théorie économique" ([52], p 166).

Paradigme de la théorie économique néo-classique microéconomique

Le développement macroéconomique, opéré depuis 1954 par les néo-classiques, en riposte au keynésianisme ambiant, néo-classicisme caractérisé à la fois par son fameux rendement d'échelle constant inhérent à l'homogénéité d'ordre un de sa fonction de production de type Cobb-Douglas : $Y(K, L) = aK^\alpha L^\beta$ tel que $\alpha + \beta = 1$, mais aussi par sa théorie du taux d'intérêt classique et le rôle central attribué à sa capacité capitalistique $\frac{K}{L}$ au cœur de sa représentation, a permis "d'installer définitivement le paradigme classique dans une perspective dynamique qui pourra lui assurer sa longévité tout en affirmant et en apportant la preuve du déterminisme économique de l'évolution "10. Le classicisme a pris une forme plus actuelle avec Debreu et Hicks, le premier pour sa "nouvelle théorie de la valeur en 1959, où il s'efforce d'établir une identité formelle entre l'économie du certain et de l'incertain... (via l'introduction) de variables d'écart "11, le second pour son introduction de biens et services objet(s) de négociation de livraison différée et conditionnelle selon des mécanismes semblables à des marchés à terme ou de futures" dès lors "La représentation classique se réduit à de l'algèbre, et la science économique... (à) une axiomatique algébrique qui s'opère sur des êtres algébriques "12. Ainsi, et "au terme de (tous) ces efforts, le paradigme classique est totalement constitué dans l'ordre théorique et statistique tout en affirmant sa fidélité aux principes de base ... , et ce, toujours sans faire la moindre référence à la monnaie "13, sans agrégation aucune, puisque tout le raisonnement classique est basé sur la rationalité de deux agents représentatifs : l'un pour tous les consommateurs, l'autre pour tous les producteurs. L'État, dans la vision classique et néo-classique n'étant qu'un agent parmi d'autres.

9. école néoclassique, Wikipedia avec adaptation

10. Les quatre piliers de la science économique, Cotta A. et Calvet C., Ed Fayard 2005, p 101-102

11. idem p 106-107

12. idem p 107

13. idem, p 107

Paradigme de la théorie économique de Keynes

Dans *“Le paradigme keynésien :... l'économiste s'intéresse de plus près aux quantités globales : revenu national, consommation nationale, épargne globale. Le succès remporté par les idées de Keynes à partir des années trente est assimilé à... (une) révolution scientifique (selon la conception kühnienne) : on parle alors de... révolution keynésienne”*¹⁴

Les principes fondateurs du paradigme keynésien, nés *“après la parution de La Théorie générale (1936)”*¹⁵ et conçus pour le court terme. En effet *“la cinématique keynésienne est du domaine du court terme. Elle s'intéresse au passage d'une situation déterminée à une autre, entre deux temps, et non pas à l'évolution générale, à la trajectoire ou à la dynamique de long-terme”*¹⁶. Le paradigme keynésien s'oppose, à la fois, au paradigme classique (statique¹⁷, du laisser-faire sans faire de politique économique), classicisme qui stipule que toute économie libérale, évolue inéluctablement vers un équilibre général, avec résorption, à l'équilibre, du chômage. Il est aussi et surtout, en opposition au paradigme marxiste dynamique¹⁸, qui stipule que toute économie capitaliste, évolue inéluctablement, par le biais de *“la loi de la baisse tendancielle du taux de profit”*¹⁹, sur le long-terme vers le communisme égalitaire, avec résorption, à coup de révolutions, expressions de la dialectique historique matérialiste, résorption à la fois des intérêts privés et des classes et la fin du capitalisme. Ainsi *“Le positionnement de la politique de Keynes (il en fallait une, devenue une nécessité, pour endiguer l'activité économique et résorber le chômage) se situe entre le laisser-faire (classique) et la préparation de la révolution (Marxiste)”*²⁰. C'est *“au nom de l'efficacité et de la liberté”*²¹ que Keynes va *“construire un autre paradigme justifiant l'action”*²² de l'État dans le domaine économique, monétaire et comptable. D'où :

1. La nécessité pour l'État d'avoir une politique économique.
2. La nécessité pour l'État d'avoir un contrôle étroit à la fois de l'émission et de la création de la monnaie, centrale et bancaire et une politique directrice des taux d'intérêt pour toutes les échéances. En effet *“Le monisme de Keynes s'exprime d'une façon simple. C'est la quantité de monnaie qui est responsable du chômage, et la relation causale est en sens unique”*²³.
3. La nécessité pour l'État d'avoir une comptabilité nationale :
“Keynes invente le concept mesurable... en économie fermée, ... (et) retient six concepts (dont il faut suivre étroitement l'évolution, d'abord deux variables monétaires : La masse monétaire

14. Hinti Saïd, La science économique : une réflexion épistémologique et méthodologique, imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p 167

15. idem p 243

16. idem p 251

17. idem p 251

18. idem p 251

19. idem, p 245

20. idem, p 245

21. idem, p 244

22. idem, p 244

23. idem, p 247

M, le taux d'intérêt r , puis quatre variables réelles que sont : le produit effectif ou produit national Y , la consommation effective des biens C , l'investissement ou formation du capital fixe I , l'épargne ou le revenu non consommé S ²⁴ avec des hypothèses keynésiennes, sur la nature de chacune de ces six variables pour pouvoir mener à bien une politique économique et monétaire efficaces via ses fameux multiplicateurs à la fois monétaires inhérents à l'offre de monnaie et budgétaire inhérents à (l'offre) des dépenses publiques.

Synthèse keynésienne-néoclassique

La synthèse keynésienne-néoclassique, en vogue depuis les années 1970, bien que hors de notre propos, parce que, hybride en terme de paradigmes, ne traite pas de l'évolution du système économique, hormis l'analyse purement technique inhérente à l'analyse macroéconométrique des séries temporelles et plus particulièrement des modèle VAR ou Vectoriels AutoRégressifs comme le montre l'exemple traité dans la *section 10.4 Chocs et multiplicateurs* de l'ouvrage de Gouriéroux et Monfort ([42], p385 et suite).

Carence de dynamique de long-terme au sein des théories : classique, néoclassique, keynésienne et néokeynésienne

Comme il a été mentionné précédemment, dans les **Remarques : 5 et 6**, la théorie classique et néoclassique étant fondamentalement statiques, il ne peut y avoir en leur sein de dynamiques ni de court terme ni de long-terme. Hormis le courant hétérodoxe, la théorie keynésienne et néokeynésienne étant fondamentalement dynamique, de court terme, le long-terme est donc absent de l'analyse keynésienne et néokeynésienne. Les anticipations des agents pour un agrégat y sont de simples moyennes de longue période, de l'agrégat en question. Il ne peut y avoir, donc, en cas de présence de mémoire longue effective ou fallacieuse, d'analyse vigilante et perspicace de l'évolution dynamique, à la suite d'un choc exogène, de la trajectoire, des divers équilibres transitoires successifs possibles, du système économique, tendant vers son équilibre final, selon les unes et les autres des théories, aussi bien classique, néoclassique, keynésienne que néokeynésienne.

Retour sur la scientificité de l'économie

Dans une vision critique sceptique, l'économie reste, selon ses détracteurs, en dépit de tout, une simple traduction théorique de "*débats d'idées*". Les approches, à supposer méthodologiques des économistes, selon ces sceptiques, ne sont pas nécessairement fondatrices de sciences ou de structures scientifiques, en dépit de leur parenté apparente avec la recherche optimale. Les économistes ne sont jamais parvenus réellement, sur de longues périodes, disent leurs détracteurs, au bien-être d'une nation, en biens et services pour ses citoyens, sous contrainte de rareté de ses ressources.

24. idem, 247-248

Or si on admet avec Donald Black ([13]) que *“la science est une question de degré : la scienticité”* en tant que *“fonction curviligne de la distance sociale du sujet”*²⁵ attribuée aussi bien à une idée, à une méthodologie, à une théorie ou à une fondation et que *“la scienticité d’une idée croît avec sa testabilité, sa généralité, sa simplicité, sa validité et son originalité”* ([13]) dans un espace social déterminé, muni d’une structure topologique adéquate, dès lors comment peut-on définir la *“scienticité”* d’une telle science ? Par son objet, par sa méthode, par son affiliation, par sa théorie de causalité, entre cause et effet, ou selon Duhem, par la classification, sur l’échelle de signification, de sa représentation mathématique²⁶ ? Comment peut-on définir la scienticité d’une théorie ? Que peut-on espérer d’une théorie scientifique ? *“Nous fait-elle connaître les régularités de la nature (physique, sociale, économique) ou cherche-t-elle également leurs causes ? Les lois scientifiques sont-elles immuables ou évoluent-elles au cours du temps ? Quel rôle les mathématiques jouent-elles dans la connaissance de la nature (physique, sociale, économique) ? Comment les connaissances se rapportent-elles à ce qui nous est donné, ou à ce que nous éprouvons dans une expérience vécue ? Comment concevoir la possibilité d’une vérification des connaissances scientifiques ? Qu’est ce que l’interprétation d’une théorie ? Les entités théoriques sont-elles réelles ou ne sont-elles que des fictions commodes ? Comment les théories peuvent-elles être comparées les unes aux autres ? La philosophie des sciences (naturelles, sociales ou humaines) produit-elle elle-même des connaissances ou a-t-elle pour but d’analyser le sens des énoncés de la science ?”*²⁷.

Le grand économiste Schumpeter soulève la question pertinente en tout temps de la scienticité de l’économie dans sa formulation du débat passionnant suivant : *“La réponse à la question (l’économie est-elle une science ?) dépend de ce que l’on entend par “science”. Ainsi, dans le langage quotidien aussi bien que dans le parler académique... , le mot est souvent employé pour renvoyer à la physique mathématique. Ce qui exclut évidemment toutes les sciences sociales et aussi l’économie. L’économie dans son ensemble, n’est pas non plus une science si nous faisons de l’emploi de méthodes semblables à celles de la physique mathématique le caractère spécifique de la science. Dans ce cas, seule une faible partie de l’économie est scientifique, en certaines de ses parties et non en d’autres. Les susceptibilités concernant le “rang” ou la “dignité” ne devraient pas trouver là à s’exercer : appeler un domaine donné “science” ne devrait pas faire figure de compliment ni de reproche”* ([88], p 30). Schumpeter donne plusieurs définitions de la science, reprises reformulées et appuyées par Hinti dans son ouvrage *“La science économique : une réflexion épistémologique et méthodologique”* ([52], p 82 et suite) ainsi :

1. *La science est une connaissance transmissible, extensible, systématique.*

25. idem

26. Dictionnaire d’Histoire et Philosophie des Sciences, Sous la direction de Lecourt D., Ed PUF 1999, concept=théorie, p 941

27. Philosophie des sciences, Naturalismes et réalismes , Tome 2, Textes réunis par S. Laugier et P. Wagner, Ed Vrin, 2004, Préface p :7-8

2. *La connaissance scientifique est objective.*
3. *La science étudie les phénomènes .*
4. *La science établit des relations universelles : les lois .*
5. *La science permet des prévisions .*

À la lumière des cinq définitions Schumpéteriennes, présentées plus loin à la sous section 2.1.2 du chapitre 2, sensées être toutes équivalentes, Schumpeter va examiner l'économie ainsi. "*Puisque l'économie emploie des techniques dont le grand public ignore le maniement, et puisque les économistes sont là pour les cultiver , l'économie est évidemment une science, en conformité avec notre définition du mot (science)*" ([88], p 30).

Mémoire longue en finance

En dépit des avancées théoriques spectaculaires, relatives à la notion de mémoire longue depuis pas moins de trois décennies (Granger et Joyeux (1980, [45]), Hosking (1981, [54]), Porter-Hudak(1983, [35]), Fox et Taquq(1986, [34]), Dahlhaus(1989, [26]), Guégan et al(1991, [15]), Robinson(1995, [82], [81]), Abry et al (1998, [2]) etc... , il n'en demeure pas moins qu'*en finance il reste (essentiellement) un sujet de recherche empirique*²⁸. En effet, rares sont les économistes qui ont tenté de justifier théoriquement son existence en économie et en finance par des mécanismes purement économiques, hormis les deux modes techniques de sa génération : l'agrégation et l'intégration stochastique par rapport à un mouvement brownien usuel. Pourtant d'éminents spécialistes de l'économie s'accordent aujourd'hui à souligner son importance, ainsi : "*Les travaux empiriques en macroéconomie et en finance ont été témoins d'une renaissance aussi bien théorique qu'empirique de l'économétrie de la mémoire longue et de l'intégration fractionnaire*"[48] parce qu'"*Il y a une évidence substantielle que les processus de longue mémoire décrivent les données financières plutôt bien, tels que les primes à l'avance, la différentielle : de taux d'intérêt, de taux d'inflation et de taux de change*" [48]. D'autres affirment encore qu'il s'agit, plus encore, de l'émergence, dans un sens kühnien, même relativisé, d'un "*nouveau paradigme de la macroéconomie et de la finance*"([29], pp 417-438)".

Cette difficulté à expliquer la mémoire longue, par des mécanismes intrinsèques n'est pas spécifique à l'économie mais constitue plutôt un fait général qu'on peut retrouver ailleurs dans tout type de données à dépendance forte.

En effet, on n'a toujours pas de réponses définitives aux débats qui animent les spécialistes de la mémoire longue sur l'effectivité d'une nature intrinsèque spécifique aux données à dépendance forte. Dès lors "*la bonne qualité d'ajustement d'une LRD sur les données n'a pas résolu (encore) le débat sur la présence effective d'une telle LRD dans ces données*"([29], p 469), ou que la persistance est-ce bien le fait d'une mémoire longue seule ou en contraste avec d'autre processus gouvernant les changements de régimes stochastiques latents. En d'autres termes "*l'effet de Hurst peut il provenir*

28. "Long-memory in finance is still an empirical research topic" [http ://fedc.wiwi.hu-berlin.de/xplore/tutorials/xaghtmlframe209.html](http://fedc.wiwi.hu-berlin.de/xplore/tutorials/xaghtmlframe209.html)

d'un processus non stationnaire caché"[48]. "Cependant l'histoire du phénomène de Hurst paraît loin d'être parachevée"[48]. De tels débats se retrouvent transposés, à leur tour, en économétrie comme l'évoquent Mikosch et Stărcia ([29], p 456) ainsi :

"Néanmoins, le sens commun peut permettre de se poser au moins deux questions : Quelles sont les raisons économiques qui justifient l'existence d'une LRD en valeur absolue des log de rendements ? Qu'est ce que nous aurions gagné si nous avions connu (au préalable) qu'il existe une LRD en valeur absolue des log de rendements ?"

L'investigation statistique de la mémoire longue (exposée plus loin, dans les sous-sections 2.1.7-2.1.8 du chapitre 2) a débuté vers les années 1950 en hydrologie avec Hurst, et portée par les probabilités, grâce aux travaux pionniers de Mandelbrot dès les années 1960, puis ceux de Taqqu, son disciple depuis les années 1970. Depuis les années 1980, c'est grâce aux travaux précurseurs de Granger, que la mémoire longue est devenue, de facto, un centre d'intérêt de l'économétrie et de la finance. Or l'économétrie, ce lieu de rencontre entre la statistique mathématique, et la théorie économique au service de l'économie à vocation prévisionnelle est devenue un instrument majeur de l'analyse économique, même si la macroéconométrie n'a pas eu tout le succès escompté qu'on attendait d'elle. L'analyse des séries temporelles, qui est aujourd'hui au cœur de nombreuses analyses dynamiques s'est avérée une science à part entière très fructueuse indépendamment de la véracité des doctrines économiques qui les adossaient.

Cela reste également vrai, pour la finance, cette branche très dynamique de la recherche en économie de marché dont le diagnostic des plus pertinents, a émis de sérieuses réserves, sur l'efficacité des marchés financiers, nationaux et internationaux. En effet, spécialisée de nos jours, dans le négoce des instruments multiples et le transfert des anticipations diversifiées de revenus et de risques, la gestion des nouveaux produits financiers contraste de plus en plus avec l'hypothèse fondamentale d'efficacité des marchés financiers, de plus en plus hypothéquée. De fait : *"En finance, plus précisément, la présence de mémoire longue a d'importantes implications sur la théorie financière moderne et notamment sur la théorie d'efficacité des marchés financiers. L'existence d'une mémoire de long-terme permet d'expliquer la possibilité d'une déconnexion durable entre les cours et les fondamentaux et de rendre ainsi compte des délais d'ajustement des prix à l'information"*²⁹.

1.2.3 Le paradigme de la mémoire longue

Le paradigme de l'invariance d'échelles

Peut-on attribuer à la branche de la mémoire longue inhérente à tous ces phénomènes critiques à invariance d'échelle ou auto-similaires ([60], p 51), un paradigme unifié spécifique qu'on peut appeler *"paradigme d'invariance d'échelle"*, entendu dans le sens kühnien de *matrice disciplinaire* dont les éléments stratégiques sont à la fois irréductibles et fondateurs, selon la théorie générale de Roth et Bourguine ([17] et [86]), de *communautés épistémiques* dans toutes les activités

29. Mérièm CHAOUACHI, work paper, ERMES, 2005

scientifique et dans toutes les sciences et savoirs d’une époque. En effet, le passage suivant identifie clairement ce paradigme d’invariance d’échelle sous le vocable de “*principes unificateurs*” ainsi : “*Les lois d’échelle sont omniprésentes dans la nature, et elles envahissent les activités neuronales, comportementales et linguistiques. Une loi d’échelle suggère l’existence de processus ou de motifs qui se répètent à différentes échelles d’analyse. Bien que les variables qui expriment une loi d’échelle puissent varier d’un type d’activité à l’autre, la récurrence des lois d’échelle à travers autant de systèmes différents a conduit à rechercher des principes unificateurs. Dans les systèmes biologiques, des lois d’échelle peuvent refléter des processus adaptatifs de différents types et sont souvent liés à des systèmes complexes qui s’apprêtent à proximité des points critiques. La même chose est vraie pour la perception, la mémoire, le langage et d’autres phénomènes cognitifs. Les résultats de lois d’échelle dans les sciences cognitives sont à titre indicatif de l’invariance d’échelle dans les mécanismes cognitifs et les interactions multiplicatives entre des composantes interdépendantes de la cognition*” ([18], Abstract).

Indépendamment de la *permanence d’un doute irréductible*³⁰ quant à la *scientificité de l’économie politique*³¹, voir d’une *conviction mal partagée*³², entre “*les communs*” des économistes qui opèrent en son sein, et qui ont toujours eu une distinction conceptuelle et méthodologique canoniques claire et nette *du positif du normatif et de l’art*³³ d’un côté, et leurs adversaires, les économistes positivistes, qui préfèrent, eux, se référer aux pères fondateurs de l’économie politique et au plus célèbre d’entre eux : Keynes, qui préconise, quant à lui, que : *la science positive... corps de savoir systématisé concernant ce qui est ; la science normative ou régulatrice... corps de savoir systématisé discutant les critères de ce qui devrait être ; et l’art... système de règles pour l’obtention d’une fin donnée*³⁴. Deux conceptions contradictoires de l’économie traduisant la nécessité d’un compromis, en quête d’une homogénéité de l’ensemble du dispositif, deux courants de pensée économique ne pouvant ni s’ignorer ni s’exclure, parce que complémentaires, au sein d’un même *paradigme* selon la conception de Thomas Kühn ou *programme scientifique de recherche* selon la conception d’Imre Lakatos ou *épistémé* selon la conception de Michel Foucault, deux courants de pensée économique articulés autour d’une même stratégie globale et au service d’un même système qui les a toutes les deux, sinon planifiées, du moins pensées, en “multitâches”.

1.2.4 La mémoire longue à l’oeuvre en Économie : une fusion intégrale est en perspective

Comme on a pu le constater précédemment, la mémoire longue temporelle et spatiale est à l’oeuvre aujourd’hui en économie et plus particulièrement en finance, identifiable par ses aspects caractéristiques

30. Benetti, Carlo et Cartelier, Jean in L’économie devient-elle une science dure? Série : grands débats, Ed Economica 1995, p224

31. idem p224

32. idem p 216

33. Hubert Brochier in L’économie normative, Série : grands débats, Ed Economica 1997, p9

34. idem p9

comme la persistance et autres faits caractéristiques propres. Parmi les mécanismes qui la génèrent, deux mécanismes sont bien identifiés que sont l'agrégation et l'intégration stochastique par rapport à un mouvement brownien usuel. Il s'agit d'un processus statistiquement identifiable (par la vitesse de convergence de sa somme partielle comme l'a suggéré Diebold et Inoue(2000) par exemple) par rapport à d'autres phénomènes persistants semblables comme certains changements de régimes stochastiques ou certains types de chaos à comportement de mémoire longue. Toutes les grandeurs macroéconomiques manifestent cette structure de dépendance forte ou de longue portée y compris les démographies et les ressources naturelles, en flux et en stocks à l'image de la modélisation de la capacité de stockage des flux du Nil opérée par Hurst ou celle des files d'attente inhérentes aux réseaux Ethernet opérée par Taqqu. Néanmoins ces approches statistiques réussies n'ont pas encore eu d'équivalent en théorie, ou en pensée et doctrines économiques. A ce jour, ni les modèles DSGE, ni les modèles macroéconométriques à grande échelle, ni même les modèles computationnels multi-agent, des plus pertinents, n'ont intégré, à ce jour, en leur sein le paradigme de la mémoire longue d'une manière satisfaisante et définitive. Cela aurait pu donner, en effet, un éclairage précieux sur la persistance des crises aussi bien économiques que financières et leur récurrence. Cela aurait pu, aussi, alerter sur l'épuisement en cours de certaines ressources naturelles irremplaçables. Les nouvelles approches économiques : écosystémiques, environnementales, écologiques, de développement durable, d'échauffement climatique, pourraient l'adopter à bon escient comme paradigme fondateur de la pensée économique dans la préservation de l'éco-système avec, en filigrane, l'idée qu'un choc exogène qui perturberait l'ordre naturel des écosystèmes éloignerait ceux ci de leurs sentiers d'équilibre, pour de très longues périodes.

Or l'utilitarisme doctrinal intégral de John Stuart Mill amendé par une dynamique structurelle de long-terme, pourrait être une meilleure infrastructure théorique porteuse de projets écosystémiques tant espérés. Les approches aussi bien des fracatals de Mandelbrot que de l'agrégation et de la cointégration de Granger pourraient à leur tour lui donner sa cohérence doctrinale et combler les carences et les insuffisances des différents modèles économiques et macroéconomiques actuels, en vogue tels que les DSGE, et autres où le paradigme utilitariste, amputé, fragmentaire de John Stuart Mill a été complètement dénaturé et défiguré, sans donner la moindre références à ses autres dimensions, les plus intéressantes et les plus pertinentes qui auraient pu donner à la fois du sens et de la substance à l'utilitarisme réduit et mutilé ambiant.

1.3 Estimation et tests adaptatifs pour des processus LRD

Il s'agit de la synthèse de deux articles : un publié dans JMVA [6], l'autre soumis. Dans cette partie, nous supposons que les processus LRD vérifient les hypothèses $SIG(d, \beta)$.

1.3.1 Présentation de la statistique unidimensionnelle de Ratios d'Incréments : $IR(m)$

Soit $(X_1, \dots, X_N)$ une trajectoire de X . Soit $IR_N(m)$ la variable aléatoire définie, pour tout $m \in \mathbb{N}^*$ par :

$$IR_N(m) := \frac{1}{N-3m} \sum_{k=0}^{N-3m-1} \frac{|\sum_{t=k+1}^{k+m} (X_{t+m} - X_t)| + \sum_{t=k+m+1}^{k+2m} (X_{t+m} - X_t)|}{|\sum_{t=k+1}^{k+m} (X_{t+m} - X_t)| + |\sum_{t=k+m+1}^{k+2m} (X_{t+m} - X_t)|}.$$

Il s'agit d'un balayage de ratio normalisé de valeur absolue de la somme et de la somme de la valeur absolue, de deux sommes partielles d'incrément, de longueur m chacune, sur l'ensemble de la trajectoire du processus gaussien de longueur $N-3m$. En faisant tendre à la fois N , m et N/m vers ∞ , ce ratio de variables aléatoires va tendre vers une fonction continue et dérivable, ψ de mouvement brownien fractionnaire B_H , avec $H = d+0.5$. Ce travail a déjà été réalisé par Surgailis *et al*, (2008), permettant d'estimer le paramètre d de la mémoire longue.

$$\begin{aligned} IR_N(m) &\xrightarrow[N/m, m \rightarrow \infty]{\mathcal{P}} \Lambda_0(d) \\ (N/m)^{\frac{1}{2}} \left(IR_N(m) - \mathbb{E} IR_N(m) \right) &\xrightarrow[N/m, m \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma^2(d)) \end{aligned}$$

où $\sigma^2(d)$ est défini par 3.2. La fonction $d \mapsto \Lambda_0(d)$ définie par l'expression :

$$\Lambda_0(d) := \Lambda(\rho(d)) = \frac{2}{\pi} \arctan \left(\sqrt{\frac{1+\rho(d)}{1-\rho(d)}} \right) + \frac{1}{\pi} \sqrt{\frac{1+\rho(d)}{1-\rho(d)}} \log \left(\frac{2}{1+\rho(d)} \right) \quad \text{pour } |\rho(d)| \leq 1$$

est une fonction continue, croissante, dérivable et inversible pour tout $d \in (-0.5, 15)$ de même que $\sigma^2(d)$ pour $d \in (-0.5, 125)$. $\rho(d)$ étant la fonction de corrélation définie par :

$$\rho(d) := \begin{cases} \frac{4^{d+1.5} - 9^{d+0.5} - 7}{2(4 - 4^{d+0.5})} & \text{pour } d \in (-0.5, 0.5) \cup (0.5, 1.5) \\ \frac{9 \log(3)}{8 \log(2)} - 2 & \text{pour } d = 0.5 \end{cases}$$

Au vu des propriétés de la fonction limite $\Lambda_0(d)$, l'estimateur $\hat{d} = \Lambda_0^{-1}(IR_N)$ est consistant et satisfait le CLT (1.5), défini plus loin pour $j=1$, néanmoins il exige la connaissance du comportement du second ordre de la densité spectrale au voisinage de la fréquence 0. En plus il est empiriquement moins performant que les autres estimateurs tels que le local Whittle de Robinson (1995) [81], le Fractionnel Exponentiel de Moulines et Soulier (2003) [70], les ondelettes de Bardet *et al* (2008), [10].

1.3.2 Estimateur issu de la Statistique Multidimensionnelle de Ratios d'Incréments

Soit $SIG(d, \beta)$: pour $-0.5 < d < 1.5$ and $\beta > 0$ une classe de processus semi-paramétriques et soit $m_j = j m$, $j = 1, \dots, p$ avec $p \in \mathbb{N}^*$ et $m \in \mathbb{N}^*$, et soit le vecteur aléatoire $(IR_N(m_j))_{1 \leq j \leq p}$. Nous étendons les résultats de Surgailis *et al.*, (2008), obtenus pour $m \in \mathbb{N}^*$ à $m \in (0, \infty)$ avec la convention : $(IR_N(j m))_{1 \leq j \leq p} = (IR_N(j [m]))_{1 \leq j \leq p}$ (qui ne change rien aux résultats asymptotiques). Nous montrons le théorème central limite multidimensionnel satisfait par $(IR_N(j m))_{1 \leq j \leq p}$ pour tout $d \in (-0.5, 1.25)$:

Propriété 1. *Supposons que l'**Hypothèse** $SIG(d, \beta)$ est vérifiée avec $-0.5 < d < 1.25$ et $\beta > 0$. Alors*

$$\sqrt{\frac{N}{m}} \left(IR_N(j m) - \mathbb{E}[IR_N(j m)] \right)_{1 \leq j \leq p} \xrightarrow[N/m \wedge m \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_p(d)) \quad (1.5)$$

avec $\Gamma_p(d) = (\sigma_{i,j}(d))_{1 \leq i,j \leq p}$ où $\sigma_{i,j}(d)$ pour $i, j \in \{1, \dots, p\}$, est définie par l'équation 3.5 du chapitre 3.

Nous supposons, dans la suite, que $\Gamma_p(d)$ est une matrice définie positive pour tout $d \in (-0.5, 1.25)$. Dès lors ce CLT peut être utilisé pour estimer d .

Propriété 2. *Soit X satisfaisant l'**Hypothèse** $SIG(d, \beta)$ avec $-0.5 < d < 1.5$ et $0 < \beta \leq 2$. Alors il existe une constante $K(d, \beta)$ non nulle, dépendant uniquement de d et de β tel que pour m suffisamment grand, on a :*

$$\mathbb{E}IR_N(m) = \Lambda_0(d) + \begin{cases} K(d, \beta) m^{-\beta} (1 + o(1)) & \text{si } \beta < 2d + 1 \\ K(d, 2d + 1) m^{-\beta} \log m (1 + o(1)) & \text{si } \beta = 2d + 1 \text{ (si } d = 0.5 \Leftrightarrow \beta = 2) \\ O(m^{-2d-1}) & \text{si } \beta > 2d + 1 \end{cases}$$

Dès lors en choisissant m et N tel que $(\sqrt{N/m}) m^{-\beta} \log m \rightarrow 0$ lorsque $m, N \rightarrow \infty$, le terme $\mathbb{E}[IR(jm)]$ peut être remplacé par $\Lambda_0(d)$ dans la **Propriété 1**. Puis, en utilisant la Delta-method avec la fonction $(x_i)_{1 \leq i \leq p} \mapsto (\Lambda_0^{-1}(x_i))_{1 \leq i \leq p}$ (la fonction $d \in (-0.5, 1.5) \rightarrow \Lambda_0(d)$ est une fonction $\mathcal{C}^\infty$ croissante), nous obtenons :

Théorème 3. *Soit $\hat{d}_N(j m) := \Lambda_0^{-1}(IR_N(j m))$ pour $1 \leq j \leq p$. Supposons l'**Hypothèse** $SIG(d, \beta)$ vérifiée, avec $-0.5 < d < 1.25$ et $0 < \beta \leq 2$. Alors si $m \sim C N^\alpha$ avec $C > 0$ et $(1 + 2\beta)^{-1} \vee (4d + 3)^{-1} < \alpha < 1$,*

$$\sqrt{\frac{N}{m}} \left(\hat{d}_N(j m) - d \right)_{1 \leq j \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left(0, (\Lambda'_0(d))^{-2} \Gamma_p(d) \right). \quad (1.6)$$

Soit

$$\widehat{\Sigma}_N(m) := (\Lambda'_0(\widehat{d}_N(m))^{-2} \Gamma_p(\widehat{d}_N(m))). \quad (1.7)$$

La fonction $d \in (-0.5, 1.5) \mapsto \sigma(d)/\Lambda'(d)$ est $\mathcal{C}^\infty$ et donc, sous les hypothèses du **Théorème 3**,

$$\widehat{\Sigma}_N(m) \xrightarrow[N \rightarrow \infty]{\mathcal{P}} (\Lambda'_0(d))^{-2} \Gamma_p(d).$$

Donc, une estimation pseudo-moindre carrée généralisée PMCG de d peut être définie par

$$\widetilde{d}_N(m) := (J'_p(\widehat{\Sigma}_N(m))^{-1} J_p)^{-1} J'_p(\widehat{\Sigma}_N(m))^{-1} (\widehat{d}_N(m_i))_{1 \leq i \leq p}$$

avec $J_p := (1)_{1 \leq j \leq p}$ et J'_p sa transposée. On déduit à partir du théorème de Gauss-Markov, que la variance asymptotique de $\widetilde{d}_N(m)$ est plus petite que celle de $\widehat{d}_N(jm)$, $j = 1, \dots, p$. Nous obtenons donc, sous les hypothèses du **Théorème 3** :

$$\sqrt{\frac{N}{m}} (\widetilde{d}_N(m) - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, \Lambda'_0(d)^{-2} (J'_p \Gamma_p^{-1}(d) J_p)^{-1}\right). \quad (1.8)$$

et $\Lambda'_0(d)^{-2} (J'_p \Gamma_p^{-1}(d) J_p)^{-1} \leq \Lambda'_0(d)^{-2} \sigma^2(d)$.

Maintenant, considérons le problème de test suivant : pour $(X_1, \dots, X_n)$ une trajectoire d'un processus X gaussien, choisir entre les hypothèses

- H_0 : la densité spectrale de X satisfait l'Hypothèse $S(d, \beta)$ avec $-0.5 < d < 0.5$ et $\beta > 0$;
- H_1 : la densité spectrale de X ne satisfait pas un tel comportement.

Nous déduisons à partir du CLT multidimensionnel (1.6) une statistique de test de type qualité d'ajustement χ^2 définie par :

$$\widehat{T}_N(m) := \frac{N}{m} (\widetilde{d}_N(m) - \widehat{d}_N(jm))'_{1 \leq j \leq p} (\widehat{\Sigma}_N(m))^{-1} (\widetilde{d}_N(m) - \widehat{d}_N(jm))_{1 \leq j \leq p}.$$

Alors le théorème limite suivant peut être déduit à partir du **Théorème 3**.

Proposition 1. *Sous les hypothèses du **Théorème 3** alors*

$$\widehat{T}_N(m) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p-1).$$

1.3.3 La version adaptative de l'estimateur

Pour être appliqués, le **Théorème 3** et le **CLT**(1.8) exigent la connaissance préalable de β . Mais en pratique β est inconnu. La procédure suivante consiste à estimer β à partir des données par l'emploi d'une séquence optimale $(\widetilde{m}_N)$ pour $d \in (-0.5, 1.25)$ et $\alpha \in (0, 1)$. Définissons

$$Q_N(\alpha, d) := (\widehat{d}_N(j N^\alpha) - \widetilde{d}_N(N^\alpha))'_{1 \leq j \leq p} (\widehat{\Sigma}_N(N^\alpha))^{-1} (\widehat{d}_N(j N^\alpha) - \widetilde{d}_N(N^\alpha))_{1 \leq j \leq p}, \quad (1.9)$$

qui correspond à la somme des carrés des distances pseudo-généralisées entre les points $(\widehat{d}_N(j N^\alpha))_j$ et l'estimateur PMCG de d . Notons que les conventions adoptées précédemment permettent d'écrire

à la fois $\widehat{d}_N(j N^\alpha) = \widehat{d}_N(j [N^\alpha])$ et $\widetilde{d}_N(N^\alpha) = \widetilde{d}_N([N^\alpha])$. Donc la fonction $\widehat{Q}_N(\alpha)$ peut être minimisée par une discrétisation judicieuse de l'intervalle $(0, 1)$ comme suit :

$$\widehat{\alpha}_N := \operatorname{Argmin}_{\alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha) \quad \text{avec} \quad \mathcal{A}_N = \left\{ \frac{2}{\log N}, \frac{3}{\log N}, \dots, \frac{\log[N/p]}{\log N} \right\}.$$

D'où la proposition fondamentale suivante :

Proposition 2. *Supposons que l'**Hypothèse SIG**(d, β) est vérifiée avec $-0.5 < d < 1.25$ et $0 < \beta \leq 2$. Alors,*

$$\widehat{\alpha}_N \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \alpha^* = \frac{1}{(1 + 2\beta)} \vee \frac{1}{(4d + 3)}.$$

Finalement définissons :

$$\widetilde{m}_N := N^{\widetilde{\alpha}_N} \quad \text{avec} \quad \widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6 \widehat{\alpha}_N}{(p - 2)(1 - \widehat{\alpha}_N)} \cdot \frac{\log \log N}{\log N}.$$

et l'estimateur

$$\widetilde{d}_N^{(IR)} := \widetilde{d}_N(\widetilde{m}_N) = \widetilde{d}_N(N^{\widetilde{\alpha}_N}). \quad (1.10)$$

On déduit à partir du théorème central limite (**Théorème3, CLT 1.8**) le théorème limite suivant qui donne en plus le comportement asymptotique de l'estimateur $\widetilde{d}_N^{(IR)}$:

Théorème 4. *Sous les hypothèses de la **Proposition 2**,*

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}} (\widetilde{d}_N^{(IR)} - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0; \Lambda'_0(d)^{-2} (J'_p \Gamma_p^{-1}(d) J_p)^{-1}\right). \quad (1.11)$$

$$\text{De plus, } \forall \rho > \frac{2(1 + 3\beta)}{(p - 2)\beta}, \quad \frac{N^{\frac{\beta}{1+2\beta}}}{(\log N)^\rho} \cdot |\widetilde{d}_N^{(IR)} - d| \xrightarrow[N \rightarrow \infty]{\mathcal{P}} 0.$$

Dans ce cadre semi-paramétrique, la vitesse de convergence de $\widetilde{d}_N^{(IR)}$ est la même (à un facteur logarithmique multiplicatif près) que celle de l'estimateur minimax de d (see Giraitis *et al.*, 1997).

1.3.4 Application à des tests statistiques

Le théorème central limite (1.11), satisfait par $\widetilde{d}_N^{(IR)}$, constitue un avantage supplémentaire de $\widetilde{d}_N^{(IR)}$ par rapport à d'autres types d'estimateurs adaptatifs de d (voir par exemple Moulines et Soulier, 2003, pour une vue d'ensemble des estimateurs d dans le domaine fréquentiel). De plus $\widetilde{d}_N^{(IR)}$ peut être utilisé pour $d \in (-0.5, 1.25)$, *i.e.* tout aussi bien pour des processus stationnaires que non-stationnaires, sans modifications dans sa définition. Ces deux avantages conjoints permettent de définir un test de stationnarité basé sur cet estimateur $\widetilde{d}_N^{(IR)}$.

Test d'adéquation

On teste H_0 : , avec

$$\hat{T}_N(m) := \frac{N}{m} (\tilde{d}_N(m) - \hat{d}_N(jm))'_{1 \leq j \leq p} (\hat{\Sigma}_N(m))^{-1} (\tilde{d}_N(m) - \hat{d}_N(jm))_{1 \leq j \leq p}.$$

$$\tilde{T}_N := \hat{T}_N(N^{\tilde{\alpha}_N})$$

alors :

Proposition 3. *Sous les hypothèses de la Proposition 2 alors,*

$$\tilde{T}_N \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p-1).$$

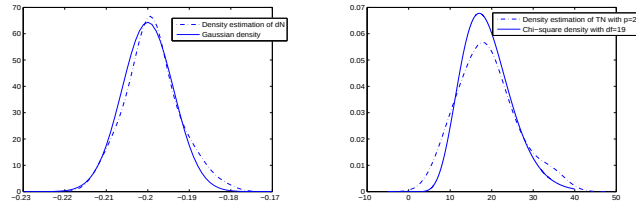


FIGURE 1.1 – *Density estimations and corresponding theoretical densities of $\tilde{d}_N$ and $\tilde{T}_N$ for 100 samples of FARIMA(1, d, 1) with $d = -0.2$ for $N = 10^5$ and $p = 20$.*

Test de stationnarité

Grâce au théorème central limite établi, nos deux tests de stationnarité et de non stationnarité seront construits autour de $d = 0.5$ et pour différentes dimensions p de $IR_N(p)$ auxquelles correspond $\sigma_p(0.5) = \left(\Lambda'_0(0.5)^{-2} (J'_p \Gamma_p^{-1}(0.5) J_p)^{-1} \right)^{1/2}$. Les valeurs $\sigma_p(0.5)$ relatives à $p=5, 10, 15, 20$ sont respectivement $\simeq 0.9082; 0.8289; 0.8016; 0.7861$. On note $q_{1-\alpha}$ le quantile d'ordre $(1-\alpha)$ d'une loi normale standard $\mathcal{N}(0, 1)$.

1. Une première statistique de test de stationnarité est établie avec les deux règles de décision H_0 : “ X processus stationnaire quand $\tilde{S}_N = 1$ ” contre H_1 : “ X processus non stationnaire $\tilde{S}_N = 0$ ” via la la statistique :

$$\tilde{S}_N := 1_{\tilde{d}_N^{(IR)} > 0.5 + \sigma_p(0.5) q_{1-\alpha} N^{(\tilde{\alpha}_N - 1)/2}}$$

On compare alors ce test aux deux tests classiques de stationnarité : KPSS (Kwiatkowski, Phillips, Schmidt, Shin, [49] p. 514) et LMC (voir Leybourne MacCabe *et al.* [64] or [65])

2. Une deuxième statistique de test de non stationnarité est établie avec les deux règles de décision H'_0 : “*X processus non stationnaire quand $\widetilde{NS}_N = 1$* ” contre H'_1 : “*X processus stationnaire quand $\widetilde{NS}_N = 0$* ” via la statistique :

$$\widetilde{NS}_N := 1_{\widehat{d}_N^{(IR)} < 0.5 - \sigma_p(0.5) q_{1-\alpha} N^{(\bar{\alpha}_N - 1)/2}}$$

On compare alors ce test aux deux tests classiques de non stationnarité : Dickey-Fuller “augmenté” (voir [49], p. 516-528) et Phillips et Perron (voir [31], p. 137-146)

La mise en œuvre de ces 6 statistiques respectivement : $\widetilde{S}_N, KPSS, LMC$ d’un côté et $\widetilde{NS}_N, ADF, PP$ de l’autre consiste à mener une expérience de Monte Carlo sur un échantillon de taille N , issu d’un modèle semi-paramétrique donné, en mesurant la fréquence de réalisations des trajectoires qui satisfont l’une ou l’autre des hypothèses : H_0, H_1, H'_0, H'_1 .

1.4 Mise en œuvre empirique des différentes procédures d’estimation et de tests

1.4.1 Simulations

Les simulations réalisées sont menées sur différents modèles semi-paramétriques gaussiens stationnaires et non stationnaires (FARIMA(p,d,q), GARMA) et des méthodes de simulation de type Paxson(d,1), en présence ou non de (Trend, Saisonnalité, ...), avec $d \in I_d = (-0.5, 1.25)$, mais aussi, à titre exploratoire, d’autres types de lois (uniforme, Burr). Les objectifs de ces simulations sont multiples :

1. estimer à la fois la fenêtre m mais aussi le paramètre d de ces modèles avec son $\sqrt{MSE}$ et le comparer avec d’autres estimateurs disponibles.
2. tester l’adéquation de la qualité d’ajustement de cet estimateur adaptatif avec un χ^2_{p-1} théorique d’ordre $1 - \alpha$.
3. tester la stationnarité de la trajectoire du processus en comparaison avec les tests de stationnarité disponibles standards.
4. tester la non stationnarité de la trajectoire du processus en comparaison avec les tests de non stationnarité disponibles standards.

Résultats concernant l’estimation

Les 4 principaux processus générés sont respectivement : ARIMA(1,d,0), FARIMA(0,d,0), FARIMA(1,d,1), Paxson(d,1) stationnaires et non stationnaires. Dans la majorité des cas de figures (dans plus de 95% des cas) aussi bien pour $d \in I_d$ que pour $(p = 5, 10, 15, 20; N = 1000, 5000, 10000)$, les $\sqrt{MSE}$ relatifs aux $\widehat{d}_N^{IR}$ que nous obtenons pour les différents modèles sont, à plus de 50%, plus

particulièrement pour $p=15$, globalement meilleures, que tous les autres $\sqrt{MSE}$ des différents estimateurs, tels que $\hat{d}_N^{MS}$ de Moulines et Soulier, $\hat{d}_N^{ROB}$ du local Whittle de Robinson, ou $\hat{d}_N^{WAV}$ des ondelettes de Bardet et Bibi.

Résultats concernant le test d'adéquation

Le test d'adéquation de la qualité d'ajustement de cet estimateur adaptatif $\hat{d}_N^{IR}$ avec un χ_{p-1}^2 théorique d'ordre 95% donne de très bons résultats, respectivement de plus de (91% pour $p=5$; 88% pour $p=10$; 85% pour $p=15$; 80% pour $p=20$) pour les différents estimateurs $\hat{d}_N^{IR}$ inhérents aux différents modèles testés.

Résultats concernant le test de stationnarité

Les 4 principaux processus testés sont respectivement : ARIMA(1,d,0), FARIMA(0,d,0), FARIMA(1,d,1), Paxson(d,1) stationnaires et non stationnaires. Notre test de stationnarité $\hat{S}_N$ (respectivement de non stationnarité $\widehat{NS}_N$) est comparé aux deux autres tests de stationnarité : KPSS et LMC (respectivement aux deux tests de non stationnarité ADF et PP). On constate que quelque soit le processus, les résultats des tests, dans leur ensemble, sont très satisfaisants pour nos deux statistiques $\hat{S}_N$ et $\widehat{NS}_N$, sur tout l'intervalle $d \in I_d$, et restent globalement satisfaisants pour les 4 autres statistiques concurrentes : ADF, PP, KPSS, LMC, avec un léger avantage pour l'ADF et le PP sur le KPSS ou le LMC.

1.4.2 Applications aux séries financières

Parmi 16 séries financières disponibles de rendements (en logarithmes différenciés) de notre base de données, seules 5 séries présentent un nombre minimal de ruptures en distribution=2 selon l'algorithme de Lavielle (choix 4 : rupture en distribution).

Ces 5 séries financières retenues et analysées ici, sont respectivement les rendements r : du taux de change du dollar américain en Deutsch Mark, des quatre indices boursiers, que sont les deux Dow Jones (transportation et utilities), le Nasdaq et le Nikkei225. On vérifie empiriquement les faits statistiques suivants, et ce pour chaque série de rendement (avec son signe) dans sa totalité (pour les 5 séries) :

- $\hat{d}^{IR}(r) = 0$, confirmé par les 3 autres estimateurs : $\tilde{d}^{MS}$, $\tilde{d}^{ROB}$, $\tilde{d}^{WAV}$, quelque soit la série r avec son signe prise dans sa totalité parmi les 5 séries précédentes.
- Les 6 tests de *stationnarité/nonstationnarité* affirment tous à tort que les 5 séries sont stationnaires, sur l'ensemble de leurs trajectoires respectives.
- Pourtant le test de rupture en distribution de Lavielle identifie 2 ruptures et donc 3 sous-séries (r_1, r_2, r_3) pour chaque série r avec son signe, à des dates différentes et bien localisées.
- $\hat{d}^{IR}(|r_i|^\theta) \gg 0$ avec $\theta \in (0, 3)$, confirmé par les 3 autres estimateurs, quelques soient les sous-séries de la série mère r prise parmi les 5 séries financières précédentes.

- La valeur critique de la statistique V/S de test de mémoire courte contre mémoire longue est de 0.1869, largement dépassée par les $|r_i|^\theta$, avec, en plus, une forte valeur de la kurtosis, synonyme de queues épaisses, dans la distribution de la sous-série r_i correspondante, et ce pour chacune des 5 séries mères .

1.5 Conclusion

Au niveau épistémologique, la problématique des “*deux sciences*”, proposée pour l’économie, ne fait que reproduire, en projection sur l’échelle de la scientificité, des problématiques révélées par Quine dans “*Deux dogmes de l’empirisme*”, relatives à deux savoirs : *a priori* et *a posteriori*³⁵. En effet, les deux dogmes de l’Empirisme logique stipulent³⁶ que :

1. Il existe un critère de distinction fixe entre les propositions synthétiques et analytiques.
2. Toute proposition, ayant une signification quelconque, peut être réduite à une seule et unique proposition, portant sur l’expérience immédiate. Cette expérience immédiate constitue alors sa méthode de vérification.

Ce schéma à double entrées est réfuté par Quine ainsi :

1. Il n’y a pas de *frontière fixe* et définitive entre les *propositions synthétiques* et *analytiques*.
2. Il n’y a pas d’*expérience cruciale* qui mettrait à défaut la *signification* d’une proposition.

En transposant à notre contexte, le quartet (*signification* , *proposition* , *synthétique* , *analytique*), appréhendé dans ce qui précède, en (*scientificité* , *savoir* , *a posteriori* , *a priori*), dès lors les deux dogmes évoqués par Quine (bien qu’ils demeurent toujours extrêmement complexes à approcher) prendraient désormais une forme plus proche de notre quête ainsi³⁷ :

1. Il n’y aurait pas de *frontière fixe et définitive* entre les savoirs *a posteriori* et *a priori*.
2. Il n’y aurait pas d’*expérience cruciale* qui mettrait à défaut la *scientificité* d’un *savoir*.

La **rationalité**, à son tour, attribuée discrétionnairement à l’*homo-sapiens* et à son équivalent économique l’*homo-oeconomicus*, dans une vision inhérente au contrôle optimal adaptatif, trop mécaniste, sacrifie arbitrairement l’autre dimension, conceptuelle et interprétative, de l’homme³⁸.

Force est de constater, en définitive, que le “*paradigme de la mémoire longue*” tant espéré, avec son infrastructure construite pourtant avec succès par Mandelbrot et par Granger, n’a pu avoir en macroéconomie, finalement, tout l’impact qu’il a eu en finance, encore moins en théorie ou en pensée économiques.

35. “Deux dogmes de l’empirisme” :Cinquante ans après, Paolo Parrini, [http ://www.cairn.info/revue-diogene-2006-4-page-108.htm](http://www.cairn.info/revue-diogene-2006-4-page-108.htm)

36. [http ://blog.philotropes.org/post/2009/10/22/Les-dogmes-quinien](http://blog.philotropes.org/post/2009/10/22/Les-dogmes-quinien)s

37. [http ://www.claudegagnon.net/josette/questionjugement.htm](http://www.claudegagnon.net/josette/questionjugement.htm)

38. Les animaux sont-ils des rationnels ? Phd de Philosophie, Benoît Hardy-Vallée, 2006

Nous proposons, en fin de thèse, en conclusion et perspectives, et aux vus des travaux et diagnostics de Stiglitz, une réhabilitation de la doctrine utilitariste intégrale (sans mutilation de ses autres dimensions des plus intéressantes comme elle l'a eue par le passé) de John Stuart Mill, seul grand économiste qui a pensé sa théorie économique en écosystème. La réhabilitation de l'utilitarisme intégral de John Stuart Mill pourrait constituer cette infrastructure doctrinale de la future fondation économique, avec néanmoins deux volets techniques intégrés des plus pertinents : l'analyse fractale de Mandelbrot et l'agrégation et la cointégration de Granger qui intègrent de facto la mémoire longue.

Au niveau statistique l'ensemble de nos résultats est concluant, sans recourir à trop d'hypothèses restrictives, autres que les hypothèses $SIG(d, \beta)$. Les temps de calculs restent convenables et ce pour l'ensembles des séries réelles et simulées. La seule limitation de notre méthode adaptative reste son incapacité à détecter les ruptures au sein d'une série chronologique.

Chapitre 2

Épistémologie de la mémoire longue en économie et son “*paradigme*” associé

Sommaire

2.1	L'épistémologie économique, ce labyrinthe des concepts problématiques	46
2.1.1	Problématiques et controverses autour du concept d'épistémologie	49
2.1.2	Controverses autour de la <i>scientificité</i> de l'économie	54
2.1.4	Paradigme de l'économétrie des séries temporelles en économie	60
2.1.5	Problématique de la cointégration des séries temporelles multivariées . . .	62
2.1.6	Les origines diverses et multiples de la mémoire longue	63
2.1.7	Contexte général de l'apparition du phénomène de mémoire longue	66
2.1.8	Auto-similarité, accroissements indépendants, accroissements stationnaires 69	
2.1.9	Méthodes, groupes et opérateurs de renormalisation	71
2.2	Economie et mémoire longue	73
2.2.1	Les théories économiques revisitées, eu égard au paradigme de la mémoire longue	73
2.2.2	Projet de Stiglitz de refonte intégrale de l'économie	79
2.3	Appréhension d'une réhabilitation LRD de l'utilitarisme de John Stuart Mill	83

Résumé : La scientificité de l'économie, loin d'être unanimement admise, continue à faire débat. En effet, “*calquée sur le modèle des sciences de la nature*”¹, la fondation économique, s'est construite, à partir de la sociologie, des mathématiques,

1. A. Diemer (2011) : économie Générale, Introduction p 5

et d'un paradigme économique fondateur, rattaché à une école de pensée : classique, keynésienne, néoclassique, néokeynésienne, hétérodoxe. Or cette approche "*physicaliste*", pose plus de problèmes qu'elle ne résout. Le questionnement "*le temps dans l'économie versus l'économie dans le temps*"[14], connu des classiques, ignore les problématiques des processus stochastiques. C'est à partir des années 1970, sous l'impulsion de la modélisation probabiliste de Box et Jenkins, que les acceptions déterministes et stochastiques des évolutions dynamiques, s'opposent. En dépit de la révolution économétrique qu'ont pu avoir les séries temporelles, en finance, paradoxalement, rien de comparable n'a eu lieu en analyse économique. L'agrégation des variables microéconomiques du modèle d'équilibre général économique (GE), s'installe avec de nouvelles variables agrégées, en macroéconomie, initialement statique, puis dynamique (dit DGE), par indogénéisation formelle du temps, dans les années 80 (sans une réussite franche, pour les modèles macroéconométriques, bien connus), puis stochastique (dit SDGE ou DSGE), dans les années 2000. Les modèles alternatifs (que se soient les modèles macroéconométriques à grande échelle, ou les modèles computationnels multi-agents) ne sont pas meilleurs. Hormis la finance, le statut des séries temporelles, aussi bien à mémoire courte qu'à mémoire longue, resté jusqu'à aujourd'hui économiquement précaire, ne leur confère, aucunement, une prééminence substantielle en économie. En effet, elles ont toujours été considérées, comme des outils techniques neutres, à portée universelle. De ce fait, elles ne figurent, directement, dans aucun des paradigmes économiques standards, déterministes par construction, dans lesquels, il y a toujours eu "*impossibilité de penser le temps et l'argent*"². Un réexamen à la fois des problématiques, mais aussi, des paradigmes fondateurs de l'économie, s'impose à nous, dans cette première partie. Nous essayerons d'y réhabiliter, à la fois les mécanismes et les processus générateurs de la dynamique économique, eu égard des séries temporelles à mémoire longue et de leur "*paradigme associé*".

Keywords : Scientificité ; Paradigme ; Séries temporelles ; Mémoire longue ; Agrégation ; Intégration fractionnaire ; Intégration stochastique ; Économie classique ; Économie keynésienne ; Courant hétérodoxe ; Économie néoclassique ; Économie néokeynésienne ; DSGE.

Introduction

"Calquée sur le modèle des sciences de la nature"^{3 4 5}, l'économie⁶ avec sa rationalité, ses "lois" et ses théories est, de tout temps, affiliée au physicalisme initié par les physiocrates, repris, par Adam Smith, consolidé par David Ricardo, et confirmé par John Stuart Mill. En effet, pour Adam Smith, le philosophe et père fondateur de l'économie classique, l'économie est "*la sagesse dans la gestion du patrimoine d'une nation, en quête d'un équilibre harmonieux et stable*"⁷, alors que pour David Ricardo, *l'érudit autodidacte polyvalent, et le maître à penser de Marx*, l'économie est la "*science de la distribution*"⁸. John Stuart Mill, *l'utilitariste* voit quant à lui, l'économie comme "*l'action permanente de manière à produire le plus grand bonheur global entre tous les êtres vivants, dans des limites raisonnables*"⁹. Les théories économiques, quelque soient leurs affiliations aux différentes

2. *Les trous noirs de la science économique : Essai sur l'impossibilité de penser le temps et l'argent*. Sapir, J. 2000

3. Le sociologue en train de se faire : G. Tiffon et V.Garcia, Logique sociale, l'Harmattan 2009, p109

4. L'épistémologie des sciences sociales, Denis Collin, Pages perso, 2006

5. A. Diemer(2011) : Economie Générale, Introduction p 5

6. concept polysémique à formulations et attributs multiples dont : analyse,théorie,sciences,politique,sans paler des préfixes : micro,macro et des terminaisons : métrie,logie,etc...

7. Wikipedia, avec adaptation

8. Il écrit en 1820 à Malthus : "l'économie politique est selon vous une enquête sur la nature et les causes de la richesse. J'estime au contraire qu'elle doit être définie comme : une enquête sur la distribution... De jour en jour, je suis plus convaincu que la première étude est vaine et décevante et que la seconde constitue l'objet propre de la science. Wikipedia

9. Wikipedia avec adaptation

écoles de pensées respectives, sont toutes animées, de fait, à l'image des théories scientifiques de la physique, par la recherche de lois, essentiellement des lois d'équilibre (général, partiel, tendanciel, de court ou de long-terme, en concurrence pure et parfaite ou non). Or le physicalisme, ce courant de pensée issu du positivisme logique, prend la physique comme l'étalon de prédilection suprême et le référentiel idéal dans toute construction de fondation scientifique. Il n'en demeure pas moins qu'un tel physicalisme est problématique, en économie, qui est "*biaisé(e) par des présuppositions liées aux valeurs qui orientent l'activité...de l'économiste*"¹⁰. Le contraste, en science économique, entre déterminisme latent véhiculé et standards normatifs dissimulés, s'avère, en l'état, inéluctablement problématique, vis à vis d'une science économique, de nature sociale et de surcroît empirique. La "*théorie des deux sciences*" : à la fois de la nature et de l'homme en société, apporte plus de confusion qu'elle n'en résout.

Nous essayerons, tout le long de cette partie d'épistémologie, de définir le sens et la substance du vocable "*science économique*" et son statut exact, sur l'échelle des valeurs standards de la scientificité d'une fondation à caractère scientifique ([23]). A dire vrai, ce débat n'est jamais clos, et devient récurrent en période de crise, où l'on cherche un bouc émissaire qui porterait la responsabilité de l'échec des politiques économiques mises en œuvre. Les conséquences de la négation de la scientificité de l'économie entraîne la mise en question aussi bien de sa méthodologie, que de ses outils d'investigation, ou de ses anticipations rationnelles ou encore de ses prévisions à court, moyen et long-terme. Pourtant les philosophes des sciences, tels que Popper, Kühn ou Lakatos ont élaboré, chacun, une ligne frontière entre science et non science, dont aucune n'a jamais été aussi précise pour statuer quant à la scientificité ou non, de toute ou partie de l'économie (selon Quine [77], avec adaptation). Une telle négation, en bloc, de la scientificité de l'économie, sans aménagement, serait paradoxale avec les faits scientifiques, avérés dans beaucoup de domaines économiques. Sinon, où peut-on classer la part de la production scientifique des mathématiciens, probabilistes, statisticiens, analystes des données, qui s'investissent dans les différentes branches de l'économie et qui produisent et diffusent leurs résultats, dans des revues scientifiques des plus prestigieuses ? Il serait paradoxal de nier la scientificité de cette part scientifique et technique quelque soit son apport direct et son utilité immédiate et de continuer à affirmer, à tort, que cette partie scientifique est à proprement parler extra-économique, du fait qu'elle n'apporterait pas, directement, de sens et de substance, aux paradigmes¹¹ économiques standards, passés et présents. Ce paradoxe est encore plus flagrant, pour tous ces modèles scientifiques de références, en finance, tels que : ARCH,

10. <http://denis.collin.pagesperso-orange.fr/Epsh.html>

11. Le *paradigme de la mémoire longue* que nous exposons ici, peut inclure d'autres sous-paradigmes latents ou sous-jacents comme celui de l'*invariance d'échelles*, proposé par Taqqu et al ([29], p 528, ligne 3), appelé aussi sous-paradigme de *renormalisation*, appréhendé par les physiciens Laguës et Lesne en ces termes : "*Les méthodes de renormalisation..., incontournables dès qu'on s'intéresse à des propriétés asymptotiques de systèmes...(sont) fondées sur l'invariance d'échelles*" ([60], p 51)". Les dictionnaires de philosophie des sciences le confirment une troisième fois "*Les théories à invariance de jauge (d'échelles) sont (toutes) renormalisables*" ([16], p543-544, mot=invariance de jauge).

GARCH, Volatilité Stochastique, Changement Stochastique de régimes, etc..., développés par des économistes de renommée tels que Mandelbrot, Granger, Engle, et tant d'autres économistes, sans que ces modèles n'intègrent, pour autant, aucun des principaux paradigmes économiques standards.

On peut objecter que les modèles issus de la macroéconométrie ([32]), mais aussi ceux issus de la macroéconomie financière ([22]), sont décevants. On doit cependant reconnaître que l'apport des investigations économétriques, dans toutes les branches de l'économie, depuis les années trente, est concluant, avec des contributions, parfois, de grande qualité scientifique.

Ainsi, il semblerait, en premier abord, que la modélisation, d'obédience physicaliste, sans dynamique fractale, ni loi d'échelle constant, aussi bien macroéconomique que macroéconométrique, pourrait être une première source d'incohérence qui aurait causé les mauvais pouvoirs explicatifs et prédictifs de la fondation économique pour une nation, ce qui aurait contribué, à la mise en cause de sa scientificité.

La présence de la mémoire longue au sein de beaucoup de variables économiques intégrées crée, en cas d'un choc exogène, une déconnexion durable entre les valeurs de ces variables et leurs fondamentaux, à l'instar des actifs financiers. Cette déconnexion hypothèque, par la même, le retour, supposé jusqu'alors inéluctable, à l'équilibre de ces variables avec en filigrane, la mise à défaut de l'hypothèse fondamentale de l'efficience des marchés. Si en plus, ces variables sont cointégrées, elles ont une tendance stochastique commune, instaurant un équilibre de long-terme en dépit de leur variabilité individuelle inhérente à leur non stationnarité respective.

L'extension du concept de mémoire longue aux processus fractionnaires de paramètre $d \geq 0.5$ permet de faire la jonction, par différences successives, entre mémoire longue et racine unitaire. Hormis, les changements de régimes, les cycles et les saisonnalités, hors de notre propos, à ce niveau de l'analyse, la présence, d'une racine unitaire est relativement fréquente, dans les variables économiques, rendant celles ci non stationnaires, avec des propriétés statistiques non standards ([42], chapitre 14).

Les variables macroéconomiques¹² sont toutes obtenues par agrégation pondérée de variables microéconomiques, avec leurs lots latents de mémoire longue effective ou fallacieuse induite de facto pour toute "série agrégée de séries autorégressives d'ordre 1 associée à une loi Bêta $B(p, 1-p)$, $0 < p < 1$ " ([42], p422), fait nouveau que les paradigmes économiques standards n'évoquent guère.

L'ensemble des problématiques soulevées plus haut ne sont que les préalables d'un vaste examen de l'ensemble des tenants et aboutissants d'une analyse épistémologique que nous voulons minutieuse eu égard aux paradigmes fondateurs de l'économie mais aussi un éclairage nécessaire sur

12. La macroéconomie est fondamentalement la microéconomie des variables agrégées (globales) et des agents représentatifs économiquement rationnels (l'homo-oeconomicus vis à vis des "lois de la nature économique" prolonge l'homo-sapiens vis à vis des "lois de la nature physique") dans un même programme d'optimisation d'une fonction objectif sous les mêmes contraintes de ressources disponibles limitées

les concepts peu précis ou mal définis que nous rencontrerons sur notre chemin d'investigation épistémologique, toutes choses égales par ailleurs.

Nous donnerons ainsi les définitions d'un paradigme puis nous spécifierons les principaux fondements d'un paradigme économique et ses traits caractéristiques, plus particulièrement sa pérennité et sa survie en dépit de la résurgence de nouveaux paradigmes alternatifs, à l'inverse des paradigmes des sciences dites exactes.

Nous faisons par la suite l'inventaire des divers paradigmes économiques et leur carence en mémoire longue.

Nous mettrons en exergue, les nuances que porte le vocable, épistémologie, à la fois, avec l'histoire des sciences mais aussi avec la philosophie des sciences, afin de l'utiliser à bon escient, par la suite, dans notre quête inhérente à l' "*épistémologie de la mémoire longue en économie et son paradigme associé*".

Les bases de la fondation économique sont, dans leur ensemble, des reproductions de modèles microéconomiques d'obédience physicaliste, essentiellement déterministe avec emploi de statistique élémentaire dépourvue d'inférences et de tests statistiques. Or la sophistication d'instruments de gestion planifiée, à la fin de la seconde guerre mondiale, a exigé, pour des fins de comptabilité nationale, la résurgence des modes d'agrégation divers, sans aucune référence, pour autant à la loi d'échelle constant. L'introduction relativement tardive de la statistique inférentielle joignant le courant économétrique déjà lancé dans les années trente, n'a pu donner à travers la macroéconométrie des années 1970 un apport significatif majeur. Les modèles d'analyse simplifiée des phénomènes statistiques majeurs en macroéconomie comme la non stationnarité inhérente à la racine unitaire, l'agrégation, la cointégration, la causalité entre facteurs, la persistance, la saisonnalité, les cycles, les relations d'équilibre de long-terme (plus particulièrement entre variables cointégrées), ont tous montré leurs insuffisances. Pourtant bon nombre de ces phénomènes statistiques de non stationnarité sont étroitement liés, d'une manière ou d'une autre, à la mémoire longue et aux processus fractionnaires associés.

Les travaux précurseurs de Granger, dans son effort soutenu d'investigation probabiliste en quête de réformer l'analyse économique, depuis les années 1960, ont pu établir une jonction durable, en macroéconomie et en finance, entre probabilité et statistique. Grâce à sa persévérance, en quête de faire évoluer probabilité et statistique, conjointement entre deux écueils : théorique et empirique indissociables, ses travaux ont pu avoir, plus particulièrement, en séries temporelles, une portée économique et économétrique considérables. Dès lors l'appréhension et la compréhension des phénomènes économiques, précités, jusqu'alors ignorés ou passés sous silence, sont désormais élucidés et mis à profit, par les théoriciens modélisateurs en économétrie, les décideurs politiques en macroéconomie et les gestionnaires de portefeuilles en finance.

2.1 L'épistémologie économique, ce labyrinthe des concepts problématiques

“*Les effets et les conséquences de long-terme*” sont inscrits dans toute politique économique comme inscrites “*la poule dans l’œuf, la fleur dans la graine*”, selon l’expression consacrée de Henry Hazlitt, ce grand économiste avisé. Expression on ne peut plus explicite où la “*mémoire de long-terme*”¹³, largement mise en évidence, aussi bien pour une et même série intégrée relativement à son propre passé lointain, que pour des séries cointégrées, très fréquentes en macroéconomie avec leur équilibre de long-terme sous-jacent. Cette *mémoire longue* ou *mémoire de long-terme*, inhérente aux processus de dépendance de longue portée, objet de notre quête, constitue, de toute évidence, la composante principale d’une interférence entre vagues successives d’effets du passé lointain conjuguée au présent. Ces vagues successives émanent de processus de dépendance de longue portée, jusqu’alors, ignorés, ou passés sous silence. Les processus indépendants ou de dépendance de courte portée, eux, étaient les mieux appréhendés, jusqu’alors, grâce aux outils statistiques satisfaisants disponibles. Les “*effets et conséquences de long-terme*” inscrits dans toute politique économique, émanent dans une forme de reproduction “*génétique*”, en tout ou partie, d’une auto-similarité temporelle inhérente à cette politique économique. Dès lors cette mémoire de long-terme (ou de longue portée) est-elle spécifique au domaine temporel économique, ou se manifeste-t-elle, aussi bien dans d’autres types de topologies, spatiales et autres, mais aussi dans d’autres disciplines ? Peut-on établir une taxinomie de la mémoire longue selon les topologies et les disciplines ? Ces taxinomies fondent-elles une épistémologie de la mémoire longue ? Comment peut-on distinguer la mémoire longue de l’“*effet papillon (chaotique déterministe)*” et son fameux questionnement métaphorique « *le battement d’ailes d’un papillon au Brésil peut-il provoquer une tornade au Texas ?* » questionnement dû au météorologue Edward Lorenz inhérent à l’extrême sensibilité aux conditions initiales, des chaos. Ce questionnement est de plus en plus pertinent sur fond de crise financière mondiale qui secoue le monde depuis le second semestre de 2008, et ce en dépit de son apparente accalmie relative actuelle.

Dès lors peut-on espérer opérer une appréhension d’une épistémologie de la mémoire longue, à travers le processus d’accumulation du savoir relatif à ce même phénomène relativement récent de mémoire longue, aussi bien dans le domaine économique que dans d’autres disciplines et/ou savoirs, arts et/ou sciences avec un effet de miroir. Or “*L’indépendance de longue portée (ou de mémoire longue), tout comme les queues épaisses, a acquis une telle signification, presque religieuse, et généré une telle controverse*” (Resnick (2007), [78]). Sa projection en économie ne fait qu’amplifier la discorde puisqu’elle met en cause, d’une manière frontale non seulement l’efficacité des marchés mais aussi d’autres hypothèses fondamentales comme la dynamique (dichotomique) supposée soit $I(0)$ soit $I(1)$ reliant les variables économiques entre elles ou par rapport à leurs propres passés respectifs. S’agit-il d’un mécanisme structurel, substantiel, intrinsèque ou s’agit-il d’un phénomène purement distributionnel par effet d’agrégation de lois comme la loi bêta et l’effet fallacieux de leur

13. Mériem CHAOUACHI, work paper ERMES 2005

mémoire longue purement distributionnelle ? Et avant toute chose, qu'entend-on par épistémologie de la mémoire longue au sein de l'épistémologie de la science économique ? *“En quoi consiste alors une épistémologie de la science économique, endossant toute politique économique, au sens Hazlittien, avec sa dimension “génétique reproductrice” de long-terme ? Est-ce pour aider à comprendre l'origine et l'élaboration progressive de cette science sociale ? Est-ce pour expliquer ses crises et ses révolutions ? Est-ce pour spécifier la relation entre l'économie positive, l'économie normative et l'art économique, ou encore, entre la science économique et l'économie politique”*¹⁴.

Or, hormis l'agrégation ou l'intégration stochastique par rapport à un mouvement brownien usuel, les mécanismes causaux intrinsèques de production de tels processus à mémoire longue (ou LRD), ne sont nulle part élucidés et ce à ce jour, dans aucune discipline, hormis leurs manifestations et effets générés aussi bien temporels que fréquentiels.

Pour entamer notre quête, sur un tel sujet épistémologique aussi périlleux, sans vaciller, nous devons avoir des repères sur notre chemin constitués d'écueils à la fois théoriques, mais aussi empiriques, à l'intérieur des frontières de la *“fondation économique”* qu'on désire élucider et connaître son statut exact eu égard aux standards des sciences. Ceci permettra de mener ensuite un prolongement, le plus naturel possible, à ce qui a été déjà établi en épistémologie économique *“classique”*, de ce nouveau fait économétrique et statistique majeur : *la mémoire longue*.

La synthèse des dictionnaires de la philosophie des sciences relativement à la fondation scientifique quelqu'elle soit, et ce en dépit des divergences entre les écoles de pensées philosophiques et les doctrines épistémologiques, relèvent quatre critères fondamentaux pour qu'une discipline (en l'occurrence l'économie) puisse revendiquer le statut de fondation scientifique :

1. *Premièrement, l'existence d'un ensemble de lois objectives* au sein de la discipline.
2. *Deuxièmement, une théorie qui adosse la connaissance scientifique* au sein de la discipline.
3. *Troisièmement, une méthodologie scientifique* dans l'établissement de ses lois.
4. *Quatrièmement, un ensemble de mécanismes hypothético-déductifs* de validation des hypothèses et de délibération eu égard aux attributions initiales de l'enquête, même s'il n'y a pas toujours eu de consensus entre les philosophes des sciences à propos de tous ces critères réunis. Le dernier critère, pour ne citer que celui là, constitue un point de discordance entre l'école poppérienne et l'école kühnienne.

Ainsi on peut lire : *“Pour... (Kühn) la démarche de la science n'est pas hypothético-déductive”*¹⁵ : ce qui met en cause, la valeur des tests économétriques qui s'y réfèrent. En conséquence, notre quête de ces critères relativement à l'économie constitue la ligne directrice de notre approche et le plan de route que nous suivrons, dans notre entreprise qui doit être *“méthodologique scientifique”*, une *“méthodologie scientifique”* entendue dans le sens donné par les auteurs Doreian et Hummon, dans

14. La science économique : une réflexion épistémologique et méthodologique, Hinti Saïd, Imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p 6

15. La philosophie des sciences, Lecourt D. Ed Puf 2002, p 84

leur ouvrage “Modelling Social Processes”¹⁶. En effet ceux-ci proposent une définition synthétique, valable pour toute méthodologie inhérente aux processus sociaux (ou méthodologie des sciences sociales) :

Définition 1 (Méthodologie inhérente aux processus sociaux). *Une méthodologie (scientifique) est un système de règles qui guident l'enquête scientifique. Un manquement à l'une des composantes de la méthodologie dépend de la nature de l'identification de la substance (ou nature) de l'enquête, de son utilisation, de son objet et des préférences de l'investigateur... Nous identifions quatre composantes d'une méthodologie inhérentes à la modélisation des processus sociaux :*

1. *Une (première) composante d'identification de la substance (ou nature) du processus social investi.*
2. *Une (deuxième) composante épistémologique.*
3. *Une (troisième) composante technique.*
4. *Une (quatrième) composante de délibération.*

Les auteurs explicitent ces quatre composantes une à une et élucident le détail de leurs contenus de la manière suivante :

*“La première composante de notre méthodologie (en l'occurrence : l'identification de la substance) n'est pas incluse explicitement, normalement dans les discussions conventionnelles de la méthodologie, néanmoins la méthodologie doit exiger (au préalable) une identification du contexte. Cette composante offre (donc) au minimum un vocabulaire et un ensemble de définitions organisés d'une façon selon qu'une théorie est construite ou non. Une telle organisation constitue la mission principale de cette (première) composante (d'identification)”*¹⁷.

Quant à la composante épistémologique du processus social, les auteurs précisent ses attributs :

“La méthodologie exige en second (lieu) une composante épistémologique. Cette composante fournit les critères de vérités et d'évaluations des connaissances.” Ils reconnaissent ici *“l'impossibilité de fournir, (en cela) , des critères ultimes, dans le sens où chaque fois que ces critères sont utilisés ils généreront l'ultime vérité”*¹⁸, à cause des imbrications philosophiques qui en découlent. *“Néanmoins, il est nécessaire que les hypothèses épistémologiques soient établies et utilisées au sein d'une méthodologie viable”*¹⁹.

Pour la composante technique de la méthodologie, les auteurs affirment que : *“La troisième composante de notre méthodologie concerne les techniques, instruments et procédures qui sont appliqués au problème (inhérent à l'enquête). Traditionnellement, de telles techniques, telles les inférences statistiques, l'analyse des données, les techniques de génération de données... ont formé les composantes des (quêtes et) procédures de la méthodologie des sciences sociales”*²⁰.

16. Modelling Social Processes, Doreian, Hummon Collection : Progress in Mathematical Social Science, 1996, p 1

17. idem

18. idem

19. idem

20. idem

Pour la composante de délibération ils la définissent comme le but ultime de l'enquête qui doit fournir au décideur :

l'“*identification des modes d'implémentation et d'intervention dans la conduite et le comportement des processus sociaux... Une méthodologie qui contient une telle composante tente de découvrir non seulement ce qui se passe mais aussi comment (le cours de) quelque chose peut être modifié.*”.

Mais qu'entend-t-on, au juste, par épistémologie dans la définition précédente ?

2.1.1 Problématiques et controverses autour du concept d'épistémologie

Le *paradigme*, en tant que *matrice disciplinaire* chez Kühn, un *programme scientifique de recherche* chez Lakatos, un nouvel *épistémé* ou *corps de principes* chez Michel Foucault, est un référentiel de logiques et de normes alternatives propres en rupture avec l'état passé de la science et de sa méthodologie scientifique, dans leurs évolutions, dans une époque donnée, “*une sorte de distance en terme de sens, entre la réalité et sa représentation*”²¹. Ainsi tout “*nouveau paradigme*” est à la fois “*une nouvelle conception de la réalité et une nouvelle logique par laquelle cette réalité est comprise*”

Mais comment définir la ligne frontière (ou “*critère de démarcation*”²²), entre science et non science ? S'agit-il d'un même profil de frontière, au sein des sciences de la nature, des sciences sociales, des sciences humaines ? Si on admet avec le sociologue américain Donald Black (2000,[13]) que “*la science est une question de degré : la scienticité ... en tant que fonction curviligne de la distance sociale du sujet...attribuée aussi bien à une idée, à une méthodologie, à une théorie ou à une fondation et que la scienticité d'une idée croît avec sa testabilité, sa généralité, sa simplicité, sa validité et son originalité*”²³ dans un espace social déterminé, muni d'une structure topologique adéquate, dès lors comment peut-on définir la “*scienticité*”^{24 25} d'une fondation “*prétendument scientifique*” ? par son

21. *Entre des modèles qui se contentent d'analyser statistiquement et ceux qui produisent des modèles qui font plus “sens” dans un établissement, la distance peut se mesurer en termes de paradigmes, in : Apprivoiser l'épistémologie, Gérard Fourrez, Ed De Boek & Larcier 2003, p78*

22. *En épistémologie contemporaine, on entend habituellement par critère de démarcation le signe, le caractère ou la propriété permettant de distinguer la science (empirique) de la non-science (ou de la pseudo-science)...les épistémologues ont suggéré divers critères de démarcation, sans qu'aucun cependant ne parvienne à arracher l'assentiment général.*“ Vocabulaire technique et analytique de l'épistémologie, Robert Nadeau, Ed PUF 1999, Vocabulaire=critère de démarcation, p 125.

23. Black D., *Dreams of Pure Sociology*, in : *Sociological Theory*, N° 18 (November) 2000 , pp 343-367

24. Nous parlerons de scienticité au lieu de scientificité, pour désigner non pas le caractère scientifique d'une discipline, mais le fondement ontologique qui fait d'une idée, d'un savoir, d'une connaissance, d'une science , avec des degrés divers, une science en soi, dont on cherchera à élucider, par la suite, le sens le plus exhaustif, relativement à notre domaine d'investigation : l'économie

25. Black D., *Dreams of Pure Sociology*, in : *Sociological Theory*, N° 18 (November) 2000, pp 343-367. Dans l'une de ses correspondances, le Professeur et Sociologue Américain Donald Black, de l'Université de Virginie, explique les raisons de son choix et de ses préférences, pour le mot scienticité, au lieu et à la place du mot scientificité, ainsi : “*Le mot scientificité,...est un mot d'usage rare .(De plus il) est à la fois très maladroit et laid... (ce n'est) même pas sûr que le mot scientificité soit prévu pour capter la signification...recherchée (en terme de degré de science)....d'ailleurs,...c'est un mot pratiquement inconnu en anglais...(Depuis la première*

objet, par sa méthode, par son affiliation, par sa théorie de causalité, entre cause et effet, ou comme le suggère Duhem, par l'existence d'“*un système de propositions mathématiques, déduites (à partir) d'un petit nombre de principes (généraux de la discipline concernée),... (propositions représentant) aussi simplement, aussi complètement, aussi exactement,... (les) lois expérimentales (de cette discipline)*”²⁶ ? Comment peut-on définir par la même la scientificité d'une théorie ? Que peut-on espérer d'une théorie scientifique ? “*Nous fait-elle connaître les régularités de la nature (physique, sociale, économique) ou cherche-t-elle également leurs causes ? Les lois scientifiques sont-elles immuables ou évoluent-elles au cours du temps ? Quel rôle les mathématiques jouent-elles dans la connaissance de la nature (physique, sociale, économique) ? Comment les connaissances se rapportent-elles à ce qui nous est donné, ou à ce que nous éprouvons dans une expérience vécue ? Comment concevoir la possibilité d'une vérification des connaissances scientifiques ? Qu'est ce que l'interprétation d'une théorie ? Les entités théoriques sont-elles réelles ou ne sont-elles que des fictions commodes ? Comment les théories peuvent-elles être comparées les unes aux autres ? La philosophie des sciences (naturelles, sociales ou humaines) produit-elle elle-même des connaissances ou a-t-elle pour but d'analyser le sens des énoncés de la science ?*”²⁷.

Ces mêmes préoccupations et approches, relatives aux problématiques des théories scientifiques, se retrouvent reconduites en philosophie des sciences. En effet “*La philosophie des sciences (parfaissant, avec l'histoire des sciences, toute l'ambivalence de l'épistémologie) pose plusieurs problèmes :*

1. *Le problème de la démarcation : À quoi reconnaît-on qu'une connaissance est scientifique ?...*
2. *L'unité de la science : Doit-on parler de la science ou des sciences ?...*
3. *Le problème de la classification : S'il y a plusieurs sciences (reconnaissables par leur unité méthodologique, s'il y en a une), existe-il une classification de ces sciences, par leur objet par exemple ?...*

“²⁸

Ces mêmes préoccupations se retrouvent en épistémologie des sciences économiques, bien que le concept ambivalent d'épistémologie, mérite d'être élucidé davantage. Dès lors “*En quoi consiste alors une épistémologie de la science économique ? Est-ce pour aider à comprendre l'origine et l'élaboration progressive de cette science sociale ? Est-ce pour expliquer ses crises et ses révolutions ? Est-ce pour spécifier la relation entre l'économie positive, l'économie normative et l'art économique, ou encore, entre la science économique et l'économie politique*”²⁹. La dichotomie entre “*vérification-*

apparition de ce concept de scientificité en 2000), les gens semblent aimer...(ce) mot scientificité. Ils commencent à l'employer dans leurs écrits...”

26. Dictionnaire d'Histoire et Philosophie des Sciences, Sous la direction de Lecourt D., Ed PUF 1999, concept=théorie, p 941

27. Philosophie des sciences, Naturalismes et réalismes, Tome 2, Textes réunis par S. Laugier et P. Wagner, Ed Vrin, 2004, Préface p :7-8

28. La philosophie de A à Z, de Clément et al, ed Hatier 1997, p 322-323 avec une légère adaptation stylistique du troisième alinéa.

29. La science économique : une réflexion épistémologique et méthodologique, Hinti Saïd, Imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p 6

nisme(logique)/réfutationnisme (Poppérien) , hyper-empirisme/anti-empirisme, explicatif/normatif, intentionnisme/déterminisme, explication totale/explication locale, purisme/anarchisme méthodologique“³⁰, est-elle réelle ou constitue-t-elle une pure gymnastique intellectuelle et un simple jeu de l'esprit ?

Ces questions légitimes, au demeurant, restent, de tout temps, chargées d'acuité et de soubresauts. Néanmoins et en dépit des réserves, les philosophes des sciences, estiment que les frontières, entre sciences et non sciences d'une époque, ne sont ni clairement définies, ni définitives, ne serait ce que par la nature complexe et fluctuante du rapport entre le langage et les faits réels. En effet *“il n'y aurait guère de sens, de toute façon, à vouloir établir une stricte séparation entre ce qui relève du langage et ce qui relève des faits ou de l'expérience, la science elle-même, jusqu'à nouvel ordre, se présentant comme un ensemble de propositions sur le réel”*³¹. Comme on le voit, deux problématiques distinctes s'opposent : *“celle de la philosophie du langage (et de la logique) et celle de la philosophie des sciences, sans qu'on puisse décider ce qui est premier”*³². Elles sont en mutation permanente, l'une et l'autre et au sein desquelles et avec elles les théories changent, ainsi que les variables qui les traduisent où *“tout changement de théorie devrait conduire inéluctablement à un changement de signification de ces (ses) variables”*³³.

Aussi assiste t-on à des remises en question récurrentes de ces frontières, d'une époque à l'autre, débouchant sur des débats animés, sur la nature scientifique de telle ou telle théorie, de telle ou telle méthodologie.

Dans sa quête épistémologique, sur la possibilité de la science, Jean Piaget préconise une classification préalable, très éclairante, sur ce qu'est l'épistémologie. Parmi tout l'éventail des théories de la connaissance qui se rapportent à la science, il affirme *“...nous classerons les théories de la connaissance en trois grandes catégories :*

1. *Celles qui partent d'une réflexion sur les sciences et tendent à la prolonger en une théorie générale de la connaissance.*
2. *Celles qui, s'appuyant sur une critique des sciences, cherchent à atteindre un mode de connaissance distinct de la connaissance scientifique (en opposition avec celle-ci et non plus en son prolongement).*
3. *Celles qui demeurent à l'intérieur d'une réflexion sur les sciences.*

*...Nous appellerons "métascientifiques" les théories de la connaissance du type (1), "parascientifiques" celles du type (2) et "scientifiques" tout court celle du type (3)"*³⁴. Il n'en demeure pas moins,

30. épistémologies et sciences de Gestion, Martinet, A.C et al, Ed Economica 1990, pp 18 et suite

31. Philosophie des sciences : Naturalisme et réalismes , Textes réunis par Laugier S. et Wagner P., Ed Librairie Philosophique, J. VRIN, 2004, p 23

32. Philosophie des sciences : Théories , expériences et méthodes, Textes réunis par Laugier S. et Wagner P., Ed Librairie Philosophique, J. VRIN, 2004, p 13

33. Philosophie des sciences : Naturalisme et réalismes, Textes réunis par Laugier S. et Wagner P., Ed Librairie Philosophique, J. VRIN, 2004, p 65

34. Logique et connaissance scientifique, Volume publié sous la direction de Jean Piaget, Ed Gallimard 1967, p 15-16.

que cette conception de l'épistémologie, chez Piaget, a été taxée d'épistémologie "génétique" voire de "psychologisme" par certains épistémologues modernes comme Henri Bareau³⁵ en opposition à l'épistémologie historique qu'ils revendiquent. Ces deux visions contradictoires de l'épistémologie, ne sont que le reflet d'une vieille querelle conceptuelle, autour du vocable ambivalent d'épistémologie, non seulement entre épistémologues d'une même tradition linguistique, mais également entre ceux de langue française et ceux de langue anglaise. Il s'agit d'une confusion entre épistémologie et philosophie des sciences, comme en témoigne le passage suivant "*La philosophie des sciences - ou plutôt, comme Reichenbach dans "les trois tâches de l'épistémologie" de 1938 l'appelle épistémologie - sert trois buts : (1) décrire les relations logiques entre les énoncés de la science, (2) critiquer ces relations, c'est-à-dire en évaluer la validité, et (3) conseiller les scientifiques sur les décisions à prendre*"³⁶. Les dictionnaires de philosophie ne font que reproduire ces confusions et les perpétuer, entre d'un côté l'histoire des sciences, la philosophie des sciences et l'épistémologie de l'autre. En effet, dans la tradition française, ce vocable *épistémologie* désigne, selon une définition synthétique, obtenue à partir des trois dictionnaires de philosophie^{37 38 39}, tantôt la discipline qui prend la science pour objet, sans être à proprement parler, une "*philosophie des sciences*", tantôt la philosophie des sciences (c'est-à-dire l'analyse critique des sciences considérées comme données, dans leurs développements et résultats), voire même l'analyse de l'esprit scientifique, l'étude : des méthodes, des crises, de l'histoire des sciences modernes, la philosophie d'une science particulière. Dans la tradition anglo-saxonne, il désigne plutôt la philosophie (voire la théorie) de la connaissance ou *Gnoséologie*. Dans le "*Vocabulaire Européen des philosophies*"⁴⁰, bien que le vocable d'épistémologie soit plus explicite, cette ambiguïté persiste. En effet, on peut lire, à propos du concept "*épistémologie*" : "*Le terme français d'épistémologie, de même que l'allemand Wissenschaftstheorie, absorbe simplement dans une harmonie quelque peu de façade une multiplicité d'approches : - théorie générale de la connaissance, analyse technique et logique des théories scientifiques, analyse historique de leur développement - que l'anglais tend pour sa part à distinguer (epistemology, philosophy of science, history of science). Mais en réalité il n'existe plus ni doctrine fondatrice ni orientation unitaire dans le domaine de la théorie de la connaissance et de la science. L'expérience de la traduction est devenue, corrélativement, celle de la prolifération de termes "intraduisibles"*".

Il n'en demeure point que le Dictionnaire de Sociologie : Le Robert/Seuil⁴¹, en donne une définition synthétique, en apparence, plus précise et plus exhaustive, sur son origine, d'abord, "*L'épistémologie*

35. auteur chez PUF de l'épistémologie (Que sais-je, p13)

36. Philosophie des sciences : théories, expériences et méthodes, texte réunis par Laugier S. et Wagner P., Ed Librairie Philosophique J. VRIN 2004, Présentation de Hans Reichenbach par Alexis Bienvenu, p 295

37. La philosophie de A à Z d'Élisabeth Clément et al, Hatier(1994), mot=épistémologie, p 110

38. Vocabulaire des études philosophiques d'Auroux et Weil, Hachette (1995), mot=épistémologie p 69

39. Dictionnaire de philosophie de Durozoi et Roussel, Nathan (1997), mot=épistémologie, p 132

40. Vocabulaire Européen des philosophies, sous la direction de Barbara Cassin, Ed Seuil Le Robert 2004, mot=épistémologie, pp 358-364

41. Dictionnaire de Sociologie Le Robert/Seuil, Ed 1999, Sous la direction conjointe d'André Akoun et Pierre Ansart, mot = épistémologie, p 191-192

s'est détachée de la philosophie pour constituer au 20^e siècle une discipline partiellement autonome.⁴² Ensuite, sur son champ d'investigation en général “Elle est associée à la philosophie (théorie de la connaissance) et à l'histoire des sciences ainsi qu'à la logique”⁴³. Dans ce dictionnaire l'auteur distinguera au moins sept spécifications différentes de l'épistémologie dans ses diverses manifestations. Il s'agit de :

1. *L'épistémologie générale, s'attachant aux traits communs des diverses sciences (valeur, origine logique et portée).*
2. *Les épistémologies régionales, attentives à la spécificité des disciplines (et particulièrement de leur méthode) ou d'ensembles disciplinaires.*
3. *Les épistémologies internes, produites par autoréflexion au sein des champs scientifiques.*
4. *L'épistémologie externe, constituée en (une) discipline autonome.*
5. *L'épistémologie normative, visant à tracer une ligne de partage claire entre science et non science.*
6. *L'épistémologie analytique (ou descriptive), cherchant à saisir les procédures de construction de la connaissance scientifique.*
7. *Dans un sens large mais critiqué, le terme (épistémologie) désigne l'étude des processus psychiques de connaissance.*⁴³

L'épistémologie contemporaine, attelée au départ, au programme des positivistes logiques du cercle de Vienne, s'est vue opter plutôt pour une *épistémologie externe* (spécification 4), *générale* (spécification 1) et *normative* (spécification 5). Mais elle s'est vue doublée, par “l'histoire et la philosophie des sciences, attentives au développement concret des disciplines”⁴⁴. Le stimulus de ces deux disciplines (d'histoire et de philosophie des sciences) a opéré un renversement de tendances et de perspectives. Le renfort logistique d'une approche sociologique radicale (ou programme fort), a permis à cette tendance sociologique de manifester une volonté affichée de “faire descendre la science de son piédestal”⁴⁵ avec le risque non négligeable des “dérapages idéologiques”⁴⁶ et d’“aporie relativiste”⁴⁷. En effet, “Les sciences sociales et plus particulièrement la sociologie entretiennent des rapports difficiles avec l'épistémologie”⁴⁸. Même si à la base, la fondation des sciences sociales était construite, à partir d'un modèle épistémologique général, profondément marqué par la thématique de l'épistémologie des sciences de la nature (physique), ces sciences sociales ont pris “très vite leur écart par rapport à ce modèle”⁴⁹ et plus particulièrement la sociologie, elle-même. Celle-ci ne peut s'identifier à un tel statut, marginalisé par la mathématisation, à outrance, de ce modèle, issu des sciences de la nature,

42. idem

43. idem

44. idem

45. idem

46. idem

47. idem

48. idem

49. idem

modèle qu'elle estime dégradant, pour son statut originel, en tant que discipline appartenant, à part entière, aux *sciences de l'esprit*, plutôt qu'aux *sciences de la nature*. Néanmoins, "*D'une façon plus contemporaine, le pluralisme épistémologique, manifesté par la coexistence au sein de la sociologie de divers programmes et paradigmes favorise*"⁵⁰ à la fois, le relativisme, dans ses approches et le rationalisme, dans ses méthodes. Ainsi s'est déployée une sorte de retrouvaille avec certaines des spécifications épistémologiques, de cette sociologie, à la fois plus proches des ses aspirations, et plus ouvertes à sa méthodologie. Une telle retrouvaille ne provoque pas toujours, le consentement, ni l'enthousiasme des épistémologues les plus intransigeants, qui se sont rangés, en bon nombre, du côté des scientifiques et des positivistes.

On tâchera d'éviter, donc, autant que possible, l'usage sans l'avoir endigué, au préalable, de ce vocable ambivalent d'épistémologie, évitant par la même, toute source de confusion.

2.1.2 Controverses autour de la *scientificité* de l'économie

Les sciences économiques, comme le montre leur histoire, de par leur affiliation aux sciences sociales, et surtout du fait qu'elles sont des "*science(s) de l'homme*"⁵¹ à la fois en société mais aussi de "*l'homme...placé dans un flux historique*"⁵², de par leur infrastructure philosophique aussi, n'ont pu échapper aux retombées des problématiques des autres activités scientifiques, et aux questionnements inéluctables que toute entité vivante est amenée, un moment ou un autre, à se poser. Par la même "*un économiste avisé ne peut...se laisser obstruer dans sa quête du savoir par le dogmatisme ou par les canons analytiques standards, mais il doit mettre à profit sa propre réflexion sur les phénomènes qu'il observe, construire sa propre connaissance par un jeu d'essais et d'erreurs, tout en respectant la logique du raisonnement,..., développer sa propre réflexion sur la matière scientifique à laquelle on s'intéresse, ne serait-ce que par curiosité intellectuelle, (tout cela) constitue une voie féconde pour faire progresser le savoir (et la science)*"⁵³.

En dépit de la multiplicité des courants de pensée philosophique sur la nature de la science, on peut néanmoins retenir trois courants principaux, de pensées philosophiques, qui ont pu déboucher sur un questionnement de la pensée économique. D'un côté les inconditionnels de la scientificité de la science économique (tels les positivistes), les autres, soit en opposition à cette scientificité (tels les sceptiques), soit en quête de preuve de cette scientificité (tels les relativistes dits aussi réalistes), avec de nouvelles problématiques et de nouvelles perspectives d'investigations méthodologiques. En effet face à "*la gravité de la crise que traverse la science économique*"⁵⁴ disent les réalistes, et plus particulièrement la "*crise méthodologique*"⁵⁵, il ne peut y avoir, ni de fondation scientifique, digne

50. idem

51. Histoire de la théorie économique, Jessua C., Ed PUF 1991, p 18

52. idem, p 19

53. La science économique : une réflexion épistémologique et méthodologique, Hinti Saïd, Imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p3

54. Les trous noirs de l'économie, Sapir J., Ed Albin Michel 2000 et du Seuil 2003, p 13

55. idem, p 18

de ce nom “*hors d'une méthodologie robuste, (et) il n'est pas d'attitude scientifique possible*”⁵⁶, ni de solutions, en dehors de cette perspective méthodologique robuste, du moins dans l'esprit du courant “*réaliste*”, eu égard à toute scientificité de l'économie, dans son état actuel.

Parce que les sciences économiques prennent le comportement économique de l'homme en société , “*en tant qu'objet de connaissance...(même si) les situations évoluent et ce ne sont pas nécessairement les mêmes réalités humaines que l'on doit se proposer d'expliquer*”⁵⁷. A moins d'un usage justifié de l'expression : “*toutes choses égales par ailleurs*”, les sciences économiques ne peuvent, ni ignorer, ni faire obstruction, à la nature humaine de l'homme, ni aux flux de l'histoire qui le conditionnent. L'homme se trouve, ainsi, au centre de trois pesanteurs ou champs de forces, simultanées, spontanées : les pesanteurs de la nature (physique), avec leurs lois (ou régularités) naturelles (physiques), les pesanteurs de la société, avec leurs lois (ou régularités) sociales, les flux de l'histoire, avec leurs lois (ou régularités) historiques, conditionnant tout ou partie, tout comportement de l'homme en société et plus particulièrement son comportement économique.

La stimulation de lois économiques administrées, autres que les lois spontanées (à l'instar des préférences révélées), par le biais d'un dirigisme étatique, ou par une régulation publique, font partie, plutôt de l'économisme doctrinal et normatif de politiques économiques, inhérent aux différentes écoles de pensée économiques, à la frontière des sciences économiques, du côté des sciences politiques et des politiques organisationnelles et sociales.

Il s'agit, en fait, de la projection, au sein de l'économie, de trois courants philosophiques :

- en premier lieu, le courant qu'on peut baptiser d'économisme-positiviste. Ce courant combine la vision scientiste depuis les physiocrates, et ce jusqu'aux néo-classiques, avec en filigrane, la vision positiviste appliquée à une science sociale particulière, de surcroît, la science économique. Dès lors la science économique apparaît pour ce courant positiviste comme le versant naturel de la sociologie, fondée par Auguste Comte, avec son scientisme exacerbé allant jusqu'à “*prophétiser*” sa sociologie comme la nouvelle “*religion de l'humanité*”. Or la vision des économistes bien que scientiste, elle reste néanmoins, fort apaisée, aussi bien des physiocrates, des classiques et même des néo-classiques, tous convaincus en l'existence de lois objectives (qu'Auguste Comte appelle faits généraux) de la nature économique. “*Nous trouvons en effet couramment exprimée sous la plume des Physiocrates, de Smith, de Ricardo, l'affirmation que la tâche de l'économiste est de révéler des lois qui s'imposeraient aux hommes aussi impérativement que les lois de la nature ou de la physique*”⁵⁸. C'est ainsi que Quesnay (le physiocrate) se réfère aux lois de la physiologie, tandis que c'est l'astronomie qui sert de référence à Adam Smith (le classique), comme le montre le passage suivant : “*Adam Smith, souvent crédité comme étant le fondateur de la discipline (de l'économie politique), était et s'était considéré lui même (avant tout) comme philosophe...(il) a clairement planifié de faire des sciences sociales ce que Newton a fait pour la science naturelle*

56. idem, p 13

57. idem, p 19

58. Histoire de la théorie économique, Claude Jessua, Ed PUF 1991, p 235

(la physique) “⁵⁹.

- En second lieu, une projection au sein de l'économique du scepticisme de Hume, à l'égard des sciences empiriques, adopté par ses disciples, hostiles à la scientificité de toute science empirique, notamment l'économie.
- En troisième lieu, une projection au sein de l'économique du relativisme scientifique (ou historicisme), principalement celui de Kühn et de ses disciples, eu égard à toute science, donc à toute science sociale, même si Kühn a toujours refusé ce relativisme et s'en est toujours démarqué.

Ce sont donc trois visions contradictoires, de par leurs approches mutuellement exclusives qui ont largement contribué, en dépit de leurs lignes de fracture, à la richesse, à la fois de la pensée économique et de la gnoséologie⁶⁰, ainsi qu'à la philosophie des sciences et à leur histoire.

Il n'en demeure pas moins que les trois courants puisent, en empruntant à la philosophie grecque, tout ou partie, de ses certitudes, de ses doutes, de ses inquiétudes, et de ses questionnements, avec leur mise préalable, au goût du jour. Ces emprunts sont multiples et variés, répartis entre sophistes, platoniciens ou épicuriens, avec des adages et des filiations hybrides, des uns et des autres, et plus particulièrement de Platon et d'Aristote car “*Platon et Aristote sont les inventeurs de la philosophie politique occidentale, et également par les descriptions jointes à leurs analyses, de la sociologie politique*”⁶¹. Cette filiation à la culture grecque, en quête d'accéder à la sensibilité et à “*l'oreille*” grecques, est jugée indispensable, dans tout projet de résurrection de l'Occident où toute construction de “*science(s) de l'homme*”⁶² se doit d'ancrer “*l'homme...dans un flux historique*”⁶³ qui est le sien. Ainsi “*l'âge classique européen ne pouvait que tendre à les imiter (les Grecs) en ce domaine (philosophie de l'histoire, plus particulièrement de l'homme en société) comme en d'autres, ainsi que l'avaient fait les historiens latins...Cette tendance philosophique de l'histoire (à la manière des historiens grecs) va s'accroître au 19^e siècle...(associée à) une autre tendance...(chez des écrivains) qui se prend d'intérêt pour le passé (grec) comme tel et vise sa résurrection intégrale.*”⁶⁴

Vérification de la scientificité des principes de la fondation économique

Les économistes ont toujours eu conscience, depuis les physiocrates et jusqu'à nos jours, de la problématique de la scientificité, de ses conséquences et de ses répercussions sur la pertinence et la véracité de leur discipline, considérée tantôt comme une pensée économique, tantôt comme une théorie économique, tantôt encore comme une analyse économique.

Les deux sources de conflit semblent être concentrées autour de deux problématiques : d'un côté la problématique de l'existence ou non de lois en économie et la problématique de la modélisation de ces

59. Economics and the philosophy of Science, Deborah A. Redman, Oxford University Press, 1991, p 91-92

60. gnoseologie=épistémologie, dans le sens anglo-saxon de philosophie ou de théorie de la connaissance

61. épistémologie, Barreau H., Ed PUF 2002 (Que sais-je), p 89

62. Histoire de la théorie économique, Jessua C., Ed PUF 1991, p 18

63. idem, p 19

64. idem p 90

lois (ou faits économiques généraux) au sein de l'économie, traversée, en tout temps, “d'incertitude profonde sur la distance qui sépare ce qu'elle est censée représenter et la “réalité“ dont elle souhaite rendre compte...La science économique ne peut au mieux en donner qu'une représentation partielle. Ses constructions théoriques laissent forcément de côté nombreuses dimensions de la réalité, mais en toute connaissance de cause, du moins pour la bonne théorie “⁶⁵. On retrouve ici la projection au sein de l'économie des vieilles querelles entre philosophes depuis l'antiquité grecque autour de l'existence ou non de vérités ou de lois en dehors de l'empirisme et de l'expérience d'un côté et le degré de restitution plus ou moins intégrale de la réalité par sa modélisation reproductrice.

Dès lors, un diagnostic minutieux, au préalable, des mécanismes d'investigation et d'approche dans chacune des étapes de modélisation au sein de la fondation économique s'impose inéluctablement à nous. Ce diagnostic portera sur la véracité à la fois des théories mais aussi des infrastructures inhérentes à la fondation économique, diagnostic qui constitue une partie intégrante de l'épistémologie économique, vocable qu'on élucidera plus loin.

Dès lors, en vue d'un diagnostic judicieux et pertinent, nous constatons d'emblée, que :

Remarque 1. *La majorité des modèles économétriques généralisés ensuite à la macroéconométrie sont des modèles issus des sciences physiques ou des disciplines apparentées telles que la mécanique statistique, la mécanique des fluides, l'analyse des systèmes ou la théorie du signal.*

Remarque 2. *Les modèles économétriques sont des expressions d'hypothèses économiques formulées à priori.*

Remarque 3. *La part des nouvelles critiques adressées à la scientificité de l'économie vient essentiellement de la problématique de l'inférence statistique inhérente à l'aléatoire avec son lot de tâtonnement sur le futur et de marge d'erreur prévisionnelle, le plus souvent, de piètre qualité.*

Le grand économiste Schumpeter (considéré par Nème comme l'un des cinq plus grands économistes avec Marx et Keynes)⁶⁶ soulève la question pertinente en tout temps de la scientificité de l'économie dans sa formulation du débat passionnant suivant⁶⁷ : “La réponse à la question (l'économie est-elle une science ?) dépend de ce que l'on entend par “science“. Ainsi, dans le langage quotidien aussi bien que dans le parler académique..., le mot est souvent employé pour renvoyer à la physique mathématique. Ce qui exclut évidemment toutes les sciences sociales et aussi l'économie. L'économie dans son ensemble, n'est pas non plus une science si nous faisons de l'emploi de méthodes semblables à celles de la physique mathématique le caractère spécifique de la science. Dans ce cas, seule une faible partie de l'économie est scientifique, en certaines de ses parties et non en d'autres. Les susceptibilités concernant le “rang“ ou la “dignité“ ne devraient pas trouver là à s'exercer : appeler

65. La démocratie et le marché, Fitoussi J-P, Col Nouveau collège de philosophie, Ed Grasset 2004, 50-51

66. selon Colette Nème, dans son ouvrage : la pensée économique contemporaine depuis Keynes

67. Histoire de l'analyse économique, Schumpeter J.A, tome 1 : L'âge des fondateurs, Editions Gallimard 1983, p

un domaine donné “science” ne devrait pas faire figure de compliment ni de reproche “. Schumpeter donne au fur et à mesure de son analyse, étape par étape, une définition de la science à chaque étape. En voici sa première définition de toute science :

Définition 1. *Toute espèce de connaissance qui a fait l’objet d’efforts conscients pour l’améliorer est une science.*

Dans son commentaire de cette définition Schumpeter s’explique en ces termes “*De semblables efforts engendrent des habitudes d’esprit - méthodes ou “techniques” - et une maîtrise des faits créés de toutes pièces par ces techniques qui passent la portée des habitudes mentales et de la connaissance concrète de la vie quotidienne* “. Il propose, dans la foulée une seconde définition de la science , censée être équivalente à la première :

Définition 2. *Est une science tout domaine de connaissance qui a mis au jour des techniques spécialisées de recherche des faits et d’interprétation ou d’inférence-analyse.*

Dans l’étape suivante l’auteur met l’accent sur les aspects sociologiques des praticiens de la science et donne encore une troisième définition censée être équivalente, à son tour, aux deux premières :

Définition 3. *Est une science tout domaine de connaissance où des hommes, nommés chercheurs, hommes de sciences ou spécialistes, se vouent à l’amélioration du capital existant des faits et des méthodes et, au long de ce processus, acquièrent en ces deux points une maîtrise qui les différencie du “profane” et finalement aussi du simple “praticien”.*

Et l’auteur ajoute “*Beaucoup d’autres définitions pourraient être tout aussi bonnes. En voici deux que j’ajoute sans entrer en de plus amples explications* “

Définition 4. *La science est raffinement du sens commun.*

Définition 5. *La science est une connaissance outillée.*

L’auteur commente cette dernière définition ainsi “*La science étant une connaissance outillée, c’est-à-dire qui se définit par l’usage de techniques particulières, on a le sentiment qu’il nous faut y inclure, par exemple, la magie pratiquée dans une tribu primitive si elle emploie des techniques qui ne sont pas généralement accessibles et qui ont été développées et pratiquées dans un cercle de magiciens professionnels...Bien sûr nous devrions l’y inclure par principe ...(car si on ne l’incluait pas) l’exclusion de toute espèce de connaissance outillée reviendrait à déclarer que nos propres normes doivent être absolument valables pour tous les temps et pour tous les lieux. Mais nous ne pouvons le faire* “

Après avoir donné ces cinq définitions équivalentes de la science, Schumpeter va examiner l’économie à la lumière de ces définitions. Ainsi “*Puisque l’économie emploie des techniques dont le grand public ignore le maniement, et puisque les économistes sont là pour les cultiver , l’économie est évidemment une science, en conformité avec notre définition du mot (science)* “

Ces définitions schumpeteriennes de la science sont reformulées et appuyées par Hinti dans son ouvrage *“La science économique : une réflexion épistémologique et méthodologique”*. Ainsi : une science, quelle qu'elle soit, est identifiée selon Hinti (2000) par ses caractéristiques propres que sont ([52], p 82 et suite) :

1. *La science est une connaissance transmissible, extensible, systématique.*
2. *La connaissance scientifique est objective.*
3. *La science étudie les phénomènes .*
4. *La science établit des relations universelles : les lois .*
5. *La science permet des prévisions .*

2.1.3 Problématiques et controverses autour de l'apport de l'économétrie⁶⁸ à l'économie

La méthode hypothético-déductive due à Popper, on le sait, est une *“méthode de raisonnement consistant à raisonner déductivement à partir d'hypothèses dont la validité est appréciée a posteriori à partir de leurs conséquences”*⁶⁹. Elle est le propre des mathématiques *“elle est surtout évoquée à propos des mathématiques pour désigner la démarche qui tire de l'axiomatique toutes les conséquences possibles”*⁷⁰ et des sciences expérimentales *“La méthode hypothético-déductive se rencontre dans les sciences expérimentales, la conclusion est alors soumise à la vérification de l'hypothèse de départ”*⁷¹ comme les sciences physiques et naturelles car *“il n'existe d'ensemble cohérent de la nature dans les sciences physiques et naturelles que grâce à des raisonnements qui complètent les données de l'expérience au moyen d'une combinaison d'hypothèses”*⁷². La mise en œuvre de la méthode hypothético-déductive consiste en la construction au préalable de modèles théoriques censés représenter et reproduire la réalité du phénomène étudié et en second lieu, la mise en place d'expériences réelles ou simulées sur ces modèles théoriques pour pouvoir délibérer quant à la validité ou non de la théorie sous-jacentes. La construction de variables aléatoires ou de processus stochastiques est censée représenter le hasard probabilisable de la nature dans ses manifestations spontanées. Dès lors l'économétrie *“c'est à dire l'étude mathématique, sur la base de données statistiques des relations économiques particulières (sciences des lois de la production, de la distribution et de la consommation)”*⁷³ répond pleinement aux exigences scientifiques où l'expérience prend la forme d'un ensemble de réalisations de processus stochastiques. La validité d'un modèle théorique

68. Un préalable à toute analyse de l'apport des séries temporelles, ainsi que celui des séries temporelles à mémoire longue plus particulièrement

69. Vocabulaire des études philosophiques, S Auroux et Y Weil, collection faire le point, Ed Hachette 1993, p101

70. Dictionnaire de philosophie, G. Durozoi et A. Roussel, Ed Nathan 1997, p 191, mot=Hypothético-déductif

71. idem

72. W. Dilthey 1894 dans : Philosophie comme débat entre les textes, J. Médina, C. Morali, A. Sénik, Ed Magnard 1987 p 506

73. Dictionnaire de philosophie, G. Durozoi et A. Roussel, Ed Nathan 1997, p 191, mot = économie

est alors quantifiable via sa qualité d'ajustement par rapport aux données réelles ou simulées, pour un degré de signification donné.

L'une des conséquences de la scientificité de la méthode hypothético-déductive est la possibilité de remettre en cause toute théorie scientifique : *"C'est en effet parce qu'elle peut être fausse qu'une affirmation (hypothèse) est scientifique"*⁷⁴.

Comment est née l'économétrie qui constitue, de fait, le vrai laboratoire d'expertise et de validation des théories économiques, leur salle d'examen en quelque sorte et le moment de vérité de leurs modèles économétriques, et en quoi diffère-t-elle de l'analyse économique ?

Sa genèse est décrite ainsi : *" Tandis que Keynes préparait et publiait la « Théorie Générale », une mutation radicale s'amorçait , dont on ne prendrait toute la mesure que plus tard : la mathématisation de la discipline (l' économie)...(une mathématisation qui a commencé déjà il y a un siècle avec) un philosophe, Augustin Cournot, qui publie en 1838 le premier véritable traité d'économie mathématique intitulé, les « Recherches sur les principes mathématiques de la théorie des richesses » "*⁷⁵

Dans les années 1930, toutes les conditions favorables sont alors réunies pour l'émergence d'une quantification des *faits généraux* statistiques économiques. Dès lors l'ère économétrique s'annonce prometteuse et porteuse de grands projets. *"C'est un économiste..., Ragnan Frish, premier titulaire du prix de science économique en mémoire d'Alfred Nobel, qui joue un rôle déterminant dans la naissance et l'organisation de la nouvelle discipline qu'il baptise économétrie"*⁷⁶. Ainsi le projet se réalise. *" La réunion de fondation de la (première) société d'économétrie s'est tenue à Cleveland le 29 décembre 1930 (Présidée par Schumpeter)...Son objectif : promouvoir (au sein de la théorie économique)...un raisonnement constructif et rigoureux tel que celui qui en est venu à dominer dans les sciences naturelles (physiques)...l'éditorial du premier numéro d'Econometrica, leur revue, écrit par Ragnan Frish, stipule :...L'expérience a montré que chacun de ces trois points de vue, celui des statistiques, celui de la théorie économique et celui des mathématiques, est une condition nécessaire, mais non suffisante en elle-même, à une compréhension réelle des relations quantitatives dans la vie économique. C'est l'unification des trois qui est puissante. Et c'est cette unification qui constitue l'économétrie, (Econometrica, volume 1, 1933, p 2)"*⁷⁷.

2.1.4 Paradigme de l'économétrie des séries temporelles en économie

Dans notre quête épistémologique de la mémoire longue en économie, en macroéconomie et en finance, le paradigme d'intérêt pour notre quête, est le paradigme de l'économétrie des séries temporelles avec ses problématiques et ses anomalies. Il s'agit de la mise en œuvre de la théorie

74. L'économie : une science normale, dans : Alternatives économiques N°057 (07/2003)

75. La pensée économique depuis Keynes , Historique et dictionnaire des principaux auteurs, M. Beaud et G. Dostalier, Ed du Seuil 1993, p 89

76. idem p 92

77. idem p 92

de réfutation de Popper dont l'algorithme est présenté dans l'ouvrage de référence ([42], pp 179-181). Cet algorithme poppérien consiste à procéder étape par étape : d'abord par identification a priori, ensuite par estimation des modèles retenus, puis par vérification, puis par sélection du (des) meilleur(s) modèle(s), parmi tous les modèles valides eu égard des critères statistiques usuels, ces deux dernières étapes constituent l'identification a posteriori, puis, enfin, faire de la prévision. Il n'en demeure pas moins que les modèles macroéconométriques n'ont pas donné encore tout ce qu'on espérait d'eux en terme de valeurs ajoutées à l'analyse économique. Le bilan de l'investigation macroéconométrique, en dehors de son utilité technique, reste, somme toute mitigé, et a donné lieu parfois même, à des difficultés et des problématiques extra-économiques.

La fondation de l'économétrie des séries temporelles quant à elle a réalisé des avancées spectaculaires en terme de la mise en œuvre des modèles économétriques nouveaux et des processus statistiques novateurs.

Il n'en demeure pas moins que les problématiques de l'économétrie des séries temporelles restent des problématiques propres à cette fondation relativement jeune. Il s'agit principalement ([42], p13-16) de :

1. la problématique de la prévision et de sa qualité. En effet :

La qualité de toute prévision dépend, de fait, de la régularité de l'évolution dans le temps de la série analysée et de l'horizon de la prévision : plus la série est régulière dans le temps plus il est facile de prévoir ses valeurs futures pour les horizons de plus en plus petits. Les prévisions seront bonnes en général, dans les périodes de croissance ou pour des variables à variations régulières, par exemple de type linéaire ou exponentiel, pour les petits horizons, moins bonnes en périodes de récession ou pour des évolutions de type autre que linéaire ou exponentiel et des horizons de plus en plus lointains. Elles seront erronées dans le cas d'un changement structurel que rien dans le passé ne permettait de supposer. L'ajustement étant, à l'inverse, une prévision dans le passé, pour régénérer les valeurs manquantes de la série.

2. la problématique de la présence d'une tendance, car :

La présence d'une tendance (ou trend) non stochastique est un signe de non stationnarité de la série puisque sa moyenne empirique, dans ce cas, dépend du temps. La présence d'une même tendance (un même trend) de deux variables économiques traduit une forte corrélation entre ces deux variables sans qu'elles aient, nécessairement, un lien explicatif entre elles. Un tel lien explicatif entre deux variables ne peut être affirmé qu'après avoir retiré de la série analysée, au préalable, sa tendance (ou trend).

3. la problématique des variations saisonnières. En effet :

Tout comme la tendance (trend) non stochastique, la saisonnalité peut être un effet pervers dont il faut s'en débarrasser pour stationnariser la série avant tout calcul prévisionnel. En effet la saisonnalité apparaît dans la modélisation probabiliste de Box et Jenkins avec un facteur de la forme $(1 - L^S)^{D_S}$, S étant un nombre entier interdisant l'inversion du polynôme

autoregressif ce qui se traduit par la non stationnarité de la série. Ceci se manifeste clairement dans le domaine des fréquences à travers la densité spectrale $f(\omega)$ avec au dénominateur le facteur $|1 - e^{iS\omega}|^{2D_S}$ qui rendrait celle-ci de carré non intégrable.

4. la problématique de(s) rupture(s). De fait :
les ruptures dans l'évolution des variables économiques sont soit en niveaux (la moyenne est une fonction en escalier, sur des sous périodes, dans le temps) soit en pente (la moyenne est en dents de scie, sur des sous périodes, dans le temps). Elles peuvent être l'effet d'un changement de politique via des variables exogènes soit à la suite d'un choc exogène (par exemple chocs pétroliers et assimilés), soit d'une manière endogène par l'effet de changement de structure au niveau des relations structurelles entre variables endogènes (processus à changement d'états, comme le SETAR multiplicatifs et processus à changement de régimes)([47] p 27).
5. la problématique de la causalité et du décalage temporel. En effet :
l'observation simultanée dans le temps de plusieurs variables peut permettre de répondre à des questions liées à la causalité. Une fois déterminé le sens de la causalité, si cela est désormais possible, il faut savoir ensuite quel délai et pendant combien de temps (décalage temporel) la variable explicative influe sur la variable expliquée.
6. la problématique de la séparation du court et du long-terme due au fait que :
les influences entre variables prennent plus ou moins de temps : elles sont plus ou moins persistantes. L'une des problématiques les plus importantes de la macroéconométrie est de séparer ces relations persistantes (de long-terme) de celles (de court terme) qui ne le sont pas. Ces dernières s'interprètent souvent en terme d'ajustement.
7. la problématique de l'étude de la qualité des anticipations des agents car :
les agents économiques comme certains responsables d'administrations ou des chefs d'entreprises prennent leurs décisions, le plus souvent, sur la base d'anticipations inhérentes à leur jugement de la conjoncture aidés en cela par leurs départements d'études et de prévisions. Ils disposent ainsi d'outils d'analyses comparatives entre les valeurs prévues et les valeurs réalisées pour affiner leurs prévisions futures et établir leurs anticipations en conséquence. Cela nous permet de comprendre comment ils calculent implicitement leurs anticipations.

2.1.5 Problématique de la cointégration des séries temporelles multivariées

La cointégration est un fait statistique majeur inhérent à l'existence d'une tendance stochastique commune entre deux ou plusieurs processus $y_{1,t}, \dots, y_{k,t}$ intégrés non stationnaires : $I(d_1), \dots, I(d_k)$ avec $k \geq 2$ et $d_i \geq \frac{1}{2}$, $\forall i \in (1, k)$. Elle est le propre des variables économiques reliées, par des coefficients aléatoires à long-terme, comme la demande de monnaie et le revenu national, pour un seul et même pays, ou les taux de change de plusieurs pays. Si la cointégration a été le propre des processus intégrés non stationnaires, elle est aujourd'hui utilisée entre des processus intégrés

stationnaires et non stationnaires, paramétriques et non paramétriques.

2.1.6 Les origines diverses et multiples de la mémoire longue

Le phénomène de mémoire longue est au croisement de plusieurs phénomènes de nature et d'aspect divers et multiples. Il peut découler de l'auto-similarité (ou fractalité, stabilité) de certaines distributions et de certains processus. Il peut découler de l'intégration fractionnaire de type Box et Jenkins (ou intégration fractionnaire stochastique de certains Browniens). Il peut découler de l'agrégation de certains processus de mémoire courte, etc...

Les définitions multiples de la mémoire longue traduisent cette multitude de manifestations et d'aspects.

Ce phénomène de la mémoire longue se retrouve également à la fois au croisement des différentes anomalies en économétrie et en statistique (liées à la racine unitaire, à la cointégration etc...) de même qu'au croisement des méthodes de résolutions (groupes de renormalisation, estimations locales, etc...)

Paradoxes des lois des extrêmes et mémoire longue

Les lois statistiques modélisent parfois des phénomènes mystérieux. On montre dans l'ouvrage "La loi de l'arc sinus ou l'injustice fondamentale de la nature"⁷⁸ comment la loi arc sinus est une modélisation de l'injustice naturelle. Or la loi arc sinus est une loi $Beta(\mu, \nu)$ particulière ([33], p 50) avec $\mu = \frac{1}{2}$ et $\nu = \frac{1}{2}$, dont l'agrégation aboutit à la mémoire longue non stationnaire selon la définition de Granger⁷⁹. On appelle $Beta(\mu, 1-\mu)$ la loi de l'arc sinus généralisé⁸⁰ dont l'agrégation aboutit à la mémoire longue ([42], p 442) stationnaire si $0 < \mu < \frac{1}{2}$ et non stationnaire si $\frac{1}{2} \leq \mu < 1$.

Dans ses "Remarques sur l'épistémologie des phénomènes rares"⁸¹, l'auteur relève les remarques suivantes :

1. La pression sociale envers la science pour que celle-ci donne son verdict relativement aux phénomènes rares comme les séismes, les crues etc...pousse les scientifiques aux errements modélisateurs de la rareté et des phénomènes rares sans que ces modèles puissent être justifiées, ni techniquement réfutables "*de sorte que les fabricants de conclusions douteuses ne prennent aucun risque au sens épistémologique que Popper a donné à ce terme (en agissant ainsi)*".⁸²

78. La probabilité le hasard et la certitude, P. Deheuvels, Que Sai-je, Ed Puf 1996, p 120

79. La définition de Granger de la mémoire longue est étendue aux processus fractionnaires non stationnaires pour $\frac{1}{2} \leq d < 1$ in Journal of Econometrics, 2002

80. Feller, An introduction to Probability Theory and Its Applications, Volume 2, p 50

81. Splendeurs et misères des lois de valeurs extrêmes, Bouleau N., Centre de Mathématiques appliquées de l'école Nationale des Ponts et Chaussées de Paris

82. idem

2. De fait, les événements et phénomènes rares relèvent beaucoup plus du hasard fortuit que du hasard probabilisable “à la limite, les événements intrinsèquement uniques, peuvent être du domaine de l’incertain, sans être probabilisables”⁸³. Dès lors, on s’interroge “sur le sens même du concept de probabilité (et de loi) lorsqu’on l’applique à des événements exceptionnels”⁸⁴
3. Ainsi “les phénomènes rares de la nature (tempêtes, séismes etc...) ou de la vie économique ont des lois de probabilité nécessairement mal connues”⁸⁵
4. Ainsi “Un principe épistémologique se dégage : toute démarche attribuant une valeur numérique précise pour la probabilité (a posteriori) d’un phénomène rare est suspecte, sauf si les lois physiques régissant le phénomène en question sont explicitement et exhaustivement connues (a priori)”⁸⁶.

Le paradigme de la renormalisation

Hormis la forme explicite, définie partout, de la fonction d’autocovariance du bruit fractionnaire gaussien FGN, la mémoire longue pour les autres processus se présente comme une propriété asymptotique et ce, pour la plupart des processus ⁸⁷. Les méthodes de renormalisation employées dans l’estimation du paramètre de la mémoire longue sont fondées sur “le paradigme de l’invariance d’échelle”⁸⁸, sous l’un des six aspects suivants : l’intégration économétrique fractionnaire de type Box et Jenkins, l’auto-similarité des processus stochastiques, la fractalité pour les processus déterministes (comme les chaos), l’ α -stabilité de la distribution, l’asymétrie de l’information et le pas du temps à travers l’agrégation. Aussi attribue-t-on à la mémoire longue inhérente à tous ces phénomènes critiques à invariance d’échelle ou auto-similaires ([60], p51), sous l’angle de groupes de renormalisation bien choisis, un paradigme unifié qu’on pourrait appeler “paradigme de la renormalisation”⁸⁹. Force est de constater que “les théories à invariance de jauge (de Lorentz) sont (toutes) renormalisables”⁸⁹, et que d’après les physiciens Laguës et Lesne ([60] p77) : “Les méthodes de renormalisation sont utilisées dans tous les domaines de la physique. Incontournables dès que l’on s’intéresse à des propriétés asymptotiques de systèmes où des fluctuations existent à toutes les échelles spatiales et/ou temporelles, elles donnent accès à des propriétés intrinsèques, universelles, indépendantes des détails microscopiques du système. Leurs résultats, fondés sur l’invariance d’échelle, sont insensibles à de nombreuses simplifications et lacunes d’un modèle particulier. Elles ne sont pas nées en 1971 avec la résolution des comportements critiques [Wilson 1971], mais ont une longue histoire”. Néanmoins le paradigme de la mémoire longue est plus global que celui de la

83. idem

84. idem

85. idem

86. idem

87. la définition de la mémoire longue est une définition asymptotique“([61] p330)

88. Theory and Application of Long-Range Dependence, Taqqu et al, p 528, ligne 3

89. Dictionnaire d’histoire et philosophie des sciences, sous la direction de Lecourt D, Ed PUF 1999, mot=invariance de jauge, p 543-544

renormalisation, strictement dédié à la loi d'échelle constant.

Le *paradigme de la mémoire longue* est entendu, en économie, non pas dans le sens kühnien pur et dur, ni même une *matrice disciplinaire* dont les éléments stratégiques sont à la fois irréductibles et fondateurs de toutes les activités scientifique et dans toutes les sciences et savoirs d'une époque, mais plutôt comme un "épistémé"⁹⁰ inspirateur de bien de "programmes scientifiques de recherche"⁹¹.

Mémoire courte contre mémoire longue et problématique de l'agrégation

La mémoire longue, comme l'a si bien démontré Beran dans son ouvrage de référence en la matière ([12], p14), a deux aspects : physique et statistique. Seul l'aspect physique est réel et effectif, alors que l'aspect statistique en dépit de son apparence peut être fallacieux (spurious). Beran([12], p14), Gourieroux et Monfort ([42], p 441) reprennent l'exemple désormais célèbre de Granger(1980) d'agrégation d'une série de mémoire courte de type AR(1) issue d'une loi *Bêta* : $B(p, 1-p)$. Cette agrégation donne étrangement une série agrégée de mémoire longue. D'où la propriété ([42], p 442, 11.44) suivante :

Propriété 3. *La série agrégée des séries autoregressives d'ordre 1, associées à une loi $Beta(p, 1-p)$ avec $0 < p < 1$ est un processus fractionnaire. En particulier :*

- (a) *Elle est stationnaire à mémoire longue si $p < \frac{1}{2}$*
- (b) *Elle est non stationnaire à mémoire longue si $p \geq \frac{1}{2}$*

Mémoire courte contre mémoire longue et problématique du changement de régimes Markoviens

Tout comme pour l'agrégation des processus aurogressifs d'ordre 1 issus de la loi $B(p, 1-p)$, vu précédemment et qui aboutit à une mémoire longue, Rioublanc ([79], pp 94-95-97-107) a montré la présence de comportement de mémoire longue pour les modèles de chaîne de Markov sous la condition suivante : l'une au moins des deux probabilités de rester entre l'instant t et $t+1$ au même état : (p_{11} et/ou p_{22}) est voisine de 1. En effet dans son étude des deux modèles SETAR suivants, avec changement de régimes, en niveaux, régi à la date t par une chaîne de Markov à deux états 1 et 2 par $S_t = 1$, $S_t = 2$ et ε_t un bruit blanc gaussien :

$$(Modèle 1) X_t = \begin{cases} \mu_1 + 0.6 X_{t-1} + \varepsilon_t \text{ si } S_t = 1 \\ \mu_2 + 0.3 X_{t-1} + \varepsilon_t \text{ si } S_t = 2 \end{cases} \quad (Modèle 2) X_t = \begin{cases} \mu_1 + \varepsilon_t \text{ si } S_t = 1 \\ \mu_2 + \varepsilon_t \text{ si } S_t = 2 \end{cases}$$

Rioublanc a pu établir, en premier lieu, que la mémoire longue semble caractériser les processus stochastiques avec changement de régimes Markoviens, en niveaux, avec le constat (ou fait) empirique suivant : avec une forte probabilité (p_{11} et/ou p_{22}), plus les niveaux moyens μ_1 et μ_2 sont faibles, plus le comportement du type mémoire longue est manifeste. Ce résultat spécifique aux

90. terminologie de Michel Foucault

91. terminologie de Lakatos

processus stochastiques à changement de régimes semble être beaucoup plus général et concerne tous les processus stochastiques à événement rares, d'après les auteurs que Rioublanc cite comme Granger et Terasvirta (1999, [46]), Gouriéroux et Jasiak (2001, [43]), Granger et Hyung (2004, [44]). Effectivement, les auteurs Granger et Dittmann (2002, [27], pp 131-132, paragraphe 6) aboutissent à des conclusions similaires.

2.1.7 Contexte général de l'apparition du phénomène de mémoire longue

Le phénomène de mémoire longue (ou de dépendance de longue portée ou de dépendance forte ou de bruit en $\frac{1}{f}$) ([29], ligne 13-14), est connu des scientifiques, depuis le dix-neuvième siècle, dans des disciplines aussi diverses et variées que l'astronomie, l'agronomie, la physique, la chimie, ou la turbulence, tantôt sous le vocable d'*erreurs systématiques* chez Peirce (1837), d'*erreurs semi-systématiques* chez l'astronome Newcomb (1895) et de Karl-Pearson (1902), d'*erreurs semi-constantes* chez le chimiste Student (1927), ou de décroissance de proximité spatiale lente dans les essais d'uniformité agricole de Smith (1938) ([12], p34) et par Whittle en 1956.

Toutes ces investigations précoces récusent la formule $\frac{\sigma}{\sqrt{n}}$ pour la moyenne empirique d'un échantillon de taille n . “*Wilson et Hilferty (1929) affirment clairement que les données de Peirce (et leurs erreurs systématiques) illustrent que la dépendance en $\frac{\sigma}{\sqrt{n}}$ n'est pas scientifiquement satisfaisante en pratique, même pour estimer l'indépendance des moyennes*”. Mosteller et Tukey arrivent à la conclusion que “*Même en traitant avec une statistique, la plus simple de toutes, comme la moyenne arithmétique, il est souvent vital d'utiliser, en guise d'évaluation, la possibilité d'une incertitude latente. Obtenir une mesure valide de l'incertitude n'est pas une simple question de respect d'une formule*” ([12], p 39). Les travaux pionniers de Kolmogorov sur les processus auto-similaires, (1940), ont donné une impulsion forte aux études asymptotiques.

Le mouvement Brownien H-fractal (ou d-fractionnaire) est un outil fondamental dans l'analyse des divers aspects des processus à mémoire longue comme l'indique ce passage de l'article d'El-Nouty (1991, ([15], p 94)) : “*Le mouvement Brownien fractal entre de plusieurs manières dans le cadre des processus longue-mémoire. D'une part, ses propriétés élémentaires justifient sa place dans cette catégorie de processus; d'autre part, le mouvement Brownien fractal, (en tant que) processus Gaussien, illustre les difficultés techniques qui existent pour les processus longue-mémoire*”.

Il est désormais connu, dans l'approche classique ([34], Préface), que bon nombre de théorèmes limites, qui ont toujours été étudiés sous l'hypothèse d'indépendance des variables aléatoires sous-jacentes continuent de fonctionner, merveilleusement bien, pour certains types de structures de dépendances inhérentes à ces variables aléatoires. L'indépendance n'est donc pas nécessaire pour établir ces théorèmes limites classiques. Néanmoins, au fur et à mesure que la dépendance est de plus en plus forte, de plus en plus, apparaissent de nouveaux phénomènes imprévus. Dès lors, deux questions fondamentales se posent : la première, à partir de quelle valeur et de quelle manière cette dépendance conserve-t-elle ces théorèmes central-limites classiques intacts ? la seconde : existe-t-il des théorèmes limites analogues dans le cas de dépendance forte ?

La compréhension des relations complexes entre les hypothèses de dépendance et les procédures de calcul des limites s'est considérablement développée au cours de ces trois dernières décennies. La grande variété des structures de dépendance, parmi lesquels on compte les processus à mémoire longue, permet l'établissement d'une grande variété de théorèmes limites correspondants.

Le principe d'invariance d'échelle a constitué dans les années 1950, à proprement parlé, le théorème central limite fonctionnel. Entre temps de nouveaux outils plus performants sont apparus tel que le principe d'invariance forte d'échelle. Il s'agit, de fait, d'un théorème d'approximation presque sûre. *“Dans ce cas (de dépendance de longue portée = de mémoire longue), il y a, en principe, suffisamment de dépendance pour que le théorème central limite (fonctionnel) reste valide “*([34], Préface). C'est le cas des fonctions linéaires et non linéaires de variables aléatoires gaussiennes par usage des polynômes d'Hermite. Dans le cas général des variables aléatoires quelconques, le théorème central limite fonctionnel subsiste moyennant l'utilisation des polynômes d'Appell et la limite de la U-Statistique (Statistique d'ordre) débouchant sur des intégrales stochastiques multiples d'Itô qui jouent un rôle central dans l'établissement de ces théorèmes central-limites fonctionnels. Des théorèmes central-limites fonctionnels existent pour les lois infiniment divisibles par utilisation non plus des polynômes (de types Appell) mais directement via la mesure stochastique associée. Quand la variance du processus est finie (donc d'espérance également finie) et que les variables sont dépendantes, la covariance, on le sait, fournit déjà une description, même partielle de cette dépendance, pour une moyenne mobile quelconque. Ce n'est plus le cas lorsque la variance est infinie. Il faut utiliser alors les fonctions poids de la moyenne mobile concernée, et non plus la covariance, pour caractériser cette dépendance. Des théorèmes central-limites fonctionnels existent pour certains types de moyennes mobiles dont les domaines d'attractions (ou lois) sont stables, moyennant une utilisation d'une topologie non usuelle.

Les processus vérifiant la propriété d'auto-similarité, cet autre versant de la dépendance de longue portée ou de mémoire longue , constituent, depuis leur apparition en 1940 avec Kolmogorov, des sujets d'intérêt et d'investigation majeurs en probabilité, en statistique, en physique ou en turbulence.

Le lien étroit entre le paramètre H d'auto-similarité des processus auto-similaires Gaussiens ou α -stables d'indice α d'un côté et le paramètre d d'intégration fractionnaire ($d = H - \frac{1}{2}$ ou $d = H - \frac{1}{\alpha}$ respectivement) permet de préconiser le lien étroit entre l'auto-similarité et la mémoire longue ainsi que le passage entre les intégrales stochastiques par rapport à la mesure α -stable de Lévy d'un côté et l'intégrale fractionnaire de Riemann-Liouville de l'autre. Pour la physique, les processus auto-similaires constituent des outils précieux dans la compréhension des phénomènes critiques via les groupes de renormalisation de la théorie des champs.

Émergence de la statistique de la mémoire longue

Néanmoins c'est en hydrologie avec Hurst en 1951 que fut le développement de la statistique de la mémoire longue le plus abouti avec une investigation statistique novatrice faisant intervenir pour la première fois la statistique de la portée (standardisée) entre les extrêmes, tirant profit des avancées statistiques majeures dans les deux branches phares de la statistique des années 1940 : la loi des sommes partielles d'un côté et la loi des extrêmes de l'autre. Les deux théories étant unifiées par Gnedenko en 1943.

Historiquement, c'est Bortkiewicz, le premier en 1922, qui s'est intéressé à l'étude des valeurs extrêmes. En 1923 Von Mises avait introduit, sans la nommer, la notion fondamentale de caractéristique de la valeur la plus grande et sa relation asymptotique avec la moyenne des plus grandes valeurs de la loi normale correspondante. Cette étude poursuivie par Dodd (1923), suivis sur leur lancée, en 1925 par Tippet qui a tabulé ces plus grandes valeurs, pour la loi normale, pour des échantillons de taille allant de 2 à 1000. Néanmoins *“Le premier papier (pourtant passé inaperçu) basé sur le concept de type de distributions différentes de la (loi) normale est du à Fréchet (1927)...Fisher et Tippet (reprenant cette même approche de Fréchet) ont publié (en 1928) le papier (qui va devenir) basique sur les valeurs extrêmes...et trouvèrent en plus de la distribution asymptotique de Fréchet deux autres (distributions asymptotiques) valides pour d'autres types...En 1936 Von Mises ...donna les conditions suffisantes sous lesquelles les trois distributions asymptotiques sont valides. En 1943 Gnedenko (en) donna les conditions nécessaires et suffisantes”*⁹²

Les travaux de Mandelbrot des années 1960 ont porté sur la justification du phénomène de Hurst par l'application de la théorie de l'auto-similarité des distributions, introduites par Kolmogorov (1940) conceptualisées par Pinsker et Yaglom (1954). Mandelbrot et Van Ness (1968) vont définir le Bruit Gaussien Fractionnaire, comme la dérivée, au sens des distributions, du Mouvement Brownien Fractionnaire qui va s'avérer déterminante, par la suite, dans le développement de la statistique de la mémoire longue. En plus de la statistique des valeurs extrêmes et des lois associées respectivement de : Fréchet, Weibull, et Gumbel et de leur importance dans la modélisation de la mémoire longue, la statistique de la somme partielle fut la seconde grande découverte très féconde en statistique de la mémoire longue qui a pu mettre en avant les polynômes d'Hermite et les lois α -Stables de Lévy dans la modélisation de la p-value $P(|X| > x)$ en $x^{-\alpha}L(x)$ des queues des distributions épaisses ou leptokurtiques (il s'agit d'un fait stylisé majeur avec la grande variabilité et la ségmentation de la volatilité des séries financières). La statistique de la mémoire longue rompt avec la statistique habituelle construite autour de l'indépendance ou de la normalité asymptotique. La leptokurticité de la distribution exclut la normalité asymptotique non conditionnelle, de même que le comportement asymptotique de la fonction d'autocorrélation (à décroissance hyperbolique lente, dans le domaine temporelle), ou autour de zéro de la densité spectrale (son équivalent, dans le domaine fréquentiel), exclut à son tour l'indépendance de la suite des variables aléatoires sous-jacentes. La fractalité en tant qu'expression de l'auto-similarité a permis l'établissement de comparaison des comportements

92. Gumbel E.J., Statistics of Extremes, ED Dover publications INC, 1958 reprint 2004, p 3

dynamiques non linéaires de la mémoire longue fractales et chaotiques via l'exposant de stabilité de Lyapounov. Les relations entre la théorie des valeurs extrêmes et la théorie des sommes de variables aléatoires ont été établies depuis les travaux de Gnedenko dans son théorème relatif à la convergence en loi de la suite des maxima d'un échantillon de variables aléatoires indépendantes identiquement distribuées (notées v.a iid) $X_1, X_2, \dots, X_n$ de même fonction de répartition F

Théorème 5 (Théorème de Gnedenko 1943). *Soient $X_1, X_2, \dots, X_n$: n variables aléatoires iid de même f.r. F alors les deux assertions suivantes sont équivalentes :*

i) Il existe un réel $\alpha > 0$ et une fonction $L(x)$ à variation lente à l'infini tq

$$1 - F(x) = x^{-\alpha} L(x)$$

. ii) Il existe une suite de réels $(a_n)_{n \geq 1}$ et $\alpha > 0$ tel que :

$$\lim_{n \rightarrow +\infty} Pr\{ \max_{i=1, \dots, n} X_i < a_n x \} = \exp(-x^{-\alpha}) \quad \forall x \in \mathbb{R}^+$$

de plus, si l'une des deux assertions est vraie, la valeur de α est commune aux deux.

Si $F(0)=0$ c'est-à-dire si $(X_i)_{1 \leq i \leq n} \in \mathbb{R}^+$ et si $0 < \alpha < 1$ la théorie de la convergence en loi de la somme des v.a. X_i fournit le théorème suivant :

Théorème 6. *L'assertion i) du théorème ci-dessus est équivalente à :*

Il existe une suite de réels $(\alpha_n)_{n \geq 1}$ tq :

$$\lim_{n \rightarrow +\infty} F^{n*}(\alpha_n x) = G_\alpha(x) \quad \forall x \in \mathbb{R}^+, \text{ où } G_\alpha \text{ est la loi stable sur } \mathbb{R}^+ \text{ d'indice } \alpha$$

2.1.8 Auto-similarité, accroissements indépendants, accroissements stationnaires

Le processus aléatoire $B_H(t)$ dont la structure de la fonction d'autocovariance est en puissance du type $\mathbb{E}|B_H(t+s) - B_H(s)|^2 = |t|^2$, avec $0 < H < 1$ est devenu relativement familier aux statisticiens, au cours des trois dernières décennies, grâce aux travaux de Mandelbrot et Van Ness (1968). Le mathématicien Shoenberg est un habitué de cette fonction en puissance : $K(t, s) = |t-s|^{2H}$ depuis 1938 (en tant que noyau défini négatif), en dehors de toute considération probabiliste. Néanmoins il faut attendre les investigations probabilistes, de Kolmogorov de 1940 à 1941 pour mener à son terme l'analyse des différentes facettes de cette structure. En effet Kolmogorov l'a introduite dans sa modélisation théorique complète des phénomènes de turbulences, inhérents aux champs aléatoires à accroissements homogènes et isotropes. Ces champs aléatoires avaient deux types de structure : soit auto-similaires soit à auto-covariance en puissances, pour les grandes distances. En 1940 il présente, via une analyse spectrale la première classe de ces processus : les processus à accroissements stationnaires (notés si). En cette même année 1940 il décrit les processus

auto-similaires à accroissements indépendants (notés sssi) qui ont leur second moment fini. Ces processus (sssi) du second ordre peuvent avoir une représentation spectrale du type :

$$B_H(t) - B_H(0) = \int_{-\infty}^{+\infty} (e^{i\lambda t} - 1) |\lambda|^{-H-\frac{1}{2}} dB(\lambda) \quad (2.1)$$

avec $B(\lambda)$ un processus stochastique à accroissements orthogonaux homogènes.

Pinsker et Yaglom (1954) propose pour ce processus $B_H(t)$ une décomposition du type Wold en terme d'accroissements orthogonaux dB de bruit blanc comme suit :

$$B_H(t) - B_H(s) = \int_{-\infty}^t \left((t-p)_+^{H-\frac{1}{2}} - (s-p)_+^{H-\frac{1}{2}} \right) dB(p), \quad t > s, x_+ = \max(0, x) \quad (2.2)$$

Plus tard en 1968, Mandelbrot et Van Ness vont interpréter la deuxième partie de l'équation (2.2) comme une intégration fractionnaire de Weyl d'un bruit blanc et suggèrent d'appeler le processus $B_H(t)$: Mouvement Brownien Fractionnaire dans le cas où la mesure dB(p) est Gaussienne. Yaglom va développer à partir de 1955 des formules de prévisions linéaires explicites pour les processus $\{B_H(t)\}_{t \geq 0}$ à accroissements indépendants sur l'intervalle infini $\Delta =]-\infty, t[$ $t > 0$, $B_H(0) = 0$ de la manière suivante :

$$\hat{B}_H(t + \tau) = B_H(t) - \frac{\cos(\pi H)}{\pi} \int_0^{+\infty} \left(B_H(t) - B_H(t-p) \right) \left(\frac{\tau}{p} \right)^{H+\frac{1}{2}} \cdot \frac{dp}{p + \tau} \quad (2.3)$$

Pour trouver une prévision linéaire explicite sur un intervalle borné du type $[-a, a]$ Krein propose dès 1955 une méthode originale pour calculer ces intégrales en passant par le calcul de leurs inverses faisant intervenir des opérateurs $\varphi(x)$ de type Volterra appliqués à la fonction d'auto-covariance $K(t, s) = |t - s|^{2H-2}$ prise comme noyau. Le problème précédent se trouve transformé en un problème de recherche des fonctions $\varphi_a(x)$ associées à une fonction $f_a(t)$ donnée, satisfaisant l'équation :

$$\int_{-a}^a |t - s|^{2H-2} \varphi_a(s) ds = f_a(t) \quad \forall |t| \leq a \quad (2.4)$$

dont la solution $\varphi_a(s)$ associée à $f_a(t) = 1$, $\frac{1}{2} < H < 1$ et au Bruit Fractionnaire Gaussien (FGN) définie par Mandelbrot et Van Ness comme : $\xi_H(t) = \frac{dB_H(t)}{dt}$ est la famille des solutions :

$$\varphi_a(s) = \frac{\cos(\pi H)}{\pi} (a^2 - s^2)^{\frac{1}{2}-H}$$

Dans le cas où $0 < H < \frac{1}{2}$ il faut attendre 1967 pour que Grigoriev généralise l'équation (2.3) pour des intervalles finis du type $\Delta =]t_1, t_2[$.

$$\hat{B}_H(t) = \frac{\cos(\pi H)}{\pi} \int_{t_1}^{t_2} \left(\frac{(t-t_2)(t-t_1)}{(t_2-s)(s-t_1)} \right)^{H+\frac{1}{2}} \left(\frac{1}{t-s} - \frac{a(t)}{|s|} - b(t) \right) \cdot B_H(s) ds$$

avec $t \in]-\infty, t_1[\cup]t_2, +\infty[$ et a et b deux fonctions hypergéométriques dépendant de la position relative des points $(0, t, t_1, t_2)$.

La méthode de résolution explicite de Kreio utilisant les opérateurs de Volterra s'est avérée pratique dans la détermination du processus d'innovation $\{M_H(t)\}_{t>0}$ inhérent au processus $\{B_H(t)\}_{t>0}$, $B_H(0) = 0$ aussi bien dans le cas $0 < H < \frac{1}{2}$ que dans le cas $\frac{1}{2} < H < 1$. Cette approche de recherche de l'innovation $\{M_H(t)\}_{t>0}$ du processus $\{B_H(t)\}_{t>0}$, $B_H(0) = 0$ fut utilisée en 1969, d'abord par Molchan puis par Molchan et Golosov en la même année 1969, pour donner en définitive le résultat fondamental suivant :

$$M_H(t) = \int_0^t \left((t-s)s \right)^{\frac{1}{2}-H} \cdot dB_H(s) \quad (2.5)$$

Dans le cas d'un Mouvement Brownien Fractionnaire (FBM) le processus $M_H(t)$ défini par l'équation (2.5) a des accroissements gaussiens indépendants du type :

$$E|dM_H(t)|^2 = c|t|^{1-2H} \cdot dt$$

ce qui n'est rien d'autres qu'une intégration fractionnaire de type Liouville.

L'équation (2.5) permet de résoudre le problème de la norme dans l'espace de Hilbert via le noyau de reproduction :

$$EB_H(t)B_H(s) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$$

Selon les auteurs Molchan et Golosov (1969) cette norme prend la forme :

$$\|m\|^2 = c \int |\tilde{m}(t)|^2 |t|^{2H-1} dt \quad (2.6)$$

Avec

$$\tilde{m}(t) = \frac{d}{dt} \int_0^t \frac{(t-s)^{\frac{1}{2}-H}}{\Gamma(\frac{3}{2}-H)} s^{\frac{1}{2}-H} m'(s) ds$$

La distribution du bruit $\xi_H(t) = \frac{dB_H(t)}{dt}$ est invariante par rapport à une transformation linéaire fractionnaire due au fait que $B_H(t)$ est auto-similaire. Ainsi

$$B_H(at + a_0) \stackrel{d}{=} |a|^H B_H(t)$$

$$\left\{ \xi_H\left(\frac{at+b}{ct+d}\right) \left| \frac{ct+d}{\sqrt{|ad-cb|}} \right|^{2H-2} \right\} \stackrel{d}{=} \{ \xi_H(t) \}$$

Dès lors, la prévision linéaire $\hat{\xi}_H(t)$ pour $s \in [t_1, t_2]$ est la fonction filtrée :

$$\hat{\xi}_H(t) = \frac{\cos(\pi H)}{\pi} \int_{[t_1, t_2]} \left(\frac{(t-t_2)(t-t_1)}{(t_2-s)(s-t_1)} \right)^{H-\frac{1}{2}} \cdot \frac{dB_H(s)}{|t-s|}$$

2.1.9 Méthodes, groupes et opérateurs de renormalisation

Les méthodes de renormalisation ont leur origine en théorie quantique des champs, classées traditionnellement en électrodynamique quantique. Le terme de "groupe de renormalisation" fut introduit par Stueckelberg et Petermann ([1953]), et en 1954, Gell-Mann et Low proposèrent une

"renormalisation" de la masse de l'électron pour manipuler des séries divergentes, en combinant à la fois une sommation par paquets et une transformation des paramètres. Bien qu'elle se rapporte à une dynamique des systèmes complexes, dans leurs phases critiques, elle s'est avérée de portée pluridisciplinaire et a constitué et constitue encore l'un des outils les plus féconds et les plus pertinents dans la manipulation des séries autosimilaires divergentes ([63] p 26).

Dès 1966, Kadanoff a fait émerger l'importance de la notion d'invariance d'échelles sous-jacente à toutes les méthodes de renormalisation. Il proposa l'idée des "blocs de spins" pour étudier analytiquement les transitions se produisant dans les milieux magnétiques, point de départ d'une transposition systématique des idées de renormalisation à l'étude des transitions de phase critiques observées dans les systèmes de spins, les plus divers, mais également dans les fluides (Wilson 1971, 1975, Wilson et Kogut 1974, Fisher 1974).

L'importance des résultats apportés par cette méthode entre l'étude des phénomènes critiques rencontrés en mécanique statistique et les techniques de l'électrodynamique quantique a valu le prix Nobel de physique à Wilson (avec Kenneth) en 1982.

Les approches numériques introduites par Brezin et al (1974,1976), tout comme les comptes rendus des divers congrès sur "les phénomènes critiques et les transitions de phase", ont pu élaborer les méthodes perturbatives de la mécanique quantique et leur représentation diagrammatique, facilement transposables à l'analyse d'autres systèmes hors d'équilibre. De même l'émergence des méthodes de renormalisation numériques a permis de contourner les difficultés liées aux effets de taille finie dans la simulation des phénomènes critiques pour obtenir les exposants critiques par une simple analyse statistique des configurations renormalisées. Cette approche s'est avérée fructueuse dans l'étude de la percolation (inhérente aux milieux globalement désordonnés et localement ordonnés binaires en deux espèces), (Reynolds et al (1980)). Ainsi les conformations des polymères (ces molécules complexes formées à partir de l'assemblage de motifs moléculaires identiques appelés monomères) et dont l'étude physique se rattache à la théorie des processus stochastiques) sont modélisées comme des marches aléatoires sans recouvrement. L'extension de ces méthodes à des systèmes hors d'équilibre a fourni des résultats complets aussi bien sur les phénomènes critiques dynamiques, de même que sur les classes d'universalité associées ainsi que sur les propriétés à grande échelle de la turbulence décrite par des équations hydrodynamiques bruitées. Ce champ d'application est loin d'être clos, comme en témoignent de récents travaux sur d'autres équations aux dérivées partielles (Bricmont et Kupiainen 1992).

À la suite de la découverte des propriétés universelles du fameux scénario vers les chaos via l'accumulation de doublages de périodes, les méthodes de renormalisation ont été appliquées avec succès à la théorie des systèmes dynamiques dissipatifs pour décrire la transition vers les chaos déterministes suivant ce même scénario (Feigenbaum 1977, Couillet et Tresser 1978) ou par intermittence (Eckmann et al 1981, Hu et Rudnick 1982). Parallèlement à ces travaux, des méthodes adaptées à l'évolution vers le chaos dans des systèmes dynamiques hamiltoniens, ont été élaborés (travaux de Kadanoff 1981, Shenker 1982, Rand et al 1982, 1983, Lanford 1984) dans le cadre

de la théorie KAM (Kolmogorov(1954), Arnold (1963), Moser(1962)). Ce type de renormalisation temporelle s'emploie plus généralement dans l'étude de la dépendance par rapport aux paramètres ou au bruit des régimes asymptotiques (Collet et Lesne 1989). L'extension des idées de renormalisation aux systèmes spatio-temporels ouvre un nouveau champ d'applications : celui des processus stochastiques ainsi que les phénomènes décrits dans ce formalisme. Dès lors on peut aussi bien aborder l'étude des marches aléatoires de la sorte, que modéliser des polymères ou des phénomènes de diffusion en milieu désordonné (Bricmont et Kupiainen 1991) ou fractal, tels les amas de percolation selon la même approche. L'utilisation de la renormalisation s'est poursuivie en mécanique statistique telle que l'étude des transitions magnétiques au sein des systèmes de spins de plus en plus complexes, comme les verres de spins ou les systèmes frustrés⁹³ des transitions vers des phases, encore mal connues de la matière, comme les transitions des cristaux liquides. Les recherches se sont orientées également vers l'étude par renormalisation des structures et des mesures fractales. De nos jours, les approches numériques de renormalisation directe ou analytique se sont à la fois proliférées et généralisées.

Bien que le statut mathématique de la renormalisation soit précisé ouvrant la voie à des outils de la théorie des groupes, il n'en demeure pas moins que la renormalisation reste un des pôles les plus attrayants de la recherche en mathématiques.

2.2 Economie et mémoire longue

2.2.1 Les théories économiques revisitées, eu égard au paradigme de la mémoire longue

Si "*les effets et les conséquences de long-terme*" sont inscrits dans toute politique économique comme inscrites "*la poule dans l'œuf, la fleur dans la graine*", il n'en demeure pas moins qu'il est crucial, dans une perspective d'anticipation économique rigoureuse, sinon de juguler de tels effets de long-terme, du moins de les endiguer. Dès lors le paradigme de la mémoire longue, acquiert de fait, tout son droit, et prend de fait tout son sens.

À de rares exceptions, le phénomène de mémoire longue, ou de long-terme a été, par le passé, soit complètement ignoré des économistes, soit passé sous silence, quand cela restait non préjudiciable à la qualité d'ajustement ou de prévision de la conjoncture, pour le court et le moyen terme, moyennant des corrections au préalable des variations saisonnières, ou des cycles économiques latents. Mais dès qu'on raisonne à long-terme comme c'est le cas pour les firmes dans leur stratégie d'investissement, l'impact de la mémoire longue devient crucial. C'est Granger le premier, dans les années (1980), qui a pu ouvrir la première brèche en économétrie des séries temporelles à

93. le phénomène de "frustration" (est un mécanisme qui) empêche les éléments d'un système complexe de satisfaire simultanément leurs tendances individuelles à minimiser l'énergie de leurs interactions, in "Ordre et désordre : le magnétisme frustré montre l'exemple" CLEFS CEA - N° 56 - HIVER 2007-2008

mémoire longue, en introduisant l'intégration fractionnaire aux modèles ARIMA de Box et Jenkins, suivi sur son élan par Hosking(1981) puis par Geweke et Porter-Hudak (1983), pour une analyse économétrique complètement dédiée à la mémoire longue, dans le domaine fréquentiel. Il faut néanmoins attendre les années 1990 pour voir les prémises d'une étude économique intégrant la mémoire longue comme il a été mentionné dans ([61], p 328), ainsi : *“Dans le domaine économique..., on peut citer : - l'étude du comportement du produit réel par Diebold et Rudebusch (1989) et Sowell(1992), - l'étude de la dynamique des prix (Tieslau1992,...), -l'analyse du comportement du revenu disponible et de l'hypothèse de revenu permanent (Diebold et Rudebusch (1991), - l'étude de la consommation des ménages (Diebold et Rudebusch (1991)...) -les comportements des salaires réels et l'hypothèse de substitution inter-temporelle (Hassett(1990)), - l'analyse de l'évolution de la masse monétaire(Porter-Hudak(1990)), - la question de la prévisibilité de la rentabilité des titres et de l'hypothèse de marchés efficients (Greene et Fielitz (1977))....”*

L'indicateur d'une opportunité d'introduction du paradigme de renormalisation au sein de la théorie économique se reporte sur les tendances, les saisonnalités et les cycles en économie où peuvent s'opérer les non stationnarités modélisables soit par des FARIMA du type Boxet Jenkins soit par des GARMA du type Gegenbauer, puis sur la cointégration fractionnaire, puis sur l'agrégation de variables du type AR(1) qui peut engendrer une mémoire longue fallacieuse (ou artificielle).

Quant aux faits stylisés établis pour toutes les séries de rendements d'actifs financiers quelqu'ils soient, ils se caractérisent par :

- Une distribution des rendements de ces actifs à queues épaisses.
- Une ségmentation de la volatilité (rattachée à des cycles économiques latents comme le cycle des affaires et autres) avec hausse et basse volatilité.
- Une asymétrie des réponses face aux chocs.
- Une persistance entre les puissances des valeurs absolues de ces rendements.
- Une covolatilité entre les actifs d'un même secteur économique.

Les paradigmes standards des sciences économiques

L'épistémologie de l'économie se réduit de fait à l'épistémologie de ses paradigmes et de leurs communautés épistémiques de l'économie car même si *“l'esprit critique est indispensable pour la recherche et la démarche scientifique. La réflexion épistémologique (réflexion sur les paradigmes et sur les groupes de la science normale socialement constitués autour de ces paradigmes) en est un élément incontournable...(elle) intervient...afin d'éclaircir les idées , d'orienter le(s) débat(s) et d'enrichir les connaissances ...l'épistémologie et la méthodologie (ces deux éléments de base) sont nécessaires pour mieux cerner le statut et le contenu de la connaissance économique*⁹⁴

D'où la définition épistémologique d'un paradigme : ([52], p 163)

Définition 2. *Un paradigme est un ensemble de concepts, de croyances, de lois et de théories qui*

94. La science économique, Une réflexion épistémologique et méthodologique, Hinti S., Imprimerie Najah Al Jadida, Casablanca, 2000, p 3

servent de modèle de référence à une communauté scientifique..., à un moment donné. Autrement dit un paradigme est un ensemble de principes généraux pour un cadre intellectuel ou idéologique à l'intérieur duquel se déroule un débat scientifique. C'est aussi une sorte de constellation complète des valeurs, techniques, croyances...communes aux membres d'une collectivité scientifique donnée et sur la base desquelles, théories et modèles peuvent se développer.

Nous aurons, ainsi, autant de paradigmes que de *théories-cadres économiques*, basées chacune sur un "référentiel épistémique" et une "communauté épistémique" propres, depuis l'Antiquité et ce jusqu'à nos jours ([52] p 166) . En "adopt(ant) l'idée de matrice disciplinaire de Kühn, il est possible de dégager..., six étapes majeures, (en économie) qui peuvent ressembler au paradigme".⁹⁵ on a :

1. *Le paradigme à connotation philosophique, allant de l'antiquité au moyen âge : l'économie est issue de la philosophie, de la morale..., la richesse n'est pas une finalité en soi.*
2. *Le paradigme des mercantilistes, du 16^e au 17^e siècles : les sources de la richesse sont la monnaie et les métaux précieux : or, diamant, etc... ce paradigme fut dominant en Espagne, au Portugal, en France et en Angleterre.*
3. *Le paradigme des physiocrates, à partir du 18^e siècle : la terre agricole est la principale source de richesse. Quesnay, médecin, voit l'économie comme un organisme vivant dans lequel circule le surplus agricole. D'où la conception du circuit économique*⁹⁶.

Il n'en demeure pas moins qu'"Il est de coutumes de présenter la science économique comme comportant quatre courants (paradigmes) principaux : classiques, néoclassiques, keynésiens et contemporains."⁹⁷

Paradigme de la théorie économique classique

Dans "Le paradigme des classiques : l'économie politique est la recherche des lois de la reproduction et la consommation des richesses. Autrement dit, tout ce qui satisfait les besoins de la consommation et de la production. L'appropriation de la nature pour la transformer et l'adapter à nos besoins dépend du travail. Leur paradigme est composé de matrices telles que la recherche et l'harmonisation des intérêts individuels, le travail comme facteur de production (valeur-travail)...Les idées classiques étaient dominantes en Angleterre pendant longtemps, puis contestées et remplacées vers la fin du 19^e par le marginalisme (néoclassique). (Dans) Le paradigme des néoclassiques : l'économie politique est la recherche des lois du calcul optimal des agents économiques pour une meilleure allocation des richesses. L'essentiel de leurs analyses a pour but la détermination des prix relatifs des biens, des services et des facteurs de production. Leur paradigme fait référence à des éléments tels que la valeur-utilité, la rémunération des facteurs de production à leur productivité

⁹⁵. Hinti Saïd, La science économique : une réflexion épistémologique et méthodologique, Imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p 166

⁹⁶. idem

⁹⁷. idem

*marginale, l'optimisation...La révolution marginaliste est une vraie révolution (dans une optique) kühnienne parce qu'elle a provoqué des changements profonds dans la théorie économique.*⁹⁸

Les principes fondateurs (ou postulats) du paradigme classique, dans sa première version, profondément “*optimiste(s) et progressiste(s) (quant) à l'intelligence humaine*”⁹⁹ sont : l'invariance, l'exclusion de la causalité, la souveraineté du monde réel sur le monde monétaire, la souveraineté des choix conscients¹⁰⁰, ou d'une manière plus explicite :

1. Il existe une nature humaine invariante.
2. La monnaie a un caractère totalement négligeable.
3. Tous les êtres humains participent ensemble à déterminer l'état du monde économique et social en l'absence de tout pouvoir.
4. La nature humaine, si elle est éclairée, peut parvenir à l'état optimal qui s'impose au pouvoir.

La deuxième version du paradigme classique, qui va compléter la première, s'intéresse plus au marché et “*ne s'intéresse pas à l'individu en tant que tel. Le marché (régulateur dynamique (idem p 85) permettant de déterminer des situations d'équilibre (idem p 68)) est un éclairage donné en présentant d'emblée les résultats de choix collectifs*”¹⁰¹ “*Ainsi, la seconde représentation classique par l'intermédiaire du marché ouvre un passage spontané à une analyse dynamique en termes d'élasticité de l'offre et de la demande*”¹⁰² sur tous les marchés des biens et services concourant en cela à un équilibre général toujours accessible, constituant “*un point de référence ou d'attraction. (En effet) Même si les équilibres partiels ne s'obtiennent pas simultanément, l'équilibre général (lui) continuerait de s'imposer comme un appel permanent au fonctionnement d'une économie nationale*”¹⁰³. “*Les hypothèses de Walras (instituant l'équilibre général au sein d'une économie d'échange) sont celles des modèles classiques : rationalité des agents (dont le postulat de non illusion monétaire (idem De Mourgues, p 253), avenir certain, élasticité unitaire des anticipations, prix flexibles, (distributions de biens...déterminées une fois pour toute, la monnaie...(étant) un simple instrument de transaction (idem De Mourgues, p 253).*”¹⁰⁴. Or “*La loi de Walras ($\sum_{i=1}^n p_i x_i^{XD} = 0$) et la théorie quantitative (de Pigou) ($M = PkY$) sont en contradiction, dès lors que l'on accepte la loi de Say (demande de biens = offre des biens : $\sum_{i=1}^n p_i x_i^d = \sum_{i=1}^n p_i x_i^s$)*”¹⁰⁵. C'est alors grâce à l'apport de Patinkin qui, en s'appuyant sur l'effet d'encaisse réelle et le postulat de non illusion monétaire, va rendre le système classique cohérent, en sacrifiant au passage la Loi de Say, à l'origine de cette contradiction.

Le développement macroéconomique, opéré depuis 1954 par les néo-classiques, en riposte au keynésianisme ambiant, néo-classicisme avec son fameux rendement d'échelle constant inhérent à l'ho-

98. idem

99. Les quatre piliers de la science économique, Cotta A. et Calvet C., Ed Fayard 2005, p 42

100. idem, p 41

101. idem, p 66

102. idem, p 86

103. idem, p 88

104. La monnaie, système financier et théorie monétaire, De Mourgues, 2e Ed Economica 1990, p 354

105. idem, p 358

mogénéité d'ordre un de sa fonction de production de type Cobb-Douglas : $Y(K, L) = aK^\alpha L^\beta$ tel que $\alpha + \beta = 1$, sa théorie du taux d'intérêt classique et le rôle central attribué à sa capacité capitalistique $\frac{K}{L}$ au cœur de sa représentation, a permis "d'installer définitivement le paradigme classique dans une perspective dynamique qui pourra lui assurer sa longévité tout en affirmant et en apportant la preuve du déterminisme économique de l'évolution" ¹⁰⁶. Le classicisme a pris une forme plus actuelle avec Debreu et Hicks, le premier pour sa "nouvelle théorie de la valeur en 1959, où il s'efforce d'établir une identité formelle entre l'économie du certain et de l'incertain... (via l'introduction) de variables d'écart" ¹⁰⁷, le second pour l'introduction de biens et services objet(s) de négociation de livraison différée et conditionnelle selon des mécanismes semblables à des marchés à terme ou de futures "dès lors "La représentation classique se réduit à de l'algèbre, et la science économique... (à) une axiomatique algébrique qui s'opère sur des êtres algébriques" ¹⁰⁸. Ainsi, et "au terme de (tous) ces efforts, le paradigme classique est totalement constitué dans l'ordre théorique et statistique tout en affirmant sa fidélité aux principes de base ..., et ce, toujours sans faire la moindre référence à la monnaie" ¹⁰⁹, sans agrégation aucune, puisque tout le raisonnement classique est basé sur la rationalité de deux agents représentatifs : l'un pour tous les consommateurs, l'autre pour tous les producteurs. L'État, dans la vision classique et néo-classique n'étant qu'un agent, parmi d'autres.

Paradigme de la théorie économique de Keynes

Dans "Le paradigme keynésien :...l'économiste s'intéresse de plus près aux quantités globales : revenu national, consommation nationale, épargne globale. Le succès remporté par les idées de Keynes à partir des années trente est assimilée à...(une) révolution scientifique (selon la conception kühnienne) : on parle alors de...révolution keynésienne" ¹¹⁰

Les principes fondateurs du paradigme keynésien, nés "après la parution de *La Théorie générale* (1936)" ¹¹¹ et conçus pour le court terme car "*La cinématique keynésienne est du domaine du court terme. Elle s'intéresse au passage d'une situation déterminée à une autre, entre deux temps, et non pas à l'évolution générale, à la trajectoire ou à la dynamique de long-terme*" ¹¹², en opposition, à la fois, au paradigme classique (statique ¹¹³, du laisser-faire sans faire de politique économique), classicisme qui stipule que toute économie libérale, évolue inéluctablement vers un équilibre général, avec résorption, à l'équilibre, du chômage, mais aussi et surtout, en opposition au paradigme marxiste dynamique ¹¹⁴, qui stipule que toute économie capitaliste, évolue inéluctablement, par le

106. Les quatre piliers de la science économique, Cotta A. et Calvet C., Ed Fayard 2005, p 101-102

107. idem p 106-107

108. idem p 107

109. idem, p 107

110. Hinti Saïd, La science économique : une réflexion épistémologique et méthodologique, imprimerie Najah Al Jadida, 2000, Casablanca(Maroc), p 167

111. idem p 243

112. idem p 251

113. idem p 251

114. idem p 251

biais de “la loi de la baisse tendancielle du taux de profit”¹¹⁵, sur le long-terme vers le communisme égalitaire, avec résorption, à coup de révolutions, expressions de la dialectique historique matérialiste, résorption à la fois des intérêts privés et des classes et la fin du capitalisme. Ainsi “Le positionnement de la politique de Keynes (il en fallait une, devenue une nécessité, pour endiguer l’activité économique et résorber le chômage) se situe entre le laisser-faire (classique) et la préparation de la révolution (Marxiste)”¹¹⁶. C’est “au nom de l’efficacité et de la liberté”¹¹⁷ que Keynes va “construire un autre paradigme justifiant l’action”¹¹⁸ de l’État dans le domaine économique, monétaire et comptable. D’où :

1. La nécessité pour l’État d’avoir une politique économique. Mais aussi
2. La nécessité pour l’État d’avoir un contrôle étroit à la fois de l’émission et de la création de la monnaie, centrale et bancaire et une politique directrice des taux d’intérêt pour toutes les échéances : En effet “Le monisme de Keynes s’exprime d’une façon simple. C’est la quantité de monnaie qui est responsable du chômage, et la relation causale est en sens unique”¹¹⁹.
3. La nécessité pour l’État d’avoir une comptabilité nationale :
 “Keynes invente le concept mesurable...En économie fermée,...(et) retient six concepts (dont il faut suivre étroitement l’évolution, d’abord deux variables monétaires : La masse monétaire M , le taux d’intérêt r , puis quatre variables réelles que sont : le produit effectif ou produit national Y , la consommation effective des biens C , l’investissement ou formation du capital fixe I , l’épargne ou le revenu non consommé S ”¹²⁰

La nature de chacune des six variables est conforme aux hypothèses keynésiennes standards. Les politiques économiques : monétaires et budgétaires s’opèrent via les fameux multiplicateurs, monétaire de l’offre de monnaie et budgétaire de (l’offre) des dépenses publiques. La première hypothèse concerne la fonction d’offre globale des entreprises Y_{S-E} (sans l’État), supposée ne dépendre que du niveau de l’emploi N . Autrement dit $Y_{S-E} = f(N) \equiv C + I$. L’introduction de l’État comme producteur de services, avec un revenu Y_E . Ce revenu ne dépend, lui aussi, que du niveau N de l’emploi. Il est dépensé intégralement, dans le seul poste G des dépenses publiques, autrement dit $Y_E \equiv G$. Le revenu Y_S de toutes les entreprises, y compris celui l’État, en tant que producteur de services, est $Y_S = Y_{S-E} + Y_E$. Le revenu global est intégralement dépensé en $C + I + G$. À l’équilibre (de court terme, sans le marché extérieur), la fonction de demande globale effective, satisfait l’identité comptable (*source*) $\equiv$ (*emploi*) ou $Y_D \equiv C + S + T = Y_S \equiv C + I + G$. Une politique économique consiste alors permet à agir sur la fonction d’offre Y_S via ses postes comptables que sont C , I et G . L’investissement global (sur le long-terme) est inhérent aux seules décisions des entreprises qui s’opèrent dans un certain climat psychologique,...(ou) “état des anticipations à long-terme”, en fonction duquel, les

115. idem, p 245

116. idem, p 245

117. idem, p 244

118. idem, p 244

119. idem, p 247

120. idem, 247-248

agents (les entreprises) déterminent un taux d'intérêt moyen de longue période, qu'ils considèrent comme normal " ¹²¹ avec $I=I(r)$. Toute la logique du paradigme keynésien est de stimuler l'activité économique, par le jeu (ou politique) de l'offre de monnaie et/ou des dépenses publiques, en quête du plein emploi, via les multiplicateurs de la demande effective, égale, à l'équilibre, aux différents postes de l'offre, multiplicateurs qui ne sont que des dérivées partielles, de cette demande effective, prises une à une, toutes choses égales par ailleurs, par rapport aux différentes variables des différents postes de l'offre. À l'équilibre $I=S$, $L=M$, $N_s = N_d$ pour l'offre et la demande d'emploi, les multiplicateurs $\frac{\Delta Y_s}{\Delta C}$, $\frac{\Delta Y_s}{\Delta I}$, $\frac{\Delta Y_s}{\Delta G}$ avec, à l'équilibre, $I = S$ et $L = M$, en utilisant les fonctions explicites, de toutes ces variables agrégées et globales I , S , L , M , en fonctions, par exemple de Y et de r , le taux d'intérêt, ou de ω , le taux de salaires, sur les trois marchés : des biens, de la monnaie, et de l'emploi. Les néokeynésiens étendent les concepts keynésiens originaux, relatifs à une économie fermée(modèle jugé naïf), à une économie ouverte, ¹²², en introduisant un quatrième équilibre, sur le marché des échanges avec l'extérieur, entre exportations X par les entreprises et importations M à destination des ménages, de telle sorte qu'on ait $Y_D \equiv C + S + T + M = Y_S \equiv C + I + G + X$, avec la possibilité d'utiliser le nouveau multiplicateur $\frac{\Delta Y_s}{\Delta X}$ qui stimule l'exportation.

Remarquons qu'hormis les théories de la croissance et assimilées les théorie classique et néoclassique de l'économie restent somme toute fondamentalement statiques. Il ne peut y avoir en leur sein ni de dynamiques fractales ni à mémoire longue. La théorie keynésienne qui reste fondamentalement de court terme, n'a de dynamiques ni fractale ni à mémoire longue. Les anticipations des agents y sont de simples moyennes de longue période, de l'agrégat. Il ne peut y avoir, donc, d'analyse fractale de l'évolution dynamique, suite à un choc exogène, de la trajectoire, des divers équilibres transitoires successifs, du système économique, tendant vers son équilibre final, selon l'une et l'autre des deux théories, aussi bien classique qu'keynésienne. La synthèse keynésienne-néoclassique, en vogue depuis les années 1970, bien que hors de notre propos, parce que, hybride en terme de paradigmes, ne traite pas de l'évolution fractale du système économique, hormis l'analyse purement technique inhérente à l'analyse macroéconométrique des séries temporelles et plus particulièrement des modèle VAR ou Vectoriels AutoRégressifs comme le montre l'exemple traité dans la *section 10.4 Chocs et multiplicateurs* de l'ouvrage de Gouréroux et Monfort ([42], p 385 et suite)

2.2.2 Projet de Stiglitz de refonte intégrale de l'économie

Paradigme de Stiglitz de l'économie monétaire et équilibre général du crédit de 2005

Le paradigme de Stiglitz s'inscrit dans le cadre de la théorie de l'équilibre général dans un régime de concurrence bancaire. Les principes de ce paradigme sont donnés par Stiglitz ([91],p6) ¹²³ sous

121. La monnaie, système financier et théorie monétaire, De Mourgues, 2e Ed Economica 1990, p 368

122. Dictionnaire des sciences économiques, ss la direction de Jessua C., Labrousse C., Vitry D., avec Gaumont D.(conseiller scientifique), Ed PUF 2001, Les canaux de transmissions, p 598

123. économie monétaire : un nouveau paradigme, Ed Economica 2005, p 65 et suite

la forme de principes de base. Ainsi pour qu'une banque de crédit prospère grâce à la progression de son activité de prêt, qu'elle soit de dépôts ou d'investissement, elle doit appliquer les cinq recommandations suivantes, érigées en principes d'un seul et unique paradigme d'équilibre général de crédit :

- La banque n'emprunte jamais pour souscrire des bons du trésor.
- Une baisse de la valeur nette de la banque entraîne une diminution de l'activité de prêt.
- Un accroissement du risque, préservant le rendement moyen, réduit l'activité de prêt.
- une augmentation des réserves obligatoires entraîne, une réduction de l'activité de prêt.
- une augmentation de taux sur les bons du trésor engendre moins de prêts.

Pourtant ce paradigme n'a pas servi, à prémunir les banques de la crise de 2008. Sa carence, aussi bien, vis à vis des modèles financiers usuels, que des modèles de mémoire longue, démontre son impertinence pour être envisagé comme paradigme solvateur de la macroéconomie.

Stiglitz et son projet de réforme des sciences économiques

La responsabilité de la crise économique de 2008 incombe, en partie, selon Stiglitz (2010,[89]), au manquement grave de perspicacité, dû à un diagnostic fallacieux, de la part des économistes adeptes de l'équilibre général de marché et du libre échange, qui n'ont rien vu venir, selon lui, avant que le désastre de 2008 ne puisse survenir. Non seulement ils *“ont échoué dans leur mission fondamentale de prédiction et de prévoyance”* dit-il, pis encore, ce sont les sciences économiques en l'état, elles mêmes qui ont servi d'infrastructures théoriques pitoyables et de préalables calamiteux à ce cataclysme mondial. Dès lors il faut repenser les sciences économiques, dit-il, de fond en comble. *“L'économie (néo-classique) est censée être une science prédictive ; pourtant, beaucoup de prédictions cruciales de l'économie néoclassique peuvent être rejetée d'emblée. La plus évidente est (l'assertion néoclassique qui prétend) que le chômage n'existe pas. De même que l'équilibre du marché implique que l'offre et la demande de pommes (ou de n'importe quel autre bien ou service) soient égales, de même que l'offre et la demande de travail sont égales... Ce (ne sont pas) le(s) seule(s) conclusion(s) bizarre(s) de l'économie néoclassique...”* ([89], p 432-433)... *“Ses adeptes (néoclassiques) soutiennent aussi que le rationnement de crédit n'existe pas : tout le monde peut emprunter autant qu'il veut, à un taux d'intérêt qui reflète comme il convient le risque de défaut de paiement...(Ils soutiennent encore que) la structure financière des entreprises (est indépendante des modes de financement) en empruntant ou en vendant des actions (en cédant des parts de son capital).* “ Or les réformes des politiques économiques proposées par Stiglitz, le keynésien, dépassent de très loin le cadre interventionniste régulateur keynésien standard de l'État. Il s'agit d'une stratégie globale (doctrinale, d'abord, avec l'instauration d'un nouvel ordre capitaliste) assortie d'une logistique (la réforme de la science économique). Stiglitz opte, cette fois-ci, plutôt pour une politique économique agressive, de la part des États, où le long-terme prime sur le court et le moyen terme. Il planifie la stratégie de batailles mondiales, face à des défis clairement identifiés, à relever : *“le (premier) défi le plus flagrant est l'écart entre la demande mondiale et l'offre mon-*

diale...le (deuxième) plus grand défi est ... le réchauffement de la planète...(le troisième) se sont les déséquilibres (économiques) mondiaux (entre les pays développés et les pays en développement)...Le quatrième défi (est) le casse tête industriel (qui a fait que) les progrès de la productivité ont été tels que, même quand l'industrie se développe, l'emploi diminue, et ce modèle va probablement persister. Le cinquième défi est celui des inégalités (dans) la répartition des revenus et de la fortune dans le monde...Le (sixième et le) dernier défi est (le défi de) la stabilité...des institutions financières mondiales)". Dès lors les sciences économiques deviennent des machines à outils théoriques, au sein d'un système économique mécanisé, assujéti, organisé pour servir des plans de batailles rangées dans une guerre totale contre tout déraillement aux normes économiques requises. Il s'agit d'une vision économique agressive qui rompt de fait avec le libéralisme, aussi bien dans ses fondements que dans ses méthodes, et qui rejoint par beaucoup de ses aspects autoritaires le paradigme économique des régimes socialistes. Stiglitz identifie les carences de la science économique actuelle en plusieurs niveaux, après l'avoir clairement définie : *"La science économique est une science sociale. Elle étudie les interactions d'individus pour produire des biens et services. Pour comprendre leurs interactions, il faut avoir une idée plus large de leur façon de se comporter"* ([89] p 438). Or le modèle économique néoclassique (d'obédience microéconomique) a toujours fonctionné selon Stiglitz sur la base de plusieurs hypothèses érigées en dogmes absurdes telle que :

1. la rationalité des agents économiques : *"La croyance de la rationalité est bien ancrée dans la science économique. L'introspection - et, plus encore, un regard sur mes semblables - m'a convaincu que c'était une absurdité"* ([89] p 438).
2. l'avidité de l'agent économique *"qui sous-tend les modèles économiques dominants ...(et qui n'est ni plus ni moins qu'un) individu calculateur, rationnel, égoïste et intéressé. Aucune place n'est faite (ni) à la sensibilité humaine (chez cet agent économique), (ni) au civisme, (ni) à l'altruisme"* ([89] p 439).

La conséquence de cette vision systémique mécaniste des sciences économiques par Stiglitz, dans une optique de gestion publique autoritaire globale, ne peut permettre d'y avoir d'autocorrélations spontanées, par le simple fait d'endogénéisation des dépendances fortes intrinsèques, susceptibles de favoriser l'émergence d'un quelconque processus de mémoire longue.

Projet de Stiglitz pour repenser la macroéconomie

Le constat le plus récent de Stiglitz vis à vis des modèles macroéconomiques en vigueur est sans appel ([90], Abstract). Il l'affirme sans nuance, que : *"Les modèles macroéconomiques standard ont échoué, par tous les tests les plus importants de la théorie scientifique. Ils n'ont pas prévu que la crise financière (de 2008) se produirait, et quand elle a lieu, ils ont minimisé ses effets. Les autorités monétaires ont permis à des bulles (spéculatives) de se développer et se sont concentrées sur le maintien du taux d'inflation faible, en partie, parce que les modèles standard suggéraient qu'une inflation faible était nécessaire et presque suffisante à l'efficacité et à la croissance (en conformité avec les conclusions inhérentes à la courbe de Phillips pour résorber le chômage). Une fois la crise*

neutralisée, les décideurs s'appuient (de nouveau) sur les (mêmes) modèles (macroéconomiques qui ont déjà) échoué. Malgré la diversité (des modèles et les approches inhérentes à) la macroéconomie (standard), la somme de ces échecs souligne la nécessité (à la fois) d'un réexamen fondamental (et profond) de (ces) modèles (et de ces approches mais aussi) une réaffirmation des leçons (tirées) de la théorie de l'équilibre général moderne (le fameux Dynamic Stochastic General Equilibrium DSGE) qui ont, apparemment été oubliées, dans les années qui ont précédé la crise (de 2008)“. Ce constat est partagé par Fitoussi en ces termes :*“ On est entré dans un nouveau monde, cela ne fait pas de doute. Mais cela fait un quart de siècle que ce changement est à l'oeuvre. La brutalité de la crise (de 2008) a servi de révélateur. Entrer dans un nouveau monde n'implique pas nécessairement l'avènement d'un nouveau capitalisme car celui-ci ne peut être que d'ordre doctrinal. Pour l'instant, nous vivons toujours sur les règles du capitalisme libéral. La preuve, c'est que, à peine guérie de son extinction de voix qui l'a conduit à laisser aux institutions intergouvernementales le monopole de l'action, la Commission européenne est repartie dans les postulats d'avant-crise. Elle a remis en selle le Pacte de stabilité et mis vingt pays sur vingt-sept en procédure pour déficit excessif! Le plus grand obstacle à l'éclosion d'un nouveau capitalisme est donc d'ordre intellectuel. Notre vision du monde s'est révélée erronée. Nous vivions sur le postulat que les marchés étaient efficaces. Or, ils se sont révélés d'une totale inefficacité“*¹²⁴ Dans un article récent, Chatelain et al (2012) [22], révèlent les faits stylisés, bien identifiés, inhérents aux crises financières récurrentes et aux mécanismes qui les génèrent. Ils récusent par la même les modèles DSGE que Stiglitz estime capables s'ils étaient mis en oeuvre en l'état, de nous avoir préservés de la crise de 2008.

Le diagnostic de Stiglitz pointe du doigt les raisons endogènes (propres aux modèles macroéconomiques standard eux mêmes), des ralentissements de l'activité économique d'avant la crise. Il émet aussi des recommandations afin de procéder à des changements structurels du système économique, pour acquérir l'immunité nécessaire face à la crise, contrairement à ce que les modèles standard ont toujours soutenu, selon lui.

Rapport Stiglitz pour une réforme radicale du système monétaire et financier international

Il s'agit du rapport de la commission des 18 experts internationaux, désignés par l'ONU et présidés par Stiglitz, en forme d'états généraux restreints, pour se pencher sur les causes de la crise mondiale de 2008 et les modalités de préventions à mettre en oeuvre par les états membres. Les conclusions de la commission [25] se situent en trois niveaux : le diagnostic sur l'état des lieux et les dysfonctionnements du système, le projet de sauvetage salutaire et les recommandations à mettre en oeuvre pour y parvenir. La commission constate ([25], p 231) que :

1. Les institutions nationales à vocation économique et leurs cadres réglementaires associés, ne sont plus adaptés au monde globalisé actuel, *“pour (être en mesure de) garantir la concurrence*

124. Jean-Paul Fitoussi : "l'avènement d'un nouveau capitalisme ne peut être que d'ordre doctrinal", In la tribune du 07/01/2010

(nationale), protéger les consommateurs et les investisseurs (nationaux), gérer les faillites (nationales), faire respecter les contrats et assurer la stabilité de l'économie (nationale)“.

2. *“la mondialisation économique est allée plus vite que l'élaboration d'institutions mondiales capables de contribuer à la gérer“.*
3. *“l'expansion de l'activité économique transfrontière“* est sans précédent avec des courroies de transmission incontrôlables.

Dans ses conclusions, la commission préconise de :

1. concevoir des ripostes macroéconomiques mondiales.
2. concevoir un régime réglementaire mondial strict.
3. créer, à l'image de ce qui a été créé au niveau national, des institutions internationales à vocation économique et des cadres juridiques internationaux *“pour (être en mesure de) garantir la concurrence (internationale), protéger les consommateurs et les investisseurs (internationaux), gérer les faillites (internationales), faire respecter les contrats (internationaux) et assurer la stabilité de l'économie mondiale“*
4. réformer, de fond en comble, les institutions internationales existantes actuellement.
5. créer un *“Conseil de coordination économique mondiale, pour rendre la gestion de l'économie mondiale plus cohérente“*.

2.3 Appréhension d'une réhabilitation LRD de l'utilitarisme de John Stuart Mill

Le réexamen attentif et approfondi de la pensée économique depuis les classiques, nous surprend à plus d'un égard par son originalité et sa pertinence. En effet on est surpris par la découverte de la doctrine utilitariste multidimensionnelle de John Stuart Mill et de son étonnante actualité, dans le marasme économique actuel. Cette théorie utilitariste de John Stuart Mill dans sa version intégrale préconise que toute pensée, toute politique, tous programmes économiques doivent obéir à cinq exigences érigés en (relevés au rang de) principes fondamentaux, dont l'épicentre est la recherche du bonheur, quantitatif et du bien-être qualitatif, à la fois des agents économiques mais aussi des espèces de l'environnement écosystémique entier, dans des proportions raisonnables.

Cela se traduit dans notre langage par :

1. l'optimisation quantitative de l'utilité (et/ou le profit) de l'agent économique (qu'il soit consommateur, firme, voire même l'État dans un cadre néoclassique).
2. Ceci reste vrai pour une optimisation qualitative de l'agrément(c'est-à-dire le bien-être)... de l'agent.
3. l'optimisation du bonheur quantitatif environnemental pour toutes les espèces.

4. l'optimisation de la qualité environnementale (c'est-à-dire le bien-être) pour toutes les espèces.
5. tous ces programmes sont contraints par une optimisation dans des limites raisonnables, qui ne mettraient pas en péril aussi bien l'écosystème que la biodiversité.

Or le dernier cri d'alarme de Stiglitz est un message d'alerte en détresse stipulant que : “*Le reste du monde s’efforçait d’imiter l’Amérique, mais, s’il y parvenait pleinement, la planète n’y survivrait pas*” ([89], p 342), un message d'alerte en détresse, qui en dit long. Il n'en demeure pas moins que, presque toutes les propositions alternatives de Stiglitz sont (et le demeurent) de l'ordre technique de la gestion efficace et bien informée de l'économie nationale et internationale et non doctrinale qui donnerait à toute action économique d'un agent à la fois un sens et une substance éthique et politique.

L'adoption de l'utilitarisme écosystémique intégral originel de John Stuart Mill doté d'une dynamique structurelle de longue portée, donne à l'action économique sa dimension doctrinale, et constitue par la même une alternative plus que plausible pour une refonte intégrale de l'économie autour des axes tels :

1. la reconciliation de l'utilitarisme originel de John Stuart Mill (l'utilité et le profit, étant vus comme des objectifs finaux) avec le paradigme altruiste, écosystémique et raisonnable de développement durable, à la fois en tant qu'infrastructure mais aussi en tant que référentiel, à tout programme de calcul optimal quantitatif mais aussi qualitatif. En effet les quatre des cinq composantes principale de la pensée économique de John Stuart Mill (altruisme, écosystémisme, optimisation raisonnable et recherche de la qualité, dont seule l'utilité a été conservée), sont omises dans presque tous les modèles économiques standards alors qu'elles faisaient partie prenante du système Stuart Millien global.
2. l'agrégation des variables microéconomiques en macroéconomie et leurs cointégrées, doit s'opérer, pour la même période, sur la même échelle de fréquence (toutes journalières, hebdomadaires, mensuelles, trimestrielles ou annuelles).
3. La valeur absolue en puissance de tout rendement logarithmique des prix relatifs (des denrées ou de leurs actifs correspondants cotés en bourse, pour deux cotations différées successives) est un processus de mémoire longue. Ceci mettrait à défaut, a priori, la présupposée efficience des marchés et avec elle l'hypothèse de concurrence pure et parfaite puisque la présence de la mémoire longue rend ces rendements en valeur absolue autocorrélés.
4. Par agrégation, les taux de croissance, a priori, de tous les agrégats macroéconomiques suivent des lois similaires avec différentes mémoire longues sous-jacentes (ou multifractal).
5. Par agrégation, a priori, le taux de croissance, aussi bien du PIB que de la population d'un pays suit une loi de puissance similaire.
6. Hormis les relations comptables, toutes les relations standards qui lient les agrégats entre eux sont remplacés par des relations de cointegration de long-terme. Les agrégats en question sont

désormais cointégrés.

7. Les changements de régimes et des ruptures, en présence de mémoire longue, pour les variables économiques sont dus fort probablement soit à des faits endogènes comme les cycles latents des affaires et autres sources de non stationnarités, soit à des faits exogènes comme les chocs pétroliers, les guerres, etc...
8. Les agrégats (ou variables) de flux et les agrégats (ou variables) de stocks peuvent présenter en l'état ou transformés, à leur tour, de la mémoire longue, avec possibilité de changements de régimes stochastique ou déterministes (avec des points de rupture fixes).

La vérification de ces principes a priori, peut être appréhendée, soit directement sur des données réelles, soit par simulations, soit en dernier ressort, en se rapportant à des études déjà faites par d'autres chercheurs, établissant directement ou indirectement de tels principes. Certains faits macroéconomique rapportés par Stiglitz correspondent à certains faits stylisés rencontrés en finance. En effet *“Ces nouveaux résultats montraient que l'hypothèse de l'efficacité (ou l'efficience) des marchés n'avait aucune base scientifique. Certes les marchés fournissaient des incitations, mais leurs échecs étaient omniprésents et il y avait des écarts persistants entre bénéfices sociaux (ou publics) et bénéfices privés”* ([89], p 430). Tout porte à croire que cette persistance décrite par Stiglitz est le fait soit d'une mémoire longue soit d'un changement de régimes. Or les imperfections de l'information au sein d'un marché, selon lui fussent-elles *“... de petites asymétries d'information pouvaient avoir de trop gros effets”* ([89], p 431). Un tel effet rappelle, par son ampleur disproportionné, le comportement des processus chaotiques.

Dans leur diagnostic quantifié des différentes dimensions de la crise de 2008, Stiglitz-Sen-Fitoussi préconisent des recommandations (et des normes) assorties de tableaux de bord (ou boîtes à outils internationales centralisées) à mettre en place, des stratégies globales à mettre en œuvre pour garantir le développement durable tant espéré. Un des tableaux incorporé dans le rapport est issu des travaux de l'UNECE/OCDE/Eurostat¹²⁵ [24]. Il préconise d'établir des indicateurs (de mesurabilité de durabilité) de développement durable en bien-être par domaine (bien-être fondamental, bien-être économique, bien-être politique, bien-être social). Deux classes d'indicateurs : de stock et de flux sont ainsi suivies régulièrement pour chaque pays. La liste des stocks contient tout particulièrement : les réserves de ressources énergétiques, les réserves de ressources minérales, les stocks de ressources en bois, les stocks de ressources marines. La liste des flux contient l'épuisement des ressources énergétiques, l'épuisement des ressources minérales, l'épuisement des ressources en bois, l'épuisement des ressources marines. Or l'analyse de ces mêmes ressources fait de plus en plus appel aux approches de la mémoire longue et à ses modèles eu égard à ces nouvelles disciplines écosystémiques, comme en témoignent, des articles multiples (voir [57], [58], [72], [41], [68], [5]), mentionnés au premier chapitre de synthèse.

125. UNECE/OECD/Eurostat (2008)

Chapitre 3

An adaptive MIR estimator of the memory parameter and the goodness-of-fit test

Sommaire

3.1	Introduction	88
3.2	The multidimensional increment ratio statistic and its statistical applications	89
3.3	Adaptive versions of the estimator and goodness-of-fit test	92
3.4	Simulations and Monte-Carlo experiments	95
3.4.1	Application of the IR estimator and tests applied to generated data . . .	95
3.4.2	Comparison with other adaptive semiparametric estimator of the memory parameter	96
3.5	Proofs	99

Abstract : The increment ratio (IR) statistic was first defined and studied in Surgailis *et al.* (2008) for estimating the memory parameter either of a stationary or an increment stationary Gaussian process. Here three extensions are proposed in the case of stationary processes. First, a multidimensional central limit theorem is established for a vector composed by several IR statistics. Second, a goodness-of-fit χ^2 -type test can be deduced from this theorem. Finally, this theorem allows to construct adaptive versions of the estimator and the test which are studied in a general semiparametric frame. The adaptive estimator of the long-memory parameter is proved to follow an oracle property. Simulations attest of the interesting accuracies and robustness of the estimator and the test, even in the non Gaussian case.

Keywords : Long-memory Gaussian processes; goodness-of-fit test; estimation of the memory parameter; minimax adaptive estimator.

3.1 Introduction

After almost thirty years of intensive and numerous studies, the long-memory processes now form an important topic of the time series study (see for instance the book edited by Doukhan *et al* [29], 2003). The most famous long-memory stationary time series are the fractional Gaussian noises (fGn) with Hurst parameter H and FARIMA(p, d, q) processes. For both these time series, the spectral density f in 0 follows a power law : $f(\lambda) \sim C |\lambda|^{-2d}$ where $H = d + 1/2$ in the case of the fGn. In the case of a long memory process $d \in (0, 1/2)$ but a natural expansion to $d \in (-1/2, 0]$ (short memory) implied that d can be considered more generally as a memory parameter.

There are a lot of statistical results relative to the estimation of this memory parameter d . First and main results in this direction have been obtained for parametric models with the essential articles of Fox and Taqqu [34], (1986) and Dahlhaus [26], (1989) for Gaussian time series, Giraitis and Surgailis [36], (1990) for linear processes and Giraitis and Taqqu [37], (1999) for non linear functions of Gaussian processes.

However parametric estimators are not really robust and can induce no consistent estimations. Thus, the research is now rather focused on semiparametric estimators of the memory parameter. Different approaches were considered : the famous R/S statistic (see Hurst [55], 1951), the log-periodogram estimator (studied first by Geweke and Porter-Hudack [35], 1983, notably improved by Robinson [82], 1995a, and Moulines and Soulier [70], 2003), the local Whittle estimator (see Robinson [81], 1995b) or the wavelet based estimator (see Veitch *et al* [3], 2003, Moulines *et al* [71], 2007 or Bardet *et al* [10], 2008). All these estimators require the choice of an auxiliary parameter (frequency bandwidth, scales, etc.) but adaptive versions of these estimators are generally built for avoiding this choice. In a general semiparametric frame, Giraitis *et al* [40], (1997) obtained the asymptotic lower bound for the minimax risk in the estimation of d , expressed as a function of the second order parameter of the spectral density expansion around 0. Several adaptive semiparametric estimators are proved to follow an oracle property up to multiplicative logarithm term. But simulations (see for instance Bardet *et al* [9], 2003 or [10], 2008) show that the most accurate estimators are local Whittle, global log-periodogram and wavelet based estimators.

In this paper, we consider the IR (Increment Ratio) estimator of a long-memory parameter (see its definition in the next section) for Gaussian time series recently introduced in Surgailis *et al.* [92], (2008) and we propose three extensions. First, a multivariate central limit theorem is established for a vector of IR statistics with different “windows” (see Section 3.2) and this induces to consider a pseudo-generalized least squares estimator of the parameter d . Second, this multivariate result allows us to define an adaptive estimator of the memory parameter d based on IR statistics : an “optimal” window is automatically computed (see Section 3.3). This notably improves the results of Surgailis *et al.* [92], (2008) in which the choice of m is either theoretical (and cannot be applied to data) or guided by empirical rules without justifications. Third, an adaptive goodness-of-fit test is deduced and its convergence to a chi-square distribution is established (see Section 3.3).

In Section 3.4, several Monte Carlo simulations are realized for optimizing the adaptive estimator and exhibiting the theoretical results. Then some numerical comparisons are made with the 3 semiparametric estimators previously mentioned (local Whittle, global log-periodogram and wavelet based estimators) and the results are even better than the theory seems to indicate : as well in terms of convergence rate than in terms of robustness (notably in case of trend or seasonal component), the adaptive IR estimator and goodness-of-fit test provide efficient results. Finally, all the proofs are grouped in Section 3.5.

3.2 The multidimensional increment ratio statistic and its statistical applications

Let $X = (X_k)_{k \in \mathbb{N}}$ be a Gaussian time series satisfying the following Assumption $S(d, \beta)$:

Assumption $S(d, \beta)$: *There exist $\varepsilon > 0$, $c_0 > 0$, $c'_0 > 0$ and $c_1 \in \mathbb{R}$ such that $X = (X_t)_{t \in \mathbb{Z}}$ is a stationary Gaussian time series having a spectral density f satisfying for all $\lambda \in (-\pi, 0) \cup (0, \pi)$*

$$f(\lambda) = c_0 |\lambda|^{-2d} + c_1 |\lambda|^{-2d+\beta} + O(|\lambda|^{-2d+\beta+\varepsilon}) \quad \text{and} \quad |f'(\lambda)| \leq c'_0 |\lambda|^{-2d-1}. \quad (3.1)$$

Remark 1. *Note that here we only consider the case of stationary processes. However, as it was already done in Surgailis et al.[92], (2008), it could be possible, mutatis mutandis, to extend our results to the case of processes having stationary increments.*

Let $(X_1, \dots, X_N)$ be a path of X . For $m \in \mathbb{N}^*$, define the random variable $IR_N(m)$ such as

$$IR_N(m) := \frac{1}{N-3m} \sum_{k=0}^{N-3m-1} \frac{|(\sum_{t=k+1}^{k+m} X_{t+m} - \sum_{t=k+1}^{k+m} X_t) + (\sum_{t=k+m+1}^{k+2m} X_{t+m} - \sum_{t=k+m+1}^{k+2m} X_t)|}{|(\sum_{t=k+1}^{k+m} X_{t+m} - \sum_{t=k+1}^{k+m} X_t)| + |(\sum_{t=k+m+1}^{k+2m} X_{t+m} - \sum_{t=k+m+1}^{k+2m} X_t)|}.$$

From Surgailis et al. [92], (2008), with m such that $N/m \rightarrow \infty$ and $m \rightarrow \infty$,

$$\sqrt{\frac{N}{m}} (IR_N(m) - \mathbb{E}IR_N(m)) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(d)),$$

where

$$\sigma^2(d) := 2 \int_0^\infty \text{Cov} \left(\frac{|Z_d(0) + Z_d(1)|}{|Z_d(0)| + |Z_d(1)|}, \frac{|Z_d(\tau) + Z_d(\tau+1)|}{|Z_d(\tau)| + |Z_d(\tau+1)|} \right) d\tau \quad (3.2)$$

$$\text{and } Z_d(\tau) := \frac{1}{\sqrt{|4^{d+0.5} - 4|}} (B_{d+0.5}(\tau+2) - 2B_{d+0.5}(\tau+1) + B_{d+0.5}(\tau)) \quad (3.3)$$

with B_H a standardized fractional Brownian motion (FBM) with Hurst parameter $H \in (0, 1)$.

Remark 2. *This convergence was obtained for Gaussian processes in Surgailis et al.[92], (2008), but there also exist results concerning a modified IR statistic applied to stable processes (see Vaiciulis[94],*

2009) with a different kind of limit theorem. We may suspect that it is also possible to extend the previous central limit theorem to long memory linear processes (since a Donsker type theorem with FBM as limit was proved for long memory linear processes, see for instance Ho and Hsing [53], 1997) but such a result requires to prove a non obvious central limit theorem for a functional of a multidimensional linear process. Surgailis et al. [92], (2008) also considered the case of i.i.d.r.v. in the domain of attraction of a stable law with index $0 < \alpha < 2$ and skewness parameter $-1 \leq \beta \leq 1$ and concluded that $IR_N(m)$ converges to almost the same limit. Finally, in Bardet and Surgailis [7], (2011) a “continuous” version of the IR statistic is considered for several kind of continuous time processes (Gaussian processes, diffusions and Lévy processes).

Now, instead of this univariate IR statistic, define a multivariate IR statistic as follows : let $m_j = j m$, $j = 1, \dots, p$ with $2 \leq p[N/m] - 4$, and define the random vector $(IR_N(j m))_{1 \leq j \leq p}$. Thus, p is the number of considered window lengths of this multivariate statistic. In the sequel we naturally extend the results obtained for $m \in \mathbb{N}^*$ to $m \in (0, \infty)$ by the convention : $(IR_N(j m))_{1 \leq j \leq p} = (IR_N(j [m]))_{1 \leq j \leq p}$ (which change nothing to the asymptotic results).

We can establish a multidimensional central limit theorem satisfied by $(IR_N(j m))_{1 \leq j \leq p}$.

Property 1. Assume that Assumption $S(d, \beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Then

$$\sqrt{\frac{N}{m}} \left(IR_N(j m) - \mathbb{E}[IR_N(j m)] \right)_{1 \leq j \leq p} \xrightarrow[N/m \wedge m \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_p(d)) \quad (3.4)$$

with $\Gamma_p(d) = (\sigma_{i,j}(d))_{1 \leq i,j \leq p}$ where for $t \in \mathbb{R}$

$$\begin{aligned} \sigma_{i,j}(d) &:= \int_{-\infty}^{\infty} \text{Cov} \left(\frac{|Z_d^{(i)}(0) + Z_d^{(i)}(i)|}{|Z_d^{(i)}(0)| + |Z_d^{(i)}(i)|}, \frac{|Z_d^{(j)}(\tau) + Z_d^{(j)}(\tau + j)|}{|Z_d^{(j)}(\tau)| + |Z_d^{(j)}(\tau + j)|} \right) d\tau \\ \text{and } Z_d^{(j)}(\tau) &= \frac{1}{\sqrt{|4^{d+0.5} - 4|}} (B_{d+0.5}(\tau + 2j) - 2B_{d+0.5}(\tau + j) + B_{d+0.5}(\tau)). \end{aligned} \quad (3.5)$$

The proof of this property as well as all the other proofs are given in Appendix. Moreover we will assume in the sequel that $\Gamma_p(d)$ is a definite positive matrix for all $d \in (-0.5, 0.5)$.

Remark 3. Numerical experiments show that $\Gamma_p(d)$ is always invertible for several values of p and d ($d \in (-0.5, 0.5)$).

Remark 4. Note that Assumption $S(d, \beta)$ is a little stronger than the conditions required in Surgailis et al. [92], (2008) where f is supposed to satisfy $f(\lambda) = c_0 |\lambda|^{-2d} + O(|\lambda|^{-2d+\beta})$ and $|f'(\lambda)| \leq c'_0 |\lambda|^{-2d-1}$. Note that Property 1 and following Theorem 1 and Proposition 1 are as well checked under these assumptions of Surgailis et al. (2008) even if $\beta \geq 2d + 1$ (a case which is not considered in their Theorem 2.4). However our automatic procedure for choosing an adaptive scale $\tilde{m}_N$ requires to specify the second order of the expansion of f and we prefer to already give results under such assumption.

As in Surgailis *et al.*[92], (2008), for $r \in (-1, 1)$, define the function $\Lambda(r)$ by

$$\Lambda(r) := \frac{2}{\pi} \arctan \sqrt{\frac{1+r}{1-r}} + \frac{1}{\pi} \sqrt{\frac{1+r}{1-r}} \log\left(\frac{2}{1+r}\right). \quad (3.6)$$

and for $d \in (-0.5, 1.5)$ let

$$\Lambda_0(d) := \Lambda(\rho(d)) \quad \text{where} \quad \rho(d) := \frac{4^{d+1.5} - 9^{d+0.5} - 7}{2(4 - 4^{d+0.5})}. \quad (3.7)$$

The function $d \in (-0.5, 1.5) \rightarrow \Lambda_0(d)$ is a $\mathcal{C}^\infty$ increasing function. Now, Property 2 (see in Section 3.5) provides the asymptotic behavior of $\mathbb{E}[IR(m)]$ when $m \rightarrow \infty$, which is $\mathbb{E}[IR(m)] \sim \Lambda_0(d) + Cm^{-\beta}$ if $\beta < 2d + 1$, $\mathbb{E}[IR(m)] \sim \Lambda_0(d) + Cm^{-\beta} \log m$ if $\beta = 2d + 1$ and $\mathbb{E}[IR(m)] \sim \Lambda_0(d) + O(m^{-(2d+1)})$ if $\beta > 2d + 1$ (C is a non vanishing real number depending on d and β). Therefore by choosing m and N such as $(\sqrt{N/m})m^{-\beta} \rightarrow 0$, $(\sqrt{N/m})m^{-\beta} \log m \rightarrow 0$ and $(\sqrt{N/m})m^{-(2\beta+1)} \rightarrow 0$ (respectively) when $m, N \rightarrow \infty$, the term $\mathbb{E}[IR(jm)]$ can be replaced by $\Lambda_0(d)$ in Property 1. Then, using the Delta-method with function $(x_i)_{1 \leq i \leq p} \mapsto (\Lambda_0^{-1}(x_i))_{1 \leq i \leq p}$, we obtain the following theorem.

Theorem 1. *Let $\hat{d}_N(jm) := \Lambda_0^{-1}(IR_N(jm))$ for $1 \leq j \leq p$. Assume that Assumption $S(d, \beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Then if $m \sim C N^\alpha$ with $C > 0$ and $(1+2\beta)^{-1} \vee (4d+3)^{-1} < \alpha < 1$ then*

$$\sqrt{\frac{N}{m}} \left(\hat{d}_N(jm) - d \right)_{1 \leq j \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, (\Lambda'_0(d))^{-2} \Gamma_p(d)\right). \quad (3.8)$$

Remark 5. *If $\beta < 2d+1$, the estimator $\hat{d}_N(m)$ is a semiparametric estimator of d and its asymptotic mean square error can be minimized with an appropriate sequence (m_N) reaching the well-known minimax rate of convergence for memory parameter d in this semiparametric setting (see for instance Giraitis *et al.* [40], 1997). Indeed, under Assumption $S(d, \beta)$ with $d \in (-0.5, 0.5)$ and $\beta > 0$ and if $m_N = \lfloor N^{1/(1+2\beta)} \rfloor$, then the estimator $\hat{d}_N(m_N)$ is rate optimal in the minimax sense, i.e.*

$$\limsup_{N \rightarrow \infty} \sup_{d \in (-0.5, 0.5)} \sup_{f \in S(d, \beta)} N^{\frac{2\beta}{1+2\beta}} \cdot \mathbb{E}[(\hat{d}_N(m_N) - d)^2] < \infty.$$

From the multidimensional CLT (3.8) a pseudo-generalized least squares estimation (LSE) of d is possible by defining the following matrix :

$$\hat{\Sigma}_N(m) := (\Lambda'_0(\hat{d}_N(m)))^{-2} \Gamma_p(\hat{d}_N(m)). \quad (3.9)$$

Since the function $d \in (-0.5, 1.5) \mapsto \sigma(d)/\Lambda'(d)$ is $\mathcal{C}^\infty$ it is obvious that under assumptions of Theorem 1 then

$$\hat{\Sigma}_N(m) \xrightarrow[N \rightarrow \infty]{\mathcal{P}} (\Lambda'_0(d))^{-2} \Gamma_p(d).$$

Then with the vector $J_p := (1)_{1 \leq j \leq p}$ and denoting J_p' its transpose, the pseudo-generalized LSE of d is

$$\tilde{d}_N(m) := (J_p' (\hat{\Sigma}_N(m))^{-1} J_p)^{-1} J_p' (\hat{\Sigma}_N(m))^{-1} (\hat{d}_N(m_i))_{1 \leq i \leq p}$$

It is well known (Gauss-Markov Theorem) that the Mean Square Error (MSE) of $\tilde{d}_N(m)$ is smaller or equal than all the MSEs of $\hat{d}_N(jm)$, $j = 1, \dots, p$. Hence, we obtain under the assumptions of MCLT Theorem 3.8 :

$$\sqrt{\frac{N}{m}} (\tilde{d}_N(m) - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, \Lambda_0'(d)^{-2} (J_p' \Gamma_p^{-1}(d) J_p)^{-1}\right), \quad (3.10)$$

and $\Lambda_0'(d)^{-2} (J_p' \Gamma_p^{-1}(d) J_p)^{-1} \leq \Lambda_0'(d)^{-2} \sigma^2(d)$.

Now, consider the following test problem : for $(X_1, \dots, X_n)$ a path of X a Gaussian time series, choose between

- H_0 : the spectral density of X satisfies Assumption $S(d, \beta)$ with $-0.5 < d < 0.5$ and $\beta > 0$;
- H_1 : the spectral density of X does not satisfy such a behavior.

We deduce from the multidimensional CLT (3.8) a χ^2 -type goodness-of-fit test statistic defined by :

$$\hat{T}_N(m) := \frac{N}{m} (\tilde{d}_N(m) - \hat{d}_N(jm))'_{1 \leq j \leq p} (\hat{\Sigma}_N(m))^{-1} (\tilde{d}_N(m) - \hat{d}_N(jm))_{1 \leq j \leq p}.$$

Then the following limit theorem can be deduced from Theorem 1.

Proposition 1. *Under the assumptions of Theorem 1 then*

$$\hat{T}_N(m) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p-1).$$

3.3 Adaptive versions of the estimator and goodness-of-fit test

Theorem 1 and Proposition 1 are interesting but they require the knowledge of β to be used (and therefore an appropriated choice of m). We now suggest a procedure (see also Bardet *et al.* [10], 2008) for obtaining a data-driven selection of an optimal sequence (m_N) . For $d \in (-0.5, 1.5)$ and $\alpha \in (0, 1)$, define

$$Q_N(\alpha, d) := (\hat{d}_N(j N^\alpha) - d)'_{1 \leq j \leq p} (\hat{\Sigma}_N(N^\alpha))^{-1} (\hat{d}_N(j N^\alpha) - d)_{1 \leq j \leq p}. \quad (3.11)$$

Note that by the previous convention, $\hat{d}_N(j N^\alpha) = \hat{d}_N(j [N^\alpha])$ and $\tilde{d}_N(N^\alpha) = \tilde{d}_N([N^\alpha])$. Thus $Q_N(\alpha, d)$ corresponds to the sum of the pseudo-generalized squared distance. From previous computations, it is obvious that for a fixed $\alpha \in (0, 1)$, Q is minimized by $\tilde{d}_N(N^\alpha)$ and therefore for $0 < \alpha < 1$ define

$$\hat{Q}_N(\alpha) := Q_N(\alpha, \tilde{d}_N(N^\alpha)).$$

It remains to minimize $\widehat{Q}_N(\alpha)$ on $(0, 1)$. However, since $\widehat{\alpha}_N$ has to be obtained from numerical computations, the interval $(0, 1)$ can be discretized as follows,

$$\widehat{\alpha}_N \in \mathcal{A}_N = \left\{ \frac{2}{\log N}, \frac{3}{\log N}, \dots, \frac{\log[N/p]}{\log N} \right\}.$$

Hence, if $\alpha \in \mathcal{A}_N$, it exists $k \in \{2, 3, \dots, \log[N/p]\}$ such that $k = \alpha \log N$. Consequently, define $\widehat{\alpha}_N$ by

$$\widehat{Q}_N(\widehat{\alpha}_N) := \min_{\alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha).$$

From the central limit theorem (3.8) one deduces the following proposition.

Proposition 2. *Assume that Assumption $S(d, \beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Moreover, if $\beta > 2d + 1$, suppose that c_0, c_1, c_2, d, β and ε are such that Condition (3.32) or (3.33) holds. Then,*

$$\widehat{\alpha}_N \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \alpha^* = \frac{1}{(1 + 2\beta) \wedge (4d + 3)}.$$

Remark 6. *The choice of the set of discretization $\mathcal{A}_N$ is implied by our proof of convergence of $\widehat{\alpha}_N$ to α^* . If the interval $(0, 1)$ is stepped in N^c points, with $c > 0$, the used proof cannot attest this convergence. However $\log N$ may be replaced in the previous expression of $\mathcal{A}_N$ by any negligible function of N compared to functions N^c with $c > 0$ (for instance, $(\log N)^a$ or a $\log N$ can be used).*

Remark 7. *The reference to Condition (3.32) or (3.33) is necessary because our proof of the convergence of $\widehat{\alpha}_N$ to α^* requires to know the exact convergence rate of $\mathbb{E}[IR_N(N^\alpha)] - \Lambda_0(d)$ when $\alpha < \alpha^*$. When $\beta \leq 2d + 1$, since we replaced the conditions on the spectral density of Surgailis et al. (2008) by a second order condition (Assumption $S(d, \beta)$), this convergence rate can be obtained by computations (see Property 2). But if $\beta > 2d + 1$, we can only obtain $\mathbb{E}[IR_N(N^\alpha)] - \Lambda_0(d) = O(m^{-2d-1})$ under Assumption $S(d, \beta)$: the convergence rate could be slower than m^{-2d-1} and then $\widehat{\alpha}_N$ could converge to $\alpha' < \alpha^*$ (from the proof of Proposition 2). Condition (3.32) and (3.33), which are not very strong, allow to obtain a first order bound for $\mathbb{E}[IR_N(N^\alpha)] - \Lambda_0(d)$ (see Property 3) and hence to prove $\widehat{\alpha}_N \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \alpha^*$.*

From a straightforward application of the proof of Proposition 2, the asymptotic behavior of $\widehat{\alpha}_N$ can be specified, that is,

$$\Pr \left(\frac{N^{\alpha^*}}{(\log N)^\lambda} \leq N^{\widehat{\alpha}_N} \leq N^{\alpha^*} \cdot (\log N)^\mu \right) \xrightarrow[N \rightarrow \infty]{} 1, \quad (3.12)$$

for all positive real numbers λ and μ such that $\lambda > \frac{2\alpha^*}{(p-2)(1-\alpha^*)}$ and $\mu > \frac{12}{p-2}$. Consequently, the selected window $\widehat{m}_N = N^{\widehat{\alpha}_N}$ asymptotically grows as N^{α^*} up to a logarithm factor.

Finally, Proposition 2 can be used to define an adaptive estimator of d . First, define the straightforward estimator $\widetilde{d}_N(N^{\widehat{\alpha}_N})$, which should minimize the mean square error using $\widehat{\alpha}_N$. However, the

estimator $\tilde{d}_N(N^{\hat{\alpha}_N})$ does not satisfy a CLT since $\Pr(\hat{\alpha}_N \leq \alpha^*) > 0$ and therefore it cannot be asserted that $\mathbb{E}(\sqrt{N/N^{\hat{\alpha}_N}}(\tilde{d}_N(N^{\hat{\alpha}_N}) - d)) = 0$. To establish a CLT satisfied by an adaptive estimator of d , a (few) shifted sequence of $\hat{\alpha}_N$, so called $\tilde{\alpha}_N$, has to be considered to ensure $\Pr(\tilde{\alpha}_N \leq \alpha^*) \xrightarrow{N \rightarrow \infty} 0$.

Hence, consider the adaptive scale sequence $(\tilde{m}_N)$ such as

$$\tilde{m}_N := N^{\tilde{\alpha}_N} \quad \text{with} \quad \tilde{\alpha}_N := \hat{\alpha}_N + \frac{6\hat{\alpha}_N}{(p-2)(1-\hat{\alpha}_N)} \cdot \frac{\log \log N}{\log N}.$$

and the estimator

$$\tilde{d}_N^{(IR)} := \tilde{d}_N(\tilde{m}_N) = \tilde{d}_N(N^{\tilde{\alpha}_N}).$$

The following theorem provides the asymptotic behavior of the estimator $\tilde{d}_N^{(IR)}$.

Theorem 2. *Under assumptions of Proposition 2,*

$$\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}(\tilde{d}_N^{(IR)} - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0; \Lambda'_0(d)^{-2} (J'_p \Gamma_p^{-1}(d) J_p)^{-1}\right). \quad (3.13)$$

Moreover, if $\beta \leq 2d+1$, $\forall \rho > \frac{2(1+3\beta)}{(p-2)\beta}$, $\frac{N^{\frac{\beta}{1+2\beta}}}{(\log N)^\rho} \cdot |\tilde{d}_N^{(IR)} - d| \xrightarrow[N \rightarrow \infty]{\mathcal{P}} 0$.

Remark 8. *When $\beta \leq 2d+1$, the adaptive estimator $\tilde{d}_N^{(IR)}$ converges to d with a rate of convergence rate equal to the minimax rate of convergence $N^{\frac{\beta}{1+2\beta}}$ up to a logarithm factor (this result being classical within this semiparametric framework). Thus there exists $\ell < 0$ such that*

$$N^{\frac{2\beta}{1+2\beta}} (\log N)^\ell \mathbb{E}(\tilde{d}_N^{(IR)} - d)^2 < \infty.$$

Therefore $\tilde{d}_N^{(IR)}$ satisfies an oracle property for the considered semiparametric model.

If $\beta > 2d+1$, the estimator is not rate optimal. However, simulations (see Section 3.4) will show that even if $\beta > 2d+1$, the rate of convergence of $\tilde{d}_N^{(IR)}$ can be better than the one of the best known rate optimal estimators (local Whittle or global log-periodogram estimators).

Moreover an adaptive version of the previous goodness-of-fit test can be derived. Thus define

$$\tilde{T}_N^{(IR)} := \hat{T}_N(N^{\tilde{\alpha}_N}). \quad (3.14)$$

Then we have the following proposition.

Proposition 3. *Under the assumptions of Proposition 2 then,*

$$\tilde{T}_N^{(IR)} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p-1).$$

3.4 Simulations and Monte-Carlo experiments

In the sequel, the numerical properties (consistency, robustness, choice of the parameter p) of $\tilde{d}_N^{(IR)}$ are investigated. Then the simulation results of $\tilde{d}_N^{(IR)}$ are compared to those obtained with the best known semiparametric long-memory estimators.

Remark 9. *Note that all the softwares (in Matlab language) used in this section are available with a free access on <http://samm.univ-paris1.fr/~Jean-Marc-Bardet>.*

To begin with, the simulation conditions have to be specified. The results are obtained from 100 generated independent samples of each process belonging to the following "benchmark". The concrete procedures of generation of these processes are obtained from the circulant matrix method, as detailed in Doukhan *et al.* (2003). The simulations are realized for different values of d , N and processes which satisfy Assumption $S(d, \beta)$:

1. the fractional Gaussian noise (fGn) of parameter $H = d + 1/2$ (for $-0.5 < d < 0.5$) and $\sigma^2 = 1$. Such a process is such that Assumption $S(d, 2)$ holds ;
2. the FARIMA $[p, d, q]$ process with parameter d such that $d \in (-0.5, 0.5)$, the innovation variance σ^2 satisfying $\sigma^2 = 1$ and $p, q \in \mathbb{N}$. A FARIMA $[p, d, q]$ process is such that Assumption $S(d, 2)$ holds ;
3. the Gaussian stationary process $X^{(d, \beta)}$, such as its spectral density is

$$f_3(\lambda) = \frac{1}{|\lambda|^{2d}}(1 + |\lambda|^\beta) \quad \text{for } \lambda \in [-\pi, 0) \cup (0, \pi], \quad (3.15)$$

with $d \in (-0.5, 0.5)$ and $\beta \in (0, \infty)$. Therefore the spectral density f_3 is such as Assumption $S(d, \beta)$ holds.

A "benchmark" which will be considered in the sequel consists of the following particular cases of these processes for $d = -0.4, -0.2, 0, 0.2, 0.4$:

- fGn processes with parameters $H = d + 1/2$;
- FARIMA $[0, d, 0]$ processes with standard Gaussian innovations ;
- FARIMA $[1, d, 1]$ processes with standard Gaussian innovations and AR coefficient $\phi = -0.3$ and MA coefficient $\theta = 0.7$;
- $X^{(d, \beta)}$ Gaussian processes with $\beta = 1$.

3.4.1 Application of the IR estimator and tests applied to generated data

Choice of the parameter p : This parameter is important to estimate the "beginning" of the linear part of the graph drawn by points $(i, IR(im))_i$. On the one hand, if p is a too small a number (for instance $p = 3$), another small linear part of this graph (even before the "true" beginning N^{α^*}) may be chosen. On the other hand, if p is a too large a number (for instance $p = 50$ for $N = 1000$),

the estimator $\tilde{\alpha}_N$ will certainly satisfy $\tilde{\alpha}_N < \alpha^*$ since it will not be possible to consider p different windows larger than N^{α^*} . Moreover, it is possible that a "good" choice of p depends on the "flatness" of the spectral density f , *i.e.* on β . We have proceeded to simulations for several values of p (and N and d). Only $\sqrt{MSE}$ of estimators are presented. The results are specified in Table 3.1.

Conclusions from Table 3.1 : It is clear that $\tilde{d}_N^{(IR)}$ converges to d for the four processes, the faster for fGn and FARIMA(0, d , 0). The optimal choice of p seems to depend on N for the four processes : $\hat{p} = 10$ for $N = 10^3$, $\hat{p} = 15$ for $N = 10^4$ and $\hat{p} \in [15, 20]$ for $N = 10^5$. The flatness of the spectral density of the process does not seem to have any influence, as well as the value of d (result obtained in the detailed simulations). We will adopt in the sequel the choice $\hat{p} = [1.5 \log(N)]$ reflecting these results. On the contrary to the choice of m , this choice of p only depends on N and even if the adaptive scale $\tilde{m}_N$ depends on p its value does not change a lot when $p \in \{10, \dots, 20\}$ for $10^3 \leq N \leq 10^5$.

Concerning the adaptive choice of m , the main point to be remarked is that the smoother the spectral density the smaller m ; thus $\tilde{m}_N$ is smaller for a trajectory of a fGn or a FARIMA(0, d , 0) than for a trajectory of a FARIMA(1, d , 1) or $X^{(d,1)}$. The choice of p does not appear to significantly affect the value of $\tilde{m}_N$. More detailed results show that the larger d included in $(-0.5, 0.5)$ the smaller $\tilde{m}_N$: for instance, for the fGn, $N = 10^4$ and $p = 15$, the mean of $\tilde{m}_N$ is respectively equal to 23.9, 8.3, 4.5, 4.2 and 3.8 for d , respectively, equal to -0.4 , -0.2 , 0 , 0.2 and 0.4 . This phenomena can be deduced from the theoretical study since $\alpha^* = (4d + 3)^{-1}$ in this case and therefore $\tilde{m}_N$ almost grows as $N^{(4d+3)^{-1}}$.

Finally, concerning the goodness-of-fit test, we remark that it is too conservative for $p = 5$ or 10 but close to the expected results for $p = 15$ and 20 , especially for FARIMA(1, d , 1) or $X^{(d,1)}$.

Asymptotic distributions of the estimator and test : Figure 3.1 provides the density estimations of $\tilde{d}_N^{(IR)}$ and $\tilde{T}_N^{(IR)}$ for 100 independent samples of FGN processes with $d = 0.2$ with $N = 10^4$ for $p = 15$. The goodness-of-fit to the theoretical asymptotic distributions (respectively Gaussian and chi-square) is satisfying.

3.4.2 Comparison with other adaptive semiparametric estimator of the memory parameter

Consistency of semiparametric estimators : Here we consider the previous "benchmark" and apply the estimator $\tilde{d}_N^{(IR)}$ and 3 other semiparametric estimators of d known for their accuracies are considered :

- $\hat{d}_{MS}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier [70](2003),

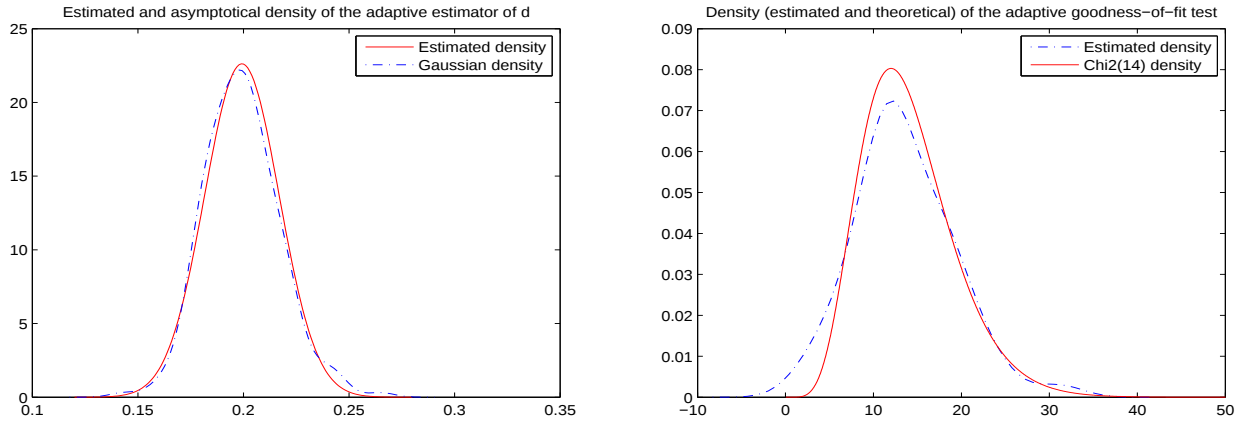


FIGURE 3.1 – Density estimations and corresponding theoretical densities of $\tilde{d}_N^{(IR)}$ and $\tilde{T}_N^{(IR)}$ for 100 samples of fGn with $d = 0.2$ with $N = 10^4$ and $p = 15$.

also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;

- $\hat{d}_R$ is the local Whittle estimator introduced by Robinson [81] (1995b). The trimming parameter is $m = N/30$. Other trimming parameter choices exist in the literature as $m = N^{\frac{4}{5}}$ or $m = N^{0.65}$;
- $\hat{d}_W$ is an adaptive wavelet based estimator introduced in Bardet *et al.* [10], (2008) using a Lemarie-Meyer type wavelet (another similar choice could be the adaptive wavelet estimator introduced in Veitch *et al.* [3], 2003, using a Daubechies wavelet, but its robustness property are quite less interesting).
- $\tilde{d}_N^{(IR)}$ defined previously with $p = [1.5 * \log(N)]$.
- $\hat{d}_N(10)$ and $\hat{d}_N(30)$ which are the (univariate) IR estimator with $m = 10$ and $m = 30$ respectively, considered in Surgailis *et al.* [92], (2008).

Simulation results are reported in Table 3.2.

Conclusions from Table 3.2 : The adaptive IR estimator $\tilde{d}_N^{(IR)}$ numerically shows a convincing convergence rate with respect to the other estimators.

The estimators $\hat{d}_N(10)$ and $\hat{d}_N(30)$ are clearly the worst estimators of d . This can be explained by two facts.

1. The numerical expression of the matrix $\hat{\Sigma}_N(m)$ is almost a diagonal matrix, and therefore a least squares regression using several window lengths provides better estimations than an estimator using only one window length.
2. $\hat{d}_N(10)$ and $\hat{d}_N(30)$ use a fixed window length ($m = 10$ and $m = 30$) for any process and N while we know that $m \simeq N^{\alpha^*}$ is the optimal choice which is approximated by $\tilde{m}_N$.

Both the “spectral” estimators $\hat{d}_R$ and $\hat{d}_{MS}$ provide more stable results that do not depend very much on d and the process, while the wavelet based estimator $\hat{d}_W$ and $\tilde{d}_N^{(IR)}$ are more sensible to the flatness of the spectral density. But, especially for “smooth processes” (fGn and FARIMA(0, d , 0)), $\tilde{d}_N^{(IR)}$ is a very accurate semiparametric estimator and is globally more efficient than the other estimators.

Robustness of the different semiparametric estimators : To conclude with the numerical properties of the estimators, five different processes not satisfying Assumption $S(d, \beta)$ are considered :

- a FARIMA(0, d , 0) process with innovations satisfying a uniform law ;
- a FARIMA(0, d , 0) process with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x) = 1 - \frac{1}{2} \frac{1}{1+x^2}$ for $x \geq 0$ and $F(x) = \frac{1}{2} \frac{1}{1+x^2}$ for $x \leq 0$ (and therefore $\mathbb{E}|X_i|^2 = \infty$ but $\mathbb{E}|X_i| < \infty$) ;
- a FARIMA(0, d , 0) process with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x) = 1 - \frac{1}{2} \frac{1}{1+|x|^{3/2}}$ for $x \geq 0$ and $F(x) = \frac{1}{2} \frac{1}{1+|x|^{3/2}}$ for $x \leq 0$ (and therefore $\mathbb{E}|X_i|^2 = \infty$ but $\mathbb{E}|X_i| < \infty$) ;
- a Gaussian stationary process with a spectral density $f(\lambda) = ||\lambda| - \pi/2|^{-2d}$ for all $\lambda \in [-\pi, \pi] \setminus \{-\pi/2, \pi/2\}$: this is a GARMA(0, d , 0) process. The local behavior of f in 0 is $f(|\lambda|) \sim (\pi/2)^{-2d} |\lambda|^{-2d}$ with $d = 0$, but the smoothness condition for f in Assumption $S(0, \beta)$ is not satisfied.
- a trended fGn with parameter $H = d + 0.5$ and an additive linear trend ;
- a fGn ($H = d + 0.5$) with an additive linear trend and an additive sinusoidal seasonal component of period $T = 12$.

The results of these simulations are given in Table 3.3.

Conclusions from Table 3.3 : The main advantage of $\hat{d}_W$ and $\tilde{d}_N^{(IR)}$ with respect to $\hat{d}_{MS}$ and $\hat{d}_R$ is exhibited in this table : they are robust with respect to smooth trends, especially in the case of long memory processes ($d > 0$). This has already been observed in Bruzaite and Vaiciulis [20], (2008) for IR statistic (and even for certain discontinuous trends). Both those estimators are also robust with respect to seasonal component and this robustness would have been improved if we had chosen m (or scales) as a multiple of the period (which is generally known).

The second good surprise of these simulations is that the adaptive IR estimator $\tilde{d}_N^{(IR)}$ is also consistent for non Gaussian distributions even if the function Λ in (3.6) and therefore all our results are typically obtained for Gaussian distributions. The case of finite-variance processes is not surprising (see Remark 2). But this is more surprising for infinite variance processes. A first explanation of this was given in Surgailis *et al.* [92], (2008) in the case of i.i.d.r.v. in the domain of attraction of a stable law with index $0 < \alpha < 2$ and skewness parameter $-1 \leq \beta \leq 1$: they

concluded that $IR_N(m)$ converges to almost the same limit. The extension to α -stable linear processes of this first explanation should require technical developments but the expression of the IR statistic (which is bounded in $[0, 1]$ for any processes) could allow to apply it to infinite variance processes. Note that the other semiparametric estimators are also consistent in such frames with faster convergence rates notably for the local Whittle estimator.

3.5 Proofs

Proof of Property 1. We proceed in two steps.

Step 1 : First, we compute the limit of $\frac{N}{m} \text{Cov}(IR_N(jm), IR_N(j'm))$ when N , m and $N/m \rightarrow \infty$. As in Surgailis *et al* [92], (2008), define also for all $j = 1, \dots, p$ and $k = 1, \dots, N - 3m_j$ (with $m_j = jm$) :

$$Y_{m_j}(k) := \frac{1}{V_{m_j}} \sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t) \quad , \quad \text{with} \quad V_{m_j}^2 := \mathbb{E} \left[\left(\sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t) \right)^2 \right] \quad (3.16)$$

$$\text{and} \quad \eta_{m_j}(k) := \frac{|Y_{m_j}(k) + Y_{m_j}(k + m_j)|}{|Y_{m_j}(k)| + |Y_{m_j}(k + m_j)|}. \quad (3.17)$$

Note that $Y_{m_j}(k) \sim \mathcal{N}(0, 1)$ for any k and j and

$$IR_N(m_j) = \frac{1}{N - 3m_j} \sum_{k=0}^{N-3m_j-1} \eta_{m_j}(k) \quad \text{for all } j = 1, \dots, p.$$

$$\begin{aligned} \text{Cov}(IR_N(m_j), IR_N(m_{j'})) &= \frac{1}{N - 3m_j} \frac{1}{N - 3m_{j'}} \sum_{k=0}^{N-3m_j-1} \sum_{k'=0}^{N-3m_{j'}-1} \text{Cov}(\eta_{m_j}(k), \eta_{m_{j'}}(k')) \\ &= \frac{1}{(\frac{N}{m_j} - 3)(\frac{N}{m_{j'}} - 3)} \int_{\tau=0}^{\frac{N-1}{m_j}-3} \int_{\tau'=0}^{\frac{N-1}{m_{j'}}-3} \text{Cov}(\eta_{m_j}([m_j\tau]), \eta_{m_{j'}}([m_{j'}\tau'])) d\tau d\tau'. \end{aligned} \quad (3.19)$$

Now according to (5.20) of the same article, with $\rightarrow_{FDD}$ denoting the finite distribution convergence when $m \rightarrow \infty$,

$$Y_m([m\tau]) \rightarrow_{FDD} Z_d(\tau)$$

where Z_d is defined in (3.3). Now

$$\begin{aligned} Y_{jm}(k) &= \frac{1}{V_{m_j}} \sum_{t=1}^{jm} X_{t+jm+1} - \sum_{t=1}^{jm} X_{t+1} X_t \\ &= \frac{1}{V_{m_j}} \sum_{i=-(j-1)}^{j-1} (j - |i|) V_m Y_m(t + (j + i - 1)m). \end{aligned}$$

But $V_m^2 \sim c_0 V(d)m^{2d+1}$ when $m \rightarrow \infty$ (see (2.20) in Surgailis *et al*, 2008). Therefore we obtain $Y_{jm}([mj\tau]) \sim \frac{1}{j^{d+1/2}} \sum_{i=-(j-1)}^{j-1} (j - |i|) Y_m([mj\tau] + (j + i - 1)m)$ when $m \rightarrow \infty$ (in distribution) and more generally,

$$(Y_{jm}([mj\tau]), Y_{j'm}([mj'\tau'])) \xrightarrow{FDD} \left(\frac{1}{j^{d+1/2}} \sum_{i=-(j-1)}^{j-1} (j - |i|) Z_d(j\tau + j + i - 1), \frac{1}{(j')^{d+1/2}} \sum_{i'=-(j'-1)}^{j'-1} (j' - |i'|) Z_d(j'\tau' + j' + i' - 1) \right), \quad (3.20)$$

when $m \rightarrow \infty$. Hence, obvious computations lead to define for $t \in \mathbb{R}$

$$Z_d^{(j)}(t) := \sum_{i=-(j-1)}^{j-1} (j - |i|) Z_d(t + j + i - 1) = \frac{B_{d+0.5}(t + 2j) - 2B_{d+0.5}(t + j) + B_{d+0.5}(t)}{\sqrt{|4^{d+0.5} - 4|}} \quad (3.21)$$

$$\gamma_d^{(j,j')}(t) := \text{Cov}(\psi(Z_d^{(j)}(0), Z_d^{(j)}(j)), \psi(Z_d^{(j')}(t), Z_d^{(j')}(t + j'))). \quad (3.22)$$

Now, as the function $\psi(x, y) = \frac{|x+y|}{|x|+|y|}$ is a continuous (on $\mathbb{R}^2 \setminus \{0, 0\}$) and bounded function (with $0 \leq \psi(x, y) \leq 1$) and since $\eta_{m_j}([mj\tau]) = \psi(Y_{m_j}([mj\tau]), Y_{m_j}([mj(\tau + 1)]))$, then from (3.20),

$$\begin{aligned} \text{Cov}(\eta_{m_j}([mj\tau]), \eta_{m_{j'}}([mj'\tau'])) &\xrightarrow{m \rightarrow \infty} \text{Cov}(\psi(Z_d^{(j)}(j\tau), Z_d^{(j)}(j(\tau + 1))), \psi(Z_d^{(j')}(j'\tau'), Z_d^{(j')}(j'(\tau' + 1)))) \\ &\xrightarrow{m \rightarrow \infty} \gamma_d^{(j,j')}(j'\tau' - j\tau), \end{aligned} \quad (3.24)$$

using the stationarity of the process Z_d and therefore of processes $Z_d^{(j)}$ and $Z_d^{(j')}$. Hence, when N, m and $N/m \rightarrow \infty$,

$$\begin{aligned} \frac{N}{m} \text{Cov}(IR_N(jm), IR_N(j'm)) &\sim \frac{N}{m(\frac{N}{jm} - 3)(\frac{N}{j'm} - 3)} \\ &\times \int_0^{\frac{N-1}{jm}-3} \int_0^{\frac{N-1}{j'm}-3} \text{Cov}(\psi(Z_d^{(j)}(j\tau), Z_d^{(j)}(j\tau + j)), \psi(Z_d^{(j')}(j'\tau'), Z_d^{(j')}(j'\tau' + j'))) d\tau d\tau' \\ &\sim \frac{mN}{(N - 3jm)(N - 3j'm)} \int_0^{\frac{N-1}{m}-3j} \int_0^{\frac{N-1}{m}-3j'} \gamma_d^{(j,j')}(s' - s) ds ds' \\ &\sim \frac{m}{N} \int_{-\frac{N}{m}}^{\frac{N}{m}} (\frac{N}{m} - |u|) \gamma_d^{(j,j')}(u) du \\ &\rightarrow \int_{-\infty}^{\infty} \gamma_d^{(j,j')}(u) du =: \sigma_{j,j'}(d). \end{aligned} \quad (3.25)$$

This last limit is obtained, *mutatis mutandis*, from the relation (5.23) Surgailis *et al* [92], (2008), and thus $\gamma_d^{(j,j')}(u) = C(u^{-2} \wedge 1)$, implying $\frac{m}{N} \int_{-\frac{N}{m}}^{\frac{N}{m}} |u| \gamma_d^{(j,j')}(u) du \xrightarrow{N, m, \frac{N}{m} \rightarrow \infty} 0$. It achieves the first step of the proof.

Step 2 : It remains to prove the multidimensional central limit theorem. Then consider a linear combination of $(IR_N(m_j))_{1 \leq j \leq p}$, i.e. $\sum_{j=1}^p u_j IR_N(m_j)$ with $(u_1, \dots, u_p) \in \mathbb{R}^p$. For ease of notation, we will

restrict our purpose to $p = 2$, with $m_i = r_i m$ where $r_1 \leq r_2$ are fixed positive integers. Then with the previous notations and following the notations and results of Theorem 2.5 of Surgailis *et al.* [92], (2008) :

$$\begin{aligned} u_1 IR_N(r_1 m) + u_2 IR_N(r_2 m) &= u_1 (\mathbb{E}[IR_N(r_1 m)] + S_K(r_1 m) + \tilde{S}_K(r_1 m)) \\ &\quad + u_2 (\mathbb{E}[IR_N(r_2 m)] + S_K(r_2 m) + \tilde{S}_K(r_2 m)). \end{aligned}$$

From (5.31) of Surgailis *et al.* [92], (2008), we have $\tilde{S}_K(m_1) = o(S_K(m_1))$ and $\tilde{S}_K(m_2) = o(S_K(m_2))$ when $K \rightarrow \infty$ and from a Hermitian decomposition $(N/m)^{1/2}(u_1 S_K(m_i) + u_2 S_K(m_2)) \rightarrow_D \mathcal{N}(0, \gamma_K^2)$ as N, m and $N/m \rightarrow \infty$ since the cumulants of $(N/m)^{1/2}(u_1 S_K(m_i) + u_2 S_K(m_2))$ of order greater or equal to 3 converge to 0 (since this result is proved for each $S_K(m_i)$). Moreover, from the previous computations, $\gamma_K^2 \rightarrow (u_1^2 \sigma_{r_1, r_1}(d) + 2u_1 u_2 \sigma_{r_1, r_2}(d) + u_2^2 \sigma_{r_2, r_2}(d))$ when $K \rightarrow \infty$. Therefore the multidimensional central limit theorem is established. $\square$

Property 2. *Let X satisfy Assumption $S(d, \beta)$ with $-0.5 < d < 0.5$ and $\beta > 0$. Then, there exists a constant $K(d, \beta) < 0$ depending only on d and β such as*

$$\begin{aligned} \mathbb{E}[IR_N(m)] &= \Lambda_0(d) + K(d, \beta) \times m^{-\beta} + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log(m)) & \text{if } -2d + \beta < 1, \\ &= \Lambda_0(d) + K(d, \beta) \times m^{-\beta} \log(m) + O(m^{-\beta}) & \text{if } -2d + \beta = 1; \\ &= \Lambda_0(d) + O(m^{-2d-1}) & \text{if } -2d + \beta > 1. \end{aligned}$$

Proof of Property 2. As in Surgailis *et al* (2008), we can write :

$$\mathbb{E}[IR_N(m)] = \mathbb{E}\left(\frac{|Y^0 + Y^1|}{|Y^0| + |Y^1|}\right) = \Lambda\left(\frac{R_m}{V_m^2}\right) \quad \text{with} \quad \frac{R_m}{V_m^2} := 1 - 2 \frac{\int_0^\pi f(x) \frac{\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx}{\int_0^\pi f(x) \frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx}.$$

Therefore an expansion of R_m/V_m^2 will provide an expansion of $\mathbb{E}[IR_N(m)]$ when $m \rightarrow \infty$ and the multidimensional CLT (3.8) will be deduced from the Delta-method.

Step 1 Let f satisfy Assumption $S(d, \beta)$. Then we are going to establish that there exist positive real numbers C_1 and C_2 specified in (3.26) and (3.27) and such that :

1. if $-1 < -2d < 1$ and $-2d + \beta < 1$, $\frac{R_m}{V_m^2} = \rho(d) + C_1(-2d, \beta) m^{-\beta} + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log m)$
2. if $-1 < -2d < 1$ and $-2d + \beta = 1$, $\frac{R_m}{V_m^2} = \rho(d) + C_2(1 - \beta, \beta) m^{-\beta} \log m + O(m^{-\beta})$
3. if $-1 < -2d < 1$ and $-2d + \beta > 1$, $\frac{R_m}{V_m^2} = \rho(d) + O(m^{-2d-1})$.

Indeed under Assumption $S(d, \beta)$ and with $J_j(a, m)$, $j = 4, 6$, defined in (3.42) of Lemma 1 (see below), it is clear that,

$$\frac{R_m}{V_m^2} = 1 - 2 \frac{J_6(-2d, m) + \frac{c_1}{c_0} J_6(-2d + \beta, m) + O(J_6(-2d + \beta + \varepsilon))}{J_4(-2d, m) + \frac{c_1}{c_0} J_4(-2d + \beta, m) + O(J_4(-2d + \beta + \varepsilon))},$$

since $\int_0^\pi O(x^{-2d+\beta+\varepsilon}) \frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = O(J_j(-2d+\beta+\varepsilon))$ for $j = 4, 6$. Now we follow the results of Lemma 1.

1. Let $-1 < -2d + \beta < 1$. Then for any $\varepsilon > 0$,

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{C_{61}(-2d)m^{1+2d} + C_{62}(-2d) + \frac{c_1}{c_0}(C_{61}(-2d+\beta)m^{1+2d-\beta} + C_{62}(-2d+\beta)) + O(m^{1+2d-\beta-\varepsilon+\log m})}{C_{41}(-2d)m^{1+2d} + C_{42}(-2d) + \frac{c_1}{c_0}(C_{41}(-2d+\beta)m^{1+2d-\beta} + C_{42}(-2d+\beta)) + O(m^{1+2d-\beta-\varepsilon+\log m})} \\ &= 1 - \frac{2}{C_{41}(-2d)} \left[C_{61}(-2d) + \frac{c_1}{c_0} C_{61}(-2d+\beta)m^{-\beta} \right] \left[1 - \frac{c_1}{c_0} \frac{C_{41}(-2d+\beta)}{C_{41}(-2d)} m^{-\beta} \right] + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log m) \\ &= 1 - \frac{2C_{61}(-2d)}{C_{41}(-2d)} + 2 \frac{c_1}{c_0} \left[\frac{C_{61}(-2d)C_{41}(-2d+\beta)}{C_{41}(-2d)C_{41}(-2d)} - \frac{C_{61}(-2d+\beta)}{C_{41}(-2d)} \right] m^{-\beta} + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log m). \end{aligned}$$

As a consequence, with $\rho(d)$ defined in (3.7) and C_{j1} defined in Lemma 1,

$$\begin{aligned} \frac{R_m}{V_m^2} &= \rho(d) + C_1(-2d, \beta) m^{-\beta} + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log m) \quad (m \rightarrow \infty), \quad \text{with} \\ C_1(-2d, \beta) &:= 2 \frac{c_1}{c_0} \frac{1}{C_{41}^2(-2d)} [C_{61}(-2d)C_{41}(-2d+\beta) - C_{61}(-2d+\beta)C_{41}(-2d)], \quad (3.26) \end{aligned}$$

and numerical experiments prove that $C_1(-2d, \beta)/c_1$ is negative for any $d \in (-0.5, 0.5)$ and $\beta > 0$.

2. Let $-2d + \beta = 1$.

Again with Lemma 1,

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{[C_{61}(-2d)m^\beta + C'_{61} \frac{c_1}{c_0} \log(m\pi) + C_{62}(-2d) + \frac{c_1}{c_0} C'_{62} + O(1)]}{[C_{41}(-2d)m^\beta + C'_{41} \frac{c_1}{c_0} \log(m\pi) + C_{42}(-2d) + \frac{c_1}{c_0} C'_{42} + O(1)]} \\ &= 1 - \frac{2}{C_{41}(a)} [C_{61}(-2d) + (C'_{61} \frac{c_1}{c_0} \log(m))m^{-\beta}] \left[1 - \left(\frac{C'_{41}}{C_{41}(a)} \frac{c_1}{c_0} \log(m) \right) m^{-\beta} \right] + O(m^{-\beta}) \\ &= 1 - \frac{2}{C_{41}(-2d)} \left[C_{61}(-2d) - \frac{c_1}{c_0} \left(\frac{C_{61}(-2d)C'_{41}}{C_{41}(-2d)} - C'_{61} \right) \log(m) m^{-\beta} \right] + O(m^{-\beta}). \end{aligned}$$

As a consequence,

$$\begin{aligned} \frac{R_m}{V_m^2} &= \rho(d) + C_2(-2d, \beta) m^{-\beta} \log m + O(m^{-\beta}) \quad (m \rightarrow \infty), \quad \text{with} \\ C_2(-2d, \beta) &:= 2 \frac{c_1}{c_0} \frac{1}{C_{41}^2(-2d)} (C'_{41}C_{61}(-2d) - C'_{61}C_{41}(-2d)), \quad (3.27) \end{aligned}$$

and numerical experiments prove that $C_2(-2d, \beta)/c_1$ is negative for any $d \in (-0.5, 0.5)$ and $\beta = 1 - 2d$.

3. Let $-2d + \beta > 1$.

Once again with Lemma 1 :

$$\begin{aligned}
\frac{R_m}{V_m^2} &= 1 - 2 \frac{[C_{61}(-2d)m^{1+2d} + C_{62}(-2d) + \frac{c_1}{c_0}C_{61}''(-2d+\beta) + \frac{c_1}{c_0}C_{62}''(-2d+\beta)m^{1+2d-\beta} + O(1)]}{C_{41}(-2d)m^{1+2d}[1 + \frac{C_{42}(-2d)}{C_{41}(-2d)}m^{-2d-1} + \frac{c_1}{c_0}\frac{C_{41}''(-2d+\beta)}{C_{41}(-2d)}m^{-2d-1} + \frac{c_1}{c_0}\frac{C_{42}''(-2d+\beta)}{C_{41}(-2d)}m^{-\beta} + O(m^{-2d-1})]} \\
&= 1 - \frac{2}{C_{41}(-2d)}[C_{61}(-2d) + O(m^{-2d-1})][1 - O(m^{-2d-1})] \\
&= 1 - \frac{2C_{61}(-2d)}{C_{41}(-2d)} + O(m^{-2d-1}).
\end{aligned}$$

Note that it is not possible to specify the second order term of this expansion as in both the previous cases. As a consequence,

$$\frac{R_m}{V_m^2} = \rho(d) + O(m^{-2d-1}) \quad (m \rightarrow \infty). \quad (3.28)$$

Step 2 : A Taylor expansion of $\Lambda(\cdot)$ around $\rho(d)$ provides :

$$\Lambda\left(\frac{R_m}{V_m^2}\right) \simeq \Lambda(\rho(d)) + \left[\frac{\partial \Lambda}{\partial \rho}\right](\rho(d))\left(\frac{R_m}{V_m^2} - \rho(d)\right) + \frac{1}{2} \left[\frac{\partial^2 \Lambda}{\partial \rho^2}\right](\rho(d))\left(\frac{R_m}{V_m^2} - \rho(d)\right)^2. \quad (3.29)$$

Note that numerical experiments show that $\left[\frac{\partial \Lambda}{\partial \rho}\right](\rho) > 0.2$ for any $\rho \in (-1, 1)$. As a consequence, using the previous expansions of R_m/V_m^2 obtained in Step 1 and since $\mathbb{E}[IR_N(m)] = \Lambda(R_m/V_m^2)$, then

$$\mathbb{E}[IR_N(m)] = \Lambda_0(d) + \begin{cases} c_1 C_1'(d, \beta) m^{-\beta} + O(m^{-\beta-\varepsilon} + m^{-2d-1} \log m + m^{-2\beta}) & \text{if } \beta < 1 + 2d \\ c_1 C_2'(\beta) m^{-\beta} \log m + O(m^{-\beta}) & \text{if } \beta = 1 + 2d \\ O(m^{-2d-1}) & \text{if } \beta > 1 + 2d \end{cases},$$

with $C_1'(d, \beta) < 0$ for all $d \in (-0.5, 0.5)$ and $\beta \in (0, 1 + 2d)$ and $C_2'(\beta) < 0$ for all $0 < \beta < 2$. □

Proof of Theorem 1. Using Property 2, if $m \simeq C N^\alpha$ with $C > 0$ and $(1+2\beta)^{-1} \vee (4d+3)^{-1} < \alpha < 1$ then $\sqrt{N/m} (\mathbb{E}[IR_N(m)] - \Lambda_0(d)) \xrightarrow{N \rightarrow \infty} 0$ and it implies that the multidimensional CLT (3.4) can be replaced by

$$\sqrt{\frac{N}{m}} \left(IR_N(m_j) - \Lambda_0(d) \right)_{1 \leq j \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_p(d)). \quad (3.30)$$

It remains to apply the Delta-method with the function Λ_0^{-1} to CLT (3.30). This is possible since the function $d \rightarrow \Lambda_0(d)$ is an increasing function such that $\Lambda_0'(d) > 0$ and $(\Lambda_0^{-1})'(\Lambda_0(d)) = 1/\Lambda_0'(d) > 0$ for all $d \in (-0.5, 0.5)$. It achieves the proof of Theorem 1. □

Proof of Proposition 1. For ease of writing we will note $\widehat{\Sigma}_N$ instead of $\widehat{\Sigma}_N(N^\alpha)$ in the sequel. We have $(\widetilde{d}_N(m) - \widehat{d}_N(jm))_{1 \leq j \leq p} = \widehat{M}_N(\widehat{d}_N(jm) - d)_{1 \leq j \leq p}$ with $\widehat{M}_N$ the orthogonal (for the Euclidean

norm $\|\cdot\|_{\widehat{\Sigma}_N}$) projector matrix on $((1)_{1 \leq i \leq p})^\perp$ (which is a linear subspace with dimension $p - 1$ included in $\mathbb{R}^p$) in $\mathbb{R}^p$, i.e. $\widehat{M}_N = J_p(J'_p \widehat{\Sigma}_N^{-1} J_p)^{-1} J'_p \widehat{\Sigma}_N^{-1}$. Now, by denoting $\Sigma_N^{1/2}$ a symmetric matrix such as $\Sigma_N^{1/2} \Sigma_N^{1/2} = \Sigma_N$,

$$\begin{aligned} \|(\widehat{d}_N(m) - \widehat{d}_N(jm))_{1 \leq j \leq p}\|_{\widehat{\Sigma}_N}^2 &= (\widehat{d}_N(jm) - d)_{1 \leq j \leq p}' \widehat{M}_N \widehat{\Sigma}_N^{-1} \widehat{M}_N (\widehat{d}_N(jm) - d)_{1 \leq j \leq p} \\ &= Z'_N \widehat{\Sigma}_N^{1/2} \widehat{M}_N \widehat{\Sigma}_N^{-1} \widehat{M}_N \widehat{\Sigma}_N^{1/2} Z_N \\ &= (\widehat{A}_N Z_N)' (\widehat{A}_N Z_N) \end{aligned}$$

with $\widehat{A}_N = \Sigma_N^{-1/2} \widehat{M}_N \widehat{\Sigma}_N^{1/2}$ and Z_N a random vector such as $\sqrt{N/m} Z_N \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}_p(0, I_p)$ from Theorem 1. But we also have $\widehat{A}_N = \Sigma_N^{-1/2} J_p(J'_p \widehat{\Sigma}_N^{-1} J_p)^{-1} J'_p \widehat{\Sigma}_N^{-1/2} = \widehat{H}_N (\widehat{H}'_N \widehat{H}_N)^{-1} \widehat{H}'_N$ with $\widehat{H}_N = \Sigma_N^{-1/2} J_p$ a matrix of size $(p \times (p - 1))$ with rank $p - 1$ (since the rank of J_p is $(p - 1)$). Hence $\widehat{A}_N$ is an orthogonal projector to the linear subspace of dimension $p - 1$ generated by the matrix $\widehat{H}_N$. Now using Cochran Theorem (see for instance Anderson and Styan, 1982), $\sqrt{N/m} \widehat{A}_N Z_N$ is asymptotically a Gaussian vector such as $N/m (\widehat{A}_N Z_N)' (\widehat{A}_N Z_N) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p - 1)$. $\square$

In Property 2, a second order expansion of $\mathbb{E}[IR_N(m)]$ cannot be specified in the case $\beta > 2d + 1$. In the following Property 3, we show some inequalities satisfied by $\mathbb{E}[IR_N(m)]$ which will be useful for obtaining the consistency of the adaptive estimator in this case.

Property 3. *Let X satisfy Assumption $S(d, \beta)$ with $-0.5 < d < 0.5$, $\beta > 1 + 2d$. Moreover, suppose that the spectral density of X satisfies Condition (3.32) or (3.33). Then there exists a constant $L > 0$ depending only on $c_0, c_1, c_2, d, \beta, \varepsilon$ such that*

$$|\mathbb{E}[IR_N(m)] - \Lambda_0(d)| \geq L m^{-2d-1}. \quad (3.31)$$

Proof of Property 3. Using the expansion of $J_j(a, m)$, $j = 4, 6$, for $a > 1$ (see Lemma 1) and the same computations than in Property 2, we obtain

$$\begin{aligned} & -\frac{2}{C_{41}^2(-2d)} \left[(C_{62}(-2d)C_{41}(-2d) - C_{42}(-2d)C_{61}(-2d)) + \frac{c_1}{c_0} (C_{61}''(-2d+\beta)C_{41}(-2d) - C_{41}''(-2d+\beta)C_{61}(-2d)) \right. \\ & \quad \left. + \frac{|c_2|}{c_0} (C_{61}''(-2d+\beta+\varepsilon)C_{41}(-2d) + C_{41}''(-2d+\beta+\varepsilon)C_{61}(-2d)) \right] m^{-2d-1} (1 + o(1)) \\ & \leq \frac{R_m}{V_m^2} - \rho(d) \leq \\ & -\frac{2}{C_{41}^2(-2d)} \left[(C_{62}(-2d)C_{41}(-2d) - C_{42}(-2d)C_{61}(-2d)) + \frac{c_1}{c_0} (C_{61}''(-2d+\beta)C_{41}(-2d) - C_{41}''(-2d+\beta)C_{61}(-2d)) \right. \\ & \quad \left. - \frac{|c_2|}{c_0} (C_{61}''(-2d+\beta+\varepsilon)C_{41}(-2d) + C_{41}''(-2d+\beta+\varepsilon)C_{61}(-2d)) \right] m^{-2d-1} (1 + o(1)). \end{aligned}$$

Now, denote

$$\begin{aligned} D_0(d) &:= C_{62}(-2d)C_{41}(-2d) - C_{42}(-2d)C_{61}(-2d) = \frac{C_{42}(-2d)C_{41}(-2d)}{48(1-2^{-1+2d})} (2^{4+2d} - 5 - 3^{2+2d}), \\ D_1(d, \beta) &:= C_{62}(-2d + \beta)C_{41}(-2d) - C_{42}(-2d + \beta)C_{61}(-2d) = \frac{C_{42}(-2d + \beta)C_{41}(-2d)}{128(1-2^{-1+2d})} (2^{4+2d} - 5 - 3^{2+2d}), \\ D_2(d, \beta, \varepsilon) &:= C_{61}''(-2d + \beta + \varepsilon)C_{41}(-2d) + C_{41}''(-2d + \beta + \varepsilon)C_{61}(-2d). \end{aligned}$$

Since $-0.5 < d < 0.5$, $2^{4+2d} - 5 - 3^{2+2d} > 0$ and $1 - 2^{-1+2d} > 0$. Moreover, from the sign of the constants presented in Lemma 1, we have $D_0(d) \neq 0$ except for $d = 0$, $D_1(d, \beta) \neq 0$ except for $d = 2\beta$ and $D_2(d, \beta, \varepsilon) > 0$ for all $d \in (-0.5, 0.5)$, $\beta > 0$ and $\varepsilon > 0$. Therefore, if $c_0, c_1, c_2, d, \beta, \varepsilon$ are such that

$$K_1 := D_0(d) + \frac{c_1}{c_0} D_1(d, \beta) - \frac{|c_2|}{c_0} D_2(d, \beta, \varepsilon) > 0 \quad (3.32)$$

$$\text{or } K_2 := D_0(d) + \frac{c_1}{c_0} D_1(d, \beta) + \frac{|c_2|}{c_0} D_2(d, \beta, \varepsilon) < 0. \quad (3.33)$$

and from the signs of $D_0(d)$, $D_1(d, \beta)$ and $D_2(d, \beta, \varepsilon)$, when (d, β, ε) is fixed, these conditions are not impossible but hold following the values of $\frac{c_1}{c_0}$ and $\frac{|c_2|}{c_0}$. Then $\frac{R_m}{V_m^2} - \rho(d) \leq -\frac{K_1}{C_{41}^2(-2d)} m^{-2d-1}$ or $\frac{R_m}{V_m^2} - \rho(d) \geq -\frac{K_2}{C_{41}^2(-2d)} m^{-2d-1}$ for m large enough following (3.32) or (3.33) holds. Then, if (3.32) holds, since $\mathbb{E}[IR_N(m)] = \Lambda(\frac{R_m}{V_m^2})$, since the function $r \rightarrow \Lambda(r)$ is an increasing and $\mathcal{C}^1$ function and since $\mathbb{E}[IR_N(m)] = \Lambda(\frac{R_m}{V_m^2})$ then when m large enough, from a Taylor expansion,

$$\mathbb{E}[IR_N(m)] \leq \Lambda\left(\rho(d) - \frac{K_1}{C_{41}^2(-2d)} m^{-2d-1}\right) \implies \mathbb{E}[IR_N(m)] \leq \Lambda_0(d) - \frac{1}{2} \Lambda'(\rho(d)) \frac{K_1}{C_{41}^2(-2d)} m^{-2d-1}.$$

Now following the same process if (3.33) holds, we deduce inequality (3.31). $\square$

Proof of Proposition 2. Let $\varepsilon > 0$ be a fixed positive real number, such that $\alpha^* + \varepsilon < 1$.

I. First, a bound of $\Pr(\hat{\alpha}_N \leq \alpha^* + \varepsilon)$ is provided. Indeed,

$$\begin{aligned} \Pr(\hat{\alpha}_N \leq \alpha^* + \varepsilon) &\geq \Pr\left(\hat{Q}_N(\alpha^* + \varepsilon/2) \leq \min_{\alpha \geq \alpha^* + \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \hat{Q}_N(\alpha)\right) \\ &\geq 1 - \Pr\left(\bigcup_{\alpha \geq \alpha^* + \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \hat{Q}_N(\alpha^* + \varepsilon/2) > \hat{Q}_N(\alpha)\right) \\ &\geq 1 - \sum_{k=\lceil(\alpha^* + \varepsilon) \log N\rceil}^{\log[N/p]} \Pr\left(\hat{Q}_N(\alpha^* + \varepsilon/2) > \hat{Q}_N\left(\frac{k}{\log N}\right)\right). \end{aligned} \quad (3.34)$$

But, for $\alpha \geq \alpha^* + \varepsilon$,

$$\begin{aligned} & \Pr \left(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha) \right) \\ &= \Pr \left(\left\| (\widehat{d}_N(i N^{\alpha^* + \varepsilon/2}))_{1 \leq i \leq p} - \widetilde{d}_N(N^{\alpha^* + \varepsilon/2}) \right\|_{\widehat{\Sigma}_N(N^{\alpha^* + \varepsilon/2})}^2 > \left\| (\widehat{d}_N(i N^\alpha) - \widetilde{d}_N(N^\alpha))_{1 \leq i \leq p} \right\|_{\widehat{\Sigma}_N(N^\alpha)}^2 \right) \\ & \text{with } \|X\|_\Omega^2 = X' \Omega^{-1} X. \text{ Set } Z_N(\alpha) = \frac{N}{N^\alpha} \left\| (\widehat{d}_N(i N^\alpha))_{1 \leq i \leq p} - \widetilde{d}_N(N^\alpha) \right\|_{\widehat{\Sigma}_N(N^\alpha)}^2. \text{ Then,} \\ & \Pr \left(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha) \right) = \Pr \left(Z_N(\alpha^* + \varepsilon/2) > N^{\alpha - (\alpha^* + \varepsilon/2)} Z_N(\alpha) \right) \\ & \leq \Pr \left(Z_N(\alpha^* + \varepsilon/2) > N^{(\alpha - (\alpha^* + \varepsilon/2))/2} \right) + \Pr \left(Z_N(\alpha) < N^{-(\alpha - (\alpha^* + \varepsilon/2))/2} \right). \end{aligned}$$

From Proposition 1, for all $\alpha > \alpha^*$, $Z_N(\alpha) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(p-1)$. As a consequence, for N large enough,

$$\Pr \left(Z_N(\alpha) \leq N^{-(\alpha - (\alpha^* + \varepsilon/2))/2} \right) \leq \frac{2}{2^{(p-1)/2} \Gamma((p-1)/2)} \cdot N^{-(\frac{p-1}{2}) \frac{(\alpha - (\alpha^* + \varepsilon/2))}{2}}.$$

Moreover, from Markov inequality and with N large enough,

$$\begin{aligned} \Pr \left(Z_N(\alpha^* + \varepsilon/2) > N^{(\alpha - (\alpha^* + \varepsilon/2))/2} \right) &\leq 2 \Pr \left(\exp(\sqrt{\chi^2(p-1)}) > \exp(N^{(\alpha - (\alpha^* + \varepsilon/2))/4}) \right) \\ &\leq 2 \mathbb{E}(\exp(\sqrt{\chi^2(p-1)})) \exp(-N^{(\alpha - (\alpha^* + \varepsilon/2))/4}). \end{aligned}$$

We deduce that there exists $M_1 > 0$ not depending on N , such that for large enough N ,

$$\Pr \left(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha) \right) \leq M_1 \exp(-N^{(\alpha - (\alpha^* + \varepsilon/2))/4}).$$

since $\mathbb{E}(\exp(\sqrt{\chi^2(p-1)})) < \infty$ does not depend on N . Thus, inequality (3.34) becomes, with $M_2 > 0$ and for N large enough,

$$\begin{aligned} \Pr(\widehat{\alpha}_N \leq \alpha^* + \varepsilon) &\geq 1 - M_1 e^{-N^{\varepsilon/8}} \sum_{k=0}^{\log[N/p] - [(\alpha^* + \varepsilon) \log N]} \exp(-N^{\frac{k}{4 \log N}}) \\ &\geq 1 - M_2 e^{-N^{\varepsilon/8}}. \end{aligned} \quad (3.35)$$

II. Second, a bound of $\Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon)$ can also be computed. Following the previous arguments and notations,

$$\begin{aligned} \Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon) &\geq \Pr \left(\widehat{Q}_N(\alpha^* + \frac{1 - \alpha^*}{2\alpha^*} \varepsilon) \leq \min_{\alpha \leq \alpha^* - \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha) \right) \\ &\geq 1 - \sum_{k=2}^{[(\alpha^* - \varepsilon) \log N] + 1} \Pr \left(\widehat{Q}_N(\alpha^* + \frac{1 - \alpha^*}{2\alpha^*} \varepsilon) > \widehat{Q}_N\left(\frac{k}{\log N}\right) \right), \end{aligned} \quad (3.36)$$

and as above, with $Z_N(\alpha) = \frac{N}{N^\alpha} \left\| (\widehat{d}_N(i N^\alpha) - \widetilde{d}_N(N^\alpha))_{1 \leq i \leq p} \right\|_{\widehat{\Sigma}_N(N^\alpha)}^2$,

$$\Pr \left(\widehat{Q}_N(\alpha^* + \frac{1 - \alpha^*}{2\alpha^*} \varepsilon) > \widehat{Q}_N(\alpha) \right) = \Pr \left(Z_N(\alpha^* + \frac{1 - \alpha^*}{2\alpha^*} \varepsilon) > N^{\alpha - (\alpha^* + \frac{1 - \alpha^*}{2\alpha^*} \varepsilon)} Z_N(\alpha) \right). \quad (3.37)$$

- if $\beta \leq 2d + 1$, with $\alpha < \alpha^* = (1 + 2\beta)^{-1}$, from Property 2 and with $C \neq 0$, for $1 \leq i \leq p$,

$$\begin{aligned} \sqrt{\frac{N}{N^\alpha}} (\mathbb{E}[IR(i N^\alpha)] - \Lambda_0(d)) &\simeq C i^{-(1-\alpha^*)/2\alpha^*} N^{(\alpha^*-\alpha)/2\alpha^*} (\log N)^{\beta=2d+1} \\ \implies \sqrt{\frac{N}{N^\alpha}} (\Lambda_0^{-1}(\mathbb{E}[IR(i N^\alpha)]) - d) &\simeq C' i^{-(1-\alpha^*)/2\alpha^*} N^{(\alpha^*-\alpha)/2\alpha^*} (\log N)^{\beta=2d+1} \end{aligned} \quad (3.38)$$

with $C' \neq 0$, since $\Lambda_0(d) > 0$ for all $d \in (-0.5, 0.5)$. We deduce

$$\sqrt{\frac{N}{N^\alpha}} (\widehat{d}_N(i N^\alpha) - d)_{1 \leq i \leq p} \simeq C'' N^{(\alpha^*-\alpha)/2\alpha^*} (\log N)^{\beta=2d+1} (i^{-(1-\alpha^*)/2\alpha^*})_{1 \leq i \leq p} + (\widehat{\varepsilon}_N(i N^\alpha))_{1 \leq i \leq p},$$

with $C'' \neq 0$ and $(\widehat{\varepsilon}_N(i N^\alpha))_{1 \leq i \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, (\Lambda'_0(d))^{-2} \Gamma_p(d)\right)$ from Proposition 1. Now from the definition of $\widetilde{d}_N(N^\alpha)$, we have $(\widehat{d}_N(i N^\alpha) - \widetilde{d}_N(N^\alpha))_{1 \leq i \leq p} = \widehat{M}_N(\widehat{d}_N(i N^\alpha) - d)_{1 \leq i \leq p}$ with $\widehat{M}_N$ the orthogonal projector matrix on $(1)_{1 \leq i \leq p}^\perp$.

a consequence, for $\alpha < \alpha^* - \varepsilon$ and with the inequality $\|a - b\|^2 \geq \frac{1}{2}\|a\|^2 - \|b\|^2$,

$$Z_N(\alpha) \geq \frac{1}{2}(C'')^2 N^{\frac{\alpha^*-\alpha}{\alpha^*}} (\log^2 N)^{\beta=2d+1} \left\| \widehat{M}_N(i^{-\frac{1-\alpha^*}{2\alpha^*}})_{1 \leq i \leq p} \right\|_{\widehat{\Sigma}_N(N^\alpha)}^2 - \|\widehat{M}_N \widehat{\varepsilon}_N(i N^\alpha)\|_{\widehat{\Sigma}_N(N^\alpha)}^2.$$

Now, it is clear that $\|\widehat{M}_N \widehat{\varepsilon}_N(i N^\alpha)\|_{\widehat{\Sigma}_N(N^\alpha)}^2 \leq \|\widehat{\varepsilon}_N(i N^\alpha)\|_{\widehat{\Sigma}_N(N^\alpha)}^2 \leq C_1$ when N large enough, with $C_1 > 0$ not depending on N . Moreover the vector $(i^{-\frac{1-\alpha^*}{2\alpha^*}})_{1 \leq i \leq p}$ is not in the subspace $(1)_{1 \leq i \leq p}$ and therefore $\left\| \widehat{M}_N(i^{-\frac{1-\alpha^*}{2\alpha^*}})_{1 \leq i \leq p} \right\|_{\widehat{\Sigma}_N(N^\alpha)} \geq C_2$ for N large enough with $C_2 > 0$. We deduce that there exists $D > 0$ such that for N large enough and $\alpha < \alpha^* - \varepsilon$,

$$Z_N(\alpha) \geq D N^{\frac{\alpha^*-\alpha}{\alpha^*}} (\log^2 N)^{\beta=2d+1}.$$

Therefore, since $N^{\frac{\alpha^*-\alpha}{\alpha^*}} \xrightarrow[N \rightarrow \infty]{} \infty$ when $\alpha < \alpha^* - \varepsilon$,

$$\Pr\left(Z_N(\alpha) \geq \frac{1}{2} D N^{\frac{\alpha^*-\alpha}{\alpha^*}}\right) \xrightarrow[N \rightarrow \infty]{} 1.$$

Then, relation (3.37) becomes for $\alpha < \alpha^* - \varepsilon$ and N large enough,

$$\begin{aligned} \Pr\left(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\right) &\leq \Pr\left(\chi^2(p-1) \geq \left(\frac{1}{2} D N^{\frac{\alpha^*-\alpha}{\alpha^*}}\right) N^{\alpha-(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon)}\right) \\ &\leq \Pr\left(\chi^2(p-1) \geq \frac{D}{2} N^{\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)}\right) \\ &\leq M_2 N^{-(\frac{p-1}{2})\frac{1-\alpha^*}{2\alpha^*}\varepsilon}, \end{aligned}$$

with $M_2 > 0$, because $\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon) \geq \frac{1-\alpha^*}{2\alpha^*}\varepsilon$ for all $\alpha \leq \alpha^* - \varepsilon$. Hence, from inequality (3.36), for large enough N ,

$$\Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon) \geq 1 - M_2 \log N N^{-(p-1)\frac{1-\alpha^*}{4\alpha^*}\varepsilon}. \quad (3.39)$$

• if $\beta > 2d + 1$, with $\alpha < \alpha^* = (4d + 3)^{-1}$ and from Property 3, we obtain an inequality instead of (3.38) :

$$\left| \Lambda_0^{-1}(\mathbb{E}[IR_N(m)]) - d \right| \geq \frac{1}{2}(\Lambda_0(d))^{-1} L m^{-2d-1}$$

since the function $x \mapsto \Lambda_0^{-1}(x)$ is an increasing and $\mathcal{C}^1$ function, using a Taylor expansion. Therefore for $1 \leq i \leq p$,

$$\sqrt{\frac{N}{N^\alpha}} \left| \Lambda_0^{-1}(\mathbb{E}[IR(i N^\alpha)]) - d \right| \geq \frac{1}{2}(\Lambda_0(d))^{-1} L i^{-(1-\alpha^*)/2\alpha^*} N^{(\alpha^*-\alpha)/2\alpha^*}. \quad (3.40)$$

Now, as previously and with the same notation,

$$(\widehat{d}_N(i N^\alpha) - \widetilde{d}_N(N^\alpha))_{1 \leq i \leq p} \simeq \widehat{M}_n(\Lambda_0^{-1}(\mathbb{E}[IR(i N^\alpha)]) - d)_{1 \leq i \leq p} + \widehat{M}_n(\widehat{\varepsilon}_N(i N^\alpha))_{1 \leq i \leq p}. \quad (3.41)$$

Now plugging (3.40) in (3.41) and following the same steps of the proof in the case $\beta \leq 2d + 1$, the same kind of bound (3.39) can be obtained.

Finally, inequalities (3.35) and (3.39) imply that $\Pr(|\widehat{\alpha}_N - \alpha^*| \geq \varepsilon) \xrightarrow{N \rightarrow \infty} 0$. □

Proof of Theorem 2. The results of Theorem 2 can be easily deduced from Theorem 1 and Proposition 2 (and its proof) by using conditional probabilities. □

Proof of Proposition 3. Proposition 3 can be deduced from Theorem 2 using the same kind of proof than in Proposition 1 and conditional distributions. □

Lemma 1. For $j = 4, 6$, denote

$$J_j(a, m) := \int_0^\pi x^a \frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx. \quad (3.42)$$

Then, we have the following expansion when $m \rightarrow \infty$:

1. if $-1 < a < 1$, $J_j(a, m) = C_{j1}(a) m^{1-a} + C_{j2}(a) + O(m^{-1-(a \wedge 0)})$;
2. if $a = 1$, $J_j(a, m) = C'_{j1} \log(m) + C'_{j2} + O(m^{-1})$;
3. if $a > 1$, $J_j(a, m) = C''_{j1}(a) + O(m^{1-a} + m^{-2})$,

where constants $C_{j1}(a)$, $C_{j2}(a)$, $C'_{j1}(a)$, $C'_{j2}(a)$ and $C''_{j1}(a)$ are specified in the following proof.

Proof of Lemma 1. 1. Let $-1 < a < 1$.

We begin with the expansion of $J_4(a, m)$. First, decompose $J_4(a, m)$ as follows

$$J_4(a, m) = 2^{a+1} \int_0^{\frac{\pi}{2}} y^a \sin^4(my) \left[\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right] dy + \int_0^\pi \frac{x^a}{(\frac{x}{2})^2} \sin^4(\frac{mx}{2}) dx. \quad (3.43)$$

Using integrations by parts and $\sin^4(\frac{x}{2}) = \sin^2(\frac{x}{2}) - \frac{1}{4} \sin^2(x) = \frac{1}{8}(3 - 4\cos(y) + \cos(2y))$, we obtain for $m \rightarrow \infty$:

$$\begin{aligned} \int_0^\pi \frac{x^a}{(\frac{x}{2})^2} \sin^4(\frac{mx}{2}) dx &= 4m^{1-a} \left(\left(1 - \frac{1}{2^{1+a}}\right) \int_0^\infty \frac{\sin^2(\frac{y}{2})}{y^{2(\frac{1-a}{2})+1}} dy - \frac{1}{8} \int_{m\pi}^\infty y^{a-2} (3 - 4\cos(y) + \cos(2y)) dy \right) \\ &= \frac{\pi(1 - \frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})} m^{1-a} - 3 \frac{1}{2(1-a)} \pi^{a-1} + O(m^{-1}) \end{aligned}$$

where the left right side term of the last relation is obtained by integration by parts and the left side term is deduced from the following relation (see Doukhan *et al.* [29], 2003, p. 31)

$$\int_0^\infty y^{-\alpha} \sin(y) dy = \frac{1}{2} \frac{\pi}{\Gamma(\alpha) \sin(\pi(\frac{\alpha}{2}))} \quad \text{for } 0 < \alpha < 2. \quad (3.44)$$

Moreover, with the linearization of $\sin^4 u$ and Taylor expansions $\frac{1}{\sin^2(y)} - \frac{1}{y^2} \underset{y \rightarrow 0}{\sim} \frac{1}{3}$ and $\frac{\cos(y)}{\sin^3(y)} \underset{y \rightarrow 0}{\sim} \frac{y}{15}$,

$$2^{a+1} \int_0^{\frac{\pi}{2}} y^a \sin^4(my) \left[\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right] dy = 3 \frac{2^{a+1}}{8} \int_0^{\frac{\pi}{2}} y^a \left[\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right] dy + O(m^{-1-(a \wedge 0)}). \quad (3.45)$$

Finally, by replacing this expansion in (3.43), one deduces

$$\begin{aligned} J_4(a, m) &= \int_0^\pi x^a \frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = C_{41}(a) m^{1-a} + C_{42}(a) + O(m^{-1-(a \wedge 0)}) \quad (m \rightarrow \infty), \text{ with} \\ C_{41}(a) &:= \frac{\pi(1 - \frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})} \quad \text{and} \quad C_{42}(a) := \frac{3}{2^{2-a}} \int_0^{\frac{\pi}{2}} y^a \left[\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right] dy - \frac{3}{2(1-a)} \pi^{a-1}. \end{aligned} \quad (3.46)$$

Note that $C_{41}(a) > 0$ and $C_{42}(a) < 0$ for all $0 < a < 1$, $C_{42}(a) > 0$ for all $-1 < a < 0$, $C_{42}(0) = 0$.

A similar expansion procedure of $J_6(a, m)$ with $\sin^6(\frac{mx}{2})$ instead of $\sin^4(\frac{mx}{2})$ can be provided. As previously with $\sin^6(\frac{y}{2}) = \frac{1}{32}(10 - 15\cos(y) + 6\cos(2y) - \cos(3y))$, when $m \rightarrow \infty$,

$$\begin{aligned} J_6(a, m) &= C_{61}(a) m^{1-a} + C_{62}(a) + O(m^{-1-(a \wedge 0)}), \\ \text{with } C_{61}(a) &:= \frac{\pi(15 + 3^{1-a} - 2^{1-a}6)}{16(1-a)\Gamma(1-a)\sin(\frac{\pi}{2}(1-a))} \quad \text{and} \quad C_{62}(a) := \frac{5}{6} C_{42}(a). \end{aligned}$$

Moreover it is clear that $C_{61}(a) > 0$.

2. let $a = 1$.

When $m \rightarrow \infty$ we obtain the following expansion :

$$\int_0^\pi \frac{x \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = \frac{1}{2} \left(\int_0^{m\pi} \frac{\sin(2x) - 2x}{2x^2} dx - 4 \int_0^{m\pi} \frac{\sin(x) - x}{x^2} dx \right) + 4 \int_0^{\frac{\pi}{2}} y \sin^4(my) \left(\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right) dy$$

But,

$$\int_0^{m\pi} \frac{\sin(2x) - 2x}{2x^2} dx - 4 \int_0^{m\pi} \frac{\sin(x) - x}{x^2} dx = \frac{3}{2} \left(\log(m\pi) + \int_1^\infty \frac{\sin y}{y^2} dy + \int_0^1 \frac{\sin y - y}{y^2} dy \right) + O(m^{-1}).$$

Moreover from previous computations (see the case $a < 1$),

$$\int_0^{\frac{\pi}{2}} y \sin^4(my) \left(\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right) dy = \frac{3}{8} \int_0^{\frac{\pi}{2}} y \left(\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right) dy + O(m^{-1}).$$

As a consequence, when $m \rightarrow \infty$,

$$\begin{aligned} \int_0^\pi \frac{x \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx &= C'_{41} \log(m) + C'_{42} + O(m^{-1}), \quad \text{with } C'_{41} := \frac{3}{2} \quad \text{and} \\ C'_{42} &:= \frac{3}{2} \left(\log(\pi) + \int_0^{\frac{\pi}{2}} y \left(\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right) dy + \int_1^\infty \frac{\sin y}{y^2} dy + \int_0^1 \frac{\sin y - y}{y^2} dy \right). \end{aligned}$$

Note that $C'_{41} > 0$ and $C'_{42} \simeq 2.34 > 0$.

In the same way, we obtain the following expansions when $m \rightarrow \infty$,

$$\begin{aligned} \int_0^\pi \frac{x \sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx &= C'_{61} \log(m) + C'_{62} + O(m^{-1}) \quad \text{with } C'_{61} := \frac{5}{4} \quad \text{and} \\ C'_{62} &:= \frac{5}{4} \log(\pi) + \frac{5}{4} \int_0^{\frac{\pi}{2}} y \left(\frac{1}{\sin^2(y)} - \frac{1}{y^2} \right) dy + \frac{1}{8} \int_1^\infty \frac{1}{y} \left(-\cos(3y) + 6\cos(2y) - 15\cos(y) \right) dy + 4 \int_0^1 \frac{1}{y} \sin^6\left(\frac{y}{2}\right) dy. \end{aligned}$$

Note again that $C'_{61} > 0$ and numerical experiments show that $C'_{62} > 0$.

3. Let $a > 1$. Then, with the linearization of $\sin^4(u)$,

$$\begin{aligned} \int_0^\pi \frac{x^a \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx &= \frac{3}{8} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} dx - \frac{1}{2} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} \cos(mx) dx + \frac{1}{8} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} \cos(2mx) dx \\ &= C''_{41}(a) + \frac{1}{m} \int_0^\pi \left(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16} \right) \left(g(x) + h(x) \right) dx, \end{aligned} \quad (3.47)$$

with : $g(x) = \left(\frac{ax^{a-1}}{\sin^2(\frac{x}{2})} - 4ax^{a-3} \right) - \left(\frac{x^a \cos(\frac{x}{2})}{\sin^3(\frac{x}{2})} - 8x^{a-3} \right)$ and $h(x) = (4a - 8)x^{a-3}$.

First, if $1 < a$, with an integration by parts,

$$\frac{1}{m} \int_0^\pi \left(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16} \right) h(x) dx = O(m^{1-a} + m^{-2}). \quad (3.48)$$

Moreover,

$$\begin{aligned} \frac{1}{m} \int_0^\pi \left(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16} \right) g(x) dx \\ = \left(\frac{1}{32} - \frac{(-1)^m}{2} \right) (a\pi^2 - 4a + 8) \pi^{a-3} \frac{1}{m^2} - \frac{1}{m^2} \int_0^\pi \left(-\frac{\cos(mx)}{2} + \frac{\cos(2mx)}{32} \right) g'(x) dx \end{aligned}$$

since $g(x) \underset{x=0+}{\sim} \frac{a}{3} x^{a-1}$ and $g'(x) \underset{x=0+}{\sim} \frac{a(a-1)}{3} x^{a-2}$. Therefore, if $1 < a$,

$$\frac{1}{m} \int_0^\pi \left(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16} \right) g(x) dx = O(m^{-2}).$$

In conclusion, for $1 < a$ we deduce,

$$\int_0^\pi \frac{x^a \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = C''_{41}(a) + O(m^{1-a} + m^{-2}) \quad \text{with} \quad C''_{41}(a) := \frac{3}{8} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} dx > 0.$$

Similarly, for $1 < a$ we deduce,

$$\int_0^\pi \frac{x^a \sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = C''_{61}(a) + O(m^{1-a} + m^{-2}) \quad \text{with} \quad C''_{61}(a) := \frac{5}{16} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} dx = \frac{5}{6} C''_{41}(a) > 0.$$

□

Acknowledgments. The authors are grateful to both the referees for their very careful reading and many relevant suggestions and corrections that strongly improve the content and the form of the paper.

		Model	Estimates	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$N = 10^3$	fGn ($H = d + 1/2$)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.088*	0.094	0.101	0.111	
		$\text{mean}(\tilde{m}_N)$	11.8	12.5	16.0	19.4	
		$\widehat{proba}$	0.93	0.89	0.86	0.85	
	FARIMA(0, d , 0)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.112	0.099	0.094*	0.107	
		$\text{mean}(\tilde{m}_N)$	13.9	12.5	14.6	17.9	
		$\widehat{proba}$	0.94	0.92	0.88	0.86	
	FARIMA(1, d , 1)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.141	0.136*	0.140	0.149	
		$\text{mean}(\tilde{m}_N)$	15.2	15.0	18.2	21.1	
		$\widehat{proba}$	0.94	0.89	0.86	0.82	
$X^{(d,\beta)}, \beta = 1$	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.122	0.112*	0.121	0.123		
	$\text{mean}(\tilde{m}_N)$	14.1	13.8	16.2	20.0		
	$\widehat{proba}$	0.91	0.90	0.87	0.85		

		Model	Estimates	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$N = 10^4$	fGn ($H = d + 1/2$)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.030	0.022	0.019	0.018*	
		$\text{mean}(\tilde{m}_N)$	13.7	10.3	9.4	8.9	
		$\widehat{proba}$	0.95	0.89	0.87	0.84	
	FARIMA(0, d , 0)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.039	0.034	0.033	0.031*	
		$\text{mean}(\tilde{m}_N)$	11.5	9.0	8.0	7.2	
		$\widehat{proba}$	0.95	0.90	0.88	0.82	
	FARIMA(1, d , 1)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.067	0.062	0.061*	0.061*	
		$\text{mean}(\tilde{m}_N)$	18.1	15.9	13.8	13.3	
		$\widehat{proba}$	0.95	0.90	0.84	0.78	
$X^{(d,\beta)}, \beta = 1$	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.071	0.068	0.067*	0.071		
	$\text{mean}(\tilde{m}_N)$	15.2	13.6	11.7	10.9		
	$\widehat{proba}$	0.92	0.88	0.85	0.80		

		Model	Estimates	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$N = 10^5$	fGn ($H = d + 1/2$)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.012	0.008	0.007	0.006*	
		$\text{mean}(\tilde{m}_N)$	14.0	9.8	6.9	7.9	
		$\widehat{proba}$	0.92	0.90	0.87	0.85	
	FARIMA(0, d , 0)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.021	0.019*	0.019*	0.019*	
		$\text{mean}(\tilde{m}_N)$	15.8	12.7	11.1	9.8	
		$\widehat{proba}$	0.96	0.94	0.92	0.89	
	FARIMA(1, d , 1)	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.039	0.037	0.035*	0.035*	
		$\text{mean}(\tilde{m}_N)$	25.7	21.8	21.4	20.4	
		$\widehat{proba}$	0.98	0.98	0.94	0.93	
$X^{(d,\beta)}, \beta = 1$	$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.042	0.042	0.040*	0.041		
	$\text{mean}(\tilde{m}_N)$	22.3	19.9	19.7	16.9		
	$\widehat{proba}$	0.99	0.97	0.93	0.90		

TABLE 3.1 – $\sqrt{MSE}$ of the estimator $\tilde{d}_N^{(IR)}$, sample mean of the estimator $\tilde{m}_N$ and sample frequency that $\widehat{T}_N \leq q_{\chi^2(p-1)}(0.95)$ following p from simulations of the different processes of the benchmark. For each value of N (10^3 , 10^4 and 10^5), of d (-0.4 , -0.2 , 0 , 0.2 and 0.4) and p (5 , 10 , 15 , 20), 100 independent samples of each process are generated. The values $\sqrt{MSE} \tilde{d}_N^{(IR)}$, $\text{mean}(\tilde{m}_N)$ and $\widehat{proba}$ are obtained from sample mean on the different values of d .

$N = 10^3 \longrightarrow$

Model	$\sqrt{MSE}$	$d = -0.4$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$
fGn ($H = d + 1/2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.102	0.088	0.094 *	0.095	0.098
	$\sqrt{MSE} \hat{d}_R$	0.091	0.108	0.106	0.117	0.090
	$\sqrt{MSE} \hat{d}_W$	0.215	0.103	0.078	0.073*	0.061*
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.074*	0.087*	0.102	0.084	0.110
	$\sqrt{MSE} \hat{d}_N(10)$	0.096	0.135	0.154	0.158	0.154
	$\sqrt{MSE} \hat{d}_N(30)$	0.112	0.192	0.246	0.270	0.252
FARIMA(0, d , 0)	$\sqrt{MSE} \hat{d}_{MS}$	0.096	0.096	0.098	0.096	0.093
	$\sqrt{MSE} \hat{d}_R$	0.094	0.113	0.107	0.112	0.084
	$\sqrt{MSE} \hat{d}_W$	0.069*	0.073*	0.074*	0.082*	0.085*
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.116	0.085	0.103	0.094	0.101
	$\sqrt{MSE} \hat{d}_N(10)$	0.139	0.133	0.148	0.146	0.156
	$\sqrt{MSE} \hat{d}_N(30)$	0.157	0.209	0.232	0.247	0.243
FARIMA(1, d , 1)	$\sqrt{MSE} \hat{d}_{MS}$	0.098	0.092*	0.089*	0.088*	0.094
	$\sqrt{MSE} \hat{d}_R$	0.093*	0.110	0.115	0.110	0.089*
	$\sqrt{MSE} \hat{d}_W$	0.108	0.120	0.113	0.117	0.095
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.153	0.131	0.135	0.138	0.123
	$\sqrt{MSE} \hat{d}_N(10)$	0.212	0.188	0.173	0.157	0.155
	$\sqrt{MSE} \hat{d}_N(30)$	0.197	0.228	0.250	0.265	0.280
$X^{(D, D')}, D' = 1$	$\sqrt{MSE} \hat{d}_{MS}$	0.092	0.089*	0.113*	0.107*	0.100*
	$\sqrt{MSE} \hat{d}_R$	0.093	0.111	0.129	0.124	0.111
	$\sqrt{MSE} \hat{d}_W$	0.217	0.209	0.211	0.201	0.189
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.075*	0.101	0.121	0.122	0.131
	$\sqrt{MSE} \hat{d}_N(10)$	0.109	0.143	0.163	0.168	0.180
	$\sqrt{MSE} \hat{d}_N(30)$	0.109	0.177	0.228	0.249	0.247

 $N = 10^4 \longrightarrow$

Model	$\sqrt{MSE}$	$d = -0.4$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$
fGn ($H = d + 1/2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.040	0.031	0.032	0.035	0.035
	$\sqrt{MSE} \hat{d}_R$	0.040	0.027	0.029	0.031	0.030
	$\sqrt{MSE} \hat{d}_W$	0.129	0.045	0.026	0.022	0.020
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.019*	0.019*	0.017*	0.016*	0.019*
	$\sqrt{MSE} \hat{d}_N(10)$	0.036	0.038	0.049	0.043	0.048
	$\sqrt{MSE} \hat{d}_N(30)$	0.043	0.070	0.086	0.081	0.076
FARIMA(0, d , 0)	$\sqrt{MSE} \hat{d}_{MS}$	0.036	0.030	0.031	0.035	0.032
	$\sqrt{MSE} \hat{d}_R$	0.031	0.028	0.027	0.029	0.029
	$\sqrt{MSE} \hat{d}_W$	0.020*	0.018*	0.023	0.025	0.028*
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.066	0.031	0.018*	0.020*	0.028*
	$\sqrt{MSE} \hat{d}_N(10)$	0.076	0.047	0.043	0.053	0.038
	$\sqrt{MSE} \hat{d}_N(30)$	0.074	0.085	0.073	0.086	0.073
FARIMA(1, d , 1)	$\sqrt{MSE} \hat{d}_{MS}$	0.035	0.033	0.032	0.036	0.031
	$\sqrt{MSE} \hat{d}_R$	0.031*	0.029*	0.030*	0.032*	0.027*
	$\sqrt{MSE} \hat{d}_W$	0.054	0.054	0.050	0.052	0.048
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.099	0.066	0.052	0.047	0.046
	$\sqrt{MSE} \hat{d}_N(10)$	0.141	0.095	0.075	0.055	0.051
	$\sqrt{MSE} \hat{d}_N(30)$	0.111	0.085	0.094	0.090	0.074
$X^{(D, D')}, D' = 1$	$\sqrt{MSE} \hat{d}_{MS}$	0.029	0.037*	0.035*	0.041*	0.038*
	$\sqrt{MSE} \hat{d}_R$	0.032	0.041	0.037	0.041*	0.039
	$\sqrt{MSE} \hat{d}_W$	0.110	0.115	0.115	0.112	0.114
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.018*	0.064	0.092	0.084	0.081
	$\sqrt{MSE} \hat{d}_N(10)$	0.035	0.093	0.102	0.106	0.094
	$\sqrt{MSE} \hat{d}_N(30)$	0.039	0.088	0.084	0.074	0.077

TABLE 3.2 – Comparison of the different log-memory parameter estimators for processes of the benchmark. For each process and value of d and N , $\sqrt{MSE}$ are computed from 100 independent generated samples.

$N = 10^3 \longrightarrow$						
Model+Innovation	$\sqrt{MSE}$	$d = -0.4$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$
FARIMA(0, d , 0) Uniform	$\sqrt{MSE} \hat{d}_{MS}$	0.189	0.090	0.091	0.082*	0.092
	$\sqrt{MSE} \hat{d}_R$	0.171	0.104	0.109	0.102	0.086*
	$\sqrt{MSE} \hat{d}_W$	0.111*	0.066*	0.072*	0.118	0.129
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.186	0.081	0.083	0.112	0.093
FARIMA(0, d , 0) Burr ($\alpha = 2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.174	0.087	0.092	0.084	0.091*
	$\sqrt{MSE} \hat{d}_R$	0.183	0.104	0.097	0.107	0.079
	$\sqrt{MSE} \hat{d}_W$	0.149*	0.086*	0.130	0.101	0.129
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.221	0.119	0.076*	0.082*	0.139
FARIMA(0, d , 0) Burr ($\alpha = 3/2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.188	0.087*	0.063*	0.099*	0.075
	$\sqrt{MSE} \hat{d}_R$	0.183*	0.110	0.079	0.125	0.072*
	$\sqrt{MSE} \hat{d}_W$	0.219	0.108	0.138	0.146	0.159
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.264	0.134	0.094	0.155	0.187
GARMA(0, d , 0)	$\sqrt{MSE} \hat{d}_{MS}$	0.149	0.109	0.086	0.130	0.172
	$\sqrt{MSE} \hat{d}_R$	0.098*	0.104	0.090	0.132	0.125*
	$\sqrt{MSE} \hat{d}_W$	0.117	0.074*	0.081*	0.182	0.314
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.124	0.121	0.110	0.102*	0.331
Trend	$\sqrt{MSE} \hat{d}_{MS}$	1.307	0.891	0.538	0.290	0.150
	$\sqrt{MSE} \hat{d}_R$	0.900	0.700	0.498	0.275	0.087
	$\sqrt{MSE} \hat{d}_W$	0.222*	0.103*	0.083	0.071	0.059*
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	1.65	0.223	0.079*	0.050*	0.076
Trend + Seasonality	$\sqrt{MSE} \hat{d}_{MS}$	1.178	0.803	0.477	0.238	0.123
	$\sqrt{MSE} \hat{d}_R$	0.900	0.700	0.498	0.284	0.091*
	$\sqrt{MSE} \hat{d}_W$	0.628*	0.407*	0.318	0.274	0.283
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	1.54	1.01	0.311*	0.158*	0.145

$N = 10^4 \longrightarrow$						
Model+Innovation	$\sqrt{MSE}$	$d = -0.4$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$
FARIMA(0, d , 0) Uniform	$\sqrt{MSE} \hat{d}_{MS}$	0.177	0.039	0.033	0.034	0.034
	$\sqrt{MSE} \hat{d}_R$	0.171	0.032	0.030	0.028	0.032*
	$\sqrt{MSE} \hat{d}_W$	0.125*	0.027*	0.025	0.028	0.035
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.165	0.042	0.017*	0.027*	0.032*
FARIMA(0, d , 0) Burr ($\alpha = 2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.180	0.036	0.041	0.033	0.032
	$\sqrt{MSE} \hat{d}_R$	0.169	0.031*	0.030	0.031*	0.029*
	$\sqrt{MSE} \hat{d}_W$	0.138*	0.068	0.065	0.076	0.066
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.219	0.067	0.018*	0.039	0.074
FARIMA(0, d , 0) Burr ($\alpha = 3/2$)	$\sqrt{MSE} \hat{d}_{MS}$	0.18	0.038	0.026*	0.030	0.021*
	$\sqrt{MSE} \hat{d}_R$	0.174	0.033*	0.031	0.023*	0.023
	$\sqrt{MSE} \hat{d}_W$	0.126*	0.058	0.149	0.124	0.090
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.264	0.113	0.030	0.099	0.159
GARMA(0, d , 0)	$\sqrt{MSE} \hat{d}_{MS}$	0.063	0.041	0.028	0.032	0.060
	$\sqrt{MSE} \hat{d}_R$	0.037*	0.033*	0.025	0.026*	0.030*
	$\sqrt{MSE} \hat{d}_W$	0.061	0.052	0.021	0.078	0.081
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.074	0.040	0.016*	0.055	0.109
Trend	$\sqrt{MSE} \hat{d}_{MS}$	1.16	0.785	0.450	0.171	0.072
	$\sqrt{MSE} \hat{d}_R$	0.900	0.700	0.431	0.192	0.067
	$\sqrt{MSE} \hat{d}_W$	0.135	0.046	0.021*	0.019	0.021
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.019*	0.021*	0.021*	0.016*	0.020*
Trend + Seasonality	$\sqrt{MSE} \hat{d}_{MS}$	1.219	0.841	0.474	0.194	0.099
	$\sqrt{MSE} \hat{d}_R$	0.900	0.700	0.431	0.189	0.063
	$\sqrt{MSE} \hat{d}_W$	0.097*	0.073*	0.063	0.065	0.051
	$\sqrt{MSE} \hat{d}_N^{(IR)}$	0.671	0.382	0.049*	0.047*	0.041*

TABLE 3.3 – Comparison of the different long-memory parameter estimators for processes of the benchmark. For each process and value of d and N , $\sqrt{MSE}$ are computed from 100 independent generated samples.

Chapitre 4

New semiparametric stationarity tests based on adaptive MIR statistics

Sommaire

4.1	Introduction	116
4.2	The multidimensional increment ratio statistic	118
4.3	The adaptive version of the estimator	121
4.4	Stationarity and nonstationarity tests	123
4.4.1	A stationarity test	123
4.4.2	A new nonstationarity test	124
4.5	Results of simulations and application to Econometric and Financial data	125
4.5.1	Numerical procedure for computing the estimator and tests	125
4.5.2	Monte-Carlo experiments on several time series	125
4.5.3	Application to the the Stocks and the Exchange Rate Markets	127
4.6	Proofs	131

Abstract : In this paper, we show that the adaptive multidimensional increment ratio estimator of the long range memory parameter defined in Bardet and Dola (2012, [6]) satisfies a central limit theorem (CLT in the sequel) for a large semiparametric class of Gaussian fractionally integrated processes with memory parameter $d \in (-0.5, 1.25)$. Since the asymptotic variance of this CLT can be computed, tests of stationarity or nonstationarity distinguishing the assumptions $d < 0.5$ and $d \geq 0.5$ cases are constructed. These tests are also consistent tests of unit root. Simulations done on a large benchmark of short memory, long memory and non stationary processes show the accuracy of this test with respect to other usual stationarity or nonstationarity tests (KPSS, LMC, ADF and PP tests). Finally, the estimator and tests are applied to log-returns of famous financial data inherent to stocks or indexes markets and to their absolute value power laws.

Keywords : Gaussian fractionally integrated processes ; semiparametric estimators of the memory parameter ; test of long-memory ; stationarity test ; unit root test.

4.1 Introduction

Consider the set $I(d)$ of fractionally integrated time series $X = (X_k)_{k \in \mathbb{Z}}$ for $-0.5 < d < 1.5$ by :

Assumption $I(d)$: $X = (X_t)_{t \in \mathbb{Z}}$ is a time series if there exists a continuous function $f^* : [-\pi, \pi] \rightarrow [0, \infty[$ satisfying :

1. if $-0.5 < d < 0.5$, X is a stationary process having a spectral density f satisfying

$$f(\lambda) = |\lambda|^{-2d} f^*(\lambda) \quad \text{for all } \lambda \in (-\pi, 0) \cup (0, \pi), \text{ with } f^*(0) > 0. \quad (4.1)$$

2. if $0.5 \leq d < 1.5$, $U = (U_t)_{t \in \mathbb{Z}} = X_t - X_{t-1}$ is a stationary process having a spectral density f satisfying

$$f(\lambda) = |\lambda|^{2-2d} f^*(\lambda) \quad \text{for all } \lambda \in (-\pi, 0) \cup (0, \pi), \text{ with } f^*(0) > 0. \quad (4.2)$$

The case $d \in (0, 0.5)$ is the case of long-memory processes, while short-memory processes are considered when $-0.5 < d \leq 0$ and nonstationary processes when $d \geq 0.5$. ARFIMA(p,d,q) processes (which are linear processes) or fractional Gaussian noises (with parameter $H = d + 1/2 \in (0, 1)$) are famous examples of processes satisfying Assumption $I(d)$. The purpose of this paper is two-fold : firstly, we establish the consistency of an adaptive semiparametric estimator of d for any $d \in (-0.5, 1.25)$. Secondly, we use this estimator for building new semiparametric stationary tests. Numerous articles have been devoted to estimate d in the case $d \in (-0.5, 0.5)$. The books of Beran (1994, [12]) or Doukhan *et al.* (2003, [29]) provide large surveys of such parametric (mainly maximum likelihood or Whittle estimators) or semiparametric estimators (mainly local Whittle, log-periodogram or wavelet based estimators). Here we will restrict our discussion to the case of semiparametric estimators that are best suited to address the general case of processes satisfying Assumption $I(d)$. Even if first versions of local Whittle, log-periodogram and wavelet based estimators (see for instance Robinson, 1995a [82] and 1995b [81], Abry and Veitch, 1998 [2]) are only considered in the case $d < 0.5$, new extensions have been provided for also estimating d when $d \geq 0.5$ (see for instance Hurvich and Ray, 1995 [56], Velasco, 1999a [96], Velasco and Robinson, 2000 [84], Moulines and Soulier, 2003, Shimotsu and Phillips, 2005 [76], Giraitis *et al.*, 2003 [38], 2006 [39], Abadir *et al.* [1], 2007 or Moulines *et al.*, 2007 [71]). Moreover, adaptive versions of these estimators have also been defined for avoiding any trimming or bandwidth parameters generally required by these methods (see for instance Giraitis *et al.*, 2000, Moulines and Soulier, 2003 [70], or Veitch *et al.*, 2003 [95], or Bardet *et al.*, 2008 [10]). However there still no exists an adaptive estimator of d satisfying a central limit theorem (for providing confidence intervals or tests) and

valid for $d < 0.5$ but also for $d \geq 0.5$. This is the first objective of this paper and it will be achieved using multidimensional Increment Ratio (IR) statistics.

Indeed, Surgailis *et al.* (2008, [92]) first defined the statistic IR_N (see its definition in (4.5)) from an observed trajectory $(X_1, \dots, X_N)$. Its asymptotic behavior is studied and a central limit theorem (CLT in the sequel) is established for $d \in (-0.5, 0.5) \cup (0.5, 1.25)$ inducing a CLT. Therefore, the estimator $\hat{d}_N = \Lambda_0^{-1}(IR_N)$, where $d \mapsto \Lambda_0(d)$ is a smooth and increasing function, is a consistent estimator of d satisfying also a CLT (see more details below). However this new estimator was not totally satisfying : firstly, it requires the knowledge of the second order behavior of the spectral density that is clearly unknown in practice. Secondly, its numerical accuracy is interesting but clearly less than the one of local Whittle or log-periodogram estimators. As a consequence, in Bardet and Dola (2012, [6]), we built an adaptive multidimensional IR estimator $\tilde{d}_N^{(IR)}$ (see its definition in (4.15)) answering to both these points but only for $-0.5 < d < 0.5$. This is an adaptive semiparametric estimator of d and its numerical performances are often better than the ones of local Whittle or log-periodogram estimators.

Here we extend this preliminary work to the case $0.5 \leq d < 1.25$. Hence we obtain a CLT satisfied by $\tilde{d}_N^{(IR)}$ for all $d \in (-0.5, 1.25)$ with an explicit asymptotic variance depending only on d and this notably allows to obtain confidence intervals. The case $d = 0.5$ is now studied and this offers new interesting perspectives : our adaptive estimator can be used for building a stationarity (or nonstationarity) test since 0.5 is the “border number” between stationarity and nonstationarity.

There exist several famous stationarity (or nonstationarity) tests. For stationarity tests we may cite the KPSS (Kwiatkowski, Phillips, Schmidt, Shin) test (see for instance Hamilton, 1994 [49], p. 514) and LMC test (see Leybourne and McCabe, 2000). For nonstationarity tests we may cite the Augmented Dickey-Fuller test (ADF test in the sequel, see Hamilton, 1994 [49], p. 516-528) and the Philipps and Perron test (PP test in the sequel, see for instance Elder, 2001 [31], p. 137-146). All these tests are unit root tests, *i.e.* and roughly speaking, semiparametric tests based on the model $X_t = \rho X_{t-1} + \varepsilon_t$ with $|\rho| \leq 1$. A test about $d = 0.5$ for a process satisfying Assumption $I(d)$ is therefore a refinement of a basic unit root test since the case $\rho = 1$ is a particular case of $I(1)$ and the case $|\rho| < 1$ a particular case of $I(0)$. Thus, a stationarity (or nonstationarity test) based on the estimator of d provides a more sensible test than usual unit root tests.

This principle of stationarity test linked to d was also already investigated in many articles. We can notably cite Robinson (1994, [80]), Tanaka (1999, [93]), Ling and Li (2001, [67]), Ling (2003, [66]) or Nielsen (2004, [73]). However, all these papers provide parametric tests, with a specified model (for instance ARFIMA or ARFIMA-GARCH processes). More recently, several papers have been devoted to the construction of semi-parametric tests, see for instance Giraitis *et al.* (2006, [39]), Abadir *et al.* (2007, [1]) or Surgailis *et al.* (2010, [62]).

Here we slightly restrict the general class $I(d)$ to the Gaussian semiparametric class $IG(d, \beta)$ defined below (see the beginning of Section 4.2). For processes belonging to this class, we construct a

new stationarity test $\tilde{S}_N$ which accepts the stationarity assumption when $\tilde{d}_N^{(IR)} \leq 0.5 + s$ with s a threshold depending on the type I error test and N , while the new nonstationarity test $\tilde{T}_N$ accepts the nonstationarity assumption when $\tilde{d}_N^{(IR)} \geq 0.5 - s$. Note that $\tilde{d}_N^{(IR)} \leq s'$ also provides a test for deciding between short and long range dependency, as this is done by the V/S test (see details in Giraitis *et al.*, 2003, [38])

In Section 4.5, numerous simulations are realized on several models of time series (short and long memory processes).

First, the new multidimensional IR estimator $\tilde{d}_N^{(IR)}$ is compared to the most efficient and famous semiparametric estimators for $d \in [-0.4, 1.2]$; the performances of $\tilde{d}_N^{(IR)}$ are extremely convincing and globally better than the other estimators.

Secondly, the new stationarity $\tilde{S}_N$ and nonstationarity $\tilde{T}_N$ tests are compared on the same benchmark of processes to the most famous unit root tests (KPSS and LMC, ADF and PP tests). And the results are quite surprising : even on AR(1) or ARIMA(1,1,0) processes, multidimensional IR $\tilde{S}_N$ and $\tilde{T}_N$ tests provide better results than KPSS and LMC stationarity tests which are nevertheless constructed for such processes. Note however that ADF and PP tests provide results slightly better than $\tilde{S}_N$ and $\tilde{T}_N$ tests for these processes. For long-memory processes (such as ARFIMA processes), the results are clear : $\tilde{S}_N$ and $\tilde{T}_N$ tests are efficient tests of (non)stationarity while KPSS, LMC, ADF and PP tests are not relevant at all.

Finally, we studied the stationarity and long range dependency properties of Econometric data. We chose to apply estimators and tests to the log-returns of daily closing value of 5 classical Stocks and Exchange Rate Markets. After cutting the series in 3 stages using an algorithm of change detection, we found again this well known result : the log-returns are stationary and short memory processes while absolute values or powers of absolute values of log-returns are generally stationary and long memory processes. Classical stationarity or nonstationarity tests are not able to lead to such conclusions. We also remarked that these time series during the “last” (and third) stages (after 1997 for almost all) are generally closer to nonstationary processes than during the previous stages with a long memory parameter close to 0.5.

The forthcoming Section 4.2 is devoted to the definition and asymptotic behavior of the adaptive multidimensional IR estimator of d . The stationarity and nonstationarity tests are presented in Section 4.4 while Section 4.5 provides the results of simulations and application on econometric data. Finally Section 4.6 contains the proofs of main results.

4.2 The multidimensional increment ratio statistic

In this paper we consider a semiparametric class $IG(d, \beta)$: for $0 \leq d < 1.5$ and $\beta > 0$ define :

Assumption $IG(d, \beta)$: $X = (X_t)_{t \in \mathbb{Z}}$ is a Gaussian time series such that there exist $\epsilon > 0$, $c_0 > 0$, $c'_0 > 0$ and $c_1 \in \mathbb{R}$ satisfying :

1. if $d < 0.5$, X is a stationary process having a spectral density f satisfying for all $\lambda \in (-\pi, 0) \cup (0, \pi)$

$$f(\lambda) = c_0 |\lambda|^{-2d} + c_1 |\lambda|^{-2d+\beta} + O(|\lambda|^{-2d+\beta+\epsilon}) \quad \text{and} \quad |f'(\lambda)| \leq c'_0 |\lambda|^{-2d-1}. \quad (4.3)$$

2. if $0.5 \leq d < 1.5$, $U = (U_t)_{t \in \mathbb{Z}} = X_t - X_{t-1}$ is a stationary process having a spectral density f satisfying for all $\lambda \in (-\pi, 0) \cup (0, \pi)$

$$f(\lambda) = c_0 |\lambda|^{2-2d} + c_1 |\lambda|^{2-2d+\beta} + O(|\lambda|^{2-2d+\beta+\epsilon}) \quad \text{and} \quad |f'(\lambda)| \leq c'_0 |\lambda|^{-2d+1}. \quad (4.4)$$

Note that Assumption $IG(d, \beta)$ is a particular (but still general) case of the more usual set $I(d)$ of fractionally integrated processes defined above.

Remark 10. We considered here only Gaussian processes. In Surgailis *et al.* (2008) and Bardet and Dola (2012, [6]), simulations exhibited that the obtained limit theorems should be also valid for linear processes. However a theoretical proof of such result would require limit theorems for functionals of multidimensional linear processes difficult to be established.

In this section, under Assumption $IG(d, \beta)$, we establish central limit theorems which extend to the case $d \in [0.5, 1.25)$ those already obtained in Bardet and Dola (2012, [6]) for $d \in (-0.5, 0.5)$.

Let $X = (X_k)_{k \in \mathbb{N}}$ be a process satisfying Assumption $IG(d, \beta)$ and $(X_1, \dots, X_N)$ be a path of X . For any $\ell \in \mathbb{N}^*$ define

$$IR_N(\ell) := \frac{1}{N - 3\ell} \sum_{k=0}^{N-3\ell-1} \frac{\left| \left(\sum_{t=k+1}^{k+\ell} X_{t+\ell} - \sum_{t=k+1}^{k+\ell} X_t \right) + \left(\sum_{t=k+\ell+1}^{k+2\ell} X_{t+\ell} - \sum_{t=k+\ell+1}^{k+2\ell} X_t \right) \right|}{\left| \left(\sum_{t=k+1}^{k+\ell} X_{t+\ell} - \sum_{t=k+1}^{k+\ell} X_t \right) \right| + \left| \left(\sum_{t=k+\ell+1}^{k+2\ell} X_{t+\ell} - \sum_{t=k+\ell+1}^{k+2\ell} X_t \right) \right|}. \quad (4.5)$$

The statistic IR_N was first defined in Surgailis *et al.* (2008) as a way to estimate the memory parameter. In Bardet and Surgailis (2011) a simple version of IR-statistic was also introduced to measure the roughness of continuous time processes. The main interest of such a statistic is to be very robust to additional or multiplicative trends.

As in Bardet and Dola (2012, [6]), let $m_j = j m$, $j = 1, \dots, p$ with $p \in \mathbb{N}^*$ and $m \in \mathbb{N}^*$, and define the random vector $(IR_N(m_j))_{1 \leq j \leq p}$. In the sequel we naturally extend the results obtained for $m \in \mathbb{N}^*$ to $m \in (0, \infty)$ by the convention : $(IR_N(j m))_{1 \leq j \leq p} = (IR_N(j [m]))_{1 \leq j \leq p}$ (which changes nothing to the asymptotic results).

For $H \in (0, 1)$, let $B_H = (B_H(t))_{t \in \mathbb{R}}$ be a standard fractional Brownian motion, *i.e.* a centered Gaussian process having stationary increments and such as $\text{Cov}(B_H(t), B_H(s)) = \frac{1}{2} (|t|^{2H} +$

$|s|^{2H} - |t-s|^{2H}$). Now, using obvious modifications of Surgailis *et al.* (2008), for $d \in (-0.5, 1.25)$ and $p \in \mathbb{N}^*$, define the stationary multidimensional centered Gaussian processes $(Z_d^{(1)}(\tau), \dots, Z_d^{(p)}(\tau))$ such as for $\tau \in \mathbb{R}$,

$$Z_d^{(j)}(\tau) := \begin{cases} \frac{\sqrt{2d(2d+1)}}{\sqrt{|4^{d+0.5} - 4|}} \int_0^1 (B_{d-0.5}(\tau+s+j) - B_{d-0.5}(\tau+s)) ds & \text{if } d \in (0.5, 1.25) \\ \frac{1}{\sqrt{|4^{d+0.5} - 4|}} (B_{d+0.5}(\tau+2j) - 2B_{d+0.5}(\tau+j) + B_{d+0.5}(\tau)) & \text{if } d \in (-0.5, 0.5) \end{cases} \quad (4.6)$$

and by continuous extension when $d \rightarrow 0.5$:

$$\text{Cov}(Z_{0.5}^{(i)}(0), Z_{0.5}^{(j)}(\tau)) := \frac{1}{4 \log 2} (-h(\tau+i-j) + h(\tau+i) + h(\tau-j) - h(\tau)) \quad \text{for } \tau \in \mathbb{R},$$

with $h(x) = \frac{1}{2}(|x-1|^2 \log|x-1| + |x+1|^2 \log|x+1| - 2|x|^2 \log|x|)$ for $x \in \mathbb{R}$, using the convention $0 \times \log 0 = 0$. Now, we establish a multidimensional central limit theorem satisfied by $(IR_N(jm))_{1 \leq j \leq p}$ for all $d \in (-0.5, 1.25)$:

Proposition 4. *Assume that Assumption IG(d, β) holds with $-0.5 \leq d < 1.25$ and $\beta > 0$. Then*

$$\sqrt{\frac{N}{m}} \left(IR_N(jm) - \mathbb{E}[IR_N(jm)] \right)_{1 \leq j \leq p} \xrightarrow[N[m] \wedge m \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_p(d)) \quad (4.7)$$

with $\Gamma_p(d) = (\sigma_{i,j}(d))_{1 \leq i,j \leq p}$ where for $i, j \in \{1, \dots, p\}$,

$$\sigma_{i,j}(d) := \int_{-\infty}^{\infty} \text{Cov} \left(\frac{|Z_d^{(i)}(0) + Z_d^{(i)}(i)|}{|Z_d^{(i)}(0)| + |Z_d^{(i)}(i)|} \frac{|Z_d^{(j)}(\tau) + Z_d^{(j)}(\tau+j)|}{|Z_d^{(j)}(\tau)| + |Z_d^{(j)}(\tau+j)|} \right) d\tau. \quad (4.8)$$

The proof of this proposition as well as all the other proofs is given in Section 4.6. As numerical experiments seem to show, we will assume in the sequel that $\Gamma_p(d)$ is a definite positive matrix for all $d \in (-0.5, 1.25)$.

Now, this central limit theorem can be used for estimating d . To begin with,

Property 4. *Let X satisfying Assumption IG(d, β) with $0.5 \leq d < 1.5$ and $0 < \beta \leq 2$. Then, there exists a non-vanishing constant $K(d, \beta)$ depending only on d and β such that for m large enough,*

$$\mathbb{E}[IR_N(m)] = \begin{cases} \Lambda_0(d) + K(d, \beta) \times m^{-\beta} (1 + o(1)) & \text{if } \beta < 1 + 2d \\ \Lambda_0(d) + K(0.5, \beta) \times m^{-2} \log m (1 + o(1)) & \text{if } \beta = 2 \text{ and } d = 0.5 \end{cases}$$

$$\text{with } \Lambda_0(d) := \Lambda(\rho(d)) \quad \text{where } \rho(d) := \begin{cases} \frac{4^{d+1.5} - 9^{d+0.5} - 7}{2(4 - 4^{d+0.5})} & \text{for } 0.5 < d < 1.5 \\ \frac{9 \log(3)}{8 \log(2)} - 2 & \text{for } d = 0.5 \end{cases} \quad (4.9)$$

$$\text{and } \Lambda(r) := \frac{2}{\pi} \arctan \sqrt{\frac{1+r}{1-r}} + \frac{1}{\pi} \sqrt{\frac{1+r}{1-r}} \log\left(\frac{2}{1+r}\right) \quad \text{for } |r| \leq 1. \quad (4.10)$$

Therefore by choosing m and N such as $(\sqrt{N/m})m^{-\beta} \log m \rightarrow 0$ when $m, N \rightarrow \infty$, the term $\mathbb{E}[IR(jm)]$ can be replaced by $\Lambda_0(d)$ in Proposition 4. Then, using the Delta-method with the function $(x_i)_{1 \leq i \leq p} \mapsto (\Lambda_0^{-1}(x_i))_{1 \leq i \leq p}$ (the function $d \in (-0.5, 1.5) \rightarrow \Lambda_0(d)$ is a $\mathcal{C}^\infty$ increasing function), we obtain :

Theorem 3. *Let $\hat{d}_N(jm) := \Lambda_0^{-1}(IR_N(jm))$ for $1 \leq j \leq p$. Assume that Assumption IG(d, β) holds with $0.5 \leq d < 1.25$ and $0 < \beta \leq 2$. Then if $m \sim CN^\alpha$ with $C > 0$ and $(1 + 2\beta)^{-1} < \alpha < 1$,*

$$\sqrt{\frac{N}{m}} \left(\hat{d}_N(jm) - d \right)_{1 \leq j \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left(0, (\Lambda'_0(d))^{-2} \Gamma_p(d) \right). \quad (4.11)$$

This result is an extension to the case $0.5 \leq d \leq 1.25$ from the case $-0.5 < d < 0.5$ already obtained in Bardet and Dola (2012, [6]). Note that the consistency of $\hat{d}_N(jm)$ is ensured when $1.25 \leq d < 1.5$ but the previous CLT does not hold (the asymptotic variance of $\sqrt{\frac{N}{m}} \hat{d}_N(jm)$ diverges to ∞ when $d \rightarrow 1.25$ (see Surgailis *et al.*, [92], 2008).

Now define

$$\hat{\Sigma}_N(m) := (\Lambda'_0(\hat{d}_N(m)))^{-2} \Gamma_p(\hat{d}_N(m)). \quad (4.12)$$

The function $d \in (-0.5, 1.5) \mapsto \sigma(d)/\Lambda'(d)$ is $\mathcal{C}^\infty$ and therefore, under assumptions of Theorem 3,

$$\hat{\Sigma}_N(m) \xrightarrow[N \rightarrow \infty]{\mathcal{P}} (\Lambda'_0(d))^{-2} \Gamma_p(d).$$

Thus, a pseudo-generalized least square estimation (LSE) of d can be defined by

$$\tilde{d}_N(m) := (J_p^\top (\hat{\Sigma}_N(m))^{-1} J_p)^{-1} J_p^\top (\hat{\Sigma}_N(m))^{-1} (\hat{d}_N(m_i))_{1 \leq i \leq p}$$

with $J_p := (1)_{1 \leq j \leq p}$ and denoting $J_p^\top$ its transpose. From Gauss-Markov Theorem, the asymptotic variance of $\tilde{d}_N(m)$ is smaller than the one of $\hat{d}_N(jm)$, $j = 1, \dots, p$. Hence, we obtain under the assumptions of Theorem 3 :

$$\sqrt{\frac{N}{m}} (\tilde{d}_N(m) - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left(0, \Lambda'_0(d)^{-2} (J_p^\top \Gamma_p^{-1}(d) J_p)^{-1} \right). \quad (4.13)$$

4.3 The adaptive version of the estimator

Theorem 3 and CLT (4.13) require the knowledge of β to be applied. But in practice β is unknown. The procedure defined in Bardet *et al.* (2008) or Bardet and Dola (2012, [6]) can be used for obtaining a data-driven selection of an optimal sequence $(\tilde{m}_N)$ derived from an estimation of β . Since the case $d \in (-0.5, 0.5)$ was studied in Bardet and Dola (2012, [6]) we consider here $d \in [0.5, 1.25]$ and for $\alpha \in (0, 1)$, define

$$Q_N(\alpha, d) := (\hat{d}_N(jN^\alpha) - \tilde{d}_N(N^\alpha))_{1 \leq j \leq p}^\top (\hat{\Sigma}_N(N^\alpha))^{-1} (\hat{d}_N(jN^\alpha) - \tilde{d}_N(N^\alpha))_{1 \leq j \leq p}, \quad (4.14)$$

which corresponds to the sum of the pseudo-generalized squared distance between the points $(\widehat{d}_N(j N^\alpha))_j$ and PGLS estimate of d . Note that by the previous convention, $\widehat{d}_N(j N^\alpha) = \widehat{d}_N(j [N^\alpha])$ and $\widetilde{d}_N(N^\alpha) = \widetilde{d}_N([N^\alpha])$. Then $\widehat{Q}_N(\alpha)$ can be minimized on a discretization of $(0, 1)$ and define :

$$\widehat{\alpha}_N := \operatorname{Argmin}_{\alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha) \quad \text{with} \quad \mathcal{A}_N = \left\{ \frac{2}{\log N}, \frac{3}{\log N}, \dots, \frac{\log[N/p]}{\log N} \right\}.$$

Remark 11. *The choice of the set of discretization $\mathcal{A}_N$ is implied by our proof of convergence of $\widehat{\alpha}_N$ to α^* . If the interval $(0, 1)$ is stepped in N^c points, with $c > 0$, the used proof cannot attest this convergence. However $\log N$ may be replaced in the previous expression of $\mathcal{A}_N$ by any negligible function of N compared to functions N^c with $c > 0$ (for instance, $(\log N)^a$ or $a \log N$ with $a > 0$ can be used).*

From the central limit theorem (4.11) one deduces the following limit theorem :

Proposition 5. *Assume that Assumption IG(d, β) holds with $0.5 \leq d < 1.25$ and $0 < \beta \leq 2$. Then,*

$$\widehat{\alpha}_N \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \alpha^* = \frac{1}{(1 + 2\beta)}.$$

Finally define

$$\widetilde{m}_N := N^{\widetilde{\alpha}_N} \quad \text{with} \quad \widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6 \widehat{\alpha}_N}{(p-2)(1 - \widehat{\alpha}_N)} \cdot \frac{\log \log N}{\log N}.$$

and the estimator

$$\widetilde{d}_N^{(IR)} := \widetilde{d}_N(\widetilde{m}_N) = \widetilde{d}_N(N^{\widetilde{\alpha}_N}). \quad (4.15)$$

(the definition and use of $\widetilde{\alpha}_N$ instead of $\widehat{\alpha}_N$ are explained just before Theorem 2 in Bardet and Dola, 2012). The following theorem provides the asymptotic behavior of the estimator $\widetilde{d}_N^{(IR)}$:

Theorem 4. *Under assumptions of Proposition 5,*

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}} (\widetilde{d}_N^{(IR)} - d) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0; \Lambda'_0(d)^{-2} (J_p^\top \Gamma_p^{-1}(d) J_p)^{-1}\right). \quad (4.16)$$

$$\text{Moreover, } \forall \rho > \frac{2(1 + 3\beta)}{(p-2)\beta}, \quad \frac{N^{\frac{\beta}{1+2\beta}}}{(\log N)^\rho} \cdot |\widetilde{d}_N^{(IR)} - d| \xrightarrow[N \rightarrow \infty]{\mathcal{P}} 0.$$

The convergence rate of $\widetilde{d}_N^{(IR)}$ is the same (up to a multiplicative logarithm factor) than the one of minimax estimator of d in this semiparametric frame (see Giraitis *et al.*, 1997). The supplementary advantage of $\widetilde{d}_N^{(IR)}$ with respect to other adaptive estimators of d (see for instance Moulines and Soulier, 2003, for an overview about frequency domain estimators of d) is the central limit theorem (4.16) satisfied by $\widetilde{d}_N^{(IR)}$. Moreover $\widetilde{d}_N^{(IR)}$ can be used for $d \in (-0.5, 1.25)$, *i.e.* as well for stationary and non-stationary processes, without modifications in its definition. Both these advantages allow to define stationarity and nonstationarity tests based on $\widetilde{d}_N^{(IR)}$.

4.4 Stationarity and nonstationarity tests

Assume that $(X_1, \dots, X_N)$ is an observed trajectory of a process $X = (X_k)_{k \in \mathbb{Z}}$. We define here new stationarity and nonstationarity tests for X based on $\tilde{d}_N^{(IR)}$.

4.4.1 A stationarity test

There exist many stationarity and nonstationarity test. The most famous stationarity tests are certainly the following unit root tests :

- The KPSS (Kwiatkowski, Phillips, Schmidt, Shin) test (see for instance Hamilton, 1994, p. 514) ;
- The LMC (Leybourne, McCabe) test which is a generalization of the KPSS test (see for instance Leybourne and McCabe, 1994 and 1999)

Here, we consider the following problem of test :

- Hypothesis H_0 (stationarity) : $(X_t)_{t \in \mathbb{Z}}$ is a process satisfying Assumption $IG(d, \beta)$ with $d \in [-a_0, a'_0]$ where $0 \leq a_0, a'_0 < 1/2$ and $\beta \in [b_0, 2]$ where $0 < b_0 \leq 2$.
- Hypothesis H_1 (nonstationarity) : $(X_t)_{t \in \mathbb{Z}}$ is a process satisfying Assumption $IG(d, \beta)$ with $d \in [0.5, a_1]$ where $0 \leq a_1 < 1.25$ and $\beta \in [b_1, 2]$ where $0 < b_1 \leq 2$.

We use a test based on $\tilde{d}_N^{(IR)}$ for deciding between these hypothesis. Hence from the previous CLT 4.16 and with a type I error α , define

$$\tilde{S}_N := \tilde{d}_N^{(IR)} > 0.5 + \sigma_p(0.5) q_{1-\alpha} N^{(\tilde{\alpha}_N - 1)/2}, \quad (4.17)$$

where $\sigma_p(0.5) = \left(\Lambda'_0(0.5)^{-2} (J_p^T \Gamma_p^{-1}(0.5) J_p)^{-1} \right)^{1/2}$ (see (4.16)) and $q_{1-\alpha}$ is the $(1 - \alpha)$ quantile of a standard Gaussian random variable $\mathcal{N}(0, 1)$.

Then we define the following rules of decision :

- H_0 (stationarity) is accepted when $\tilde{S}_N = 0$ and rejected when $\tilde{S}_N = 1$.

Remark 12. In fact, the previous stationarity test $\tilde{S}_N$ defined in (4.17) can also be seen as a semi-parametric test $d < d_0$ versus $d \geq d_0$ with $d_0 = 0.5$. It is obviously possible to extend it to any value $d_0 \in (-0.5, 1.25)$ by defining $\tilde{S}_N^{(d_0)} := \tilde{d}_N^{(IR)} > d_0 + \sigma_p(d_0) q_{1-\alpha} N^{(\tilde{\alpha}_N - 1)/2}$.

From previous results, it is clear that :

Property 5. Under Hypothesis H_0 , the asymptotic type I error of the test $\tilde{S}_N$ is α and under Hypothesis H_1 , the test power tends to 1.

Moreover, this test can be used as a unit root test. Indeed, define the following typical problem of unit root test. Let $X_t = at + b + \varepsilon_t$, with $(a, b) \in \mathbb{R}^2$, and ε_t an ARIMA(p, d, q) with $d = 0$ or $d = 1$. Then, a (simplified) problem of a unit root test is to decide between :

- H_0^{UR} : $d = 0$ and (ε_t) is a stationary ARMA(p, q) process.
- H_1^{UR} : $d = 1$ and $(\varepsilon_t - \varepsilon_{t-1})_t$ is a stationary ARMA(p, q) process.

Then,

Property 6. *Under assumption H_0^{UR} , the type I error of this unit root test problem using $\tilde{S}_N$ decreases to 0 when $N \rightarrow \infty$ and the test power tends to 1.*

4.4.2 A new nonstationarity test

Famous unit root tests are more often nonstationarity test. For instance, between the most famous tests,

- The Augmented Dickey-Fuller test (see Hamilton, 1994, p. 516-528 for details);
- The Philipps and Perron test (a generalization of the ADF test with more lags) (see for instance Elder, 2001, p. 137-146).

Using the statistic $\tilde{d}_N^{(IR)}$ we propose a new nonstationarity test $\tilde{T}_N$ for deciding between :

- Hypothesis H'_0 (nonstationarity) : $(X_t)_{t \in \mathbb{Z}}$ is a process satisfying Assumption $IG(d, \beta)$ with $d \in [0.5, a'_0]$ where $0.5 \leq a'_0 < 1.25$ and $\beta \in [b'_0, 2]$ where $0 < b'_0 \leq 2$.
- Hypothesis H'_1 (stationarity) : $(X_t)_{t \in \mathbb{Z}}$ is a process satisfying Assumption $IG(d, \beta)$ with $d \in [-a'_1, b'_1]$ where $0 \leq a'_1, b'_1 < 1/2$ and $\beta \in [c'_1, 2]$ where $0 < c'_1 \leq 2$.

Then, the rule of the test is the following : Hypothesis H'_0 is accepted when $\tilde{T}_N = 1$ and rejected when $\tilde{T}_N = 0$ where

$$\tilde{T}_N := \tilde{d}_N^{(IR)} < 0.5 - \sigma_p(0.5) q_{1-\alpha} N^{(\tilde{\alpha}_N - 1)/2} . \quad (4.18)$$

Then as previously

Property 7. *Under Hypothesis H'_0 , the asymptotic type I error of the test $\tilde{T}_N$ is α and under Hypothesis H'_1 the test power tends to 1.*

As previously, this test can also be used as a unit root test where $X_t = at + b + \varepsilon_t$, with $(a, b) \in \mathbb{R}^2$, and ε_t an ARIMA(p, d, q) with $d = 0$ or $d = 1$. We consider here a “second” simplified problem of unit root test which is to decide between :

- $H_0^{UR'}$: $d = 1$ and $(\varepsilon_t - \varepsilon_{t-1})_t$ is a stationary ARMA(p, q) process.
- $H_1^{UR'}$: $d = 0$ and $(\varepsilon_t)_t$ is a stationary ARMA(p, q) process..

Then,

Property 8. *Under assumption $H_0^{UR'}$, the type I error of the unit root test problem using $\tilde{T}_N$ decreases to 0 when $N \rightarrow \infty$ and the test power tends to 1.*

4.5 Results of simulations and application to Econometric and Financial data

4.5.1 Numerical procedure for computing the estimator and tests

First of all, softwares used in this Section are available on <http://samm.univ-paris1.fr/-Jean-Marc-Bardet> with a free access on (in Matlab language) as well as classical estimators or tests.

The concrete procedure for applying our MIR-test of stationarity is the following :

1. using additional simulations (realized on ARMA, ARFIMA, FGN processes and not presented here for avoiding too much expansions), we have observed that the value of the parameter p is not really important with respect to the accuracy of the test (less than 10% on the value of $\tilde{d}_N^{(IR)}$). However, for optimizing our procedure we chose p as a stepwise function of n :

$$p = 5 \times \{n < 120\} + 10 \times \{120 \leq n < 800\} + 15 \times \{800 \leq n < 10000\} + 20 \times \{n \geq 10000\}$$

and $\sigma_5(0.5) \simeq 0.9082$; $\sigma_{10}(0.5) \simeq 0.8289$; $\sigma_{15}(0.5) \simeq 0.8016$ and $\sigma_{20}(0.5) \simeq 0.7861$.

2. then using the computation of $\tilde{m}_N$ presented in Section 4.3, the adaptive estimator $\tilde{d}_N^{(IR)}$ (defined in (4.15)) and the test statistics $\tilde{S}_N$ (defined in (4.17)) and $\tilde{T}_N$ (defined in (4.18)) are computed.

4.5.2 Monte-Carlo experiments on several time series

In the sequel the results are obtained from 300 generated independent samples of each process defined below. The concrete procedures of generation of these processes are obtained from the circulant matrix method, as detailed in Doukhan *et al.* (2003). The simulations are realized for different values of d and N and processes which satisfy Assumption $IG(d, \beta)$:

1. the usual ARIMA(p, d, q) processes with respectively $d = 0$ or $d = 1$ and an innovation process which is a Gaussian white noise. Such processes satisfy Assumption $IG(0, 2)$ or $IG(1, 2)$ holds (respectively) ;
2. the ARFIMA(p, d, q) processes with parameter d such that $d \in (-0.5, 1.25)$ and an innovation process which is a Gaussian white noise. Such ARFIMA(p, d, q) processes satisfy Assumption $IG(d, 2)$ (note that ARIMA processes are particular cases of ARFIMA processes).
3. the Gaussian stationary processes $X^{(d, \beta)}$, such as its spectral density is

$$f_3(\lambda) = \frac{1}{|\lambda|^{2d}}(1 + c_1 |\lambda|^\beta) \quad \text{for } \lambda \in [-\pi, 0) \cup (0, \pi], \quad (4.19)$$

with $d \in (-0.5, 1.5)$, $c_1 > 0$ and $\beta \in (0, \infty)$. Therefore the spectral density f_3 implies that Assumption $IG(d, \beta)$ holds. In the sequel we will use $c_1 = 5$ and $\beta = 0.5$, implying that the second order term of the spectral density is less negligible than in case of FARIMA processes.

Comparison of $\tilde{d}_N^{(IR)}$ with other semiparametric estimators of d

Here we first compare the performance of the adaptive MIR-estimator $\tilde{d}_N^{(IR)}$ with other famous semiparametric estimators of d :

- $\tilde{d}_N^{(MS)}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier (2003), also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$. Such an estimator was shown to be consistent for $d \in]-0.5, 1.25]$.
- $\tilde{d}_N^{(ADG)}$ is the extended local Whittle estimator defined by Abadir, Distaso and Giraitis (2007) which is consistent for $d > -3/2$. It is a generalization of the local Whittle estimator introduced by Robinson (1995b), consistent for $d < 0.75$, following a first extension proposed by Phillips (1999) and Shimotsu and Phillips (2005). This estimator avoids the tapering used for instance in Velasco (1999b) or Hurvich and Chen (2000). The trimming parameter is chosen as $m = N^{0.65}$ (this is not an adaptive estimator) following the numerical recommendations of Abadir *et al.* (2007).
- $\tilde{d}_N^{(WAV)}$ is an adaptive wavelet based estimator introduced in Bardet *et al.* ([8], 2013) using a Lemarie-Meyer type wavelet (another similar choice could be the adaptive wavelet estimator introduced in Veitch *et al.*, 2003, using a Daubechie's wavelet, but its robustness property are slightly less interesting). The asymptotic normality of such estimator is established for $d \in \mathbb{R}$ (when the number of vanishing moments of the wavelet function is large enough).

Table 4.1, 4.2, 4.3 and 4.4 respectively provide the results of simulations for ARIMA(1, d , 0), ARFIMA(0, d , 0), ARFIMA(1, d , 1) and $X^{(d, \beta)}$ processes for several values of d (or other parameters ϕ or θ) and N .

Conclusions of simulations : In more than 50% of cases, the estimator $\tilde{d}_N^{(IR)}$ provides the smallest $\sqrt{MSE}$ among the 4 semiparametric estimators. The good performances of $\tilde{d}_N^{(IR)}$ with respect to the other estimators are particularly clear for $d \in [0.5, 1.20]$. However, we can remark that the estimator $\tilde{d}_N^{(IR)}$ is slightly less efficient for ARFIMA(1, d , 1) processes and especially $X^{(d, \beta)}$ processes, that is to say, when the spectral density is slightly less smooth (*i.e.* close to a power law function).

Comparison of MIR tests $\tilde{S}_N$ and $\tilde{T}_N$ with other famous stationarity or nonstationarity tests

Monte-Carlo experiments were done for evaluating the performances of new tests $\tilde{S}_N$ and $\tilde{T}_N$ and for comparing them to most famous stationarity (KPSS and LMC) or nonstationarity (ADF and PP) tests (see more details on these tests in the previous section). For these tests we selected a classical "optimal" lag parameter k such as :

- $k = \left\lceil \frac{3}{13} \sqrt{n} \right\rceil$ for KPSS test ;
- $k = 0$ for LMC test ;
- $k = \left\lceil (n-1)^{1/3} \right\rceil$ for ADF test ;

– $k = \left\lceil 4 \left(\frac{n}{100} \right)^{1/4} \right\rceil$ for PP test ;

We also computed the results of V/S test, which is a nonparametric long-memory test (see more details Giraitis *et al.*, 2007). This test is useful when $d < 0.5$ but it can not really distinguish the stationary case $d \in (0, 0.5)$ to the nonstationary case $d \geq 0.5$.

The results of these simulations with a type I error classically chosen to 0.05 are provided in Tables 4.5, 4.6, 4.7 and 4.8.

Conclusions of simulations : For ARIMA($p, d, 0$) processes with $d = 0$ or $d = 1$ (notably AR(1) process when $d = 0$), ADF and PP tests are more accurate than our adaptive MIR tests. This is not really a surprise since those classical tests are unit root tests and are built from ARIMA processes. What is more surprising is that our MIR tests provide better results than KPSS and LMC tests (for LMC only in the case $d = 1$).

In case of ARFIMA or $X^{(d, \beta)}$ processes which are stationary long-memory processes when $0 < d < 0.5$ and nonstationary processes when $d \geq 0.5$, all the classical tests are not able to distinguish between the cases $d = 0$ and $d = 1$. Hence, if $0 < d < 1$, the asymptotic behaviors of KPSS or LMC tests are the same than for $d = 0$: they accept the stationarity assumption H_0 . In contrast, ADF and PP tests accept the nonstationarity assumption H'_0 even if the processes are stationary when $0 < d < 0.5$. This is not really a problem with the V/S test since this test is built for distinguishing between short and long memory processes. Conversely, the adaptive MIR tests (especially nonstationarity test $\tilde{T}_N$) which are based on the estimation of the memory parameter are very efficient for distinguishing between the stationary case $d < 0.5$ and the nonstationary case $d \geq 0.5$.

Finally, these simulations show that the adaptive MIR tests (especially nonstationarity test $\tilde{T}_N$) are the best trade-off for detecting the stationarity/nonstationarity of a process.

4.5.3 Application to the the Stocks and the Exchange Rate Markets

We applied the adaptive MIR statistics as well as the other famous long-memory estimators and stationarity tests to Econometric data, the Stocks and Exchange Rate Markets. More precisely, the 5 following daily closing value time series are considered :

1. The USA Dollar Exchange rate in Deusch-Mark, from 11/10/1983 to 08/04/2011 (7174 obs.).
2. The USA Dow Jones Transportation Index, from 31/12/1964 to 08/04/2011 (12072 obs.).
3. The USA Dow Jones Utilities Index, from 31/12/1964 to 08/04/2011 (12072 obs.).
4. The USA Nasdaq Industrials Index, from 05/02/1971 to 08/04/2011 (10481 obs.).
5. The Japan Nikkei225A Index, from 03/04/1950 to 8/04/2011 (15920 obs.).

We considered the log-return of this data and tried to test their stationarity properties. Since stationarity or nonstationarity tests are not able to detect (offline) changes, we first used an algorithm developed by M. Lavielle for detecting changes (this free software can be downloaded from

$N = 500$ ARIMA(1, d , 0)	$d = 0$ $\phi = -0.5$	$d = 0$ $\phi = -0.7$	$d = 0$ $\phi = -0.9$	$d = 1$ $\phi = -0.1$	$d = 1$ $\phi = -0.3$	$d = 1$ $\phi = -0.5$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.163	0.265	0.640	0.093	0.102	0.109
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.138	0.148	0.412	0.172	0.163	0.170
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.125	0.269	0.679	0.074	0.078	0.120
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.246	0.411	0.758	0.067	0.099	0.133

$N = 5000$ ARFIMA(1, d , 0)	$d = 0$ $\phi = -0.5$	$d = 0$ $\phi = -0.7$	$d = 0$ $\phi = -0.9$	$d = 1$ $\phi = -0.1$	$d = 1$ $\phi = -0.3$	$d = 1$ $\phi = -0.5$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.077	0.106	0.293	0.027	0.048	0.062
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.045	0.050	0.230	0.046	0.046	0.040
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.043	0.085	0.379	0.031	0.032	0.036
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.080	0.103	0.210	0.037	0.044	0.054

TABLE 4.1 – : Comparison between $\tilde{d}_N^{(IR)}$ and other famous semiparametric estimators of d ($\tilde{d}_N^{(MS)}$, $\tilde{d}_N^{(ADG)}$ and $\tilde{d}_N^{(WAV)}$) applied to ARIMA(1, d , 0) process $((1 - B)^d(1 + \phi B)X = \varepsilon)$, with several ϕ and N values

his homepage : <http://www.math.u-psud.fr/~lavielle/programmes-lavielle.html>). This algorithm provides the choice of detecting changes in mean, in variance, ..., and we chose to detect parametric changes in the distribution. Note that the number of changes is also estimated since this algorithm is based on the minimization of a penalized contrast. We obtained for each time series an estimated number of changes equal to 2 which are the following :

- Two breaks points for the US dollar-Deutsch Mark Exchange rate return are estimated, corresponding to the dates : 21/08/2006 and 24/12/2007. The Financial crisis of 2007-2011, followed by the late 2000 recession and the 2010 European sovereign debt crisis can cause such breaks.
- Both the breaks points estimated for the US Dow Jones Transportation Index return, of the New-York Stock Market, correspond to the dates : 17/11/1969 and 15/09/1997. The first break change can be a consequence on transportation companies difficulties the American Viet-Nam war against communist block. The second change point can be viewed as a contagion by the spread of the Thai crisis in 1997 to other countries and mainly the US stock Market.
- Both the breaks points estimated for the US Dow Jones Utilities Index return correspond to the dates : 02/06/1969 and 14/07/1998. The same arguments as above can justify the first break. The second at 1998 is probably a consequence of “*the long very acute crisis in the bond markets,..., the dramatic fiscal crisis and Russian Flight to quality caused by it, may have been warning the largest known by the global financial system : we never went too close to a definitive breakdown of relations between the various financial instruments*”(Wikipedia).
- The two breaks points for the US Nasdaq Industrials Index return correspond to the dates : 17/07/1998 and 27/12/2002. The first break at 1998 is explained by the Russian flight to quality as above. The second break at 2002 corresponds to the Brazilian public debt crisis of 2002 toward

$N = 500$ ARFIMA(0,d,0)	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.088	0.092	0.097	0.096	0.101	0.101	0.099	0.105
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.144	0.134	0.146	0.152	0.168	0.175	0.165	0.157
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.075	0.078	0.080	0.084	0.083	0.079	0.077	0.081
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.071	0.079	0.087	0.088	0.087	0.085	0.069	0.076

$N = 5000$ ARFIMA(0,d,0)	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.037	0.025	0.031	0.031	0.035	0.035	0.038	0.049
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.043	0.042	0.043	0.042	0.055	0.054	0.046	0.147
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.034	0.033	0.032	0.036	0.033	0.032	0.033	0.032
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.033	0.032	0.031	0.023	0.023	0.038	0.039	0.041

TABLE 4.2 – : Comparison between $\tilde{d}_N^{(IR)}$ and other famous semiparametric estimators of d ($\tilde{d}_N^{(MS)}$, $\tilde{d}_N^{(ADG)}$ and $\tilde{d}_N^{(WAV)}$) applied to ARFIMA(0, d , 0) process, with several d and N values

- foreign owners (mainly the U.S. and the IMF) which implicitly assigns a default of payment probability close to 100% with a direct impact on the financial markets indexes as the Nasdaq.
- Both the breaks points estimated for the Japanese Nikkei225A Index return corresponds to the dates 29/10/1975 and 12/02/1990, perhaps as consequence of the strong dependency of Japan to the middle east Oil following 1974 or anticipating 1990 oil crisis. The credit crunch which is seen as a major factor in the U.S. recession of 1990-91 can play a role in the second break point.

Data and estimated instant breaks can be seen on Figure 4.1. Then, we applied the estimators and tests described in the previous subsection on trajectories obtained in each stages for the 5 economic time series. These applications were done on the log-returns, their absolute values, their squared values and their θ -power laws with θ maximized for each LRD estimators. The results of these numerical experiments can be seen in Tables 4.9-4.13.

Conclusions of numerical experiments : We exhibited again the well known result : the log-returns are stationary and short memory processes while absolute values or power θ of log-returns are generally stationary but long memory processes (for this conclusion, we essentially consider the results of $\tilde{S}_N$, $\tilde{T}_N$ and V/S tests since the other tests have been shown not to be relevant in the cases of long-memory processes). However the last and third estimated stage of each time series provides generally the largest estimated values of the memory parameter d (for power law of log-returns) which are close to 0.5 ; hence, for Nasdaq time series, we accepted the nonstationarity assumption.

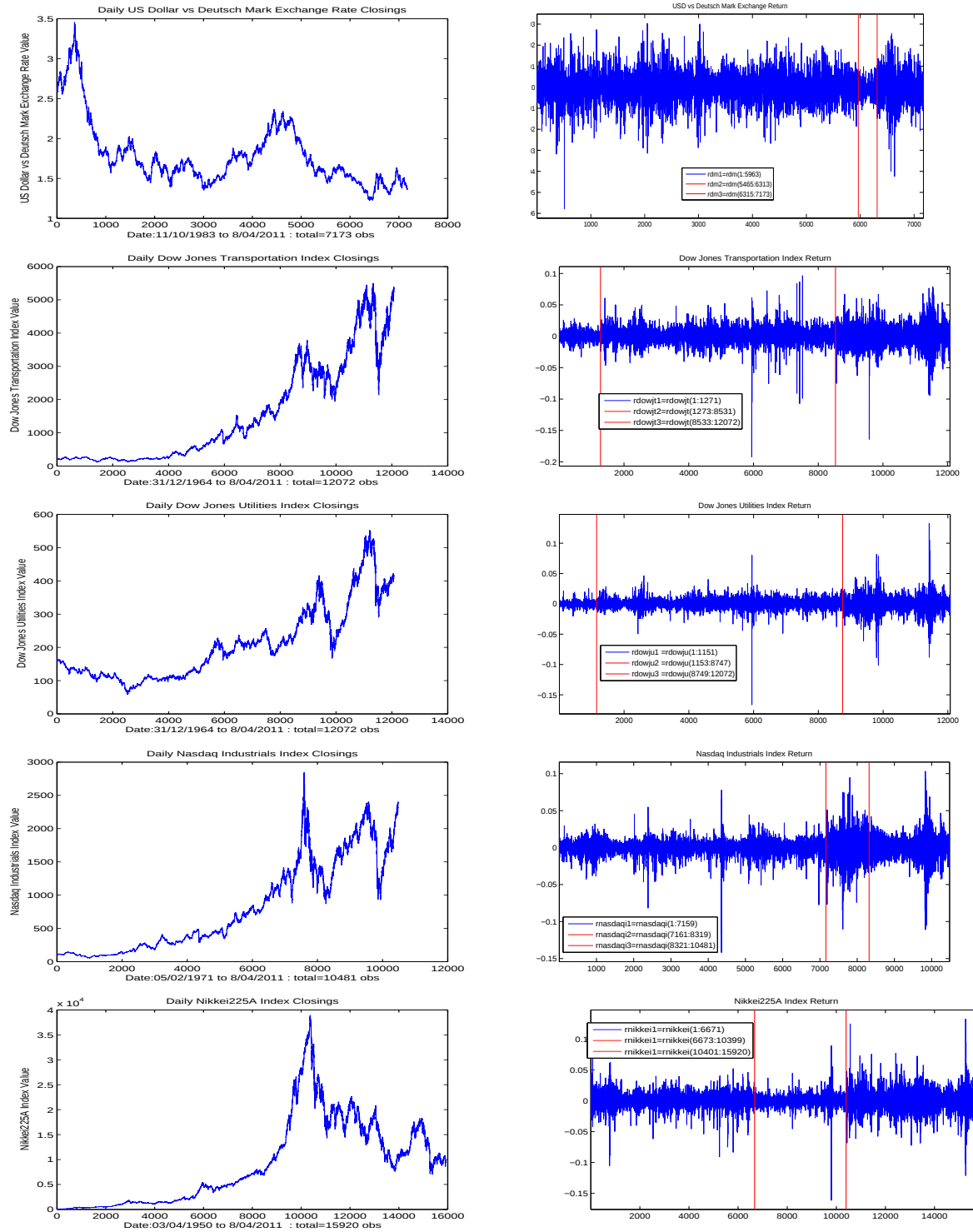


FIGURE 4.1 – The financial data (DowJonesTransportations, DowJonesUtilities, NasdaqIndustrials, Nikkei225A and US Dollar vs Deutsch Mark) : original data (left) and log-return with their both estimated breaks instants (right) occurred at the distribution Changes

$N = 500$ ARFIMA(1,d,1) $\phi = -0.3; \theta = 0.7$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.152	0.132	0.125	0.125	0.118	0.117	0.111	0.112
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.138	0.137	0.144	0.155	0.161	0.179	0.172	0.170
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.092	0.088	0.090	0.097	0.096	0.087	0.087	0.087
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.173	0.154	0.152	0.148	0.139	0.132	0.105	0.098

$N = 5000$ ARFIMA(1,d,1) $\phi = -0.3; \theta = 0.7$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.070	0.062	0.053	0.052	0.052	0.054	0.059	0.58
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.038	0.042	0.041	0.050	0.052	0.054	0.045	0.150
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.039	0.035	0.033	0.037	0.038	0.037	0.035	0.033
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.049	0.057	0.056	0.053	0.051	0.050	0.048	0.050

TABLE 4.3 – : Comparison between $\tilde{d}_N^{(IR)}$ and other famous semiparametric estimators of d ($\tilde{d}_N^{(MS)}$, $\tilde{d}_N^{(ADG)}$ and $\tilde{d}_N^{(WAV)}$) applied to ARFIMA(1, d , 1) process (with $\phi = -0.3$ and $\theta = 0.7$), with several d and N values.

4.6 Proofs

Proof of Proposition 4. This proposition is based on results of Surgailis *et al.* (2008) and was already proved in Bardet et Dola (2012) in the case $-0.5 < d < 0.5$.

Mutatis mutandis, the case $0.5 < d < 1.25$ can be treated exactly following the same steps.

The only new proof which has to be established concerns the case $d = 0.5$ since Surgailis *et al.* (2008) do not provide a CLT satisfied by the (unidimensional) statistic $IR_N(m)$ in this case. Let $Y_m(j)$ the standardized process defined Surgailis *et al.* (2008). Then, for $d = 0.5$,

$$\forall j \geq 1, \quad |\gamma_m(j)| = |\mathbb{E}(Y_m(j)Y_m(0))| = \frac{2}{V_m^2} \left| \int_0^\pi \cos(jx) x(c_0 + O(x^\beta)) \frac{\sin^4(\frac{mx}{2})}{\sin^4(\frac{x}{2})} dx \right|.$$

Denote $\gamma_m(j) = \rho_m(j) = \frac{2}{V_m^2} (I_1 + I_2)$ as in (5.39) of Surgailis *et al.* (2008). Both inequalities (5.41) and (5.42) remain true for $d = 0.5$ and

$$|I_1| \leq C \frac{m^3}{j}, \quad |I_2| \leq C \frac{m^4}{j^2} \quad \implies \quad |I_1 + I_2| \leq C \frac{m^3}{j} \quad \implies \quad |\gamma_m(j)| = |\rho_m(j)| \leq \frac{2}{V_m^2} (|I_1 + I_2|) \leq C \frac{m}{j}.$$

Now let $\eta_m(j) := \frac{|Y_m(j) + Y_m(j+m)|}{|Y_m(j)| + |Y_m(j+m)|} := \psi(Y_m(j), Y_m(j+m))$. The Hermite rank of the function ψ is 2 and therefore the equation (5.23) of Surgailis *et al.* (2008) obtained from Arcones Lemma remains valid. Hence :

$$|\text{Cov}(\eta_m(0), \eta_m(j))| \leq C \frac{m^2}{j^2}$$

$N = 500$ $X^{(d,\beta)}$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.110	0.170	0.201	0.211	0.209	0.205	0.210	0.202
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.187	0.188	0.204	0.200	0.192	0.187	0.200	0.192
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.177	0.182	0.190	0.184	0.174	0.179	0.196	0.189
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.224	0.225	0.230	0.220	0.213	0.199	0.185	0.175

$N = 5000$ $X^{(d,\beta)}$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = .6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\sqrt{MSE} \tilde{d}_N^{(IR)}$	0.110	0.152	0.150	0.151	0.152	0.153	0.152	0.142
$\sqrt{MSE} \tilde{d}_N^{(MS)}$	0.120	0.123	0.132	0.131	0.132	0.127	0.104	0.155
$\sqrt{MSE} \tilde{d}_N^{(ADG)}$	0.139	0.138	0.141	0.134	0.134	0.140	0.140	0.143
$\sqrt{MSE} \tilde{d}_N^{(WAV)}$	0.170	0.173	0.167	0.165	0.167	0.166	0.164	0.150

TABLE 4.4 – : Comparison between $\tilde{d}_N^{(IR)}$ and other famous semiparametric estimators of d ($\tilde{d}_N^{(MS)}$, $\tilde{d}_N^{(ADG)}$ and $\tilde{d}_N^{(WAV)}$) applied to $X^{(d,\beta)}$ process with several d and N values.

from Lemma (8.2) and then the equations (5.28-5.31) remain valid for all $d \in [0.5, 1.25]$. Then for $d = 0.5$,

$$\sqrt{\frac{N}{m}} \left(IR_N(m) - \mathbb{E}[IR_N(m)] \right) \xrightarrow[N/m \wedge m \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(0.5)),$$

with $\sigma^2(0.5) \simeq (0.2524)^2$. □

Proof of Property 4. As in Surgailis *et al* (2008), we can write :

$$\mathbb{E}[IR_N(m)] = \mathbb{E}\left(\frac{|Y^0 + Y^1|}{|Y^0| + |Y^1|}\right) = \Lambda\left(\frac{R_m}{V_m^2}\right) \quad \text{with} \quad \frac{R_m}{V_m^2} := 1 - 2 \frac{\int_0^\pi f(x) \frac{\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx}{\int_0^\pi f(x) \frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx}.$$

Therefore an expansion of R_m/V_m^2 provides an expansion of $\mathbb{E}[IR_N(m)]$ when $m \rightarrow \infty$.

Step 1 Let f satisfy Assumption $IG(d, \beta)$. Then we are going to establish that there exist positive real numbers C_1 , C_2 and C_3 specified in (4.20), (4.21) and (4.22) such that for $0.5 \leq d < 1.5$ and with $\rho(d)$ defined in (4.9),

1. if $\beta < 2d - 1$, $\frac{R_m}{V_m^2} = \rho(d) + C_1(2 - 2d, \beta)m^{-\beta} + O(m^{-2} + m^{-2\beta})$;
2. if $\beta = 2d - 1$, $\frac{R_m}{V_m^2} = \rho(d) + C_2(2 - 2d, \beta)m^{-\beta} + O(m^{-2} + m^{-2-\beta} \log(m) + m^{-2\beta})$;
3. if $2d - 1 < \beta < 2d + 1$, $\frac{R_m}{V_m^2} = \rho(d) + C_3(2 - 2d, \beta)m^{-\beta} + O(m^{-\beta-\epsilon} + m^{-2d-1} \log(m) + m^{-2\beta})$;
4. if $\beta = 2d + 1$, $\frac{R_m}{V_m^2} = \rho(d) + O(m^{-2d-1} \log(m) + m^{-2})$.

$N = 500$ ARIMA(1, d , 0)	$d = 0$ $\phi = -0.5$	$d = 0$ $\phi = -0.7$	$d = 0$ $\phi = -0.9$	$d = 1$ $\phi = -0.1$	$d = 1$ $\phi = -0.3$	$d = 1$ $\phi = -0.5$
$\tilde{S}_N$: Accepted H_0	1	1	0.37	0	0	0
$KPSS$: Accepted H_0	0.88	0.78	0.29	0	0	0
LMC : Accepted H_0	0.97	1	0.84	0.02	0	0
V/S : Accepted H_0	0.96	0.93	0.84	0.09	0.08	0.12
$\tilde{T}_N$: Rejected H'_0	0.99	0.77	0.08	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	0.94	0	0	0	0
ADF : Rejected H'_0	1	1	1	0.06	0.04	0.04
PP : Rejected H'_0	1	1	1	0.06	0.03	0.02
$N = 5000$ ARIMA(1, d , 0)	$d = 0$ $\phi = -0.5$	$d = 0$ $\phi = -0.7$	$d = 0$ $\phi = -0.9$	$d = 1$ $\phi = -0.1$	$d = 1$ $\phi = -0.3$	$d = 1$ $\phi = -0.5$
$\tilde{S}_N$: Accepted H_0	1	1	0.91	0	0	0
$KPSS$: Accepted H_0	0.95	0.91	0.70	0	0	0
LMC : Accepted H_0	0.95	1	1	0	0	0
V/S : Accepted H_0	0.93	0.97	0.90	0	0	0
$\tilde{T}_N$: Rejected H'_0	1	1	0.87	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	0.95	0	0	0
ADF : Rejected H'_0	1	1	1	0.09	0.01	0.04
PP : Rejected H'_0	1	1	1	0.07	0.01	0.04

TABLE 4.5 – Comparisons of stationarity and nonstationarity tests from 300 independent replications of ARIMA(1, d , 0) processes ($X_t + \phi X_{t-1} = \varepsilon_t$) for several values of ϕ and N . The accuracy of tests is measured by the frequencies of trajectories “accepted as stationary” (accepted H_0 or rejected H'_0) among the 300 replications which should be close to 1 for $d \in (-0.5, 0.5)$ and close to 0 for $d \in [0.5, 1.2]$

$N = 500$ ARFIMA(0, d , 0)	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	1	0.72	0.09	0.01	0
$KPSS$: Accepted H_0	1	0.95	0.63	0.21	0.06	0.01	0	0
LMC : Accepted H_0	0	0.06	0.75	1	1	1	0.52	0
V/S : Accepted H_0	1	0.97	0.81	0.51	0.30	0.20	0.09	0.05
$\tilde{T}_N$: Rejected H'_0	1	1	0.97	0.53	0.02	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	0.99	0.48	0.01	0	0	0
ADF : Rejected H'_0	1	1	1	0.98	0.60	0.24	0.06	0.01
PP : Rejected H'_0	1	1	1	1	0.90	0.43	0.05	0

$N = 5000$ ARFIMA(0, d , 0)	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	1	0.08	0	0	0
$KPSS$: Accepted H_0	1	0.95	0.35	0.01	0	0	0	0
LMC : Accepted H_0	0	0.05	0.97	1	1	1	0.53	0
V/S : Accepted H_0	1	0.95	0.50	0.17	0.05	0	0	0
$\tilde{T}_N$: Rejected H'_0	1	1	1	0.94	0	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	1	0.89	0	0	0	0
ADF : Rejected H'_0	1	1	1	1	0.88	0.53	0.07	0
PP : Rejected H'_0	1	1	1	1	1	0.75	0.07	0

TABLE 4.6 – Comparisons of stationarity and nonstationarity tests from 300 independent replications of ARFIMA(0, d , 0) processes for several values of d and N . The accuracy of tests is measured by the frequencies of trajectories “accepted as stationary” (accepted H_0 or rejected H'_0) among the 300 replications which should be close to 1 for $d \in (-0.5, 0.5)$ and close to 0 for $d \in [0.5, 1.2]$

$N = 500$ ARFIMA(1, d , 1) $\phi = -0.3; \theta = 0.7$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	0.95	0.47	0.11	0.01	0
$KPSS$: Accepted H_0	1	0.90	0.56	0.19	0.06	0.01	0	0
LMC : Accepted H_0	0.12	0	0	0	0	0	0	0
V/S : Accepted H_0	1	0.96	0.78	0.54	0.34	0.18	0.09	0.05
$\tilde{T}_N$: Rejected H'_0	1	1	0.84	0.23	0.01	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	0.96	0.21	0	0	0	0
ADF : Rejected H'_0	1	1	1	0.96	0.59	0.26	0.05	0.01
PP : Rejected H'_0	1	1	1	1	0.74	0.30	0.03	0

$N = 5000$ ARFIMA(1, d , 1) $\phi = -0.3; \theta = 0.7$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	0.99	0.12	0	0	0
$KPSS$: Accepted H_0	1	0.92	0.31	0.02	0	0	0	0
LMC : Accepted H_0	0	0	0	0	0	0	0	0
V/S : Accepted H_0	1	0.95	0.61	0.22	0.07	0	0.01	0
$\tilde{T}_N$: Rejected H'_0	1	1	1	0.67	0.01	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	1	0.86	0	0	0	0
ADF : Rejected H'_0	1	1	1	1	0.91	0.45	0.04	0
PP : Rejected H'_0	1	1	1	1	0.99	0.59	0.03	0

TABLE 4.7 – Comparisons of stationarity and nonstationarity tests from 300 independent replications of ARFIMA(1, d , 1) processes (with $\phi = -0.3$ and $\theta = 0.7$) for several values of d and N . The accuracy of tests is measured by the frequencies of trajectories “accepted as stationary” (accepted H_0 or rejected H'_0) among the 300 replications which should be close to 1 for $d \in (-0.5, 0.5)$ and close to 0 for $d \in [0.5, 1.2]$

$N = 500$ $X^{(d,\beta)}$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	1	0.99	0.49	0.05	0.01
$KPSS$: Accepted H_0	1	1	0.89	0.46	0.15	0.03	0.01	0
LMC : Accepted H_0	0	0	0.13	0.86	1	1	0.82	0
V/S : Accepted H_0	1	1	0.93	0.69	0.42	0.25	0.16	0.09
$\tilde{T}_N$: Rejected H'_0	1	1	1	0.93	0.37	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	1	0.98	0.23	0	0	0
ADF : Rejected H'_0	1	1	1	1	1	0.88	0.38	0.11
PP : Rejected H'_0	1	1	1	1	0.90	0.43	0.05	0

$N = 5000$ $X^{(d,\beta)}$	$d = -0.2$	$d = 0$	$d = 0.2$	$d = 0.4$	$d = 0.6$	$d = 0.8$	$d = 1$	$d = 1.2$
$\tilde{S}_N$: Accepted H_0	1	1	1	1	1	0.03	0	0
$KPSS$: Accepted H_0	1	1	0.65	0.12	0.01	0	0	0
LMC : Accepted H_0	0	0	0.39	1	1	1	1	0
V/S : Accepted H_0	1	0.99	0.79	0.29	0.11	0.04	0	0
$\tilde{T}_N$: Rejected H'_0	1	1	1	0.99	0.82	0	0	0
$\hat{T}_{ADG}$: Rejected H'_0	1	1	1	1	0.30	0	0	0
ADF : Rejected H'_0	1	1	1	1	1	0.98	0.34	0.01
PP : Rejected H'_0	1	1	1	1	1	1	0.50	0.01

TABLE 4.8 – Comparisons of stationarity and nonstationarity tests from 300 independent replications of $X^{(d,\beta)}$ processes for several values of d and N . The accuracy of tests is measured by the frequencies of trajectories “accepted as stationary” (accepted H_0 or rejected H'_0) among the 300 replications which should be close to 1 for $d \in (-0.5, 0.5)$ and close to 0 for $d \in [0.5, 1.2]$

Under Assumption $IG(d, \beta)$ and with $J_j(a, m)$ defined in (4.26) in the same Lemma 4.26, it is clear that,

$$\frac{R_m}{V_m^2} = 1 - 2 \frac{J_6(2-2d, m) + \frac{c_1}{c_0} J_6(2-2d+\beta, m) + O(J_6(2-2d+\beta+\varepsilon))}{J_4(2-2d, m) + \frac{c_1}{c_0} J_4(2-2d+\beta, m) + O(J_4(2-2d+\beta+\varepsilon))},$$

since $\int_0^\pi O(x^{2-2d+\beta+\varepsilon}) \frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = O(J_j(2-2d+\beta+\varepsilon))$. Now using the results of Lemma 4.26 and constants $C_{j\ell}$, $C'_{j\ell}$ and $C''_{j\ell}$, $j = 4, 6$, $\ell = 1, 2$ defined in Lemma 4.26,

1. Let $0 < \beta < 2d-1 < 2$, i.e. $-1 < 2-2d+\beta < 1$. Then

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{C_{61}(2-2d) m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0} C_{61}(2-2d+\beta) m^{1+2d-\beta} + O(m^{2d-1-\beta})}{C_{41}(2-2d) m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0} C_{41}(2-2d+\beta) m^{1+2d-\beta} + O(m^{2d-1-\beta})} \\ &= 1 - \frac{2}{C_{41}(2-2d)} \left[C_{61}(2-2d) + \frac{c_1}{c_0} C_{61}(2-2d+\beta) m^{-\beta} \right] \left[1 - \frac{c_1}{c_0} \frac{C_{41}(2-2d+\beta)}{C_{41}(2-2d)} m^{-\beta} \right] + O(m^{-2}) \\ &= 1 - \frac{2C_{61}(2-2d)}{C_{41}(2-2d)} + 2 \frac{c_1}{c_0} \left[\frac{C_{61}(2-2d)C_{41}(2-2d+\beta)}{C_{41}(2-2d)C_{41}(2-2d)} - \frac{C_{61}(2-2d+\beta)}{C_{41}(2-2d)} \right] m^{-\beta} + O(m^{-2} + m^{-2\beta}). \end{aligned}$$

As a consequence,,

$$\begin{aligned} \frac{R_m}{V_m^2} &= \rho(d) + C_1(2-2d, \beta) m^{-\beta} + O(m^{-2} + m^{-2\beta}) \quad (m \rightarrow \infty), \quad \text{with } 0 < \beta < 2d-1 < 2 \text{ and} \\ C_1(2-2d, \beta) &:= 2 \frac{c_1}{c_0} \frac{1}{C_{41}^2(2-2d)} [C_{61}(2-2d)C_{41}(2-2d+\beta) - C_{61}(2-2d+\beta)C_{41}(2-2d)], \end{aligned} \quad (4.20)$$

and numerical experiments proves that $C_1(2-2d, \beta)/c_1$ is negative for any $d \in (0.5, 1.5)$ and $\beta > 0$.

2. Let $\beta = 2d-1$, i.e. $2-2d+\beta = 1$. Then,

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{C_{61}(2-2d) m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0} C'_{61}(1) m^{1-2d} + O(\log(m))}{C_{41}(2-2d) m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0} C'_{41}(1) m^{1-2d} + O(\log(m))} \\ &= 1 - \frac{2}{C_{41}(2-2d)} \left[C_{61}(2-2d) + \frac{c_1}{c_0} C'_{61}(1) m^{1-2d} \right] \left[1 - \frac{c_1}{c_0} \frac{C'_{41}(1)}{C_{41}(2-2d)} m^{1-2d} \right] + O(m^{-2} + m^{-2d-1} \log(m)) \\ &= 1 - \frac{2C_{61}(2-2d)}{C_{41}(2-2d)} + 2 \frac{c_1}{c_0} \left[\frac{C_{61}(2-2d)C'_{41}(1)}{C_{41}(2-2d)C_{41}(2-2d)} - \frac{C'_{61}(1)}{C_{41}(2-2d)} \right] m^{1-2d} + O(m^{-2} + m^{-2d-1} \log(m) + m^{2-4d}). \end{aligned}$$

As a consequence,

$$\begin{aligned} \frac{R_m}{V_m^2} &= \rho(d) + C_2(2-2d, \beta) m^{-\beta} + O(m^{-2} + m^{-2-\beta} \log(m) + m^{-2\beta}) \quad (m \rightarrow \infty), \quad \text{with } 0 < \beta = 2d-1 < 2 \text{ and} \\ C_2(2-2d, \beta) &:= 2 \frac{c_1}{c_0} \frac{1}{C_{41}^2(2-2d)} [C_{61}(2-2d)C'_{41}(1) - C'_{61}(1)C_{41}(2-2d)], \end{aligned} \quad (4.21)$$

and numerical experiments proves that $C_2(2-2d, \beta)/c_1$ is negative for any $d \in [0.5, 1.5)$ and $\beta > 0$.

3. Let $2d-1 < \beta < 2d+1$, i.e. $1 < 2-2d+\beta < 3$. Then,

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{C_{61}(2-2d)m^{1+2d} + \frac{c_1}{c_0}C'_{61}(2-2d+\beta)m^{1+2d-\beta} + O(m^{1+2d-\beta-\epsilon} + \log(m))}{C_{41}(2-2d)m^{1+2d} + \frac{c_1}{c_0}C'_{41}(2-2d+\beta)m^{1+2d-\beta} + O(m^{1+2d-\beta-\epsilon} + m^{-2d-1}\log(m))} \\ &= 1 - \frac{2}{C_{41}(2-2d)} \left[C_{61}(2-2d) + \frac{c_1}{c_0}C'_{61}(2-2d+\beta)m^{-\beta} \right] \left[1 - \frac{c_1}{c_0} \frac{C'_{41}(2-2d+\beta)}{C_{41}(2-2d)} m^{-\beta} \right] + O(m^{-\beta-\epsilon} + m^{-2d-1}\log(m)) \\ &= 1 - \frac{2C_{61}(2-2d)}{C_{41}(2-2d)} + 2\frac{c_1}{c_0} \left[\frac{C_{61}(2-2d)C'_{41}(2-2d+\beta)}{C_{41}(2-2d)C_{41}(2-2d)} - \frac{C'_{61}(2-2d+\beta)}{C_{41}(2-2d)} \right] m^{-\beta} + O(m^{-\beta-\epsilon} + m^{-2d-1}\log(m)). \end{aligned}$$

As a consequence,

$$\begin{aligned} \frac{R_m}{V_m^2} &= \rho(d) + C_3(2-2d, \beta) m^{-\beta} + O(m^{-\beta-\epsilon} + m^{-2d-1}\log(m) + m^{-2\beta}) \quad (m \rightarrow \infty), \quad \text{and} \\ C_3(2-2d, \beta) &:= 2\frac{c_1}{c_0} \frac{1}{C_{41}^2(2-2d)} [C_{61}(2-2d)C'_{41}(2-2d+\beta) - C'_{61}(2-2d+\beta)C_{41}(2-2d)], \end{aligned} \quad (4.22)$$

and numerical experiments proves that $C_3(2-2d, \beta)/c_1$ is negative for any $d \in [0.5, 1.5)$ and $\beta > 0$.

4. Let $\beta = 2d+1$. Then, Once again with Lemma 4.26 :

$$\begin{aligned} \frac{R_m}{V_m^2} &= 1 - 2 \frac{C_{61}(2-2d)m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0}C'_{62}(3)\log(m) + O(1)}{C_{41}(2-2d)m^{1+2d} + O(m^{2d-1}) + \frac{c_1}{c_0}C'_{42}(3)\log(m) + O(1)} \\ &= 1 - \frac{2}{C_{41}(2-2d)} \left[C_{61}(2-2d) + \frac{c_1}{c_0}C'_{62}(3)m^{-\beta}\log(m) \right] \left[1 - \frac{c_1}{c_0} \frac{C'_{42}(3)}{C_{41}(2-2d)} m^{-\beta}\log(m) \right] + O(m^{-2} + m^{-2d-1}) \\ &= 1 - \frac{2C_{61}(2-2d)}{C_{41}(2-2d)} + 2\frac{c_1}{c_0} \left[\frac{C_{61}(2-2d)C'_{42}(3)}{C_{41}(2-2d)C_{41}(2-2d)} - \frac{C'_{62}(3)}{C_{41}(2-2d)} \right] m^{-\beta}\log(m) + O(m^{-2}). \end{aligned}$$

As a consequence,

$$\frac{R_m}{V_m^2} = \rho(d) + O(m^{-2d-1}\log(m) + m^{-2}) \quad (m \rightarrow \infty), \quad \text{with } 2 < \beta = 2d+1 < 4. \quad (4.23)$$

Step 2 : A Taylor expansion of $\Lambda(\cdot)$ around $\rho(d)$ provides :

$$\Lambda\left(\frac{R_m}{V_m^2}\right) \simeq \Lambda(\rho(d)) + \left[\frac{\partial \Lambda}{\partial \rho}\right](\rho(d))\left(\frac{R_m}{V_m^2} - \rho(d)\right) + \frac{1}{2} \left[\frac{\partial^2 \Lambda}{\partial \rho^2}\right](\rho(d))\left(\frac{R_m}{V_m^2} - \rho(d)\right)^2.$$

Note that numerical experiments show that $\left[\frac{\partial \Lambda}{\partial \rho}\right](\rho) > 0.2$ for any $\rho \in (-1, 1)$. As a consequence, using the previous expansions of R_m/V_m^2 obtained in Step 1 and since $\mathbb{E}[IR_N(m)] = \Lambda(R_m/V_m^2)$, then for all $0 < \beta \leq 2$:

$$\mathbb{E}[IR_N(m)] = \Lambda_0(d) + \begin{cases} c_1 C'_1(d, \beta) m^{-\beta} + O(m^{-2} + m^{-2\beta}) & \text{if } \beta < 2d-1 \\ c_1 C'_2(d, \beta) m^{-\beta} + O(m^{-2} + m^{-2-\beta}\log m + m^{-2\beta}) & \text{if } \beta = 2d-1 \\ c_1 C'_3(d, \beta) m^{-\beta} + O(m^{-\beta-\epsilon} + m^{-2d-1}\log m + m^{-2\beta}) & \text{if } 2d-1 < \beta < 2d+1 \\ O(m^{-2d-1}\log m + m^{-2}) & \text{if } \beta = 1+2d \end{cases}$$

with $C'_\ell(d, \beta) = \left[\frac{\partial \Lambda}{\partial \rho} \right](\rho(d)) C_\ell(2-2d, \beta)$ for $\ell = 1, 2, 3$ and C_ℓ defined in (4.20), (4.21) and (4.22). $\square$

Proof of Theorem 3. Using Property 4, if $m \simeq C N^\alpha$ with $C > 0$ and $(1 + 2\beta)^{-1} < \alpha < 1$ then $\sqrt{N/m} (\mathbb{E}[IR_N(m)] - \Lambda_0(d)) \xrightarrow{N \rightarrow \infty} 0$ and it implies that the multidimensional CLT (4.7) can be replaced by

$$\sqrt{\frac{N}{m}} (IR_N(m_j) - \Lambda_0(d))_{1 \leq j \leq p} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_p(d)). \quad (4.24)$$

It remains to apply the Delta-method with the function Λ_0^{-1} to CLT (4.24). This is possible since the function $d \rightarrow \Lambda_0(d)$ is an increasing function such that $\Lambda'_0(d) > 0$ and $(\Lambda_0^{-1})'(\Lambda_0(d)) = 1/\Lambda'_0(d) > 0$ for all $d \in (-0.5, 1.5)$. It achieves the proof of Theorem 3. $\square$

Proof of Proposition 5. See Bardet and Dola (2012, [6]). $\square$

Proof of Theorem 4. See Bardet and Dola (2012, [6]). $\square$

Appendix

We first recall usual equalities frequently used in the sequel :

Lemma 2. For all $\lambda > 0$

1. For $a \in (0, 2)$, $\frac{2}{|\lambda|^{a-1}} \int_0^\infty \frac{\sin(\lambda x)}{x^a} dx = \frac{4a}{2^a |\lambda|^a} \int_0^\infty \frac{\sin^2(\lambda x)}{x^{a+1}} dx = \frac{\pi}{\Gamma(a) \sin(\frac{a\pi}{2})}$.
2. For $b \in (-1, 1)$, $\frac{1}{2^{1-b} - 1} \int_0^\infty \frac{\sin^4(\lambda x)}{x^{4-b}} dx = \frac{16}{-15 + 6 \cdot 2^{3-b} - 3^{3-b}} \times \int_0^\infty \frac{\sin^6(\lambda x)}{x^{4-b}} dx = \frac{2^{3-b} |\lambda|^{3-b} \pi}{4 \Gamma(4-b) \sin(\frac{(1-b)\pi}{2})}$.
3. For $b \in (1, 3)$, $\frac{1}{1 - 2^{1-b}} \int_0^\infty \frac{\sin^4(\lambda x)}{x^{4-b}} dx = \frac{16}{15 - 6 \cdot 2^{3-b} + 3^{3-b}} \times \int_0^\infty \frac{\sin^6(\lambda x)}{x^{4-b}} dx = \frac{2^{3-b} |\lambda|^{3-b} \pi}{4 \Gamma(4-b) \sin(\frac{(3-b)\pi}{2})}$.

Proof. These equations are given or deduced (using decompositions of $\sin^j(\cdot)$ and integration by parts) from the following relation (see [29], 2003, p. 31).

$$\int_0^\infty y^{-\alpha} \sin(y) dy = \frac{1}{2} \frac{\pi}{\Gamma(\alpha) \sin(\pi(\frac{\alpha}{2}))} \quad \text{for } 0 < \alpha < 2.$$

Lemma 3. For $j = 4, 6$, denote

$$J_j(a, m) := \int_0^\pi x^a \frac{\sin^j(\frac{mx}{2})}{\sin^4(\frac{x}{2})} dx. \quad (4.25)$$

Then, we have the following expansions when $m \rightarrow \infty$:

$$J_j(a, m) = \begin{cases} C_{j1}(a) m^{3-a} + O(m^{1-a}) & \text{if } -1 < a < 1 \\ C'_{j1}(1) m^{3-a} + O(\log(m)) & \text{if } a = 1 \\ C'_{j1}(a) m^{3-a} + O(1) & \text{if } 1 < a < 3 \\ C'_{j2}(3) \log(m) + O(1) & \text{if } a = 3 \\ C''_{j1}(a) + O(m^{-(a-3) \wedge 2}) & \text{if } a > 3 \end{cases} \quad (4.26)$$

with the following real constants (which do not vanish for any a on the corresponding set) :

- $C_{41}(a) := \frac{4 \pi(1 - \frac{2^{3-a}}{4})}{(3-a)\Gamma(3-a)\sin(\frac{(3-a)\pi}{2})}$; $C_{61}(a) := \frac{\pi(15 - 6 \cdot 2^{3-a} + 3^{3-a})}{4(3-a)\Gamma(3-a)\sin(\frac{(3-a)\pi}{2})}$
- $C'_{41}(a) := \left(\frac{6}{3-a} \mathbf{1}_{\{1 \leq a < 3\}} + 16 \int_0^1 \frac{\sin^4(\frac{y}{2})}{y^{4-a}} dy + 2 \int_1^\infty \frac{1}{y^{4-a}} \left(-4 \cos(y) + \cos(2y) \right) dy \right)$;
 $C'_{61}(a) := \left[16 \int_0^1 \frac{\sin^6(\frac{y}{2})}{y^{4-a}} dy + \frac{5}{3-a} \mathbf{1}_{\{1 \leq a < 3\}} + \frac{1}{2} \int_1^\infty \frac{1}{y^{4-a}} \left(-15 \cos(y) + 6 \cos(2y) - \cos(3y) \right) dy \right]$
- $C'_{42}(a) := \left(6 \cdot \mathbf{1}_{\{a=3\}} + \mathbf{1}_{\{a=1\}} \right)$ and $C'_{62}(a) := \left(5 \cdot \mathbf{1}_{\{a=3\}} + \frac{5}{6} \cdot \mathbf{1}_{\{a=1\}} \right)$
- $C''_{41}(a) := \frac{3}{8} \int_0^\pi \frac{x^a}{\sin^4(\frac{x}{2})} dx$ and $C''_{61}(a) := \frac{5}{16} \int_0^\pi \frac{x^a}{\sin^4(\frac{x}{2})} dx$.

Proof. The proof of these expansions follows the steps than those of Lemma 1 in Bardet and Dola (2012, [6]). Hence we write for $j = 4, 6$,

$$J_j(a, m) = \tilde{J}_j(a, m) + \int_0^\pi x^a \sin^j\left(\frac{mx}{2}\right) \frac{1}{(\frac{x}{2})^4} dx + \int_0^\pi x^a \sin^j\left(\frac{mx}{2}\right) \frac{2}{3} \frac{1}{(\frac{x}{2})^2} dx \quad (4.27)$$

with

$$\tilde{J}_j(a, m) := \int_0^\pi x^a \sin^j\left(\frac{mx}{2}\right) \left(\frac{1}{\sin^4(\frac{x}{2})} - \frac{1}{(\frac{x}{2})^4} - \frac{2}{3} \frac{1}{(\frac{x}{2})^2} \right) dx.$$

The expansions when $m \rightarrow \infty$ of both the right hand sided integrals in (4.27) are obtained from Lemma 2. It remains to obtain the expansion of $\tilde{J}_j(a, m)$. Then, using classical trigonometric and Taylor expansions :

$$\begin{aligned} \sin^4\left(\frac{y}{2}\right) &= \frac{1}{8} (3 - 4 \cos(y) + \cos(2y)) \quad \text{and} \quad \frac{1}{\sin^4(y)} - \frac{1}{y^4} - \frac{2}{3} \frac{1}{y^2} \sim \frac{11}{45} \quad (y \rightarrow 0) \\ \sin^6\left(\frac{y}{2}\right) &= \frac{1}{32} (10 - 15 \cos(y) + 6 \cos(2y) - \cos(3y)) \quad \text{and} \quad \frac{1}{y^5} + \frac{1}{3} \frac{1}{y^3} - \frac{\cos(y)}{\sin^5(y)} \sim \frac{31}{945} y \quad (y \rightarrow 0), \end{aligned}$$

the expansions of $\tilde{J}_j(a, m)$ can be obtained. Numerical experiments show that $C''_{41}(a) \neq 0$, $C''_{61}(a) \neq 0$, $C''_{42}(a) \neq 0$ and $C''_{62}(a) \neq 0$.

r=(USD1 vs Deutsh-Mark Exchange Rate Return)

Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 5963]	S	S	S	S	S	S	SM	5.5	-0.2	-0.031	0.059	0.057	-0.007
[5965 : 6313]	S	S	S	S	S	S	SM	3.4	0.1	0.034	0.169	0.122	-0.015
[6315 : 7173]	S	S	S	S	S	S	SM	5.3	-0.4	0.098	0.140	0.043	0.019

 $|r| = abs(\text{USD1 vs Deutsh-Mark Exchange Rate Return})$

Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 5963]	S	S	S	S	NS	NS	LM	9.5	1.8	0.294	0.301	0.344	0.275
[5965 : 6313]	S	S	NS	S	NS	NS	LM	3.6	1.1	-0.121	0.153	0.414	-0.038
[6315 : 7173]	S	S	S	S	NS	NS	LM	9.2	1.8	0.168	0.417	0.389	0.410

 $r^2 = (\text{USD1 vs Deutsh-Mark Exchange Rate Return})^2$

Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 5963]	S	S	S	S	S	S	LM	289.5	10.7	0.081	0.258	0.298	0.078
[5965 : 6313]	S	S	S	S	NS	S	LM	8.7	2.3	-0.018	0.127	0.431	-0.096
[6315 : 7173]	S	S	S	S	NS	S	LM	81.3	7.1	0.035	0.411	0.336	0.428

 $|r|^\theta = (abs(\text{USD1 vs Deutsh-Mark Exchange Rate Return}))^\theta$

$\hat{\theta}_i^{(j)} =$ ArgMax $_{\theta}(\hat{d}(r_i ^\theta))$	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
$\hat{\theta}_1^{(IR)} = 0.32$	S	S	NS	S	NS	NS	SM	3.5	-0.5	0.321*	0.251	0.256	0.343
$\hat{\theta}_1^{(MS)} = 0.97$	S	S	S	S	NS	NS	SM	8.7	1.1	0.293	0.301*	0.343	0.275
$\hat{\theta}_1^{(ADG)} = 1.12$	S	S	S	S	NS	NS	SM	13.7	2.3	0.302	0.300	0.345*	0.273
$\hat{\theta}_1^{(WAV)} = 0.77$	S	S	S	S	NS	NS	SM	5.1	1.1	0.273	0.298	0.335	0.379*
$\hat{\theta}_2^{(IR)} = 0.05$	S	S	NS	NS	S	NS	LM	27.9	-5.0	0.246*	0.078	-0.005	0.103
$\hat{\theta}_2^{(MS)} = 1.31$	S	S	S	S	NS	NS	LM	4.9	1.5	-0.103	0.166*	0.446	-0.072
$\hat{\theta}_2^{(ADG)} = 1.50$	S	S	S	S	NS	S	LM	5.8	1.8	-0.092	0.162	0.450*	-0.082
$\hat{\theta}_2^{(WAV)} = 0.03$	S	S	NS	NS	S	NS	LM	30.9	-5.4	0.239	0.113	-0.030	0.211*
$\hat{\theta}_3^{(IR)} = 0.63$	S	S	NS	NS	NS	NS	LM	3.8	0.7	0.244*	0.354	0.333	0.097
$\hat{\theta}_3^{(MS)} = 1.44$	S	S	S	S	NS	NS	LM	27.1	3.7	0.159	0.436*	0.387	0.441
$\hat{\theta}_3^{(ADG)} = 1.19$	S	S	S	S	NS	NS	LM	14.9	2.6	0.168	0.430	0.394*	0.430
$\hat{\theta}_3^{(WAV)} = 2.90$	S	S	S	S	NS	S	LM	223.4	13.3	0.053	0.291	0.233	0.475*

TABLE 4.9 – Results of stationarity, nonstationarity and V/S tests and the 4 long memory parameter estimators applied to several functionals f of USD1 vs Deutsh-Mark Exchange Rate Return : from the top to bottom, $f(x) = x$, $f(x) = |x|$, $f(x) = x^2$ and $f(x) = |x|^\theta$ with θ maximizing the 4 different long memory parameter estimators ("S" for "stationarity" decision and "NS" for "nonstationarity" decision). Statistics are applied to the 3 estimated stages of each trajectory (obtained from a change detection algorithm).

r=Dow Jones Transportation Index Return													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\widehat{d^{IR}}$	$\widehat{d^{MS}}$	$\widehat{d^{ROB}}$	$\widehat{d^{WAV}}$
[1 : 1271]	S	S	S	S	S	NS	SM	4.6	0.0	0.218	0.174	0.098	0.198
[1273 : 8531]	S	S	S	S	S	S	SM	21.7	-0.8	0.053	0.002	0.008	-0.404
[8533 : 12071]	S	S	S	S	S	S	SM	8.3	-0.3	0.002	-0.015	-0.034	-0.038

$ r = abs(\text{Dow Jones Transportation Index Return})$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\widehat{d^{IR}}$	$\widehat{d^{MS}}$	$\widehat{d^{ROB}}$	$\widehat{d^{WAV}}$
[1 : 1272]	S	S	S	S	NS	NS	LM	6.1	1.5	0.154	0.320	0.270	0.166
[1273 : 8532]	S	S	S	S	NS	NS	LM	57.3	4.4	0.322	0.260	0.240	0.168
[8533 : 12071]	S	S	S	S	NS	NS	LM	16.3	2.5	0.405	0.476	0.496	0.374

$r^2 = (\text{Dow Jones Transportation Index Return})^2$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\widehat{d^{IR}}$	$\widehat{d^{MS}}$	$\widehat{d^{ROB}}$	$\widehat{d^{WAV}}$
[1 : 1272]	S	S	S	S	NS	S	LM	32.3	4.4	0.158	0.284	0.231	0.231
[1273 : 8532]	S	S	S	S	S	NS	LM	2301.5	39.9	0.334	0.122	0.093	0.118
[8533 : 12071]	S	NS	S	S	NS	NS	LM	459.0	15.5	0.416	0.452	0.434	0.356

$ r ^\theta = (abs(\text{Dow Jones Transportation Index Return}))^\theta$													
$\hat{\theta}_i^{(j)} =$ $\text{ArgMax}_\theta(\hat{d}(r_i ^\theta))$	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\widehat{d^{IR}}$	$\widehat{d^{MS}}$	$\widehat{d^{ROB}}$	$\widehat{d^{WAV}}$
$\hat{\theta}_1^{(IR)} = 1.83$	S	S	S	S	NS	S	LM	25.0	3.8	0.252*	0.291	0.237	0.202
$\hat{\theta}_1^{(MS)} = 0.45$	S	S	NS	S	NS	NS	LM	2.7	0.4	0.118	0.331*	0.290	0.047
$\hat{\theta}_1^{(ADG)} = 0.36$	S	S	NS	S	NS	NS	LM	2.9	-0.4	0.118	0.329	0.291*	0.237
$\hat{\theta}_1^{(WAV)} = 0.03$	S	S	NS	S	NS	NS	LM	12.8	-3.4	0.149	0.257	0.260	0.327*
$\hat{\theta}_2^{(IR)} = 2.06$	S	S	S	S	S	NS	LM	2551.6	42.6	0.355*	0.113	0.086	0.110
$\hat{\theta}_2^{(MS)} = 0.68$	S	S	S	S	NS	NS	LM	10.6	1.6	0.308	0.276*	0.261	0.135
$\hat{\theta}_2^{(ADG)} = 0.65$	S	S	S	S	NS	NS	LM	9.2	1.4	0.303	0.276	0.261*	0.129
$\hat{\theta}_2^{(WAV)} = 1.29$	S	S	S	S	NS	NS	LM	246.8	10.1	0.330	0.227	0.200	0.504*
$\hat{\theta}_3^{(IR)} = 0.66$	S	NS	S	S	NS	NS	LM	5.4	1.1	0.444*	0.435	0.461	0.374
$\hat{\theta}_3^{(MS)} = 1.38$	S	S	S	S	NS	NS	LM	64.8	5.2	0.402	0.492*	0.499	0.391
$\hat{\theta}_3^{(ADG)} = 1.22$	S	S	S	S	NS	NS	LM	36.2	3.8	0.400	0.489	0.502*	0.387
$\hat{\theta}_3^{(WAV)} = 2.75$	S	S	S	S	NS	NS	LM	1698.7	35.8	0.407	0.315	0.287	0.466*

TABLE 4.10 – Results of stationarity, nonstationarity and V/S tests and the 4 long memory parameter estimators applied to several functionals f of DowJones Transportation Index Return : from the top to bottom, $f(x) = x$, $f(x) = |x|$, $f(x) = x^2$ and $f(x) = |x|^\theta$ with θ maximizing the 4 different long memory parameter estimators ("S" for "stationarity" decision and "NS" for "nonstationarity" decision). Statistics are applied to the 3 estimated stages of each trajectory (obtained from a change detection algorithm).

r=Dow Jones Utilities Index Return													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 1152]	S	S	S	S	S	S	SM	7.3	0.6	0.191	0.037	-0.132	0.222
[1153 : 8748]	S	S	S	S	S	S	SM	43.2	-1.3	0.094	0.025	0.001	0.043
[8749 : 12071]	S	S	S	S	S	S	SM	13.0	0.0	0.026	0.024	0.001	-0.032

$ r = abs(\text{Dow Jones Utilities Index Return})$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 1152]	S	S	S	S	NS	NS	LM	11.9	2.4	0.283	0.287	0.316	0.225
[1153 : 8748]	S	S	S	S	NS	NS	LM	127.4	5.9	0.134	0.301	0.304	0.184
[8749 : 12071]	S	S	S	S	NS	NS	LM	25.5	3.4	0.417	0.559	0.484	0.595

$r^2 = (\text{Dow Jones Utilities Index Return})^2$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 1152]	S	S	S	S	NS	S	LM	63.1	6.7	0.250	0.253	0.212	0.270
[1153 : 8748]	S	S	S	S	NS	NS	LM	5322.4	67.8	0.130	0.100	0.100	0.100
[8749 : 12071]	S	NS	S	S	NS	NS	LM	289.6	14.0	0.510	0.468	0.423	0.513

$ r ^\theta = (abs(\text{Dow Jones Utilities Index Return}))^\theta$													
$\hat{\theta}_i^{(j)} =$ $\text{ArgMax}_\theta(\hat{d}(r_i ^\theta))$	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
$\hat{\theta}_1^{(IR)} = 0.39$	S	S	NS	S	NS	NS	LM	3.3	-0.1	0.354*	0.262	0.327	0.145
$\hat{\theta}_1^{(MS)} = 1.09$	S	S	S	S	NS	NS	LM	14.6	8.4	0.215	0.288*	0.308	0.234
$\hat{\theta}_1^{(ADG)} = 0.60$	S	S	NS	S	NS	NS	LM	4.5	0.8	0.311	0.276	0.336*	0.396
$\hat{\theta}_1^{(WAV)} = 0.63$	S	S	S	S	NS	NS	LM	4.8	1.0	0.310	0.278	0.336	0.398*
$\hat{\theta}_2^{(IR)} = 3.00$	S	S	S	S	S	S	LM	7320.6	84.9	0.165*	0.015	0.017	0.040
$\hat{\theta}_2^{(MS)} = 0.61$	S	S	S	S	NS	NS	LM	9.1	1.2	0.113	0.330*	0.327	0.113
$\hat{\theta}_2^{(ADG)} = 0.67$	S	S	S	S	NS	NS	LM	13.0	1.6	0.117	0.330	0.327*	0.113
$\hat{\theta}_2^{(WAV)} = 1.84$	S	S	S	S	NS	NS	LM	4386.6	59.1	0.125	0.130	0.129	0.377*
$\hat{\theta}_3^{(IR)} = 2.69$	S	NS	S	S	NS	NS	LM	683.1	22.5	0.527*	0.394	0.344	0.426
$\hat{\theta}_3^{(MS)} = 0.95$	S	S	S	S	NS	NS	LM	21.6	3.1	0.415	0.560*	0.483	0.544
$\hat{\theta}_3^{(ADG)} = 1.10$	S	S	S	S	NS	NS	LM	35.2	4.2	0.421	0.557	0.485*	0.364
$\hat{\theta}_3^{(WAV)} = 1.03$	S	S	S	S	NS	NS	LM	28.2	3.7	0.419	0.559	0.484	0.723*

TABLE 4.11 – Results of stationarity, nonstationarity and V/S tests and the 4 long memory parameter estimators applied to several functionals f of Dow Jones Utilities Index Return : from the top to bottom, $f(x) = x$, $f(x) = |x|$, $f(x) = x^2$ and $f(x) = |x|^\theta$ with θ maximizing the 4 different long memory parameter estimators (“S“ for ”stationarity“ decision and ”NS“ for ”nonstationarity“ decision). Statistics are applied to the 3 estimated stages of each trajectory (obtained from a change detection algorithm).

r=Nasdaq Industrials Index Return													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 7160]	S	S	S	S	S	S	SM	20.7	-1.5	0.141	0.073	0.092	-0.202
[7161 : 8320]	S	S	S	S	S	S	SM	4.6	0.0	0.012	0.070	0.116	0.014
[8321 : 10480]	S	S	S	S	NS	S	SM	10.4	-0.3	0.045	0.078	0.082	-0.045

$ r = abs(\text{Nasdaq Industrials Index Return})$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 7160]	S	S	S	S	S	NS	SM	52.4	4.4	0.361	0.309	0.287	0.274
[7161 : 8320]	S	S	S	S	NS	NS	LM	7.4	1.6	0.284	0.532	0.504	0.385
[8321 : 10480]	S	NS	S	NS	NS	NS	LM	18.3	3.0	0.516	0.761	0.606	0.668

$r^2 = (\text{Nasdaq Industrials Index Return})^2$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
[1 : 7160]	S	S	S	S	S	NS	SM	1356.8	31.4	0.381	0.146	0.114	0.100
[7161 : 8320]	S	S	S	S	NS	NS	LM	49.0	5.4	0.304	0.466	0.378	0.432
[8321 : 10480]	S	NS	S	S	NS	NS	LM	140.0	10.0	0.498	0.786	0.544	0.708

$ r ^\theta = (abs(\text{Nasdaq Industrials Index Return}))^\theta$													
$\hat{\theta}_i^{(j)} =$ $\text{ArgMax}_\theta(\hat{d}(r_i ^\theta))$	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\hat{d}$			
	$\tilde{S}_N$	$\tilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\tilde{\kappa}$	$\tilde{s}$	$\hat{d}^{IR}$	$\hat{d}^{MS}$	$\hat{d}^{ROB}$	$\hat{d}^{WAV}$
$\hat{\theta}_1^{(IR)} = 1.04$	S	S	S	S	S	NS	LM	63.7	4.9	0.396*	0.304	0.281	0.325
$\hat{\theta}_1^{(MS)} = 0.67$	S	S	S	S	NS	NS	LM	10.2	0.7	0.188	0.329*	0.325	0.293
$\hat{\theta}_1^{(ADG)} = 0.56$	S	S	S	S	NS	NS	LM	6.3	0.9	0.178	0.326	0.328*	0.275
$\hat{\theta}_1^{(WAV)} = 0.83$	S	S	S	S	NS	NS	LM	22.4	2.6	0.199	0.324	0.311	0.587*
$\hat{\theta}_2^{(IR)} = 2.83$	S	S	S	S	NS	S	LM	142.0	9.8	0.317*	0.374	0.276	0.263
$\hat{\theta}_2^{(MS)} = 1.03$	S	S	S	S	NS	NS	LM	7.9	1.7	0.284	0.532*	0.501	0.388
$\hat{\theta}_2^{(ADG)} = 0.73$	S	S	NS	S	NS	NS	LM	4.2	0.9	0.300	0.517	0.517*	0.340
$\hat{\theta}_2^{(WAV)} = 1.87$	S	S	S	S	NS	NS	LM	39.8	4.8	0.299	0.479	0.395	0.432*
$\hat{\theta}_3^{(IR)} = 2.60$	S	NS	S	S	NS	S	LM	256.9	14.3	0.548*	0.272	0.479	0.669
$\hat{\theta}_3^{(MS)} = 1.70$	S	NS	S	S	NS	NS	LM	89.1	7.7	0.504	0.801*	0.575	0.739
$\hat{\theta}_3^{(ADG)} = 1.13$	S	NS	S	S	NS	NS	LM	26.0	3.7	0.526	0.772	0.608*	0.671
$\hat{\theta}_3^{(WAV)} = 1.26$	S	NS	S	S	NS	NS	LM	36.0	4.5	0.532	0.782	0.606	0.760*

TABLE 4.12 – Results of stationarity, nonstationarity and V/S tests and the 4 long memory parameter estimators applied to several functionals f of Nasdaq Industrials Index Return : from the top to bottom, $f(x) = x$, $f(x) = |x|$, $f(x) = x^2$ and $f(x) = |x|^\theta$ with θ maximizing the 4 different long memory parameter estimators ("S" for "stationarity" decision and "NS" for "nonstationarity" decision). Statistics are applied to the 3 estimated stages of each trajectory (obtained from a change detection algorithm).

r=Nikkei 225A Index Return													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\widehat{d}$			
	$\widetilde{S}_N$	$\widetilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\widetilde{\kappa}$	$\widetilde{s}$	$\widehat{d}^{IR}$	$\widehat{d}^{MS}$	$\widehat{d}^{ROB}$	$\widehat{d}^{WAV}$
[1 : 6672]	S	S	S	S	S	S	SM	12.6	-0.6	0.083	0.067	0.084	0.022
[6673 : 10400]	S	S	S	S	S	S	SM	63.7	-2.3	-0.021	-0.016	-0.013	-0.039
[10401 : 15919]	S	S	S	S	S	S	SM	9.0	-0.1	0.033	0.047	-0.005	-0.015

$ r = abs(\text{Nikkei 225A Index Return})$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\widehat{d}$			
	$\widetilde{S}_N$	$\widetilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\widetilde{\kappa}$	$\widetilde{s}$	$\widehat{d}^{IR}$	$\widehat{d}^{MS}$	$\widehat{d}^{ROB}$	$\widehat{d}^{WAV}$
[1 : 6672]	S	S	S	S	NS	NS	LM	26.1	3.3	0.302	0.343	0.313	0.218
[6673 : 10400]	S	S	S	S	NS	NS	LM	150.9	7.5	0.196	0.346	0.304	0.321
[10401 : 15919]	S	S	S	S	NS	NS	LM	17.0	2.6	0.413	0.415	0.431	0.335

$r^2 = (\text{Nikkei 225A Index Return})^2$													
Segments Breaks	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\widehat{d}$			
	$\widetilde{S}_N$	$\widetilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\widetilde{\kappa}$	$\widetilde{s}$	$\widehat{d}^{IR}$	$\widehat{d}^{MS}$	$\widehat{d}^{ROB}$	$\widehat{d}^{WAV}$
[1 : 6672]	S	S	S	S	NS	NS	LM	427.8	16.9	0.275	0.241	0.267	0.080
[6673 : 10400]	S	S	S	S	S	NS	LM	2610.8	48.0	0.230	0.146	0.154	0.117
[10401 : 15919]	S	S	S	S	NS	NS	LM	235.5	12.3	0.381	0.396	0.363	0.377

$ r ^\theta = (abs(\text{Nikkei 225A Index Return}))^\theta$													
$\widehat{\theta}_i^{(j)} =$ $\text{ArgMax}_\theta(\widehat{d}(r_i ^\theta))$	(Non)Stationarity Test						LRD	Kurtosis	Skewness	$\widehat{d}$			
	$\widetilde{S}_N$	$\widetilde{T}_N$	ADF	PP	KPSS	LMC	V/S	$\widetilde{\kappa}$	$\widetilde{s}$	$\widehat{d}^{IR}$	$\widehat{d}^{MS}$	$\widehat{d}^{ROB}$	$\widehat{d}^{WAV}$
$\widehat{\theta}_1^{(IR)} = 1.53$	S	S	S	S	NS	NS	LM	149.8	9.0	0.323*	0.296	0.296	0.201
$\widehat{\theta}_1^{(MS)} = 0.86$	S	S	S	S	NS	NS	LM	15.4	4.4	0.286	0.345*	0.311	0.213
$\widehat{\theta}_1^{(ADG)} = 1.01$	S	S	S	S	NS	NS	LM	27.1	3.4	0.303	0.342	0.313*	0.218
$\widehat{\theta}_1^{(WAV)} = 1.30$	S	S	S	S	NS	NS	LM	74.8	6.1	0.273	0.322	0.307	0.622*
$\widehat{\theta}_2^{(IR)} = 3.00$	S	S	S	S	S	S	LM	3487.4	58.3	0.252*	0.037	0.042	0.045
$\widehat{\theta}_2^{(MS)} = 0.84$	S	S	S	S	NS	NS	SM	55.5	4.1	0.180	0.353*	0.304	0.154
$\widehat{\theta}_2^{(ADG)} = 0.91$	S	S	S	S	NS	NS	SM	87.1	5.3	0.186	0.352	0.305*	0.035
$\widehat{\theta}_2^{(WAV)} = 1.64$	S	S	S	S	S	NS	SM	1697.1	35.8	0.221	0.220	0.222	0.465*
$\widehat{\theta}_3^{(IR)} = 1.23$	S	NS	S	S	NS	NS	LM	35.7	4.1	0.467*	0.412	0.426	0.386
$\widehat{\theta}_3^{(MS)} = 0.87$	S	S	S	S	NS	NS	LM	11.0	2.0	0.429	0.415*	0.428	0.351
$\widehat{\theta}_3^{(ADG)} = 1.00$	S	S	S	S	NS	NS	LM	17.0	2.6	0.413	0.415	0.431*	0.335
$\widehat{\theta}_3^{(WAV)} = 1.27$	S	NS	S	S	NS	NS	LM	40.3	4.4	0.467	0.411	0.424	0.425*

TABLE 4.13 – Results of stationarity, nonstationarity and V/S tests and the 4 long memory parameter estimators applied to several functionals f of Nikkei 225A Index Return : from the top to bottom, $f(x) = x$, $f(x) = |x|$, $f(x) = x^2$ and $f(x) = |x|^\theta$ with θ maximizing the 4 different long memory parameter estimators (“S” for “stationarity” decision and “NS” for “nonstationarity” decision). Statistics are applied to the 3 estimated stages of each trajectory (obtained from a change detection algorithm).

Chapitre 5

Conclusion et Perspectives

Le projet de restructuration de l'économie par l'introduction du paradigme de mémoire longue en son sein, à l'instar de la finance, trouve, en pratique, comme en théorie, à priori, toute sa légitimité et toute sa justification en économie quelque'elle soit : classique, keynésienne, néo-classique ou néo-keynésienne dans une refonte dynamique stochastique globale réintégrant les effets et les conséquences lointains. La légitimité pratique proviendrait à priori des politiques économiques elles mêmes comme le préconisait Henry Hazlett dans sa formule, désormais célèbre : *“Les effets et les conséquences de long terme sont inscrits dans toute politique économique (quelqu'elle soit) comme inscrites la poule dans l'œuf, la fleur dans la graine”*. La légitimité théorique proviendrait à priori des modalités mêmes de collectes en haute ou en basse fréquence, de toutes ces masses d'informations, dans leurs structure, comme dans leur chronologie, transformées en bases de données, d'agrégats catégoriels, sectoriels, chronologiques, collectes devenues indispensables à l'élaboration de plans et de politiques économiques. Or ces données peuvent être fortement corrélées entre elles, dans une analyse longitudinale ou transversale et donc source de processus de longue mémoire, tendanciels, saisonniers, ou cycliques, gouvernant tout ou partie de ces séries chronologiques et non chronologiques. Le marché du capital sous quelque paradigme qui le sous tend que se soit, assure quant à lui les niveaux optimaux pour un échange intra et inter sectoriels indépendamment des doctrines économiques, échange à la fois quantifiable en terme de valeurs quelque soit le standard de mesure utilisé, mais aussi chronologiquement repérable, dans l'espace et dans le temps. Il n'en demeure pas moins que ce qui est vrai pour le marché du capital restera vrai aussi, pour tout autre type de marché (de biens et services, d'emploi, de monnaie, de change...ou de toute autre forme de marché en conformité avec sa doctrine et son paradigme fondateurs).

Il nous paraît inéluctable l'introduction d'effet de structure d'auto-similarité en analyse économique, et par voie de conséquence, en pensée économique, en conformité avec les découvertes de lois d'échelles en économie par Mandelbrot modélisées ensuite par Granger et Joyeux(1980)([12], p : 35-36) avec les prémisses d'une investigation plutôt d'ordre économique, comme il a déjà été men-

tionné par Mignon et Lardic ([61], p 328), sans investigation épistémologique accompagnatrice, pour autant, qui la sous tend, la justifie et l'encadre.

L'internationalisation et la délocalisation des marchés, via l'interconnexion de leurs structures détentrices d'actifs internationaux, comme de leurs sources d'informations, constituent de fait un relais de transmission, de tout dérèglement ou dysfonctionnement, à l'ensemble des marchés internationaux, accompagnés, le plus souvent, par des effets amplificateurs voire multiplicateurs de paniques généralisées. Trois questions restent en suspens :

1. Comment endiguer la spéculation débridée, source des dérèglements financiers chroniques avec leurs réactions en chaîne sur tous les marchés financiers ?
2. Comment consolider la confiance des agents économiques contre les affolements délibérés et les déroutes injustifiés ?
3. Comment garantir la fiabilité de toute assurance de convertibilité juste et équitable de biens, services et monnaie, sur tous les marchés, sans qu'il y ait ni pénurie ni excès d'abondance (Certains économistes ont vu dans l'effondrement du marché des biens immobiliers aux états Unis débutant en 2006 et atteignant son apogée en 2008 comme le vrai déclencheur de la crise financière actuelle) ?

Aussi bien dans son livre de référence [89] que dans le rapport de sa commission [25], Stiglitz, le néokeynésien, prix Noble en Sciences Economiques en 2001, évoque plusieurs autres problématiques plus aigües, plus promptes et plus globales et y propose à la fois des solutions qui leurs sont dédiées, mais aussi et surtout l'ébauche de nouveaux principes fondateurs d'une nouvelle science économique, plus à même d'y répondre correctement, théoriquement, du moins, à toutes les incohérences connues par le passé jusqu'à aujourd'hui, en son sein, et qui hypotéquaient très sérieusement, autres fois, son fonctionnement normal.

Il n'en demeure pas moins que toutes ces propositions de stiglit ne concernent de fait que les agents et les institutions ou les États, sans avoir aucun volet écosystémique doctrinal global explicite. Il semble pourtant que la réhabilitation de la théorie utilitariste intégrale, de bonheur et de bien-être quantitatifs et qualitatifs, de John Stuart-Mill, au vu de ses exigences doctrinales, constitue une entreprise très prometteuse pour un avenir écosystémique meilleur, au prix d'une dynamisation structurelle de longue portée. Cela semble constituer à l'avenir un vrai défi à relever par l'économie et les économistes.

Au niveau statistique l'ensemble de nos résultats est concluant, sans recourir à trop d'hypothèses restrictives, autres que les hypothèses $SIG(d, \beta)$. Les temps de calculs restent convenables et ce pour l'ensembles des séries réelles et simulées. La seule limitation de notre méthode adaptative reste son incapacité à détecter les ruptures au sein d'une série chronologique.

D'autres idées demeurent pourtant en friche, constituant de vrais défis à relever. Elles concernent, entre autres, la possibilité ou non d'étendre notre approche semi-paramétrique adaptative à d'autres types des structures de processus comme les processus linéaires, ou les processus vectoriels gaussiens cointégrés et/ou cointégrés fractionnaires.

Bibliographie

- [1] Abadir K.M., W. Distaso and L. Giraitis (2007)., *Nonstationarity-extended local whittle estimation.*, Journal of Econometrics.
- [2] Abry P. and D. Veitch (1998), *Wavelet analysis of long-range-dependent traffic*, IEEE Trans. Inform. Theory, Vol. 44, 2-15.
- [3] Abry P., M.S. Taqqu and D. Veitch (2003), *On the automatic selection of the onset of scaling*, Fractals, Vol. 11, 377-390.
- [4] Baillie R.T. and M.L. King (1996) (ed.), *Fractional differencing and long memory process*, Journal of econometrics, Vol. 73, 1.
- [5] Barbosa S. M., E. Caporappli, S. Fatichi and M.E. Silva (2009), *Deterministic versus stochastic trends : Detection and challenges*, J. Geophys. Res., Vol. 114.
- [6] Bardet J.M. and B. Dola (2012) , *Adaptive estimator of the memory parameter and goodness-of-fit test using a multidimensional increment ratio statistic*, Journal of Multivariate Analysis, Vol. 105, Issue 1, 222-240.
- [7] Bardet J.M. and D. Surgailis (2011), *Measuring the roughness of random paths by increment ratios*, Bernoulli Vol. 17, Issue 2, 749-780.
- [8] Bardet J.M. and H. Bibi (2013)., *Adaptive semiparametric wavelet estimator and goodness-of-fit test for long-memory linear processes.*, Electronic Journal of Statistics.
- [9] Bardet J.M., G. Lang, G. Oppenheim, A. Philippe, S. Stoev and Taqqu, M.S. (2003), *Semiparametric estimation of the long-range dependence parameter : a survey.*
- [10] Bardet J.M., H. Bibi and A. Jouini (2008), *Adaptive wavelet-based estimator of the memory parameter for stationary gaussian processes*, Bernoulli, 14, 691-724.
- [11] Barros C.P, L.A. Gil-Alana and J.E. Payne (2012), *Evidence of long memory behavior in u.s. renewable energy consumption*, Energy Policy, Vol. 41, 822-826.
- [12] Beran J. (1994), *Statistics for long-memory processes*, Chapman and Hall . One Penn Plaza . New York. NY 10119.
- [13] Black D. (2000), *Dreams of pure sociology*, Sociological Theory, Vol. 18, Issue 1, 343-367.

- [14] Boland L.A. (1978), *Time in economics vs. economics in time : the hayek problem*, Canadian Journal of Economics, 11, 24062.
- [15] Borgard F., D. Bosq, M. Carbon, C. El-Nouty, D. Guégan, E. Guerre and D. Smili (1991), *Processus longue mémoire*, Tech. Report XXXVI-Fascicule 1-2, Institut de Statistique de l'Université de Paris.
- [16] Bourgeois T. (1999), *Dictionnaire d'histoire et philosophie des sciences, s/s direction : Dominique lecourt*, PUF.
- [17] Bourguine P. and C. Roth (2005), *Epistemic communities : description and hierarchic categorization*, Internet Adress Site <http://arxiv.org/abs/nlin.AO/0409013>.
- [18] Brown G.D.A., R.F. Cancho, J.G. Holden, C.T. Kello, K.L.H.T Rhodes and G.C. Van Orden (2010), *Scappling laws in cognitive sciences*, Trends In Cognitive Science, Vol. 14, Issue 5, Pages 223-232.
- [19] Brown J., W. Rea and M. Reale (2011), *Long memory in temperature reconstructions*, Climate Change, Vol. 107, 247-265.
- [20] Bruzaite, K. and Vaiciulis, M. (2008). , *The increment ratio statistic under deterministic trends.*, Lith. Math. J., 48, 256-269.
- [21] Calvet C. and A. Cotta A. (2005), *Les quatre piliers de la science économique*, Fayard.
- [22] Chatelain J.B and K. Ralf (2012), *The failure of financial macroeconomics and what to do about it*, W.P.2012.30 CES Univ. Paris1.
- [23] Collectif (1984), *Epistémologie des sciences sociales*, Tech. Report Revue trimestrielle, Vol. X X X V I, Issue 4, Revue Internationale des sciences sociales, 102.
- [24] Commission présidée par Stiglitz J.E. , A. Sen and J.P Fitoussi (2009), *la mesure des performances économiques et du progrès social*, www.stiglitz-sen-fitoussi.fr.
- [25] Commission Stiglitz J.E. (2010)-Traduit par Chemla(2012), *Le rapport stiglitz pour une vraie réforme du système monétaire et financier international après la crise mondiale*, Actes Sud.
- [26] Dahlhaus R. (1989, 2006), *Efficient parameter estimation for self-similar processes.*, Annals of Statistics, 17, 1749-1766. Corrected in Vol. 34 (2006), 2, 1045-1047.
- [27] Davidson J. and T. Teräsvirta (2002) (ed.), *Long memory and nonlinear time series*, Journal of econometrics, Vol. 110, 2.
- [28] Doukhan P. (2006), *Processus empiriques et séries temporelles*, ENSAE.
- [29] Doukhan P. and G. Oppenheim and M.S. Taqqu (2003) (ed.), *Theory and applications of long-range dependence*, Birkhäuser.
- [30] Dülger F. and Z.A. Ozdemir (2011) , *Current account sustainability in seven developed countries*, Journal of Economic and Social Research, Vol. 7, Issue 2, 47-80.
- [31] Elder J. and P.E. Kennedy (2001), *Testing for unit roots : What should students be taught ?*, Journal of Economic Education, Vol. 32, 137-146.

- [32] Favero C.A. (2007), *Model evaluation in macroeconometrics : from early empirical macroeconomics models to dsge models*, W.P.2007.327 IGIER Univ. Bocconi Italy.
- [33] Feller W. (1970) , *An introduction to probability theory and its applications, vol.2, second edition*, John Wiley sons, New York, USA.
- [34] Fox R. and M.S. Taqqu (1986), *Large-sample properties of prameter estimates for strongly dependent stationary gaussian time series*, Annals of Statistics, 14, 517-532.
- [35] Geweke J. and S. Porter-Hudak (1983), *The estimation and application of long-memory time series models*, Journal of Time Series Analysis, 4, 221-238.
- [36] Giraitis L. and D. Surgailis (1990). , *A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of whittle estimate.*, Prob. Th. and Rel. Field., 86, 87-104.
- [37] Giraitis L. and M.S. Taqqu (1999)., *Whittle estimator for finite-variance non-gaussian time series with long memory.*, Ann. Statist., 27, 178-203.
- [38] Giraitis L., P. Kokoszka, R. Leipus and G. Teyssière (2003)., *Rescaled variance and related tests for long memory in volatility and levels.*, Journal of Econometrics, 112, 265-294.
- [39] Giraitis L., R. Leipus and A. Philippe (2006)., *A test for stationarity versus trends and unit roots for a wide class of dependent errors.*, Econometric Theory, 22, 989-1029.
- [40] Giraitis, L., Robinson P.M., and Samarov, A. (1997)., *Rate optimal semiparametric estimation of the memory parameter of the gaussian time series with long range dependence.*, J. Time Ser. Anal., 18, 49-61.
- [41] Gneiting T., D.B. Percival, D.A. Rothrock and A. S. Thorndike (2008) , *The variance of mean sea-ice thickness : Effect of long-range dependence*, J. Geophys. Res., Vol. 113, C01004.
- [42] Gouriéroux C. and A. Monfort (1995), *Séries temporelles et modèles dynamiques*, Economica, 2e édition.
- [43] Gouriéroux C. and J. Jasiak (2001), *Memory and infrequent breaks*, Economic Letters, 70, 29-47.
- [44] Granger C.W.J. and N. Hyung (2004), *Occasional structural breaks and long memory with an application to the sp 500 absolute stock returns*, Journal of Empirical Finance, 11, 399-421.
- [45] Granger C.W.J. and R. Joyeux (1980), *An introduction to long-memory time series models and fractional differencing*, Journal of Time Series Analysis, 1, 15-29.
- [46] Granger C.W.J. and T. Terasvirta (1999), *A simple nonlinear time series model with misleading linear properties*, Economic Letters, 62, 161-165.
- [47] Guégan D. (2003), *Les chaos en finance*, Economica, Paris, France.
- [48] Guégan D. (2005), *How can we define the concept of long memory? an econometric survey*, Econometric Reviews, Vol. 24, Issue 2, 113149.

- [49] Hamilton J.D. (1994), *Time series analysis*, Princeton University Press, Princeton, New Jersey.
- [50] Heyde C.C. (2002), *On modes of long-range dependence*, Journal of Applied Probability, Vol. 39, 882-888.
- [51] Heyde C.C. and Y. Yang (1997), *On defining long-range dependence*, Journal of Applied Probability, Vol. 34, 939-944.
- [52] Hinti S. (2000), *La science économique : une réflexion épistémologique et méthodologique*, Najah Al Jadida, Casablanca, Maroc.
- [53] Ho H.C. and T. Hsing (1997). , *Limit theorems for functionals of moving averages.*, Ann. Probab., 25, 1636-1669.
- [54] Hosking J.M.R (1981), *Fractional differencing*, Biometrika, Vol. 68, 1, 165-176.
- [55] Hurst H.E. (1951), *long-terme storage capacity of reservoirs*, Transactions of the American Society Civil Engineers, Vol. 116, 770-799.
- [56] Hurvich C.M. and B.K Ray (1995), *Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes*, Journal of Time Series Analysis, Vol. 16, 17-41.
- [57] Könning T., M. Lohre and P. Sibbertsen (2003), *Modeling water flow of the rhine river using seasonal long memory*, Water Resour. Res., Vol. 39, 1132.
- [58] Koutsoyiannis D. and A. Montanari (2007) , *Statistical analysis of hydroclimatic time series : Uncertainty and insights*, Water Resour. Res., Vol. 43, W05429.
- [59] Kropp J.P. and H.J. Schellnhuber (2011) , *In extremis : disruptive events and trends in climate and hydrology*, isbn : 9783642148620, Springer, Berlin.
- [60] Laguës M. and A. Lesne (2003), *Invariances d'échelle, des changements d'états à la turbulence*, Belin.
- [61] Lardic S. and V. Mignon (2002), *économétrie des séries temporelles macroéconomiques et financières*, Economica, Paris, France.
- [62] Lavancier F., A. Philippe and D. Surgailis (2010), *A two-sample test for comparison of long memory parameters*, Journal of Multivariate Analysis archive, Vol. 101, 9, 2118-2136.
- [63] Lesne A. (1996), *Méthodes de renormalisation*, Eyrolles.
- [64] Leybourne S. J. and B. P. M. McCabe (1994), *A consistent test for a unit root*, Journal of Business and Economic Statistics, Vol. 12, 157-166.
- [65] Leybourne S. J. and B. P. M. McCabe (1999), *Modified stationary tests with data dependent model selection rules*, Journal of Business and Economic Statistics, Vol. 17, 264-270.
- [66] Ling S. (2003), *Adaptive estimators and tests of stationary and non-stationary short and long memory arima-garch models*, J. Amer. Statist. Assoc., Vol. 92, 1184-1192.
- [67] Ling S. and W.K. Li (2001), *Asymptotic inference for nonstationary fractionally integrated autoregressive moving-average models*, Econometric Theory, Vol. 17, 738-765.

- [68] Mestre O., H.W. Rust and V. K. C. Venema (2008), *Fewer jumps, less memory : Homogenized temperature records and long memory*, J. Geophys. Res., Vol. 113, D19110.
- [69] Montanari A. (2012), *Hydrology of the po river : looking for changing patterns in river discharge*, Hydrol. Earth Syst. Sci. Discuss., 9, 6689 – 6713.
- [70] Moulines E. and Ph. Soulier (2003), *Semiparametric spectral estimation for fractional processes, 251-301*, P. Doukhan and M.S. Taqqu.
- [71] Moulines E., F. Roueff and M.S. Taqqu (2007)., *On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter.*, J. Time Ser. Anal., 28, 155-187.
- [72] Mudelsee M. (2007), *Long memory of rivers from spatial aggregation*, Water Resour. Res., Vol. 43, W01202.
- [73] Nielsen M.O. (2004), *Efficient likelihood inference in nonstationary univariate models*, Econometric Theory, Vol. 20, 116-146.
- [74] Niquidet K. and L. Sun (2012), *Do forest products prices display long memory ?*, Canadian Journal of Agricultural Economics, forthcoming.
- [75] Petherick A. (2011), *Long term temperature*, Climatic Change, Vol. 107, 247-265.
- [76] Phillips P.C.B and K. Shimotsu (2005), *The exact local whittle estimation of fractional integration*, The Annals of Statistics, Vol. 33, Issue 4, 1890-1933.
- [77] Quine W.V.O. (1951), *Two dogmas of empiricism*, The Philosophical Review 60 : 20-43. Reprinted in his 1953 From a Logical Point of View. Harvard University Press.
- [78] Resnick S. I. (2007), *Heavy-tail phenomena, probabilistic and statistical modeling*, Springer.
- [79] Rioublanc S. (2005), *Modèles à changements de régimes markoviens : Détection des régimes, mémoire courte ou mémoire longue et prévision*, Ph.D. thesis, école Normale Supérieure de Cachan.
- [80] Robinson P.M. (1994), *Efficient tests of nonstationary hypotheses*, Journal of the American Statistical Association, Vol. 89, 1420-37.
- [81] Robinson P.M. (1995) , *Gaussian semiparametric estimation of long-range dependence*, The Annals of Statistics, Vol. 23, 1630-1661.
- [82] Robinson P.M. (1995), *Log-periodogram regression of time series with long range dependence*, The Annals of Statistics, Vol. 23, 1048-1072.
- [83] Robinson P.M. (2003) (ed.), *Time series with long memory, advanced texts in econometrics*, Oxford University Press.
- [84] Robinson P.M. and C. Velasco (2000), *Whittle pseudo-maximum likelihood estimation for nonstationary time series*, Journal of the American Statistical Association, December 2000, Vol. 95, Issue 452, Theory and Methods, 1229-1243.

- [85] Robinson P.M. and M. Rosenblatt (1996) (ed.), *Athens conference on applied probability and time series*, Lecture Notes in Statistics Journal 115, Vol. 2.
- [86] Roth C. (2005) , *Coévolution in epistemic networks restructuring social complex system*, Ph.D. thesis, école Polytechnique.
- [87] Samorodnitsky G. (2006), *Long range dependence*, Now Publishers Inc. PO Box 1024, Hanover, MA 02339, USA.
- [88] Schumpeter J.A. (1983), *Histoire de l'analyse économique, tome1 : L'âge des fondateurs*, Editions Gallimard.
- [89] Stiglitz J.E. (2010), *Le triomphe de la cupidité*, W.W. Norton & Company, Inc., New York, Copyright J. E. Stiglitz.
- [90] Stiglitz J.E. (2011), *Rethinking macroeconomics : What failed, and how to repair it*, Journal of the European Economic Association, 9, Vol. 4, 591-645.
- [91] Stiglitz J.E. and B. Greenwald (2005), *économie monétaire : un nouveau paradigme*, Economica.
- [92] Surgailis D., G. Teyssière and M. Vaičiulis (2008), *The increment ratio statistic*, Journal of Multivariate Analysis, 99, 510-541.
- [93] Tanaka K. (1999), *The nonstationary fractional unit root*, Econometric Theory, Vol. 15, 549-582.
- [94] Vaiciulis M. (2009)., *An estimator of the tail index based on increment ratio statistics.*, Lith. Math. J., 49, 222-233.
- [95] Veitch D., P. Abry and M.S. Taqqu (2003)., *On the automatic selection of the onset of scaling.*, Fractals, 11, 377-390.
- [96] Velasco C. (1999) , *Non-stationary log-periodogram regression*, Journal of Econometrics, Vol. 91, 325-371.

Résumé : Les modalités d'investigation, de notre quête des solutions, aux "*Problèmes économétriques d'analyse des séries temporelles à mémoire longue*", objet de cette thèse, en sciences économiques, sont menées ici, sous trois angles, dans trois parties distinctes, inhérentes aux approches : *épistémologique*, *statistique* et *économique*. En première partie de la thèse, dans une *approche épistémologique*, nous spécifions en quoi le concept de mémoire longue peut apparaître, pour les sciences économiques, même dans une acception modérée, comme un nouveau *paradigme* kühnien, ou une nouvelle *matrice disciplinaire*, un nouvel *épistémé* dans le sens de Michel Foucault, un nouveau *programme scientifique de recherche* dans le sens de Lakatos, pour la macroéconomie et la finance. En deuxième partie de la thèse, dans une *approche statistique*, semi-paramétrique, nous proposons trois extensions de la statistique IR (Increment Ratio) de Surgailis et al, (2008), la plus élaborée de la mémoire longue, connue à ce jour, extensions relatives à des processus, stationnaires ou à accroissements stationnaires, gaussiens. Premièrement, un théorème central limite multidimensionnelle est établi pour un vecteur composé de plusieurs statistiques IR. Deuxièmement, un test d'adéquation de qualité d'ajustement de type χ^2 est déduit de ce théorème. Troisièmement, ce théorème nous a permis de construire des versions adaptatives de l'estimateur et du test d'adéquation étudiés dans un cadre semi-paramétrique général. Nous prouvons que l'estimateur adaptatif du paramètre de la mémoire longue suit une propriété d'Oracle. Les simulations que nous avons menées attestent de la précision et de la robustesse de l'estimateur et du test d'adéquation, même dans le cas non gaussien. En troisième partie de la thèse, nous en déduisons deux tests respectivement de stationnarité et de non stationnarité pour les processus $I(d)$ stationnaires et non stationnaires, pour tout réel d tel que $(-0.5 < d < 1.25)$. Nos deux tests s'avèrent nettement meilleurs sur l'ensemble des valeurs de d de cet interval $(-0.5 < d < 1.25)$ que l'ensemble des tests standards de stationnarité (comme le kpsstest ou le lmctest de Leybourne et Mac Cabe) ou de non stationnarité (comme l'adftest augmenté de Dickey-Fuller, ou le pptest de Phillips et Perron). De plus l'implémentation du test V/S, une amélioration du test R/S modifié de Lo, nous a permis, de distinguer la mémoire courte de la mémoire longue pour n'importe quel processus LRD. Dans une *approche économique*, au sein de cette troisième partie de la thèse, nous mettons en oeuvre les résultats théoriques précédents comparés à ceux issus d'autres méthodes statistiques : paramétriques, semi-paramétriques ou non paramétriques (ou heuristiques) appliquées à des séries économiques et financières.

Mots clés : Econométrie ; Séries chronologiques ; Statistique mathématique ;
Stiglitz, Joseph Eugene(1943-...) ; Mill, John Stuart(1806-1873).

Abstract : The investigation modes, to solve the "*Econometric Problems of the long memory time series analysis*", object of this thesis, in economic sciences, are carried out here, under three angles, in three distinct parts, inherent to *epistemological*, *statistical* and *economical* approaches. In the first part of the thesis, we specify, in an *epistemological approach*, in what the concept of long-memory can appear, for economics, even in a moderate meaning, as a new kühnian "*paradigm*" or a new "*disciplinary matrix*", a new "*épistémé*" in the Michel Foucault meaning, or a new "*scientific research program*" in the Lakatos sense, for macroeconomics and finance. In the second part of the thesis, we propose, in a semiparametric *statistical approach*, three extensions of the most elaborate IR (Increment Ratio) statistics of Surgailis et al, (2008), among the long-memory statistics, relating to stationary or stationary increments Gaussian processes. First, a multidimensional central limit theorem is established, for a vector made up of several IR statistics. Second, a χ^2 goodness of fit test is deduced from this Multidimensional central limit theorem. Lastly, this Multidimensional central limit theorem allowed us to build adaptive versions of the estimator and the goodness of fit test, studied within a general semiparametric framework. We prove that the adaptive estimator of the long memory parameter follows an Oracle property. Simulations which we carried out attest of the precision and the robustness of the estimator and the goodness of fit test, even in the non Gaussian case. In the third part of the thesis, we deduce from our adaptive Multidimensional Increment Ratio statistics, two tests respectively of stationarity and nonstationarity, for any stationary or nonstationary $I(d)$ process, and for any real number d such that $(-0.5 < d < 1.25)$. Our Two tests are clearly better on the whole interval $(-0.5 < d < 1.25)$ than the well known standard stationary tests (as the kpsstest or the lmctest of Leybourne and Mac Cabe) or the nonstationary tests (as the Augmented adftest of Dickey-Fuller, or the pptest of Phillips and Perron). Moreover the implementation of V/S test, an improvement of the modified test R/S of Lo, allowed us to distinguish between the short and the long memory for any LRD process. In an empirical econometric approach, we implement these previous theoretical results, and compare them with those resulting from other statistical methods : parametric, semiparametric or nonparametric (or heuristic) applied to economical and financial series.

Keywords : Econometrics ; Time Series ; Mathematical statistics ;
Stiglitz, Joseph Eugene(1943-...) ; Mill, John Stuart(1806-1873).