UNIVERSITÉ PARIS I PANTHÉON-SORBONNE
U.F.R. MATHÉMATIQUES

N^o attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|

THÈSE DE DOCTORAT

Présentée par

Hatem BIBI

Pour obtenir le grade de

DOCTEUR EN SCIENCES
SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

Construction des bases d'ondelettes de $L^2([0,1])$ & Estimation du paramètre de longue mémoire par la méthode des ondelettes

Sous la Direction de

Pr. Jean-Marc BARDET & **Pr. Abdellatif JOUINI**

Soutenue le 4 Novembre 2011

Composition du jury

Antoine Ayache	Examinateur	Professeur de l'Université Lille I
Jean-Marc Bardet	Co-directeur	Professeur de l'Université Paris 1
Paul Doukhan	Président	Professeur de l'Université Cergy-Pontoise
Abdellatif Jouini	Co-directeur	Professeur de l'Université de Tunis
François Roueff	Rapporteur	Professeur TELECOM ParisTech
Mohamed Sifi	Rapporteur	Professeur de l'Université de Tunis

A ma grand mère TIMA et mon beau père Si El HEDI

Remerciements

Je tiens tout d'abord à adresser mes plus vifs remerciements à mes deux encadreurs, Pr Jean-Marc Bardet et Pr Abdellatif Jouini, ils ont été présents et patients pendant toutes ces années de thèse, leur aide m'était vraiment inestimable.
Je remercie les membres du jury pour l'intérêt qu'ils ont porté à ma thèse. Pr.Sifi et Pr.Roueff d'avoir accepté d'être rapporteurs, Pr Doukhan qui a présidé ce jury et Pr.Ayache d'y avoir été examinateur.

Je remercie Marie Cottrell qui m'a permis d'intégrer le laboratoire du SAMM. Pour son soutien et sa gentillese, et grâce à qui, règne une convivialité sans pareil au SAMM.

A les membres du SAMM, merci à vous tous : Xavier, Annie, Patrice, Patrick, Cecile, Ciprian, Giles, Joseph, Joel, Corrine, Omar, Olivier, Imen, Bechir, Mohamed, Madalina, Cecile, Vincent, Xavier, Charles, Ciprian, Sandie, Riadh, Catherine, Fania, Khalifa, Aurelien, Lotfi, Hugo, Jihed, Denis, Kader, Dhaou, Moez, Solesnes, Nathanael, Lionel.

A Hédi qui m'a beaucoup aidé à Paris, Enak et Mr Hammadi.
A ma femme qui a beaucoup sacrifié. Elle m'a épaulé et a su être patiente pendant ces années, à mes enfants Hamza et Khalil. A mes parents à qui je dois tout, à ma belle famille qui m'a soutenu, et enfin à tous ceux qui ont contribué à cette thèse de prés ou de loin.

Table des matières

Résumé

L'analyse par ondelettes est présentée dans cette thèse comme alternative à l'analyse de Fourier dans deux domaines, à savoir la construction des bases sur l'intervalle d'une part et l'estimation du paramètre de longue mémoire d'autre part.

La construction des bases d'ondelettes débuta au début des années 90. Ces constructions utilisaient les bases splines jusqu'a la découverte des bases de Daubechies et l'introduction de l'analyse multirésolution (AMR).

La méthode AMR donne toutes les bases d'ondelettes connues à ce jour pour l'espace des fonctions de carré sommable $L^2(\mathbb{R})$. Une telle analyse consiste à décomposer le signal sur une gamme très étendue d'échelles, opération que l'on peut comparer à une cartographie

Mais qu'en est il de l'espace des fonctions de carré sommable sur $[0,1]$? Il ne suffit pas de prendre les restriction de ces bases à $[0,1]$. Cette initiative a été prise par Yves Meyer [9]. Il montra que les restrictions des fonctions d'échelles à l'intervalle formaient un système libre alors que les restrictions des ondelettes associées formaient un système lié.

Nous présentons une solution à cette dépendance dans le premier volet de cette thèse et généralisons la méthode d'Yves Meyer en partant d'une Analyse Multirésolution Orthogonale arbitraire à support compact. Pour aboutir au cas biorthogonale on utilise la méthode de dérivation et d'intégration. Comme applications, on étudie les espaces de Sobolev $H^s([0,1])$ et $H_0^s([0,1])$ pour $s \in \mathbb{N}$.

Pour ce qui est du second volet, nous présentons dans un cadre semiparamétrique, un estimateur adaptatif du paramètre de longue mémoire basé sur les ondelettes dans le cas de processus stationnaire gaussien puis linéaire. La propriété de comportement de la variance du coefficient d'ondelette en puissance suggère un estimateur obtenu par une simple régression dans un schéma log-log des échelles sur la variance empirique. L'estimateur obtenu vérifie alors un théorème limite central, pour lequel on estime l'échelle la vitesse maximale de convergence ainsi que l'échelle minimale à partir de laquelle les propriétés sont valides. Nous procédons à des différents ajustements afin d'obtenir des estimateurs adaptés à certaines conditions. Nous étudions alors les propriétés de consistance et robustesse de ces estimateurs. Les comparaisons avec les estimateurs existants confirment la performance de nos estimateurs et enfin un test d'adéquation est établi pour chaque cas.

Plan de la thèse

L'idée d'utiliser des bases d'ondelettes s'est imposée depuis que ces bases ont fait la preuve de leur efficacité dans le traitement du signal. Y. Meyer et P.G. Lemarié [8] découvrirent la première base orthonormale d'ondelettes dans la classe de Schwartz , puis I. Daubechies avec une famille de bases orthonormales d'ondelettes à support compact. Finalement grâce à la notion d'analyse multirésolution (ou AMR) introduite par S. Mallat, des algorithmes rapides d'analyse dans ces bases ont été mis en œuvre. De plus elles forment des bases inconditionnelles pour les espaces de Sobolev. Afin que les choses soient claires, on aura pris soin avant de définir cette AMR de présenter dans le chapitre 1 les notions d'ondelettes et leurs transformées.

L'AMR est alors présentée. Cette analyse de l'espace des fonctions de carrés sommables $L^2(\mathbb{R})$ que l'on présente dans une première partie suivie de l'analyse biorthogonale consiste à découper cet espace en une suite croissante de sous-espaces vectoriels fermés . Ces sous-espaces sont d'intersection égale à $\{0\}$ et de réunion dense dans l'espace. Chaque sous-espace est l'ensemble de toutes les approximations possibles d'un même signal à l'échelle associée au sous-espace. Le signal à analyser sera approximé par une succession de projections orthogonales sur les sous-espaces, on obtient ainsi une cartographie de ce signal. Les bases de $L^2(\mathbb{R})$ sont bien définies et ne présentent plus de difficultés quand à leurs constructions, cependant quand on entame l'espace $L^2([0,1])$, il ne suffit pas de prendre la restrictions de ces familles à l'intervalle [0,1], des effets de bord interfèrent dans les propriétés d'indépendance nécessaires à la construction de ces bases. Nous présentons alors la construction de bases d'ondelettes sur l'intervalle décrite par Y. Meyer [9]. Ce dernier montre en effet que dan le cas des fonctions d'échelles, leurs restrictions à l'intervalle conserve l'aspect de l'indépendance ce qui n'est pas le cas des familles d'ondelettes. Nous nous proposons dans le premier article alors de généraliser sa méthode en partant d'une Analyse Multirésolution Orthogonale arbitraire, pour aboutir par extension au cas biorthogonale obtenue par la méthode de dérivation et d'integration. L'étude des espaces fonctionnels de Sobolev $H^s([0,1])$ et $H^s_0([0,1])$ s'impose alors de lui même. Un résumé des principaux résultats s'ensuit. Ce travail [5] ayant fait l'objet d'une publication .

Le chapitre 2 est consacrée à l'estimation du paramètre de longue mémoire par ondelettes. Cette démarche s'inscrit dans la continuité naturelle des travaux qui débutèrent avec Abry *et al.* [2, 1] pour les processus autosimilaires. Bardet *et al.* [3] montra la consistance de cet estimateur dans un cadre semiparamétrique pour le cas gaussien, Moulines *et al.* [10] montrèrent l'optimalité de cet estimateur au sens du critère minimax. Le cas linéaire fut entrepris par Roueff et Taqqu [11]. Nous y apportons notre contribution en proposant un estimateur semi-paramétrique adaptatif pour des processus stationnaires à longue mémoire gaussien puis linéaire. On commence par présenter les notions de stationnarité, nous donnons la définition d'un processus à longue mémoire puis des exemples tels que les processus farima, brownien fractionnaire et le bruit gaussien fractionnaire. Dans la section 2.1, on présente les méthodes d'estimation du paramètre de longue mémoire suivies de la méthode des ondelettes et un résumé des principaux résultats suivis des articles dans leurs versions originale. L'article [4] traitant du cas gaussien a fait l'objet d'une publication, le cas linéaire a été soumis. On adoptera dans les cas la même procédure de construction de l'estimateur. Celle ci est basée sur la propriété de linéarité (après application du logarithme) de la variance du coefficient d'ondelette par rapport aux échelles, donc de la variance empirique des coefficients d'ondelettes. L'estimateur de D (paramètre de la longue mémoire) déduit par les moindres carrés vérifie alors un théorème limite central et à une vitesse qui dépendra d'un paramètre D'. C'est ce paramètre qui joue un rôle important dans l'estimation de la vitesse de convergence de l'estimateur qui sera soumis à différents ajustement adaptatif pour vérifier les bonnes propriétés que l'on exigera. Pour les cas gaussien puis linéaires. On procédera par des simulations à des vérifications de consistance et de robustesse des estimateurs correspondants. Un test d'adéquation pour le cas linéaire sera établi.

Bibliographie

[1] Abry, P., Veitch, D., Flandrin, P. Long-Range Dependence : Revisiting Aggregation with Wavelets, *Journal of Time Series Analysis*, 19(3) : 253-266, Mai 1998.

[2] Abry, P. and Veitch, D. Wavelet analysis of long-range-dependent traffic, *IEEE Trans.on Info.Theory*, 44, 2–15,1998.

[3] Bardet, J.M., Lang, G., Moulines, E. and Soulier, P. Wavelet estimator of long range-dependant processes, *Statist. Infer. Stochast. Processes*, Vol. 3,85-99, 2000.

[4] Bardet, J.M., Bibi, H. and Jouini, A. Adaptive wavelet-based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli*,(2008), 14, 691-724.

[5] Bibi, H., Jouini, A., Kratou, M. More general constructions of wavelets on the interval . *Communications in Mathematical Analysis*, vol.4 (1),pp.45-57,2008

[6] Jouini, A., Lemarié-Rieusset, P. G.
Analyses multirésolutions biorthogonales sur l'intervalle et applications, *Annales de L'I.H.P, Analyses non linéaire*, Vol.10, N^{o}4 (1993), pp453-476.

[7] Jouini, A. Constructions de bases d'ondelettes sur les variétés, Thesis, ORSAY, 1993.

[8] Lemarié, P.G. and Meyer, Y. Ondelettes et bases hilbertiennes. *Rev.Mat. Iberoamericana*, 2(1-2) :1–18, 1986.

[9] Meyer, Y. Ondelettes sur l'intervalle, *Revista Mathematica Ibero-americana*, Vol 7 (1991), pp115-134.

[10] Moulines, E., Roueff, F and Taqqu, M.S. Central limit theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context, *fractals*,15(4),301-313,2007.

[11] Roueff, F. and Taqqu, M.S. . Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *J. Time Ser. Anal.*, 30, 534–558(2009).

Première partie

Préambule & Principaux résultats

Chapitre 1

Construction de bases d'ondelettes sur l'intervalle

1.1 Introduction

En 1873, Dubois-Reymond construit une fonction continue de la variable réelle x et 2π-périodique dont la série de Fourier diverge en un point donné. Ce contre exemple amena A. Haar à se poser, puis à résoudre le problème de l'existence d'une base orthonormée $h_o(x), h_1(x), \ldots, h_m(x), \ldots$ de $L^2([0,1])$ ayant la propriété que, pour toute fonction continue $f(x)$, la série $\sum_0^\infty \langle f, h_m \rangle\ h_m(x)$ converge uniformément vers $f(x)$. Mais la construction du système de Haar ne convient pas à l'analyse et à la synthèse des espaces de Holder C^s, pour un certain s $\in]0,1[$.

Ce problème a été étudié depuis le travail de Pionnier de Haar. G. Faber et J. Schauder ont commencé par remplacer les fonctions $h_m(x)$ du système de Haar par leurs primitives $\Delta_m(x)$ puis par approximer une fonction continue sur [0,1] par les sommes partielles de la série $a + bx + \sum_1^\infty \alpha_m \Delta_m\ (x)$. Si $f(x)$ appartient à C_o^s ([0,1]), on a $\alpha_m = o\,(m^{-s})$ et réciproquement, si cette condition est vérifiée, la série $a + bx + \sum_1^\infty \alpha_m \Delta_m\ (x)$ converge vers $f(x)$ en norme C^s ([0,1]). Par contre, le système de Schauder (complété par 1 et x) ne peut plus servir à l'analyse de l'espace $L^2([0,1])$ car le coefficient α_m se calcule par

$$\alpha_m = f\Big((\frac{k+1}{2})2^{-j}\Big) - \frac{1}{2}\Big[f(k2^{-j}) + f((k+1)2^{-j})\Big]$$

n'a plus de sens si $f(x) \in L^2([0,1])$.

Pour corriger ce défaut de la base de Schauder, Ph. Franklin a eu l'idée d'orthogonormaliser la suite $1, x, \Delta_1(x), \ldots, \Delta_m(x), \ldots$ en utilisant le procédé de Gram-Schmidt. Mais le système de Franklin est un peu tombé dans l'oubli parce que les fonctions obtenues ne sont pas fournies par un algorithme aussi simple que celui des fonctions $h_m(x)$ du système de Haar.

1.2 Les ondelettes

La transformée en ondelettes est une solution à certaines difficultés que posaient la transformée de Fourier. Les transformées de Fourier et de Fourier à fenêtre glissante

sont respectivement des transformations globales et locales mais de résolution temporelles fixes. A l'opposé de la transformée d'ondelettes qui est à représentation temporelle variable et qui revêt d'autres aspects tels que l'inversion de l'analyse et la recherche de représentations parcimonieuses. Les transformées d'ondelettes sont obtenues par intégration d'un signal multiplié par des fonctions analysantes de base. Une question naturelle se pose alors. Peut-on reconstruire le signal d'origine à partir de sa transformée. Sous certaines hypothèses la réponse est oui. Il est même possible de reconstruire le signal à partir de valeurs discrètes de la transformée. Ainsi s'introduit la notion de la transformée discrète en ondelettes. En cherchant à minimiser le nombre d'informations discrètes nécessaires à la reconstruction du signal on est conduit à la notion de bases d'ondelettes. On introduit alors les bases d'ondelettes en partant de la notion d'analyse multirésolution qui fournit un cadre de décomposition d'un signal sous la forme d'une suite d'approximation croissante complétée par une suite de détails.
Enfin, on introduit les bases biorthogonales d'ondelettes dont l'idée est de relâcher les fortes contraintes que doit vérifier une ondelette engendrant une base orthonormée. La clé est de considérer deux ondelettes au lieu d'une seule avec un lien de dualité entre ces deux ondelettes.

La transformée en ondelettes

Définition (Ondelette mère). Une ondelette est une fonction ψ de $L^2(\mathbb{R})$ (appelée ondelette mère) vérifiant

$$\int_o^{+\infty} \mid \overset{\curlywedge}{\psi}(t\xi) \mid^2 \frac{dt}{t} = C_\psi, \quad \text{pour} \quad \xi \neq 0$$

où $\overset{\curlywedge}{\psi}$ est la transformée de Fourier classique de ψ donnée par :

$$\overset{\curlywedge}{\psi}(\lambda) = \int_{\mathbb{R}} \psi(x) e^{-i\lambda x} dx,$$

avec $0 < C_\psi < +\infty$ et C_ψ est indépendante de ξ.

Pour toute échelle $a \in \mathbb{R}_+^*$ et toute position $b \in \mathbb{R}$, on défini un atome de la transformée par :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi(\frac{t-b}{a}).$$

La famille $\{\psi_{a,b}\}$ est la famille d'ondelettes associées à ψ. La transformée continue en ondelettes de la fonction f est la famille des coefficients $C_f(a,b)$ définis par

$$\text{Analyse} \quad : \quad C_f(a,b) = \int_{\mathbb{R}} f(t)\overline{\psi_{a,b}(t)} dt = \langle f, \psi_{a,b} \rangle_{L^2(\mathbb{R})}, \quad a \in \mathbb{R}_+^*, b \in \mathbb{R}.$$

La formule de synthèse ou de reconstruction sous certaines conditions dites d'admissibilité est :

$$\text{Synthèse} \quad : \quad \int_{]0,+\infty[\times\mathbb{R}} C_f(a,b) \psi_{a,b}(t) \frac{dadb}{a^2} \quad \text{dans} \quad L^2(\mathbb{R}), \quad t \in \mathbb{R}.$$

La transformée continue en ondelettes (C_f) associe à un signal f une infinité de coefficients doublement indicés par $a \in \mathbb{R}_+^*$ et $b \in \mathbb{R}$. Il y a redondance de l'information.

La transformée discrète est la solution à cette redondance. On restreint le balayage des valeurs de (a,b) non plus à $\mathbb{R}_+^* \times \mathbb{R}$, mais à un sous ensemble discret. En se fixant $a_0 > 1$ et $b_0 > 0$, $p, n \in \mathbb{Z}$ et en prenant $a \in \{a_0^p\}_{p\in\mathbb{Z}}$ et $b \in \{na_0^p b_0\}_{p,n\in\mathbb{Z}}$. On se sert alors de la famille dénombrable d'ondelettes :

$$\psi_{n,p} = a_0^{p/2}\psi(a_0^p t - nb_0).$$

Le choix usuel de $a = 2$ et $b = 1$ est donné par le théorème de Shannon [3]. Nous notons :

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k).$$

Exemples d'ondelettes

- **L'ondelette de Haar :** C'est la plus simple des ondelettes. Elle vaut :

$$H(x) = \begin{cases} +1, & x \in [0, \frac{1}{2}[, \\ -1, & x \in]\frac{1}{2}, 1]. \end{cases}$$

- **Les dérivées de gaussiennes :** Soit une gaussienne $G(x) = e^{-\pi x^2}$. Si on note $\psi_n = \frac{\partial^n}{\partial x^n}G$ pour $n \in \mathbb{N}^*$. Alors l'ondelette ψ_n est $\mathcal{C}^\infty$ et admet n moments nuls. Sa transformée de Fourier est aussi une gaussienne.
- **L'ondelette de Morlet :** Il s'agit d'une gaussienne modulée $\psi(t) = e^{-i\pi t^2}e^{2i\pi k_0 t}$ Sa transformée de Fourier est une gaussienne décalée de k_0 et vaut $\hat{\psi}(\omega) = e^{-\pi(\omega - k_0)}$.

Dans la plupart des cas, les ondelettes sont définies par leurs filtres associés. L'ondelette n'aura pas alors une formule analytique mais c'est par un algorithme de reconstruction (du type algorithme de Mallat) que l'on peut y acceder. Comme par exemple les ondelettes de I. Daubechies (dbN) qui, au début des années 90, ont marqué une étape décisive dans l'histoire des ondelettes. Les ondelettes de Daubechies ont un support de longueur $2N - 1$ (N étant le nombre de moments nuls). Leurs régularités augmentent avec l'ordre, quand N est grand alors ψ appartient à $\mathcal{C}^{\mu N}$ avec $\mu \sim 0,206$.

Question : On peut se demander sous quelles conditions la famille $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}}$ est une base orthonormée de $L^2(\mathbb{R})$.
Réponse : La notion d'analyse multirésolution orthogonale répond à cette attente.

1.3 L'analyse multirésolution et bases orthonormées d'ondelettes

Un analyse multirésolution (ou AMR) de $L^2(\mathbb{R})$ est une famille $M = \{V_j\}_{j\in\mathbb{Z}}$ de sous espaces fermés de $L^2(\mathbb{R})$ vérifiant les propriétés suivantes :

i) $V_j \subset V_{j+1}$.

ii) $(f(x) \in V_j) \Leftrightarrow (f(2x) \in V_{j+1})$.

iii) $\bigcap_j V_j = \{0\}$ et $\overline{\bigcup_j V_j} = L^2(\mathbb{R})$.

iv) $(f(x) \in V_0) \Leftrightarrow (f(x) \in V_0)$ pour tout $k \in \mathbb{Z}$.

v) Il existe une fonction $g(x)$ dans V_0 tels que $\{g(x-k)\}_{k\in\mathbb{Z}}$ soit une base de Riesz pour V_0.

Si on désigne par P_j le projecteur orthogonal de $L^2(\mathbb{R})$ sur V_j c'est à dire : $P_jf = \sum_{k\in\mathbb{Z}} C_{j,k}g_{j,k}$ avec $g_{j,k}(x) = 2^{j/2}g(2^jx-k)$. Alors on a :

$$\lim_{j\to-\infty}\|P_jf\| = 0 \quad \text{et} \quad \lim_{j\to+\infty}\|f-P_jf\| = 0.$$

On introduit la notion de fonction d'échelle. La fonction d'échelle $\varphi \in L^2(\mathbb{R})$ est définie par :

$$\hat{\varphi}(\xi) = \frac{\hat{g}(\xi)}{\left(\sum_{k\in\mathbb{Z}}|\hat{g}(\xi+2k\pi)|^2\right)^{1/2}}.$$

φ vérifie :

i) $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ est une base orthonormée de V_0.

ii) Si on note $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^jx-k)$, alors $(\varphi_{j,k})_{k\in\mathbb{Z}}$ est une base orthonormée de V_j.

iii) $(\varphi_{j,k})_{j,k\in\mathbb{Z}}$ est une base hilbertienne de $L^2(\mathbb{R})$.

iv) On note W_j (L'espace d'ondelettes) tel que $V_{j+1} = V_j \oplus W_j$.

v) Le projecteur P_j s'exprime par $P_jf = \sum\limits_{k\in\mathbb{Z}} < f, \varphi_{j,k} > \varphi_{j,k}$.

vi) On a la relation $P_{j+1}oP_j = P_j = P_joP_{j+1}$.

La propriété 1.2.1) dans la définition d'une analyse multirésolution (A.M.R) donne l'existence d'une unique suite $\{\alpha_n\}_{n\in\mathbb{Z}}$ d'équation fonctionnelle vérifiée par φ :

$$\frac{1}{2}\varphi(\frac{x}{2}) = \sum_{n\in\mathbb{Z}}\alpha_n\varphi(x-n).$$

Si on pose $Q_j = P_{j+1} - P_j$. Alors l'opérateur Q_j est un projecteur orthogonal de $L^2(\mathbb{R})$ sur $W_j = V_{j+1} \cap (V_j)^{\perp}$. De plus. L'espace W_0 possède une base de Riesz $\{\psi(x-k)\}_{k\in\mathbb{Z}}$ où l'ondelette ψ est donnée par :

$$\widehat{\psi}(2\xi) = e^{-i\xi}\overline{m(\xi+\Pi)}\widehat{\varphi}(\xi).$$

Le projecteur Q_j s'exprime par :

$$Q_jf = \sum_{k\in\mathbb{Z}} < f, \psi_{j,k} > \psi_{j,k} \quad \text{avec} \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^jx-k).$$

La densité devient :

$$\bigoplus_j^{\perp} W_j = L^2(\mathbb{R}).$$

1.4 L'analyse multirésolution biorthogonale

Les ondelettes orthogonales engendrent des bases orthonormées et constituent des familles faciles à manier. Cependant, elles ne sont pas évidentes à construire du fait de leurs régularités ou de leurs définitions implicites. Relâcher la contrainte d'orthogonalité permet d'améliorer certaines caractéristiques des ondelettes tels que la forme ou la régularité tout en disposant des formules explicites pour les ondelettes. Il est alors possible de construire des ondelettes présentant des propriétés plus attractives au prix de l'introduction d'une difficulté supplémentaire dans le calcul. On construit alors deux ondelettes (en dualité) que l'on note ψ et ψ^*. Elles sont appelées ondelettes biorthogonales, d'ou l'analyse multirésolution biorthogonale.
Une analyse multirésolution biorthogonale de $L^2(\mathbb{R})$ est la donnée d'un couple d'analyses multirésolutions (V_j, V_j^*) de $L^2(\mathbb{R})$ tel que $L^2(\mathbb{R}) = V_o \oplus (V_o^*)^{\perp}$.

On a :

i) Soit P_j le projecteur oblique de $L^2(\mathbb{R})$ sur V_j parallèlement à $(V_j^*)^{\perp}$. Pour tout $f \in L^2(\mathbb{R})$, on a alors :

$$P_j o P_{j+1} = P_{j+1} o P_j = P_j, \quad \lim_{j \to -\infty} \|P_j f\|_2 = 0, \quad \lim_{j \to +\infty} \|P_j f - f\|_2 = 0.$$

ii) V_o et V_o^* ont respectivement les bases de Riesz $\{g(x-k)\}_{k\in\mathbb{Z}}$ et $\{g^*(x-k)\}_{k\in\mathbb{Z}}$ telles que :

$$< g(x), g^*(x-k) >= \delta_{o,k}. \tag{1.4.1}$$

iii) Le projecteur P_j s'écrit alors $P_j f = \sum_{k\in\mathbb{Z}} < f, g^*_{j,k} > g_{j,k}$.

iv) (V_j) et (V_j^*) sont des analyses multirésolutions, donc on a :

$$\begin{cases} \widehat{g}(2\xi) = m(\xi)\widehat{g}(\xi), \\ \widehat{g}^*(2\xi) = m^*(\xi)\widehat{g}^*(\xi). \end{cases}$$

v) $m(\xi)\overline{m^*(\xi)} + m(\xi+\pi)\overline{m^*(\xi+\pi)} = 1$, ou encore $\sum_{k\in\mathbb{Z}} \widehat{g}(\xi+2k\pi)\overline{\widehat{g^*}(\xi+2k\pi)} = 1$.

vi) Si on pose $Q_j = P_{j+1} - P_j$, alors Q_j est un projecteur sur $W_j = V_{j+1} \cap (V_j^*)^{\perp}$. parallèlement à $(W_j^*)^{\perp}$ avec $W_j^* = V_{j+1}^* \cap (V_j)^{\perp}$.

vii) L'espace W_o possède une base de Riesz

$$\{\gamma(x-k)\}_{k\in\mathbb{Z}} \quad \text{avec} \quad \widehat{\gamma}(2\xi) = e^{-i\xi}\overline{m^*(\xi+\pi)}\widehat{g}(\xi),$$

et l'espace $W_o^* = V_1^* \cap (V_o)^{\perp}$ a une base de Riesz

$$\{\gamma^*(x-k)\}_{k\in\mathbb{Z}} \quad \text{avec} \quad \widehat{\gamma^*}(2\xi) = e^{-i\xi}\overline{m(\xi+\pi)}\widehat{g}^*(\xi).$$

viii) Q_j s'écrit alors $Q_j f = \sum_{k\in\mathbb{Z}} < f, \gamma^*_{j,k} > \gamma_{j,k}$.

Une des propriétés fondamentales des analyses multirésolutions biorthogonales est leur compatibilité avec la dérivation. En effet, on considère g et g^* deux fonctions d'échelle conjuguées (g et g^* avec $g \in H^1(\mathbb{R})$), alors il existe deux fonctions d'échelle conjuguées $\widetilde{g}$ et $\widetilde{g}^*$ vérifiant 1.4.1 telles que l'on ait :

$$\begin{cases} g'(x) = \widetilde{g}(x) - \widetilde{g}(x-1), \\ \widetilde{g}^{*'}(x) = g^*(x+1) - g^*(x). \end{cases}$$

Il est clair que si g et g^* sont à support compact alors $\widetilde{g}$ et $\widetilde{g}^*$sont aussi à support compact. Cette méthode est appelée méthode de *dérivation et d'intégration* et a été introduite par P.G. Lemarié. Si on définit $\widetilde{P}_j$ par :

$$\widetilde{P}_j f = \sum_{k\in\mathbb{Z}} < f, \widetilde{g}^*_{j,k} > \widetilde{g}_{j,k},$$

alors on a la formule de commutation suivante :

$$\frac{d}{dx} o P_j = \widetilde{P}_j o \frac{d}{dx}.$$

1.5 Analyse Multirésolution orthogonale sur l'intervalle $[0,1]$

Définition Une suite $\{V_j\}_{j\geq j_0}$ des sous espaces fermés de $L^2([0,1])$ est dite analyse multirésolution de $L^2([0,1])$ associée à $V_j(\mathbb{R})$ si on a :

i) $\forall j \geq j_0, v_j([0,1]) \subset V_j \subset V_j([0,1])$.

ii) $\forall j \geq j_0, V_j \subset V_{j+1}$.

où $V_j([0,1])$ l'espace des restrictions à $[0,1]$ des fonctions de $V_j(\mathbb{R})$.

Nous présentons la construction de Meyer [**?**] d'une base orthonormée d'ondelettes sur l'intervalle [0,1].

Nous partons toujours de l'analyse multirésolution orthogonale $(V_j(\mathbb{R}))_{j\in\mathbb{Z}}$ de Daubechies et on désigne par :

- $S(j)$ l'intervalle d'entiers k définis par $-2N+2 \leq k \leq 2^j - 1$. Ce qui est équivalent au fait que le support de la fonction $\varphi_{j,k} = 2^{\frac{j}{2}}\varphi(2^j x - k)$ rencontre l'intervalle $]0,1[$.
- j_o le plus petit entier j tel que $2^j \geq 4N-4$ (pour séparer les fonctions des bords 0 et 1).

On rappelle les équations d'échelle :

- $\frac{1}{2}\varphi(x) = \sum_0^{2N-1} \alpha_k \varphi(2x-k)$ avec $\alpha_o \neq 0$ et $\alpha_{2N-1} \neq 0$.
- $\frac{1}{2}\psi(x) = \sum_0^{2N-1} \beta_k \varphi(2x-k)$ avec $\beta_o \neq 0$ et $\beta_{2N-1} \neq 0$.

Soit le lemme :

Lemme 1.1 *Soit $f(x) = \sum_{-\infty}^{+\infty} c_k \varphi(x-k)$ une fonction de $V_o(\mathbb{R})$. Supposons que $f(x) = 0$ pour $x \leq 0$, alors $c_k = 0$ pour $k \leq -1$*

Le lemme nous permet de conclure que $j \geq j_o$ et $f(x) = \sum_{-\infty}^{+\infty} c_k \varphi\left(2^j x - k\right)$ une fonction arbitraire de $V_j(\mathbb{R})$ telle que $f(x) = 0$ pour $0 \leq x \leq 1$, alors $c_k = 0$ pour tout $k \in S(j)$.

Nous pouvons alors en déduire une base de Riesz pour $V_j([0,1])$.

- $\{\varphi_{j,k|[0,1]}, k \in S(j)\}, \quad j \geq j_o\}$ est une Base de Riesz de $V_j([0,1])$
- Si $\{\varphi_{j+1,k|[0,1]}, k \in S(j+1)\}$ est une base de $V_{j+1}([0,1])$, alors $\{\varphi_{j,k|[0,1]}, \quad k \in S(j)\}$ est une base de $V_j([0,1])$.

Une base orthonormée de $V_j([0,1])$ est donnée par le corollaire suivant.

Corollaire 1.1 *Pour* $j \geq j_o$*, il existe* $(2N-2)$ *fonctions* $\varphi_i^\alpha, \quad (1 \leq i \leq 2N-2)$ *et* $(2N-2)$ *fonctions* $\varphi_i^\beta, \quad (1 \leq i \leq 2N-2)$ *tels que les fonctions*

- $\varphi_{i,j}^\alpha = 2^{j/2}\varphi_i^\alpha\left(2^j x\right)$, $(1 \leq i \leq 2N-2)$,
- $\varphi_{j,k} = 2^{j/2}\varphi\left(2^j x - k\right)$, $(0 \leq k \leq 2^j - 2N + 1)$,
- $\varphi_{i,j}^\beta = 2^{j/2}\varphi_i^\beta\left(2^j x - 2^j\right)$, $(1 \leq i \leq 2N-2)$,

forment une base orthonormée de $V_j([0,1])$.

1.5.1 Les Ondelettes orthogonales sur $[0,1]$

Nous disposons déjà d'une base orthonormée de $V_j([0,1])$, $j \geq j_0$. Nous présentons la construction d'une base orthonormée de l'espace d'ondelettes

$$W_j([0,1]) = V_{j+1}([0,1]) \cap V_j([0,1])^\perp.$$

Désignons par $V_o([0,+\infty[)$ l'espace des restrictions à $[0,+\infty[$ des fonctions de $V_o(\mathbb{R})$. Alors nous avons :

- les fonctions $\psi(x-k)|_{[0,+\infty[}, -2N+2 \leq k \leq -N$, appartiennent à $V_o([0,+\infty[)$.
- les fonctions $\psi(2^j x-k)|_{[0,1]}, -2N+2 \leq k \leq -N$, appartiennent à $V_j([0,1])$.

On peut remarquer que les fonctions $\psi(2^j x-k)|_{[0,1]}, 2^j - N + 1 \leq k \leq 2^j - 1$, sont dans l'espace $V_j([0,1])$. Ce qui nous permet de déduire le théorème qui suit :

Théorème Pour tout $j \geq 0$, une base de $V_{j+1}([0,1])$ est constituée de la réunion de la base $\varphi_{j,k}$ de $V_j([0,1])$ et des fonctions $\psi_{j,k}$ telles que $-N+1 \leq k \leq 2^j - N$.

Pour construire une base orthonormée de $W_j([0,1])$ pour $0 \leq j \leq j_o$, il suffit de projeter orthogonalement sur W_j les fonctions $\psi_{j,k}$ telles que $-N+1 \leq k \leq 2^j - N$. Puisque nous disposons déjà d'une base orthonormée de $V_j([0,1])$. L'opérateur de projection orthogonale sur $V_j([0,1])$ est explicite. Une fois projetés sur $W_j([0,1])$ les $\psi_{j,k}$ deviennent des fonctions $h_{j,k}$ qu'il convient ensuite d'orthonormaliser entre elles pour $-N+1 \leq k \leq 2^j - N$. Et on a le corollaire suivant :

Corollaire Pour $j \geq j_o$, il existe $(N-1)$ fonctions ψ_i^α $(1 \leq i \leq N-1)$ et $(N-1)$ fonctions ψ_i^β $(1 \leq i \leq N-1)$ telles que les fonctions :

- $\psi_{i,j}^\alpha = 2^{j/2}\psi_i^\alpha(2^j x), (1 \leq i \leq N-1)$,
- $\psi_{j,k} = 2^{j/2}\psi(2^j x - k), (0 \leq k \leq 2^j - 2N + 1)$,
- $\psi_{i,j}^\beta = 2^{j/2}\psi_i^\beta(2^j x - 2^j), (1 \leq i \leq N-1)$,

forment une base orthonormée de $W_j([0,1])$.

Ainsi, nous avons présenté la méthode de Meyer [7] pour la construction d'une base orthonormée d'ondelettes sur l'intervalle $[0,1]$ en partant de l'analyse multirésolution orthogonale de I. Daubechies. Le résultat principal de ce travail **est que les restrictions des fonctions d'échelle à l'intervalle forment un système linéairement indépendant alors que les restrictions des ondelettes associées forment un système**

lié. On ne peut donc pas définir de la même manière une analyse multirésolution biorthogonale (A.M.R.O) sur un domaine borné. Nous généralisons ces résultats dans notre article dont nous présentons les principaux résultats.

1.6 Principaux résultats

Nous partons d'une AMR orthogonale $V_j(\mathbb{R})$ de $L^2(\mathbb{R})$ dont la fonction d'échelle φ est à support compact $[N_1, N_2]$.

On désigne par :

- j_o le plus petit entier j tel que $2^{j_0} \geq 2\ (N_2 - N_1 - 1)$
- $S\,(j) = \{k \in \mathbb{Z}, -N_2 + 1 \leq k \leq 2^j - N_1 - 1\}$
- $v_j([0,1]) = \text{Vect}\{\varphi_{j,k}, \text{supp}\varphi_{j,k} \subset [0,1]\}$
- $V_j([0,1]) = \text{Vect}\{\varphi_{j,k/[0,1]}, \varphi_{j,k} \in V_j(\mathbb{R})\}$

A partir de ces deux analyses considérées comme minimale $(v_j([0,1]))$ et maximale $(V_j([0,1]))$, on définit alors l'AMR sur l'intervalle comme étant comprise entre ces deux analyses. Plus exactement, pour une suite $\{V_j\}_{j\geq j_0}$ des sous espaces fermés de $L^2([0,1])$ est dite analyse multirésolution de $L^2([0,1])$ associée à $V_j(\mathbb{R})$ si on a :

i) $\forall j \geq j_0, v_j([0,1]) \subset V_j \subset V_j([0,1])$.

ii) $\forall j \geq j_0, V_j \subset V_{j+1}$.

Comme premier résultat important dans la construction de bases orthonormées de $V_j([0,1])$, nous avons :

Corollaire 1.2 *i) Il existe* $(N_2 - N_1 - 1)$ *fonctions* $\varphi_i^\alpha, (1 \leq i \leq N_2 - N_1 - 1)$ *et* $(N_2 - N_1 - 1)$ *fonctions* φ_i^β $(1 \leq i \leq N_2 - N_1 - 1)$ *telles que les fonctions*

a) $\varphi_{i,j}^\alpha = 2^{j/2}\varphi_i^\alpha(2^j x)|_{[0,1]}, (1 \leq i \leq N_2 - N_1 - 1)$,

b) $\varphi_{j,k} = 2^{j/2}\varphi(2^j x - k)\ ,\ (-N_1 \leq k \leq 2^j - N_2)$,

c) $\varphi_{i,j}^\beta = 2^{j/2}\varphi_i^\beta(2^j x - 2^j)|_{[0,1]},\ (1 \leq i \leq N_2 - N_1 - 1)$,

forment une base orthonormée de $V_j([0,1])$.

ii) Soit V_j, $j \geq j_o$, *une analyse multirésolution de* $L^2([0,1])$ *associée à* $V_j(\mathbb{R})$, *alors il existe* N_o *fonctions* $\varphi_i^\alpha(1 \leq i \leq N_o)$ *et* N_o *fonctions* φ_i^β $(1 \leq i \leq N_o)$ *telles que les fonctions*

a) $\varphi_{i,j}^\alpha = 2^{j/2}\varphi_i^\alpha(2^j x)|_{[0,1]}, (1 \leq i \leq N_o)$,

b) $\varphi_{j,k} = 2^{j/2}\varphi(2^j x - k)\ ,(-N_1 \leq k \leq 2^j - N_2)$,

c) $\varphi_{i,j}^\beta = 2^{j/2}\varphi_i^\beta(2^j x - 2^j)|_{[0,1]}, (1 \leq i \leq N_o)$,

forment une base orthonormée de V_j.

Ainsi l'espace V_j contient un système orthonormée $\varphi_{j,k} = 2^{j/2}\varphi(2^j x - k)\ , (-N_1 \leq k \leq 2^j - N_2)$, auquel on ajoute les fonctions des bords $\{0\}$ et $\{1\}$ des collections $\varphi_{i,j}^\alpha$ et $\varphi_{i,j}^\beta$. Nous disposons donc d'une base orthonormée de $V_j([0,1])$.

Afin de présenter la construction d'une base orthonormée de $W_j([0,1]) = V_{j+1}([0,1]) \cap (V_j([0,1]))^\perp$.

On note :

$$V_j([N_1,+\infty[) = \mathrm{Vect}\{\varphi_{j,k/[N_1,+\infty[}, \varphi_{j,k} \in V_j(\mathbb{R})\}.$$

Il est important de remarquer que $W_j([0,1])$ n'est pas l'espace des restrictions à $[0,1]$ des fonctions de $W_j(\mathbb{R})$. La construction des ondelettes sur l'intervalle $[0,1]$ repose sur l'énoncé suivant qui permet de compléter la base $\varphi_{j,k}$, $k \in S(j)$, en une base de $V_{j+1}([0,1])$ et qui spécifie que Les fonctions $2^{j/2}\psi(2^jx-k)_{/[0,1]}$ tel que $-\frac{1}{2}(N_2+N_1-1) \leq k \leq 2^j-\frac{1}{2}(N_2+N_1+1)$, forment une base de Riesz de l'espace $W_j([0,1])$.

La notion d'analyse multirésolution biorthogonale sur l'intervalle $[0,1]$ introduite par A.Jouini P.G. Lemarié [4] se présente comme suit :

Définition Une suite (V_j, V_j^*) des sous espaces fermés de $L^2([0,1])$ associée à une analyse multirésolution biorthogonale $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ de $L^2(\mathbb{R})$ est dite analyse multirésolution biorthogonale de $L^2([0,1])$ si

- $v_j([0,1]) \subset V_j \subset V_j([0,1])$ et $v_j^*([0,1]) \subset V_j^* \subset V_j^*([0,1])$.
- $V_j \subset V_{j+1}$ et $V_j^* \subset V_{j+1}^*$.
- $L^2([0,1]) = V_j \oplus (V_j^*)^\perp$.

Soit $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ une analyse multirésolution biorthogonale de $L^2(\mathbb{R})$ associée aux fonctions d'échelle conjugué g et g^* avec supp$g = [N_1, N_2]$.

On note $P_i^\alpha(x) = \sum_{k\leq -N_1-1} k^i g(x-k)$, et $P_i^\beta(x) = \sum_{k\geq -N_2-1} k^i g(x-k)$.

Nous pouvons énoncer ce premier résultat sur la propriété de commutation entre les projecteurs obliques et la dérivation.

Théorème 1.1 *Soit $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ une analyse multirésolution biorthogonale de $L^2(\mathbb{R})$, (g, g^*) sont les fonctions d'échelle à support compact et (V_j, V_j^*) est l'analyse multirésolution biorthogonale de $L^2([0,1])$ associée à $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$. On suppose que*

i) g est différentiable et $g'(x) = \widetilde{g}(x) - \widetilde{g}(x-1)$.

ii) V_j contient les fonctions $P_{0,j}^\alpha(x) = P_0^\alpha(2^jx)_{/[0,1]}$ et $P_{0,j}^\beta(x) = P_0^\beta(2^jx-2^j)_{/[0,1]}$.

Si on désigne par

$$\begin{aligned}\widetilde{V_j} &= \{f \in L^2([0,1]) \backslash \exists h \in V_j, f = h'\},\\ V_j^* &= \{f \in L^2([0,1]) \backslash f' \in V_j^*, f(0) = f(1) = 0\}.\end{aligned}$$

Alors, $(\widetilde{V_j}, \widetilde{V_j}^*)$ est une analyse multirésolution biorthogonale de $L^2([0,1])$. De plus, si on désigne par P_j (resp $\widetilde{P}_j$) le projecteur oblique de $L^2([0,1])$ dans V_j (resp. $\widetilde{V}_j$) parallèlement à $(V_j^*)^\perp$ (resp $(\widetilde{V_j}^*)^\perp$), alors on a la formule de commutation

$$\frac{d}{dx} o P_j = \widetilde{P}_j o \frac{d}{dx}.$$

Corollaire 1.3 *Soit $V_j(\mathbb{R})$ une analyse multirésolution orthogonale de $L^2(\mathbb{R})$ associée à une fonction d'échelle g de classe C^m $(m \in \mathbb{N}^*)$. On désigne par $(V_j^{(m)}(\mathbb{R})\ , V_j^{*(m)}(\mathbb{R}))$ l'analyse multirésolution biorthogonale construite par m dérivations et m intégrations.*

Alors $V_j^{(m)}([0,1])$ *et* $V_j^{*(m)}([0,1]) \cap H_0^m([0,1])$ *forment une analyse multirésolution biorthogonale de* $L^2([0,1])$. *De plus, si on désigne par* $P_j^{(m)}$ *le projecteur oblique sur* $V_j^{(m)}([0,1])$ *parallèlement à* $[V_j^{*(m)}([0,1]) \cap H_0^m([0,1])]^\perp$, *on a*

$$\frac{d}{dx} o P_j^{(m)} = P_j^{(m+1)} o \frac{d}{dx}.$$

Proposition 1.1 *Soit* $P_j^{(m)}$ *le projecteur oblique sur* $V_j^{(m)}$ *parallèlement à* $V_j^{*(m)}$ *et* $P^{(m)*}$ *son adjoint. On définit* $Q_j^{(m)} = P_j^{(m+1)} - P_j^{(m)}$, $Q_j^{(m)*} = P_{j+1}^{(m)*} - P_j^{(m)*}$ *et* j_o *un entier satisfaisant* $2^{j_o} - 1 \geq 2N_2 - 2N_1 - 2 + 2m$. *Alors on a les formules de commutation suivantes :*

$$\begin{array}{rcl} si \quad f \in H^1([0,1]), \quad \dfrac{d}{dx}(P_j^{(m)} f) & = & P_j^{(m+1)}(\dfrac{df}{dx}), \\ si \quad f \in H_o^1([0,1]), \quad \dfrac{d}{dx}(P^{(m+1)^*} f) & = & P_j^{(m)^*}(\dfrac{df}{dx}) \end{array}$$

On a une caractérisation des espaces $H^s([0,1])$ et $H_o^s([0,1])$ en termes de normes :

Théorème 1.2 *On suppose que la fonction d'échelle* φ *est de classe* $C^{p+\varepsilon}$, $p \in N^*$, $p \geq m$, $\varepsilon > 0$ *et* j_o *un entier satisfaisant* $2^{j_o} - 1 \geq 2N_2 - 2N_1 - 2 + 2p$. *Alors on a :*

i) Pour $f \in L^2([0,1])$, $\|f\|_2 \approx \|P_{j_o}^{(m)} f\|_2 + (\sum_{j \geq j_o} \|Q_j^{(m)} f\|_2^2)^{\frac{1}{2}}$.

ii) Pour $f \in L^2([0,1])$, $\|f\|_2 \approx \|P_{j_o}^{(m)*} f\|_2 + (\sum_{j \geq j_o} \|Q_j^{(m)*} f\|_2^2)^{\frac{1}{2}}$.

iii) Pour $s \in \mathbb{Z}$ *tel que* $-m \leq s \leq p - m$, *on a*

- $f \in H^s([0,1]) \Leftrightarrow P_{j_o}^{(m)} f \in L^2([0,1])$ *et* $\sum_{j \geq j_o} 4^{js} \|Q_j^{(m)} f\|_2^2 < +\infty$.
- $f \in H_o^{-s}([0,1]) \Leftrightarrow P_{j_o}^{(m)*} f \in L^2([0,1])$ *et* $\sum_{j \geq j_o} 4^{-js} \|Q_j^{(m)*} f\|_2^2 < +\infty$.

Bibliographie

[1] Cohen, A., Daubechies, I., Jawert, B. and Vial, P. Multirésolution Analysis, wavelets and fast algorithms on an interval, *C.R.A.S, 1992.*

[2] Daubechies, I. Ten lectures on wavelets, SIAM,1992.

[3] Mallat, S. A wavelet tour of signal processing, Academic Press, San Diego, California, 1998.

[4] Jouini, A.P. et G. Lemarié-Rieusset, Analyse multirésolution biorthogonale sur l'intervalle et applications, *Annales de l'I.H.P.* Section C 10 (1993) 453-476.

[5] Meyer, Y. Ondelettes et opérateurs,tome 2 : Opérateurs de Calderon-Zygmund, *Hermann,1997.*

[6] Meyer, Y. et Jaffard, S. Bases d'ondelettes dans des ouverts de $\mathbb{R}^n$. *J. Math Pures et Appl*, 68, pp95-108, 1989.

[7] Meyer, Y. Ondelettes sur l'intervalle, *Revista Mathematica Ibero-americana*, Vol 7 (1991), pp115-13

Chapitre 2

Estimation du paramètre de longue mémoire

Préambule mathématique

On s'intéresse aux processus $\{X_k\}_{k\geq 1}$ stationnaires, à variances finies dont la densité spectrale $f(\lambda)$ pour $\lambda \in (-\pi, \pi)$ se comporte comme une loi de puissance aux basses fréquences, c'est à dire en $|\lambda|^{-2d}$ quand $\lambda \to 0^+$. Le cas $d > 0$ correspond à la mémoire longue, $d = 0$ à la mémoire courte et $d < 0$ à la *dépendance négative*. Pour que X_k soit stationnaire, il est nécessaire que $\int_{-\pi}^{\pi} f(\lambda)d\lambda < \infty$ et donc que $d < 1/2$.

Dans la majorité des cas, la densité spectrale est affectée d'une fonction de nuisance que l'on note $f^*(\lambda)$ jouissant d'une certaine régularité au voisinage de l'origine. La forme générale de la densité spectrale est alors $|\lambda|^{-2d} f^*(\lambda)$ et le but est alors d'estimer le paramètre d en présence de $f^*(\lambda)$.

La nécessité de stationnariser le processus est importante. En effet la faible stationnarité est le premier aspect que tout statisticien cherchera à vérifier. A titre d'exemple dans la modélisation linéaire du type Box et Jenkins, il est nécessaire que le processus étudié soit faiblement stationnaire. Dans ce cas le processus $(X_t)_{t\in\mathbb{Z}}$ est intégré d'ordre 0, sinon on supposera qu'il existe un $k \in \mathbb{N}^*$ tel que $(I - B)^k X_t$ soit asymptotiquement faiblement stationnaire [1]. Le processus $(X_t)_{t\in\mathbb{Z}}$ est dit intégré d'ordre k. La majorité des cas étudiés présentent un ordre d'intégration d'ordre l'unité, (travaux de Fuller [25], Dickey et Fuller [21]). Un intérêt vers les valeurs factionnaires $d \in]0, 1[$ s'ensuivit, ce qui apporta une grande souplesse à la modélisation.
Plus généralement, on dit qu'un processus $(X_t)_{t\in\mathbb{Z}}$ est un processus intégré d'ordre $d \in (0, 1)$ si $(I - B)^d X_t$ est asymptotiquement faiblement stationnaire. Il est toutefois important de définir mathématiquement ces différentes notions de stationnarité.

2.0.1 Les notions de stationnarité

Définition 2.1 *Le processus X_t est dit strictement ou fortement stationnaire si pour tout $k, n \in \mathbb{N}^*$ le n-uplet $t_1 < t_2 < \ldots < t_n$, tel que $t_i \in \mathbb{Z}$, la suite $(X_{t_1+k}, \ldots, X_{t_n+k})$ a la même loi de probabilité que la suite $(X_{t_1}, \ldots, X_{t_n})$.*

1. B étant l'opérateur retard défini pour $b \in \mathbb{N}$ par : $B^b X_t = X_{t-b}$

Définition 2.2 *Un processus* $X = (X_t)_{t\in\mathbb{Z}}$ *est dit stationnaire d'ordre deux ou stationnaire au sens faible, si on a :*

- $\forall t \in \mathbb{Z}, \mathbb{E}(X_t^2) < \infty$, *et* $\mathbb{E}(X_t) = m$ *indépendant de* t,
- $\forall t, h \in \mathbb{Z}$, $cov(X_t, X_{t+h}) = r(h)$ *indépendant de* t.

2.0.2 Les processus à longue mémoire

Soit un processus stationnaire $(X_t)_{t\in\mathbb{Z}}$ de second ordre. X_t est dit à longue mémoire si l'une de trois propositions suivantes est vérifiée.

- La suite des covariances n'est pas sommable :
  $\sum_{-\infty}^{+\infty} |r(k)| = \infty$, où $r(.)$ est la fonction d'autocovariance,
- La suite des covariances tend vers zéro lentement et de façon régulière :

$$\mathrm{cov}(X_1, X_{n+1}) = n^{-D}L(n),\ 0 < D < 1$$

  L étant une fonction à variations lentes à l'infini [2].
- La densité spectrale de X_t admet une singularité en λ_0 :

$$f(\lambda) = |\lambda - \lambda_0|^{D-1} L\Big(\frac{1}{|\lambda - \lambda_0|}\Big),\ 0 < D < 1,\ \lambda \to \lambda_0.$$

  $L(\lambda)$ [3] étant à variations lentes en 0.

Le bruit gaussien fractionnaire (fgn)

Rappelons auparavant la définition du mouvement brownien fractionnaire. $B_H = \{B_H(t), t \in \mathbb{R}^+\}$ est un brownien fractionnaire lorsque B_H est gaussien centré, continu, à accroissements stationnaires et tel que $\mathbb{E}(B_H(t) - B_H(s))^2 = \sigma^2|t-s|^{2H}$ pour tout $(t,s) \in \mathbb{R}^2$ et $H \in (0,1)$. Le bruit gaussien fractionnaire $X_H = \{X_H(t), t \in \mathbb{R}^+\}$ est défini alors comme étant les accroissements unitaire de B_H soit :

$$X_H(t) = B_H(t+1) - B_H(t), \quad t \in \mathbb{R}_+$$

Sa fonction d'autocovariance est :

$$\begin{aligned} r_H(k) &= \frac{\sigma^2}{2}\big(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}\big), \forall k \in \mathbb{N} \qquad (2.0.1) \\ &\sim H(2H-1)|k|^{2H-2}, k \to \infty. \end{aligned}$$

Le processus *FARIMA*(p,d,q)

Soit $(\varepsilon_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires centrées, iid [4] et de variance finie.
Un processus $X = \{X_t, t \in \mathbb{Z}\}$ est un $FARIMA(p,d,q)$ où $d \in (-\frac{1}{2}, \frac{1}{2})$ s'il vérifie :

$$\phi(B)(1-B)^d(X_t) = \theta(B)\varepsilon_t,$$

où ε_t est un bruit blanc centré de variance σ^2, B l'opérateur de retard et $(\phi(.), \theta(.))$ étant des polynômes à coefficients réels à racines en dehors du cercle unité non communes aux deux polynômes.

2. L est bornée sur les intervalles finis et pour tout $t > 0$ à l'infini. $\frac{L(tx)}{L(x)} \to 1, \quad x \to +\infty$
3. Si L est oscillante on parle alors de longue mémoire saisonnière.
4. indépendantes et identiquement distribuées.

- $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^P$.
- $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_p B^q$.

La fonction $(1-B)^d$ est défini par

$$(1-B)^d = \sum_{k\geq 0} b_k(d) B^k \quad \text{òu} \quad b_k(d) = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)}$$

La densité spectrale du processus *FARIMA(p,d,q)* est :

$$f(\lambda) = \frac{\sigma^2 |\theta(e^{i\lambda})|^2}{2\pi |\phi(e^{i\lambda})|^2} |1 - e^{i\lambda}|^{-2d} : \quad \lambda \in [\pi, 0[\cup]0, \pi].$$

Pour $0 < d < \frac{1}{2}$, $\{X_t\}_{t\in\mathbb{Z}}$ est alors stationnaire à longue mémoire.

2.1 Les méthodes d'estimation du paramètre de longue mémoire

C'est en 1951 qu'un ingénieur hydraulicien du nom de Hurst [33] publia un article sur les crues du Nil. C'était le point de départ des travaux sur les processus à longue mémoire . Diverses méthodes ont été élaborées. Tant dans le domaine spectral que temporel. Les méthodes spectrales ont pour but l'estimation de l'exposant de la puissance de la densité spectrale exprimée au point de singularité qui généralement est zéro. Ces méthodes sont dites locales lorsque les hypothèses de régularité de f sont considérées au voisinage de zéro et globales en dehors des voisinages de zéro.
Les méthodes les plus connues sont basées sur la transformée de Fourier. Elles furent développées par Peter Robinson ([42, 43] et sont basées sur le principe suivant :

Si on néglige la fonction de nuisance f^* dans l'expression de la densité spectrale alors $f(\lambda) = |\lambda|^{-2d}$, en appliquant le log on a alors $\log f(\lambda) \sim -2d \log(|\lambda|)$, ainsi le paramètre d'intérêt d peut etre estimé par une régression linéaire sur le périodogramme (notion développée ultérieurement). Cette méthode dite GPH (Geweke et Porter-Hudack) est présentée dans [26] dans un cadre paramétrique. Le cadre semiparamétrique a été considéré par Künsch [35] et développé par Robinson [43]. La méthode de Whittle basée sur Fourier (ou LWF [5]) est une méthode basée sur le pseudo-maximum de vraisemblance . Elle a été développée par Fox et Taqqu [23] dans un cadre paramétrique puis étendue au cadre semi-paramétrique par Robinson [42]. Moulines et *et al* dans [38, 40, 41] ont repris les mêmes travaux en utilisant les ondelettes. Les ondelettes présentent plusieurs avantages dont la robustesse aux tendances polynomiales. La méthode d'estimation par ondelettes d a vu le jour avec Abry et Veitch [1] sous l'hypothèse de la décorrélation des coefficients d'ondelettes, s'ensuivirent d'autres développements en 1999 [2], en 2000 [7] et en 2003 [1]. Veitch *et al* [47] dans le choix de la fonction d'échelle et la sélection automatique de la fréquence de rupture dans [46]. Bardet *al.* (2000) ont montré des résultats asymptotiques pour ce type d'estimateur (noté LRW [6]) dans le cas gaussien, et Bardet (2002) a considéré le cas particulier du mouvement Brownien fractionnaire [12]. Dans un cadre semi-paramétrique et en considérant le cas des observations continues, des résultats sur la consistance de ces estimateurs furent présentes par Bardet *et al.* dans

5. LWF : local whittle Fourier
6. LRW : Local régression Wavelets

[13], ces derniers résultats ont été améliorés par Moulines et *et al* en 2007 [40] en donnant une vitesse de convergence optimale au sens du critère minimax. Finalement Roueff et Taqqu [17] présentèrent e des résultats analogues à ces derniers pour la cas des processus linéaires. En 2009 Abry *et al.* [5] considérèrent le cas non gaussien.
Nous présentons dans ce qui suit des méthodes d'estimations locales et globales tant dans le domaine temporel que spectral.

2.1.1 Les méthodes spectrales

<u>**Les Méthodes locales**</u>

La construction des estimateurs par les méthodes locales supposent (outre l'intégrabilité sur $(-\pi, \pi])$ que le comportement de la densité spectrales en zéro soit de la forme $f(x) \sim Cx^{-2d}$ quand $x \to 0^+$ avec $-1/2 < d < 1/2$. Les paramètres à estimer sont alors C et d. Ces méthodes en questions utilisent le log-périodogramme de Geweke Porter Hudack [26]. Elles ont été améliorées par Kunsch [35]. Les versions adaptatives ont été proposées par Hurvich *et al.*.

L'estimateur Geweke Porter Hudack (GPH)

Soit m un entier fixé, pour tout n on pose $n_m = 2M[\frac{N}{2m}]$ et $K_n = [\frac{N}{2m}]$. On définit alors le périodogramme $I_n(x)$ de $\{X_1, X_2, \cdots, X_{n_m}\}$ comme suit :

$$I_n(x) = |\omega_n(x)|^2, \quad \text{où} \quad \omega_n(x) = (2\pi n_m)^{-1/2} \sum_{t=1}^{n_m} X_t e^{itx}$$

Ces quantités sont évaluées aux fréquences de Fourier $x_s = \frac{2\pi s}{n}, \quad 1 \le s \le n$. Le domaine fréquentiel est alors subdivisé en segments disjoints de longueur m et la moyenne du périodogramme est calculée sur chaque segment. Plus exactement, pour $k = 1, \cdots, K_n$ on note $J_k = \{m(k-1)+1, \cdots, mk\}$ et

$$Y_{n,k} = Y_{n,2K_n-k+1} = \log\Big(2e^{-\tau_m} \sum_{i \in J_k} I_n(x_i)\Big).$$

avec $\tau_m = \psi(m)$, $\sigma_m^2 = m\psi\prime(m)$ où $\psi(z)$ la fonction digamma $\psi(z) = \frac{\Gamma\prime(z)}{\Gamma(z)}$, la fonction $\Gamma(z)$ étant la fonction usuelle gamma. De plus on pose :

$$g(x) = -2\log|1 - e^{ix}| \quad \text{et} \quad y_k = \frac{(2k-1)\pi}{2K_n}, \quad 1 \le k \le K_n.$$

Pour $0 < L < M \le K_n$ avec L et M convenablement choisis. On définit alors :

$$\hat{d}(L, M) \triangleq \frac{\sum_{j=L+1}^{M}\Big(g(y_i) - (M-L)^{-1}\sum_{j=L+1}^{M} g(y_i)Y_{n,j}\Big)}{\sum_{j=L+1}^{M}\Big(g(y_i) - (M-L)^{-1}\sum_{j=L+1}^{M} g(y_i)\Big)^2}$$

c'est l'estimateur par moindres carrés ordinaires de d pour le modèle linéaire

$$Y_{n,j} = C + d.g(y_j) + \eta_j, \quad l+1 \le j \le m$$

Une remise en question des travaux de Geweke et Porter Hudack 1983 [26] par Robinson sur la validité de certaines approximations non valides dans le cas de la longue mémoire le mena à poser dans [42] les hypothèses suivants :

i) Il existe $0 < C < \infty$, $-1/2 < d < 1/2$ et $0 < \alpha \leq 2$ tels que $f(x) = Cx^{-2d}(1+O(x^{\alpha}))$

ii) $\{X_t\}_{t\in\mathbb{Z}}$ est un processus gaussien.

iii) quand $n \to \infty$ alors on a

$$M^{1/2}\frac{\log(M)}{L} + L\frac{\log^2(n)}{M} + \frac{M^{1+1/(2\alpha)}}{n} \to 0.$$

Sous ces trois hypothèses, il établit le théorème suivant :

$$\frac{M^{1/2}}{\log(n)}(\hat{d}(L,M) - d) \underset{\mathcal{L}}{\to} \mathcal{N}(0, m\psi'(m)).$$

De plus la vitesse de convergence de l'estimateur GPH est de l'ordre de n^{-h} avec $h < 2/5$

L'estimateur local de Whitle (LWF)

On considère la fonction objective suivante :

$$Q(C,d) \triangleq M^{-1}\sum_{j=1}^{M}\{\log(Cx_j^{-2d}) + x_j^{2d}I_n(x_j)/C\}$$

On définit $\Theta = [\triangle_1, \triangle_2] \subset (-1/2, 1/2)$ l'ensemble d'admissibilité de d . On définit $(\hat{d}, \hat{C}) = \arg\min_{0<C<\infty} Q(C,d)$. Les propriétés statistiques de cet estimateur sont présentées dans Robinson [44]. Sous la condition que M tende vers l'infini plus lentement que n. Les hypothèses suivantes ont été posées :

(W1) Pour $x \to 0^+$, $f(x) \simeq C_0x^{-2d_0}(1+O(x^{\beta}))$. De plus $f(x)$ est différentiable dans un voisinage de l'origine.

(W2) X_t est un processus linéaire,

$$X_t = \mu + \sum_{j=0}^{\infty}\psi_j Z_{t-j}$$

où Z_t est tel que $\mathbb{E}(Z_t/\mathfrak{F}_{t-1}) = 0$, $\mathbb{E}(Z_t^2/\mathfrak{F}_{t-1}) = 1$, $\mathbb{E}(Z_t^3/\mathfrak{F}_{t-1}) = \mu_3$, $\mathbb{E}(Z_t^4) = \mu_4$, et $\mathfrak{F}_t = \sigma(Z_t, s \leq t)$.

(W3) Quand $n \to \infty$ alors $\frac{1}{M} + M^{1+2\beta}\frac{(\log(M))^2}{n^{2\beta}} \to 0$

Sous ces trois hypothèses on a :

$$M^{1/2}(\hat{d} - d) \to \mathcal{N}(0, 1/4)$$

Il est à noter que la vitesse de convergence de cet estimateur est n^{-h} avec $h < 2/5$.

Les Méthodes globales

Les méthodes globales supposent des hypothèses supplémentaires de régularité de la densité spectrale sur le domaine fréquentiel. Alors de meilleures vitesses de convergence peuvent être obtenues. Les estimateurs ainsi construits ont des erreurs quadratiques de l'ordre de $O(\log(n)/n)$.

La régression par log-périodogramme

On note $h_0(x) = 1/\sqrt{2\pi}$ et $h_j(x) = \frac{1}{\sqrt{\pi}}\cos(jx)$ pour $j > 0$. On suppose :

(G1) $f^*(x)$ est bornée sur $[-\pi, \pi]\backslash\{0\}$ et differentiable. De plus, les coefficients $\{\theta_j\}_{j\geq 0}$ du développement en série de Fourier de $\log(f^*(x)) = \sum_{j=0}^{\infty}\theta_j h_j(x)$ sont tel que $\sum_{j=0}^{\infty} j^{\beta}|\theta_j| < \infty$. Un estimateur semi-paramétrique $\hat{d}$ de d peut être obtenu en tronquant le développement en série de Fourier de $\log(f^*(x))$ aux p-premiers termes, puis en faisant tendre p vers l'infini à une certaine vitesse. Cette technique est souvent utilisée en estimation fonctionnelle. Un estimateur semi-paramétrique de d est alors obtenu comme suit :

$$\hat{d}_{p,n} \triangleq \Big(\sum_{k=1}^{2K_n}\tilde{g}^2_{p,n}(y_k)\Big)^{-1}\Big(\sum_{k=1}^{2K_n}\tilde{g}_{p,n}(y_k)Y_{n,k}\Big)$$

où $\tilde{g} = g\sum_{j=0}^{p}\tilde{\alpha}_j h_j(x)$, avec $\tilde{\alpha}_j = \frac{\pi}{K_n}\sum_{k=1}^{2K_n} g(y_k)h_j(y_k)$

(G2) $\{X_t\}$ est un processus gaussien.

(G3) pour $n\to\infty$, on a $p^3\frac{\log^2(n)}{n^2} + \frac{n}{p^{1+2\beta}} \to 0$

L'hypothèse la plus répandue est celle de considérer que les θ_j sont à décroissance exponentielle vers zéro. Dans ce cas on pourra poser $p = \log(n)$. Sous cette hypothèse ainsi que les hypothèses (G1-G3) Les résultats de Moulines et Soulier [39] s'ensuivent :

$$\begin{aligned} \frac{n}{p}(\hat{d}-d) &\underset{d}{\to} \mathcal{N}(0, m\sigma_m^2) \\ \lim_{n\to\infty}\mathbb{E}(\hat{d}-d)^2 &= m\sigma_m^2 \end{aligned}$$

Ils donnent l'expression du biais et de la variance de l'estimation. Sous l'hypothèse (G1), ce biais est majoré par $\frac{\log(n)}{p^{\beta}}$. La vitesse optimale de convergence est $n^{-\beta/(2\beta+1)}$

2.1.2 Les méthodes temporelles

La singularité en zéro de la densité spectrale est exprimée de façon équivalente dans le domaine temporel en une condition sur les coefficients de Fourier de f ou en termes de coefficient d'autocovariance $\gamma(h) = \mathrm{cov}(X_i, X_{i+h})$. On considère un processus à longue mémoire stationnaire, sa fonction d'autocovariance s'écrit $\gamma(h) = h^{2d-1}L(h)$ avec $0 < d < 1/2$ et $L(h)$ une fonction à décroissance lente pour $|h| \to \infty$

La méthode R/S

Cette méthode est basée sur la statistique $Q(n) = \frac{R(n)}{S_n}$ où n désigne la taille de l'échantillon et $Q(n) > 0$. Elle permet de détecter des cycles non périodiques. On note :

$$R(n) = \max_{1\leq k\leq n}\sum_{i=1}^{k}(X_i - \bar{X}_n) - \min_{1\leq k\leq n}\sum_{i=1}^{k}(X_i - \bar{X}_n),$$

$\overline{X}_n, S_n^2$ étant respectivement la moyenne et la variance empirique.
Cette méthode se présente comme suit :

On détermine une suite d'entiers k_i vérifiant certaines conditions définies par Davis et Harte [20].

On trace la droite $log(Q(k_i)) = a + b\log(k_i) + u$, la méthode des moindres carrés ordinaires nous donne alors une estimation des coefficients a et b, le paramètre de Hurst est alors $\hat{H}=\hat{b}$.

Le seul avantage de cette méthode est qu'elle permet d'obtenir un estimateur qui possède des bonnes propriétés de robustesse [37] mais comporte un inconvénient important au niveau de la distribution asymptotique de la statistique R/S .

La méthode de Lo

Parmi les inconvénients de la statistique R/S proposée par Hurst, sa sensibilité à la présence de la mémoire courte. Pour surmonter ce problème, Lo [9] a proposé une statistique "R/S modifiée". Sa distribution limite est invariante aux différentes formes des processus à mémoires courtes. Cette méthode permet de proposer un test dont l'hypothèse nulle est l'absence de dépendance de long terme. La statistique R/S modifiée de LO possède la forme suivante :

$$\bar{Q}_q = \frac{1}{\sqrt{n}}\frac{R(n)}{S_q(n)}.$$

où $S_q(n)$ est exprimée par

$$S_q(n) = \{S_n^2 + \frac{2}{n}\sum_{j=1}^{q} w_j(q)[\sum_{i=j+1}^{n}(X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n)]\}^{\frac{1}{2}}.$$

Les w_j sont des poids.

Lo et Mackinlay [36] et Andrews [8] ont montré par simulations que lorsque q est relativement grand par rapport à la taille de l'échantillon, l'estimateur était biaisé.

La méthode des variances agrégées

La méthode des variances agrégées [48] est une méthode simple qui permet d'estimer le paramètre d'autosimilarité H en plusieurs étapes :

1. Pour m donné et sous certaines conditions, on divise la série d'origine $X = \{X_i, i \geq 1\}$ en $\frac{n}{m}$ sous-séries, chacune étant de taille m.
2. Pour chaque sous-série, on calcule $X^m(k)$ et $var(\hat{X}^{(m)})$

$$\begin{aligned} X^m(k) &= \frac{1}{m}\sum_{i=(k-1)m+1}^{km} X_i, \quad \text{pour} \quad k = 1, 2, .., n/m, \\ var(\hat{X}^{(m)}) &= \frac{1}{n/m}\sum_{k=1}^{n/m}(\hat{X}^m(k))^2 - (\frac{1}{n/m}\sum_{k=1}^{n/m}\hat{X}(m)(k))^2. \end{aligned}$$

3. Enfin, on estime alors par les moindres carrés ordinaires la pente $b = 2H - 2$ dans le modèle :

$$\log(var(\hat{X}^{(m)})) = \log c + b\log(m) + u.$$

Cette méthode présente les mêmes inconvénients que ceux de la méthode R/S de Hurst. De plus à partir de la représentation graphique, on ne peut discriminer entre le comportement de longue et courte mémoire.

Bien d'autres méthodes existent : la méthode du Log-variogramme de Guyon et Leon dans [30] basée sur les variations quadratiques, la méthode des variations quadratiques généralisées par Istas et Lang [34], l'estimateur FEXP, etc.... Nous pensons avoir présenté les estimateurs les plus connus. Mais qu'en est-il de la vitesse de convergence de ces différents estimateurs ?

Les vitesses de convergence

Les estimateurs du paramètre de queue de distribution dont la forme est du type $f(x) = Cx^{-\alpha}\{1 + O(x^{-\alpha p})\}$ pour $x \to \infty$ ont une vitesse de convergence optimale de l'ordre de $n^{p/(2p+1)}$ (voir Hall et Welsh dans [31]). Cette forme fait penser aux densités des processus à mémoire longue qui, généralement s'écrivent sous la forme $f(\lambda) = |\lambda|^{-\alpha} L(\lambda)$ [7] inspira certains auteurs dans la méthodologie de recherche des vitesses de convergence pour l'estimateur du paramètre de longue mémoire.

Robinson [43] a traité les propriétés asymptotiques de l'estimateur semiparamétrique de Gweeke Porter-Hudack (GPH) dans une version modifiée de cet estimateur. Il conclut que sous certaines faibles conditions locales incluant la condition :

$$L(\lambda) = C + O(|\lambda|^{\beta}) \quad \text{quand} \quad \lambda \to 0, \qquad C \in (C, \infty), \quad \beta \in (0, 2) \tag{2.1.1}$$

L'estimateur avait une vitesse de convergence (au sens du critère du minimax) de l'ordre $m(r) = n^{-r} M_n$ pour $\alpha \in (-1, 1)$ où $r = r(\beta) = \beta/(2\beta + 1)$ et $M_N \to \infty$ une suite à variation lente.

La normalité asymptotique et la vitesse de convergence furent traités par Robinson dans [44] pour le cadre gaussien. Il montra que sous 2.1.1 et des conditions plus faibles que le premier cas, la vitesse était $n^{-r} M_n$ où M_n est tel que $\log^{-1/(1+2\beta)}(n)/M_n = O(1)$.

Quand β est inconnu, une borne inférieure pour le risque quadratique des estimateurs adaptatifs est obtenue. La méthodologie adaptative est décrite par Giraitis et al dans [27] et plus récemment dans [32].

2.1.3 La méthode d'estimation par ondelettes

Soit un processus $\{X_t\}_{t \in \mathbb{R}}$ un processus stationnaire à longue mémoire de paramètre α, une ondelette ψ. La densité spectrale vérifie au voisinage de zéro la relation suivante :

$$f(\lambda) \underset{\lambda \to 0}{\sim} |\lambda|^{-2d}$$

On définit le coefficient d'ondelette non nécessairement issue d'une analyse multirésolution :

$$\begin{aligned} d_X(j,k) &= \langle X, \psi_{j,k} \rangle \\ &\equiv \int_{\mathbb{R}} X(t) \psi_{j,k}(t) dt \end{aligned}$$

7. $L(\lambda) \to C, C \in (0, \infty)$ $\alpha \in (-1, 1), \lambda \in [-\pi, \pi]$

où $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k), \quad j,k \in \mathbb{Z}$.
On montre sous certaines conditions que

$$\mathbb{E}(d_X(j,.)^2) \underset{j\to+\infty}{\simeq} 2^{j\alpha}c_f C(\alpha,\psi), \text{ où } C_{(\alpha,\psi)} = \int \frac{|\hat{\psi}(\lambda)|^2}{|\lambda|^\alpha} d\lambda \qquad (2.1.2)$$

Cette relation log-linéaire suggère un estimateur du paramètre α par les moindres carrés ordinaires. La variance empirique [8] $\frac{1}{n_j}\sum_{k=1}^{n_j}(d_X(j,k)^2)$ s'écrivant suivant un modèle linéaire des échelles j.

Cette méthode qui a été introduite par Flandrin [22], fut ensuite développée numériquement par Abry *et al* [3] puis Veitch *et al* [46], et théoriquement par Bardet *et al* [6], avec les contributions de Moulines *et al* [38] dans le cas gaussien et Roueff et Taqqu [17] dans le cas linéaire, tout ceci dans le cadre d'une analyse par ondelette discrète.
Le but de notre travail a été de donner également des résultats de convergence dans le cas plus général des ondelettes continues, puis de mettre en place une procédure adaptative pour sélectionner automatiquement les échelles choisies pour la régression.

2.2 Principaux résultats

Dans toute la suite on considère un un processus stationnaire, centré, de second ordre et de densité spectrale f vérifiant

$$f(\lambda) \simeq \frac{1}{\lambda^{2d}}(c_d + c_{d'}\lambda^{d'}) \quad \text{quand} \quad \lambda \to 0.$$

2.2.1 Le cas des processus gaussiens

Nous exposons ici les principaux résultats obtenus dans un article Bardet *et al.* [14] On considère ici le cas où $X = (X_t)_{t\in\mathbb{Z}}$ est un processus gaussien.

On émet des hypothèses [9] locales sur le terme de nuisance f^* de la densité spectrale ainsi que sur la régularité de l'ondelette.

Pour $(a,b) \in \mathbb{R}_+^* \times \mathbb{R}$. Le coefficient d'ondelette considéré est :

$$e(a,b) = \frac{1}{\sqrt{a}}\sum_{k=1}^{a}\psi(\frac{k}{a})X_{k+ab}.$$

On montre que sous certaines conditions, le comportement asymptotique de de la variance de $e(a,b)$ est en loi de puissance des échelles a. C'est-à-dire :

$$\mathbb{E}(e^2(a,0)) \underset{a\to\infty}{\sim} K_{(\psi,2d)}\, a^{2d}, \text{ avec } K_{(\psi,\alpha)} = \int_{-\infty}^{\infty} |\widehat{\psi}(u)|^2 \cdot |u|^{-\alpha} du > 0 \qquad (2.2.1)$$

Après normalisation. Pour $a \in \mathbb{N}^*$ et $b \in \mathbb{Z}$, on pose :

8. $n_j = \lfloor \frac{n}{2^j} \rfloor$ est le nombre de coefficients disponibles d'ondelettes à l'octave j
9. voir section 2.1

$$\tilde{e}(a,b) = \frac{e(a,b)}{\left(f^*(0)K_{(\psi,2d)}\, a^{2d}\right)^{1/2}} \qquad (2.2.2)$$

$$\tilde{T}_N(a) = \frac{1}{[\frac{N}{a}]}\sum_{k=1}^{[\frac{N}{a}]} \tilde{e}^2(a,k-1). \qquad (2.2.3)$$

Le comportement de la variance des coefficients d'ondelettes dans la relation 2.1.3 est à la base de nos estimations. Cette propriété de loi de puissance suggère un estimateur déduit à partir d'une régression log-log. En prenant des échelles de la forme $(r_1 a_N, \ldots, r_\ell a_N)$ où les r_i sont des entiers, une régression linéaire de $(\log(\widehat{T}_N(r_i a_N))_i$ par $(\log(r_i a_N))_i$ fournit un estimateur $\widehat{d}(a_N)$ qui satisfait le théorème limite central (TCL) suivant à la vitesse $\sqrt{\frac{N}{a_N}}$.

Proposition 2.1 *Soient*
- $\ell \in \mathbb{N} \setminus \{0,1\}$ *et* $(r_1,\cdots,r_\ell) \in (\mathbb{N}^*)^\ell$.
- $(a_n)_{n\in\mathbb{N}}$ *tel que* $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ *et* $a_N \cdot N^{-1/(1+2d')} \underset{N\to\infty}{\longrightarrow} \infty$.

Sous certaines hypothèses (voir article), on a :

$$\sqrt{\frac{N}{a_N}}\Big(\log \tilde{T}_N(r_i a_N)\Big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,d)\big), \qquad (2.2.4)$$

avec $\Gamma(r_1,\cdots,r_\ell,\psi,d) = (\gamma_{ij})_{1\le i,j\le \ell}$ *et* $d_{ij} = PGCD(r_i,r_j)$,

$$\gamma_{ij} = \frac{8(r_i r_j)^{2-2d}}{K^2_{(\psi,2d)} d_{ij}} \sum_{m=-\infty}^{\infty}\Big(\int_0^\infty \frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D}\cos(u\, d_{ij} m)\, du\Big)^2. \qquad (2.2.5)$$

De ceci, on déduit (sous les mêmes hypothèses) pour $\widehat{d}(a_N)$ le théorème limite central suivant :

$$\sqrt{\frac{N}{a_N}}\left(\begin{pmatrix}\widehat{D}(a_N)\\ \widehat{K}(a_N)\end{pmatrix} - \begin{pmatrix} D \\ K\end{pmatrix}\right) \xrightarrow[N\to\infty]{D} \mathcal{N}_2(0,\Omega), \qquad (2.2.6)$$

avec $\Omega = \Omega(r_1,\cdots,r_\ell,\psi,D)$ spécifiée dans l'article.

Afin de parvenir à construire un estimateur effectif (le paramètre d' est a priori inconnu), d'autres ajustement sont nécessaires. Le choix d'une suite d'échelles minimales rendant maximale la vitesse de convergence serait de la forme $a_n \sim N^{\frac{1}{2d'+1}}$ et dépend donc du paramètre d'. L'estimation de D' est alors nécessaire.

On peut se référer à la méthode introduite dans et basée sur un test d'adéquation du Khi-deux. Ici,

On considère alors des échelles a_N de la forme $a_N = N^\alpha$. On note $\alpha^* = \frac{1}{1+2d'}$. L'échelle optimale est donc N^{α^*}. En se basant sur la méthode de Veitch *et al.*[46]. Cette échelle peut être estimée par la minimisation d'une fonctionnelle $Q_N(\alpha, d(a_N), K(a_N))$ dépendant de trois paramètres. Cette fonction correspond au carré de la distance entre les ℓ points $\big(\log(i\cdot N^\alpha)\,,\,\log T_N(i\cdot N^\alpha)\big)_i$ et la droite de régression. Un estimateur $\widehat{\alpha}$ de

α^* est obtenu en minimisant $Q_N(\alpha, d, K)$ par rapport à α. On montre la consistance de cette estimation sous certaines conditions et on a alors :

$$\widehat{\alpha}_N = \frac{\log \widehat{a}_N}{\log N} \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^* = \frac{1}{1+2d'}.$$

L'estimateur obtenu $\widehat{\widehat{d}} = \widehat{d}(\widehat{a}_N)$ où $\widehat{a}_N = N^{\widehat{\alpha}_N}$ ne vérifie pas le TLC 2.1. On considère alors un nouvel estimateur adaptatif $\tilde{d}_N = \widehat{d}(\tilde{a}_N)$ de d, en posant comme estimateur des échelles $\tilde{a}_N = N^{\tilde{\alpha}_N}$ où :

$$\tilde{\alpha}_N = \widehat{\alpha}_N + \frac{3}{(\ell-2)\widehat{d'}_N} \cdot \frac{\log\log N}{\log N}, \quad \tilde{a}_N = N^{\tilde{\alpha}_N} = N^{\widehat{\alpha}_N} \cdot \big(\log N\big)^{\frac{3}{(\ell-2)\widehat{d'}_N}}$$

On obtient alors sous certaines conditions :

$$\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big(\tilde{d}_N - d\big) \xrightarrow[N\to\infty]{D} \mathcal{N}(0\,;\sigma_d^2) \quad \text{avec} \quad \tilde{d}_N = \widehat{d}(\tilde{a}_N). \tag{2.2.7}$$

$$\text{et} \quad \forall \rho > \frac{2(1+3d')}{(\ell-2)d'}, \quad \frac{N^{\frac{d'}{1+2d'}}}{(\log N)^{\rho}} \cdot \big|\tilde{d}_N - d\big| \xrightarrow[N\to\infty]{\mathcal{P}} 0. \tag{2.2.8}$$

L'estimateur ainsi obtenu a une vitesse de convergence égale à $N^{d'/(1+2d')}$ (à un logarithme prés).

Simulations

On procède à des simulations afin de vérifier dans un premier temps les propriétés de consistance et de robustesse des estimateurs $\widehat{\widehat{d}}_N$ et $\tilde{d}_N$ puis à des comparaison avec d'autres estimateurs semiparamétriques.

Etude de la consistance et la robustesse

On commence par choisir une ondelette vérifiant nos hypothèses sur les ondelettes. Puis on génère des échantillons de trois types de processus choisis (fgn, $FARIMA$ et processus gaussien stationnaire à densité spectrale donnée) vérifiant les hypothèses locales. Les résultats numériques indiquent que la convergence et la normalité asymptotique des estimateurs $\widehat{\widehat{d}}_N$ et $\tilde{d}_N$ sont vérifiées. La robustesse des estimateurs $\widehat{\widehat{d}}_N$ et $\tilde{d}_N$ sont vérifiées. Quatre processus à mémoire courte sont choisis et les valeurs de $\sqrt{MSE}$ indiquent la convergence rapide de nos estimations prouvant la robustesse de ces estimateurs.

La comparaison avec d'autres estimateurs

La comparaison avec d'autres estimateurs est effectuée dans les conditions suivantes : Pour plusieurs valeurs de d et différentes tailles d'échantillons, on donne les valeurs de $\sqrt{MSE}$ pour les estimateurs suivant :

1. $\hat{d}_{BGK}$: l'estimateur paramétrique *optimal* de Whittle de Banshali *et al*[10].
2. $\hat{d}_{GRS}$: l'estimateur adaptatif du périodogramme de Giraitis *et al*[12]
3. $\hat{d}_{MS}$: l'estimateur adaptatif global du périodogramme de Moulines et Soulier [41] (estimateur FEXP).

4. $\hat{d}_{ATV}$: l'estimateur de Whittle de Robinson [44].
5. $\hat{d}_R$: l'estimateur adaptatif basé sur les ondelettes de Veitch *et al* [46].
6. $\widehat{\hat{d}}_N$.

Il en ressort que $\hat{d}_{BGK}$ et $\hat{d}_{GRS}$ présentent des inconvénients respectivement en estimation et en vitesse de convergence. Des similarités existent entre $\hat{d}_{MS}$ et $\hat{d}_R$ nous avons les mêmes vitesses de convergence pour tous types de processus. Entre $\hat{d}_{ATV}$ et $\hat{d}_N$, les estimateurs par ondelettes, les résultats sont aussi très proches avec des vitesses de convergence plus rapides pour les processus dont la densité spectrale est "régulière" (proche d'une simple loi de puissance), et moins rapides lorsque la densité spectrale est plus irrégulière. Cependant, notre estimateur offre globalement les meilleurs résultats en terme de robustesse et de vitesse de convergence.

2.2.2 Le cas des processus linéaires

On considère un processus linéaire défini par

$$X_t = \sum_{s \in \mathbb{Z}} \alpha(t-s)\xi_s, \quad t \in \mathbb{Z}$$

où (ξ_t) est un bruit blanc fort, $(\alpha(k))_k$ une famille de réels qui vérifie

$$\begin{aligned} |\widehat{\alpha}(\lambda)|^2 &= \frac{1}{\lambda^{2d}}\big(c_d + c_{d'}\lambda^{d'}(1+\varepsilon(\lambda))\big) \quad \text{pour tout} \quad \lambda \in [-\pi,\pi] \quad \text{et} \quad \varepsilon(\lambda) \to 0 (\lambda \to 0) \\ \widehat{\alpha}(\lambda) &:= \frac{1}{2\pi}\sum_{k \in \mathbb{Z}} \alpha(k) e^{-ik\lambda}. \end{aligned}$$

La densité spectrale f du processus admet donc le même développement asymptotique puisqu'elle s'écrit $f(\lambda) = 2\pi|\widehat{\alpha}(\lambda)|^2$. X est donc un processus stationnaire, centré, d'ordre deux et à longue mémoire.

Pour une échelle et $(a,b) \in \mathbb{N}_+^* \times \mathbb{Z}$, on définit le coefficient d'ondelette à partir d'un échantillon $(X_1, \dots, X_N)$ du processus X comme suit :

$$e_N(a,b) := \frac{1}{\sqrt{a}}\sum_{t=1}^{N} X_t \psi(\frac{t-b}{a}) = \frac{1}{\sqrt{a}}\sum_{t=1}^{N}\sum_{s \in \mathbb{Z}} \alpha(t-s)\psi(\frac{t-b}{a})\xi_s \qquad (2.2.9)$$

Sous certaines conditions, la variance de ce coefficient vérifie :

$$\mathbb{E}(e^2(a,0)) = 2\pi\, c_d \left(K_{(\psi,2d)}\, a^{2d} + \frac{c_{d'}}{c_d} K_{(\psi,2d-d')}\, a^{2d-d'}\right) + o\big(a^{2d-d'}\big) \quad \text{quand} \quad a \to \infty, \qquad (2.2.10)$$

On considère la La variance empirique des coefficients d'ondelettes comme suit :

$$T_N(a) := \frac{1}{N-a}\sum_{k=1}^{N-a} e^2(a,k) \quad \text{avec} \quad v1 \leq a < N.$$

Cette différence avec le cas gaussien s'explique par le fait qu'une telle expression conduit à une matrice de covariance asymptotique bien plus simple et une meilleure vitesse de convergence de la variance empirique vers la variance théorique.

Comme premier résultat, nous avons sous certaines hypothèses ce théorème limite central :

Proposition 2.2 *Soient $\ell \in \mathbb{N} \setminus \{0,1\}$, des échelles $(r_1, \cdots, r_\ell) \in (\mathbb{N}^*)^\ell$. et $(a_n)_{n\in\mathbb{N}}$ tel que :*

$$\begin{aligned} N/a_N &\underset{N\to\infty}{\longrightarrow} \infty \\ a_N\, N^{-1/(1+2d')} &\underset{N\to\infty}{\longrightarrow} \infty. \end{aligned}$$

Alors

$$\sqrt{\frac{N}{a_N}}\Big(\log T_N(r_i a_N) - 2d\log(r_i a_N) - \log\big(\frac{c_d}{2\pi}K_{(\psi,2d)}\big)\Big)_{1\leq i\leq \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,d)\big), \tag{2.2.11}$$

avec $\Gamma(r_1, \cdots, r_\ell, \psi, d) = (\gamma_{ij})_{1\leq i,j\leq \ell}$ et

$$\gamma_{ij} = 4\pi\,\frac{(r_i r_j')^{1-2d}}{K^2_{(\psi,2d)}}\int_{-\infty}^{\infty}\frac{\big|\widehat{\psi}(r_i\lambda)\big|^2\big|\widehat{\psi}(r_j\lambda)\big|^2}{\lambda^{4d}}d\lambda. \tag{2.2.12}$$

La même procédure que dans le cas gaussien sera adoptée. L'estimateur de l'échelle optimale $\widehat{\alpha}_N$ est obtenu par minimisation de la même fonctionnelle $Q_N(\alpha, c, d)$. Cet estimateur vérifie aussi la propriété :

$$\widehat{\alpha}_N = \frac{\log \widehat{a}_N}{\log N} \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^* = \frac{1}{1+2d'}.$$

On définit alors le premier estimateur de d par $\widehat{\widehat{d}} := \widehat{d}(N^{\widehat{\alpha}_N})$. Les mêmes raisons que dans le cas gaussien nous mènent cependant à lui préférer l'estimateur adaptatif $\widetilde{d}$, construit de la manière suivante :

on définit d'abord un nouvel estimateur de l'échelle minimale par :

$$\widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6\widehat{\alpha}_N}{(\ell-2)(1-\widehat{\alpha}_N)}\,\frac{\log\log N}{\log N}.$$

Le nouvel estimateur adaptatif de d est obtenu par les pseudo-moindres carrés généralisés de matrice de covariance associée : $\widehat{\Gamma}_N := \Gamma(1, \cdots, \ell, \widehat{\widehat{d_N}}, \psi)$ qui converge vers $\Gamma(1, \cdots, \ell, d, \psi)$. L'estimateur $\widetilde{d}_N$ vérifie sous les conditions du théorème limite centrale précédent :

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}}\big(\widetilde{d}_N - d\big) \xrightarrow[N\to\infty]{D} \mathcal{N}(0\,;\,\sigma_d^2(\ell)) \text{ et } \forall \rho > \frac{2(1+3d')}{(\ell-2)d'},\quad \frac{N^{\frac{d'}{1+2d'}}}{(\log N)^\rho}\cdot\big|\widetilde{d}_N - d\big| \xrightarrow[N\to\infty]{\mathcal{P}} 0.$$

Un test d'adéquation adaptatif est proposé. Il est basé sur la somme des carrés des résidus obtenus par la régression PGLS [10] munie de la métrique associée à la matrice $\widehat{\Gamma}_N$. Plus précisément, soit $\tilde{e}_{\tilde{\alpha}}$ le vecteur résidu des ℓ points alors

$$\tilde{T_N} = \frac{N}{N\tilde{\alpha}_N}\Big(\tilde{e}_{\tilde{\alpha}}\Big)'\hat{\Gamma}_N^{-1}\Big(\tilde{e}_{\tilde{\alpha}}\Big)$$

Cette statistique (sous conditions) suit une loi de Khi-deux (à $(\ell-2)$dl). Ce test peut etre vu comme un test de longue mémoire pour les processus linéaires.

10. Pseudo Generalized least square

Simulations

Des simulations ont ensuite été effectuées :

- Permettant de vérifier la consistance et la robustesse de $\widetilde{d}_N$.
- Des comparaisons avec d'autres estimateurs semiparamétriques.
- La consistance et la robustesse du test d'adéquation adaptatif de la statistique $\widetilde{T}_N$.

On choisit trois types de processus de référence (FGN, FARIMA et un processus à densité donnée) pour lesquels on génère 100 trajectoires de différentes tailles N pour différentes valeurs de d.

Une première étape consistera à montrer que la valeur de la variance asymptotique dépend très peu de d ce qui est exhibé par une figure donnant pour différentes de N le graphe de $\sigma_d^2(l)$ en fonction de d.

les simulations montrent que notre estimateur $\widetilde{d}_N$ est performant de une vitesse de convergence rapide comparée aux autres estimateurs.

La robustesse de $\widetilde{d}_N$

Trois processus ne vérifiant pas les hypothèses émis au départ sont sélectionnés. Le premier étant gaussien à densité spectrale donnée, le second $FARIMA$ avec tendance linéaire et le troisième $FARIMA$ avec tendance linéaire et composante saisonnière sinusoïdale. Pour différentes valeurs de d et N on donne les valeurs de $\sqrt{MSE}$ pour les trois estimateurs $\widetilde{d}_N, \widehat{d}_M S, \widehat{d}_R$ ainsi que la probabilité d'acceptation du test d'adéquation. Il en résulte que notre estimateur adaptatif $\widetilde{d}_N$ est robuste à toutes les tendance et composantes saisonnière.

La consistance et la robustesse du test d'adéquation

La consistance du test à déjà été vérifiée dans les tableaux précédents. On vérifie sa robustesse en sélectionnant trois processus qui ne vérifient pas la condition de stationnarité ou celle de la forme de la densité spectrale exigée par les hypothèses. Pour différentes valeurs de N le degré d'acceptation (pour un risque de 0.05) est donné. Il en résulte que l'hypothèse de longue mémoire est rejetée pour deux des processus mais pas le troisième (qui présente une rupture en le paramètre d : dans ce cas c'est le comportement “moyen” qui l'emporte car notre statistique considère le processus sur sa globabilité).

Bibliographie

[1] Abry, P., Veitch, D. and Flandrin, P. (1998). Long-range dependent : revisiting aggregation with wavelets, *J. Time Ser. Anal.* , 19, 253–266.

[2] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377–390.

[3] Abry, P and Veitch, D A wavelet based joint estimator of the parameter of long-range dependence traffic, *IEEE Trans. on Info. Theory, Special Issue on Multiscale Statistical Signal Analysis and its Application*, 45, 878–897, 1999.

[4] Abry, P. and Veitch, D. Wavelet analysis of long-range-dependent traffic, *IEEE Trans.on Info.Theory*, 44, 2–15, 1998.

[5] Abry, P., Helgason, H. and Pipiras, V. Wavelet-based Analysis og Non-Gaussian Long-Range Dependant processes and Estimation of the Hurst Parameter. Preprint, Submitted to Elsevier.

[6] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. Self-similarity and long-range dependence through the wavelet lens, In *Long-range Dependence : Theory and Applications*,P. Doukhan, G. Oppenheim and M.S. Taqqu editors, Birkhäuser, 2003.

[7] Abry, P, Flandrin, P., Taqqu, M.S. and Veitch, D. Wavelets for the analysis estimation and synthesis of scaling data, In *Self-similar Network Traffic and Performance Evaluation*,K.Park and W.Willinger, eds. Willey (Interscience Division), 2000, New York.

[8] Andrews, D.W.K, Sun, Y. Adaptive local polynomial Whittle estimation of long-range dependence, *Econometrica*, 72, 569–614,2004.

[9] Andrews, W, Lo, A. Long-Terms Memory In Stock Market Prices *Econometrica*, 59 : 1279-1313, 1991

[10] Bhansali, R. Giraitis L. and Kokoszka, P.S. Estimation of the memory parameter by fitting fractionally-differenced autoregressive models, *J. Multivariate Analysis*, 97, 2101–2130, (2006)

[11] Bardet, J.-M. Testing for the presence of self-similarity of Gaussian time series having stationary increments. *Journal of Time Series Analysis.* 21 497-515. (2000).

[12] Bardet, J.-M. Statistical study of the wavelet analysis of fractional Brownian motion. IEEE Trans. Inform. Theory 48 991-999, (2002)

[13] Bardet, J.-M., Lang, G., Moulines, E. and Soulier, P. Waveletest esimator of long-range dependent processes. *Stat. Inference Stoch.Process.* 3 85-99. 19th Rencontres Franco-Belges de Statisticiens (Marseille, 1998).

[14] Bardet, J.M,. Bibi, H. and Jouini, A. Adaptive wavelet based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli Society*,691-724,vol. 14,3, 2008.

[15] Bardet, J.M., Lang, G., Moulines, E. and Soulier, P. Wavelet estimator of long range-dependant processes, *Statist. Infer. Stochast. Processes*, Vol. 3,85-99, 2000.

[16] Bardet, J.M. Statistical Study of the Wavelet Analysis of Fractional Brownian Motion, *IEEE Trans. Inf. Theory*, Vol. 48, No. 4, 991-999, 2002.

[17] Bardet, J.M and Bertrand, P. Identification of the multiscale fractional Brownian motion with biomechanical applications, *J. Time Ser. Anal.*, Vol. 28, 1-52,2007.

[18] Bardet, J.M. and Tudor, C. *A wavelet analysis of the Rosenblatt process : Chaos expansion and estimation of the self-similarity parameter*, Preprint, 2009.

[19] Bardet, J.M. Testing for the presence of self-similarity of Gaussian time series having stationary increments, *J. Time Ser. Anal.* , 21, 497–516,2000.

[20] Davis, R.B., Harte, D.S. Test for Hurst effect. *Biometrika*, 74 :95-101, 1987.

[21] Dickey, D.A and Fuller, W.A, Distribution of the estimatorfor the autoregressive form time series with unit roots, Journal of the American Association, vol 74 : issue 366, pp427-431, Jun 1979

[22] Flandrin, P. Wavelet analysiss and synthesis of fractionnal Brownian motion, *IEEE Trans. Inf. Theory,* 38, 910-917, MR1162229. 1992

[23] Fox, R. and Taqqu, M.S. Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14(2) : 517–532, 1986.

[24] Fox, R., Taqqu, M.S. CentraI limit theorems for quadratic forms in random variables having Iong-range dependence. *Probab. Th. Rel. Fields*. 74, 213-240 (1987)

[25] Fuller, W.A, Introduction to statistical time series New york, Wiley,1976.

[26] Geweke, J and Porter-Hudak, S. The estimation and application of long-memory time series models. *J. Time Ser. Anal*, 4 :221–238, 1983.

[27] Giraitis, L., Robinson, P.M., Samarov, A. Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependance. *Journal of Time Series Analysis*, 18 :49-61,1997.

[28] Giraitis, L., Robinson P.M., and Samarov, A. Adaptive semiparametric estimation of the memory parameter, *J.* Multivariate Anal., 72, 183–207(2000).

[29] Granger, C. and Joyeux, R. An introduction to long-memory time series models and fractional differencing. *J.* Time Ser. Anal., 1 :15–29,1980

[30] Guyon, X., Leon, J. Convergence en loi des H-varaitions d'un processus gaussien stationnaire. *Ann. Inst. Poincaré*, 25,265-282

[31] Hall, P., Welsh, A.H. Best attainable rates of convergence for estimates of parameters of regular variation. *The Annals of Statistics*,12(3) :1079-1084,1984.

[32] Henry, H. Bandwidth choicie, optimal rates and adaptivity in semiparametric estimation of long memory. in *Applications of Long Memory in Economics*, A. Kirman and G. Teisseyre eds., pp 157-172, Springer : Heidelberg (2007)

[33] Hurst, H.E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, Vol.116, pp 770–799, 1951.

[34] Istas, J., Lang, G. Quadratic variations and estimation of the local Hölder index of a gaussian process. *Ann. Inst. Poincaré*, 33,407-436.

[35] Künsch, H. R. , Statistical aspects of self-similar processes, in *Proceedings of the First World Congress of the Bernoulli Society*,Y. Prokhorov V. Sazanov, eds,Vol. 1, VNU Science Press, Utrecht, pp. 67.74.(1987)

[36] Lo, A and Mackinlay, C. The Size and Power of the Variance Ratio Test in Finite Samples : A Monte Carlo Investigation. *Journal of Econometrics*, 4 : 230-238, 1989

[37] Mandelbrot, B. and Taqqu, M.S. Robust R/S analysis of long-run serial correlation. *Bullletin of the International Statistical Institute*, 48 : 69-99,1979

[38] Moulines, E., Roueff, F. and Taqqu, M.S. Central limit theorem for the log-régression wavelet estimation of the memory parameter in the Gaussian semi-parametric context, *fractals*, 15(4),301-313,2007.

[39] Moulines, E., Soulier, P. Data driven order sélection for projection estimator of the spectral density of time series with long-range dependence, *J. Time Series Anal.* Vol 21, (2), pp 193-218, March 2000.

[40] Moulines, E., Roueff, F. and Taqqu, M.S. On the spectral density estimation of the wavelet coefficients of long memory time series with application to the log-régression estimation of the memory parameter, *J. Time Series Anal.* Vol(28),2007.

[41] Moulines, E., Soulier, P. Semi-parametric spectral estimation for fractionnal processes, In *Theory and applications of long range dependance* , Birkhäuser Boston, MA 251-30, 2003.

[42] Robinson, P.M. Non parametric and semi-parametric methods in economics and statistics, *Proceedings of the Fifth international synmposium in economic theory and econometrics*, W.Barnett, J.Powell and G.Tauchen eds., 1991, pp.437-457.

[43] Robinson, P.M. Log-periodogram régression of time series with long range dependence. *The Annals of Statistics*, 23 : 1048-1072, 1995.

[44] Robinson, P.M. Gaussian semiparametric estimation of long-range dependence. *The Annals of Statistics*, 23(5) : 1630-1661, 1995.

[45] Roueff, F and Taqqu, M. S. (2009a). Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *J.* Time Ser. Anal., 30, 534–558.

[46] Veitch, D., Abry, P. and Taqqu, M. S. On the automatic sélection of the onset of scaling. *Fractals*. 11 377-390, 2003

[47] Veitch, D., Taqqu, M. S. and Abry, P. . Meaningful MRA initialisation for discrete time series. *Signal Processing* 80 1971-1983.(2000)

[48] Willinger, W., Taqqu, M.S, and Teverovsky, V. Stock market price and long-range dependence, *Finance and statistics*, 1 : 1-14, 1999.

Deuxième partie

Articles

Chapitre 1

More General Constructions of Wavelets on the Interval

More General Constructions of Wavelets on the Interval

Published in Communications in Mathematical Analysis,. vol.4 -1(2008), pp.45-57.

Hatem Bibi[a,b], Abdellatif Jouini[b], Mouna Kratou[b]

hatem.bibi@malix.univ-paris1.fr, Abdellatif.jouini@fst.mu.tn, Kratoumouna@yahoo.fr

[a] Samos-Matisse-CES, Université Paris 1, CNRS UMR 8174,90 rue de Tolbiac, 75013 Paris, FRANCE.
[b] Département de Mathématiques, Faculté des Sciences de Tunis,1060 Tunis, TUNISIE.

abstract

In this paper we present general constructions of orthogonal and biorthogonal multiresolution analysis on the interval. In the first one, we describe a direct method to define an orthonormal multiresolution analysis. In the second one, we use the integration and derivation method for constructing a biorthogonal multiresolution analysis. As applications, we prove that these analyses are adapted to study regular functions on the interval.

1.1 Introduction

The search for wavelet bases on a bounded domain has been an active field for many years, since the beginning of the 1990's. All these constructions use either the basis of I. Daubechies or the spline basis. In his fundamental paper on wavelets on the interval [14], Y. Meyer proved that one can take restrictions of the orthonormal multiresolution analysis of I. Daubechies to the interval $[0, 1]$ and then we can study functions known only on the interval. More precisely, he proves that the restrictions of Daubechies scaling functions on the interval are linearly independent but the restrictions of associated wavelets on the interval are not linearly independent.

In 1992, we have constructed multiresolution analysis on the interval by using Daubechies wavelets [9]. The associated bases have compact support and allow also the study of divergence-free vector functions on $[0,1]^n$.

There are related constructions as well by A. Canuto and coworkers [1] and by A. Jouini and P. G. Lemarié ([8] and [10]).

In this paper we aim to generalize the result for every orthonormal multiresolution analysis. Next, we present orthogonal and biorthogonal systems on $[0, 1]$ which are constructed by means of dyadic translations and dilatations from a finite number of basic functions and are well-adapted to study Sobolev spaces $H^s([0, 1])$ and $H^s_0([0, 1])$ $(s \in \mathbb{Z})$.

The contents of this paper is the following.

In Section 1.2, we at first define and construct new orthogonal multiresolution analysis on the interval $[0, 1]$. Next, we prove the Meyer's lemma [14] for the general case of an orthonormal multiresolution analysis with compact support. Then,we construct the associated wavelet bases which are more technical. In section 1.3, we study biorthogonal multiresolution analysis (V_j, V_j^*) $(j \in \mathbb{Z})$ on the interval $[0, 1]$. By a derivation on V_j and an integration on V_j^*, we get a new biorthogonal multiresolution analysis $(\widetilde{V_j}, \widetilde{V_j}^*)$ of the space $L^2([0, 1])$. If we denote P_j the projector from $L^2([0, 1])$ on V_j parallel to $(V_j^*)^{\perp}$and $\widetilde{P_j}$ be the projector in $\widetilde{V_j}$ parallel to $(\widetilde{V_j}^*)^{\perp}$, then we have the following commutation property

$$\frac{d}{dx} o P_j = \widetilde{P_j} o \frac{d}{dx}.$$

The section 1.4 is devoted to applications. We prove that the biorthogonal multiresolution analysis constructed in section 1.3 is adapted to study Sobolev spaces $H^s([0, 1])$ and $H^s_0([0, 1])$ for $s \in \mathbb{Z}$.

1.2 Orthogonal multiresolution analysis on the interval [0,1]

It is clear that if we consider an orthogonal multiresolution analysis, and if we take its restriction to $[0,1]$, we do not get an orthogonal multiresolution analysis of $L^2([0,1])$. Moreover, for the orthogonal multiresolution analysis $V_j(\mathbb{R})$ of I. Daubechies, if we consider the associated scaling functions $\varphi_{j,k}(x)_{[0,1]}$, we have an independent system which is not orthogonal. However, if we consider the associated wavelets $\psi_{j,k}(x)_{[0,1]}$, we get a dependent system (see [14]) and the support of the wavelet ψ is very important in this case. Then, the construction of an orthogonal multiresolution analysis in $[0,1]$ (or biorthogonal) is technical especially near the boundaries 0 and 1.

In this section, we shall prove the precedent result for any orthogonal multiresolution analysis with compact support. More precisely, we use a direct method based on the result described in [14] to construct orthogonal multiresolution analysis on the interval $[0,1]$ which are generated by a finite number of basic functions. These analyses are regular and have compact support.

For this purpose, we consider an orthogonal multiresolution analysis $V_j(\mathbb{R})$ of $L^2(\mathbb{R})$ where the scaling function φ has a compact support $[N_1, N_2]$. We recall first the scaling equations for this analysis. The inclusion $V_0 \subset V_1$ gives the two following equations

$$\varphi(\frac{x}{2}) = \sum_{k=N_1}^{N_2} a_k \varphi(x-k) \quad \text{where } a_{N_1} a_{N_2} \neq 0 \tag{1.2.1}$$

and

$$\hat{\varphi}(2\xi) = m_0(\xi)\hat{\varphi}(\xi) \quad \text{where} \quad m_0(\xi) = \frac{1}{2}\sum_{k=N_1}^{N_2} a_k e^{-ik\xi}. \tag{1.2.2}$$

We assume that the associated wavelet ψ has the same support (by a simple translation) and then is defined by

$$\psi(\frac{x}{2}) = \sum_{k=N_1}^{N_2} b_k \varphi(x-k) \quad \text{where} \quad b_{N_1} b_{N_2} \neq 0. \tag{1.2.3}$$

Note that we cannot define in the same manner as classical wavelet theory the notion of multiresolution analysis in the interval because we do not have the invariance and dilatation properties in a bounded domain. Then, we present differently this notion. Let j_0 be an integer such that $2^{j_0} \geq 2(N_2 - N_1 - 1)$ (we can separate the boundaries functions). We denote

$$V_j([0,1]) = Vect\{\varphi_{j,k[0,1]}, \varphi_{j,k} \in V_j(\mathbb{R})\}, \tag{1.2.4}$$
$$v_j([0,1]) = Vect\{\varphi_{j,k}, supp\varphi_{j,k} \subset [0,1]\}. \tag{1.2.5}$$

Definition 1.2.1 *A sequence $\{V_j\}_{j \geq j_0}$ of closed subspaces of $L^2([0,1])$ is called a multiresolution analysis on $L^2([0,1])$ associated with $V_j(\mathbb{R})$ if*

i) $\forall j \geq j_0, v_j([0,1]) \subset V_j \subset V_j([0,1])$

ii) $\forall j \geq j_0, V_j \subset V_{j+1}$.

It is clear that these spaces contain a finite number of functions due to compacity of the support and then the Gram-Schmidt method gives orthonormal systems if these systems are linearly independent. We now proceed to prove an elementary lemma which will be useful in analysis for functions defined on the interval $[0,1]$. We begin by the case of the interval $]-\infty,0]$. In fact, we prove that only the functions $\varphi_{j,k}$ whose support intersects the interval $]-\infty,0[$ occur in the analysis of an arbitrary function in $V_0(\mathbb{R})$ and with support in $]-\infty,0]$.

Lemma 1.1 *If* $f(x)=\sum_{-\infty}^{+\infty} c_k\varphi(x-k)$ *is a function of* $V_0(\mathbb{R})$ *such that* $f(x)=0$ *for* $x\leq 0$. *Then* $c_k=0$ *for* $k\leq -N_1-1$.

Proof *The support of the function* $\varphi(x-k)$ *is* $[k+N_1,k+N_2]$ *and then is included in* $]-\infty,0]$ *for* $k\leq -N_2$. *We have* $c_k=\int_{-\infty}^{+\infty} f(x)\overline{\varphi(x-k)}dx=0$ *for* $k\leq -N_2$.

Let p *be the smallest integer of* k *such that* $c_k\neq 0$. *If* $p\geq -N_1$, *then we have the result. If* $p<-N_1$, *then* $f(x)=0$ *on the interval* $[p+N_1,p+N_1+1]$ *Because the support of the scaling function* φ *is equal to* $[N_1,N_2]$. *Using the hypothesis that* f *is a function of* $V_0(\mathbb{R})$, *we obtain* $f(x)=c_p\varphi(x-p)$. *Then, we have a contradiction.*

The following result generalizes the result of Y. Meyer in [14] and gives an other multiresolution analysis of $L^2([0,1])$.

Theorem 1.1 *Let* $j\geq j_0$ *and* $f(x)=\sum_{-\infty}^{+\infty} c_k\varphi(2^jx-k)$ *be a function of* $V_j(\mathbb{R})$ *such that* $f(x)=0$ *for* $0\leq x\leq 1$. *Then* $c_k=0$ *for* $-N_2+1\leq k\leq 2^j-N_1-1$.

Proof *Let* $j\geq j_0$ *and* $2^{j_0}\geq 2(N_2-N_1-1)$, *we can consider three cases.*

1. *If* $-N_2+1\leq k\leq -N_1-1$, *the support of the scaling functions* $\varphi_{j,k}$ *is included in* $]-\infty,N_2-N_1-1]\subset]-\infty,\frac{1}{2}]$.
2. *If* $-N_1\leq k\leq 2^j-N_2$, *the support of the scaling functions* $\varphi_{j,k}$ *is included in* [0,1].
3. *If* $2^j-N_2+1\leq k\leq 2^j-N_1-1$, *the support of the scaling functions* $\varphi_{j,k}$ *is included in* $[N_2-N_1-1,+\infty[\subset[\frac{1}{2},+\infty[$.

We see that in the case 2 , we have

$$c_k=\int_{-\infty}^{+\infty} f(x)\overline{\varphi(x-k)}dx=0.$$

Applying Lemma 1.1 to the first case and the third case, we get $c_k=0$. *This yields* $c_k=0$ *for* $-N_2+1\leq k\leq 2^j-N_1-1$.

Theorem 1.1 is the basis for our strategy : to get the bases on the interval. As a consequence, we have the following result.

Corollary 1.1 *For* $j\geq j_0$, *the functions* $\varphi_{j,k/[0,1]}$, $-N_2+1\leq k\leq 2^j-N_1-1$, *form a Riesz basis of the space* $V_j([0,1])$.

Remark 1.1 *The results described above are true for every integer* j *by using an iteration and Lemma 2 in [14].*

Corollary 1.2 *For $j \geq j_0$,*

i) there exist $(N_2 - N_1 - 1)$ functions φ_i^α $(1 \leq i \leq N_2 - N_1 - 1)$ and $(N_2 - N_1 - 1)$ functions φ_i^β $(1 \leq i \leq N_2 - N_1 - 1)$ such that the functions

$$\begin{aligned} \varphi_{i,j}^\alpha &= 2^{j/2}\varphi_i^\alpha(2^j x)|_{[0,1]}, (1 \leq i \leq N_2 - N_1 - 1), \\ \varphi_{j,k} &= 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2), \\ \varphi_{i,j}^\beta &= 2^{j/2}\varphi_i^\beta(2^j x - 2^j)|_{[0,1]}, (1 \leq i \leq N_2 - N_1 - 1), \end{aligned}$$

form an orthonormal basis of $V_j([0,1])$.

ii) Let $V_j, j \geq j_0$ (for large value j), be a multiresolution analysis of $L^2([0,1])$ associated with $V_j(\mathbb{R})$, then there exist N_α functions φ_i^α $(1 \leq i \leq N_\alpha)$ and N_β functions φ_i^β $(1 \leq i \leq N_\beta)$ such that the functions

$$\begin{aligned} \varphi_{i,j}^\alpha &= 2_i^{j/2}\varphi_i^\beta(2^j x)|_{[0,1]}, (1 \leq i \leq N_\alpha) \\ \varphi_{j,k} &= 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2), \\ \varphi_{i,j}^\beta &= 2^{j/2}\varphi_i^\beta(2^j x - 2^j)|_{[0,1]}, (1 \leq i \leq N_\beta) \end{aligned}$$

form an orthonormal basis of V_j.

Proof *It is clear now how to get an orthogonal basis of $V_j([0,1])$. It is enough to apply Gram-Schmidt to functions $\varphi_{j,k/[0,1]}$, $-N_2 + 1 \leq k \leq -N_1 - 1$ (near the boundary 0) and then to functions $\varphi_{j,k/[0,1]}$, $2^j - N_2 + 1 \leq k \leq 2^j - N_1 - 1$ (near the boundary 1). In every case, we have $(N_2 - N_1 - 1)$ functions. We obtain new functions $\varphi_{i,j}^\alpha = 2^{j/2}\varphi_i^\alpha(2^j x)|_{[0,1]}$, $(1 \leq i \leq N_2 - N_1 - 1)$ near the boundary 0 and in the same way new functions $\varphi_{i,j}^\beta = 2^{j/2}\varphi_i^\beta(2^j x - 2^j)|_{[0,1]}$, $(1 \leq i \leq N_2 - N_1 - 1)$ near the boundary 1. To prove ii, we apply the method described above to every multiresolution analysis on the interval defined as Definition 1.2.1.*

Remark 1.2 *It is easy to see that the space V_j contains the orthonormal system*

$$\varphi_{j,k} = 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2)$$

and we add boundaries functions near 0 and 1 from the collections $\varphi_{i,j}^\alpha$ and $\varphi_{i,j}^\beta$.

We define

$$V_j^T([0,1]) = \{f \in V_j([0,1])/f|_T = 0\},$$

where $T \subset \{0,1\}$ and $j \geq j_0$. We obviously have

$$V_j^T([0,1]) \subset V_{j+1}^T([0,1]).$$

The corresponding spaces $V_j^T([0,1]$ are generated by the functions $(\varphi_{j,k})|_{[0,1]}$, $k \in D_j^T$ where the set D_j^T is defined by

* $D_j^T = \{k| -N_1 \leq k \leq 2^j - N_2\}$ if $T = \{0,1\}$.
* $D_j^T = \{k| -N_1 \leq k \leq 2^j - N_1 - 1\}$ if $T = \{0\}$.

* $D_j^T = \{k| -N_2+1 \leq k \leq 2^j - N_2\}$ if $T=\{1\}$.
* $D_j^T = \{k| -N_2+1 \leq k \leq 2^j - N_1 - 1\}$ if $T=\emptyset$.

Using Corollary 2.2, we obtain an orthonormal basis of $V_j^T([0,1])$.

Theorem 1.2 *The space $V_j^T([0,1])$ has orthonormal basis $(\varphi_{j,k}^T)$, $k \in D_j^T$ where*

i) *if $T=\{0,1\}$*

$$\varphi_{j,k}^T = \varphi_{j,k} = 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2)$$

.

ii) *if $T=\{0\}$*

$$\begin{aligned}\varphi_{j,k}^T &= \varphi_{j,k} = 2^{j/2}\varphi(2^j x - k)(-N_1 \leq k \leq 2^j - N_2) \\ &= \varphi^{\beta}_{j,k-2^j+N_2}, (2^j - N_2 + 1 \leq k \leq 2^j - N_1 - 1)\end{aligned}$$

iii) *if $T=\{1\}$*

$$\begin{aligned}\varphi_{j,k}^T &= \varphi^{\alpha}_{j,k+N_2}, (-N_2+1 \leq k \leq -N_1 - 1) \\ &= \varphi_{j,k} = 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2)\end{aligned}$$

iv) *if $T=\emptyset$.*

$$\begin{aligned}\varphi_{j,k}^T &= \varphi^{\alpha}_{j,k+N_2}, (-N_2+1 \leq k \leq -N_1 - 1), \\ &= \varphi_{j,k} = 2^{j/2}\varphi(2^j x - k), (-N_1 \leq k \leq 2^j - N_2) \\ &= \varphi^{\beta}_{j,k-2^j+N_2}(2^j - N_2 + 1 \leq k \leq 2^j - N_1 - 1)\end{aligned}$$

We conclude that the orthogonal projector P_j^T from $L^2([0,1])$ into $V_j^T([0,1])$ is given by

$$P_j^T f = \sum_{k \in D_j^T} \langle f|\varphi^T_{(j,k)}\rangle \varphi^T_{(j,k)},$$

and satisfies $P_j^T o P_{j+1}^T = P_{j+1}^T \, o \, P_j^T = P_j^T$.

In the following, we establish the second goal of this paper. In fact, we should construct a wavelet basis of the space $W_j([0,1]) = V_{j+1}([0,1]) \cap (V_j([0,1]))^{\perp}$. We denote by

$$V_j([N_1,+\infty[) = Vect\{\varphi_{j,k/[N_1,+\infty[}, \varphi_{j,k} \in V_j(\mathbb{R})\}. \tag{2.9}$$

Recall that the QMF condition gives that the mask of an orthonormal scaling function must have an even number of coefficients. This means that $N_2 - N_1$ is odd. We have the first important result.

Lemma 1.2 *The functions $\psi(x-k)_{/[N_1,+\infty[}$, $N_1 - N_2 + 1 \leq k \leq -\frac{1}{2}(N_2 - N_1 + 1)$, belong to $V_0([N_1,+\infty[)$.*

Proof *The relations 1.2.1 and 1.2.3 gives*

$$\begin{aligned}\varphi(2x) &= \overline{a_{N_1}}\varphi(x+\frac{1}{2}N_1)+\overline{a_{N_1+2}}\varphi(x+\frac{1}{2}N_1+1)+\ldots+\overline{a_{N_2-1}}\varphi(x+\frac{1}{2}N_2-\frac{1}{2}) \quad (1.2.6)\\ &+ \overline{b_{N_1}}\psi(x+\frac{1}{2}N_1)+\overline{b_{N_1+2}}\psi(x+\frac{1}{2}N_1+1)+\ldots+\overline{b_{N_2-1}}\psi(x+\frac{1}{2}N_2-\frac{1}{2})\end{aligned}$$

and

$$\begin{aligned}\varphi(2x-1) &= \overline{a_{N_1+1}}\varphi(x+\frac{1}{2}N_1)+\overline{a_{N_1+3}}\varphi(x+\frac{1}{2}N_1+1)+\ldots+\overline{a_{N_2-1}}\varphi(x+\frac{1}{2}N_2-\frac{1}{2}) (1.2.7)\\ &+ \overline{b_{N_1+1}}\psi(x+\frac{1}{2}N_1)+\overline{b_{N_1+3}}\psi(x+\frac{1}{2}N_1+1)+\ldots+\overline{b_{N_2-1}}\psi(x+\frac{1}{2}N_2-\frac{1}{2}).\end{aligned}$$

We replace now x by $x+N_2-\frac{3}{2}N_1-1$ in 1.2.6, then by $x+N_2-\frac{3}{2}N_1-2,\ldots$ and finally, by $x+\frac{1}{2}(N_2-2N_1+2)$. Recall that support of φ and ψ is $[N_1,N_2]$, then, we obtain

$$\overline{a_{N_1}}\varphi(x+N_2-N_1-1)+\overline{b_{N_1}}\psi(x+N_2-N_1-1)=0$$

Next

$$\overline{a_{N_1}}\varphi(x+N_2-N_1-2)+\overline{a_{N_1+1}}\varphi(x+N_2-N_1-1)+\overline{b_{N_1}}\psi(x+N_2-N_1-2)+\overline{b_{N_1+1}}\psi(x+N_2-N_1-1)=0$$

until the last equation. We conclude that the functions $\psi(x-k)_{[N_1,+\infty[}$, $N_1-N_2+1\le k\le -\frac{1}{2}(N_2-N_1+1)$, belong to $V_0([N_1,+\infty[)$.

Lemma 1.3 *The functions $\psi(2^jx-k)_{[0,1]}$, $-N_2+1\le k\le -\frac{1}{2}(N_2+N_1+1)$, belong to $V_j([0,1])$.*

Proof *By replacing x by $2^j(x-N_1)$ and using Lemma 1.2, we obtain the result.*

We reach the main result of this section

Theorem 1.3 *For $j\ge j_0$, the functions*

$$\begin{aligned}\varphi_{j,k/[0,1]} \quad &, \quad -N_2+1\le k\le 2^j-N_1-1\\ 2^{j/2}\psi(2^jx-k)_{/[0,1]} \quad &, \quad -\frac{1}{2}(N_2+N_1-1)\le k\le 2^j-\frac{1}{2}(N_2+N_1+1)\end{aligned}$$

form a Riesz basis of the spaces $V_{j+1}([0,1])$.

Proof *Lemmas 1.2 and 1.3 immediately imply the main result of this section.*

Remark 1.3 *If we apply the results described above to the orthogonal multiresolution of I. Daubechies, we obtain the Meyer's lemma in [14].*

We can obtain an orthogonal basis of $W_j([0,1])$. In fact, we do corrections to the functions $\psi_{j,k[0,1]}$, $-\frac{1}{2}(N_2+N_1-1)\le k\le -N_1-1$ to get orthogonality to $\varphi^{\alpha}_{i,j}$, $(1\le i\le (N_2+N_1-1)$. Then, by using Gram-Schmidt for new functions we get wavelets near 0.We do the same thing for the functions $\psi_{j,k/[0,1]}$, $2^j-N_2+1\le k\le 2^j-\frac{1}{2}(N_2+N_1-1)$ to get wavelets near 1. Moreover, the results result of A. Jouini and P. G. Lemarié given in [9] allows to construct the basis for every space W_j (orthogonal complement of V_j in V_{j+1}).

We will now construct an orthonormal basis of the space $W_j^T([0,1])$. We remark first that $\dim W_j^T([0,1]) = 2^j$. We denote by $\Delta_j^T = \{d \in D_{j+1}^T / d \notin D_j^T\}$. The space $W_j^T([0,1])$ contains the functions $\psi_{j,k}, -N_1 \leq k \leq 2^j - N_2$. We have $(2^j - (N_2 - N_1 - 1))$ functions in $W_j^T([0,1])$. Then, we must construct $(N_2 - N_1 - 1)$ functions. We denote by $A_j^T(I) = V_j^T(([0,1]) \oplus Vect\{\psi_{j,k} \; -N_1 \leq k \leq 2^j - N_2\}$. We see that, for $-N_1 \leq k \leq N_2 - 2N_1 - 2$, $\varphi_{j+1,k} \in A_j^T(I)$. We have the same treatment for $\varphi_{j+1,2^{j+1}-N_2}$. We conclude that $\varphi_{j+1,-N_1+2k}$ and $\varphi_{j+1,2^{j+1}-N_2-(2k+1)}$, $0 \leq k \leq \frac{N_2-N_1-1}{2} - 1$ form a generating system of a supplement of $A_j^T(I)$ in $V_{j+1}^T(I)$. Using Gram-Schmidt, we obtain an orthonormal basis $\psi_{j,k}^T$, $k \in \Delta_j^T$.

We conclude that the orthogonal projector Q_j^T from $L^2(I)$ into $W_j^T([0,1])$ is given by

$$Q_j^T f = \sum_{k \in \Delta_j^T} \langle f | \psi_{(j,k)}^T \rangle \psi_{(j,k)}^T,$$

and satisfies

$$Q_j^T o Q_{j+1}^T = Q_{j+1}^T \, o \, Q_j^T = Q_j^T.$$

1.3 Biorthogonal multiresolution analysis on the interval [0,1]

First we give some definitions of biorthogonal multiresolution analysis on the interval $[0,1]$, and then we describe constructions on this interval.

Définition 1.1 *A sequence (V_j, V_j^*) of closed subspaces of $L^2([0,1])$ associated with a biorthogonal multiresolution analysis $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ of $L^2(\mathbb{R})$ is called a biorthogonal multiresolution analysis of $L^2([0,1])$ if*

$$v_j([0,1]) \subset V_j \subset V_j([0,1]) \text{ and } v_j^*([0,1]) \subset V_j^* \subset V_j^*([0,1])$$
$$V_j \subset V_{j+1} \text{ and } V_j^* \subset V_{j+1}^*$$
$$L^2([0,1]) = V_j \oplus (V_j^*)^\perp$$

Let $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ be a biorthogonal multiresolution analysis of $L^2(\mathbb{R})$ with multiscale functions (g, g^*). We assume that supp(g)=$[N_1, N_2]$, and we denote by

$$P_i^\alpha(x) = \sum_{k \leq -N_1 - 1} k^i g(x-k),$$

and

$$P_i^\beta(x) = \sum_{k \geq -N_2 - 1} k^i g(x-k).$$

Our construction is based on the following result :

Theorem 1.4 *We consider a biorthogonal multiresolution analysis $(V_j(\mathbb{R}), V_j^*(\mathbb{R}))$ of $L^2(\mathbb{R})$, (g, g^*) are the multiscale functions with compact support and (V_j, V_j^*) the associated biorthogonal multiresolution analysis of $L^2([0,1])$. We assume that*

i) g is differentiable and $g'(x) = \widetilde{g}(x) - \widetilde{g}(x-1)$.

ii) V_j *contains the functions* $P^{\alpha}_{0,j}(x)=P^{\alpha}_0(2^j x)_{[0,1]}$ *and* $P^{\beta}_{0,j}(x)=P^{\beta}_0(2^j x-2^j)_{[0,1]}$. *If we denote by*

$$\widetilde{V_j}=\{f\in L^2([0,1])/\exists g\in V_j, f=g^{'}\},$$

and

$$\widetilde{V_j^*}=\{f\in L^2([0,1])\ /\ f^{'}\in V_j^*,\quad \text{and } f(0)=f(1)=0\}.$$

Then $(\widetilde{V_j},\widetilde{V_j^*})$ *is a biorthogonal multiresolution analysis of* $L^2([0,1])$. *Moreover, if we denote by* P_j *(resp* $\widetilde{P}_j$*) the projector from* $L^2([0,1])$ *into* V_j *(resp.* $\widetilde{V_j}$*) parallel to* $(V_j^*)^{\perp}$ *(resp* $(\widetilde{V_j}^*)^{\perp}$*), then we have the following commutation property*

$$\frac{d}{dx}oP_j=\widetilde{P}_j o\frac{d}{dx}.$$

Proof *We clearly have*

$$\begin{aligned}\widetilde{g}(x-k) &= (\sum_{p=0}^{\infty} g(x-k-p))^{'}\\ &\text{and}\\ (\widetilde{g}^*(x-k))^{'} &= g^*(x-k+1)-g^*(x-k).\end{aligned}$$

Then $\widetilde{v_j}\subset\widetilde{V_j}([0,1])$ *and* $\widetilde{v_j}^*\subset\widetilde{V_j}^*([0,1])$. *Moreover, since* V_j *contains the functions* $P^{\beta}_{0,j}(x)$, *we have* $\widetilde{V_j}([0,1])\subset\widetilde{V_j}$ *and* $\widetilde{V_j}^*([0,1])\subset\widetilde{V_j}^*$. *In the same way, we have*

$$\begin{aligned}\widetilde{V_j} &\subset \widetilde{V_{j+1}}\\ &\text{and}\\ \widetilde{V_j^*} &\subset \widetilde{V_{j+1}^*}.\end{aligned}$$

To see the duality between $\widetilde{V_j}$ *and* $\widetilde{V_j^*}$, *we consider a basis* $(\alpha_0=1,\alpha_1,\ldots,\alpha_n)$ *of* V_j *with* $dimV_j=n+1$ *and a dual basis* $(\beta_0,\beta_1,\ldots,\beta_n)$ *of* V_j^*. *Then the derivation is an isomorphism from* $\widetilde{V_j^*}$ *onto Vect*$(\beta_1,\ldots,\beta_n)$ *and from Vect*$(\alpha_1,\ldots,\alpha_n)$ *onto* $\widetilde{V_j}$. *If we define*

$$\widetilde{\alpha}_i=\frac{d}{dx}\,\alpha_i\qquad\text{and}\qquad\widetilde{\beta}_i=-\int_0^x\beta_i(t)dt,$$

then, by integration, we conclude that the bases $(\widetilde{\alpha_i})$ *and* $(\widetilde{\beta_i})$ *are biorthogonal and we have a duality between* $\widetilde{V_j}$ *and* $\widetilde{V_j^*}$. *Finally, the commutation property is satisfied. In fact, we have*

$$\begin{aligned}
\frac{d}{dx}\, o\,(P_j f) &= \frac{d}{dx}\langle f,\beta_0\rangle 1 + \sum_{i=1}^{n}\langle f,\beta_i\rangle \frac{d}{dx}\,\alpha_i \\
&= \sum_{i=1}^{n}\langle f,\beta_i\rangle \widetilde{\alpha}_i \\
&\text{and} \\
\widetilde{P}_j\, o\,(\frac{d}{dx}\, f) &= \sum_{i=1}^{n}\langle \frac{d}{dx}\, f, \widetilde{\beta}_i\rangle \widetilde{\alpha}_i \\
&= \sum_{i=1}^{n}(\Big[f\widetilde{\beta}_i\Big]_0^1 + \langle f,\beta_i\rangle)\widetilde{\alpha}_i, \\
&= \sum_{i=1}^{n}\langle f,\beta_i\rangle \widetilde{\alpha}_i.
\end{aligned}$$

Corollary 1.3 *Let $V_j(\mathbb{R})$ be the orthogonal multiresolution analysis of $L^2(\mathbb{R})$ with the scaling function φ of class C^m $(m \in \mathbb{N}^*)$. We denote by $V_j^{(m)}(\mathbb{R})$ and $V_j^{*(m)}(\mathbb{R})$ the multiresolution analysis constructed by m derivations and m integrations. Then $V_j^{(m)}([0,1])$ and $V_j^{*(m)}([0,1]) \cap H_0^m([0,1])$ form a biorthogonal multiresolution analysis of $L^2([0,1])$. Moreover, if we denote by $P_j^{(m)}$ the projector on $V_j^{(m)}([0,1])$ parallel to $[V_j^{*(m)}([0,1]) \cap H_0^m([0,1])]^\perp$, we have*

$$\frac{d}{dx} o P_j^{(m)} = P_j^{(m+1)} o \frac{d}{dx}.$$

We can apply the method described in section 1.2 to construct Riesz bases of the spaces $V_j^{(m)}([0,1])$ and $V_j^{*(m)}([0,1]) \cap H_0^m([0,1])$. In fact, we define g and g^* by

$$(1-e^{-i\xi})^m \overset{\curlywedge}{g}(\xi) = (i\xi)^m \overset{\curlywedge}{\varphi}(\xi) \tag{1.3.1}$$

and

$$(i\xi)^m \overset{\curlywedge}{g^*}(\xi) = (e^{i\xi}-1)^m \overset{\curlywedge}{\varphi}(\xi). \tag{1.3.2}$$

The functions $g_{j,k[0,1]}$ form a basis of $V_j^{(m)}([0,1])$. To construct a basis of $V_j^{*(m)}([0,1]) \cap H_0^m([0,1])$ we take the functions $g_{j,k}$ with support in $[0,1]$ and the boundaries functions defined by

$$g_{j,k}^{\alpha *} = \sum_{p=-N_2+1}^{-N_1+m-1} \alpha_{i,j,p}\, g^*_{j,p/[0,1]}, 1 \le i \le N_2 - N_1 - 1, \tag{1.3.3}$$

and

$$g_{j,k}^{\beta *} = \sum_{p=2^j-N_2+1}^{2^j-N_1+m-1} \alpha_{i,j,p}\, g^*_{j,p/[0,1]}, 1 \le i \le N_2 - N_1 - 1. \tag{1.3.4}$$

The real constants $\alpha_{i,j,p}$ are determined by the following conditions : for $1 \le i \le N_2 - N_1 - 2)$

$$\int_0^{+\infty} (\sum_{p=-N_2+1}^{-N_1+m-1} \alpha_{i,j,p} 2^{\frac{j}{2}} g^*(2^j x - p) 2^{\frac{j}{2}} g(2^j x + N_2 - N_1 - m - q) dx = \delta_{i,q}.$$

We define

$$V_j^T([0,1]) = \{f \in V_j([0,1]) / \ f|_T = 0\}$$

where $T \subset \{0,1\}$ and $j \geq j_0$. We obviously have

$$V_j^T([0,1]) \subset V_{j+1}^T([0,1]).$$

We shall construct a subspace $V_j^{*T}([0,1])$ of $V_j^*([0,1])$ such that $V_j^{*T}([0,1]) \subset V_{j+1}^{*T}([0,1])$ and $V_j^T([0,1])$ and $V_j^{*T}([0,1])$ are in duality for the scalar product on $[0,1]$. A direct method as in the previous section gives the basis of $V_j^T([0,1])$. A basis of $V_j^{*T}([0,1])$ is given by the functions $\varphi_{j,k}^*$ with compact support in $[0,1]$ and we add boundaries functions in a way similar in [**?**]. Using the Gram-Schmidt orthogonalization, we obtain biorthogonal bases of $V_j^T([0,1])$ and $V_j^{*T}([0,1])$. More precisely, Theorem 1.4 and Corollary 1.3 give a biorthogonal multiresolution analysis $(V_j^{(m),T}([0,1]),\ V_j^{*(m),T}([0,1]))$ of $L^2([0,1])$ and furthermore a straightforward computation yields

$$\frac{d}{dx} o P_j^{(m),T} = P^{(m+1),T} o \frac{d}{dx}.$$

A method similar to that used in the previous section shows that dual bases of $W_j^T([0,1])$ and $W_j^{*T}([0,1])$ are given by : $\psi_{(j,k)}^Z$ and $\psi_{(j,k)}^{*Z}$ for $k \in \Delta_j^T$.

1.4 The study of regular spaces of functions on the interval [0,1]

In this section, we give some applications of the multiresolution analysis on the interval $[0,1]$ described above. In fact, we study regular spaces of functions (Sobolev spaces) on the interval $[0,1]$.

We denote by

- $V_j(\mathbb{R})$: an orthogonal multiresolution analysis of $L^2(\mathbb{R})$ with the associated scaling function φ of class $C^{m+\varepsilon}$ on $\mathbb{R}$ $(m \in \mathbb{N}^*)$.
- $V_j^{(m)}(\mathbb{R})$: the multiresolution analysis constructed by derivation and g the function in $V_0^{(m)}(\mathbb{R})$ defined by

$$(1 - e^{-i\xi})^m \hat{g}(\xi) = (i\xi)^d \overset{\wedge}{\varphi}(\xi).$$

- $V_j^{*(m)}(\mathbb{R})$: the multiresolution analysis constructed by integration and g^* the function in $V_0^{*(m)}(\mathbb{R})$ defined by

$$(i\xi)^m \overset{\wedge}{g^*}(\xi) = (e^{i\xi} - 1)^d \overset{\wedge}{\varphi}(\xi).$$

- $V_j^{(m)} = V_j^{(m)}([0,1])$ and $V_j^{*(m)} = V_j^{*(m)}([0,1]) \cap H_0^m([0,1])$.
- $(V_j^{(m)}, V_j^{*(m)})$ forms a biorthogonal multiresolution analysis of $L^2([0,1])$.
- $W_j^{(m)} = V_{j+1}^{(m)} \cap (V_j^{*(m)})^\perp$ and $W_j^{*(m)} = V_{j+1}^{*(m)} \cap (V_j^{(m)})^\perp$,

Proposition 1.1 *Let $P_j^{(m)}$ be the projector on $V_j^{(m)}$ parallel to $(V_j^{*(m)})^\perp$ and $P_j^{*(m)}$ its adjoint. We define $Q_j^{(m)} = P_j^{(m+1)} - P_j^{(m)}$, $Q_j^{*(m)} = P_{j+1}^{*(m)} - P_j^{*(m)}$ and let j_0 be an* *integer satisfying $2^{j_0}-1 \geq 2N_2-2N_1-2+2m$. Then we have the following commutation properties*

$$\frac{d}{dx}(P_j^{(m)} f) = P_j^{(m+1)}(\frac{df}{dx}) \quad \textit{if} \quad f \in H^1([0,1]), \tag{1.4.1}$$

and

$$\frac{d}{dx}(P^{*(m+1)} f) = P_j^{*(m)}(\frac{df}{dx}) \quad \textit{if} \quad f \in H_0^1([0,1]). \tag{1.4.2}$$

Proof *To prove this Proposition, it is enough to remark that if $f \in H^1([0,1])$ and $g \in H_0^1([0,1])$, then we have*

$$\langle P_j f, g\rangle_{L^2([0,1])} = \langle f, P_j^* g\rangle$$

and

$$\langle \frac{df}{dx}, g\rangle = -\langle f, \frac{dg}{dx}\rangle.$$

We can now establish the main result of this section.

Theorem 1.5 *Assume that φ is a $C^{p+\varepsilon}$ -function, $p \in N^*$, $p \geq m$, $\varepsilon > 0$ and let j_0 be an integer satisfying $2^{j_0}-1 \geq 2N_2-2N_1-2+2p$. Then we have*

i) for $f \in L^2([0,1])$, $\|f\|_2 \approx \|P_{j_0}^{(m)} f\|_2 + (\sum_{j\geq j_0} \|Q_j^{(m)} f\|_2^2)^{\frac{1}{2}}$.

ii) For $f \in L^2([0,1])$, $\|f\|_2 \approx \|P_{j_0}^{*(m)} f\|_2 + (\sum_{j\geq j_0} \|Q_j^{*(m)} f\|_2^2)^{\frac{1}{2}}$.

iii) For $s \in \mathbb{Z}$ such that $-m \leq s \leq p-m$, we have

- $f \in H^s([0,1]) \Leftrightarrow P_{j_0}^{(m)} f \in L^2([0,1])$ *and* $\sum_{j\geq j_0} 4^{js}\|Q_j^{(m)} f\|_2^2 < +\infty$.
- $f \in H_0^{-s}([0,1]) \Leftrightarrow P_{j_0}^{*(m)} f \in L^2([0,1])$ *and* $\sum_{j\geq j_0} 4^{js}\|Q_j^{*(m)} f\|_2^2 < +\infty$.

Proof *The proof of this Theorem is classical in the wavelet theory. We obtain the direct inequalities in i and ii from the vaguelette Lemma [?] and the inverse inequalities by duality. The equivalences in iii are immediate because if $f \in H^s([0,1])$ then its norm is equal to $||f||_2 + ||f^{(s)}||_2$. We set*

$$f = P_{j_0}^{(m)} f + \sum_{j=j_0}^{\infty} Q_j^{(m)} f,$$

then, we have

$$||f||_2 \approx ||P_{j_0}^{(m)} f||_2 + (\sum_{j=j_0}^{\infty} ||Q_j^{(m)} f||_2^2)^{frac12},$$

$$f^{(s)} = (\frac{d}{dx})^s(P_{j_0}^{(m)} f + \sum_{j=j_0}^{\infty} Q_j^{(m)} f) = P_{j_0}^{(m+s)} f^{(s)} + \sum_{j=j_0}^{\infty} Q_j^{(m+s)} f^{(s)},$$

and

$$||f^{(s)}||_2 \approx ||p_{j_0}^{(m+s)}{}_{j_0} f^{(k)}||_2 + (\sum_{j=j_0}^{\infty} ||Q_j^{(m+s)} f^{(s)}||_2^2)^{\frac{1}{2}}.$$

thus, we obtain

$$||P_{j_0}^{(m+s)} f^{(s)}||_2 = ||(\frac{d}{dx})^s(P_{j_0}^{(m)} f)||_2 \leq C||P_{j_0}^{(m)} f||_2,$$

and

$$\|Q_j^{(m+s)} f^{(s)}\|_2 = \|(\frac{d}{dx})^s (Q_j^{(m)} f)\|_2 \approx 2^{js} ||Q_j^{(m)} f||_2.$$

Then the characterization of $H^s([0,1])$ *is immediate. We characterize in the same way the spaces* $H_0^s([0,1])$.

If we apply the same method described above for the biorthogonal multiresolution analysis $(V_j^{(m),T}([0,1]), V_j^{*(m),T}([0,1]))$, *then corollary 3.1 and classical wavelet theory give the same result for the space* $H^{s,T}([0,1]) = \{f \in H^s([0,1]) \ / \ f^{(p)}|_T = 0, \ 0 \leq p \leq s-1\}$.

We have the equivalent results for Besov spaces.

1.5 Conclusion

In this paper, we have described more general constructions of compact wavelet bases on the interval. More precisely, we have constructed orthogonal and biorthogonal systems on $[0,1]$ which are provided by dyadic translations and dilatations of a finite number of basic functions. By derivation and integration, we obtain new regular multiresolution analyses on the interval $[0,1]$ which satisfy the commutation properties 1.4.1 and 1.4.2. We then deduced that these analyses are well adapted to study Sobolev spaces $H^s([0,1])$ and $H^{s,T}([0,1])$ $(s \in \mathbb{Z})$.

Bibliography

[1] Canuto, C., Tabacco, A. Urban, K., The wavelet element method I, Construction and analysis, *Appl. Comp. Harmon.* 6 (1999), no. 1,pp1-52.

[2] Ciesielski, Z., Figiel, T., Spline bases in classical function spaces on compact C^∞ manifolds, Part I and II, *Studia Mathematical*, T.LXXVI (1983), pp1-58, pp95-136.

[3] Ciesielski, Z., Figiel, T., Spline approximation and Besov spaces on compact manifolds, *Studia Mathematical*, T.LXXV (1982), pp13-36.

[4] Cohen,A., Dahmen, W., De Vore, R., Multiscale decomposition on bounded domains, IGPM, Technical report (2003).

[5] Dahmen, W., Schneider, R.,Composite wavelet basesfor operator equations, *Acta Numer.* 9 (1997), pp155-228.

[6] Dahmen, W., Schneider, R., Wavelets on manifolds, Construction and domain decomposition, *SIAM J. Math Annal.* 31 (2000), pp184-230.

[7] Daubechies, I., Orthonormal bases of wavelets with compact support, *Commm. Pure and Appl. Math*, 42 (1988), pp906-996.

[8] Jouini, A., Constructions de bases d'ondelettes sur les variétés, Thesis, ORSAY, 1993.

[9] Jouini , A., Lemarié-Rieusset, P. G., Analyses multirésolutions biorthogonales sur l'intervalle et applications, *Annales de L'I.H.P, Analyses non linéaires*,Vol.10, N^{o}4 (1993), pp453-476.

[10] Jouini, A., Lemarié-Rieusset, P.G., Ondelettes sur un ouvert borné du plan, Prepublication 46 (1992), ORSAY.

[11] Jouini, A., Lemarié-Rieusset, P.G., Wavelets on the L-shaped domain, Prepublication 178, 2003, Evry.

[12] Kahane, J. P., Lemarié-Rieusset, P.G. Fourier seriers and Wavelets, *Gordon and Breach Publishers*, Vol 3, 1995.

[13] Masson, R. Méthodes d'ondelettes en simulation numérique pour les problèmes elliptiques et point selle, Thesis, Université Paris 6, Jan. 1999.

[14] Meyer, Y. Ondelettes sur l'intervalle, *Revista MathematicaIbero-americana*, Vol 7 (1991), pp115-134.

[15] Meyer, Y.,Wavelets and operators. Analysis at Urbana, Vol.1, edited by E.Berkson, N.T.Peck and J.Uhl, *London Math.Society Lecture Notes*, Series 1, 1987.

[16] Meyer, Y., Jaffard, S. Bases d'ondelettes dans des ouverts de $\mathbb{R}^n$. *J. Math Pures et Appl*, 68, pp95-108, 1989.

[17] Sweldens, W. The lifting scheme : A custom design construction of biorthogonal wavelets, *Appl. Comp. Harm. Anal*, 3 (1996), pp186-200.

Chapitre 2

Wavelet based estimator of D for stationary Gaussian processes

Adaptive wavelet based estimator of the memory parameter for stationary Gaussian processes

Jean-Marc Bardet[a], Hatem Bibi[a,b], Abdellatif Jouini[b]

Published in Bernoulli 14, 3 (2008) pp.691-724

bardet@univ-paris1.fr, hatem.bibi@malix.univ-paris1.fr, Abdellatif.jouini@fst.mu.tn,

[a] Samos-Matisse-CES, Université Paris 1, CNRS UMR8174,90 rue de Tolbiac, 75013 Paris, FRANCE.
[b]Département de Mathématiques, Faculté des Sciences de Tunis,1060 Tunis, TUNISIE.

Abstarct : This work is intended as a contribution to a wavelet-based adaptive estimator of the memory parameter in the classical semi-parametric framework for Gaussian stationary processes. In particular we introduce and develop the choice of a data-driven optimal bandwidth. Moreover, we establish a central limit theorem for the estimator of the memory parameter with the minimax rate of convergence (up to a logarithm factor). The quality of the estimators are attested by simulations.

2.1 Introduction

Let $X = (X_t)_{t\in\mathbb{Z}}$ be a second-order zero-mean stationary process and its covariogram be defined

$$r(t) = \mathbb{E}(X_0 \cdot X_t), \quad \text{for } t \in \mathbb{Z}.$$

Assume the spectral density f of X, with

$$f(\lambda) = \frac{1}{2\pi} \cdot \sum_{k\in\mathbb{Z}} r(k) \cdot e^{-ik},$$

exists and represents a continuous function on $[-\pi, 0)[\cup]0, \pi]$. Consequently, the spectral density of X should satisfy the asymptotic property,

$$f(\lambda) \sim C \cdot \frac{1}{\lambda^D} \quad \text{when } \lambda \to 0,$$

with $D < 1$ called the "memory parameter" and $C > 0$. If $D \in (0, 1)$, the process X is a so-called long-memory process, if not X is called a short memory process (see [8], for more details).

This paper deals with two semi-parametric frameworks which are :

- **Assumption A1 :** X is a zero mean stationary Gaussian process with spectral density satisfying

$$f(\lambda) = |\lambda|^{-D} \cdot f^*(\lambda) \;\; \text{for all} \;\; \lambda \in [-\pi, 0)[\cup]0, \pi],$$

  with $f^*(0) > 0$ and $f^* \in \mathcal{H}(D', C_{D'})$ where $0 < D'$, $0 < C_{D'}$ and

$$\mathcal{H}(D', C_{D'}) = \Big\{ g : [-\pi, \pi] \to \mathbb{R}^+ \text{ such that } |g(\lambda) - g(0)| \le C_{D'} \cdot |\lambda|^{D'} \text{ for all } \lambda \in [-\pi, \pi] \Big\}.$$

- **Assumption A1' :** X is a zero-mean stationary Gaussian process with spectral density satisfying

$$f(\lambda) = |\lambda|^{-D} \cdot f^*(\lambda) \;\; \text{for all} \;\; \lambda \in [-\pi, 0)[\cup]0, \pi],$$

  with $f^*(0) > 0$ and $f^* \in \mathcal{H}'(D', C_{D'})$ where $0 < D'$, $C_{D'} > 0$ and

$$\mathcal{H}'(D', C_{D'}) = \Big\{ g : [-\pi, \pi] \to \mathbb{R}^+ \text{ such that } g(\lambda) = g(0) + C_{D'} \, |\lambda|^{D'} + o\big(|\lambda|^{D'}\big) \text{ when } \lambda \to 0 \Big\}.$$

Remark 2.1 *A great number of earlier works concerning the estimation of the long range parameter in a semi-parametric framework (see for instance [11, 12]) are based on Assumption A1 or equivalent assumption on f. Another expression (see [16, 14] or [13]) is $f(\lambda) = |1-e^{i\lambda}|^{-2d} \cdot f^*(\lambda)$ with f^* a function such that $|f^*(\lambda) - f^*(0)| \leq f^*(0) \cdot \lambda^{\beta}$ and $0 < \beta$). It is obvious that for $\beta \leq 2$ such an assumption corresponds to Assumption A1 with $D' = \beta$. Moreover, following arguments developed in [11, 12] , if $f^* \in \mathcal{H}(D', C_{D'})$ with $D' > 2$ is such that f^* is $s \in \mathbb{N}^*$ times differentiable around $\lambda = 0$ with $f^{*(s)}$ satisfying a Lipschitzian condition of degree $0 < \ell < 1$ around 0, then $D' \leq s + \ell$.So for our purpose, D' is a more pertinent parameter than $s + \ell$ (which is often used in no-parametric literature). Finally, the Assumption A1' is a necessary condition to study the following adaptive estimator of D.*

We have $\mathcal{H}'(D', C_{D'}) \subset \mathcal{H}(D', C_{D'})$. Fractional Gaussian noises (with $D' = 2$) and FARIMA[p,d,q] processes (with also $D' = 2$) represent the first and well known examples of processes satisfying Assumption A1' (and therefore Assumption A1).

Remark 2.2 *In [2] , an adaptive procedure covers a more general class of functions than $\mathcal{H}(D', C_{D'})$, i.e. $\mathcal{H}_{AS}(D', C_{D'})$ defined by :*

$$\mathcal{H}_{AS}(D', C_{D'}) = \left\{ \begin{array}{c} g : [-\pi, \pi] \to \mathbb{R}^+ \text{ such that, as } \lambda \to 0 \\ g(\lambda) = g(0) + \sum_{i=0}^{k} C_i' \lambda^{2i} + C_{D'} |\lambda|^{D'} + o\big(|\lambda|^{D'}\big) \quad \text{with } 2k < D' \leq 2k+2 \end{array} \right\}.$$

Unfortunately, the adaptive wavelet based estimator defined below, as local or global log-periodogram estimators, is unable to be adapted to such a class (and therefore, when $D' > 2$, its convergence rate will be the same than if the spectral density is included in $\mathcal{H}_{AS}(2, C_2)$, at the contrary to Andrew and Sun estimator).

This work is to provide a wavelet-based semi-parametric estimation of the parameter D. This method has been introduced by Flandrin [3] and numerically developed by Abry *et al.* [?] and Veitch *et al.* [20]. Asymptotic results are reported in Bardet *et al.* [6] and more recently in Moulines *et al.* [13]. Taking into account these papers, two points of our work can be highlighted : first, a central limit theorem based on conditions which are weaker than those in Bardet *et al.* [6]. Secondly, we define an auto-driven estimator $\tilde{D}_n$ of D (its definition being different than in Veitch *et al.*, [20]). This results in a central limit theorem followed by $\tilde{D}_n$ and this estimator is proved rate optimal up to a logarithm factor (see below). Below we shall develop this point.

Define the usual Sobolev space $\tilde{W}(\beta, L)$ for $\beta > 0$ and $L > 0$,

$$\tilde{W}(\beta, L) = \left\{ g(\lambda) = \sum_{\ell \in \mathbb{Z}} g_\ell e^{2\pi i \ell \lambda} \in \mathbb{L}^2([0,1]) \,/\, \sum_{\ell \in \mathbb{Z}} (1+|\ell|)^{\beta} |g_\ell| < \infty \;\text{ and }\; \sum_{\ell \in \mathbb{Z}} |g_\ell|^2 \leq L \right\}.$$

Let ψ be a "mother" wavelet satisfying the following assumption :
Assumption $W(\infty)$ **:** $\psi : \mathbb{R} \mapsto \mathbb{R}$ *with* $[0,1]$*-support and such that*

1. ψ *is included in the Sobolev class* $\tilde{W}(\infty, L)$ *with* $L > 0$;
2. $\int_0^1 \psi(t)\, dt = 0$ *and* $\psi(0) = \psi(1) = 0$.

A consequence of the first point of this Assumption is : for all $p>0$, $\sup_{\lambda\in\mathbb{R}}|\widehat{\psi}(\lambda)|(1+|\lambda|)^p<\infty$, where $\widehat{\psi}(u)=\int_0^1\psi(t)\,e^{-iut}dt$ is the Fourier transform of ψ. A useful consequence of the second point is $\widehat{\psi}(u)\sim C\,u$ for $u\to 0$ with $|C|<\infty$ a real number not depending on u.

The function ψ is a smooth compactly supported function (the interval $[0,1]$ is meant for better readability, but the following results can be extended to another interval) with its m first vanishing moments. If $D'\leq 2$ and $0<D<1$ in Assumptions A1, Assumption $W(\infty)$ can be replaced by a weaker assumption :

Assumption $W(5/2)$ **:** $\psi:\ \mathbb{R}\mapsto\mathbb{R}$ *with* $[0,1]$*-support and such that*

1. ψ *is included in the Sobolev class* $\widetilde{W}(5/2,L)$ *with* $L>0$*;*
2. $\displaystyle\int_0^1\psi(t)\,dt=0$ *and* $\psi(0)=\psi(1)=0$.

Remark 2.3 *The choice of a wavelet satisfying Assumption* $W(\infty)$ *is quite restricted because of the required smoothness of* ψ*. For instance, the function* $\psi(t)=(t^2-t+a)\exp(-1/t(1-t))$ *and* $a\simeq 0.23087577$ *satisfies Assumption* $W(\infty)$*. The class of "wavelet" checking Assumption* $W(5/2)$ *is larger. For instance,* ψ *can be a dilated Daubechies "mother" wavelet of order* d *with* $d\geq 6$ *to ensure the smoothness of the function* ψ*.It is also possible to apply the following theory to "essentially" compactly supported "mother" wavelet like the Lemarié-Meyer wavelet. Note that it is not necessary to choose* ψ *being a "mother" wavelet associated to a multi-resolution analysis of* $\mathbb{L}^2(\mathbb{R})$ *as in the recent paper of Moulines* et al. *(2007). The whole theory can be developed without this assumption, in which case the choice of* ψ *is larger.*

If $Y=(Y_t)_{t\in\mathbb{R}}$ is a continuous-time process for $(a,b)\in\mathbb{R}_+^*\times\mathbb{R}$, the "classical" wavelet coefficient $d(a,b)$ of the process Y for the scale a and the shift b is

$$d(a,b)=\frac{1}{\sqrt{a}}\int_{\mathbb{R}}\psi(\frac{t}{a}-b)Y_t\,dt. \tag{2.1.1}$$

However, this formula (2.1.1) of a wavelet coefficient cannot be computed from a time series. The support of ψ being $[0,1]$, let us take the following approximation of formula (2.1.1) and define the wavelet coefficients of $X=(X_t)_{t\in\mathbb{Z}}$ by

$$e(a,b)\ =\ \frac{1}{\sqrt{a}}\sum_{k=1}^{a}\psi(\frac{k}{a})X_{k+ab}, \tag{2.1.2}$$

for $(a,b)\in\mathbb{N}_+^*\times\mathbb{Z}$. Note that this approximation is the same as the wavelet coefficient computed from Mallat algorithm for an orthogonal discrete wavelet basis (for instance with Daubechies mother wavelet).

Remark 2.4 *Here a continuous wavelet transform is considered. The discrete wavelet transform where* $a=2^j$*, in other words numerically very interesting (using Mallat cascade algorithm) is just a particular case. The main point in studing a continuous transform is to offer a larger number of "scales" for computing the data-driven optimal bandwidth (see below).*

Under Assumption A1, for all $b \in \mathbb{Z}$, the asymptotic behavior of the variance of $e(a,b)$ is a power law in scale a (when $a \to \infty$). Indeed, for all $a \in \mathbb{N}^*$, $(e(a,b))_{b\in\mathbb{Z}}$ is a Gaussian stationary process and (see Section more details in 3.2) :

$$\mathbb{E}(e^2(a,0)) \sim K_{(\psi,D)} \cdot a^D \quad \text{when } a \to \infty, \tag{2.1.3}$$

with a constant $K_{(\psi,D)}$ such that,

$$K_{(\psi,\alpha)} = \int_{-\infty}^{\infty} |\widehat{\psi}(u)|^2 \cdot |u|^{-\alpha} du > 0 \quad \text{for all } \alpha < 1, \tag{2.1.4}$$

where $\widehat{\psi}$ is the Fourier transform of ψ (the existence of $K_{(\psi,\alpha)}$ is established in Section 2.5). Note that (2.1.3) is also checked without the Gaussian hypothesis in Assumption A1 (the existence of the second moment order of X is sufficient).
The principle of the wavelet-based estimation of D is linked to this power law a^D. Indeed, let $(X_1,\dots,X_N)$ be a sampled path of X and define $\widehat{T}_N(a)$ a sample variance of $e(a,.)$ obtained from an appropriate choice of shifts b, *i.e.*

$$\widehat{T}_N(a) \quad = \quad \frac{1}{[N/a]} \sum_{k=1}^{[N/a]} e^2(a,k-1). \tag{2.1.5}$$

Then, when $a = a_N \to \infty$ satisfies $\lim_{N\to\infty} a_N \cdot N^{-1/(2D'+1)} = \infty$, a central limit theorem for $\log(\widehat{T}_N(a_N))$ can be proved. More precisely we get

$$\log(\widehat{T}_N(a_N)) = D\log(a_N) + \log(f^*(0)K_{(\psi,D)}) + \sqrt{\frac{a_N}{N}} \cdot \varepsilon_N,$$

with $\varepsilon_N \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2_{(\psi,D)})$ and $\sigma^2_{(\psi,D)} > 0$. As a consequence, using different scales $(r_1 a_N,\dots,r_\ell a_N))$ where $(r_1,\dots,r_\ell) \in (\mathbb{N}^*)^\ell$ with a_N a "large enough" scale, a linear regression of $(\log(\widehat{T}_N(r_i a_N))_i$ by $(\log(r_i a_N))_i$ provides an estimator $\widehat{D}(a_N)$ which satisfies at the same time a central limit theorem with a convergence rate $\sqrt{\frac{N}{a_N}}$.
But the main problem is : how to select a large enough scale a_N considering that the smaller a_N, the faster the convergence rate of $\widehat{D}(a_N)$. An optimal solution would be to chose a_N larger but closer to $N^{1/(2D'+1)}$, but the parameter D' is supposed to be unknown. In Veitch *et al.* [20], an automatic selection procedure is proposed using a chi-squared goodness of fit statistic. This procedure is applied successfully on a large number of numerical examples without any theoretical proofs however. Our present method is close to the latter. Roughly speaking, the "optimal" choice of scale (a_N) is based on the "best" linear regression among all the possible linear regressions of ℓ consecutive points $(a, \log(\widehat{T}_N(a)))$, where ℓ is a fixed integer number. Formally speaking, a contrast is minimized and the chosen scale $\tilde{a}_N$ satisfies :

$$\frac{\log(\tilde{a}_N)}{\log N} \xrightarrow[N\to\infty]{\mathcal{P}} \frac{1}{2D'+1}.$$

Thus, the adaptive estimator $\tilde{D}_N$ of D for this scale $\tilde{a}_N$ is such that :

$$\sqrt{\frac{N}{\tilde{a}_N}}(\tilde{D}_N - D) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2_D),$$

with $\sigma_D^2 > 0$. Consequently, the minimax rate of convergence $N^{D'/(1+2D')}$, up to a logarithm factor, for the estimation of the long memory parameter D in this semi-parametric setting (see Giraitis *et al.*, [11]) is given by $\tilde{D}_N$.

Such a rate of convergence can also be obtained by other adaptive estimators (for more details see below). However, $\tilde{D}_N$ has several "theoretic" advantages : firstly, it can be applied to all $D < -1$ and $D' > 0$ (which are very general conditions covering long and short memory, in fact larger conditions than those usually required for adaptive log-periodogram or local Whittle estimators) with a nearly optimal convergence rate. Secondly, $\tilde{D}_N$ satisfies a central limit theorem and sharp confidence intervals for D can be computed (in such a case, the asymptotic σ_D^2 is replaced by $\sigma^2_{\tilde{D}_N}$, for more details see below). Finally, under additive assumptions on ψ (ψ is supposed to have its first m vanishing moments), $\tilde{D}_N$ can also be applied to a process with a polynomial trend of degree $\leq m-1$.

We then give a several simulations in order to appreciate empirical properties of the adaptive estimator $\tilde{D}_N$. First, using a benchmark composed of 5 different "test" processes satisfying Assumption A1' (see below), the central limit theorem satisfied by $\tilde{D}_N$ is empirically checked. The empirical choice of the parameter ℓ is also studied. Moreover, the robustness of $\tilde{D}_N$ is successfully tested. Finally, the adaptive wavelet-based estimator is compared with several existing adaptive estimators of the memory parameter from generated paths of the 5 different "test" processes (Giraitis-Robinson-Samarov adaptive local log-periodogram, Moulines-Soulier adaptive global log-periodogram, Robinson local Whittle, Abry-Taqqu-Veitch data-driven wavelet based, Bhansali-Giraitis-Kokoszka FAR estimators). The simulations results of $\tilde{D}_N$ are convincing. The convergence rate of $\tilde{D}_N$ is often ranges among the best of the 5 test processes (however the Robinson local Whittle estimator $\widehat{D}_R$ provides more uniformly accurate estimations of D). Three other numerical advantages are offered by the adaptive wavelet-based estimator (and not by $\widehat{D}_R$). Firstly, it is a very low consuming time estimator. Secondly it is a very robust estimator : it is not sensitive to possible polynomial trends and seems to be consistent in non-Gaussian cases. Finally, the graph of the log-log regression of sample variance of wavelet coefficients is meaningful and may lead us to model data with more general processes like locally fractional Gaussian noise (see Bardet [8]).

The central limit theorem for sample variance of wavelet coefficient is subject of section 3.2.Section 2.3 is concerned with the automatic selection of the scale as well as the asymptotic behavior of $\tilde{D}_N$. Finally simulations are given in section 2.4 and proofs in section 2.5.

2.2 A central limit theorem for the sample variance of wavelet coefficients

The following asymptotic behavior of the variance of wavelet coefficients is the basis of all further developments. The first point that explains all that follows is the

Property 2.1 *Under Assumption A1 and Assumption* $W(\infty)$*, for* $a \in \mathbb{N}^*$*,* $(e(a,b))_{b\in\mathbb{Z}}$ *is a zero mean Gaussian stationary process and it exists* $M > 0$ *not depending on* a *such* *that, for all $a \in \mathbb{N}^*$,*

$$\Big|\mathbb{E}(e^2(a,0)) - f^*(0)K_{(\psi,D)} \cdot a^D\Big| \leq M \cdot a^{D-D'}. \tag{2.2.1}$$

Please see Section 2.5 for the proofs. The paper of Moulines *et al.* [13] gives similar results for multi-resolution wavelet analysis. The special case of long memory process can also be studied with weaker Assumption $W(5/2)$,

Property 2.2 *Under Assumption $W(5/2)$ and Assumption A1 with $0 < D < 1$ and $0 < D' \leq 2$, for $a \in \mathbb{N}^*$, $(e(a,b))_{b\in\mathbb{Z}}$ is a zero mean Gaussian stationary process and (2.2.1) holds.*

Two corollaries can be added to both those properties. First, under Assumption A1' a more precise result can be established.

Corollary 2.1 *Under :*
- *Assumption A1' and Assumption $W(\infty)$;*
- **or** *Assumption A1' with $0 < D < 1$, $0 < D' \leq 2$ and Assumption $W(5/2)$;*

then $(e(a,b))_{b\in\mathbb{Z}}$ is a zero mean Gaussian stationary process and

$$\mathbb{E}(e^2(a,0)) = f^*(0)\Big(K_{(\psi,D)} \cdot a^D + C_{D'}K_{(\psi,D-D')} \cdot a^{D-D'}\Big) + o\big(a^{D-D'}\big) \quad \textit{when} \quad a \to \infty. \tag{2.2.2}$$

This corollary is key point for the estimation of an appropriated sequence of scale $a = (a_N)$. Indeed, when $f^* \in \mathcal{H}'(D', C_{D'})$, then $f^* \in \mathcal{H}(D'', C_{D''})$ for all D'' satisfying $0 < D'' \leq D'$. Therefore, Assumption A1' is required for obtaining the optimal choice of a_N, *i.e.* $a_N \simeq N^{1/(2D'+1)}$ (see below for more details). The following corollary generalizes the above Properties 2.1 and 2.2.

Corollary 2.2 *Properties 2.1 and 2.2 are also checked when the Gaussian hypothesis of X is replaced by $\mathbb{E}X_k^2 < \infty$ for all $k \in \mathbb{Z}$.*

Remark 2.5 *In this paper, the Gaussian hypothesis has been taken into account merely to insure the convergence of the sample variance (2.1.5) of wavelet coefficients following a central limit theorem (see below). Such a convergence can also be obtained for more general processes using a different proof of the central limit theorem, for instance for linear processes (see a forthcoming work).*

As mentioned in the introduction, this property allows an estimation of D from a log-log regression, as soon as a consistant estimator of $\mathbb{E}(e^2(a,0))$ is provided from a sample $(X_1, \ldots, X_N)$ of the time series X. Define then the normalized wavelet coefficient such that

$$\tilde{e}(a,b) = \frac{e(a,b)}{\big(f^*(0)K_{(\psi,D)} \cdot a^D\big)^{1/2}} \quad \text{for } a \in \mathbb{N}^* \text{ and } b \in \mathbb{Z}. \tag{2.2.3}$$

From property 2.1, it is obvious that under Assumptions A1 it exists $M' > 0$ satisfying for all $a \in \mathbb{N}^*$,

$$\Big|\mathbb{E}(\tilde{e}^2(a,0)) - 1\Big| \leq M' \cdot \frac{1}{a^{D'}}.$$

To use this formula to estimate D by a log-log regression, an estimator of the variance of $e(a,0)$ should be considered (let us remember that a sample $(X_1,\dots,X_N)$ of is supposed to be known, but parameters $(D,\ D',\ C_{D'})$ are unknown). Consider the sample variance and the normalized sample variance of the wavelet coefficient, for $1 \le a < N$,

$$\widehat{T}_N(a) = \frac{1}{[\frac{N}{a}]}\sum_{k=1}^{[\frac{N}{a}]} e^2(a,k-1) \quad \text{and} \quad \tilde{T}_N(a) = \frac{1}{[\frac{N}{a}]}\sum_{k=1}^{[\frac{N}{a}]} \tilde{e}^2(a,k-1). \tag{2.2.4}$$

The following proposition specifies a central limit theorem satisfied by $\log \tilde{T}_N(a)$, which provides the first step for obtaining the asymptotic properties of the estimator by log-log regression. More generally, the following multidimensional central limit theorem for a vector $(\log \tilde{T}_N(a_i))_i$ can be established.

Proposition 2.1 *Define $\ell \in \mathbb{N}\setminus\{0,1\}$ and $(r_1,\cdots,r_\ell)\in(\mathbb{N}^*)^\ell$. Let $(a_n)_{n\in\mathbb{N}}$ be such that $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N \cdot N^{-1/(1+2D')} \underset{N\to\infty}{\longrightarrow} \infty$. Under Assumption A1 and Assumption $W(\infty)$,*

$$\sqrt{\frac{N}{a_N}}\Big(\log \tilde{T}_N(r_i a_N)\Big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big), \tag{2.2.5}$$

with $\Gamma(r_1,\cdots,r_\ell,\psi,D)=(\gamma_{ij})_{1\le i,j\le \ell}$ the covariance matrix such that

$$\gamma_{ij} = \frac{8(r_i r_j)^{2-D}}{K^2_{(\psi,D)} d_{ij}} \sum_{m=-\infty}^{\infty}\Big(\int_0^\infty \frac{\widehat{\psi}(u r_i)\overline{\widehat{\psi}}(u r_j)}{u^D}\cos(u\, d_{ij} m)\, du\Big)^2. \tag{2.2.6}$$

The same result under weaker assumptions on ψ can be also established when X is a long memory process.

Proposition 2.2 *Define $\ell \in \mathbb{N}\setminus\{0,1\}$ and $(r_1,\cdots,r_\ell)\in(\mathbb{N}^*)^\ell$. Let $(a_n)_{n\in\mathbb{N}}$ be such that $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N \cdot N^{-1/(1+2D')} \underset{N\to\infty}{\longrightarrow} \infty$. Under Assumption $W(5/2)$ and Assumption A1 with $D\in(0,1)$ and $D'\in(0,2)$, the CLT (3.2.7) holds.*

These results can be easily generalized for processes with polynomial trends if ψ is considered having its first m vanishing moments. i.e,

Corollary 2.3 *Given the same hypothesis as in Proposition 3.1 or 2.2 and if ψ is such that $m\in\mathbb{N}\setminus\{0,1\}$ is satisfying, $\int t^p\psi(t)\,dt = 0$ for all $p\in\{0,1,\dots,m-1\}$ the CLT (3.2.7) also holds for any process $X'=(X'_t)_{t\in\mathbb{Z}}$ such that for all $t\in Z$, $\mathbb{E}X'_t = P_m(t)$ with $P_m(t)=a_0+a_1t+\cdots+a_{m-1}t^{m-1}$ is a polynomial function and $(a_i)_{0\le i\le m-1}$ are real numbers.*

2.3 Adaptive estimator of memory parameter using data driven optimal scales

The CLT (3.2.7) implies the following CLT for the vector $(\log \widehat{T}_N(r_i a_N))_i$,

$$\sqrt{\frac{N}{a_N}}\Big(\log \widehat{T}_N(r_i a_N) - D\log(r_i a_N) - \log(f^*(0)K_{(\psi,D)})\Big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big).$$

and therefore,

$$\big(\log \widehat{T}_N(r_i a_N)\big)_{1\le i\le \ell} = A_N \cdot \begin{pmatrix} D \\ K \end{pmatrix} + \frac{1}{\sqrt{N/a_N}}\,\big(\varepsilon_i\big)_{1\le i\le \ell},$$

with $A_N = \begin{pmatrix} \log(r_1 a_N) & 1 \\ \vdots & \vdots \\ \log(r_\ell a_N) & 1 \end{pmatrix}$, $K = -\log(f^*(0)\cdot K_{(\psi,D)})$

and

$$\big(\varepsilon_i\big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big).$$

Therefore, a log-log regression of $\big(\widehat{T}_N(r_i a_N)\big)_{1\le i\le \ell}$ on scales $\big(r_i a_N\big)_{1\le i\le \ell}$ provides an estimator $\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix}$ of $\begin{pmatrix} D \\ K \end{pmatrix}$ such that

$$\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix} = (A_N'\cdot A_N)^{-1}\cdot A_N'\cdot Y_{a_N}^{(r_1,\ldots,r_\ell)} \quad \text{with} \quad Y_{a_N}^{(r_1,\ldots,r_\ell)} = \big(\log \widehat{T}_N(r_i a_N)\big)_{1\le i\le \ell}, \tag{2.3.1}$$

which satisfies the following CLT,

Proposition 2.3 *Under the Assumptions of the Proposition 3.1,*

$$\sqrt{\frac{N}{a_N}}\Big(\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix} - \begin{pmatrix} D \\ K \end{pmatrix}\Big) \xrightarrow[N\to\infty]{D} \mathcal{N}_2(0\,;\,(A'\cdot A)^{-1}\cdot A'\cdot\Gamma(r_1,\cdots,r_\ell,\psi,D)\cdot A\cdot(A'\cdot A)^{-1}), \tag{2.3.2}$$

with $A = \begin{pmatrix} \log(r_1) & 1 \\ \vdots & \vdots \\ \log(r_\ell) & 1 \end{pmatrix}$ *and* $\Gamma(r_1,\cdots,r_\ell,\psi,D)$ *given by (3.2.8).*

Moreover, under Assumption A1' and if $D\in(-1,1)$, $\widehat{D}(a_N)$ is a semi-parametric estimator of D and its asymptotic mean square error can be minimized with an appropriate scales sequence (a_N) reaching the well-known minimax rate of convergence for memory parameter D in this semi-parametric setting (see for instance Giraitis *et al.*, [11, 12]). Indeed,

Proposition 2.4 *Let X satisfy Assumption A1' with $D\in(-1,1)$ and ψ the assumption $W(\infty)$. Let (a_N) be a sequence such that $a_N = [N^{1/(1+2D')}]$. Then, the estimator $\widehat{D}(a_N)$ is rate optimal in the minimax sense,* i.e.

$$\limsup_{N\to\infty}\ \sup_{D\in(-1,1)}\ \sup_{f^*\in\mathcal{H}(D',C_{D'})} N^{\frac{2D'}{1+2D'}}\cdot\mathbb{E}[\widehat{D}(a_N)-D)^2] < +\infty.$$

Remark 2.6 *As far as we know, there are no theoretic results of optimality in case of $D\le -1$, but according to the usual following non-parametric theory, such minimax results can also be obtained. Moreover, in case of long-memory processes (if $D\in(0,1)$), under Assumption A1' for X and Assumption $W(5/2)$ for ψ, the estimator $\widehat{D}(a_N)$ is also rate optimal in the minimax sense.*

In the previous Propositions 3.1 and 2.3, the rate of convergence of scale a_N obeys to the following condition,

$$\frac{N}{a_N} \underset{N\to\infty}{\longrightarrow} \infty \quad \text{and} \quad \frac{a_N}{N^{1/(1+2D')}} \underset{N\to\infty}{\longrightarrow} \infty \quad \text{with} \quad D' \in (0,\infty).$$

Now, for better readability, take $a_N = N^\alpha$. Then, the above condition goes as follow :

$$a_N = N^\alpha \quad \text{with} \quad \alpha^* < \alpha < 1 \quad \text{and} \quad \alpha^* = \frac{1}{1+2D'}. \tag{2.3.3}$$

Thus an optimal choice (leading to a faster convergence rate of the estimator) is obtained for $\alpha = \alpha^* + \varepsilon$ with $\varepsilon \to 0+$. But α^* depends on D' which is unknown. To solve this problem, Veitch *et al.* [20] suggest a chi-square-based test (constructed from a distance between the regression line and the different points $(\log \widehat{T}_N(r_i a_N), \log(r_i a_N))$. It seems to be an efficient and interesting numerical way to estimate D, but without theoretical proofs (contrary to global or local log-periodogram procedures which are proved to reach the minimax convergence rate, see for instance Moulines and Soulier, [14]).

We suggest a new procedure for the data-driven selection of optimal scales, *i.e.* optimal α. Let us consider an important parameter, the number of considered scales $\ell \in \mathbb{N} \setminus \{0,1,2\}$ and set $(r_1,\ldots,r_\ell) = (1,\ldots,\ell)$. For $\alpha \in (0,1)$, define also

- the vector $Y_N(\alpha) = \big(\log \widehat{T}_N(i\cdot N^\alpha)\big)_{1\le i\le \ell}$;
- the matrix $A_N(\alpha) = \begin{pmatrix} \log(N^\alpha) & 1 \\ \vdots & \vdots \\ \log(\ell\cdot N^\alpha) & 1 \end{pmatrix}$;
- the contrast, $Q_N(\alpha,D,K) = \Big(Y_N(\alpha) - A_N(\alpha)\cdot\begin{pmatrix} D \\ K\end{pmatrix}\Big)'\cdot\Big(Y_N(\alpha) - A_N(\alpha)\cdot\begin{pmatrix} D \\ K\end{pmatrix}\Big)$.

$Q_N(\alpha,D,K)$ corresponds to a squared distance between the ℓ points $\big(\log(i\cdot N^\alpha)\,,\,\log T_N(i\cdot N^\alpha)\big)_i$ and a line. The point is to minimize this contrast for these three parameters. It is obvious that for a fixed $\alpha \in (0,1)$ Q is minimized from the previous least square regression and therefore,

$$Q_N(\widehat{\alpha}_N, \widehat{D}(a_N), \widehat{K}(a_N)) = \min_{\alpha\in(0,1), D<1, K\in\mathbb{R}} Q_N(\alpha,D,K).$$

with $(\widehat{D}(a_N), \widehat{K}(a_N))$ obtained as in relation (2.3.1). However, since $\widehat{\alpha}_N$ has to be obtained from numerical computations, the interval $(0,1)$ can be discretized as follows,

$$\widehat{\alpha}_N \in \mathcal{A}_N = \Big\{\frac{2}{\log N}, \frac{3}{\log N}, \ldots, \frac{\log[N/\ell]}{\log N}\Big\}.$$

Hence, if $\alpha \in \mathcal{A}_N$, it exists $k \in \{2,3,\ldots,\log[N/\ell]\}$ such that $k = \alpha\cdot\log N$.

Remark 2.7 *This choice of discretization is implied by the following proof of the consistency of $\widehat{\alpha}_N$. If the interval $(0,1)$ is stepped in N^β points, with $\beta > 0$, the used proof cannot attest this consistency. Finally, it is the same framework as the usual discrete wavelet transform (see for instance Veitch et al., [20]) but less restricted since $\log N$ may be replaced in the previous expression of $\mathcal{A}_N$ by any negligible function of N compared to functions N^β with $\beta > 0$ (for instance, $(\log N)^d$ or $d\log N$ can be used).*

Consequently, take

$$\widehat{Q}_N(\alpha) = Q_N(\alpha, \widehat{D}(a_N), \widehat{K}(a_N));$$

then, minimize Q_N for variables (α, D, K) is equivalent to minimize $\widehat{Q}_N$ for variable $\alpha \in \mathcal{A}_N$, that is

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

From this central limit theorem derives

Proposition 2.5 *Let X satisfy Assumption A1' and ψ Assumption $W(\infty)$ (or Assumption $W(5/2)$ if $0 < D < 1$ and $0 < D' \leq 2$). Then,*

$$\widehat{\alpha}_N = \frac{\log \widehat{a}_N}{\log N} \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^* = \frac{1}{1+2D'}.$$

This proves also the consistency of an estimator $\widehat{D'}_N$ of the parameter D',

Corollary 2.4 *Taking the hypothesis of Proposition 2.5, we have*

$$\widehat{D'}_N = \frac{1-\widehat{\alpha}_N}{2\widehat{\alpha}_N} \xrightarrow[N\to\infty]{\mathcal{P}} D'.$$

The estimator $\widehat{\alpha}_N$ defines the selected scale $\widehat{a}_N$ such that $\widehat{a}_N = N^{\widehat{\alpha}_N}$. From a straightforward application of the proof of Proposition 2.5 (see the details in the proof of Theorem 3.1), the asymptotic behavior of $\widehat{a}_N$ can be specified, that is,

$$\Pr\Big(\frac{N^{\alpha^*}}{(\log N)^{\lambda}} \leq N^{\widehat{\alpha}_N} \leq N^{\alpha^*} \cdot (\log N)^{\mu}\Big) \xrightarrow[N\to\infty]{\mathcal{P}} 1, \qquad (2.3.4)$$

for all positive real numbers λ and μ such that $\lambda > \frac{2}{(\ell-2)D'}$ and $\mu > \frac{12}{\ell-2}$. Consequently, the selected scale is asymptotically equal to N^{α^*} up to a logarithm factor.

Finally, Proposition 2.5 can be used to define an adaptive estimator of D. First, define the straightforward estimator

$$\widehat{\widehat{D}}_N = \widehat{D}(\widehat{a}_N),$$

which should minimize the mean square error using $\widehat{a}_N$. However, the estimator $\widehat{\widehat{D}}_N$ does not attest a CLT since $\Pr(\widehat{\alpha}_N \leq \alpha^*) > 0$ and therefore it can not be asserted that $\mathbb{E}(\sqrt{N/\widehat{a}_N}(\widehat{\widehat{D}}_N - D)) = 0$. To establish a CLT satisfied by an adaptive estimator $\tilde{D}_N$ of D, an adaptive scale sequence $(\tilde{a}_N) = (N^{\tilde{\alpha}_N})$ has to be defined to ensure $\Pr(\tilde{\alpha}_N \leq \alpha^*) \xrightarrow[N\to\infty]{} 0$. The following theorem provides the asymptotic behavior of such an estimator,

Theorem 2.1 *Let X satisfy Assumption A1' and ψ Assumption $W(\infty)$ (or Assumption $W(5/2)$ if $0 < D < 1$ and $0 < D' \leq 2$). Define,*

$$\tilde{\alpha}_N = \widehat{\alpha}_N + \frac{3}{(\ell-2)\widehat{D'}_N} \cdot \frac{\log\log N}{\log N}, \quad \tilde{a}_N = N^{\tilde{\alpha}_N} = N^{\widehat{\alpha}_N} \cdot \big(\log N\big)^{\frac{3}{(\ell-2)\widehat{D'}_N}} \quad \textit{and} \quad \tilde{D}_N = \widehat{D}(\tilde{a}_N).$$

Then, with $\sigma_D^2 = (1\ 0) \cdot (A' \cdot A)^{-1} \cdot A' \cdot \Gamma(1, \cdots, \ell, \psi, D) \cdot A \cdot (A' \cdot A)^{-1} \cdot (1\ 0)'$,

$$\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}} \big(\tilde{D}_N - D \big) \underset{N \to \infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}(0\,;\, \sigma_D^2) \qquad (2.3.5)$$

$$\text{and} \quad \forall \rho > \frac{2(1+3D')}{(\ell-2)D'}, \quad \frac{N^{\frac{D'}{1+2D'}}}{(\log N)^\rho} \cdot \big| \tilde{D}_N - D \big| \underset{N \to \infty}{\overset{\mathcal{P}}{\longrightarrow}} 0. \qquad (2.3.6)$$

Remark 2.8 *Both the adaptive estimators* $\widehat{\widehat{D}}_N$ *and* $\tilde{D}_N$ *converge to* D *with a rate of convergence rate equal to the minimax rate of convergence* $N^{\frac{D'}{1+2D'}}$ *up to a logarithm factor (this result being classical within this semi-parametric framework). Unfortunately, our method cannot prove that the mean square error of both these estimators reaches the optimal rate and therefore to be oracles.*

To conclude this theoretic approach, the main properties satisfied by the estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can be summarized as follows :

1. Both the adaptive estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ converge at D with a rate of convergence rate equal to the minimax rate of convergence $N^{\frac{D'}{1+2D'}}$ up to a logarithm factor for all $D < -1$ and $D' > 0$ (this being very general conditions covering long and short memory, even larger than usual conditions required for adaptive log-periodogram or local Whittle estimators) whith X considered a Gaussian process.
2. The estimator $\tilde{D}_N$ satisfies the CLT (3.3.4) and therefore sharp confidence intervals for D can be computed (in which case, the asymptotic matrix $\Gamma(1, \ldots, \ell, \psi, D)$ is replaced by $\Gamma(1, \ldots, \ell, \psi, \widehat{D}_N)$). This is not applicable to an adaptive log-periodogram or local Whittle estimators.
3. The main Property 2.1 is also satisfied without the Gaussian hypothesis. Therefore, adaptive estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can also be interesting estimators of D for non-Gaussian processes like linear or more general processes (but a CLT similar to Theorem 3.1 has to be established...).
4. Under additive assumptions on ψ (ψ is supposed to have its first m vanishing moments), both estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can also be used for a process X with a polynomial trend of degree $\leq m-1$, which again cannot be yielded with an adaptive log-periodogram or local Whittle estimators.

2.4 Simulations

The adaptive wavelet basis estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ are new estimators of the memory parameter D in the semi-parametric frame. Different estimators of this kind are also reported in other research works to have proved optimal. In this paper, some theoretic advantages of adaptive wavelet basis estimators have been highlighted. But what about concrete procedure and results of such estimators applied to an observed sample ? The following simulations will help to answer this question.

First, the properties (consistency, robustness, choice of the parameter ℓ and mother wavelet function ψ) of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ are investigated. Secondly, in cases of Gaussian long-memory processes (with $D \in (0,1)$ and $D' \leq 2$), the simulation results of the estimator

$\widehat{\widehat{D}}_N$ are compared to those obtained with the best known semi-parametric long-memory estimators.

To begin with, the simulations 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.* [8]. The simulations are realized for different values of D, N and processes which satisfy Assumption A1' and therefore Assumption A1 (the article of Moulines *et al.*[13], gives a lot of details on this point) :

1. the fractional Gaussian noise (fGn) of parameter $H = (D+1)/2$ (for $-1 < D < 1$) and $\sigma^2 = 1$. The spectral density f_{fGn} of a fGn is such that f^*_{fGn} is included in $\mathcal{H}(2, C_2)$ (thus $D' = 2$) ;
2. the FARIMA[p,d,q] process with parameter d such that $d = D/2 \in (-0.5, 0.5)$ (therefore $-1 < D < 1$), the innovation variance σ^2 satisfying $\sigma^2 = 1$ and $p, q \in \mathbb{N}$. The spectral density f_{FARIMA} of such a process is such that f^*_{FARIMA} is included in the set $\mathcal{H}(2, C_2)$ (thus $D' = 2$) ;
3. the Gaussian stationary process $X^{(D,D')}$, such that its spectral density is
$$f_3(\lambda) = \frac{1}{\lambda^D}(1 + \lambda^{D'}) \quad \text{for } \lambda \in [-\pi, \pi], \tag{2.4.1}$$
with $D \in (-\infty, 1)$ and $D' \in (0, \infty)$. Therefore $f_3^* = 1 + \lambda^{D'} \in \mathcal{H}(D', 1)$ with $D' \in (0, \infty)$.

In the long memory frame, a "benchmark" of processes is considered for $D = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$:

- fGn processes with parameters $H = (D+1)/2$ and $\sigma^2 = 1$;
- FARIMA[0,d,0] processes with $d = D/2$ and standard Gaussian innovations ;
- FARIMA[1,d,0] processes with $d = D/2$, standard Gaussian innovations and AR coefficient $\phi = 0.95$;
- FARIMA[1,d,1] processes with $d = D/2$, standard Gaussian innovations and AR coefficient $\phi = -0.3$ and MA coefficient $\phi = 0.7$;
- $X^{(D,D')}$ Gaussian processes with $D' = 1$.

2.4.1 Properties of adaptive wavelet basis estimators from simulations

Below, we give the different properties of the adaptive wavelet based method.

Choice of the mother wavelet ψ : For short memory processes ($D \leq 0$), let the wavelet ψ_{SM} be such that $\psi_{SM}(t) = (t^2 - t + a)\exp(-1/t(1-t))$ with $a \simeq 0.23087577$. It satisfies Assumption $W(\infty)$. Lemarié-Meyer wavelets can be also investigated but this will lead to quite different theoretic studies since its support is not bounded (but "essentially" compact).
For long memory processes ($0 < D < 1$), let the mother wavelet ψ_{LM} be such that $\psi_{LM}(t) = 100 \cdot t^2(t-1)^2(t^2 - t + 3/14)\mathbb{I}_{0 \leq t \leq 1}$ which satisfies Assumption $W(5/2)$. Note that Daubechies mother wavelet or ψ_{SM} lead to "similar" results (but not as good).

Choice of the parameter ℓ : This parameter is very important to estimate the "beginning" of the linear part of the graph drawn by points $(\log(a_i), \log \widehat{T}(a_i))_i$. On the one hand, if ℓ is a too small a number (for instance $\ell = 3$), another small linear part of this graph (even before the "true" beginning N^{α^*}) may be chosen ; consequently, the $\sqrt{MSE}$ (square root of the mean square error) of $\widehat{\alpha}_N$ and therefore of $\widehat{\widehat{D}}_N$ or $\tilde{D}_N$ will be too large. On the other hand, if ℓ is a too large a number (for instance $\ell = 50$ for $N = 1000$), the estimator $\widehat{\alpha}_N$ will certainly satisfy $\widehat{\alpha}_N < \alpha^*$ since it will not be possible to consider ℓ different scales larger than N^{α^*} (if $D' = 1$ therefore $\alpha' = 1/3$, then a_N has to satisfy : $N/(50a_N) = 20/a_N$ is a large number and ($a_N > N^{1/3} = 10$; this is not really possible). Moreover, it is possible that a "good" choice of ℓ depends on the "flatness" of the spectral density f, *i.e.* on D'. We have proceeded to simulations for each different values of ℓ (and N and D). Only $\sqrt{MSE}$ of estimators are presented. The results are specified in Table 1.

In Table 1, two phenomena can be distinguished : the detection of α^* and the estimation of D :

- To estimate α^*, ℓ has to be small enough, especially because of "D' close to 0" and so "α' close to 1" is possible. However, our simulations indicate that ℓ must not be too small (for instance $\ell = 5$ leads to an important MSE for $\widehat{\alpha}_N$ implying an important MSE for $\widehat{\widehat{D}}_N$) and seems to be independent of N (cases $N = 1000$ and $N = 10000$ are quite similar). Hence, our choice is $\ell_1 = 15$ **to estimate** α^* **for any** N.
- To estimate D, once α^* is estimated, a second value ℓ_2 of ℓ can be chosen. We use an adaptive procedure which, roughly speaking, consists in determining the “end” of the acceptable linear zone. Firstly, we use again the same procedure than for estimating $\widehat{a}_N$ but with scales $(a_N/i)_{1 \leq i \leq \ell_1}$ and $\ell_1 = 15$. It provides an estimator $\widehat{b}_N$ corresponding to the maximum of acceptable (for a linear regression) scales. Secondly, **the adaptive number of scales** ℓ_2 **is computed from the formula** $\ell_2 = \widehat{\ell} = [\widehat{b}_N/\widehat{a}_N]$. The simulations carried out with such values of ℓ_1 and ℓ_2 are detailed in Table 1.

As it may be seen in Table 1, the choice of parameters $(\ell_1 = 15, \ell_2 = \widehat{\ell})$ provides the best results for estimating D, almost uniformly for all processes.

Consistency of the estimators $\widehat{\alpha}_N$ and $\tilde{\alpha}_N$: the previous numerical results (here we consider $\ell_1 = 15$) show that $\widehat{\alpha}_N$ and $\tilde{\alpha}_N$ converge (very slowly) to the optimal rate α^*, that is 0.2 for the first four processes and 1/3 for the fifth. Figure 1 illustrates the evolution with N of the log-log plotting and the choice of the onset of scaling.
Figure 1 shows that $\log T_N(i \cdot N^\alpha)$ is not a linear function of the logarithm of the scales $\log(i \cdot N^\alpha)$ when N increases and $\alpha < \alpha^*$ (a consequence of Property 2.1 : it means there is a bias). Moreover, if $\alpha > \alpha^*$ and α increases, a linear model appears with an increasing error variance.

Consistency and distribution of the estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$: The results of Table 1 show the consistency with N of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ only by using $\ell_1 = 15$. Figure 2 provides the histograms of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ for 100 independent samples of FARIMA$(1, d, 1)$ processes with $D = 0.5$ and $N = 10^5$. Both the histograms of Figure 2 are similar to

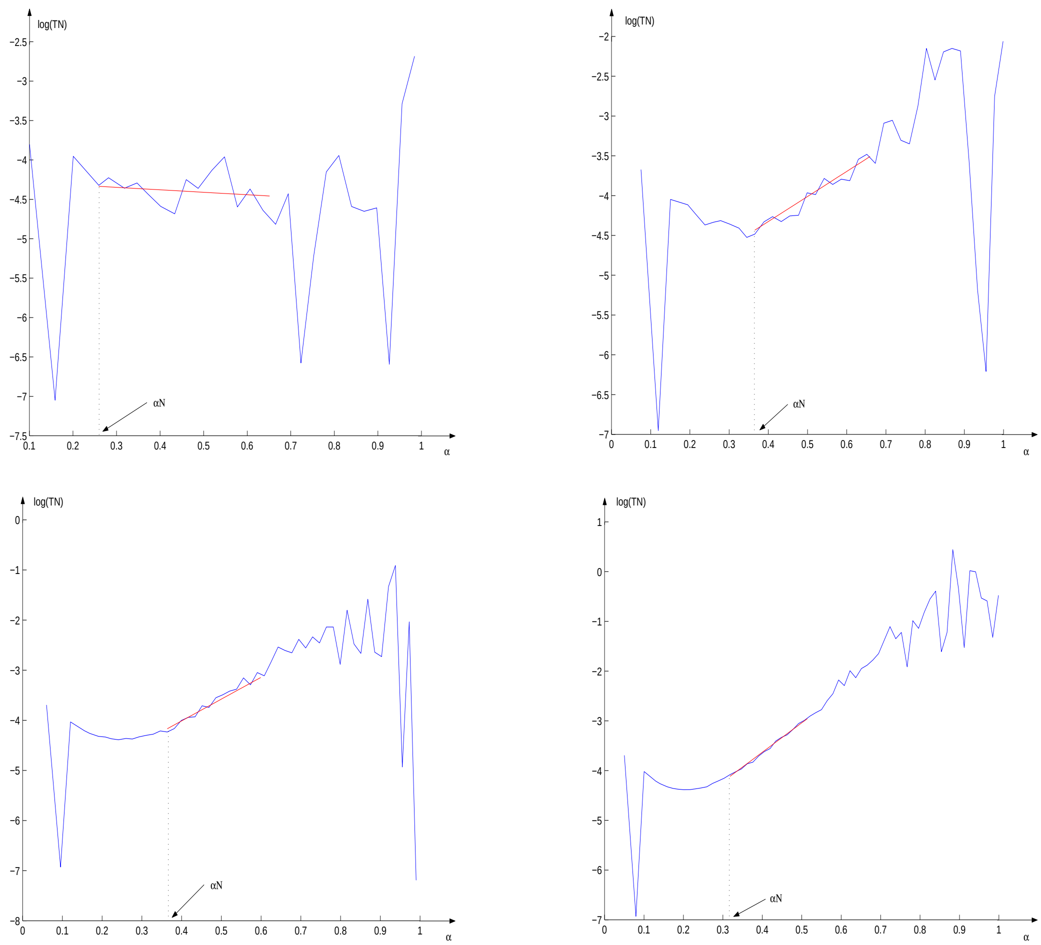

FIGURE 2.1 – *Log-log graphs for different samples of* $X^{(D,D')}$ *with* $D = 0.5$ *and* $D' = 1$ *when* $N = 10^3$ *(up and left,* $\widehat{\widehat{D}}_N \simeq 1.04$*),* $N = 10^4$ *(up and right,* $\widehat{\widehat{D}}_N \simeq 0.66$*),* $N = 10^5$ *(down and left,* $\widehat{\widehat{D}}_N \simeq 0.62$*) and* $N = 10^6$ *(down and right,* $\widehat{\widehat{D}}_N \simeq 0.54$*).*

Gaussian distribution histograms. It is not surprising for $\tilde{D}_N$ since Theorem 3.1 shows that the asymptotic distribution of $\tilde{D}_N$ is a Gaussian distribution with mean equal to D. The asymptotic distribution of $\widehat{\widehat{D}}_N$ and the Gaussian distribution seem also to be similar. A Cramer-von Mises test of normality indicates that both distributions of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can be considered a Gaussian distribution (respectively $W \simeq 0.07$, $p - value \simeq 0.24$ and $W \simeq 0.05$, $p - value \simeq 0.54$).

Consistency in case of short memory : The following Table 2 provides the behavior of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ if $D \leq 0$ and $D' > 0$. Two processes are considered in such a frame : a FARIMA$(0, d, 0)$ process with $-0.5 < d < 0$ and therefore $-1 < D \leq 0$ (always with $D' = 2$) and a process $X^{(D,D')}$ and $D < 0$ and $D' > 0$. The results are displayed in Table 2.4.1 (here $N = 1000$, $N = 10000$ and $N = 100000$, $\ell_1 = 15$ and $\ell_2 = [5\,N^{0.1}]$)

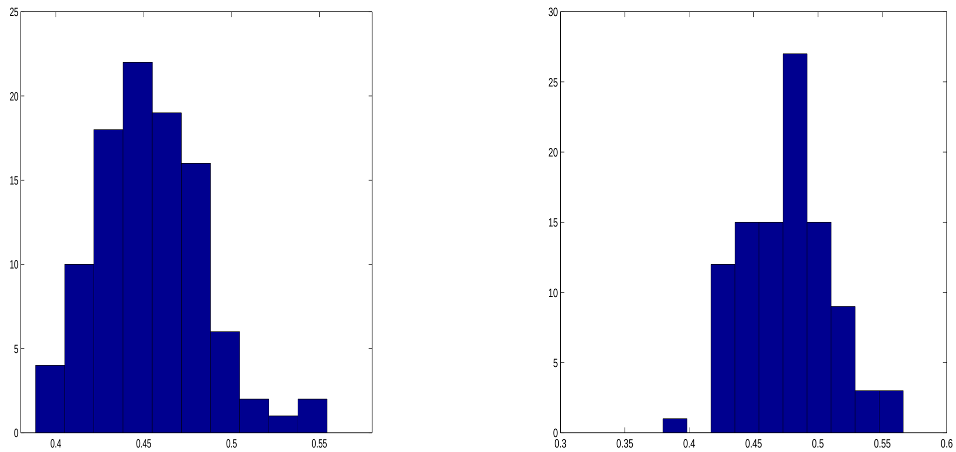

FIGURE 2.2 – *Histograms of* $\widehat{\widehat{D}}_N$ *and* $\tilde{D}_N$ *for* 100 *samples of FARIMA*$(1, d, 1)$ *with* $D = 0.5$ *for* $N = 10^5$.

for different choices of D and D'. Thus it appears that $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can be successively applied to short memory processes as well. Moreover, the larger D', the faster their convergence rates.

Robustness of $\widehat{\widehat{D}}_N$**,** $\tilde{D}_N$ **:** To conclude with the numerical properties of the estimators, four different processes not satisfying Assumption $A1'$ are considered :

- a FARIMA$(0, d, 0)$ process (denoted $P1$) with innovations satisfying a uniform law (and $\mathbb{E}X_i^2 < \infty$) ;
- a FARIMA$(0, d, 0)$ process (denoted $P2$) with innovations satisfying a distribution with density w.r.t. Lebesgue measure $f(x) = 3/4 * (1 + |x|)^{-5/2}$ for $x \in \mathbb{R}$ (and therefore $\mathbb{E}|X_i|^2 = \infty$ but $\mathbb{E}|X_i| < \infty$) ;
- a FARIMA$(0, d, 0)$ process (denoted $P3$) with innovations satisfying a Cauchy distribution (and $\mathbb{E}|X_i| = \infty$) ;
- a Gaussian stationary process (denoted $P4$) with a spectral density $f(\lambda) = (|\lambda| - \pi/2)^{-1/2}$ for all $\lambda \in [-\pi, \pi] \backslash \{-\pi/2, \pi/2\}$. The local behavior of f in 0 is $f(|\lambda|) \sim \sqrt{\pi/2}\, |\lambda|^D$ with $D = 0$, but the smoothness condition for f in Assumption A1 is not satisfied.

For the first 3 processes, D is varies in $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ and 100 independent replications are taken into account. The results of these simulations are given in Table 3.

As outlined in the theoretical part of this paper, the estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ seem also to be accurate for $\mathbb{L}^2$-linear processes. For $\mathbb{L}^\alpha$-linear processes with $1 \leq \alpha < 2$, they are also convergent with a slower rate of convergence. Despite the spectral density of process $P4$ does not satisfies the smoothness hypothesis requires in Assumptions A1 or A1', the convergence rates of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ are still convincing. These results confirm the robustness of wavelet based estimators.

2.4.2 Comparisons from simulations

Comparisons with other semi-parametric long-memory parameter estimators from simulations Here we consider only long-memory Gaussian processes ($D \in (0,1)$) based on the usual hypothesis $0 < D' \leq 2$. More precisely, the "benchmark" is : 100 generated independent samples of each process with length $N = 10^3$ and $N = 10^4$ and different values of D, $D = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Several different semi-parametric estimators of D are considered :

- $\widehat{D}_{BGK}$ is an "optimal" parametric Whittle estimator obtained from a BIC criterium model selection of fractionally differenced autoregressive models (introduced by Bhansali it et al., [10]). The required confidence interval of the estimation $\widehat{D}_{BGK}$ is $[\widehat{D}_R - 2/N^{1/4}\,,\, \widehat{D}_R - 2/N^{1/4}]$;
- $\widehat{D}_{GRS}$ is an adaptive local periodogram estimator introduced by Giraitis *et al* [12]. It requires two parameters : a bandwidth parameter m, with a procedure of determination provided in this article, and a number of low trimmed frequencies l (satisfying different conditions but without being fixed in this paper; after a number of simulations, $l = \max(m^{1/3}, 10)$ is chosen);
- $\widehat{D}_{MS}$ is an adaptive global periodogram estimator introduced by Moulines and Soulier [14], also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;
- $\widehat{D}_R$ is a local Whittle estimator introduced by Robinson [43]. The trimming parameter is $m = N/30$;
- $\widehat{D}_{ATV}$ is an adaptive wavelet based estimator introduced by Veitch *et al.* [20] using a Db4 wavelet (and described above);
- $\widehat{\widehat{D}}_N$ defined previously with $\ell_1 = 15$ and $\ell_2 = N^{1-\widehat{\alpha}_N}/10$ and a mother wavelet $\psi(t) = 100 \cdot t^2(t-1)^2(t^2 - t + 3/14)\mathbb{I}_{0 \leq t \leq 1}$ satisfying assumption $W(5/2)$.

Softwares (using Matlab language) for computing some of these estimators are available on Internet (see the website of D. Veitch [1] for $\widehat{D}_{ATV}$ and the homepage of E. Moulines [2] for $\widehat{D}_{MS}$ and $\widehat{D}_R$). The other softwares are available on [3]. Simulation results are reported in Table 4.

Comments on the results of Table 4 : These simulations allow to distinguish four "clusters" of estimators.

- $\widehat{D}_{BGK}$ is obtained from a BIC-criterium hierarchical model selection (from 2 to 11 parameters, corresponding to the length of the approximation of the Fourier expansion of the spectral density) using Whittle estimation. For these simulations, the BIC criterion is generally minimal for 5 to 7 parameters to be estimated. Simulation results are not very satisfactory except for $D = 0.1$ (close to the short memory). Moreover, this procedure is rather time-consuming.
- $\widehat{D}_{GRS}$ offers good results for fGn and FARIMA$(0, d, 0)$. However, this estimator does not converge fast enough for the other processes.
- Estimators $\widehat{D}_{MS}$ and $\widehat{D}_R$ have similar properties. They (especially $\widehat{D}_R$) are very interesting because they offer the same fairly good rates of convergence for all processes of the benchmark.

1. `http://wwww.cubinlab.ee.mu.oz.au/~darryl/`
2. `http://www.tsi.enst.fr/~moulines/`
3. `http://samos.univ-paris1.fr/spip/-Jean-Marc-Bardet`

- Being built on similar principles, estimators $\widehat{D}_{ATV}$ and $\widehat{\widehat{D}}_N$ have similar behavior as well. Their convergence rates are the fastest for fGn and FARIMA$(0,d,0)$ and are almost close to fast ones for the other processes. Their times of computing, especially for $\widehat{D}_{ATV}$ for which the computations of wavelet coefficients with that the Mallat algorithm, are the shortest.

Conclusion : Which estimator among those studied above has to be chosen in a practical frame, *i.e.* an observed time series ? We propose the following procedure for estimating an eventual long memory parameter :

1. Firstly, since this procedure is very low time consuming and applicable to processes with smooth trends, draw the log-log regression of wavelet coefficients' variances onto scales. If a linear zone appears in this graph, consider the estimator $\widehat{\widehat{D}}_N$ (or $\widehat{D}_{ATV}$) of D.
2. If a linear zone appears in the previous graph and if the observed time series seems to be without a trend, compute $\widehat{D}_R$.
3. Compare both the estimated value of D from confidence intervals (available for $\widehat{\widehat{D}}_N$ or $\widehat{D}_{ATV}$ and $\widehat{D}_R$).

2.5 Proofs

Proof [Property 2.1] The arguments of this proof are similar to those of Abry *et al.* [**?**] or Moulines *et al.* [13]. First, for $a \in \mathbb{N}^*$,

$$\begin{aligned}
&\mathbb{E}(e^2(a,0)) \\
&= \frac{1}{a}\sum_{k=1}^{a}\sum_{k'=1}^{a}\psi(k/a)\psi(k'/a)\mathbb{E}(X_k X_{k'}) \qquad (2.5.1)\\
&= \frac{1}{a}\sum_{k=1}^{a}\sum_{k'=1}^{a}\psi(k/a)\psi(k'/a)r(k-k') \\
&= \frac{1}{a}\sum_{k=1}^{a}\sum_{k'=1}^{a}\psi(k/a)\psi(k'/a)\int_{-\pi}^{\pi}f(\lambda)e^{i\lambda(k-k')}d\lambda \\
&= \int_{-a\pi}^{a\pi}f\big(\frac{u}{a}\big)\times\frac{1}{a^2}\sum_{k=1}^{a}\sum_{k'=1}^{a}\psi\big(\frac{k}{a}\big)\psi\big(\frac{k'}{a}\big)e^{iu\left(\frac{k}{a}-\frac{k'}{a}\right)}du \\
&= \int_{-a\pi}^{a\pi}f\big(\frac{u}{a}\big)\times\Big\{\Big(\frac{1}{a}\sum_{k=1}^{a}\psi\big(\frac{k}{a}\big)\cos\big(\frac{k}{a}u\big)\Big)^2+\Big(\frac{1}{a}\sum_{k=1}^{a}\psi\big(\frac{k}{a}\big)\sin\big(\frac{k}{a}u\big)\Big)^2\Big\}\,du \qquad (2.5.2)
\end{aligned}$$

Now, it is well known that if $\psi \in \tilde{W}(\beta,L)$ the Sobolev space with parameters $\beta > 1/2$ and $L > 0$, then

$$\sup_{|u|\le a\pi}\Delta_a(u)\le C_{\beta,L}\frac{1}{a^{\beta-1/2}} \quad \text{with} \quad \Delta_a(u):=\Big|\frac{1}{a}\sum_{k=1}^{a}\psi\big(\frac{k}{a}\big)e^{-iu\frac{k}{a}}-\int_0^1\psi(t)e^{-iut}dt\Big|, \qquad (2.5.3)$$

with $C_{\beta,L} > 0$ only depending on β and L (see for instance Devore and Lorentz, [11]). Therefore if ψ satisfies Assumption $W(\infty)$ and X Assumption A1, for all $\beta > 1/2$, since

$\sup_{u\in\mathbb{R}} |\widehat{\psi}(u)| < \infty$,

$$\begin{aligned}\Big|\mathbb{E}(e^2(a,0)) - \int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big) \times |\widehat{\psi}(u)|^2\, du\Big| &\leq 2C_{\beta,L}\,\frac{2}{a^{\beta-3/2}}\int_0^{a\pi} f\big(\frac{u}{a}\big)|\widehat{\psi}(u)|\, du \\ &+ C^2_{\beta,L}\,\frac{2}{a^{2\beta-2}}\int_0^{a\pi} f\big(\frac{u}{a}\big)\, du \\ &\leq 2\cdot C^2_{\beta,L}\,\frac{2}{a^{2\beta-3}}\int_0^{\pi} f(v)\, dv, \qquad (2.5.4)\end{aligned}$$

since $\sup_{u\in\mathbb{R}}(1+u^n)|\widehat{\psi}(u)| < \infty$ for all $n \in \mathbb{N}$. Consequently, if ψ satisfies Assumption $W(\infty)$, for all $n > 0$, for all $a \in \mathbb{N}^*$, there exists $C(n) > 0$ not depending on a such that

$$\Big|\mathbb{E}(e^2(a,0)) - \int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big) \times |\widehat{\psi}(u)|^2\, du\Big| \leq C(n)\,\frac{1}{a^n}. \qquad (2.5.5)$$

But from Assumption $W(\infty)$, for all $c < 1$,

$$K_{(\psi,c)} = \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(u)|^2}{|u|^c}\, du < \infty,$$

because Assumption $W(\infty)$ implies that $|\widehat{\psi}(u)| = O(|u|)$ when $u \to 0$ and there exists $p > 1-c$ such that $\sup_{u\in\mathbb{R}} |\widehat{\psi}(u)|^2(1+|u|)^p < \infty$. Moreover, for all $p > 1-c$,

$$\begin{aligned}\Big|\int_{-a\pi}^{a\pi} \frac{|\widehat{\psi}(u)|^2}{|u|^c}\, du - K_{(\psi,c)}\Big| &= 2\int_{a\pi}^{\infty} \frac{|\widehat{\psi}(u)|^2}{u^c}\, du \\ &\leq C\cdot\int_{a\pi}^{\infty} \frac{1}{u^{p+c}}\, du \\ &\leq C'\cdot\frac{1}{a^{p+c-1}},\end{aligned}$$

with $C > 0$ and $C' > 0$ not depending on a. As a consequence, under Assumption A1, for all $p > 1-D$, all $n \in \mathbb{N}$ and all $a \in \mathbb{N}^*$,

$$\begin{aligned}\Big|\mathbb{E}(e^2(a,0)) - f^*(0)\cdot\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(u)|^2}{|u/a|^D}\, du\Big| &\leq 2f^*(0)a^D\int_{a\pi}^{\infty} \frac{|\widehat{\psi}(u)|^2}{u^D}\, du \\ &+ C_{D'}a^{D-D'}\int_{-a\pi}^{a\pi} \frac{|\widehat{\psi}(u)|^2}{|u|^{D-D'}}\, du + C(n)\,\frac{1}{a^n} \\ \Longrightarrow \quad \Big|\mathbb{E}(e^2(a,0)) - f^*(0)K_{(\psi,D)}\cdot a^D\Big| &\leq C' f^*(0)\cdot a^{1-p} + C_{D'}K_{(\psi,D-D')}\cdot a^{D-D'}.\end{aligned}$$

Now, by choosing p such that $1-p < D-D'$, the inequality (2.2.1) is obtained. □

Proof [Property 2.2] Using the proof of previous Property 2.1, with Assumption $W(5/2)$, ψ is included in a Sobolev space $\widetilde{W}(5/2, L)$, inequality (3.5.1) is checked with $\beta = 5/2$ and (2.5.4) is replaced by

$$\Big|\mathbb{E}(e^2(a,0)) - a\int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big) \times |\widehat{\psi}(u)|^2\, du\Big| \leq 2\cdot C^2_{5/2,L}\,\frac{2}{a^2}\int_0^{\pi} f(v)\, dv, \qquad (2.5.6)$$

since $\sup_{u\in\mathbb{R}}(1+u^{3/2})|\widehat{\psi}(u)| < \infty$. Therefore, inequality (3.5.12) is replaced by

$$\Big|\mathbb{E}(e^2(a,0)) - a\int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du\Big| \quad\leq\quad C(2)\,\frac{1}{a^2}.$$

The end of the proof is similar to the end of the previous proof, but now $K_{(\psi,c)}$ exists for $-2<c<1$ and

$$\Big|\int_{-a\pi}^{a\pi}\frac{|\widehat{\psi}(u)|^2}{|u|^c}\,du - K_{(\psi,c)}\Big| \leq C'\cdot\frac{1}{a^{2+c}}.$$

Finally, under Assumption A1', for all $a\in\mathbb{N}^*$, since $-2<D-D'<1$,

$$\Big|\mathbb{E}(e^2(a,0)) - f^*(0)K_{(\psi,D)}\cdot a^D\Big| \quad\leq\quad C_{D'}K_{(\psi,D-D')}\cdot a^{D-D'} + C'\,\frac{1}{a^2},$$

which achieves the proof. □

Proof [Corollary 3.1] Both these proofs provide main arguments to establish (3.2.3). For better readability , we will consider only Assumption A1' and Assumption $W(\infty)$ (the long memory process being similar). The main difference consists in specifying the asymptotic behavior of $\displaystyle\int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du$. But,

$$\int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du = \int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du + 2\int_{\sqrt{a}}^{a\pi} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du. \quad (2.5.7)$$

The asymptotic behavior of $\widehat{\psi}(u)$ when $u\to\infty$ (ψ is considered to satisfy Assumption $W(\infty)$), this behavior induces that

$$\int_{\sqrt{a}}^{a\pi} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du \leq Ca^D\int_{\sqrt{a}}^{\infty} u^{-D}\times|\widehat{\psi}(u)|^2\,du \leq \frac{C(n)}{a^n}, \quad (2.5.8)$$

for all $n\in\mathbb{N}$. Moreover,

$$\begin{aligned}\int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du \quad&=\quad f^*(0)\int_{-\sqrt{a}}^{\sqrt{a}}\Big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\Big)|\widehat{\psi}(u)|^2\,du\\ &+\quad \int_{-\sqrt{a}}^{\sqrt{a}}\Big(f\big(\frac{u}{a}\big) - f^*(0)\Big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\Big)\Big)|\widehat{\psi}(u)|^2\,du. \quad (2.5.9)\end{aligned}$$

From computations of previous proofs,

$$\int_{-\sqrt{a}}^{\sqrt{a}}\Big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\Big)|\widehat{\psi}(u)|^2\,du = K_{(\psi,D)}\cdot a^D + C_{D'}K_{(\psi,D-D')}\cdot a^{D-D'} + \Lambda(a), \quad (2.5.10)$$

and $|\Lambda(a)| \leq \dfrac{C(n)}{a^n}$. Finally, using $f(\lambda) = f^*(0)\big(|\lambda|^{-D} + C_{D'}|\lambda|^{D'-D}\big) + o\big(|\lambda|^{D'-D}\big)$ when $\lambda\to 0$, we obtain

$$\begin{aligned}&\int_{-\sqrt{a}}^{\sqrt{a}}\Big(f\big(\frac{u}{a}\big) - f^*(0)\big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\big)\Big)|\widehat{\psi}(u)|^2\,du\\ &\qquad = \int_{-\sqrt{a}}^{\sqrt{a}}\Big|\frac{u}{a}\Big|^{D-D'}\Big(f\big(\frac{u}{a}\big) - f^*(0)\big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\big)\Big)|\widehat{\psi}(u)|^2\Big|\frac{u}{a}\Big|^{D'-D}\,du\\ &\qquad = a^{D-D'}\int_{-\sqrt{a}}^{\sqrt{a}} g(u,a)|\widehat{\psi}(u)|^2|u|^{D'-D}du,\end{aligned}$$

with for all $u \in [-\sqrt{a}, \sqrt{a}]$, $g(u,a) \to 0$ when $a \to \infty$. Therefore, from Lebesgue Theorem (checked from the asymptotic behavior of $\widehat{\psi}$),

$$\lim_{a\to\infty} a^{D-D'} \int_{-\sqrt{a}}^{\sqrt{a}} \Big(f\big(\frac{u}{a}\big) - f^*(0)\Big(\Big|\frac{u}{a}\Big|^{-D} + C_{D'}\Big|\frac{u}{a}\Big|^{D'-D}\Big)\Big) |\widehat{\psi}(u)|^2 \, du = 0. \quad (2.5.11)$$

As a consequence, from (2.5.7), (2.5.8), (2.5.9), (2.5.10) and (2.5.11), the corollary is proven. □

Proof [Proposition 3.1] This proof can be decomposed into three steps :**Step 1**, **Step 2** and **Step 3**.

Step 1. In this part, $\dfrac{N}{a_N} \cdot \mathrm{Cov}(\widetilde{T}_N(r_i a_N), \widetilde{T}_N(r_j a_N))_{1\le i,j\le \ell}$ is proven to converge at an asymptotic covariance matrix Γ. First, for all $(i,j) \in \{1,\dots,\ell\}^2$,

$$\mathrm{Cov}(\widetilde{T}_N(r_i a_N), \widetilde{T}_N(r_j a_N)) = \\ 2\frac{1}{[N/r_i a_N]}\frac{1}{[N/r_j a_N]} \sum_{p=1}^{[N/r_i a_N]} \sum_{q=1}^{[N/r_j a_N]} \Big(\mathrm{Cov}(\tilde{e}(r_i a_N, p), \tilde{e}(r_j a_N, q))\Big)^2, \quad (2.5.12)$$

because X is a Gaussian process. Therefore, by considering only $i = j$ and $p = q$, for N and a_N large enough,

$$\mathrm{Cov}(\widetilde{T}_N(r_i a_N), \widetilde{T}_N(r_i a_N)) \geq \frac{1}{r_i}\frac{N}{a_N}. \quad (2.5.13)$$

Now, for $(p,q) \in \{1,\dots,[N/r_i a_N]\} \times \{1,\dots,[N/r_i a_N]\}$,

$$\begin{aligned}
&\mathrm{Cov}\big(\tilde{e}(r_i a_N, p), \tilde{e}(r_j a_N, q)\big) \\
&= \frac{a_N^{1-D}(r_i r_j)^{(1-D)/2}}{f^*(0)K_{(\psi,D)}} \frac{1}{r_i a_N}\frac{1}{r_j a_N} \sum_{k=1}^{r_i a_N}\sum_{k'=1}^{r_j a_N} \psi\big(\frac{k}{r_i a_N}\big)\psi\big(\frac{k'}{r_j a_N}\big) r\big(k - k' + a_N(r_i p - r_j q)\big) \\
&= \frac{a_N^{1-D}(r_i r_j)^{(1-D)/2}}{f^*(0)K_{(\psi,D)}} \frac{1}{r_i a_N}\frac{1}{r_j a_N} \sum_{k=1}^{r_i a_N}\sum_{k'=1}^{r_j a_N} \psi\big(\frac{k}{r_i a_N}\big)\psi\big(\frac{k'}{r_j a_N}\big) \int_{-\pi}^{\pi} d\lambda\, f(\lambda) e^{-i\lambda(k-k'+a_N(r_i p - r_j q))} \\
&= \frac{(r_i r_j)^{(1-D)/2}}{a_N^D f^*(0)K_{(\psi,D)}} \frac{1}{r_i a_N}\frac{1}{r_j a_N} \sum_{k=1}^{r_i a_N}\sum_{k'=1}^{r_j a_N} \psi\big(\frac{k}{r_i a_N}\big)\psi\big(\frac{k'}{r_j a_N}\big) \int_{-\pi a_N}^{\pi a_N} du\, f\big(\frac{u}{a_N}\big) e^{-iu(\frac{k}{a_N}-\frac{k'}{a_N}+r_i p - r_j q)}.
\end{aligned}$$

Using the same expansion as in (3.5.12), under Assumption $W(\infty)$ the previous equality becomes, for all $n \in \mathbb{N}^*$,

$$\begin{aligned}
&\Big|\mathrm{Cov}\big(\tilde{e}(r_i a_N, p), \tilde{e}(r_j a_N, q)\big) - \frac{(r_i r_j)^{(1-D)/2}}{a_N^D f^*(0)K_{(\psi,D)}} \int_{-\pi a_N}^{\pi a_N} du\, \widehat{\psi}(u r_i)\overline{\widehat{\psi}}(u r_j) f\big(\frac{u}{a_N}\big) e^{-iu(r_i p - r_j q)}\Big| \\
&\qquad \le \frac{C(n)}{a_N^{n+D}} \int_{-\pi a_N}^{\pi a_N} du\, \big|\widehat{\psi}(u r_i)\overline{\widehat{\psi}}(u r_j) f\big(\frac{u}{a_N}\big)\big| \\
&\qquad \le \frac{C'(n)}{a_N^n} \int_{-\infty}^{\infty} du\, |u|^{-D} \big|\widehat{\psi}(u r_i)\overline{\widehat{\psi}}(u r_j)\big| \\
&\qquad \le \frac{C''(n)}{a_N^n}, \quad (2.5.14)
\end{aligned}$$

with $C(n), C'(n), C''(n) > 0$ not depending on a_N and due the asymptotic behaviors of $\widehat{\psi}(u)$ when $u \to 0$ and $u \to \infty$. Now, under Assumption A1,

$$\Big|\int_{-\pi a_N}^{\pi a_N} du\, \widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j) f\big(\frac{u}{a_N}\big) e^{-iu(r_ip-r_jq)} - a_N f^*(0)\int_{-\pi}^{\pi} du\, \frac{\widehat{\psi}(ur_ia_N)\overline{\widehat{\psi}}(ur_ja_N)}{|u|^D} e^{-iua_N(r_ip-r_jq)}\Big|$$
$$\le a_N^{D-D'} f^*(0) C_{D'} \int_{-\pi a_N}^{\pi a_N} du\, \frac{\big|\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)\big|}{|u|^{D-D'}}$$
$$\le a_N^{D-D'} f^*(0) C_{D'} \int_{-\infty}^{\infty} du\, \frac{\big|\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)\big|}{|u|^{D-D'}}, \qquad (2.5.15)$$

since $\displaystyle\int_{-\infty}^{\infty} du\, \frac{|\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)|}{|u|^{D-D'}} < \infty$ from Assumption $W(\infty)$. Finally, from (2.5.14) and (2.5.15), we have $C > 0$ not depending on N such that for all $a_N \in \mathbb{N}^*$,

$$\Big|\mathrm{Cov}\big(\tilde{e}(r_ia_N,p), \tilde{e}(r_ja_N,q)\big) - \frac{a_N^{1-D}(r_ir_j)^{(1-D)/2}}{K_{(\psi,D)}} \int_{-\pi}^{\pi} du\, \frac{\widehat{\psi}(ur_ia_N)\overline{\widehat{\psi}}(ur_ja_N)}{|u|^D} e^{-iua_N(r_ip-r_jq)}\Big| \le C\, a_N^{-D'}. \quad (2.5.16)$$

It remains to evaluate

$$a_N^{1-D}\int_{-\pi}^{\pi} du\, \frac{\widehat{\psi}(ur_ia_N)\overline{\widehat{\psi}}(ur_ja_N)}{|u|^D} e^{-iua_N(r_ip-r_jq)} = \int_{-\pi a_N}^{\pi a_N} du\, \frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{|u|^D} e^{-iu(r_ip-r_jq)}$$

. Thus, if $|r_ip - r_jq| \ge 1$, using an integration by parts,

$$\Big|\int_{-\pi a_N}^{\pi a_N} du\, \frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{|u|^D} e^{-iu(r_ip-r_jq)}\Big| = \Big|\frac{1}{-i(r_ip-r_jq)}\Big[\frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D} e^{-iu(r_ip-r_jq)}\Big]_{-\pi a_N}^{\pi a_N}$$
$$+ \frac{1}{i(r_ip-r_jq)}\int_{-\pi a_N}^{\pi a_N} du\, \frac{\partial}{\partial u}\Big(\frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D}\Big) e^{-iu(r_ip-r_jq)}\Big|$$
$$\le \frac{1}{|r_ip-r_jq|}\int_{-\infty}^{\infty}\Big(\frac{D}{|u|^{D+1}}\big|\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)\big| + \frac{1}{|u|^D}\Big|\frac{\partial}{\partial u}\Big(\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)\Big)\Big|\Big)\, du$$
$$\le C\frac{1}{|r_ip-r_jq|} \qquad (2.5.17)$$

with $C < \infty$ not depending on N, since :

- $\widehat{\psi}(\pi r_ia_N)\overline{\widehat{\psi}}(\pi r_ja_N) = \widehat{\psi}(-\pi r_ia_N)\overline{\widehat{\psi}}(-\pi r_ja_N)$ and $\sin(\pi a_N(r_ip - r_jq)) = 0$;
- from Assumption $W(\infty)$, $\limsup_{u\to 0} u^{-1}\,|\widehat{\psi}(u)| < \infty$, $\limsup_{u\to 0}\big|\frac{\partial}{\partial u}\widehat{\psi}(u)\big| < \infty$

$$\Longrightarrow \limsup_{u\to 0} u^{-1}\Big|\frac{\partial}{\partial u}\Big(\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)\Big)\Big| < \infty;$$

- from Assumption $W(\infty)$,
  for all $n \in \mathbb{N}$, $\sup_{u\in\mathbb{R}}(1+|u|)^n\,|\widehat{\psi}(u)| < \infty$ and $\sup_{u\in\mathbb{R}}(1+|u|)^n\,\big|\frac{\partial}{\partial u}\widehat{\psi}(u)\big| < \infty$.

Moreover, if $|r_ip - r_jq| = 0$, from Cauchy-Schwartz Inequality and Property 2.1, for a_N large enough

$$\Big|\mathrm{Cov}\big(\tilde{e}(r_ia_N,p), \tilde{e}(r_ja_N,q)\big)\Big| \le \Big(\mathbb{E}(\tilde{e}^2(r_ia_N,p))\cdot\mathbb{E}(\tilde{d}^2(r_ja_N,q))\Big)^{1/2} \le 2. \qquad (2.5.18)$$

Therefore, using (2.5.16), (2.5.17) and (2.5.18) and the inequality $(x+y)^2 \leq 2(x^2+y^2)$ for all $(x,y) \in \mathbb{R}^2$, we have $C>0$ such that for a_N large enough,

$$\mathrm{Cov}^2\big(\tilde{e}(r_i a_N,p),\tilde{e}(r_j a_N,q)\big) \leq C\Big(\frac{1}{(1+|r_i p - r_j q|)^2}+\frac{1}{a_N^{2D'}}\Big) \qquad (2.5.19)$$

Hence, with (3.5.19),

$$\begin{aligned}\Big|\mathrm{Cov}(\tilde{T}_N(r_i a_N) \quad &, \quad \tilde{T}_N(r_j a_N))\Big| \\ &\leq C\frac{1}{[N/r_i a_N]}\frac{1}{[N/r_j a_N]}\sum_{p=1}^{[N/r_i a_N]}\sum_{q=1}^{[N/r_j a_N]}\Big(\frac{1}{(1+|r_i p - r_j q|)^2}+\frac{1}{a_N^{2D'}}\Big)\end{aligned}$$

But, from the theorem of comparison between sums and integrals,

$$\begin{aligned}\sum_{p=1}^{[N/r_i a_N]}\sum_{q=1}^{[N/r_j a_N]}(1+|r_i p - r_j q|)^{-2} &\leq \frac{1}{r_i r_j}\int_0^{N/a_N}\int_0^{N/a_N}\frac{du\,dv}{(1+|u-v|)^2} \\ &\leq \frac{2}{r_i r_j}\int_0^{N/a_N}\frac{N/a_N\,dw}{(1+w)^2} \\ &\leq \frac{2}{r_i r_j}\cdot\frac{N}{a_N}.\end{aligned}$$

As a consequence, if a_N is such that $\limsup_{N\to\infty}\dfrac{N}{a_N}\dfrac{1}{a_N^{2D'}}<\infty$ then

$$\limsup_{N\to\infty}\frac{N}{a_N}\Big|\mathrm{Cov}(\tilde{T}_N(r_i a_N),\tilde{T}_N(r_j a_N))\Big|<\infty$$

. More precisely, since this covariance is a sum of positive terms, if $\limsup_{N\to\infty}\dfrac{N}{a_N}\dfrac{1}{a_N^{2D'}}=0$,

$$\lim_{N\to\infty}\frac{N}{a_N}\Big(\mathrm{Cov}(\tilde{S}_N(r_i a_N),\tilde{S}_N(r_j a_N))\Big)_{1\leq i,j\leq \ell}=\Gamma(r_1,\cdots,r_\ell,\psi,D), \qquad (2.5.20)$$

a non null (from (2.5.13)) symmetric matrix with $\Gamma(r_1,\cdots,r_\ell,\psi,D)=(\gamma_{ij})_{1\leq i,j\leq \ell}$ that can be specified. Indeed, from the previous computations, if $\limsup_{N\to\infty}\dfrac{N}{a_N}\dfrac{1}{a_N^{2D'}}=0$,

$$\begin{aligned}\gamma_{ij} &= \lim_{N\to\infty}\frac{8r_i r_j a_N}{N}\sum_{p=1}^{[N/r_i a_N]}\sum_{q=1}^{[N/r_j a_N]}\Big(\frac{(r_i r_j)^{(1-D)/2}}{K_{(\psi,D)}}\int_0^\infty du\,\frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D}\cos(u(r_i p - r_j q))\Big)^2 \\ &= \lim_{N\to\infty}\frac{8(r_i r_j)^{2-D}a_N}{K_{(\psi,D)}^2 N}\sum_{m=-[N/d_{ij}a_N]+1}^{[N/d_{ij}a_N]-1}(\frac{N}{d_{ij}a_N}-|m|)\Big(\int_0^\infty du\,\frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D}\cos(u\,d_{ij}m)\Big)^2 \\ &= \frac{8(r_i r_j)^{2-D}}{K_{(\psi,D)}^2 d_{ij}}\sum_{m=-\infty}^{\infty}\Big(\int_0^\infty\frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{u^D}\cos(u\,d_{ij}m)\,du\Big)^2,\end{aligned}$$

with $d_{ij}=GCD(r_i\,;\,r_j)$. Therefore, the matrix Γ depends only on on $r_1,\cdots,r_\ell,\psi,D$.

Step 2.Generaly speaking, the above result is not sufficient to obtain the central limit theorem,

$$\sqrt{\frac{N}{a_N}}\Big(\tilde{T}_N(r_i a_N)-\mathbb{E}(\tilde{e}^2(r_i a_N,0)\Big)_{1\leq i\leq \ell}\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}_\ell(0,\Gamma(r_1,\cdots,r_\ell,\psi,D)). \quad (2.5.21)$$

However, each $\tilde{T}_N(r_i a_N)$ is a quadratic form of a Gaussian process. *Mutatis mutandis*, it is exactly the same framework (*i.e.* a Lindeberg central limit theorem) as that of Proposition 2.1 in Bardet [6], and (2.5.21) is checked. Moreover, if $(a_n)_n$ is such that $\limsup_{N\to\infty}\frac{N}{a_N^{1+2D'}}=0$ then using the asymptotic behavior of $\mathbb{E}(\tilde{e}^2(r_i a_N,0)$ provided in Property 2.1,

$$\sqrt{\frac{N}{a_N}}\Big(\mathbb{E}(\tilde{e}^2(r_i a_N,0)\Big)\ \xrightarrow[N\to\infty]{}\ 0.$$

As a consequence, under those assumptions,

$$\sqrt{\frac{N}{a_N}}\Big(\tilde{T}_N(r_i a_N)-1\Big)_{1\leq i\leq \ell}\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}_\ell(0,\Gamma(r_1,\cdots,r_\ell,\psi,D)). \quad (2.5.22)$$

Step 3. The logarithm function $(x_1,..,x_\ell)\in(0,+\infty)^\ell\mapsto(\log x_1,..,\log x_m)$ is $\mathcal{C}^2$ on $(0,+\infty)^\ell$. As a consequence, using the Delta-method, the central limit theorem (3.2.7) for the vector $\Big(\log\tilde{T}_N(r_i a_N)\Big)_{1\leq i\leq\ell}$ follows with the same asymptotical covariance matrix $\Gamma(r_1,\cdots,r_\ell,\psi,D)$ (because the Jacobian matrix of the function in $(1,..,1)$ is the identity matrix). □

Proof [Proposition 2.2] There is a perfect identity between this proof and that of Proposition 3.1, both of which are based on the approximations of Fourier transforms provided in the proof of Property 2.2. □

Proof [Corollary 2.3] It is clear that $X'_t=X_t+P_m(t)$ for all $t\in\mathbb{Z}$, with $X=(X_t)_t$ satisfying Proposition 3.1 and 2.2. But, any wavelet coefficient of $(P_m(t))_t$ is obviously null from the assumption on ψ. Therefore the statistic $\widehat{T}_N$ is the same for X and X'. □

Proof [Proposition 2.5] Let $\varepsilon>0$ be a fixed positive real number, such that $\alpha^*+\varepsilon<1$.

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

$$\begin{aligned}\Pr\left(\widehat{\alpha}_N\leq\alpha^*+\varepsilon\right) &\geq \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2)\leq\min_{\alpha\geq\alpha^*+\varepsilon\text{ and }\alpha\in\mathcal{A}_N}\widehat{Q}_N(\alpha)\Big)\\ &\geq 1-\Pr\Big(\bigcup_{\alpha\geq\alpha^*+\varepsilon\text{ and }\alpha\in\mathcal{A}_N}\widehat{Q}_N(\alpha^*+\varepsilon/2)>\widehat{Q}_N(\alpha)\Big)\\ &\geq 1-\sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/\ell]}\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2)>\widehat{Q}_N(\frac{k}{\log N})\Big). \quad (2.5.23)\end{aligned}$$

But, for $\alpha\geq\alpha^*+1$,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2)>\widehat{Q}_N(\alpha)\Big)=\Pr\Big(\Big\|P_N(\alpha^*+\varepsilon/2)\cdot Y_N(\alpha^*+\varepsilon/2)\Big\|^2>\Big\|P_N(\alpha)\cdot Y_N(\alpha)\Big\|^2\Big)$$

with $P_N(\alpha) = I_\ell - A_N(\alpha)\cdot\big(A'_N(\alpha)\cdot A_N(\alpha)\big)^{-1}\cdot A_N(\alpha)$ for all $\alpha\in(0,1)$, *i.e.* $P_N(\alpha)$ is the matrix of an orthogonal projection on the orthogonal subspace (in $\mathbb{R}^\ell$) generated by $A_N(\alpha)$ (and I_ℓ is the identity matrix in $\mathbb{R}^\ell$). From the expression of $A_N(\alpha)$, it is obvious that for all $\alpha\in(0,1)$,

$$P_N(\alpha) = P = I_\ell - A\cdot\big(A'\cdot A\big)^{-1}\cdot A,$$

with the matrix $A = \begin{pmatrix}\log(r_1) & 1\\ \vdots & \vdots \\ \log(r_\ell) & 1\end{pmatrix}$ as in Proposition 2.3. Thereby,

$$\begin{aligned}
&\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big)\\
= \; &\Pr\Big(\Big\|P\cdot Y_N(\alpha^*+\varepsilon/2)\Big\|^2 > \Big\|P\cdot Y_N(\alpha)\Big\|^2\Big)\\
= \; &\Pr\bigg(\Big\|P\cdot\sqrt{\frac{N}{N^{\alpha^*+\varepsilon/2}}}\,Y_N(\alpha^*+\varepsilon/2)\Big\|^2 > N^{\alpha-(\alpha^*+\varepsilon/2)}\Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2\bigg)\\
\le \; &\Pr\Big(V_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) + \Pr\Big(V_N(\alpha)\le N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big)
\end{aligned}$$

with $V_N(\alpha) = \Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2$ for all $\alpha\in(0,1)$. From Proposition 3.1, for all $\alpha>\alpha^*$, the asymptotic law of $P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)$ is a Gaussian law with covariance matrix $P\cdot\Gamma\cdot P'$. Moreover, the rank of the matrix is $P\cdot\Gamma\cdot P'$ is $\ell-2$ (this is the rank of P) and we have $0<\lambda_-$, not depending on N) such that $P\cdot\Gamma\cdot P' - \lambda_- P\cdot P'$ is a non-negative matrix ($0<\lambda_-<\min\{\lambda\in \mathrm{Sp}(\Gamma)\}$). As a consequence, for a large enough N,

$$\begin{aligned}
\Pr\Big(V_N(\alpha)\le N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) &\le 2\cdot\Pr\Big(V_-\le N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big)\\
&\le \frac{1}{2^{\ell/2-2}\Gamma(\ell/2)}\cdot\Big(\frac{N}{\lambda_-}\Big)^{-(\frac{\ell}{2}-1)\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}},
\end{aligned}$$

with $V_-\sim\lambda_-\cdot\chi^2(\ell-2)$. Moreover, from Markov inequality,

$$\begin{aligned}
\Pr\Big(V_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) &\le 2\cdot\Pr\Big(\exp(\sqrt{V_+}) > \exp\big(N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)\Big)\\
&\le 2\cdot\mathbb{E}(\exp(\sqrt{V_+}))\cdot\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)
\end{aligned}$$

with $V_+\sim\lambda_+\cdot\chi^2(\ell-2)$ and $\lambda_+>\max\{\lambda\in\mathrm{Sp}(\Gamma)\}>0$. Like $\mathbb{E}(\exp(\sqrt{V_+}))<\infty$ does not depend on N, we obtain that $M_1>0$ not depending on N, such that for large enough N,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big)\le M_1\cdot N^{-(\frac{\ell}{2}-1)\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}},$$

and therefore, the inequality (2.5.23) becomes, for N large enough,

$$\begin{aligned}
\Pr\big(\widehat{\alpha}_N\le\alpha^*+\varepsilon\big) &\ge 1 - M_1\cdot\sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/\ell]} N^{-\frac{(\ell-2)}{4}\Big(\big(\frac{k}{\log N}\big)-(\alpha^*+\varepsilon/2)\Big)}\\
&\ge 1 - M_1\cdot\log N\cdot N^{-\frac{(\ell-2)}{12}\varepsilon}. \qquad (2.5.24)
\end{aligned}$$

II. Secondly, a bound of $\Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon)$ is provided. Following the above arguments and notations ,

$$\begin{aligned} & \Pr\left(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\right) \\ \geq \;& \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) \leq \min_{\alpha \leq \alpha^*-\varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \qquad (2.5.25) \\ \geq \;& 1 - \sum_{k=2}^{[(\alpha^*-\varepsilon)\log N]+1} \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\frac{k}{\log N})\Big), \qquad (2.5.26) \end{aligned}$$

and as above,

$$\begin{aligned} & \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) \\ & = \Pr\Bigg(\Big\|P \cdot \sqrt{\frac{N}{N^{\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon}}}\, Y_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon)\Big\|^2 > N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)} \Big\|P \cdot \sqrt{\frac{N}{N^{\alpha}}}\, Y_N(\alpha)\Big\|^2\Bigg). \qquad (2.5.27) \end{aligned}$$

Now, in the case $a_N = N^\alpha$ with $\alpha \leq \alpha^*$, the sample variance of wavelet coefficients is biased. In this case, from the relation of Corollary 3.1 under Assumption A1',

$$\Big(Y_N(\alpha)\Big)_{1\leq i\leq \ell} = \Big(\frac{C_{D'}K_{(\psi,D-D')}}{f^*(0)K_{(\psi,D)}}(iN^\alpha)^{-D'}(1+o_i(1))\Big)_{1\leq i\leq \ell} + \Big(\sqrt{\frac{N^\alpha}{N}} \cdot \varepsilon_N(\alpha)\Big)_{1\leq i\leq \ell},$$

with $o_i(1) \to 0$ when $N \to \infty$ for all i and $\mathbb{E}(Z_N(\alpha)) = 0$. As a consequence, for large enough N,

$$\begin{aligned} \Big\|P \cdot \sqrt{\frac{N}{N^\alpha}}\, Y_N(\alpha)\Big\|^2 &= \Big\|P \cdot \varepsilon_N(\alpha)\Big\|^2 + N^{\frac{\alpha^*-\alpha}{\alpha^*}} \Big\|P \cdot \Big(\frac{C_{D'}K_{(\psi,D-D')}}{f^*(0)K_{(\psi,D)}} i^{-D'}(1+o_i(1))\Big)_{1\leq i\leq \ell}\Big\|^2 \\ &\geq D \cdot N^{\frac{\alpha^*-\alpha}{\alpha^*}}, \end{aligned}$$

with $D > 0$, because the vector $(i^{-D'})_{1\leq i\leq \ell}$ is not in the orthogonal subspace of the subspace generated by the matrix A. Then, the relation (2.5.27) becomes,

$$\begin{aligned} \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) &\leq \Pr\Big(\Big\|P \cdot \sqrt{\frac{N}{N^{\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon}}}\, Y_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon)\Big\|^2 \\ &\geq D \cdot N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)} \cdot N^{\frac{\alpha^*-\alpha}{\alpha^*}}\Big) \\ &\leq \Pr\Big(V_+ \geq D \cdot N^{\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)}\Big) \\ &\leq M_2 \cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon}, \end{aligned}$$

with $M_2 > 0$, because $V_+ \sim \lambda_+ \cdot \chi^2(\ell-2)$ and $\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon) \geq \frac{1-\alpha^*}{2\alpha^*}\varepsilon$ for all $\alpha \leq \alpha^* - \varepsilon$. Hence, from the inequality (2.5.26), for large enough N ,

$$\Pr\left(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\right) \geq 1 - M_2 \cdot \log N \cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon}. \qquad (2.5.28)$$

The inequalities (2.5.24) and (2.5.28) imply that $\Pr\left(|\widehat{\alpha}_N - \alpha| \geq \varepsilon\right) \underset{N\to\infty}{\longrightarrow} 0.$ □

Proof [Theorem 3.1] The central limit theorem of (3.3.4) can be established from the following arguments. First, $\Pr(\tilde{\alpha}_N > \alpha^*) \underset{N\to\infty}{\longrightarrow} 1$. Following the previous proof, there is for all $\varepsilon > 0$,

$$\Pr\big(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\big) \geq 1 - M_2 \cdot \log N \cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon}.$$

Consequently, if $\varepsilon_N = \lambda \cdot \dfrac{\log\log N}{\log N}$ with $\lambda > \dfrac{2}{(\ell-2)D'}$ then,

$$\begin{aligned}
\Pr\big(\widehat{\alpha}_N \geq \alpha^* - \varepsilon_N\big) &\geq 1 - M_2 \cdot \log N \cdot N^{-\lambda\frac{(\ell-2)D'}{2}\cdot\frac{\log\log N}{\log N}} \\
&\geq 1 - M_2 \cdot \big(\log N\big)^{1-\lambda\frac{(\ell-2)D'}{2}} \\
\Longrightarrow \quad & \Pr\big(\widehat{\alpha}_N + \varepsilon_N \geq \alpha^*\big) \underset{N\to\infty}{\longrightarrow} 1.
\end{aligned}$$

Now, from Corollary 2.4, $\widehat{D'}_N \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} D'$. Therefore, $\Pr\big(\widehat{D'}_N \leq \frac{4}{3}D'\big) \underset{N\to\infty}{\longrightarrow} 1$. Thus, with $\lambda \geq \dfrac{9}{4(\ell-2)D'}$, $\Pr\big(\tilde{\alpha}_N + (\varepsilon_N - \dfrac{3}{(\ell-2)\widehat{D'}_N} \cdot \dfrac{\log\log N}{\log N}) \geq \alpha^*\big) \underset{N\to\infty}{\longrightarrow} 1$ which implies $\Pr(\tilde{\alpha}_N > \alpha^*) \underset{N\to\infty}{\longrightarrow} 1$.
Secondly, for $x \in_R$,

$$\begin{aligned}
\lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}(\tilde{D}_N - D) \leq x\Big) &= \lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}(\tilde{D}_N - D) \leq x \bigcap \tilde{\alpha}_N > \alpha^*\Big) \\
&\quad + \lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}(\tilde{D}_N - D) \leq x \bigcap \tilde{\alpha}_N \leq \alpha^*\Big) \\
&= \lim_{N\to\infty} \int_{\alpha^*}^1 \Pr\Big(\sqrt{\frac{N}{N^{\alpha}}}(\tilde{D}_N - D) \leq x\Big) f_{\widehat{\alpha}_N}(\alpha)\, d\alpha \\
&= \lim_{N\to\infty} \Pr\Big(Z_\Gamma \leq x\Big) \cdot \int_{\alpha^*}^1 f_{\widehat{\alpha}_N}(\alpha)\, d\alpha \\
&= \Pr\Big(Z_\Gamma \leq x\Big),
\end{aligned}$$

with $f_{\widehat{\alpha}_N}(\alpha)$ the probability density function of $\widehat{\alpha}_N$ and $Z_\Gamma \sim \mathcal{N}(0\,;\,(A'\cdot A)^{-1}\cdot A'\cdot\Gamma\cdot A\cdot(A'\cdot A)^{-1})$.

To prove the second part of (3.3.4), we infer deduces from above that

$$\Pr\Big(\alpha^* < \tilde{\alpha}_N < \alpha^* + \frac{3}{(\ell-2)\widehat{D'}_N} \cdot \frac{\log\log N}{\log N} + \mu \cdot \frac{\log\log N}{\log N}\Big) \underset{N\to\infty}{\longrightarrow} 1,$$

with $\mu > \frac{12}{\ell-2}$. Therefore, $\nu < \frac{4}{(\ell-2)D'} + \frac{12}{\ell-2}$,

$$\Pr\Big(N^{\alpha^*} < N^{\tilde{\alpha}_N} < N^{\alpha^*} \cdot (\log N)^\nu\Big) \underset{N\to\infty}{\longrightarrow} 1.$$

This inequality and the previous central limit theorem result in : for all $\rho > \nu/2$, and $\varepsilon > 0$,

$$\begin{aligned}
\Pr\Big(\frac{N^{\frac{D'}{1+2D'}}}{(\log N)^\rho} \cdot \big|\tilde{D}_N - D\big| > \varepsilon\Big) &= \Pr\Big(\frac{N^{\frac{1}{2}(\widehat{\alpha}_N - \alpha^*)}}{(\log N)^\rho} \cdot \sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big|\tilde{D}_N - D\big| > \varepsilon\Big) \\
&\underset{N\to\infty}{\longrightarrow} 0. \ \square
\end{aligned}$$

$N = 10^3$

	$\sqrt{MSE}$	$\ell = 5$	$\ell = 10$	$\ell = 15$	$\ell = 20$	$\ell = 25$	$\begin{cases} \ell_1 = 15 \\ \ell_2 = \widehat{\ell} \end{cases}$
fGn ($H = \frac{D+1}{2}$)	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.16, 0.75	0.14, 0.19	0.13, 0.17	**0.14, 0.15**	**0.14, 0.15**	0.15, 0.18
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.12, 0.32	0.07, 0.13	0.05, 0.08	0.04, 0.05	**0.04, 0.04**	0.05, 0.08
FARIMA$(0, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.21, 0.81	0.15, 0.20	0.14, 0.17	**0.15, 0.15**	**0.15, 0.15**	0.15, 0.19
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.14, 0.34	0.07, 0.13	0.05, 0.09	0.05, 0.06	**0.04, 0.04**	0.05, 0.09
FARIMA$(1, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.30, 0.96	0.28, 0.35	**0.27, 0.29**	**0.29, 0.27**	0.30, 0.30	0.31, 0.35
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.19, 0.44	0.15, 0.24	0.12, 0.17	0.11, 0.15	**0.11, 0.12**	0.12, 0.17
FARIMA$(1, \frac{D}{2}, 1)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.60, 0.92	0.43, 0.41	0.39, 0.35	0.36, 0.35	0.32, 0.33	**0.21, 0.20**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.17, 0.38	0.11, 0.18	0.09, 0.12	0.07, 0.09	**0.06, 0.07**	0.09, 0.12
$X^{(D,D')}$, $D' = 1$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.33, 0.68	0.29, 0.28	0.27, 0.26	0.26, 0.27	**0.25, 0.25**	0.29, 0.30
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.10, 0.22	**0.10, 0.07**	0.11, 0.07	0.12, 0.12	0.13, 0.13	0.11, 0.07

$N = 10^4$

	$\sqrt{MSE}$	$\ell = 5$	$\ell = 10$	$\ell = 15$	$\ell = 20$	$\ell = 25$	$\begin{cases} \ell_1 = 15 \\ \ell_2 = \widehat{\ell} \end{cases}$
fGn ($H = \frac{D+1}{2}$)	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.08, 0.26	0.05, 0.05	0.05, 0.05	**0.04, 0.04**	**0.04, 0.04**	**0.04, 0.04**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.08, 0.22	0.05, 0.06	**0.04, 0.05**	**0.04, 0.05**	0.05, 0.05	**0.04, 0.05**
FARIMA$(0, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.08, 0.31	0.06, 0.06	**0.05, 0.05**	**0.05, 0.05**	**0.05, 0.05**	**0.05, 0.05**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.09, 0.24	0.05, 0.07	**0.04, 0.05**	**0.04, 0.05**	0.05, 0.05	**0.04, 0.05**
FARIMA$(1, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.13, 0.57	0.10, 0.10	**0.09, 0.08**	**0.09, 0.08**	0.09, 0.09	**0.09, 0.08**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.15, 0.36	0.09, 0.16	0.08, 0.11	0.07, 0.09	**0.06, 0.08**	0.08, 0.11
FARIMA$(1, \frac{D}{2}, 1)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.22, 0.63	0.17, 0.15	0.16, 0.13	0.15, 0.14	0.15, 0.14	**0.09 , 0.09**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.16, 0.38	0.11, 0.17	0.08, 0.11	0.07, 0.09	**0.06, 0.07**	0.08, 0.11
$X^{(D,D')}$, $D' = 1$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.23, 0.36	0.19, 0.15	0.18, 0.17	0.17, 0.17	**0.15, 0.14**	**0.15, 0.14**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.10, 0.18	**0.12, 0.08**	0.13, 0.12	0.14, 0.14	0.15, 0.15	0.13, 0.12

$N = 10^5$

	$\sqrt{MSE}$	$\ell = 5$	$\ell = 10$	$\ell = 15$	$\ell = 20$	$\ell = 25$	$\begin{cases} \ell_1 = 15 \\ \ell_2 = \widehat{\ell} \end{cases}$
fGn ($H = \frac{D+1}{2}$)	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.04, 0.09	0.03, 0.03	0.02, 0.03	**0.02, 0.02**	**0.02, 0.02**	**0.02, 0.02**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.07, 0.16	**0.06, 0.04**	0.06, 0.06	0.07, 0.07	0.07, 0.07	0.06, 0.06
FARIMA$(0, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.03, 0.13	**0.02, 0.02**	**0.02, 0.02**	**0.02, 0.02**	**0.02, 0.02**	**0.02, 0.02**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.07, 0.18	0.04, 0.05	**0.04, 0.03**	0.04, 0.04	0.05, 0.05	**0.04, 0.03**
FARIMA$(1, \frac{D}{2}, 0)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.05, 0.25	0.05, 0.04	0.04, 0.03	0.04, 0.03	0.04, 0.04	**0.03, 0.02**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.12, 0.30	0.07, 0.12	0.05, 0.07	0.04, 0.06	**0.04, 0.05**	0.05, 0.07
FARIMA$(1, \frac{D}{2}, 1)$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.08, 0.30	0.06, 0.04	0.05, 0.04	0.05, 0.04	0.05, 0.05	**0.04, 0.03**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.13, 0.33	0.09, 0.15	0.08, 0.11	0.07, 0.09	**0.06, 0.08**	0.08, 0.11
$X^{(D,D')}$, $D' = 1$	$\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.13, 0.19	0.11, 0.08	0.10, 0.08	0.09, 0.09	0.09, 0.09	**0.08, 0.07**
	$\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.09, 0.15	**0.10, 0.07**	0.11, 0.09	0.12, 0.11	0.13, 0.13	0.11, 0.09

TABLE 2.1 – *Consistency of estimators $\widehat{\widehat{D}}_N$, $\tilde{D}_N$, $\widehat{\alpha}_N$, $\tilde{\alpha}_N$ following ℓ from simulations of the different long-memory processes of the benchmark. For each value of N (10^3, 10^4 and 10^5), of D (0.1, 0.3, 0.5, 0.7 and 0.9) and ℓ (5, 10, 15, 20, 25 and $(15, \widehat{\ell})$), 100 independent samples of each process are generated. The $\sqrt{MSE}$ of each estimator is obtained from a mean of $\sqrt{MSE}$ obtained for the different values of D.*

		FARIMA(0, −0.25, 0)	$X^{(-1,1)}$	$X^{(-1,3)}$	$X^{(-3,1)}$	$X^{(-3,3)}$
$N = 10^3$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.15, 0.20	0.30, 0.30	0.38, 0.37	0.36, 0.37	0.39, 0.38
$N = 10^4$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.04, 0.04	0.15, 0.14	0.08, 0.08	0.13, 0.14	0.13, 0.13
$N = 10^5$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.03, 0.03	0.06, 0.05	0.04, 0.03	0.04, 0.04	0.03, 0.03

TABLE 2.2 – Estimation of the memory parameter from 100 independent samples in case of short memory ($D \leq 0$).

		$P1$	$P2$	$P3$	$P4$
$N = 10^3$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.22, 0.23	0.32, 0.41	0.47, 0.76	0.40, 0.41
$N = 10^4$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.06, 0.06	0.18, 0.28	0.24, 0.65	0.13, 0.13
$N = 10^5$	$\sqrt{MSE}$ $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.02, 0.02	0.02, 0.02	0.14, 0.47	0.03, 0.04

TABLE 2.3 – *Estimation of the long-memory parameter from* 100 *independent samples in case of processes* $P1 - 4$ *defined above.*

			$D = 0.1$	$D = 0.3$	$D = 0.5$	$D = 0.7$	$D = 0.9$
$N = 10^3$ ⟶	fGn ($H = (D+1)/2$)	$\widehat{D}_{BGK}$	**0.089**	0.171	0.259	0.341	0.369
		$\widehat{D}_{GRS}$	0.114	**0.132**	0.147	0.155	0.175
		$\widehat{D}_{MS}$	0.163	0.169	0.181	0.195	0.191
		$\widehat{D}_{R}$	0.211	0.220	0.215	0.218	**0.128**
		$\widehat{D}_{ATV}$	0.176	0.153	0.156	0.164	0.162
		$\widehat{\widehat{D}}_N$	0.139	0.147	**0.133**	**0.140**	0.150
	FARIMA(0, $\frac{D}{2}$, 0)	$\widehat{D}_{BGK}$	**0.094**	0.138	0.239	0.326	0.413
		$\widehat{D}_{GRS}$	0.131	0.139	0.150	0.150	0.162
		$\widehat{D}_{MS}$	0.172	0.167	0.174	0.197	0.188
		$\widehat{D}_{R}$	0.246	0.189	0.223	0.234	0.181
		$\widehat{D}_{ATV}$	0.128	**0.107**	**0.081**	**0.074**	**0.065**
		$\widehat{\widehat{D}}_N$	0.161	0.146	0.149	0.149	0.161
	FARIMA(1, $\frac{D}{2}$, 0)	$\widehat{D}_{BGK}$	**0.146**	**0.203**	0.239	0.236	0.212
		$\widehat{D}_{GRS}$	0.519	0.545	0.588	0.585	0.830
		$\widehat{D}_{MS}$	0.235	0.258	0.256	0.252	0.249
		$\widehat{D}_{R}$	0.242	0.241	**0.234**	**0.202**	0.144
		$\widehat{D}_{ATV}$	0.248	0.267	0.280	0.268	0.375
		$\widehat{\widehat{D}}_N$	0.340	0.319	0.314	0.315	0.334
	FARIMA(1, $\frac{D}{2}$, 1)	$\widehat{D}_{BGK}$	0.204	0.253	0.342	0.363	0.384
		$\widehat{D}_{GRS}$	0.901	0.894	0.866	0.870	0.893
		$\widehat{D}_{MS}$	0.181	**0.175**	**0.180**	**0.185**	0.181
		$\widehat{D}_{R}$	0.204	0.200	0.200	0.191	**0.130**
		$\widehat{D}_{ATV}$	0.392	0.380	0.371	0.343	0.355
		$\widehat{\widehat{D}}_N$	**0.170**	0.218	0.225	0.226	0.213
	$X^{(D,D')}$, $D' = 1$	$\widehat{D}_{BGK}$	**0.090**	**0.139**	0.261	0.328	0.388
		$\widehat{D}_{GRS}$	0.342	0.339	0.331	0.300	0.315
		$\widehat{D}_{MS}$	0.176	0.178	0.182	**0.166**	0.177
		$\widehat{D}_{R}$	0.219	0.232	0.231	0.173	**0.167**
		$\widehat{D}_{ATV}$	0.153	0.161	**0.168**	0.176	0.176
		$\widehat{\widehat{D}}_N$	0.284	0.294	0.293	0.292	0.288

$N = 10^4 \longrightarrow$

		$D = 0.1$	$D = 0.3$	$D = 0.5$	$D = 0.7$	$D = 0.9$
fGn ($H = (D+1)/2$)	$\widehat{D}_{BGK}$	0.062	0.143	0.182	0.171	0.182
	$\widehat{D}_{GRS}$	0.040	0.047	0.054	0.068	0.066
	$\widehat{D}_{MS}$	0.069	0.064	0.061	0.071	0.063
	$\widehat{D}_{R}$	0.063	0.055	0.058	0.063	0.052
	$\widehat{D}_{ATV}$	**0.036**	**0.042**	**0.041**	0.047	0.045
	$\widehat{\widehat{D}}_{N}$	0.050	0.040	**0.041**	**0.039**	**0.040**
FARIMA$(0, \frac{D}{2}, 0)$	$\widehat{D}_{BGK}$	0.059	0.141	0.195	0.187	0.178
	$\widehat{D}_{GRS}$	0.042	0.048	0.050	0.046	0.057
	$\widehat{D}_{MS}$	0.072	0.055	0.066	0.059	0.065
	$\widehat{D}_{R}$	0.073	0.053	0.064	0.057	0.059
	$\widehat{D}_{ATV}$	**0.026**	**0.038**	**0.039**	**0.032**	**0.022**
	$\widehat{\widehat{D}}_{N}$	0.053	0.050	0.056	0.055	0.044
FARIMA$(1, \frac{D}{2}, 0)$	$\widehat{D}_{BGK}$	0.085	0.148	0.146	0.164	0.120
	$\widehat{D}_{GRS}$	0.179	0.175	0.182	0.192	0.190
	$\widehat{D}_{MS}$	0.109	0.105	0.099	0.100	0.094
	$\widehat{D}_{R}$	**0.063**	**0.059**	**0.057**	**0.054**	**0.054**
	$\widehat{D}_{ATV}$	0.118	0.101	0.088	0.120	0.081
	$\widehat{\widehat{D}}_{N}$	0.095	0.085	0.093	0.081	0.097
FARIMA$(1, \frac{D}{2}, 1)$	$\widehat{D}_{BGK}$	0.111	0.201	0.189	0.202	0.181
	$\widehat{D}_{GRS}$	0.308	0.321	0.306	0.314	0.311
	$\widehat{D}_{MS}$	0.070	0.064	0.065	**0.064**	0.069
	$\widehat{D}_{R}$	**0.063**	**0.057**	**0.060**	**0.064**	**0.052**
	$\widehat{D}_{ATV}$	0.114	0.118	0.103	0.102	0.093
	$\widehat{\widehat{D}}_{N}$	0.095	0.099	0.087	0.101	0.090
$X^{(D,D')}$, $D' = 1$	$\widehat{D}_{BGK}$	0.069	0.110	0.204	0.190	0.197
	$\widehat{D}_{GRS}$	0.192	0.185	0.172	0.177	0.190
	$\widehat{D}_{MS}$	0.083	0.059	0.071	0.066	0.068
	$\widehat{D}_{R}$	**0.066**	**0.057**	**0.068**	**0.054**	**0.064**
	$\widehat{D}_{ATV}$	0.124	0.131	0.139	0.147	0.153
	$\widehat{\widehat{D}}_{N}$	0.158	0.143	0.152	0.158	0.155

TABLE 2.4 – *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.*

Bibliography

[1] Abry, P. and Veitch, D. (1998) Wavelet analysis of long-range-dependent traffic, *IEEE Trans. on Info. Theory*, 44, 2–15.

[2] Abry, P. and Veitch, D. (1999) A wavelet based joint estimator of the parameter of long-range dependence traffic, *IEEE Trans. on Info. Theory, Special Issue on "Multiscale Statistical Signal Analysis and its Application"*, 45, 878–897.

[3] Abry, P., Veitch, D. and Flandrin, P. (1998). Long-range dependent : revisiting aggregation with wavelets, *J. Time Ser. Anal.* , 19, 253–266.

[4] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens, In P. Doukhan, G. Oppenheim and M.S. Taqqu editors, *Long-range Dependence : Theory and Applications*, Birkhäuser.

[5] Andrews, D.W.K. and Sun, Y. (2004). Adaptive local polynomial Whittle estimation of long-range dependence, *Econometrica*, 72, 569–614.

[6] Bardet, J.M. (2000). Testing for the presence of self-similarity of Gaussian time series having stationary increments, *J. Time Ser. Anal.* , 21, 497–516.

[7] Bardet, J.M. (2002). Statistical Study of the Wavelet Analysis of Fractional Brownian Motion, *IEEE Trans. Inf. Theory*, 48, 991–999.

[8] Bardet J.M. and Bertrand, P. (2007). Identification of the multiscale fractional Brownian motion with biomechanical applications, *J. Time Ser. Anal.*, 28, 1–52.

[9] Bardet, J.M., Lang, G., Moulines, E. and Soulier, P. (2000). Wavelet estimator of long range-dependant processes, *Statist. Infer. Stochast. Processes*, 3, 85–99.

[10] Bhansali, R. Giraitis L. and Kokoszka, P.S. (2006). Estimation of the memory parameter by fitting fractionally-differenced autoregressive models, *J. Multivariate Analysis*, 97, 2101–2130.

[11] Devore P. and Lorentz, G. (1993). *Constructive approximations*, Springer.

[12] Doukhan, P., Oppenheim, G. and Taqqu M.S. (Editors) (2003). *Theory and applications of long-range dependence*, Birkhäuser.

[13] Giraitis, L., Robinson P.M., and Samarov, A. (1997). Rate optimal semi-parametric estimation of the memory parameter of the Gaussian time series with long range dependence, *J. Time Ser. Anal.*, 18, 49–61.

[14] Giraitis, L., Robinson P.M., and Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter, *J. Multivariate Anal.*, 72, 183–207.

[15] Moulines, E., Roueff, F. and Taqqu, M.S. (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.

[16] Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractionnal processes, In P. Doukhan, G. Openheim and M.S. Taqqu editors, *Theory and applications of long-range dependence*, 251–301, Birkhäuser, Boston.

[17] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence, *The Annals of statistics*, 23, 1630–1661.

[18] Taqqu, M.S., and Teverovsky, V. (1996). Semiparametric graphical estimation technics for long memeory data, *Lectures Notes in Statistics*, 115, 420–432.

[19] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377–390.

Chapitre 3

Wavelet based estimator of D for stationary linear processes

Adaptive semiparametric wavelet estimator and goodness-of-fit test for long memory linear processes

Accepted in Journal of Time series Analysis (2011)

Jean-Marc Bardet[a] and *Hatem Bibi*[a,b],

bardet@univ-paris1.fr, hatem.bibi@malix.univ-paris1.fr

[a] : SAMM, Université Paris 1, 90 rue de Tolbiac, 75013 Paris, FRANCE.
[b] : Département de Mathématiques, Faculté des Sciences de Tunis,1060 Tunis, TUNISIE.

abstract This paper is first devoted to the study of an adaptive wavelet-based estimator of the long-memory parameter for linear processes in a general semiparametric frame. As such this is an extension of the previous contribution of Bardet *et al.* (2008) which only concerned Gaussian processes. Moreover, the definition of the long-memory parameter estimator has been modified and the asymptotic results are improved even in the Gaussian case. Finally an adaptive goodness-of-fit test is also built and easy to be employed : it is a chi-square type test. Simulations confirm the interesting properties of consistency and robustness of the adaptive estimator and test.

3.1 Introduction

Presently, long memory processes have become a widely-studied subject area and find frequent applications (see for instance Doukhan *et al.*, 2003). The best known long-memory stationary time series are the fractional Gaussian noises (fGn) with Hurst parameter H and the 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. This behavior of the spectral density is frequently considered as a definition of a stationary long-memory (or long-range-dependent) process where d is the long memory parameter.

In this paper, we study a general case of linear process with a memory parameter d and we propose an adaptive wavelet-based estimator of this parameter. Hence for $d < 1/2$ and $d' > 0$, we consider the following semiparametric framework :

Assumption A(d,d') **:** $X = (X_t)_{t\in\mathbb{Z}}$ *is a zero mean stationary linear process, i.e.*

$$X_t = \sum_{s\in\mathbb{Z}} \alpha(t-s)\xi_s, \quad t \in \mathbb{Z}, \quad \textit{where}$$

- $(\xi_s)_{s\in\mathbb{Z}}$ *is a sequence of independent identically distributed random variables following a symmetric distribution, i.e. for all* $M \in \mathbb{R}$, $\Pr(\xi_0 > M) = \Pr(\xi_0 < -M)$, *and satisfying* $\mathbb{E}\xi_0 = 0$, $Var\xi_0 = 1$ *and* $\mu_4 := \mathbb{E}\xi_0^4 < \infty$ *;*
- $(\alpha(t))_{t\in\mathbb{Z}}$ *is a sequence of real numbers such that there exist* $c_d > 0$ *and* $c_{d'} \in \mathbb{R}$ *satisfying*

$$|\widehat{\alpha}(\lambda)|^2 = \frac{1}{\lambda^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1+\varepsilon(\lambda))\right) \quad \textit{for any} \quad \lambda \in [-\pi,0)\cup(0,\pi], \qquad (3.1.1)$$

where $\widehat{\alpha}(\lambda) := \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\alpha(k)e^{-ik\lambda}$ *for* $\lambda \in [-\pi,0)\cup(0,\pi]$ *and* $\varepsilon(\lambda) \to 0$ *(*$\lambda \to 0$*).*

Consequently, if X satisfies Assumption A(d,d'), the spectral density f of X is such that

$$f(\lambda) = 2\pi\,|\widehat{\alpha}(\lambda)|^2 = \frac{2\pi}{\lambda^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1+\varepsilon(\lambda))\right) \quad \text{for any} \quad \lambda \in [-\pi,0)\cup(0,\pi], \qquad (3.1.2)$$

with $\varepsilon(\lambda) \to 0$ $(\lambda \to 0)$. Thus, if $d \in (0,1/2)$, the process X is a long-memory process, and if $d \leq 0$, it is a short-memory process (see Doukhan *et al.*, 2003).

After preliminary studies devoted to self-similar processes, Abry *et al.* (1998) were the first to propose the use of a wavelet-based estimator for estimating the parameter d of a long memory process by computing the log-log regression slope for different scales of wavelet coefficient sample variances. Bardet *et al.* (2000) provided proofs of the consistency of such an estimator in a Gaussian semiparametric frame. Moulines *et al.* (2007) improved these results and established a central limit theorem (CLT in the sequel) for the estimator of d which they proved rate optimal for the minimax criterion. Finally, Roueff and Taqqu (2009a) yielded similar results in a semiparametric frame for linear processes.

All of these studies used a wavelet analysis based on a discrete multi-resolution wavelet transform, which in particular allows to compute the wavelet coefficients with the fast Mallat's algorithm. However, these results are inferred from a semiparametric frame such as to (3.1.2) and consider the "optimal" scale used for the wavelet analysis (which depends on the second order expansion d') to be known although, in fact it is unknown. Two studies present automatic selection method for this "optimal" scale in the Gaussian semiparametric frame. A procedure based on a chi-square test was introduced in Veitch *et al.* (2003) but despite convincing numerical results, it lacks proofs of its consistency. Whereas, Bardet *et al.* (2008) proved the consistency of a procedure for choosing optimal scales based on the detection of the "most linear part" of the log-variogram graph. Moreover, the considered wavelet function is not necessarily associated with a multi-resolution analysis : although the computation cost is more important, this offers a larger wavelet function choice and scales are not limited to powers of 2.

The present paper is an extension of this previous study of Bardet *et al.* (2008). Improvements concern three following central issues :

1. The semiparametric Gaussian framework of Bardet *et al.* (2008) is extended to the semiparametric framework Assumption A(d, d') for linear processes. The same automatic procedure of the optimal scale selection can also be used and thus we obtain adaptive estimators.
2. As in Bardet *et al.* (2008), the "mother" wavelet is not necessarily associated with a discrete multi-resolution transform. We also slightly modified the definition of the wavelet coefficient sample variance ("variogram"). The result of both these changes is a multidimensional central limit theorem satisfied by the logarithms of variograms with a very simple asymptotic covariance matrix (see (3.2.8) for its definition) depending only on d and the Fourier transform of the wavelet function. Hence it is easy to compute an adaptive pseudo-generalized least square estimator (PGLSE in the sequel) of d, satisfying a CLT with an asymptotic variance which is smaller than the adaptive ordinary least square estimator of d. Simulations confirm the good performance of this PGLSE.
3. Finally, we used this PGLSE to perform an adaptive goodness-of-fit test. It represents a normalized sum of the squared PGLS-distance between the PGLS-regression line and the points. We proved that this test statistic converges in distribution to a chi-square distribution. Since the asymptotic covariance matrix is easily approximated, the test is very simple test to compute. When $d > 0$ this test is a long-memory test. Moreover, simulations show that this test provides good properties of consistency under H_0 and reasonable properties of robustness under H_1.

In the light of these results, this paper is a conclusion to the study of Bardet *et al.* (2008), and the adaptive PGLS estimator and test interesting extensions of Roueff and Taqqu (2009a).

The present paper is organized into four sections as follows. Assumptions, definitions and a first multidimensional central limit theorem are the subject matter of Section 3.2. Section 3.3 is devoted to the construction and consistency of the adaptive PGLS estimator and goodness-of-fit test. In Section 3.4 features a Monte Carlo simulations-based demonstration of the convergence of the adaptive estimator, followed by comparisons with other efficient semiparametric estimators and investigations into the consistency and robustness properties of the adaptive goodness-of-fit test. Proofs figure in Section 3.5.

3.2 A central limit theorem for the sample variance of wavelet coefficients

Let $\psi : \mathbb{R} \to \mathbb{R}$ be a function (called the wavelet function) and $k \in \mathbb{N}^*$. We shall consider the following assumption on ψ :

Assumption $\Psi(k)$ **:** *the function* $\psi : \mathbb{R} \to \mathbb{R}$ *is such that*

1. *the support of* ψ *is included in* $(0,1)$ *;*
2. $\displaystyle\int_0^1 \psi(t)\, dt = 0$ *;*
3. $\psi \in \mathcal{C}^k(\mathbb{R})$*, the set of* k*-times continuously differentiable functions on* $\mathbb{R}$*.*

Straightforward implications of Assumption $\Psi(k)$ are :

- $\psi^{(j)}(0) = \psi^{(j)}(1) = 0$ for any $0 \le j \le k$, where $\psi^{(j)}$ is the j-th derivative of ψ.
- If $\widehat{\psi}(u)$ is the Fourier transform of ψ, *i.e.*

$$\widehat{\psi}(u) := \int_0^1 \psi(t)\, e^{-iut} dt,$$

  then $\widehat{\psi}(u) \sim C\, u^k$ $(u \to 0)$ with C a real number not depending on u.
- Moreover,

$$\sup_{u\in\mathbb{R}} \big| u^k\, \widehat{\psi}(u) \big| \le \sup_{x\in[0,1]} |\psi^{(k)}(x)|. \tag{3.2.1}$$

If $Y = (Y_t)_{t\in\mathbb{R}}$ is a continuous-time process, for $(a,b) \in \mathbb{R}_+^* \times \mathbb{R}$, the "classical" wavelet coefficient $d(a,b)$ of the process Y for the scale a and the shift b is $d(a,b) := \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi(\frac{t-b}{a}) Y_t\, dt$. However, since the process X satisfying Assumption $\mathrm{A}(d,d')$ is a discrete-time process, we define the wavelet coefficients of X by

$$e(a,b) := \sum_{j=1}^{a} \Big(\frac{1}{\sqrt{a}}\, \psi(\frac{j}{a}) \Big)\, X_{b+j} \tag{3.2.2}$$

for $(a,b) \in \mathbb{N}^* \times \mathbb{Z}$. Note that if a path $(X_1,\ldots,X_N)$ is observed, for $a \in \mathbb{N}^*$ and $b = 1,\ldots,N-a$ we can also write $e(a,b) = \frac{1}{\sqrt{a}} \sum_{t=1}^N \psi(\frac{t-b}{a}) X_t$, which is more directly implied by the definition of $d(a,b)$.

In the sequel, we will use the usual convention $y = o(g(x))$ $(x \to \infty)$ when $\lim_{x\to\infty} y/g(x) = 0$,

Property 3.1 *Under Assumption $A(d,d')$ with $d < 1/2$ and $d' > 0$, and if ψ satisfies Assumption $\Psi(k)$ with $k > d' - d + 1/2$, for $a \in \mathbb{N}^*$, then $(e(a,b))_{b\in\mathbb{Z}}$ is a zero mean stationary linear process and*

$$\mathbb{E}(e^2(a,0)) = 2\pi\big(c_d K_{(\psi,2d)}\, a^{2d} + c_{d'} K_{(\psi,2d-d')}\, a^{2d-d'}\big) + o\big(a^{2d-d'}\big) \quad \textit{when} \quad a \to \infty, \tag{3.2.3}$$

$$\textit{with} \qquad K_{(\psi,\alpha)} := \int_{-\infty}^{\infty} |\widehat{\psi}(u)|^2\, |u|^{-\alpha} du > 0 \quad \textit{for all } \alpha < 1. \tag{3.2.4}$$

Refer to Section 3.5 for all the proofs of this paper.
Let $(X_1,\ldots,X_N)$ be an observed path of X satisfying Assumption A(d,d'). As soon as a consistent estimator of $\mathbb{E}(e^2(a,0))$ is provided, Property 3.1 allows to make a log-log regression-based estimation of $2d$. Hence, for $a \in \{1,\ldots,N-1\}$, consider the sample variance of the wavelet coefficients,

$$T_N(a) := \frac{1}{N-a} \sum_{b=1}^{N-a} e^2(a,b). \tag{3.2.5}$$

Remark 3.1 *In Bardet et al. (2000), (2008) or in Moulines et al. (2007) or Roueff and Taqqu (2009), the considered sample variance of wavelet coefficients is*

$$V_N(a) := \frac{1}{[N/a]} \sum_{b=1}^{[N/a]} e^2(a,ab) \tag{3.2.6}$$

(with $a = 2^j$ in case of multiresolution analysis). Definition (3.2.5) has both a drawback and two advantages with respect to the usual definition (3.2.6). On the one hand, $T_N(a)$ is not adapted to the fast Mallat's algorithm and therefore its use is more time consuming than the one of $V_N(a)$. Its advantage twofold : if γ and γ' respectively denote the asymptotic variances of $\sqrt{N/a}\, T_N(a)$ and $\sqrt{N/a}\, V_N(a)$ when a, $N \to \infty$, then the expression of γ is clearly simpler than the one of γ' since

$$\begin{cases} \gamma &= 4\pi \dfrac{1}{K^2_{(\psi,2d)}} \displaystyle\int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(\lambda)\big|^4}{|\lambda|^{4d}} d\lambda \quad \textit{(see (3.2.8) below)} \\ \gamma' &= \dfrac{2}{K^2_{(\psi,2d)}} \displaystyle\sum_{m=-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(u)\big|^2}{|u|^{2d}} \cos(u\,m)\, du \Big)^2 \quad \textit{(see Bardet et al., 2008)}, \end{cases}$$

and this will have consequences to the computation of PGLS estimators below. Furthermore, as inferred from numerical approximations not reported here, for our choice of ψ (see Section 3.4), γ is nearly twice smaller than γ' (following d). This confers the same advantage to the variance of the wavelet-based estimators of d computed from $(T_N(ar_i))_i$ with respect to the one computed from $(V_N(ar_i))_i$ (see below).

The following proposition specifies a multidimensional central limit theorem for a vector $(\log \widetilde{T}_N(ar_i))_i$, which provides the first step towards obtaining a CLT for the estimator of d computed from an ordinary least square regression :

Proposition 3.1 *Define $\ell \in \mathbb{N}\backslash\{0,1\}$ and $(r_1,\cdots,r_\ell) \in (\mathbb{N}^*)^\ell$ with $0<r_1<r_2<\cdots<r_\ell$. Under Assumption $A(d,d')$ with $d<1/2$ and $d'>0$, if ψ satisfies Assumption $\Psi(k)$ with $k \geq d'-d+1/2$ and if $(a_n)_{n\in\mathbb{N}}$ is such as $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N\, N^{-1/(1+2d')} \underset{N\to\infty}{\longrightarrow} \infty$, then*

$$\sqrt{\frac{N}{a_N}}\Big(\log T_N(r_i a_N) - 2d\log(r_i a_N) - \log\big(2\pi\, c_d\, K_{(\psi,2d)}\big)\Big)_{1\leq i\leq \ell} \underset{N\to\infty}{\overset{D}{\longrightarrow}} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big), \tag{3.2.7}$$

with $\Gamma(r_1,\ldots,r_\ell,\psi,d)=(\gamma_{ij})_{1\leq i,j\leq\ell}$ the asymptotic covariance matrix such as

$$\gamma_{ij} \;=\; 4\pi\,\frac{(r_i r_j)^{1-2d}}{K^2_{(\psi,2d)}}\int_{-\infty}^{\infty}\frac{\big|\widehat{\psi}(r_i\lambda)\big|^2\big|\widehat{\psi}(r_j\lambda)\big|^2}{\lambda^{4d}}\,d\lambda. \tag{3.2.8}$$

Since it is not easy to minimize $\Gamma(r_1,\ldots,r_\ell,\psi,d)$ in terms of $(r_1,\ldots,r_\ell)$ and for simplifying the following results, we chose now only to consider the case $(r_1,r_2,\ldots,r_\ell)=(1,2,\ldots,\ell)$.

3.3 Adaptive estimator of the memory parameter and adaptive goodness-of-fit test

The CLT of Proposition 3.1 opens a certain number of perspectives. As we shall see, the simple expression of the asymptotic covariance matrix reveals to be very advantageous as compared to the complicated expression of the asymptotic covariance obtained in the case of a multiresolution analysis (see Roueff and Taqqu, 2009a). Proposition 3.1 confirms the consistency of estimator $\widehat{d}_N$ of d. Hence, we define

$$\widehat{d}_N(a_N) := \big(0\ \tfrac{1}{2}\big)\,(Z'_{a_N}\,Z_{a_N})^{-1}Z'_{a_N}\big(\log T_N(r_i a_N)\big)_{1\leq i\leq\ell} \quad\text{with}\quad Z_{a_N}=\begin{pmatrix}1 & \log(a_N)\\ 1 & \log(2a_N)\\ \vdots & \vdots\\ 1 & \log(\ell a_N)\end{pmatrix}. \tag{3.3.1}$$

where A' denotes the transpose of a matrix A. Then, it can be clearly inferred from Proposition 3.1 that $\widehat{d}_N(a_N)$ converges to d following a CLT with convergence rate $\sqrt{N/a_N}$ when a_N satisfies the condition $a_N\,N^{-1/(1+2d')} \underset{N\to\infty}{\longrightarrow} \infty$.

But d' is actually unknown. Bardet *et al.* (2008) presented an automatic procedure for choosing an "optimal" scale a_N. We shall presently apply this procedure. Here a brief recall of its principle : for $\alpha\in(0,1)$, define

$$Q_N(\alpha,c,d)=\Big(Y_N(\alpha)-Z_{N^\alpha}\begin{pmatrix}c\\2d\end{pmatrix}\Big)'\cdot\Big(Y_N(\alpha)-Z_{N^\alpha}\begin{pmatrix}c\\2d\end{pmatrix}\Big),\quad\text{with}\quad Y_N(\alpha)=\big(\log T_N(iN^\alpha)\big)_{1\leq i\leq\ell}.$$

$Q_N(\alpha,c,d)$ corresponds to a squared distance between the ℓ points $\big(\log(iN^\alpha)\,,\,\log T_N(iN^\alpha)\big)_i$ and the line of slope $2d$ and intercept c. It can be minimized in terms of α, c and d first by defining for $\alpha \in (0,1)$

$$\widehat{Q}_N(\alpha) = Q_N\big(\alpha, \widehat{c}(N^\alpha), 2\widehat{d}(N^\alpha)\big) \quad \text{with} \quad \begin{pmatrix} \widehat{c}(N^\alpha) \\ 2\widehat{d}(N^\alpha) \end{pmatrix} = \big(Z'_{N^\alpha} Z_{N^\alpha}\big)^{-1} Z'_{N^\alpha} Y_N(\alpha);$$

and then by defining $\widehat{\alpha}_N$ by :

$$\widehat{Q}_N(\widehat{\alpha}_N) = \min_{\alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha) \quad \text{where} \quad \mathcal{A}_N = \Big\{ \frac{2}{\log N}, \frac{3}{\log N}, \ldots, \frac{\log[N/\ell]}{\log N} \Big\}.$$

Remark 3.2 *As outlined in Bardet et al. (2008) in the definition of the set $\mathcal{A}_N$, $\log N$ can be replaced by any sequence negligible with respect to any power law of N. Hence, in numerical applications we will use $10 \log N$ which significantly increases the precision of $\widehat{\alpha}_N$.*

Under the assumptions of Proposition 3.1, we obtain (see the proof in Bardet *et al.*, 2008),

$$\widehat{\alpha}_N = \frac{\log \widehat{a}_N}{\log N} \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^* = \frac{1}{1+2d'}.$$

We then define :

$$\widehat{\widehat{d}}_N := \widehat{d}(N^{\widehat{\alpha}_N}) \quad \text{and} \quad \widehat{\Gamma}_N := \Gamma(1, \ldots, \ell, \widehat{\widehat{d}}_N, \psi). \tag{3.3.2}$$

It is clear that $\widehat{\widehat{d}}_N \xrightarrow[N\to\infty]{\mathcal{P}} d$ (for a convergence rate see also Bardet *et al.*, 2008) and $\widehat{\Gamma}_N \xrightarrow[N\to\infty]{\mathcal{P}} \Gamma(1, \ldots, \ell, d, \psi)$ (from the expression of Γ in (3.2.8) which is a continuous function of the variable d). We will prefer to consider :

$$\widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6\widehat{\alpha}_N}{(\ell-2)(1-\widehat{\alpha}_N)} \, \frac{\log\log N}{\log N}.$$

rather than $\widehat{\alpha}_N$ for technical reasons (*i.e.* $\Pr(\widetilde{\alpha}_N \leq \alpha^*) \xrightarrow[N\to\infty]{} 0$ which is not satisfied by $\widehat{\alpha}_N$, see Bardet *et al.*, 2008). Consequently, with the usual expression of PGLSE, the adaptive estimators of c and d can be defined as follows :

$$\begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} := \big(Z'_{N^{\widetilde{\alpha}_N}} \widehat{\Gamma}_N^{-1} Z_{N^{\widetilde{\alpha}_N}}\big)^{-1} Z'_{N^{\widetilde{\alpha}_N}} \widehat{\Gamma}_N^{-1} Y_N(\widetilde{\alpha}_N). \tag{3.3.3}$$

The following theorem provides the asymptotic behavior of the estimator $\widetilde{d}_N$,

Theorem 3.1 *Under the assumptions of Proposition 3.1,*

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}}\big(\widetilde{d}_N - d\big) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}\big(0\,;\,\sigma_d^2(\ell)\big) \tag{3.3.4}$$

$$\textit{with} \qquad \sigma_d^2(\ell) := \big(0\ \tfrac{1}{2}\big)\big(Z_1'\big(\Gamma(1, \ldots, \ell, d, \psi)\big)^{-1} Z_1\big)^{-1}\big(0\ \tfrac{1}{2}\big)' \tag{3.3.5}$$

$$\textit{and for all} \quad \rho > \frac{2(1+3d')}{(\ell-2)d'}, \quad \frac{N^{\frac{d'}{1+2d'}}}{(\log N)^\rho} \times \big|\widetilde{d}_N - d\big| \xrightarrow[N\to\infty]{\mathcal{P}} 0. \tag{3.3.6}$$

Remark 3.3 *1. From Gauss-Markov Theorem, the asymptotic variance of $\widetilde{d}_N$ is smaller or equal to the one of $\widehat{\widehat{d}}_N$. Moreover $\widetilde{d}_N$ satisfies the CLT (3.3.4) which provides confidence intervals which can be easily computed.*

2. In the Gaussian case, the adaptive estimator $\widetilde{d}_N$ converges to d with a rate of convergence being equal to the minimax rate of convergence $N^{\frac{d'}{1+2d'}}$ up to a logarithm factor (see Giraitis et al., 1997). Thus, this estimator is comparable to adaptive log-periodogram or local Whittle estimators (see respectively Moulines and Soulier, 2003, and Robinson, 1995).

3. Under additive assumptions on ψ (ψ is supposed to have its first m vanishing moments), the estimator $\widetilde{d}_N$ can also be applied to a process X with an additive polynomial trend of degree $\leq m-1$. Then the trend is being "vanished" by the wavelet function in the expression of the wavelet coefficient and the value of $\widetilde{d}_N$ is the same as the result obtained without this additive trend. No such robustness property can be obtained with the cited adaptive log-periodogram or local Whittle estimator (however an adaptive version of the local Whittle estimator robust for polynomial trends was defined in Andrews and Sun, 2004).

Finally an adaptive goodness-of-fit test can be deduced from the previous PGLS regression. It consists on a sum of the PGLS squared distances between the PGLS regression line and the points. To be precise, consider the statistic :

$$\widetilde{T}_N := \frac{N}{N^{\widetilde{\alpha}_N}} \Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} \Big)' \, \widehat{\Gamma}_N^{-1} \Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} \Big). \qquad (3.3.7)$$

Then, using the previous results, we obtain :

Theorem 3.2 *Under the assumptions of Proposition 3.1,*

$$\widetilde{T}_N \xrightarrow[N\to\infty]{D} \chi^2(\ell-2). \qquad (3.3.8)$$

This (adaptive) goodness-of-fit test is therefore very simple to be computed and used. In the case where $d > 0$, which can be tested easily using Theorem 3.1, this test can also be seen as a test of long memory for linear processes.

3.4 Simulations

We then examined the numerical consistency and robustness of $\widetilde{d}_N$. We proceeded to simulations and we compared the values of $\widetilde{d}_N$ with those of the more accurate semiparametric long-memory estimators. To conclude we examined the numerical properties of the test statistic $\widetilde{T}_N$.

Remark 3.4 *Note that all softwares (in Matlab language) used in this section are freely available access on* `http://samm.univ-paris1.fr/-Jean-Marc-Bardet`.

First of all we need to specify the simulation conditions. The results are based on 100 generated independent samples of each process belonging to the following "benchmark". The concrete generation procedures of these processes are based on the circulant matrix method in case of Gaussian processes and the truncation of an infinite sum if the process is non-Gaussian (see Doukhan *et al.*, 2003). The simulations carried out for $d = 0$, 0.1, 0.2, 0.3 and 0.4, for $N = 10^3$ and 10^4 for all the following processes which satisfy Assumption A(d, d') :

1. the fractional Gaussian noise (fGn) of parameter $H = d + 1/2$ for $d \in [0, 0.5)$ and $\sigma^2 = 1$. A fGn is such that Assumption A$(d, 2)$ holds (even if a fGn is rarely presented as a Gaussian linear process) ;
2. a FARIMA(p, d, q) process with parameter d such that $d \in [0, 0.5)$, $p, q \in \mathbb{N}$. A FARIMA(p, d, q) process is such that Assumption A$(d, 2)$ holds if $(\xi_i)_i$ the innovation process is such that $E\xi_i = 0$, $\mathbb{E}\xi_i^4 < \infty$ and ξ_i symmetric random variables.
3. The centered Gaussian stationary process $X^{(d,d')}$, with spectral density is

$$f_3(\lambda) = \frac{1}{\lambda^{2d}}(1 + \lambda^{d'}) \quad \text{for } \lambda \in [-\pi, 0) \cup (0, \pi], \tag{3.4.1}$$

with $d \in [0, 0.5)$ and $d' \in (0, \infty)$. $X^{(d,d')}$ being a Gaussian process with spectral density f_3, it is considered a linear process within the Wold decomposition Theorem, thus confirming Assumption $A(d, d')$ holds.

The "benchmark" referred to below include the following particular processes for

$$d = 0,\ 0.1,\ 0.2,\ 0.3,\ 0.4 :$$

- X_1 : fGn processes with parameters $H = d + 1/2$;
- X_2 : FARIMA$(0, d, 0)$ processes with standard Gaussian innovations ;
- X_3 : FARIMA$(0, d, 0)$ processes with innovations following a uniform $\mathcal{U}[-1, 1]$ distribution ;
- X_4 : FARIMA$(0, d, 0)$ processes 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 $\mathbb{E}|X_i| < \infty$) ;
- X_5 : FARIMA$(0, d, 0)$ processes 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 $\mathbb{E}|X_i| < \infty$) ;
- X_6 : FARIMA$(1, d, 1)$ processes with standard Gaussian innovations, MA coefficient $\phi = -0.3$ and AR coefficient $\phi = 0.7$;
- X_7 : FARIMA$(1, d, 1)$ processes with innovations following a uniform $\mathcal{U}[-1, 1]$ distribution, MA coefficient $\phi = -0.3$ and AR coefficient $\phi = 0.7$;
- X_8 : $X^{(d,d')}$ Gaussian processes with $d' = 1$.

Note that the processes X_4 and X_5 do not satisfy the condition $\mathbb{E}\xi_0^4$ required in Theorems 3.1 and 3.2. However, considering the logarithm of wavelet coefficient sample variance and not only the wavelet coefficient sample variance, we should be able to prove the consistency of $\widetilde{d}_N$ under $\mathbb{E}\xi_0^r$ with $r \geq 2$.

3.4.1 Comparison of the wavelet-based estimator with other estimators

The wavelet-based estimator has been computed using the following parameters :

Choice of the function ψ : A wavelet function ψ associated with a multi-resolution analysis being not mandatory, as mentioned above, we use function $\psi(x) = x^4(1-x)^4\left(x^2-x+\frac{5}{22}\right)\mathbb{I}_{x\in[0,1]}$

which satisfies Assumption $\Psi(3)$ (and therefore in any cases $3 = k > d' - d + 1/2$ which is required for theoretical limit theorems).

Choice of the parameter ℓ : This parameter largely determines the "beginning" of the linear part of the graph drawn by points $(\log(ia_N), \log T_N(ia_N))_{1 \le i \le \ell}$ and hence the data-driven estimator $\widetilde{a}_N$. We adopted on this point a two step procedure :

1. According to numerical study (not detailed here), $\ell = [2 * log(N)]$ (therefore $\ell = 13$ for $N = 1000$ and $\ell = 18$ for $N = 10000$) seems an appropriate first step : the computation of $\widetilde{\alpha}_n$.
2. Concerning the computation of $\tilde{d}_N$, $\widehat{\Gamma}_N$ seems not be influenced a lot by d. For illustrating this point and using classical approximations of the integrals defined in $\Gamma(1, \dots, \ell, d, \psi)$, we computed $\sigma_d^2(\ell) = \big(0 \ \frac{1}{2}\big)\big(Z_1'\big(\Gamma(1, \dots, \ell, d, \psi)\big)^{-1} Z_1\big)^{-1}\big(0 \ \frac{1}{2}\big)'$ taking into account several values of d and ℓ. For the results of these numerical experiments refer to Figure 2. It can be inferred that for any $d \in [0, 0.5)$, $\sigma_d^2(\ell)$ is almost independent on d and decreases as ℓ increases. Then we chose to select for the second step the "largest" possible value of ℓ, *i.e.* $\ell = N^{1-\widetilde{\alpha}_N}(\log N)^{-1}$ which induces that the larger considered scale is $N(\log N)^{-1}$ (which is negligible with respect to N, confirming the CLT (3.2.7).

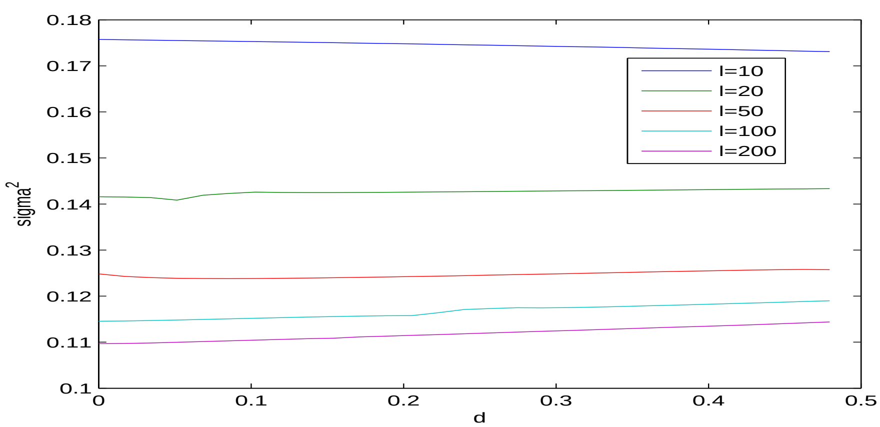

FIGURE 3.1 – *Graph of the approximated values of* $\sigma_d^2(\ell)$ *defined in (3.3.4) for* $d \in [0, 0.5]$ *and* $\ell = 10, 20, 50, 100$ *and* 200.

We applied $\widetilde{d}_N$ to the above mentioned benchmark but also both the following semiparametric d-estimators (see Bardet *et al*, 2003 or 2008) , we obtain :

- $\widehat{d}_{MS}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier (1998, 2003), also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;
- $\widehat{d}_R$ is the local Whittle estimator introduced by Robinson (1995). The trimming parameter is $m = N/30$.

For simulation results see Table 3.1.

Conclusions from Table 3.1 : For each process and value of d and N, $\sqrt{MSE}$ takes into account 100 independently generated samples. The frequency of acceptation of the

$N = 10^3 \longrightarrow$

Model	$\sqrt{MSE}$	$d=0$	$d=0.1$	$d=0.2$	$d=0.3$	$d=0.4$
X_1	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.089	0.091	0.096	0.090	0.100
	$\sqrt{MSE}\ \widehat{d}_R$	0.102	0.114	0.116	0.106	0.102
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.047**	**0.045**	**0.039**	**0.044**	**0.048**
	$\widetilde{p}_n$	0.91	0.82	0.79	0.73	0.68
X_2	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.091	0.094	0.086	0.091	0.099
	$\sqrt{MSE}\ \widehat{d}_R$	0.107	0.105	0.112	0.110	0.097
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.047**	**0.048**	**0.052**	**0.057**	**0.066**
	$\widetilde{p}_n$	0.85	0.86	0.80	0.74	0.68
X_3	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.092	0.094	0.080	0.099	0.096
	$\sqrt{MSE}\ \widehat{d}_R$	0.113	0.113	0.100	0.112	0.095
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.046**	**0.049**	**0.055**	**0.055**	**0.070**
	$\widetilde{p}_n$	0.87	0.85	0.79	0.84	0.71
X_4	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.088	0.079	0.079	0.093	0.104
	$\sqrt{MSE}\ \widehat{d}_R$	0.096	0.100	0.103	0.097	0.095
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.052**	**0.053**	**0.055**	**0.062**	**0.063**
	$\widetilde{p}_n$	0.86	0.82	0.79	0.74	0.71
X_5	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.069	0.067	0.077	0.121	0.143
	$\sqrt{MSE}\ \widehat{d}_R$	0.072	0.078	0.093	0.087	0.074
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.053**	**0.054**	**0.060**	**0.064**	**0.071**
	$\widetilde{p}_n$	0.82	0.80	0.78	0.75	0.74
X_6	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.096**	**0.091**	**0.090**	**0.086**	**0.093**
	$\sqrt{MSE}\ \widehat{d}_R$	0.111	0.102	0.100	0.101	0.101
	$\sqrt{MSE}\ \widetilde{d}_N$	0.154	0.153	0.143	0.168	0.141
	$\widetilde{p}_n$	0.69	0.68	0.68	0.65	0.57
X_7	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.085**	**0.096**	**0.086**	**0.093**	0.098
	$\sqrt{MSE}\ \widehat{d}_R$	0.106	0.116	0.097	0.099	**0.092**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.145	0.148	0.150	0.161	0.148
	$\widetilde{p}_n$	0.70	0.69	0.71	0.67	0.60
X_8	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.097**	**0.104**	**0.097**	**0.094**	**0.101**
	$\sqrt{MSE}\ \widehat{d}_R$	0.120	0.116	0.117	0.113	0.110
	$\sqrt{MSE}\ \widetilde{d}_N$	0.181	0.182	0.180	0.179	0.175
	$\widetilde{p}_n$	0.82	0.80	0.80	0.76	0.73

$N = 10^4 \longrightarrow$

Model	$\sqrt{MSE}$	$d=0$	$d=0.1$	$d=0.2$	$d=0.3$	$d=0.4$
X_1	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.032	0.029	0.031	0.031	0.036
	$\sqrt{MSE}\ \widehat{d}_R$	0.028	0.028	0.029	0.029	0.032
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.016**	**0.018**	**0.018**	**0.024**	**0.024**
	$\widetilde{p}_n$	0.99	0.97	0.97	0.97	0.95
X_2	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.034	0.030	0.029	0.032	0.028
	$\sqrt{MSE}\ \widehat{d}_R$	0.027	0.027	0.029	0.028	0.023
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.017**	**0.019**	**0.021**	**0.022**	**0.022**
	$\widetilde{p}_n$	0.96	0.97	0.96	0.98	0.94
X_3	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.034	0.034	0.033	0.030	0.031
	$\sqrt{MSE}\ \widehat{d}_R$	0.029	0.028	0.028	0.028	0.029
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.016**	**0.017**	**0.019**	**0.020**	**0.020**
	$\widetilde{p}_n$	0.96	0.97	0.95	0.96	0.96
X_4	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.029	0.060	0.036	0.031	0.031
	$\sqrt{MSE}\ \widehat{d}_R$	0.025	0.027	0.029	0.031	0.029
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.016**	**0.021**	**0.022**	**0.022**	**0.024**
	$\widetilde{p}_n$	0.96	0.94	0.92	0.93	0.95
X_5	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.093	**0.046**	0.039	0.073	0.047
	$\sqrt{MSE}\ \widehat{d}_R$	0.040	0.046	0.035	0.032	**0.024**
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.039**	**0.019**	**0.024**	**0.025**	0.025
	$\widetilde{p}_n$	0.93	0.93	0.90	0.92	0.91
X_6	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.031	0.032	0.033	0.032	0.029
	$\sqrt{MSE}\ \widehat{d}_R$	**0.029**	**0.028**	**0.028**	**0.028**	**0.028**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.044	0.044	0.044	0.042	0.048
	$\widetilde{p}_n$	0.94	0.88	0.90	0.92	0.86
X_7	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.030	0.031	0.037	0.030	0.029
	$\sqrt{MSE}\ \widehat{d}_R$	**0.027**	**0.027**	**0.032**	**0.028**	**0.027**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.044	0.043	0.047	0.045	0.049
	$\widetilde{p}_n$	0.95	0.94	0.89	0.90	0.88
X_8	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.038**	0.040	**0.040**	**0.035**	0.037
	$\sqrt{MSE}\ \widehat{d}_R$	0.039	**0.038**	**0.040**	0.036	**0.035**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.084	0.084	0.085	0.083	0.086
	$\widetilde{p}_n$	0.97	0.95	0.93	0.91	0.92

TABLE 3.1 – *Comparison of the different long-memory parameter estimators.*

adaptive goodness-of-fit test is $\widetilde{p}_n = \frac{1}{n}\#\big(\widetilde{T}_N < q_{\chi^2(\ell-2)}(0.95)\big)$ with $n = 100$. Compared to other estimators, $\widetilde{d}_N$ shows numerically convincing convergence rate. With both the "spectral" estimator $\widehat{d}_R$ and $\widehat{d}_{MS}$, the results are quiet stable and not sensible to d and to the flatness of the spectral density of the process. However the spectral density of the process notably effects the convergence rate of $\widetilde{d}_N$. As compared to the other estimators, $\widetilde{d}_N$ is a very accurate and even more efficient for "smooth" spectral densities (fGn and FARIMA$(0, d, 0)$), $\widetilde{d}_N$.

Remark 3.5 *A previous comparison (Bardet et al., 2008) of two adaptive wavelet-based estimators (respectively defined in Veitch et al., (2003) and in Bardet et al., 2008) with $\widehat{d}_{MS}$ and $\widehat{d}_R$ (as well as with two further estimators as defined respectively in Giraitis et al., (2000), and Giraitis et al., (2006) neither of which display good numerical properties of consistency) shows that $\sqrt{MSE}$ of $\widetilde{d}_N$ obtained in Table 3.1 is generally smaller to $\sqrt{MSE}$ of Bardet et al.'s (2008)-based estimator because we opted for definition (3.2.5) instead of (3.2.6) and PGLS regression instead of LS regression.*

Comparison of the robustness of the different semiparametric estimators : To study the robustness of the estimator $\widetilde{d}_N$, take the three different processes not satisfying Assumption $A(d, d')$ as follows :

- A Gaussian stationary process with a spectral density $f(\lambda) = \big||\lambda| - \pi/2\big|^{-2\delta}$ for all $\lambda \in [-\pi, \pi] \setminus \{-\pi/2, \pi/2\}$ so called a GARMA$(0, \delta, 0)$ process. The local behavior of f in 0 is $f(|\lambda|) \sim (\pi/2)^{-2\delta}\,|\lambda|^{-2d}$ with $d = 0$ but it does not satisfy Assumption $A(0, d')$ (here d' should be 2) since $f(\lambda) \to \infty$ $(\lambda \to \pi/2)$.
- A Gaussian FARIMA$(0, d, 0)$ with an additive linear trend $(X_t = FARIMA_t + (1-2t/N)$ for $t = 1, \ldots, N$ and therefore the mean value of $(X_1, \ldots, X_N) \simeq 0)$;
- A Gaussian FARIMA$(0, d, 0)$ with an additive linear trend and an additive sinusoidal seasonal component of period $T = 12$ $(X_t = FARIMA_t + (1-2t/N) + \sin(\pi\, t/6)$ for $t = 1, \ldots, N$ hence the mean value of $(X_1, \ldots, X_N) \simeq 0)$.

For results of these simulations see Table 3.2.

Conclusions from Table 3.2 : The main advantage of $\widetilde{d}_N$ with respect to $\widehat{d}_{MS}$ and $\widehat{d}_R$, as listed in this table, is the robustness with respect to smooth trends (or seasonality). Note that the sample mean value of $\widehat{d}_{MS}$ and $\widehat{d}_R$ for processes with trend or with trend and seasonality is almost 0.5 for any choice of d.

3.4.2 Consistency and robustness of the adaptive goodness-of-fit test :

Tables 3.1 and 3.2 also provide informations concerning the adaptive goodness-of-fit test. The consistency properties of this test are clearly satisfactory when N is large en-

$N = 10^3 \longrightarrow$

Model	$\sqrt{MSE}$	$d(=\delta)=0$	$d(=\delta)=0.1$	$d(=\delta)=0.2$	$d(=\delta)=0.3$	$d(=\delta)=0.4$
GARMA$(0,\delta,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.089	0.091	0.123	0.132	0.166
	$\sqrt{MSE}\ \widehat{d}_R$	0.112	0.111	0.119	0.106	**0.106**
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.052**	**0.050**	**0.080**	**0.079**	0.154
	$\widetilde{p}_n$	0.90	0.91	0.85	0.83	0.77
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.548	0.411	0.292	0.190	0.142
	$\sqrt{MSE}\ \widehat{d}_R$	0.499	0.394	0.279	0.167	0.091
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.044**	**0.045**	**0.040**	**0.044**	**0.041**
	$\widetilde{p}_n$	0.88	0.92	0.90	0.83	0.86
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.479	0.347	0.233	0.142	0.112
	$\sqrt{MSE}\ \widehat{d}_R$	0.499	0.393	0.279	0.167	**0.091**
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.216**	**0.215**	**0.215**	**0.217**	0.185
	$\widetilde{p}_n$	0.35	0.26	0.18	0.21	0.18

$N = 10^4 \longrightarrow$

Model	$\sqrt{MSE}$	$d(=\delta)=0$	$d(=\delta)=0.1$	$d(=\delta)=0.2$	$d(=\delta)=0.3$	$d(=\delta)=0.4$
GARMA$(0,\delta,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.031	0.035	0.039	0.049	0.062
	$\sqrt{MSE}\ \widehat{d}_R$	0.028	0.031	**0.030**	**0.030**	**0.034**
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.016**	**0.029**	0.032	0.038	0.039
	$\widetilde{p}_n$	0.96	0.96	0.95	0.96	0.90
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.452	0.286	0.167	0.096	0.056
	$\sqrt{MSE}\ \widehat{d}_R$	0.433	0.308	0.191	0.100	0.051
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.021**	**0.016**	**0.020**	**0.018**	**0.023**
	$\widetilde{p}_n$	0.96	0.98	0.97	0.97	0.95
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.471	0.307	0.196	0.123	0.076
	$\sqrt{MSE}\ \widehat{d}_R$	0.432	0.305	0.191	0.100	0.052
	$\sqrt{MSE}\ \widetilde{d}_N$	**0.042**	**0.046**	**0.043**	**0.048**	**0.052**
	$\widetilde{p}_n$	0.89	0.86	0.87	0.82	0.72

TABLE 3.2 – *Robustness of the different long-memory parameter estimators. For each process and value of* d *and* N, $\sqrt{MSE}$ *takes into account* 100 *independent generated samples. The frequency of acceptation of the adaptive goodness-of-fit test is* $\widetilde{p}_n = \frac{1}{n}\#\big(\widetilde{T}_N < q_{\chi^2(\ell-2)}(0.95)\big)$ *with* $n = 100$.

ough ($N = 1000$ seems to be too small to correctly using this goodness-of-fit test).

In order to appreciate the behavior of the test statistic under H_1, we consider a process which satisfying neither the stationarity condition nor relation (3.1.2). We have selected 3 particular cases :

1. a process X denoted MFARIMA defined as a succession of two independent Gaussian FARIMA processes. More precisely, we consider $X_t = FARIMA(0, 0.1, 0)$ for $t = 1, \dots, N/2$ and $X_t = FARIMA(0, 0.4, 0)$ for $t = N/2 + 1, \dots, N$.
2. a process X denoted MGN defined by the increments of a multifractional Brownian motion (introduced in Peltier and Lévy-Vehel, 1995). Using the harmonizable representation, define $Y = (Y_t)_t$ by
$$Y_t := C(t) \int_{\mathbb{R}} \frac{e^{itx} - 1}{|x|^{H(t)+1/2}} dW(x)$$
where $H(\cdot)$ as well as $C(\cdot)$ are functions (the case $H(\cdot) = H$ with $H \in (0,1)$ is the case of fBm) and the complex isotropic random measure dW satisfies $dW = dW1 + i\,dW2$

Model	$N = 10^3$	$N = 10^4$
MFARIMA	$\widetilde{p}_n = 0.42$	$\widetilde{p}_n = 0.90$
MGN	$\widetilde{p}_n = 0.13$	$\widetilde{p}_n = 0.07$
MFGN	$\widetilde{p}_n = 0.03$	$\widetilde{p}_n = 0.06$

TABLE 3.3 – *Robustness of the adaptive goodness-of-fit test. The frequency of acceptation of the adaptive goodness-of-fit test is* $\widetilde{p}_n = \frac{1}{n}\#\big(\widetilde{T}_N < q_{\chi^2(\ell-2)}(0.95)\big)$ *(with* $n = 100$ *independent replications).*

with $dW1$ and $dW2$ two independent real-valued Brownian measures (see more details on this part in section 7.2.2 of Samorodnitsky and Taqqu, 1994). When $g = g_1 + i\,g_2$ and $h = h_1 + i\,h_2$ where g_1, h_1 and g_2, h_2 are respectively even and odd real-valued functions such that $\int_{\mathbb{R}}(g_i^2(x))dx < \infty$ and $\int_{\mathbb{R}}(h_i^2(x))dx < \infty$ $(i = 1, 2)$, then $\mathbb{E}\big[\big(\int_{\mathbb{R}} g(\xi)\,dW(\xi)\big)\big(\int_{\mathbb{R}} h(\xi)\,dW(\xi)\big)\big] = \int_{\mathbb{R}} g(x)\overline{h(x)}\,dx$. Here we chose $H(t) = 0.5 + 0.4\sin(t/10)$ and $C(t) = 1$. Then with $X_t = Y_{t+1} - Y_t$ for $t \in \mathbb{Z}$, the process X is not a stationary process, it rather behaves "locally" as a fGn with a parameter $H(t)$ (therefore depending on t).

3. a process X denoted MFGN and defined by the increments of a multiscale fractional Brownian motion (introduced in Bardet and Bertrand, 2007). Let $Z = (Z_t)_t$ be such that

$$Z_t := \int_{\mathbb{R}} \sigma(x)\, \frac{e^{itx} - 1}{|x|^{H(x)+1/2}} dW(x)$$

with dW previously defined, $H(\cdot)$ and $\sigma(\cdot)$ being piecewise constant functions. We chose $\sigma(x) = \mathbb{I}_{0.001 \leq |x| \leq 0.1}$ and $H(x) = 0.9$ for $0.001 \leq |x| \leq 0.04$ and $H(x) = 0.1$ for $0.04 \leq |x| \leq 3$ (such a choice was done for modeling heartbeat signals in the paper Bardet *al.*, 2011). Define $X_t := Z_{t+1} - Z_t$ for $t \in \mathbb{Z}$; then $X = (X_t)_{t\in\mathbb{Z}}$ is a Gaussian stationary process which can be written as a Gaussian linear process (Wold decomposition Theorem) and behaving as a fGn of parameter 0.9 for low frequencies (large time) and as a fGn of parameter 0.1 for high frequencies (small time).

We applied the test statistic based on $\widetilde{T}_N$ to 100 independent replications of these processes. The results figure in Table 3.3. This goodness-of-fit test is rejected for processes MGN and MFGN. Whereas for the process MFARIMA which actually does not satisfy the assumptions of Theorem 3.2 it is not rejected. It is due to the fact the test calculates the average behavior of the sample whereas in case of change (for example MFARIMA) it calculates the average of LRD parameter (a sample mean of 0.28 for $\widetilde{d_N}$ and a standard deviation 0.03 are obtained for $N = 10^4$).

3.5 Proofs

First, we will use many times the following lemmas :

Lemma 3.1 *If g is a function satisfying Assumption $\Psi(k)$ with $k \geq 1$, then for all $\lambda \in \mathbb{R}$,*

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda\,t}dt\Big| \;\leq\; C_g(k)\,\min\Big(\frac{1+|\lambda|^k}{a^k}\,,\,1\Big) \tag{3.5.1}$$

$$\text{with} \quad C_g(k) \;=\; 2\sum_{p=0}^{k}\binom{k}{p}\sup_{x\in[0,1]}|g^{(p)}(x)|. \tag{3.5.2}$$

Proof 3.1 (Proof of Lemma 3.1) *1/ We first prove that if h is a $\mathcal{C}^k(\mathbb{R})$ function such as $h(x)=0$ for $x \notin [0,1]$ with $k \geq 1$, then for all $a>0$:*

$$\Big|\frac{1}{a}\sum_{j=1}^{a} h\big(\frac{j}{a}\big) - \int_0^1 h(t)dt\Big| \leq \sup_{x\in[0,1]}|h^{(k)}(x)|\,\frac{1}{a^k}. \tag{3.5.3}$$

This proof is established by induction on k. If $k=1$, the classical approximation of an integral by a Riemann sum implies

$$\Big|\frac{1}{a}\sum_{j=1}^{a} h\big(\frac{j}{a}\big) - \int_0^1 h(t)dt\Big| \leq \sup_{x\in[0,1]}|h'(x)|\,\frac{1}{a}.$$

Now assume that the relationship (3.5.3) is true for any $k \leq n$ with $n \in \mathbb{N}^*$. We are going to prove that (3.5.3) is also true for $k=n+1$. Indeed, assume that h satisfies Assumption $\Psi(n+1)$. Then, with the usual Taylor expansion

$$\Big|h(t)-h(u)-\sum_{k=1}^{n}\frac{(t-u)^k}{k!}h^{(k)}(u)\Big| \leq \frac{|t-u|^{n+1}}{(n+1)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)| \quad \text{for} \quad (t,u)\in[0,1]^2$$

,

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a} h\big(\frac{j}{a}\big) - \int_0^1 h(t)dt\Big| &\leq \Big|\sum_{j=1}^{a}\int_{(j-1)/a}^{j/a}\sum_{k=1}^{n}\frac{(j/a-t)^k}{k!}h^{(k)}(j/a)dt\Big| + \frac{1}{(n+2)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\frac{1}{a^{n+1}} \\ &\leq \sum_{k=1}^{n}\frac{1}{a^k(k+1)!}\Big|\frac{1}{a}\sum_{j=1}^{a}h^{(k)}(j/a)dt\Big| + \frac{1}{(n+2)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\frac{1}{a^{n+1}}.\end{aligned}$$

Using (3.5.3) for $h^{(k)}$ and $k=1,\ldots,n$, we have

$$\Big|\frac{1}{a}\sum_{j=1}^{a} h^{(k)}(j/a)dt - \int_0^1 h^{(k)}(t)dt\Big| \leq \frac{1}{(n-k+1)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\frac{1}{a^{n+1-k}}$$

since $h^{(k)}$ *satisfies Assumption* $\Psi(n+1-k)$. *But* $\int_0^1 h^{(k)}(t)dt = \left[\frac{1}{(k+1)!}h^{(k+1)}(t)\right]_0^1 = 0$.

Therefore,

$$\Big|\frac{1}{a}\sum_{j=1}^{a} h(\frac{j}{a}) - \int_0^1 h(t)dt\Big| \leq \Big(\sum_{k=1}^{n}\frac{1}{(k+1)!}\frac{1}{(n-k+1)!} + \frac{1}{(n+2)!}\Big) \sup_{x\in[0,1]}|h^{(n+1)}(x)|\,\frac{1}{a^{n+1}}$$
$$\leq (e-2) \sup_{x\in[0,1]}|h^{(n+1)}(x)|\,\frac{1}{a^{n+1}},$$

and thus (3.5.3) is true for $k=n+1$ *and therefore for any* $k\in\mathbb{N}^*$.

2/ Now, we apply (3.5.3) for $h(t)=g(t)e^{-it\lambda}$ *when* $\lambda\in[a,a]$. *Since* $|h^{(k)}(t)|\leq\sum_{p=0}^{k}\binom{k}{p}|\lambda|^p|g^{(k-p)}(t)|$,

and for all $\lambda\in[a,a]$, $\sup_{x\in[0,1]}|h^{(k)}(x)|\leq\max(1,|\lambda|^k)\sum_{p=0}^{k}\binom{k}{p}\sup_{x\in[0,1]}|g^{(p)}(x)|$ *and (3.5.1) follows.*

Now when $|\lambda|>a$, *it is clear that*

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g(\frac{j}{a})e^{-i\lambda\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda t}dt\Big| \leq 2\sup_{x\in[0,1]}|g(x)|$$

and (3.5.1) follows. Moreover, if g *is not the null function, we can not expect a really smaller bound. Indeed, if we denote* λ' *such as* $\int_0^1 g(t)e^{-i\lambda' t}dt\neq 0$ *(if* λ' *does not exist,* $g(x)=0$ *for all* $x\in\mathbb{R}$*). Then, for* $a>\lambda'$ *and for* $\lambda=\lambda'+2n\pi a$ *with* $n\in\mathbb{Z}^*$, *then* $\frac{1}{a}\sum_{j=1}^{a} g(j/a)e^{-i\lambda j/a} = \frac{1}{a}\sum_{j=1}^{a} g(j/a)e^{-i\lambda' j/a} = \int_0^1 g(t)e^{-i\lambda' t} + O(a^{-k})$ *when* $a\to\infty$ *from the previous case* $|\lambda'|\leq a$. *But we also have* $\int_0^1 g(t)e^{-i\lambda t} = O(|\lambda|^{-k}) = O(a^{-k})$ *from* k *integrations by parts since* g *satisfies Assumption* $\Psi(k)$. *Therefore, for any* $\lambda=\lambda'+2n\pi a$ *with* $n\in\mathbb{Z}^*$,

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g(\frac{j}{a})e^{-i\lambda\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda t}dt\Big| = \Big|\int_0^1 g(t)e^{-i\lambda' t}\Big| + O\big(a^{-k}\big)$$

that induces that we cannot expect a better bound than $O(1)$ *when* $\lambda\in\mathbb{R}$.

Lemma 3.2 *If* g *is a function satisfying Assumption* $\Psi(k)$ *with* $k\geq 0$, *then for all* $a\geq 1$ *and* $\lambda\in[-a\pi,0)\cup(0,a\pi]$,

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g(\frac{j}{a})e^{-i\lambda\frac{j}{a}}\Big| \leq D_g(k)\,\frac{1}{|\lambda|^k} \quad \textit{with} \quad D_g(k)=10^k\sup_{x\in[0,1]}|g^{(k)}(x)|. \tag{3.5.4}$$

Proof 3.2 (Proof of Lemma 3.2) *This proof is also established by induction on k. If $k=0$, it is obvious that :*

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}\Big| \le \sup_{x\in[0,1]}|g(x)|,$$

and (3.5.4) is satisfied. Now assume that property (3.5.4) is true for any $k\le n$ with $n\in\mathbb{N}^*$. We are going to prove that (3.5.4) is also true for $k=n+1$. Indeed, assume that g satisfies Assumption $\Psi(n+1)$. Then, with

$$S_j(a,\lambda) := \sum_{\ell=0}^{j} e^{-i\lambda\ell/a} = \frac{1}{2i\sin(\lambda/2a)}\big(e^{i\lambda/2a}-e^{-i\lambda/2a}e^{-ij\lambda/a}\big) \quad \textit{for} \quad j\in\{0,1,\dots,a\},$$

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}\Big| &= \Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)\big(S_j(a,\lambda)-S_{j-1}(a,\lambda)\big)\Big| \\ &\le I_a(\lambda)+\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big| \quad \textit{with} \quad I_a(\lambda):=\Big|\frac{1}{a}\sum_{j=1}^{a-1}\big(g\big(\frac{j}{a}\big)-g\big(\frac{j+1}{a}\big)\big)S_j(a,\lambda)\Big|. \end{aligned} \quad (3.5.5)$$

But since g satisfies Assumption $\Psi(n+1)$ and $a\ge 1$,

$$\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big| \le \sup_{x\in[0,1]}|g^{(n+1)}(x)|\,\frac{1}{a^{n+1}(n+1)!}. \quad (3.5.6)$$

Now, with the usual Taylor expansion

$$\big|g\big(\frac{j+1}{a}\big)-g\big(\frac{j}{a}\big)-\sum_{k=1}^{n}\frac{1}{a^k k!}g^{(k)}\big(\frac{j}{a}\big)\big| \le \frac{1}{a^{n+1}(n+1)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)| \quad \textit{for} \quad j\in\{0,1,\dots,a-1\}.$$

Therefore,

$$I_a(\lambda) \le \sum_{k=1}^{n}\frac{1}{a^k k!}\Big|\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)S_j(a,\lambda)\Big| + \frac{1}{a^{n+1}(n+1)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)|.$$

From the definition of $S_j(a,\lambda)$ and with the inequality $\frac{2}{\pi}u\le\sin(u)\le u$ for $u\in[0,\pi/2]$, we have for $\lambda\in[-a\pi,0)\cup(0,a\pi]$ and $k\in\{1,\dots,n\}$:

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)S_j(a,\lambda)\Big| &\le \frac{1}{2|\sin(\lambda/2a)|}\Big(\Big|\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)\Big|+\Big|\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}\Big|\Big) \\ &\le \frac{\pi a}{2|\lambda|}\Big(\frac{1}{a^{n+1-k}(n+1-k)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)|+D_{g^{(k)}}(n+1-k)\frac{1}{|\lambda|^{n+1-k}}\Big),\end{aligned}$$

using (3.5.3) for bounding $\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)$ *and the induction hypothesis for bounding* $\frac{1}{a}\sum_{j=1}^{a-1} g^{(k)}\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}$. *Hence, with (3.5.6),*

$$\begin{aligned} I_a(\lambda)+\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big| &\le \frac{1}{a^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=0}^{n+1}\frac{1}{(n+1-k)!\,k!} \\ &+ \frac{\pi a}{2|\lambda|}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n}\frac{10^{n+1-k}}{a^k k!}\frac{1}{|\lambda|^{n+1-k}} \qquad (3.5.7)\\ &\le \frac{(2\pi)^{n+1}}{(n+1)!\,|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|+\frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n}\frac{1}{k!}\big(\frac{\pi}{10}\big)^k \qquad (3.5.8)\\ &\le \frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n+1}\frac{1}{k!}\big(\frac{\pi}{5}\big)^k \\ &\le \frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\,\big(e^{\pi/5}-1\big), \qquad (3.5.9)\end{aligned}$$

since $a^{-k}\le\pi^k\,|\lambda|^{-k}$ *for all* $\lambda\in[-a\pi,0)\cup(0,a\pi]$ *and* $k\in\{0,1,\dots,n+1\}$. *Thus since* $e^{\pi/5}-1<1$ *and from (3.5.5) and (3.5.9), we deduce that (3.5.4) is true for* $k=n+1$ *and therefore for any* $k\in\mathbb{N}$.

Proof 3.3 (Proof of Property 3.1) *First, since* $(X_t)_{t\in\mathbb{Z}}$ *is a stationary centered linear process,* $e(a,b)=\sum_{j=1}^{a}\Big(\frac{1}{\sqrt{a}}\,\psi(\frac{j}{a})\Big)X_{b+j}$ *for any* $(a,b)\in\mathbb{N}^*\times\mathbb{Z}$ *from (3.2.2) and* $\sum_{j=1}^{a}\frac{1}{\sqrt{a}}\big|\psi\big(\frac{j}{a}\big)\big|<\infty$, *it is clear that for* $a\in\mathbb{N}^*$, $(e(a,b))_{b\in\mathbb{Z}}$ *is a stationary centered linear process.*

Now following similar computations to those performed in Bardet et al. (2008) [Proof of Property 1], we obtain with f *the spectral density of* X *and for* $a\in\mathbb{N}^*$,

$$\mathbb{E}(e^2(a,0)) = \int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big)\times\Big|\frac{1}{a}\sum_{j=1}^{a}\psi\big(\frac{j}{a}\big)e^{-i\frac{j}{a}u}\Big|^2\,du.$$

Now, since ψ *satisfies Assumption* $\Psi(k)$ *and therefore (3.2.1), from Lemma 3.1, for* $u\in[-\sqrt{a},\sqrt{a}]$ *and* a *large enough,*

$$\begin{aligned}\Big|\big|\frac{1}{a}\sum_{j=1}^{a}\psi\big(\frac{j}{a}\big)e^{-i\frac{j}{a}u}\big|^2-|\widehat{\psi}(u)|^2\Big| &\le 2C_\psi(k)\frac{|u|^k}{a^k}\,|\widehat{\psi}(u)|+C_\psi^2(k)\frac{|u|^{2k}}{a^{2k}} \\ &\le \Big(2C_\psi(k)\sup_{x\in[0,1]}|\psi^{(k)}(x)|+C_\psi^2(k)\Big)\,\frac{1}{a^k}. \qquad (3.5.10)\end{aligned}$$

Moreover, for $|u| \in [\sqrt{a}\,,\, a\pi]$*, from Lemma 3.2 and* $a \in \mathbb{N}^*$*, we have,*

$$\Big|\frac{1}{a}\sum_{j=1}^{a}\psi\big(\frac{j}{a}\big)e^{-i\frac{j}{a}u}\Big|^2 \;\leq\; D^2_\psi(k)\frac{1}{|u|^{2k}}, \tag{3.5.11}$$

Now, using (3.5.10) and (3.5.11), since there exists $c_f > 0$ *satisfying* $f(\lambda) \leq c_f|\lambda|^{-2d}$ *for all* $\lambda \in [-\pi,\pi]$*, we deduce with* $F_\psi(k) = 2C_\psi(k)\sup_{x\in[0,1]}|\psi^{(k)}(x)| + C^2_\psi(k)$ *and for all* $d < 1/2$*,*

$$\begin{aligned}\Big|\mathbb{E}(e^2(a,0)) - \int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)\times|\widehat{\psi}(u)|^2\,du\Big| &\leq \frac{F_\psi(k)}{a^k}\int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)du + 2D^2_\psi(k)\int_{\sqrt{a}}^{a\pi}\frac{1}{|u|^{2k}}f\big(\frac{u}{a}\big)\,du \\ &\leq a^{2d}\Big(\frac{2c_f\,F_\psi(k)}{1-2d} + \frac{2D^2_\psi(k)}{2k+2d-1}\Big)\frac{1}{a^{k+d-1/2}}. \qquad (3.5.12)\end{aligned}$$

Now, using again (3.2.1), for a large enough,

$$\begin{aligned}\Big|\int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du - \int_{-\infty}^{\infty} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du\Big| &\leq \big(2c_f\sup_{x\in[0,1]}|\psi^{(k)}(x)|\big)a^{2d}\int_{\sqrt{a}}^{\infty}\frac{1}{u^{2d+2k}}\,du \\ &\leq a^{2d}\Big(\frac{2c_f\sup_{x\in[0,1]}|\psi^{(k)}(x)|}{2k+2d-1}\Big)\frac{1}{a^{k+d-1/2}}. \qquad (3.5.13)\end{aligned}$$

Finally, from Assumption $A(d,d')$ *and using the definition (3.2.4) of* $K_{(\psi,\alpha)}$*, we obtain the following expansion :*

$$\begin{aligned}\int_{-\infty}^{\infty} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du &= 2\pi\int_{-\infty}^{\infty}\Big(c_d\Big|\frac{u}{a}\Big|^{-2d} + c_{d'}\Big|\frac{u}{a}\Big|^{d'-2d} + \Big|\frac{u}{a}\Big|^{d'-2d}\varepsilon\big(\frac{u}{a}\big)\Big)\,|\widehat{\psi}(u)|^2\,du \\ &= 2\pi\,c_d\,K_{(\psi,2d)}\,a^{2d} + 2\pi\,c_{d'}\,K_{(\psi,2d-d')}\,a^{2d-d'} + o(a^{2d-d'}) \qquad (3.5.14)\end{aligned}$$

because $\lim_{\lambda\to 0}\varepsilon(\lambda) = 0$ *and applying Lebesgue Theorem. Then, using (3.5.12), (3.5.13) and (3.5.14), we obtain that there exists* C *only depending on* ψ *and* k *such as for a large enough,*

$$\Big|\mathbb{E}(e^2(a,0)) - 2\pi\,c_d\,K_{(\psi,2d)}\,a^{2d} - 2\pi\,c_{d'}\,K_{(\psi,2d-d')}\,a^{2d-d'}\Big| \;\leq\; a^{2d}\left(C\,a^{-k-d+1/2} + o(a^{-d'})\right). \tag{3.5.15}$$

When $k > d' - d + 1/2$ *implying* $k + d - 1/2 > d'$*, then (3.2.3) holds.*

Proof 3.4 (Proof of Theorem 3.1) *We decompose this proof in 4 steps. First define the normalized wavelet coefficients of X by :*

$$\widetilde{e}_N(a,b) := \frac{e(a,b)}{\sqrt{\mathbb{E}(e^2(a,0))}} \quad \text{for } a \in \mathbb{N}^* \text{ and } b \in \mathbb{Z}, \tag{3.5.16}$$

and the normalized sample variance of wavelet coefficients by :

$$\widetilde{T}_N(a) := \frac{1}{N-a}\sum_{k=1}^{N-a} \widetilde{e}^2(a,k). \tag{3.5.17}$$

Step 1 *We prove in this part that $\big(N\, Cov(\widetilde{T}_N(r_i a_N), \widetilde{T}_N(r_j a_N)\big)_{1\le i,j\le \ell}$ converges to the asymptotic covariance matrix $\Gamma(r_1,\dots,r_\ell,\psi,d)$ defined in (3.2.8). First for $\lambda \in \mathbb{R}$, denote*

$$S_a(\lambda) := \frac{1}{a}\sum_{t=1}^{a} \psi(\frac{t}{a}) e^{i\lambda t/a}.$$

Then for $a \in \mathbb{N}^*$ and $b = 1,\cdots,N-a$, since ψ is $[0,1]$-supported function and $\widehat{\alpha} \in \mathbb{L}^2([-\pi,\pi])$ inducing $\alpha(k) = \int_{-\pi}^{\pi} \widehat{\alpha}(\lambda) e^{ik\lambda} d\lambda$,

$$\begin{aligned} \sum_{t=1}^{N} \alpha(t-s)\psi(\frac{t-b}{a}) &= \sum_{t=0}^{a} \psi(\frac{t}{a}) \int_{-\pi}^{\pi} \widehat{\alpha}(\lambda) e^{i\lambda(t-s+b)} d\lambda \\ &= \int_{-\pi}^{\pi} a S_a(a\lambda)\widehat{\alpha}(\lambda) e^{i(b-s)\lambda} d\lambda \\ &= \int_{-a\pi}^{a\pi} S_a(\lambda)\widehat{\alpha}(\frac{\lambda}{a}) e^{i(b-s)\frac{\lambda}{a}} d\lambda. \end{aligned} \tag{3.5.18}$$

But, for $a, a' \in \mathbb{N}^*$,

$$\begin{aligned} Cov(\widetilde{T}_N(a), \widetilde{T}_N(a')) &= \frac{1}{N-a}\frac{1}{N-a'} \sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} Cov(\widetilde{e}^2(a,b), \widetilde{e}^2(a',b')) \\ &= \frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{4\pi^2(N-a)(N-a')} \sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} Cov(e^2(a,b), e^2(a',b')). \end{aligned} \tag{3.5.19}$$

Now,

$$\begin{aligned} Cov(e^2_{(a,b)}, e^2_{(a',b')}) = \frac{1}{a\,a'} & \sum_{t_1,t_2,t_3,t_4=1}^{N} \sum_{s_1,s_2,s_3,s_4 \in \mathbb{Z}} \Big(\prod_{i=1}^{2} \alpha(t_i - s_i)\psi(\frac{t_i - b}{a})\Big) \\ & \times \Big(\prod_{i=1}^{2} \alpha(t_i - s_i)\psi(\frac{t_i - b'}{a'})\Big) Cov\big(\xi_{s_1}\xi_{s_2}, \xi_{s_3}\xi_{s_4}\big) \qquad (3.5.20) \\ &= C_1 + C_2, \qquad (3.5.21) \end{aligned}$$

since there are only two nonvanishing cases : $s_1 = s_2 = s_3 = s_4$ (Case 1 => C_1), $s_1 = s_3 \neq s_2 = s_4$ and $s_1 = s_4 \neq s_2 = s_3$ (Case 2 => C_2).

** Case 1 : in such a case, $Cov\big(\xi_{s_1}\xi_{s_2}, \xi_{s_3}\xi_{s_4}\big) = \mu_4 - 1$ and*

$$C_1 = \frac{\mu_4 - 1}{a\,a'} \sum_{s\in\mathbb{Z}} \Big|\sum_{t=1}^{N} \alpha(t-s)\psi(\frac{t-b}{a})\Big|^2 \Big|\sum_{t=1}^{N} \alpha(t-s)\psi(\frac{t-b'}{a'})\Big|^2$$

$$C_1 = (\mu_4 - 1)\,a\,a' \lim_{M\to\infty} \int_{[-\pi,\pi]^4} d\lambda d\lambda' d\mu d\mu' e^{i[b(\lambda-\lambda')+b'(\mu-\mu')]} \times \sum_{s=-M}^{M} e^{is[(\lambda-\lambda')+(\mu-\mu')]} S_a(a\lambda)\widehat{\alpha}(\lambda)\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')} S_{a'}(a'\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}$$

using the relation (3.5.18), with $\overline{z}$ denoting the conjugate of $z \in \mathbb{C}$. From the usual asymptotic behavior of Dirichlet kernel, for g a 2π-periodic function such as $g \in \mathcal{C}^1((-\pi,\pi))$, $\lim_{M\to\infty} \int_{-\pi}^{\pi} D_M(z)g(x+z)dz = g(x)$ uniformly in x with

$$D_M(z) := \frac{1}{2\pi} \sum_{k=-M}^{M} e^{ikz} = \frac{1}{2\pi} \frac{\sin\big((2M+1)z/2\big)}{\sin\big(z/2\big)}. \qquad (3.5.22)$$

Thus with $h : \mathbb{R}^4 \mapsto \mathbb{R}$ a continuously differentiable function 2π-periodic for each component,

$$\lim_{M\to\infty} \int_{[-\pi,\pi]^4} 2\pi\, D_M((\lambda-\lambda')+(\mu-\mu'))h(\lambda,\lambda',\mu,\mu')d\lambda d\lambda' d\mu d\mu' = 2\pi \int_{[-\pi,\pi]^3} h(\lambda'-\mu+\mu',\lambda',\mu,\mu')d\lambda' d\mu d\mu';$$

Therefore,

$$C_1 = 2\pi\,(\mu_4 - 1)\,a\,a' \int_{[-\pi,\pi]^3} d\lambda' d\mu d\mu' e^{i(\mu-\mu')(b'-b)} \times S_a(a(\lambda'-\mu+\mu'))\widehat{\alpha}(\lambda'-\mu+\mu')\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')} S_{a'}(a'\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}. \qquad (3.5.23)$$

* *Case 2 : in such a case, with* $s_1 \neq s_2$, $Cov\big(\xi_{s_1}\xi_{s_2}, \xi_{s_1}\xi_{s_2}\big) = 1$ *and*

$$\begin{aligned}
C_2 = & \frac{2}{a\,a'} \sum_{(s,s')\in\mathbb{Z}^2, s\neq s'} \sum_{t_1=1}^{N} \alpha(t_1-s)\psi(\frac{t_1-b}{a}) \sum_{t_2=1}^{N} \alpha(t_2-s)\psi(\frac{t_2-b'}{a'}) \\
& \qquad \times \sum_{t_3=1}^{N} \alpha(t_3-s')\psi(\frac{t_3-b}{a}) \sum_{t_4=1}^{N} \alpha(t_4-s')\psi(\frac{t_4-b'}{a'}) \\
= & -\frac{2C_1}{\mu_4-1} + \frac{1}{a\,a'} \sum_{(s,s')\in\mathbb{Z}^2} \sum_{t_1=1}^{N} \alpha(t_1-s)\psi(\frac{t_1-b}{a}) \sum_{t_2=1}^{N} \alpha(t_2-s)\psi(\frac{t_2-b'}{a'}) \\
& \times \qquad \sum_{t_3=1}^{N} \alpha(t_3-s')\psi(\frac{t_3-b}{a}) \sum_{t_4=1}^{N} \alpha(t_4-s')\psi(\frac{t_4-b'}{a'}) \\
C_2 = & -\frac{2C_1}{\mu_4-1} + 2\,a\,a' \lim_{M\to\infty} \lim_{M'\to\infty} \int_{[-\pi,\pi]^4} d\lambda d\lambda' d\mu d\mu' e^{i[b(\lambda-\mu)-b'(\lambda'-\mu')]} \\
& \times \sum_{s=-M}^{M} \sum_{s=-M'}^{M'} e^{is(\lambda'-\lambda)+is'(\mu'-\mu)} S_a(a\lambda)\widehat{\alpha}(\lambda)\overline{S_{a'}(a'\lambda')\widehat{\alpha}(\lambda')} S_a(a\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')} \\
= & -\frac{2C_1}{\mu_4-1} + 8\pi^2\,a\,a' \int_{[-\pi,\pi]^2} e^{i(\lambda-\mu)(b-b')} S_a(a\lambda)\overline{S_{a'}(a'\lambda)} S_a(a\mu)\overline{S_{a'}(a'\mu)} \times \big|\widehat{\alpha}(\lambda)\big|^2 \big|\widehat{\alpha}(\mu)\big|^2 d\lambda d\mu,
\end{aligned}$$

using the asymptotic behaviors of two Dirichlet kernels.

Now we have to compute $\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}(C_1+C_2)$. *In both cases* ($C_1$ *and* C_2), *one again obtains a function of a Dirichlet kernel :*

$$F_N(a,a',v) := \sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} e^{i\,v\,(b-b')} = e^{iv(a-a')/2} \frac{\sin((N-a)v/2)\,\sin((N-a')v/2)}{\sin^2(v/2)}. \qquad (3.5.24)$$

For a continuous function $h : [-\pi,\pi] \mapsto \mathbb{R}$,

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\int_{-\pi}^{\pi} h(v)F_N(a,a',v)dv &= \lim_{N\to\infty}\frac{1}{N^2}\int_{-\pi N}^{\pi N} h(\frac{v}{N})F_N(a,a',\frac{v}{N})dv \\
&= 4h(0)\int_{-\infty}^{\infty}\frac{\sin^2(v/2)}{v^2}dv = 2\pi h(0),
\end{aligned}$$

thanks to Lebesgue Theorem and with $a/N \to 0$ *($N \to 0$). Then, from (3.5.23),*

$$N \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1 \sim 4\pi^2 (\mu_4 - 1) a a' \int_{[-\pi,\pi]^2} d\lambda' d\mu' |S_a(a\lambda')|^2 \, |S_{a'}(a'\mu')|^2 \, |\widehat{\alpha}(\lambda')|^2 |\widehat{\alpha}(\mu')|^2$$

$$\sim 4\pi^2 (\mu_4 - 1) \int_{-a\pi}^{a\pi} |S_a(\lambda)|^2 \, |\widehat{\alpha}(\lambda/a)|^2 d\lambda \int_{-a'\pi}^{a'\pi} |S_a(\mu)|^2 \, |\widehat{\alpha}(\mu/a')|^2 d\mu$$

$$\Longrightarrow N \frac{(\mathbb{E}(e^2(a,0)))^{-1} \mathbb{E}(e^2(a',0)))^{-1}}{4\pi^2 (N-a)(N-a')} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1 \underset{N\to\infty}{\longrightarrow} (\mu_4 - 1) \qquad (3.5.25)$$

and therefore

$$\frac{N}{a_N} \frac{(r a_N r' a_N)^{-2d} (c_d K_{(\psi,2d)})^{-2}}{4\pi^2 (N - r a_N)(N - r' a_N)} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_1 \underset{N\to\infty}{\longrightarrow} 0, \qquad (3.5.26)$$

with $a = ra_N$ *and* $a' = r'a_N$*, using the same arguments than in Property 3.1 since* $a_N \to \infty$.

Moreover, always with $a_N \to \infty$ *and* $N/a_N \to \infty$,

$$N \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_2 \sim 16\pi^3 a a' \int_{-\pi}^{\pi} \big|S_a(a\lambda)\big|^2 \big|S_{a'}(a'\lambda)\big|^2 \big|\widehat{\alpha}(\lambda)\big|^4 d\lambda$$

$$- \frac{2N}{\mu_4 - 1} \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1$$

$$\sim 16\pi^3 r r' a_N \int_{-a_N\pi}^{a_N\pi} \big|S_{ra_N}(r\lambda)\big|^2 \big|S_{r'a_N}(r'\lambda)\big|^2 \big|\widehat{\alpha}(\lambda/a_N)\big|^4 d\lambda$$

$$- \frac{2N}{\mu_4 - 1} \frac{1}{N - ra_N} \frac{1}{N - r'a_N} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_1$$

$$\Longrightarrow \frac{N}{a_N} \frac{(r\, r'\, a_N^2)^{-2d} (c_d K_{(\psi,2d)})^{-2}}{4\pi^2 (N - ra_N)(N - r'a_N)} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_2 \underset{N\to\infty}{\longrightarrow} 4\pi \frac{(rr')^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r\lambda)\big|^2 \big|\widehat{\psi}(r'\lambda)\big|^2}{\lambda^{4d}} d\lambda,$$

always using the same trick than in Property 3.1 since $a_N \to \infty$ *and* $N/a_N \to \infty$. *Therefore, with (3.5.26), one deduces that :*

$$\frac{N}{a_N} \, Cov(\widetilde{T}_N(r\, a_N), \widetilde{T}_N(r'\, a_N)) \underset{N\to\infty}{\longrightarrow} 4\pi \frac{(rr')^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r\lambda)\big|^2 \big|\widehat{\psi}(r'\lambda)\big|^2}{\lambda^{4d}} d\lambda. \quad (3.5.27)$$

Note that if $r = r'$ *then* $\dfrac{N}{r\, a_N} \, Var(\widetilde{T}_N(r\, a_N)) \underset{N\to\infty}{\longrightarrow} \sigma^2_\psi(d) = 64\pi^5 \dfrac{K_{(\psi*\psi,4d)}}{K^2_{(\psi,2d)}}$ *only depending on* ψ *and* d.

Step 2 *We prove here that if the distribution of the innovations* $(\xi_t)_t$ *is such that there exists* $r>0$ *satisfying* $\mathbb{E}\big(e^{r\xi_0}\big)\leq\infty$ *(condition so-called the Cramèr condition), then for any* $a\in\mathbb{N}^*$, $(\widetilde{T}_N(r_i\,a_N))_{1\leq i\leq \ell}=\Big(\frac{1}{N-r_ia_N}\sum_{k=1}^{N-r_ia_N}\widetilde{e}^2(r_ia_N,k)\Big)_{1\leq i\leq \ell}$ *satisfies a central limit theorem. Such theorem is implied by proving* $\sqrt{\frac{N}{a_N}}\sum_{i=1}^{\ell}\frac{u_i}{N-r_ia_N}\sum_{k=1}^{N-r_ia_N}\widetilde{e}^2(r_ia_N,k)$ *asymptotically follows a Gaussian distribution for any vector* $(u_i)_{1\leq i\leq \ell}\in\mathbb{R}^\ell$.

For establishing this result we are going to adapt a proof of Giraitis (1985) where central limit theorems for function of linear processes are proved using a decomposition with Appell polynomials. Indeed since X *satisfies Assumption* $A(d,d')$ *and can be a two-sided linear process, martingale type results as in Wu (2002) or Furmanczyk (2007) can not be applied. Moreover, since* $(a_N)_N$ *is a sequence depending on* N *it is required to prove a central limit theorem for triangular arrays. Unfortunately the recent paper of Roueff and Taqqu (2009) dealing with central limit theorems for arrays of decimated linear processes, and which can be applied to establish a multidimensional central limit for the variogram of wavelet coefficients associated to a multi-resolution analysis can not be applied here because in this paper this variogram is defined as in (3.2.6) with coefficients taken every* n/n_j $(\simeq a_N$ *with our notation) and the mean of* n_j $(N/a_N$ *with our notation) coefficients is considered (with a convergence rate* $\sqrt{n_j}$*). Our definition of the wavelet coefficient variogram (3.2.5) is an average of* $N-a_N$ *terms and the convergence rate is* N/a_N. *Then we chose to adapt the method and results of Giraitis (1985).*

More precisely, consider the case $\ell=1$. *For* $a>0$, $(\widetilde{e}(a,b))_{1\leq b\leq N-a}$ *is a stationary linear process satisfying assumptions of the paper of Giraitis (called* X_t *in this article). Now we consider* $H_2(x)=x^2-1$ *the second-order Hermite polynomial and we would like to prove* *that*

$$\Big(\frac{N}{a_N}\Big)^{1/2}\frac{1}{N-a_N}\sum_{b=1}^{N-a_N}\big(\widetilde e^2(a_N,b)-1\big) \;\simeq\; \Big(\frac{1}{Na_N}\Big)^{-1/2}\sum_{b=1}^{N-a_N} H_2(\widetilde e(a_N,b)) \;\xrightarrow[N\to\infty]{\mathcal{L}}\; \mathcal{N}\big(0,\sigma_\psi^2(d)\big). \quad (3.5.28)$$

Now since the distribution of ξ_0 is supposed to satisfy the Cramèr condition, following the proof of Proposition 6 (Giraitis, 1985), define $S_N^{(n)}=\sum_{b=1}^{N-a_N}A_n^{(a_N)}(\widetilde e(a_N,b))$ where $A_n^{(a_N)}$ is the Appell polynomial of degree n corresponding to the probability distribution of $\widetilde e(a_N,\cdot)$. We are going to prove that the cumulants of order $k\geq 3$ are such as

$$\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=o\big((Na_N)^{k/2}\big) \quad (3.5.29)$$

for any $n(1),\cdots,n(k)\geq 2$ (the computation of the cumulants of order 2 is induced by Step 1 of this proof) and then (3.5.28) holds. Indeed, $\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=\sum_{\gamma\in\Gamma_0(T)}d_\gamma I_\gamma(N)$ where $\Gamma_0(T)$ is the set of possible diagrams and the definition of $I_\gamma(N)$ is provided in (34) of Giraitis (1985).

In the case of Gaussian diagrams, $I_\gamma(N)=o\big((Na_N)^{k/2}\big)$, since this case is induced by the Gaussian case and the second order moments.

If γ is a non-Gaussian diagram, mutatis mutandis, *we are going to follow the notation and proof of Lemma 2 of Giraitis (1985). Note first from Step 1, we can write :*

$$\widetilde e(a,b)=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s \quad \textit{with} \quad \beta_a(s)=\frac{\sqrt{a}}{\sqrt{\mathbb{E}e^2(a,b)}}\int_{-\pi}^{\pi}S_a(a\lambda)\widehat\alpha(\lambda)e^{i\lambda s}d\lambda. \quad (3.5.30)$$

Then for $u\in[-\pi,\pi]$,

$$\begin{aligned}
\widehat\beta_a(u) &= \frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\beta_a(s)e^{-isu}\\
&= \frac{\sqrt{a}}{2\pi\sqrt{\mathbb{E}e^2(a,b)}}\lim_{m\to\infty}\int_{-\pi}^{\pi}\sum_{s=-m}^{m}S_a(a\lambda)\widehat\alpha(\lambda)e^{is(\lambda-u)}d\lambda\\
&= \frac{\sqrt{a}}{\sqrt{\mathbb{E}e^2(a,b)}}S_a(au)\widehat\alpha(u),
\end{aligned}$$

with the asymptotic behavior of Dirichlet kernel. Now, in the case a/ of Lemma 2 of Giraitis (1985), consider the diagram $V_1 = \{(1,1),(2,1),(3,1)\}$ *and assume that for the rows* L_j *of the array* T, $j = 1, \cdots, k$ $(k \geq 3)$, $|V_1 \cap L_j| \geq 1$ *for at least* 3 *different rows* L_j. *Then the inequality (39) can be repeated, and on the hyperplane* x_{V_1}, *a part of the integral (34) provides*

$$\Big|\int_{\{x_{11}+x_{21}+x_{31}=0\}\cap[-\pi,\pi]^3} dx_{11}dx_{21}dx_{31}\prod_{j=1}^{3} D_N((x_{j1}+\cdots+x_{jn(j)})\widehat{\beta}_a(x_{j1})\Big| \leq C\,\alpha_1(u_1)\alpha_2(u_2)\alpha_3(u_3),$$

with $u_i = x_{i2}+\cdots+x_{in(i)}$ *and the same expressions of* α_i *provided in Giraitis (1985). It remains to bound* $\alpha_i(u)$. *But, with the same approximations as in the proof of Property 3.1, for* a_N *and* N *large enough*

$$\begin{aligned}
\alpha_1(u) &= \int_{-\pi}^{\pi}\big|\widehat{\beta}_{a_N}(u)\,D_N(x+u)\big|dx \sim \sqrt{2\pi}\,\frac{1}{\sqrt{a_N}}\int_{-a_N\pi}^{a_N\pi}\Big|\frac{\widehat{\psi}(x)}{|x|^d}\Big|\,\Big|D_N\big(\frac{x}{a_N}+u\big)\Big|du \\
&\leq 2\sqrt{a_N}\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|}{|x|^d}\Big\}\int_{-\pi}^{\pi}|D_N(x+u)|dx \\
&\leq 2C\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|}{|x|^d}\Big\}\sqrt{a_N}\,\log N,
\end{aligned}$$

since there exists $C>0$ *such as* $\int_{-\pi}^{\pi}|D_N(x+u)|dx \leq C\log N$ *for any* $u \in [-\pi,\pi]$. *Now for* $i = 2,3$, a_N *and* N *large enough,*

$$\begin{aligned}
\alpha_i^2(u) &= \|\widehat{\beta}_{a_N}(\cdot)\,D_N(u+\cdot)\|_2^2 \\
&\leq 2\int_{-a_N\pi}^{a_N\pi}\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}D_N^2\big(\frac{x}{a_N}+u\big)du \\
&\leq 2C\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\Big\}\,a_N\int_{-\pi}^{\pi}|D_N^2(x+u)|dx \\
&\leq C'\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\Big\}\,Na_N.
\end{aligned}$$

Then $\alpha_1(u_1)\alpha_2(u_2)\alpha_3(u_3) = o((Na_N)^{3/2})$.

For the $k-3$ *other terms, a result corresponding to Lemma 1 of Giraitis (1985) can also* *be obtained. Indeed, for* a_N *and* N *large enough,*

$$\begin{aligned}
\|g_{N,j}\|_2^2 &= \int_{[-\pi,\pi]^{n(j)}} dx\, D_N^2(x_1+\cdots+x_{n(j)}) \prod_{i=1}^{n(j)} |\widehat{\beta}_{a_N}(x_i)|^2 \\
&\leq C \int_{[-a_N\pi,a_N\pi]^{n(j)}} dx\, D_N^2(\frac{1}{a_N}(x_1+\cdots+x_{n(j)})) \prod_{i=1}^{n(j)} \frac{|\widehat{\psi}(x_i)|^2}{|x_i|^{2d}} \\
&\leq C \,\big|\sup_{x\in\mathbb{R}} \{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\}\big|^{n(j)} a_N \, \|D_N(\,\cdot\,)\|_2^2 \\
&\leq C'\, N a_N
\end{aligned}$$

with $C' \geq 0$ *not depending on* N *and* a_N*. Thus* $\|g_{N,j}\|_2 \leq C\,(Na_N)^{1/2}$ *with* $C \geq 0$*. Using the same reasoning, there also exists* $C' \geq 0$ *such as* $\|g'_{N,j}\|_2 \leq C\,(Na_N)^{1/2}$ *for* $j \geq 2$ *while* $\|g'_{N,1}\|_2 = O(\sqrt{a_N}\,\log N) = o((Na_N)^{1/2})$*. As a consequence, for* γ *such as* $|V_1 \cap L_j| \geq 1$ *for at least 3 different rows* L_j*, and more generally with* $|V_1| \geq 3$*,*

$$I_\gamma(N) = o\big((Na_N)^{k/2}\big). \qquad (3.5.31)$$

For other γ*, it remains to bound the function* $h(u_1,u_2)$ *defined in Giraitis (1985, p. 32) as follows (with* $x = x_{11}+x_{12}$*) and with* $u_1+u_2 \neq 0$ *:*

$$\begin{aligned}
h(u_1,u_2) &= \Big(\int_{-\pi}^{\pi} |\widehat{\beta}_{a_N}(-x)\, D_N(u_1+x)\, D_N(u_2-x)\big|dx\Big)\Big(\int_{-\pi}^{\pi} |\widehat{\beta}_{a_N}(x)\big|^2 dx\Big) \\
&\leq \big|\sup_{x\in\mathbb{R}} \{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\}\big|\, a_N \Big(\int_{-\pi}^{\pi} \big|D_N(u_1+x)\, D_N(u_2-x)\big|dx\Big)\Big(2\pi \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\, dx\Big).
\end{aligned}$$

But

$$\begin{aligned}
\int_{-\pi}^{\pi} \big|D_N(u_1+x)\, D_N(u_2-x)\big|dx &\leq 2\int_{-2\pi N}^{2\pi N} \Big|\frac{\sin(x)}{x}\,\frac{\sin(\frac{N}{2}(u_1+u_2)-x)}{\sin(\frac{1}{2}(u_1+u_2)-\frac{x}{N})}\Big|dx \\
&\leq \begin{cases} C\,\log N \big|\sin(\frac{1}{2}(u_1+u_2))\big|^{-1} & \textit{if } |u_1+u_2| \geq (N\log N)^{-1} \\ C\,N & \textit{if } |u_1+u_2| < (N\log N)^{-1} \end{cases}.
\end{aligned}$$

Therefore,

$$\|h(u_1,u_2)\|_2^2 = \int_{[-\pi,\pi]^2} h^2(u_1,u_2)du_1du_2 \;\leq\; C\,a_N^2\Big(\log^2 N\int_{(N\log N)^{-1}}^{\pi}(\sin x)^{-2}\,dx + N^2\int_0^{(N\log N)^{-1}} dx\Big) \;\leq\; C\,a_N^2\big(N\log^3 N + N\log N\big),$$

and hence $\|h(u_1,u_2)\|_2 = o(Na_N)$. *Finally, (3.5.31) holds for all* γ *and it implies (3.5.29) and therefore (3.5.28).*

If $\ell > 1$, *the same proof can be repeated from the linearity properties of cumulants. Thus,* $(\widetilde{T}_N(r_i\,a_N))_{1\leq i\leq \ell}$ *satisfies the following central limit :*

$$\sqrt{\frac{N}{a_N}}\big(\widetilde{T}_N(r_i\,a_N)-1\big)_{1\leq i\leq \ell}\xrightarrow[N\to\infty]{D}\mathcal{N}\big(0\,,\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big), \tag{3.5.32}$$

with $\Gamma(r_1,\ldots,r_\ell,\psi,d) = (\gamma_{ij})_{1\leq i,j\leq \ell}$ *given in (3.2.8).*

Step 3 *Now we extend the central limit obtained in Step 2 for linear processes with an innovation distribution satisfying a Cramèr condition* ($\mathbb{E}\big(e^{r\xi_0}\big) < \infty$) *to the weaker condition* $\mathbb{E}\xi_0^4 < \infty$ *using a truncation procedure. Thus assume now that* $\mathbb{E}\xi_0^4 < \infty$. *Let* $M > 0$ *and define* $\xi_t^- = \xi_t\,\mathbb{I}_{|\xi|\leq M}$ *and* $\xi_t^+ = \xi_t\,\mathbb{I}_{|\xi|>M}$, $\widetilde{e}^-(a,b) = \sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^-$ *and* $\widetilde{e}^+(a,b) = \sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^+$ *using (3.5.30). Clearly* $\widetilde{e}(a,b) = \widetilde{e}^+(a,b)+\widetilde{e}^-(a,b)$. *We are going to prove that (3.5.32) holds. For this, we begin by writing*

$$\widetilde{T}_N(r_i\,a_N)-1 \;=\; \frac{1}{N-r_ia_N}\Big(\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1 \tag{3.5.33}$$

$$-2\widetilde{e}^+(r_ia_N,b)\widetilde{e}^-(r_ia_N,b)+\big(\widetilde{e}^+(r_ia_N,b)\big)^2\Big) \tag{3.5.34}$$

We first prove that $\big(\widetilde{T}_N^-(r_i\,a_N)-1\big)_{1\leq i\leq\ell} = \big(\frac{1}{N-r_ia_N}\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1\big)_{1\leq i\leq\ell}$ *also satisfies (3.5.32). Indeed,* $(\widetilde{e}^-(r_ia_N,b))$ *is a linear process with innovations* (ξ_t^-) *satisfying the Cramèr condition and it is obvious that* $\Big(\frac{\mathbb{E}\big(\widetilde{e}(r_ia_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}^-(r_ia_N,b)\big)^2}\Big)^{1/2}\widetilde{e}^-(r_ia_N,b)_{b,i}$

has exactly the same distribution than $\widetilde{e}(r_i a_N, b)_{b,i}$. *Therefore it remains to prove that* $\sqrt{\frac{N}{a_N}}\Big(\frac{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}-1\Big)$ *converges to* 0. *We have* $\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2 = \big(\sum_{s\in\mathbb{Z}} \beta_a^2(s)\big)\,\mathbb{E}(\xi_0)^2 = 1$ *and* $\mathbb{E}\xi_0^2 = 1$ *(from Property 3.1). Then*

$$\Big|\frac{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}-1\Big| \;\le\; 2\left(\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2\right)^{1/2} + \mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2.$$

We have $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2 = \big(\sum_{s\in\mathbb{Z}} \beta_a^2(s)\big)\,\mathbb{E}(\xi_0^+)^2 = \mathbb{E}(\xi_0^+)^2$ *from previous arguments and since we assume that the distribution of* ξ_0 *is symmetric. But using Hölder's and Markov's inequalities :*

$$\mathbb{E}(\xi_0^+)^2 \le (\mathbb{E}\xi_0^4)^{1/2}(\Pr(|\xi_0|>M))^{1/2} \le (\mathbb{E}\xi_0^4)\,M^{-2}.$$

Hence, there exists $C>0$ *not depending on* M *and* N,

$$\sqrt{\frac{N}{a_N}}\Big|\frac{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}-1\Big| \le \frac{C}{M}\sqrt{N}a_N \underset{N\to\infty}{\longrightarrow} 0$$

when $M=N$ *(for instance). Therefore* $\big(\widetilde{T}_N^-(r_i\,a_N)-1\big)_{1\le i\le \ell}$ *satisfies the CLT (3.5.32). From (3.5.33), it remains to prove that*

$$\sqrt{\frac{N}{a_N}}\,\frac{1}{N-r_i a_N}\Big(\sum_{b=1}^{N-r_i a_N} -2\widetilde{e}^+(r_i a_N,b)\widetilde{e}^-(r_i a_N,b) + \big(\widetilde{e}^+(r_i a_N,b)\big)^2\Big) \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0.$$

From Markov's and Hölder inequalities, this is implied when

$$\sqrt{\frac{N}{a_N}}\left(\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2 + 2\sqrt{\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2}\right) \underset{N\to\infty}{\longrightarrow} 0 \quad \text{with} \quad \mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2 = 1$$

. *Using* $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2 \le (\mathbb{E}\xi_0^4)\,M^{-2}$ *obtained above, we deduce that this statement holds when* $M=N$ *(for instance). As a consequence, from (3.5.33), the CLT (3.5.32) holds even if the distribution of* ξ_0 *is symmetric and such that* $\mathbb{E}\xi_0^4<\infty$.

Step 4 *It remains to apply the Delta-method to (3.5.32) with the function*

$$(x_1,\ldots,x_\ell)\mapsto(\log x_1,\ldots,\log x_\ell):$$

$$\sqrt{\frac{N}{a_N}}\Big(\log\big(T_N(r_i\,a_N)\big)-\log\big(\mathbb{E}\big(e^2(a_N,0)\big)\big)\Big)_{1\le i\le \ell}\xrightarrow[N\to\infty]{D}\mathcal{N}\big(0\,,\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big),$$

With $\mathbb{E}\big(e^2(a_N,0)\big)$ *provided in Property 3.1, we obtain*

$$\log\mathbb{E}\big(e^2(a_N,0)\big)=2d\,\log(a_N)+\log\big(2\pi\,c_d K_{(\psi,2d)}\big)+\frac{c_{d'}K_{(\psi,2d-d')}}{c_d K_{(\psi,2d)}}\,\frac{1}{a_N^{d'}}\big(1+o(1)\big)$$

Therefore, when $\sqrt{\frac{N}{a_N}}\frac{1}{a_N^{d'}}\underset{N\to\infty}{\longrightarrow}0$*, i.e.* $N^{\frac{1}{1+2d'}}=o(a_N)$*, the CLT (3.2.7) holds.*

Proof 3.5 (Proof of Theorem 3.1) *Here we use Theorem 1 of Bardet et al. (2008) where it was proved that the CLT (3.2.7) is still valid when* a_N *is replaced by* $N^{\widetilde{\alpha}_N}$*. Then, since* $\widetilde{d}_N=\widetilde{M}_N\,Y_N(\widetilde{\alpha}_N)$ *with* $\widetilde{M}_N=\big(0\ 1/2\big)\big(Z_1'\widehat{\Gamma}_N^{-1}Z_1\big)^{-1}Z_1'\widehat{\Gamma}_N^{-1}$ *we deduce that* $\sqrt{N/N^{\widetilde{\alpha}_N}}\big(\widetilde{d}_N-d\big)$ *is asymptotically Gaussian with asymptotic variance the limit in probability of* $\widetilde{M}_N\,\Gamma(1,\ldots,\ell,d,\psi)\,\widetilde{M}_N'$*, that is* σ^2*.*

The relation (3.3.6) is also an obvious consequence of Theorem 1 of Bardet et al. (2008).

Proof 3.6 (Proof of Theorem 3.2) *The theory of linear models can be applied :* $Z_{N^{\widetilde{\alpha}_N}}\begin{pmatrix}\widetilde{c}_N\\ 2\widetilde{d}_N\end{pmatrix}$ *is an orthogonal projector of* $Y_N(\widetilde{\alpha}_N)$ *on a subspace of dimension* 2*, therefore* $Y_N(\widetilde{\alpha}_N)-Z_{N^{\widetilde{\alpha}_N}}\begin{pmatrix}\widetilde{c}_N\\ 2\widetilde{d}_N\end{pmatrix}$ *is an orthogonal projector of* $Y_N(\widetilde{\alpha}_N)$ *on a subspace of dimension* $\ell-2$*. Moreover, using the CLT (3.2.7) where* a_N *is replaced by* $N^{\widetilde{\alpha}_N}$*, we deduce that* $\sqrt{N/N^{\widetilde{\alpha}_N}}\widehat{\Gamma}_N^{-1}Y_N(\widetilde{\alpha}_N)$ *asymptotically follows a Gaussian distribution with asymptotic covariance matrix* I_ℓ *(identity matrix). Hence from the usual Cochran Theorem we deduce (3.3.8).*

Bibliography

[1] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens, In Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Editors), *Theory and applications of long-range dependence*, 527–556, Birkhäuser, Boston.

[2] Andrews, D.W.K. and Sun, Y. (2004). Adaptive local polynomial Whittle estimation of long-range dependence, *Econometrica*, **72**, 569–614.

[3] Bardet, J.M. and Bertrand, P. (2007) Definition, properties and wavelet analysis of multiscale fractional Brownian motion, *Fractals* **15**, 73–87.

[4] Bardet, J.M., Bibi, H. and Jouini, A. (2008). Adaptive wavelet-based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli*, 14, 691-724.

[5] Bardet, J.M., Billat V. and Kammoun, I. (2011). A new process for modeling heartbeat signals during exhaustive run with an adaptive estimator of its fractal parameters, to appear in *J. of Applied Statistics*.

[6] Bardet, J.M., Lang, G., Moulines, E. and Soulier, P. (2000). Wavelet estimator of long range-dependant processes, *Statist. Infer. Stochast. Processes*, 3, 85–99.

[7] Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S. and Taqqu, M.S. (2003). Semiparametric estimation of the long-range dependence parameter : a survey. In *Theory and applications of long-range dependence*, Birkhäuser Boston, 557-577.

[8] Doukhan, P., Oppenheim, G. and Taqqu M.S. (Editors) (2003). *Theory and applications of long-range dependence*, Birkhäuser.

[9] Furmanczyk, K. (2007). Some remarks on the central limit theorem for functionals of linear processes under short-range dependence. *Probab. Math. Statist.*, 27, 235–245.

[10] Giraitis, L. (1985). The central limit theorem for functionals of a linear process. *Lith. Mat. J.*, 25, 43–57.

[11] 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.

[12] Giraitis, L., Robinson P.M., and Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter. *J. Multivariate Anal.*, 72, 183–207.

[13] Moulines, E., Roueff, F. and Taqqu, M.S. (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.

[14] Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractionnal processes, In Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Editors), *Theory and applications of long-range dependence*, 251–301, Birkhäuser, Boston.

[15] Peltier, R. and Lévy Véhel, J. (1995) Multifractional Brownian Motion : definition and preliminary results. Preprint INRIA, available on `http://hal.inria.fr/docs/00/07/40/45/PDF/RR-2645.pdf`.

[16] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence, *Ann. Statist.*, 23, 1630–1661.

[17] Roueff, F. and Taqqu, M.S. (2009a). Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *J. Time Ser. Anal.*, 30, 534–558.

[18] Roueff, F. and Taqqu, M.S. (2009b). Central limit theorems for arrays of decimated linear processes. *Stochastic Process. Appl.*, 119, 3006–3041

[19] Samorodnitsky G. and Taqqu, M.S. (1994). Stable non-Gaussian random processes, London, U.K., Chapman & Hall.

[20] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377–390.

[21] Wu, W.B. (2002). Central limit theorems for functionals of linear processes and their applications. *Statist. Sinica*, 12, 635–649.