ORSAY
N$^o$ d'ordre : 5099

# THÈSE

présentée devant

L' UNIVERSITÉ DE PARIS-SUD
U.F.R. scientifique d'ORSAY

pour obtenir le titre de

Docteur en sciences de l'Université de PARIS-SUD
Spécialité : Mathématiques et applications des mathématiques

par

Jean-Marc Bardet

Sujet de la thèse

# Tests d'autosimilarité des processus gaussiens. Dimension fractale et dimension de corrélation.

Soutenue le 18 décembre 1997 devant la commission d'examen composée de :

| | | | |
|---|---|---|---|
| Mrs. | Jean-Marc | AZAÏS | Rapporteur |
| | Didier | DACUNHA-CASTELLE | Président |
| | Paul | DOUKHAN | |
| | Francis | HIRSCH | |
| | Stéphane | MALLAT | |
| | Georges | OPPENHEIM | |
| | Donatas | SURGAÏLIS | Rapporteur |

# Abstract

This thesis is composed of four parts connected by the notion of fractality. The first part is an overview of chaotic dynamical systems. An important place is given to different dimensions and Lyapunov exponents. Statistic treatments of these notions are then exposed, with a special interest for correlation dimension and local Lyapunov exponents. The second part is a study of a correlation dimension, said local, for continuous stochastic processes. This dimension is given by the asymptotic behaviour of self-intersection occupation measure. We compare first this dimension to Hausdorff one. We compute then local correlation dimension of multiparameter fractional Brownian motion and give new bounds for Hausdorff dimension. The third part presents a method for testing self-similarity of discretized observations of a Gaussian process with stationary increments. The test is based on the estimation of a distance between the process and a set of processes containing all the fractional Brownian motions. This distance is based on two estimations of given variances of increments. The second one requires regression estimates of the self-similarity index $H$. Two estimators of $H$ are then introduced. They present good robustness and computing time properties. The test also gives a method to find a sampling period where fractional Brownian motion modeling should be possible. The fourth part studies a second self-similarity test for time series using wavelet analysis. The test statistic is a distance between two estimations of variances of wavelet coefficients for given scales. Two estimators of self-similarity index are also introduced and they present good robustness and computing time properties.

# Remerciements

Mes premiers remerciements vont à Didier Dacunha-Castelle sans qui je n'aurais pas eu la joie de me remettre à l'étude des mathématiques. Son expérience, ses connaissances et ses qualités intuitives m'ont nourri ainsi que sa présence chaque fois opportune dans mon travail. Ma gratitude va également vers l'homme dont la vie même, et sa façon de l'appréhender, sont un enseignement.

Je suis très reconnaissant envers Jean-Marc Azaïs et Donatas Surgailis pour avoir accepté d'être les rapporteurs de cette thèse. Leurs remarques m'ont été précieuses pour achever sa rédaction.

Je remercie vivement Paul Doukhan, Francis Hirsch, Stéphane Mallat et Georges Oppenheim de m'avoir fait l'honneur et le plaisir de participer au jury de cette thèse.

Je tiens à remercier Jean-Michel Poggi avec qui j'ai eu de nombreuses et fructueuses discussions. Je remercie aussi Georges Oppenheim, Xavier Guyon, Mario Wschebor, Jean-Pierre Kahane et Marielle Noto pour les "pistes" qu'ils m'ont fournies.

Je remercie toute l'équipe du Laboratoire de Modélisation Stochastique et Statistique d'Orsay, et plus particulièrement Fabrice, Gonzalo, Paul, Bernard, Marc et Patrick pour leurs "coups de pouce". Merci encore à Catherine et Sabine.

Je remercie aussi tous les membres de l'équipe de mathématiques de l'Université d'Evry qui m'ont offert d'excellentes conditions de travail. Merci à Valérie.

Merci enfin à Angélique pour son soutien et son affection.

# Table des Matières

# Chapitre 0

# Introduction

## 0.1 Fractales

L'apparition dans le monde mathématique des premières fractales date d'un siècle. Cantor, puis Julia et Hausdorff se sont intéressés à ces " monstres ", ces figures géométriques dont la construction consistait en l'itération d'un même motif à des échelles de plus en plus petites. Cependant, ces travaux sont restés marginaux car semblant sans possibles débouchées et ils ne trouvèrent pas de relais dans la communauté mathématique. Au début des années soixantes, Mandelbrot a su redonner vie à l'étude des fractales en exhibant leurs nombreuses manifestations naturelles, ce qui permit une médiatisation au delà de la sphère scientifique.
Mais d'abord, qu'est-ce qu'une fractale? Avec Mandelbrot, on pense à la carte de la côte bretonne ou au flocon de neige de Von Koch. Plus généralement, on vise par ce terme des objets au tracé fractionné et irrégulier, ce qui comprend les figures géométriques de dimension non entière et les ensembles de points invariants par changement d'échelle. Les fractales sont aussi bien des objets dont la construction est entièrement déterminée que des trajectoires de champs aléatoires. En fait, il manque pour l'instant une véritable définition de la notion de fractale, et beaucoup reste à étudier et comprendre. Ceci peut se faire par l'étude de notions dont les définitions sont établies. Aussi, nous sommes nous intéressé à l'autosimilarité et à la détermination de dimensions.

## 0.2 Fractales et systèmes dynamiques chaotiques

Notre premier domaine d'investigation (Chapitre 1), concerne les systèmes dynamiques, et plus particulièrement les systèmes dynamiques chaotiques. On associe souvent fractales et chaos, peut-être parce que leur développement médiatique a été concommitant. Mais aussi parce qu'il existe des liaisons entre ces notions, liaisons dont les fils ne sont

pas encore tous dénoués. Nous avons donc voulu en savoir plus sur ces liens par le biais d'une synthèse bibliographique. Il a fallu commencer par définir ce qu'est un système dynamique chaotique, ce qui nous a mené à la notion centrale d'attracteur étrange. Un attracteur étrange est un attracteur, c'est-à-dire, grossièrement, un ensemble contenant les vecteurs d'état du système quand celui-ci a suffisamment évolué, pour un système présentant une dépendance sensitive aux conditions initiales, c'est-à-dire que deux points très proches peuvent avoir un devenir très différent. Les attracteurs étranges sont très souvent des fractales, mais ce n'est pas toujours le cas. Mais si un attracteur est de dimension non entière, cet attracteur est étrange. Aussi a-t-on deux possibilités pour démontrer qu'un système est chaotique: soit on montre qu'un attracteur est de dimension non entière, soit on montre que le système est sensible aux conditions initiales au moins dans une direction tout en restant dans un domaine borné. Dans le premier cas, cela nous a demandé de revenir sur les définitions des différentes dimensions pouvant être non entières: dimension de Hausdorff et capacité de Kolmogorov, qui ne prennent en compte que la géométrie de l'attracteur, et les dimensions d'information, de corrélation, de Lyapunov,...etc, qui prennent aussi en compte le "temps" passé en chaque lieu de l'attracteur. Dans le second cas, cela nous a demandé de quantifier la divergence des trajectoires, ce qui peut souvent se faire grâce à la détermination des exposants de Lyapunov qui mesurent le taux de contraction ou de dilatation des trajectoires suivant les directions.
Mais les valeurs des dimensions ou des exposants de Lyapunov sont la plupart du temps impossibles à déterminer théoriquement, même quand on connaît explicitement l'évolution du système. On peut alors avoir recours à des simulations numériques permettant l'estimation de ces dimensions et exposants. Nous avons donc recenser les différents traitements statistiques utilisés pour estimer ces grandeurs. Ce sont les estimateurs de la dimension de corrélation et du plus grand des exposants de Lyapunov locaux qui reviennent le plus souvent. L'estimation de la dimension de corrélation est notamment très attractive car d'une grande simplicité. Elle s'effectue de la manière suivante: si on connaît $N$ points de l'attracteur, $X_1$, $X_2$,..., $X_N$, pour plusieurs rayons $r_k > 0$, on détermine la proportion $C(N, r_k)$ de bipoints $(X_i, X_j)$ dont l'inter-distance est inférieure à $r_k$, et on fait une régression des $\log C(N, r_k)$ par les $\log r_k$; le coefficient de proportionnalité estimé est une estimation de la dimension de corrélation. De très nombreux articles attestent de l'intérêt de cette méthode et de son efficacité, d'autant qu'elle permet aussi d'estimer le nombre de degrés de liberté du système, c'est-à-dire le nombre de variables nécessaires pour le paramétrer. On peut cependant constater que si l'étude statistique des systèmes dynamiques chaotiques a suscité la création de nouveaux outils comme les exposants de Lyapunov locaux ou la dimension de corrélation, les méthodes statistiques utilisées sont indifférentes au caractère chaotique du système.

## 0.3 Dimensions fractales de champs aléatoires

En second lieu (Chapitre 2), nous nous sommes intéressés à la fractalité de processus vectoriels aléatoires continus. Nous avons considéré un processus continu $X = \{X_t; t \in \mathbb{R}\}$ à valeurs dans $\mathbb{R}^d$. Le but de notre travail a été la détermination d'une dimension "fractale" d'une portion de trajectoire de ce processus. En fait, nous nous sommes restreint à la dimension de Hausdorff, et à une seconde dimension dite dimension de corrélation locale. Nous avons défini cette dernière (lorsqu'elle existe) de la manière suivante: on considère $(X_{T/N}(\omega), X_{2T/N}(\omega), .., X_T(\omega))$, $N$ vecteurs issus d'une trajectoire $X(\omega) = \{X_t(\omega); t \in [0, T]\}$. On définit la variable $C_T(N, r)$, qui correspond à la proportion de bipoints de la trajectoire, dont l'interdistance est inférieure à une longueur $r > 0$. On note $C_T(r)$ sa limite quand $N \to +\infty$, qui correspond à cette proportion pour la trajectoire continue. Si $C_T(r, \omega)$ se comporte approximativement comme une puissance de $r$, la dimension de corrélation locale $\nu(\omega)$ de $X(\omega)$ existe et est cette même puissance. Ainsi,

$$\begin{aligned} C_T(N, r, \omega) &= \frac{2}{N(N-1)} \sum_{1 \le i < j \le N} \mathbb{I}_{(\|X_{jT/N}(\omega) - X_{iT/N}(\omega)\| \le r)}. \\ C_T(r, \omega) &= \lim_{N \to +\infty} C_T(N, r, \omega) = \frac{1}{T^2} \int_0^T \int_0^T \mathbb{I}_{(\|X_t(\omega) - X_s(\omega)\| \le r)} dsdt \\ \nu(\omega) &= \lim_{r \to 0} \frac{\log C_T(r, \omega)}{\log r}. \end{aligned}$$

La dimension de corrélation locale peut aussi s'exprimer à partir de la mesure d'occupation $\mu(I)$ du champ aléatoire $X_t - X_s$, pour $(s, t) \in [0, T]^2$ et $I \subset \mathbb{R}^d$, autrement dénommée mesure d'occupation d'auto-intersection de $X$. On a $C_T(r) = T^2 \mu(B(0, r))$, où $B(0, r)$ est la boule ce centre 0 et de rayon $r$, et donc $\nu(\omega)$ existe et vaut $\nu$ lorsque $\mu(B(0, r)) \sim r^\nu L(r)$ quand $r \to 0$ avec $L(r)$ une fonction négligeable devant $\log r$. Nous avons alors obtenu les deux propriétés suivantes, qui confèrent à $\nu$ sa qualité de dimension (nous notons $dim_H S$ la dimension de Hausdorff d'un ensemble $S \subset \mathbb{R}^d$):

**Propriété 1**: *1/ Si $\nu(\omega)$ existe, alors $0 \le \nu(\omega) \le dim_H X(\omega) \le d$.*
*2/ Soient $D_d$ le volume de la boule $B_d(0, 1)$ et $V_0$ un voisinage de 0. Si $X(\omega)$ admet un temps local d'auto-intersection $\alpha(x, [0, T]^2, \omega)$ pour $x \in V_0$, continu et strictement positif en 0, alors,* $\nu(\omega)$ existe et $\nu = d = dim_H X(\omega)$.

Nous avons ensuite calculé cette dimension pour des processus vectoriels gaussiens continus à accroissements stationnaires, dont les coordonnées, indépendantes entre elles, sont des mouvements browniens fractionnaires, soit: $X = \{X_t = (X_t^1, X_t^2, ..., X_t^d), t \in [0, T]\}$, processus gaussien continu à valeurs dans $\mathbb{R}^d$ ($T > 0$, $d \in \mathbb{N}^*$), de moyenne nulle, avec $X_0 = 0$ p.s., d'index $(\alpha_1, \alpha_2, ..., \alpha_d)$, où chaque $\alpha_i \in ]0, 1[$, et tel que:

$\mathbb{E}X_t^iX_s^j = \frac{1}{2}\delta_{ij}a_i(\mid t\mid^{2\alpha_i} + \mid s\mid^{2\alpha_i} - \mid s-t\mid^{2\alpha_i})$ pour $(s,t)\in[0,T]^2$, $(i,j)\in\{1,2,...,d\}^2$. Pour faciliter l'écriture des résultats, nous supposerons que $0<\alpha_1\leq\alpha_2\leq...\leq\alpha_d<1$. Deux cas vont se présenter, d'où la proposition suivante:

**Proposition**: *1/ Si* $\sum_{k=1}^{d}\alpha_k\leq 1$, *alors* $\nu=d$ *avec une probabilité positive.*
*2/ Si* $\sum_{k=1}^{d}\alpha_k>1$, *alors,* $\nu$ *existe p.s. et* $\nu=\min_{i\in\{1,..,d\}}\left\{\frac{1+\sum_{k=1}^{i}(\alpha_i-\alpha_k)}{\alpha_i}\right\}$ *p.s.*

Ainsi, quand $\sum_{k=1}^{d}\alpha_k>1$, en notant $i_0$ l'entier tel que $\nu=\frac{1+\sum_{k=1}^{i_0}(\alpha_{i_0}-\alpha_k)}{\alpha_{i_0}}$, on montre que $\nu\in]i_0-1,i_0]$. Donc $\nu$ ne dépend que des $i_0$ coordonnées de $X$ les plus erratiques. De plus, on remarque que $\sum_{k=1}^{i_0-1}\alpha_k<1$ et $\sum_{k=1}^{i_0}\alpha_k\geq 1$. Donc, quand on considère les $i_0-1$ coordonnées les plus erratiques, on peut associer à leur trajectoire un temps local d'auto-intersection: la dimension de corrélation locale qui leur est associée vaut $i_0-1$ avec une probabilité positive. Si on rajoute une autre coordonnée, même la plus erratique (qui est tout de même moins erratique que chacune des $i_0-1$ coordonnées), on ne peut plus définir un temps local d'auto-intersection, et le fait de rajouter d'autres coordonnées n'y change rien: on a $\nu\leq i_0$.

Nous sommes revenu alors sur le calcul de la dimension de Hausdorff de l'image ($dim_HX$) et du graphe ($dim_HGr(X)$) de $X$, initialement déterminée par Cuzick dans Cuzick (1978). Nous avons ainsi montré que la démonstration de Cuzick ne permet d'obtenir qu'un encadrement dont la borne inférieure est donnée par la dimension de corrélation locale. Plus généralement, nous avons aussi obtenu un encadrement pour les champs gaussiens $(N,d)$ sur $[0,1]^N$ d'index $\alpha=(\alpha_1,..,\alpha_d)$. On alors:

**Proposition**: *Avec une probabilité de* 1,

- si $\sum_{k=1}^{d}\alpha_k\leq N$, $dim_HX=d$.
- si $\sum_{k=1}^{d}\alpha_k>N$, $\min_{i\in\{1,2,..,d\}}\left\{\frac{N+\sum_{k=1}^{i}(\alpha_i-\alpha_k)}{\alpha_i}\right\}\leq dim_HX\leq\frac{N+\sum_{k=1}^{d}(\alpha_i-\alpha_k)}{\alpha_d}$.
- $\min_{i\in\{1,2,..,d\}}\left\{\frac{N+\sum_{k=1}^{i}(\alpha_i-\alpha_k)}{\alpha_i}\right\}\leq dim_HGr(X)\leq d+N-\sum_{k=1}^{d}\alpha_k$.

## 0.4 Estimations de la dimension fractale et tests d'autosimilarité

Nous avons ensuite (Chapitres 3 et 4) considéré les problèmes statistiques suivants: connaissant $N$ données numériques $X_1, X_2, .., X_N$, comment estimer la dimension de Hausdorff du graphe formé par les $(i, X_i)$ et comment savoir si cette trajectoire est autosimilaire, c'est-à-dire "globalement" invariante quand on change d'échelle? Plus précisément, en ayant supposé que $X_1, X_2, .., X_N$ peut être modélisé par un processus gaussien centré à accroissements stationnaires, on voudrait avoir une estimation du paramètre d'autosimilarité de ce processus, ce qui permet d'avoir une estimation de la dimension de Hausdorff du graphe, et on voudrait savoir si ce processus se comporte comme un processus autosimilaire au moins pour un certain nombre d'échelles.
Si, pour l'instant, on ne trouve pas de test d'autosimilarité dans la littérature, on trouve un grand nombre de travaux sur l'estimation du paramètre de Hurst. Citons notamment les techniques par maximum de vraisemblance initiée par Fox et Taqqu (1986) et Giraitis et Surgailis (1990), par log-périodogramme introduite par Geweke et Porter-Hudack (1983) et développée par Robinson (1995), ou encore par aggrégation dont on trouve l'étude dans Taqqu *et al.* (1995,1996). Les vitesses de convergence des estimateurs du paramètre d'autosimilarité sont données par un théorème de limite centrale pour la première méthode et un théorème de limite non centrale pour la deuxième méthode. Nous avons choisi de travailler avec la troisième méthode, qui nous a permis d'obtenir des théorèmes limites pour plusieurs estimateurs, et qui nous a aussi permis de construire des tests d'autosimilarité assez naturels.

Nous l'avons dit, nous supposons que $X_1, X_2, .., X_N$ est une série chronologique issue d'un processus gaussien centré $X = \{X_t, t \in \mathbb{R}_+\}$ à accroissements stationnaires. Ce processus est dit autosimilaire de paramètre d'autosimilarité $H$, s'il vérifie: pour tout $d \in \mathbb{N}^*$, pour tout $(t_1, ..., t_d) \in \mathbb{R}_+^d$, et pour tout $c \in \mathbb{R}_+$:

$$(X_{t_1}, ..., X_{t_d}) \stackrel{\mathcal{D}}{=} (c^{-H} X_{ct_1}, ..., c^{-H} X_{ct_d}).$$

Donc toute distribution finie de $X$ doit être invariante en distribution pour tout changement d'échelle. En fait, les seuls processus gaussiens à accroissements stationnaires autosimilaires sont les mouvements browniens fractionnaires. Nous avons travaillé sur une version un peu plus faible de l'autosimilarité, que nous avons appelée autosimilarité pour un nombre $m$ d'échelles $(n_1, n_2, .., n_m) \in \mathbb{N}^m$, ce qui signifie que $X$ doit vérifier simplement: il existe $H \in [0, 1]$ et $\sigma^2 > 0$, tels que pour tout $n \in \{n_1, n_2, .., n_m\}$,

$$\mathbb{E} X_n^2 = \sigma^2 n^{2H}.$$

Les mouvements browniens fractionnaires sont évidemment autosimilaires pour toutes échelles. Notre but a été d'estimer les paramètres $H$ et $\sigma^2$ pour obtenir une projection de $X$ sur l'ensemble des processus autosimilaires pour les échelles $(n_1, n_2, .., n_m)$ qui minimise une distance entre $X$ et cet ensemble de processus. En considérant cette distance comme notre statistique de test dont nous donnons la distribution asymptotique, nous avons obtenu des tests simples permettant aussi de rejeter l'hypothèse "mouvement brownien fractionnaire".

Les premières distances (Chapitre 3) que nous avons considérées ont pour objet le vecteur $S_N$ composé des logarithmes des variances empiriques des accroissements de longueur $n_1, n_2, .., n_m$. Un point de l'ensemble des processus autosimilaires pour les échelles $(n_1, n_2, .., n_m)$ a alors pour coordonnées $(2H \log n_1 + \log \sigma^2, ..2H \log n_m + \log \sigma^2)$. Une première distance permet d'obtenir par moindres carrés ordinaires une estimation de $\theta = (H, \sigma^2)$: c'est la méthode d'estimation par aggrégation. Si $X$ est autosimilaire pour les échelles $(n_1, n_2, .., n_m)$ avec $0 < H < 3/4$, nous montrons un théorème de limite centrale pour $\theta$, et si $3/4 < H < 1$, un théorème de limite non centrale dont la vitesse de convergence est en $N^{2H-2}$. Dans le cas où $0 < H < 3/4$, on peut aussi obtenir un estimateur de $\theta$ par moindres carrés généralisés, dont nous montrons la convergence par un théroème de limite centrale dont la variance asymptotique est très proche de celle obtenue par maximum de vraisemblance. Nous obtenons ainsi, dans le cas où $3/4 < H < 1$, une seconde distance entre $S_N$ et son projeté par moindres carrés ordinaires dont nous donnons la distribution asymptotique, et dans le cas où $0 < H < 3/4$, une distance entre $S_N$ et son projeté par moindres carrés généralisés dont la distribution asymptotique est une loi du Chi-deux à $m - 2$ degrés de liberté. Cette seconde distance est très intéressante car facilement testable et de vitesse de rejet en $N$.

Nous avons ensuite (Chapitre 4) considéré une distance dont l'objet est le vecteur $I_N$ composé des logarithmes des variances empiriques des coefficients d'ondelette pour des échelles $n_1, n_2, .., n_m$. Un point de l'ensemble des processus autosimilaires pour les échelles $(n_1, n_2, .., n_m)$ a alors pour coordonnées $((2H + 1) \log n_1 + \log C, ..2H \log n_m + \log C)$ ($C$ est une constante positive). Pour $0 < H < 1$, on peut obtenir des estimateurs de $\theta' = (H, C)$ par moindres carrés ordinaires et généralisés, ainsi que des théorèmes de limite non-centrale quand les échelles choisies sont suffisamment grande. La distance entre $S_N$ et son projeté par moindres carrés généralisés suit une distribution asymptotique qui est une loi du Chi-deux à $m - 2$ degrés de liberté. Ici encore, le test donne un résultat très simple à tester avec cette fois ci, une vitesse de rejet au mieux en $N^{2/3-\varepsilon}$ où $\varepsilon > 0$.
Ce qui permet d'éviter la dichotomie en $H = 3/4$ du précédent test est le choix d'une ondelette mère ayant ses moments d'ordre 0 et d'ordre 1 nuls. Les coefficents d'ondelette ont ainsi une vitesse de décorrélation plus importante que les accroissements dont seul le moment d'ordre 0 est nul. Pour avoir une convergence suffisante des coefficients d'ondelette empiriques vers les coefficients d'ondelette, il faut avoir des échelles suffisamment grandes, et des ondelettes mères suffisamment régulières ou bien encore de

moments empiriques nuls. Toutes ces conditions n'étaient pas nécessaires avec les estimations similaires de $H$ effectuées par Abry *et al.* (1997), Abry et Veitch (1997) ou Gonçalvès et Abry (1997), car le comportement local du processus y étaient connus. Le passage par une discrétisation se paie donc de certaines restrictions sur le choix de l'ondelette.

Nous avons réalisé des simulations numériques de ces tests. En ce qui concerne les estimateurs de $H$ (donc de la dimension de Hausdorff du graphe), la méthode d'aggrégation avec estimation par moindres carrés généralisés donnent de très bons résultats lorsque $0 < H < 3/4$, très proches de ceux obtenus par maximum de vraisemblance. Cela conduit à un test très performant lorsque l'on sait que $0 < H < 3/4$. Si a priori $H$ peut être supérieur à $3/4$, le test utilisant l'analyse par ondelette est plus intéressant et donne numériquement satisfaction. Les simulations confirment certains avantages des méthodes temporelles que nous avons employées vis-à-vis des méthodes spectrales du maximum de vraisemblance ou du log-périodogramme: les calculs sont beaucoup plus rapides (il n'y a ni transformation de Fourier ni recherche d'extremum) et les estimateurs et tests sont très robustes aux bruits blancs.
Les tests effectués sur des données réelles (charge d'électricité sur le réseau français et niveau de crues du Nil) offrent des exemples des autres possibilités offerte par les méthodes choisies. Tout d'abord, la liberté dans le choix du nombre d'échelles entraîne la possibilité de faire plusieurs tests à partir des mêmes données en changeant ce nombre d'échelles. On a ainsi une plus grande certitude sur les résultats de ce test. Ensuite, en choisissant des échelles multiples d'un entier $l$, de la forme $n_i = il$, $i = 1, .., m$, et en jouant sur la valeur de $l$, on peut tester les données pour déterminer un possible $l = l_0$ pour lequel le test est accepté. On pourra en déduire alors qu'un reéchantillonage du processus par une période $l_0$ fournit une série chronologique autosimilaire pour les échelles $1, 2, .., m$. La modélisation de ce nouveau processus par un brownien fractionnaire sera alors légitime. C'est notamment ce qui se passe avec les 50400 mesures de charge électrique qui rejettent les tests pour des échelles "petites" (avec $l < 50$), mais qui les acceptent quand $l = l_0 = 50$. On pourra ainsi modéliser la série chronologique $(X_0, X_{50}, .., X_{50400})$ par un mouvement brownien fractionnaire de paramètre $H \simeq 0.29$.

## 0.5 Perspectives...

Les principaux résultats développés dans cette thèse concernent la fractalité de processus et champs gaussiens. Le traitement probabiliste (Chapitre 2) introduit une nouvelle dimension, la dimension de corrélation locale, dont on pourrait étendre la définition aux champs continus. Il serait alors naturel de déterminer sa valeur pour les champs gaussiens à index, dont on peut penser qu'elle sera égale à la borne inférieure de la dimension de Hausdorff. Pour ces mêmes champs, dans le cas de l'anisotropie, un nouveau

problème se pose pour savoir si on peut améliorer l'encadrement de la dimension de Hausdorff, voire arriver à obtenir sa valeur. Enfin, il serait intéressant de savoir si l'on peut déterminer la dimension de corrélation locale de processus vectoriels $\alpha$-stables, de Lévy ou ARIMA, pour obtenir ensuite des résultats sur la dimension de Hausdorff.
Les traitements statistiques que nous proposons dans les Chapitres 3 et 4 offrent de nouvelles estimations du paramètre d'autosimilarité, ainsi que les premiers tests d'autosimilarité pour les processus gaussiens à accroissements stationnaires. Nous pensons que l'estimation induite par l'analyse par ondelette peut être améliorée, pour donner une plus grande vitesse de convergence et sans doute, un théorème central limite. Le choix de certaines ondelettes spécifiques pourraient aussi faciliter cette tâche. Enfin, la technique par ondelette devraient pouvoir s'étendre à d'autres processus autosimilaires, comme les processus ARIMA, et la possibilité de choisir le nombre des moments nuls de l'ondelette devrait aussi pouvoir permettre de tester des processus de Lévy, ou des processus $\alpha$-stables.

# Chapitre Premier

# Les traitements statistiques des systèmes dynamiques chaotiques

## 1.1 Introduction

### 1.1.1 "Le chaos: nouveau paradigme scientifique"

Cet intitulé provocateur, illustrant la pensée de Kuhn (1961), est sous-jacent (et parfois clairement présent) dans nombre d'articles, revues scientifiques ou publications les plus récents (par exemple Prigogine et Steingers (1979), Gleick (1987), Ruelle (1991)). Dépassant le simple phénomène de mode, la science a du prendre en compte depuis une vingtaine d'année cette nouvelle composante, souvent synonyme de complexité, et ce aussi bien en physique, en biologie qu'en mathématiques (voir Bergé et Pommeau (1984), Dahan *et al.* (1992)). C'est au traitement mathématique des phénoménes chaotiques que nous allons nous restreindre ici.

Henri Poincaré, par ses recherches sur le système des trois corps, fut le premier à s'interesser à ce qu'on l'on appelera les systèmes dynamiques chaotiques. Il s'agissait à partir des lois de Newton, de prendre en compte non plus deux, comme on le faisait jusqu'alors, mais trois masses ponctuelles, et de prédire le comportement d'un tel système (on peut penser aux planètes du système solaire). Poincaré, après avoir montré l'impossibilité d'une résolution directe des équations, démontra que la moindre diffèrence dans les conditions initiales du système pouvait entraîner des évolutions sans commune mesure.

À la suite de cela, et même si Hadamard et Julia en topologie, et les grands mathématiciens et physiciens russes de ce siècle (Lyapunov, Kolmogorov, Sinai, Arnold, Landau, Asonov, ...), en probabilité et systèmes non-linéaires, avaient bati les futurs outils mathématiques propres à son étude, le chaos tel quel ne réapparut vraiment qu'en 1963, avec les prévisions météorologiques de Lorenz (le terme "chaos" n'apparut, lui, qu'au cours des

années 70). A partir d'une troncature des équations hydrodynamiques de Navier-Stockes, simulant des phénomènes de convection en trois dimensions, Lorenz s'apercut de ce qui deviendra le médiatique "effet ailes de papillon", c'est-à-dire de la dépendance sensitive aux conditions initiales de ce système (d'où l'impossibilité des prévisions météorologiques à long terme). Notons aussi que cette découverte s'est produite avec l'aide de simulations numériques par ordinateur, dont l'utilisation est devenu depuis un apport considérable pour l'étude du chaos.
Les ailes de papillons de Lorenz mettaient aussi en évidence une propriété géométrique souvent associée aux systèmes chaotiques: l'ensemble des points décrit avec le temps était une fractale, un ensemble de dimension non entière. Mandelbrot (Mandelbrot (1982)) a su médiatiser ces objets singuliers déjà étudiés au début de ce siècle par Cantor, Julia ou Hausdorff. Beaucoup de recherches fructueuses notamment ont portées sur les fractales présentant la propriété d'invariance d'échelle, qui se rencontre dans la nature (flocons de neige, fougères,etc...). Cependant, les relations entre le phénomène de chaos et celui de fractalité ne sont pas intimement liés et nous verrons que l'un va parfois sans l'autre.
On a depuis trouvé des manifestations du chaos dans de multiples phénomènes: hydrodynamique avec les problèmes de turbulence et les notions de quasipériodicité et d'apériodicité, mécanique et électromagnètisme avec la notion d'oscillation (boussole, champs magnétique terrestre, scintillement du soleil), épidémiologie (rougeole), biologie humaine (pulsations cardiaques et taux sanguins), démographie (Lynx du Canada), finances,... Cependant, les preuves théoriques de la présence d'un phénomène chaotique laissent souvent la place à des constatations numériques. On mèle des techniques ressortant de l'aléatoire et du déterministe avec la prétention de vouloir distinguer l'un de l'autre (voir Casdagli (1991), Osborne et Provenzale (1989), et sur les rapports entre aléatoire et déterminisme, Dahan *et al.* (1992)).

### 1.1.2 Les traitements statistiques du chaos

Les mathématiciens traitent les phénomènes chaotiques sur plusieurs fronts: géometrique, combinatoire, topologique, analytique et statistique. Leur point de concours est la théorie des systèmes dynamiques, qui s'est édifiée avec les apports de la théorie ergodique. Les systèmes dynamiques étant, quand ils sont expansifs, sans véritables intérêts concrets et quand ils sont conservatifs, déjà bien compris, ce sont les systémes dynamiques dissipatifs qui motiveront les recherches.
Grossièrement, un tel systéme est dit chaotique lorsque son évolution présente une dépendance sensitive aux conditions initiales, c'est-à-dire que deux positions très proches peuvent conduire à deux trajectoires très différentes. A plus ou moins long terme, la notion de trajectoire n'est plus pertinente pour étudier le système. On doit donc faire appel à des techniques stochastiques ou statistiques pour déterminer les caractéristiques générales du système (et aussi pour savoir si l'on a bien à faire avec le chaos). Ces

traitements peuvent porter sur la dynamique propre du système, son évolution, et ils seront alors relatifs à une mesure invariante pour ce système (évaluation des exposants de Lyapunov, dimensions d'information, de corrélation, de Lyapunov,..), ou bien sur la forme globale des trajectoires, et ils seront relatifs à un ensemble invariant pour le système, son attracteur (évaluation des dimensions de plongement, de Hausdorff,...).

## 1.2 Définitions et propriétés

### 1.2.1 Définition d'un système chaotique

Nous considérerons un système dynamique défini par reccurence (une autre voie aurait pu être de le définir par équation différentielle) dont l'évolution est représentée dans un espace métrique $M$ de dimension $m$, dit **espace des phases** auquel est associée la distance $d$, par un **vecteur d'état** $X_n$ et une équation de récurrence:

$$X_{n+1} = f(X_n), \tag{1}$$

où $f$ est une fonction de $M \to M$. On adoptera l'écriture: $f^n = f_O f_O ..._O f$, n fois. Pour le moment, on ne fait pas d'hypothèses sur la régularité de $f$, mais de manière générale, $f$ sera toujours au moins continue. Sommairement, deux critères sont nécessaires pour que le système soit chaotique: un phénomène de "bornitude" de la fonction, impliquant l'invariance locale asymptotique, et la caractéristique de sensibilité aux conditions initiales. On doit donc tout d'abord localiser le phénomène; un système est dit chaotique sur un ensemble, que nous appellerons l'attracteur $A$. Avant d'énoncer une définition de cet attracteur, on définit d'abord deux ensembles contenant $A$ : un ensemble attracteur $E$ et son bassin d'attraction $B$ (nous allons suivre ici les définitions proposées par Eckmann et Ruelle (1985), dont on peut trouver des versions plus rigoureuses dans Ruelle (1989b)).

#### Attracteurs

**Définition 1.2.1** *Un ensemble fermé $E \subset M$ est un* **ensemble attracteur** *pour $f$ si:*

- *$E$ est invariant par $f$, soit $f(E) = E$.*
- *il existe $B$ ouvert de $M$ tel que $B = \{x \in M / \lim_{n\to\infty} d(f^n(x), E) = 0\}$. $B$ est le* **bassin d'attraction** *de $E$.*

Intuitivement, $E$ représente un ensemble contenant $A$ tel que tous les points issus de $B$ finissent par "échouer" dans $A$ quand on laisse le système évoluer. La plupart du temps, on a $E = A$ et $E$ compact. $E$ peut être un ou des points fixes, une ou des orbites périodiques, donc un ensemble de mesure nulle. De manière plus générale, $E$ contient $A$ et d'autres ensembles. La définition suivante spécifie le caractère d'attracteur de $A$:

**Définition 1.2.2** *Un ensemble $A \subset E$ est un* **attracteur** *de $E$ si:*

- $f(A) = A$.
- *l'ensemble des points périodiques de $f$ dans $A$, c'est-à-dire l'ensemble des $x \in A$ tel qu'il existe $p \in \mathbb{N}^*$ vérifiant $f^p(x) = x$, est dense dans $A$.*

Il peut exister plusieurs attracteurs dans un ensemble attracteur. L'invariance de $A$ par $f$ n'implique pas pour autant l'existence d'un bassin d'attraction pour $A$. L'attracteur $A$ présente parfois un phénomène d'**invariance d'échelle**, effet de son invariance par $f$. Cette invariance d'échelle, combinée avec la possible non-linéarité de $f$, peut être source de la fractalité de $A$ (voir l'article de Eckmann dans Dahan *et al.* (1992)).
La définition de l'attracteur doit être étoffer pour admettre une stabilité aux petites pertubations comme celles obtenues par troncature numérique ou par un bruit de mesure. Utilisant une définition plus rigoureuse des attracteurs, Ruelle dans Ruelle (1980) montre que sous des hypothèses assez faibles sur $f$, $A$, et sur une pertubation $\rho_x^\varepsilon$ (où $x \in M$), **les attracteurs d'un système dynamique déterministe sont stables pour de petites perturbations**.

### Attracteurs étranges

Un attracteur va représenter géométriquement le comportement asymptotique d'un système; lorsque celui-ci présentera une dépendance sensitive aux conditions initiales, on parlera d'attracteur étrange (nous reprenons ici une définition de Douady dans Dahan *et al.* (1992)).

**Définition 1.2.3** *Un système dynamique présente une* **dépendance sensitive aux conditions initiales** *lorsque $f$ appliquée à certains vecteurs de $M$, se comporte comme une similitude de rapport supérieur à 1 en module:*

$$\exists l > 0,\ \exists X_0 \in M,\ \textit{tels que } \forall e > 0,\ \forall X_0' \in M \ \textit{vérifiant } \|X_0' - X_0\| = e,\ \textit{alors}$$
$$\exists N \in \mathbb{N} \ \textit{tel que } \|f^N(X_0') - f^N(X_0)\| \geq l.$$

*On parlera de* **divergence exponentielle des trajectoires** *dans les cas particuliers où $f$ présente des propriétés entrainant une "constance" itérative de l'éloignement de certaines conditions initiales originairement proches (voir plus loin avec les exposants de Lyapunov).*

Il nous faut rappeler que le système est dissipatif, ce qui signifie que $f$ contracte les volumes de manière globale. A la limite, après un nombre infini d'itérations par $f$, tout volume de $M$ doit avoir une mesure de Lebesgue nulle. La dépendance sensitive aux conditions initiales dilatera certaines directions tout en contractant les autres.

**Définition 1.2.4** *Lorsque tout point de l'attracteur est condition initiale d'un phénomène de dépendance sensitive, on l'appellera* **attracteur étrange**.

**Les attracteurs étranges sont souvent des fractales, mais ce n'est pas une obligation**; l'exemple du "chat" d'Arnold, obtenu par à partir d'un difféomorphisme sur le tore nous enseigne le contraire...(Arnold (1980)). Il nous est maintenant possible de définir ce qu'est un système chaotique:

**Définition 1.2.5 Un système dynamique dissipatif sera dit chaotique si et seulement si il admet au moins un attracteur étrange**.

Les systèmes dynamiques sont aussi le plus souvent dépendant de paramètres relatifs aux conditions empiriques de l'expérience (paramètres de contrôle, de jauge, ...). En modifiant légèrement un paramètre, le système peut présenter des comportements totalement différents; notamment passer d'un mode périodique à un mode chaotique (ou réciproquement). On parlera alors d'un phénomène de **bifurcation** (pour des exemples, voir Feigenbaum (1980), Benedicks et Carleson (1991), et pour des résultats généraux, voir Ruelle (1989b)).

### 1.2.2 Des exemples et contre-exemples de systèmes chaotiques

Nous présentons des exemples "historiques" de systèmes dynamiques chaotiques ayant une certaine valeur générique:

#### Exemple 1: fonction de Hénon

Cette transformation de $\mathbb{R}^2$, par sa généricité, son apparente simplicité et le grand nombre de publications qu'elle a suscité, depuis Hénon (1976) jusqu'à Benedicks et Carlsen (1991), va être notre principal appui empirique. Elle se définit par la transformation:

$$\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = \begin{pmatrix} 1 - ax_n^2 \\ bx_{n-1} \end{pmatrix}, \tag{2}$$

où $(x_0, x_1) \in \mathbb{R}^2$, et avec $a$ et $b$ réels. On a bien ici $X_n = f(X_{n-1})$, avec $X_n = \begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix}$.
Une autre option aurait pu être de considérer la série numérique, définie par:

$$x_{n+1} = 1 - ax_n^2 + bx_{n-1} \tag{3}$$

ce qui revient à définir une fonction $f$ de la forme:

$$x_{n+1} = f(x_n, x_{n-1}). \tag{4}$$

Quand on fait varier $a$ et $b$, apparaissent des bifurcations très étonnantes, où l'on passe de deux points fixes à des attracteurs étranges dont on ne sait toujours pas s'ils ne sont pas que de très longues orbites périodiques (pour plus de détails numériques, voir Feit (1979), théoriques, voir Benedicks et Carleson (1991)). Un cas très souvent repris est celui avec $a = 0.3$ et $b = 1.4$ pour lequel on observe un attracteur (apparemment) étrange caractéristique.

**Contre-exemple 2: fonction quadratique**

Autre fonction d'expression simple et d'utilisation fréquente, la fonction quadratique (dite aussi fonction de Feigenbaum ou fonction logistique) est une application dans $\mathbb{R}$, qui génère la suite numérique suivante:

$$X_{n+1} = aX_n(1 - X_n), \tag{5}$$

avec $a$ un réel appartenant à $[0; 4]$ et $X_0$ appartenant à $[0; 1]$. Cette suite admet un point fixe pour tout $a$, qui pour $a > 1$ est un point source. Quand $a$ croit de 0 à 4, la suite à un comportement qui passe d'un point limite à un cycle limite de taille 2, puis 4, puis 8,... avec donc à chaque fois un doublement de période, jusqu'à la valeur $a_\infty = 3,5699...$ (solution de l'équation $x(x - 2)(x + 2) = 4$), où le système devient "chaotique" ("l'attracteur" y est alors un ensemble de Cantor porté par $[0; 1]$). Ce phénomène de cascade périodique est un des trois scénari de bifurcation, de "destabilisation périodique", source d'apparition de chaos, qui existe aussi avec la fonction de Hénon (Bergé, Pomeau, Vidal (1984)). Quand $a$ croît de $a_\infty$ à 4, la suite admet sur des intervalles en $a$ de mesure strictement positive aussi bien des points limites, que des cycles limites, qu'un comportement asymptotique de type chaotique. La valeur $a = 4$ est souvent mise en avant, car la suite se comporte alors asymptotiquement comme un processus aléatoire suivant une distribution de Cauchy (pour plus de détails, Feigenbaum (1978), Eckmann et Ruelle 1985)).
En fait, cette série chronologique n'a jamais un comportement chaotique tel que nous l'avons précédemment défini. En effet, pour $a > 1$, les points 0 et 1 sont des points sources; l'intervalle $[0; 1]$ n'est pas alors un ensemble attracteur et il ne possède pas de bassin d'attraction véritable. Il ne peut donc pas exister de stabilité aux petites pertubations, et même si apparait sur ses intervalles une dépendance sensitive aux conditions initiales, le système n'est jamais chaotique.

**Exemples 3:**

Dans le cas de systèmes dynamiques continus, de nombreux exemples sont possibles, dont:
- l'équation de Navier-Stockes

$$\frac{\delta v_i}{\delta t} = -\sum_j v_j \delta_j v_i \ + \ \nu \Delta v_i - \frac{1}{d}\delta_i p + g_i,$$

avec la condition d'incompressibilité $\sum_j \delta_j v_j = 0$, décrivant les mouvements des particules de liquides ou de gaz ($v$ est la vitesse, $d$ la masse volumique, $g$ représente les forces extérieures, $p$ la pression, et $\nu$ la viscosité)

- l'équation de Lorenz (troncature de l'équation de Navier-Stockes)

$$\begin{cases} \frac{\delta x}{\delta t} & = & -\sigma x + \sigma y \\ \frac{\delta y}{\delta t} & = & -xz + rx - y \\ \frac{\delta z}{\delta t} & = & -xy - bz, \end{cases}$$

modélisant un phénomène de convection pour un fluide entre deux plaques de température, avec $x, y, z$ représentant la position d'une particule de fluide, $r$ le gradient vertical de température proportionnel au nombre de Rayleigh, $\sigma$ le nombre de Prandtl du fluide, $b$ relié au vecteur d'onde convectif. Lorenz utilisait la distribution $r = 28$, $\sigma = 10$, et $b = 8/3$ quand il remarqua la sensibilité au conditions initiales, génératrice de l'attracteur à la forme d'un papillon.

- le modèle de la régénérescence sanguine de Mackey-Glass (utilisé aussi en finance), où

$$\frac{\delta x(t)}{\delta t} = \frac{ax(t-\tau)}{1 + (x(t-\tau))^k} - bx(t)$$

est une de ses formes possibles. Ce sont des équations différentielles non linéaires et non intégrables, dont on peut, à partir de solutions approchées, créer des séries chronologiques. Les équations de Navier-Stockes et de Mackey-Glass ont, sous leur forme générale, des attracteurs de dimension infinie. Ruelle et Takens (1971) ont montré que l'obtention d'un caractère chaotique, à partir d'un système en temps continu, n'est possible que sur des espaces de phases de dimension au moins égale à trois; dans le cas discret, la dimension 2 est au minimum requise pour l'obtention du chaos.

### 1.2.3 Utilisation de la théorie ergodique

Soit le système dynamique défini par $X_{n+1} = f(X_n)$ avec $f$ une fonction continue sur $M$, espace métrique de dimension $m$, tel que $(X_n)_{n\in\mathbb{N}}$ ait un comportement asymptotique gènèrant un ensemble attracteur $E$ compact, pour $X_0$ appartenant à l'ensemble $B$, bassin d'attraction de $E$. On suppose que $A$ est un attracteur de $E$. On étudie maintenant le système à l'aide de la théorie de la mesure.

**Définition 1.2.6** *Si $\mu$ est une mesure de probabilité sur $(A, \mathcal{T})$, où $\mathcal{T}$ est une tribu associée à $A$, $\mu$ est une* **mesure invariante** *sur $A$ par rapport à $f$ si et seulement si, pour tout ensemble mesurable $K \subset A$, on ait $\mu(K) = \mu(f^{-1}(K))$.*

Beaucoup de résultats intéressants sont possibles quand on ajoute l'hypothèse d'ergodicité pour $\mu$:

**Définition 1.2.7** *Sous les hypothèses de la Définition 1.2.6, $\mu$ est dite* **ergodique** *s'il n'existe pas de décomposition non triviale de $\mu$ sous la forme $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, où $\mu_1$ et $\mu_2$ sont deux autres mesures invariantes sur $A$ par rapport à $f$ (et $\alpha_1$ et $\alpha_2$ sont deux réels).*

Une autre caractérisation de l'ergodicité de $\mu$ réside dans le fait que pour tout ensemble mesurable $B$ invariant par $f$ et inclus dans $A$, on a ou bien $\mu(B) = 0$ ou bien $\mu(B) = 1$. De cette définition, on a le résultat suivant, issu du théoréme du point fixe de Markov-Kakutani:

**Théorème 1.2.1** *Si $f : A \to A$ est continue et si $A$ est un compact de $M$, il existe une mesure de probabilité $\mu$ invariante sur $A$ par rapport à $f$, de support contenu dans $A$. On peut choisir $\mu$ ergodique.*

Il est important de noter que $A$ possède généralement une famille dénombrable de mesures de probabilité invariantes par $f$ de support inclus dans $A$. Un cas particulier très important, utilisé empiriquement pour trouver une mesure invariante, est celui de la mesure physique sur $A$, donnée par la définition suivante:

**Définition 1.2.8** *Avec $f : A \to A$, la* **mesure physique** *sur $A$ pour $f$ est la mesure $\mu_p$ telle que pour $C$ ensemble mesurable de $A$, et pour $\mu_p$-presque tout $x \in A$:*

$$\mu_p(C) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \mathbb{I}_{(f^i(x))\in C)}. \tag{6}$$

Intuitivement, cette mesure exprime la proportion de "temps" passé par le système dans "chaque partie" de $A$. Le théorème suivant va donner une place centrale à cette mesure invariante particulière:

**Théorème 1.2.2 Théorème Ergodique**: *Si $\mu$ est une mesure invariante ergodique sur $A$ pour $f$, alors pour $C$ ensemble mesurable de $A$ et pour $\mu$-presque tout $x \in A$:*

$$\mu(C) = \mu_p(C) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \mathbb{I}_{(f^i(x))\in C)}. \qquad (7)$$

On trouve des résultats intéressants sur certaines mesures ergodiques, les mesures SRB et les mesures de Kolmogorov dans Eckmann et Ruelle (1985). La présentation précédente nous permet cependant de définir de nouvelles notions particulièrement utilisées dans l'étude des systèmes dynamiques chaotiques.

## 1.3 Notions caractéristiques des systèmes dynamiques chaotiques

On peut dégager trois types de concepts permettant une étude précise des systèmes dynamiques chaotiques. Le premier type n'est pas a priori spécialement réservé aux systèmes chaotiques. Il regroupe des caractéristiques topologiques de l'espace des phases et de l'attracteur du système, lorsqu'il existe. Le second type regroupe des caractéristiques liées à l'évolution temporelle du système, à sa dynamique, dont les exposants de Lyapunov qui sont les témoins de la dépendance aux conditions initiales. Enfin, un troisième type comprend des objets qui dépendent à la fois de la géométrie des attracteurs et de la dynamique du système.

### 1.3.1 Caractéristiques géométriques du système

On associe souvent phénomène chaotique et fractale. Nous l'avons dit, ce sont deux notions distinctes, il existe des systèmes chaotiques ayant des attracteurs non fractals, et il existe des attracteurs fractals pour des systèmes non chaotiques. Il n'en demeure pas moins que le comportement asymptotique de nombreux processus chaotiques admet un attracteur fractal, souvent de dimension non entière. Il faut donc définir ce qu'est une (ou des) dimension(s) pouvant être non-entière(s). En fait, il existe de nombreuses définitions possibles de la dimension d'un ensemble (on trouvera plus de détails sur ce qui suit dans Falconer (1990), Cutler (1991), Eckman et Ruelle (1985) ou Isham (1994)). Avant de les énumérer, rappellons ce qu'est un ensemble de Cantor, premier représentant générique et historique des ensembles de dimension non-entière.

**Définition 1.3.1** *Un ensemble compact $E$ de $M$, espace métrique, est un* **ensemble de Cantor** *si:*

*(1) chaque point de $E$ est un point limite.*

*(2) les seuls sous-ensembles connexes non vides de $E$ sont les ensembles constitués d'un point de $E$.*

De la première propriété, on dira aussi que $E$ est complètement disconnecté.

### Dimension de Hausdorff

La première, historiquement, des dimensions non entières fut la dimension de Hausdorff, qui peut être déterminer pour tout ensemble $A$ d'un espace metrique non vide.

**Définition 1.3.2** *Pour $r$ réel positif, et $\sigma = (\sigma_n)_{n\in\mathbb{N}}$ un recouvrement dénombrable de $A$ tel que $\sup_{n\in\mathbb{N}} diam(\sigma_n) \leq r$, pour tout $\alpha \geq 0$, on appelle* **mesure de Hausdorff de $A$ de dimension** *$\alpha$ le réel $m^\alpha(A)$, tel que:*

$$m^\alpha(A) = \lim_{r\to 0} m_r^\alpha(A) \text{ avec } m_r^\alpha(A) = \inf_\sigma \sum_n diam(\sigma_n)^\alpha.$$

*La* **dimension de Hausdorff de** *$A$, $dim_H(A)$, vérifie:*

$$dim_H(A) = \sup\{\alpha \in \mathbb{R} \text{ tel que } m_r^\alpha(A) > 0\} = \inf\{\alpha \in \mathbb{R} \text{ tel que } m_r^\alpha(A) = 0\}$$

On peut montrer que la mesure de Hausdorff de $A$ de dimension $\alpha$ est une mesure positive. Pour $F \in \mathbb{R}^n$, une sous variété de dimension $m$ continuement dérivable, alors $dim_H(F) = m$, ce qui permet la continuité de cette dimension avec la dimension entière habituelle. De plus, si $F \subset \mathbb{R}^n$ tel que $dim_H(F) < 1$, alors F est complétement disconnecté. Il est aussi possible de ne considérer que des recouvrements finis de boules ou bien fermées ou bien ouvertes. La détermination de la dimension de Hausdorff n'est possible que dans des cas simples, notamment pour des ensembles de Cantor présentant une invariance d'échelle. Pour plus de détails, voir Falconer (1990).

### Capacité de Kolmogorov

Une approche plus intuitive de la notion de dimension non-entière consiste à travailler à partir d'éléments de volume de diamètres très petits. Appelons $N(r, A)$ le nombre minimum de boules fermées de rayon $r$ nécessaire pour recouvrir un ensemble $A$ d'un espace métrique non vide (on pourrait aussi considérer des cubes de longueur $r$, ou encore n'importe quels ensembles de diamètre infèrieur à $r$). Alors:

**Définition 1.3.3** *La* **capacité de Kolmogorov** *de $A$, $dim_K(A)$ est définie par:*

$$dim_K(A) = \lim_{r\to 0} \frac{\log[N(r, A)]}{\log(1/r)} \quad (\text{si cette limite existe}).$$

On appelle aussi cette dimension, box-counting dimension ou dimension de packing. Pour préciser l'existence de cette dimension, on pourra parler de plus grande et plus petite capacités de Komogorov, à partir des limites sup et inf au lieu de la simple limite (l'existence de cette dernière étant fondée par l'égalité des deux autres). Nous allons illustrer les relations entre la dimension de Hausdorff et la capacité de Kolmogorov par les deux propriétés suivantes:

**Proposition 1.3.1** *i/ pour tout* $F \subset \mathbb{R}^n$, $dim_H(F) \leq dim.inf_K(F) \leq dim.sup_K(F)$.
*ii/ pour* $F$ *ensemble de Cantor:* $dim_H(F) = dim_K(F)$.

Remarquons que $\nu(A) = \lim_{r\to 0} N(r, A)r^s$ n'est pas une mesure de $\mathbb{R}^n$. Cela retire à la capacité de Kolmogorov des propriétés de régularité que possède la dimension de Hausdorff. Cependant, l'application aux cas concrets de la capacité (notamment par l'utilisation de la méthode de "décompte de mailles") est incontestablement plus aisée que celle de la dimension de Hausdorff

### Dimension de plongement

Pour l'heure, nous sommes restés dans un cas idéal: avec la connaissance a priori de la dimension de l'espace des phases. Mais ce n'est pas le cas général ou, en tous cas, la situation expérimentale. On possède plutôt des données numériques, parfois multidimensionelles, provenant souvent d'appareils de mesure à champs restreints. Un cas fréquent est celui où l'on ne dispose que d'une série finie de données numériques réelles. Pourtant, le système dynamique peut avoir été généré par une équation dans un espace des phases multidimensionel, et si un attracteur existait, il serait plongé dans cet espace. On peut penser à un espace des phases en mécanique statistique, dans lequel sont pris en compte les positions et vitesses des points d'une trajectoire, et sur lequel on ne mesurerait que la norme de la vitesse. L'utilisation d'un plongement autorise à revenir dans un espace de dimension suffisante pour qu'on puisse étudier le système.

**Définition 1.3.4** *Un* **plongement** *d'une variété* $M$ *dans une autre variété* $N$*, est une fonction* $g : M \to N$ *continuement différentiable, telle que* $g(M)$ *soit homéomorphe à* $M$ *et telle que* $g$ *soit une immersion (c'est-à-dire que la matrice jacobienne de* $g$ *en tout point de* $M$ *a un déterminant de signe constant et transforme* $M$ *en un fermé). Si un plongement existe de* $M$ *dans* $N$*, on dit que* $M$ *peut être plongé dans* $N$*.*

Trois théorèmes majeurs vont permettre de revenir à une situation où l'étude du système est possible, c'est-à-dire au cas où l'espace des phases considéré a une dimension suffisante pour que les attracteurs puissent y être plongés. En premier lieu, le **théorème de Whitney**:

**Théorème 1.3.1** *Si $M$ est une variété séparable compacte différentiable de dimension $m$, il existe un plongement $g : M \to \mathbb{R}^{2m+1}$ tel que $g(M)$ soit une sous-variété fermée différentiable de $\mathbb{R}^{2m+1}$.*

Ce théorème signifie que quelque soit la variété compacte différentiable $M$ de dimension $m$, on pourra la décrire au minimum dans un espace des phases de dimension $2m+1$ (minimum au sens où une dimension comprise entre $m$ et $2m+1$ peut suffire). En conséquence, on est sûr de pouvoir paramétrer $M$ avec $2m+1$ coordonnées réelles. Nous aimerions trouver une propriété de même teneur valable pour des dimensions non-entières, comme celles de certains attracteurs étranges; c'est ce à quoi pourvoit le **théorème de Mané**:

**Théorème 1.3.2** *Si $A$ est un compact de $E$, sous-espace de dimension finie d'un espace de Banach $B$, tel que $dimE > 2dim_K(A)+1$, alors l'ensemble des projections $\pi : B \to E$ telles que $\pi_{/A}$ soient injectives, est dense dans l'ensemble des projections de $B$ dans $E$.*

Ainsi, un attracteur $A$ compact pourra toujours être plongé dans un espace de dimension supèrieure à $2dim_K(A)+1$, et donc dans un tel espace, il existera un paramètrage pour $A$. Ainsi, une partie de notre travail de "reconstruction" est effectué: si on a un espace des phases de dimension suffisante, on pourra y plonger notre attracteur, quelle que soit sa dimension si tant est qu'elle soit finie. Mais comment, à partir d'une suite de données qui peuvent être simplement de dimension 1? Le **théorème de Takens** répond à une telle question:

**Théorème 1.3.3** *Si $M$ est une variété compacte de dimension $m$, si $f : M \to M$ est un $C^2$-difféomorphisme, et si $y$: $M \to \mathbb{R}$ est une fonction numérique $C^2$, alors $\Phi : M \to \mathbb{R}^{2m+1}$ telle que: $\Phi(x) = [y(x), y(f(x)), y(f^2(x)), ..., y(f^{2m}(x))]$ est un plongement de $M$ dans $\mathbb{R}^{2m+1}$.*

Ainsi, ne possédant que des scalaires, il est possible de déterminer un attracteur de dimension finie quelconque. L'utilisation d'une telle fonction $\Phi$ constitue la **méthode des retards**. Il convient donc de chercher la dimension de l'espace des phases permettant à un attracteur d'y être plongé, et donc d'être paramètré. En conséquence, on définit:

**Définition 1.3.5** *La* **dimension de plongement** *associée à une série numérique $(x_n)_{n\in\mathbb{N}}$ est la dimension entière minimale $d_0$ telle qu'il existe une fonction $F : \mathbb{R}^{d_0} \to \mathbb{R}$ vérifiant:*

$$x_{n+1} = F(x_n, x_{n-1}, ..., x_{n-d_0+1}) \text{ pour tout } n \in \mathbb{N}, \text{ pour tout } (x_{d_0-1}, ..., x_0)$$

*Alors, le vecteur $X_n =^t (x_n, x_{n-1}, ..., x_{n-d_0+1})$ pourra vérifier une équation de type $X_{n+1} = f(X_n)$ pour tout $n$ et tout $X_0$.*

La recherche de cette dimension sera souvent la première étape obligée dans l'étude d'un système dynamique dont on ne connaît pas la dimension de l'espace des phases et dont on ne dispose que de données partielles issues des vecteurs d'états permettant la compréhension du système.

## 1.3.2 Notions liées à la dynamique du système

### Exposants de Lyapunov

La dépendance sensitive aux conditions initiales, requise dans notre définition d'un système chaotique était restée sous une forme assez qualitative. Elle peut, sous certaines hypothèses, être démontrée et mesurée. Plus précisèment, si on considère un système dynamique $X_{n+1} = f(X_n)$ avec $f : M \to M$, où $M$ est une variété différentiable de dimension $m$, tel que $f$ soit différentiable sur $M$. On appelle $\mu$ une mesure invariante ergodique sur $A$, attracteur de $M$, pour $f$, et on note $\| \ \|$ une norme sur $M$, $\dfrac{\delta f(x_0)}{\delta x}$ la matrice jacobienne de $f$ en $x_0$.

**Théorème 1.3.4** *Sous les hypothèses précèdentes, et en supposant que*

$$\int_A \max(0, \log \| \frac{\delta f(x)}{\delta x} \|) d\mu(x) < +\infty,$$

*alors il existe $m$ réels $\lambda_i$ tels que:*

$$\lambda_i \in Sp \ \{ \lim_{n\to\infty} \frac{1}{2n} \log( \frac{\delta f^n(x_0)}{\delta x})^t (\frac{\delta f^n(x_0)}{\delta x}))\}, \ \textit{pour } \mu - \textit{presque tout } x_0 \in A. \qquad (8)$$

Les $\lambda_i$ sont appelés les **exposants de Lyapunov** du système, généralisations des exposants caractéristiques imaginés par Poincaré il y a un siècle. On notera $E_i$ les sous-espaces associés aux $\lambda_i$. L'existence de ces réels est garantie par le **théorème multiplicatif ergodique** de Oseledec, généralisation du théorème ergodique de Birkhoff à plusieurs dimensions. De plus, ce théorème implique l'indépendance des $\lambda_i$ par rapport aux conditions initiales en raison de l'ergodicité de la mesure. Cependant, les exposants de Lyapunov sont fonction de la mesure choisie et choisir une mesure non ergodique les rend dépendants des conditions initiales. Les conditions d'existence de ces exposants sont en fait toujours satisfaites sur un ensemble attracteur, donc a fortiori sur un attracteur inclus dans un ensemble attracteur. Ils représentent la "mesure" de la divergence (ou convergence) des trajectoires, à savoir que deux trajectoires très proches, suivant une direction parallèle à un $E_i$, divergeront si $\lambda_i$ est positif, convergeront s'il est négatif.
**L'existence d'un exposant de Lyapunov strictement positif atteste de la dépendance sensitive aux conditions initiales au moins dans une direction,**

**et donc du caractère chaotique du système.** Ainsi, l'exposant de valeur maximale dans $\mathbb{R}$ joue un rôle important. S'il est positif strictement, on est certain que le système est chaotique, $A$ étant alors un attracteur étrange. S'il est nul, l'attracteur pourra être un cycle limite ou un attracteur quasipériodique (sans pouvoir être " étrange"), s'il est strictement négatif, l'attracteur sera un unique point fixe (pour plus de details, voir Ruelle (1989a)).

### Variétés stables, variétés instables

Nous l'avons évoqué précédemment, il n'y a pas divergence exponentielle des trajectoires pour toutes les directions possibles de l'espace des phases, ou alors le système ne serait plus dissipatif. D'ailleurs, l'aspect dissipatif se retrouve dans la relation:

$$\lim_{n\to\infty} \frac{1}{n} \log \det \frac{\delta f^n(x_0)}{\delta x} = \sum_{i=1}^{m} \lambda_i \ < 0, \quad \text{pour } \mu - \text{ presque tout } x_0 \in A, \tag{9}$$

qui traduit la contraction des $m$-volumes. Il existe donc des directions vectorielles singulières dans l'espace des phases. On en déduit ainsi:

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \log[(\frac{\delta f^n(x)}{\delta x}).u] \ , \quad \text{pour } \mu - \text{ presque tout } u \in E_i^x,$$

avec $E_i^x$ sous-espace associé à $\lambda_i$, pour le point $x$ (en effet, en chaque point $x$ de $A$ la matrice jacobienne "change", les sous-espaces sont donc dépendants de $x$, même si les exposants ne le sont pas). On peut aussi travailler avec des variétés, définies comme suit:

**Définition 1.3.6** *On appelle* **variété stable en $x$ associée à** $\lambda_i$*, l'ensemble* $VS_i^x$*, tel que, pour* $\lambda_i \leq 0$*:*

$$VS_i^x = \{y \in M \ / \lim_{n\to\infty} \frac{1}{n} \log d(f^n(x), f^n(y) \leq \lambda_i.$$

*De la même manière, on appelle* **variété instable en $x$ associée à** $\lambda_i$*, l'ensemble* $VI_i^x$*, tel que pour* $\lambda_i > 0$*:*

$$VI_i^x = \{y \in M \ /\exists y_{-n}, x_{-n} \in M/\forall n \in \mathbb{N} \ y = f^n y_{-n}, \ x = f^n x_{-n} \ \ et \ \lim_{n\to\infty} \frac{1}{n} \log d(x_{-n}, y_{-n}) \leq -\lambda_i\}.$$

*On appellera simplement* **variété instable en** $x$*, l'ensemble*

$$VI^x = \{y \in M \ /\exists y_{-n}, x_{-n} \in M/\forall n \in \mathbb{N} \ y = f^n y_{-n}, \ x = f^n x_{-n} \ et \ \lim_{n\to\infty} \frac{1}{n} \log d((x_{-n}, y_{-n}) < 0\}.$$

Ces variétés sont tangentes en $x$ aux sous-espaces précédemment définis. On a encore:

**Théorème 1.3.5** *Si $A$ est aussi un ensemble attracteur, et si $x \in A$, alors $VI^x \subset A$.*

(voir encore Eckmann et Ruelle (1985)).

### 1.3.3 Notions liées à la dynamique d'un système et à sa structure géométrique

Dans les paragraphes précédents, nous nous sommes attachés à décrire des dimensions géomètriques, c'est-à-dire des dimensions ne prenant en compte que la structure géomètrique de l'attracteur. Il existe aussi des dimensions qui sont dépendantes à la dynamique du système. Plus précisément, ces dimensions sont fonction d'une mesure de probabilité invariante sur l'attracteur considéré, et prennent donc en compte le "temps" passé en chaque lieu de cet attracteur. A cet effet, Rényi a introduit tout un spectre de dimensions, définies par:

**Définition 1.3.7** *Soit $\mu$ une mesure de probabilité sur $M$ un espace métrique séparable muni d'une distance $d$. Si $B(x, r)$ est la boule de centre $x$ et de rayon $r$, et pour $q \in \mathbb{R}$, on appelle dimension de Rényi d'ordre $q$, le réel $D(q)$, si il existe, tel que:*

$$D(q) = \frac{1}{q-1} \lim_{r \to 0} \frac{\log \int_X \mu(B(x,r))^{q-1} d\mu(x)}{\log r}.$$

En fait, ici encore on peut définir des limites supérieures et inférieures à la place de la limite précédente, pour obtenir $\bar{D}(q)$ et $\underline{D}(q)$ (Pour plus de détails sur ces dimensions, voir Cutler (1991), Ding *et al.* (1993), ou encore Pesin (1993)). Nous allons présenter les dimensions les plus utilisées parmi ces dimensions de Rényi, en remarquant que pour $q = 0$, on retrouve une dimension qui ne témoigne que de la structure géométrique de l'attracteur, la capacité de Kolmogorov. En revanche, pour $q = 1$ (en fait, $q \to 1$) et $q = 2$, on obtient respectivement les dimensions d'information et de corrélation, qui sont liées à la dynamique du système.

#### Dimension d'information d'une mesure de probabilité

La dimension d'information peut donc se définir à partir de la définition 1.3.7. On peut aussi la définir à partir de la dimension de Hausdorff, et:

**Définition 1.3.8** *Si $P$ est une mesure de probabilité invariante sur l'ensemble $F \subset M$ où $M$ est un espace métrique, la* **dimension d'information** *de $P$ se définit par:*

$$dim_H(P) = \inf_{B \subset M} \{dim_H(B)/P(B) = 1\}$$

Pour un attracteur $A$ sur lequel est définie la mesure de probabilité invariante $\mu$, on a:

**Proposition 1.3.2** $dim_H(\mu) \leq dim_H(supp(\mu)) \leq dim_H(A)$

Une autre propriété intéressante va mettre en relation la dimension d'information et une autre dimension, la dimension ponctuelle, qui est beaucoup plus facilement calculable.

### Dimension ponctuelle

**Définition 1.3.9** *Si $\mu$ est une mesure de probabilité sur $M$ espace métrique de dimension finie, et si $B_x(r)$ est la boule fermée de centre $x$ et de rayon $r$, on appelle* **dimension ponctuelle** *de $\mu$ en $x \in M$, le réel $\alpha_x$, s'il existe, tel que:*

$$\alpha_x = \lim_{r \to 0} \frac{\log \mu(B_x(r))}{\log(r)}.$$

L'intérêt de cette dimension empirique est sa fonction d'estimateur de la dimension d'information. En effet, Young (1982) montre le théorème suivant:

**Théorème 1.3.6** *Si $\mu$ est une mesure de probabilité sur $M$, variété de dimension finie, et avec $B_x(r)$ boule fermée de centre $x$ et de rayon $r$, alors:*

$$si \ \lim_{r \to 0} \frac{\log[\mu(B_x(r))]}{\log(r)} = \alpha, \ pour \ \mu\text{−}presque \ tout \ x \in M, \ alors \ dim_H(\mu) = \alpha = dim_K(supp(\mu)).$$

Cette égalité entre les deux dimensions est très utiles lorsque l'on a choisi la mesure ergodique sur l'attracteur, auquel cas il suffit de montrer l'existence de la limite en un point pour vérifier le théorème. Nous verrons que la dimension ponctuelle peut facilement être estimée dans la section suivante.

### Dimension de corrélation

Cette dimension, qui est sans doute la plus utilisée expérimentalement, est véritablement apparue dans Grassberger et Proccacia (1983), même si elle est la dimension de Rényi $D(2)$. Nous suivons ici sa définition "rigoureuse" instituée par Pesin (1993):

**Définition 1.3.10** *Si $\mu$ est une mesure de probabilité sur $M$ espace métrique de dimension finie, et si $B_x(r)$ est la boule fermée de centre $x$ et de rayon $r$, on appelle* **dimension de corrélation** *de $\mu$, le réel $\nu$, s'il existe, tel que:*

$$\nu = \lim_{r \to 0} \frac{\log \int_M \mu(B_x(r)) d\mu(x)}{\log(r)}.$$

La dimension de corrélation présente l'avantage de pouvoir être estimée facilement et plus efficacement que la dimension ponctuelle (voir prochaine section). Pesin dans Pesin (1993) montre une généralisation du Théorème 1.3.6, qui relie l'ensemble des dimensions de Rényi dans le cas où la dynamique du système s'inscrit dans la géométrie de l'attracteur:

**Théorème 1.3.7** *Si $\mu$ est une mesure de probabilité borélienne sur $M$ variété de dimension finie, et si $\lim_{r\to 0} \frac{\log[\mu(B_x(r))]}{\log(r)} = \alpha$ pour $\mu$-presque tout $x \in M$, avec $B_x(r)$ boule fermée de centre $x$ et de rayon $r$, alors pour tout $q \in \mathbb{R}$,*

$$dim_H(\mu) = \alpha = \nu = dim_K(supp(\mu)) = D(q).$$

**Dimension de Lyapunov**

Il est aussi possible d'envisager une dimension calculée à l'aide des exposants de Lyapunov du système. La dynamique du sysème semble ici prendre le dessus sur ses caractéristiques géomètriques:

**Définition 1.3.11** *Soient $\lambda_1, ..., \lambda_m$ tels que $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m$, les exposants de Lyapunov d'un système dynamique sur une variété différentiable de dimension $m$, dont une mesure invariante ergodique est $\mu$. La* **dimension de Lyapunov** *de $\mu$ se définit par:*

$$dim_\Lambda(\mu) = k + \frac{\sum_{i=1}^{k} \lambda_i}{\mid \lambda_{k+1} \mid}, \text{ et } k = max\{j \in \mathbb{N} / \sum_{i=1}^{j} \lambda_i \geq 0\}$$

*(si $\lambda_1 < 0$ alors $dim_\Lambda(\mu) = 0$).*

Un théorème (Ledrappier) va nous permettre de comparer cette dimension avec la dimension d'information.

**Théorème 1.3.8** *Si $f$ est deux fois continuement différentiable, et si $\mu$ est une mesure ergodique de support compact, alors:*

$$dim_H(\mu) \leq dim_\Lambda(\mu).$$

Une autre relation de Ledrappier permet de lier la dimension de Lyapunov aux dimensions de Hausdorff et d'information, puisque:

**Théorème 1.3.9** *Si $A$ est un compact invariant pour la fonction $f$ différentiable, et si $M(A)$ est l'ensemble des mesures ergodiques de support $A$, alors:*

$$sup\{dim_H(\rho) : \rho \in M(A)\} \leq dim_H(A) \leq \sup\{dim_\Lambda(\rho) : \rho \in M(A)\}.$$

**Entropie de Kolmogorov**

Ce que l'on appelle invariant de Kolmogorov-Sinai, ou taux moyen de création d'information, ici nommée entropie de Kolmogorov, n'est pas sans rapport avec la notion de dimension. Si $\mu$ est une mesure invariante pour $f$, alors l'entropie de Kolmogorov de $\mu$ mesure le taux asymptotique de création d'information issue de l'itération de $f$. De manière plus rigoureuse:

**Définition 1.3.12** *Soit $\mu$ une mesure invariante pour $f : M \to M$ où $M$ est un espace metrique, de support compact, dont $S = (S_1, S_2, ..., S_k)$ est une partition ($\mu$ -mesurable) finie. Si $f^{-l}(S_i)$ désigne l'image réciproque de $S_i$ par $f^l$, appelons $S^{(j)}$ la partition composée des $(S_{i_1} \cap f^{-1}(S_{i_2}) \cap ... \cap f^{-j+1}(S_{i_j}))$ avec $i_s = 1, 2, .., k$ pour $s = 1, 2, .., j$. Alors* **l'entropie de Kolmogorov** *de $\mu$, $h(\mu)$, est donnée par:*

$$h(\mu) = \lim_{diam(S) \to 0} \lim_{j \to \infty} \frac{H(S^{(j)})}{j} \quad avec\ H(S) = \quad - \quad \sum_{i=1}^{k} \mu(S_i) \log(\mu(S_i)).$$

Les deux théorèmes suivants (le premier de Ruelle, le second de Ledrappier et Young) illustrent Les rapports entre l'entropie de Kolmogorov et les autres dimensions:

**Théorème 1.3.10** *Si $\mu$ est une mesure ergodique de support compact associée à $f$, fonction différentiable sur une variété de dimension finie, et si les $\lambda_i$ sont les exposants de Lyapunov pour $\mu$, alors:*

*(1)* $h(\mu) \leq \sum_i (\lambda_i > 0)$.

*(2) si $f$ est un difféomorphisme deux fois différentiable, et si $\mu$ est absolument continue sur les variétés instables, alors: $h(\mu) = \sum_i (\lambda_i > 0)$.*

## 1.4 Traitements statistiques du chaos

Un travail statistique sur un système dynamique pouvant être chaotique consiste essentiellement en l'estimation des grandeurs le caractérisant (essentiellement les exposants de Lyapunov et les dimensions présentées précédemment). L'intervention de techniques statistiques répond aux difficultés que représente une détermination théorique de ces grandeurs. Ainsi, la démonstration rigoureuse de la chaoticité du systéme, à savoir l'existence d'un attracteur et d'un exposant de Lyapunov strictement positif, est souvent impossible (en témoigne l'article de Benedicks et Carleson (1991) sur l'attracteur de Hénon). De même, sauf en de rares cas découlant d'une autosimilarité reconnue, le calcul des différentes dimensions n'est pas réalisable.

Aussi doit-on se contenter d'estimer ces notions à partir de simulations numériques de trajectoires. Le fondement commun aux différentes techniques employées est la simulation de trajectoires issues de différentes conditions initiales, reparties suivant une distribution uniforme dans un ensemble contenu dans le bassin d'attraction. Ce choix est légitime lorsqu'on choisit d'étudier le système à partir d'une mesure invariante ergodique. Les différentes estimations se font sur chaque trajectoire (prise suffisamment "longue"), puis elles sont moyennées sur l'ensemble des trajectoires.
Cependant, nous verrons que de nombreux auteurs identifient les données issues de chaque trajectoire à des variables aléatoires. Ils appliquent alors des techniques probabilistes pour estimer exposants ou dimensions, en faisant certaines hypothèses (indépendance,..) sur ces variables. Dans ce cas, comme dans le précédent dans lequel on distribue les conditions initiales aléatoirement, nous verrons que la tentation est grande d'utiliser les résultats obtenus pour distinguer les systèmes dynamiques dont la trajectoire est déterministe, d'une série de variables alétoires.

### 1.4.1 Estimation des exposants de Lyapunov

Deux approches distinctes permettent l'estimation des exposants de Lyapunov. Le plus souvent, on se contente d'estimer le plus grand de ces réels, dont le signe, relatif au comportement de deux trajectoires initialement voisines, décidera de la nature du système: la positivité de cet exposant est une condition nécessaire, dans le cas d'un système suffisamment régulier, à son caractère chaotique.

#### Estimation directe

La première démarche est une estimation directe des exposants de Lyapunov. Son point de départ est la donnée de $X_1, X_2, .., X_N$, $N$ vecteurs de l'espace des phases de dimension $m$, et l'utilisation de la définition du plus grand des exposants de Lyapunov comme l'espérance de $\log \| \frac{\delta f(x)}{\delta x} \|$ pour $X$ distribué selon une mesure invariante ergodique sur l'attracteur. Wolff, dans Wolff (1992), construit une estimation de cette espérance distincte de la moyenne empirique qui n'a pas de sens en raison des taux de variations sur les variétés stables et de l'aspect local des résultats. Ainsi, dans le cas où la dimension $m = 1$, Wolff utilise la propriété:

$$\lambda(X_0) = \lim_{n\to\infty} \frac{1}{n} \log | \frac{\delta f^n}{\delta x}(X_0)|,$$

et il en déduit, sur un petit domaine, l'expression d'un **exposant de Lyapunov local** $\lambda(X_i, m, h)$ tel que:

$$\lambda(X_i, n, h) = \frac{1}{n} \frac{1}{Card(S_i)} \sum_{j\in S_i} \log \| \frac{X_{i+n} - X_{j+n}}{X_i - X_j} \|,$$

pour $n$ entier inférieur à $N$, $h$ positif, $i = 1, ..., N-n$ et avec $S_i = \{j/0 < |X_i - X_j| \leq h\}$. Wolff simplifie cette expression en utilisant $\lambda'(X_i, n, h)$ tel que:

$$\lambda'(X_i, n, h) = \frac{1}{n} \frac{1}{Card(S_i)} \sum_{j \in S_i} \log \| \frac{X_{i+n} - X_{j+n}}{h} \|.$$

Il justifie cette substitution par le fait que pour des données réelles $X_i$ considérées comme des réalisations de variables aléatoires indépendantes et identiquement distribuées avec une densité deux fois dérivables, alors:

$$E(\lambda(X_i, n, h)) - E(\lambda'(X_i, n, h)) = \log(2) + \frac{1}{12} + O(h^2),$$

pour $h$ pris suffisamment petit. Wolff, dans ses applications numériques, prend des modèles stochastiques, et un "temps" de dérivation $n$ très petit (1 ou 2). Il peut alors tracer pour chaque modèle la courbe donnant les exposants locaux en fonction de la position dans l'attracteur, qu'il reprend en faisant varier $n$. Il donne une méthode de détermination de la non-linéarité par le tracé de la courbe représentative de la diffèrence théorique entre les exposants locaux pour un système déterministe et pour le même système avec un bruit, en fonction de la diffèrence empirique; on obtiendrait ainsi, quand il y a linéarité l'assurance de la non-linéarité du système... Vis-à-vis du chaos, Wolff indique une méthode utilisant les exposants locaux permettant de distinguer système déterministe et processus stochastique. Il propose sur la suite d'observations une division en deux parts. On découpe chacune de ces parts en petits intervalles réguliers sur lesquels on calcule les exposants locaux. Enfin, on identifie les deux exposants venant d'un même intervalle, que l'on place sur un graphe donnant l'un en fonction de l'autre. Il semble (?) que les systèmes déterministes présentent une linéarité non observée pour les processus stochastiques. Cela dit, il reste à étendre ces résultats à des dimensions supérieures et à donner des justifications théoriques aux observations précédentes.

Eckmann et Ruelle (1985) suggère un autre mode d'estimation "directe" des exposants de Lyapunov. Leur propos repose sur l'utilisation d'un moindre carré pour estimer la matrice jacobienne exprimant les divergences/convergences de trajectoires voisines après quelques itérations du système. Ne sachant pas si les $X_i$ sont bien des vecteurs d'états (c'est-à-dire que l'on ne sait pas si la dimension de plongement est inférieure à $m$), une première étape consiste ùtiliser le théorème des retards de Tackens pour travailler avec des données de dimension suffisante pour paramétrer le système. Pour ce faire, ils considèrent les vecteurs $Y_{i\tau}^d$ de $\mathbb{R}^{dm}$, où $Y_{i\tau}^d = (X_{(i+d-1)\tau}, ..., X_{i\tau})$ en choisissant un temps "d'échantillonage" $\tau$ entier, ni trop petit (pour des raisons de calcul numérique), ni trop grand (pour que la divergence de trajectoire est une forme continue, que ce ne soit pas un "billard"). Fixant un $i$ quelconque, ils recherchent l'ensemble des $Y_{j\tau}^d$ tels que $\|Y_{j\tau}^d - Y_{i\tau}^d\| \leq r$ pour un $r$ suffisamment petit, et tels que $\|Y_{(j+1)\tau}^d - Y_{(i+1)\tau}^d \| \leq r$.

Ils recherche alors une matrice $T^d_{i\tau}$ d'ordre $dm$ telle que, pour les $j$ sélectionnés, on ait:

$$T^d_{i\tau}(Y^d_{j\tau} - Y^d_{i\tau}) \simeq (Y^d_{(j+1)\tau} - Y^d_{(i+1)\tau}).$$

Si $N$ est suffisamment grand, on peut supposer qu'il y a un nombre de $j$ plus grand que $(dm)^2$. Eckmann et Ruelle choisissent donc d'estimer $T^d_{i\tau}$ par moindres carrés. Les résultats ne sont justifiés que pour les exposants positifs, l'estimation des autres étant, par nature de la démarche, sans valeur.

**Estimation indirecte**

Le deuxième mode d'estimation des exposants de Lyapunov, nous le qualifierons d'indirect. Il se fonde sur une reconstruction de la fonction $f$ à partir de $X_1, X_2, .., X_N$, qui sont $N$ réels (nous reprenons ici les articles de Mac Caffrey *et al.* (1992) et Nychka *et al.* (1992)). Comme précdemment, ne connaissant pas la dimension de plongement $d_0$, une première étape est de considérer les $Y^d_i$ tels que $Y^d_i = (X_{(i+d-1)}, ..., X_i)$. En fait, le but est l'estimation de la matrice jacobienne de $f$ en tout vecteur $Y^d_n$, ce qui permet une estimation consistante du plus grand des exposants de Lyapunov:

**Théorème 1.4.1** *On suppose:*

- *il existe $f^{(d)}$, une fonction différentiable admettant un attracteur et une suite de vecteurs aléatoires $(E^d_n)_{n\in\mathbb{N}}$ avec $E^d_n =^t (e_n, ..., e_{n-d+1})$, les $e_n$ étant des variables aléatoires réelles indépendantes et uniformément distribuées, telles que les $Y^d_n$ forment une chaîne de Markov telle que $Y^d_{n+1} = f^{(d)}(Y^d_n) + \sigma E^d_n$, avec $\sigma \in \mathbb{R}^+$.*
- *$Y^d_0$ est distribué suivant $\mu$, mesure invariante de cette chaîne.*
- *les $J^d_n = \frac{\delta f^{(d)}(Y^d_n)}{\delta x}$ forment un processus stochastique stationnaire ergodique, tel que $\sup_{n\in\mathbb{N}} \|J^d_n\| < +\infty$ et $\liminf_{n\to\infty} \dfrac{1}{n} \log \|J^d_n J^d_{n-1} ... J^d_0\| > -\infty$ avec une probabilité de 1.*
- $\displaystyle\int_A \max(0, \log \|\frac{\delta f^{(d)}(x)}{\delta x}\|) d\mu(x) < +\infty.$
- *il existe une suite $(\beta_n)$ tendant vers zéro, et une estimation consistante de $J^d_n$, $G^d_n$, telle que $\max_{n\leq N} \|G^d_n - J^d_n\| = O(\beta_N)$ quand $N \to +\infty$.*

*Alors, si $\lambda^d_M = \dfrac{1}{M} \log \|J^d_M J^d_{M-1} ... J^d_0\|$, et $\lambda'^d_M = \dfrac{1}{M} \log \|G^d_M G^d_{M-1} ... G^d_0\|$, on a:*

$$\lambda^d_M - \lambda'^d_M = O(\beta_N^{1/M}) \text{ quand } M, N \to +\infty \text{ et } M\beta_N \to 0. \tag{10}$$

Pour donner un ordre de consistance plus exact, se déduit le corrolaire suivant:

**Corollaire 1.4.1** *Sous les hypothèses du théorème précédent, et si* $\alpha = \sup_{n\in\mathbb{N}} \|J_n^d\|$ *vérifie* $M\beta_N(\alpha/\gamma)^M \to 0$ *quand* $M, N \to +\infty$ *pour* $0 < \gamma < \sup_{k\in\mathbb{N}} e^{\lambda'^d_k}$*, alors:*

$$\lambda_M^d - \lambda'^d_M = O(\beta_N(\alpha/\gamma)^M) \text{ quand } M, N \to +\infty.$$

dff Il est important de remarquer que les $\lambda_M^d$ convergent, quand $M \to +\infty$, vers le plus grand des exposants de Lyapunov pour la fonction $f^{(d)}$ considérée. Il suffit alors d'utiliser des méthodes classiques d'estimation non-paramètriques pour reconstruire la fonction $f$, et donc obtenir une estimation de sa matrice jacobienne, connaissant les $Y_n^d$ pour $n \in \{0, 1, ..., N - d_0 + 1\}$. Mac Caffrey *et al.* (1992) proposent ainsi quatre approches différentes, une "locale" et trois "globales". La reconstruction de $f$ par splines locaux à plateaux étroits est la méthode de reconstruction locale. $f_{est}^{(d)}$ est une combinaison linéaire de monômes de degré (somme des degrés de chaque variable) inférieur à $k$ et de $N$ fonctions de base radiales, les coefficients étant déterminés par résolution de $N$ équations linéaires. La reconstruction de $f$ par fonctions de bases radiales est une globalisation de la méthode précédente. La méthode de reconstruction de $f$ à l'aide d'une poursuite par projection s'effectue de la manière suivante: on considère $(\alpha_1, .., \alpha_K) \in (\mathbb{R}^d)^K$ et on définit $h : (\mathbb{R}^d)^{K+1} \to \mathbb{R}$ par:

$$h_{est}(Y^d, \alpha_1, ..., \alpha_K) = \sum_{k=1}^{K} G_k({}^t\alpha_k Y^d),$$

où les $G_k : \mathbb{R}^d \to \mathbb{R}$ sont des fonctions minimisant:

$$L(h) = \sum_{i=1}^{N} [X_i - h(Y_i^d, \alpha_1, ..., \alpha_K)]^2 + \mu \sum_{k=1}^{K} \int_{\mathbb{R}^d} \left( \frac{\delta^2 G_k(u)}{\delta x^2} \right)^2 du,$$

en choisissant les $\alpha_k$ tels que $\sum_{i=1}^{N} [X_i - h_{est}(Y_i^d, \alpha_1, ..., \alpha_K)]^2$ soit minimum. Enfin, la reconstruction de $f$ peut se faire à l'aide d'un réseaux de neurones. Il s'agit alors de reconstituer la fonction comme un "système" transformant les entrées $Y_n$ en une fonction $f_{est}^{(d)}$ par l'intermédiaire de variables cachées, $u_n$ qui sont des combinaisons linéaires des entrées et dont le nombre reste à choisir, et d'une fonction non-linéaire $G$. Pour tout $Y \in \mathbb{R}^d$, on a:

$$f_{est}^{(d)}(Y) = \sum_{j=1}^{K} \beta_j G(u_j + \mu_j) \ = \sum_{j=1}^{K} \beta_j G(\gamma_j^t Y + \mu_j),$$

avec $G(u) = \dfrac{e^u}{1 + e^u}$, et $\gamma_j \in \mathbb{R}^d$. Les coefficients $\beta_j$, $\gamma_j$ et $\mu_j$ sont déterminés par moindres carrés non-linéaires (on minimalise les sorties par rapport aux observations), le nombre $K$ étant fixé a priori. L'avantage notoire de cette méthode est son indépendance vis-à-vis de la dimension de l'espace d'état quand celle-ci est suffisamment grande. De plus, le choix du nombre de variables cachées $K$ n'intervient que lorsque l'échantillon des données observées est maigre.

Dans la continuation de ces travaux, Bailey *et al.* (1997) estiment indirectement les exposants de Lyapunov locaux pour un système dont l'espace des phases est de dimension $d$ (ce qui généralise l'approche unidimensionnel de Wolff (1992)). Ainsi, quand le système s'écrit encore $Y_{n+1}^d = f^{(d)}(Y_n^d) + \sigma E_n^d$ avec $E_n^d = (\varepsilon_n, 0, .., 0)$ où les $\varepsilon_n$ sont des v.a.i.i.d telles que $\mathbb{E}\varepsilon_n = 0$ et $\text{var}\varepsilon_n = \sigma^2$, alors l'exposant de Lyapunov local au rang $n$ et après $m$ itérations, se définit par:

$$\lambda_m(n) = \frac{1}{m} \log \parallel J_{n+m-1} J_{n+m-2} .. J_n U_0 \parallel,$$

avec $J_n$ la matrice jacobienne de $f^d$ calculée en $Y_n^d$, et $U_0$ un vecteur unitaire pour la norme euclidienne $\parallel . \parallel$. Lorsque $m \to +\infty$, $\lambda_m(n)$ converge en probabilité vers le plus grand des exposants de Lyapunov $\lambda$ du système. Bailey *et al.* (1997) donne une possibilité d'avoir des intervalles de confiance pour $\lambda$ en deux temps. Ils montrent d'abord un théorème centrale limite en $m$ pour $\lambda_m$, à partir de résultats sur les chaînes de Markov. Ensuite, ils utilisent une estimation de la matrice jacobienne $J_n$ à l'aide de la méthode de reconstruction par réseaux de neurones de la fonction $f^{(d)}$, dont les paramètres sont estimés par maximum de vraisemblance. L'application de ces techniques est assez concluante sur des données numériques constituées par le nombre de cas de rougeoles à New-York et à Copenhague.
Lu et Smith (1997) proposent aussi d'estimer les exposants de Lyapunov locaux en utilisant une estimation de la matrice jacobienne après $m$ itérations pour un modèle similaire à celui étudié précédemment. Pour un nombre d'itérations $m$, ils définissent $d$ exposants de Lyapunov locaux qui sont des estimations des $d$ exposants de Lyapunov du système (on ne s'intéresse plus seulemnt au plus grand des exposants). L'estimation de la matrice jacobienne après $m$ itérations est obtenue par des techniques de régressions linéaires et quadratiques locales, qui ne nécessitent pas la reconstruction de $f^{(d)}$. Cela leur permet aussi de déterminer le biais et la variance des estimateurs des exposants de Lyapunov locaux, qui tendent vers 0 lorsque $m$ et $n$ tendent vers l'infini ($n$ étant le nombre d'observations).
Ajoutons enfin la démarche de Bosq (1995) et Bosq et Guégan (1995) consistant en l'estimation de la densité d'une mesure de probabilité invariante associée au système. L'estimation se fait par des méthodes non paramétriques notamment avec la méthode d'estimation par noyaux. Ceci permet de donner en plus des vitesses de convergence. Bosq et Guégan (1995) proposent encore l'estimation de la fonction $f^d$, ce qui permet d'obtenir des résultats sur les exoposants de Lyapunov.

### 1.4.2 Estimation d'une dimension "géomètrique" de l'attracteur

Nous avons mis en évidence deux types de dimensions géométriques: la dimension de Hausdorff et la capacité de Kolmogorov. Pour estimer cette dernière, la méthode qui prévaut est l'estimation de la dimension "box-counting", qui est directement calculable à partir de $X_1, .., X_N$, $N$ vecteurs d'états (nous suivons ici Falconer (1990)). On recouvre les $X_i$, que l'on suppose asymptotiquement contenu dans un attracteur $A$, par un maillage rectangulaire de taille $r$, et on compte le nombre $N(r)$ de mailles contenant au moins un $X_i$. On effectue ce travail pour un certain nombre de $r$ pris suffisamment petits et on estime la dimension par régression des $\log N(r)$ par les $\log(1/r)$. On espère ainsi pouvoir obtenir un estimateur de $dim_K(A)$.
En réalité se trouve posé un problème lié à la définition même de la capacité, à savoir que le calcul de la limite précédente présuppose que $N(r) \sim r^{dim_K(A)}$, mais la démarche proposée donne le même résultat si on a $N(r) \sim Cr^{dim_K(A)}$ et même si $N(r) \sim H(r)r^{dim_K(A)}$, avec $C$ une constante quelconque, et $H(r)$ une fonction à variations lentes. C'est ce que l'on appelle le cas lacunaire, qui modèlise d'ailleurs mieux les "fractales", qui possédent souvent de larges régions "vides". Un second problème, plus empirique, se pose avec les tailles des mailles de plus en plus petites qui nécessitent un nombre croissant d'observations, non toujours disponibles.
Pour dépasser ces difficultés, le maillage peut être transformé. Ainsi, Hall et Wood (1993), dans le cas d'un espace d'état de dimension deux, et avec une modèlisation du système par un processus gaussien, redécoupent le maillage initial deux fois pour finalement effectuer un moindre carré sur des estimateurs locaux de la dimension. De par la forme du modèle utilisé, ils en déduisent aussi un estimateur de la variance du processus en certains points du maillage et le biais de l'estimateur de la dimension.

### 1.4.3 Estimation d'une dimension liée à une mesure invariante

#### Estimation de la dimension d'information: dimension ponctuelle

Nous l'avons vu, la dimension ponctuelle peut être identique à la dimension d'information d'une mesure, sous certaines conditions (voir Théorème de Young 1.3.6). L'intêret d'utiliser la dimension ponctuelle est qu'elle peut être facilement estimée par le procédé suivant. On considère toujours $(X_1, .., X_N)$, et les $Y_n^d$ obtenus par la méthode des retards. Si $\mu^d$ est la mesure de probabilité physique invariante induite par les $Y_n^d$, i.e. $\mu^d(\Gamma) = \frac{1}{N}\sum_{i=1}^{N-d+1} \mathbb{I}_{Y_i^d \in \Gamma}$ pour $\Gamma$ un fermé de l'espace des phases, alors la dimension ponctuelle de $\mu^d$ en $Y_i^d$ peut être estimée par:

$$\alpha_i^d(r) = \frac{\log \mu^d(B_r(Y_i^d))}{\log r}, \text{ quand } r \to 0.$$

On obtient une estimation $\hat{\alpha}_i^d$ de la dimension ponctuelle en $Y_i^d$ en calculant $\alpha_i^d(r)$ pour des $r$ suffisamment petits et en effectuant un moindres carrés.
Quand on connaît $d_0$ la dimension de plongement, on estime bien la dimension ponctuelle, et donc la dimension d'information de $\mu = \mu^{d_0}$, qui est une mesure invariante ergodique sur l'attracteur du système. Eckmann et Ruelle (1985) indique une méthode ne nécessitant pas cette connaissance supplémentaire. Leur argument consiste en ce que pour $d \geq dim_H(\mu)$ alors $\hat{\alpha}^d = dim_H(\mu^d) = dim_H(\mu)$, tandis que pour $d \leq dim_H(\mu^d)$ on a $\hat{\alpha}^d = d$. Cette constatation se retrouve expérimentalement: $\hat{\alpha}^d$ devient constant lorsque $d$ croît suffisamment.

### Estimation de la dimension de corrélation

Comme pour la dimension d'information, il est possible d'estimer la dimension de corrélation par des méthodes très simple; en fait, l'estimateur de la dimension de corrélation a d'abord été appelé dimension de corrélation lors de son apparition dans Grassberger et Proccacia (1983), avant que Pesin (1993) nomme dimension de corrélation la dimension de Rényi $D(2)$. L'estimateur proposé par Grassberger et Proccacia se calcule à partir de $(X_1, .., X_N)$, $N$ vecteurs de $M$ (variété de dimension $m$), supposés connus, et en utilisant les $(Y_1^d, .., Y_{n-d+1}^d)$ obtenus par la méthode des retards. On définit alors l'entier de corrélation, $C^d(N, r)$, par:

$$C^d(N, r) = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbb{I}_{(\|Y_i^d - Y_j^d\| \leq r)}, \tag{11}$$

où $\| \, . \, \|$ est une norme sur $M^d$. On veut alors obtenir l'exposant de corrélation, noté $\nu^d$, définit par:

$$\nu^d = \lim_{r \to 0} \lim_{N \to +\infty} \frac{\log C^d(N, r)}{\log r}, \tag{12}$$

lorsque cette limite existe (le choix de la norme n'intervient donc pas). L'exposant de corrélation est une "moyenne" des estimations des dimensions ponctuelles, et il fournit ainsi des résultats plus fiables que la dimension ponctuelle pour l'estimation de la dimension d'information. Mais, cet exposant est aussi une estimation de la dimension de corrélation, par ce théorème de Pesin:

**Théorème 1.4.2** *Soit $\mu^d$ une mesure ergodique invariante pour $f^d$, telle que $Y_{n+1}^d = f^d(Y_n^d)$. Alors il existe un ensemble dénombrable $Q \subset \mathbb{R}$ tel que si $\lim_{N \to +\infty} C^d(N, r)$ existe pour $r > 0$, pour tout $r \in \mathbb{R}^+ - Q$ on ait $\mu^d$-presque sûrement:*

$$\lim_{N \to +\infty} C^d(N, r) = \int_{M^d} \mu^d(B(y, r)) d\mu^d(y). \tag{13}$$

Ainsi, l'exposant de corrélation $\nu^d$ est une estimation de la dimension de corrélation de $\mu^d$, lorsque cette dimension existe. Pour déterminer cet exposant, on calcule $C^d(N,r)$ pour un certain nombre de $r_i$ suffisament petits, puis on obtient $\nu^d$ par régression des $\log C^d(N,r_i)$ en fonction de $\log r_i$ et par moindres carrés (on suppose que $N$ est suffisament grand). Grassberger et Proccacia (1983) proposèrent cette méthode, très souvent reprise depuis (voir Malraison *et al.* (1983), Bergé *et al.* (1984),..), et dont la validation théorique se trouve dans Denker et Keller (1986).
Cutler dans Cutler (1991) propose une méthode pour étudier le cas où la dimension de corrélation de $\mu^d$ existe, mais où:

$$C^d(N,r) \simeq r^{\nu^d} L(r) \text{ quand } r \to 0,$$

avec $L(r)$ une fonction à variations lentes non réduite à une constante (phénomènes de lacunarité). Utilisant un résultat de Denker et Keller (1986) sur la consistence de la méthode de Grassberger et Proccacia, Cutler déterminer un estimateur consistent par moindres carrés de $L(r)$ pour les $r$ considérés. Plusieurs exemples viennent alors confirmer l'intérêt de cette méthode.
Serinko dans Serinko (1994) propose un estimateur de l'exposant de corrélation par moindres carrés pour certains systèmes dynamiques vérifiant une condition de mélange dite " weak Bernoulli " (dont la fonction de Cantor ou de Feigenbaum). En choisissant la vitesse de convergence de $r$ vers 0 en fonction de celle de $N$ vers $+\infty$, Serinko obtient un théorème limite qui ne nécessite pas l'existence a priori de la dimension de corrélation.

Il est aussi possible de concevoir l'exposant de corrélation comme une variable aléatoire déterminée à partir des $Y_n^d$, qui sont des variables aléatoires de mesure de probabilité $\mu^d$. En effet, par le théorème de Pesin, la dimension de corrélation (quand elle existe) se déduit du comportement asymptotique de $P(\| Y - X \| \leq r)$ quand $r \to 0$, avec $X$ et $Y$ deux variables indépendantes identiquement distribuées suivant la mesure de probabilité $\mu^d$. Certains auteurs travaillent alors directement sur cette probabilité en faisant des hypothèses sur la mesure $\mu^d$.
Ainsi, Theiler dans Theiler (1990) propose de déterminer la dimension de corrélation à partir de deux modèles d'indépendance des variables $Y_n^d$. Le premier suppose que les vecteurs $Y_n^d$ sont identiquement distribués et indépendants et c'est le modèle APM. Le second suppose que les interdistances entre les $Y_n^d$ sont identiquement distribuées et indépendantes et c'est le modèle IDH (les vecteurs $Y_n^d$ étant supposés situés sur un ensemble attracteur de $\mathbb{R}^d$, de dimension permettant le plongement d'un attracteur). Les deux modèles sont jugés a priori par l'auteur comme naifs et faux et ne permettent que d'apporter des cas extrèmes aux calculs de dimension. En moyennant les dimensions ponctuelles obtenues sur un grand nombre d'échantillons, Theiler obtient des résultats sur la consistance des exposants de corrélation comme estimateurs de la dimension de corrélation dans les deux cas. Le modèle IDH prédit une erreur statistique (écart type) proportionnelle à $1/N$, alors qu'elle est en $1/\sqrt{N}$ pour le modèle APM. Il en déduit

alors celle pour un modèle dit "exact" (avec des vecteurs aléatoires corrélés), qui est en $\sqrt{A/N + B/N^2}$. En fait, le coefficient A peut s'annuler (cas IDH où l'on a $N^2$ variables indépendantes), ce qui ne rend plus vraiment utilisable l'application directe du théorème central limite, cas se présentant pour des "fractales" (la divergence exponentielle des trajectoires créant une "indépendance" des interdistances) quand $N$ est très grand.
Dans la continuation de ceci, Smith dans Smith (1992) commence par supposer que, pour $d$ suffisamment grand, $C^d(r) = ar^p$ pour $r \leq r_0$. Cela lui permet d'évaluer $C^d(r_i)$ avec $r_i = r_o\phi^i$ et $0 < \phi < 1$, en utilisant le modèle des interdistances indépendantes IDH et un modèle binomial de distribution du nombre d'interdistances $N_{r_i}$ suivant $r_i$. Cela le mène à estimer la dimension de corrélation par maximum de vraisemblance à partir de la distribution des $N_{r_1}, ..., N_{r_k}$ connaissant $N_{r_0}$. Il obtient ainsi un premier estimateur, appelé estimateur binomial, prenant une forme simplifiée pour $\phi \to 1$ et $k \to \infty$, appelée estimateur de Hill. Il évoque encore le cas lacunaire pour lequel Theiler avait déjà montré que l'estimateur proposé ne converge que si la fonction à variations lentes (précédemment évoquée) converge vers une constante quand $r$ tend vers zéro (voir Theiler (1988)). Recherchant les propriétés statistiques de ces estimateurs, il montre ensuite que si les $Y_n^d$ sont des variables aléatoires identiquement distribuées de densité de probabilité deux fois différentiable et échantillonnées avec un retard les "rendant" indépendantes (?), alors $C^d(N, r) = ar^d(1 + gr^2 + o(r^2))$ pour $r \to 0$, et donc la dimension de corrélation vaut $d \in \mathbb{N}$. Quand il utilise un système dynamique déterministe avec un bruit de mesure, un bruit "externe", qu'il choisit blanc et gaussien, la dimension de corrélation est encore entière et de même valeur que la dimension de plongement considérée.
On peut aussi signaler la contribution de Cutler dans Cutler (1994), qui fonde une théorie pour la dimension de corrélation pour des sèries chronologiques issues aussi bien de systèmes déterministes que de processus gaussiens stationnaires continus. Pour cela, Cutler définit pour l'ensemble des processus de $[0, T]$ dans $\mathbb{R}$, noté $\mathcal{C}([\prime, \mathcal{T}])$, une intégrale de corrélation spatiale pour une mesure $\rho$ telle que:

$$\nu_T = \lim_{r \to 0} \frac{\log P(\| X - Y \|_T \leq r)}{\log r},$$

où $X$ et $Y$ sont deux processus de $\mathcal{C}([\prime, \mathcal{T}])$, indépendants, et distribués selon $\rho$, et avec une distance entre deux processus définie par: $\| X - Y \|_T = \sup_{t \in [0,T]} \mid X_t - Y_t \mid$. Lorsque $\nu_T = \nu_{T'}$ pour tout $(T, T') \in \mathbb{R}^2$, alors on appellera dimension de corrélation cette valeur. Cela permet à Cutler de vérifier que cette dimension n'existe pas pour les processus gaussiens stationnaires dont la mesure spectrale n'est pas discrète, qu'elle existe et vaut $2/(\alpha - 1)$ pour les bruits gaussiens de densité spectrale en $\lambda^{-\alpha}$ avec $1 < \alpha < 3$.

La méthode de Grassberger et Proccacia a semblé un temps capable de distinguer les systèmes dynamiques chaotiques de systèmes dynamiques bruités par une perturbation stochastique. En effet, comme le suggère le travail de Smith, et comme semblent

le montrer de nombreuses simulations numériques, quand on augmente la valeur de $d$, l'exposant de corrélation croît avec $d$ pour un système dynamique bruité (de type $Y_{n+1}^d = f^d(Y_n^d) + \varepsilon_n$), alors qu'il reste constant dès que $d$ devient plus grand qu'un $d_0$ dans le cas d'un système dynamique non bruité (de type $Y_{n+1}^d = f^d(Y_n^d)$). Osborne et Provenzale (1989) observèrent que ceci n'est pas toujours vrai, notamment si le bruit a une densité spectrale en $1/f^\alpha$ avec $1 < \alpha < 3$ (phénomène de longue dépendance), ce qui conduit à une dimension fractale non entière et non dépendante de $d$ quand $d \geq d_0$ (nous verrons que $d_0$ correspond à la dimension de plongement). Theiler dans Theiler (1991) apporte une contribution théorique à cette observation, dans le cas où le bruit est gaussien.
Cependant, si la détermination de l'exposant de corrélation ne semble pas permettre de distinguer entre "aléatoire" et "déterministe" (puisque la dimension de corrélation dépend dans les deux cas d'une mesure invariante), elle permet en revanche d'obtenir la dimension de corrélation du système quand un plongement de dimension suffisante est effectué, c'est-à-dire, la dimension de corrélation intéressante pour l'étude du système et relative à un attracteur de ce système. C'est ce que montre Ding *et al.* (1993) par le théorème suivant:

**Théorème 1.4.3** *Si $A$ est un ensemble de $\mathbb{R}^n$ muni d'une probabilité invariante $\mu$. Si $m \geq \nu$, où $\nu = D_2(\mu)$ est la dimension de corrélation de $\mu$, alors pour presque tout plongement $F : \mathbb{R} \to \mathbb{R}^m$, la dimension de corrélation est conservée, soit $D_2(F(\mu)) = \nu$.*

Une conséquence directe de ce résultat est bien le fait que lorsque on augmente la valeur de $d$, la dimension de corrélation de $\mu^d$ croît jusqu'à $\nu$, obtenu pour $d = d_0$, puis reste égale à $\nu$ pour tout $d > d_0$.

### 1.4.4 Estimation de la dimension de plongement

Plusieurs démarches sont possibles pour estimer de la dimension de plongement, que nous appellerons $d_0$. Ces méthodes auront pour point commun de n'être aucunemment spécifiques aux systèmes chaotiques.
La première démarche est une application directe du thréoème 1.4.3. En effet, supposons que l'on ne posséde du système dont l'espace des phases peut être de dimension $m$ quelconque, que des données $(X_1, .., X_N)$ pouvant être réelles. On calcule alors l'exposant de corrélation pour des valeurs croissantes de $d$, et on observera un palier à partir de $d = d_0$, pour lequel l'exposant reste constant: $d_0$ est la dimension de plongement. Cette méthode a été consacrée par l'expérience et elle correspond aussi à la recherche du nombre de "degrés de liberté" du système étudié (voir les travaux de Babloyantz *et al.* (1985) sur le nombre de degrés de liberté du cerveau pendant le sommeil, léveil et la crise d'épilepsie...).
Une seconde méthode possible est celle proposée par Tong et Chang (1992) qui considère

les $X_i$ (choisis réels) comme des variables aléatoires. En fait, ils postulent que le système est régi par une équation de type $Y_{n+1}^{d_0} = F^{(d_0)}(Y_n^{d_0}) + E_{n+1}^{d_0}$ avec $Y_n^{d_0} = (X_n, .., X_{n-d_0+1})$, $E_n^{d_0} = (e_{n+d_0-1}, 0, ..., 0)$, les $e_i$ étant une suite de martingales dont la distribution a un support borné, et $F^{d_0} : \mathbb{R}^{d_0} \to \mathbb{R}$ étant une fonction continue. Une telle interprétation est justifiée par le fait que si la variance de la séquence de martingales tend vers zéro, la distribution invariante du système stochastique converge vers la distribution invariante du processus déterministe (?). La démarche classique alors mise en place est la méthode des noyaux permettant d'obtenir une estimation $G_N^d$ de l'espérance $E(Y_j^d / Y_{j-1}^d, ..., Y_{j-d}^d)$ et une estimation $G_{N,i}^d$ de l'espérance $E(Y_j^d / Y_{j-1}^d, ..., Y_{j-d}^d)$ quand on ne prend pas en compte le vecteur $Y_i^d$ pour calculer cette estimation. Par "cross-validation", on construit une fonction $CV(d)$ telle que:

$$CV(d) = \frac{1}{N-k+1} \sum_{i=k}^{N} (Y_i^d - G_{N,i}^d(Y_{i-d}^d))^2 w(Y_{i-d}^d),$$

avec $w : \mathbb{R}^d \to \mathbb{R}^+$, une fonction poids. En faisant varier $d$ dans une fourchette convenable, $CV(d)$ admet un minimum: c'est $\widehat{d_0}$, estimateur de $d_0$, que l'on démontre consistant quand $N$ tend vers l'infini. Les applications numériques confirment la validité de ces résultats, même pour de petits échantillons, quand la dimension de plongement n'est pas trop grande...

# Bibliographie


[1] Adler, R. (1981). *Geometry of random fields.* Wiley, New York.

[2] Arnold, V. et Avez, A. (1967). *Problèmes ergodiques de la mécanique classique..* Gauthier-Villars, Paris.

[3] Babloyantz, A., Salazar, J.M. et Nicolis, C. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Phys. Letters* **111** 152-156.

[4] Bailey, B., Ellner, S. et Nychka, D. (1997). Chaos with confidence: asymptotics and applications of local Lyapunov expnonents. In *Fields Institute Communications* **11** 115-133

[5] Bartlett, M.S. (1990). Chance or chaos? *J. R. Statist. Soc. A* **153** 321-347.

[6] Benedicks, M. and Carleson, L. (1991). The dynamics of the Hénon map. *Ann. Math.* **133** 73-169.

[7] Bergé, P., Pomeau, Y., et Vidal, C. (1984). *L'ordre dans le chaos.* Herman, Paris.

[8] Bosq, D. (1995). Optimal asymptotic quadratic error of density estimators for strong mixing or chaotic data. *J. Stat. Probab. Lett.* **22** 339-347.

[9] Bosq, D. et Guégan, D. (1995). Non parametric estimation of the chaotic function and the invariant measure of a dynamical system. *J. Stat. Probab. Lett.* **25** 201-212.

[10] Casdagli, M. (1992). Chaos and deterministic versus stochastic non-linear modelling. *J. R. Statist. Soc. B* **54** 303-328.

[11] Cheng, B. et Tong, H. (1992). On consistent nonparametric order determination and chaos. *J. R. Statist. Soc. B* **54** 427-449.

[12] Cutler, C.D. (1991). Some results on the behavior and estimation of the fractal dimensions of distributions on attractors. *J. Statist. Phys.* **62** 651-708.

[13] Cutler, C.D. (1994). A theory of correlation dimension for stationary time series. *Philos. Trans. R. Soc. Lond. A* **348** 343-355.

[14] Dahan Dalmenico, A., Chabert, J.L. et Chemla, K. (1992). *Chaos et déterminisme.* Seuil: Points sciences.

[15] Denker, M. et Keller, G. (1986). Rigorous statistical procedure for data from dynamical systems. *J. Statist. Phys.* **49** 67-93.

[16] Ding, M., Gregobi, C., Ott, E., Sauer, T., and Yorke, J. (1993). Estimating correlation dimension from a chaotic time series: when does plateau onset occur? *Physica D* **69** 404-424.

[17] Eckmann, J.-P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57** 617-656.

[18] Falconer, K. (1990). *Fractal geometry.* Chichester: Wiley.

[19] Feigenbaum, M.J. (1980). The transition to aperiodic behavior in turbulent systems. *Comm. Math. Phys.* **77** 65-86.

[20] Feigenbaum, M.J., Kadanoff, L.P. and Shenker, S. J. (1982). Quasiperiodicity in dissipative systems: a renormalizion group analysis. *Physica D* **5** 370-386.

[21] Fraedrich, K. and Risheng, W. (1993). Estimating the correlation dimension of an attractor from noisy and small datasets based on re-embedding. *Physica D* **65** 373-398.

[22] Gleick, J. (1991). *La théoris du chaos.* Flammarion: Champs.

[23] Grassberger, P. and Proccacia, I. (1983). Measuring the strangeness of strange attractors. *Physica D* **9** 189-208.

[24] Guckeinheimer, J. and Holmes, P. (1986). *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields.*

[25] Hènon, M. (1976). A two dimensional mapping with a strange attractor. *Comm. Math. Phys.* **50** 69-77.

[26] Isham, V. (1993). Statistical aspects of chaos: a review. *Networks and Chaos - Statistical and probabilistic aspects* - Chapman and Hall, London, 124-200.

[27] Jensen, J.L.. (1993). Chaotic dynamical systems with a view towards statistics: a review. *Networks and Chaos- Statistical and probabilistic aspects* Chapman and Hall, London, 201-250.

[28] Kuhn, T.S. (1983). *La structure des révolutions scientifiques.* Flammarion: Champs.

[29] Lu, Z.Q. et Smith, R.L. (1997). Estimating local Lyapunov exponents. In *Fields Institute Communications* **11** 135-151

[30] McCaffrey, D., Ellner, S., Gallant, A.R. et Nychka, D. (1992). Estimating the Lyapunov exponent of a chaotic system with non-parametric regression. *J. Am. Statist. Ass.* **86** 682-695

[31] Mandelbrot, B.B. (1995). *Les objects fractals.* Flammarion: Champs.

[32] Malraison, B., Atten, P., Bergé, P. et Dubois, M. (1984). Dimension of strange attractors: an experimental determination for the chaotic regime of two convective systems. *J. Phys. Lett.* **44** L-897.

[33] Marteau, P.F. and Abarbanel, H.D.I. (1991). Noise reduction in chaotic time series using scaled probabilistic methods. *J. Nonlinear Sci.* **1** 313-343.

[34] Mayer-Kress, G. (1986). *Dimensions and entropies in chaotic systems.* New-York: Springer-Verlag.

[35] Nychka, D., Ellner, S., Gallant, A.R. and McCaffrey, D. (1992). Finding chaos in noisy systems. *J. R. Stat. Soc. B* **54** 399-426.

[36] Ogata, Y. and Katsura, K. (1991). Maximum likelihood estimates of the fractal dimension for random spatial patterns. *Biometrika* **78** 463-474.

[37] Ornstein, D.S. and Weiss, B. (1991). Statistical properties of chaotic systems. *Bull. Am. Math. Soc.* **24** 11-116.

[38] Osborne, A.R. and Provenzale, A. (1989). Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* **35** 357-381.

[39] Pesin, Y.B. (1993). On rigorous mathematical definitions of correlation dimension and generalized spectrum for dimensions. *J. Statist. Phys.* **71** 529-547.

[40] Prigogine, I et Stengers, I. (1986). *La nouvelle alliance.* Gallimard: Folio essais..

[41] Provenzale, A. and Osborne, A.R. (1991). Deterministic chaos versus random noise: finite correlation dimension for colored noises with power-law power spectra. 260-275.

[42] Rényi, A. (1970). *Probability theory.* .

[43] Ruelle, D. et Takens, (1971). . *Comm. Math. Phys.* **82** 137-151.

[44] Ruelle, D. (1981). Small random perturbations of dynamical systems and the definition of attractors. *Comm. Math. Phys.* **82** 137-151.

[45] Ruelle, D. (1989a). *Chaotic evolution and strange attractors.* Cambridge: Cambridge University Press.

[46] Ruelle, D. (1989b). *Elements of differentiable dynamics and bifurcation theory.* New-York: Academic Press.

[47] Ruelle, D. (1990). Deterministic chaos: the science and the fiction. *Proc. R. Soc. Lond. B* **427** 241-248.

[48] Ruelle, D. (1991). *Hasard et chaos* Odile Jacob: Points.

[49] Scheinkman, J. (1994). Nonlinear dynamics in economics and finance. *Philos. Trans. R. Soc. Lond. A* **346** 235-250.

[50] Serinko, R.J. (1994). A consistent approach to least squares estimation of correlation dimension in weak Bernoulli dynamical systems. *Annals Appl. Probab.* **4** 1234-1254.

[51] Smith, R.L. (1992). Optimal estimation of fractal dimension. In Nonlinear modeling and forecasting, Proc. Vol. XII, Eds. M. Casdagli and S. Eubank, Addison-Wesley.

[52] Smith, R.L. (1992). Estimating dimension in noisy chaotic time series. *J. R. Stat. Soc. B* **54** 329-351.

[53] Theiler, J. (1988). Lacunarity in a best estimator of fractal dimension. *Physics Letters A* **133** 195-200.

[54] Theiler, J. (1990). Statistical precision of dimension estimators. *Physical Review A* **41** 3038-3051.

[55] Theiler, J. (1991). Some comments on the correlation dimension of $1/f^\alpha$ noise. *Physics Letters A* **155** 480-493.

[56] Tong, H. (1990). *Nonlinear Times Series.* Oxford, Oxford Univ. Press.

[57] Wolf, A., Swift, J.B., Swinney, H.L. and Vastano, J.A. (1985). Determining Lyapunov exponents from a time series. *Physica D* **16** 285-315.

[58] Wolff, R.C.L. (1992). Local Lyapunov exponents: looking closely at chaos. *J. R. Stat. Soc. B* **54** 353-371.

[59] Young, L.S. (1982). Dimension, entropy and Lyapunov exponents. *Ergod. Theory Dynam. Syst.* **2** 109-124.

# Chapitre 2

# Dimension de corrélation locale et dimension de Hausdorff de champs aléatoires continus

Local correlation and Hausdorff dimensions of continuous random fields

By Jean-Marc Bardet

*Université Paris-Sud and Université Evry-Val d'Essonne*

Local correlation is a new dimension for continuous random field. This dimension is given by the asymptotic behaviour of intersection occupation measure. We compare first this dimension to Hausdorff one. We compute then local correlation dimension of multiparameter fractional Brownian motion and give new bounds for Hausdorff dimension.

## 2.1 Introduction

We consider a $\mathbb{R}^d$-valued continuous random vector process $X = \{X_t; t \in \mathbb{R}\}$ and for $J$ compact set of $\mathbb{R}^2$ we define its self-intersection occupation measure $\mu(I, J)$ in a compact set $I \in \mathbb{R}^d$, i.e.

$$\mu(I, J) = \int_J \mathbb{I}_{(X_t - X_s \in I)} dsdt,$$

*Key words and phrases.* Fractional Brownian motion, correlation dimension, Hausdorff dimension, intersection local time.

*AMS 1991 subject classifications.* Primary 60G15; secondary 60G17, 60J55.

If $\mu(.,J)$ is absolutely continuous with respect to $\mathbb{R}^d$-Lebesgue measure $\lambda_d$ then $X$ has a self-intersection local time (S.I.L.T.) denoted $\alpha(.,J)$ which is the Radon-Nikodym derivative $\frac{d\mu}{d\lambda_d}$. This definition of S.I.L.T. was first done by Rosen in [18] and [19]. But if $J$ contains the diagonal $\{s=t\}$ then lots of random fields (for instance, planar Brownian motions) don't have S.I.L.T. Varadhan's renormalization technics are then used to study the behavior of $\alpha(x,J)$ when $x \to 0$ (see Rosen [21]) or of $\alpha_\varepsilon(0,J)$ when $\varepsilon \to 0$, where $\alpha_\varepsilon(.) = \int_J \delta_\varepsilon(X_t - X_s)dsdt$ and $\delta_\varepsilon$ is a sequence of regularizing kernels when $\varepsilon \to 0$ (see Le Gall [16], Dynkin [7], Rosen [20], or recently Imkeller and al. [13]).

Working only with $\mu(I,J)$ avoids these difficulties. One can then be interested in the asymptotic behavior of $\mu(B_d(0,r),[0,T]^2)$ when $r \to 0$, with $T > 0$ and $B_d(0,r)$ a $d$-ball with center 0 and radius $r$. It is well known that if it exists $C > 0$ such that $\mu(B_d(0,r),[0,T]^2) \sim Cr^d$ when $r \to 0$ then $X$ has a S.I.L.T. in 0 (see Geman and Horowitz [11], Theorem 7.2). Now assume for a path of $X$ that $\mu(B_d(0,r),[0,T]^2) \sim r^\nu L(r)$ when $r \to 0$ where $\nu \geq 0$ and $L(r)$ verifies $\lim_{r\to 0} \log L(r)/\log r = 0$. First, we easily prove that $0 \leq \nu \leq d$ and secondly that $\nu$ is smaller than Hausdorff dimension of this path. As a consequence, we give new relations between the existence of a S.I.L.T. for $X$ (or a sub-process of $X$) and the value of Hausdorff dimension of $X$.
$\nu$ has all the properties generally required to define a dimension and we call it (whenever it exists) local correlation dimension. Usual correlation dimension was introduced by Grassberger and Proccacia [12] for easily computing fractal dimension of strange attractors. If $X_1, X_2, ..., X_N$ are $N$ vectors of $\mathbb{R}^d$ and if $C(N,r)$ is the ratio of the number of $X_1, X_2, ..., X_N$ with interdistances less than $r > 0$, then correlation dimension $\nu$ whenever it exists is defined by:

$$\nu = \lim_{r\to 0} \lim_{N\to+\infty} \frac{\log C(N,r)}{\log r}.$$

This dimension is connected to the dynamic of the time series $X_1, X_2, ..., X_N$, unlike Hausdorff dimension which only depends on geometry of the set $\{X_1, X_2, ..., X_N\}$ (see Eckmann and Ruelle [8]). But if $X = \{X_t, t > 0\}$ is a continuous stochastic process, $\nu$ often either doesn't exist or merely equals $d$ (for instance, see [22] or [3]). By considering $X_{T/N}(\omega), X_{2T/N}(\omega), ..., X_T(\omega)$ instead of $X_1, X_2, ..., X_N$, we go from a global study to a local one. Computing "local" correlation dimension of $X$ can then yield pertinent results.

The main result is given in Section 3 which is devoted to multiparameter fractional Brownian motions on $\mathbb{R}^d$ with index $(\alpha_1, \alpha_2, ..., \alpha_d)$ where each $\alpha_i \in ]0,1[$. If $\alpha_1+\alpha_2+...+\alpha_d < 1$, we show that $\nu$ exists and $\nu = d$ with positive probability. If $\alpha_1+\alpha_2+...+\alpha_d = 1$, we show that $\mu(B_d(0,r),[0,T]^2) \sim Cr^d \log(1/r)$ when $r \to 0$ with positive probability, implying $\nu$ exists and $\nu = d$ with positive probability. If $\alpha_1+\alpha_2+...+\alpha_d > 1$, we show

$\mu(B_d(0,r),[0,T]^2) \sim Cr^\nu(\log(1/r))^e$ when $r \to 0$ a.s., with

$$\nu = \min_{i\in\{1,..,d\}} \left\{ \frac{1+\sum_{k=1}^{i}(\alpha_i - \alpha_k)}{\alpha_i} \right\},$$

where $\nu \in [1,d[$ and $e=1$ when $\nu \in \mathbb{N}$ or else $e=0$. In this case, local correlation dimension exists a.s. On the other hand, one can go back on the value of Hausdorff dimension of $X$ and $X$-graph given by Cuzick [4]. By the way, we correct an inexact computation of [4] and show that computings only provide bracketings for these dimensions with lower bounds giving by local correlation dimensions (see Propositions 2.3.4 and 2.3.5). In this case, computing local correlation dimension provides more precise informations on "fractal" dimension of $X$.

## 2.2 Definitions and properties

In all this Section, we consider a $\mathbb{R}^d$-valued continuous random vector process $X = \{X_t, t \in [0,T]\}$ ($T>0$, $d \in \mathbb{N}^*$). We will use the following notations:

- $\| \,.\, \|$ is a norm on $\mathbb{R}^d$, $B_d(0,r)$ is $d$-ball with center 0 and radius $r$, $\mathbb{I}$ is the indicator function and $\lambda_n$ is the Lebesgue measure on $\mathbb{R}^n$.
- $Y = \{Y_{s,t} = X_t - X_s, (s,t) \in [0,T]^2\}$ is a continuous $\mathbb{R}^d$-valued field on $[0,T]^2$.
- $Z = \{Z_{s,t} = \| X_t - X_s \|, (s,t) \in [0,T]^2\}$ is a continuous $\mathbb{R}^+$-valued field on $[0,T]^2$.

Let $(X_{T/N}(\omega), X_{2T/N}(\omega), ..., X_T(\omega))$ be $N$ vectors of path $X(\omega)$. We define random variable $C_T(N,r)$ corresponding to the ratio of the number of discretized trajectory points with distances less than $r>0$ to the total number of distincts pairs formed by $N$ points. In the limit $N \to +\infty$, $C_T(N,r) \to C_T(r)$ with probability 1, where $C_T(r)$ is the probability that a pair $(X_t, X_s)$ ($t$ and $s$ are chosen at random (uniformly) on $[0,T]$) is separated by a distance less than $r$. If $C_T(r,\omega)$ roughly grows like a $r$-power when $r \to 0$ then the local correlation dimension $\nu(\omega)$ of $X(\omega)$ exists and is equal to this power. More precisely,

**Definitions**: *For $r>0$ one can define $[0,1]$-valued random variables $C_T(N,r)$ and $C_T(r)$, and if it exists the local correlation dimension $\nu$, such as:*

$$\begin{aligned} C_T(N,r,\omega) &= \frac{2}{N(N-1)} \sum_{1\le i<j\le N} \mathbb{I}_{(\|X_{jT/N}(\omega)-X_{iT/N}(\omega)\|\le r)}. & (1)\\ C_T(r,\omega) &= \lim_{N\to+\infty} C_T(N,r,\omega) = \frac{1}{T^2}\int_0^T\int_0^T \mathbb{I}_{(\|X_t(\omega)-X_s(\omega)\|\le r)} dsdt. & (2)\\ \nu(\omega) &= \lim_{r\to 0} \frac{\log C_T(r,\omega)}{\log r}. & (3) \end{aligned}$$

**Remark 1**: The existence of $C_T(r,\omega)$ and its expression in (2) result from $X$ (and thus $Y$) continuity. So $C_T(N,r,\omega)$ can be seen as a Riemann sumn converging to $C_T(r,\omega)$ when $N \to +\infty$.

Local correlation dimension can also be expressed from occupation measures. Denote

- $\mu_Y(I,[0,T]^2,\omega)$ is the occupation measure of $Y_{s,t}(\omega)$ in a Borel set $I \in \mathbb{R}^d$ for $(s,t) \in [0,T]^2$. So $\mu_Y(I,[0,T]^2,\omega) = \lambda_2\{(s,t) \in [0,T]^2 / Y_{s,t}(\omega) \in I\}$.
- $\mu_Z(J,[0,T]^2,\omega)$ is the occupation measure of $Z_{s,t}(\omega)$ in a Borel set $J \in \mathbb{R}^+$ for $(s,t) \in [0,T]^2$. So $\mu_Z(J,[0,T]^2,\omega) = \lambda_2\{(s,t) \in [0,T]^2 / Z_{s,t}(\omega) \in J\}$.

As a consequence, for any $r > 0$ and almost every $\omega$,

$$C_T(r,\omega) = \frac{1}{T^2}\mu_Y(B_d(0,r),[0,T]^2,\omega) = \frac{1}{T^2}\mu_Z([0,r],[0,T]^2,\omega). \tag{4}$$

This property will allow us to use known results about occupation measures (we essentially use results of [1] and [11]). First, we will use $\mu_Z$ to give a relation between local correlation and Hausdorff dimensions. We denote by $dim_H S$ the Hausdorff dimension of a set $S \subset \mathbb{R}^d$.

**Property 2.2.1** *If $\nu(\omega)$ exists then*

$$\nu(\omega) \le dim_H X(\omega) \le d. \tag{5}$$

**Proof**: (we follow an idea of [6]).
Let us define an energy integral denoted $dim_C X$ such as:

$$dim_C X(\omega) = \sup\{\alpha \in \mathbb{R}_+ / \int_0^T \int_0^T \frac{dsdt}{\parallel X_t(\omega) - X_s(\omega) \parallel^\alpha} < +\infty\}. \tag{6}$$

For a compact in $\mathbb{R}^d$, capacity and Hausdorff dimensions are equals (see [10] and [14]). Now, it is obvious from capacity dimension definition that $dim_C X(\omega)$ is smaller than capacity dimension of $X(\omega)$. If $\nu(\omega)$ exists then it will be enough to show $dim_C X(\omega) = \nu(\omega)$ for proving (5).

If $\nu(\omega)$ exists then $C_T(r,\omega) = r^\nu l(r,\omega)$ with $\lim_{r\to 0} \dfrac{\log l(r)}{\log r} = 0$. Let $\alpha$ be positive real number. Applying a usual result to the occupation measure $\mu_Z(.,[0,T]^2,\omega)$, one obtains:

$$\begin{aligned}
\int_0^T \int_0^T \frac{dsdt}{\parallel X_t(\omega) - X_s(\omega) \parallel^\alpha} &= \int_0^{+\infty} \frac{d\mu_Z(r,[0,T]^2,\omega)}{r^\alpha} \\
&= \int_0^{R(\omega)} \frac{\mu_Z(r,[0,T]^2,\omega)dr}{r^{\alpha+1}} \qquad (7) \\
&= T^2 \int_0^{R(\omega)} \frac{r^\nu l(r,\omega)dr}{r^{\alpha+1}} \qquad (8)
\end{aligned}$$

,$\mu_Z(0,[0,T]^2,\omega)=0$ and $\mu_Z(+\infty,[0,T]^2,\omega)=T^2$ for almost every $\omega$ implies (7), and $R(\omega)=\sup\{\parallel X_t(\omega)-X_s(\omega)\parallel / (s,t)\in[0,T]^2\}$. Then, for $\epsilon>0$ and $\alpha=\nu+2\epsilon$,

$$\int_0^{+\infty}\frac{r^\nu l(r,\omega)dr}{r^{\alpha+1}}\geq\int_0^{r_0}\frac{m(\omega)dr}{r^{1+\epsilon}} \tag{9}$$

because there exists $m>0$ such as $l(r)\geq mr^\epsilon$ for $r\in[0,r_0]$ ($r_0>0$). One deduces that (8)$=+\infty$ thus $dim_C X(\omega)\leq\nu(\omega)$.
Similarly, if $\alpha=\nu-2\epsilon$ then

$$\int_0^{r_1}\frac{r^\nu l(r,\omega)dr}{r^{\alpha+1}}\leq\int_0^{r_1}\frac{M(\omega)dr}{r^{1-\epsilon}} \tag{10}$$

because there is $M>0$ such as $l(r)\leq Mr^{-\epsilon}$ for $r\in[0,r_1]$. Consequently, (8)$<+\infty$ and so $dim_C X(\omega)\geq\nu(\omega)$. Thus, $dim_C X(\omega)=\nu(\omega)$ and this proves (5).■

If $\mu_Y(.,[0,T]^2)$ is absolutely continuous with respect to Lebesgue measure $\lambda_d$, i.e., $Y$ has a local time, then local correlation dimension of $X$ can be merely deduced:

**Property 2.2.2** *Let $V_0$ be a 0-neighborhood of $\mathbb{R}^d$. If $Y(\omega)$ has a local time $\alpha(x,[0,T]^2,\omega)$ for $x\in V_0$, which is continuous and positive for $x=0$ then*

$$\nu(\omega)\ \textit{exists and}\ \nu(\omega)=d=dim_H X(\omega). \tag{11}$$

**Proof**: We have: $\mu_Y(B_d(0,r),[0,T]^2,\omega)=\int_{B_d(0,r)}\alpha(x,[0,T]^2,\omega)dx$. From conditions on $\alpha(.,[0,T]^2,\omega)$, we obtain for $r$ small enough:

$$\frac{1}{2}\alpha(0,[0,T]^2,\omega)\lambda_d(B_d(0,r))\leq\mu_Y(B_d(0,r),[0,T]^2,\omega)\leq 2\alpha(0,[0,T]^2,\omega)\lambda_d(B_d(0,r)).$$

As a consequence, with property (4),

$$\log C_T(r,\omega)\sim\log r^d \text{ when } r\to 0.$$

From Definition (3) one obtains (11). ■

**Remark 1**: A local time for $Y$ is an intersection local time for $X$. Thus, Property 2.2.2 shows that if $X$ has an intersection local time satisfying certain weak conditions then Hausdorff dimension of $X$ is equal to $d$. In this sense, local correlation dimension provides relations between intersection local time and Hausdorff dimension of $X$.

Property 2.2.2 allows us to precise a bracketing for $\nu$.

**Proposition 2.2.1** *Let $X_t={}^t(X_t^1,..,X_t^d)$ and assume that ${}^t(X_t^{i_1},..,X_t^{i_k})$ has an intersection local time $\alpha_k(x,[0,T]^2,\omega)$ for $x$ in a 0-neighborhood ($k\in\mathbb{N}^*$ and $(i_1,..,i_k)\in\{1,..,d\}^k$). If the local correlation dimension of $X(\omega)$ exists and if $\alpha_k(x,[0,T]^2,\omega)$ is continuous and positive for $x=0$ then*

$$0\leq k\leq\nu(\omega)\leq dim_H X(\omega)\leq d. \tag{12}$$

## 2.3 Fractional Brownian motion in $\mathbb{R}^d$

Now we suppose that $X$ is a fractional Brownian motion in $\mathbb{R}^d$ (f.B.m.d) with index $(\alpha_1, \alpha_2, ..., \alpha_d)$ and defined as follow: $X = \{X_t = (X_t^1, X_t^2, ..., X_t^d), t \in [0,T]\}$ is a continuous $\mathbb{R}^d$-valued Gaussian field such as each $X^i$ has zero-mean (and $X_0^i = 0$ a.s.), stationary increments, independent coordinate fields and with each $\alpha_i \in ]0,1[$ for $(i,j) \in \{1,2,...,d\}^2$ and $(s,t) \in [0,T]^2$:

$$\mathbb{E}X_t^i X_s^j = \frac{1}{2}\delta_{ij}a_i(\mid t \mid^{2\alpha_i} + \mid s \mid^{2\alpha_i} - \mid s-t \mid^{2\alpha_i}).$$

For ease of writing we assume that $0 < \alpha_1 \leq \alpha_2 \leq .. \leq \alpha_d < 1$. We want to find the local correlation dimension of $X$. Two cases have to be distinguished:

**- case** 1: if $\sum_{k=1}^{d} \alpha_k < 1$ we apply Property 2.2.2 then

**Proposition 2.3.1** *$\nu$ exists and $\nu = d$ with positive probability.*

**Proof**: Berman [2] and Pitt [17] have given results on existence and joint continuity for Gaussian field local times. Consequently, one verifies that $Y$ has a continuous local time in a 0-neighborhood for almost every $\omega$ if $\sum_{k=1}^{d} \alpha_k < 1$. Moreover, Rosen in [19] has shown that this local time is positive in 0 with positive probability for $\alpha_1 = .. = \alpha_d < 1/d$. By Cuzick Theorem 1 of [5], we deduce an extension of Rosen's result for $0 \leq \alpha_1 \leq \alpha_2 \leq .. \leq \alpha_d < 1$ and $\sum_{k=1}^{d} \alpha_k < 1$. So, it's possible to apply Property 2.2.2, giving Proposition 2.3.1.■

**- case** 2: if $\sum_{k=1}^{d} \alpha_k \geq 1$, we turn to establish asymptotic behavior of $C_T(r)$ when $r \to 0$ from asymptotic behavior of occupation measure $\mu$, where

$$\mu(x, I, \omega) = \lambda_2\{(s,t) \in I / \parallel X_t(\omega) - X_s(\omega) \parallel^2 \in [0,x]\}$$

for all $x \in \mathbb{R}_+$ and all Borelian set $I \in [0,T]^2$. For ease of computing we choose norm $\parallel . \parallel$ such as

$$\parallel z \parallel^2 = (z_1^2/2a_1) + ... + (z_d^2/2a_d) \text{ for } z = (z_1, ..., z_d) \in \mathbb{R}^d.$$

We have again:

$$C_T(r, \omega) = \frac{1}{T^2}\mu(r^2, [0,T]^2, \omega).$$

If $f$ is a measurable function on $\mathbb{R}_+$ then

$$\int_0^{+\infty} f(x)d\mu(x, I, \omega) = \int_I f(\| X_t(\omega) - X_s(\omega) \|^2)dsdt.$$

Let $f(x) = e^{-px}$ and denote $\hat{\mu}(p, \omega)$ the Laplace transform of $\mu(x, [0, T]^2, \omega)$. Thus,

$$\hat{\mu}(p, \omega) = \int_0^T \int_0^T \exp(-p \| X_t(\omega) - X_s(\omega) \|^2)dsdt. \tag{13}$$

One can then establish the behavior of $\hat{\mu}(p, \omega)$ when $p \to +\infty$. The behavior of $C_T(r, \omega)$ when $r \to 0$ is provided by a Tauberian theorem. So, we obtain:

**Proposition 2.3.2** *if* $\sum_{k=1}^{d} \alpha_k = 1$ *then* $\nu$ *exists and* $\nu = d$ *with positive probability.*

**Proposition 2.3.3** *if* $\sum_{k=1}^{d} \alpha_k > 1$ *then* $\nu$ *exists a.s. and*

$$\nu = \min_{i \in \{1,..,d\}} \left\{ \frac{1 + \sum_{k=1}^{i}(\alpha_i - \alpha_k)}{\alpha_i} \right\} \quad a.s. \tag{14}$$

If $\sum_{k=1}^{d} \alpha_k > 1$, there is $i_0 \in \{2, .., d\}$ such as $\nu = \dfrac{1 + \sum_{k=1}^{i_0}(\alpha_{i_0} - \alpha_k)}{\alpha_{i_0}}$ and $\nu \in ]i_0 - 1, i_0]$. So, $\nu$ is only depending on the $i_0$ most irregular $X$ coordinates. Moreover, one notices that $\sum_{k=1}^{i_0-1} \alpha_k < 1$ and $\sum_{k=1}^{i_0} \alpha_k \geq 1$. If one considers the $(i_0 - 1)$ most irregular $X$ coordinates it exists an intersection local time for their trajectory (see Proposition 2.3.1). By inequality (12) it is only natural that local correlation dimension is greater than $i_0 - 1$. If one adds another coordinate then one can not define an intersection local time for $X$ and $\nu \leq i_0$.

**Remark 2**: For the case of Brownian motion ($\alpha_i = 1/2$), $\nu = 2$ with positive probability when $d = 2$ and $\nu = 2$ with probability 1 when $d \geq 3$. One finds similar results for Hausdorff dimension.

**Remark 3**: Local properties of $X$ are only used in determination of local correlation dimension. So, Proposition 2.3.1, 2.3.2 and 2.3.3 can be extended for centered Gaussian fields with stationary increments verifying:

(1) for each $i = 1, .., d$ there exists a $\alpha_i$ such that

$$\alpha_i = \sup\{\alpha : \sigma_i(t) = o(|t|^\alpha), t \to 0\} = \inf\{\alpha : |t|^{2\alpha} = o(\sigma_i(t)), t \to 0\},$$

with $\sigma_i^2(t) = \mathbb{E}(X_{t+s}^i - X_s^i)^2$ ($X$ is called $\alpha$-index Gaussian field).

(2) if $\Sigma_t$ is the covariance matrix of the vector $X_t - X_0$, there exists an $\epsilon > 0$ such that, for all $t \in [0, T]$,
$$\det \Sigma_t \geq \epsilon \prod_{i=1}^{d} \sigma_i^2(t).$$

Those conditions are used by Cuzick [4] and Adler [1] for the following determination of Hausdorff dimension.

Cuzick [4] has shown that with probability 1,
$$dim_H X = \min\left\{d, \frac{1+\sum_{k=1}^{d}(\alpha_d - \alpha_k)}{\alpha_d}\right\}. \tag{15}$$

The proof is a two-step procedure. Upper bound of $dim_H X$ is obtained with a Hölder condition for $X$: it is the second term of (15). Lower bound of $dim_H X$ comes from computation of $dim_C X$ (see (6)) and Cuzick finds the second term of (15) again. But we have shown that $\nu = dim_C X$ when $\nu$ exists (see Property 2.2.1) and by Proposition 2.3.3, $\nu$ exists and is given by (14). We have found a fault in Cuzick computation of $dim_C X$ in Theorem 1 of [4] (he uses an equivalent in (8) of $\int_0^{+\infty} (y_1^2 + a)^{-\lambda/2} dy_1$ which is $C(\lambda)a^{(-\lambda+1)/2}$ only if $\lambda > 1$; so he has to stop the iteration after $i_0$ steps instead of $d$ steps). After correction we find the same value as in (14). Now one can only give the following proposition:

**Proposition 2.3.4** *With probability* 1,

$$- \quad \textit{if } \sum_{k=1}^{d} \alpha_k \leq 1,\ dim_H X = d \tag{16}$$
$$- \quad \textit{if } \sum_{k=1}^{d} \alpha_k > 1,\ \frac{1+\sum_{k=1}^{i_0}(\alpha_{i_0} - \alpha_k)}{\alpha_{i_0}} \leq dim_H X \leq \frac{1+\sum_{k=1}^{d}(\alpha_d - \alpha_k)}{\alpha_d}. \tag{17}$$

If $\sum_{k=1}^{d} \alpha_k > 1$, $dim_H X$ has an exact value in the following particular case:

**Corollary 2.3.1** *If* $\sum_{k=1}^{d} \alpha_k > 1$ *and if* $\alpha_d = \alpha_{i_0}$ *then with probability* 1,
$$\nu = dim_H X = \frac{1+\sum_{k=1}^{d}(\alpha_d - \alpha_k)}{\alpha_d}. \tag{18}$$

Similarly, the Hausdorff dimension of the graph of $X$ ($Gr(X) = \{(t, X_t), t \in [0,T]\}$) has to be revised. As above, this dimension only can be bracketed by lower and upper bounds. One determines these bounds by adding a new f.B.m coordinate (with parameter $\alpha_{d+1} \to 1$) to $X$ (see [4] again).

**Proposition 2.3.5** *With probability* 1*,*

$$\frac{1 + \sum_{k=1}^{i_0}(\alpha_{i_0} - \alpha_k)}{\alpha_{i_0}} \leq dim_H Gr(X) \leq d + 1 - \sum_{k=1}^{d} \alpha_k. \tag{19}$$

**Remark 4**: These Hausdorff dimension bracketings can be generalized for $(N, d)$ Gaussian fields on $[0,1]^N$ of index $\alpha$ (see [1]). Instead of Theorem 1 of [4], we have, with probability 1,

- if $\displaystyle\sum_{k=1}^{d} \alpha_k \leq N$, $dim_H X = d$.
- if $\displaystyle\sum_{k=1}^{d} \alpha_k > N$, $\displaystyle\min_{i \in \{1,2,..,d\}} \left\{ \frac{N + \sum_{k=1}^{i}(\alpha_i - \alpha_k)}{\alpha_i} \right\} \leq dim_H X \leq \frac{N + \sum_{k=1}^{d}(\alpha_i - \alpha_k)}{\alpha_d}$.
- $\displaystyle\min_{i \in \{1,2,..,d\}} \left\{ \frac{N + \sum_{k=1}^{i}(\alpha_i - \alpha_k)}{\alpha_i} \right\} \leq dim_H Gr(X) \leq d + N - \sum_{k=1}^{d} \alpha_k$.

## 2.4 Proofs

For $0 < \alpha_1 \leq \alpha_2 \leq ... \leq \alpha_d < 1$, denote:

$$\begin{aligned} S(j) &= \alpha_1 + \alpha_2 + ... + \alpha_j \quad \text{et} \\ V(j) &= \frac{1 + \sum_{k=1}^{j}(\alpha_j - \alpha_k)}{2\alpha_j}. \end{aligned}$$

If $S(d) \geq 1$, let integer $i_0 \in \{2, ..., d\}$ be such that $S(i_0 - 1) < 1$ and $S(i_0) \geq 1$. To establish Proposition 2.3.2 and 2.3.3, we show the following lemmas:

**Lemma 2.4.1** *For large enough $p$, there exist positive real numbers $m_0(\alpha_1, ..., \alpha_d, T)$ and $d_1(\alpha_1, ..., \alpha_d, T)$ such that*

$$\text{-if } S(i_0) = 1, \quad \mathbb{E}\hat{\mu}(p) = m_0 \frac{\log p}{p^{V(i_0)}} + O\left(\frac{\log p}{p^{V(i_0)+d_1}}\right). \tag{20}$$

$$\text{-if } S(i_0) > 1, \quad \mathbb{E}\hat{\mu}(p) = m_0 \frac{1}{p^{V(i_0)}} + O\left(\frac{\log p}{p^{V(i_0)+d_1}}\right). \tag{21}$$

**Proof**: We have:

$$\begin{aligned}
\mathbb{E}\hat{\mu}(p) &= \int_0^T \int_0^T \mathbb{E}e^{-p\|X_t-X_s\|^2} dsdt \\
&= \int_0^T \int_0^T \prod_{i=1}^d \mathbb{E}\left[\exp\left(-\frac{p(X_t^i-X_s^i)^2}{2a_i}\right)\right] dsdt \quad \text{(independence)} \\
&= \int_0^T \int_0^T \frac{1}{\prod_{i=1}^d (1+p \mid t-s \mid^{2\alpha_i})^{1/2}} dsdt \\
&= 2\int_0^T \frac{T-u}{\prod_{i=1}^d (1+pu^{2\alpha_i})^{1/2}} du. \qquad (22)
\end{aligned}$$

Let's give an asymptotic development of (22). We know:

$$(1+x)^{-1/2} = \sum_{k=0}^{+\infty} \frac{2k!}{4^k (k!)^2} (-x)^k \quad \text{if } x \in [0,1]. \qquad (23)$$

Moreover, if $x \in [0,1]$,

$$0 < a_k(x) = \left(\frac{4k!}{4^{2k}(2k!)^2} - \frac{2(2k+1)!x}{4^{2k+1}((2k+1)!)^2}\right)x^{2k} < \frac{2(4k)!}{4^{2k+1}(2k+1)!(2k)!} = b_k \qquad (24)$$

and $\sum_{k=0}^{+\infty} b_k < +\infty$: $\sum_{k=0}^{+\infty} a_k(x)$ uniformly converges on $[0,1]$.

Denote $B(p) = \int_0^T \frac{T-u}{\prod_{i=1}^d (1+pu^{2\alpha_i})^{1/2}} du = \sum_{k=0}^{d} B_k(p)$, with

$$\begin{aligned}
B_k(p) &= \int_{p^{-1/2\alpha_k}}^{p^{-1/2\alpha_{k+1}}} \frac{T-u}{\prod_{i=1}^d (1+pu^{2\alpha_i})^{1/2}} du \quad \text{for } k \in \{0,..,d-1\}, \\
B_d(p) &= \int_{p^{-1/2\alpha_d}}^{T} \frac{T-u}{\prod_{i=1}^d (1+pu^{2\alpha_i})^{1/2}} du
\end{aligned}$$

(with $1/\alpha_0 = +\infty$). If $k < d$ and $u \in [p^{-1/2\alpha_k}, p^{-1/2\alpha_{k+1}}]$ then $pu^{2\alpha_i} \geq 1$ for $i \in \{0,...,k\}$, and $pu^{2\alpha_i} \leq 1$ for $i \in \{k+1,...,d\}$. Using (23) and (24), one obtains:

$$\begin{aligned}
B_k(p) &= \int_{p^{-1/2\alpha_k}}^{p^{-1/2\alpha_{k+1}}} \frac{(T-u)}{p^{k/2}u^{S(k)}} \prod_{i=1}^k (1+p^{-1}u^{-2\alpha_i})^{-1/2} \prod_{i=k+1}^d (1+pu^{2\alpha_i})^{-1/2} du \\
&= \int_{p^{-1/2\alpha_k}}^{p^{-1/2\alpha_{k+1}}} \frac{(T-u)}{p^{k/2}u^{S(k)}} \prod_{i=1}^k \left(\sum_{l_i=0}^{+\infty} a_{l_i}(p^{-1}u^{-2\alpha_i})\right) \prod_{i=k+1}^d \left(\sum_{l_i=0}^{+\infty} a_{l_i}(pu^{2\alpha_i})\right) du.
\end{aligned}$$

Each series uniformly converges on $[p^{-1/2\alpha_k}, p^{-1/2\alpha_{k+1}}]$ and so, their products uniformly converges on $[p^{-1/2\alpha_k}, p^{-1/2\alpha_{k+1}}]$. Sumns and integrals can be reversed and

$$B_k(p) = \sum_{l_1=0}^{+\infty} \dots \sum_{l_d=0}^{+\infty} \int_{p^{-1/2\alpha_k}}^{p^{-1/2\alpha_{k+1}}} \frac{(T-u)}{p^{k/2}u^{S(k)}} \prod_{i=1}^k a_{l_i}(p^{-1}u^{-2\alpha_i}) \prod_{i=k+1}^d a_{l_i}(pu^{2\alpha_i}) du.$$

For each $(l_1, ..., l_d) \in \mathbb{N}^d$,

$$(T-u)p^{-k/2}u^{-S(k)}a_{l_1}(p^{-1}u^{-2\alpha_1})...a_{l_k}(p^{-1}u^{-2\alpha_k})a_{l_{k+1}}(pu^{2\alpha_{k+1}})...a_{l_d}(pu^{2\alpha_d})$$

is a sumn of $2d$ terms in $u$-power. By their integration on $[p^{-1/2\alpha_k}, p^{-1/2\alpha_{k+1}}]$, one obtains a sumn with terms in

$$\begin{array}{ll} & p^{-V(k)+\sum_{i=1}^{k} g_i(\alpha_i/\alpha_k-1)+\sum_{i=k+1}^{d} g_i(1-\alpha_i/\alpha_k)}, \\ & p^{-(V(k)+1/2\alpha_k)+\sum_{i=1}^{k} g_i(\alpha_i/\alpha_k-1)+\sum_{i=k+1}^{d} g_i(1-\alpha_i/\alpha_k)}, \\ \log p & p^{-d/2-\sum_{i=1}^{k} g_i+\sum_{i=k+1}^{d} g_i}, \\ & p^{-V(k+1)+\sum_{i=1}^{k} g_i(\alpha_i/\alpha_{k+1}-1)+\sum_{i=k+1}^{d} g_i(1-\alpha_i/\alpha_{k+1})}, \\ & p^{-(V(k+1)+1/2\alpha_{k+1})+\sum_{i=1}^{k} g_i(\alpha_i/\alpha_{k+1}-1)+\sum_{i=k+1}^{d} g_i(1-\alpha_i/\alpha_{k+1})} \end{array}$$

(with $g_i \in \mathbb{N}$). One proceeds as previously for all $(l_1, ..., l_d) \in \mathbb{N}^d$. The smallest power in $p$ is given for $g_i = 0$ and for all $i \in \{1, ..., d\}$; its coefficient can be computed by integration of $Tp^{-k/2}u^{-S(k)}$ on $[p^{-1/2\alpha_k}, p^{-1/2\alpha_{k+1}}]$. Finally, for $k \in \{0, 1, .., d-1\}$, $B_k(p)$ has the following asymptotic development when $p$ is large enough:

$$\text{-if } \alpha_k = \alpha_{k+1} \qquad B_k(p) = 0, \tag{25}$$

$$\text{-if } \alpha_k \neq \alpha_{k+1} \qquad B_k(p) = \left(\frac{T(\alpha_{k+1}-\alpha_k)}{2\alpha_{k+1}\alpha_k}\right)\frac{\log p}{p^{k/2}} + O\left(\frac{\log p}{p^{k/2+d_k}}\right) \text{ if } S(k) = 1, \tag{26}$$

$$B_k(p) = \left(\frac{T}{S(k)-1}\right)\frac{1}{p^{V(k)}} + O\left(\frac{\log p}{p^{V(k)+d_k}}\right) \text{ if } S(k) > 1, \tag{27}$$

$$B_k(p) = \left(\frac{T}{1-S(k)}\right)\frac{1}{p^{V(k+1)}} + O\left(\frac{\log p}{p^{V(k+1)+d_k}}\right) \text{ if } S(k) < 1, \tag{28}$$

where $d_k(\alpha_1, ..., \alpha_k)$ is a positive real number.

We proceed similarly for $B_d(p)$. If $u \in [p^{-1/2\alpha_d}, T]$ then $pu^{2\alpha_i} \geq 1$ for all $i \in \{1, 2, ..., d\}$ and then

$$\begin{aligned} B_d(p) &= \int_{p^{-1/2\alpha_d}}^{T} \frac{(T-u)}{p^{d/2}u^{S(d)}} \prod_{i=1}^{d}(1+p^{-1}u^{-2\alpha_i})^{-1/2} du \\ &= \int_{p^{-1/2\alpha_d}}^{T} \frac{(T-u)}{p^{d/2}u^{S(d)}} \prod_{i=1}^{d}\left(\sum_{l_i=0}^{+\infty} a_{l_i}(p^{-1}u^{-2\alpha_i})\right) du \\ &= \sum_{l_1=0}^{+\infty} ... \sum_{l_d=0}^{+\infty} \int_{p^{-1/2\alpha_d}}^{T} \frac{(T-u)a_{l_1}(p^{-1}u^{-2\alpha_1})...a_{l_d}(p^{-1}u^{-2\alpha_d})}{p^{d/2}u^{S(d)}} du. \end{aligned}$$

By the same way as above, we finally show:

$$\text{-if } S(d) = 1, \qquad B_d(p) = \left(\frac{T}{2\alpha_d}\right)\frac{\log p}{p^{d/2}} + O\left(\frac{\log p}{p^{d/2+d_d}}\right), \tag{29}$$

$$\text{-if } \; S(d) > 1, \qquad B_d(p) = \left(\frac{T}{S(d)-1}\right)\frac{1}{p^{V(d)}} + O\left(\frac{\log p}{p^{V(d)+d_k}}\right), \tag{30}$$

where $d_d(\alpha_1, ..., \alpha_d)$ is a positive real number.

If $S(d) \geq 1$ then $V(i_0) = \inf\{V(1), ..., V(d)\}$ where $i_0$ verifying $S(i_0 - 1) < 1$ and $S(i_0) \geq 1$. Using (25), (26), (27), (28), (29) and (30), one deduces that the main term of $B(p)$ asymptotic development is $\log p.p^{-V(i_0)}$ if $S(i_0) = 1$ or $p^{-V(i_0)}$ if $S(i_0) > 1$. It gives asymptotic developments (20) and (21). ■

From (20) and (21), it is reasonable to introduce the $\mathbb{R}^+$-valued random variable $m(p)$ such that:

- $m(p, \omega) = \dfrac{\hat{\mu}(p, \omega)p^{V(i_0)}}{\log p}$ when $S(i_0) = 1$
- $m(p, \omega) = \hat{\mu}(p, \omega)p^{V(i_0)}$ when $S(i_0) > 1$.

Lemma 2.4.1 implies $\mathbb{E}m(p) \underset{p\to+\infty}{\longrightarrow} m_0(\alpha_1, ..., \alpha_d, T)$. We now show quadratic mean convergence of $m(p)$:

**Lemma 2.4.2** *If $S(d) > 1$ there exist two positive real numbers $\zeta(\alpha_1, ..., \alpha_d, T)$ and $M(\alpha_1, ..., \alpha_d, T)$, such that for large enough $p$:*

$$\mathbb{E}(m(p) - m_0)^2 \leq \frac{M}{p^{2\zeta}}$$

**Proof**: We have:

$$\mathbb{E}(\hat{\mu}(p))^2 = \mathbb{E}\left[\int_0^T \int_0^T e^{-p\|X_t - X_s\|^2} dsdt\right]^2$$

$$= \int_0^T \int_0^T \int_0^T \int_0^T \prod_{i=1}^d \mathbb{E}\left[\exp\left(-\frac{p((X_t^i - X_s^i)^2 + (X_{t'}^i - X_{s'}^i)^2)}{2a_i}\right)\right] dsdtds'dt'$$

Letting $W^i_{s,t,s',t'} = \dfrac{1}{\sqrt{2a_i}}\begin{pmatrix} X_t^i - X_s^i \\ X_{t'}^i - X_{s'}^i \end{pmatrix}$, then:

$$\mathbb{E}\left[\exp\left(-\frac{p((X_t^i - X_s^i)^2 + (X_{t'}^i - X_{s'}^i)^2)}{2a_i}\right)\right] = \frac{1}{\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'}))^{1/2}}$$

(det is determinant, $I_2$ is the identity matrix of $\mathbb{R}^2$ and cov is covariance matrix). Kono showed in proof of Theorem 2 [15] existence of $0 < \epsilon_i \leq 1$ such that, for all $(s, t, s', t') \in [0, T]^4$:

$$\det(\mathrm{cov}(W^i_{s,t,s',t'})) \geq \epsilon_i^2 \mathbb{E}(X_t^i - X_s^i)^2 \mathbb{E}(X_{t'}^i - X_{s'}^i)^2.$$

Thus, $\forall (s,t,s',t') \in [0,T]^4$,

$$\begin{aligned}\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'})) &\geq 1 + p(|t-s|^{2\alpha_i} + |t'-s'|^{2\alpha_i}) + \epsilon_i^2(|t-s|^{2\alpha_i}|t'-s'|^{2\alpha_i}) \\ &\geq \epsilon_i^2(1+p(|t-s|^{2\alpha_i})(1+p(|t'-s'|^{2\alpha_i}).\end{aligned} \tag{31}$$

Moreover, $\forall (s,t,s',t') \in [0,T]^4$,

$$\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'})) = 1 + p(|t-s|^{2\alpha_i} + |t'-s'|^{2\alpha_i})$$
$$+p^2(|t-s|^{2\alpha_i}|t'-s'|^{2\alpha_i} - \frac{1}{4}(|t-s'|^{2\alpha_i} + |t'-s|^{2\alpha_i} - |t-t'|^{2\alpha_i} - |s-s'|^{2\alpha_i})^2).$$

If $P(u,v,w) = (|w+u|^{2\alpha_i} + |w-v|^{2\alpha_i} - |u-v+w|^{2\alpha_i} - |w|^{2\alpha_i})^2$ then, for $|w| \geq p^{-1/4}$, $|u| \leq p^{-1/2}$ and $|v| \leq p^{-1/2}$, and large enough $p$,

$$\begin{aligned}P(u,v,w) &= |w|^{4\alpha_i}(|1+\frac{u}{w}|^{2\alpha_i} + |1-\frac{v}{w}|^{2\alpha_i} - |1+\frac{u-v}{w}|^{2\alpha_i} - 1)^2 \\ &= 4\alpha_i^2(2\alpha_i-1)^2|w|^{4\alpha_i-4}\left(4u^2v^2 + O(\frac{u^3}{w}) + O(\frac{v^3}{w})\right) \\ &\leq 4\eta_i p^{1-\alpha_i}u^2v^2,\end{aligned}$$

with $0 < \eta_i \leq 8\alpha_i^2(2\alpha_i-1)^2$ for large enough $p$. This implies:

$$\begin{aligned}|u|^{2\alpha_i}|v|^{2\alpha_i} - \frac{1}{4}P(u,v,w) &\geq |u|^{2\alpha_i}|v|^{2\alpha_i}(1-\eta_i p^{1-\alpha_i}|u|^{2\alpha_i-2}|v|^{2\alpha_i-2}) \\ &\geq (1-\eta_i p^{\alpha_i-1})|u|^{2\alpha_i}|v|^{2\alpha_i}.\end{aligned}$$

Thus, if $|t-s| \leq p^{-1/2}$, $|t'-s'| \leq p^{-1/2}$ and $|s-s'| \geq p^{-1/4}$, when $p$ is large enough,

$$\begin{aligned}\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'})) &\geq 1 + p(|t-s|^{2\alpha_i} + |t'-s'|^{2\alpha_i}) + p^2(1-\eta_i p^{\alpha_i-1})|t-s|^{2\alpha_i}|t'-s'|^{2\alpha_i} \\ &\geq (1-\eta_i p^{\alpha_i-1})(1+p(|t-s|^{2\alpha_i})(1+p|t'-s'|^{2\alpha_i})\end{aligned} \tag{32}$$

Now,

$$\begin{aligned}\mathbb{E}(\hat{\mu}(p))^2 &= \int_0^T\int_0^T\int_0^T\int_0^T \prod_{i=1}^d \frac{1}{\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'}))^{1/2}} ds dt ds' dt' \\ &= \int_{[0,T]^4} \mathbb{I}_{|t-s|\leq p^{-1/2}}\mathbb{I}_{|t'-s'|\leq p^{-1/2}}\mathbb{I}_{|s-s'|\geq p^{-1/4}} \prod_{i=1}^d \frac{1}{\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'}))^{1/2}} ds dt ds' dt' \\ &+ \int_{[0,T]^4}(1-\mathbb{I}_{|t-s|\leq p^{-1/2}}\mathbb{I}_{|t'-s'|\leq p^{-1/2}}\mathbb{I}_{|s-s'|\geq p^{-1/4}}) \prod_{i=1}^d (\frac{1}{\det(I_2 + p\mathrm{cov}(W^i_{s,t,s',t'}))^{1/2}} ds dt ds' dt'\end{aligned}$$

Denote $I_1$ and $I_2$ both previous integrals. With (32),

$$I_1 \leq \int_{[0,T]^4} \mathbb{I}_{|t-s|\leq p^{-1/2}}\mathbb{I}_{|t'-s'|\leq p^{-1/2}}\mathbb{I}_{|s-s'|\geq p^{-1/4}} \prod_{i=1}^d \frac{(1-\eta_i p^{\alpha_i-1})^{-1/2}}{(1+p(|t-s|^{2\alpha_i})^{1/2}(1+p|t'-s'|^{2\alpha_i})^{1/2}}$$

$$\begin{aligned}
&\leq \quad 4\prod_{i=1}^{d}(1-\eta_i p^{\alpha_i-1})^{-1/2}\left(\int_0^{p^{-1/2}}\frac{T-u}{\prod_{i=1}^{d}(1+pu^{2\alpha_i})^{1/2}}du\right)^2\\
&\leq \quad \prod_{i=1}^{d}(1-\eta_i p^{\alpha_i-1})^{-1/2}(\mathbb{E}\hat{\mu}(p))^2. \qquad (33)
\end{aligned}$$

As $1-\mathbb{I}_{|t-s|\leq p^{-1/2}}\mathbb{I}_{|t'-s'|\leq p^{-1/2}}\mathbb{I}_{|s-s'|\geq p^{-1/4}} \leq \mathbb{I}_{|t-s|\geq p^{-1/2}}+\mathbb{I}_{|t'-s'|\geq p^{-1/2}}+\mathbb{I}_{|s-s'|\leq p^{-1/4}}$ and from (31), one obtains

$$\begin{aligned}
I_2 &\leq \int_{[0,T]^4}\mathbb{I}_{|t-s|\geq p^{-1/2}}\prod_{i=1}^{d}\frac{\epsilon_i^{-1}}{(1+p(|t-s|^{2\alpha_i})^{1/2}(1+p|t'-s'|^{2\alpha_i})^{1/2}}dsdtds'dt'\\
&+ \int_{[0,T]^4}\mathbb{I}_{|t'-s'|\geq p^{-1/2}}\prod_{i=1}^{d}\frac{\epsilon_i^{-1}}{(1+p(|t-s|^{2\alpha_i})^{1/2}(1+p|t'-s'|^{2\alpha_i})^{1/2}}dsdtds'dt'\\
&+ \int_{[0,T]^4}\mathbb{I}_{|s-s'|\leq p^{-1/4}}\prod_{i=1}^{d}\frac{\epsilon_i^{-1}}{(1+p(|t-s|^{2\alpha_i})^{1/2}(1+p|t'-s'|^{2\alpha_i})^{1/2}}dsdtds'dt'\\
&\leq 2(\prod_{i=1}^{d}\epsilon_i^{-1})(\mathbb{E}\hat{\mu}(p))\int_{p^{-1/2}}^{T}\frac{2(T-u)}{\prod_{i=1}^{d}(1+pu^{2\alpha_i})^{1/2}}du\\
&+ 4(\prod_{i=1}^{d}\epsilon_i^{-1})\int_{[-T,T]^4}\mathbb{I}_{|w|\leq p^{-1/4}}\frac{(T-u)(T-v)}{\prod_{i=1}^{d}(1+pu^{2\alpha_i})^{1/2}(1+pv^{2\alpha_i})^{1/2}}dudvdwdz\\
&\leq 2(\prod_{i=1}^{d}\epsilon_i^{-1})(\mathbb{E}\hat{\mu}(p))\frac{\log p}{p^{d/2}}+\frac{16T\prod_{i=1}^{d}\epsilon_i^{-1}}{p^{1/4}}(\mathbb{E}\hat{\mu}(p))^2 \qquad (34)
\end{aligned}$$

(one bounds $\int_{p^{-1/2}}^{T}\frac{(T-u)}{\prod_{i=1}^{d}(1+pu^{2\alpha_i})^{1/2}}du$ by the same computation than $B_d(p)$ one in lemma 2.4.1). (33) and (34) imply that:

$$\mathbb{E}(\hat{\mu}(p))^2 \leq (\mathbb{E}\hat{\mu}(p))^2\left(\prod_{i=1}^{d}(1-\eta_i p^{\alpha_i-1})^{-1/2}+\frac{16T\prod_{i=1}^{d}\epsilon_i^{-1}}{p^{1/4}}+2(\prod_{i=1}^{d}\epsilon_i^{-1})\frac{\log p}{\mathbb{E}\hat{\mu}(p)p^{d/2}}\right)$$

As $S(d)>1$, $\mathbb{E}\hat{\mu}(p)\sim\frac{m_0}{p^{V(i_0)}}$ and $V(i_0)<d$, thus $\frac{\log p}{\mathbb{E}\hat{\mu}(p)p^{d/2}}\leq\frac{1}{p^{(d/4-V(i_0)/4)}}$ for large enough $p$. So, there are two positive real numbers $\zeta_1(\alpha_1,...,\alpha_d,T)$ and $M_1(\alpha_1,...,\alpha_d,T)$ such that:

$$\begin{aligned}
\mathbb{E}(\hat{\mu}(p))^2-(\mathbb{E}\hat{\mu}(p))^2 &\leq \frac{M_1}{p^{2\zeta_1}}(\mathbb{E}\hat{\mu}(p))^2,\\
\mathbb{E}(m(p)-\mathbb{E}m(p))^2 &\leq \frac{M_1}{p^{2\zeta_1}}. \qquad (35)
\end{aligned}$$

Using Lemma 2.4.1, there are two positive real numbers $\zeta_2(\alpha_1,...,\alpha_d,T)$ and $M_2(\alpha_1,...,\alpha_d,T)$ such that

$$(\mathbb{E}m(p)-m_0)^2\leq\frac{M_2}{p^{2\zeta_2}}. \qquad (36)$$

By (35) and (36), one deduces Lemma 2.4.2.■

**Lemma 2.4.3** *If $S(d)=1$ there is a positive real number $M(\alpha_1,...,\alpha_d,T)$, such that for large enough $p$,*

$$\mathbb{E}(m(p)-m_0)^2 \leq M\frac{\log\log p}{\log p}.$$

**Proof**: We proceed as above, with $|s-s'|\geq 1/\log p$ in (32). For large enough $p$, $\log p^{-1}\geq p^{-1/4}$. So, for all $(s,t,s',t')\in[0,T]^4$ verifying $|t-s|\leq p^{-1/2}$, $|t'-s'|\leq p^{-1/2}$ and $|s-s'|\geq 1/\log p$, there exist $0<\eta_i\leq 8\alpha_i^2(2\alpha_i-1)^2$ for $i\in\{1,2,...,d\}$, such as,

$$\det(I_2+p\mathrm{cov}(W^i_{s,t,s',t'}))\geq(1-\eta_i p^{\alpha_i-1})(1+p(|t-s|^{2\alpha_i})(1+p|t'-s'|^{2\alpha_i}). \qquad (37)$$

As a consequence, if we denote $\mathbb{E}(\hat{\mu}(p))^2=I'_1+I'_2$, where $I'_1$ and $I'_2$ are $I_1$ and $I_2$ with $\mathbb{I}_{|s-s'|\geq(\log p)^{-1}}$ instead of $\mathbb{I}_{|s-s'|\geq p^{-1/4}}$, then,

$$\begin{aligned}
I'_1 &\leq \prod_{i=1}^d(1-\eta_i p^{\alpha_i-1})^{-1/2}(\mathbb{E}\hat{\mu}(p))^2\\
I'_2 &\leq 2(\prod_{i=1}^d\epsilon_i^{-1})(\mathbb{E}\hat{\mu}(p))\int_{(\log p)^{-1}}^T\frac{2(T-u)}{\prod_{i=1}^d(1+pu^{2\alpha_i})^{1/2}}du\\
&+ 4(\prod_{i=1}^d\epsilon_i^{-1})\int_{[-T,T]^4}\mathbb{I}_{|w|\leq(\log p)^{-1}}\frac{(T-u)(T-v)}{\prod_{i=1}^d(1+pu^{2\alpha_i})^{1/2}(1+pv^{2\alpha_i})^{1/2}}dudvdwdz\\
&\leq 8(\prod_{i=1}^d\epsilon_i^{-1})(\mathbb{E}\hat{\mu}(p))\frac{\log\log p}{p^{d/2}}+\frac{16T\prod_{i=1}^d\epsilon_i^{-1}}{\log p}(\mathbb{E}\hat{\mu}(p))^2
\end{aligned}$$

As $S(d)=1$, $\mathbb{E}\hat{\mu}(p)\sim m_0\frac{\log p}{p^{d/2}}$, thus $\frac{\log\log p}{\mathbb{E}\hat{\mu}(p)p^{d/2}}\leq\frac{2\log\log p}{m_0\log p}$ for large enough $p$. One then obtains:

$$\mathbb{E}(\hat{\mu}(p))^2\leq(\mathbb{E}\hat{\mu}(p))^2\left(\prod_{i=1}^d(1-\eta_i p^{\alpha_i-1})^{-1/2}+\frac{16T\prod_{i=1}^d\epsilon_i^{-1}}{\log p}+16(\prod_{i=1}^d\epsilon_i^{-1})\frac{\log\log p}{m_0\log p}\right)$$

As a result, there exists a positive real number $L(\alpha_1,...,\alpha_d,T)$ such that for large enough $p$,

$$\begin{aligned}
\mathbb{E}(\hat{\mu}(p))^2-(\mathbb{E}\hat{\mu}(p))^2 &\leq \frac{Lm_0^2\log\log p}{\log p}(\mathbb{E}\hat{\mu}(p))^2\\
\text{and thus, }\ \mathbb{E}(m(p)-\mathbb{E}m(p))^2 &\leq L\frac{\log\log p}{\log p}.
\end{aligned}$$

We have then shown Lemma 2.4.3 after a similar conclusion to proof of Lemma 2.4.2. ■

If $S(d)>1$ then one can obtain a.s. convergence of $m(p)$ to $m_0$.

**Lemma 2.4.4** *If $S(d) > 1$ then $m(p) \xrightarrow[p\to+\infty]{a.s.} m_0$.*

**Proof**: Denote $u_n = m(n^{1/\zeta})$. By Lemma 2.4.2, $\mathbb{E}(u_n - \mathbb{E}u_n)^2 \leq \frac{M}{n^2}$. Thus, with Tchebychev Inequality, Borel-Cantelli Lemma and using $\mathbb{E}u_n \to m_0$ when $n \to +\infty$, one obtains:

$$u_n \xrightarrow[n\to+\infty]{a.s.} m_0. \tag{38}$$

$(\hat{\mu}(p))_p$ is a decreasing sequence for all $\omega$ (because $\hat{\mu}(p,\omega) = \int_0^T \int_0^T e^{-p\|X_t(\omega)-X_s(\omega)\|^2} dsdt$ and $\| X_t(\omega) - X_s(\omega) \|^2$ is a $\mathbb{R}_+$-valued field). So, for $p \in [n^{1/\zeta}, (n+1)^{1/\zeta}]$ and for all $\omega$,

$$\frac{n^{V(i_0)/\zeta}}{(\log(n+1)^{1/\zeta})^{e_0}}\hat{\mu}((n+1)^{1/\zeta},\omega) \leq \frac{p^{V(i_0)}}{(\log p)^{e_0}}\hat{\mu}(p,\omega) \leq \frac{(n+1)^{V(i_0)/\zeta}}{(\log n^{1/\zeta})^{e_0}}\hat{\mu}(n^{1/\zeta},\omega),$$
$$\left(\frac{n}{n+1}\right)^{V(i_0)/\zeta}\left(\frac{\log n}{\log(n+1)}\right)^{e_0} u_{n+1} \leq m(p) \leq \left(\frac{n+1}{n}\right)^{V(i_0)/\zeta}\left(\frac{\log(n+1)}{\log n}\right)^{e_0} u_n.$$

One then uses $n \to +\infty$ and (38) and it achieves to prove

$$m(p) \xrightarrow[p\to+\infty]{a.s.} m_0. \blacksquare \tag{39}$$

It's now possible to show Propositions 2.3.2 and 2.3.3:

**Proof of Proposition 2.3.3**: If $S(d) > 1$, one has just shown that $m(p) \xrightarrow[p\to+\infty]{a.s.} m_0$. So, for almost every $\omega$,

$$\hat{\mu}(p,\omega) \sim \frac{m_0(\log p)^{e_0}}{p^{V(i_0)}} \text{ when } p \to +\infty, \tag{40}$$

where $e_0 = 1$ if $V(i_0) \in \mathbb{N}$ else $e_0 = 0$. For each $\omega$ verifying (40), one applies a Karamata Tauberian theorem to $\hat{\mu}(p,\omega)$. More precisely, we have conditions for applying Theorem 3, chap. XIII, of Feller (1971) (the possible logarithm of (40) is a slow varying function). As a consequence, for almost every $\omega$,

$$\mu(r,[0,T]^2,\omega) \sim \frac{m_0}{\Gamma(V(i_0)+1)}(\log\frac{1}{r})^{e_0} r^{V(i_0)} \text{ when } r \to 0, \tag{41}$$
$$\text{thus } C_T(r,\omega) \sim \frac{m_0}{\Gamma(V(i_0)+1)T^2}(2\log\frac{1}{r})^{e_0} r^{2V(i_0)} \text{ when } r \to 0. \tag{42}$$

One deduces the almost sure convergence of $\nu$ to $2V(i_0)$ and (14) is shown (using $2V(i_0) < d$ when $S(d) > 1$). $\blacksquare$.

**Proof of Proposition 2.3.2**: If $S(d) = 1$, we cannot prove the almost sure convergence as in previous proof. Nevertheless, we have shown in Lemma 2.4.3 that $m(p)$ converges to $m_0$ when $p \to +\infty$ in quadratic mean. Thus, there exists a $\mathbb{R}^+$-valued random variable such that for all $\omega$,

$$\frac{p^{d/2}}{\log p}\hat{\mu}(p,\omega) \underset{p\to+\infty}{\longrightarrow} \beta(\omega). \tag{43}$$

Moreover, $\mathbb{E}\beta = m_0$ and $\mathbb{E}\beta^2 = m_0^2$. For each $\omega$, one uses the same Tauberian theorem as above, and

$$C_T(r,\omega) \quad \sim \quad \frac{2\beta(\omega)}{\Gamma(d/2+1)T^2} r^d \log\frac{1}{r} \text{ when } r \to 0. \tag{44}$$

If $\beta > 0$ then $\nu$ exists and $\nu = d$. As $\beta = 0$ with probability 1 is impossible, then $\nu$ exists and $\nu = d$ with positive probability. ∎

**Acknowledgment.** The author expresses his gratitude to Prof. D. Dacunha-Castelle. He thanks Prof. M. Wschebor and J.P. Kahane for several stimulating conversations.

# Bibliographie


[1] Adler, R. (1981). *Geometry of random fields.* Wiley, New York.

[2] Berman, S. (1973). Local non-determinism and local times of Gaussian processes. *Indiana Math. J.* **23** 69-94.

[3] Cutler, C. (1994). A theory of correlation dimension for stationary time series. *Philos. Trans. R. Soc. Lond. A* **348** 343-355.

[4] Cuzick, J. (1978). Some local properties of Gaussian vector fields. *Ann. Probab.* **6** 984-994.

[5] Cuzick, J. (1982). Multiple points of a Gaussian vector fields. *Z. Warsch. Verw. Gebiete* **61** 431-436.

[6] Ding, M., Gregobi, C., Ott, E., Sauer, T., and Yorke, J. (1993). Estimating correlation dimension from a chaotic time series: when does plateau onset occur? *Physica D* **69** 404-424.

[7] Dynkin, E. (1988). Self-intersection gauge for random walks and for Brownian motion. *Ann. Probab.* **16** 1-57.

[8] Eckmann, J.-P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57** 617-656.

[9] Feller, W. (1971). *An introduction to probability theory and its applications.* Vol. 2, Wiley.

[10] Frostman, O. (1935). Potemtiel d'équilibre et capacité des ensembles, avec quelques applications à la théorie des fonctions. *Medd. Lunds. Univ. Mat. Semin.* **3**.

[11] Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* **8** 1-67.

[12] Grassberger, P. and Proccacia, I. (1983). Measuring the strangeness of strange attractors. *Physica D* **9** 189-208.

[13] Imkeller, P., Perez-Abreu, V. and Vives, J. (1995). Chaos expansions of double intersection local time of Brownian motion in $\mathbb{R}^d$ and renormalization. *Stoch. Pro. Appl.* **56** 1-34.

[14] Kametani, S. (1944). On Hausdorff's measures and generalized capacities with some of their applications to the theory of functions. *Japanese J. Math.* **19** 217-257.

[15] Kono, N. (1978). Double points of a Gaussian path. *Z. Warsch. Verw. Gebiete* **45** 175-180.

[16] Le Gall, J.F. (1985). Sur le temps local d'intersection du mouvement brownien plan et la méthode de renormalisation de Varadhan. *Séminaire de Probabilités XIX. Lecture Notes in Math.* **1123** 314-331.

[17] Pitt, L. (1978). Local times for Gaussian vector fields. *Indiana Univ. Math. J.* **27** 309-330.

[18] Rosen, J. (1983). A local time approach to the self-intersections of Brownian paths in space. *Comm. Math. Phys.* **88** 327-338.

[19] Rosen, J. (1984). Self-intersections of random fields. *Ann. Probab.* **12** 108-119.

[20] Rosen, J. (1987). The intersection local time of fractional Brownian motion in the plane. *J. Mult. Anal.* **23** 37-46.

[21] Rosen, J. (1988). Continuity and singularity of the intersection local time of stable processes in $\mathbb{R}^2$. *Ann. Probab.* **16** 75-79.

[22] Smith, R.L. (1992). Estimating dimension in noisy chaotic time series. *J. R. Stat. Soc. B* **54** 329-351.

# Chapitre 3

# Tester l'autosimilarité des processus gaussiens à accroissements stationnaires

Testing self-similarity of Gaussian processes with stationary increments

By Jean-Marc Bardet

*Université Paris-Sud and Université Evry-Val d'Essonne*

We present a method for testing self-similarity of discretized observations of a Gaussian process with stationary increments. The test is based on the estimation of a distance between the process and a set of processes containing all the fractional Brownian motions. This distance is based on two estimations of given variances of increments. The second one requires regression estimates of the self-similarity index $H$. Two estimators of $H$ are then introduced. They present good robustness and computing time properties. The test also gives a method to find a sampling period where fractional Brownian motion modeling should be possible.

## 3.1 Introduction

The self-similarity property has led to a big variety of important results, in deterministic settings as well as in stochastic ones. For instance, many physical or financial phenomena

*Key words and phrases.* Self-similarity; Fractional Brownian Motion; Gaussian Processes; Estimation of the self-similarity index.

*AMS 1991 subject classifications.* Primary 62G10; secondary 60F05, 60G18, 62F10.

have been modeled with self-similar (S.S.) processes (see Mandelbrot and Van Ness (1968), Mandelbrot (1983), Beran (1994)). The aim of our work is to test the self-similarity of discretized Gaussian processes with stationary increments.
A scalar stochastic process $X = \{(X_t), t \in \mathbb{R}\}$ is a S.S. process with index $H > 0$ if, for any $d \in \mathbb{N}^*$, for any $(t_1, .., t_d) \in \mathbb{R}^d$, and for any $c \in \mathbb{R}_+$:

$$(X_{t_1}, .., X_{t_d}) \stackrel{\mathcal{D}}{=} (c^{-H} X_{ct_1}, .., c^{-H} X_{ct_d}).$$

Fractional Brownian motions (F.B.M.), fractional ARIMA processes, some $\alpha$-stable processes, etc ..., are S.S. processes (for more details, see the book of Samorodnitsky and Taqqu (1994)). Nevertheless, $X$ is a centered S.S. Gaussian process (with $X_0 = 0$ a.s.), with stationary increments if and only if $X$ is a F.B.M. with index $H \in ]0, 1]$. As a consequence, for any $t \in \mathbb{R}$:

$$\mathbb{E} X_t^2 = \mid t \mid^{2H} \mathbb{E}(X_1^2). \tag{1}$$

The following test of self-similarity is based on this property.

Assume that an observation $X_0, X_1, X_2, .., X_N$ is given by a discretization of a F.B.M. with index $H \in ]0, 1[$. Let $n_1, .., n_l$ be distinct positive integers, and let $\gamma(n_i)$ be the increment variances of length $n_i$ (for $i = 1, .., l$). The test statistic is a distance between two estimations of the vector $(\log \gamma(n_1), .., \log \gamma(n_l))$. The first one is deduced from an estimation of $(\gamma(n_1), .., \gamma(n_l))$ with usual sample variances. The second one is a consequence of property (1): if $\hat{H}$ and $\hat{K}$ are estimators of $H$ and $\log \gamma(1)$, then vector $(2\hat{H} \log n_1 + \hat{K}, .., 2\hat{H} \log n_l + \hat{K})$ is an estimation of $(\log \gamma(n_1), .., \log \gamma(n_l))$. The considered distance has a known asymptotic distribution and is also proportional to a positive power of $N$. As a result, the test is only accepted by a set of processes containing all F.B.M.

We have chosen to estimate $\gamma(n_i)$ with empirical variances $S_N(n_i)$, computed with $X_{kn_i}$, where $k = 0, 1, .., [N/n_i]$. If $X$ is F.B.M., $(S_N(n_1), .., S_N(n_l))$ converges to $(\gamma(n_1), .., \gamma(n_l))$. We show central limit (when $0 < H < 3/4$) and non-central limit (when $3/4 < H < 1$) theorems for $(S_N(n_1), .., S_N(n_l))$ and $(\log S_N(n_i), .., \log S_N(n_l))$. These results are deduced from Taqqu (1975), Fox and Taqqu (1987), Giraitis and Surgailis (1990) and Poggi and Viano (1996).
We estimate $H$ and $\log \gamma(1)$ from previous results by regression of $\log S_N(n_i)$ in terms of $\log n_i$. We obtain ordinary least square estimators (O.L.S.) of $H$ and $K$. It is the aggregated variance method, also called log-variance plot method (see Beran (1994), Istas and Lang (1997) or Poggi and Viano (1996)). Consequently, if $X$ is F.B.M. then $(2\hat{H} \log n_1 + \hat{K}, .., 2\hat{H} \log n_l + \hat{K})$ and $(\log S_N(n_1), .., \log S_N(n_l))$ have linearly equivalent asymptotic behaviours. A normalized distance $LM(N)$ between both estimations of $(\log \gamma(n_1), .., \log \gamma(n_l))$ has known asymptotic distribution only depending on $(n_1, .., n_l)$

and $H$.
If $X$ is F.B.M. with $0 < H < 3/4$, $(\sqrt{N} \log S_N(n_1), .., \sqrt{N} \log S_N(n_l))$ is asympotically Gaussian with identifiable asymptotic variance. Therefore, we obtain a new regression line by a generalized least squares estimation of $H$ and $K$. We construct a new distance $SM(N)$ between this second regression line and each $\log S_N(n_i)$, weighted by $N$. The distribution of $SM^2(N)$ converges with $N$ to a $\chi^2_{l-2}$ distribution. In this case, the test is very simple to use.
We also present an extension of this test. In fact, $(\log S_N(n_1), .., \log S_N(n_l))$ converges to $(\log \gamma(n_1), .., \log \gamma(n_l))$ if $X$ is in a set of centered Gaussian processes with stationary increments. The tested property is then: there exist two real numbers $H'$ and $K'$ such as for any $i = 1, .., l$: $\log \gamma(n_i) = 2H' \log n_i + K'$. We call this property: self-similarity for scales $(n_1, .., n_l)$. We use the same test statistics $LM$ and $SM$ as above and show that they diverge if $X$ is not S.S. for scales $(n_1, .., n_l)$.

We have chosen to work with on the time-parameter space instead of spectral one for simplicity and robustness reasons. The S.S. property is then directly tested. But a dichotomy ($0 < H < 3/4$ and $3/4 < H < 1$) is introduced. This problem is practically solved by the continuity of the asymptotic distribution of $LM$ or $SM$. Moreover, one easily obtains confidence intervals for L.S. estimators. It provides informations about the choice of test ($SM$ or $LM$).
This dichotomy could be avoided with other estimations of $H$. For $0 < H < 1$ we obtain central limit theorems with regression of log-periodogram (see Geweke and Porter-Hudak (1983), Yajima (1989) or Robinson (1995)), or with maximum likelihood (M.L.) estimation (see Fox and Taqqu (1986), Giraitis and Surgailis (1990) or Beran (1994)). These other estimations could be useful if we change the test statistic. For example, the test statistic can be a distance between two estimations of spectral density $X_{kn_i}$ with $i = 1, .., l$ (see Robinson (1995)). Such a test avoid the dichotomy but it is also less robust than variance test. It requires more of data and chosen frequencies have to be close to zero. Moreover, asymptotic covariance computations take a long time (L.S. are done with $m$ data and $m \to +\infty$ when $N \to +\infty$). Furthermore, the best convergent rate of test is less than $N^{4/5}$ instead of $N$ for variance test when $0 < H < 3/4$.

Simulations show test simplicity and robustness. In fact, as estimations of $H$ are linear and explicit, computing time is very short in comparison with M.L. approach. Moreover, these estimations of $H$ are robust in comparison with M.L. estimation. Asymptotical properties of the test will be confirmed by numerical simulations using a F.B.M. generator. We have also tested real data that seem to be a skeleton of a F.B.M. (because of aggregation phenomena for example). S.S. test allows us to find a new sampling period for which F.B.M. modeling is adapted.

The following section introduces the first estimation of $(\log \gamma(n_1), .., \log \gamma(n_l))$. Section 3 presents the second estimation of $(\log \gamma(n_1), .., \log \gamma(n_l))$ and both L.S. estimators of $H$. Test statistics are intoduced in Section 4 and their extensions are presented in Section 5. Simulation results appear in Section 6. Proofs of main propositions are displayed in Section 7.

## 3.2 First estimation of $(n_i)$-length increment variances

For all following sections, we consider $(X_0, .., X_N)$ a time series where $X = \{(X_t), t \in \mathbb{R}\}$ is a stochastic process and $n_1, .., n_l$, some distinct positive integers. We will make the following assumption:

**F.B.M. assumption**: *Let $X$ be a fractional Brownian motion (F.B.M.) with index $H \in ]0,1[-\{3/4\}$.*

Set $\gamma(t) = \mathbb{E}(X_{t+s} - X_s)^2$ for any $(t,s) \in \mathbb{R}_+^2$. With property (1) and under F.B.M. assumption,

$$\forall t \in \mathbb{R}, \quad \gamma(t) = \gamma(1) \mid t \mid^{2H} . \tag{2}$$

We use empirical variances $S_N(n)$ for estimating $n$-length increments of $X$ $(n \in \mathbb{N})$ and:

$$S_N(n) = \frac{1}{[N/n]} \sum_{j=1}^{[N/n]} (X_{jn} - X_{(j-1)n})^2. \tag{3}$$

**Remark**: Another $\gamma(n)$ estimation is $S_N^{(2)}(n)$ such as $S_N^{(2)}(n) = \dfrac{1}{N-n+1} \displaystyle\sum_{j=0}^{N-n} (X_{j+n} - X_j)^2$. $S_N^{(2)}(n)$ is interessant because its computation takes into account all data. But we have $Var(S_N^{(2)}(n)) \geq Var(S_N(n))$ and the use of $S_N(n)$ allows us to choose a new sampling period.

Then we have:

**Proposition 3.2.1** *Under F.B.M. assumption, and for any $n \in \mathbb{N}$:*

*(1)* $\mathbb{E}S_N(n) = \gamma(n)$

*(2) For $N \to +\infty$, $S_N(n) \to \gamma(n)$ in q.m. and a.s.*

**Proof**: $m_n(j) = X_{jn} - X_{(j-1)n}$ are Gaussian variables such as:

$$\lim_{|j-j'|\to\infty} cov(m_n(j), m_n(j')) = 0.$$

Consequently, $(m_n(j))_j$ is a Gaussian stationary ergodic process (for instance, see Azencott and Dacunha-Castelle (1984)). Ergodic theorem implies the previous proposition for $S_N(n) = \frac{1}{[N/n]} \sum_{j=1}^{[N/n]} m_n^2(j)$ when $N \to +\infty$. ■

It is possible to give limit theorems for vector $(S_N(n_1), .., S_N(n_l))$. Two cases have to be distinguished. If $H \in ]0, 3/4[$, we can use results of Giraitis and Surgailis (1990) and Poggi and Viano (1996) and have a central limit theorem. If $H \in ]3/4, 1[$, we can use results of Taqqu (1975) or Fox and Taqqu (1987) and have a non-central limit theorem.

**Proposition 3.2.2** *Under F.B.M. assumption:*
*1- if* $H \in ]0, 3/4[$: $\sqrt{N}\,(S_N(n_i) - \gamma(n_i))_{1\le i\le l} \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}_l(0; F)$,
*where* $F = (f_{ij})_{1\le i,j\le l}$ *is the matrix such as, with* $d_{ij} = GCD(n_i, n_j)$:

$$f_{ij} = 2d_{ij}\gamma(1)^2 \sum_{k=-\infty}^{+\infty} \left( | \, kd_{ij} - n_j \, |^{2H} + | \, kd_{ij} + n_i \, |^{2H} - | \, kd_{ij} + n_i - n_j \, |^{2H} - | \, kd_{ij} \, |^{2H} \right)^2.$$

*2- if* $H \in ]3/4, 1[$:

$$N^{2-2H}\,(S_N(n_i) - \gamma(n_i))_{1\le i\le l} \xrightarrow[N\to+\infty]{\mathcal{D}} H(2H-1)\gamma(1)^t(n_1^2\mathcal{Z}_1, .., n_l^2\mathcal{Z}_l),$$

*where* $\mathcal{Z}_1, .., \mathcal{Z}_l$ *are identically distributed centered random variables (characteristic function of* $(\mathcal{Z}_1, .., \mathcal{Z}_l)$ *is given in proposition 3.5.1).*

Proposition 3.2.2 can be extended if $X$ is disturbed by a white noise:

**Corollary 3.2.1** *Under F.B.M. assumption, one considers* $(X'_0, X'_1, .., X'_N)$, *where* $X'_i = X_i + \epsilon_i^N$ *with* $(\epsilon_i^N)$ *a white noise of variance* $\sigma^2(N)$, *independent of* $(X_i)$. *Let* $(S'_N(n_i))_{i\in\{1,2,..,l\}}$ *be the vector computed from* $(X'_0, X'_1, .., X'_N)$. *Then,* $(S'_N(n_1), .., S'_N(n_l))$ *has the same asymptotic convergence as* $(S_N(n_1), .., S_N(n_l))$ *in proposition 3.2.2, if:*
*1- for* $H \in ]0, 3/4[$, $\lim_{N\to+\infty} \sigma^2(N) = 0$.
*2- for* $H \in ]3/4, 1[$, $\sigma^2(.)$ *is bounded on* $\mathbb{N}$.

This last result will be interesting in practical applications, because of the robustness of statistic $(S_N(n_1), .., S_N(n_l))$ to measurement or observation errors. Proposition 3.2.2 also gives an estimation of $(\log\gamma(n_1), .., \log\gamma(n_l))$. With regularity on $\mathbb{R}_+^*$ of logarithm function, one obtains:

**Proposition 3.2.3** *Under F.B.M. assumption:*
*1- if* $H \in ]0,3/4[$: $\sqrt{N}\left(\log S_N(n_i) \ - \ \log\gamma(n_i)\right)_{1\leq i\leq l} \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}_l(O;G)$,
*where* $G=(g_{ij})_{1\leq i,j\leq l}$ *is the matrix such as:*

$$g_{ij}=\frac{2d_{ij}}{(n_in_j)^{2H}}\sum_{k=-\infty}^{+\infty}\Big(\mid kd_{ij}-n_j\mid^{2H}+\mid kd_{ij}+n_i\mid^{2H}-\mid kd_{ij}+n_i-n_j\mid^{2H}-\mid kd_{ij}\mid^{2H}\Big)^2.$$

*2- if* $H\in]3/4,1[$:

$$\left(\left[\frac{N}{n_i}\right]^{2-2H}\left(\log S_N(n_i) \ - \ \log\gamma(n_i)\right)\right)_{1\leq i\leq l} \xrightarrow[N\to+\infty]{\mathcal{D}} H(2H-1)^t(\mathcal{Z}_1,..,\mathcal{Z}_l),$$

*where* $\mathcal{Z}_1,..,\mathcal{Z}_l$ *are the same as in proposition 3.2.2.*

## 3.3 Second estimation of $n_i$-length increment variances based on log-log regression estimates

Denote: $I_l={}^t(1,1,..,1)$, $A=I-\dfrac{I_l^tI_l}{{}^tI_lI_l}$, and

$$S_N=(\log S_N(n_1),..,\log S_N(n_l)),\ ES=(\log\gamma(n_1),..,\log\gamma(n_l)),\ L=(\log n_1,..,\log n_l).$$

Under F.B.M. assumption, it exists $\theta={}^t(2H,K)$, where $K=\log\gamma(1)$, such as $ES=2H.L+K.I_l=J.\theta$ (with $J=(L,I_l)$). Using proposition 3.2.3, one obtains that $S_N=J.\theta+\beta_N$, where $\beta_N$ is a remainder which is asymptotically Gaussian ($H\in]0,3/4[$) or asymptotically non-Gaussian ($H\in]3/4,1[$). By the linearity of this model, one obtains an estimation $\widehat{\theta_1}(N)$ of $\theta$ by regression of $S_N$ on $J$ and ordinary least squares (O.L.S.) (this is the aggregated variance method of $H$ estimation, exposed in Beran (1994) or in Taqqu *et al.* (1995)). $\widehat{\theta_1}(N)$ is selected to minimize:

$$\parallel S_N-J.\theta\parallel^2_{O.L.S.}={}^t(S_N-J.\theta)(S_N-J.\theta).$$

By this way we obtain an estimator $\widehat{H_1}(N)$ of $H$:

**Proposition 3.3.1** *Suppose F.B.M. assumption. The O.L.S. estimator of* $H$ *is* $\widehat{H_1}$ *such as* $\widehat{H_1}(N)=\dfrac{{}^t(AL)S_N}{2{}^t(AL)(AL)}$, *and*
*1- if* $H\in]0,3/4[$ *with* $G$ *the matrix defined in proposition 3.2.3:*

$$\sqrt{N}(\widehat{H_1}(N) \ - \ H) \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}(O;\sigma_1^2),\ \ with\ \ \sigma_1^2=\frac{1}{4}\frac{{}^t(AL)G(AL)}{({}^t(AL)(AL))^2}.$$

*2- if $H \in ]3/4, 1[$, with $(\mathcal{Z}_1, .., \mathcal{Z}_l)$ the process defined in proposition 3.2.3:*

$$N^{2-2H}(\widehat{H_1}(N) \; - \; H) \xrightarrow[N\to+\infty]{\mathcal{D}} \frac{H(2H-1)}{2} \frac{{}^t(AL)}{{}^t(AL)(AL)} {}^t(n_1^{2H-2}\mathcal{Z}_1, .., n_l^{2H-2}\mathcal{Z}_l).$$

**Proof:** We obtain the estimator $\widehat{H_1}(N)$ by applying O.L.S., and we have:

$$\widehat{H_1}(N) - H = \frac{{}^t(AL)}{2^t(AL)(AL)} \beta_N \;.\blacksquare$$

One deduces that $J.\widehat{\theta_1}(N)$ is a second estimation of $(\log \gamma(n_i))_{1\leq i\leq l}$. $J.\widehat{\theta_1}(N)$ corresponds to estimates on the regression line of $(L, S_N)$.

If $0 < H < 3/4$, we can identify the $S_N$ asymptotic covariance matrix. In fact, this matrix $G$ has the form $G = G(n_1, .., n_l, H)$ and the matrix $\widehat{G_1}(N) = G(n_1, .., n_l, \widehat{H_1}(N))$ converges in probability to $G$ (if $0 < H$, $G(n_1, .., n_l, H)$ is differentiable in $H$). So, it is possible to determine an estimation $\widehat{\theta_2}(N)$ of $\theta$ by generalized least squares (G.L.S.) of the log-log régression and by minimizing

$$\| S_N - J.\theta \|^2_{\widehat{G_1}(N)} =^t (S_N - J.\theta)\widehat{G_1}(N)^{-1}(S_N - J.\theta).$$

Denote $B = I - \dfrac{I_l^t I_l G^{-1}}{{}^tI_l G^{-1} I_l}$ and $B_N = I - \dfrac{I_l^t I_l \widehat{G_1}(N)^{-1}}{{}^tI_l \widehat{G_1}(N)^{-1} I_l}$. We obtain the following central limit theorem:

**Proposition 3.3.2** *Suppose F.B.M. assumption with $0 < H < 3/4$. The G.L.S. estimator of $H$ is $\widehat{H_2}$ such as:*

$$\widehat{H_2}(N) = \frac{{}^t(B_N L)\widehat{G_1}(N)^{-1} S_N}{2^t(B_N L)\widehat{G_1}(N)^{-1}(B_N L)}, \quad \textit{and} \tag{4}$$

$$\sqrt{N}(\widehat{H_2}(N) \; - \; H) \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}(O; \sigma_2^2), \quad \textit{with } \sigma_2^2 = \frac{1}{4^t(BL)G^{-1}(BL)} \leq \sigma_1^2. \tag{5}$$

**Proof**: $\widehat{G_1}(N)^{-1} \xrightarrow[N\to+\infty]{\mathcal{P}} G^{-1}$ because $G(n_1, .., n_l, H)$ is a positive symmetric matrix and differentiable in $H$ for $H \in ]0, 1[$. G.L.S. theory gives

$$\widehat{H_2}(N) = \frac{{}^t(B_N L)\widehat{G_1}(N)^{-1} S_N}{2^t(B_N L)\widehat{G_1}(N)^{-1}(B_N L)}$$

and $\widehat{H_2}(N)$ exists because $B_N L$ can't be null vector as $B_N$ is a positive symmetric matrix if $\widehat{H_1}(N) \in ]0, 3/4[$ (and $\widehat{H_1}(N) \xrightarrow[N\to+\infty]{\mathcal{P}} H \in ]0, 3/4[$). As ${}^tB_N \widehat{G_1}(N)^{-1} = \widehat{G_1}(N)^{-1} B_N$ and

$B_N I_l = 0_l$, we have:

$$\widehat{H_2}(N) - H = \frac{{}^t(B_N L)\widehat{G_1}(N)^{-1}}{2\,{}^t(B_N L)\widehat{G_1}(N)^{-1}(B_N L)}\beta_N.$$

Proposition 3.2.3 gives $\sqrt{N}\beta_N \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}_l(O;G)$ and so we obtain (5). We deduce $\sigma_2^2 \leq \sigma_1^2$ from the G.L.S. Gauss-Markov theorem. ■

If $H \in ]0,3/4[$, this second regression has two advantages. First, $\widehat{H_2}(N)$ is better than $\widehat{H_1}(N)$. Second, $J.\widehat{\theta_2}(N)$ is another convergent estimation of $(\log\gamma(n_1),..,\log\gamma(n_l))$ giving a simplified form of the following self-similarity test.

## 3.4 Self-similarity test

We have two increments variance estimations, for scales $(n_1,..,n_l)$. In fact, $S_N$ and $J.\widehat{\theta_2}(N)$ are $ES$ estimations. The test is based on the determination of a distance between those both estimations. We begin to compute $\widehat{H_1}(N)$ from $X_0, X_1,..,X_N$. We form two tests:

1- Consider the distance $LM(N)$ (Long-Memory) defined by:

$$LM^2(N) = N^{4-4\widehat{H_1}(N)} \parallel S_N - J.\widehat{\theta_1}(N) \parallel^2_{O.L.S.}. \tag{6}$$

**Proposition 3.4.1** *Under F.B.M. assumption, and if* $H \in ]3/4,1[$*, then the distribution of* $LM^2(N)$ *converges in law to a distribution only depending on* $(n_1,..,n_l)$ *and* $H$*.*

2- Consider the distance $SM(N)$ (Short-Memory) defined by:

$$SM^2(N) = N \parallel S_N - J.\widehat{\theta_1}(N) \parallel^2_{\widehat{G_1}(N)^{-1}} \tag{7}$$

**Proposition 3.4.2** *Under F.B.M. assumption, and if* $H \in ]0,3/4[$*, then*

$$SM^2(N) \xrightarrow[N\to+\infty]{\mathcal{D}} \chi^2(l-2). \tag{8}$$

**Remark 1**: Proposition (3.4.2) may be explained with heuristic arguments. Remainders are turned white, thus it's only natural for the sum of the second regression remainder squares to asymptotically form a $\chi^2$ process. The number of freedom degrees is $l-2$ because one loses two freedom degrees after the twice estimation of $\theta$.

**Remark 2**: The test presents a dichotomy around $H = 3/4$. The choice between $LM$ and $SM$ will be delicate if $\hat{H} > 3/4$ but $H < 3/4$ or vice versa. But, asymptotic distributions of $LM$ and $SM$ practically seem to be continuous in $H$ for $0 < H < 1$. In fact, $(\sqrt{N}\log S_N(n_1), .., \sqrt{N}\log S_N(n_l))$ has continuous asymptotic variance for $H \in ]0,1[$. Thus $LM$ (and $SM$) has approximatively the same distribution when $\hat{H}$ is a bit less than 3/4 or a bit more than 3/4. Moreover, one has theoretical and practical confidence intervals for L.S. estimators. Then, the choice of test ($SM$ or $LM$) can be confirmed or be questioned.

Distances $LM(N)$ and $SM(N)$ are estimations of a distance between process $X$ and a set of processes containing all the F.B.M. On the previous form, the test can only be used to reject the F.B.M. hypothesis. Nevertheless, it's possible to extend its use and giving more informations about the S.S. of $X_0, X_1, .., X_N$.

## 3.5 Extensions of previous results

Propositions 3.2.2 and 3.2.3 can be extended to a set of Gaussian processes. Now, we consider a process $X$ such as:

**G assumption**: *Let $X$ be a zero mean Gaussian process ($X_0 = 0$ a.s.), with stationary increments, such as $\mathbb{E}(X_{t+s} - X_s)^2 = \gamma(t)$ for any $(t,s) \in \mathbb{R}_+^2$, and such as increment covariance $\rho(t) = \mathbb{E}(X_{t+s+1} - X_{t+s})(X_{s+1} - X_s)$ verifies: $\lim_{t\to+\infty}\rho(t) = 0$.*

More, we will use one of the following assumptions:

- **G1 assumption**: *Let $X$ verify G assumption and $\rho(.) \in \ell^2$.*
- **G2 assumption**: *Let $X$ verify G assumption and $\rho(t) \sim t^{-D}L(t)$ when $t \to +\infty$, with $0 < D < 1/2$ and $L(.)$ a slow varying function.*

**Remark**: If $X$ verifies the F.B.M. assumption, then $X$ verifies G assumption. More, if $H \in ]0, 3/4[$, then $X$ verifies G1 assumption, and if $H \in ]3/4, 1[$, then $X$ verifies G2 assumption (with $D = 2H - 2$).

Corresponding to propositions 3.2.2 and 3.2.3, we have:

**Proposition 3.5.1** *1- if $X$ verifies G1 assumption, then*

$$\sqrt{N}\left(S_N(n_i) \; - \; \gamma(n_i)\right)_{1\le i\le l} \; \xrightarrow[N\to+\infty]{\mathcal{D}} \; \mathcal{N}_l(0; F),$$

*where* $F = (f_{ij})_{1 \le i,j \le l}$ *is the matrix such as, with* $d_{ij} = GCD(n_i, n_j)$*:*

$$f_{ij} = 2d_{ij} \sum_{k=-\infty}^{+\infty} \left(\gamma(kd_{ij} - n_j) + \gamma(kd_{ij} + n_i) - \gamma(kd_{ij} + n_i - n_j) - \gamma(kd_{ij})\right)^2.$$

*2- if* $X$ *verifies G2 assumption, then*

$$\frac{N^D}{L(N)} \left(S_N(n_i) \ - \ \gamma(n_i)\right)_{1 \le i \le l} \ \xrightarrow[N \to +\infty]{\mathcal{D}} \ {}^t(n_1^2 \mathcal{Z}_1, .., n_l^2 \mathcal{Z}_l)$$

*where* $\mathcal{Z}_1, .., \mathcal{Z}_l$ *are the same as in proposition 3.2.2.* $(\mathcal{Z}_1, .., \mathcal{Z}_l)$ *characteristic function verifies for* $u_1, .., u_l$ *small:*

$$f(u_1, u_2, .., u_l) = exp\left(\frac{1}{2} \sum_{k=2}^{+\infty} \frac{(2i)^k}{k} \sum_{i_1, .., i_k = 1}^{l} u_{i_1} .. u_{i_k} I(n_{i_1}, .., n_{i_k})\right) \quad \textit{and}$$

$$I(n_{i_1}, .., n_{i_k}) = \int_0^1 .. \int_0^1 \frac{\mid y_1 - y_2 \mid^{-D} .. \mid y_k - y_1 \mid^{-D}}{n_{i_1}^{1-D} .. n_{i_k}^{1-D}} dy_1 .. dy_k.$$

**Corollary 3.5.1** *Under G assumption, one considers* $(X'_0, X'_1, .., X'_N)$*, where* $X'_i = X_i + \epsilon_i^N$ *with* $(\epsilon_i^N)$ *a white noise of variance* $\sigma^2(N)$*, independent of* $(X_i)$*. Let* $(S'_N(n_1), .., S'_N(n_l))$ *be the vector computed from* $(X'_0, .., X'_N)$*. Then,* $(S'_N(n_1), .., S'_N(n_l))$ *has the same asymptotic convergence as* $(S_N(n_1), .., S_N(n_l))$ *in proposition 3.5.1, if:*
*1- under G1 assumption,* $\lim_{N \to +\infty} \sigma^2(N) = 0$.
*2- under G2 assumption,* $\sigma^2(.)$ *is bounded on* $\mathbb{N}$.

**Proposition 3.5.2** *1- if* $X$ *verifies G1 assumption, then*

$$\sqrt{N} \left(\log S_N(n_i) \ - \ \log \gamma(n_i)\right)_{1 \le i \le l} \ \xrightarrow[N \to +\infty]{\mathcal{D}} \ \mathcal{N}_l(O; G),$$

*where* $G = (g_{ij})_{1 \le i,j \le l}$ *is the matrix such as:*

$$g_{ij} = \frac{2d_{ij}}{\gamma(n_i)\gamma(n_j)} \sum_{k=-\infty}^{+\infty} \left(\gamma(kd_{ij} - n_j) + \gamma(kd_{ij} + n_i) - \gamma(kd_{ij} + n_i - n_j) - \gamma(kd_{ij})\right)^2.$$

*2- if* $X$ *verifies G2 assumption, then*

$$\frac{N^D}{L(N)} \left(\log S_N(n_i) \ - \ \log \gamma(n_i)\right)_{1 \le i \le l} \ \xrightarrow[N \to +\infty]{\mathcal{D}} \ {}^t\left(\frac{n_1^2}{\gamma(n_1)} \mathcal{Z}_1, .., \frac{n_l^2}{\gamma(n_l)} \mathcal{Z}_l\right),$$

*where* $\mathcal{Z}_1, .., \mathcal{Z}_l$ *are the same as in proposition 3.5.1.*

So, we can estimate $ES = (\log \gamma(n_1), .., \log \gamma(n_l))$ with the vector $S_N$ when $X$ verifies G1 or G2 assumption. In both cases, we can determine $\widehat{\theta_1}(N)$ from $S_N$, and $\widehat{\theta_1}(N)$ converges with a central limit or a non central limit theorem to a vector depending on $ES$. The second estimation of $ES$, $J.\widehat{\theta_1}(N)$, converges to $ES$ only if there exist two real numbers $H'$ and $K'$ (thus a vector $\theta'$), such as $ES = 2H'.L + K'.I_l = J.\theta'$. We will use this property to test a restrict self-similarity form defined by:

$SS(n_1, .., n_l)$ **property**: *A process $X$ with stationary increments $\gamma(t) = \mathbb{E}(X_{t+s} - X_s)^2$, for any $(t, s) \in \mathbb{R}_+^2$ will be said self-similar for the scales $n_1, .., n_l$ ($SS(n_1, .., n_l)$), if there are two real numbers $H$ and $\sigma^2$ such as, for any $n = n_1, .., n_l$: $\gamma(n) = \sigma^2 n^{2H}$.*

We can remark that F.B.M. are $SS(n_1, .., n_l)$ for any scales $n_1, .., n_l$. From the previous test $LM$, it is possible to form a test of $SS(n_1, .., n_l)$:

**Proposition 3.5.3** *Under G2 assumption and if $X$ is $SS(n_1, .., n_l)$, then the distribution of $LM(N)$ converges to a distribution only depending on $(n_1, .., n_l)$ and $D$. If $X$ is not $SS(n_1, .., n_l)$, then the distribution of $LM(N)$ diverges with $N^D$ law power.*

**Proof of proposition 3.5.3**: See the proof of proposition 3.5.4.

It's also possible to form a test $SM$ under G1-assumption. For this, we define a weak F.B.M. property for scales $(n_1, .., n_l)$:

$WFBM(n_1, .., n_l)$ **property**: *A process $X$ with stationary increments will be said a weak F.B.M. of index $H$, if there exists $H \in ]0, 1[$ such as, for any $(i, j) \in \{1, 2, .., l\}^2$ and with $d_{ij} = GCD(n_i, n_j)$ and $\gamma(t) = \mathbb{E}(X_{t+s} - X_s)^2$, for any $(t, s) \in \mathbb{R}_+^2$,*

$$\sum_{k=-\infty}^{+\infty} \frac{\left(\gamma(kd_{ij} - n_j) + \gamma(kd_{ij} + n_i) - \gamma(kd_{ij} + n_i - n_j) - \gamma(kd_{ij})\right)^2}{\gamma(n_i)\gamma(n_j)} =$$
$$\sum_{k=-\infty}^{+\infty} \frac{\left(| kd_{ij} - n_j |^{2H} + | kd_{ij} + n_i |^{2H} - | kd_{ij} + n_i - n_j |^{2H} - | kd_{ij} |^{2H}\right)^2}{n_i^{2H} n_j^{2H}}. \quad (9)$$

Defining $\widehat{H_2}(N)$ as in proposition 3.3.2, with matrix $\widehat{G_1}(N) = G(n_1, .., n_l, \widehat{H_1}(N))$, we have:

**Proposition 3.5.4** *Under G1 assumption and if $X$ is $SS(n_1, .., n_l)$ and $WFBM(n_1, .., n_l)$, then $SM^2(N) \xrightarrow[N \to +\infty]{\mathcal{D}} \chi^2(l - 2)$. If $X$ is not $SS(n_1, .., n_l)$ then the distribution of $SM^2(N)$ diverges with $N$.*

**Proof of proposition 3.5.4**: Under G1 assumption, and if $X$ is $SS(n_1,..,n_l)$ with index $H$, then $\widehat{H_1}(N)$ converges to $H$ and we have a central limit theorem. If $\Lambda = (\lambda_{ij})_{1\leq i,j\leq l}$ is the limit covariance matrix of the vector $S_N$, then:

$$\lambda_{ij} = \frac{2d_{ij}}{\gamma(n_i)\gamma(n_j)} \sum_{k=-\infty}^{+\infty} \left(\gamma(kd_{ij}-n_j)+\gamma(kd_{ij}+n_i)-\gamma(kd_{ij}+n_i-n_j)-\gamma(kd_{ij})\right)^2.$$

Now $G(n_1,..,n_l,\widehat{H_1}(N))$ converges in probability to $G(n_1,..,n_l,H)$, and then, with property (9), $G(n_1,..,n_l,\widehat{H_1}(N))$ converges in probability to $\Lambda$. Thus, $\widehat{H_2}(N)$ is again a G.L.S. estimator of $H$, and so $SM(N) \xrightarrow[N\to+\infty]{\mathcal{D}} \chi^2(l-2)$.

If $X$ verifies G1 assumption but if it isn't $SS(n_1,..,n_l)$, then $\widehat{\theta_1}(N)$ converges to a vector $\theta' = (2H',K')$ and there exists $\epsilon > 0$ and $i_0 \in \{1,..,l\}$ such as:
$|\log\gamma(n_{i_0}) - 2H'\log n_{i_0} - K'| \leq \epsilon$ (if $\epsilon$ doesn't exist, then $X$ is $SS(n_1,..,n_l)$). So, there exists $C > 0$ such as: $\| S_N - J\theta \|^2_{\widehat{G_1}(N)} \geq C\epsilon^2$ for $N$ large enough. Thus, $SM^2(N) \geq NC\epsilon^2$; with probability 1 and $SM^2(N)$ diverges with $N$. ■

**Remark 1**: More $l$ is great, more the test is precise and more $SS(n_1,..,n_l)$ is close to the set of all F.B.M.

**Remark 2**: The set of processes verifying $SS(n_1,..,n_l)$ and $WFBM(n_1,..,n_l)$ might be very close to the set of F.B.M. It should be legitimate to model data accepting the test of proposition 3.5.4 with a skeleton of F.B.M.

## 3.6 Results of simulations

We generate F.B.M. with arbitrary parameter $H$ with an algorithm using wavelets synthesis, following the idea of Abry and Sellan (1996). We also examine real data and test its self-similarity.

### 3.6.1 Estimators of H

First, we compute $\widehat{H_1}(N)$, from the simulated path of a F.M.B. We work with 200 to 50000 data, $H$ in $]0,1[$, $l = 5,6,..,30$, and $(n_1,..,n_l)$ proportional to $(1,2,..,l)$. As $\widehat{G_1}(N)$ (see proposition 3.2.3) is defined by a $\ell^2$-convergent series, we truncate it out of $-200, 200$ and then, we compute $\widehat{H_2}(N)$. We compare these estimations with maximum likelihood (M.L.) estimator of $H$ following Whittle's approximation (see Beran (1994) for example). We repeat this operation 50 times for fixed $H$, $N$, $l$, and $(n_1,..,n_l)$. So, we can determine empirical bias, and empirical standard deviation of each estimator.

Figure 1 exhibits two examples (we have pointed $\widehat{H_2}(N)$ out even if $H > 3/4$). Figure 2 exhibits the same results when a white noise is added to the signal.

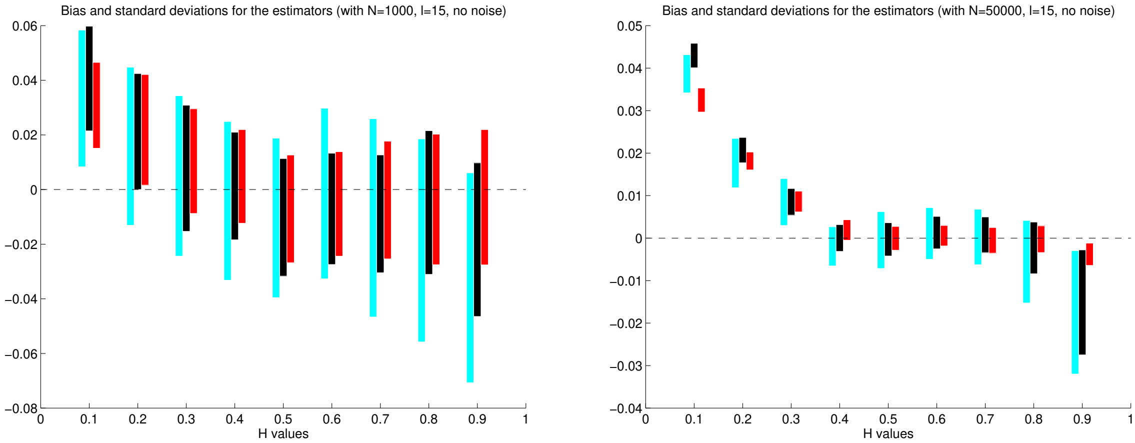



Figure 3.1 : Estimators of $H$ for F.B.M. (from left to right $\widehat{H_1}$, $\widehat{H_2}$ and $\widehat{H}_{M.L.}$) for each value of $H$

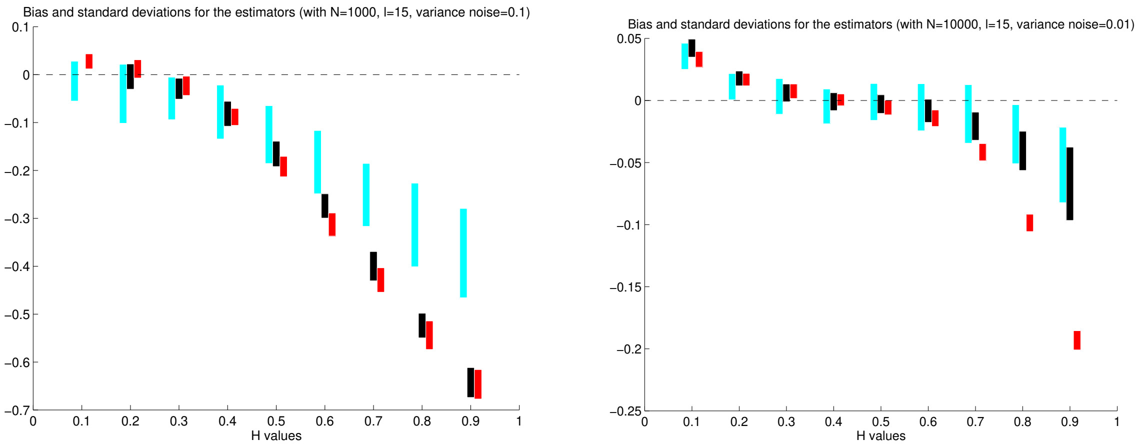



Figure 3.2 : Estimators of $H$ for noisy F.B.M. (from left to right $\widehat{H_1}$, $\widehat{H_2}$ and $\widehat{H}_{M.L.}$) for each value of $H$

It appears that the M.L. estimator gives the best estimations of $H$ (for convergence speed and precision) when the signal is a F.B.M. However, if $H < 3/4$, $\widehat{H_2}(N)$ gives nearly similar results. $\widehat{H_1}(N)$ converges in mean to $H$ as quickly as $\widehat{H_2}(N)$, but it is less precise than $\widehat{H_2}(N)$. So, we prefer to give confidence intervals of $\widehat{H_2}(N)$ to estimate $H$. But, the M.L. estimator is not robust when a Gaussian white noise is added to a F.B.M., even if its amplitude is small. $\widehat{H_1}(N)$ is the most robust of estimators and $\widehat{H_2}(N)$ is robust and precise (when $H < 3/4$). Both L.S. estimators are very advantageous for time computing also. Computing $\widehat{H_1}(N)$ requires only l-matrix multiplications, when the about $lN$ operations to compute $S_N$ have been done. Computing $\widehat{H_2}(N)$ requires the truncature of $l^2$ infinite sums to determine the covariance matrix, and, for large $N$, these calculations are negligible. So, computing $\widehat{H_2}(N)$ requires again about $lN$ operations. But computing M.L. estimator, even with Whittle approximation, requires to find an extremum of a function composed with Fourier transforms: about $kN \log N$ operations are necessary ($k$ depending on the truncature and the required precision of the extremum). For example, when $N = 50000$, computing $\widehat{H_1}(N)$ and $\widehat{H_2}(N)$ needs less than 5 seconds, but for the M.L. estimator, we have a result with precision of $10^{-6}$ after one hour. Thus, when there is a lot of data, and when immediate estimations are asked (finance..), the important rapidity of $\widehat{H_1}(N)$ and $\widehat{H_2}(N)$ calculations is a great advantage.

### 3.6.2 The tests

Before testing real data, we apply the tests to simulated F.B.M. We repeat this operation 50 times for fixed $H$, $N$, $l$, and $(n_1, .., n_l)$. Then, we reconstruct the $SM(N)$ distribution density with a kernel method. If $H < 3/4$, this function should be nearly a $\chi^2_{l-2}$ distribution density. Figure 3 exhibits two examples of the comparison of empirical and theoretical densities, the first one for a F.B.M. and the second one for a noisy F.B.M.

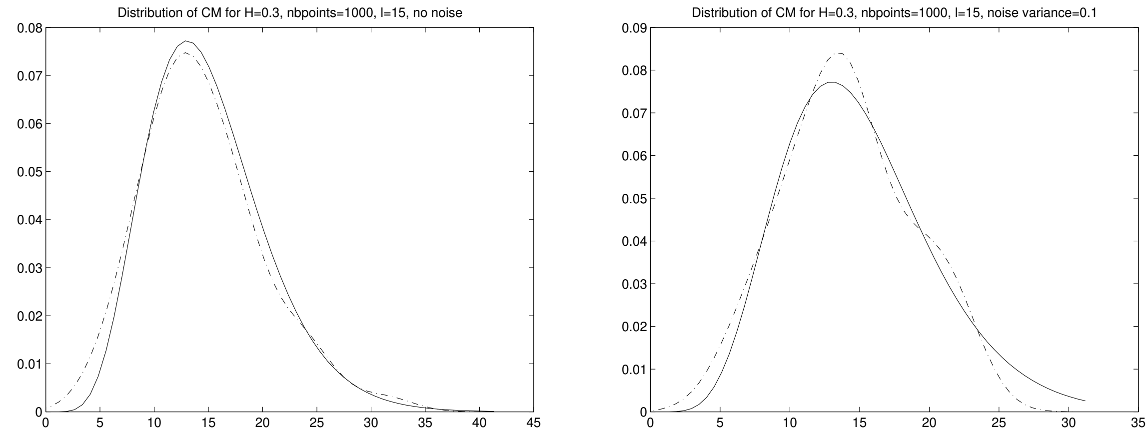



Figure 3.3 : Empirical statistics compared with the corresponding $\chi^2_{l-2}$

Empirical statistics fit well to the theoretical ones. By this way, we have a proof of the efficiency of the F.B.M. generator. Then, we have tested with 95%-forecast intervals two series of real data which seem to be self-similar. Because a great freedom of choice of $l$ and $(n_1, .., n_l)$ is left, we could select parameters presenting the most "self-similar" distance $SM(N)$ ($SM$ nearest to the confidence interval). In another words, we chose a new range of samples with the "best" self-similarity property.

First tested data are 50400 measures of load consumption computed each minute during five weeks on the French national grid system (see Poggi and Viano (1996)). Because of aggregation phenomena, this data could be considered as self-similar one. In fact, $H$ estimations are nearly the same for any scales used (we always find $H \simeq 0.3$). The regression seems to be adapted, $S_N$ and $L$ are lined up (see figure 4). But it is not the case if we compare with the expected linearity of a F.B.M with index $H = 0.3$ and for $N = 50000$ (see figure 4 also). Moreover, for any scales, $SM$ test always rejects self-similarity hypothesis. So, these data should not be a discretized F.B.M. Nevertheless, one can find the smallest $SM$ with a new sampling period of 20. For $n_i = 20i, i = 1, .., 15$, $SM \simeq 29$ and $\chi^2_{13}(95\%) \simeq 22.7$: $(X_{20k})$ is "nearly" self-similar.

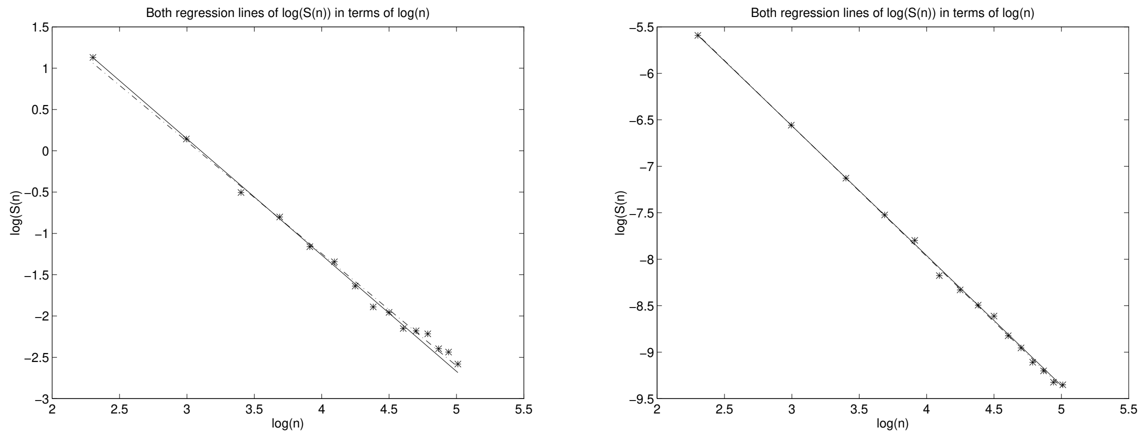



Figure 3.4 : Regression lines with both estimators, for load consumption data and for corresponding F.B.M.

The second tested series of data are the yearly minimum water levels of the Nile River between year 722 and year 1281 (see Beran (1994)). So, N=560 (we have not taken into account the first hundred referenced data), and for $n_i = i$ with $i = 1, .., 15$, we obtain $\widehat{H_1}(N) = 0.86$ and $LM(N) = 0.45$. With the F.B.M. simulator, we find empirical 95%-percentils, $q(H)$, of the asymptotic distribution of $LM(N)$, for $H \in ]3/4, 1[$ (we used 50 $LM$ computations). The function q(.) is nearly regular, strictly decreasing, and we find estimates $q(0.80) \simeq 15$, $q(0.825) \simeq 3$, $q(0.85) \simeq 1.5$, $q(0.875) \simeq 0.3$, $q(0.90) \simeq 0.15$. Thus, the hypothesis "$SS(1, 2, .., 15)$" is reasonable. It's possible that Nile River data are the skeleton of a F.B.M.

## 3.7 Annexe: proofs

**Proof of proposition 3.5.1**:
1- Under G1 assumption, we use a result of Giraitis and Surgailis (1990) and Poggi and Viano (1996). Theorem 2.2 of Poggi and Viano (1996) requires condition on increment spectral densities: they must be in $\mathbb{L}^2[-\pi,\pi[$. If $\rho \in \ell^2$, this condition is verified. So, we can use this theorem under the assumptions of proposition 3.5.1.

2- We proceed as in Taqqu (1975).
One applies proposition **6.1.** of Taqqu (1975) for each $n_j$ to $(m'_{n_j}(i))_{i\in\{1,..,[N/n_j]\}}$, where $m'_{n_j}(i)=\gamma(n_j)^{-1/2}(X_{in_j}-X_{(i-1)n_j})$. One obtains
$\mathbb{E}m'_{n_j}(i+k)m'_{n_j}(i)) \sim k^{-D}\left[\dfrac{n_j^{2-D}L(k)}{\gamma(n_j)}\right]$, and thus:

$$\frac{\gamma(n_j)}{n_j^{2-D}N^{1-D}L(N)}\sum_{i=1}^{[N/n_j]}(m'_{n_j}(i)^2-1) \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{Z}(1/n_j)$$

where $\mathcal{Z}$ is a $(1-D)$ self-similar Rosenblatt process. Consequently, it implies that for each $n_j$:

$$\frac{N^D}{L(N)}(S_N(n_j)\ -\ \gamma(n_j)) \xrightarrow[N\to+\infty]{\mathcal{D}} n_j^{3-D}\mathcal{Z}(1/n_j). \qquad (10)$$

We now turn to show the weak convergence of the vector $(S_N(n_1),..,S_N(n_l))$ and will determine its asymptotic distribution. Denote:

$$f_N(u_1,..,u_l)=\mathbb{E}exp\left(i\sum_{j=1}^{l}u_j\frac{1}{n_j^{3-D}N^{-D}L(N)}(S_N(n_j)-\gamma(n_j))\right) \text{ with } (u_1,..,u_l)\in\mathbb{R}^l.$$

We are going to prove the convergence of the function $f_N(u_1z,..,u_lz)$ when $N\to+\infty$ and when $z$ is close enough to zero. Define:
1/ $N_j=[N/n_j]$, for $j\in\{1,2,..,l\}$,
2/ $V(N)=N_1+..+N_l$,
3/ $M_{V(N)}$ the vector $(m'_{n_1}(1),..,m'_{n_1}(N_1),m'_{n_2}(1),..,m'_{n_l}(1),..,m'_{n_l}(N_l))$, and $R_{V(n)}$ it's covariance matrix,
4/ $D_{V(N)}(j)$ be the $V(N)$-matrix, null everywhere except for diagonal terms between rows $N_1+..+N_{j-1}+1$ and $N_1+..+N_j$, which are all equal to 1.

Then, it exists $\epsilon>0$, such as for $|z|<\epsilon$:

$$f_N(u_1z,..,u_lz)=exp\left(\frac{1}{2}\sum_{k=2}^{+\infty}\frac{(2iz)^k}{k}Tr\left[\sum_{j=1}^{l}v_jD_{V(N)}(j)R_{V(N)}\right]^k\right)$$

with $v_j = \dfrac{u_j \gamma(n_j)}{n_j^{2-D} N^{1-D} L(N)}$ (for details, see Taqqu (1975)). But,

$$Tr\left[\sum_{j=1}^{l} v_j D_{V(N)}(j) R_{V(N)}\right]^k = \sum_{i_1,..,i_k=1}^{l} v_{i_1}..v_{i_k} Tr\left[(D_{V(N)}(i_1)R_{V(N)})..(D_{V(N)}(i_k)R_{V(N)})\right]$$

We determine $Tr\left[(D_{V(N)}(i_1)R_{V(N)})..(D_{V(N)}(i_k)R_{V(N)})\right]$ with a decomposition in packages. In fact, let $R(i,j)$ be the submatrix of $R_{V(N)}$, take between rows $N_1+N_2+..+N_{i-1}+1$ and $N_1+N_2+..+N_i$, and between columns $N_1+N_2+..+N_{j-1}+1$ and $N_1+N_2+..+N_j$. Let $r(i,j)$ be the general term of $R_{V(N)}$. Then:

$$\begin{aligned} Tr\left[(D_{V(N)}(i_1)R_{V(N)})..(D_{V(N)}(i_k)R_{V(N)})\right] &= Tr\left[R(i_1,i_2)R(i_2,i_3)..R(i_k,i_1)\right] \\ &= \sum_{j_1=N_1+..+N_{i_1-1}+1}^{N_1+..+N_{i_1}} .. \sum_{j_k=N_1+..+N_{i_k-1}+1}^{N_1+..+N_{i_k}} r(j_1,j_2)r(j_2,j_3)..r(j_k,j_1). \end{aligned}$$

$r(j_p,j_q) = cov(m'_{n_p}(j_p - N_1+..+N_{i_p}), m'_{n_q}(j_q - N_1+..+N_{i_q}))$, and when $(ni - n'i') \to +\infty$, then:

$$r(i,i') \sim \frac{n.n'}{(\gamma(n)\gamma(n'))^{1/2}} \mid ni - n'i' \mid^{-D} L(ni - n'i').$$

So, with Riemann series, one obtains:

$$\begin{aligned} &\lim_{N\to+\infty} \left(\frac{1}{N^{1-D}L(N)}\right)^k Tr\left[(D_{V(N)}(i_1)R_{V(N)})..(D_{V(N)}(i_k)R_{V(N)})\right] \\ &= \lim_{N\to+\infty} \frac{(n_{i_1}..n_{i_k})^2}{\gamma(n_{i_1})..\gamma(n_{i_1})} \frac{1}{N^k} \sum_{j_1=1}^{N_{i_1}} .. \sum_{j_k=1}^{N_{i_k}} \left(\frac{\mid n_{i_1}j_1 - n_{i_2}j_2 \mid}{N}\right)^{-D} \frac{L(n_{i_1}j_1 - n_{i_2}j_2)}{L(N)} ... \\ &\qquad .. \left(\frac{\mid n_{i_k}j_k - n_{i_1}j_1 \mid}{N}\right)^{-D} \frac{L(n_{i_k}j_k - n_{i_1}j_1)}{L(N)} \\ &= \frac{n_{i_1}^2..n_{i_k}^2}{\gamma(n_{i_1})..\gamma(n_{i_k})} \int_0^{\frac{1}{n_{i_1}}} .. \int_0^{\frac{1}{n_{i_k}}} \mid n_{i_1}x_1 - n_{i_2}x_2 \mid^{-D} .. \mid n_{i_k}x_k - n_{i_1}x_1 \mid^{-D} dx_1..dx_k \end{aligned}$$

(with $1/n_0 = 0$), and thus $f_N(u_1 z,..,u_l z)$ has a limit when $N \to +\infty$, that is:

$$f(u_1 z,..,u_l z) = exp\left(\frac{1}{2}\sum_{k=2}^{+\infty} \frac{(2iz)^k}{k} \sum_{i_1,..,i_k=1}^{l} u_{i_1}..u_{i_k} I(n_{i_1},..,n_{i_k})\right) \quad \text{with :}$$

$$I(n_{i_1},..,n_{i_k}) = \int_0^1 .. \int_0^1 \frac{\mid y_1 - y_2 \mid^{-D} .. \mid y_k - y_1 \mid^{-D}}{n_{i_1}^{1-D}..n_{i_k}^{1-D}} dy_1..dy_k.$$

(one shows that $f$ is an analytical function as in Taqqu (1975)). $\blacksquare$

**Proof of corollary 3.5.1**: We consider the process $X'_i$, that verifies G assumption, with a variance $\gamma'(i) = \gamma(i) + 2\sigma^2(N)(1-\delta(i))$ and an increment covariance $\rho'(k) = \rho(k) + \sigma^2(N)(2\delta(k) - \delta(k+1) - \delta(k-1))$ (function $\delta(x)$ is equal to 1 if $x=0$, or else 0).

If $X$ verifies G1 assumption then $\rho(.) \in \ell^2$, thus $\rho'(.) \in \ell^2$ if $\lim_{N\to+\infty}\sigma^2(N) = 0$. One can apply proposition 3.5.1 and one obtains $\mathbb{E}S'_N(n) = (\gamma(i) + 2\sigma^2(N))$, and if $F'(N)$ is the covariance matrix of $\sqrt{N}S'_N(n_1), .., \sqrt{N}S'_N(n_l))$, we show that $\lim_{N\to+\infty} F'(N) = F$ when $\lim_{N\to+\infty}\sigma^2(N) = 0$

Under G2 assumption, $\rho(k) \sim k^{-D}L(k)$ when $k \to +\infty$, then $\rho'(k)$ has the same asymptotic behaviour that $\rho(k)$, if $\sigma(.)$ is bounded on $\mathbb{N}$. But the proof of proposition 3.5.1 only needs the asymptotic behaviour of $\rho(.)$. Then $(S'_N(n_1), .., S'_N(n_l))$ asymptotically behaves as $(S_N(n_1), .., S_N(n_l))$. ∎

Before showing propositions 3.4.1 and 3.4.2, we recall the following result:

**Lemma 3.7.1** *If $V$ is a non-null vector of $\mathbb{R}^l$ then the matrix $W(V) = (I - \frac{V\,{}^tV}{{}^tVV})$ is a symmetric matrix of $(l-1)$ rank, and it's a diagonalisable matrix with eigenvalues* 1 *and* 0. *$W(V)$ is the orthogonal projection matrix on $(\perp V)$.*

**Proof of proposition 3.4.1**: One defines $\widehat{ES_N} = J\widehat{\theta_1}(N) = \widehat{H_1}(N).L + \widehat{K_1}(N).I_l$. So, we have:

$$\widehat{ES_N} = (I - A)S_N + \frac{(AL)^t(AL)S_N}{{}^t(AL)(AL)}.$$

$$\text{Thus: } S_N - \widehat{ES_N} = \left(AS_N - \frac{(AL)^t(AL)S_N}{{}^t(AL)(AL)}\right) = A\left(I - \frac{(AL)^t(AL)}{{}^t(AL)(AL)}\right)S_N.$$

Consequently, $S_N - \widehat{ES_N} = W(I_l)W(AL)S_N = W(I_l)W(AL)(S_N - ES)$, because $ES = 2H.L + K.I_l$ and $W(I_l)W(AL)L = W(I_l)W(AL)I_l = 0$. Denote $\Pi = W(I_l)W(AL)$.

$$\begin{aligned} \text{So } LM(N) &= N^{4-4\widehat{H_1}(N)}\ {}^t(S_N - \widehat{ES_N})(S_N - \widehat{ES_N}) \\ &= {}^t(N^{2-2\widehat{H_1}(N)}\Pi(S_N - ES))(N^{2-2\widehat{H_1}(N)}\Pi(S_N - ES)) \\ &= {}^t(N^{2-2\widehat{H_1}(N)}(S_N - ES))\Pi(N^{2-2\widehat{H_1}(N)}(S_N - ES)). \qquad (11) \end{aligned}$$

Moreover, $N^{2-2\widehat{H_1}(N)} = N^{2-2H}N^{2H-2\widehat{H_1}(N)} = N^{2-2H}\exp(2(H - \widehat{H_1}(N))\log N)$. With proposition 3.3.1, one has:

$$2N^{2-2H}(H - \widehat{H_1}(N)) = H(2H-1)\frac{{}^t(AL)}{{}^t(AL)(AL)}{}^t(n_1^{2H-2}\mathcal{Z}_1, .., n_l^{2H-2}\mathcal{Z}_l).$$

Thus, $N^{2-2\widehat{H_1}(N)} = N^{2-2H}e^{Y(N)}$, and $Y(N)$ is a random variable such as:

$$Y(N) \underset{N\to+\infty}{\xrightarrow{\mathcal{P}}} 0, \quad \text{thus } e^{Y(N)} \underset{N\to+\infty}{\xrightarrow{\mathcal{P}}} 1.$$

With (11), and proposition 3.2.3, one deduces:

$$LM(N) \underset{N\to+\infty}{\xrightarrow{\mathcal{D}}} H^2(2H-1)^2 \left(n_1^{2-2H}\mathcal{Z}_1, .., n_l^{2-2H}\mathcal{Z}_l\right) \Pi^t \left(n_1^{2-2H}\mathcal{Z}_1, .., n_l^{2-2H}\mathcal{Z}_l\right).$$

$LM(N)$ converges to a random variable which only depends on $n_1, .., n_l$ and $H$. ■

**Proof of proposition 3.4.2**: One defines $\widehat{ES_N} = J.\widehat{\theta_2}(N) = 2\widehat{H_2}(N).L + \widehat{K_2}(N).I_l$. Then:

$$\widehat{ES_N} = \left(\frac{(B_N L)^t (B_N L)\widehat{G_1}(N)^{-1}}{{}^t(B_N L)\widehat{G_1}(N)^{-1}(B_N L)} + I - B_N\right) S_N = (I - C_N)S_N.$$

So: $S_N - \widehat{ES_N} = C_N(S_N - ES)$, because $C_N.ES = 0$, with

$$C_N = \widehat{G_1}(N)^{1/2} W(\widehat{G_1}(N)^{-1/2} I_l) W(\widehat{G_1}(N)^{-1/2} B_N L)\widehat{G_1}(N)^{1/2}.$$

Denote $D_N = W(\widehat{G_1}(N)^{-1/2}I_l)W(\widehat{G_1}(N)^{-1/2}B_N L)$. The rank of $D_N$ is $(l-2)$ since $\widehat{G_1}(N)^{-1/2}I_l$ and $\widehat{G_1}(N)^{-1/2}B_N L$ are linearly independent and they are in $D_N$ kernel. $D_N$ is also diagonalisable, and its eigenvalues are $(l-2)$ ones and 2 zeros; $D_N$ is the orthogonal projection matrix on $(\perp (\widehat{G_1}(N)^{-1/2}I_l, \widehat{G_1}(N)^{-1/2}B_N L))$. If we define $S'_N = \sqrt{N}(S_N - ES)$, we have:

$$SM(N) = {}^t(D_N\widehat{G_1}(N)^{-1/2}S'_N)(D_N\widehat{G_1}(N)^{-1/2}S'_N).$$

With $\widehat{G_1}(N)^{-1} \underset{N\to+\infty}{\xrightarrow{\mathcal{D}}} G^{-1}$ and $S'_N \underset{N\to+\infty}{\xrightarrow{\mathcal{D}}} \mathcal{N}_l(0, G)$, then we have:

$$\widehat{G_1}(N)^{-1/2}S'_N \underset{N\to+\infty}{\xrightarrow{\mathcal{D}}} \mathcal{N}_l(0, I).$$

In an orthogonal base constructed with eigenspaces of $D_N$, the vector $D_N\widehat{G_1}(N)^{-1/2}S'_N$ is written with $l-2$ random independent contents following a standard normal distribution, and the two last contents are zeros. So, we obtain:

$$SM(N) \underset{N\to+\infty}{\xrightarrow{\mathcal{D}}} \chi^2(l-2).■$$

**Acknowledgment.** The author expresses his gratitude to Prof. D. Dacunha-Castelle, X. Guyon and J.M. Poggi.

# Bibliographie


[1] Abry, P. and Sellan, F. (1996). The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: remarks and fast implementation. *Applied and Computational Harmonic Analysis* 3, 377-383.

[2] Azencott, R. and Dacunha-Castelle, D. (1984). *Séries d'observations irrégulières.* Masson, Paris.

[3] Beran, J (1994). *Statistics for long memory processes.* Monographs on Statist. and Appl. Probab. 61. Chapman & Hall, 315 p.

[4] Breuer, P. and Major, P. (1983). Central limit theorem for non-linear functionals of Gaussian fields. *J. Multivariate Anal.* 13, 425-441.

[5] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent Gaussian time series. *Ann. Statist.* 14, 517-532.

[6] Fox, R. and Taqqu, M.S. (1987). Central limits theorems for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields* 74, 213-240.

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

[8] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle estimate. *Probab. Theory Related Fields* 86, 87-104.

[9] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Poincaré* 33, 407-436.

[10] Mandelbrot, B.B. (1983). *The fractal geometry of nature.* Freeman, San Francisco.

[11] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 422-437.

[12] Poggi, J.M. and Viano, M.C. (1996). An estimate of the fractal index using multiscale aggregates. To appear in *J. Time Ser. Anal.*

[13] Robinson, P.M. (1995). Log-periodogram regression of time series with long-range dependence. *Ann. Statist.* 23, 1048-1072.

[14] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable non-Gaussian random processes*. Stochastic modeling. Chapman & Hall, 636 p.

[15] Taqqu, M.S. (1975). Weak Convergence to Fractional Brownian Motion and to Rosenblatt Process. *Zeit. Wahr. verw. Geb.* 31, 287-302.

[16] Taqqu, M.S., Teverovsky, V. and Willinger, W. (1995). Estimators for long-range dependence: an empirical study. *Fractals* 3, 785-798.

[17] Yajima, Y. (1989). Asymptotic properties of the LSE in a regression model with long-range stationary errors. *Ann. Statist.* 19, 158-177.

# Chapitre 4

# Un test d'autosimilarité utilisant l'ananlyse par ondelettes

A test of self-similarity using wavelet analysis

By Jean-Marc Bardet

*Université Paris-Sud and Université Evry-Val d'Essonne*

We present a simple and performing test of self-similarity for time series. The test statistic is a distance between two estimations of variances of wavelet coefficients for given scales, where these coefficients are calculated for a class of mother wavelets. Two estimators of self-similarity index are also introduced and they present good robustness and computing time properties.

## 4.1 Introduction

The set of fractional Brownian motions (f.B.m.) takes an important place among self-similar processes (see Samorodnitsky and Taqqu (1994)). f.B.m. are the only self-similar Gaussian processes with stationary increments and they present long-range dependences. Then, their study is quite possible and they provide an interesting modeling of self-similar processes (see Beran (1994) or Mandelbrot (1982)).
Self-similarity property means that any part of the path is invariant in distribution by

*Key words and phrases.* Self-similarity; Fractional Brownian Motion; Estimation of the self-similarity index; Wavelet Decomposition.

*AMS 1991 subject classifications.* Primary 62G10; secondary 60F05, 60G18, 62F10.

scaling change. It makes natural its analysis by wavelet transformation which is given from scaling modification of a mother wavelet. In concrete terms, the variance of coefficients of given f.B.m. preserve self-similarity property for all shifts. It provides some methods to generate f.B.m. (see Abry and Sellan (1996)) or to estimate scale parameter $H$ (see Flandrin (1992), Abry *et al* (1995), Abry and Veitch (1997) or Abry *et al.* (1997)) or local behaviour of $H$ (see Goncalvès and Abry (1997)).

All of this is done with the knowledge of a continuous portion of f.B.m. and with a local view. In this article, we suppose that only a time series of f.B.m., i.e. a skeleton $(X_0, X_1, .., X_N)$, is known. By this way, we can only compute approximations $d_a(i)$ of coefficients $T(a,i)$, giving by expression:

$$d_a(i) = \frac{1}{\sqrt{a}} \sum_{k=-\infty}^{+\infty} \psi(\frac{k}{a} - i) X_k \text{ instead of } T(a,i) = \frac{1}{\sqrt{a}} \int \psi(\frac{t}{a} - i) X_t dt,$$

with $\psi$ a mother wavelet. As a consequence, the using of $d_a(i)$ leads to more important technical difficulties than with $T(a,i)$, and it implies to make assumptions in the choice on chosen wavelets which have not to be made with the use of $T(a,i)$ (see Abry *et al.* (1997)).
We consider a vector $(I_{a_1}(N), .., I_{a_m}(N))$ composed with sample variance of coefficients $d_a(i)$ for scales $(a_1, a_2, .., a_m)$, computed for each $a_k$ through all possible shifts $i$. We show that when each $a_k = a_k(N)$ is a suitable sequence, this vector converges when $N \to +\infty$ to a vector $(K a_1^{2H+1}, .., K a_m^{2H+1})$ where $K > 0$, with a central limit theorem where asymptotic covariance matrix is identifiable. Then, an estimation of $(H, K)$ by a log-log regression provides ordinary and generalized least square estimators of $(H, K)$. By this way, we obtain two estimation of the vector of multiscale coefficient variance logarithms. The first one is $(\log I_{a_1}(N), .., \log I_{a_m}(N))$ and the second one is vector $((2\hat{H}+1)\log a_1 + \widehat{L}, .., (2\hat{H}+1)\log a_m + \widehat{L})$ where $\hat{H}$ and $\widehat{L}$ are given by generalized least squares (with $L = \log K$). A distance between both these estimations is the test statistic and it corresponds to a normalized sum of regression remainders squared. We show that the distribution of this distance is asymptotically a Chi-square with $(m-2)$ freedom degrees when $X$ is a f.B.m. The test is thus very simple to use.

This test is essentially the same than self-similarity test proposed in Bardet (1997). The major improvement is the avoidance of a dichotomy around $H = 3/4$. In fact, the test and the estimations of $H$ are valid for all $H \in ]0,1[$. In a sense, there is a correspondence between sample variance vectors of coefficients and periodogram, where wavelet transformation takes place of Fourier transformation. In the same way, log-periodogram technics used in Geweke and Porter-Hudak (1983) or Robinson (1995), work with small frequencies and we work with great scales. The major interest of wavelet method is the time treatment of data instead of spectral one. It allows more simple computations and greater robustness.

Simulations confirm these properties and previous limit theorems. We also remark that we can not use any wavelets when a time series is only provided. In fact, we observe unusable estimations and test when scales are not large enough or when mother wavelet is not regular enough. We have also tested empiric data. By this way, we valid the modeling of electricity consumption on the French national grid system with a f.B.m. proposed by Poggi and Viano (1996) for large scales.

## 4.2 Central limit theorems

We suppose that $(X_0, X_1, .., X_N)$ is a time series where $N \in \mathbb{N}$ and $X$ is a fractional Brownian motion with index $H$ and defined as follow: $X = \{X_t, t \in \mathbb{R}\}$ is a continuous $\mathbb{R}$-valued zero-mean ($X_0 = 0$ a.s.) Gaussian process with stationary increments such that, for $(s,t) \in \mathbb{R}^2$:

$$\mathbb{E}X_t X_s = \frac{1}{2}\sigma^2(\mid t \mid^{2H} + \mid s \mid^{2H} - \mid s-t \mid^{2H}). \tag{1}$$

One supposes that $\psi : \mathbb{R} \to \mathbb{R}$ is a function with support included in $[-1,1]$ and verifying one of the two following assumptions:

**Assumption An**: *$\psi$ is a $\mathcal{C}^{n+1}[\mathbb{R}]$ function with both first vanishing moments, i.e.*

$$\int \psi(t)dt = 0 \quad \textit{and} \quad \int t\psi(t)dt = 0 \tag{2}$$

*(the domain of integration is $]-\infty,+\infty[$ or $[-1,1]$).*

For example, we will use $\psi_1$ such that:

$$\psi_1(x) = \begin{cases} 13(1-x^2)^6 - 12(1-x^2)^5 & \text{when } |x| \leq 1 \\ 0 & \text{else.} \end{cases} \tag{3}$$

**Assumption F($n_o$)**: *$\psi$ is a measurable function such that:*

*1/ it exists $n_0 \in \mathbb{N}^*$ such that for all $m \in \mathbb{N}^*$,*

$$\sum_{k=-n_0 m}^{n_0 m} \psi(\frac{k}{n_0 m}) = 0 \quad \textit{and} \quad \sum_{k=-n_0 m}^{n_0 m} \frac{k}{n_0 m}\psi(\frac{k}{n_0 m}) = 0. \tag{4}$$

*2/ for $p$ large enough,*

$$\frac{1}{p^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\psi(\frac{k}{p})\psi(\frac{k'}{p})\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H} \;=\; \int\int \psi(t)\psi(t')|t-t'|^{2H}dtdt' \;+\; O(\frac{1}{p}).$$

Under Assumption F($n_o$), one can remark that both $\psi$ first moments are null again. $\psi$ may be an even function with a finite number of rational discontinuity points and bounded variations by intervals. For example, we will use $\psi_F$ such that:

$$\psi_F(x) = \begin{cases} -1 & \text{when } 1/3 < |x| \leq 2/3 \\ 1 & \text{when } 2/3 < |x| \leq 1 \\ 0 & \text{else.} \end{cases} \tag{5}$$

We will use the conventional notations, namely, $O(1)$ for a bounded function, $o(1)$ for a function converging to 0. Now we define coefficients $d_a(i)$ of $(X_0, X_1, .., X_N)$ where $a \in \mathbb{N}^*$ and $i \in \{1, 2, .., [N/a] - 1\}$ by:

$$d_a(i) = \frac{1}{\sqrt{a}} \sum_{k=-\infty}^{+\infty} \psi(\frac{k}{a} - i) X_k = \frac{1}{\sqrt{a}} \sum_{k=-a}^{+a} \psi(\frac{k}{a}) X_{k+ai}. \tag{6}$$

For each $(a, i)$, $d_a(i)$ is a zero-mean Gaussian variable and for large enough $a$, its variance is a self-similar deterministic function in $a$, independent of the shift $i$:

**Proposition 4.2.1** *Hyp1 - if $\psi$ verifies Assumption A0 and $i \in \{1, 2, .., [N/a] - 1\}$ where $a$ is a large enough integer, or*
*Hyp2 - if it exists $n_0$ such that $\psi$ verifies Assumption F($n_0$) and $i \in \{1, 2, .., [N/a] - 1\}$ where $a$ is a large enough $n_0$ multiple integer,*

$$\mathbb{E} d_a^2(i) = a^{2H+1}(C_\psi(H) + O(\frac{1}{a})) \text{ where } C_\psi(H) = -\frac{\sigma^2}{2} \int\int \psi(t)\psi(t')|t - t'|^{2H} dt dt'. \tag{7}$$

**Remark 1**: $C_\psi(H)$ can be written with wavelet transform of the wavelet itself $A_\psi(\alpha, \tau)$, so-called the reproducing kernel (see Flandrin (1992)), where: $A_\psi(\alpha, \tau) = \sqrt{\alpha} \int \psi(t)\psi(\alpha t - \tau) dt$ and we have:

$$C_\psi(H) = \int A_\psi(1, u)|u|^{2H} du.$$

Proposition 4.2.1 leads us to define $\tilde{d}_a(i)$ for $a \in \mathbb{N}^*$ and for $i \in \{1, 2, .., [N/a] - 1\}$, such that

$$\tilde{d}_a(i) = \frac{d_a(i)}{a^{H+1/2}\sqrt{C_\psi(H)}} \tag{8}$$

and then $\lim_{a \to 0} \mathbb{E} \tilde{d}_a^2(i) = 1$. Now we consider random variable $I_N(a)$ and $\tilde{I}_N(a)$ where

$$I_N(a) = \frac{1}{[N/a] - 1} \sum_{k=1}^{[N/a]-1} d_a^2(k) \text{ and } \tilde{I}_N(a) = \frac{1}{[N/a] - 1} \sum_{k=1}^{[N/a]-1} \tilde{d}_a^2(k) \tag{9}$$

(we will assume that $[N/a] \geq 2$). $I_N(a)$ and $\tilde{I}_N(a)$ are respectively empirical variances of $d_a(i)$ and $\tilde{d}_a(i)$ through all possible shifts $i$. Using (7), one obtains

$$\mathbb{E}\tilde{I}_N(a) = \mathbb{E}\tilde{d}_a^2(i) = 1 + O(\frac{1}{a}), \tag{10}$$

under assumptions of Proposition 4.2.1. Using properties of $\psi$ and particularly condition $\int t\psi(t)dt = 0$, we can show that $\lim_{|i-j|\to+\infty} |\mathrm{cov}(\tilde{d}_a(i), \tilde{d}_a(j))| = 0$ and limit theorems for $\tilde{I}_N(a)$. Before this, we have to specify the behaviour of $a = a(N)$, a sequence verifying the following condition:

**Conditions C**: $(u(k))_{k\in\mathbb{N}} \in (\mathbb{N}^*)^{\mathbb{N}}$ is such that:

$$[\frac{k}{u(k)}] \geq 2, \quad \lim_{k\to+\infty} u(k) = +\infty, \text{ and } \lim_{k\to+\infty} \frac{k}{u^3(k)} = 0.$$

**Proposition 4.2.2** *Hyp1 - if $\psi$ verifies Assumption A3 and if $a(k)$ and $b(k)$ are two sequences verifying Conditions C, or*
*Hyp2 - if it exists $n_0$ such that $\psi$ verifies Assumption F($n_0$) and if $a(k)$ and $b(k)$ are two sequences of $n_0$ multiple integers verifying Conditions C, then*

*1/ There exists $K > 0$ such that for large enough $N$ and for $(i, j) \in \{1, 2, .., [N/a(N)] - 1\}^2$ with $|i - j| > 2$,*

$$|cov(\tilde{d}_{a(N)}(i), \tilde{d}_{a(N)}(j))| \leq K(a(N)^{-4} + |i - j|^{2H-4}). \tag{11}$$

*2/ There exists $K > 0$ such that for $(i, j) \in \{1, 2, .., [N/a(N)]-1\}\times\{1, 2, .., [N/b(N)]-1\}$ with $|a(N)i - b(N)j| > a(N) + b(N)$ and for large enough $N$,*

$$|cov(\tilde{d}_{a(N)}(i), \tilde{d}_{b(N)}(j))| \leq K\left(a(N)^H b(N)^{-4-H} + b(N)^H a(N)^{-4-H} + \frac{a(N)^{2-H}b(N)^{2-H}}{|a(N)i - b(N)j|^{2H-4}}\right). \tag{12}$$

Using (11), it is possible to give central limit theorem for $\tilde{I}_N(a(N))$ when $a(k)$ is a sequence of integers verifying Conditions C. More generally, if $a_1, a_2, .., a_m$ are $m$ sequences of integers verifying Conditions C, we obtain the following central limit theorem for vector $(\tilde{I}_N(a_i(N)))_{i\in\{1,2,..,m\}}$:

**Proposition 4.2.3** *Let $(b(k))_{k\in\mathbb{N}}$ be a sequence verifying conditions C and $(n_1, n_2, .., n_m)$ be $m$ given positive distinct integers. Denote $a_i(k) = n_i b(k)$ for $i \in \{1, 2, .., m\}$ and $k \in \mathbb{N}$.*
*Hyp1 - if $\psi$ verifies Assumption A3, or*

*Hyp2 - if it exists $n_0$ such that $\psi$ verifies Assumption F($n_0$) and if $b(k)$ is a sequence of $n_0$ multiple integers, then*

$$\sqrt{\frac{N}{b(N)}}\left(\tilde{I}_N(a_i(N)) \ -1\right)_{i\in\{1,2,..,m\}} \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}_m(0;F) \tag{13}$$

*where $F=(f_{ij})_{1\leq i,j\leq m}$ is the matrix such that with $D_{ij}=GCD(n_i,n_j)$,*

$$f_{ij}=\frac{\sigma^4 D_{ij}}{2C_\psi^2(H)n_i^{2H}n_j^{2H}}\sum_{k=-\infty}^{+\infty}\left(\int\int \psi(t)\psi(t')\,|kD_{ij}+n_it-n_jt'|^{2H}\,dtdt'\right)^2. \tag{14}$$

**Remark 2**: If $\psi$ have just verifyed the admissibility condition (i.e. $\int\psi(t)dt=0$) then Proposition 4.2.3 would only be valid for $0<H<3/4$. For instance, it is what happens if we use Haar wavelet $\psi_H$ such that

$$\psi_H(x)=\begin{cases}1 & \text{when } 0\leq x<1/2\\ -1 & \text{when } 1/2\leq x<1\\ 0 & \text{else.}\end{cases} \tag{15}$$

In the same way, if $\psi$ verifies $\int t^k\psi(t)dt=0$ for all $k\leq k_0$ with $k_0\geq 2$ then $|\mathrm{cov}(\tilde{d}_a(i),\tilde{d}_a(j))|$ converges faster to 0 when $|i-j|\to+\infty$ than in Proposition 4.2.2 and thus the vector $(\tilde{I}_N(a_1(N)),..,\tilde{I}_N(a_m(N)))$ is "closer" to a vector composed with independent variables.

**Remark 3**: It is possible to use a larger class of wavelets if we restrict conditions for sequence $(b(k))$ and thus sequences $(a_i(k))$. In fact, from the Proposition 4.2.2 proof it follows that central limit theorem (13) is again valid if sequence $(b(k))$ verifies Conditions C and $\lim\limits_{N\to+\infty}\frac{N}{b^2(N)}$ and if $\psi$ verifies Assumption A1.

Proposition 4.2.3 can be extended if $X$ is disturbed by a white noise:

**Corollary 4.2.1** *Let us take the same assumptions as in Proposition 4.2.3. Consider $(X'_0,..,X'_N)$, where $X'_i=X_i+\varepsilon_i$ with $(\varepsilon_i)$ a white noise of variance $\gamma^2(i)$, independent of $(X_i)$. Let $(\tilde{I}'_N(a_1),..,\tilde{I}'_N(a_m))$ be the vector computed from $(X'_0,..,X'_N)$. Then, $(\tilde{I}'_N(a_1),..,\tilde{I}'_N(a_m))$ has the same asymptotic convergence as $(\tilde{I}_N(a_1),..,\tilde{I}_N(a_m))$ in proposition 4.2.3 if $\gamma^2(.)$ is bounded on $\mathbb{N}$.*

Using Proposition 4.2.3 and regularity of logarithm function on $]0,+\infty[$, we finally obtain the following central limit theorem:

**Proposition 4.2.4** *Under the same assumptions as in Proposition 4.2.3, with same matrix $F$,*

$$\sqrt{\frac{N}{b(N)}}\left(\log I_N(a_i(N)) \ -(2H+1)\log a_i(N)-\log C_\psi(H)\right)_{i\in\{1,2,..,m\}} \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}_m(0;F).$$

**Remark 4**: For ease of writing, we have chosen a support in $[-1, 1]$ for $\psi$. But all propositions of this article are still true if we work with any compactly supported wavelets verifying Assumptions A3 or F($n_0$). For instance, $\psi$ can be Daubechies-$N$ wavelets, with $N \geq 20$. Another choice can be infinite width supported wavelets, but with compact effective support. It is the case with Meyer, Morlet or Mexican Hat wavelets, which all are infinite derivable (for details, see Daubechies (1992)).

## 4.3 Estimation of $H$ and test of self-similarity

Proposition 4.2.4 provides a method to estimate $H$ by regression. In fact, if we denote:

$I_m =^t (1, 1, .., 1)$, $Y_N = (\log I_N(a_i(N)))_{1\leq i\leq m}$, $L = (\log n_i)_{1\leq i\leq m}$, $L_N = (\log a_i(N))_{1\leq i\leq m}$,

and if $X$ is a f.B.m., under assumption of Proposition 4.2.3, it exists $\theta = {}^t(2H + 1, K)$, where $K = \log C_\psi(H)$, such as $Y_N = J_N.\theta + \beta_N$, with $J_N = (L_N, I_m))$ and where $\beta_N$ is a remainder which is asymptotically Gaussian. By the linearity of this model, one obtains an estimation $\widehat{\theta_1}(N)$ of $\theta$ by regression of $Y_N$ on $J_N$ and ordinary least squares (O.L.S.). $\widehat{\theta_1}(N)$ is selected to minimize:

$$\parallel Y_N - J_N.\theta \parallel^2_{O.L.S.} = {}^t(Y_N - J_N.\theta)(Y_N - J_N.\theta).$$

But we can also identify the $\beta_N$ asymptotic covariance matrix. In fact, this matrix $F$ has the form $F = F_\psi(n_1, .., n_m, H)$ (it does not depend on $\sigma^2$) and the matrix $\widehat{F}(N) = F_\psi(n_1, .., n_m, \widehat{H_1}(N))$ converges in probability to $F$ (if $0 < H$, $F_\psi(n_1, .., n_l, H)$ is differentiable in $H$). So, it is possible to determine an estimation $\widehat{\theta_2}(N)$ of $\theta$ by generalized least squares (G.L.S.) and by minimizing

$$\parallel Y_N - J_N.\theta \parallel^2_{\widehat{F}(N)^{-1}} =^t (Y_N - J_N.\theta)\widehat{F}(N)^{-1}(Y_N - J_N.\theta).$$

Denote $A = I - \dfrac{I_m^t I_m}{{}^t I_m I_m}$, $B = I - \dfrac{I_m^t I_m F^{-1}}{{}^t I_m G^{-1} I_m}$ and $B_N = I - \dfrac{I_m^t I_m \widehat{F}(N)^{-1}}{{}^t I_m \widehat{F}(N)^{-1} I_m}$. Then:

**Proposition 4.3.1** *Suppose assumptions of Proposition 4.2.3.*
*1/ The O.L.S. estimator of $H$ is $\widehat{H_1}$ such as $\widehat{H_1}(N) = -\dfrac{1}{2} + \dfrac{{}^t(AL)Y_N}{2^t(AL)(AL)}$ and*

$$\sqrt{\frac{N}{b(N)}}(\widehat{H_1}(N) - H) \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}(O; \sigma_1^2), \text{ with } \sigma_1^2 = \frac{1}{4}\frac{{}^t(AL)F(AL)}{({}^t(AL)(AL))^2}. \tag{16}$$

*2/ The G.L.S. estimator of $H$ is $\widehat{H_2}$ such as* $\widehat{H_2}(N) = -\dfrac{1}{2} + \dfrac{{}^t(B_N L)\widehat{F}(N)^{-1}Y_N}{2^t(B_N L)\widehat{F}(N)^{-1}(B_N L)}$ *and*

$$\sqrt{\frac{N}{b(N)}}(\widehat{H_2}(N) \; - \; H) \xrightarrow[N\to+\infty]{\mathcal{D}} \mathcal{N}(O;\sigma_2^2), \ \textit{with } \sigma_2^2 = \frac{1}{4^t(BL)F^{-1}(BL)} \leq \sigma_1^2. \tag{17}$$

**Proof of Proposition 4.3.1**: We refer to Proposition of Bardet (1997) except in expressions of $\widehat{H_1}(N)$ and$\widehat{H_2}(N)$, where $L$ takes place of $L_N$ because $AL_N = AL$ and $BL_N = BL$. ■

$Y_N$ and $J_N.\widehat{\theta_2}(N)$ are two estimations of $J_N.\theta$. The computation of a distance between those both estimations provides the statistic of a self-similarity test. Consider the distance $T(N)$ defined by:

$$T^2(N) = \frac{N}{b(N)} \parallel Y_N - J_N.\widehat{\theta_2}(N) \parallel^2_{\widehat{F}(N)^{-1}} . \tag{18}$$

**Proposition 4.3.2** *Under assumptions of Proposition 4.2.3, then*

$$T^2(N) \xrightarrow[N\to+\infty]{\mathcal{D}} \chi^2(m-2). \tag{19}$$

**Remark 5**: Proposition (4.3.2) may be explained with heuristic arguments. Remainders are turned white, thus it's only natural for the sum of G.LS. remainder squares to asymptotically form a $\chi^2$ process. The number of freedom degrees is $l-2$ because one loses two freedom degrees after the twice estimation of $\theta$.

**Remark 6**: Distance $T(N)$ is an estimation of a distance between process $X$ and a set of processes containing all the f.B.m. By this way, the test can be used to reject with $N/b(N)$ convergence rate all processes which are not self-similar.

## 4.4 Simulations

We first test data which are generated by the algorithm giving in Abry and Sellan (1996). By this way, we have to our service fractional Brownian motions with arbitrary $H$ and length $N$. In a first step, we compute $\widehat{H_1}(N)$ from calculates of coefficients. We choose $b(N) \sim N^{1/3}$ and $a_i(N) = ib(N)$ for $i = 1, 2, .., m$ with $m$ depending on $N$, $m \in \{5, 6, .., 20\}$. In a second step, we compute $\widehat{F}(N)$ from its explicit expression giving in (14), with $\widehat{H_1}(N)$ instead of $H$. We repeat these two steps for 50 different f.B.m., when $H$ and $N$ are given. For example, with mother wavelet $\psi_1$ giving in (3), $N = 10000$ and $m = 10$, we obtain results showed in Figure 1:

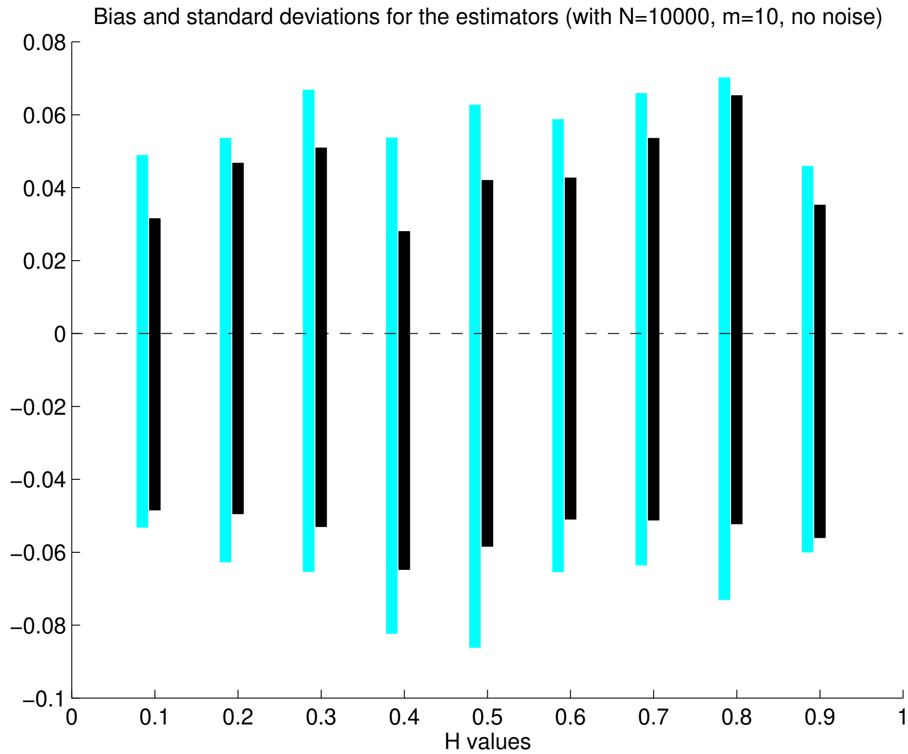



Figure 4.1 : Estimators of $H$ for f.B.m. (from left to right $\widehat{H_1}$ and $\widehat{H_2}$) for each value of $H$

With all using wavelets verifying Assumptions A3 or $\mathrm{F}(n_0)$, we observe the same distribution of $H$ estimators than in Figure 1. We note that convergence rates of $\widehat{H_1}(N)$ and $\widehat{H_2}(N)$ are not so good that maximum likelihood estimator (M.L.E.) one. Wavelet estimators need almost $N = 1000$ to provide interesting results. But $\widehat{H_1}(N)$ and $\widehat{H_2}(N)$ are very robust (see below) and very fast to be computed. Moreover, for all $H$, $\widehat{H_1}(N)$ and $\widehat{H_2}(N)$ seem to have no bias and their standard deviations are nearly similar (it is not the case with M.L.E.).
We notice same phenomena with usual wavelets Morlet, Meyer or Mexican Hat, which have both first vanishing moments, good regularity but only a compact effective support. When we use Daubechies wavelets, it appears that regularity conditions are necessary to obtain efficient estimations. For example, we remark that estimators are unusable if we use a 2-Daubechies wavelet (both first vanishing moments but it has no derivatives) when $H \geq 0.3$. These problems disappear when we use a 10-Daubechies wavelet which

verifies Assumption A1 (see Figure 2).

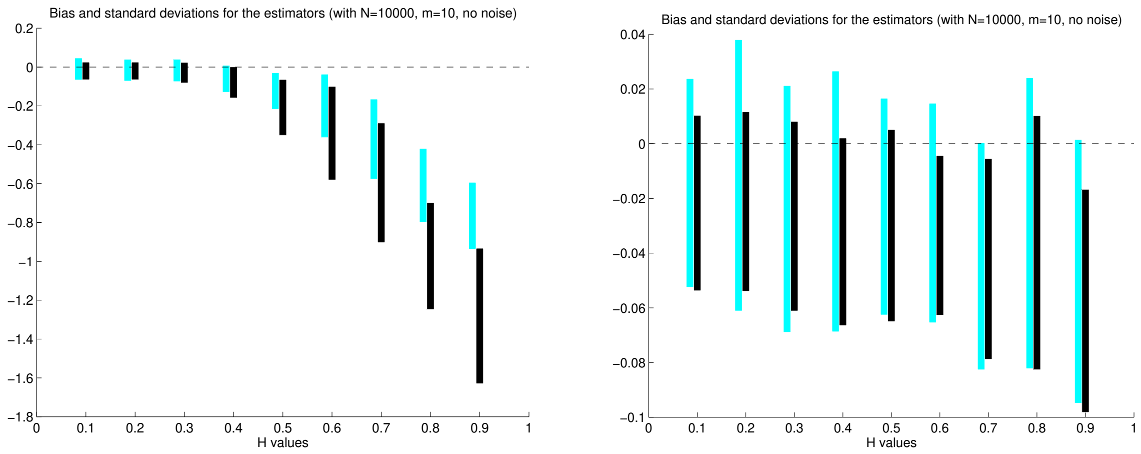



Figure 4.2 : Estimators of $H$ for f.B.m. (from left to right $\widehat{H_1}$ and $\widehat{H_2}$) for each value of $H$, with db2 and db10 wavelets

Then, it is possible to apply the test to simulated f.B.m. With the same procedure as above, we obtain 50 values of $T^2(N)$ for each $H$ and each wavelet. From Proposition 4.3.2, we know that $T^2(N)$ have nearly the same distribution as $\chi^2_{m-2}$, when $N$ is large enough. We estimate the density of $T^2(N)$ with a kernel density estimator. For example, for $N = 10000$ and $m = 10$, and with mother wavelets $\psi_1$ and $\psi_F$, we compare $T^2(N)$ distribution and $\chi^2_8$ distribution on the Figure 3 (for $\psi_F$, we choose $b(N) = 60$ because $\psi_F$ verifies Assumption $F(6)$):

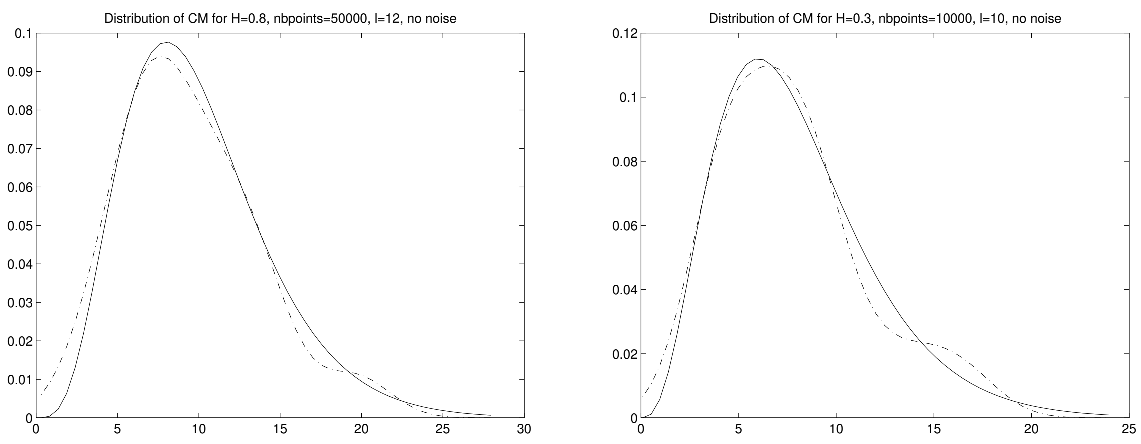
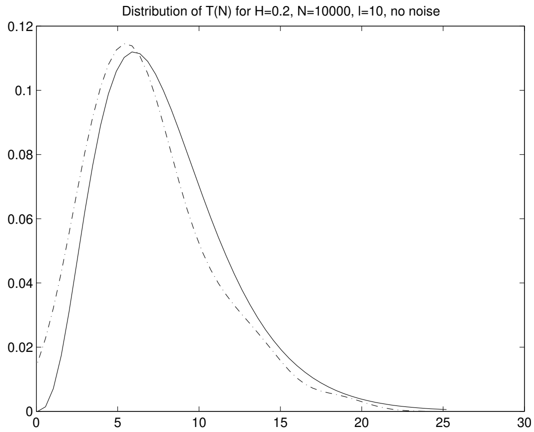



Figure 4.3 : Empirical statistics for $\psi_1$ and $\psi_F$ compared with the corresponding $\chi^2_8$

More generally, we obtain expected distributions with all wavelets verifying Assumptions $A3$ or $F(n_0)$ that we test. We also try usual Morlet and Mexican Hat wavelets.

These wavelets verify Assumption A3, except for their support which is only effectively compact. We also obtain expected results, even if we add a Gaussian white noise to the simulated f.B.m. (see an example with Mexican Hat wavelet and $N = 10000$ on Figure 4). We can note that Mexican Hat wavelet provides the best results among wavelets we use.



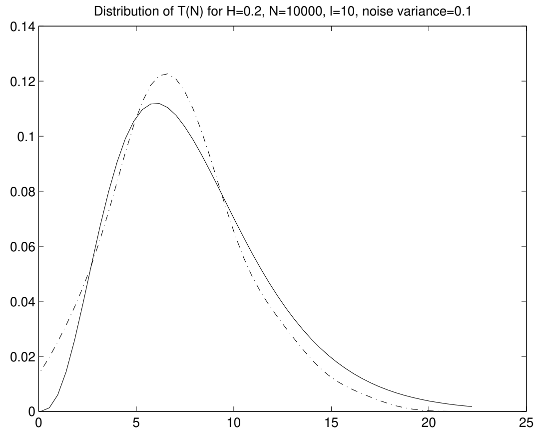




Figure 4.4 : Empirical statistics for $\psi_m$ compared with the corresponding $\chi^2_8$ for a f.B.m. and a noisy f.B.m. and with Mexican Hat wavelet

In a second time we test empiric data already tested in Bardet (1997), which are 50400 measures of load consumption computed each minute during five weeks on the French national grid system. In Bardet (1997), we have rejected self-similarity hypothesis for small scales. But with wavelet test, we test data with large scales: with $b(N) = 100$ and $m = 12$, we obtain $T^2(N) \simeq 11.37$ with Mexican Hat wavelet (and $\chi^2_{10}(0.05) \simeq 3.940$, $\chi^2_{10}(0.95) \simeq 18.307$). As a conclusion, we can say that it is possible to modelize $(X_{100k})$ with a self-similar process.

## 4.5 Proofs

We will use the two following lemmas:

**Lemma 4.5.1** *If $\psi$ verifies Assumption An then for large enough $p$*

$$\sum_{k=-\infty}^{+\infty} \psi(\frac{k}{p}) = O(\frac{1}{p^n}). \tag{20}$$

**Proof**: $\left|\dfrac{1}{p}\sum\limits_{k=-\infty}^{+\infty}\psi(\dfrac{k}{p})-\int\psi(t)dt\right| = \left|\sum\limits_{k=-p}^{+p}\int_{\frac{k}{p}}^{\frac{k+1}{p}}\left(\psi(\dfrac{k}{p})-\psi(t))dt\right)\right|$
$\leq \left|\dfrac{1}{2p^2}\sum\limits_{k=-\infty}^{+\infty}\psi'(\dfrac{k}{p})\right|+\left|\dfrac{1}{6p^3}\sum\limits_{k=-\infty}^{+\infty}\psi''(\dfrac{k}{p})\right|+..+\left|\dfrac{1}{(n+1)!p^{n+1}}\sum\limits_{k=-\infty}^{+\infty}\psi^{(n)}(\dfrac{k}{p})\right|+\dfrac{2Mp}{(n+2)!p^{n+2}}$,
where $M=\sup|\psi^{(n+1)}|$. Like $\psi$ verifies Assumption An, $\psi^{(i)}(1)=\psi^{(i)}(-1)=0$ for $i\in\{0,1,..,n+1\}$. Then $\int\psi^{(i)}(t)dt=0$ for $i\in\{0,1,..,n+1\}$ and the iteration of the decomposition for $\dfrac{1}{p}\sum\limits_{k=-\infty}^{+\infty}\psi^{(i)}(\dfrac{k}{p})$ provides (20). ■

**Lemma 4.5.2** *If $\psi$ verifies Assumption A0 then for large enough $p$*

$$\frac{1}{p^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\psi(\frac{k}{p})\psi(\frac{k'}{p})\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H} = \int\int\psi(t)\psi(t')|t-t'|^{2H}dtdt' + O(\frac{1}{p}). \quad (21)$$

**Proof**: If we denote $M_0=\sup|\psi|$ and $M_1=\sup|\psi'|$, then:

$$\begin{aligned}
&\left|\frac{1}{p^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\psi(\frac{k}{p})\psi(\frac{k'}{p})\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H}-\int\int\psi(t)\psi(t')|t-t'|^{2H}dtdt'\right|\\
&\leq\left|\sum_{k=-ps}^{+p}\sum_{k'=-p}^{+p}\int_{\frac{k}{p}}^{\frac{k+1}{p}}\int_{\frac{k'}{p}}^{\frac{k'+1}{p}}\psi(t)\psi(t')\left(\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H}-|t-t'|^{2H}\right)dtdt'\right|\\
&\qquad+\left|\sum_{k=-p}^{+p}\sum_{k'=-p}^{+p}\int_{\frac{k}{p}}^{\frac{k+1}{p}}\int_{\frac{k'}{p}}^{\frac{k'+1}{p}}\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H}\left(\psi(\frac{k}{p})\psi(\frac{k'}{p})-\psi(t)\psi(t')\right)dtdt'\right|\\
&\leq\sum_{k=-p}^{+p}\sum_{k'=-p}^{+p}\int_{\frac{k}{p}}^{\frac{k+1}{p}}\int_{\frac{k'}{p}}^{\frac{k'+1}{p}}\frac{4HM_0^2}{p^{2H}}\left|k-k'\right|^{2H-1}dtdt'\\
&\qquad+\sum_{k=-p}^{+p}\sum_{k'=-p}^{+p}\int_{\frac{k}{p}}^{\frac{k+1}{p}}\int_{\frac{k'}{p}}^{\frac{k'+1}{p}}\left(2M_0M_1(t-\frac{k}{p})+M_1^2(t-\frac{k}{p})(t'-\frac{k'}{p})\right)\left|\frac{k}{p}-\frac{k'}{p}\right|^{2H}dtdt'\\
&\leq\frac{8HM_0^2}{p^2}\sum_{k=-p}^{+p}|\frac{k}{p}|^{2H-1}+\left(\frac{2M_0M_1}{p^2}+\frac{M_1^2}{p^3}\right)\sum_{k=-p}^{+p}|\frac{k}{p}|^{2H}\\
&\leq\frac{C}{p}\ \text{ with } C>0.\ ■
\end{aligned}$$

**Proof of Proposition 4.2.1**: Using (6) and (1), one obtains:

$$\mathbb{E}d_a^2(i) = \frac{\sigma^2}{2a}\sum_{k=-a}^{+a}\sum_{k'=-a}^{+a}\psi(\frac{k}{a})\psi(\frac{k'}{a})(|k+ai|^{2H}+|k'+ai|^{2H}-|k-k'|^{2H})$$

$$= \frac{\sigma^2 a^{2H}}{2a}\left(2i^{2H}\sum_{k=-a}^{+a}\psi(\frac{k}{a})\sum_{k'=-a}^{+a}\psi(\frac{k'}{a})\left(1+2H\frac{k'}{ai}+H(2H-1)\frac{k'^2}{a^2i^2}\theta(a,i,k)^{2H-2}\right)\right.$$
$$\left. - \sum_{k=-a}^{+a}\sum_{k'=-a}^{+a}\psi(\frac{k}{a})\psi(\frac{k'}{a})\left|\frac{k}{a}-\frac{k'}{a}\right|^{2H}\right) \quad \text{with } \theta(a,i,k)\in[1,2].$$

In case of Hyp1 and with Lemmas 4.5.1 and 4.5.2, we have

$$\mathbb{E}d_a^2(i) = \frac{\sigma^2 a^{2H-1}}{2}\left[0(a^{1-n}i^{2H-2}) - a^2\left(\int\int \psi(t)\psi(t')|t-t'|^{2H}dtdt' \ + \ O(\frac{1}{a})\right)\right],$$

and then (7) is proved. In case of Hyp2, we have directly (7). ■

**Proof of Proposition 4.2.2**: We prove (12) which implies (11). If $\psi$ verifies Assumption An, for large enough $a$ and $b$, and for $(i,j)$ such that $|ai-bj|>a+b$,

$$|\mathrm{cov}(\tilde{d}_a(i),\tilde{d}_b(j))| =$$
$$= \frac{\sigma^2}{2a^{H+1}b^{H+1}C_\psi(H)}\left|\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})(|k+ai|^{2H}+|k'+bj|^{2H}-|k-k'+ai-bj|^{2H})\right|$$
$$= \left|O(a^{-n}b^{-1}i^{2H-2}+b^{-n}a^{-1}j^{2H-2}) - \frac{\sigma^2|ai-bj|^{2H}}{2a^{H+1}b^{H+1}C_\psi(H)}\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})\left(1+\frac{k-k'}{ai-bj}\right)^{2H}\right|$$
$$\leq J_1+J_2+J_3 \quad \text{where:}$$

$$J_1 \leq K_1(a^{-n}b^{-1}+b^{-n}a^{-1}), \tag{22}$$

$$J_3 = \frac{\sigma^2|ai-bj|^{2H-5}}{2a^{H+1}b^{H+1}C_\psi(H)}\left|\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\frac{2H..(2H-4)}{5!}\theta^{2H-5}(k-k')^5\right| \text{ and so:}$$

$$J_3 \leq K_3|ai-bj|^{2H-5}\frac{(a^6+b^6)}{a^{H+1}b^{H+1}}, \tag{23}$$

with $K_1>0$ and $K_3>0$, and

$$J_2 = \frac{\sigma^2|ai-bj|^{2H}}{2a^{H+1}b^{H+1}C_\psi(H)}\left|\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})\left(1+\sum_{q=1}^{4}\frac{2H..(2H+1-q)}{q!}\left|\frac{k-k'}{ai-bj}\right|^q\right)\right|.$$

Denote $S_p(a)=\sum_{k=-a}^{+a}\psi(\frac{k}{a})\left(\frac{k}{a}\right)^p$. $\psi(t)$ and $t\psi(t)$ verifies Lemma 4.5.1, so $S_0(a)=O(a^{-n})$, $S_1(a)=O(a^{-n})$ and $S_q(a)=O(a)$ for $q\geq 2$. Then:

$$\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b}) = S_0(a)S_0(b) = O(a^{-n}b^{-n}),$$
$$\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})(k-k') = aS_1(a)S_0(b) - S_0(a)bS_1(b) = O(a^{-n}b^{-n}(a+b)),$$
$$\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})(k-k')^2 = a^2S_2(a)S_0(b) - 2abS_1(a)S_1(b) + S_0(a)b^2S_2(b)$$
$$= O(a^3b^{-n} + a^{-n}b^3 + a^{1-n}b^{1-n}),$$
$$\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})(k-k')^3 = a^3S_3(a)S_0(b) + 3a^2bS_2(a)S_1(b) + 3ab^2S_1(a)S_2(b) +$$
$$b^3S_0(a)S_3(b) = O((a^3b^{-n} + a^{-n}b^3)(a+b)),$$
$$\sum_{k=-a}^{+a}\sum_{k'=-b}^{+b}\psi(\frac{k}{a})\psi(\frac{k'}{b})(k-k')^4 = a^4S_4(a)S_0(b) + 4a^3bS_3(a)S_1(b) + 6a^2b^2S_2(a)S_2(b)$$
$$+4ab^3S_1(a)S_3(b) + b^4S_0(a)S_4(b) = O((a^4b^{-n} + a^{-n}b^4)(a+b) + a^3b^3).$$

Now, suppose that $\psi$ verifies Assumption A3. After computations, with $|ai-bj| \leq N$ and $a$, $b$, verifying Condition C, we show that the dominant term of $J_2$ is given by $O(a^3b^3)$. Then, it exists $K_2 > 0$ such that:

$$J_2 \leq K_2|ai-bj|^{2H-4}a^{2-H}b^{2-H}. \tag{24}$$

Finally, we have:

$$|\mathrm{cov}(\tilde{d}_a(i), \tilde{d}_b(j))| \leq K_1(a^Hb^{-1-n-H} + b^Ha^{-1-n-H}) + K_3|ai-bj|^{2H-5}\frac{(a^6+b^6)}{a^{H+1}b^{H+1}} +$$
$$+K_2|ai-bj|^{2H-4}a^{2-H}b^{2-H}$$

If $\psi$ verifies Assumption A3, if $a$ and $b$ verify condition C and with $|ai-bj| \leq N$ then there exists $K > 0$ such that for large enough $N$:

$$|\mathrm{cov}(\tilde{d}_a(i), \tilde{d}_b(j))| \leq K\left(a^Hb^{-4-H} + b^Ha^{-4-H} + \frac{a^{2-H}b^{2-H}}{|ai-bj|^{2H-4}}\right).$$

In case of Hyp2, we have $J_1 = 0$ instead of (22), and we have again (23) and (24). As a result, (12) is proved again. ■

**Proof of Proposition 4.2.3**: 1/ We first show $\lim_{N\to+\infty}\frac{N}{b(N)}\mathrm{cov}(\tilde{I}_N(a_i(N)), \tilde{I}_N(a_j(N)) = f_{ij}$, where $f_{ij}$ is given in (14). Denote $N_i = [\frac{N}{a_i(N)}] - 1$ for $i \in \{1,2,..,m\}$. We have:

$$\frac{N}{b(N)}\mathrm{cov}(\tilde{I}_N(a_i(N)), \tilde{I}_N(a_j(N)) = \frac{2N}{b(N)N_iN_j}\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\mathrm{cov}^2(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)).$$

From previous computations we have:

$$\operatorname{cov}(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) = O(a_i(N)^H a_j(N)^{-4-H} + a_i(N)^H a_j(N)^{-4-H})$$

$$-\frac{\sigma^2}{2a_i(N)^{H+1}a_j(N)^{H+1}C_{\psi(H)}} \sum_{k=-a_i(N)}^{a_i(N)} \sum_{k'=-a_j(N)}^{a_j(N)} \psi(\frac{k}{a_i(N)})\psi(\frac{k'}{a_j(N)})|k-k'+a_i(N)l-a_j(N)m|^{2H}$$

$$= O(b(N)^{-4}) - \frac{\sigma^2}{2n_i^H n_j^H C_{\psi(H)}} J \quad \text{where}$$

$$J = \frac{1}{a_i(N)}\frac{1}{a_j(N)} \sum_{k=-a_i(N)}^{a_i(N)} \sum_{k'=-a_j(N)}^{a_j(N)} \psi(\frac{k}{a_i(N)})\psi(\frac{k'}{a_j(N)}) \left|\frac{n_i k}{a_i(N)} - \frac{n_j k'}{a_j(N)} + n_i l - n_j m\right|^{2H}.$$

Using Lemma 4.5.2, one obtains for large enough $N$, and for $|k-k'| \geq |a_i(N)l - a_j(N)m|$:

$$J = \int\int \psi(t)\psi(t')|n_i t - n_j t' + n_i l - n_j m|^{2H} + O(\frac{1}{b(N)}).$$

Using Lemma 4.5.1, one obtains for large enough $N$, and for $|k-k'| < |a_i(N)l - a_j(N)m|$:

$$J = \int\int \psi(t)\psi(t')|n_i t - n_j t' + n_i l - n_j m|^{2H} + O(\frac{1}{b^8(N)}).$$

Finally, we obtain for $(l,m)$ such that $|n_i l - n_j m| > n_i + n_j$,

$$\operatorname{cov}^2(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) =$$

$$\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C^2_{\psi(H)}} \left(O(b(N)^{-4}) - \int\int \psi(t)\psi(t')|n_i t - n_j t' + n_i l - n_j m|^{2H} dt dt'\right)^2,$$

else, for $(l,m)$ such that $|n_i l - n_j m| \leq n_i + n_j$,

$$\operatorname{cov}^2(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) =$$

$$\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C^2_{\psi(H)}} \left(O(b(N)^{-1}) - \int\int \psi(t)\psi(t')|n_i t - n_j t' + n_i l - n_j m|^{2H} dt dt'\right)^2.$$

By this way, we deduce:

$$\sum_{l=1}^{N_i} \sum_{m=1}^{N_j} \operatorname{cov}^2(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) =$$

$$\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C_{\psi(H)}} \sum_{l=1}^{N_i} \sum_{m=1}^{N_j} \mathbb{I}_{|n_i l - n_j m| \leq n_i + n_j} \left(O(\frac{1}{b(N)}) - \int\int \psi(t)\psi(t')|n_i t - n_j t' + n_i l - n_j m|^{2H} dt dt'\right)^2$$

$$+\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C^2_{\psi(H)}}\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\mathbb{I}_{|n_il-n_jm|>n_i+n_j}\left(O(\frac{1}{b(N)^4})-\int\int\psi(t)\psi(t')|n_it-n_jt'+n_il-n_jm|^{2H}dtdt'\right)^2$$
$$=\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C^2_{\psi(H)}}\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\left(\int\int\psi(t)\psi(t')|n_it-n_jt'+n_il-n_jm|^{2H}dtdt'\right)^2+O(\frac{N}{b(N)^4}+\frac{N^2}{b(N)^6}),$$

because $\#\{(l,m)/|n_il-n_jm|\leq n_i+n_j\}=O(\frac{N}{b(N)})$ when $(l,m)\in\{1,..,n_ib(N)\}\times\{1,..,n_jb(N)\}$, and $\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}O(b(N)^{-4})=O(\frac{N^2}{b(N)^6})$.

We know that $\lim_{N\to+\infty}\frac{N}{b(N)}\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\mathbb{I}_{|n_il-n_jm|>n_i+n_j}|n_il-n_jm|^{2H-4}$ exists and is a positive real number. As a consequence and with (12), one obtains

$$\lim_{N\to+\infty}\frac{N}{b(N)}\mathrm{cov}(\tilde I_N(a_i(N),\tilde I_N(a_j(N))<+\infty.$$

This limit is given by $f_ij$ where:

$$f_{ij}=\lim_{N\to+\infty}\frac{2N}{b(N)N_iN_j}\frac{\sigma^4}{4n_i^{2H}n_j^{2H}C^2_{\psi(H)}}\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\left(\int\int\psi(t)\psi(t')|n_it-n_jt'+n_il-n_jm|^{2H}dtdt'\right)^2.$$

Using a similar method to Poggi and Viano (1996) and denoting $D_{ij}=GCD(n_i,n_j)$, we obtain the same expression for $f_{ij}$ than in (14).

In case of Hyp2, the proof is the same with simplifications giving by (4).

2/ Let us prove in a second time the central limit theorem (13). We follow a similar process as in proof of Theorem 6.1 of Taqqu (1975), already used in Bardet (1997). Denote:

$$f_N(u_1,..,u_m)=\mathbb{E}exp\left(i\sum_{j=1}^m u_j\frac{N}{b(N)}\left(\tilde I_N(a_j(N)-\mathbb{E}\tilde I_N(a_j(N))\right)\right)\text{ with }(u_1,..,u_m)\in\mathbb{R}^m.$$

We are going to prove the convergence of the function $f_N(u_1z,..,u_mz)$ when $N\to+\infty$ and when $z$ is close enough to zero. It is shown in Bardet (1997) that it exists $\varepsilon>0$, such as for $|z|<\varepsilon$:

$$f_N(u_1z,..,u_mz)=\exp\left(\frac{1}{2}\sum_{k=2}^{+\infty}\frac{(2iz)^k}{k}\sum_{i_1,..,i_k=1}^m u_{i_1}..u_{i_k}I_N(n_{i_1},..,n_{i_k})\right)\text{ where}$$

$$I_N(n_{i_1},..,n_{i_k})=\frac{(\frac{N}{b(N)})^{k/2}}{N_{i_1}..N_{i_k}}\sum_{j_1=1}^{N_{i_1}}..\sum_{j_k=1}^{N_{i_k}}\mathrm{cov}(\tilde d_{a_{i_1}(N)}(j_1),\tilde d_{a_{i_2}(N)}(j_2))..\mathrm{cov}(\tilde d_{a_{i_k}(N)}(j_k),\tilde d_{a_{i_1}(N)}(j_1)).$$

As a consequence of (12),

$$\begin{aligned}
\operatorname{cov}(\tilde{d}_{a_{i_p}(N)}(j_p), \tilde{d}_{a_{i_q}(N)}(j_q)) &\leq K\left(b(N)^{-4} + |n_{i_p}j_p - n_{i_q}j_q|^{2H-4}\right) && (25)\\
&\leq \frac{2K}{|n_{i_p}j_p - n_{i_q}j_q|^{4/3}}, && (26)
\end{aligned}$$

when $|n_{i_p}j_p - n_{i_q}j_q| \geq 2(n_{i_p} + n_{i_q})$. Let us denote $\Delta$ the set of $(j_1, .., j_k) \in \{1, .., N_{i_1}\} \times .. \times \{1, .., N_{i_k}\}$ such that

$$|n_{i_p}j_p - n_{i_{p+1}}j_{p+1}| \geq 2(n_{i_p} + n_{i_{p+1}}) \text{ for all } p \in \{1, 2, .., k\}, \quad \text{with } i_{k+1} = i_1.$$

It is easy to show that there exists $D(n_1, .., n_k) > 0$ such that $\#\Delta \leq D(n_1, .., n_k)\dfrac{N}{b(N)}$.
By this way, we deduce, on the one hand,

$$\frac{(\frac{N}{b(N)})^{k/2}}{N_{i_1}..N_{i_k}} \sum_{j_1=1}^{N_{i_1}} .. \sum_{j_k=1}^{N_{i_k}} \mathbb{I}_{(j_1,..,j_k)\in\Delta} |\operatorname{cov}(\tilde{d}_{a_{i_1}(N)}(j_1), \tilde{d}_{a_{i_2}(N)}(j_2))|..|\operatorname{cov}(\tilde{d}_{a_{i_k}(N)}(j_k), \tilde{d}_{a_{i_1}(N)}(j_1))| \quad (27)$$

$$\leq \frac{n_1...n_k b(N)^{k/2}}{N^{k/2}} D(n_1, .., n_k)\frac{N}{b(N)}$$

because $|\operatorname{cov}(\tilde{d}_a(i), \tilde{d}_b(j))| \leq 1$ for all $(a, b, i, j) \in \mathbb{N}^{*4}$. Then the limit of (27) when $N \to +\infty$ is 0 from $k \geq 3$.
On the other hand, from (25),

$$\frac{(\frac{N}{b(N)})^{k/2}}{N_{i_1}..N_{i_k}} \sum_{j_1=1}^{N_{i_1}} .. \sum_{j_k=1}^{N_{i_k}} \mathbb{I}_{(j_1,..,j_k)\notin\Delta} |\operatorname{cov}(\tilde{d}_{a_{i_1}(N)}(j_1), \tilde{d}_{a_{i_2}(N)}(j_2))|..|\operatorname{cov}(\tilde{d}_{a_{i_k}(N)}(j_k), \tilde{d}_{a_{i_1}(N)}(j_1))| \quad (28)$$

$$\begin{aligned}
&\leq \frac{n_1...n_k b(N)^{k/2}}{N^{k/2}} \sum_{j_1=1}^{N_{i_1}} .. \sum_{j_k=1}^{N_{i_k}} \mathbb{I}_{(j_1,..,j_k)\notin\Delta} \frac{2K}{|n_{i_1}j_1 - n_{i_2}j_2|^{4/3}} .. \frac{2K}{|n_{i_k}j_k - n_{i_1}j_1|^{4/3}}\\
&\leq \frac{n_1...n_k b(N)^{k/2}}{N^{k/2}} D'(n_1, .., n_k)\frac{N}{b(N)}, \text{ where } D'(n_1, .., n_k) = 2\sum_{m_1=1}^{+\infty} ... \sum_{m_k=1}^{+\infty} \frac{2K}{|m_1|^{4/3}} .. \frac{2K}{|m_k|^{4/3}}.
\end{aligned}$$

Thus, the limit of (28) when $N \to +\infty$ is 0 from $k \geq 3$.
Finally, we have shown that $\lim_{N\to+\infty} I_N(n_{i_1}, .., n_{i_k}) = 0$ from $k \geq 3$, and then $f_N(u_1z, .., u_mz)$ has a limit when $N \to +\infty$, that is:

$$f(u_1z, .., u_mz) = exp\left(\frac{1}{2}\frac{(2iz)^2}{2}\sum_{i,j=1}^{m} u_iu_jf_{ij}\right),$$

and like $f_N(u_1z, .., u_mz)$ is an analytical function, then $\dfrac{N}{b(N)}\left(\tilde{I}_N(a_j(N)) - \mathbb{E}\tilde{I}_N(a_j(N))\right)_{1\leq j\leq m}$ asymptotically behaves as a Gaussian distribution given in Proposition 4.2.3. Thus, with

(10), $\mathbb{E}\tilde{I}_N(a_j(N)) = 1 + O(\frac{1}{b(N)})$, Proposition 4.2.3 is proved. ■

**Proof of Corollary 4.2.1**: Let $d'_a(i)$ and $\tilde{d}'_a(i)$ be respectively coefficients and normalized coefficients computed from $(X'_0, .., X'_N)$. We obtain:

$$\begin{aligned}
\mathbb{E}(d'_a(i))^2 &= \mathbb{E}d_a^2(i) + \frac{1}{a}\sum_{k=-a}^{a}\sum_{k'=-a}^{a}\psi(\frac{k}{a})\psi(\frac{k'}{a})\mathbb{E}\varepsilon_{k+ia}\varepsilon_{k'+ia} \\
&= \mathbb{E}d_a^2(i) + \frac{1}{a}\sum_{k=-a}^{a}\psi^2(\frac{k}{a})\gamma^2(k+ia) \\
&= \mathbb{E}d_a^2(i) + O(1),
\end{aligned}$$

because $\psi$ and $\gamma$ are bounded. By this way, for large enough $a$,

$$\mathbb{E}(\tilde{d}'_a(i))^2 = 1 + O(\frac{1}{a}) + O(\frac{1}{a^{2H+1}}) = 1 + O(\frac{1}{a}). \tag{29}$$

We also obtain:

$$\begin{aligned}
&\text{cov}(\tilde{d}'_{a_i(N)}(l), \tilde{d}'_{a_j(N)}(m)) = \text{cov}(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) + \\
&- \frac{\sigma^2}{a_i(N)^{H+1}a_j(N)^{H+1}C_{\psi(H)}}\sum_{k=-a_i(N)}^{a_i(N)}\sum_{k'=-a_j(N)}^{a_j(N)}\psi(\frac{k}{a_i(N)})\psi(\frac{k'}{a_j(N)})\mathbb{E}\varepsilon_{k+la_i(N)}\varepsilon_{k'+ma_j(N)},
\end{aligned}$$

with same notations as in Proposition 4.2.3. For $(l, m)$ such that $|ln_i - mn_j| > |n_i + n_j|$, we have $\mathbb{E}\varepsilon_{k+la_i(N)}\varepsilon_{k'+ma_j(N)} = 0$ and thus,

$$\text{cov}(\tilde{d}'_{a_i(N)}(l), \tilde{d}'_{a_j(N)}(m)) = \text{cov}(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)). \tag{30}$$

For $(l, m)$ such that $|ln_i - mn_j| \le |n_i + n_j|$, we have

$$\frac{1}{b(N)^2}\sum_{k=-a_i(N)}^{a_i(N)}\sum_{k'=-a_j(N)}^{a_j(N)}\psi(\frac{k}{a_i(N)})\psi(\frac{k'}{a_j(N)})\mathbb{E}\varepsilon_{k+la_i(N)}\varepsilon_{k'+ma_j(N)} = O(1),$$

because $\psi$ and $\gamma$ are bounded, and thus,

$$\text{cov}(\tilde{d}'_{a_i(N)}(l), \tilde{d}'_{a_j(N)}(m)) = \text{cov}(\tilde{d}_{a_i(N)}(l), \tilde{d}_{a_j(N)}(m)) + O(\frac{1}{b(N)^{2H}}). \tag{31}$$

With (29), (30) and (31), we show that nothing changing in previous proofs by working from $(X'_0, .., X'_N)$ instead of $(X_0, .., X_N)$. Thus, Corollary 4.2.1 is proved. ■

**Acknowledgment.** The author thanks Prof. J.M. Poggi, G. Oppenheim and D. Dacunha-Castelle for several stimulating conversations.

# Bibliographie

[1] Abry, P. and Sellan, F. (1996). The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: remarks and fast implementation. *Applied and Computational Harmonic Analysis* 3, 377-383.

[2] Abry, P., Goncalves, P. and Flandrin, P. (1995). Wavelets, spectrum analysis and $1/f$ processes. Wavelets and Statistics, Lectures Note in Statistics, 103, 15-29.

[3] Abry, P. and Veitch D. (1997). Wavelet analysis of long range dependent traffic. To appear in *IEEE Trans. on Inform. Theory.*

[4] Abry, P., Veitch D. and Flandrin, P. (1997). Long-range dependence: revisiting aggregation with wavelets. To appear in *J. Time Ser. Anal.*

[5] Bardet, J.M. (1997). Testing self-similarity of Gaussian processes with stationary increments. *Preprint Orsay.*

[6] Beran, J (1994). *Statistics for long memory processes.* Monographs on Stat. and Appli. Prob. 61. Chapman & Hall, 315 p.

[7] Daubechies, I. (1992). *Ten lectures on wavelets.* CBMS, SIAM 61, Philadelphia.

[8] Flandrin, P. (1989). On the spectrum of fractional Brownian motions. *IEEE Trans. on Info. Theory* 35 197-199.

[9] Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. on Inform. Theory* 38 910-917.

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

[11] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle estimate. *Probab. Theory Related Fields* 86, 87-104.

[12] Goncalvès, P. and Abry, P. (1997).Multiple-window wavelet transform and local scaling exponent estimation. *Preprint INRIA*.

[13] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Poincaré* 33, 407-436.

[14] Mandelbrot, B.B. (1983). *The fractal geometry of nature*. Freeman, San Francisco.

[15] Poggi, J.M. and Viano, M.C. (1996). An estimate of the fractal index using multiscale aggregates. To appear in *J. Time Ser. Anal.*

[16] Robinson, P.M. (1995). Log-periodogram regression of time series with long-range dependence. *Ann. Statist.* 23, 1048-1072.

[17] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable non-Gaussian random processes*. Stochastic modeling. Chapman & Hall, 636 p.

[18] Taqqu, M.S. (1975). Weak Convergence to Fractional Brownian Motion and to Rosenblatt Process. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 31, 287-302.

[19] Taqqu, M.S., Teverovsky, V. and Willinger, W. (1995). Estimators for long-range dependence: an empirical study. *Fractals* 3, 785-798.

[20] Yajima, Y. (1989). Asymptotic properties of the LSE in a regression model with long-range stationary errors. *Ann. Statist.* 19, 158-177.

# Tests d'autosimilarité des processus gaussiens. Dimension fractale et dimension de corrélation locale.

Jean-Marc Bardet

RESUME: Cette thèse se compose de quatre parties ayant pour dénominateur commun la notion de fractalité. La première partie est une présentation des systèmes dynamiques chaotiques sous la forme d'une revue bibliographique. On s'intéresse aux différentes dimensions d'un attracteur et aux exposants de Lyapunov. On recense les différents traîtements statistiques utilisés pour estimer ces données, notamment les notions d'exposants de Lyapunov locaux et de dimension de corrélation. La seconde partie définit et étudie une dimension de corrélation dite locale pour les processus stochastiques continus. Cette dimension se détermine à partir du comportement asymptotique de la mesure d'occupation d'auto-intersection. On compare cette dimension avec la dimension de Hausdorff, et on détermine ces deux dimensions dans le cas des champs gaussiens à index. La troisième partie présente un test d'autosimilarité pour une série d'observations provenant d'un processus gaussien à accroissements stationnaires. Ce test repose sur l'estimation d'une distance entre le processus considéré et un ensemble de processus contenant tous les mouvements browniens fractionnaires. Cette distance est construite à partir d'une double estimation, pour une gamme d'échelles temporelles, de la variance des accroissements du processus. La seconde de ces estimations fait appel à une double estimation du paramètre d'autosimilarité par des méthodes de régression. Ces estimateurs présentent une vitesse de calcul avantageuse et une grande robustesse en comparaison de l'estimation par maximum de vraisemblance. La quatrième partie considère un second test d'autosimilarité construit sur le même principe, mais utilisant l'analyse par ondelettes. La statistique de test est alors une distance entre deux estimations différentes de la variance des coefficients d'ondelettes pour certaines échelles. Deux estimateurs robustes et simples du paramètre d'autosimilarité sont encore introduits.

MOTS CLES:
- Fractional Brownian motion
- Gaussian fields
- Hausdorff dimension
- Correlation dimension
- Estimation of the fractal dimension
- Self-similarity
- Wavelet analysis
- Non-parametric test