

# Méthodes d'analyse du mouvement brownien fractionnaire : théorie et résultats comparatifs

## Analysis methods for fractional brownian motion: theory and comparative results

par Rachid JENNANE\*, Rachid HARBA\*, Gérard JACQUET\*\*

\*Laboratoire d'Electronique, Signaux, Images EA 1715 et GDR ISIS du CNRS ESPEO, Université d'Orléans, BP 6744, 45067 Orléans Cedex 2.

Tél : 33 2 38 49 45 38 Fax : 33 2 38 41 72 45 e-mail : Rachid.Jennane@univ-orleans.fr e-mail : Rachid.Harba@univ-orleans.fr

\*\*Laboratoire Traitement du Signal et Instrumentation UMR 5516 et GDR ISIS du CNRS Université de Saint-Etienne, 23, rue Paul Michelon 42023 Saint-Etienne Cedex 2.

Tél : 33 4 77 48 50 24 Fax : 33 4 77 48 50 39 e-mail : Jacquet@univ-st-etienne.fr

### *résumé et mots clés*

Cet article a pour objectif de comparer les performances des différents estimateurs du paramètre  $H$  du mouvement brownien fractionnaire (fBm) sur des signaux théoriquement exacts synthétisés par la méthode de Cholesky. Des ensembles de 100 signaux de taille  $N = 32$  à 1024 par puissance de 2 sont générés pour des valeurs du paramètre  $H$  variant de 0.1 à 0.9 par pas de 0.1. Les principales méthodes d'estimation ont été regroupées en quatre classes : les méthodes géométriques, les méthodes temporelles, les méthodes fréquentielles, et enfin les méthodes basées sur une décomposition multi-échelles. Chaque technique a été évaluée en termes de biais et de variance, cette dernière étant comparée à la borne de Cramer-Rao (BCR). Des tests statistiques montrent que seul l'estimateur par maximum de vraisemblance (EMV) est efficace (non biaisé et atteignant la BCR) pour toutes les valeurs de  $H$  et de  $N$  testées. Ce résultat expérimental complète les résultats théoriques d'efficacité de l'EMV démontrés par Dahlhaus uniquement dans le cas asymptotique [1]. D'un point de vue pratique, l'implémentation de l'EMV devient d'un coût calculatoire important pour des signaux de grande longueur. L'approche de Whittle du maximum de vraisemblance dans le domaine fréquentiel permet de lever ces limitations, cette technique étant asymptotiquement efficace. Enfin, des résultats montrent que l'approche EMV permet de traiter le cas des signaux fBm pollués par un bruit.

Mouvement brownien fractionnaire, fBm, maximum de vraisemblance, méthode de Whittle.

### *abstract and key words*

In this paper, several analysis methods for fractional Brownian motion are studied using reference test signals generated by the Cholesky procedure. Several sets of 100 signals having sample size ranging from  $N = 32$  to 1024 by power of 2 are generated for  $H = 0.1$  to 0.9 by steps of 0.1. Analysis techniques of the  $H$  parameter among the most well known in signal processing are studied. They are grouped in four categories: frequency based methods, geometrical methods, time methods, and multi-resolution methods. Quality of these estimators is assessed in terms of bias and variance. The variance is compared to the Cramer-Rao lower bound (CRLB). Statistical tests show that only the maximum likelihood estimator (MLE) is efficient (unbiased and reaches the CRLB) for every  $H$  and  $N$  tested. This experimental result about the efficiency of MLE extends that demonstrated by Dahlhaus only in the

asymptotical case [1]. But MLE has a high computer burden in case of signals having a lot of samples. The Whittle approximation in the frequency domain of the MLE allows to get ride of these limitations, this technique being asymptotically efficient. Finally, results show that MLE approach can be used to measure the  $H$  parameter on signals embedded in a white noise.

Fractional Brownian motion, fBm, maximum likelihood estimator, Whittle method.

## 1. introduction

Le mouvement brownien fractionnaire (fBm pour fractional Brownian motion) de paramètre  $H$  tel que  $0 < H < 1$  est un modèle non stationnaire de signaux fractals stochastiques [2]. Les incréments de ce processus, nommés bruits gaussiens fractionnaires (fGn pour fractional Gaussian noises), sont stationnaires. Ces deux modèles ont tout d'abord été utilisés pour modéliser des signaux issus de phénomènes physiques tels que des bruits en  $1/f$  [3], des séries économiques [4], et plus récemment des données de trafic Ethernet [5]. En deux dimensions, le paramètre  $H$  permet de quantifier la notion intuitive de rugosité d'une image [6]. Il est alors employé pour caractériser des textures [7] ou pour analyser des surfaces [8].

Lors de l'application du modèle fBm à un cas concret, il convient dans un premier temps de se soucier du caractère fractal des signaux expérimentaux (signaux discrets de longueur finie). Des méthodes ont déjà été proposées pour s'assurer de l'adéquation entre le modèle fBm et des signaux réels [9][10]. Dans un deuxième temps, il est nécessaire d'estimer aussi précisément que possible le paramètre  $H$ . Il existe de nombreuses méthodes qui permettent de réaliser cette tâche et il est difficile d'en choisir une *a priori*. Théoriquement, même si des résultats d'efficacité asymptotique sont démontrés pour certains estimateurs [1], rien ne permet de connaître leur comportement dans toutes les circonstances (signaux de longueur réduite, sensibilité au bruit...). Il est donc nécessaire de mener des études expérimentales pour mieux cerner les potentialités des estimateurs. Ces études reviennent à tester ces diverses méthodes sur des signaux fBm de synthèse [11][12]. Dans la première référence, une analyse croisée synthèse-analyse n'a pas permis de tirer de conclusions définitives. En effet, certaines méthodes de synthèse ne sont qu'approximatives et peuvent interférer avec l'analyse. Dans la deuxième étude, seuls quelques estimateurs ont été évalués uniquement pour  $H > 0.5$ . Pour mener une étude générale, il faut au préalable disposer d'une méthode de synthèse générant de vrais signaux fBm. La synthèse par décomposition de Cholesky de la matrice de covariance des incréments du fBm est théoriquement exacte [13]. En effet, les statistiques d'ordre 2 des incréments de tels signaux sont celles des incréments gaus-

siens du fBm. L'appréciation de la qualité d'un estimateur revient à comparer la valeur du paramètre  $H$  mesurée sur ces signaux à celle injectée lors de la synthèse.

L'objectif du travail présenté dans cet article est d'étudier les méthodes d'estimation du paramètre  $H$  du fBm les plus utilisées en traitement du signal. Cette évaluation sera faite sur des données synthétiques théoriquement exactes, c'est-à-dire qu'aucun doute ne subsiste quant à la valeur du paramètre  $H$  injectée lors de leur génération. Pour chaque méthode, la qualité des estimateurs sera évaluée par la mesure de leur biais et de leur variance. La variance sera comparée à la borne de Cramer-Rao (BCR).

Le plan de cet article est le suivant : la section 2 définit le cadre de cette étude. En premier lieu, elle est consacrée à un bref rappel des principales propriétés du fBm en insistant sur le cas des signaux discrets de longueur finie. Puis, la méthode de Cholesky est décrite et il est montré comment les divers estimateurs seront évalués. La section 3 décrit successivement les différents estimateurs retenus pour l'analyse de signaux fBm et présente les résultats obtenus. La section 4 étudie plus particulièrement le cas des signaux de grande longueur ou bruités.

## 2. cadre de l'étude

Les propriétés du fBm et de ses incréments sont brièvement rappelées en insistant sur celles utilisées pour estimer le paramètre  $H$  de signaux discrets composés de  $N$  échantillons. Elles sont présentées en particulier avec plus de détails dans [2] en continu et dans [14] en échantillonné (dans la suite du texte, les parenthèses seront utilisées en temps continu et les crochets en temps discret). Enfin, nous décrirons la méthode qui permet de synthétiser des traces fBm exactes et exposerons les critères qui permettront de qualifier un estimateur donné (biais et variance).

### 2.1. propriétés du fBm

Le fBm de paramètre  $H$  ( $0 < H < 1$ ), noté  $B_H$ , a été défini par Mandelbrot et Van Ness [2] comme une extension du mouvement brownien (pour  $H = 0.5$ ,  $B_H$  est le mouvement brownien ordinaire  $B$ ) :

$$\begin{aligned}
 B_H(t) - B_H(0) &= \frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^0 \left( (t-s)^{H-1/2} - (s)^{H-1/2} \right) dB(s) \right. \\
 &\quad \left. + \int_0^t (t-s)^{H-1/2} dB(s) \right\} \quad (1)
 \end{aligned}$$

D'un point de vue pratique, c'est un processus gaussien, continu, centré, et non stationnaire au deuxième ordre. Avec la condition initiale  $B_H(0) = 0$ , sa fonction d'autocorrélation, notée  $r$ , le décrit entièrement :

$$\begin{aligned}
 r_{B_H}(i, j) &= E(B_H(i) B_H(j)) \\
 &= \frac{a^2 V_H}{2} (|i|^{2H} + |j|^{2H} - |i - j|^{2H}) \quad (2)
 \end{aligned}$$

$E(\cdot)$  est l'opérateur espérance mathématique,  $a$  est une constante qui permet de construire toute une classe de processus à un facteur d'échelle près et  $V_H$  est une fonction de  $H$  telle que [15] :

$$V_H = \Gamma(1 - 2H) \frac{\cos(\pi H)}{\pi H} \quad (3)$$

Une des propriétés notable du fBm est son auto-similarité statistique :

$$B_H(km) \equiv k^H B_H(m); \quad \forall k \text{ et } m > 0 \quad (4)$$

$\equiv$  signifie égalité en distribution.

En ce qui concerne les caractéristiques spectrales, Flandrin a utilisé une approche temps-fréquence qui prend en compte le caractère non stationnaire du fBm [16]. Il montre, à partir de (2), que le spectre de Wigner-Ville (SWV) du fBm varie comme :

$$SWV_{B_H}(t, f) \propto (1 - 2^{1-2H} \cos(4\pi ft)) |f|^{-1-2H} \quad (5)$$

$\propto$  signifie « proportionnel à ».

Ce spectre peut prendre des valeurs négatives ce qui met en évidence la difficulté d'étudier ce processus non stationnaire. Cependant, en moyennant cette dernière expression sur un grand intervalle de temps, la relation (5) peut s'écrire sous la forme d'un spectre moyen (SM) toujours positif :

$$SM_{B_H}(f) \propto |f|^{-1-2H} \quad (6)$$

Le fBm est donc un processus basse fréquence en  $1/f^\alpha$  avec  $\alpha$  dans l'intervalle  $]1, 3[$ . Il présente une singularité pour  $f = 0$ . Le fBm discret  $B_H[n]$  est tel que  $B_H[n] = B_H(nTe)$  où  $Te$  est la période d'échantillonnage. Le spectre moyen du fBm discret, avec  $fe = 1/Te$ , peut s'écrire approximativement comme :

$$SM_{B_H}[f] \propto \sum_{k=-\infty}^{+\infty} |f + kfe|^{-1-2H} \quad (7)$$

L'échantillonnage induit un repliement spectral et propage les singularités pour les multiples de  $fe$ . Le repliement sera surtout significatif pour les faibles valeurs de  $H$ .

La troncature liée à la durée limitée du signal réel introduit un lissage du spectre moyen, mais ce lissage a peu d'influence sur la forme globale du spectre.

Enfin, une décomposition en ondelettes montre que la non-stationnarité du processus se retrouve uniquement dans le signal de forme basse fréquence, les signaux de détails étant quant à eux stationnaires et quasiment décorrés [17]. Lors d'une décomposition en ondelettes dyadiques, la variance des signaux de détails  $d_j$  à une échelle  $j$  suit une loi telle que :

$$\text{Var}(d_j[n]) \propto (2^j)^{2H+1} \quad (8)$$

$\text{Var}(\cdot)$  est l'opérateur variance mathématique.

## 2.2. propriétés des incréments du fBm

Comme le fBm est continu non dérivable, on définit ses incréments pour un intervalle de temps  $m$ . Ils peuvent être assimilés à une dérivée du processus à une résolution  $m$ . Ils sont appelés bruits gaussiens fractionnaires (fGn pour fractional Gaussian noises) notés  $G_m$  et définis comme :

$$G_m(i) = B_H(i) - B_H(i - m) \quad (9)$$

Ce sont des processus gaussiens, centrés et stationnaires. Leur fonction d'autocorrélation  $r$  les décrit entièrement à une résolution  $m$  :

$$\begin{aligned}
 r_{G_m}(\tau) &= E(G_m(i) G_m(i + \tau)) \\
 &= \frac{\sigma_m^2}{2|m|^{2H}} (|\tau + m|^{2H} - 2|\tau|^{2H} + |\tau - m|^{2H}) \quad (10)
 \end{aligned}$$

$\sigma_m^2 = a^2 V_H |m|^{2H}$  représente la variance de  $G_m$ .

Pour  $0 < H < 0.5$ , la corrélation est négative. Pour  $H = 0.5$  ce processus est le bruit blanc gaussien. Quand  $0.5 < H < 1$ , la corrélation est positive. De plus, pour  $\tau$  grand, la fonction d'autocorrélation varie comme  $\tau^{2H-2}$ . De ce fait, elle n'est pas sommable : on parle alors de longue mémoire ou de longue dépendance [18].

La densité spectrale de puissance (DSP) du fGn peut se calculer directement par transformée de Fourier de l'équation (10) [16] :

$$\text{DSP}_{G_m}(f) \propto \sin^2(\pi f m) |f|^{-1-2H} \quad (11)$$

Lorsque  $m$  tend vers 0, on peut alors définir un spectre normalisé du fGn continu, dérivée du fBm :

$$\text{DSP}_{G_m}(f) \propto |f|^{1-2H} \quad (12)$$

Le fGn est donc un processus stationnaire en  $1/f^\alpha$  avec  $\alpha$  dans l'intervalle  $] -1, 1[$ . Ce processus est globalement haute fréquence si  $H < 0.5$ , basse fréquence si  $H > 0.5$ .

Concernant le fGn discret, le coefficient  $m$  s'exprime en multiple de la période d'échantillonnage  $T_e$ . La plus petite valeur possible de  $m$  étant 1, le spectre du fGn sera en  $1/f^\alpha$  uniquement aux basses fréquences. La DSP du fGn discret est alors la suivante [12] :

$$\text{DSP}_{G_m}[f] \propto \sin^2 \left( \pi m \frac{f}{f_e} \right) \sum_{k=-\infty}^{+\infty} |f + k f_e|^{-1-2H} \quad (13)$$

Un terme de lissage doit être rajoutée à la DSP du fGn discret ayant  $N$  échantillons, ce qui conduit à une expression lourde à manipuler.

En revanche, la fonction d'autocorrélation du fGn discret s'écrit plus simplement pour  $m = 1$  [14] :

$$r_{G_1}[i] = \frac{\sigma_1^2}{2} [|i+1|^{2H} - 2|i|^{2H} + |i-1|^{2H}] \quad (14)$$

*Remarque :* Cette expression est la version discrétisée de (10) signifiant que la fonction d'autocorrélation du fGn discret est égale à celle du fGn continu aux instants d'échantillonnage. Ceci explique en partie les raisons du bon comportement de certaines méthodes de synthèse ou d'analyse basées sur la fonction d'autocorrélation du fGn.

Une des propriétés notables des processus fGn, conséquence de la nature fractale du fBm, est la similarité statistique entre processus obtenus à des échelles différentes :

$$G_{km}(i) \equiv k^H G_m(i) \quad (15)$$

Cette équation est vraie pour tout  $k > 0$ .

À partir de ces diverses propriétés, on peut construire plusieurs estimateurs du paramètre  $H$ . Avant de les comparer, il convient de définir un cadre précis pour mener à bien cette étude : comment synthétiser des signaux tests de qualité, et comment évaluer précisément les divers estimateurs du paramètre  $H$ .

### 2.3. génération de signaux tests

Pour élaborer des signaux tests, il faut non seulement prendre en compte l'exactitude théorique de la méthode de synthèse mais aussi les difficultés pratiques liées à l'utilisation de signaux discrets de longueur finie. Concernant le fBm, il existe de nombreuses méthodes de synthèse [19]. Elles ont fait l'objet d'études pratiques [9][11]. Parmi ces méthodes, celle basée sur les ondelettes [20][21] et celle dite de Cholesky [13][14] peuvent être utilisées pour générer des signaux tests en raison de leurs bonnes propriétés théoriques. Pour la première méthode, une trace fBm peut être construite à partir d'un signal ARIMA fractionnaire [22][23] pour les basses fréquences et d'ondelettes

pour les hautes fréquences. Il est montré que cette construction converge vers le processus fBm [24]. Cependant, il faudra nécessairement négliger des termes lors d'une implémentation pratique, et cette méthode ne sera alors plus exacte [21]. Par contre, la méthode fondée sur la décomposition de Cholesky de la matrice de covariance des incréments du fBm est théoriquement exacte pour des signaux discrets de longueur finie. On peut expliquer ceci par le fait que cette méthode utilise la fonction d'autocorrélation du fGn discret qui est égale à celle du fGn continu aux instants d'échantillonnage (voir remarque du paragraphe 2.2). De plus, cette méthode est facile à implanter numériquement pour le cas de données à une dimension. C'est celle que nous avons utilisée.

Cette méthode consiste à filtrer un bruit blanc gaussien pour lui donner la fonction d'autocorrélation du fGn. Elle utilise la matrice de covariance  $\mathbf{R}$  du fGn discret calculée pour un incrément de  $m = 1$  :  $[\mathbf{R}]_{ij} = r[|i-j|]$  donnée par (14). Cette matrice est Toeplitz, définie positive. Elle peut être décomposée par la méthode de Cholesky comme :

$$\mathbf{R} = \mathbf{L}\mathbf{L}^T \quad (16)$$

$\mathbf{L}$  est une matrice triangulaire inférieure, définie positive. Soit  $\mathbf{x}$  un vecteur tel que :

$$\mathbf{x} = \mathbf{L}\mathbf{y} \quad (17)$$

où  $\mathbf{y}$  est un vecteur bruit blanc gaussien, alors :

$$E(\mathbf{x}\mathbf{x}^T) = E(\mathbf{L}\mathbf{y}\mathbf{y}^T\mathbf{L}^T) = \mathbf{R} \quad (18)$$

Ceci montre que  $\mathbf{x}$ , vecteur gaussien, a des statistiques d'ordre 2 identiques à celles du fGn. Pour un processus gaussien, une méthode de synthèse est exacte au sens strict si les statistiques d'ordre deux du simulateur numérique sont égales à celles du processus cible continu aux instants d'échantillonnage [25]. On en déduit que la méthode de Cholesky est théoriquement exacte pour générer des signaux numériques fGn quelle que soit la taille des données.

Le processus synthétisé  $\mathbf{x}$  est donc identique au processus fGn noté  $G_1$ . Avec la condition initiale  $B_H[0] = 0$ , le fBm discret noté  $B_H$  est obtenu en incrémentant suivant (9) pour  $m = 1$  :

$$B_H[i] = \sum_{j=1}^i G_1[j] = G_1[i] + B_H[i-1] \quad (19)$$

Trois signaux synthétisés par cette méthode sont présentés sur la figure 1.

La non-stationnarité du processus est évidente ainsi que l'évolution des traces avec  $H$  : plus  $H$  augmente, plus le signal devient lisse.

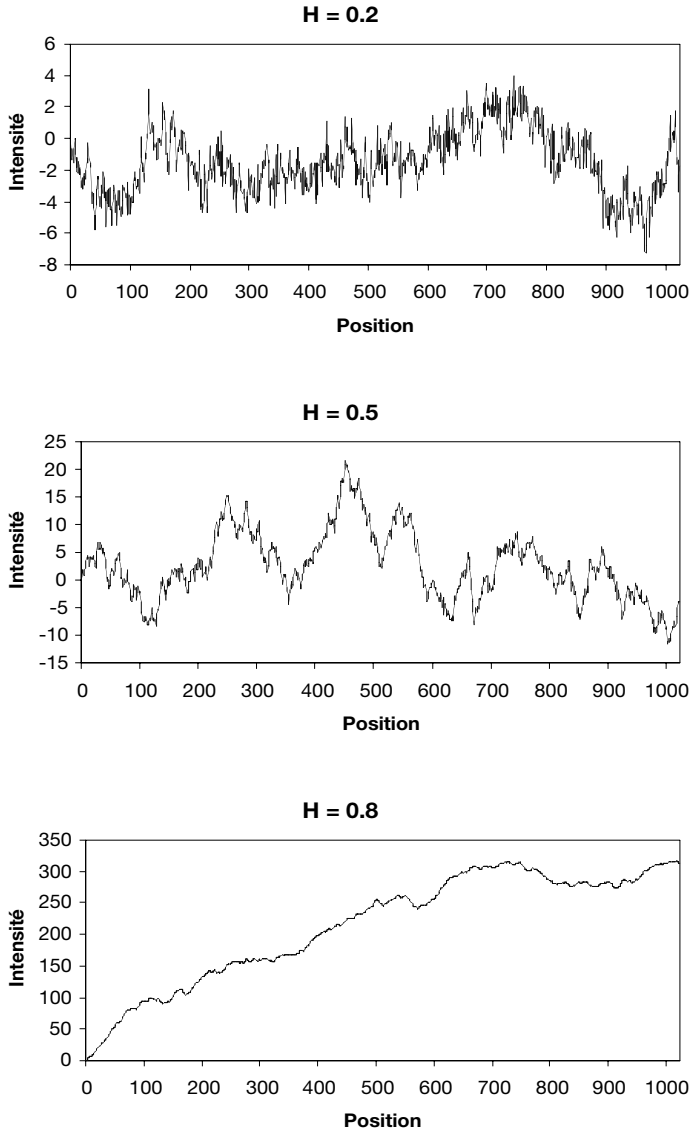


Figure 1. – Signaux fBm de synthèse type Cholesky de 1024 échantillons pour  $H = 0.2$ ,  $0.5$ , et  $0.8$ .

## 2.4. biais, variance et borne de Cramer-Rao

Deux critères doivent être pris en compte pour quantifier les performances d'un estimateur : son biais et sa variance. Ces deux quantités devant être les plus petites possibles.

**Biais :**

Un estimateur  $\hat{H}$  du paramètre  $H$  est non biaisé si :

$$E(\hat{H}) = H \quad (20)$$

En pratique, seul un estimateur de  $E(\hat{H})$  est accessible, car l'expérience est menée sur un nombre limité de réalisations. Le biais est alors qualifié par un test d'hypothèse [26].

Pour savoir si une méthode d'analyse est non biaisée, nous allons effectuer un test d'hypothèse bilatéral au niveau  $\alpha$  avec l'hypothèse  $H1 : E(\hat{H}) = H$ . Si l'hypothèse  $H1$  est refusée, on acceptera l'hypothèse alternative  $H0 : E(\hat{H}) \neq H$ . Dans le cas gaussien, l'hypothèse  $H1$  sera acceptée si la moyenne empirique  $H_{est}$  est calculée sur  $NR$  réalisations de longueur  $N$  est comprise entre deux bornes, l'une supérieure (Sup) et l'autre inférieure (Inf) :

$$\text{Sup} = H - z_{1-\alpha/2} \frac{S}{\sqrt{NR}} \leq H_{est} \leq H + z_{1-\alpha/2} \frac{S}{\sqrt{NR}} = \text{Inf} \quad (21)$$

$S$  est l'écart-type empirique de l'estimateur, et  $z_a$  est la  $100 \times a$ ème centile d'une loi gaussienne telle que :

$$\text{Prob}[X \leq z_a] = a \quad (22)$$

$X$  est une variable aléatoire gaussienne de moyenne nulle et d'écart type unité.

Si  $H_{est}$  est comprise entre les deux bornes Sup et Inf pour une valeur de  $H$  donnée, on accepte  $H1$  au niveau  $\alpha$ , l'estimateur est alors considéré comme non biaisé. Sinon  $H0$  est acceptée.

### 2.4.1. variance et borne de Cramer-Rao.

Pour un estimateur non biaisé  $\hat{H}$  du paramètre  $H$ , sous certaines conditions de régularités [27], sa variance est toujours supérieure ou égale à la borne de Cramer-Rao (BCR) qui est unique pour un paramètre donné :

$$\text{Var}(\hat{H}) \geq \frac{1}{E \left[ \left( \frac{\partial \text{Ln}(P(x; H))}{\partial H} \right)^2 \right]} = \text{BCR}(H) \quad (23)$$

Le calcul de cette borne est reporté en annexe pour l'estimation du paramètre  $H$  du fBm. Avant de présenter les résultats, il convient de noter que plusieurs auteurs ont aussi calculé d'autres bornes de Cramer-Rao pour des processus proches du fBm. Pour les processus à longue mémoire, une borne de Cramer-Rao asymptotique est calculée dans [1][18]. Pour les processus en  $\frac{1}{f}$ ,

un calcul de borne de Cramer-Rao peut-être conduit *via* une décomposition en ondelettes en supposant que les signaux de détails sont indépendants à différentes échelles [28]. Enfin, une étude portant sur l'estimation conjointe des deux paramètres de processus à longue dépendance reprend de manière comparative plusieurs de ces calculs [29].

Le tableau suivant donne les valeurs de la racine carrée de la BCR du paramètre  $H$  du fBm, quantité homogène à  $H$ , en fonction de  $H$  et  $N$ . Les valeurs de  $N$  varient de 32 à 1024 par puissance entière de 2.

**Tableau 1.** – racine carrée de la borne de Cramer-Rao pour différentes valeurs de  $H$  et différentes longueurs de vecteur  $N$ .

H/N	32	64	128	256	512	1024
0.1	0.072	0.048	0.033	0.023	0.016	0.011
0.2	0.094	0.063	0.043	0.030	0.021	0.015
0.3	0.105	0.072	0.049	0.034	0.024	0.017
0.4	0.111	0.076	0.053	0.037	0.026	0.018
0.5	0.113	0.079	0.055	0.039	0.027	0.019
0.6	0.113	0.079	0.056	0.039	0.028	0.019
0.7	0.108	0.078	0.056	0.040	0.028	0.020
0.8	0.099	0.074	0.054	0.039	0.028	0.020
0.9	0.077	0.062	0.049	0.037	0.027	0.020

La variance de chaque estimateur sera donc comparée à cette borne inférieure. De la même manière que pour la moyenne, il n'est possible d'avoir accès qu'à une estimée de  $\text{Var}(\hat{H})$ . La qualification de la variance sera donc obtenue par un test d'hypothèse [26].

Soit  $S^2$  la variance empirique non biaisée d'un estimateur du paramètre  $H$ . Pour savoir si la variance d'estimation est égale à la BCR, nous allons effectuer un test d'hypothèse unilatéral au niveau  $\alpha$  avec l'hypothèse  $H_1$  telle que  $E(S^2) = \text{BCR}(H)$ . L'hypothèse alternative  $H_0$  sera :  $E(S^2) > \text{BCR}(H)$ .  $H_1$  est acceptée si  $S^2$  est inférieure à une limite  $\varepsilon^2$  :

$$S^2 \leq \chi_{NR-1, 1-\alpha}^2 \frac{\text{BCR}(H)}{NR-1} = \varepsilon^2 \quad (24)$$

$\chi_{D,a}^2$  est la  $100 \times a^{\text{ième}}$  centile pour la loi du chi-deux à  $D$  degrés de liberté.

Si l'écart-type constaté est inférieur à  $\varepsilon$  pour une valeur de  $H$  donnée, on conclura que la variance atteint la borne de Cramer-Rao au niveau  $\alpha$ .

Si les hypothèses sur la moyenne et sur la variance sont acceptées, on pourra dire que l'estimateur est efficace au sens de Fisher [27].

Il convient finalement de choisir  $NR$ , le nombre de réalisations.  $NR$  est relié à la puissance du test notée  $1 - \beta$ . Concernant le test sur la moyenne et en supposant que les densités de probabilité des estimateurs sont gaussiennes, on trouve [30] :

$$NR = \left[ \frac{S(z_{1-\alpha/2} + z_{1-\beta})}{d} \right]^2 \quad (25)$$

$d$  est la distance entre l'hypothèse  $H_0$  et  $H_1$ . Nous choisissons  $d = 0.01$ , soit le dixième de la différence entre les valeurs de  $H$  testées. Nous choisissons également  $\alpha = \beta = 0.01$  ce qui revient à faire un test avec un niveau de 1 % et une puissance de 99 %. On en déduit alors que  $Z_{1-\alpha/2} = 2.56$  et  $Z_{1-\beta} = 2.33$ .

Pour déterminer  $NR$ , il faudrait connaître  $S$  avant de réaliser les expériences, ce qui n'est pas le cas ici. Sachant que les tests seront appliqués pour  $N = 1024$ , on peut remplacer  $S$  par la racine carrée de la borne de Cramer-Rao en se mettant dans le cas le plus défavorable ce qui conduit à  $S = 0.02$  (tableau 1, dernière colonne). En prenant en compte toutes ces valeurs, on trouve  $NR = 98$ . Nous prenons la valeur ronde supérieure  $NR = 100$ .

Maintenant que le cadre de l'étude est précisé, les divers estimateurs du paramètre  $H$  vont être étudiés.

### 3. estimateurs du paramètre $H$

Il existe de nombreuses méthodes permettant d'analyser des signaux de type fBm. Elles peuvent être classées en quatre catégories :

- les méthodes basées sur des propriétés de la géométrie fractale,
- les méthodes temporelles,
- les méthodes fréquentielles,
- les méthodes basées sur une analyse multi-échelles.

Afin d'étudier le biais et la variance de chaque estimateur, ceux-ci seront mis en œuvre sur des ensembles de  $NR = 100$  réalisations indépendantes générées par la méthode de Cholesky pour des longueurs de signaux de 32 à 1024 par puissance de 2. Ces signaux auront des valeurs de  $H$  qui varieront de 0.1 à 0.9 par pas de 0.1. La moyenne des estimations  $H_{\text{est}}$  permettra d'évaluer le biais relativement à la valeur théorique de synthèse. De même, l'écart-type empirique  $S$  de l'estimateur sera aussi évalué. La connaissance de leur moyenne et variance donne une bonne idée de la répartition statistique des mesures. Dans un premier temps, les estimateurs seront analysés qualitativement et nous dégagerons les grandes tendances. Ensuite, une analyse quantitative sera effectuée pour  $N = 1024$  en appliquant les tests d'hypothèse sur la moyenne et sur la variance.

Enfin, chaque méthode sera mise en œuvre de manière standard, en général celle qui a été présentée dans l'article de référence.

#### 3.1. estimation par mesure de paramètres géométriques

Les propriétés de base de la géométrie fractale peuvent être utilisées pour mesurer le paramètre  $H$ . On considère alors le signal fBm comme une forme géométrique sur laquelle on mesure en général une longueur ou une surface. Plusieurs méthodes reprenant les propriétés de la géométrie fractale ont été décrites [31][32]. Elles sont toutes basées sur le même schéma : mesure d'un paramètre géométrique à des résolutions différentes, puis régression linéaire du logarithme du paramètre en fonction du logarithme de l'échelle. La principale difficulté d'application de

ces méthodes est que de nombreuses approximations doivent être faites lors du passage d'une courbe continue à une suite de points discrets.

### 3.1.1. estimation par la méthode des boîtes (EMB) [33]

Cette méthode très connue consiste à découper l'espace occupé par cette forme en carrés, ou boîtes, de côté  $\eta$ . Ensuite, il s'agit de compter le nombre de boîtes  $NB$  contenant une partie de la forme. Elle peut aussi s'appliquer à un signal considéré comme une forme géométrique. Pour un signal fBm on a alors la relation suivante :

$$NB(\eta) \propto \eta^{H-2} \quad (26)$$

Le tracé de  $NB$  en fonction de  $\eta$  dans une échelle log-log est une droite de pente  $H - 2$ . Une régression linéaire permet d'obtenir  $H$ .

Pour l'implémentation de cette méthode, et afin d'éviter les phénomènes de plages décrit dans [31], nous faisons varier la taille des boîtes par puissance de 2 [33]. De plus, la taille verticale des boîtes est choisie pour correspondre à la dynamique totale du signal avec un même nombre de découpages horizontaux et verticaux.

Comme le montre le tableau 2, le biais est important pour les faibles tailles de signaux ainsi que pour les valeurs extrêmes de  $H$ . La variance est importante et n'est que peu réduite lorsque la taille des signaux augmente.

Cette méthode s'applique difficilement aux signaux échantillonnés. Si on considère la forme définie par les points d'échantillonnage, plusieurs limitations apparaissent. Pour des boîtes de côté unité,  $NB(1)$  est égal au nombre d'échantillons  $N$  quel que soit le signal (point  $B$  de la figure 2). Pour une taille  $\eta = N$ , le nombre de boîtes est de 1 (point  $A$  de la figure 2). Ceci veut dire

que les deux points extrêmes de mesure sont fixés. De plus, le nombre de boîtes ne peut évoluer qu'entre les deux limites théoriques :  $NB(\eta) = (1/\eta)^2$  pour  $H \rightarrow 0$  (signaux très agités) et  $NB(\eta) = 1/\eta$  pour  $H \rightarrow 1$  (signaux très lisses). Finalement, il existe une limite supérieure horizontale  $NB(\eta) = N$ , car les intersections entre les points d'échantillonnage et le nombre de boîtes sont au plus égales au nombre d'échantillons.

En résumé, la courbe d'évolution de  $NB(\eta)$  est donc contrainte dans le triangle  $ABC$  et doit passer par les points  $A$  et  $B$  comme le montre la figure 2.

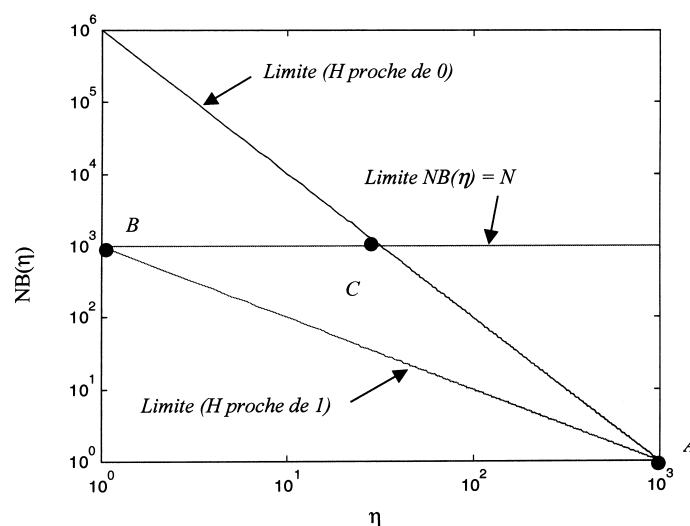


Figure 2. – limites supérieures et inférieure ainsi que les points remarquables  $A$  et  $B$  (voir texte) pour le nombre de boîtes  $NB(\eta)$  en fonction de la taille des boîtes  $\eta$  en échelle log-log pour un signal de 1024 échantillons. La courbe  $NB(\eta)$  est contrainte dans le triangle  $ABC$  et doit passer par les points  $A$  et  $B$ .

Tableau 2. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode des boîtes (EMB).

H/N	32	64	128	256	512	1024
0.1	0.409±0.212	0.289±0.179	0.249±0.148	0.206±0.160	0.173±0.142	0.137±0.143
0.2	0.448±0.196	0.383±0.197	0.336±0.177	0.312±0.174	0.260±0.164	0.257±0.168
0.3	0.491±0.231	0.403±0.202	0.425±0.210	0.400±0.200	0.384±0.224	0.388±0.231
0.4	0.531±0.221	0.510±0.195	0.483±0.197	0.495±0.191	0.460±0.226	0.460±0.224
0.5	0.628±0.221	0.503±0.212	0.532±0.204	0.519±0.232	0.560±0.230	0.530±0.223
0.6	0.629±0.223	0.545±0.229	0.606±0.248	0.551±0.245	0.623±0.225	0.579±0.251
0.7	0.684±0.232	0.648±0.263	0.672±0.234	0.608±0.231	0.655±0.239	0.644±0.222
0.8	0.749±0.255	0.693±0.220	0.671±0.244	0.666±0.217	0.688±0.231	0.686±0.224
0.9	0.728±0.205	0.737±0.206	0.667±0.220	0.693±0.227	0.729±0.190	0.736±0.190

### 3.1.2. estimation par morphologie mathématique (EMM) [32]

Comme pour la méthode des boîtes, cette estimation a d'abord été proposée pour mesurer la dimension fractale de formes géométriques [34][35], puis de celle de graphe [31].

Une mise en œuvre de cette méthode sur des signaux numériques a été présentée par Maragos *et al.* [32]. Elle consiste à éroder et dilater le signal par un élément structurant de taille  $\varepsilon$ . Ensuite, l'aire  $A$  comprise entre les signaux érodés et dilatés est calculée. Il est possible de montrer que pour un processus fractal, on a la relation suivante [32] :

$$H = \lim_{\varepsilon \rightarrow 0} \frac{\log(A)}{\log(\varepsilon)} \quad (27)$$

Le tracé en échelle log-log de  $A$  en fonction de la taille de l'élément structurant peut être considéré comme une droite de pente  $H$  ce qui permet son estimation par régression linéaire.

Même si la relation (27) s'applique à toute forme d'éléments structurant convexes, le choix souvent empirique de sa taille et de sa forme peut modifier les résultats. Afin de ne pas ajouter de variables supplémentaires à cette implémentation, nous avons choisi un élément structurant horizontal plat. Sa taille peut varier de 2 échantillons jusqu'à une taille maximale égale à la moitié de la longueur du signal. Le tableau suivant reporte les résultats. Le biais est relativement faible pour les valeurs de  $H$  proches de 1. L'erreur devient importante au fur et à mesure que  $H$  diminue. La variance est proche de la borne de Cramer-Rao pour les faibles tailles de signaux mais n'est que peu réduite lorsque la taille augmente. Cette méthode souffre des mêmes problèmes que l'EMB relativement à l'échantillonnage.

## 3.2. méthodes temporelles

La méthode présentée, basée sur les propriétés statistiques du fBm, est la seule de toutes celles étudiées dans ce travail qui n'utilise pas une régression dans une échelle log-log. Elle nécessite par contre de rechercher le maximum d'une fonction.

### 3.2.1. estimation par maximum de vraisemblance (EMV) [13]

L'estimation par maximum de vraisemblance est une méthode fréquemment utilisée en traitement du signal du fait de ses bonnes propriétés [27]. Concernant la mesure du paramètre  $H$ , Dahlhaus a montré que cette technique fournit des estimées asymptotiquement efficaces [1].

Dans la pratique, cette estimation est effectuée sur les incréments du signal [13][14]. Par la suite, nous allons noter  $\mathbf{x}$  le vecteur incréments des observations supposé être un vecteur fGn. Ce vecteur est stationnaire, gaussien, centré, et sa fonction de vraisemblance paramétrée par  $H$  et  $\sigma^2$  s'écrit :

$$\text{Prob}(\mathbf{x}; H, \sigma^2) = \frac{1}{(2\pi)^{(\frac{N}{2})} |\mathbf{R}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) \quad (28)$$

$\mathbf{R}$  est la matrice d'autocorrélation de  $\mathbf{x}$  qui dépend de  $H$  et de  $\sigma^2$  comme précisé dans (14) avec  $[\mathbf{R}]_{i,j} = r[|i-j|]$ . Il est plus simple de rechercher le maximum du logarithme de la fonction de vraisemblance noté  $LVF$  où les termes constants ne sont pas pris en compte :

$$LVF(\mathbf{x}; H, \sigma^2) = -\text{Ln}|\mathbf{R}| - \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x} \quad (29)$$

On décompose  $\mathbf{R}$  comme  $\sigma^2 \mathbf{R}'$ , puis on calcule la dérivée du LFV par rapport à  $\sigma^2$ . La valeur de  $\sigma^2$  pour laquelle l'expres-

Tableau 3. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  par morphologie mathématique (EMM).

H/N	32	64	128	256	512	1024
0.1	0.481±0.071	0.438±0.047	0.388±0.036	0.345±0.024	0.307±0.022	0.278±0.019
0.2	0.531±0.074	0.495±0.053	0.449±0.041	0.412±0.029	0.376±0.028	0.348±0.023
0.3	0.581±0.077	0.550±0.055	0.510±0.046	0.478±0.033	0.444±0.032	0.419±0.026
0.4	0.632±0.081	0.607±0.058	0.573±0.051	0.545±0.038	0.514±0.037	0.492±0.030
0.5	0.684±0.082	0.666±0.063	0.636±0.057	0.614±0.043	0.587±0.042	0.568±0.037
0.6	0.737±0.084	0.725±0.069	0.700±0.062	0.684±0.048	0.661±0.048	0.647±0.045
0.7	0.790±0.084	0.785±0.073	0.765±0.065	0.756±0.054	0.737±0.054	0.727±0.055
0.8	0.843±0.083	0.844±0.075	0.830±0.067	0.828±0.059	0.813±0.060	0.807±0.063
0.9	0.890±0.080	0.904±0.071	0.896±0.069	0.897±0.057	0.890±0.062	0.885±0.067



sion obtenue est nulle est reportée dans (29). La valeur de  $H$  par maximum de vraisemblance,  $\hat{H}_{MV}$ , est donc la suivante :

$$\hat{H}_{MV} = \underset{0 < H < 1}{\text{Max}} \left\{ -\text{Ln}|\mathbf{R}'| - N \text{Ln} \left( \frac{\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}}{N} \right) \right\} \quad (30)$$

La recherche du maximum est un problème d'optimisation sans contrainte simplifié par le fait que  $H$  est dans l'intervalle  $]0, 1[$ . Le calcul du déterminant et de l'inverse de la matrice d'autocorrélation peut être fait de manière efficace par l'algorithme de Levinson [36] car cette matrice est Toeplitz. Les résultats sont présentés tableau 4.

Les résultats pour le biais et la variance sont excellents quels que soient  $N$  et  $H$ .

Comme pour la méthode de synthèse de Cholesky, l' $EMV$  est fondée sur la fonction d'autocorrélation du fGn discret qui est égale à celle du fGn continu aux instants d'échantillonnage (voir

remarque du paragraphe 2.2). Ceci explique principalement le bon fonctionnement de cette méthode.

### 3.3. méthodes fréquentielles

Ces méthodes ont pour caractéristique commune de s'appuyer sur la forme globalement en  $\frac{1}{f}$  des spectres du fBm et de ses incréments. L'étape finale est une régression en échelle log-log qui permet d'estimer la valeur du paramètre  $H$ .

#### 3.3.1. estimation par analyse spectrale sur le fBm (ASB) [37]

Nous avons vu que le spectre moyen du fBm, défini à partir du spectre de Wigner-Ville, suit une loi proportionnelle à  $|f|^{-1-2H}$ .

Tableau 4. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  par maximum de vraisemblance ( $EMV$ ).

H/N	32	64	128	256	512	1024
0.1	0.099±0.074	0.097±0.045	0.098±0.033	0.098±0.023	0.098±0.016	0.100±0.011
0.2	0.193±0.098	0.195±0.059	0.197±0.043	0.197±0.031	0.197±0.022	0.200±0.014
0.3	0.290±0.110	0.294±0.067	0.296±0.050	0.296±0.035	0.296±0.025	0.300±0.017
0.4	0.388±0.112	0.393±0.071	0.395±0.053	0.396±0.038	0.395±0.027	0.399±0.018
0.5	0.488±0.110	0.492±0.074	0.494±0.056	0.495±0.039	0.495±0.028	0.499±0.019
0.6	0.586±0.106	0.591±0.075	0.594±0.057	0.595±0.040	0.595±0.029	0.600±0.019
0.7	0.682±0.102	0.689±0.076	0.693±0.058	0.695±0.040	0.694±0.030	0.699±0.019
0.8	0.773±0.096	0.784±0.074	0.791±0.058	0.794±0.040	0.794±0.030	0.799±0.020
0.9	0.860±0.082	0.877±0.063	0.885±0.053	0.892±0.037	0.893±0.029	0.899±0.019

Tableau 5. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode d'analyse spectrale sur le fBm (ASB).

H/N	32	64	128	256	512	1024
0.1	-.198±0.196	-.171±0.139	-.156±0.102	-.151±0.069	-.144±0.044	-.140±0.030
0.2	-.006±0.193	0.012±0.138	0.029±0.100	0.028±0.072	0.031±0.048	0.035±0.034
0.3	0.145±0.192	0.163±0.131	0.176±0.100	0.176±0.075	0.175±0.050	0.181±0.042
0.4	0.260±0.189	0.294±0.134	0.301±0.090	0.300±0.071	0.300±0.048	0.308±0.042
0.5	0.344±0.179	0.381±0.129	0.391±0.076	0.394±0.057	0.392±0.041	0.398±0.028
0.6	0.406±0.192	0.434±0.131	0.438±0.074	0.441±0.057	0.438±0.053	0.435±0.044
0.7	0.447±0.194	0.454±0.134	0.463±0.098	0.454±0.080	0.450±0.080	0.436±0.063
0.8	0.464±0.193	0.475±0.159	0.462±0.114	0.453±0.106	0.448±0.096	0.435±0.087
0.9	0.457±0.199	0.439±0.138	0.430±0.095	0.432±0.105	0.426±0.078	0.417±0.068

Dans une représentation en échelles log-log, cette quantité en fonction de la fréquence est une droite de pente  $-1 - 2H$ . Une régression linéaire permet alors de mesurer le paramètre  $H$ . En ce qui concerne l'implémentation, le spectre moyen est estimé par le carré du module de la transformée de Fourier obtenu par l'algorithme FFT. La régression linéaire est effectuée sur toute la plage de fréquence.

Le tableau 5 montre que le paramètre  $H$  est fortement sous-estimé. Ceci provient principalement de la non-stationnarité du fBm qui provoque une forte discontinuité entre le début et la fin du signal ce qui rajoute des hautes fréquences. La variance, quant à elle, est bien plus importante que la borne de Cramer-Rao.

Globalement, cette méthode est très peu performante.

### 3.3.2. estimation spectrale sur le fGn (ASG) [37]

La DSP du fGn suit une loi en  $|f|^{1-2H}$ . Comme pour l'analyse sur le fBm, nous réalisons directement la régression linéaire sur toute la plage fréquentielle de la DSP. Le tableau 6 présente les résultats. Nous constatons que ASG est biaisée pour les valeurs de  $H$  différentes de 0.5. La variance, quant à elle, reste largement plus importante que la borne de Cramer-Rao, surtout pour les faibles tailles de signaux.

## 3.4. méthodes basées sur une approche multi-échelles

Ces méthodes sont basées sur une approche temps-échelle qui s'adapte naturellement au caractère en  $\frac{1}{f}$  du processus fBm. Les ondelettes en sont l'illustration la plus évidente. D'autres méthodes antérieures à cette dernière peuvent maintenant être vues sous l'aspect multi-échelles [17].

### 3.4.1. estimation par ondelettes (EOND) [38]

$H$  peut être estimé directement en utilisant (8) de la manière suivante :

$$\text{Log}(\text{Var}(d_j[i])) = -(2H + 1)j + \text{constante} \quad (31)$$

Si le nombre de moments nuls de l'ondelette est supérieur ou égal à 2, les signaux de détails  $d_j$  sont en pratique peu corrélés [17], ce qui est préférable si  $H$  est estimé à partir de (31) afin de réduire la variance d'estimation. N'importe quelle ondelette de Daubechies [39] à plus de deux moments nuls peut convenir.

Plusieurs éléments doivent être pris en compte pour une mise en œuvre efficace de cette méthode d'analyse. Tout d'abord, les effets de bords dus au filtrage peuvent conduire à des valeurs erronées. Les échantillons ainsi pollués sont ignorés dans les calculs. Afin de réduire ces effets de bord, nous avons choisi une ondelette la plus petite possible, c'est-à-dire une ondelette de Daubechies à 4 coefficients. De plus, et c'est la principale cause d'erreur, la relation (31) n'est valable qu'en continu [17]. L'échantillonnage introduit un biais, surtout dans les premières étapes de décomposition, la dimension fractale locale (entre deux échelles) étant plus grande que celle du signal continu. Cette différence s'estompe pour les petites échelles. Ceci conduit à une sous-estimation du paramètre  $H$ , cette différence étant plus petite pour les  $H$  élevés où cet effet est moins sensible. Nous avons mis en place une correction qui améliore très sensiblement la qualité des résultats. Elle consiste à pondérer la valeur des variances par un coefficient qui dépend de l'ondelette utilisée, de l'échelle et de la valeur de  $H$  en accord avec les formules suivantes [17] :

$$\text{Var}_{\text{corrigé}}(D_j) = \frac{\text{Var}(D_j)}{\frac{C_0(j)}{2^{(2H+1)j}}} \quad (32)$$

Tableau 6. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode d'analyse spectrale sur le fGn (ASG).

H/N	32	64	128	256	512	1024
0.1	-.057±0.212	-.025±0.121	-.049±0.085	-.040±0.064	-.042±0.041	-.039±0.027
0.2	0.098±0.224	0.131±0.124	0.116±0.088	0.123±0.062	0.123±0.040	0.125±0.027
0.3	0.246±0.212	0.265±0.130	0.255±0.088	0.260±0.062	0.259±0.040	0.261±0.027
0.4	0.379±0.210	0.388±0.132	0.382±0.089	0.383±0.062	0.383±0.039	0.384±0.028
0.5	0.501±0.212	0.505±0.135	0.501±0.089	0.503±0.062	0.500±0.029	0.500±0.028
0.6	0.621±0.213	0.618±0.137	0.618±0.088	0.613±0.062	0.612±0.039	0.613±0.027
0.7	0.734±0.207	0.738±0.140	0.732±0.090	0.723±0.063	0.722±0.039	0.722±0.027
0.8	0.848±0.198	0.836±0.145	0.844±0.091	0.831±0.064	0.830±0.040	0.831±0.026
0.9	0.961±0.192	0.944±0.145	0.954±0.099	0.939±0.067	0.938±0.041	0.938±0.027

où

$$L(k) \propto k^{H-2} \quad (35)$$

$$C_0(j) = - \sum_{n=-\infty}^{\infty} \gamma_g[-n] D_{j-1}[n] \quad (33)$$

et

$$D_j[n] = \sum_{n'=-\infty}^{\infty} \gamma_h[2n-n'] D_{j-1}[n'] \text{ si } j \geq 1 \text{ avec} \\ D_0[n] = |n|^{2H} \quad (34)$$

$\gamma_h$  et  $\gamma_g$  sont les fonctions d'autocorrélation des réponses impulsionnelles des filtres  $H$  et  $G$  de la décomposition en ondelettes. Il est clair qu'une estimation précise par cette méthode ne peut se faire que de manière itérative, car la valeur de  $H$  est nécessaire pour réaliser les corrections. La convergence est très rapide, 3 ou 4 itérations suffisent. Enfin, pour améliorer la variance, il est souhaitable de faire une régression linéaire pondérée [40]. Les résultats sont présentés dans le tableau suivant. On constate que cette méthode d'estimation donne de bons résultats pour la moyenne des estimées. La variance est plus importante que la borne de Cramer-Rao. Des corrections supplémentaires peuvent être appliquées lors de la régression [41], mais elles n'ont que peu d'incidence sur la variance qui reste assez supérieure à la borne de Cramer-Rao.

### 3.4.2. estimation par la méthode de Burlaga-Klein (EBK) [42]

La longueur d'une forme fractale mesurée par un étalon de longueur  $k$  varie comme [43] :

Burlaga et Klein proposent une méthode pour l'estimation de la longueur d'un signal en fonction d'un étalon [42]. Ils suggèrent de le lisser en le moyennant localement sur des intervalles de temps jointifs de longueur  $k$ . La valeur absolue de ces incréments est proportionnelle à la longueur de la courbe lissée aux premières étapes du traitement. Plus précisément, Burlaga et Klein définissent la longueur d'une fonction  $f$  comme :

$$L_{BK}(k) = \frac{1}{\text{Ent} \left[ \frac{N}{k} \right] k^3} \sum_{i=0}^{\text{Ent} \left[ \frac{N}{k} \right] - 1} \left| \sum_{n=0}^{k-1} (f[n+ik] - f[n+(i+1)k]) \right| \quad (36)$$

$\text{Ent}[\cdot]$  désigne l'opérateur partie entière. Le tracé de  $\log(L_{BK}(k))$  en fonction de  $\log(k)$  est une droite de pente  $H - 2$ . Une régression linéaire permet alors de déterminer  $H$ .

Nous avons choisi une longueur maximale pour le lissage de 10 échantillons. La régression est effectuée sur une longueur identique.

On voit dans le tableau 7 que l'estimation de  $H$  est toujours biaisée inférieurement, résultat qui a déjà été signalé dans [11][17]. La variance, quant à elle, est très supérieure à la borne de Cramer-Rao.

Cette méthode peut aussi être envisagée sous l'aspect multi-échelles : la somme la plus interne correspondant au signal de forme à l'étape  $j$  et la différence au signal de détail à l'étape  $j + 1$ . La décomposition peut alors se comparer à une analyse par les ondelettes de Haar, la seule différence étant liée à l'usage de la valeur absolue au lieu du carré.

Tableau 7. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode des ondelettes (EOND).

H/N	32	64	128	256	512	1024
0.1	0.003±0.591	0.053±0.233	0.071±0.128	0.079±0.075	0.090±0.043	0.094±0.029
0.2	0.138±0.601	0.156±0.244	0.177±0.125	0.187±0.073	0.189±0.047	0.193±0.030
0.3	0.236±0.565	0.274±0.236	0.275±0.127	0.286±0.074	0.291±0.047	0.295±0.030
0.4	0.379±0.587	0.353±0.245	0.372±0.131	0.385±0.077	0.390±0.050	0.395±0.033
0.5	0.416±0.600	0.471±0.250	0.479±0.134	0.485±0.077	0.492±0.048	0.495±0.031
0.6	0.558±0.557	0.573±0.239	0.582±0.129	0.587±0.080	0.593±0.048	0.594±0.033
0.7	0.648±0.578	0.657±0.250	0.686±0.131	0.692±0.075	0.693±0.049	0.695±0.034
0.8	0.753±0.571	0.759±0.238	0.778±0.135	0.787±0.083	0.794±0.052	0.796±0.034
0.9	0.848±0.542	0.875±0.244	0.878±0.133	0.891±0.080	0.893±0.050	0.896±0.035

Tableau 8. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode de Burlaga-Klein (EBK).

H/N	32	64	128	256	512	1024
0.1	−.357±0.234	−.263±0.143	−.215±0.099	−.195±0.075	−.189±0.047	−.107±0.045
0.2	−.171±0.221	−.073±0.146	−.019±0.099	−.003±0.073	0.002±0.049	0.073±0.046
0.3	−.012±0.207	0.079±0.145	0.134±0.100	0.150±0.071	0.154±0.050	0.212±0.046
0.4	0.129±0.205	0.212±0.149	0.268±0.105	0.285±0.069	0.288±0.041	0.336±0.046
0.5	0.257±0.202	0.332±0.157	0.391±0.110	0.410±0.069	0.413±0.051	0.452±0.046
0.6	0.378±0.194	0.446±0.162	0.505±0.113	0.528±0.069	0.531±0.051	0.563±0.046
0.7	0.492±0.187	0.558±0.163	0.614±0.112	0.642±0.070	0.646±0.054	0.671±0.048
0.8	0.600±0.186	0.663±0.153	0.718±0.110	0.751±0.074	0.756±0.058	0.774±0.054
0.9	0.712±0.175	0.768±0.150	0.819±0.105	0.852±0.070	0.862±0.059	0.872±0.055

### 3.4.3. estimation par la méthode de la variance (EVAR) [6]

L'égalité statistique (15) permet d'écrire :

$$\text{Var}[G_k(i)] = k^{2H} \text{Var}[G_1(i)] \quad (37)$$

Le tracé de  $\text{Var}[G_k]$  en fonction de  $k$  dans une échelle log-log est une droite de pente  $2H$ . Une régression linéaire permet d'estimer  $H$ .

Cette approche comporte de nombreuses variantes [6][44], et plus récemment [45]. Le seul paramètre pouvant être choisi est la longueur de la régression linéaire : les valeurs de la variance calculées pour un  $k$  grand pouvant être moins précise du fait du faible nombre d'échantillons. Le tableau suivant présente les résultats pour cette méthode.

Le biais est relativement faible quel que soit  $H$  mais la variance est assez loin de la borne de Cramer-Rao.

De la même manière que précédemment, la méthode de la variance peut être associée à une analyse multirésolutions où le filtre de forme serait un filtre passe tout et le filtre de détail le filtre  $(-1, 1)$ .

### Conclusions concernant les résultats préliminaires

Seules EVAR, EOND et EMV donnent de bons résultats concernant le biais. Par contre, seule EMV a une variance d'estimation proche de la borne de Cramer-Rao.

Pour préciser cette analyse préliminaire, une étude quantitative va être menée. Des tests d'hypothèses sur la moyenne et sur la variance permettront de connaître si une méthode d'analyse est

Tableau 9. – valeurs de la moyenne des estimées de  $H \pm$  l'écart-type constaté en fonction de  $H$  et  $N$  pour la méthode de la variance (EVAR).

H/N	32	64	128	256	512	1024
0.1	0.074±0.086	0.095±0.061	0.097±0.057	0.100±0.036	0.099±0.027	0.101±0.020
0.2	0.147±0.121	0.187±0.088	0.188±0.086	0.199±0.055	0.198±0.041	0.201±0.030
0.3	0.220±0.149	0.279±0.109	0.279±0.106	0.297±0.068	0.296±0.050	0.300±0.037
0.4	0.295±0.170	0.370±0.125	0.371±0.118	0.395±0.076	0.394±0.056	0.399±0.041
0.5	0.376±0.186	0.461±0.139	0.463±0.123	0.423±0.057	0.492±0.060	0.498±0.044
0.6	0.461±0.200	0.551±0.149	0.554±0.124	0.589±0.084	0.589±0.061	0.596±0.047
0.7	0.551±0.213	0.639±0.154	0.642±0.125	0.682±0.087	0.685±0.063	0.692±0.050
0.8	0.645±0.218	0.725±0.154	0.727±0.126	0.770±0.089	0.776±0.064	0.785±0.053
0.9	0.751±0.205	0.814±0.141	0.813±0.128	0.854±0.077	0.863±0.063	0.873±0.053

non biaisée et si la variance atteint la borne de Cramer-Rao. Ces tests seront menés pour  $N = 1024$ . Il est en effet légitime de considérer que si une méthode n'est pas efficace pour des signaux de grande taille, elle ne le sera pas non plus pour des données comportant peu d'échantillons.

### 3.5. comparaison des estimateurs pour $N = 1024$

Nous avons reporté les résultats de la moyenne et de l'écart-type empiriques dans le tableau 10 pour  $N = 1024$  et  $NR = 100$ . Ces résultats sont calculés pour des valeurs de  $H$  allant de 0.1 à 0.9 par pas de 0.1. Nous avons adopté les conventions suivantes : si l'estimateur est non biaisé, la valeur sera grisée. Pour ce faire, il est nécessaire de calculer les deux valeurs Inf et Sup du test sur la moyenne telles que décrites dans le paragraphe 2.4 équation (21). L'écart-type mesuré  $S$  sera grisé dans le tableau 10 si la variance d'estimation atteint la borne de Cramer-Rao, c'est-à-dire si l'écart-type empirique satisfait à l'inégalité (24).

Nous pouvons séparer les méthodes suivant quatre classes en fonction de la qualité des résultats :

- celles qui donnent de très mauvais résultats : EMM, ASB, et EBK.
- celles qui donnent de bons résultats pour la moyenne lorsque  $H$  est proche de 0.5 : EMB et ASG. Pour cette valeur de  $H$ , le fBm est le mouvement brownien ordinaire et le fGn le bruit blanc.
- celles qui donnent de bons résultats uniquement pour la moyenne : EOND et EVAR. On peut raisonnablement penser que si l'on testait EOND pour des signaux plus longs, elle donnerait aussi de bons résultats pour la variance. En effet, il a été montré que sous des hypothèses de corrélation nulle des signaux de détails et pour le régime asymptotique, la variance de cet estimateur atteint la borne de Cramer-Rao [5].
- celle qui donne de très bons résultats aussi bien pour la moyenne que pour la variance quelles que soient les valeurs de  $H$  : EMV.

Seule l'EMV est un estimateur efficace pour le fBm pour des signaux de 1024 échantillons. En faisant la même étude (tests

Tableau 10. – moyennes  $\pm$  écart-types constatés des 8 différents estimateurs du paramètre  $H$  sur 100 signaux  $CHO$  de 1024 échantillons. Les valeurs grisées sont celles pour lesquelles respectivement le biais est nul ou la variance atteint la borne de Cramer-Rao.

H	EMB	EMM	EMV	ASB	ASG	EOND	EBK	EVAR
0.1	0.137	0.278	0.100	-.140	-.039	0.094	-.107	0.101
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.143	0.019	0.011	0.030	0.027	0.029	0.045	0.020
0.2	0.257	0.348	0.200	0.035	0.125	0.193	0.073	0.201
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.168	0.023	0.014	0.034	0.027	0.030	0.046	0.030
0.3	0.388	0.419	0.300	0.181	0.261	0.295	0.212	0.200
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.231	0.026	0.017	0.042	0.027	0.030	0.046	0.037
0.4	0.460	0.492	0.399	0.308	0.384	0.395	0.336	0.399
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.224	0.030	0.018	0.042	0.028	0.033	0.046	0.041
0.5	0.530	0.568	0.499	0.398	0.500	0.495	0.452	0.498
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.223	0.037	0.019	0.028	0.028	0.031	0.046	0.044
0.6	0.579	0.647	0.600	0.435	0.613	0.594	0.563	0.596
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.251	0.045	0.019	0.044	0.027	0.033	0.046	0.047
0.7	0.644	0.727	0.699	0.436	0.722	0.695	0.671	0.692
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.222	0.055	0.019	0.063	0.027	0.034	0.048	0.050
0.8	0.686	0.807	0.799	0.435	0.831	0.796	0.774	0.785
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.224	0.063	0.020	0.087	0.026	0.034	0.054	0.053
0.9	0.736	0.885	0.899	0.417	0.938	0.896	0.872	0.873
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$
	0.190	0.067	0.019	0.068	0.027	0.035	0.055	0.053

sur la moyenne et la variance) pour l'EMV sur toutes les valeurs de  $N$  étudiées (de 32 à 1024 par puissance de 2), on peut constater que l'estimateur par maximum de vraisemblance est efficace. Ce résultat expérimental complète les résultats théoriques démontrés uniquement dans le cas asymptotique [1].

En conclusion, jusqu'à des signaux de très faible longueur (quelques dizaines d'échantillons), EMV est en pratique un estimateur efficace du paramètre  $H$  du mouvement brownien fractionnaire.

## 4. étude de l'EMV du paramètre $H$

Les résultats expérimentaux précédents montrent que l'estimation par maximum de vraisemblance est efficace. Ceci laisse à penser qu'EMV devrait être systématiquement employée pour mesurer le paramètre  $H$  de données réelles. Toutefois, ces données peuvent présenter certaines particularités qui pourraient mettre en défaut cet estimateur : nombre important d'échantillons, signaux bruités. Le but de ce chapitre est de montrer comment l'EMV du paramètre  $H$  se comporte face à ces différentes situations.

### 4.1. Signaux de grande longueur : approximation de Whittle

Dans le domaine temporel, le logarithme de la fonction de vraisemblance (LFV) du vecteur observation  $\mathbf{x}$  supposé être fGn est la suivante, aux termes constants près :

$$\text{LFV}(\mathbf{x}; H, \sigma^2) = -\text{Ln}|\mathbf{R}| - \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x} \quad (38)$$

La complexité de l'algorithme EMV et la mémoire nécessaire sont en  $O(N^2)$ . Pour des signaux de taille importante ( $N$  supérieur à quelques milliers) et pour des calculs en format flottant, l'estimateur  $MV$  n'est pas techniquement portable sur la plupart des machines d'aujourd'hui. Dans ce cas, il existe une approximation dite de Whittle dans le domaine fréquentiel [46] qui est asymptotiquement efficace pour les incréments du fBm [1]. La fonction de vraisemblance dite de Whittle (LFVW) de  $\mathbf{x}$  peut alors s'exprimer de la manière suivante dans le cas discret [18] :

$$\text{LFVW}(\mathbf{x}; H, c) = \sum_{j=1}^{m^*} \left( -\text{Ln}T(j; H, c) - \frac{P(j)}{T(j; H, c)} \right) \quad (39)$$

- $m^*$  est la partie entière de  $(N - 1)/2$ ,
- $T$  est la densité spectrale de puissance théorique du processus fGn de paramètre  $H$ ,
- $P$  est le périodogramme du vecteur observation  $\mathbf{x}$ ,
- $c$ , qui est une constante de proportionnalité, permet d'ajuster la puissance du processus à celle du signal.

On pourra remarquer l'analogie entre ces deux dernières formules, l'une écrite dans le domaine temporel, et l'autre dans le domaine fréquentiel, ces deux expressions étant identiques uniquement pour des signaux de taille infinie.

Il faut donc trouver le maximum de cette fonction qui dépend de deux paramètres,  $H$  et  $c$ . On procède alors de la même manière que pour l'EMV classique. En décomposant  $T$  comme  $cT'$  et en prenant la dérivée par rapport à  $c$  égale à zéro on trouve :

$$c = \frac{\sum_{j=1}^{m^*} P(j)}{\sum_{j=1}^{m^*} T'(j; H)} \quad (40)$$

Dans une première étape, on calcule la constante  $c$ , puis en injectant cette valeur dans (39), on obtient la valeur de  $H$  estimée par maximum de vraisemblance de Whittle,  $\hat{H}_{MVW}$ , de la manière suivante :

$$\hat{H}_{MVW} = \text{Max}_{0 < H < 1} \left\{ \sum_{j=1}^{m^*} \left( -\text{Ln}cT'(j; H) - \frac{P(j)}{cT'(j; H)} \right) \right\} \quad (41)$$

$T'$  pourrait être calculée à partir de (11), ou ce qui est plus simple et plus efficace, comme la transformée de Fourier de (14). Pour plus de rapidité, on utilise l'algorithme FFT à condition que  $N$  soit une puissance entière de 2. Dans ce cas,  $m^*$  est alors égal à  $(N - 2)/2$ . La complexité est en  $O(N \log_2 N)$  et la taille mémoire nécessaire est en  $O(N)$ . On peut alors mesurer le paramètre  $H$  de signaux très longs.

On remarque (tableau 11) que les résultats sont très proches de ceux trouvés par le  $MV$  classique. En effet, cette méthode est aussi fondée sur la fonction d'autocorrélation du fGn discret ce qui explique principalement la raison du bon fonctionnement de cette méthode. On pourra toutefois noter que l'approximation de Whittle est moins précise pour  $H$  proche de zéro.

On comprend aisément que cette méthode suscite beaucoup d'intérêts eut égard à ses très bonnes performances, sa simplicité, et à sa souplesse.

### 4.2. signaux bruités

Le comportement de l'EMV en présence de bruit blanc additif a été étudié par Lundahl *et al.* sur des signaux synthétisés par la méthode de Cholesky [14]. En analysant les résultats, on constate que globalement cet estimateur est très sensible à cette perturbation, surtout pour les valeurs de  $H$  proches de 1. A titre d'exemple, sur 100 signaux de 100 échantillons chacun, pour une valeur de synthèse de  $H = 0.8$ , Lundahl *et al.* trouvent une

Tableau 11. – valeurs de la moyenne des estimées de  $H$  et de l'écart-type constaté en fonction de  $H$  et  $N$  par maximum de vraisemblance de Whittle (EMW).

H/N	32	64	128	256	512	1024
0.1	0.085±0.102	0.077±0.066	0.078±0.051	0.084±0.038	0.086±0.021	0.088±0.017
0.2	0.173±0.132	0.191±0.089	0.190±0.054	0.198±0.038	0.198±0.023	0.199±0.017
0.3	0.279±0.144	0.295±0.090	0.293±0.056	0.300±0.040	0.300±0.025	0.300±0.019
0.4	0.370±0.155	0.398±0.094	0.393±0.059	0.400±0.042	0.400±0.028	0.400±0.020
0.5	0.491±0.158	0.490±0.076	0.493±0.061	0.500±0.044	0.500±0.029	0.500±0.021
0.6	0.575±0.178	0.597±0.097	0.592±0.062	0.600±0.045	0.600±0.030	0.600±0.021
0.7	0.683±0.163	0.677±0.102	0.693±0.063	0.700±0.046	0.700±0.031	0.700±0.022
0.8	0.787±0.162	0.808±0.100	0.794±0.064	0.801±0.047	0.801±0.031	0.801±0.022
0.9	0.862±0.131	0.890±0.096	0.892±0.060	0.902±0.046	0.902±0.031	0.901±0.022

valeur estimée moyenne de 0.796 lorsqu'il n'y a pas de bruit. Cette valeur chute à 0.138 pour un rapport signal à bruit de 10 dB. Il est alors nécessaire de développer des approches adaptées pour que la mesure du paramètre  $H$  soit interprétable en présence de bruit.

S. Höfer *et al.* [47] ont montré comment prendre en compte un bruit additif blanc dans la mesure de  $H$  par EMV. Soit  $z$  la somme d'un vecteur fBm  $X$  noyé dans un bruit blanc  $\omega$  de moyenne nulle et de variance  $\sigma_\omega^2$  inconnue.  $X$  et  $\omega$  sont statistiquement indépendants. La fonction d'autocorrélation des incréments de  $z$  est alors :

$$r_{\Delta z}[k] = r_x[k] + r_{\Delta \omega}[k] \quad (42)$$

$x$  est le processus fGn dont on connaît la fonction d'autocorrélation par (14). Le deuxième terme vaut quant à lui :

$$r_{\Delta \omega}[k] = \sigma_\omega^2 \{2\delta[k] - \delta[k+1] - \delta[k-1]\} \quad (43)$$

$\delta$  est la distribution de Dirac. Il y a maintenant trois paramètres inconnus :  $H$ ,  $\sigma^2$  la variance du vecteur incrément fGn, et  $\sigma_\omega^2$  la variance du bruit. En fait, l'estimation de ces 3 paramètres peut se ramener à celle de la valeur de  $H$  et du rapport  $\sigma^2/\sigma_\omega^2$ . Il suffit alors de rechercher le couple de valeurs qui maximise la fonction de vraisemblance. La méthode du Simplex a été utilisée. De la même manière que précédemment, les moyennes et variances d'estimation par EMV pour un nombre de réalisations  $NR = 100$  ont été mesurées pour un rapport signal à bruit  $S/N$  (défini comme  $\sigma^2/\sigma_\omega^2$ ) de 0 dB. Le tableau suivant présente les résultats.

Tableau 12. – valeurs de la moyenne des estimées de  $H$  et de l'écart-type constaté en fonction de  $H$  et  $N$  par maximum de vraisemblance sur des signaux bruités tels que  $S/N$  de 0 dB.

H/N	32	64	128	256	512	1024
0.1	0.293±0.321	0.243±0.284	0.202±0.190	0.154±0.140	0.139±0.070	0.120±0.040
0.2	0.363±0.310	0.261±0.199	0.236±0.124	0.226±0.096	0.221±0.067	0.216±0.078
0.3	0.373±0.266	0.369±0.204	0.339±0.159	0.309±0.107	0.313±0.080	0.320±0.067
0.4	0.469±0.240	0.419±0.218	0.422±0.165	0.402±0.111	0.399±0.082	0.407±0.051
0.5	0.509±0.267	0.491±0.220	0.499±0.180	0.542±0.132	0.480±0.094	0.507±0.076
0.6	0.583±0.282	0.566±0.217	0.579±0.179	0.594±0.147	0.598±0.098	0.614±0.081
0.7	0.614±0.289	0.656±0.225	0.686±0.180	0.707±0.135	0.688±0.104	0.694±0.078
0.8	0.627±0.265	0.734±0.227	0.764±0.205	0.773±0.115	0.797±0.116	0.794±0.096
0.9	0.722±0.287	0.789±0.225	0.819±0.180	0.856±0.150	0.859±0.095	0.869±0.098

La prise en compte du bruit dans l'estimation par EMV améliore grandement les performances d'estimation. Par exemple, pour le rapport de 0 dB que nous avons testé, l'estimation standard est inexploitable alors que les valeurs trouvées dans le tableau 12 montrent que les mesures sont utilisables sauf pour les signaux avec peu d'échantillons et un  $H$  faible. Cependant, l'écart-type de la mesure est plus important que pour le cas non bruité. En effet, plus le nombre de paramètres à mesurer croît, plus les variances d'estimation augmentent [27].

Enfin, il convient de remarquer que cette méthode peut aussi s'adapter à la variante de Whittle lorsque des signaux de grandes tailles sont bruités.

## 5. conclusion

Cet article a consisté à évaluer expérimentalement la qualité des méthodes d'analyse du mouvement brownien fractionnaire (fBm). Cette comparaison a été effectuée sur des signaux théoriquement exacts obtenus par la décomposition de Cholesky de la matrice de covariance des incréments du fBm. Huit méthodes d'analyse parmi les plus utilisées en traitement du signal ont été comparées. Ces méthodes ont été évaluées en termes de biais et de variance, cette dernière étant comparée à la borne de Cramer-Rao. Les résultats ont montré que seule l'estimation par maximum de vraisemblance (EMV) est efficace pour des signaux de faible longueur. Ce fait expérimental complète des résultats d'efficacité démontrés uniquement dans le cas asymptotique. La méthode par ondelettes et celle basée sur la variance conduisent, quant à elles, seulement à des mesures non biaisées. Ceci laisse penser que l'EMV devrait être systématiquement employée pour estimer le paramètre  $H$ . De plus, l'approximation de Whittle dans le domaine spectral de l'EMV permet une estimation performante lorsque la taille des signaux devient importante (supérieure à quelques milliers d'échantillons). Enfin, l'EMV (classique ou variante de Whittle) permet une mesure de bonne qualité pour des signaux pollués par un bruit blanc.

A partir de ces travaux, plusieurs pistes peuvent être explorées. La première est liée au fait que le modèle fBm est souvent trop restrictif pour de nombreuses applications concrètes. De nouvelles approches telles que celles déjà entreprises par ailleurs [18][48][49] devraient permettre plus de souplesse pour l'étude de données réelles. L'utilisation des résultats précédents pour choisir un estimateur efficace sur des données non parfaitement fBm pourrait être envisagée.

Dans le domaine de l'image, une estimation performante du paramètre  $H$  serait rapidement utilisable dans de nombreux cas réels. La mise en œuvre d'un véritable estimateur 2D par maximum de vraisemblance a déjà été proposé [50], mais sa complexité est en  $O(N^4)$ . Par contre, la variante de Whittle 2D aurait une complexité en  $O(N^2 \log_2 N)$  et donc serait parfaitement utilisable sur des images de taille habituelle.

## Annexe : calcul de la borne de Cramer-Rao du paramètre $H$ du fBm

La borne de Cramer-Rao (BCR) d'un paramètre  $q$  est donnée par :

$$\text{BCR}(\theta) = \frac{1}{E \left[ \left( \frac{\partial \text{Ln}(\text{Prob}(\mathbf{x}; \theta))}{\partial \theta} \right)^2 \right]} \quad (44)$$

$\text{Prob}(\mathbf{x}; \theta)$  est la fonction de vraisemblance du vecteur observation  $\mathbf{x}$  paramétrée par  $\theta$ . Dans le cas vectoriel où  $\theta = [\theta_1, \theta_2, \dots, \theta_p]^T$ , la BCR pour une composante  $\theta_i$  peut être exprimée comme suit :

$$\text{BCR}[\theta_i] = [\mathbf{F}^{-1}]_{i,i} \quad (45)$$

$\mathbf{F}$  représente la matrice d'information de Fischer. Ses éléments s'expriment par :

$$[\mathbf{F}]_{i,j} = -E \left[ \left( \frac{\partial^2 \text{Ln}(P(\mathbf{x}; \theta))}{\partial \theta_i \partial \theta_j} \right)^2 \right] \quad (46)$$

Dans notre cas,  $\mathbf{x}$  est le vecteur fGn et  $\theta = (H, a^2)^T$ . Le vecteur observation étant gaussien centré, on trouve alors [27] :

$$[\mathbf{F}]_{i,j} = \frac{1}{2} \text{tr} \left[ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_j} \right] \quad (47)$$

$\text{tr}[\cdot]$  représente l'opérateur matriciel trace et  $\mathbf{R}$  est la matrice de covariance du vecteur observation. On rappelle que les éléments de  $\mathbf{R}$  s'expriment à l'aide de (14) :

$$[\mathbf{R}]_{i,j} = r_x[|i-j|] = r_x[k] \quad (48)$$

La matrice d'information de Fischer contient 4 éléments :

$$\begin{pmatrix} \mathbf{F}_{HH} & \mathbf{F}_{Ha^2} \\ \mathbf{F}_{a^2H} & \mathbf{F}_{a^2a^2} \end{pmatrix} \quad (49)$$

Pour simplifier les calculs, on définit  $\mathbf{R}'$  comme  $\mathbf{R}' = \mathbf{R}/a^2$ . On trouve alors :

$$\mathbf{F}_{HH} = \frac{1}{2} \text{tr} \left[ \left( \mathbf{R}'^{-1} \frac{\partial \mathbf{R}'}{\partial H} \right)^2 \right] \quad (50)$$

et

$$\mathbf{F}_{Ha^2} = \mathbf{F}_{a^2H} = \frac{1}{2a^2} \text{tr} \left[ \mathbf{R}'^{-1} \frac{\partial \mathbf{R}'}{\partial H} \right] \quad (51)$$

Enfin,

$$\mathbf{F}_{a^2a^2} = \frac{1}{2} \text{tr} \left[ \left( \frac{1}{a^2} \mathbf{I} \right)^2 \right] = \frac{N}{2a^4} \quad (52)$$



$N$  étant la taille du signal et  $I$  la matrice identité. Pour le paramètre  $H$ , la borne de Cramer-Rao est donnée par :

$$\mathbf{F}_{a^2 a^2} = \frac{1}{2} \text{tr} \left[ \left( \frac{1}{a^2} I \right)^2 \right] = \frac{N}{2a^4} \quad (53)$$

Cela s'écrit finalement :

$$\text{BCR}(H) = \frac{\mathbf{F}_{a^2 a^2}}{\mathbf{F}_{a^2 a^2} \mathbf{F}_{HH} - (\mathbf{F}_{Ha^2})^2} \quad (54)$$

$\text{BCR}(H)$  est obtenue par l'évaluation numérique de  $\partial \mathbf{R}' / \partial H$  et de  $\mathbf{R}'^{-1}$  en utilisant l'algorithme de Levinson pour cette matrice Toeplitz [36].

## BIBLIOGRAPHIE

- [1] R. Dahlhaus. « Efficient parameter estimation for self-similar processes », *The Annals of Statistics*, Vol. 17, N° 4, pp. 1749-1766, 1989.
- [2] B. B. Mandelbrot, J. W. Van Ness. « Fractional Brownian Motion, Fractional Noises and Applications », *SIAM*, Vol. 10, N° 4, pp. 422-438, 1968.
- [3] B. B. Mandelbrot. « Noises with  $1/f$  Spectrum, a Bridge Between Direct Current and White Noise », *IEEE Trans. on Information Theory*, IT-13, pp. 289-298, 1967.
- [4] C. W. J. Granger, « The Typical Spectral Shape of an Economic Variable », *Econometrica*, N° 34, pp. 150-161, 1966.
- [5] P. Abry, D. Veitch. « Wavelet Analysis of Long-Range-Dependent Traffic », *IEEE Trans. on Information Theory*, Vol. 44, N° 1, pp. 1-15, 1998.
- [6] A. P. Pentland. « Fractal-Based Description of Natural Scenes », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, N° 6, pp. 661-674, November 1984.
- [7] G. Jacquet, W. J. Ohley, M. Mont, R. Siffert, R. Schmuckler. « Measurement of Bone Structure by Use of Fractal Dimension », *Proc. of the 12th Conf. IEEE Engineering in Medicine and Biology Society*, n° 12, pp. 1402-1403, 1990.
- [8] M. Zabot, M. Veyer-Besancon, R. Harba, S. Bonnamy, H. Van Damme. « Surface Topography and Mechanical Properties of Smectite Films », *Progr. Colloid. Polym. Sci.*, Vol. 105, pp. 96-102, 1997.
- [9] R. Jennane, R. Harba, G. Jacquet. « Estimation de la Qualité des Méthodes de Synthèse du Mouvement Brownien Fractionnaire », *Traitement du Signal*, Vol. 13, N° 4, pp. 289-302, 1996.
- [10] J. M. Bardet. « Un test d'autosimilarité pour les processus gaussiens à accroissements stationnaires », *C. R. Acad. Sci. Paris, Sér. I Math.*, N° 6, pp. 521-526, 1999.
- [11] N. Gache, P. Flandrin, D. Garreau. « Fractal Dimension Estimators For Fractional Brownian Motions », *Proceedings of the ICASSP*, Vol. 5, pp. 3557-3560, 1991.
- [12] M. S. Taqqu, V. Teverovsky, W. Willinger. « Estimators for long range dependence: an empirical study », *Fractals*, Vol. 3, N° 4, pp. 785-788, 1995.
- [13] A. I. Mcleod, K. W. Hipel. « Preservation of the Rescaled Adjusted Range: A Reassessment of the Hurst Phenomenon », *Water Resources Research*, Vol. 14, N° 3, pp. 491-508, June 1978.
- [14] T. Lundahl, W. J. Ohley, S. M. Kay, R. Siffert. « Fractional Brownian Motion: A Maximum Likelihood Estimator and its Application to Image Texture », *IEEE Transactions on Medical Imaging*, Vol. MI-5, N° 3, pp. 152-161, September 1986.
- [15] R. J. Barton, H. V. Poor. « Signal Detection in Fractional Gaussian Noise », *IEEE Transactions on Information Theory*, Vol. 34, N° 5, pp. 943-959, September 1988.
- [16] P. Flandrin. « On the Spectrum of Fractional Brownian Motions », *IEEE Transactions on Information Theory*, Vol. 35, N° 1, pp. 197-199, January 1989.
- [17] P. Flandrin. « Wavelet Analysis and Synthesis of Fractional Brownian Motions », *IEEE Transactions on Information Theory*, Vol. 38, N° 2, pp. 910-917, March 1992.
- [18] J. Beran. « Statistics for Long-Memory Processes », *Chapman & Hall*, 1994.
- [19] M. F. Barnsley, R. L. Devaney, B. B. Mandelbrot, H.-O. Peitgen, D. Saupe, R. F. Voss. « The Science of Fractal Images », *Springer-Verlag*, 1988.
- [20] F. Sellan. « Synthèse de mouvements browniens fractionnaires à l'aide de la transformée par ondelettes », *CRAS*, Vol. 321, Série I, pp. 351-358, 1995.
- [21] P. Abry, F. Sellan. « The Wavelet-based synthesis for Fractional Brownian Motion Proposed by F. Sellan and Y. Meyer: Remarks and Fast Implementation », *Applied and Computational Harmonic Analysis*, Vol. 3, pp. 377-383, 1996.
- [22] C. W. Granger, R. Joyeux. « An Introduction to long Memory Time Series Models and Fractional Differencing », *Journal of Time Series Analysis*, Vol. 1, N° 1, pp. 15-29, 1980.
- [23] J. R. Hosking. « Fractional Differencing », *Biometrika*, Vol. 68, N° 1, pp. 165-176, 1981.
- [24] Y. Meyer, F. Sellan, M. S. Taqqu. « Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion », *The Journal of Fourier Analysis and Applications*, Vol. 5, pp. 466-494, 1999.
- [25] N. J. Kasdin. « Discrete Simulation of Colored Noise and Stochastic Processes and  $1/f^\alpha$  Power law Generation », *Proceedings of the IEEE*, N°5, pp. 802-827, 1995.
- [26] G. Saporta. « Probabilités, analyse des données et statistique », *Editions Technip, Paris*, 1990.
- [27] S. M. Kay. « Fundamentals of Statistical Signal Processing », *Englewood Cliffs, New Jersey : Prentice Hall International Editions*, 1992.
- [28] G. W. Wornell, A. V. Oppenheim. « Estimation of Fractal Signals from Noisy Measurements Using Wavelets », *IEEE Trans. on Signals Proc.*, Vol. 40, pp. 611-623, 1992.
- [29] D. Veitch, P. Abry. « A Wavelet-based Joint Estimator of the Parameters of Long-range Dependence », *IEEE Transactions on Info. Theory*, Vol. 45, N° 3, pp. 878-897, 1998.
- [30] J. S. Bendat, A. G. Piersol. « RANDOM DATA: Analysis and Measurement Procedures », *Wiley-Interscience*, 1971.
- [31] C. Tricot, J. F. Quiniou, D. Wehbi, C. Roques-Carmes, B. Dubuc. « Evaluation de la dimension fractale d'un graphe », *Revue Phys. Appl.*, Vol. 23, pp. 111-124, 1988.
- [32] P. Maragos, F. K. Sun. « Measuring the Fractal Dimension of Signals: Morphological Covers and Iterative Optimization », *IEEE Transactions on Signal Processing*, Vol. 41, N° 1, pp. 108-121, January 1993.
- [33] J. Feder. « Fractals », *Plenum Press, New York and London*, 1988.
- [34] J. Serra. « Image Analysis and Mathematical Morphology », *London: Academic*, Vol. 1, 1982.
- [35] M. Coster, J. L. Chermant. « Précis d'Analyse d'Images », *Editions du CNRS*, pp. 315-317, 1989.
- [36] S. M. Kay. « Modern Spectral Estimation: Theory and Application », *Englewood Cliffs : Prentice Hall International Editions*, 1988.
- [37] B. B. Mandelbrot. « The fractal geometry of nature », *W. H. Freeman*, San-Francisco, 1982.
- [38] S. G. Mallat. « A Theory For Multiresolution Signal Decomposition : The Wavelet Representation », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, N° 7, pp. 674-693, July 1989.

- [39] I. Daubechies. « Orthonormal Bases of Compactly Supported Wavelets », *Communications on Pure and Applied Mathematics*, Vol. XLI, pp. 909-996, 1988.
- [40] P. Abry. « Ondelettes et turbulences – Multirésolution, algorithmes de décomposition, invariance d'échelle et signaux de pression » Ed. Diderot, Paris, 1997.
- [41] D. Veitch, P. Abry. « A Wavelet-Based Joint Estimator of the Parameters of Long-Range Dependence », *IEEE Trans. on Information Theory*, vol. 45, n°3, pp. 878-897, April 1999.
- [42] L. F. Burlaga, L. W. Klein. « Fractal Structure of the Interplanetary Magnetic Field », *Journal of Geophysical Research*, Vol. 91, N° A1, pp. 347-350, January 1986.
- [43] B. B. Mandelbrot. « Les objets fractals : Forme, Hasard et Dimension », *Flammarion*, Paris, 1975.
- [44] T. Higuchi. « Approach to an Irregular Time Series on the Basis of the Fractal Theory », *Physica D* 31, pp. 277-283, 1988.
- [45] G. Istaş, G. Lang. « Quadratic variations and estimation of the local Hölder index of a Gaussian process », *Ann. Inst. Poincaré*, Vol. 33, pp. 407-436, 1997.
- [46] P. Whittle. « Estimation and information in stationary time series », *Ark. Mat.*, Vol. 2, pp. 423-434, 1953.
- [47] S. Höfer, R. Kumaresan, M. Pandit, W. J. Ohley. « Estimation of the fractal dimension of a stochastic fractal from noise-corrupted data », *Arch. Elec. U.*, Vol. 46, N° 1, pp. 13-21, 1992.
- [48] M. Deriche, A. H. Tewfik. « Efficient Modeling with Filtered Discrete Fractional Noise Processes », *IEEE Trans. on Signal Proc.*, Vol. 41, N° 9, pp. 2839-2849, 1993.
- [49] L. M. Kaplan, C. C. J. Kuo. « Extending Self-Similarity for Fractional Brownian Motion », *IEEE Trans. on SP*, Vol. 42, N° 12, pp. 3526-3530, 1994.
- [50] S. Höfer, H. Hannachi, M. Pandit, R. Kumaresan. « Isotropic two-dimensional fractional Brownian motion and its application in ultrasonic analysis », *IEEE EMBS-92*, pp. 1267-1269, 1992.

*Manuscrit reçu le 10 janvier 2000*

### LES AUTEURS

#### Rachid JENNANE



Rachid JENNANE a obtenu un doctorat en sciences pour l'ingénieur de l'Université d'Orléans en décembre 1995 sur la modélisation fractale de textures avec une application à l'analyse de l'architecture osseuse. Depuis 1996, il est maître de conférences à l'Université d'Orléans et effectue ses recherches au Laboratoire d'Electronique, Signaux, Images (LESI).

Visiteur de longue durée à l'Electrical Engineering Department de l'Université de Rhode Island (USA) en 1998. Ses activités de recherche portent sur le traitement du signal et de l'image appliqué aux images et volumes biomédicaux.

#### Rachid HARBA



Rachid HARBA est un ancien élève de l'ENSET Cachan où il a obtenu en 1983 son agrégation en génie électrique. En 1987, il est devenu Maître de Conférences à l'ESPEO, Université d'Orléans, puis Professeur en 1997. Il effectue ses recherches au LESI en traitement du signal et des images. Il s'est d'abord intéressé à l'analyse d'images de pierres puis à des méthodes de caractérisation de textures appliquées au domaine médical.

Il s'intéresse maintenant plus particulièrement à des méthodes fractales pour décrire des milieux tridimensionnels poreux. Il est membre du GDR ISIS depuis sa création.

#### Gérard JACQUET



Gérard JACQUET est Maître de conférences à l'Institut Supérieur des Techniques Avancées de Saint-Etienne (ISTASE) à l'Université Jean Monnet. Il fait partie de l'équipe Systèmes de Vision pour l'Etude des Mouvements au sein du Laboratoire Traitement du Signal et Instrumentation (TSI) depuis 1995, il y développe un thème de recherche relatif à la mesure du mouvement en temps réel via des architectures spécifiques. Préalablement, lorsqu'il était en

poste à Orléans, son travail au Laboratoire Electronique Signaux Images (LESI) était centré sur l'analyse de textures, en particulier en utilisant des modèles fractals, les applications principales de ces travaux se situant dans le domaine biomédical.