

Estimation Conjointe en Ondelette des Paramètres du Phénomène de Dépendance Longue

Darryl VEITCH⁽¹⁾, Patrice ABRY⁽²⁾

⁽¹⁾ SERC, Royal Melbourne Institute of Technology -
Level 2, 723 Swanston Street Carlton, Victoria 3053, - Australia <http://www.serc.rmit.edu.au/>
tel: (+613) 9282 2440 - Fax: (+613) 9282 2444 - E-mail: darryl@serc.rmit.edu.au

⁽²⁾ CNRS URA 1325 - Laboratoire de Physique - Ecole Normale Supérieure de Lyon -
46, allée d'Italie 69 364 LYON Cedex 07 - France <http://www.physique.ens-lyon.fr/ts/>
tel: (+33) 4 72 72 84 93 - Fax: (+33) 4 72 72 80 80 - E-mail: pabry@physique.ens-lyon.fr

RÉSUMÉ

Nous présentons un estimateur conjoint, reposant sur les coefficients d'une décomposition en ondelettes, des deux paramètres définissant le phénomène de dépendance à longue portée. Nous montrons que cet estimateur est caractérisé par sa simplicité à la fois conceptuelle et à la mise en œuvre : il consiste essentiellement à mesurer la pente et l'ordonnée à l'origine d'une régression linéaire. Ces excellentes performances statistiques, non biais et presque minimum de variance, sont démontrées analytiquement et illustrées par des simulations numériques. Cet outil est ensuite appliqué à l'analyse de données de télétrafic Ethernet.

ABSTRACT

We present a joint estimator for the two parameters that define the long-range dependence phenomenon, based on the coefficients of a wavelet decomposition. An important feature is its conceptual and practical simplicity, consisting essentially in measuring the slope and the intercept of a linear fit. The excellent statistical performance of the estimator, unbiased and near-minimal variance, is shown through closed form expressions and numerical simulations. We apply the tool to the analysis of Ethernet teletraffic data.

1 Motivation

Le phénomène de dépendance longue. On définit généralement le phénomène de *Dépendance à Long Terme* (DLT) par une décroissance lente, en loi de puissance pour les grands retards, de la fonction d'autocorrélation d'un processus : $\gamma(k) \sim c_\gamma |k|^{-(1-\alpha)}$, $\alpha \in (0, 1)$. De façon équivalente, on peut le définir par la divergence à l'origine en loi de puissance, de son spectre :

$$f(v) \sim c_f |v|^{-\alpha} \quad (1)$$

et $c_f = \pi^{-1} c_\gamma \Gamma(\alpha) \sin((\alpha + 1)\pi/2)$, où Γ est la fonction d'Euler. Cette définition associe la DLT aux processus stochastiques stationnaires au second ordre, ce phénomène présente néanmoins également d'intimes relations avec la propriété d'auto-similarité [9, 4] et donc les processus stochastiques autosimilaires (non stationnaires) comme le mouvement Brownien fractionnaire. L'outil proposé ici pourra également être efficacement utilisé dans ce contexte [3, 11]. Chaque définition implique deux paramètres, soit (α, c_γ) ou bien (α, c_f) respectivement. De ces deux paramètres, α est, de façon naturelle, considéré comme le plus important, car il rend compte de l'existence même du phénomène, et en le comportement à travers les (grandes) échelles d'un signal DLT, aussi bien que les propriétés des statistiques qu'on peut mesurer à partir d'une seule réalisation du processus. L'attention principale,

donc, est quasiment toujours portée sur l'estimation de α , ou d'une façon équivalente, du paramètre de Hurst H correspondant, relié à α par $H = (\alpha + 1)/2$, et sur la gamme de fréquences (ou échelles) pour laquelle le comportement en loi de puissance reste valide. L'estimation du second paramètre, c_γ (ou c_f ou d'autre choix équivalents [11, 4]), a été relativement négligée, et son importance non correctement appréciée. En fait, dans les données expérimentales et les applications, si c'est bien le paramètre α qui permet de déceler l'existence et l'étendue de la DLT, c'est bien ce deuxième paramètre qui en contrôle, pour une large mesure, amplitude, et donc l'intensité de ces conséquences. Dans l'analyse des files d'attente en télécommunications, par exemple, où des processus DLT servent de modèles pour les données de télétrafic, c'est la quantité c_γ qui gouverne la probabilité (asymptotique) que la file excède un taux de remplissage fixé a priori [5, 8]. L'estimation de c_γ devient donc essentielle dès qu'une analyse quantitative des propriétés du système ou signal étudiés est nécessaire. Or, tout comme pour l'estimation de α , les problèmes statistiques qui y sont associés sont loin d'être triviaux. Nous nous intéressons donc ici à l'estimation conjointe des paramètres qualitatif et quantitatif de la DLT.

Des estimateurs conjoints. Des estimateurs conjoints de ce couple de paramètres ont déjà été proposé dans la littérature, qui, généralement, reposent sur l'utilisation d'un principe de maximum de vraisemblance. Les deux plus connus

reparent sur une formulation de ce principe, pour l'un (Whittle, [4]) dans le domaine spectral, pour l'autre (Wornell, [10]) dans le plan des coefficients d'ondelettes. Ces deux estimateurs présentent, sous hypothèses gaussiennes, de bonnes propriétés statistiques. Ils souffrent cependant, à notre sens, de deux inconvénients majeurs : ils sont, par nature, paramétriques et leur performances d'estimation se trouveront considérablement amoindries dès que les données s'écarteront du modèle choisi, l'estimation des paramètres nécessite la mise en œuvre d'une procédure de minimisation itérative, dont la convergence peut être mal assurée et dont le coût de calcul est nécessairement très élevé.

L'estimateur ondelette. Nous proposons ici un nouvel estimateur conjoint construit à partir des coefficients d'une décomposition en ondelettes. Nous montrons que cet estimateur, quoiqu'il soit caractérisé par une grande simplicité à la fois conceptuelle et à la mise en œuvre, présente d'excellentes performances statistiques (théoriques, section 3 et expérimentales, section 4), non biais même à nombre fini d'échantillons, variance d'estimation minimale ; performances comparables à celles des estimateurs à maximum de vraisemblance. Ces bonnes performances et cette simplicité résultent de l'adéquation entre phénomène analysé et outil d'analyse, qui tous deux présentent la propriété d'invariance d'échelle [1, 3]. L'estimateur proposé présente de plus la qualité de n'être que semi-paramétrique, allégeant la contrainte de connaître un modèle qui décrive rigoureusement la totalité du spectre du signal. Le comportement $c_f|v|^{-\alpha}$ est, en effet, sur des données expérimentales, susceptible de n'être valide qu'approximativement et sur une gamme finie de fréquences, compliquant l'usage des méthodes à maximum de vraisemblance. Il présente enfin une grande robustesse vis à vis de dérives ou tendances susceptibles de se superposer aux données et de biaiser l'estimation [2].

2 L'estimateur Ondelette

Décomposition en ondelettes. On notera par $d_x(j, k) = \langle x, \psi_{j,k} \rangle$ les coefficients de la décomposition discrète (ou non redondante) en ondelettes ; les $\{\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t - k), j = 1, \dots, J, k \in \mathbb{Z}\}$ désignent la base d'ondelettes obtenue à partir de l'ondelette-mère ψ_0 , issue d'une analyse multirésolution [6]. La construction de cette famille de fonctions d'analyse repose sur deux ingrédients (I1, I2),

I1 : la base est construite à partir de l'opérateur de changement d'échelle : $\psi_{j,0}(t) = 2^{-j/2}\psi_0(2^{-j}t)$,

I2 : ψ_0 possède un nombre de moments nuls N qui peut être choisi et vérifie nécessairement $N \geq 1$. N est tel que $\forall k = 0, \dots, N-1, \int t^k \psi_0(t) dt \equiv 0$.

qui ont des conséquences fondamentales (P1, P2) sur les coefficients $d_x(j, k)$ d'un processus DLT ou à spectre en loi de puissance (ces aspects sont détaillés dans [1, 2, 3]).

P1 : L'invariance d'échelle (la loi de puissance) est capturée de façon exacte (effet de P1) :

$$\mathbb{E}d_x(j, \cdot)^2 = 2^{j\alpha} c_f C \quad (2)$$

où $C = \int |v|^{-\alpha} |\Psi_0(v)|^2 dv$,

P2 : Les $d_x(j, k)$ constitue une collection de variables quasiment décorréelées (effet de I1 et I2) [7]. En particulier, la DLT présente dans la représentation temporelle a été complètement éliminée dans le plan des coefficients d'ondelettes.

Intuition et difficulté. Dans toute la suite, \log désigne le logarithme à base 2. Réécrire la relation fondamentale (2) sous la forme, $\log(\mathbb{E}d_x(j, \cdot)^2) = j\alpha + \log(c_f C)$, suggère fortement d'utiliser la régression linéaire pour estimer (α, c_f) . La principale difficulté provient de la nécessité d'estimer la quantité statistique de second ordre $\mathbb{E}d_x(j, \cdot)^2$. Il est connu [4] que la DLT interdit l'usage de l'estimateur standard de variance et rend très délicat cette estimation. La quasi-décorrélation des $d_x(j, k)$ nous permet ici, pour estimer $\mathbb{E}d_x(j, \cdot)$, d'utiliser simplement [1, 2]

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} d_x^2(j, k) \quad (3)$$

où $n_j = n2^{-j}$ est le nombre de coefficients à l'échelle j disponible pour l'analyse d'un signal de longueur n . Notre estimateur consistera fondamentalement en une régression linéaire dans le diagramme $(x_j = \log(2^j) = j, y_j = \log(\mu_j))$.

Régression linéaire. Rappelons quelques résultats standards sur la régression linéaire pondérée (RLP) des variables aléatoires y_j par rapport aux valeurs déterministes $x_j, j = j_1 \dots j_2$. L'hypothèse fondamentale est $\mathbb{E}y_j = bx_j + a$. Définissons des quantités $S = \sum 1/\sigma_j^2, S_x = \sum x_j/\sigma_j^2, S_{xx} = \sum x_j^2/\sigma_j^2$ où les sommes sont effectuées sur $j = j_1 \dots j_2$ et σ_j^2 désigne la variance de y_j . L'estimateur (\hat{b}, \hat{a}) non biaisé habituel de (b, a) est

$$\hat{b} = \frac{\sum y_j(Sx_j - S_x)/\sigma_j^2}{SS_{xx} - S_x^2} \equiv \sum y_j w_j \quad (4)$$

$$\hat{a} = \frac{\sum y_j(S_{xx} - S_x x_j)/\sigma_j^2}{SS_{xx} - S_x^2} \equiv \sum y_j d_j \quad (5)$$

où les pondérations w_j et d_j satisfont $\sum w_j = \sum j d_j = 0, \sum j w_j = \sum d_j = 1$. Si on ajoute l'hypothèse H1 : y_j sont iid, cet estimateur est de variance minimale et dans ce cas,

$$\text{Var}(\hat{b}) = \sum \sigma_j^2 w_j^2 = S/(SS_{xx} - S_x^2) \quad (6)$$

$$\text{Var}(\hat{a}) = \sum \sigma_j^2 d_j^2 = S_{xx}/(SS_{xx} - S_x^2) \quad (7)$$

$$r = -S_x/\sqrt{SS_{xx}} < 0 \quad (8)$$

Statistique de $\log(\mu_j)$. Sous hypothèse H2 : le processus étudié est gaussien, et sous l'hypothèse H1, rendue très réaliste par la propriété P2, la variable μ_j suit une loi en $\frac{\mathbb{E}d_x(j, \cdot)^2}{n_j} X_{n_j}$ où X_v est une variable aléatoire suivant une loi de χ_v^2 à v degrés de liberté, $f_v(x) = \left(\frac{1}{2^{v/2}\Gamma(v/2)}\right) x^{v/2-1} e^{-x/2}$. Or, on peut montrer [11] que l'espérance de $\ln(X_v)$ s'écrit $\mathbb{E}\ln(X_v) = \psi(v/2) + \ln 2$, où $\psi(z) = \Gamma'(z)/\Gamma(z)$ est la fonction Psi (ou Digamma). On a donc $\mathbb{E}\log(\mu_j) = j\alpha + \log c_f C + g_j$ où le facteur $g_j = \psi(n_j/2) \log e - \log(n_j/2)$, fonction de n_j seulement, tend vers 0 pour $n_j \rightarrow \infty$, et peut être facilement calculé pour toute valeur de n_j . Comme $2^{(\sum d_j \log(\mu_j))} = \prod_{j=j_1}^{j_2} \mu_j^{d_j}$, et comme on peut aisément établir que $\mathbb{E}X_v^d = \frac{2^d \Gamma(d+v/2)}{\Gamma(v/2)}$, $(2d+v) > 0$, on peut également calculer [11] $\mathbb{E}2^{(\sum d_j \log(\mu_j))} = \prod \frac{\Gamma(d_j+n_j/2)}{\Gamma(n_j/2)}$.

Définition des estimateurs conjoints. Les calculs ci-dessus nous permettent de proposer la définition de l'estimateur conjoint $(\hat{\alpha}, \widehat{c_f C})$ **non biaisé** suivant,

$$\hat{\alpha} = \hat{b} \quad (9)$$

$$\widehat{c_f C} = p(\hat{a}, n) 2^{\hat{a}} \quad (10)$$

où $p(\hat{a}, n) = \prod \frac{\Gamma(n_j/2) \exp(\psi(n_j/2)d_j)}{\Gamma(d_j+n_j/2)}$ et \hat{b} et \hat{a} sont la pente et l'ordonnée à l'origine de la RLP effectuée dans le diagramme ($x_j = j, y_j = \log(\mu_j) - g_j$). On définit de même l'estimateur conjoint de $(\hat{\alpha}, \widehat{c_f C})$ par

$$\hat{\alpha} = \hat{b} \quad (11)$$

$$\hat{c}_f = \widehat{c_f C} / C(\hat{\alpha}, \Psi_0) \quad (12)$$

où $C(\hat{\alpha}, \Psi_0) = \int |v|^{-\hat{\alpha}} |\Psi_0(v)|^2 dv$.

Performances théoriques. Par construction, et sous les hypothèses H1 et H2, l'estimateur $(\hat{\alpha}, \widehat{c_f C})$ est non-biaisé. Le non-biais de $\widehat{c_f C}$ est toutefois obtenu à condition que l'on vérifie que pour les échelles pratiquement impliquées dans la RLP, $d_j + n_j/2 > 0$, conditions qui assurent l'existence de $\mathbb{E}\widehat{c_f C}$.

La variance de $\hat{\alpha}$ est donnée par (6). On sait que cette variance est optimale parce que celle de \hat{b} l'est. On peut également obtenir [11]

$$\text{Var} \widehat{c_f C} = (c_f C)^2 \left[\prod \frac{\Gamma(2d_j + n_j/2) \Gamma(n_j/2)}{\Gamma(d_j + n_j/2)^2} - 1 \right] \quad (13)$$

De (13), on tire les conditions qui assurent l'existence de la variance, $4d_j + n_j > 0$, qu'il faut pratiquement vérifier pour chaque octave. Le calcul de la covariance entre $\hat{\alpha}$ et $\widehat{c_f C}$ est plus compliqué et reste à l'étude. Pour l'estimateur $(\hat{\alpha}, \widehat{c_f C})$, nous ne présentons que des résultats numériques.

Borne de Cramer-Rao. Cette borne est obtenue à partir de la densité conjointe des μ_j . Définissons $vS = \sum 2^{-j}$, $vS_j = \sum j 2^{-j}$, et $vS_{jj} = \sum j^2 2^{-j}$, où $v = 2/(n \ln^2 2)$. On peut démontrer [11] sans peine sous H1 et H2 que la Borne de Cramer Rao s'écrit :

$$\text{Var}(\hat{\alpha}, \widehat{c_f C}) \geq \frac{1}{S_{jj}S - S_j^2} \begin{bmatrix} S & -(c_f C) \ln 2 S_j \\ -(c_f C) \ln 2 S_j & (c_f C)^2 \ln^2 2 S_{jj} \end{bmatrix}$$

Dans le cas $n_{j_2} \rightarrow \infty$, on a $\text{Var}(y_j) = \text{Var}(\log(\mu_j)) = v 2^{-j}$ car $g_j = 0$ et cette borne devient :

$$\text{Var}(\hat{\alpha}, \widehat{c_f C}) = \begin{bmatrix} \text{Var}(\hat{b}) & c_f C \ln 2 \cdot \text{CoV}(\hat{a}, \hat{b}) \\ c_f C \ln 2 \cdot \text{CoV}(\hat{a}, \hat{b}) & (c_f C)^2 \ln^2 2 \cdot \text{Var}(\hat{a}) \end{bmatrix}$$

Dans cette limite, les variances ci-dessus coïncident précisément avec la borne de Cramer-Rao. En pratique, l'accord numérique entre ces variances et cette borne est très bon, même pour des tailles finies et petites d'échantillons. Néanmoins, comme $\widehat{c_f C}$ est une fonction non linéaire de \hat{a} , on ne sait pas si la variance (13) est optimale quel que soit n_{j_2} . Notons, de plus, que dans cette limite, le coefficient de corrélation de $(\hat{\alpha}, \widehat{c_f C})$ est le même que celui de (\hat{b}, \hat{a}) et est donc une conséquence de la RLP. Enfin, cette borne est identique à celle obtenue dans [10], indiquant que notre estimateur atteint les performances d'un estimateur à maximum de vraisemblance.

3 Simulation Numériques

Nous avons testé numériquement biais, variance et covariance d'estimation de $(\hat{\alpha}, \widehat{c_f C})$ et $(\hat{\alpha}, \hat{c}_f)$ en fonction du nombre d'échantillons des signaux, pour différents couples (α, c_f) . Chaque point sur les graphiques ci-après est obtenu à partir de 200 réalisations indépendantes de processus aléatoires avec $(\alpha, c_f) = (0.3292, 17.435)$, $j_1 = 1$, et j_2 aussi grand que possible. Ceux-ci correspondent à la simulation (approchée), par synthèse spectrale, de bruits gaussiens fractionnaires. La décomposition discrète en ondelettes est effectuée à l'aide de l'algorithme pyramidal rapide habituel. Les ondelettes utilisées sont celles de Daubechies à six coefficients [6]. Notons cependant que l'orthogonalité ne joue aucun rôle essentiel ici et que la seule propriété déterminante de l'ondelette mère est son nombre de moments nuls.

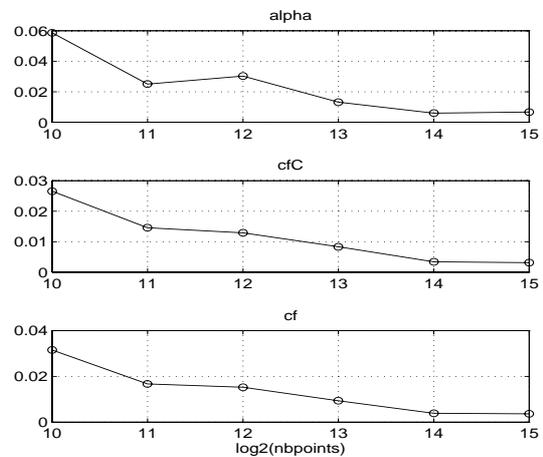


FIG. 1 — **Biais relatif.** De bas en haut, biais relatif $((\hat{\theta} - \theta)/\theta)$ pour $\theta = \hat{\alpha}$, $\widehat{c_f C}$ et \hat{c}_f en fonction du logarithme du nombre d'échantillons.

On obtient \hat{c}_f de $\widehat{c_f C}$ à partir de l'équation (12) avec $C(\hat{\alpha}, \Psi_0) = \int |v|^{-\hat{\alpha}} |\Psi_0(v)|^2 dv$. Pratiquement, cette intégrale est calculée numériquement à partir de la forme d'onde de l'ondelette-mère, qui elle-même est aisément obtenue grâce à l'algorithme pyramidal rapide.

Sur les trois figures suivantes, on constate que les résultats expérimentaux coïncident parfaitement avec les prédictions théoriques, pour l'absence de biais (figure 1) et la variance (figure 2) de $\hat{\alpha}$ et $\widehat{c_f C}$. Pour la covariance d'estimation (figure 3), l'accord avec la borne de Cramer-Rao est également très satisfaisant. Pour l'estimation de c_f , pour laquelle on ne dispose pas de prédiction théorique, on constate également une absence de biais et une décroissance de la variance comme l'inverse du nombre d'échantillons du signal. De plus, bien que l'on utilise $\hat{\alpha}$ pour obtenir \hat{c}_f , la corrélation $(\hat{\alpha}, \hat{c}_f)$ n'est que légèrement dégradée (supérieure en valeur absolue) par rapport à celle de $(\hat{\alpha}, \widehat{c_f C})$. Il est donc légitime en pratique d'assimiler les performances théoriques de \hat{c}_f à celles de $\widehat{c_f C}$.

Il est important de noter que ces bonnes performances pratiques de l'estimateur sont obtenues sous réserve d'initialiser correctement l'algorithme pyramidal. Si, dans les simulations numériques, cette étape peut facilement être réalisée [11], elle

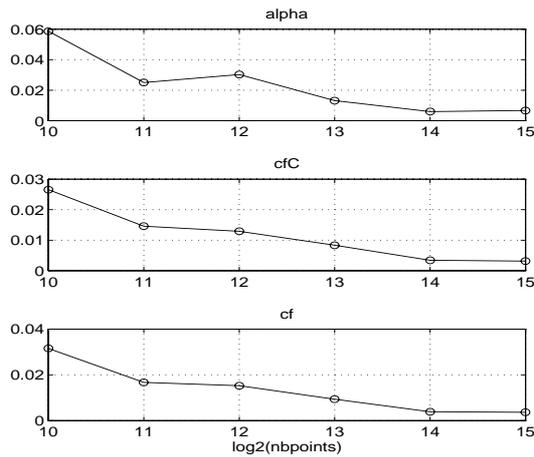


FIG. 2 — **Variance.** De bas en haut, logarithme de la variance d'estimation de $\hat{\alpha}$, $\hat{c}_f \hat{C}$ et \hat{c}_f en fonction du logarithme du nombre d'échantillons. Les cercles indiquent les simulations numériques, les astérisques désignent les valeurs théoriques.

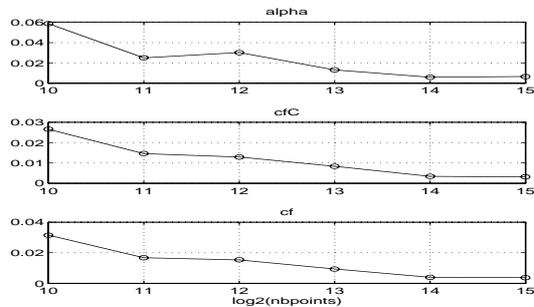


FIG. 3 — **Corrélation des estimées.** En trait plein, le coefficient de corrélation entre $\hat{\alpha}$ et $\hat{c}_f \hat{C}$, (cercles : simulation numérique, astérisques : valeurs théoriques). En trait pointillé, le coefficient de corrélation entre $\hat{\alpha}$ et \hat{c}_f .

pose pour de véritables données expérimentales des difficultés sur lesquelles nous travaillons [11]. Une absence (la séquence initiale est confondue avec la séquence d'échantillons) d'initialisation se traduit pratiquement par l'apparition d'un léger biais (biais relatif de l'ordre de 0.01.)

4 Analyse du télétrafic Ethernet

Les données de télétrafic Ethernet ont déjà fait l'objet de plusieurs études révélant l'existence de DLT [9]. Nous avons précédemment mis en œuvre l'estimateur pour le paramètre α seul [2] sur des traces collectées à Bellcore [9], dont quelques résultats pour la trace "pAug" figurent dans le tableau ci-dessous. On a noté le fait remarquable selon lequel la valeur estimée demeure à peu près identique quelle que soit la modélisation adoptée des données : A et F sont des listes des instants d'arrivée des paquets et leurs longueurs, S celle des silences, W_δ et C_δ sont respectivement les processus agrégés [9, 2, 3] de travail et de comptage (respectivement nombre d'octets ou de paquets dans une fenêtre de longueur δ). Nous complétons ici cette étude par une estimation de c_f , où plutôt

de c_f/σ_x^2 (qui caractérise la fonction de corrélation). Insistons de nouveau sur l'importance de connaître c_f pour évaluer l'impact de la DLT sur les comportements des réseaux.

pAug - D3	W_δ	C_δ	F	S	A
\hat{H}	0.792	0.846	0.822	0.799	0.832
CI	0.08	0.04	0.028	0.034	0.027
\hat{c}_f/σ_x^2	0.2459	0.1933	0.2281	0.1184	0.0626
CI	0.0111	0.0088	0.0125	0.0065	0.0034

Les valeurs rapportées dans ce tableau révèlent que les différents modèles, quoique présentant le même α , n'ont pas les mêmes caractéristiques DLT, car les estimées de c_f/σ_x^2 , varient sensiblement les unes par rapport aux autres. Les intervalles de confiance ont été calculés sous hypothèses gaussiennes.

Références

- [1] P.Abry, P.Gonçalvès and P.Flandrin, Wavelets, Spectrum estimation, $1/f$ processes. Wavelets and Statistics, *Lectures Note in Statistics*, Vol. 105, pp. 15–30, 1995.
- [2] P.Abry, D.Veitch, Wavelet Analysis of Long Range Dependent Traffic. to appear in *IEEE Trans. on Info. Theory*, (1997).
- [3] P.Abry, D.Veitch and P. Flandrin, Long-Range Dependence : revisiting Aggregation with Wavelets. to appear in *Journal of Time Series Analysis*, (1997).
- [4] J.Beran, Statistics for Long-Memory Processes, Chapman & Hall (1994)
- [5] F.Brichet, J.Roberts, A.Simonian, D.Veitch, Heavy Traffic Analysis of a Storage Model with Long Range Dependent On/Off Sources, submitted Queueing systems (1995)
- [6] I.Daubechies, Ten Lectures on Wavelets. SIAM, 1992.
- [7] P.Flandrin, Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. on Info. Theory* **IT-38**, (1992) pp.910-917.
- [8] I.Norros, A storage model with self-similar input, *Queueing Systems* 16 (1994), pp.387–396.
- [9] W.Leland, M.Taqqu, W.Willinger, D.Wilson, On the self similar nature of Ethernet traffic (extended version), *IEEE/ACM Transactions on Networking*, February (1994), pp.1-15
- [10] G.W. Wornell, A.V. Oppenheim, Estimation of Fractal Signals from Noisy Measurements Using Wavelets, *IEEE Trans. on Signal Proc.*, vol. 40, NO. 3, March 92, pp. 611–623.
- [11] D.Veitch, P.Abry, Joint wavelet-Based estimates for the parameters of the Long-Range Dependence phenomenon. Preprint (1997).