

Master des Sciences de la Matière

École Normale Supérieure de Lyon
Université Claude Bernard Lyon 1

STAGE 2005-2006
LOISEAU, Patrick
2^{ème} année
Option Physique

Ondelettes complexes pour l'analyse des lois d'échelles

Laboratoire de Physique, ENS Lyon
46, allée d'Italie ; 69364 LYON Cedex 07
[http ://www.ens-lyon.fr/PHYSIQUE](http://www.ens-lyon.fr/PHYSIQUE)

Maîtres de stage : Patrice ABRY (Laboratoire de physique)
Pierre BORGNAT (Laboratoire de physique)
Paulo GONÇALVÈS (LIP : [http ://www.ens-lyon.fr/LIP](http://www.ens-lyon.fr/LIP))

stage effectué du 2 avril au 8 août 2006

Table des matières

Introduction	3
1 Les transformations en ondelette réelle et complexe.	4
1.1 Transformation en ondelette réelle	4
1.1.1 La transformation continue	4
1.1.2 La transformation discrète et la grille dyadique	4
1.2 Transformation en ondelette complexe	6
1.2.1 Position du problème	6
1.2.2 Transformation complexe par simple transformation de Hilbert	6
1.2.3 Transformation de Hilbert complexe par la méthode Dual-Tree	7
1.2.4 Ondelettes complexes et invariance par translation temporelle	8
1.3 Analyse d'une singularité par ondelette	9
1.3.1 Régularité ponctuelle et singularités	9
1.3.2 Régularité ponctuelle et ondelette	9
1.3.3 Comportement d'une singularité de type dirac : analyse par DWT et DT-CWT	10
1.4 Autres jeux de coefficients multirésolutions et résumé	10
1.4.1 La transformation en ondelette continue dyadique (CDWT)	10
1.4.2 Les leaders ou coefficients dominants	10
1.4.3 Résumé des coefficients multirésolutions envisagées dans la suite	11
2 Étude des estimateurs de lois d'échelle	12
2.1 Le mouvement brownien fractionnaire (fBM) et l'estimation du paramètre global H par la marginale	12
2.1.1 Le mouvement brownien fractionnaire (fBM) : définition et propriétés	12
2.1.2 Estimation de H par la marginale	12
2.2 Estimation de la régularité locale	13
2.2.1 Définition de l'estimateur $\hat{h}(t)$	13
2.2.2 Résultats	15
2.3 Interprétation de la variance de $\hat{h}(t)$	15
2.3.1 Distribution des coefficients en ondelette	15
2.3.2 Réduction de l'écart type de $\hat{h}(t)$ $\sigma_{\hat{h}}$ due à la réduction de variance du \log_2 des coefficients en ondelette	17
2.3.3 Corrélation introduite par le module dans les coefficients	18
2.3.4 Conclusion : quel jeu de coefficients utiliser ?	19
2.4 Étude de régularité Hölderienne variable	19
2.4.1 Cas d'une variation douce de la régularité locale	19
2.4.2 Cas d'un saut brutal de régularité locale	19
Conclusion	20
A Régression linéaire	21
A.1 Estimation de la pente : formule des moindres carrés	21
A.2 Variance sur l'estimation de la pente dans le cas d'une estimation ne tenant pas compte de la corrélation à priori	22
B Tableaux des valeurs des écarts types de \hat{h}	23
C Logarithme de variables aléatoires gaussiennes	24
C.1 Moyennes et Variances	24
C.2 Corrélations	24

Introduction

Le but de ce travail est d'estimer, grâce à une méthode basée sur l'analyse en ondelette, la régularité locale d'un processus, caractérisée par son exposant de Hölder. La méthode consiste à utiliser la décroissance à travers les échelles des coefficients en ondelette d'une fonction à un instant t . Elle est gouvernée par l'exposant de Hölder de la fonction à cet instant, ceci dans la limite où l'échelle tend vers 0 puisqu'on s'intéresse à la régularité locale. La première idée pour estimer la régularité locale est d'utiliser la transformée en ondelette continue (CWT pour *Continuous Wavelet Transform*), puisque qu'elle permet d'avoir accès à cette décroissance des coefficients à travers les échelles. Cette méthode a déjà été étudiée ([12]). Elle est efficace mais coûteuse en ressources de calcul. Une seconde idée est d'employer la transformée en ondelette discrète (DWT pour *Discrete Wavelet Transform*) qui correspond à un échantillonnage critique de la CWT. Le problème est que la DWT ne donne pas des coefficients à tous les temps pour toutes les échelles. Ainsi, si on veut estimer la régularité en un temps t_0 , il se peut que l'on ne dispose pas de coefficients à la date t_0 pour toutes les échelles. On doit donc prendre pour certaines échelles des coefficients correspondants à des temps voisins. Ceci pose le problème de non invariance par translation temporelle de la DWT : une faible translation du signal peut induire de fortes variations des coefficients en ondelette. Une possibilité pour résoudre ce problème est d'utiliser la transformée en ondelette complexe. L'idée est d'analyser le signal avec une ondelette complexe analytique et de s'intéresser au module des coefficients. L'utilisation de deux ondelettes en quadrature permet alors de limiter fortement le problème de non invariance par translation temporelle. Une première partie de ce travail fut d'étudier comment réaliser une transformation en ondelette complexe. Deux méthodes sont introduites dans la partie 1 après quelques brefs rappels sur les transformations en ondelette continue et discrète. On discutera la méthode pour estimer l'exposant de Hölder d'une singularité grâce à une analyse en ondelette réelle ou complexe. Enfin, on mettra en œuvre la transformation en ondelette complexe et on mettra en évidence la résolution du problème de non invariance par translation temporelle. Dans la partie 2 qui présente la deuxième contribution de ce travail, nous étudierons l'estimation de régularité locale de processus stochastique, en comparant les performances des estimateurs basés sur les ondelettes réelles et de ceux basés sur les ondelettes complexes. Nous ferons ce travail dans un premier temps dans le cas où l'exposant de Hölder est constant (le fBM) avant de passer au cas où il est variable.

1 Les transformations en ondelette réelle et complexe.

La transformation en ondelette est une des solutions possibles pour faire de l'analyse temps-fréquence (ou plutôt temps-échelle dans le cas des ondelettes) de signaux non stationnaires pour lesquels la transformée de Fourier classique s'avère inappropriée à cause de la non-stationnarité ([11]).

1.1 Transformation en ondelette réelle

Une présentation détaillée de la transformation en ondelette peut être trouvée dans [6, 15, 1]. On rappelle ici brièvement les éléments essentiels.

1.1.1 La transformation continue

La transformée en ondelette d'une fonction f de L_2 est définie par le produit scalaire :

$$c(a, t) = \int_{\mathbb{R}} du f(u) \psi_{a,t}(u) = \langle f, \psi_{a,t} \rangle. \quad (1)$$

où les fonctions $\psi_{a,t}$ se déduisent toutes d'une mère fonction ψ_0 appelée ondelette mère par une opération de translation-dilatation : $\psi_{a,t}(u) = \frac{1}{\sqrt{a}} \psi_0\left(\frac{u-t}{a}\right)$ (translation d'un facteur t et dilatation d'un facteur d'échelle a). Le facteur $\frac{1}{\sqrt{a}}$ correspond au pré-facteur permettant de conserver la norme L_2 .

Pour que la transformée en ondelette corresponde effectivement à une représentation de la fonction f dans le plan temps-échelle acceptable, il faut que l'ondelette mère soit à la fois bien localisée en temps et en fréquence. Ainsi le coefficient $c(a, t)$ représente l'énergie de f à l'échelle a autour du temps t .

De plus, pour que la transformée en ondelette soit inversible, l'ondelette mère doit vérifier la condition d'admissibilité $\int_{-\infty}^{\infty} \frac{|\Psi(\nu)|^2}{|\nu|} d\nu < \infty$. Cette condition implique

$$\int_{\mathbb{R}} dt \psi_0(t) = 0, \quad (2)$$

ce qui impose à la fonction $\psi_0(t)$ d'être oscillante et localisée en temps, d'où sa dénomination d'*ondelette*. De façon plus générale, l'ondelette mère est caractérisée par son nombre de moments nuls N défini par

$$\int_{\mathbb{R}} dt t^k \psi_0(t) = 0, \text{ pour } k = 0 \text{ à } N - 1. \quad (3)$$

C'est une caractéristique importante pour l'estimation de régularité locale. Dans notre cas, nous allons chercher à estimer des exposants de Hölder compris entre 0 et 1. Nous pouvons alors en théorie nous contenter d'une ondelette de régularité 1 (1 moment nul), c'est à dire simplement admissible.

1.1.2 La transformation discrète et la grille dyadique

La transformation en ondelette continue est une représentation d'une fonction $f(t)$ (mono-dimensionnelle) dans un espace à deux dimensions (le plan temps-échelle). C'est donc une transformation redondante. Un des grands atouts de la transformée en ondelette est que l'on peut trouver des bases orthonormées d'ondelettes. On peut alors se limiter au calcul de la transformée en ondelette sur un maillage discret du plan temps-échelle appelé grille dyadique (figure 1) :

$$a \rightarrow 2^j, t \rightarrow 2^j k \text{ avec } j \in \mathbb{Z} \text{ et } k \in \mathbb{Z}.$$

La construction de ces bases orthonormées peut être associée à la notion d'*analyses multirésolutions*. Une analyse multirésolution est un ensemble $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces vectoriels de L_2 emboîtés les uns dans les autres :

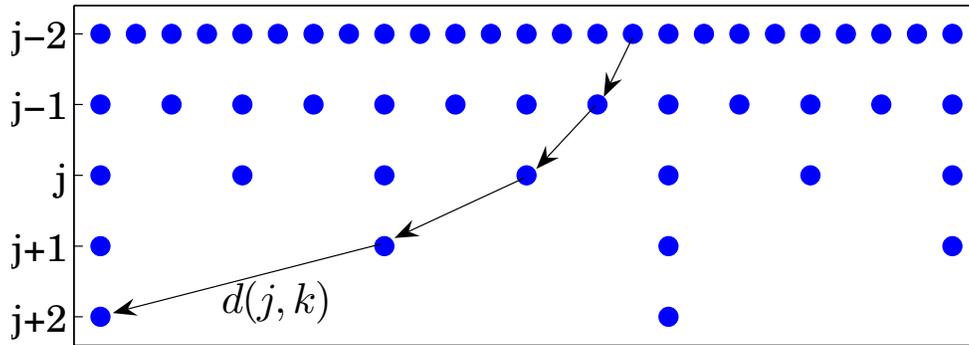
$\dots \subset V_{j+1} \subset V_j \subset V_{j-1} \subset \dots$. Chacun des espaces V_j possède une base orthonormée $\phi_{k,j}(\cdot) = 2^{-\frac{j}{2}} \phi_0(2^{-j} \cdot -k)$, $k \in \mathbb{Z}$, contruite comme pour les ondelettes par action de l'opérateur de translation-dilatation sur une fonction appelée fonction d'échelle. La projection de f dans l'espace V_j fournit les coefficients d'approximation à l'échelle j . Pour compléter ces approximations, il faut projeter dans les espaces complémentaires des V_j . On définit alors les sous espaces W_j appelés espaces de détails comme les complémentaires orthogonaux des V_j dans V_{j-1} . On a alors les propriétés suivantes :

$$\begin{cases} W_j \perp V_{j-1} \\ W_j \oplus V_j = V_{j-1} \\ \bigcup_{j \in \mathbb{Z}} V_j = L_2 \\ \bigcap_{j \in \mathbb{Z}} W_j = \emptyset \end{cases}$$

Les bases orthonormée des sous-espaces W_j sont les familles d'ondelettes $\{\psi_{j,k}\}$. Les coefficients en ondelette ou coefficients de détails sont les projections de la fonction f sur les sous espaces de détails. La définition des analyses multirésolutions permet donc la construction de bases orthonormales d'ondelette. L'utilisation de ces bases orthonormales permettent la reconstruction parfaite d'un signal à partir de sa décomposition en ondelette, à partir de la formule :

$$\hat{f}(t) := \sum_k \langle f, \phi_{Jk} \rangle \phi_{Jk}(t) + \sum_{j < J} \sum_k \langle f, \psi_{jk} \rangle \psi_{jk}(t)$$

FIG. 1: Représentation de la grille dyadique. La transformée en ondelette est calculée aux points $a = 2^j$ et $t = 2^j k$. Les flèches représente la façon de descendre dans l'arbre pour savoir quel coefficient correspond à un temp donné s'il n'y a pas de coefficients à la verticale du temps considéré. On définit ainsi $k(j, t_0) = \lceil \frac{t_0}{2^j} \rceil$

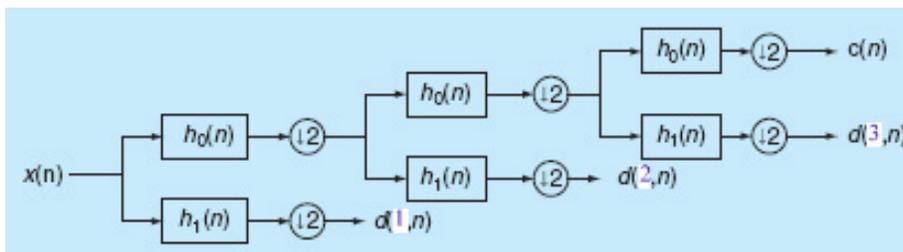


L'introduction des analyses multirésolution a débouché sur un algorithme rapide de calcul de la transformée en ondelette discrète proposé par Mallat ([15]). Supposons que l'on a affaire à un signal échantillonné $X(n)$. On va alors assimiler $X(n)$ à la projection d'une fonction sur l'espace d'approximation correspondant à l'échelle $j = 0$, V_0 . On se limite donc maintenant à $j \in \mathbb{N}$ et k devient l'échantillon. L'algorithme de Mallat montre que le calcul de la transformée en ondelette discrète se réduit à un banc de filtre en quadrature miroir. Les filtres sont alors caractéristiques de l'ondelette et de la fonction d'échelle et la recherche d'ondelette se ramène à la recherche de filtres. La figure 2 montre le fonctionnement de l'algorithme de Mallat. Il utilise deux filtres :

- h_0 : filtre passe-bas qui projète l'approximation de l'échelle $j - 1$ sur l'espace d'approximation de l'échelle j : on obtient donc l'approximation à l'échelle j .
- h_1 : filtre passe-haut qui projète l'approximation de l'échelle $j - 1$ sur l'espace d'ondelette (détail) de l'échelle j : on obtient donc les coefficients en ondelette de l'échelle j .

Le 2 avec une flèche vers le bas dans un rond symbolise une décimation par deux, c'est à dire qu'on retire un point sur deux. Cela permet d'avoir effectivement deux fois moins de coefficients en ondelette à l'échelle j qu'à l'échelle $j - 1$, comme le prévoit l'échantillonnage critique de la grille dyadique. Il existe un algorithme similaire pour la reconstruction du signal à partir de tous les détails et de la dernière approximation. Dans le cas des ondelettes orthogonales, les filtres utilisés sont les filtres h_0 et h_1 .

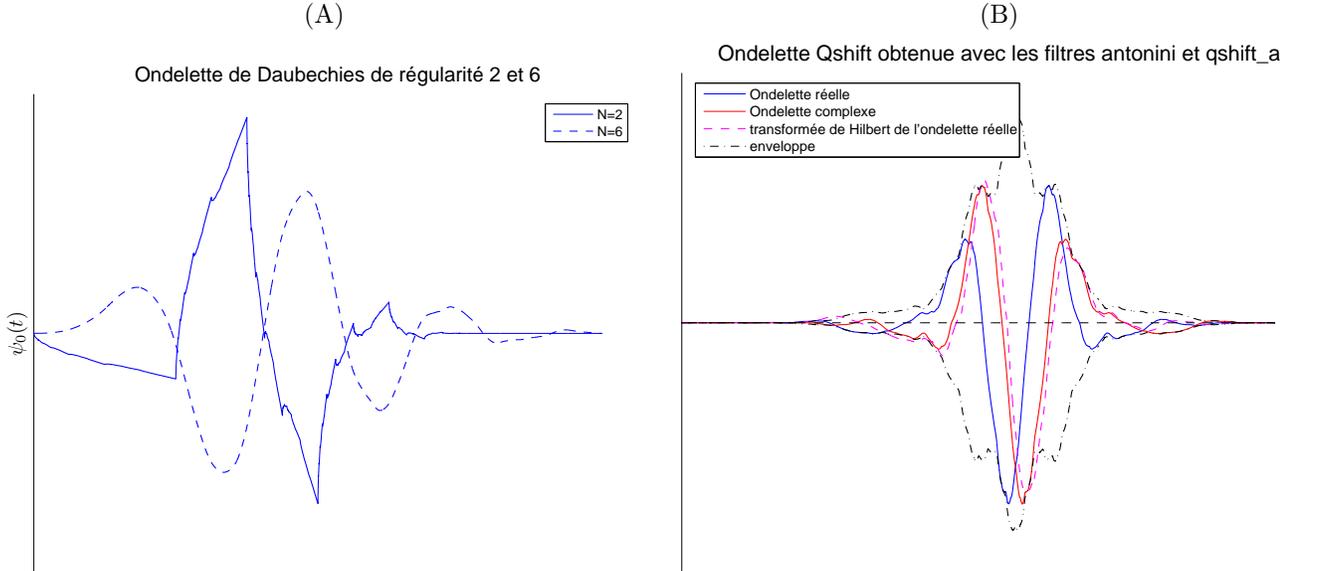
FIG. 2: Arbre représentant le fonctionnement de l'algorithme de Mallat. Figure extraite de [3]



Ingrid Daubechies a développé des ondelettes à support temporel compact dont on peut choisir la régularité. Deux

exemples de ces ondelette avec 2 et 6 moments nuls sont montrés sur le graphe (A) de la figure 3. Nous utiliserons ces ondelettes dans la suite comme base typique d'ondelettes réelles.

FIG. 3: Ondelettes de Daubechies de régularité 2 et 6 (A) et Qshift (B). Les filtres antonini et qshift_a sont le préfiltre et le filtre utilisé pour les étapes suivantes dans la procédure Dual Tree permettant l'analyse en ondelette complexe avec l'ondelette Qshift.



1.2 Transformation en ondelette complexe

1.2.1 Position du problème

La transformée en ondelette discrète pose plusieurs problèmes.

- Comme les ondelettes oscillent autour de zéro, on peut avoir des coefficients qui oscillent autour de zéro également autour d'un point où l'on analyse la régularité. Ce sera par exemple le cas lorsqu'on analyse une singularité de type dirac (voir la figure 5). On peut donc avoir de très faibles coefficients au voisinage d'une forte singularité.
- La DWT n'est pas invariante par translation temporelle à cause de la structure de la grille dyadique. Cette propriété est mise en évidence sur cette même figure 5.

La transformation en ondelette complexe est introduite pour pallier ces inconvénients. L'idée de base de la transformation en ondelette complexe est d'utiliser une ondelette analytique. Cependant, la condition de reconstruction parfaite imposée aux filtres pour pouvoir inverser la DWT, montre qu'il ne peut pas exister de filtres, même à coefficients complexes qui correspondent, à travers l'algorithme de Mallat, à l'analyse par une ondelette complexe ([3]).

1.2.2 Transformation complexe par simple transformation de Hilbert

La transformation de Hilbert d'un signal correspond dans le plan de Fourier à un filtrage de gain complexe $-i \operatorname{signe}(\nu)$ ([11]). D'après la théorie des distributions, cela correspond à une réponse impulsionnelle $\operatorname{vp}(\frac{1}{\pi t})$ où vp désigne la valeur principale au sens de Cauchy ([16]). On construit alors le signal analytique

$$z(t) = x(t) + i\mathcal{H}\{x(t)\} = x(t) + \frac{i}{\pi} \operatorname{vp} \int_{-\infty}^{\infty} \frac{x(s)}{t-s} ds, \quad (4)$$

qui ne possède que des fréquences positives. La transformée de Hilbert d'un signal réel est réelle et elle est en quadrature de phase par rapport au signal.

Plutôt que de prendre la transformée de Hilbert de l'ondelette (qui est définie par l'intermédiaire des filtres associés), nous pouvons prendre la transformée de Hilbert du signal, et l'analyser avec l'ondelette initiale, car $\langle f, \mathcal{H}\psi_{a,t} \rangle = \langle \mathcal{H}f, \psi_{a,t} \rangle$ puisque la transformation de Hilbert est un filtrage linéaire.

Le schéma est donc le suivant : on a un signal $X(n)$, on l'analyse avec une ondelette réelle par l'algorithme de Mallat pour obtenir les coefficients en ondelette $d_1(j, k)$. Ensuite, on analyse $\mathcal{H}X(n)$ par la même ondelette et on obtient les

coefficients $d_2(j, k)$. On construit les coefficients complexes : $d_{complexe}(j, k) = d_1(j, k) + i d_2(j, k)$. Dans la suite, le module de ces coefficients est nommé module Hilbert ou module par transformation de Hilbert.

Les inconvénients de cette technique sont de deux ordres :

- le support de la transformée de Hilbert d’une ondelette à support compact est infini. On perd donc en localité ;
- il y un désavantage calculatoire puisque le coût est de deux transformée en ondelette plus une transformée de Hilbert (soit deux fft).

Il est théoriquement possible de limiter l’inconvénient lié au support non compact de la transformée de Hilbert de l’ondelette en se contentant d’une transformation de Hilbert approchée. Cependant, l’approximation ne pourra pas alors être optimisée pour toutes les échelles car on ne fait la transformation de Hilbert qu’une seule fois au début.

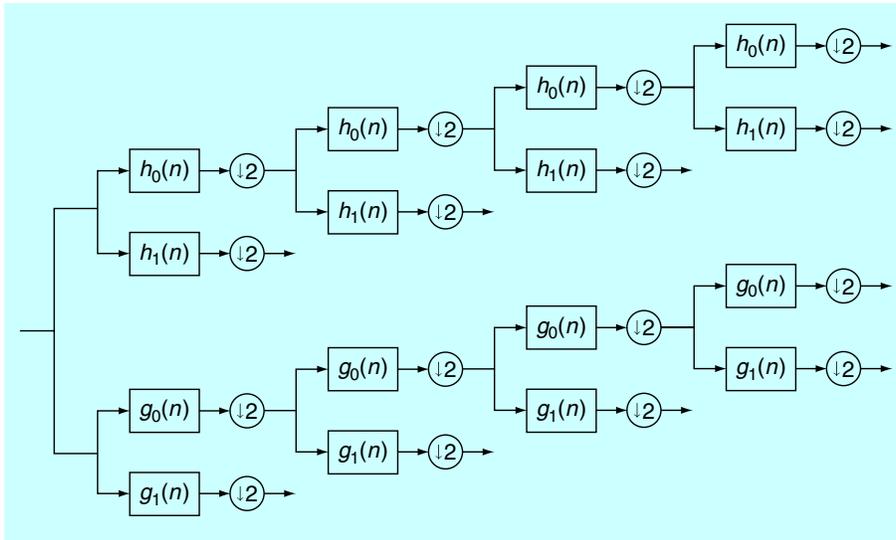
Une solution à ce problème a été proposée par Kingsbury : le Dual-Tree ([5, 3]).

1.2.3 Transformation de Hilbert complexe par la méthode Dual-Tree

La méthode Dual Tree (DT-CWT) consiste à faire l’analyse du signal par deux arbres de DWT différents, avec des filtres choisis de façon à ce qu’au final, cela revienne approximativement à la décomposition par une ondelette analytique. La figure 4 montre le fonctionnement de l’algorithme qui utilise deux jeux de filtres différents : h_1 et g_1 sont les filtres passe-haut du premier et du second arbre ; et h_0 et g_0 sont les filtres passe-bas du premier et du second arbre. Le premier arbre donne ce que l’on appellera les coefficients de la partie réelle $d_r(j, k)$ et le second arbre ceux de la partie imaginaire $d_i(j, k)$. On construit ensuite les coefficients complexes $d_{complexe}(j, k) = d_r(j, k) + i d_i(j, k)$. Le module de ces coefficient sera appelé module Dual Tree.

La synthèse des filtres adaptés à cette structure a été faite par plusieurs personnes. En particulier, Kingsbury propose des filtres appelés Qshift. Nous utiliserons par la suite l’un de ces jeux de filtres, dont l’utilisation est équivalente à l’analyse du signal par les ondelettes de la figure 3 (B). On peut voir sur cette figure que l’ondelette correspondant à l’arbre ”partie imaginaire” est très proche de la transformée de Hilbert de l’ondelette correspondant à l’arbre ”partie réelle”.

FIG. 4: Arbre représentant le fonctionnement de Dual Tree. Figure extraite de [3]



Enfin, l’utilisation de cette structure nécessite une opération de pré-filtrage, c’est à dire que les filtres utilisés à la première étape ne sont pas les mêmes que ceux utilisés aux étapes suivantes. La première échelle ne sera donc jamais prise en compte dans notre problème d’analyse.

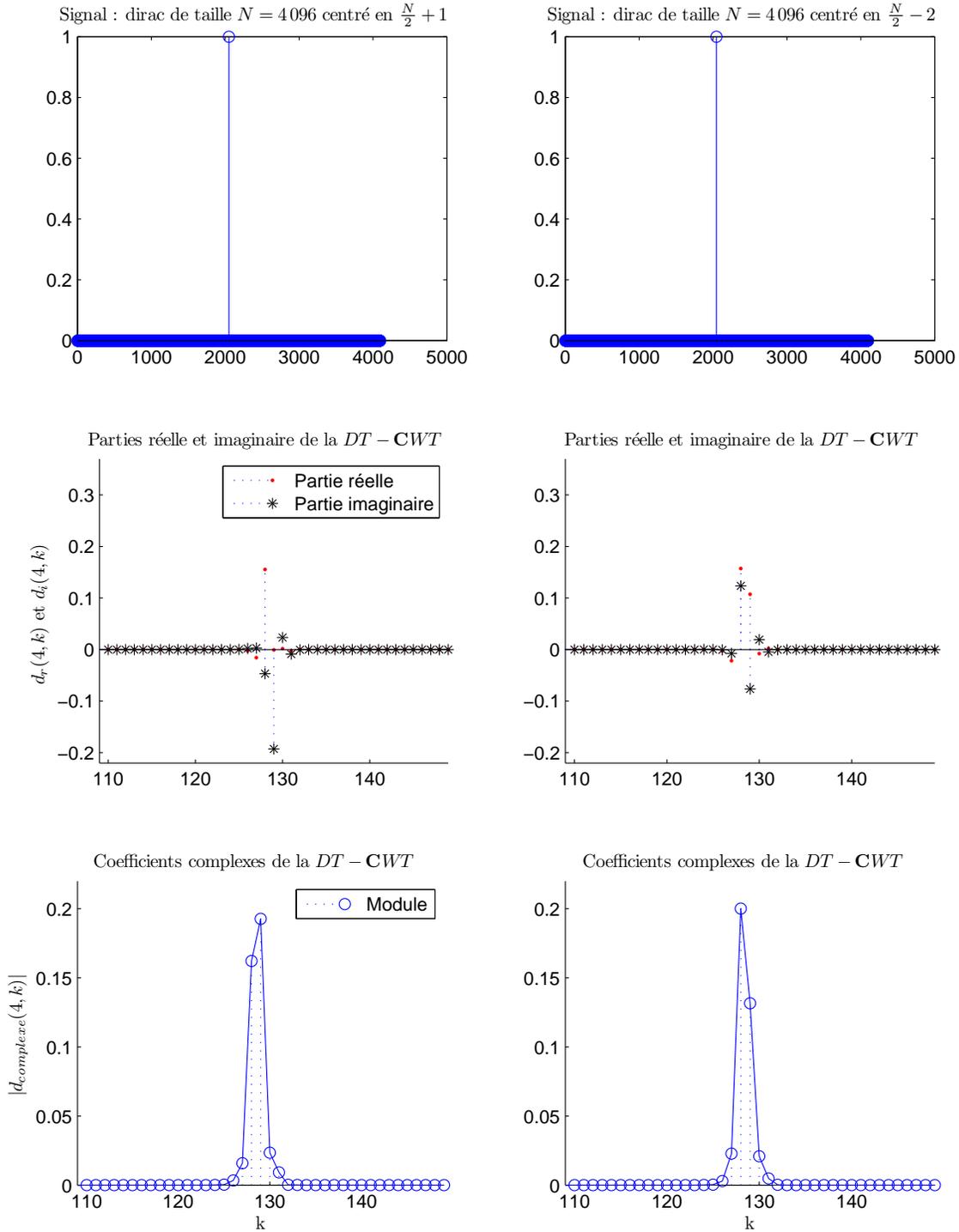
Les avantages de cette méthode par rapport à la simple transformation de Hilbert du signal (1.2.2) sont :

- un coût de calcul moindre (deux DWT seulement),
- une transformation de Hilbert approchée optimisée pour chaque échelle,
- La possibilité d’une reconstruction exacte est préservée.

L’inconvénient majeure est l’impossibilité d’utiliser les ondelettes bien connues de la DWT (ondelettes de Daubechies, spline . . .), et donc également de choisir le nombre de moments nuls (tous les filtres qshift donnent des ondelettes à deux moments nuls). Des filtres à nombre de moments nuls arbitraires adaptés à la structure Dual Tree ont été proposés par Selesnick ([2]), mais nous ne les avons pas utilisés.

1.2.4 Ondelettes complexes et invariance par translation temporelle

FIG. 5: Analyse d'un dirac placé au centre de l'intervalle (à gauche) et d'un dirac décalé vers la gauche (à droite). De haut en bas, on a : le signal ; les coefficients à l'échelle 4 de la partie réelle et de la partie imaginaire (l'échelle en x a été dilatée car les coefficients sur les côtés sont nuls) ; le module des coefficients complexes à l'échelle 4 issus de la DT-CWT.



La figure 5 montre l'analyse d'un dirac placé au centre de l'intervalle (à gauche) et d'un dirac décalé vers la gauche (à droite). À gauche, le dirac est au centre au sens où il y a des coefficients à la verticale de ce point à toutes les échelles (sauf

la dernière où il n'y a qu'un point). A droite, le signal est décalé de 3 échantillons vers la gauche. On doit donc suivre la grille en allant vers la gauche pour trouver les coefficients comme l'illustre les flèches sur la figure 1. La deuxième ligne de cette figure montre bien les oscillations autour de zéro des coefficients en ondelette de la partie réelle et de la partie imaginaire qui sont toutes deux des DWT. En comparant les deux colonnes de cette ligne, on voit bien la non invariance par translation temporelle puisque les coefficients de la partie réelle ont été beaucoup modifiés sous l'effet d'une faible translation du signal, de même pour la partie imaginaire. On voit à la troisième ligne de cette figure que ce problème est atténué si l'on considère le module des coefficients complexe. Les coefficients complexes représentés ici sont les coefficients issus de la $DT - CWT$. Le module Hilbert n'a pas été représenté, il est très proche du module DT.

1.3 Analyse d'une singularité par ondelette

1.3.1 Régularité ponctuelle et singularités

Régularité ponctuelle. La notion de *régularité ponctuelle* est une généralisation de la régularité d'une fonction (continuité, différentiabilité) à des valeurs non entières. Elle est caractérisée par l'exposant de Hölder défini dans l'encadré 1.

Encadré 1 : Régularité ponctuelle et exposant de Hölder

Régularité ponctuelle et exposant de Hölder

On dit que $f \in C^\alpha(t_0)$, avec $\alpha > 0$, s'il existe des constantes $C > 0$ et $\delta > 0$, ainsi qu'un polynôme P d'ordre strictement inférieur à α tel que

$$\text{si } |t - t_0| \leq \delta, \text{ alors } |f(t) - P(t - t_0)| \leq C|t - t_0|^\alpha.$$

Le polynôme P est alors unique, et si la fonction f admet un développement de Taylor de degré strictement inférieur à α au point t_0 et est $C^\alpha(t)$ sur un voisinage de t_0 , alors P est ce développement de Taylor. En particulier, si $\alpha < 1$, alors $P(t - t_0) = f(t_0)$.

Cette propriété est vérifiée pour tout un continuum de valeurs de α . On définit alors l'exposant de Hölder $h(t_0)$ de la fonction f au point t_0 comme la valeur maximale de α vérifiant cette propriété :

$$h(t_0) = \text{Sup}\{\alpha / f \in C^\alpha(t_0)\}.$$

L'exposant de Hölder est donc la régularité ponctuelle maximale de la fonction f autour du point t_0 .

Singularités. Il existe différents types de singularité. Le paradigme de la singularité est la singularité de type *cusp* centré en $t_0 : |t - t_0|^\alpha$. On peut montrer que l'exposant de Hölder de la fonction $f(t) = |t - t_0|^\alpha$ en t_0 est $h(t_0) = \alpha$, ce qui fait le lien entre la notion de singularité et la notion de régularité ponctuelle. Ainsi, les résultats concernant les ondelettes et la régularité ponctuelle pourront être appliqués à l'étude de singularités. La réciproque n'est pas vraie, c'est à dire qu'un exposant de Hölder h en t_0 n'implique pas un comportement en $|t - t_0|^h$ autour de t_0 .

1.3.2 Régularité ponctuelle et ondelette

Une présentation détaillée du lien entre la régularité ponctuelle et les ondelette peut être trouvée dans le chapitre de S. Jaffard dans [7].

Le résultat essentiel concerne la décroissance des coefficients en ondelette à travers les échelles au point t_0 : si le signal est de régularité ponctuelle h en t_0 , alors

$$c(t_0, a) \sim O(a^{h+\frac{1}{2}}), \quad a \rightarrow 0. \quad (5)$$

On voit donc que le passage à la DWT pose un problème dans la mesure où il est peu probable que l'abscisse t_0 coïncide avec une ligne de coefficients sur la grille dyadique. La transcription de cette loi de décroissance au cas de la DWT donne :

$$|d(j, k)| \sim 2^{j(h+\frac{1}{2})}, \quad j \rightarrow 0; \quad (6)$$

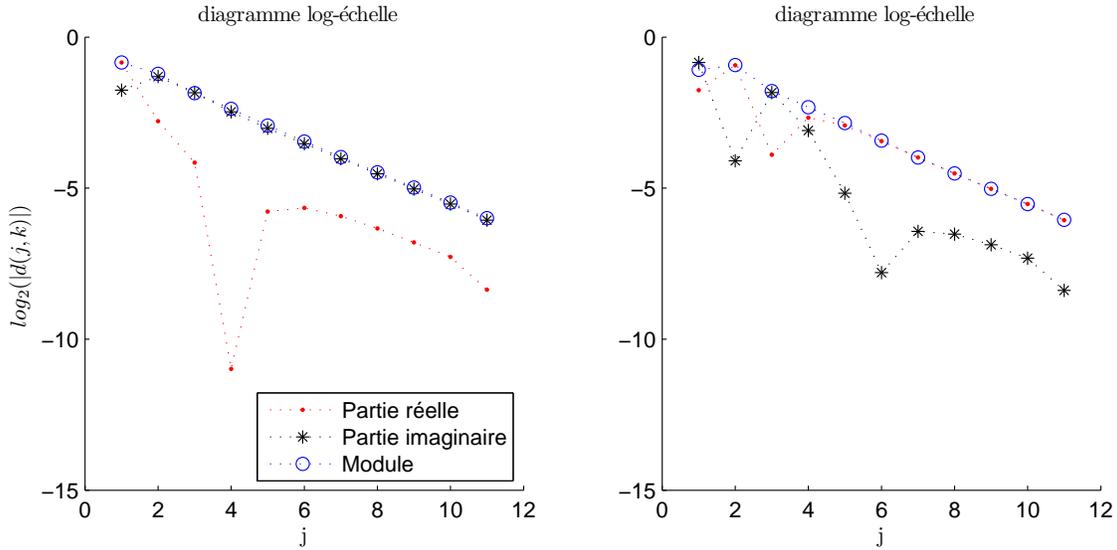
sachant que la condition $j \rightarrow 0$ est soumise à la restriction que $j \in \mathbb{N}$. C'est cette relation que nous allons utiliser malgré tout pour estimer h . Elle nous indique que la courbe $\log_2 |d(j, k)|$ contre $\log_2 2^j = j$ doit être une droite de pente $h + \frac{1}{2}$. C'est donc cette pente que nous allons estimer pour remonter à l'exposant de Hölder.

1.3.3 Comportement d'une singularité de type dirac : analyse par DWT et DT-CWT

On peut calculer théoriquement la décroissance des coefficients en ondelette d'un dirac centré en t_0 : $|c(a, t_0)| \propto \frac{1}{\sqrt{a}}$, ou en DWT $|d(j, k)| \propto 2^{j(-\frac{1}{2})}$. Ces coefficients se comportent comme ceux d'une fonction de régularité $h(t_0) = -1$ (ou d'une singularité de paramètre $\alpha = -1$ centrée en t_0).

Si l'on représente $\log_2|d(j, k)|$ en fonction de $\log_2 2^j = j$ (diagramme log-échelle), on s'attend donc à avoir une droite de pente $-\frac{1}{2}$. Ce diagramme log-échelle est présenté sur la figure 6, avec les trois type de coefficients : partie réelle, partie imaginaire et module. On voit bien ici encore la non invariance par translation temporelle de la DWT (partie réelle et partie imaginaire de la DT-CWT). En revanche, le module est bien linéaire quelle que soit la position du dirac. Le problème de non invariance par translation temporelle n'est donc plus un obstacle pour l'estimation de la pente dans cet exemple.

FIG. 6: Diagramme log-échelle pour un dirac avec partie réelle, partie imaginaire et module (c'est le module DT qui figure ici, le module Hilbert n'est pas représenté car il est très proche du module DT).



1.4 Autres jeux de coefficients multirésolutions et résumé

Dans la suite du travail, nous avons cherché à comparer différentes méthodes d'estimation de la régularité ponctuelle. En plus des transformations discrètes (réelle et complexe), nous avons regardé deux autres jeux de coefficients :

- la transformation continue-dyadique, en tant que méthode classique d'estimation ([12]),
- les leaders, connus pour être très efficaces dans l'estimation de propriétés globales de régularité (analyse multifractale), [14].

1.4.1 La transformation en ondelette continue dyadique (CDWT)

Elle correspond à un échantillonnage des échelles identique à celui de la DWT, mais on calcule les coefficients à tous les temps, c'est à dire qu'il n'y a pas de décimation.

1.4.2 Les leaders ou coefficients dominants

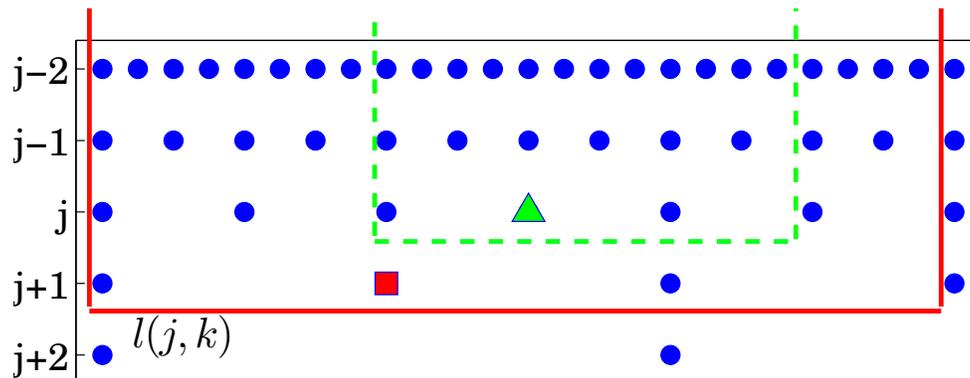
S. Jaffard a récemment proposé l'utilisation des leaders $l(j, k)$, définis à partir des coefficients en ondelette discrets $d(j, k)$ [14]. La construction des leaders s'effectue des petites vers les grandes échelles. Le leader $l(j, k)$ est défini comme le maximum, en valeur absolue, des coefficients $d(j, k-1)$, $d(j, k)$, $d(j, k+1)$, ainsi que de tous les parents de ces trois coefficients aux échelles plus fines (c'est à dire tous les coefficients à partir desquels on arrive sur un de ces trois points en descendant dans la grille dyadique comme l'indique les flèches de la figure 1). La figure 7 montre les domaines dans lesquels on prend le maximum pour deux points (le triangle et le carré) à deux échelles différentes.

TAB. 1: Résumé des différentes transformations disponibles. Les transformations disponibles sont notées par \times , les transformations non disponibles sont notées par $-$.

	Qshift		Daubechies		
	DWT	Leaders	DWT	Leaders	CDWT
partie réelle	\times arbre 1 de la DT-CWT	\times	\times DWT sur le signal	\times	\times CDWT sur le signal
partie imaginaire	\times arbre 2 de la DT-CWT	\times	\times DWT sur la \mathcal{TH} signal	\times	\times CDWT sur la \mathcal{TH} du signal
module Hilbert	\times à partir de la partie réelle	\times	\times	\times	\times
module DT	\times	\times	$-$	$-$	$-$

Prendre les leaders revient à remplacer en chaque point de la grille le coefficient en ondelette par le leader correspondant à ce point, alors $d(j, k) \rightarrow l(j, k)$. On peut prendre les leaders à partir d'un jeu de coefficient en ondelette quelconque : partie réelle, partie imaginaire, module Hilbert ou DT. Prendre les leaders revient à remplacer les coefficients par leur enveloppe convexe. Les résultats mathématique de S. Jaffard montrent que l'on peut utiliser les leaders à la place des coefficients en ondelette dans les diagrammes log-échelle et dans les régressions linéaires aboutissant à l'estimation de l'exposant de Hölder. Les leaders caractérisent même de façon plus efficace que les coefficients en ondelette discrets les propriétés de régularité ponctuelle et donc les exposants de Hölder d'une fonction.

FIG. 7: Grille dyadique et leaders. Le leader correspondant à la position du triangle vert (respectivement du carré rouge) est obtenu en prenant le maximum des valeurs absolues des coefficients dans le domaine à l'intérieur du cadre en pointillé vert (resp. en trait plein rouge) qui s'étend vers le haut jusqu'à la première échelle considérée.



1.4.3 Résumé des coefficients multirésolutions envisagées dans la suite

Nous avons utilisé deux ondelettes différentes : l'ondelette Qshift de Kingsbury, qui a une régularité 2 ; et une ondelette de Daubechies de régularité 2 également.

L'ondelette Qshift est implémentée pour la procédure DT-CWT ; alors que l'ondelette de Daubechies est implémentée pour les procédures DWT et CDWT.

Le tableau 1 résume les différentes transformations disponibles pour la construction d'estimateurs de lois d'échelle.

2 Étude des estimateurs de lois d'échelle

2.1 Le mouvement brownien fractionnaire (fBM) et l'estimation du paramètre global H par la marginale

Dans cette partie, nous définissons le fBM, que nous utiliserons par la suite comme processus de référence pour caractériser les estimateurs de régularité locale. Nous montrons aussi comment on peut estimer l'exposant global H d'autosimilarité grâce à l'analyse en ondelettes du signal : c'est l'estimation par la marginale. Cette méthode est bien connue dans le cas des ondelettes réelles ([6]), pour lesquelles des résultats théoriques justifient la méthode d'estimation. Sans garantie théorique du fonctionnement de cette méthode avec les ondelettes complexes, nous allons l'appliquer pour vérifier expérimentalement qu'elle aboutit à une valeur acceptable de H .

2.1.1 Le mouvement brownien fractionnaire (fBM) : définition et propriétés

Le fBM de paramètre H est l'unique processus aléatoire gaussien autosimilaire (de paramètre H) à accroissement stationnaire ([10]). La particularité du fBM est que l'on a $h(t) = H$ presque sûrement, c'est à dire que l'exposant de Hölder est constant, égal au paramètre d'autosimilarité H . En fait, l'ensemble des singularités qui ne sont pas de type cusp (de paramètre $\alpha = H$) est de mesure de Lebesgues nulle.

Le fBM est donc un excellent modèle où $h(t)$ est constant pour étudier le comportement des estimateurs de régularité locale, avant de passer au cas où celle-ci est variable.

Il existe de nombreuses méthodes de synthèse du fBM (voir annexe A dans [4]). Nous utilisons la synthèse par matrice circulante qui permet la synthèse rapide de signaux de grande taille. Les traces synthétisées sont de taille $N = 2^{17} = 131\,072$ points.

La caractéristique d'autosimilarité, comme la régularité Höldérienne locale (relation 6 dans la partie 1.3.2) se traduit par des propriétés particulières concernant la décroissance des coefficients en ondelette (ou de leurs moments) à travers les échelles.

En effet, pour un processus autosimilaire (de paramètre H) à accroissements stationnaires, les moments d'ordre q des coefficients en ondelette $d(j, k)$ à l'échelle j et au temps k , s'ils existent, vérifient

$$\mathbb{E}\{|d(j, k)|^q\} = 2^{jq(H+\frac{1}{2})} \underbrace{\mathbb{E}\{|d(0, 0)|^q\}}_{\text{constante}} \quad (7)$$

Cela signifie en fait que pour un processus autosimilaire à accroissements stationnaires tel que le fBM, la relation 6 n'est plus vraie seulement dans la limite des petites échelles mais est vraie partout.

C'est cette relation qui est à l'origine de la méthode d'estimation de H par la marginale. Cette méthode consiste à estimer l'exposant de la loi de puissance (7) pour $q = 2$.

2.1.2 Estimation de H par la marginale

Cette méthode est une méthode d'estimation globale du paramètre H du fBM. Elle a déjà été étudiée avec les ondelettes réelles ([9, 6] et chapitre de ABRY P., GONÇALVÈS P. et FLANDRIN P. dans [8]). C'est une méthode performante pour l'estimation du paramètre H .

Construction de l'estimateur

L'équation 7 suggère une façon simple d'estimer le paramètre H : la mesure de la pente d'un diagramme $\log_2 \mathbb{E}\{|d(j, k)|^2\}$ contre $\log_2(2^j) = j$ doit être $(2H + 1)$, elle donne donc une estimée $(2\hat{H} + 1)$. Cela implique, à partir d'une réalisation observée du fBM, d'estimer la quantité $\mathbb{E}\{|d(j, k)|^2\}$. Du fait de la propriété d'accroissements stationnaires, les coefficients en ondelette du fBM sont stationnaires à chaque échelle. Donc le moment d'ordre 2 peut être estimé par la simple moyenne temporelle : $\hat{\mathbb{E}}\{|d(j, k)|^2\} = \frac{1}{n_j} \sum_k |d(j, k)|^2$

Il ne reste plus alors qu'à faire une regression linéaire pour trouver la pente de la droite qui est $2\hat{H} + 1$ pour le moment d'ordre 2. L'annexe A explique comment faire cette régression linéaire à partir des abscisses $X_j = j$ (déterministes) et des ordonnées $Y_j = \log_2(\frac{1}{n_j} \sum_k |d(j, k)|^2)$.

Dans le cas de l'estimation par la marginale utilisant la transformée en ondelette discrète, nous pouvons utiliser une régression linéaire pondérée par le nombre de coefficients à chaque échelle. Cela signifie que nous accordons plus de confiance à l'estimation du moment si l'on dispose de d'avantage de coefficients pour l'estimer, c'est à dire pour les petites échelles

puisque le nombre de coefficients est divisé par deux chaque fois que l'échelle augmente d'une unité. On ne tient pas compte à priori de la corrélation entre les coefficients dans la régression linéaire (la matrice V de l'annexe A est diagonale).

On obtient ainsi un estimateur $\widehat{H}_{marg}(\omega_n)$ qui à chaque réalisation de fBM ω_n associe une valeur de H estimée. Pour étudier empiriquement le comportement de cet estimateur, on fait des moyennes d'ensemble sur N_{real} réalisations pour obtenir la moyenne ($\mathbb{E}\widehat{H}_{marg} = \frac{1}{N_{real}} \sum_{n=1}^{N_{real}} \widehat{H}_{marg}(\omega_n)$) et la variance ($var\widehat{H}_{marg} = \frac{1}{N_{real}} \sum_{n=1}^{N_{real}} (\widehat{H}_{marg}(\omega_n) - \mathbb{E}\widehat{H}_{marg})^2$) de l'estimateur. Nous avons utilisé $N_{real} = 100$ réalisations.

Mise en œuvre : domaine de régression linéaire et effets de bords

Cet estimateur dépend de la plage de régression linéaire choisie, c'est à dire de j_{min} et j_{max} . Pour le fBM, on a vu que la loi de puissance (7) est toujours valable, cependant nous allons nous restreindre aux petites échelles car notre but est de passer à l'estimation de régularité locale. En pratique, on a pris ici $j_{min} = 3$ et $j_{max} = 6$. Nous avons commencé la régression à l'échelle 3 seulement car le préfiltrage de la DT-CWT (1.2.3) oblige à ne pas considérer la première échelle d'une part, et d'autre part, pour que les leaders aient un sens, on doit prendre un maximum sur au moins deux échelles. Le signal étant de taille 2^{17} , on a $2^{11} = 2048$ coefficients à la dernière échelle ($j_{max} = 6$).

Cependant, les premiers coefficients subissent des effets de bords dus à la taille des filtres. Notons M la taille du plus grand filtre utilisé pour la décomposition en ondelette (pour l'ondelette Qshift, $M = 10$). Après le filtrage de l'approximation à l'échelle $j - 1$ par le filtre h_1 , il y a M points de chaque côté pollués par les effets de bords. Après la décimation d'un facteur 2, il reste $\lceil \frac{M}{2} \rceil$ points pollués dans les coefficients en ondelette à l'échelle j , c'est à dire 5 points pour l'ondelette Qshift. Cependant, le premier de ces points, correspondant à $k = \lceil \frac{M}{2} \rceil + 1$ n'est pas acceptable non plus. En effet, le calcul du leader correspondant obligerait à le comparer à son voisin de gauche ($k = \lceil \frac{M}{2} \rceil$), qui est un point pollué. Au final, les $\lceil \frac{M}{2} \rceil + 1$ premiers points de chaque échelle ne doivent pas être utilisés car ils sont affectés par les effets de bord.

Résultats

La figure 8 montre le biais et l'écart type de l'estimateur \widehat{H}_{marg} pour les différents jeux de coefficients envisagés (voir table 1) pour H allant de 0.1 à 0.9 par pas de 0.1. Ces quantités ont été estimées à partir de $N_{real} = 100$ réalisations de fBM. L'estimation à partir de la partie imaginaire n'apparaît pas sur cette figure car elle est identique à l'estimation par la partie réelle. L'estimation par la CDWT n'apparaît pas non plus car elle a des caractéristiques identiques à l'estimation par la DWT avec la même ondelette. Il en sera de même dans toute la suite, sauf pour l'estimation d'une régularité locale variable.

On constate sur cette figure que l'estimateur \widehat{H}_{marg} est d'autant plus biaisé que H est faible, et il est très peu biaisé pour les grands H . Sa variance est très faible, il est donc précis.

Du point de vue de la comparaison entre partie réelle et module, on constate que le comportement est tout à fait similaire, ce qui montre que les coefficients complexes ont un comportement analogue aux coefficients réels. Enfin le comportement du module Dual Tree et du module par transformation de Hilbert directe est strictement identique. Il apparaît donc que les bénéfices de la procédure Dual Tree, essentiellement axés sur la reconstruction parfaite, ne jouent aucun rôle pour l'analyse.

Conclusion

L'étude de l'estimateur par la marginale nous a permis de montrer que l'estimation de régularité en utilisant des coefficients en ondelette complexe est légitime. En particulier on a vu ici que, bien que rien ne prouve théoriquement que la régularité soit conservée par transformation de Hilbert (elle ne l'est pas toujours) on peut utiliser le module de Hilbert directement pour l'estimation de H .

2.2 Estimation de la régularité locale

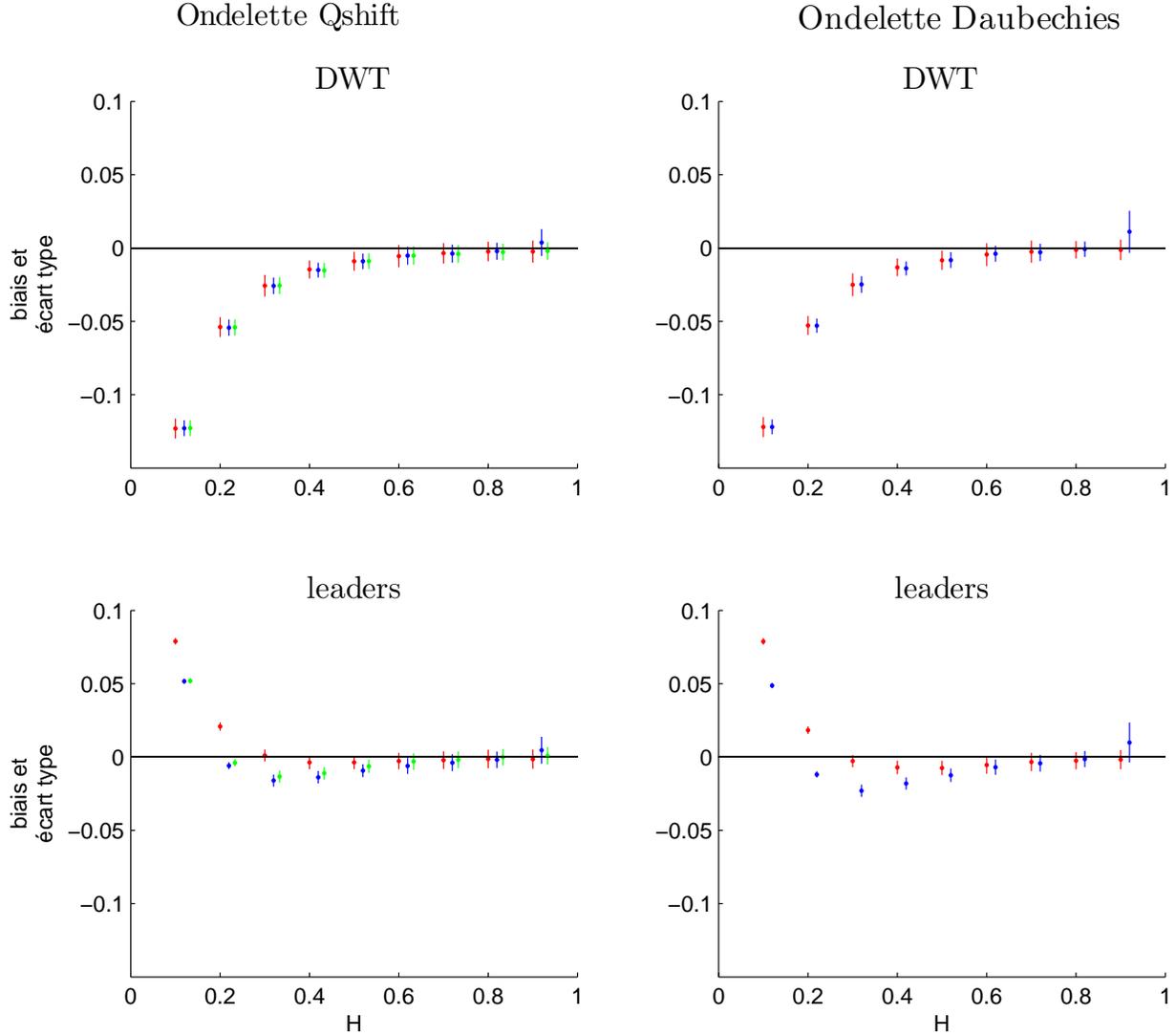
On a vu dans la partie 2.1.1 que pour le mouvement brownien fractionnaire, on a une régularité locale constante égale au paramètre H . On va ici chercher à caractériser les estimateurs de régularité locale basés sur la régression linéaire dans un diagramme log-échelle.

2.2.1 Définition de l'estimateur $\widehat{h}(t)$

L'idée pour estimer la régularité locale à partir des ondelettes est d'utiliser la décroissance à travers les échelles des coefficients en ondelette dans la limite des petites échelles. Cette décroissance est contrôlée par la relation 6 qui a été vu

FIG. 8: Biais et écart type de l'estimateur par la marginale en fonction de H pour les différentes transformations étudiés. Pour chaque valeur de H , on représente (de gauche à droite) l'estimation d'abord avec la partie réelle des coefficients (rouge), puis avec le module Hilbert (bleu), et enfin avec le module Dual Tree (vert) lorsqu'il est disponible (donc pas dans la colonne de droite). Ces deux ou trois points ont été décalés pour augmenter la lisibilité, mais ils correspondent à la même valeur de H (multiple de 0.1).

Estimation du paramètre H d'un fBM par la marginale en fonction de H : \widehat{H}_{margin}



au paragraphe 1.3.2. On peut accéder à $h(t_0)$ par la pente valant théoriquement $h + \frac{1}{2}$ dans un diagramme log-échelle. L'estimation \hat{h} est donc obtenue par regression linéaire (cf annexe A) avec comme ordonnées $Y_j = \log_2 |d(j, k)|$ (avec $k = k(j, t_0) = \lceil \frac{t_0}{2^j} \rceil$, voir figure 1) et comme abscisses $X_j = j$. Aucune pondération n'est introduite ici.

Mise en œuvre

Dans la mesure où nous cherchons à caractériser la régularité locale, nous devons plus que jamais ici nous limiter aux échelles de temps court pour la régression. Nous avons choisi de ne pas dépasser $j_{max} = 6$. De même qu'auparavant, nous avons pris $j_{min} = 3$. Cela laisse trois possibilités de plages de régression : $[3; 4]$, $[3; 5]$ ou $[3; 6]$.

Pour la gestion des effets de bords, nous avons estimé h uniquement aux points du signal dont les enfants à la sixième échelle ne sont pas pollués. Nous avons donc exclu de l'estimation $6 \cdot 2^6 = 384$ points de chaque côté du signal, ce qui laisse 130 304 points dans le signal et 2 036 coefficients non pollués à l'échelle 6.

On peut alors estimer h en chacun des points restant. Cependant, pour des raisons de coût de calcul, nous nous sommes limité à 4 000 points répartis de façon à explorer l'ensemble des positions possibles relativement aux coefficients de l'échelle 6 dans la grille dyadique.

Notre objectif est d'étudier le comportement empirique de l'estimateur \hat{h} , et de préciser l'effet du passage au module des coefficients complexes par rapport aux coefficients réels.

2.2.2 Résultats

Histogrammes des h estimés

La première étape consiste à regarder la distribution des $\hat{h}(t)$. La figure 9 montre les histogrammes des $\hat{h}(t)$ pour un fBM de paramètre $H = 0.7$, avec les différents jeux de coefficients en ondelette. Ces histogrammes ont été obtenus en régressant sur [3; 6]. Les valeurs des écarts type de ces histogrammes peuvent être consultées dans le tableau 4 de l'annexe B. Sur ces histogrammes, le module Dual Tree n'apparaît pas, même dans les cas où il est disponible car il est similaire au module par transformée de Hilbert directe, et donc pour ne pas surcharger les histogrammes, nous ne l'avons pas représenté. On voit bien sur ces histogrammes que l'écart type est plus faible avec le module (en rouge) qu'avec la partie réelle (en bleu). Cet effet est cependant moins spectaculaire avec les leaders. On voit également que l'écart type est bien plus faible avec les leaders qu'avec les deux autres méthodes. Cela est normal car les leaders étant une sous-série des coefficients en ondelette, ils sont naturellement moins variables.

Moyennes

Sur les histogrammes de la figure 9 apparaissent sous forme de traits verticaux, en noir pointillé la valeur de l'exposant H , et en bleu et rouge les valeurs des moyennes des $\hat{h}(t)$ (pour la partie réelle et le module). On voit que les moyennes de $\hat{h}(t)$ sont très proches de H puisque tous ces traits sont superposés. En fait, une étude plus poussée a montré que l'on obtient pour le biais de l'estimateur $\widehat{H_{moyenne}} = \langle \hat{h}(t) \rangle_t$ (moyenne temporelle des \hat{h} en chaque point du signal) les mêmes courbes que pour l'estimateur par la marginale (figure 8). Cela valide la possibilité d'estimer la régularité locale à l'aide des coefficients en ondelette complexes.

Écarts types

Pour préciser le gain obtenu en passant de la partie réelle au module, il faut étudier les valeurs exactes de l'écart type des distributions. Pour cela, on génère 100 réalisations de fBM de même paramètre H , puis on moyenne l'écart type de \hat{h} sur ces 100 réalisations.

Les tableaux 2, 3 et 4 de l'annexe B montrent les valeurs des écarts type de $\hat{h}(t)$ pour les trois domaines de régression [3; 4], [3; 5] ou [3; 6], pour deux valeurs de H . Dans ces tableaux, nous avons représenté l'écart type, et non la variance, de façon à pouvoir comparer avec les histogrammes. L'annexe A explique la façon dont doit se comporter la variance de la pente en fonction de la matrice de covariance (variances et corrélations) des ordonnées (les abscisses sont déterministes). Dans cette annexe, on rappelle la formulation de l'estimateur de la pente d'une droite par les moindres carrés. Puis, dans le cas où la matrice V est diagonale (c'est à dire qu'on ne tient pas compte a priori des corrélations dans les ordonnées, ce qui est notre cas), on calcule la variance de cet estimateur en fonction de la matrice de covariance effective des ordonnées. Pour interpréter la réduction de variance en passant de la partie réelle au module, nous devons donc nous intéresser à la variance des ordonnées, c'est à dire du \log en base 2 de la valeur absolue des coefficients en ondelette; c'est l'objet de la section suivante.

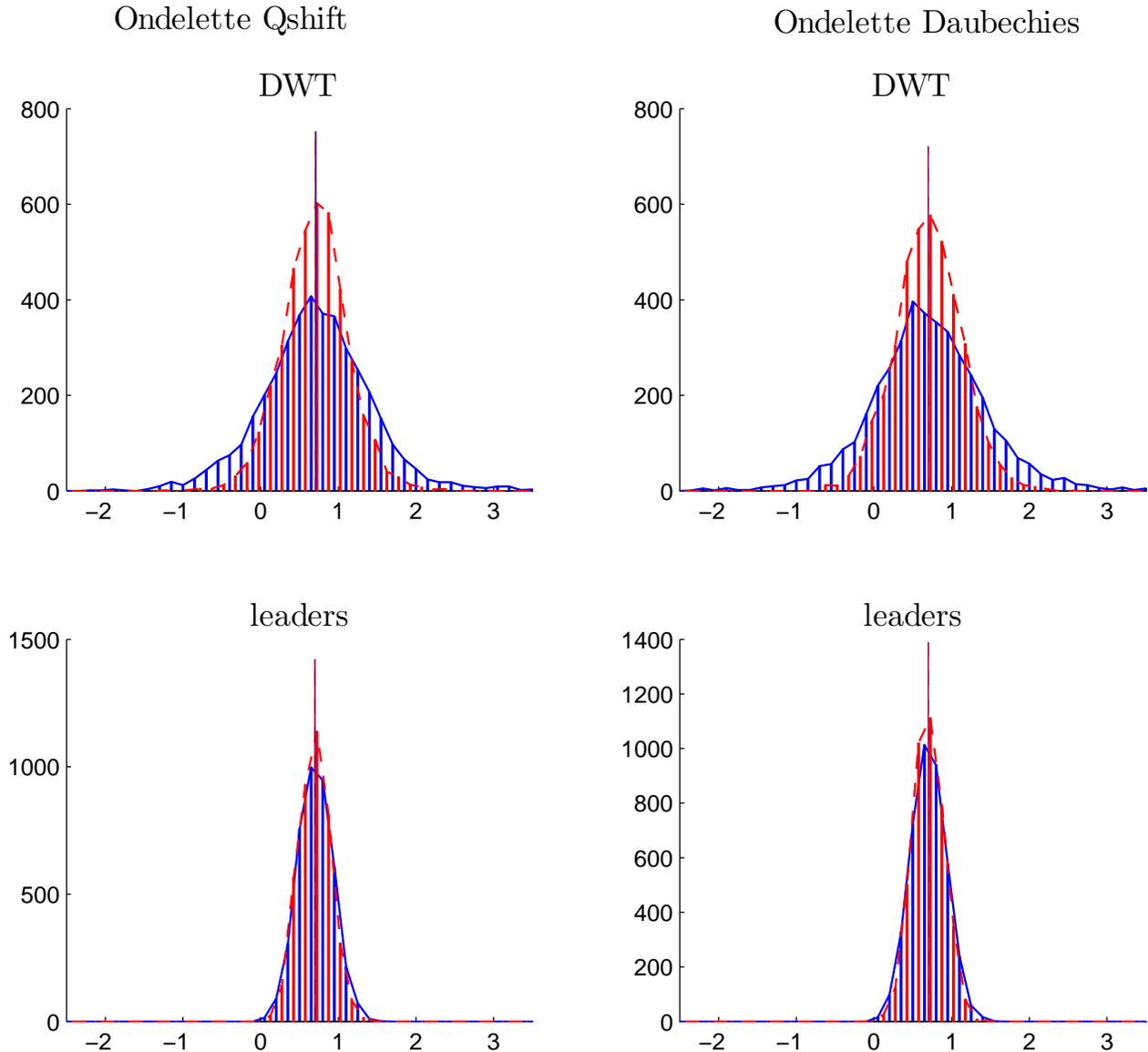
2.3 Interprétation de la variance de $\hat{h}(t)$

2.3.1 Distribution des coefficients en ondelette

Le haut de la figure 10 (page 27) montre la distribution des coefficients en ondelette $d(j, k)$ aux échelles $j = 3$ et 4, dans les différentes situations. Pour la partie réelle (en bleu, trait plein), les coefficients peuvent être positifs ou négatifs et les distributions sont centrées en 0. Par contre, les leaders sont toujours positifs puisqu'ils correspondent à des suprémums sur les valeurs absolues des coefficients en ondelettes. Les distributions des modules (en rouge, pointillé) sont également toujours à droite de 0. Le bas de la figure montre le \log_2 de la valeur absolue des distributions du haut.

FIG. 9: Histogrammes des h estimés avec les différentes méthodes. L’histogramme dont l’enveloppe est représentée en trait plein (rouge) correspond à l’estimation avec la partie réelle. L’histogramme dont l’enveloppe est représentée en trait pointillé (bleu) correspond à l’estimation avec le module Hilbert (le module DT ne figure pas car il est très proche du module Hilbert). La moyenne de h estimé par la partie réelle, la moyenne de h estimé par le module, et le paramètre H sont représentés par des traits verticaux. Les trois traits sont indiscernables, ces trois quantités sont donc très proches. Pour ces estimations, la régression linéaire a été faite de 3 à 6. Les valeurs des écarts types de ces distributions sont dans le tableau 4 de l’annexe B.

Histogrammes des $\hat{h}(t)$. $H = 0.7$



Pour la méthode DWT la distribution de la partie réelle est gaussienne. L’annexe C explique le lien entre la distribution des $d(j, k)$ et la distribution des $\log_2|d(j, k)|$. Dans la première partie de cette annexe, on étudie la moyenne et la variance du \log_2 de la valeur absolue d’une variable aléatoire gaussienne centrée en zéro. Cela décrit parfaitement les coefficients en ondelette $d(j, k)$ à une échelle j . En particulier, on peut voir que alors que la variance des coefficients $d(j, k)$ décroît à travers les échelle comme une loi de puissance dont l’exposant est fonction de H (voir l’estimation par la marginale, partie 2.1.2 et la relation 7), la variance de $\log_2|d(j, k)|$ reste constante, égale à $\frac{\pi^2}{8(\ln 2)^2} \approx 2.568$ soit un écart type de environ

1.602.

On peut constater sur les histogrammes de la figure 10 (page 27) que la variance des distributions du \log_2 du module est inférieure à la variance des distributions du \log_2 de la partie réelle ; c'est à dire pour la régression linéaire aboutissant à $\hat{h}(t)$ que la variance des ordonnées est plus faible avec le module qu'avec la partie réelle. Pour savoir si cette réduction de variance des ordonnées est la seule cause de la réduction de variance des estimées, il faut faire une étude numérique précise de la réduction de variance. Si la réduction de variance de $\log_2|d(j, k)|$ n'explique pas la réduction de variance de $\hat{h}(t)$, il faudra étudier les corrélations des ordonnées.

2.3.2 Réduction de l'écart type de $\hat{h}(t)$ $\sigma_{\hat{h}}$ due à la réduction de variance du \log_2 des coefficients en ondelette

La figure 11 (page 28) illustre la réduction de $\sigma_{\hat{h}}$ par passage de la partie réelle au module pour les trois domaines de régression linéaire envisagés. On part de $\sigma_{\hat{h}}$ pour l'estimation avec la partie réelle, ramené à 1 (points magenta d'ordonnée 1). La première source de réduction de $\sigma_{\hat{h}}$ est liée à la réduction de l'écart type des ordonnées des points de régression. Cette réduction est symbolisée sur la figure 11 par la première flèche, à gauche des points. Elle amène aux points cyan. Ces points représentent donc le pourcentage de réduction de $\sigma_{\hat{h}}$ expliqué par la réduction de l'écart type sur les ordonnées (le \log_2 des coefficients en ondelette). Pour préciser ce point, nous devons séparer le cas des leaders du cas de la DWT.

Sans les leaders : DWT

Pour cette méthode, il n'y a pas de corrélation entre les coefficients aux différentes échelles (avec la partie réelle), ou alors cette corrélation intra-échelles est négligeable. Donc la matrice de covariance des ordonnées (*COV* dans l'annexe A) est diagonale. De plus, d'après la partie 2.3.1 et l'annexe C, la variance des ordonnées est constante : $(\sigma_j^{reelle, sans leaders})^2 = contante = (\sigma_{reelle, sans leaders})^2 = \frac{\pi^2}{8(\ln 2)^2}$. D'après l'annexe A (equation 12), on a donc $var(\hat{h}_{reelle, sans leaders}) = \sigma_3^2 \cdot w_3^2 + \sigma_4^2 \cdot w_4^2 + \dots$.

Encadré 2 : w pour les différents domaines de régression linéaire

w pour les différents domaines de régression linéaire

Les coefficients w_j , définits dans l'annexe A, ne dépendent que des abscisses et sont tels que l'estimée de la pente en fonction des ordonnées Y_j s'écrit : $\hat{h} = \sum_{j=j_{min}}^{j_{max}} w_j \cdot Y_j$ (relation 11 de l'annexe A).

sur [3; 4] : $w_3 = -1$ et $w_4 = 1$.

sur [3; 5] : $w_3 = -\frac{1}{2}$, $w_4 = 0$ et $w_5 = \frac{1}{2}$.

sur [3; 6] : $w_3 = -\frac{3}{10}$, $w_4 = -\frac{1}{10}$, $w_5 = \frac{1}{10}$ et $w_6 = \frac{1}{10}$.

On voit donc que dans tous les cas $var(\hat{h}_{reelle, sans leaders}) = \left(\sum_j w_j^2\right) \cdot (\sigma_{reelle, sans leaders})^2$. Ainsi,

$$var(\hat{h}_{reelle, sans leaders}) \propto (\sigma_{reelle, sans leaders})^2. \quad (8)$$

Avec les leaders

Avec les leaders, même avec la partie réelle, les coefficients à travers les différentes échelles sont corrélés puisqu'on prend un coefficient dominant sur plusieurs échelles. Cependant, la variance du \log_2 des coefficients est toujours constante (ceci n'est plus une conséquence directe de l'annexe C car la distribution n'est plus une gaussienne centrée en 0, c'est donc ici une observation). D'ailleurs, il en va de même pour la variance des distributions du \log_2 des modules qui sont constantes également. On a donc

$$(\sigma_j^{reelle, avec leaders})^2 = contante = (\sigma_{reelle, avec leaders})^2.$$

On voit donc avec l'équation 12 de l'annexe A que dans tous les cas

$$var(\hat{h}_{reelle, avec leaders}) = \left(\sum_{(i,j)} \rho_{ij} \cdot w_j^2\right) \cdot (\sigma_{reelle, avec leaders})^2.$$

Ainsi,

$$\text{var}(\widehat{h}_{reelle, avec\ leaders}) \propto (\sigma_{reelle, avec\ leaders})^2. \quad (9)$$

Conclusion

Avec ou sans les leaders, si les corrélations des ordonnées sont inchangées, une réduction de l'écart type des ordonnées ($\log_2|d(j,k)|$) se traduit par une réduction de l'écart type de $\widehat{h}(t)$ d'un même facteur. Les points cyans de la figure 11 représente donc le facteur de réduction de $\sigma_{\widehat{h}}$ attendu en passant au module si la corrélation dans les ordonnées (c'est dire dans les \log_2 des coefficients aux différentes échelles) était inchangée.

Les derniers points (noirs) de la figure 11 correspondent aux facteurs de réduction de l'écart type de $\widehat{h}(t)$ effectivement observé. On constate que pour une régression sur [3; 4], la réduction de variance est meilleure que celle attendue par le simple effet de la réduction de variance des ordonnées (flèche descendante sur la droite), pour une régression sur [3; 5] elle est identique (pas de flèche sur la droite), et pour une régression sur [3; 6] elle est moins bonne (flèche montante sur la droite).

2.3.3 Corrélation introduite par le module dans les coefficients

La deuxième source de modification de $\sigma_{\widehat{h}}$ qui explique l'écart entre la réduction réelle de $\sigma_{\widehat{h}}$ et la réduction expliquée par la réduction de variance sur les ordonnées est l'apparition (ou l'augmentation) des corrélations $\rho_{i,j}$. Les corrélations à prendre en compte ici sont les corrélations entre les ordonnées, donc entre les \log_2 des valeurs absolus des coefficients. Dans l'annexe C, on calcule dans le cas simple de deux variables aléatoires conjointement gaussiennes centrées en 0 et corrélées, la corrélation de leurs valeurs absolus et de leurs \log_2 . On montre en particulier que les \log_2 des valeurs absolues sont moins corrélés que les valeurs absolues.

Sans les leaders

L'étude des matrices de covariance effectives des ordonnées montre que les \log_2 des coefficients (en module) présentent des corrélations constantes uniquement pour deux échelles successives, et pas au-delà. On a ainsi :

$$\rho_{34} = \rho_{45} = \rho_{56} = \rho \text{ et } \rho_{kl} \approx 0 \text{ si } |k - l| \geq 2$$

Avec l'ondelette Qshift, $\rho_{module, sans\ leaders} \approx 0.11$ ne dépend pas de H , alors qu'avec l'ondelette de Daubechies de régularité 2, $\rho_{module, sans\ leaders}$ varie de environ 0.2 pour $H = 0.1$ à environ 0.3 pour $H = 0.9$.

Ceci permet de comprendre la différence entre la réduction d'écart type attendu du fait de la réduction de variance des ordonnées, et la réduction de variance observée qui tient compte de cette corrélation. En effet, à partir des vecteurs w (cf encadré 2 page 17), on peut voir que l'expression générale

$$\frac{\text{var}(\widehat{h}_{module, sans\ leaders}(t))}{\text{var}(\widehat{h}_{reelle, sans\ leaders}(t))} = \frac{(\sigma_{module, sans\ leaders} \cdot \sum_{(i,j)} \rho_{i,j, module, sans\ leaders} w_i w_j)^2}{(\sigma_{reelle, sans\ leaders} \cdot \sum_j w_j^2)^2}$$

devient :

sur [3; 4] :

$$\frac{\text{var}(\widehat{h}_{module, sans\ leaders}(t))}{\text{var}(\widehat{h}_{reelle, sans\ leaders}(t))} = \frac{(\sigma_{module, sans\ leaders})^2}{(\sigma_{reelle, sans\ leaders})^2} \cdot (1 - \rho),$$

d'où une corrélation qui améliore la réduction de variance attendue.

sur [3; 5] :

$$\frac{\text{var}(\widehat{h}_{module, sans\ leaders}(t))}{\text{var}(\widehat{h}_{reelle, sans\ leaders}(t))} = \frac{(\sigma_{module, sans\ leaders})^2}{(\sigma_{reelle, sans\ leaders})^2} \cdot 1.$$

La corrélation ne joue aucun rôle puisqu'elle se limite à deux échelle successives et que $w_4 = 0$ (le point du milieu n'intervient pas, la seule différence par rapport au cas où on régresse sur [3; 4] est que la différence entre j_{min} et j_{max} a été multipliée par deux.

sur [3; 6] :

$$\frac{\text{var}(\hat{h}_{\text{module, sans leaders}}(t))}{\text{var}(\hat{h}_{\text{reelle, sans leaders}}(t))} = \frac{(\sigma_{\text{module, sans leaders}})^2}{(\sigma_{\text{reelle, sans leaders}})^2} \cdot \left(1 + \frac{\rho}{2}\right),$$

d'où une corrélation qui dégrade la réduction de variance attendue du fait qu'on ait une corrélation à deux échelles seulement et pas au delà.

On peut constater sur la figure 11 que ces formules sont bien corroborées expérimentalement. On voit bien en particulier qu'on "perd" pour une régression sur [3; 6] la moitié de ce qu'on avait gagné pour une régression sur [3; 4] du fait des corrélations. Les gains (ou pertes) de réduction de variance dus à la corrélation sont plus importants pour l'ondelette de Daubechies puisque la corrélation est plus forte.

Avec les leaders

Avec les leaders, les corrélations sont plus importantes et plus variables en fonction de H . En particulier, pour H faible, il y a des corrélations à plus de deux échelles. Cependant, on constate sur la figure 11 (page 28) que les conclusions sont les mêmes : la corrélation ajoutée par passage au module est favorable dans le cas d'une régression sur deux échelles ([3; 4]), indifférente dans le cas d'une régression sur trois échelles ([3; 5]) et défavorable dans le cas d'une régression sur quatre échelles ([3; 6]).

2.3.4 Conclusion : quel jeu de coefficients utiliser ?

Les tableaux de l'annexe B indiquent que les leaders donne systématiquement un $\sigma_{\hat{h}}$ plus faible que la DWT. On préférera donc toujours les leaders. De plus, l'utilisation des ondelettes complexes permet même avec les leaders de réduire cet écart type. Le choix de l'ondelette dépend de la contrainte d'estimation recherchée.

- Si on cherche à faire une estimation la plus locale possible, on fera une régression sur [3; 4]. L'introduction de corrélation par passage au module est alors favorable. On choisira donc l'ondelette qui maximise cette corrélation : l'ondelette de Daubechies.
- Si l'on souhaite faire la régression sur [3; 6] pour réduire la variabilité, la corrélation devient défavorable. On choisira alors l'ondelette Qshift (et le Dual Tree associé) qui minimise cette corrélation.

2.4 Étude de régularité Hölderienne variable

Le mouvement brownien multifractionnaire (mBM) est un modèle de signal localement autosimilaire où l'exposant $H(t)$ peut varier au cours du temps. La régularité Hölderienne $h(t)$ d'un tel processus suit les variations de $H(t)$. On pourra trouver la définition précise du mBM dans le chapitre de S. Cohen dans [7], ou dans [13]. La synthèse de ces processus est détaillé dans l'annexe C de [4]. Nous avons essayé deux types de variations de $H(t)$: variation douce et variation brusque.

2.4.1 Cas d'une variation douce de la régularité locale

La figure 12 (page 29) montre l'estimation de la régularité locale d'un mBM avec les différents jeux de coefficients en ondelettes. Les variations de $H(t)$ imposées sont obtenus à l'aide d'une fonction arctangente dont les paramètres sont ajustés pour avoir une variation douce de H de 0.1 à 0.9 sur la totalité du signal. Le signal est de taille 2^{15} , et les effets de bords sont gérés en n'estimant la régularité qu'au centre du signal comme pour le fBM. Les courbes sont obtenues en moyennant $\hat{h}(t)$ pour chaque temps sur 100 réalisations de mBM. Les régressions linéaires ont été effectuées sur [3; 6]. On constate que les courbes de $\hat{h}(t)$ suivent bien les variations de $H(t)$, y compris avec l'utilisation des ondelettes complexes. L'effet de réduction de l'écart-type de l'estimateur de régularité par passage au module semble conservé.

2.4.2 Cas d'un saut brutal de régularité locale

La figure 13 (page 30) montre l'estimation de la régularité locale d'un mBM avec les différents jeux de coefficients en ondelettes. Les variations de $H(t)$ imposées sont obtenus à l'aide d'une fonction arctangente dont les paramètres sont ajustés pour avoir un saut de H de 0.1 à 0.9 sur environ 100 points au centre du signal. De même qu'auparavant, le signal est de taille 2^{15} , et les effets de bords sont gérés en n'estimant la régularité qu'au centre du signal. Les courbes sont obtenues en moyennant $\hat{h}(t)$ pour chaque temps sur 100 réalisations de mBM. Les régressions linéaires ont été effectuées sur [3; 6]. On voit qu'au saut de H , il y a une surtension, pour toutes les méthodes. Cette surtension lors du passage rapide d'une valeur de H faible à une valeur de H élevée a déjà été observée dans [12] avec la transformation en ondelette continue. On constate qu'elle est inchangée avec les ondelettes complexes. L'effet de réduction de l'écart-type de l'estimateur de régularité par passage au module semble conservé.

Conclusion

Dans ce travail, on a tout d'abord étudié la transformation en ondelette complexe. Nous avons vu deux méthodes possibles pour utiliser des ondelettes analytiques : la transformation de Hilbert directe du signal, et la méthode Dual Tree. Nous avons vu que l'une et l'autre de ces deux méthodes permettent de limiter le problème de non invariance par translation temporelle lié à l'utilisation de la DWT. Bien que la méthode Dual Tree possède quelques avantages théoriques (elle permet surtout la reconstruction parfaite), nous n'avons pas mis en évidence de différence entre les deux modules pour l'analyse de singularités.

Ensuite, nous avons étudié les estimateurs de lois d'échelles basés sur les transformations en ondelettes réelles et complexes. Nous avons d'abord étudié ces estimateurs sur le fBM, pour lequel la régularité ponctuelle est constante. Nous avons alors observé que le comportement des coefficients en ondelette complexes est similaire au comportement des coefficients en ondelette réels (pour lesquels des résultats théoriques assurent le bien fondé de la méthode d'estimation de régularité). Cela valide la possibilité d'estimation de régularité ponctuelle à partir des coefficients en ondelette complexes. L'étude de la distribution des estimées de la régularité ponctuelle a montré que l'utilisation des modules accentue les corrélations entre les coefficients à deux échelles successives. Cette corrélation est favorable dans le cas d'une régression sur deux échelles ([3; 4]), indifférente dans le cas d'une régression sur trois échelles ([3; 5]) et défavorable dans le cas d'une régression sur quatre échelles ([3; 6]). On cherchera donc suivant les cas une ondelette qui maximise ou qui minimise cette corrélation introduite. La comparaison avec l'estimation basée sur les leaders a montré que l'on peut effectivement utiliser les leaders complexes pour l'estimation de régularité locale. Les mêmes conclusions s'appliquent alors quant à l'effet de la corrélation introduite en passant au module. Les écarts types sont toujours bien plus faibles avec les leaders qu'avec la DWT.

Enfin, nous avons très brièvement ébauché l'estimation de régularité locale variable sur le mBM. Nous avons alors observé qu'il est effectivement possible de suivre une régularité ponctuelle lentement variable grâce à un jeu de coefficients en ondelette combinant les leaders et l'utilisation des ondelettes complexes. Nous avons observé un phénomène de surtension lors du suivi d'une variation brusque de régularité locale.

Bien que l'utilisation des ondelettes complexes permette de limiter la variabilité de l'estimation de régularité locale, celle-ci reste très variable. On peut donc envisager d'étudier des méthodes de lissage pour réduire cette variabilité. Il est possible d'envisager l'étude de méthodes de lissage sur une seule réalisation de mBM.

Un autre axe pour poursuivre ce travail serait de regarder si l'utilisation des ondelettes complexes, éventuellement combinée aux leaders peut permettre d'améliorer l'estimation des spectres de singularités pour les signaux multifractals. En effet, un des problèmes de l'analyse multifractale est la présence de coefficients proches de zéro qui rend très instable l'estimation des moments d'ordre négatif. On peut envisager que l'utilisation d'une enveloppe complexe à la place d'une ondelette qui oscille autour de zéro permette de résoudre ce problème.

A Régression linéaire

Dans cette annexe, on explique la méthode utilisée pour faire les régressions linéaires qui permettent de trouver l'exposant d'autosimilarité (\widehat{H}) ou l'exposant de Hölder local ($\widehat{h}(t)$).

A.1 Estimation de la pente : formule des moindres carrés

Dans les deux cas, on dispose d'un vecteur

$$Y = \begin{pmatrix} Y_{j_{min}} \\ \vdots \\ Y_{j_{max}} \end{pmatrix}.$$

représentant les ordonnées, c'est à dire $Y_j = \log_2(\frac{1}{n_j} \sum_k |d(j, k)|^2)$, pour l'estimation par la marginale, et $Y_j = \log_2(d(j, k_0))$ pour l'estimation de l'exposant de Hölder local en k_0 . On note $X_j = \log_2(2^j) = j$ les abscisses correspondantes qui sont les échelles des coefficients en ondelette. Les abscisses sont donc déterministes alors que les ordonnées sont des variables aléatoires correspondant au \log_2 du coefficient en ondelette en k_0 pour l'estimation locale, ou de la moyenne des carrés des coefficients pour tous les k pour l'estimation globale (marginale). Le domaine de regression linéaire est $[j_{min}; j_{max}]$. On cherche à estimer les paramètres de la loi linéaire $Y_j = \widehat{\sigma} + \widehat{h} \cdot X_j$. L'estimation par les moindres carrés donne :

$$\widehat{\theta} = (A^T \cdot V^{-1} \cdot A)^{-1} \cdot A^T \cdot V^{-1} \cdot Y \quad (10)$$

avec

$$\widehat{\theta} = \begin{pmatrix} \widehat{\sigma} \\ \widehat{h} \end{pmatrix} \text{ et } A = \begin{pmatrix} 1 & X_{j_{min}} \\ \vdots & \vdots \\ 1 & X_{j_{max}} \end{pmatrix}$$

La matrice V qui intervient dans l'equation 10 est une matrice de covariance que l'on peut imposer à priori, si l'on a des informations sur la corrélations entre les Y_j , où si l'on veut introduire une pondérations entre ces coefficients. En particulier, on prendra :

$$- V = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \text{ pour n'imposer aucune corrélation et aucune pondération. Ce cas sera utilisé pour l'esti-}$$

mation locale.

$$- V = \begin{pmatrix} (\sigma_{j_{min}}^{\text{à p.}})^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (\sigma_{j_{max}}^{\text{à p.}})^2 \end{pmatrix} \text{ pour imposer la variance } (\sigma_j^{\text{à p.}})^2 \text{ à priori au coefficient } Y_j. \text{ Ce cas sera}$$

utilisé pour l'estimation par la marginale avec $(\sigma_j^{\text{à p.}})^2 = \frac{1}{n_j}$.

Nous nous plaçons maintenant dans le cas où la corrélation entre les coefficients aux différentes échelles n'est pas prise en compte à priori, et donc la matrice V introduite au paragraphe précédent est diagonale. Pour conserver la généralité, on se place dans le cas où les coefficients sont pondérés par $(\sigma_j^{\text{à p.}})^2$, sachant que le cas sans pondération peut facilement être obtenu en prenant $\sigma_j^{\text{à p.}} = 1$ pour tous les j .

Dans ce cas particulier, l'equation 10 peut s'écrire plus simplement pour \widehat{h} :

$$\widehat{h} = \sum_{j=j_{min}}^{j_{max}} w_j \cdot Y_j, \quad (11)$$

où

$$w_j = \frac{S_0 X_i - S_X}{S_0 S_{X^2} - S_X^2}$$

$$S_0 = \sum_{j=j_{min}}^{j_{max}} \frac{1}{(\sigma_j^{\text{à p.}})^2}$$

$$S_X = \sum_{j=j_{min}}^{j_{max}} \frac{X_j}{(\sigma_j^{\text{à p.}})^2}$$

$$S_{X^2} = \sum_{j=j_{min}}^{j_{max}} \frac{X_j^2}{(\sigma_j^{\text{à p.}})^2}$$

A.2 Variance sur l'estimation de la pente dans le cas d'une estimation ne tenant pas compte de la corrélation à priori

Nous sommes toujours dans la situation où la matrice V a été prise diagonale pour l'estimation de \hat{h} , et donc l'équation 11 est valable. Cependant, il est possible que les coefficients soient corrélés. On note donc COV la matrice de covariance effective de Y . Elle peut être non diagonale dans le cas où les coefficients aux différentes échelles comprise entre j_{min} et j_{max} :

$$COV = \begin{pmatrix} \sigma_{j_{min}}^2 & \rho_{j_{min} (j_{min}+1)} \sigma_{j_{min}} \sigma_{j_{min}+1} & \cdots \\ \rho_{j_{min} (j_{min}+1)} \sigma_{j_{min}} \sigma_{j_{min}+1} & \sigma_{j_{min}+1}^2 & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix},$$

où l'on a donc

$$\sigma_j^2 = \mathbb{E}_{Y_j} \{(Y_j - \mathbb{E}_{Y_j} \{Y_j\})^2\} \text{ et } \rho_{j_1 j_2} = \frac{\mathbb{E}_{Y_{j_1}, Y_{j_2}} \{Y_{j_1} \cdot Y_{j_2}\} - \mathbb{E}_{Y_{j_1}} \{Y_{j_1}\} \mathbb{E}_{Y_{j_2}} \{Y_{j_2}\}}{\sigma_{j_1} \sigma_{j_2}}$$

Par exemple pour une régression de 3 à 6,

$$COV = \begin{pmatrix} \sigma_3^2 & \rho_{34} \sigma_3 \sigma_4 & \rho_{35} \sigma_3 \sigma_5 & \rho_{36} \sigma_3 \sigma_6 \\ \rho_{34} \sigma_3 \sigma_4 & \sigma_4^2 & \rho_{45} \sigma_4 \sigma_5 & \rho_{46} \sigma_4 \sigma_6 \\ \rho_{35} \sigma_3 \sigma_5 & \rho_{45} \sigma_4 \sigma_5 & \sigma_5^2 & \rho_{56} \sigma_5 \sigma_6 \\ \rho_{36} \sigma_3 \sigma_6 & \rho_{46} \sigma_4 \sigma_6 & \rho_{56} \sigma_5 \sigma_6 & \sigma_6^2 \end{pmatrix}.$$

En utilisant l'équation 11, on peut calculer la variance de l'estimateur \hat{h} , et on trouve :

$$\begin{aligned} \text{var}(\hat{h}) &= w^T \cdot COV \cdot w \\ &= \sigma_3^2 \cdot w_3^2 + 2 \cdot \rho_{34} \sigma_3 \sigma_4 \cdot w_3 w_4 + 2 \cdot \rho_{35} \sigma_3 \sigma_5 \cdot w_3 w_5 + \cdots \end{aligned} \tag{12}$$

B Tableaux des valeurs des écarts types de \hat{h}

TAB. 2: Table des écarts types de la distribution de $\hat{h}(t)$. Les valeurs sont obtenues en moyennant sur 100 réalisations. La regression linéaire a été effectuée sur [3; 4]. Les valeurs sont présentées sous la forme *valeur réelle (fraction par rapport à la partie réelle)*.

H		Qshift		Daubechies		
		DWT	leaders	DWT	leaders	CDWT
0.2	partie réelle	2.260 (1.00)	0.381 (1.00)	2.256 (1.00)	0.378 (1.00)	2.247 (1.00)
	module Hilbert	1.237 (0.55)	0.303 (0.79)	1.169 (0.52)	0.294 (0.78)	1.200 (0.53)
	module DT	1.238 (0.55)	0.303 (0.80)	-		
0.7	partie réelle	2.267 (1.00)	0.606 (1.00)	2.214 (1.00)	0.585 (1.00)	2.188 (1.00)
	module Hilbert	1.225 (0.54)	0.495 (0.82)	1.096 (0.49)	0.456 (0.78)	1.136 (0.52)
	module DT	1.223 (0.54)	0.497 (0.82)	-		

TAB. 3: Table des écarts types de la distribution de $\hat{h}(t)$. Les valeurs sont obtenues en moyennant sur 100 réalisations. La regression linéaire a été effectuée sur [3; 5]. Les valeurs sont présentées sous la forme *valeur réelle (fraction par rapport à la partie réelle)*.

H		Qshift		Daubechies		
		DWT	leaders	DWT	leaders	CDWT
0.2	partie réelle	1.133 (1.00)	0.220 (1.00)	1.131 (1.00)	0.219 (1.00)	1.135 (1.00)
	module Hilbert	0.656 (0.58)	0.182 (0.83)	0.647 (0.57)	0.177 (0.81)	0.649 (0.57)
	module DT	0.657 (0.58)	0.182 (0.83)	-		
0.7	partie réelle	1.135 (1.00)	0.338 (1.00)	1.123 (1.00)	0.333 (1.00)	1.133 (1.00)
	module Hilbert	0.653 (0.58)	0.291 (0.86)	0.634 (0.56)	0.282 (0.85)	0.643 (0.57)
	module DT	0.654 (0.58)	0.294 (0.87)	-		

TAB. 4: Table des écarts types de la distribution de $\hat{h}(t)$. Les valeurs sont obtenues en moyennant sur 100 réalisations. La regression linéaire a été effectuée sur [3; 6]. Les valeurs sont présentées sous la forme *valeur réelle (fraction par rapport à la partie réelle)*.

H		Qshift		Daubechies		
		DWT	leaders	DWT	leaders	CDWT
0.2	partie réelle	0.715 (1.00)	0.152 (1.00)	0.717 (1.00)	0.151 (1.00)	0.720 (1.00)
	module Hilbert	0.425 (0.59)	0.128 (0.85)	0.432 (0.60)	0.125 (0.83)	0.428 (0.59)
	module DT	0.426 (0.60)	0.129 (0.85)	-		
0.7	partie réelle	0.716 (1.00)	0.225 (1.00)	0.721 (1.00)	0.228 (1.00)	0.729 (1.00)
	module Hilbert	0.426 (0.59)	0.200 (0.89)	0.436 (0.60)	0.202 (0.89)	0.435 (0.60)
	module DT	0.427 (0.60)	0.201 (0.89)	-		

C Logarithme de variables aléatoires gaussiennes

On considère deux variables aléatoires X_1 et X_2 conjointement gaussiennes de variances σ_1 et σ_2 et de moyennes nulles ; et on note ρ leur coefficient de corrélation, $p_1(x_1)$ et $p_2(x_2)$ les marginales et $p_{1,2}(x_1, x_2)$ la loi conjointe.

On a donc

$$p_i(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{x_i^2}{2\sigma_i^2}\right), \text{ et } \int_{-\infty}^{\infty} p_i(x_i) dx_i = 1, \quad i = 1, 2.$$

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad C = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$p_{1,2}(x_1, x_2) = \frac{1}{2\pi\sqrt{\text{Det}(C)}} \exp\left(-\frac{1}{2}X^T C^{-1}X\right), \text{ et } \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{1,2}(x_1, x_2) dx_1 dx_2 = 1.$$

On a bien

$$\mathbb{E}_{X_i}\{X_i\} = 0, \quad i = 1, 2$$

et

$$\rho = \frac{\mathbb{E}_{X_1, X_2}\{X_1 \cdot X_2\} - \mathbb{E}_{X_1}\{X_1\}\mathbb{E}_{X_2}\{X_2\}}{\sigma_1\sigma_2}$$

C.1 Moyennes et Variances

On cherche en fonction de σ_1, σ_2 les moyennes et variances des variables aléatoires $|X_i|$ et $\log_2|X_i|$. On obtient :

Moyennes :

$|X_i|$

$$\mu_i = \mathbb{E}_{X_i}\{|X_i|\} = \sqrt{\frac{2}{\pi}}\sigma_i, \quad i = 1, 2.$$

$\log_2|X_i|$

$$\lambda_i = \mathbb{E}_{X_i}\{\log_2|X_i|\} = -\frac{1}{2} + \log_2\sigma_i - \frac{\gamma}{2\ln 2}, \quad i = 1, 2 ;$$

où γ est la constante d'Euler.

Variance

$|X_i|$

$$\sigma_{||,i}^2 = \mathbb{E}_{X_i}\{(|X_i| - \mu_i)^2\} = \frac{\pi - 2}{\pi}\sigma_i^2, \quad i = 1, 2.$$

$\log_2|X_i|$

$$\sigma_{\log_2||,i}^2 = \mathbb{E}_{X_i}\{(\log_2|X_i| - \lambda_i)^2\} = \frac{\pi^2}{8(\ln 2)^2}, \quad i = 1, 2.$$

C.2 Corrélations

On cherche maintenant la corrélation entre les variables aléatoires $|X_1|$ et $|X_2|$. On peut faire le calcul analytiquement, et on obtient :

$$\begin{aligned} \rho_{||} &= \frac{\mathbb{E}_{X_1, X_2}\{|X_1| \cdot |X_2|\} - \mu_1\mu_2}{\sigma_{||,1}\sigma_{||,2}} \\ &= \frac{2}{\pi - 2} \left[\frac{(1 - \rho^2)^2 + \rho^2 - \rho^4}{\sqrt{1 - \rho^2}} + \rho \arctan\left(\frac{\rho}{\sqrt{1 - \rho^2}}\right) - 1 \right]. \end{aligned}$$

À l'ordre le plus bas en ρ , cela donne :

$$\rho_{||} \approx \left(\frac{1}{\pi - 2}\right)\rho^2.$$

La figure 14 en haut à gauche montre la corrélation des valeur absolue $\rho_{||}$ par rapport à la corrélation initiale ρ .

Pour la corrélation entre les variables aléatoires $\log_2|X_1|$ et $\log_2|X_2|$, un calcul analytique n'est pas possible. Mais on peut faire un calcul à l'ordre le plus bas en ρ dans la limite des faibles ρ . On obtient :

$$\rho_{\log_2||} \approx \left(\frac{4}{\pi^2}\right)\rho^2.$$

La figure 14 en haut à droite montre la corrélation des \log_2 des valeur absolue $\rho_{\log_2||}$ par rapport à la corrélation initiale ρ .

Dans la limite des faibles ρ , on a donc :

$$\rho_{\log_2||} \approx \left(\frac{4(\pi - 2)}{\pi^2}\right)\rho_{||}$$

La figure 14 en bas montre la corrélation des \log_2 des valeur absolue $\rho_{\log_2||}$ par rapport à la corrélation des valeur absolue $\rho_{||}$.

Références

- [1] DAUBECHIES I. *Ten lectures on wavelets*. SIAM, Philadelphia, 1992.
- [2] SELESNICK I. The design of approximate Hilbert transform pairs of wavelet bases. *IEEE Transactions on signal processing*, 50(5), May 2002.
- [3] SELESNICK I.W., BARANIUK R.G., and KINGSBURY N.G. The dual-tree complex wavelet transform ; a coherent framework for multiscale signal and image processing. *IEEE Signal Processing Magazine*, pages 123–151, November 2005.
- [4] COEURJOLLY J.F. *Inférence statistique pour les mouvements browniens fractionnaires et multifractionnaires*. PhD thesis, Université Joseph Fourier, 2000.
- [5] KINGSBURY N. Complex wavelets for shift invariant analysis and filtering of signals. *Journal of Applied and Computational Harmonic Analysis*, 10(3) :234–253, May 2001.
- [6] ABRY P. *ONDELETTE ET TURBULENCES. Multirésolutions, algorithmes de décomposition, invariance d'échelle et signaux de pression*. Diderot, Paris, 1997.
- [7] ABRY P., GONÇALVÈS P., and LÉVY VEHEL J. *Lois d'échelle, fractales et ondelettes*, volume 1. Hermès Sciences, Paris, 2002.
- [8] ABRY P., GONÇALVÈS P., and LÉVY VEHEL J. *Lois d'échelle, fractales et ondelettes*, volume 2. Hermès Sciences, Paris, 2002.
- [9] ABRY P., GONÇALVÈS P., and FLANDRIN P. Wavelets, spectrum estimation, and 1/f processes. In ANTONIADIS A. and OPPENHEIM G., editors, *Wavelets and statistics*. Springer-Verlag, 1995.
- [10] EMBRECHTS P. and MAEJIMA M. *Selfsimilar processes*. Princeton University Press, Princeton and Oxford, 2002.
- [11] FLANDRIN P. *Temps-fréquence*. Hermès, Paris, 2ème édition, 1998.
- [12] GONÇALVÈS P. *Représentations temps-fréquence et temps-échelle bilinéaires : synthèse et contributions*. PhD thesis, Institut national polytechnique de Grenoble, 1993.
- [13] COHEN S. From self-similar to local self-similarity : the estimation problem. In DEKKING M., LÉVY-VÉHEL J., LUTTON E., and TRICOT C., editors, *Fractals : theory and applications in engineering*. Springer, 1999.
- [14] JAFFARD S. Wavelet techniques in multifractal analysis. In LAPIDUS M. and VAN FRANKENHUYSEN M., editors, *Fractal geometry and application : *A jubilee of Benoit Mandelbrot*. Proc. of Symp. in pure mathematics, 2004.
- [15] MALLAT S. *A Wavelet tour of signal processing*. Academic Press, San Diego, 2ème édition, 1999.
- [16] APPEL W. *Mathématiques pour la physique et pour les physiciens*. H&K, Paris, 2002.

FIG. 11: Pourcentage de la réduction de $\sigma_{\hat{h}}$ par passage au module expliqué par la réduction de l'écart type des ordonnées ($\log_2|d(j, k)|$), pour les trois domaines de régression linéaire. Partant de $\sigma_{\hat{h}}$ pour la partie réelle ramené à 1, la flèche de gauche amène à $\sigma_{\hat{h}}$ attendu avec le module s'il ne modifiait pas la corrélation des ordonnées. Puis la flèche de droite amène à $\sigma_{\hat{h}}$ réellement observé (en tenant compte des corrélations).

Etude de l'écart type de $\hat{h}(t)$

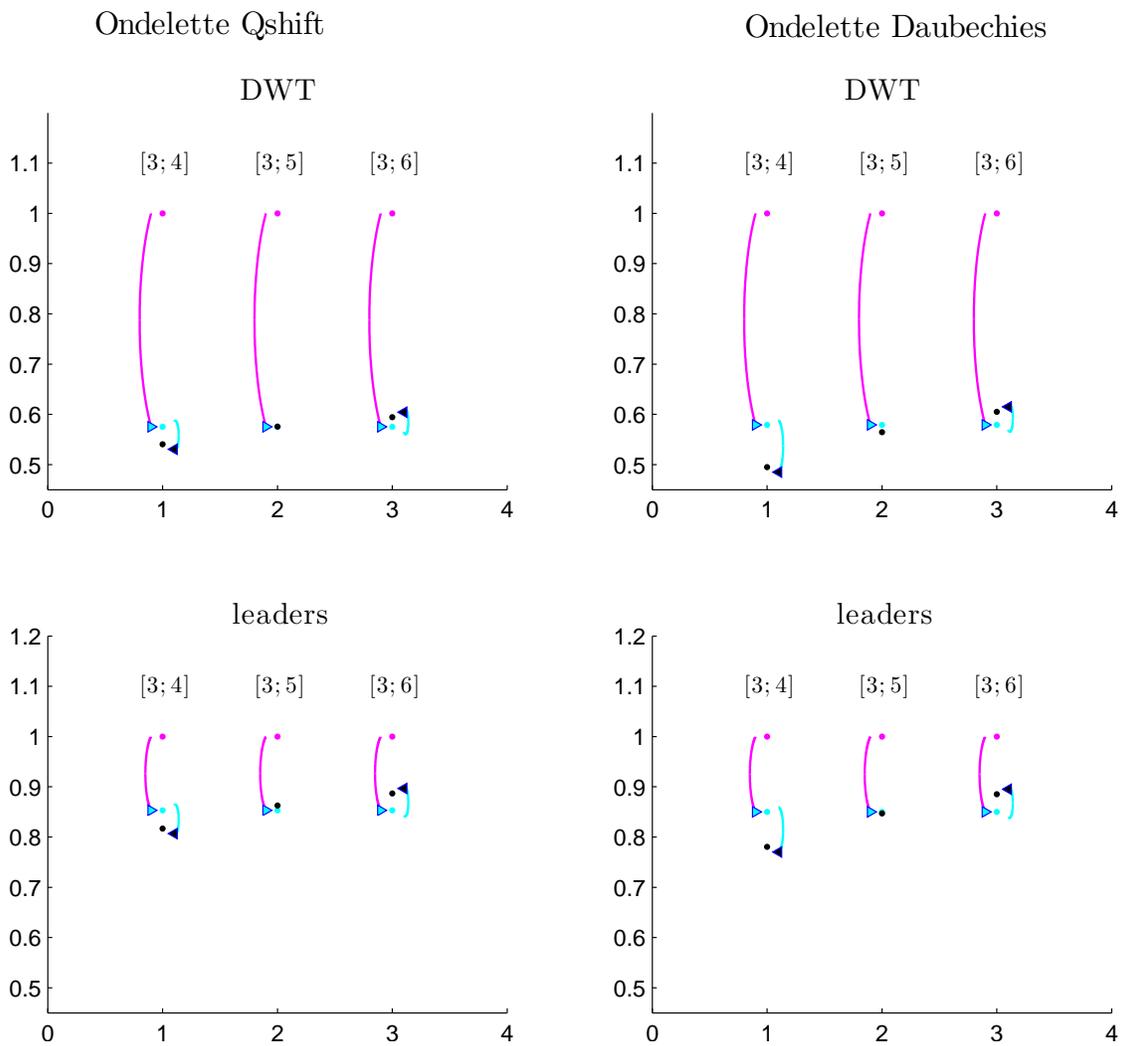


FIG. 12: Estimation de la régularité locale d'un mBM. Les variations de $H(t)$ imposées sont obtenus à l'aide d'une fonction arctangente.

Estimation de la régularité locale pour un mBM.

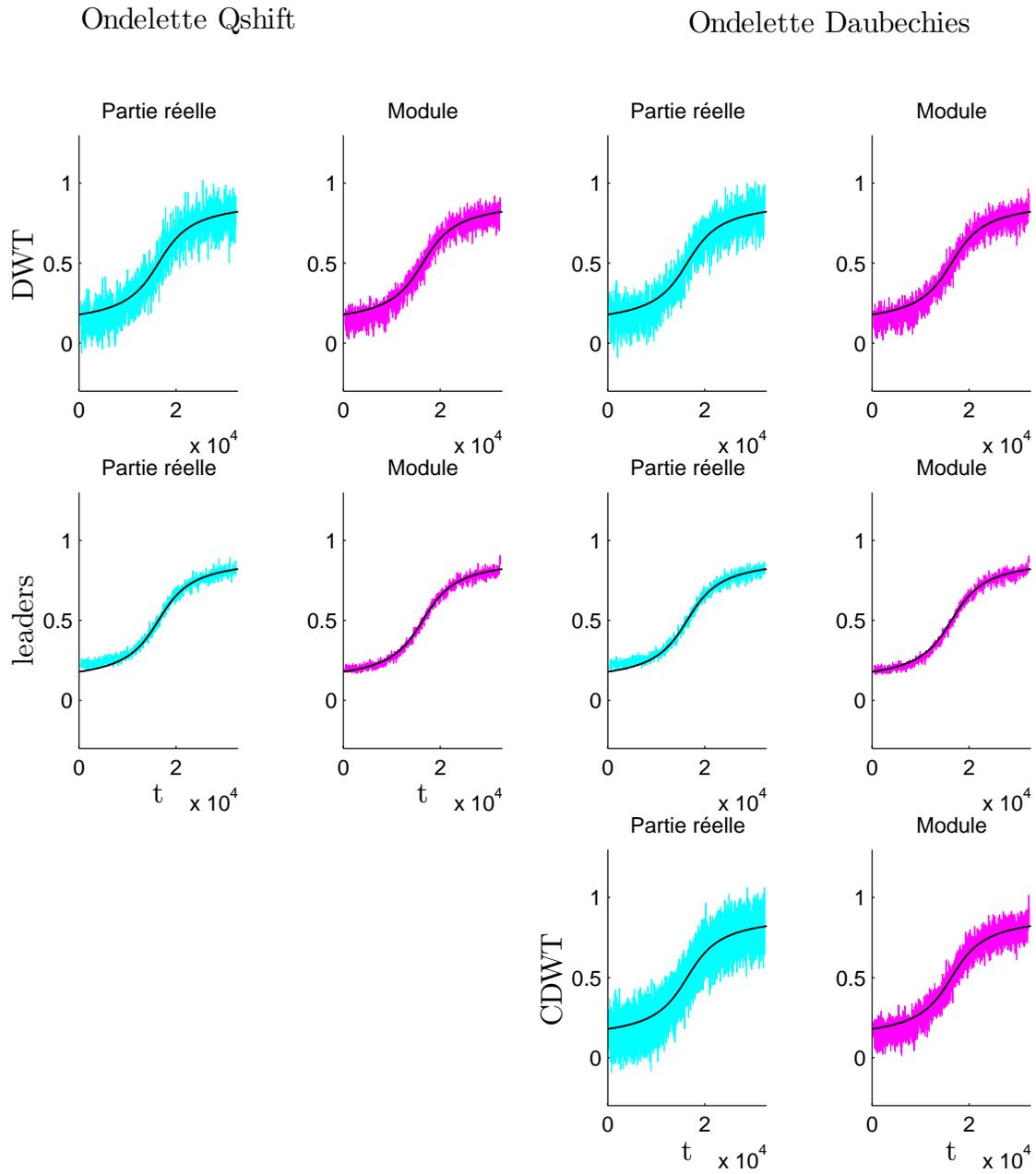


FIG. 13: Estimation de la régularité locale d'un mBM. Les variations de $H(t)$ imposées sont obtenus à l'aide d'une fonction arctangente.

Estimation de la régularité locale pour un mBM.

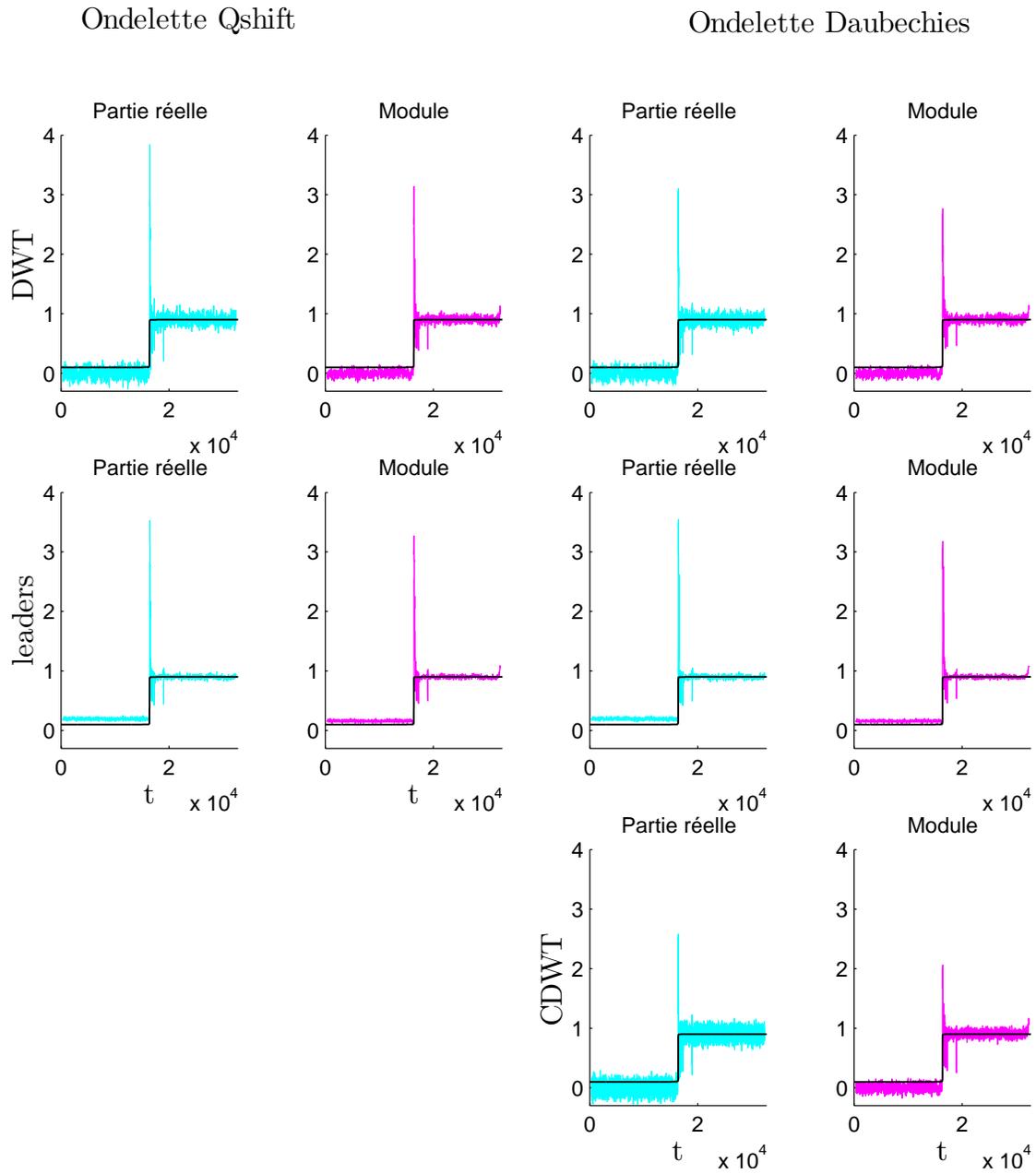


FIG. 14: Corrélations de la valeur absolue et du log de la valeur absolue de deux variables aléatoires conjointement gaussiennes corrélées. En haut à gauche : la corrélation de la valeur absolue en fonction de la corrélation initiale. En haut à droite : la corrélation du log de la valeur absolue en fonction de la corrélation initiale. En bas : la corrélation du log de la valeur absolue en fonction de la corrélation de la valeur absolue.

