

Résumé

La transformée en ondelettes a été considérée comme un outil prometteur pour l'élimination du bruit en traitement du signal. Des méthodes non linéaires ont été proposées par Donoho et Johnstone qui consiste à estimer une fonction à partir des données bruitées par un seuillage des coefficients d'ondelettes empiriques, avec comme principal objectif la minimisation de l'erreur en moyenne quadratique. Contrairement aux méthodes linéaires, ces méthodes ont de bonnes propriétés minimax sur certaines espaces fonctionnels. Cette étude porte sur un signal synthétique perturbé par un processus stationnaire.

Mots clés : Analyse en ondelettes, risque minimax, risque de bayes, bruit stationnaire, estimation minimax.

1-Introduction

Les données expérimentales X sont souvent entachées d'un bruit ε assimilable à un aléa additif. Notre modèle s'écrit sous la forme suivante:

$$X(t_i) = Y(t_i) + \varepsilon(t_i)$$

Où $X = (X(t_i))_i$ représente les données observées
 $Y = (Y(t_i))_i$ représente les données sans bruit

Notre but est d'estimer la fonction $y(t)$ en se basant sur les méthodes non linéaires proposées par Donoho et Johnstone. Plus précisément nous ferons appel à la méthode du seuillage par ondelettes.

L'estimateur obtenu est optimal dans le sens suivant:

- l'estimateur $\tilde{y}(t)$ obtenu est aussi régulier que $y(t)$
- l'estimateur $\tilde{y}(t)$ est presque minimax pour de nombreux espaces fonctionnels aussi large que l'espace de Besov à lequel probablement les données non bruitées appartiennent.

-2- Réduction du bruit par seuillage des coefficients d'ondelettes :

L'application de la transformée en ondelettes W aux données $X(t_i)$ $i=1, n$ maintient le modèle (1) inchangé :

$$w_i = \theta_i + z_i \quad \dots \quad (2)$$

$$i \in I = \{ (j,k) \mid 1 \leq j \leq m \text{ et } 1 \leq k \leq 2^{j-1} \}$$

L.DONOHO et M.JONSTONE ont établi le résultat suivant.

Soit le modèle (2) avec z_i $i \in I$ un bruit blanc gaussien

On suppose que θ_i $i \in I$ est une variable aléatoire de mesure a priori $\mathbb{J}_i \in M_{p,q}^\alpha$

$$M_{p,q}^\alpha = \{ \mathbb{J} / \|\tau\|_{b,p,q}^\alpha \leq c \}$$

Où

$$\tau = (\tau_i)_i, \quad \tau_i = (E_{\pi_i} |\theta_i|^{p \wedge q})^{1/p \wedge q}$$

L'estimateur $\theta^* = (\theta_i^*)_i$, $i \in I$ qui réalise le risque minimax bayésien B^* défini par :

$$B^*(n, \Theta_{p,q}^\alpha) = \inf_{\theta^*} \sup_{\theta} E_{\pi} E_{\theta} \|\theta^* - \theta\|^2$$

est de la forme

$$\theta_i^* = \delta^*(w_i)$$

Où

δ^* est une fonction non linéaire

De plus

L'utilisation de la fonction non linéaire « seuillage doux » :

$$\delta_\lambda(w) = \text{sig}w(|w| - \lambda)_+ \quad \text{avec un seuil } \lambda$$

Où bien de la fonction non linéaire « seuillage dur »

$$\delta_\mu(w) = w \mathbf{1}_{\{|w| \geq \mu\}} \quad \text{avec un seuil } \mu$$

fournit un estimateur bayésien presque minimax ce qui est exprimé par le théorème suivant

Théorème –1-

considérons le modèle (2)

Il existe deux constantes $N(p \wedge q)$ et $M(p \wedge q)$ telles que

$$B_{\lambda}^*(n, \Theta_{p,q}^\alpha) \leq N(p \wedge q) B^*(n, \Theta_{p,q}^\alpha)$$

$$B_{\mu}^*(n, \Theta_{p,q}^\alpha) \leq M(p \wedge q) B^*(n, \Theta_{p,q}^\alpha)$$

• A partir des évaluations numériques, il a été constaté que les constantes $N(p \wedge q)$ et $M(p \wedge q)$ sont inférieures à 2.22 et que la constante $N(p \wedge q)$ tend à être plus petite que la constante $M(p \wedge q)$

• $\theta^* = (\theta_i^*) \quad \theta_i^* = \delta_{\lambda}^*(w_i) \quad i \in I$

$$\hat{Y} = W^{-1}(\theta^*) \quad W^{-1} \text{ étant la T.O inverse}$$

Seuil universel :

Le seuil optimal λ^* est donnée par la valeur du seuil qui minimise le risque $r(\theta^*, \theta)$:

$$r(\theta^*, \theta) = E(\|\theta_{\lambda^*}^* - \theta\|^2)$$

Asymptotiquement,

$$\lambda^* \approx \sigma (2 \log n)^{1/2} \quad \sigma \text{ l'écart type du bruit}$$

ce seuil est appelé seuil universel

- DONOHO et JHONSTONE ont montré que l'utilisation du seuil universel fournit un estimateur θ^* dont le risque est à un facteur logarithmique près du risque benchmark (relatif à la sélection idéale des coefficients) ce qui est énoncé par le théorème suivant :

Théorème -2-

soit le modèle (2)

l'estimateur

$$\theta^* = \delta(w, \sigma (2\log n)^{1/2})$$

satisfait

$$r(\theta^*, \theta) \leq (2\log n + 1)(\sigma^2 + \sum_{i=1}^{i=n} (\theta_i^2 \wedge \sigma^2))$$

- le théorème suivant affirme l'estimateur θ^* obtenu atteint asymptotiquement le meilleur risque relativement au risque benchmark

Théorème -3-

Soit le modèle (2)

l'estimateur $\theta^* = \delta(w, \sigma (2\log n)^{1/2})$

satisfait

$$\inf_{\theta^*} \sup_{\theta} E(\|\theta^* - \theta\|^2) / (2\log n + 1)(\sigma^2 + \sum_{i=1}^{i=n} (\theta_i^2 \wedge \sigma^2)) \rightarrow 1 \quad \text{qd } n \rightarrow +\infty$$

- l'estimateur est d'un risque qui est à un facteur logarithmique près du risque minimax .en effet,

Théorème -4-

soit $B_{p,q}^\alpha(c)$ une boule de Besov

L'estimateur \hat{y}_n^u est presque minimax c.a.d.

$$\sup_{\theta} E(\|\hat{y}_n^u - y\|^2) \leq k \log n \inf_y \sup_y E(\|\hat{y} - y\|^2)$$

Avec $k \in \mathbb{R}$ et $y \in B_{p,q}^\alpha(c)$

Seuil sure

C.STEIN a proposé un seuil λ^s dit « seuil sure » basé sur la minimisation de l'estimateur sans biais de l'EQM :

$$\text{SURE}(\lambda, w) = 1/n (\|\theta_{\lambda}^* - w\|^2 + \sigma^2 (n - 2n_0))$$

$$n_0 = \text{card} \{ \theta_i / |\theta_i| < \lambda \}$$

l'estimateur

$$\theta^{*s} = \delta(w, \lambda^s)$$

est à une constante près du risque minimax : $\sup_{\theta} E \|\theta^{*s} - \theta\|^2 \leq c. \inf_{\theta^*} \sup_{\theta} E \|\theta^* - \theta\|^2$

Avec $c \in \mathbb{R}$ et $\theta \in \Theta_{p,q}^\alpha(c)$

3-Extension au cas d'un bruit corrélé :

W.SILVERMAN et M.JOHNSTONE ont étendu la procédure d'élimination du bruit au cas d'un bruit stationnaire corrélé (AR(1),MA(1),ARMA(1,1)...) en effectuant un seuillage dépendant d'échelle. L'estimateur obtenu est :

$$\theta^* = (\theta_{jk}^*)_{jk} \quad \theta_{jk}^* = \delta^*(w_{jk}, \lambda_j) \quad (j,k) \in I$$

λ_j est le seuil appliqué au $j^{\text{ème}}$ niveau

Théorème -5-

Soit le modèle (1) et soit Ψ une ondelette à r moments nuls et r fois continûment différentiable alors

$$\sup_y E(\|\hat{y}_n^s - y\|^2) \approx \inf_{\hat{y}} \sup_y E(\|\hat{y} - y\|^2) \quad , \quad y \in B_{p,q}^\alpha(c)$$

pour tout $p,q \in [1, +\infty[$, $c \in [0, +\infty[$ et $\alpha < r$

4-Etude sur un signal synthétique

Le signal synthétique

$$S(t) = \sum_{i=1}^{i=44} R_i(t-t_i)$$

est générée par la fonction

$$R(t-t_i) = a_i (t-t_i)_+^{b_i} \exp\{-c_i(t-t_i)\}$$

Où $a_i, b_i \in \mathbb{R}$ et $c_i \in \mathbb{R}_+^*$ $i=1, \dots, 44$

Notons que $R(t-t_i)$ est de régularité b_i en t_i

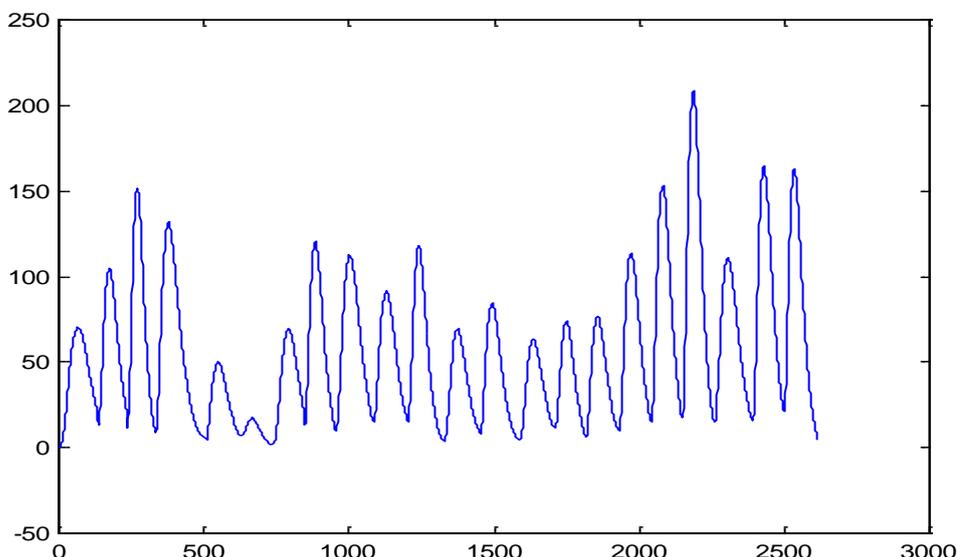


Figure 4.1 :signal synthétique

notre étude consiste à appliquer les méthodes de réduction de bruit au signal synthétique a lequel on additionne différents types de bruit (bb ,AR(1)) et par la suite .

4-1 Réduction du bruit par seuillage des coefficients d'ondelettes en utilisant les seuils sure et universel:

Les résultats du seuillage (global ou dependant d'echelles) sont illustres dans les figures suivantes:

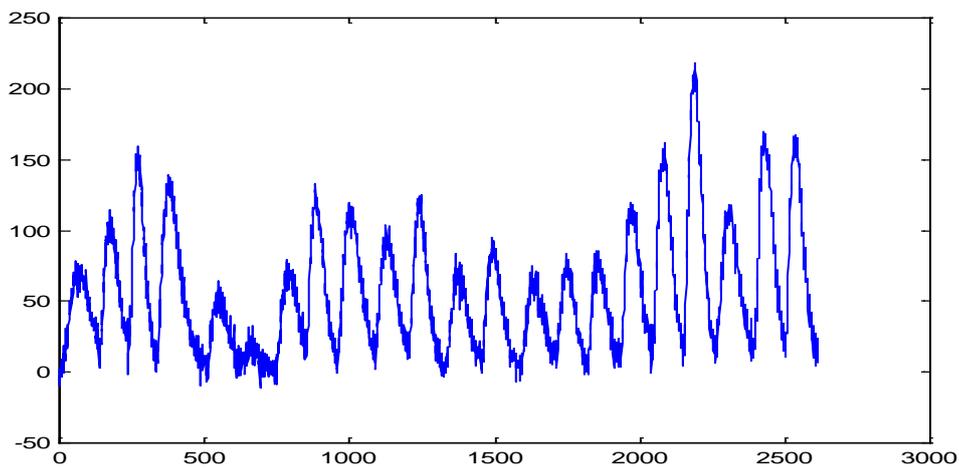


Figure4.2 : Synthétique +bruit blanc avec $\sigma=6$

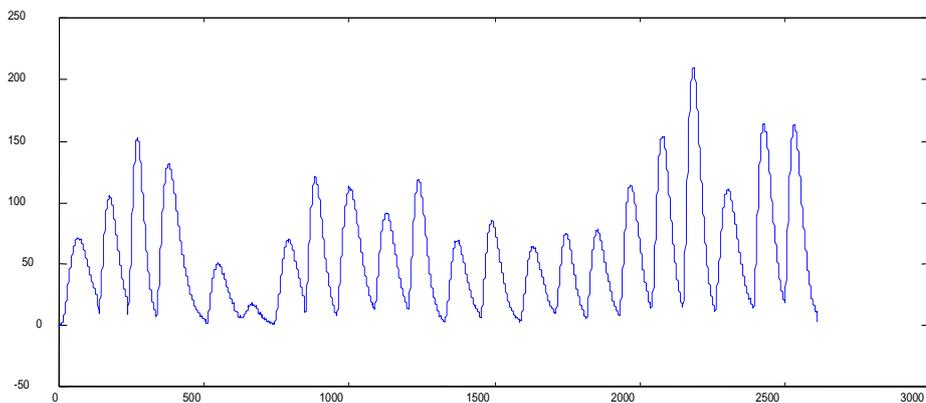


Figure4.3: Synthétique +bb débruité (sure, global)

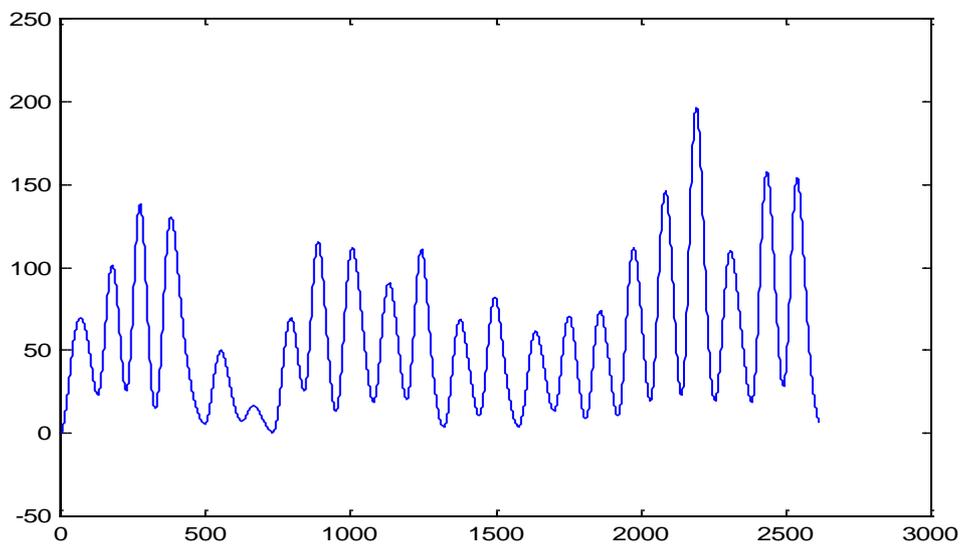


Figure4.4 : Synthétique +bb débruité (universel,global)

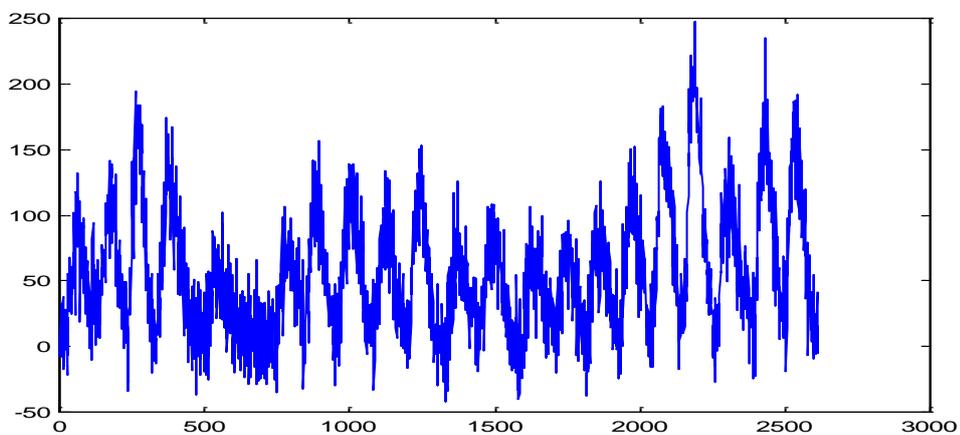


Figure4.5 :Synthétique+ar(1) avec $\phi=0.2$

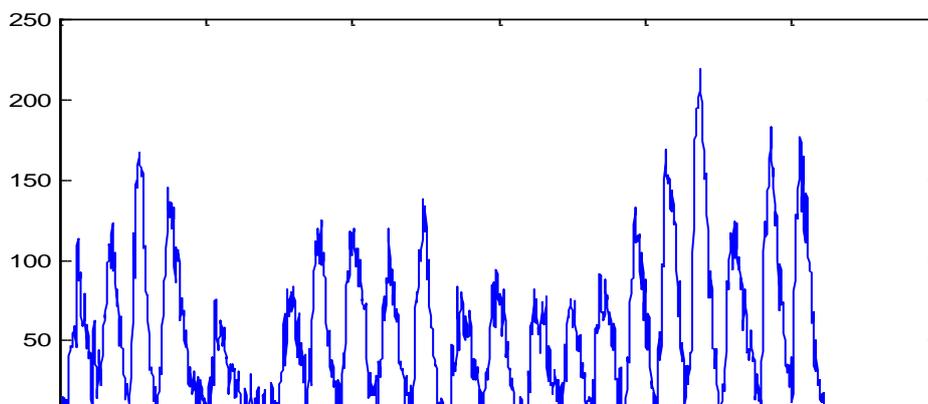


Figure4.6 :Synthétique+ar(1) ($\varphi=0.2$) après seuillage (sure global)

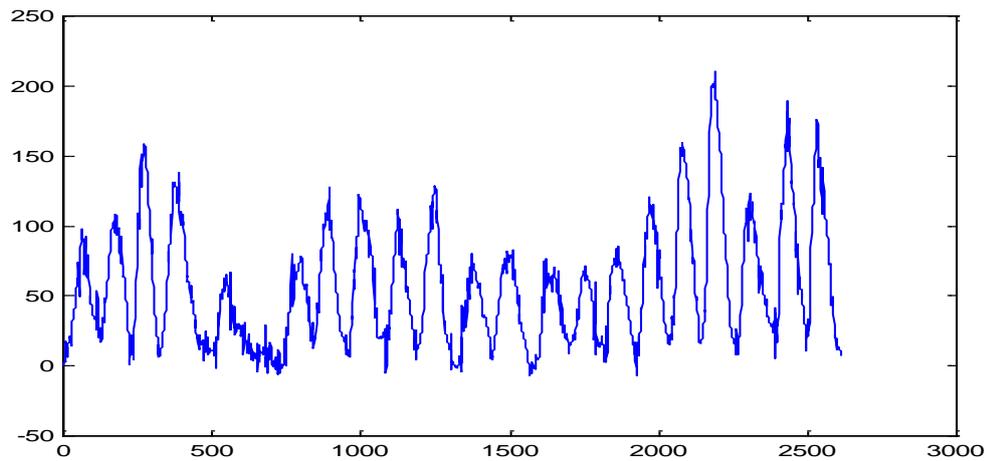


Figure4.7 :Synthétique+ar(1) ($\varphi=0.2$) après seuillage (universel global)

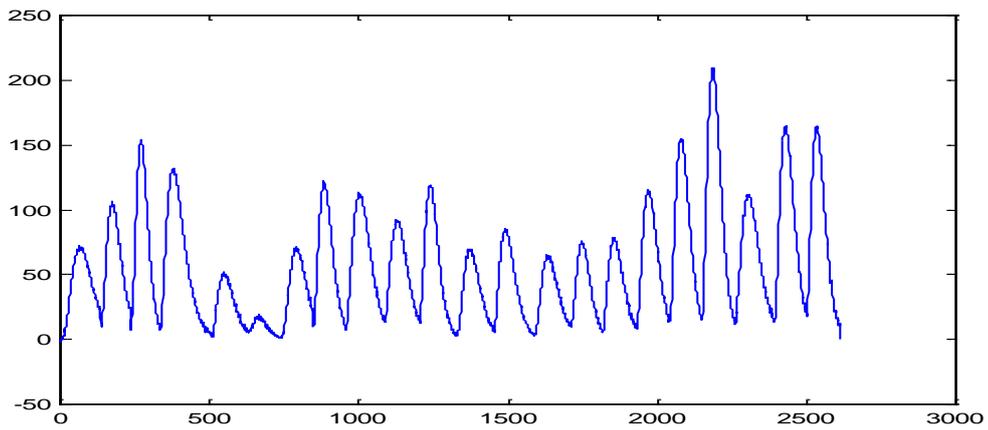


Figure4.8:Synthétique+ar(1) ($\varphi=0.2$) après seuillage (sure dépendant d'échelle)

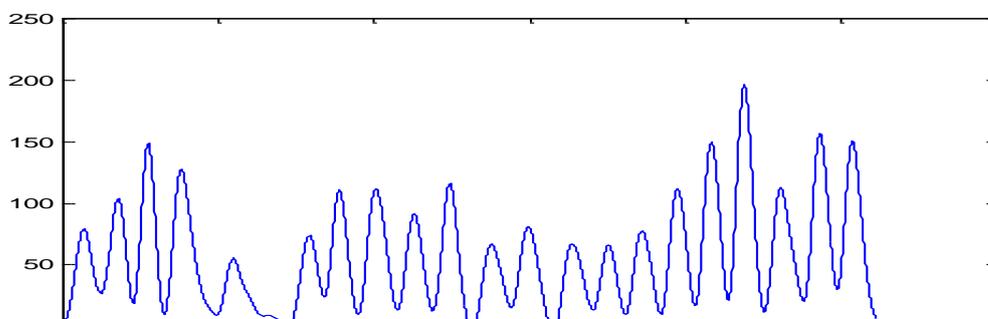


Figure4.9: Synthétique+ar(1) ($\varphi=0.2$) après seuillage (universel dépendant d'échelle)

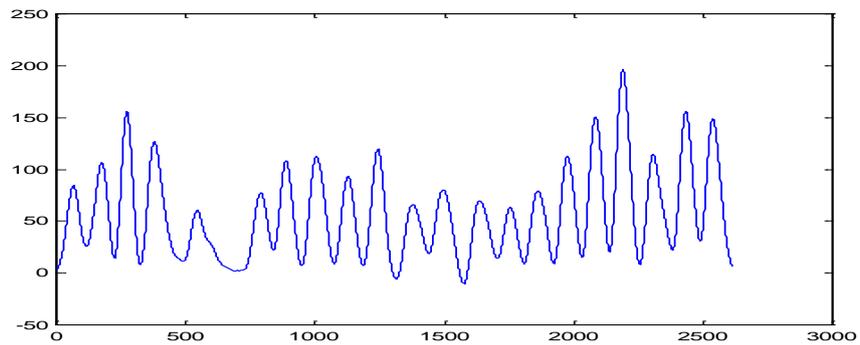


Figure4.10 :synthétique+ar(1) ($\varphi=0.5$) après seuillage (universel dépendant d'échelle)

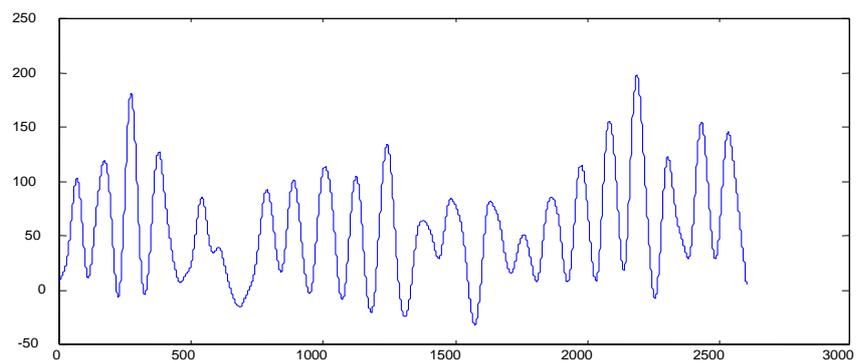


Figure4.11:synthétique+ar(1) ($\varphi=0.8$) après seuillage (universel dépendant d'échelle)

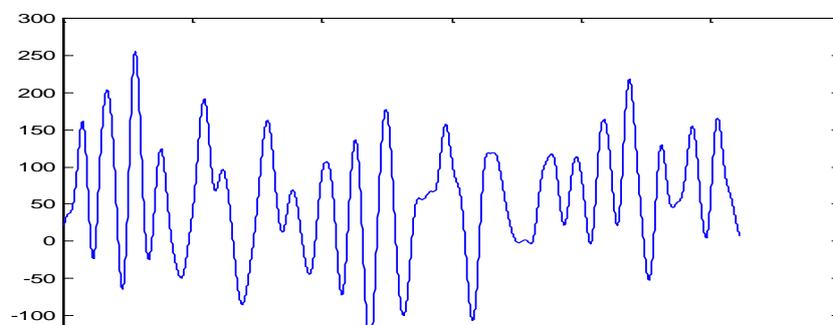


Figure4.12 :synthétique+ar(1) ($\phi=0.95$) après seuillage (universel dépendant d'échell

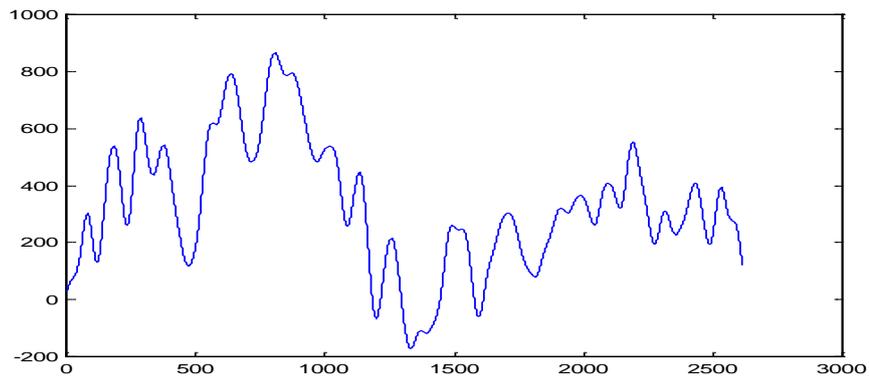


Figure4.13 :synthétique+ar(1) ($\phi=1$) après seuillage (universel dépendant d'échelle)

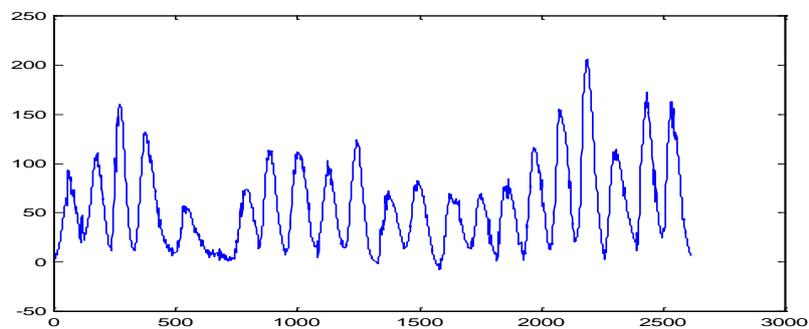


Figure4.14 :synthétique+ar(1) ($\phi=0.5$) après seuillage (sure dépendant d'échelles)

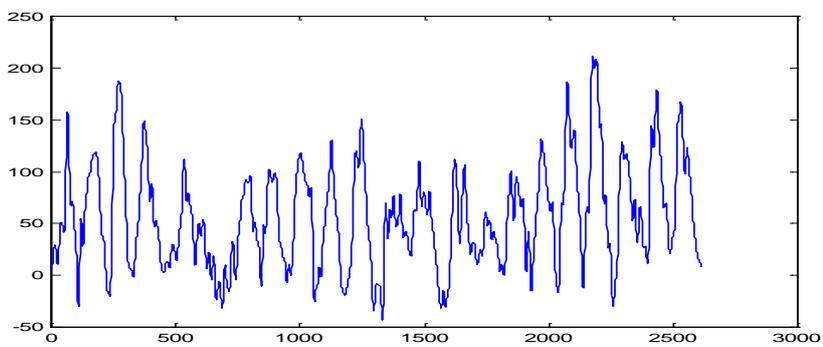


Figure4.15 :synthétique+ar(1) ($\phi=0.8$) après seuillage (sure dépendant d'échelles)

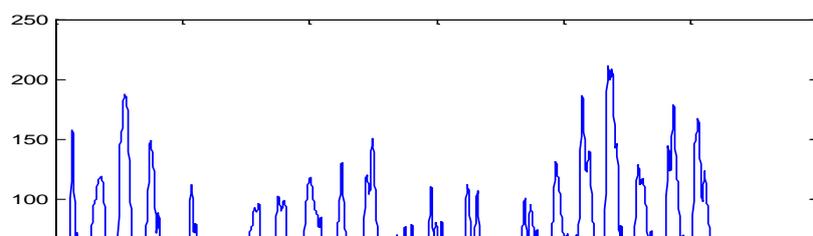


Figure4.16 :synthétique+ar(1) ($\phi=0.95$) après seuillage (sure dépendant d'échelle)

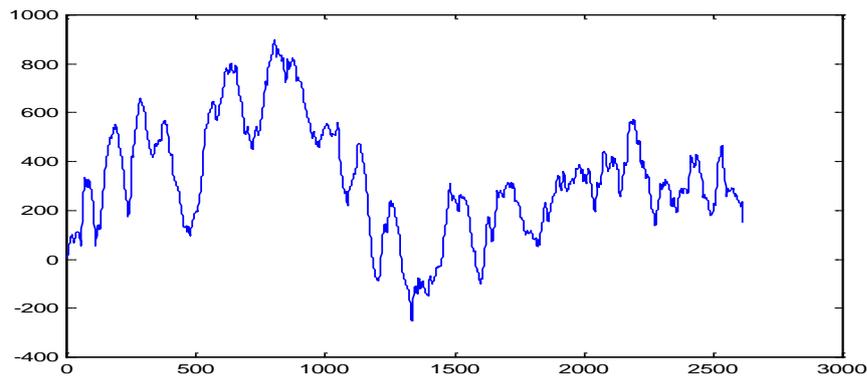


Figure4.17 : synthétique+AR(1) ($\phi=1$) après seuillage (sure dépendant d'échelle)

4-1-1 évaluation qualitative

Dans le cas d'un bruit blanc, nous remarquons qu'une grande partie du bruit a été éliminé (figure 4.3, 4.4) que ça soit par l'utilisation d'un seuillage sure ou universel. De plus, l'apparition de certaines fluctuations qui sont dues au bruit dans le résultat du seuillage sure. Quant au cas d'un bruit corrélé tel que AR(1) avec $\phi=0.2$, un seuillage global n'effectue pas un bon seuillage car on a la présence d'un remarquable taux de bruit et le passage à un seuillage dépendant d'échelle nous permet d'avoir une net amélioration des résultats.

4-1-1 évaluation quantitative

L'évaluation quantitative se fait par le biais de l'erreur en moyenne quadratique entre le signal original et le seuillage obtenu après seuillage dont les valeurs sont synthétisées dans les tableaux ci dessous (_tab4.1 et tab 4.2)

| Paramètre | $\sigma=1$ | | $\sigma=2$ | | $\sigma=4$ | | $\sigma=6$ | | $\sigma=8$ | |
|-----------|-------------|------------------|-------------|------------------|-------------|------------------|-------------|------------------|-------------|------------------|
| | Sure global | Universel global |
| EQM | 0.082 | 0.194 | 0.095 | 0.210 | 0.099 | 0.287 | 1.04 | 1.87 | 2.11 | 2.64 |

Tab4.1 : les erreurs en moyenne quadratique du signal synthétique bruité (bruit blanc) Après seuillage

| | |
|--------|-------------------|
| modèle | synthétique+AR(1) |
|--------|-------------------|

| Paramètre | $\varphi=0.2$ | | | | $\varphi=0.8$ | | | | $\varphi=0.95$ | | | |
|-----------|---------------|-------------|-----------------|-----------------|---------------|-------------|----------------|-----------------|----------------|------------------|----------------|---------------------|
| | Sure global | Univ global | Sure dép. d'éch | Univ dép. d'éch | Sure global | Univ global | Sure dép d'éch | Univ dép. d'éch | Sure global | Universel global | Sure dép d'éch | Univ dép. d'échelle |
| EQM | 8.53 | 9.56 | 5.81 | 5.29 | 32.69 | 33.26 | 30.73 | 31.05 | 65.48 | 66.10 | 61.13 | 62.92 |

**Tab4.2 :les erreurs en moyenne quadratique du signal synthétique bruité (bruit ar(1))
Après seuillage**

une remarque importante a faire, est que les erreurs obtenus après un seuillage sure sont inférieures a celles obtenues après utilisation d'un seuillage universel. Dans le cas le cas d'un bruit blanc les erreurs sont très petites et augmentent légèrement avec le paramètre σ . Dans le cas d'un bruit AR(1), nous constatons que les erreurs obtenues après utilisation d'un seuillage dépendant d'échelle sont inférieures a celles obtenues après un seuillage global (Tab4.2) et que celles-ci sont relativement petites lorsque le paramètre φ est petit et atteignent de grandes valeurs lorsque le paramètre φ s'approche de la valeur 1.

5- conclusion

Le travail décrit dans cet exposé nous a permis en premier lieu de déterminer un estimateur d'une fonction sans connaissance a priori sue cette fonction

Ceci nous a induit a révéler les avantages des méthodes de seuillage basées sur la décomposition en ondelettes orthogonales tant sur le plan théorique que sur le plan pratique. En effet, on a constaté que l'estimateur obtenu a partir des données bruitées par l'utilisation d'un seuil sure est d'un risque qui est a une constante près du risque minimax et qu'il lui est assyptotiquement équivalent dans le cas d'un seuillage dépendant d'échelle .contrairement au seuil universel dont le risque de l'estimateur qui est facteur de logn près du risque minimax.

A partir de l'étude qui a été faite sur le signal, synthétique, on a constaté que l'utilisation d'un seuillage global donne de bons résultats dans le cas d'un bruit blanc gaussien, contrairement au cas d'un bruit corrélé tel que AR (1), l'apport d'un seuillage dépendant d'échelle est positif car il permet d'améliorer les résultat en particulier dans le cas d'un bruit stationnaire. En comparant les erreurs calculées a partir de l'utilisation des seuils sure et universel, on dira que les méthodes basées sur le seuil sure sont optimales.

6- Bibliographie

Cohen A. (1992): ondelettes et traitement numérique du signal, édition Masson.

Daubechies I.: orthonormal bases of compactly supported wavelet, pure app Math, 41,909-996, 1988.

Donoho D.L. & Johnston I.M.: Minimax risk over IL_p balls for IL_q errors. Department of statistics, Stanford university, ca 94305.July 20, 1994.

Donoho D.L. & Jonstone I.M.: Minimax estimation via wavelet shrinkage. Technical report n^0 402. Department of statistics, Stanford University, 1992.

Donoho D.L. & Johnston I.M.: wavelet shrinkage asymptotia, 1995.

Donoho D.L. & Johnston I.M.: Adapting to unknown smoothness via wavelet shrinkage. Department of statistics, Stanford University, July, 1994.

Donoho D.L. & Johnston I.M.: Ideal adaptation by wavelet shrinkage. Department of statistics, Stanford University, ca 94305-4065, USA.

Donoho D.L. & Silverman B.W.: Wavelet threshold estimations for data with correlated noise. Department of statistics, Stanford University, Decembre1994.