

# Approche par blocs du filtrage des signaux piloté par les données

Abdel-Ouahab BOUDRAA<sup>1</sup>, Jean-Christophe CEXUS<sup>2</sup>

<sup>1</sup>IRENav (EA 3634), Ecole Navale/Arts-Métiers ParisTech, BCRM Brest, CC 600, 29240 Brest Cedex 9, France.

<sup>2</sup>Lab-STICC (UMR 6285), ENSTA Bretagne, 2 rue François Verny, 29806 Brest Cedex 9, France.

boudra@ecole-navale.fr, jean-christophe.cexus@ensta-bretagne.fr

**Résumé** – Dans cet article on s’intéresse au problème du filtrage des signaux par une méthode pilotée par les données, à savoir la décomposition modale empirique (EMD). Le signal filtré est obtenu par une reconstruction partielle des modes empiriques (IMFs) extraits par l’EMD et ce en utilisant un critère énergétique mesurant l’écart quadratique moyen entre deux reconstructions consécutives, le CMSE (Consecutive MSE). L’approche CMSE, largement utilisée dans la littérature du domaine, fournit un moyen simple pour estimer des signaux, en présence de bruit. Cela étant, ce filtrage est global. Ainsi, pour les signaux d’intérêt caractérisés, par exemple, par des détails fins, des événements courts ou peu contrastés (non stationnaires) cette approche a tendance à mal les reconstruire au détriment des grandes structures du signal informatif. On propose dans cet article, une approche locale ou par blocs de la méthode CMSE qui permet d’améliorer les performances de la reconstruction du signal d’intérêt. Des résultats de simulation sont présentés et comparés à ceux de l’approche par ondelettes.

**Abstract** – The aim of this work is the signals filtering by a data driven approach, namely the empirical mode decomposition (EMD). The filtered signal est estimated by the partial reconstruction of a reduced number of sifted empirical modes (IMFs) using an energetic criterion, the Consecutive MSE (CMSE). This CMSE-based filtering, largely used in the literature of the domain, is a simple way to estimate signals of interest embedded in noise. However, this strategy is a global one and thus informative events such as fine details or short events are not well reconstructed at the expense of large structures of the signal of interest. To improve the performances of the filtering, a local filtering version is proposed where CMSE strategy is applied block wise. Simulation results are presented and compared to those of wavelet-based approach.

## 1 Introduction

Le filtrage est une opération fondamentale, en particulier, en traitement des signaux, des images, ou encore en automatique. L’objectif de cette opération consiste à extraire des informations d’intérêt à partir de données bruitées. Cette opération peut être linéaire ou non-linéaire [1]. Couramment, des méthodes linéaires telles que le filtrage de Wiener sont utilisées car relativement faciles à mettre en œuvre [2]. Cependant, ces méthodes ne sont pas aussi efficaces lorsque les signaux présentent des arêtes vives (ruptures) ou encore des impulsions de courte durée. De plus, les signaux sont souvent non-stationnaires et/ou multi-composantes. Afin de pallier ces lacunes, les ondelettes fournissent des moyens simples pour estimer de façon non-linéaire les structures pertinentes des signaux, en présence de bruit [3]. Le postulat sous-jacent dans les méthodes de débruitage par ondelettes est le choix de l’ondelette mère la mieux adaptée au signal à débruiter [1]. Néanmoins, l’une des limites des ondelettes repose sur le fait que les fonctions de base sont fixes et peuvent ne pas correspondre totalement aux caractéristiques du signal. Pour contourner ce problème, des méthodes de type Matching-Pursuit permettant de décomposer le signal en une combinaison de différents atomes peuvent être utilisées [4]. Comme pour les ondelettes, si le dictionnaire est suffisamment important et riche, il est possible de

représenter une très grande classe de signaux. Cependant, les fonctions de base doivent être précisées (fonctions de Gabor, sinusoides amorties...).

Huang *et al.* proposent une approche non-linéaire pilotée par les données appelée décomposition modale empirique (EMD) permettant l’analyse des données issues de processus non-stationnaires et non-linéaires [5]. Il s’agit d’une méthode itérative, locale et auto-adaptative contrairement aux méthodes traditionnelles. L’EMD est basé sur un algorithme permettant d’extraire les diverses échelles (composantes appelées *Intrinsic Mode Functions* - IMF) présentes dans le signal en commençant par les échelles les plus fines (modes de plus hautes fréquences) pour terminer par les échelles les plus grossières (modes de plus basses fréquences) [6; 7]. Cette décomposition a donné lieu à plusieurs applications [8–18].

Dans ce travail on s’intéresse au problème de filtrage des signaux par reconstruction partielle par un nombre réduit d’IMFs sélectionnés en utilisant un critère énergétique, appliqué bloc par bloc. Ce filtrage est une extension de l’approche développée dans [12].

L’article est structuré comme suit. La section 2 est un rappel du principe de l’EMD. La section 3 présente l’ensemble de la méthode de filtrage. Des résultats de simulations sont proposés dans la section 4. Enfin, la dernière section 5 donne quelques conclusions et perspectives.

## 2 EMD

L'EMD décompose un signal  $x(t)$  en un ensemble fini de modes oscillatoires (IMFs) via un processus itératif appelé algorithme de tamisage [5]. L'IMF représente un mode oscillatoire qui peut être à la fois modulé en amplitude et en fréquence (AM-FM). Par construction, l'EMD sélectionne itérativement et localement les oscillations de fréquence la plus élevée qui reste dans le signal. Ainsi, localement, chaque IMF contient des oscillations de fréquence inférieure à celle extraite juste avant. L'EMD décompose un signal en une somme d'IMFs respectant principalement deux contraintes : (R1) Une IMF a le même nombre de passages par zéros et d'extrema ; et (R2) Une IMF est symétrique par rapport à la moyenne locale [5]. Le processus d'extraction d'une IMF est décrit comme suit :

---

**Algorithme :** EMD (méthode classique).

---

- Step 1.** Trouver les maxima/minima de  $x(t)$ .
  - Step 2.** Calculer l'enveloppe supérieure  $U(t)$  et inférieure  $L(t)$  en interpolant respectivement les maxima et minima à l'aide d'une spline cubique.
  - Step 3.** Calculer l'enveloppe moyenne locale :  

$$m(t) = [U(t) + L(t)]/2 .$$
  - Step 4.** Soustraire la moyenne locale pour obtenir le mode oscillant :  $h(t) = x(t) - m(t)$ .
  - Step 5.** Calculer le critère d'arrêt  $SD$  [5] :  
 si  $h(t)$  respecte le critère d'arrêt  $SD$ , alors  $h(t)$  est une IMF sinon définir  $x(t) = h(t)$  et répéter le processus de tamisage à partir de **Step 1**.
- 

Une fois la première IMF estimée, le même processus est appliqué au résidu  $x(t) - h(t)$  afin d'extraire les IMFs suivantes. Le processus de tamisage est également itéré plusieurs fois afin d'obtenir  $h(t)$  comme étant une IMF satisfaisant aux conditions (R1) et (R2). Au final,  $x(t)$  est décomposé en  $IMF_j(t); j = 1; \dots; C$  et un résidu  $r_C(t)$  :

$$x(t) = \sum_{j=1}^C IMF_j(t) + r_C(t) ; \quad (1)$$

où  $C$  est le nombre de modes déterminé automatiquement. Pour garantir que les IMFs conservent suffisamment de sens physique, le critère d'arrêt  $SD$  (fixé habituellement entre 0,2 et 0,3) est calculé initialement entre deux itérations successives de tamisage [5]. La somme des IMFs extraites permet de reconstruire le signal initial et assure que la décomposition est complète sans aucune perte d'informations [5].

## 3 Principe du filtrage

Soit un signal  $y(t)$  entaché d'un bruit additif  $n(t)$  :

$$x(t) = y(t) + n(t) ; \quad (2)$$

L'objectif est de trouver une approximation  $\tilde{y}(t)$  du signal  $y(t)$  qui minimise l'erreur quadratique moyenne (MSE) :

$$MSE(y; \tilde{y}) \triangleq \frac{1}{N} \sum_{i=1}^N (y(t_i) - \tilde{y}(t_i))^2 ; \quad (3)$$

avec  $N$  le nombre d'échantillons de  $y(t)$  (ou  $x(t)$ ).

Pour cela, le signal  $x(t)$  est tout d'abord décomposé en IMFs (Eq. 1) suivi par la reconstruction partielle de  $\tilde{y}(t)$  en sélectionnant les  $(C - k + 1)$  dernières IMFs [12] :

$$\tilde{y}_k(t) = \sum_{j=k}^C IMF_j(t) + r_C(t) ; \quad (4)$$

Le principe du filtrage par EMD revient à identifier l'indice  $k = j_s$  permettant de minimiser la quantité  $MSE(y; \tilde{y})$ . En pratique, le MSE (Eq. 3) ne peut pas être estimé directement car le signal d'origine  $y(t)$  est inconnue. Pour contourner cette difficulté, il est possible d'estimer une mesure de distorsion entre deux reconstructions consécutives  $\tilde{y}_k$  et  $\tilde{y}_{k+1}$  (Consecutive MSE - CMSE) ne nécessitant pas la connaissance de  $y(t)$  [12] :

$$CMSE(\tilde{y}_k; \tilde{y}_{k+1}) \triangleq \frac{1}{N} \sum_{i=1}^N (\tilde{y}_k(t_i) - \tilde{y}_{k+1}(t_i))^2 ; \quad (5)$$

$$\triangleq \frac{1}{N} \sum_{i=1}^N IMF_k^2(t_i) ; \quad (6)$$

Ainsi, le CMSE correspond à l'énergie (ou variance empirique) de la  $k^{th}$  IMF. Enfin, l'indice  $j_s$  est donné par :

$$j_s = \arg \min_{1 \leq k \leq C-1} [CMSE(\tilde{y}_k; \tilde{y}_{k+1})] ; \quad (7)$$

Le critère CMSE permet d'identifier l'IMF $_j$  à partir duquel il y a le premier changement significatif d'énergie [12]. Le principe du filtrage basé sur le CMSE est une stratégie globale. Ainsi, pour les signaux d'intérêt caractérisés, par exemple, par des détails fins ou des événements locaux cette méthode a tendance à mal les reconstruire au détriment des grandes structures présentes dans le signal.

Une approche locale, par blocs de la méthode CMSE permet d'améliorer les performances de la reconstruction du signal d'intérêt. Le CMSE est appliqué par bloc en découpant l'ensemble des  $C$  IMF $(t)$  en  $P$  blocs (définis arbitrairement) sans chevauchement de  $D$  échantillons chacun tel que  $DP \leq N$ . Pour une IMF $(t)$  donnée, le  $p^{eme}$  bloc est alors formé des échantillons :

$$IMF_j^{(p)}(t) = IMF_j(pD + t) ; \quad (8)$$

avec  $0 \leq t \leq D - 1$  et l'exposant  $(p)$  désigne le numéro du bloc. Il est possible d'estimer indépendamment les quantités  $\tilde{y}_k^{(p)}$  et  $\tilde{y}_{k+1}^{(p)}$  ainsi que l'indice  $j_s^{(p)}$  associé au  $p^{eme}$  bloc à l'aide du CMSE :

$$j_s^{(p)} = \arg \min_{1 \leq k \leq C-1} [CMSE(\tilde{y}_k^{(p)}; \tilde{y}_{k+1}^{(p)})] ; \quad (9)$$

Ainsi pour chaque bloc  $(p)$  du signal  $x(t)$ , il est possible de reconstruire une version filtrée  $\tilde{y}_{j_s^{(p)}}^{(p)}(t)$  :

$$\tilde{y}_{j_s^{(p)}}^{(p)}(t) = \bigotimes_{j=j_s^{(p)}}^{\mathcal{X}} \text{IMF}_j^{(p)}(t) + \mathbf{r}_C^{(p)}(t) : \quad (10)$$

Pour finir, le signal  $\tilde{y}(t)$  est le résultat de la concaténation des  $P$  blocs :

$$\tilde{y}(t) = \bigotimes_{p=0}^{P-1} \tilde{y}_{j_s^{(p)}}^{(p)}(t) : \quad (11)$$

avec  $\bigotimes$  le symbole de concaténation. Le fait de traiter les  $C$  IMFs par bloc pour déduire l'indice  $j_s^{(p)}$  permet d'obtenir un "masque binaire temporel" spécifique à chacune des IMFs qui a la capacité de mettre en lumière les zones pertinentes propre à chacune des IMFs. Les différentes étapes du filtrage CMSE par blocs sont décrites ci-dessous.

---

**Algorithme :** Filtrage par CMSE sur  $P$  blocs.

---

- Step 1.** Exécuter l'algorithme EMD pour obtenir les  $\text{IMF}_j(t)$  avec  $j = 1; \dots; C$  de  $x(t)$ .
- Step 2.** Diviser le support temporel des  $\text{IMF}_j(t)$  en  $P$  blocs non-superposés de  $D$  échantillons.
- Step 3.** Pour chaque  $\text{IMF}_j^{(p)}(t)$  avec  $p = 0; \dots; P-1$  :
- Step 3.1** Calculer  $\tilde{y}_k^{(p)}$  pour  $k = 1; \dots; C-1$  (Eq. 4)
- Step 3.2** Calculer CMSE( $\tilde{y}_k^{(p)}; \tilde{y}_{k+1}^{(p)}$ ) via l'équation (5).
- Step 3.3.** Trouver l'indice  $j_s^{(p)}$  via l'équation (9).
- Step 3.4.** Reconstruire le signal filtré du  $p^{\text{ème}}$  bloc :  $\tilde{y}_{j_s^{(p)}}^{(p)}(t)$  via l'équation (10).
- Step 4.** Reconstruire le signal filtré  $\tilde{y}(t)$  (Eq. 11).
- 

## 4 Résultats et Discussions

Le filtrage est testé sur un signal  $x(t)$  de taille  $N = 5120$  composé de plusieurs lois de fréquences variant au court du temps (Figure 1). La fréquence d'échantillonnage est de  $f_s=1024$  Hz. Un bruit ( $n(t)$ ) additif blanc Gaussien (RSB= 6.9dB) complète le signal. Le critère d'arrêt  $SD$  de l'EMD est fixé à 0.25 [5].

La figure 2 montre les 11 composantes (10 IMFs plus le résidu) obtenues à partir de  $x(t)$ . Les 3 premières IMFs (hautes fréquences) peuvent être associées essentiellement à du bruit. Certaines IMFs possèdent (voir notamment l'IMF<sub>4</sub>) des intervalles spécifiques avec du *signal pertinent* (d'autres parties de l'IMF étant moins intéressantes pour la reconstruction du signal filtré  $\tilde{y}(t)$ ) (Figure 2(b)).

La figure 3(a) montre la reconstruction IMF par IMF basé sur le CMSE classique. L'indice  $j_s$  estimé (Eq. 7) permet de conserver les 8 dernières composantes (7 IMFs plus le résidu) et fournit un RSB de 16.1dB. La figure 3(b) montre la reconstruction du signal par la méthode proposée (avec  $P= 64$ ,  $D= 80$ ) permettant d'obtenir un RSB supérieur de 18.6dB. Les trois dernières lignes de la figure 3(b) permet d'illustrer qu'aucune parties des IMFs est prise en compte. Quelques portions des IMF<sub>5</sub> et IMF<sub>4</sub> (en plus des autres IMFs) sont exploitées pour reconstruire le signal (illustré par les courbes de couleur magenta représentant

la concaténation des différentes valeurs  $j_s^{(p)}$  pour les  $P$  blocs). La figure 4 permet de montrer l'effet du paramètre  $P$  ainsi que d'illustrer l'efficacité de la méthode comparativement à d'autres : Moyenneur (taille 5), Médian (taille 9) ou encore l'approche ondelettes (Symlets de Daubechies d'ordre 3) avec un seuillage doux [1].

Les figures 5 et 6 illustrent la méthode appliquée à un signal Doppler<sup>1</sup> ( $N=8192$ ,  $f_s=1\text{Hz}$ ) compromis par un bruit additif blanc Gaussien (RSB=2.23dB). Les résultats confirment l'intérêt d'exploiter le CMSE par blocs par rapport à la méthode CMSE global (notamment dans la partie des hautes fréquences) ainsi que son efficacité comparativement à des méthodes plus classique.

## 5 Conclusion

Dans cet article une version par blocs du filtrage des signaux piloté par les données est présentée. Le critère énergétique CMSE est appliqué sur des blocs, sans chevauchement, du signal d'entrée permettant de conserver ses structures locales, mal reconstruites au détriment des grandes structures du signal informatif par la version globale. Alors que l'indice  $j_s$  de reconstruction partielle de modes est fixe dans le filtrage global, les résultats obtenus montrent que ce dernier est bloc dépendant, ce qui permet de tenir compte des informations locales du signal. Les résultats montrent l'intérêt et le potentiel de la version CMSE par blocs pour le filtrage des signaux. Pour confirmer l'efficacité de cette méthode, celle-ci doit être évaluée avec une plus large classe de signaux et dans différentes conditions expérimentales telles que les niveaux de bruit, les taux d'échantillonnage et les tailles d'échantillon, ainsi que la taille des blocs avec et sans chevauchement. Comme travaux futurs, nous prévoyons d'étendre cette étude à d'autres types de bruit tels que le bruit Gaussien fractionnaire.

## Références

- [1] D. Donoho, "De-noising by soft-thresholding," *IEEE TIT*, vol. 41(3), 1995.
- [2] J. Proakis *et al.*, *Digital Signal Processing : Principles, Algorithms, and Applications*. Prentice-Hall, 1996.
- [3] D. Donoho *et al.*, "Ideal spatial adaptation by Wavelet shrinkage," *Biometrika*, vol. 81(3), 1994.
- [4] S. Mallat *et al.*, "Matching Pursuits with Time-Frequency dictionaries," *IEEE TSP*, 1993.
- [5] N. Huang *et al.*, "The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Royal Soc. London. Series A : Mathematical, Physical and Engineering Sciences*, 1998.
- [6] P. Flandrin *et al.*, "Empirical Mode Decomposition as a filter bank," *IEEE SPL*, 2004.
- [7] Z. Wu *et al.*, "A study of the characteristics of white noise

---

1. Obtenue via le logiciel WAVELAB, disponible auprès de Stanford Statistics Department, avec l'aimable autorisation de D.L. Donoho and I.M. Johnstone.

