

Débruitage Aveugle par Décompositions Parcimonieuses et Aléatoires

Manuel MOUSSALLAM¹, Alexandre GRAMFORT¹, Laurent DAUDET², Gaël RICHARD¹

¹Institut Mines-Telecom - Telecom ParisTech - CNRS/LTCI UMR 5141. 75014 Paris

²Institut Langevin- CNRS UMR 7587 - Univ. Paris Diderot - ESPCI ParisTech. 75005 Paris

manuel.moussallam@espci.fr

Résumé – On s’intéresse au débruitage de séries temporelles sans connaissance du niveau de bruit ni de sa nature exacte. La partie informative d’un mélange est supposée parcimonieuse dans un dictionnaire de formes d’onde, tandis que la composante de bruit y est incohérente. Une décomposition à l’aide d’algorithmes gloutons peut permettre une séparation, mais nécessite l’optimisation de paramètres (taille du dictionnaire, critère d’arrêt, etc) qui requiert généralement une connaissance, au minimum, du niveau de bruit. Nous travaillons sur les projections normalisées du bruit sur les éléments du dictionnaire. Nous proposons un critère d’arrêt automatique, basée sur une modélisation originale de l’étape de sélection des algorithmes gloutons. Nous introduisons ensuite un double aléa dans cette étape de sélection pour contrer les effets de biais introduits par la structure du dictionnaire. Notre algorithme présente des performances en aveugle comparables à celles obtenues par des méthodes de l’état de l’art en “oracle” (optimisées par la connaissance du niveau de bruit). La méthode est également pertinente pour le débruitage de signaux réels M/EEG.

Abstract – This work addresses the problem of denoising time series in the absence of precise knowledge on the noise level or nature. The signal’s informative part is assumed sparse in a dictionary of small oscillatory waveforms on which the additive noise is expected to be incoherent. Denoising can be achieved, e.g. by greedy algorithms, however, those typically require some parameter adaptation (e.g. the stopping condition) that generally require some knowledge on the noise level. We propose a blind approach that uses an assumption on the noise normalized projections distribution to derive a self-stopping criterion. Multiple runs of randomized pursuits are then use to reduce biais introduce by the dictionary elements. The proposed approach reaches denoising performances comparable to state of the art methods informed by the noise level. We also show that real signals such as M/EEG recording can be efficiently processed.

1 Débruitage par décomposition parcimonieuse

L’acquisition de données expérimentales telles que des séries temporelles - p. ex. la magnétoencéphalographie et l’électroencéphalographie (M/EEG) - est fréquemment perturbée par du bruit. Le débruitage des signaux ainsi acquis peut alors être effectué à partir d’hypothèses sur la nature du signal d’intérêt, noté x , et la nature et/ou la puissance du bruit, noté b . Dans le cas d’un bruit additif, le modèle du signal mesuré s’écrit

$$y = x + b \in \mathbb{R}^N$$

. Une hypothèse sur le signal x consiste à supposer qu’il peut être efficacement approché par une combinaison linéaire d’un nombre réduit k d’atomes, pris dans un ensemble (redondant) appelé dictionnaire $\Phi \in \mathbb{R}^{N \times M}$. Le débruitage consiste en la résolution d’un problème de la forme :

$$\min_{\alpha} \|\alpha\|_0 \text{ sous la contrainte } \|y - \Phi\alpha\| \leq \epsilon \quad (1)$$

où $\|\alpha\|_0$ est la pseudo-norme ℓ_0 qui compte le nombre d’entrées non nulle d’un vecteur $\alpha \in \mathbb{R}^M$. Une estimation du signal débruité est $\hat{x} = \Phi\alpha$ pour lequel $\|\alpha\|_0 = k$

Différentes méthodes de décomposition parcimonieuses peuvent être utilisées, en particulier les algorithmes gloutons de

type Matching Pursuit (MP) [6]. Malheureusement, ces méthodes nécessitent en pratique la connaissance, soit de la parcimonie des données k , soit de la nature et la puissance du bruit, afin de fixer ϵ . De plus, la décomposition présente une dépendance forte au choix du dictionnaire utilisé [3]. Dans ce travail, nous proposons des solutions à ces deux limitations : 1) Nous utilisons une modélisation de la convergence de MP pour s’affranchir de la connaissance préalable de la parcimonie, et 2) nous introduisons de l’aléatoire à différentes étapes pour contrer le biais introduit par le choix du dictionnaire.

Sur le premier point, une modélisation à base de statistiques d’ordre de la convergence des algorithmes gloutons est proposée, qui permet de stopper la sélection d’atomes lorsque la corrélation de ceux-ci passe sous un seuil prédéfini, indépendant du signal et de la puissance du bruit.

Pour le deuxième point, nous étendons une idée déjà présente dans la littérature sous le nom de décompositions stochastiques. Celle-ci stipule qu’une moyenne de représentations sous-optimales (en termes de parcimonie) peut donner de meilleurs résultats (en termes d’erreur quadratique) qu’une seule représentation, fût-elle la plus parcimonieuse. Durka et al [3], et plus récemment Elad et al [2] ont ainsi proposé des algorithmes qui construisent un ensemble de J représentations $\{\alpha^j\}_{j=0..J-1}$ obtenues par J décompositions gloutonnes stochastiques, dans

lesquelles un aléa est introduit à l'étape de sélection d'un nouvel atome. Dans [2], soit R^{n-1} le signal résiduel à l'itération n , l'atome d'indice γ^n est choisi parmi les M éléments ϕ_γ de Φ avec une probabilité d'autant plus grande que la projection $|\langle R^{n-1}, \phi_{\gamma^n} \rangle|$ est grande.

2 Approche proposée

2.1 Double aléa dans l'étape de sélection

Nous proposons une approche différente qui reprend une stratégie utilisée en apprentissage statistique par la méthode *Random Forest* [4]. A chaque itération l'atome ajouté est le meilleur parmi un sous-ensemble de $m < M$ atomes de Φ pris aléatoirement. Le sous-ensemble varie à chaque itération de chaque décomposition. Ceci augmente l'aspect stochastique de la décomposition et permet de parcourir à moindre coût des dictionnaires de taille importante, vu que seules m projections sont calculées à chaque fois. Cette stratégie est d'autant plus efficace que les éléments du dictionnaire sont finement localisés en temps et en fréquence (voir [1]). Le signal débruité \hat{x} est ensuite obtenue par moyennage $\hat{x} = \frac{1}{J} \Phi \sum_j \alpha^j$.

2.2 Critère d'arrêt stochastique

Le second apport de ce travail est la proposition d'un critère d'arrêt original, aveugle, basé sur une modélisation de la convergence des algorithmes gloutons. Que les propriétés statistiques du signal résiduel changent au cours d'une exécution d'un algorithme MP est un phénomène connu, déjà étudié dans [6] à travers la variable :

$$\lambda_\Phi(R^{n-1}) = \sup_{\phi \in \Phi} \frac{|\langle R^{n-1}, \phi \rangle|}{\|R^{n-1}\|}$$

dénommée cohérence normalisée. Si le produit $|\langle R^{n-1}, \phi \rangle|$ est utilisé comme critère de sélection, on peut lier la décroissance de l'énergie du signal résiduel à cette grandeur :

$$\frac{\|R^n\|^2}{\|R^{n-1}\|^2} = 1 - \lambda_\Phi^2(R^{n-1}) \quad (2)$$

Cette expression est généralement utilisée pour garantir la convergence de MP (en bornant $\lambda_\Phi^2(R^{n-1})$). Mais l'on peut également s'en servir pour modéliser la décomposition moyenne d'un signal dans un dictionnaire, en particulier d'un bruit. Le processus MP se comporte, en effet, après les premières itérations comme un processus chaotique et le résiduel tend vers un attracteur nommé bruit de dictionnaire, noté W . Dans une optique de débruitage, il est intéressant de stopper alors la décomposition. Si on connaît l'espérance de la cohérence normalisée $\Lambda_W(\Phi) = \mathbb{E}[\lambda_\Phi(W)]$ de ce bruit de dictionnaire, seuls les atomes dont la cohérence normalisée est significativement plus importante que $\Lambda_W(\Phi)$ sont susceptibles d'être conservés [6].

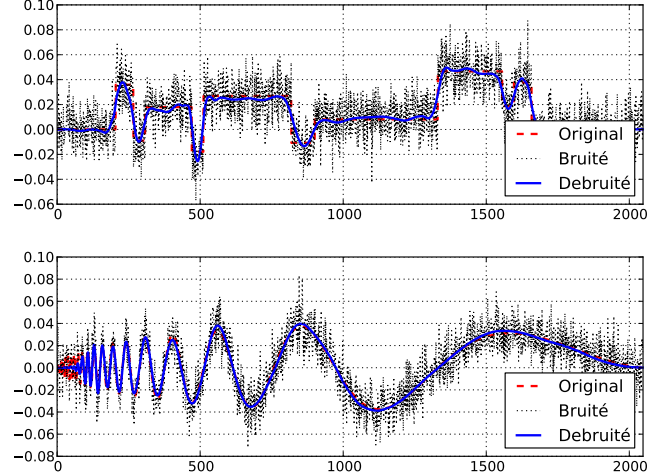


FIGURE 1 – Débruitage avec RPDN de signaux synthétiques *blocks* et *doppler* avec bruit additif (SNR de 5dB) et dictionnaire multi-échelle MDCT.

2.3 Estimation de $\Lambda_W(\Phi)$

L'originalité de notre approche consiste en la proposition d'une méthode générique pour le calcul approché de $\Lambda_W(\Phi)$. Nous modélisons les projections $|\langle R^{n-1}, \phi \rangle|$ comme les réalisations d'une variable aléatoire Z de densité f_Z et estimons ensuite la cohérence normalisée à l'aide de statistiques d'ordre : $\Lambda_W(\Phi) = \mathbb{E}[\lambda_\Phi(W)] = \mu_{M:M}$ où M est le nombre de tirages c'est-à-dire le nombre d'atomes du dictionnaire Φ_M . La statistique $\mu_{M:M}$, qui décrit le maximum parmi M échantillons, s'obtient à partir de la densité de probabilité $f_{M:M}^Z(z) = M f_Z(z)^{M-1} f_Z(z)$. En particulier si l'on suppose une distribution normale des projections de W dans Φ :

$$\frac{\langle W, \phi_i \rangle}{\|W\|} \sim \mathcal{N}(0, \sigma_W^2)$$

alors Z suit une loi demi-normale de paramètre σ_Z^2 . Comme on considère les projections normalisées, la variance σ_Z^2 est indépendante de la puissance de W et peut s'obtenir analytiquement. Pour une variable demi-normale, le calcul de l'espérance $\mu_{M:M}$ est complexe. En revanche, ces distributions sont unimodales et on peut vérifier que la médiane $\nu_{M:M}$ est proche de l'espérance. Or l'estimation de la médiane est beaucoup plus simple :

$$\nu_{M:M} = \sigma_Z \sqrt{2} \operatorname{erf}^{-1} \left(0.5^{\frac{1}{M}} \right) \approx \mu_{M:M}$$

Dès lors, connaissant la taille M du dictionnaire, le débruitage consiste à stopper chaque décomposition dès lors que

$$\sqrt{1 - \frac{\|R^n\|^2}{\|R^{n-1}\|^2}} \leq \Lambda_W(\Phi) \simeq \nu_{M:M} \quad (3)$$

On s'affranchit ainsi à la fois de la connaissance du degré de parcimonie, mais également de la nature et la puissance du bruit. La seule hypothèse est que les projections *normalisées* de

celui-ci dans le dictionnaire considéré soit distribuées normalement, ce qui est un modèle pertinent pour tout signal incohérent dans le dictionnaire choisit. L'algorithme est dénommé RPDN pour *Random Pursuit Denoising* et est résumé ci-dessous :

Input : $y, \Phi, J, \Lambda_W(\Phi)$

Output : \tilde{y}

for $j = 1..J$ **do**

initialization : $n = 1, R^0 = y$;

while condition do

Tirer au hasard $\Phi_n \subset \Phi$;

Selection $\gamma^n = \arg \max_{\phi \in \Phi_n} |\langle R^{n-1}, \phi \rangle|^2$;

Mise à jour $\hat{y}^j \leftarrow \hat{y}^j + \langle R^{n-1}, \phi_{\gamma^n} \rangle$;

et $R^n = y - \hat{y}^j$;

condition = $\lambda_{\Phi}(R^{n-1}) > \Lambda_W(\Phi)$;

Moyenne des décompositions : $\tilde{y} = \frac{1}{J} \sum_j \hat{y}_j$;

Algorithme 1: Random Pursuit Denoising (RPDN)

Dans la pratique, on observe que l'hypothèse de normalité des projections du bruit se vérifie dans de nombreux cas de figures. Cette hypothèse se révèle donc moins restrictive que, par exemple, une hypothèse de bruit blanc Gaussien.

3 Applications

3.1 Comparaison sur signaux synthétiques

Nous avons comparé notre algorithme avec des méthodes existantes, telles que le *Wavelet Shrinkage* avec invariance temporelle [5], sur des signaux synthétiques usuels. Ces méthodes sont appliquées sur des transformées en ondelettes ainsi que sur une transformée de Fourier à court terme, avec seuillage doux ou dur. Pour chaque configuration, nous effectuons une optimisation du seuil. Nous présentons le meilleur résultat obtenus parmi toutes les configurations testées. Nous comparons également avec les décompositions stochastiques [3, 2] (libellées SMP), Celles-ci utilisent un dictionnaire composé d'une union de bases MDCT. De la même façon, différents niveaux d'approximation cible ϵ sont testés et seul le meilleur score obtenu est présenté.

Une telle optimisation nécessitant la connaissance du signal original, nous parlons de méthodes *Oracle*. Nous utilisons comme métrique l'erreur quadratique moyenne (EQM) de reconstruction

$$EQM(\hat{y}, y) = 10 \log_{10} \frac{\|\hat{y} - y\|_2}{\|y\|_2}$$

Les résultats (Table 1) montrent que notre méthode permet, en aveugle, des niveaux de débruitage comparables à ces méthodes Oracle. Le type de résultat obtenus est illustré par la Figure 1.

Il est intéressant de voir que notre méthode en aveugle permet d'obtenir des résultats proches des scores optimaux. Sur certains des signaux considérés, RPDN en aveugle réalise même

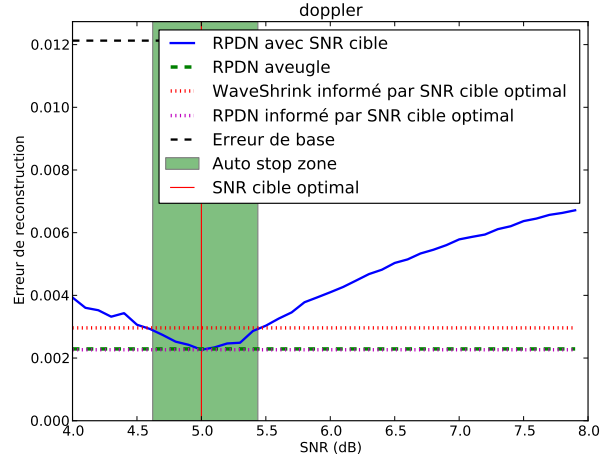


FIGURE 2 – Erreur quadratique de reconstruction et zone d'arrêt automatique des multiples décompositions de RPDN pour un signal *doppler* avec 5 dB de bruit.

les meilleurs scores (plus petite erreur quadratique de reconstruction).

On constate expérimentalement (Figure 2) que chaque décomposition s'arrête automatiquement dans une zone de SNR cohérente. Le critère d'arrêt automatique est donc pertinent. En fonction des signaux considérés et des niveaux de bruit, cette zone est automatiquement adaptée et reste toujours pertinente (autour du niveau de SNR Oracle). Notons également que la conjonction de ces multiples décompositions permet d'obtenir une erreur tout à fait équivalente à celle obtenue avec des méthodes informées.

3.2 Débruitage de signaux M/EEG

Dans le cas des signaux M/EEG, le SNR n'est pas connu, dès lors notre méthode aveugle présente un intérêt évident. La Figure 3 illustre les résultats obtenus sur 2 capteurs de MEG. Dans les deux cas présentés, 100 décompositions aléatoires sont lancées et stoppées automatiquement entre la 8e et la 19e itération de MP. Le niveau de bruit moyen estimé est d'environ 9 dB. Dans cette configuration réelle, en l'absence de connaissance précise sur la nature du bruit, on peut constater que l'algorithme fournit encore des résultats pertinents. Les paramètres des méthodes traditionnelles (*WaveShrink* ou *SMP*) sont, dans ce cas de figure, difficiles à optimiser.

Autres types de signaux L'application de cette méthode à d'autres types de signaux (en particulier les sons et les images) est possible. Nos futurs travaux tenteront de traiter des bruits complexes (bruits colorés, environnementaux,...) à l'aide de dictionnaires adaptés (ondelettes..) pour lesquels l'hypothèse de normalité de la distribution des projections du bruit reste vérifiée.

Méthode	Oracle WaveShrink			Oracle SMP			RPDN		
	-5 dB	0 dB	5 dB	-5 dB	0 dB	5 dB	-5 dB	0 dB	5 dB
Doppler	-9.5 (4)	-12.7 (0)	-16.4 (5)	-10.9 (78)	-14.8 (95)	-15.7 (0)	-10.2 (18)	-13.7 (5)	-17.5 (95)
Blocks	-9.6 (89)	-12.2 (65)	-14.6 (98)	-8.9 (11)	-12.0 (35)	-13.7 (2)	-7.6 (0)	-10.7 (0)	-13.3 (0)
Bumps	-6.3 (0)	-10.7 (0)	-14.3 (57)	-8.5 (100)	-12.2 (100)	-13.9 (13)	-6.0 (0)	-10.7 (0)	-14.1 (30)
HeaviSine	-11.4 (0)	-15.6 (0)	-19.6 (1)	-14.3 (27)	-17.5 (1)	-16.5 (0)	-15.2 (73)	-18.9 (99)	-21.5 (95)

TABLE 1 – Erreur Quadratique Moyenne (dB) atteinte et proportion de meilleurs scores (%) avec 100 répétitions d’expériences pour divers signaux synthétiques et niveaux de bruit. En gras les meilleures moyennes pour un couple expérimental Signal-niveau de bruit.

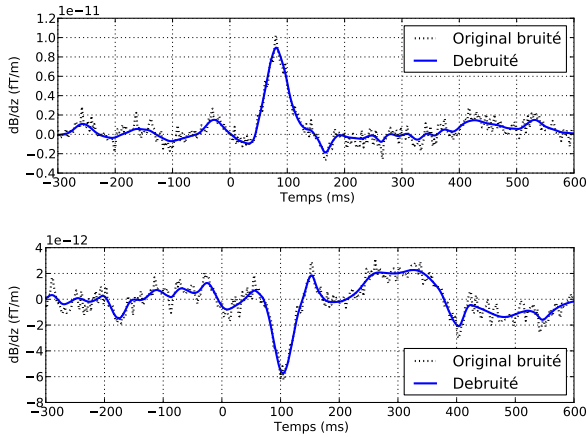


FIGURE 3 – Exemples de débruitage de données réelles de M/EEG (dictionnaire multi-échelle MDCT).

Références

- [1] Moussallam, M., Daudet, L., and Richard, G. Matching Pursuits with Random Sequential Subdictionaries. *Signal Processing*, 92, 2532-2544. 2012
- [2] Elad, M., and Yavneh, I. A Plurality of Sparse Representations Is Better Than the Sparsest One Alone. *IEEE Transactions on Information Theory*, 55(10), 4701-4714. 2009
- [3] Durka, P. J., Ircha, D., and Blinowska, K. J. Stochastic time-frequency dictionaries for matching pursuit. *IEEE Transactions on Signal Processing*, 49(3), 507-510. 2001
- [4] Breiman, L. Random forests. *Machine learning*, 45(1), 5-32. 2001.
- [5] Coifman, R., and Donoho, D. Translation-invariant denoising. Technical Report. 1995, Dept of Statistics Stanford University.
- [6] Mallat, S. and Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12), 3397-3415.(1993)