

Déconvolution aveugle dans le domaine fréquentiel fondée sur le taux d'information mutuelle

Anthony LARUE, Jérôme MARS, Christian JUTTEN

Laboratoire des Images et des Signaux,
961 rue de la Houille Blanche 38402 Saint Martin d'Hères BP 46, France
larue@lis.inpg.fr, mars@lis.inpg.fr, juttten@lis.inpg.fr

Résumé – Nous introduisons le taux d'information mutuelle comme mesure de blancheur regroupant l'ensemble des statistiques d'ordre supérieur. Ensuite, nous l'utilisons pour construire un algorithme de déconvolution aveugle fréquentiel (DAF) d'un système SISO incorporant aussi une régularisation limitant l'amplification du bruit. Enfin, le traitement de données réelles d'explosions sous-marines et de sismologie est proposé.

Abstract – We introduce the mutual information rate as a whiteness measure using all higher order statistics. Then, we propose a new frequency blind deconvolution algorithm of a SISO system including a regularization to avoid the noise amplification. Finally, we process real data of underwater explosions recordings and seismovolcanic signals.

1 Introduction

Le problème de déconvolution aveugle se rencontre dans de nombreux domaines d'applications (télécommunications, contrôle non-destructif, détection de choc, géophysique...). L'observation résulte de la convolution entre un signal d'entrée $r(t)$ et un filtre direct w plus un bruit additif $b(t)$ soit: $d(t) = (w \star r)(t) + b(t) = \sum_{i=-\infty}^{+\infty} w(i)r(t-i) + b(t)$. \star est le produit de convolution. Dans le contexte de la déconvolution aveugle, seule l'observation $d(t)$ est connue, alors que le filtre direct et le signal d'entrée sont inconnus. Mais sous l'hypothèse d'un processus d'entrée iid et non Gaussien, le signal d'entrée est unique à un rapport d'amplitude et un retard près. Le modèle de convolution et l'algorithme de déconvolution sont résumés par la figure 1.

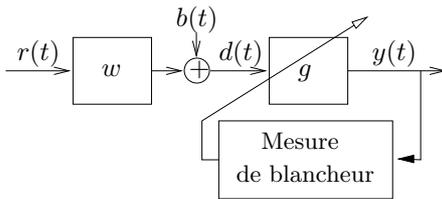


Figure 1: Modèle de convolution et déconvolution

Généralement, les algorithmes de déconvolution sont réalisés en ajustant un filtre inverse g suivant un critère de blancheur portant sur le signal de sortie $y(t)$. Pour mesurer le caractère iid de la sortie, on peut utiliser les statistiques d'ordre 2 (densité spectrale et fonction de corrélation). Ces méthodes (égalisation spectrale, Yule-Walker, Durbin) sont simples mais ne permettent pas l'estimation de la phase du filtre. Soit, on fixe la phase du filtre par des informations *a priori*, soit on utilise les statistiques d'ordre supérieur (à deux). Classiquement, on préfère les statistiques d'ordre 4 (tricorrelation [1, 2], kurtosis [3] et trispectre) à celles d'ordre 3 (bicorrelation, bispectre) car ces dernières sont nulles pour des distributions symétriques.

Les performances de ces méthodes sont très influencées par la qualité des estimateurs des statistiques d'ordre supérieur (SOS). De plus, la sélection arbitraire des SOS n'assure pas le choix optimal des statistiques utilisées pour les signaux traités. Ainsi, nous proposons de mesurer la blancheur par une mesure plus générale : le taux d'information mutuelle (TIM) de la sortie.

2 Algorithme de déconvolution

L'information mutuelle d'un vecteur aléatoire $\mathbf{z} = (z_1, \dots, z_n)$ de dimension n est définie par:

$$I(\mathbf{z}) = \sum_{i=1}^n H(z_i) - H(z_1, z_2, \dots, z_n) \quad (1)$$

où $H(z_i)$ est l'entropie de Shannon marginale de z_i , i.e. $H(z_i) = -\int_{\mathbb{R}} p_{z_i}(u) \log p_{z_i}(u) du$ et $H(z_1, \dots, z_n)$ l'entropie de Shannon conjointe $H(\mathbf{z}) = -\int_{\mathbb{R}^n} p_{\mathbf{z}}(\mathbf{u}) \log p_{\mathbf{z}}(\mathbf{u}) d\mathbf{u}$. L'information mutuelle $I(\mathbf{z})$ est un critère d'indépendance pour la séparation de sources car elle est toujours positive et $I(\mathbf{z})$ s'annule si et seulement si les composantes de \mathbf{z} sont mutuellement indépendantes. En déconvolution aveugle, on doit mesurer l'indépendance entre les échantillons d'un processus, soit une séquence de longueur infinie. Ainsi, on définit le taux d'entropie [4] d'un processus aléatoire $Z = \{Z_t\}$ par :

$$\mathcal{H}(Z) = \lim_{T \rightarrow \infty} \frac{1}{T} H(Z_1, \dots, Z_T) \quad (2)$$

Cette limite existe pour un processus stationnaire [4]. Puis, le taux d'information mutuelle (TIM) du processus aléatoire Z est défini par:

$$I(Z) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{t=T} H(Z_t) - \mathcal{H}(Z) \quad (3)$$

où $H(Z_t)$ représente l'entropie marginale du processus Z_t et $\mathcal{H}(Z)$ le taux d'entropie de Z défini en (2). Pour

plus de clarté, Z représente le processus contenant tous les échantillons, et Z_t est le processus avec un seul échantillon $z(t)$. Le taux d'information mutuelle $I(Z)$ est toujours positif et s'annule si et seulement si Z est un processus iid [4]. Ainsi, le TIM est une mesure de blancheur acceptable pour construire des algorithmes de déconvolution aveugle (Cf figure 1). Sous hypothèses de stationnarité, l'entropie marginale du processus Z_t est indépendante du temps. Tous les termes de la somme de (3) sont égaux, donc nous avons:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{t=T} H(Z_t) = H(Z_\tau) \quad \forall \tau \quad (4)$$

Par conséquent, en ajoutant l'hypothèse de stationnarité, le taux d'information mutuelle (3) se réécrit simplement:

$$I(Z) = H(Z_\tau) - \mathcal{H}(Z) \quad (5)$$

Dans (5), τ est arbitrairement choisi. Mais, il n'y a pas de raison de privilégier une valeur de τ . De plus, plusieurs réalisations du processus Z_τ sont nécessaires pour estimer convenablement l'entropie marginale $H(Z_\tau)$. Ainsi, en pratique, pour estimer l'entropie de Z_τ , nous utilisons tous les échantillons $z(1), \dots, z(T)$ comme T réalisations du processus Z_τ . Dans la suite, nous considérerons comme équivalent le processus Z_t et une réalisation de ce dernier $z(t)$. Dans le contexte de la déconvolution aveugle, on peut obtenir un critère plus simple [5, 6], en remarquant que le taux d'entropie de la sortie $y(t) = (g \star d)(t)$ s'exprime de la façon suivante:

$$\mathcal{H}(Y) = \mathcal{H}(D) + \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) \exp(-jt\theta) \right| d\theta \quad (6)$$

Avec (6), le TIM de la sortie Y est:

$$I(Y) = H(y(\tau)) - \mathcal{H}(D) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) \exp(-jt\theta) \right| d\theta \quad (7)$$

Puis, comme le taux d'entropie des données $\mathcal{H}(D)$ est indépendant du filtre inverse g , il est équivalent de mesurer la blancheur de la sortie avec (7) ou $\tilde{I}(Y)$:

$$\tilde{I}(Y) = H(y(\tau)) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) \exp(-jt\theta) \right| d\theta \quad (8)$$

qui est minimum quand le processus Y est iid. Ce critère utilise toutes les statistiques d'ordre supérieur par l'intermédiaire de l'entropie. Avec cette mesure de blancheur, on peut envisager différents algorithmes de déconvolution. Ces algorithmes sont équivalents à des méthodes de maximum de vraisemblance, dans lesquelles on remplace la distribution du signal de réflectivité supposée connue par le maximum de vraisemblance par la distribution de la sortie estimée à chaque itération. Taleb *et al.* [5] propose dans le cas d'un système de Wiener, de minimiser (8) par rapport à la réponse impulsionnelle du filtre $g(t)$. En prenant, un nombre fini de coefficients pour $g(t)$, la méthode est adaptée à l'utilisation d'un modèle MA pour le filtre inverse donc un filtre direct w autorégressif. Cet

algorithme (MAMV) n'est pas très robuste au bruit. En présence de bruit additif Gaussien, des expériences montrent que ses performances sont identiques à la méthode d'ordre 2 de Yule-Walker. **Nous proposons de construire un critère de déconvolution dans le domaine fréquentiel à l'aide de la mesure de blancheur (8) égale à une constante près au TIM.** Nous travaillerons en fréquence, car l'amplification du bruit (supposé large bande) sur la sortie est contrôlée simplement en limitant les grandes valeurs de $G(f)$. Ensuite, on ajoute une contrainte de continuité sur la réponse en fréquence du filtre inverse. Ainsi, le critère de déconvolution en fonction de la réponse en fréquence $\mathbf{G} = [G(0), \dots, G(T-1)]$ est:

$$J(\mathbf{G}) = H(y(\tau)) - \frac{1}{T} \sum_{f=0}^{T-1} \log |G(f)| + \lambda_1 \sum_{f=0}^{T-1} |G(f) - G(f+1)|^2 + \lambda_2 \sum_{f=0}^{T-1} |G(f)|^p \quad (9)$$

Les deux premiers termes sont l'estimation de la mesure de blancheur (8). Le terme pondéré par l'hyperparamètre λ_1 est la contrainte de continuité sur le spectre du filtre inverse. Le dernier terme est l'équivalent du facteur de bruit pour le filtrage de Wiener, pénalisant avec la norme \mathbb{L}^p les grandes valeurs de $G(f)$ qui amplifient des bandes de fréquences représentant uniquement du bruit. Ces deux termes de régularisation peuvent être analysés comme deux *a priori* sur la distribution de la réponse en fréquence du filtre (i) une loi conditionnelle gaussienne entre $G(f)$ et $G(f+1)$, (ii) une loi Gaussienne généralisée d'ordre p pour la loi marginale de $G(f)$. Le minimum de (9) est obtenu par une descente de gradient [7] par rapport aux coefficients complexes $G(f)$, $f = 0, \dots, T-1$. Le gradient de (9) par rapport à $G(f)$ est :

$$\begin{aligned} \widehat{\nabla J}(\mathbf{G}) &= \frac{1}{2T^2} \Psi_Y(f) D^*(f) - \frac{1}{2T} \frac{1}{G^*(f)} + \\ &+ \lambda_1 (2G(f) - G(f+1) - G(f-1)) \\ &+ \lambda_2 \frac{p}{2} \frac{|G(f)|^p}{G^*(f)} \end{aligned} \quad (10)$$

où $\Psi_Y(f)$ est la transformée de la fonction score définie par $\psi_Y(u) = -\frac{d}{du} \log p_Y(u)$. L'information des SOS est contenue dans cette fonction. **L'algorithme de déconvolution aveugle en fréquence (DAF)** est le suivant:

1. *initialisation du filtre inverse $G(f)$ et de $y(t)$*
2. *estimation de la fonction score ψ_Y ;*
3. *calcul du gradient (10);*
4. *mise à jour de $G(f) \leftarrow G(f) - \mu \widehat{\nabla J}(\mathbf{G})$;*
5. *calcul de la sortie $y(t)$;*
6. *étape de normalisation.*

μ est le pas du gradient. Nous itérons entre les étapes 2 et 6 jusqu'à la convergence de l'algorithme. L'étape de normalisation permet de lever l'indétermination d'amplitude sur la sortie. La fonction score est estimée par une méthode à noyaux développée par Pham [8] dans le cadre de la séparation de sources. Cet estimateur a de bonnes performances quelque soit la distribution du signal car il est basé sur une méthode très générale à noyaux. L'utilisation d'un noyau à support borné (splines cubiques) donne un coût de calcul assez faible.

3 Exemples

Le but des premiers tests est de montrer la capacité de notre algorithme à estimer des filtres quelque soit leurs phases. L'expérience des figures 2 et 3 utilise comme signal d'entrée (en (a)) un signal de réflectivité (applications : contrôle non-destructif, géophysique) modélisé par un signal Bernoulli-Gaussien avec 50 réflecteurs pour 400 échantillons. Les observations en (b) sont le résultat du filtrage de ce signal d'entrée par un filtre ARMA(10,10) à phase minimale (respectivement nulle) pour la figure 2 (resp. 3). La qualité de la déconvolution peut être analysée dans le domaine temporel en comparant le signal déconvolué en (c) au signal d'entrée en (a), et dans le domaine fréquentiel, on compare en (d) (resp. (e)) le module (resp. la phase) du filtre inverse estimé (en trait continu) à la valeur théorique (trait pointillé).

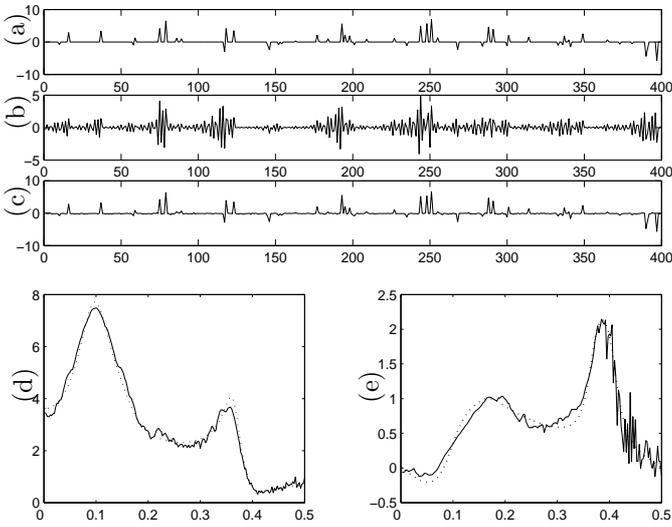


Figure 2: (a) Signal Bernoulli-Gaussien d'entrée. (b) Observation après filtre ARMA(10,10) à phase minimale. (c) Signal déconvolué par algorithme DAF. (d) (resp. (e)) Module (resp. phase) du filtre inverse estimé (trait continu) comparé avec la valeur théorique (trait pointillé)

Cette expérience avec des signaux simulés montre que l'estimation de filtre à phase non-minimale est réalisable avec l'algorithme DAF grâce à l'utilisation des SOS par le biais du TIM. L'erreur d'estimation sur la phase est due aux grands nombres de paramètres à optimiser. De plus, l'erreur est la plus grande dans la plage [0.4 0.5] qui est la partie la moins énergétique du spectre des données.

La figure 4 représente en (a) un enregistrement réel en piscine de la pression lors d'une explosion sous-marine. L'observation est constituée de trois ondes : (i) l'onde directe, (ii) une réflexion sur le fond, (iii) une réflexion à la surface. En appliquant le modèle de convolution à ces données, le signal d'entrée contient les informations relatives entre l'arrivée directe de l'onde et les réflexions. Le filtre direct représente l'onde émise. Cette onde est caractéristique de l'effet bulle avec une croissance exponentielle de la pression suivie d'une décroissance exponentielle. Pour les méthodes paramétriques, on retient un

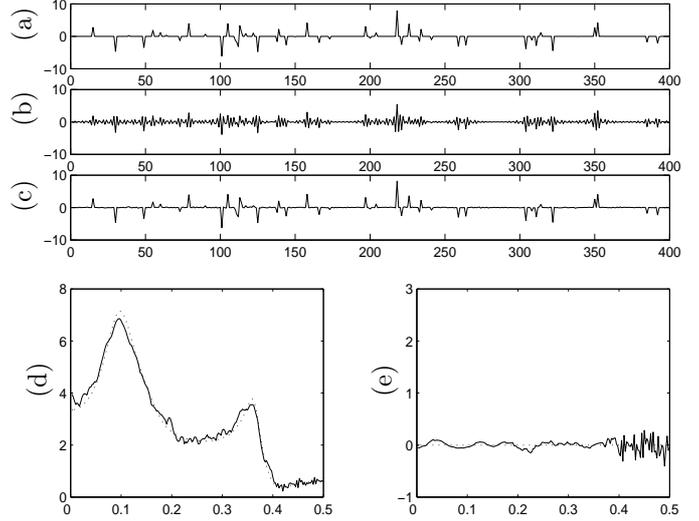


Figure 3: (a) Signal Bernoulli-Gaussien d'entrée. (b) Observation après filtre ARMA(10,10) à phase nulle. (c) Signal déconvolué par algorithme DAF. (d) (resp. (e)) Module (resp. phase) du filtre inverse estimé (trait continu) comparé avec la valeur théorique (trait pointillé)

modèle MA non causal pour cette onde. Dans la figure 4, nous comparons trois algorithmes: (b) l'égalisation spectrale, (c) la méthode de Durbin, (d) l'algorithme DAF.

L'algorithme DAF donne le meilleur compromis entre la qualité de la déconvolution (pics étroits, sans pics secondaires) et le niveau de bruit (100 premiers échantillons).

L'algorithme a une approche très générale (hypothèse de blancheur, non paramétrique), une autre application est la sismologie [9]. Le signal enregistré figure 5(a) par un géophone à la surface du volcan Purace (Colombie) présente une très forte résonance. Dans le modèle de la figure 1, le signal d'entrée est l'excitation due à une explosion à l'intérieur du volcan. La cheminée du volcan joue le rôle de filtre résonant. Le but de la déconvolution est de séparer les effets résonants (caractéristiques de la géométrie du volcan) et l'excitation liée à l'activité du volcan. L'excitation liée à une explosion sera supposée blanche. La figure 5 compare quatre algorithmes. Nous comparons la sortie estimée (colonne de gauche) correspondant au signal d'excitation et la réponse en fréquence du filtre direct estimée (colonne de droite) pour les fréquences réduites comprises entre 0 et 0.15 pour quatre algorithmes : (b) et (c) égalisation spectrale (ordre 2, fréquentiel); (d) et (e) Yule-Walker (ordre 2, temporel); (f) et (g) MAMV [5] (TIM, temporel); (h) and (i) DAF (TIM, fréquentiel). L'algorithme MAMV minimise le TIM par rapport au coefficient de la réponse impulsionnelle du filtre inverse. Ainsi, nous avons les quatre combinaisons possible de méthodes avec le choix de deux domaines de travail et l'utilisation uniquement des statistiques d'ordre 2 ou de l'ensemble des statistiques d'ordre supérieur par le biais du TIM.

Pour juger la qualité de la déconvolution, sur les allures temporelles, on doit obtenir une excitation très localisée temporellement, avec aucun effet de résonance. Le com-

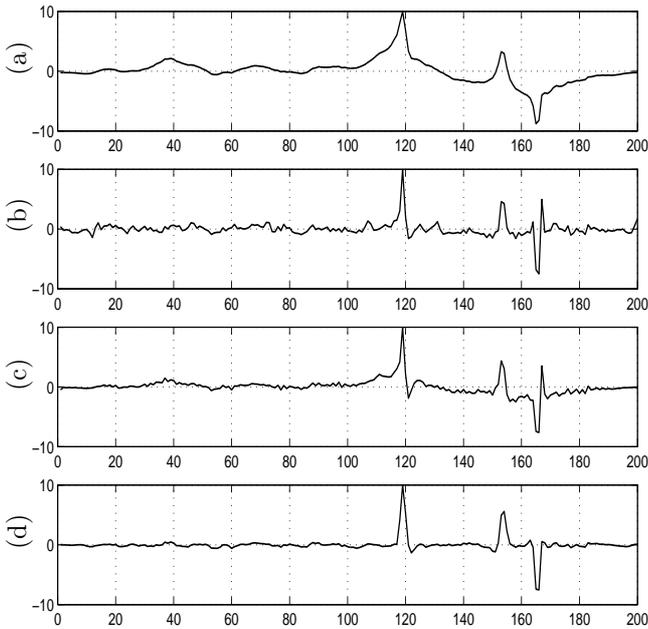


Figure 4: (a) Observation de l'explosion sous-marine (b) Egalisation spectrale, (c) Méthode de Durbin, (d) Algorithme DAF avec $\mu = 0.01$, $\lambda_1 = 0.1$, $\lambda_2 = 0.1$ et $p = 3$

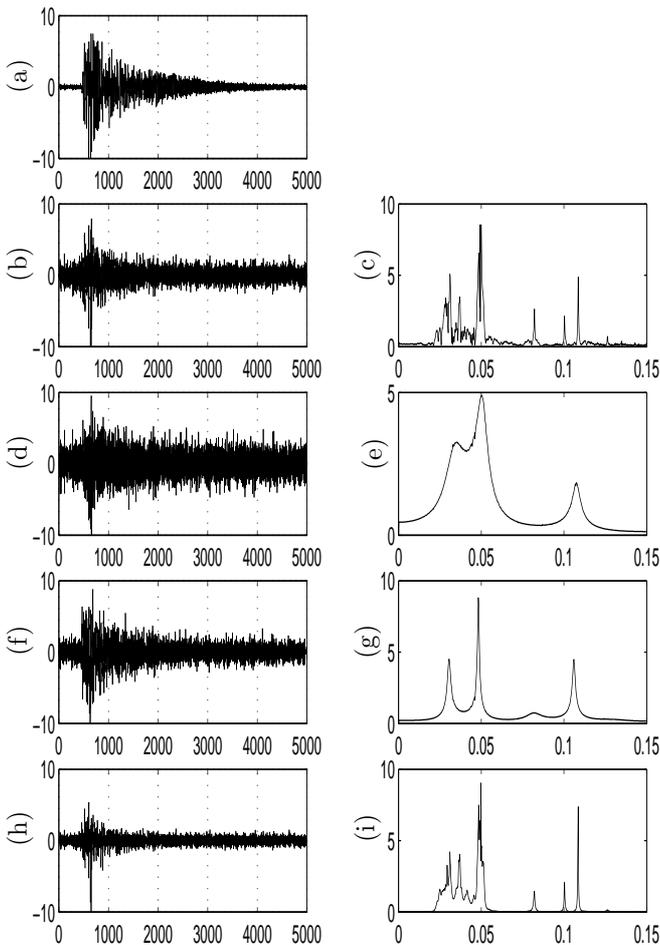


Figure 5: (a) Observations sismovolcaniques, sortie et filtre direct estimés :(b)(c) égalisation spectrale; (d) et (e) Yule-Walker; (f) et (g) MAMV; (h) et (i) DAF

promis entre la qualité de la déconvolution et l'amplification du bruit est tout simplement liée au niveau de bruit sur la sortie. Le spectre du filtre direct estimé permet de donner des informations sur les fréquences de résonance présentes dans le signal. Ces fréquences caractérisent de la géométrie du volcan. La méthode doit réaliser un compromis entre extraire le maximum de fréquences de résonance tout en ayant un spectre assez lisse. En suivant tous ces critères, on observe que l'algorithme DAF donne les meilleurs résultats en terme de localisation de l'excitation, de niveau de bruit et de nombre de fréquences de résonance extraites.

4 Conclusion

Nous avons proposé un nouvel algorithme de déconvolution dans le domaine fréquentiel. Il est particulièrement intéressant pour des applications réelles, grâce à la régularisation vis à vis du bruit. Il s'adapte aux données en sélectionnant les meilleures statistiques d'ordre supérieur pour le traitement par l'intermédiaire de l'estimation de la fonction score de la sortie.

References

- [1] P. Comon, "MA identification using fourth order cumulants," *Signal Processing*, vol. 26, 1992.
- [2] J. A. Cadzow, "Blind deconvolution via cumulant extrema," *IEEE Signal Processing Magazine*, pp. 24–42, May 1996.
- [3] M. Boumahdi, "Blind identification using the kurtosis with applications to field data," *Signal Processing*, vol. 48, no. 3, pp. 205–216, 1996.
- [4] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.
- [5] A. Taleb, J. Solé i Casals, and C. Jutten, "Quasi-nonparametric blind inversion of Wiener systems," *IEEE Trans. on Signal Processing*, vol. 49, no. 5, pp. 917–924, May 2001.
- [6] D.-T. Pham, "Contrast functions for blind separation and deconvolution sources," in *Proc. of the Int. Conf. ICA '01*, San-Diego, December 2001, pp. 37–42.
- [7] A. Van Den Bos, "Complex gradient operator and hessian," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 141, pp. 380–382, 1994.
- [8] D.-T. Pham, "Fast algorithm for mutual information based independent component analysis," *IEEE Trans. on Signal Processing*, vol. 52, no. 10, pp. 2690–2700, October 2004.
- [9] P. Lesage, F. Glangeaud, and J. Mars, "Applications of autoregressive models and time-frequency analysis to the study of a volcanic tremor and long-period events," *Journal of Volcanology and Geothermal Research*, vol. 114, no. 3, pp. 391–417, 2002.