

Un algorithme rapide séquentiel de déconvolution L^2

A Fast Sequential Algorithm for L^2 Deconvolution

par S. MOTTELET*, P. SIMARD**

* Division Mathématiques appliquées

** URA CNRS 817 Heudiasyc, Université de Technologie de Compiègne,
BP 649, 60206 Compiègne Cedex

Résumé

Dans cet article nous décrivons un algorithme séquentiel de déconvolution aux moindres carrés particulièrement bien adapté à l'estimation de séquences Bernoulli-Gaussiennes. Cet algorithme permet une recherche des événements par ordre croissant d'importance d'un critère qui sera défini, et donne d'excellents résultats dans un grand nombre de cas. Les performances font l'objet d'une étude approfondie qui permet d'apprécier la méthode vis-à-vis de différentes sources d'erreurs.

Mots clés : Contrôle non destructif, Déconvolution, Normes L^p , Méthodes de Descente

Abstract

In this paper we describe a sequential algorithm to perform a least squares deconvolution particularly well fitted to the estimation of Bernoulli-Gaussian sequences. This algorithm performs a search of events by increasing order of importance with respect to a criterion which is described in detail. It gives good results over a wide range of applications. An exhaustive analysis of performance aspects shows the robustness of the method with respect to various sources of error.

Key words : Non destructive testing, Deconvolution, L^p norms, Descent Methods

1. Introduction

Nous abordons ici le problème de la restauration d'une suite d'impulsions h dont on connaît une observation z en sortie d'un système linéaire connu sous la forme de sa réponse impulsionnelle w , observation perturbée par un bruit additif n . Ce type de problème apparaît dans tous les cas où une onde w se propage à travers un système stratifié. C'est un problème classique en géophysique où cette question fait l'objet de nombreuses publications. Ce type de préoccupation concerne également le contrôle non destructif par échographie ultrasonore, pour peu que le défaut recherché joue un rôle de réflecteur bien localisé dans l'espace (inclusions, fissures, délaminages, ...).

Les travaux publiés récemment sur la restauration de la séquence de réflectivité h proposent souvent des algorithmes *myopes*, c'est-à-dire considérant le système w inconnu. Dans ce cas il convient d'identifier ce système avant, ou conjointement à l'estimation de h .

Or en échographie ultrasonore, il n'est pas très difficile en général d'obtenir une copie assez juste de la réponse impulsionnelle du "système", puisqu'il s'agit dans ce cas simplement du signal émis par le transducteur, c'est-à-dire de la forme du champ de pression acoustique à la sortie de ce transducteur. Il existe d'ailleurs d'autres cas où le système est connu. C'est pourquoi nous proposons dans cet article de revisiter les algorithmes basés sur la minimisation de normes L^p , qui nécessitent la connaissance *a priori* du système w .

Avant de poursuivre, il convient de bien préciser que cette hypothèse n'est qu'à moitié justifiée, car on sait par exemple qu'au cours de sa propagation, une ondelette acoustique opère des rotations de phase aux interfaces, et donc, qu'en toute rigueur, cette application doit être classée dans les cas *myopes*. Nous avons toutefois quelques raisons de penser que l'utilisation de ces algorithmes simples peut être efficace dans un grand nombre d'applications.

Le problème de l'échographie ultrasonore consiste à essayer de reconstruire le plus précisément possible la séquence de réflectivité, c'est-à-dire la dérivée de l'impédance acoustique le long du

trajet parcouru par l'onde acoustique. En théorie, cette séquence peut être reconstruite de manière exacte par déconvolution des mesures. Il s'agit d'une séquence nulle presque partout, sauf aux points de rupture d'impédance acoustique, où elle présentera un pic. On décrit habituellement ce type de signal comme une séquence peu dense de pics ou d'impulsions.

2. Déconvolution par minimisation d'une norme L^p

Considérant le cadre décrit plus haut, nous adoptons pour les mesures le modèle de convolution suivant

$$z_k = \sum_i w_i h_{k-i} + n_k, \quad (1)$$

où $w \in \mathbf{R}^m$, $h \in \mathbf{R}^n$ et $z, n \in \mathbf{R}^{n+m-1}$. On peut écrire (1) de façon matricielle

$$z = Wh + n,$$

où W est la "matrice de convolution", construite à partir des échantillons de w . Le bruit n est supposé blanc, sans que cela soit une condition de fonctionnement de l'algorithme. Le principe de la déconvolution par minimisation d'une norme L^p consiste à résoudre le problème

$$\min_{y \in \mathbf{R}^n} J_p(y) = \|z - Wy\|_p^p + \lambda \|y\|_p^p, \quad (2)$$

où $\|x\|_p^p = (\sum_i |x_i|^p)$, p étant un nombre réel compris entre 1 et 2. Le vecteur y réalisant le minimum de $J_p(y)$ sera noté \hat{h} . La fonction $J_p(y)$ comporte deux termes. Le premier représente un résidu entre l'observation z et la reconstruction Wy . Le second est un terme de pénalisation (parfois appelé "préblanchiment" ou "perturbation de la diagonale"), qui permet de limiter l'influence du bruit n sur l'estimation de h .

Le procédé le plus naturel consiste à résoudre (2) pour $p = 2$. On est alors en présence d'un problème classique de moindres carrés, dont la solution est

$$\hat{h} = (W^t W + \lambda I)^{-1} W^t z. \quad (3)$$

La solution obtenue par (3) est parfaite en absence de bruit, mais devient inexploitable en présence de bruit. Dans le cas non pénalisé $\lambda = 0$, la présence de bruit additif introduit un nombre considérable de pics parasites : on obtient une estimation de h diffuse et répartie sur tous les échantillons, qui ne correspond pas à ce que l'on cherche (une séquence d'impulsions peu dense). Dans le cas pénalisé $\lambda \neq 0$, le terme de pénalisation conduit à un étalement des pics là encore incompatible avec la solution recherchée, et qui nuit à la résolution.

Partant de cette constatation, Taylor et al. [4] ont montré la possibilité d'utiliser la norme L^1 dans ce cas. On cherche alors à minimiser

$$J_1(y) = \|z - Wy\|_1 + \lambda \|y\|_1$$

Ce problème peut se mettre facilement sous la forme d'une minimisation d'une fonction linéaire sous contraintes linéaires (au prix d'un doublement du nombre de variables mises en jeu), problème classiquement résolu par l'algorithme du simplexe. Cet algorithme est très lourd et très gourmand en mémoire; c'est pourquoi il a été proposé des méthodes itératives ne tenant pas compte du caractère linéaire du problème. Deux de ces méthodes ont été proposées par Yarlagadda et al. [5], [1], il s'agit de l'IRLS (Iterative Reweighted Least Squares) et du RSD (Residual Steepest Descent). Plus récemment Mottelet [3] a proposé un algorithme rapide particulièrement efficace du point de vue numérique, et dont les résultats expérimentaux étaient équivalents à ceux obtenus avec l'IRLS.

Cependant ces algorithmes posent de nombreux problèmes lorsque l'on essaye de justifier leurs performances d'un point de vue théorique. L'idée que nous allons développer dans cet article consiste à utiliser le même type d'algorithme itératif en utilisant cette fois la norme L^2 . Le cadre hilbertien, perdu avec L^1 , est retrouvé, et permet une meilleure justification des résultats. Après avoir présenté cet algorithme, et l'avoir replacé dans le contexte plus large des méthodes de gradient, nous présenterons quelques résultats numériques permettant de tester les performances de cette méthode en termes de robustesse vis-à-vis du rapport signal à bruit, et vis-à-vis d'imprécisions dans la connaissance de l'ondelette.

3. Algorithme séquentiel de déconvolution L^2

3.1. LA MÉTHODE DE GRADIENT À PAS OPTIMAL (STEEPEST DESCENT)

Soit le problème

$$\min_{y \in \mathbf{R}^n} J_p(y), \quad (4)$$

La solution peut être obtenue à l'aide de la méthode du gradient à pas optimal, dont l'algorithme est le suivant :

$$y^{(k+1)} = y^{(k)} + \rho^{(k)} d^{(k)}, \quad \text{pour } k > 0 \quad (5)$$

où $y^{(0)} \in \mathbf{R}^n$ est donné, $d^{(k)} = -\nabla_y J_p(y^{(k)})$. Le pas optimal $\rho^{(k)}$ est solution du problème de minimisation unidimensionnelle

$$\min_{\rho} J(y^{(k)} + \rho^{(k)} d^{(k)}), \quad (6)$$

Remarque :

Si cet algorithme est utilisé sous sa forme classique, dans le cas $p = 2$, il conduit à la même solution que celle obtenue avec la méthode directe. L'algorithme de déconvolution en norme L^1 présenté dans [2], [3] reprend le principe de cette méthode mais introduit une modification quant au choix de la direction de descente $d^{(k)}$. Plutôt que d'agir simultanément sur toutes les composantes de $y^{(k)}$, on choisit d'agir uniquement sur la composante présentant la plus forte dérivée partielle en valeur absolue. L'algorithme que nous allons présenter lui est similaire, mais il utilise cette fois la norme L^2 , ce qui permet d'obtenir une expression formelle du pas optimal ρ , alors que dans le cas L^1 , ce pas optimal doit être recherché à l'aide d'un algorithme de minimisation (dichotomie ou autre).

3.2. DESCRIPTION DE L'ALGORITHME

Nous utilisons maintenant la fonction coût

$$J_2(y) = \frac{1}{2} (\|z - Wy\|^2 + \lambda \|y\|^2)$$

Pour des questions de clarté nous allons un instant omettre le numéro d'itération (k) .

Notons

$$g = \nabla J(y) = -W^t r + \lambda y$$

Il s'agit, pour $\lambda = 0$, du produit d'intercorrélacion entre w et le résidu $r = z - Wy$.

Conformément à la remarque précédente, on définit la direction de descente

$$d = [0 \ 0 \ \dots \ 0 \ -g_{j_{opt}} \ 0 \ \dots \ 0]^t$$

soit $d_j = -g_{j_{opt}} \delta_{j, j_{opt}}$, où j_{opt} est le numéro de la composante présentant la plus grande dérivée partielle

$$j_{opt} = \max_j |g_j|.$$

la recherche du pas optimal revient à trouver ρ solution du problème

$$\min_{\rho} J(y + \rho d).$$

Ce pas peut être obtenu de façon formelle. En effet en écrivant

$$J(y + \rho d) = \frac{1}{2} (\|z - W(y + \rho d)\|^2 + \lambda \|(y + \rho d)\|^2)$$

et en annulant sa dérivée par rapport à ρ on obtient

$$\rho = \frac{d^t d}{d^t (W^t W + \lambda I) d}$$

qui, compte tenu de la structure particulière de d , se met sous la forme

$$\rho = \frac{1}{w^t w + \lambda}$$

où w est le vecteur contenant les échantillons de l'ondelette. Ce résultat remarquable exprime le fait que le pas optimal est constant et égal, dans le cas $\lambda = 0$, à l'inverse de l'énergie de l'ondelette. Il peut donc être calculé une fois pour toutes au début de l'algorithme. Nous avons donc l'algorithme suivant :

– (i) initialisation

$$y^{(0)} \equiv 0, \quad \rho = \frac{1}{w^t w + \lambda}$$

– (ii) calcul du résidu et du gradient

$$r^{(k)} = z - Wy^{(k)}, \quad g^{(k)} = -W^t r^{(k)} + \lambda y^{(k)}$$

– (iii) déterminer la direction optimale j_{opt}

$$\max_j |g_j^{(k)}|$$

– (iv) mettre à jour $y^{(k)}$

$$\begin{aligned} y_{j_{opt}}^{(k+1)} &= y_{j_{opt}}^{(k)} - \rho g_{j_{opt}}^{(k)} \\ y_j^{(k+1)} &= y_j^{(k)}, \quad j \neq j_{opt} \end{aligned}$$

– retour en (ii)

Remarque :

On peut déterminer de façon arbitraire le nombre de boucles à effectuer ou spécifier un critère d'arrêt quelconque. Chaque passage détermine l'événement qui a le plus d'importance pour expliquer le résidu. En particulier, si l'on prend $\lambda = 0$, on peut montrer que les amplitudes des pics obtenus sont décroissantes. Lorsque l'on choisit la valeur $\lambda = 0$ (pas de pénalisation), l'algorithme conduit à une méthode intuitive : il s'agit d'un filtrage optimal (corrélacion de y avec w) suivi d'une détection. Le filtrage optimal permet de conserver un rapport signal à bruit raisonnable, et la détection évite une estimation diffuse de h . Cette méthode intuitive trouve ici sa justification du point de vue de l'optimisation.

La valeur $\lambda = 0$ apparaît être la plus naturelle au vu des résultats que nous avons obtenu. En effet, même en présence de bruit additif, l'algorithme se comporte très bien, pour peu que l'on limite le nombre d'itérations pour ne pas faire apparaître trop de pics parasites. Le choix d'un λ non nul se justifie normalement dès que l'on est en présence de bruit additif. Les résultats que nous avons obtenus montrent que si on ne limite pas le nombre d'itérations par un test d'arrêt adéquat, on obtient une estimation diffuse de h , problème identique rencontré avec la méthode L^2 directe. Comme test d'arrêt pour éviter ce problème, on peut par exemple stopper l'algorithme dès qu'il propose un pic d'amplitude supérieure à celui obtenu précédemment.

On a donc le choix dans le traitement du bruit. Soit on prend $\lambda = 0$ et on limite le nombre d'itérations, soit on choisit λ en rapport avec le rapport signal à bruit et on utilise le test d'arrêt proposé ci-dessus. Pour nos essais nous avons utilisé la première de ces deux options.

4. Etude des performances

Compte tenu du type de problème que nous cherchons à résoudre, il semble raisonnable de juger de la qualité des résultats obtenus selon trois critères : La précision de l'instant où l'événement est détecté, la précision de la valeur estimée de l'amplitude, et éventuellement la qualité de la reconstruction. Il est à noter que ce dernier critère ne fait pas l'unanimité car dans le contexte des problèmes mal posés qui est le notre ici, la solution n'est pas nécessairement unique. Ce critère n'a donc de valeur qu'associé aux deux précédents.

Pour apprécier ces critères, l'algorithme a été testé sur des signaux synthétiques réalisés par convolution de 50 séquences blanches Bernoulli-Gaussienne avec l'ondelette de Kramer d'ordre 4. Cette ondelette est maintenant bien connue des spécialistes, et ce choix permettra au lecteur de comparer les résultats présentés ici avec d'autres publications. Cette ondelette n'est pas à phase minimale, ce qui ne constitue pas un véritable handicap puisque elle sera supposée connue. La déconvolution s'opère donc en construisant W à l'aide des échantillons de cette ondelette. Cependant dans une seconde série d'essais, nous utiliserons des ondelettes modifiées pour tester la robustesse de l'algorithme. Puis nous présenterons quelques résultats sur signaux réels.

4.1. COMPORTEMENT DE L'ALGORITHME EN FONCTION DU RAPPORT SIGNAL À BRUIT

Les essais comportent donc à chaque fois une série de 50 signaux convenablement bruités par addition d'une séquence gaussienne

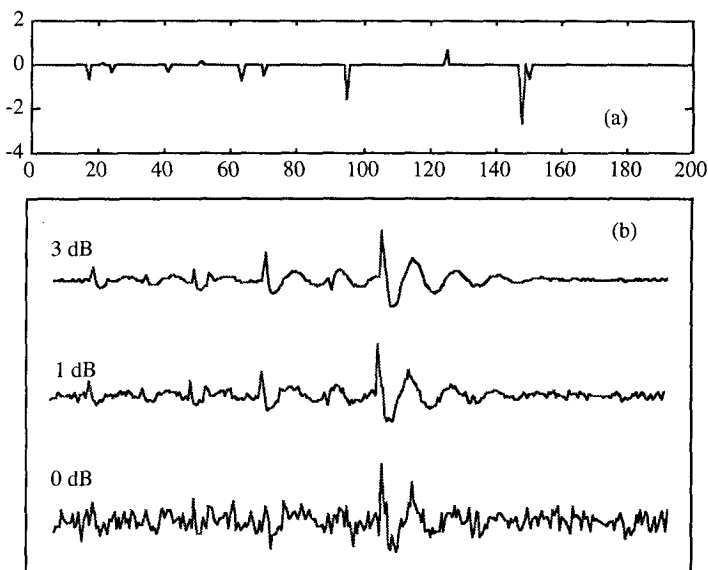


Figure 1. - (a) Séquence Bernoulli-Gaussienne à estimer. (b) Signaux obtenus par convolution de la séquence (a) avec l'ondelette de Kramer d'ordre 4, pour différents rapports signal à bruit.

pour obtenir des rapports signal à bruit variant de 0 à 3 dB (fig 1). A partir de ces essais, nous avons estimé les probabilités de bonne détection et de fausse alarme, et construit la courbe COR correspondante. Nous avons pris le parti de ne pas sur-interpréter les résultats : une détection est bonne si le pic est retrouvé à la position d'un pic dans la séquence d'origine. Dans le cas contraire, nous le considérons comme une fausse alarme.

On observe (fig 2) que dès 3 dB, les courbes convergent vers une sorte de plateau assurant 90 % de bonnes détections pour un taux négligeable de fausse alarme. Les détections manquées correspondent à des pics de petites amplitudes. Les résultats pour $RSB=1dB$ sont également convenables puisque l'on peut espérer détecter 80% des événements avec un taux de fausses alarmes inférieur à 5%. La situation se dégrade nettement à partir de 0dB, mais permet encore de détecter 50% des événements avec 5% de fausses alarmes. Si l'on compare ces résultats avec ceux fournis par l'inversion L^2 classique, on note d'une part que le comportement asymptotique correspond à un taux de bonnes détections qui ne dépasse pas 85%, et d'autre part que les courbes obtenues s'éloignent d'avantage de la courbe idéale. Pour le cas

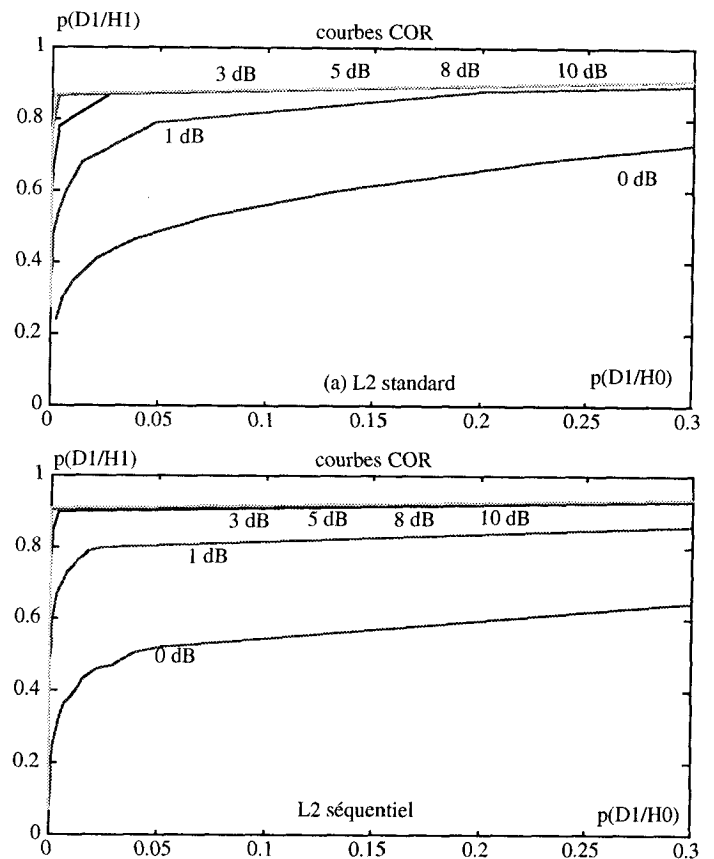


Figure 2. - Performances de la déconvolution.

Courbes COR liant le taux de fausses alarmes au taux de bonnes détections obtenues :

- (a) par une déconvolution classique au sens des moindres carrés.
- (b) par l'algorithme présenté dans cet article.

0dB, il est clair qu'à taux de détection identique, il faut accepter un taux beaucoup plus important de fausses alarmes. Il faut noter de plus que ces dernières courbes sont obtenues avec un seuillage, ce qui explique la progression de bonne détection lorsqu'on accepte un fort taux de fausses alarmes.

4.2. ROBUSTESSE DE LA MÉTHODE VIS-À-VIS DE L'ESTIMATION DE L'ONDELETTE

La méthode décrite concerne la déconvolution dite de signature, mais il n'est pas inutile de savoir dans quelle mesure la qualité des résultats dépend des erreurs commises sur l'ondelette utilisée. Pour cela, nous avons généré quatre ondelettes modifiées à partir de l'ondelette de Kramer d'ordre 4 de la façon suivante : (fig 3.1)

- Onde-1 : ondelette filtrée passe bas $f_c=0.3$
- Onde-2 : ondelette filtrée passe bas $f_c=0.8$
- Onde-3 : ondelette bruitée par addition d'une séquence blanche gaussienne
- Onde-4 : ondelette déphasée

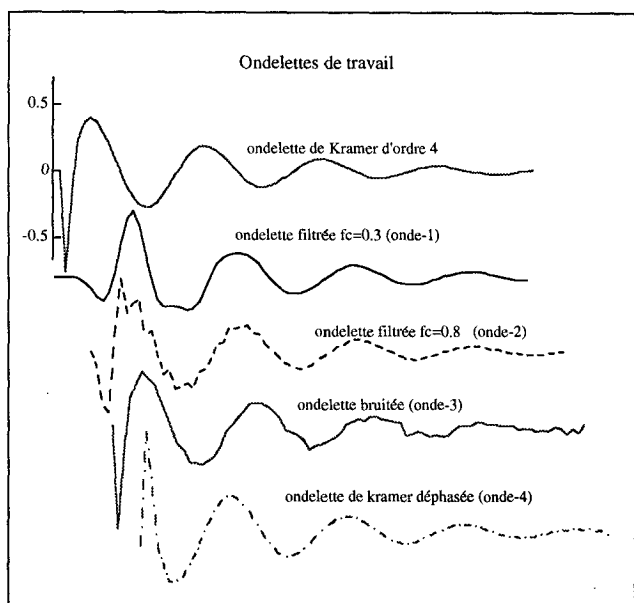


Figure 3.1. - Ondelettes utilisées pour la déconvolution.

Le signal est obtenu par convolution numérique d'une séquence Bernoulli Gaussienne comportant 11 pics, convenablement bruitées pour obtenir un rapport signal à bruit de 1dB (fig 3.2). Les résultats sont résumés dans le tableau 1

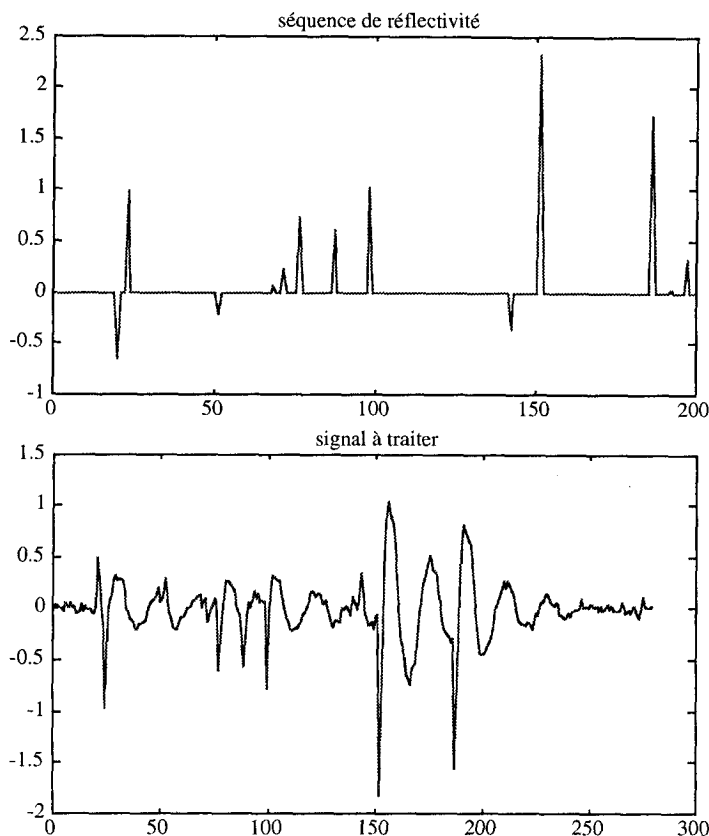


Figure 3.2. - Séquence à restaurer et signal synthétique obtenu par convolution avec une ondelette de Kramer d'ordre 4.

Ce que l'on observe n'a rien de très surprenant (fig 3.3). En ce qui concerne les ondelettes filtrées (Onde-1 et Onde-2), les performances se dégradent lorsque le contenu fréquentiel se détériore, et on retrouve le retard de groupe du filtre au niveau des instants de détection. L'addition d'un terme de phase (Onde-4) affecte le signe des pics, mais les performances restent bonnes. Quant à l'addition de bruit blanc, elle n'affecte pas vraiment la qualité des résultats.

Tableau 1

	Nombre d'événements	Bonnes Détections	Fausses alarmes	détections manquées	observations
Onde-1	11	8	4	3	retard de détection
Onde-2	11	9	0	2	léger retard de détection
Onde-3	11	10	0	1	peu d'effet
Onde-4	11	8	1	3	déphasage des pics

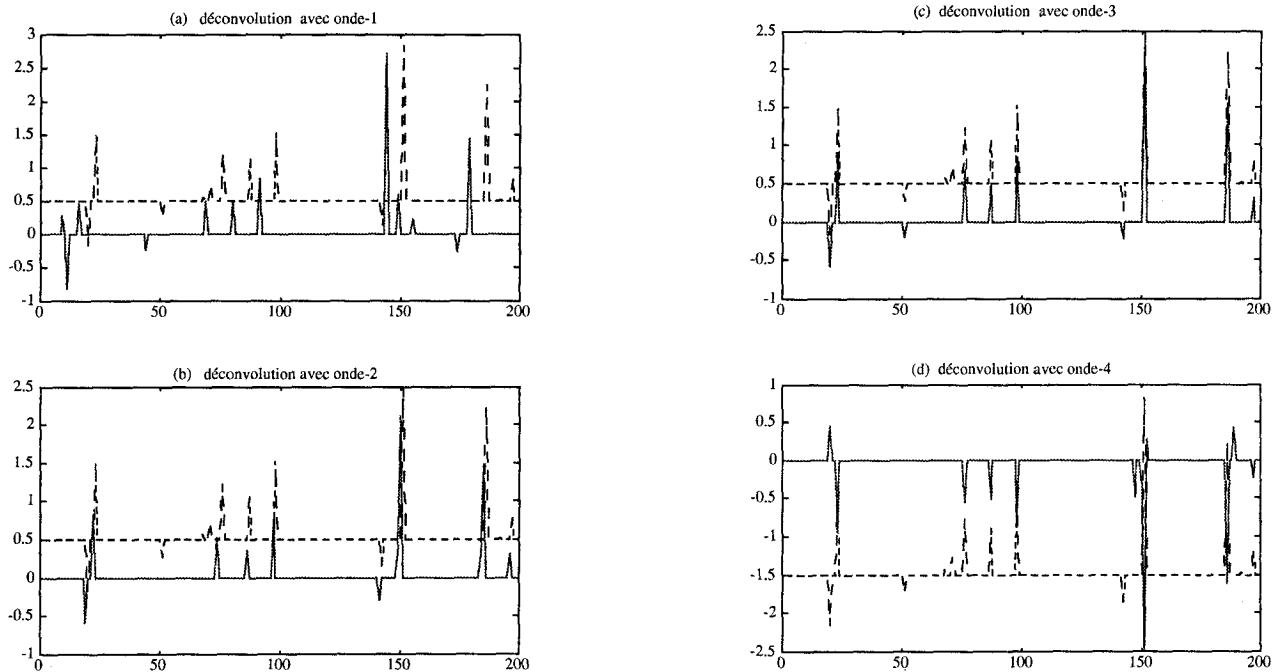


Figure 3.3. – Résultats de la déconvolution du signal de la figure 3.2 en utilisant les ondelettes onde-1 (a), onde-2 (b), onde-3 (c), onde-4 (d).

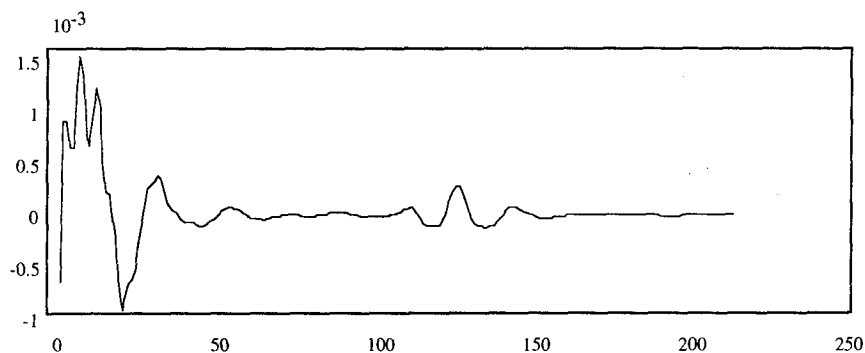


Figure 4 (a). – Contrôle d'un défaut situé à 2.75 mm de surface.

4.3. DÉCONVOLUTION DE SIGNAUX RÉELS

Les signaux ont été recueilli par échographie d'un bloc d'acier (Z2CN 18-10) dans lequel on a percé des trous parallèlement à la surface contrôlée. Ces trous d'un diamètre de $0.5 \pm 0.05\text{mm}$, sont percés à des distances de 5.25 à 1.75mm de la surface. Les enregistrements ont été effectués sur 8 bits à 100MHz et il a été nécessaire de saturer l'écho d'interface pour que la quantification laisse apparaître les échos de défauts. Cette partie saturée a été supprimée avant le traitement.

Nous présentons ici les résultats concernant la recherche des défauts situés à 2.75mm (fig 4a) et à 4.25mm (fig 5a). Dans ce dernier cas, l'écho est visible à l'oeil nu, et le début peut être apprécié un peu avant le point n°100 (fig 5b). La déconvolution fait apparaître un pic au point 95, ce qui semble donc raisonnable. En ce qui concerne le défaut situé à 2.75mm, un calcul rapide

permet d'attendre un écho situé dans le voisinage du point 40, et la déconvolution de la séquence fait apparaître un pic au point 38 (fig 4b).

5. Conclusions

Nous avons présenté ici un algorithme de déconvolution utilisant la minimisation d'une norme L^2 . Cette approche par l'analyse numérique se justifie par la nature du problème à résoudre, c'est-à-dire la restauration d'une propriété se présentant sous la forme d'une séquence peu dense d'impulsions. De nombreux essais ont permis de montrer que cet algorithme permettait une bonne localisation des pics, et une bonne restitution des amplitudes pourvu que l'ondelette utilisée corresponde bien à celle qui

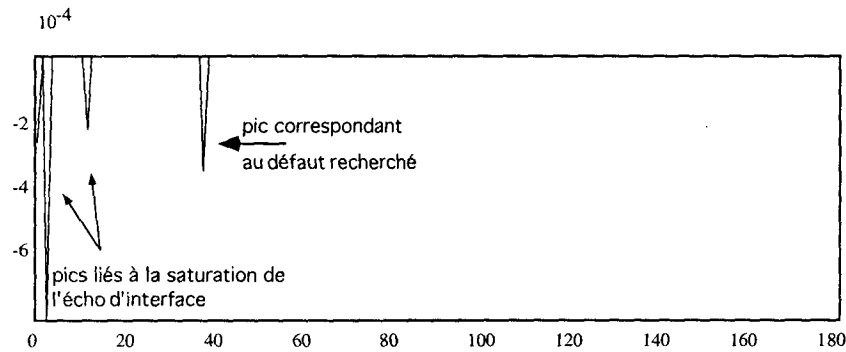


Figure 4 (b). – Résultat de la déconvolution.

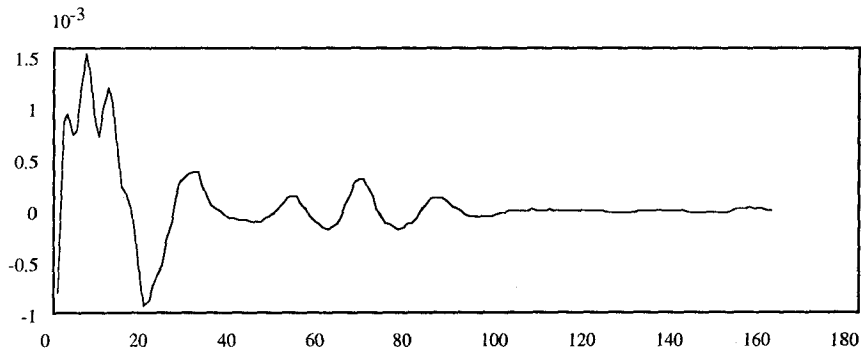


Figure 5 (a). – Contrôle d'un défaut situé à 4.25 mm de la surface.

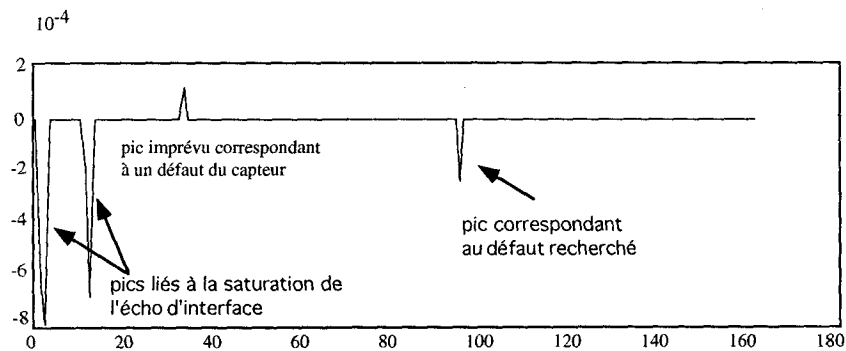


Figure 5 (b). – Résultat de la déconvolution.

intervient dans la convolution des mesures. Notre expérience nous permet d'affirmer que des problèmes pourront apparaître lorsque la densité des pics augmente. En particulier, lorsque deux pics d'amplitude voisine se succèdent, l'algorithme a tendance à déterminer un événement entre les deux. Les performances seront également moins bonnes lorsque le contenu fréquentiel de l'ondelette s'appauvrit et que celle-ci devient très oscillante; mais ce problème est commun à tous les algorithmes de déconvolution. Cet algorithme présente à notre avis quelques avantages décisifs : Tout d'abord, il permet d'effectuer la déconvolution dans le cadre L^2 , ce qui permet une interprétation physique simple des quantités calculées (en terme d'intercorrélation ou d'énergie par exemple), et fournit un résultat meilleur que l'inversion L^2 classique sans utiliser de terme de pénalisation. Il est par ailleurs extrêmement

simple à mettre en oeuvre. Enfin, non seulement il est d'une complexité réduite, et donc rapide, mais il est possible de limiter la recherche au nombre de pics souhaités, ce qui lui confère une grande souplesse d'utilisation.

BIBLIOGRAPHIE

- [1] J. B. Bednar, T. L. Watt, and R. Yarlagadda. L^1 deconvolution and its application to seismic processing for P^p deconvolution. *IEEE trans. on ASSP*, 34, 1986.
- [2] M. Bulo, S. Mottelet, and P. Simard. Comparaison d'algorithmes de déconvolution pour l'analyse et la classification des signaux d'échographie ultrasonore. In *Actes du 13^{ème} colloque sur le traitement du signal et des images, Juan Les pins*, 1991.

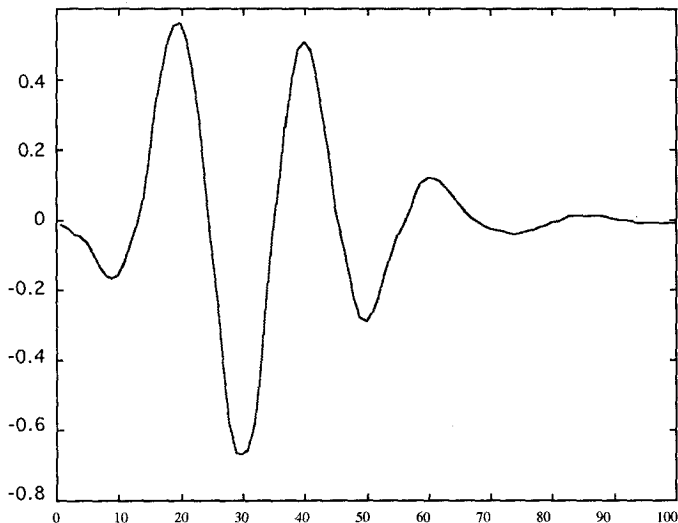


Figure 6. – Ondelette de référence obtenue par échographie d'un matériau épais sans défauts.

[3]S. Mottelet. Etude de deux méthodes de déconvolution : Normes l^p et entropie minimum, mémoire de DEA contrôle des systèmes. *Université de Technologie de Compiègne*, 1991.

[4]H. L. Taylor, S. C. Bank, and J. F. Mc Coy. Deconvolution with the l^1 norm. *Geophysics*, 44, 1979.

[5]R. Yarlagadda, J. B. Bednar, and T. L. Watt. Fast algorithms for l^p deconvolution. *IEEE trans. on ASSP*, 33, 1985.

Manuscrit reçu le 28 septembre 1993.

LES AUTEURS

Patrice SIMARD

Ingénieur en Génie Mécanique, option Acoustique et Vibrations Industrielles en 1983. Thèse de 3e cycle (1983 :84). Actuellement Maître de Conférence au département de Génie Informatique de l'Université de Technologie de Compiègne, et chercheur de l'URA CNRS n° 817 "Heuristique et diagnostic des systèmes complexes".

Stéphane MOTTELET

Né le 4 janvier 1967 à Compiègne. Ingénieur en Génie Informatique (filiale Calcul Scientifique) de l'Université de technologie de Compiègne en 1989. DEA de Contrôle des Systèmes à l'UTC en 1990. Thèse réalisée à la Division Mathématiques Appliquées de l'UTC. Titre de la thèse "Quelques aspects théoriques et numériques du contrôle d'un bassin de carènes" (soutenance en décembre 1994). Le problème étudié est celui du contrôle actif de la houle dans un bassin. Domaine de recherche actuel : contrôle des systèmes gouvernés par des équations aux dérivées partielles. Actuellement ATER au sein de l'URA CNRS 817 "Heudiasyc" au département Génie informatique de l'UTC.