BLIND DECONVOLUTION IN A NOISY AND BAND-LIMITED CONTEXT

Pham Dinh Tuan

Laboratoire Jean Kuntzman, CNRS/INPG, BP 53, 38041 Grenoble Cedex, France phone +(33) 4 76 51 44 23, fax: +(33) 4 76 63 12 63, email: Dinh-Tuan.Pham@imag.fr web: www-ljk.imag.fr/membres/Dinh-Tuan.Pham/

ABSTRACT

New modified mutual information criteria have been proposed for the deconvolution problem in the presence of noise and with non invertible convolution filter. Closed form formula for the derivative of the criteria are obtained. Simulations results are provided showing the good performance of the method.

1. INTRODUCTION

This work is motivated by seismic applications, in which a recorded seismic trace is often modelled as a convolution of a wavelet w(t) with the reflectivity series r(t) plus a superposed noise n(t):

$$x(t) = \sum_{u = -\infty}^{\infty} w(u)r(t - u) + n(t) = (w \star r)(t) + n(t)$$
 (1)

where \star denotes the convolution. It is generally assumed that the reflectivity is a white super-Gaussian process, the noise is a white Gaussian process, and the wavelet is a bandlimited filter. Seismic deconvolution consists in recovering the reflectivity from a given seismic trace. This is often done through a deconvolution filter g:

$$y(t) = \sum_{u=-\infty}^{\infty} g(u)x(t-u) = (g \star x)(t)$$
 (2)

As the wavelet and the reflectivity distribution are unknown, we have the classical *blind* deconvolution problem, *but with the added difficulty that the observation is contaminated with noise and the convolution filter (i.e. the wavelet) is band-limited, hence not invertible.*

Works on blind deconvolution has assumed (implicitly) that the convolution filter *w* is invertible, that is there exists a filter *g* such that $(g \star w)(u) = 1$ if u = 0 and 0 otherwise. Then, neglecting the noise, the reflectivity r(t) can be recovered from the observation by convolving it with the filter *g* as in (2). In the blind context *w* is unknown so *g* must be estimated, usually by maximizing a contrast function, based on the two basic properties of the source: its temporal independence and its non Gaussianity.

In seismic applications however, the noise level can be high and more importantly the wavelet is often quite band-limited: its Fourier series $W(\omega) = \sum_{u=-\infty}^{\infty} w(u)e^{-iu\omega}$ is (nearly) zero for ω outside a small (low) frequency band. Thus this filter is not invertible since the inverse filter gshould have Fourier series $G(\omega) = \sum_{u=-\infty}^{\infty} g(u)e^{-iu\omega}$ equal to $1/W(\omega)$. To our knowledge, there has not been work dealing specifically with this situation, apart from works on the Bayesian approach [1,2], which require a parametric model with prior on the parameters and hence are not really blind. The usual practice is to continue using deconvolution methods designed for the noiseless and invertible convolution filter case, hoping that the noise effect can be neglected or contained. In [3], some theoretical and experimental studies on the noise effect on some deconvolution criteria have been presented. It turns out that the mutual information contrast, which has been shown to yield optimal performance in the noiseless and invertible convolution filter case [4], is much more sensitive to noise than the (non optimal) kurtosis

The purpose of this work is to propose a modification of the mutual information criterion, so that it can handle effectively the case of noisy data and band-limited filter.

Section 2 analyzes the effect of noise and the bandlimited property of the wavelet on the mutual information criterion. Such analysis has appeared in [3], but as it provides the rational for our modified criterion, it is summarized here. The new criterion is developed in section 3. The empirical counterparts and its gradient are derived in next section. Finally section 5 provides some simulation results in a realistic setting, showing the good performance of the method.

2. NOISE INFLUENCE ON THE MUTUAL INFORMATION CRITERION

In the presence of noise, it is not possible to recover exactly the reflectivity through any deconvolution filter: the output of (2) always contains some noise. Further, if one tries to match $g \star (w \star r)$ to r, g must have its the Fourier transform very large at the frequencies outside the pass-band of w, which would blow up the noise. Thus a compromise between signal recovery and noise reduction must be made. In the *non-blind* context, where the wavelet and the noise and reflectivity variances are known, the Wiener filter realizes the best compromise at it minimizes the mean square error between y(t) and r(t). It is given in the frequency domain by

$$G_{\text{Wiener}}(\omega) = \sigma^2 W^*(\omega) / [\sigma^2 |W(\omega)|^2 + \sigma_n^2] \qquad (3)$$

where σ^2 is the variance of the reflectivity series, σ_n^2 that of the noise and * denotes complex conjugate. This formula shows that G_{Wiener} is nearly the inverse of W inside the pass-

Thank to British Gas, BP, Chevron, the Department of Trade and Industry and Shell for financial support of the project Blind Identification of Seismic Signals (BLISS), and to members of this project (Ch. Jutten, A. Larue, J. Mars, M. Van der Bann) for stimulating discussions leading to this work.

band of w (where $W\sigma^2/\sigma_n^2$ is large) and nearly zero outside this band (where $W\sigma^2/\sigma_n^2$ is nearly zero).

The mutual information approach to blind deconvolution consists in adjusting the deconvolution filter g such that the output $\{y(t)\}$ defined in (2) is the least temporally dependent. Here this dependence is measured by the mutual information rate, which can be expressed as [4]

$$H(y) - \int_0^{2\pi} \log |G(\omega)| \frac{d\omega}{2\pi} + \text{constant}$$

where H(y) is the Shannon entropy of y(t), which does not depend on *t* by stationarity. Denoting by $H^{-}(y) = \frac{1}{2}\log[2\pi e \operatorname{var}(y)] - H(y)$, the negentropy of *y*, the mutual information criterion can be written as

$$-H^{-}(y) + \frac{1}{2} \int_{0}^{2\pi} \log \frac{\operatorname{var}(y)}{f_{y}(\omega)} \frac{d\omega}{2\pi} + \operatorname{constant}$$
(4)

where f_y is the spectral density of the y(t) process, which is related to that of the x(t) process by $f_y(\omega) = |G(\omega)|^2 f_x(\omega)$.

The negentropy is a non-Gaussianity measure. It has been argued in [3] that such measure would have good behavior under noise contamination. The noise sensitivity of the criterion comes from the second term in (4). Indeed, this term is a measure of flatness of f_y , since from the equality $\operatorname{var}(y) = \int_0^{2\pi} f_y(\omega) d\omega/(2\pi)$ it can be written as

$$\frac{1}{2} \left[\log \int_0^{2\pi} f_y(\omega) \frac{d\omega}{2\pi} - \int_0^{2\pi} \log f_y(\omega) \frac{d\omega}{2\pi} \right], \qquad (5)$$

which, by Jensen's inequality, is non-negative and can be zero if and only if f_y is constant. This term will be referred to as the second order whiteness term since it tends to force the output of the deconvolution filter to be second order white. If we minimize only this term, we would get a deconvolution filter with gain:

$$|G(\omega)| = \frac{\text{Constant}}{f_x^{1/2}(\omega)} = \frac{\text{Constant}}{[\sigma^2 |W(\omega)|^2 + \sigma_n^2]^{1/2}}.$$

The constant factor comes from the scale ambiguity of the deconvolution filter. If we normalize this filter so that its output has unit variance then this constant becomes 1. Thus the above gain nearly equal $1/\sigma_n$ outside the pass-band of the wavelet (where $|W| \approx 0$) while the gain of the optimal filter is nearly zero there. Therefore the second order whiteness term in (4) is strongly affected by noise. Due to this term, the noise attenuation would be inadequate outside the pass-band of the wavelet. This finding has been supported by simulation experiments reported in [3].

3. THE MODIFIED MUTUAL INFORMATION CRITERIA

In the noiseless and invertible convolution filter case the mutual information criterion is known to yield the best performance [4]. The above analysis shows that its bad behavior in the noisy and band-limited case is due to the second term in (4). In fact, using only the first term or some other non Gaussianity measure such as the kurtosis, can yield better performance in this case [3]. But the use of a non Gaussianity criterion ignores information provided by second order statistics, which can be helpful especially for estimating the gain of the deconvolution filter, as can be seen in [4]. Therefore, our idea is to only modify the second term in (4) to correct its bad behavior in the presence of noise.

3.1 The modified criterion

We have seen that the second term in (4) is a measure of flatness of f_y . But the deconvolution output in the noisy non invertible case, *should not be flat*, as can be seen from that of the Wiener filter. Since the theoretical spectral density of the observed process $\{x(t)\}$ is $\sigma^2 |W(\omega)|^2 + \sigma_n^2$, it is the ratio

$$\frac{f_x}{|W\sigma/\sigma_n|^2 + 1} = \frac{f_x}{|W'|^2 + 1}$$
(6)

where $W' = W\sigma/\sigma_n$, which should be flat (as it is theoretically equal to the constant σ_n^2). Thus, we shall replace the second term in (4) by (5) with $f_y(\omega)$ replaced by (6). Further, the deconvolution filter which produces the $\{y(t)\}$ process involved in the negentropy term $H^-(y)$ in (4) is now taken as the Wiener filter associated with *w*. Since H^- is scale invariant, $H^-(y) = H^-(y')$ where $y'(t) = y(t)\sigma_n/\sigma$, which can be determined from $\{x(t)\}$ through the rescaled Wiener filter with frequency response

$$G'(\omega) = W'^{*}(\omega) / [|W'(\omega)|^{2} + 1]$$
(7)

Therefore, we are led to the new deconvolution criterion

$$C(W') = -H^{-}(y') + \frac{1}{2} \left[\log \int_{0}^{2\pi} \frac{f_{x}(\omega)}{|W'(\omega)|^{2} + 1} \frac{d\omega}{2\pi} - \int_{0}^{2\pi} \log \frac{f_{x}(\omega)}{|W'(\omega)|^{2} + 1} \frac{d\omega}{2\pi} \right]$$
(8)

where $y' = g' \star x$ with g' having Fourier transform G' given by (7). Minimizing this criterion would yield the Fourier transform W' of the rescaled wavelet $w' = w\sigma/\sigma_n$. Note that $|W'(\omega)|^2$ represents the signal-to-noise ratio (SNR) at the (angular) frequency ω and y' is the recovered reflectivity rescaled to have the same variance as the noise.

Once W' has been obtained, the noise variance σ_n^2 can be obtained by noting that (6) is theoretically equal to σ_n^2 . Thus we take

$$\sigma_n^2 = \int_0^{2\pi} \frac{f_x(\omega)}{|W'(\omega)|^2 + 1} \frac{d\omega}{2\pi}.$$
(9)

Note that the above criterion does not allow to determine W and σ separately but only the product $W\sigma = W'\sigma_n$. This is easy to understand: a multiplication of the reflectivity by a constant factor can always be offset by a division of the wavelet by the same factor so that the output of the convolution remains the same.

3.2 Exploiting the band-limited property

The second term of (8) does not actually provide enough information to estimate even |W'|. Indeed, minimizing this term alone would yields $|W'|^2 = cf_x - 1$ for *arbitrary* constant c > 0 and $\sigma_n^2 = 1/c$. Of course, some information on |W'| may be extracted from the first term of (8), but this term contains mainly information on the phase of the wavelet, not its amplitude. Thus criterion (8) may not provide good estimate of σ_n^2 .

To overcome the above problem, our idea is to exploit the fact that the wavelet energy is mainly concentrated on a small low frequency band. We propose two approaches.

3.2.1 Special parameterization

By forcing W' to vanish at (angular) frequency π , one avoids the above ambiguity problem since the above constant cshould then (theoretically) satisfy $cf_x(\pi) = 1$. A possible such parameterization is

$$W'(\omega) = (e^{i\omega} + 2 + e^{-i\omega})\tilde{W}_{\theta}(\omega) = 2[1 + \cos(\omega)]\tilde{W}_{\theta}(\omega)$$

where $\tilde{W}_{\theta}(\omega)$ is some smooth function depending on a vector parameter θ . It can be for example be the Fourier transform of a finite impulse response filter with coefficients being the components of θ . Then W' vanishes at π and due to its continuity, it would be nearly zero in the vicinity of π . To make W' vanish on a wider range of frequencies, one may consider the parameterization

$$W'(\omega) = (e^{i\omega} + 2 + e^{-i\omega})[e^{i\omega} - 2\cos(\omega_0) + e^{-i\omega}]\tilde{W}_{\theta}(\omega)$$

for some given ω_0 , $\tilde{W}_{\theta}(\omega)$ being as before. The factor $e^{i\omega} - 2\cos(\omega_0) + e^{-i\omega}$ vanishes for $\omega = \pm \omega_0$. Hence W' would be nearly zero on some interval containing $[\omega_0, \pi]$ if ω_0 is chosen not too far from π .

3.2.2 Pre-estimation of the noise variance: another criterion

If the pass-band of the wavelet is more or less known, one can estimate σ_n^2 by taking the average of f_x outside this band. Then one may consider σ_n^2 as known and equal to this estimate. To force the ratio (6) to be close to the "known" σ_n^2 , we add the term

$$\frac{1}{2}\int_0^{2\pi} \left[\frac{f_x(\omega)/\sigma_n^2}{|W'(\omega)|^2+1} - 1 - \log\frac{f_x(\omega)/\sigma_n^2}{|W'(\omega)|^2+1}\right] \frac{d\omega}{2\pi}$$

to the criterion (8). The above term is non negative and can be zero if and only if $(f_x/\sigma_n^2)/(|W'|^2+1) = 1$, since for a > 0, $a-1-\log a \ge 0$ with equality if and only if a = 1. Further, as $a-1-\log a \approx (a-1)^2/2$ for a near 1, this term would have the similar effect as the 1/4 the L^2 squared distance between the function $(f_x/\sigma_n^2)/(|W'|^2+1)$ and the constant 1. Thus, noting that $\sigma_n W' = \sigma W$ and $H^-(y') = H^-(\sigma_n y') =$ $H^-(y/\sigma)$, we are led to the criterion

$$-H^{-}\left(\frac{y}{\sigma}\right) + \frac{1}{2} \left[\int_{0}^{2\pi} \frac{f_{x}(\omega)}{|\sigma W(\omega)|^{2} + \sigma_{n}^{2}} \frac{d\omega}{2\pi} - 1 - \int_{0}^{2\pi} \log \frac{f_{x}(\omega)}{|\sigma W(\omega)|^{2} + \sigma_{n}^{2}} \frac{d\omega}{2\pi} \right]$$

As we have noted, only the product σW can be estimated and there is a scale ambiguity in the estimated reflectivity. Thus we may assume that $\sigma = 1$, and rewrite the above criterion as, dropping the constant -1/2,

$$C^{\dagger}(W) = -H^{-}(y) + \frac{1}{2} \left[\int_{0}^{2\pi} \frac{f_{x}(\omega)}{|W(\omega)|^{2} + \sigma_{n}^{2}} \frac{d\omega}{2\pi} - \int_{0}^{2\pi} \log \frac{f_{x}(\omega)}{|W(\omega)|^{2} + \sigma_{n}^{2}} \frac{d\omega}{2\pi} \right].$$
(10)

4. THE EMPIRICAL CRITERIA AND THEIR GRADIENT

4.1 The empirical criteria

In practice, the criteria (8) and (10) must be replaced by their empirical versions, in which the terms $H^-(y)$ (or $H^-(y')$) and f_x are replaced by their estimates. For the spectral density f_x , a natural estimate is the periodogram $\hat{f}_x(\omega) = n^{-1} |\sum_{t=0}^{n-1} x(t)e^{-i\omega t}|^2$, x_0, \ldots, x_{n-1} being the observations. This is a raw unsmoothed estimate, but the integrations involved in (8) and (10) provide an implicit smoothing, and it has a low bias (of the order 1/n). For numerical calculation, these integrations are replaced by Riemann sums based on the points $0, 2\pi/n, \ldots, 2\pi(n-1)/n$. Thus, let $\hat{H}^$ be a negentropy estimator, the empirical version of (8) is, up to a constant,

$$\hat{C}(W') = -\hat{H}^{-}(y') + \frac{1}{2} \left\{ \log \frac{1}{n} \sum_{k=0}^{n-1} \frac{\hat{f}_{x}(2\pi k/n)}{|W'(2\pi k/n)|^{2} + 1} + \frac{1}{n} \sum_{k=0}^{n-1} \log \left[\left| W'\left(\frac{2\pi k}{n}\right) \right|^{2} + 1 \right] \right\}$$
(11)

and that of (10) is, up to a constant,

$$\hat{C}^{\dagger}(W) = -\hat{H}^{-}(y) + \frac{1}{2} \left\{ \frac{1}{n} \sum_{k=0}^{n-1} \frac{\hat{f}_{x}(2\pi k/n)}{|W(2\pi k/n)|^{2} + \sigma_{n}^{2}} + \frac{1}{n} \sum_{k=0}^{n-1} \log\left[\left| W\left(\frac{2\pi k}{n}\right) \right|^{2} + \sigma_{n}^{2} \right] \right\}.$$
(12)

For the criterion (8), the noise variance will be estimated by $v_n(\hat{W}')$, given by

$$v_n(W') = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\hat{f}_x(2\pi k/n)}{|W'(2\pi k/n)|^2 + 1}.$$
 (13)

where \hat{W}' is the estimator of W' which minimizes this criterion. Compare with equation (9).

To estimate $H^{-}(y)$ one needs to be able to compute y', but the formula (2) involved all values of x(t) and those of index outside $\{0, n - 1\}$ are not observed. To overcome this difficulty, a simple approach is to extend the data $\{x(0, ..., x(n-1))\}$ periodically and thus compute y(t) as

$$y(t) = \sum_{u = -\infty}^{\infty} g(u)x(t - u \bmod n)$$
(14)

This amounts to replacing ordinary convolution by circular convolution. Since *G* is smooth, its Fourier coefficients should decay rapidly at infinity, therefore y(t) computed as above should be not much different from the one computed by (2), for $t \in \{0, ..., n-1\}$ and far from 0 and *n*.

The entropy estimator $\hat{H}(y)$ of *y* can now be constructed from $y(0), \ldots, y(n-1)$, using the method in [4] (for ex.). Finally, the negentropy estimator of *y* is estimated by

$$\hat{H}^{-}(y) = \frac{1}{2}\log[2\pi e \widehat{\operatorname{var}}(y)] - \hat{H}(y)$$

where

$$\widehat{\operatorname{var}}(y) = \frac{1}{n} \sum_{t=0}^{n-1} y^2(t) = \frac{1}{n} \sum_{k=0}^{n-1} \left| G\left(\frac{2\pi k}{n}\right) \right|^2 \widehat{f}_x\left(\frac{2\pi k}{n}\right).$$

The same method applies to the estimation of $H^{-}(y')$, replacing y, G by y', G'.

4.2 Minimization of the criteria

We assume that W' has been parameterized by a vector parameter θ . We choose to use the BFGS (Broyden-Fletcher-Goldfard-Shano) method [5] to minimize the criteria. This is a quasi Newton type algorithm, but it needs only the gradient of the criterion (and the criterion itself) since it computes an approximation to the Hessian during the iteration, using gradient and criterion values collected from previous steps. The advantage of BFGS over simple gradient descent is that the former has quadratic convergence near the solution.

In the sequel, we provide analytic formulas for the gradient of each of the criteria. Following [4], we introduce the score estimator of *y*, defined as the partial derivative of the entropy estimator:

$$\hat{\psi}_{y}[y(t)] = n \frac{\partial \hat{H}(y)}{\partial y(t)}, \qquad t = 0, \dots, n-1.$$

Then

$$\frac{\partial H(y)}{\partial \theta_{\mu}} = \frac{1}{n} \sum_{t=0}^{n-1} \frac{\partial y(t)}{\partial \theta_{\mu}} \hat{\psi}_{y}[y(t)],$$

hence by (14):

$$\frac{\partial H(y)}{\partial \theta_{\mu}} = \sum_{u=-\infty}^{\infty} \frac{\partial g(u)}{\partial \theta_{\mu}} \left\{ \frac{1}{n} \sum_{t=0}^{n-1} x(t-u \bmod n) \hat{\psi}_{y}[y(t)] \right\}.$$

The expression inside the above curly bracket {} is the sample circular cross covariance, at lag -u, between the process $\{x(t)\}$ and $\{\hat{\psi}_y[y(t)]\}$, which we denote by $\hat{c}_{x\psi_y}(-u)$. Define the cross periodogram between these processes as

$$\hat{f}_{x\psi_{y}}\left(\frac{2\pi k}{n}\right) = \sum_{u=0}^{n-1} \hat{c}_{x\psi_{y}}(u)e^{-i2\pi ku/n} \\ = \frac{1}{n} \left[\sum_{t=0}^{n-1} x(t)e^{-i2\pi kt/n}\right] \left[\sum_{t=0}^{n-1} \hat{\psi}_{y}[y(t)]e^{i2\pi kt/n}\right].$$

Then it can be shown that

$$\frac{\partial H(y)}{\partial \theta_{\mu}} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\partial G(2\pi k/n)}{\partial \theta_{\mu}} \hat{f}_{x\psi_{y}} \left(\frac{2\pi k}{n}\right).$$

One can then deduce the following result:

$$\frac{\partial \hat{H}^{-}(\mathbf{y})}{\partial \theta_{\mu}} = -\frac{1}{n} \sum_{k=0}^{n-1} \frac{\partial G(2\pi k/n)}{\partial \theta_{\mu}} \hat{f}_{\mathbf{x} \phi_{\mathbf{y}}} \Big(\frac{2\pi k}{n} \Big)$$

where

$$\hat{f}_{x\phi_y}\left(\frac{2\pi k}{n}\right) = \hat{f}_{x\psi_y}\left(\frac{2\pi k}{n}\right) - \frac{f_x(2\pi k/n)G'^*(2\pi k/n)}{\widehat{\operatorname{var}}(y)}$$

is the cross-periodogram between the processes $\{x(t)\}$ and $\{\hat{\varphi}_{y}[y(t)] = \hat{\psi}_{y}[y(t)] - y(t)/\widehat{\operatorname{var}}(y)\}$. For $G = W^{*}/(|W|^{2} + \sigma_{n}^{2})$, one can show

$$\frac{\partial \hat{H}^{-}(y)}{\partial \theta_{\mu}} = -\frac{1}{n} \sum_{k=0}^{n-1} \frac{\partial W^{*}(2\pi k/n)}{\partial \theta_{\mu}} \left\{ \frac{\hat{f}_{x\phi_{y}(y)}(2\pi k/n)}{|W(2\pi k/n)|^{2} + \sigma_{n}^{2}} - 2G^{*}\left(\frac{2\pi k}{n}\right) \Re \left[G\left(\frac{2\pi k}{n}\right)\hat{f}_{x\phi_{y}}\left(\frac{2\pi k}{n}\right)\right] \right\}$$

where \Re denotes the real part.

By the same calculation, one gets a similar formula for $\partial \hat{H}(y')/\partial \theta_m u$, with y, G, replaced by i', G', σ_n^2 replaced by 1 and ψ_y replaced by $\psi_{y'}$ defined similarly as ψ_y but with y' in place of y.

Finally, it can be shown that the gradient of the criterion (12) is

$$\begin{aligned} \frac{\partial \hat{C}^{\dagger}(W)}{\partial \theta_{\mu}} &= \frac{1}{n} \sum_{k=0}^{n-1} \frac{\partial W^{*}(2\pi k/n)}{\partial \theta_{\mu}} \left\{ \frac{f_{x\phi_{y}(y)}(2\pi k/n)}{|W(2\pi k/n)|^{2} + \sigma_{n}^{2}} \right. \\ &\left. - 2G^{*}\left(\frac{2\pi k}{n}\right) \Re \left[G\left(\frac{2\pi k}{n}\right) f_{x\phi_{y}(y)}\left(\frac{2\pi k}{n}\right) \right] \right. \\ &\left. + G^{*}\left(\frac{2\pi k}{n}\right) \left[1 - \frac{f_{x}(2\pi k/n)}{|W(2\pi k/n)|^{2} + \sigma_{n}^{2}} \right] \right\} \end{aligned}$$

and that of the criterion (11) is

$$\begin{aligned} \frac{\partial \hat{C}(W')}{\partial \theta_{\mu}} &= \frac{1}{n} \sum_{k=0}^{n-1} \frac{\partial W'^{*}(2\pi k/n)}{\partial \theta_{\mu}} \bigg\{ \frac{f_{x \phi_{y'}}(2\pi k/n)}{|W'(2\pi k/n)|^{2} + 1} \\ &- 2G'^{*} \bigg(\frac{2\pi k}{n} \bigg) \Re \bigg[G' \bigg(\frac{2\pi k}{n} \bigg) f_{x \phi_{y'}} \bigg(\frac{2\pi k}{n} \bigg) \bigg] \\ &+ G'^{*} \bigg(\frac{2\pi k}{n} \bigg) \bigg[1 - \frac{f_{x}(2\pi k/n)/\nu_{n}(W')}{|W'(2\pi k/n)|^{2} + 1} \bigg] \bigg\}. \end{aligned}$$

5. SIMULATION

We have simulated a seismic trace of length 1024. The wavelet is taken as a mixture of two Ricker waves, the first one has a phase 58.31° and a central frequency 60Hz and the second one has zero phase and a central frequency 120Hz. The reflectivity is generated as $sign(z)|z(t)|^3/4$ where z(t) are independent standard normal random variables, and the noise is generated as independent normal random variables of zero mean and variance $0.4^2 = 0.16$. This yields a signal-to-noise ratio of 17.5db.



Figure 1: Simulation results with special parameterization

In a first experiment we consider the special parameterization: $W'(\omega) = (e^{i\omega} + 2 + e^{-i\omega})\tilde{W}_{\theta}(\omega)$ with \tilde{W}_{θ} having 19 consecutive nonzero Fourier coefficients (which are the components of θ), so that the estimated wavelet is a 21-taps filter. The results are plotted in Figure 1. The wavelet is well estimated, as can be seen from the top two graphs, and also from the correlation between the estimated and true wavelets (after alignment) which is 0.9789. However, the estimated wavelet has more power in the high frequency range. As a result, the gain of the estimated deconvolution filter is much higher than the optimal Wiener gain in this range, since the gain is proportional to $|W'(\omega)|/[|W'(\omega)|^2 + 1]$ with $|W'(\omega)|^2$ the signal-to-noise ratio (SNR) at ω . An over-estimation of $|W'|^2$, hence of this gain would lead to an insufficient noise reduction. The lower right part of Figure 1 shows that the SNR+1 = $|W'|^2 + 1$ can be over-estimated up to 50%. This prevents the filter from effectively reducing the noise.

The over-estimation of $|W'|^2$ can be explained. One can write its estimator $|\hat{W}'|^2$ as $|W'|^2 + (\hat{W} - W')W'^* + W'(\hat{W}' - W')^* + |\hat{W}' - W'|^2$, hence if the estimator \hat{W} is unbiased, $|\hat{W}|^2$ would be biased upward by $E(|\hat{W}' - W'|^2)$. This problem is most serious outside the pass-band of the wavelet as the true value of $|W'|^2$ is nearly zero there. Note that the power spectral density (PSD) of the observation equals (noise variance)*($|W'|^2 + 1$). In this experiment, the noise variance is corrected estimated, 0.1631 versus the true value 0.16, hence the PSD is also over-estimated (as can be seen on the lower right part of figure 1).

The correlation after alignment between the recovered and true reflectivities is found to be 0.6868 (Figure 3). This is rather low, but one should note that even with the optimal Wiener filter, the correlation is only 0.7304. This low value can be explained by the presence of noise and the fact that the wavelet is band-limited so that the high frequency components of the reflectivity can never be recovered.



Figure 2: Simulation results with pre-estimation of noise variance

In a second experiment we pre-estimate the noise variance. By taking the average PSD of the observation over the upper half frequency range, we get an estimate of 0.1817 which is higher than the true value. This over-estimation may be explained by the fact that there are still some signal power in this range. But it could be a good thing. Indeed, as the PSD of the seismogram equals (noise variance)*(SNR+1), an over-estimation of the noise variance would at least reduce



Figure 3: From top to bottom: true reflectivity, synthetic seismogram (reflectivity convolved with the wavelet + noise), Wiener deconvolution result, and results from test 1 and 2.

the over-estimation of SNR+1, in order that the PSD is not overly over-estimated. In this experiment we set the noise variance to 0.1817 and minimize the criterion (12) to estimate the wavelet, which is parameterized simply as a 21-taps filter. It can be seen in Figure 2 that the PSD of the wavelet is under-estimated at low frequency but is better estimated at high frequency. Overall there is a correlation of 0.9790 between the estimated and the true wavelet, which is almost the same as in previous experiment (Figure 1). Since the noise variance is over-estimated, the SNR+1 is only slightly over-estimated in the high frequency range (but it is underestimated in the medium frequency range). As a result, the deconvolution filter is somewhat more effective in reducing the noise. The correlation between the recovered and true reflectivities is 0.6892, which is slightly better than in the previous experiment.

Figure 3 compares the three deconvolution outputs: the (non blind) Wiener filtering and the two test experiments. One can see that the peaky nature of the reflectivity are well recovered.

REFERENCES

- M. Lavielle, "Bayesian deconvolution of Bernoulli-Gaussian process," *Signal Processing*, vol. 33, pp. 67– 79, 1993.
- [2] M. Lavielle, "A stochastic algorithm for parametric and nonparametric estimation in the case of incomplete data," *Signal Processing*, vol. 42, pp. 3–17, 1995.
- [3] A. Larue and D. T. Pham, "Comparison of supergaussianity and whiteness assumptions for blind deconvolution in noisy context," in *Proc. EUSIPCO 2006 conference*, Florence, Italia, Sept. 4–9, 2006.
- [4] D. T. Pham, "Generalized Mutual Information Approach to Multichanel Blind Deconvolution," *Signal Processing*, vol. 87, 9,2045–2060, 2007.
- [5] Yu-Hong Dai, "Convergence properties of the BFGS algorithm," *Siam J. on Optimization*, vol. 31, no. 3, pp. 693–701, 2002