14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

# COMPARISON OF SUPERGAUSSIANITY AND WHITENESS ASSUMPTIONS FOR BLIND DECONVOLUTION IN NOISY CONTEXT

*Anthony Larue*[1], *and Dinh Tuan Pham*[2]

[1] Images and Signal Laboratory of Grenoble, INPG
BP 46, 38402 St Martin d'Hère Cedex, France
phone: + (33) 4 76 82 64 22,
fax: + (33) 4 76 82 63 84 ,
email: anthony.larue@lis.inpg.fr
web: www.lis.inpg.fr

[2] Laboratory of Modeling and Computation, 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-lmc.imag.fr.

**ABSTRACT**

We propose a frequency blind deconvolution algorithm based on mutual information rate as a measure of whiteness. In the case of seismic data, the algorithm of Wiggins [11] based on kurtosis, which is a supergaussianity criterion, is often used. We study the robustness in noisy context of these two algorithms, and compare them with Wiener filtering. We provide some theoretical explanations on the effect of the additive noise. The theoretical arguments are illustrated with a simulation of seismic signals. For such signal, the supergaussianity criterion appears more robust to noise contamination than the whiteness criterion.

## 1. INTRODUCTION

This paper is motivated by seismic applications, in which a recorded seismic trace is often modeled as a convolution of a wavelet $w(t)$ with the reflectivity series $r(t)$ plus added superposed noise $n(t)$, It is generally assumed that the reflectivity is a white supergaussian 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 or at least boosting its high frequencies content attenuated by the bandpass wavelet. This is often done through a deconvolution filter $g$. Figure 1 depicts the convolution-deconvolution system:

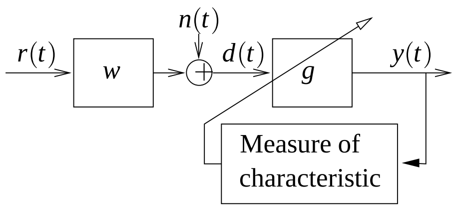


Figure 1: The convolution-deconvolution system

In seismic data processing, one rarely knows the wavelet, and the probability distribution of the reflectivity is also unknown. Our problem thus fits the description of the classical *blind* deconvolution problem, with the added difficulty that the observation is contaminated with noise and the convolution filter (i.e. the wavelet) is bandlimited. In the (more favorable) situation where there is no noise and the convolution filter is not bandlimited, there exists an unique inverse filter $g$ such that the output $y(t) = g \star d(t)$ equals $r(t)$, $\star$ denoting the convolution product. The idea is then to find an inverse filter $g$ to create an output $y(t) = g \star d(t)$ which looks like to the reflectivity series $r(t)$. Even though the above conditions are not met here, we shall stick, for simplicity, to the deconvolution procedure via a time invariant linear filter $g$. This filter will be adjusted by optimizing some criterion related to the output characteristic. The influence of the noise and the bandlimitness of the convolution filter on the adopted criterion, will be studied in a subsequent section. As the reflectivity series is assumed to satisfy some (broad) assumptions, the idea is to design criteria to force the output to tend toward the same assumptions. The most popular approach is to assume the whiteness of the reflectivity series, then, one adjusts $g$ to maximize a whiteness measure of the output. The simplest algorithm used the autocorrelation function (which is equal to a Dirac delta function) or the power spectrum (which is constant) of the reflectivity [9]. These methods employ second order statistics, which unfortunately do not contain phase information. Therefore, the phase of the wavelet remains unknown and has to be specified a priori. Wiggins [11] introduced in his Minimum Entropy Deconvolution (MED) algorithm, an entirely new concept based on the maximization of the kurtosis. The kurtosis is a fourth order statistic that measures the deviation from Gaussianity. If a white reflectivity series is convolved with a filter then the output will become more Gaussian [4]. A white reflectivity series can therefore be recovered by creating an inverse filter that render the output $y(t)$ as non Gaussian as possible. One way of doing this is to maximize the kurtosis of the created output signal. Since the kurtosis is a higher (than 2) order statistics, phase information is retained and no added a priori information is needed to recover the reflectivity series. Wiggins thus proposed the first blind deconvolution algorithm. The kurtosis is therefore traditionally interpreted as a whiteness criterion in deconvolution problems [1, 2]. It has however many disadvantages. Since it is based on the fourth order statistics, it is sensitive to the presence of outliers. Further, the use of such statistics is not optimal for the problem at hand.

Other criteria have been proposed to measure whiteness using all higher order statistics. One such criterion, recently introduced, is the mutual information rate [10, 6]. It is related to entropy and is a good general purpose measure of whiteness of a process [3]. It has been shown in the *noiseless* case that blind deconvolution based on minimization of the mutual information rate is optimal [8]. However, in the context blind deconvolution of seismic data, our numerical simulations show that the use of kurtosis as a supergaussianity constraint outperforms a general purpose whiteness constraint such as the mutual information rate. This can be explained by the presence of noise and the bandlimitness of the wavelet, as discussed in section 3. Section 2 describes the MED algorithm and a frequency blind deconvolution algorithm based on mutual information rate [6]. Simulation results are presented and discussed in section 4.

## 2. THEORETICAL ASSESSMENTS

### 2.1 MED algorithm

Wiggins [11] proposes to solve the deconvolution problem by maximization of the varimax norm. In one channel problem, it is equivalent to maximize the deconvolution output process kurtosis $K(y)$ estimated by:

$$\hat{K}(y) = \frac{T\sum_{t=1}^{T} y^4(t)}{[\sum_{t=1}^{T} y^2(t)]^2} - 3. \tag{1}$$

The simplest approach is to maximize (1) with respect of the filter coefficients $\mathbf{g} = [g_0, \ldots, g_N]^T$ where $N$ denotes the length of the filter. For this end, one equates to zero the derivative of $\hat{K}(y)$ with respect to the $g_k$, which yields a system of estimating equations: $\mathbf{Rg} = \mathbf{f}$, where $\mathbf{R}$ is a $N \times N$ autocorrelation matrix defined with the $N$ first delays of the autocorrelation function of the data $d(t)$, and the vector $\mathbf{f}$ is defined with the correlation between $y^3(t)$ and the data $d(t)$. The equation as it stands is highly nonlinear so that it cannot be solved directly. It can, however, be solved iteratively in a straightforward fashion. One starts with a value of $\mathbf{g}$, computing $\mathbf{f}$, solving the system $\mathbf{Rg} = \mathbf{f}$ for a new value of $g$, then recomputing $\mathbf{f}$ and so on ... Some extensions of this method were presented in [5], in particular a frequency approach and a modification of the norm in use.

### 2.2 Frequency domain blind deconvolution algorithm

The mutual information of a random vector $\mathbf{z} = (z_1, \ldots, z_n)$ of dimension $n$ is defined by:

$$I(\mathbf{z}) = \sum_{i=1}^{n} H(z_i) - H(z_1, z_2, \ldots, z_n) \tag{2}$$

where $H(z_i)$ denotes the Shannon marginal entropy of $z_i$: $H(z_i) = -\int_{\mathbb{R}} p_{z_i}(u) \log p_{z_i}(u) du$ and $H(z_1, z_2, \ldots, z_n)$ the Shannon joint entropy: $H(\mathbf{z}) = -\int_{\mathbb{R}^n} p_{\mathbf{z}}(\mathbf{u}) \log p_{\mathbf{z}}(\mathbf{u}) d\mathbf{u}$. The mutual information $I(\mathbf{z})$ has the nice property of being positive and vanishes if and only if the components of $\mathbf{z}$ are mutually independent. It is thus a measure of dependence of random variables. However stochastic processes are involved here so that we consider a related measures, called mutual information rate (MIR). The MIR of a stationary process $Z = \{Z_t\}$ is defined by:

$$I(Z) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{t=T} H(Z_t) - \mathscr{H}(Z) \tag{3}$$

where

$$\mathscr{H}(Z) = \lim_{T\to\infty} \frac{1}{T} H(Z_1, \ldots, Z_T) \tag{4}$$

which exists and is called the entropy rate of the (stationary) process $Z = \{Z_t\}$ (see [3]).

The mutual information rate $I(Z)$ is always positive and vanishes if and only if $Z$ is an iid process [3]. Thus it can be used as a deconvolution criterion. By stationarity, $H(Z_t)$ does not depend on $t$, hence we shall drop the index $t$ in the first term of (3). In practice to estimate $H(Z)$ one would use all samples $z(1), \ldots, z(T)$ as $T$ realizations of the random variable $Z_\tau$, for any $\tau$. To simplify the notation and to be homogeneous with the equation (1), in the following, we will write $H(z)$ for $H(Z)$.

The estimation of the entropy rate is however problematic, but fortunately, one can avoid it by noting that the entropy rate $\mathscr{H}(Y)$ of the deconvolution output $y(t) = (g \star d)(t)$ equals

$$\mathscr{H}(g \star D) = \mathscr{H}(D) + \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) e^{-jt\omega} \right| d\omega \tag{5}$$

(see [8]). Hence, the mutual information rate of the deconvolution output can be written as:

$$I(Y) = H(y) - \mathscr{H}(D) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) e^{-jt\omega} \right| d\omega \tag{6}$$

Then, since the entropy rate $\mathscr{H}(D)$ is independent of the inverse filter $g$, one can consider instead of (6), the simplified criterion [12, 10]:

$$\tilde{I}(Y) = H(y) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} g(t) e^{-jt\omega} \right| d\omega \tag{7}$$

which, like (6), is minimum when the process $\{y(t)\}$ is iid.

The above criterion has been used in [10], for a Wiener systems composed of a cascade of a direct filter $w$ and an invertible non linear distortion. The authors estimate the inverse filter $g$ in the time domain, by minimizing (7) with respect to the impulse response $g(t)$, which is done by a gradient technique. They use a finite number of coefficients for the estimation of $g(t)$, so it is equivalent to choosing a Moving Average (MA) model for $g$. Thus, the method is well adapted to the inversion of an autoregressive (AR) direct filter $w$. One can show that the algorithm is equivalent to a maximum likelihood (ML) method, replacing the source distribution, supposed known in the ML method, by the distribution of the deconvolution output estimated at each iteration. The application field of this method is limited by the parametric model: for example, in seismology, the direct filter impulse response can be a chirp or a Ricker wavelet which can not be modeled by a MA filter. Moreover, the method does not take into account the additive noise $n(t)$ of the model. With Gaussian additive noise, one can experimentally see that the method achieves the same performances as the second order method *i.e.* the Yule-Walker algorithm. To overcome these limitations, we propose to use a criterion in the frequency domain which avoids parametric approaches like the MA, AR or ARMA models, whose parameter number can be very large. Moreover, in the frequency domain, it is easy to add a regularization term for limiting noise amplification as it is usually done in Wiener filtering.

So, we propose to minimize (7) with respect to the discrete frequency response $G = [G_0, \ldots, G_{T-1}]$ of the filter $g$. Then, we have to estimate $T/2$ complex parameters (due to the hermitian symmetry of the real filter) using $T$ temporal samples of the output signal. Thus, without adding a smoothing constraint on the inverse filter frequency response, a trivial solution is found. Then, priors (regularization) are necessary on inverse filter to avoid trivial solutions. We add a smoothing constraint on $G_\nu$ *i.e.* which controls the difference $|G_\nu - G_{\nu+1}|$. Further, in seismic applications, the wavelet is often a bandpass filter, and there is a frequency band with poor information about reflectivity called the "null space". Thus, for providing a good output signal to noise ratio, the inverse filter must not amplify this frequency band. Therefore large values of the frequency response of the inverse filter are often prohibited, except if they give an important output independence improvement. To avoid noise amplification in the output, we limit the largest value of $|G_\nu|$ by a second regularization term.

Based on these remarks and approximating the integral in (7) with the rectangle method, , we propose the following frequency domain blind deconvolution (FBD) criterion:

$$J(\mathbf{G}) = H(y) - \frac{1}{T}\sum_{\nu=0}^{T-1}\log|G_\nu| + \lambda_1\sum_{\nu=0}^{T-1}|G_\nu - G_{\nu+1}|^2 + \lambda_2\sum_{\nu=0}^{T-1}|G_\nu|^p \qquad (8)$$

where $\lambda_1$ and $\lambda_2$ denote two hyperparameters. In the third right-side term, the sum is fully defined by using periodicity of $G_\nu$, *i.e.* $G_T = G_0$. The first regularization term, balanced by $\lambda_1$, constrains the frequency response of the inverse filter to be smooth enforcing $|G_\nu - G_{\nu+1}|$ to be Gaussian, *i.e.* with a maximum density for $|G_\nu - G_{\nu+1}| = 0$. Practically, we notice that this term also improves the stability and the performance of the minimization algorithm, because, it is a strong prior on the frequency response: the smoothness constraint reduces the freedom degree number. The last term penalizes (with the $\mathbb{L}^p$ norm) the largest values of the spectrum of $g$. For instance, with $p = 2$, it would enforce $|G_\nu|$ to have a Gaussian distribution. Thus, this term is equivalent to the noise factor usual in Wiener filtering: it allows a trade-off between the deconvolution quality and the noise amplification. We can interpret this criterion in a Maximum *a Posteriori* (MAP) framework: indeed, it is equivalent to take a Gaussian prior distribution of $G_\nu$ conditional to $G_{\nu-1}$ and a generalized Gaussian prior distribution (parameterized by $p$) for the marginal pdf of $|G_\nu|$.

To minimize the criterion (8) with respect to the complex-valued vector $\mathbf{G}$ according to a gradient iterative procedure, we compute the gradient [6] of the cost function (8) with respect to $G_\nu$:

$$\widehat{\nabla J(\mathbf{G})} = \frac{1}{2T^2}\Psi_y(\nu)D_\nu^* - \frac{1}{2T}\frac{1}{G_\nu^*} + \lambda_1(2G_\nu - G_{\nu+1} - G_{\nu-1}) + \lambda_2\frac{p}{2}\frac{|G_\nu|^p}{G_\nu^*} \qquad (9)$$

where $\Psi_y(\nu)$ is the Fourier transform of $\psi_y$ the score function of the random variable $y(\tau)$, which does not depend on $\tau$, defined as $\psi_y(u) = -\frac{d}{du}\log p_y(u)$ where $p_y$ is the density of $y(\tau)$.

The frequency blind deconvolution (FBD) algorithm is as follows.

1. *initialization of the inverse filter $G_\nu$ and of the deconvolution output $y(t)$;*
2. *estimation of the score function $\psi_y$;*
3. *computation of the gradient estimate (9);*
4. *updating of $G_\nu \leftarrow G_\nu - \mu\widehat{\nabla J(\mathbf{G})}$;*
5. *computation of the deconvolution output $y(t)$;*
6. *normalization step.*

We iterate the main loop (steps 2 to 6) until convergence. The normalization step 6 is required for taking into account scale indeterminacy in $G_\nu$. Here, we just normalize the inverse filter to obtain an unit power deconvolution output, but other normalization can be used. $\mu$ denotes the gradient step size (a real positive constant). To estimate the score function, we use a kernel based estimator with a low computing cost developed by Pham [7].

## 3. NOISE INFLUENCE STUDY

In this section, we are going to study the influence of the additive noise $n(t)$ in the model of the figure 1. More precisely, we try to analyze its effect on the estimation of $G(f)$ in the frequency band dominated by the noise. Indeed, we measure the characteristic on $y(t) = g\star w\star r(t) + g\star n(t)$, so we try to study what is the bias due to the noise presence, and how the criterion limited the noise amplification. In a noisy but non blind context where the wavelet is known, the minimum square error criterion yields the optimal Wiener filter $g_{\text{Wiener}}$, defined by its transfer function as:

$$\sum_{t=-\infty}^{\infty} g_{\text{Wiener}}(t)e^{-jt\omega} \overset{\text{def}}{=} G_{\text{Wiener}}(\omega) = \frac{W^*(\omega)}{|W(\omega)|^2 + \sigma_n^2} \qquad (10)$$

where $W(\omega) = \sum_{t=-\infty}^{\infty} w(t)e^{-jt\omega}$ is the Fourier series of the wavelet and $\sigma_n^2$ is the noise variance (we assume that the reflectivity series has been normalized to have unit variance). $\omega$ denotes the pulsation.

The Wiener filtering does make a trade-off between the data fitting and noise amplification. Indeed, in the passband of the wavelet $|W(\omega)|$ will be large with respect to $\sigma_n^2$ so that $G_{\text{Wiener}}(\omega) \approx 1/W(\omega)$, which means that the Wiener filter is close to the inverse of the wavelet filter ($g_{\text{Wiener}}\star w\star r \approx r$). On the other hand, in the null space we have $G_{\text{Wiener}}(\omega) = W^*(\omega)/\sigma_n^2 \approx 0$, so that the noise is strongly attenuated in this region. The Wiener filter provides the best compromise between noise reduction and fidelity of the reflectivity recovering, and will be used as reference for the comparison.

### 3.1 Minimum entropy deconvolution

For minimum entropy deconvolution algorithm, one can write the kurtosis of the deconvolution output $y$ as a function of the kurtosis of the "signal" part $g\star w\star r$:

$$K(y) = K(g\star w\star r)\left[\frac{\text{var}(g\star w\star r)}{\text{var}(y)}\right]^2$$

where $\text{var}(\cdot)$ denotes the variance. As $\text{var}(g\star w\star r) = \int_0^{2\pi}|G(\omega)W(\omega)|^2 d\omega/(2\pi)$, since $r(t)$ has unit variance by assumption, and $\text{var}(y) = \int_0^{2\pi}|G(\omega)|^2[|W(\omega)|^2 + \sigma_n^2]d\omega/(2\pi)$, one gets

$$K(y) = K(g\star w\star r)\left\{\int_0^{2\pi}\Delta(\omega)\frac{|W(\omega)|^2}{|W(\omega)|^2 + \sigma_n^2}d\omega\right\}^2 \qquad (11)$$

where

$$\Delta(\omega) = \frac{|G(\omega)|^2[|W(\omega)|^2 + \sigma_n^2]}{\int |G(\lambda)|^2[|W(\lambda)|^2 + \sigma_n^2]d\lambda}. \qquad (12)$$

Formula (11) shows that the kurtosis of the output $y(t)$ can be written as the product of two terms: the kurtosis of $g \star w \star r$ which represents the quality of the estimation of the reflectivity (or the fitting to the data) and is independent of the noise, and a second term which concentrates on the noise reduction. Naturally, to maximize the product of these terms, one must make a trade-off between maximizing each of them. The first term tries to estimate the inverse filter $g = w^{-1}$ while the second term has the effect of noise regularization. Indeed, $\Delta(\omega)$ can be viewed as a barycenter weight because $\int \Delta(\omega)d\omega = 1$, hence to maximize the second factor in (11) would lead to concentrate all the weight $\Delta(\omega)$ around the frequencies $\omega$ for which $|W(\omega)|^2/[|W(\omega)|^2 + \sigma_n^2]$ is maximum, or equivalently for which $|W(w)|$ is maximum. Thus this factor has the effect of pulling $|G(w)|$ toward 0 except at those $\omega$ for which $|W(\omega)|$ is maximum. Therefore we have in the MED algorithm a natural regularization to avoid noise amplification.

### 3.2 Mutual information rate based algorithm

Since the second term in the blind frequency deconvolution criterion (8) only enforces the continuity of the $G$ and has no effect on noise regularization, we will drop it and consider the continuous analogue of this criterion:

$$H(y) - \int_0^{2\pi} \log|G(\omega)| \frac{d\omega}{2\pi} + \lambda_2 \int_0^{2\pi} |G(\omega)|^p \frac{d\omega}{2\pi}.$$

Denote by $H^-(y) = \frac{1}{2}\log(2\pi e \sigma_y^2) - H(y)$, the negentropy of $y$, where $\sigma_y$ is the standard deviation of $y$, the above criterion can be written *up to an additive constant* as

$$J(G) = -H^-(y) + \frac{1}{2}\int \log \frac{\sigma_y^2}{f_y(\omega)} \frac{d\omega}{2\pi} + \lambda_2 \int |G(\omega)|^p \frac{d\omega}{2\pi} \qquad (13)$$

where $f_y$ is the power spectral density of the $y(t)$ process, which is related to that of the observed process $d(t)$ by $f_y(\omega) = |G(\omega)|^2 f_d(\omega)$. As we have explained in the construction of the cost function (8) in the subsection (2.2), the last term should is only meant to avoid large value of $|G|$ outside the pass-band of the wavelet. The second term can be viewed as a measure of flatness of $f_y$. Indeed, this term is minimum and zero if and only if $f_y$ is constant, meaning that $y(t)$ is a second order white process.

Since $f_y(\omega) = |G(\omega)|^2[|W(\omega)|^2 + \sigma_n^2]$, if we minimize only the above "second order whiteness" term, we would get a deconvolution filter with gain:

$$G(\omega) = \frac{\text{Constant}}{[|W(\omega)|^2 + \sigma_n^2]^{1/2}} = |G_{\text{Wiener}}(\omega)| \left[1 + \frac{\sigma_n^2}{|W(\omega)|^2}\right]^{1/2}.$$

This gains equals the Wiener gain times a factor which can be very large at frequencies $\omega$ for which $\sigma_n^2/|W(\omega)|^2$ is large, or equivalently for which $|W(\omega)|^2$ is small. The gain is however bounded: the "second order whiteness" term does prevent the noise to blow up. But its noise reduction is far smaller than that of the Wiener filter. This term tends to pull $|G(\omega)|$ to a constant (instead of 0) in the null space.

For the negentropy term, as $y(\tau) = g \star w \star r(\tau) + g \star n(\tau)$ and the random variables in this sum are independent and the second is Gaussian, we have:

$$H^-(y(\tau)) < H^-(g \star w \star r(\tau))$$

Further the larger the variance of $g \star n(\tau)$, the smaller $H^-(y(\tau))$ is. Thus, maximizing the negentropy would also prevent $g \star n(\tau)$ to become very large.

## 4. SIMULATION RESULTS AND DISCUSSIONS

Fig. 2 plots the result of a simulation experiment. The reflectivity (a) is simulated by a white process with supergaussian distribution. The supergaussianity is confirmed by the histogram on (e). On (d), the spectral density of the reflectivity series confirms the whiteness of this process. The wavelet on (b) is chosen as the sum of two Ricker wavelets of central frequencies 60 Hz and 120 Hz with respectively phase of 0 and 45 degree. This is a bandlimited and non minimum phase wavelet. The observation of (c) is obtained by convolution of the reflectivity sequence and the wavelet, with added Gaussian white noise. The signal to noise ratio is set to $8dB$.

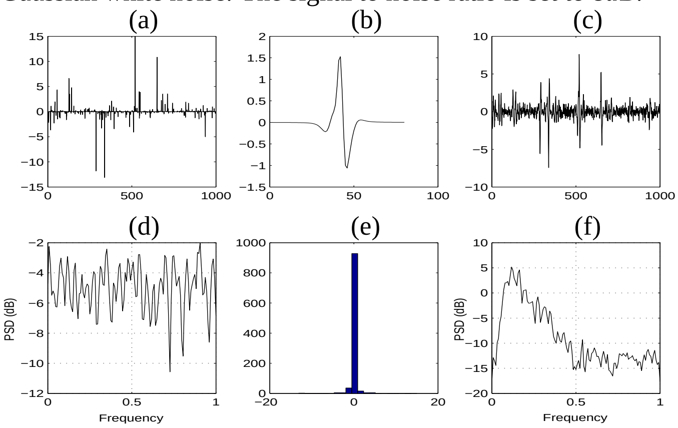


Figure 2: Simulated data for comparison of MED and FBD algorithm: (a) supergaussian reflectivity sequence, (b) wavelet, (c) observation with a SNR=8dB, (d) reflectivity power spectral density (psd) (e) reflectivity histogram (f) observation psd in dB.

One can note, by comparing the power spectrum density of the reflectivity (d) which is constant and the power spectrum density of the observation (f), that the wavelet is a bandlimited filter. Approximately, we can divide the frequency axis in two bands. The first from 0 to 0.5 is the passband of the wavelet and the second from 0.5 to 1, which is dominated by the noise, is the null space because it is a frequency band containing very poor information on the reflectivity due to the filtering done by the wavelet. It is the frequencies that one should not amplify, if one wants a good signal to noise ratio on the output $y(t)$.

On Fig. 3, we compare the Wiener filtering, the MED algorithm based on the kurtosis maximization and the FBD algorithm using mutual information rate in frequency domain, on the simulated data of Fig. 2. On the first row, we plot the deconvolution output resulted from (a) the Wiener filtering, (b) the MED algorithm and (c) the FBD deconvolution method. On the second row, we plot, on (d), (e) and (f), the power spectral density of the output of the above three algorithms.

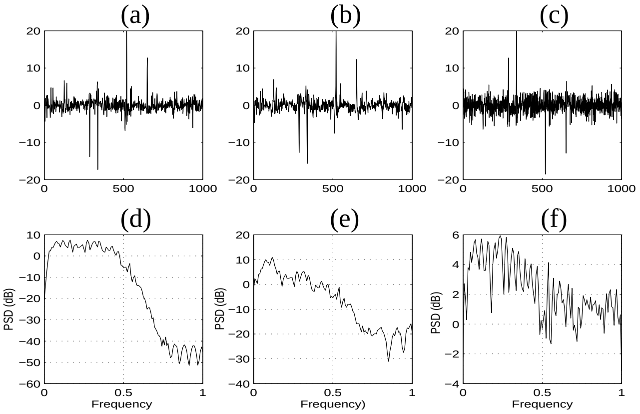


Figure 3: Comparison of Wiener filtering, MED and FBD algorithm: (a) Wiener filtering output, (b) MED deconvolution, (c) FBD deconvolution, (d) Wiener filtering output psd, (e) MED output psd (f) FBD output psd.

Concerning the deconvolution output (first row of Fig. 3), since the Wiener filtering minimize the mean square error, it provides the best SNR. However, it requires the knowledge of the wavelet. For the two other methods, we note that we can identify most of largest reflectors in the MED solution but not with FBD algorithm. Moreover, the noise level is lower for MED than FBD. This is confirmed by the spectral density. For the three algorithms, the output spectral density is quite constant in the wavelet passband $[0\ 0.5]$, so it is equivalent to a whitening in this frequency band. The noise level on the deconvolution output appears on the spectral density in the frequency band $[0.5\ 1]$. We note, on the MED spectral density difference of 20 $dB$ between the two frequency band, and only 3 $dB$ on the FBD spectral density. Note also that the FBD algorithm yields a more iid output than the MED. Nevertheless, the MED algorithm output is better because one has a better trade-off between the deconvolution quality (sharped spikes) and the noise amplification. Finally, Wiener filtering gives the best trade-off between data fitting and noise amplification with a difference between the two frequency band of 50 $dB$. To corroborate these finding, we plot the three inverse filter gains on Fig. 4.

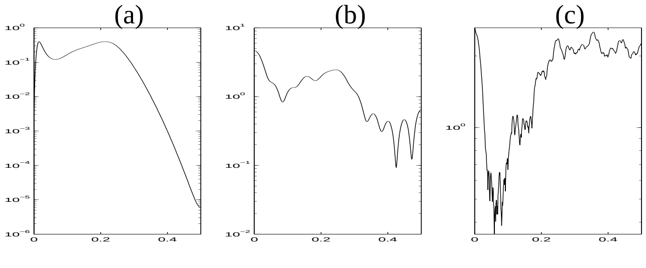


Figure 4: Frequency inverse filter gain: (a) Wiener filtering, (b) MED deconvolution, (c) FBD deconvolution

The plots show that the Wiener filtering has the lowest gain in the no space (frequency band $[0.5\ 1]$), following by the MED algorithm and the FBD algorithm. The last gain tends toward a constant gain as it is proved in the subsection 3.2. For the MED algorithm, the gain in the no space is not very large, which provides a good trade-off between deconvolution and noise amplification, nevertheless it is worst than Wiener filtering.

Finally, this simulation show that with the supergaussianity criterion, we have a natural trade-off between the data fitting and the noise amplification. Even if the kurtosis of the output is sensitive to the noise presence on the data, its effect is less important than on the whiteness measure as the mutual information rate. In fact, with supergaussianity we have a discriminant characteristic between the reflectivity and the noise. Whereas, with whiteness measure, we have an ambiguity with the whiteness of the reflectivity and of the noise.

## 5. CONCLUSION

We propose a new blind frequency algorithm based on mutual information of the output. On seismic data, we show that the supergaussianity based algorithm outperforms our algorithm. We provide some theoretical explanations of this phenomenon. Although our results concern seismic data processing, they may be applied to the blind deconvolution of supergaussian signal in a noisy context with a bandlimited convolution filter. The kurtosis is the simplest measure of the deviation from the Gaussianity, and it provides an interesting noise robustness.

## 6. ACKNOWLEDGMENTS

We would like to thank Industry Technological Facilitator (ITF) for support our work in the BLind Identification Seismic Signal (BLISS) project.

## REFERENCES

[1] M. Boumahdi, *Blind identification using the kurtosis with applications to field data*, Signal Processing, 48 (1996), pp. 205–216.

[2] J. A. Cadzow, *Blind deconvolution via cumulant extrema*, IEEE Signal Processing Magazine, (1996), pp. 24–42.

[3] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.

[4] D. L. Donoho, *On minimum entropy deconvolution*, Applied Time Series Analysis II, New York, 1981.

[5] W. Gray, *Variable norm deconvolution*, PhD thesis, Stanford University, 1979.

[6] A. Larue, J. Mars, and C. Jutten, *Frequency-domain blind deconvolution based on mutual information rate*, IEEE Trans. on Signal Processing, (Accept).

[7] D.-T. Pham, *Fast algorithm for mutual information based independent component analysis*, IEEE Trans. on Signal Processing, 52 (2004), pp. 2690–2700.

[8] D. T. Pham, *Mutual information approach to blind separation-deconvolution*, in Proceeding of the Thirteen European Signal Procesing Conference (EUSIPCO-2005), Antalia, Turquie, 2005.

[9] E. Robinson and S. Treitel, *Geophysical Signal Analysis*, Prentice Hall, 1980.

[10] A. Taleb, J. Solé i Casals, and C. Jutten, *Quasi-nonparametric blind inversion of Wiener systems*, IEEE Trans. on Signal Processing, 49 (2001), pp. 917–924.

[11] R. Wiggins, *Minimum entropy deconvolution*, Geoexploration, 16 (1978), pp. 21–35.

[12] L. Zhang and A. Cichocki, *Multichannel blind deconvolution of nonminimum-phase systems using filter decomposition*, IEEE Trans. on Signal Processing, 52 (2004), pp. 1430–1441.