Blind frequency deconvolution : A new approach using mutual information rate

Anthony Larue and Jérôme I. Mars and Christian Jutten, Laboratoire des Images et des Signaux, Grenoble*

Summary

In this paper, we purpose a blind frequency deconvolution method using higher order statistics by minimizing the mutual information rate of the deconvolved output. We add to the criterion a regularization term to limit noise amplification. Then, we compare on real data (underwater explosions and seismovolcanic phenomena) our deconvolution algorithm with existing methods.

Introduction

In reflection seismology, the sensor antenna recordings allow to describe the substratum using the reflection of the short-duration wave transmitted in the earth at each impedances changes. Under simplifying assumptions, one such trace can be modeled as a 1-D convolution:

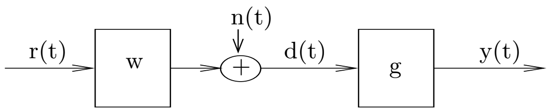

Fig. 1: Setup of the convolution and deconvolution problem

The vertical earth reflectivity $r(t)$ is convolved by the unknown wavelet w:

$$d(t) = (w * r)(t) + n(t) = \sum_{i=-\infty}^{+\infty} w(i)r(t-i) + n(t) \quad (1)$$

where $d(t)$ is the recorded signal and $n(t)$ is the additive sensor noise signal. Some methods (Champagnat et al., 1996; Lavielle, 1993) used in Bayesian formulation the prior hypothesis that the reflectivity signal $r(t)$ is a Bernouilli-Gaussian process. The first step is a detection of the reflectors and it follows by a magnitude estimation. The high noise level on recordings limits the performance of the detection step. The deconvolution problem can be also applied to the seismovolcanic phenoma. Then, the recording is the result of a convolution between the excitation $r(t)$ and the filter w, which is a resonant filter giving information about the volcano geometry. This data can be processed with a blind deconvolution algorithm, in whose only $d(t)$ is accessible to the algorithm, whereas r, n and w are unknown parameters. In a blind deconvolution problem we aim at finding a deconvolution filter g for computing the output of deconvolution process $y(t) = (g*d)(t)$. Assuming the source signal $r(t)$ is iid (Independently and Identically Distributed) and non Gaussian, the solution set of the blind deconvolution problem is generated by an only solution with a delay and magnitude modifications. Filter phase determination need to use higher order statistics (HOS). Boumahdi (Boumahdi, 1996) proposes to use the simplest HOS-the kurtosis-for estimating non-minimum phase Moving Average (MA) or autoregressive (AR) or ARMA models. These methods come up against the same problem of the noise. So, in the following part, we present our criterion and our algorithm using more general HOS, then in the last part we show tests about simulated and real data.

Deconvolution algorithm

We define for a T sample stationary stochastic process $Z = \{z(t)\}$ the mutual information rate by (Cover and Thomas, 1991):

$$I(Z) = H(z(\tau)) - H(Z) \quad (2)$$

where H denotes the entropy rate, then $H(Z)$ is the joint entropy rate of all samples $z(t)$ and $H(z(\tau))$ is the marginal entropy rate of the sample $z(\tau)$ which is the same for each τ under stationary hypothesis. We shall notice that $I(Z)$ is always positive ans vanishes if Z is iid process. With, the notation defined above, we can define with the mutual information rate of the deconvolution output signal $y(t)$ a deconvolution criterion with respect to the inverse filter g. In (Taleb et al., 2001), Taleb *et al.* define a criterion with respect to the impulse response $g(t)$, then their method is dedicated to the inversion of autoregressive model of the wavelet w. We can show that the algorithm is equivalent to a maximum likelihood (ML) method replacing the source distribution supposed known in ML method by the distribution of deconvolution output estimated at each iteration. We decide to use a frequency criterion in order to avoid a parametric approach like MA, AR or ARMA models. And, we note that it is easier in frequency domain to add a regularization like the Wiener filtering. We define the criterion with respect to $\mathbf{G} = [G(0), \dots, G(T-1)]$ the discrete frequency response of the inverse filter g:

$$J(\mathbf{G}) = H(y(\tau)) - \frac{1}{T}\sum_{\nu=0}^{T-1} \log|G(\nu)|$$
$$+\lambda_1 \sum_{\nu=0}^{T-1} |G(\nu) - G(\nu+1)|^2 + \lambda_2 \sum_{\nu=0}^{T-1} |G(\nu)|^p \quad (3)$$

The two first terms of (3) come from the mutual information rate of $y(t)$ expect for a constant, which equal to the joint entropy rate of the observed process $d(t)$. We add two regularization terms balanced by the hyperparameters λ_1 and λ_2. The first term constrains the frequency response of the inverse filter to be continuous. Practically, we notice that it also improves the stability of the minimization algorithm. The last term penalizes with the norm $\mathbb{L}^p$ the most important values of the spectrum of g. Thus, this term is equivalent to the noise factor in the Wiener filtering, it allows a trade-off between the deconvolution quality and the noise amplification.

Our goal is to minimize the criterion (3) with respect to the complex coefficient vector $\mathbf{G}$ according to a gradient iterative method (Brandwood, 1983; Van Den Bos, 1994). We compute the gradient of the criterion (3) with respect the frequency response coefficient $G(\nu)$:

$$\begin{aligned}\nabla_{G^*(\nu)} J(\mathbf{G}) &= \frac{1}{2T^2}\Psi_Y(\nu)D^*(\nu) - \frac{1}{2T}\frac{1}{G^*(\nu)} \\ &+ \lambda_1(2G(\nu) - G(\nu+1) - G(\nu-1)) \\ &+ \lambda_2 \frac{p}{2}\frac{|G(\nu)|^p}{G^*(\nu)} \qquad (4)\end{aligned}$$

where $D(\nu)$ is the spectrum of the observation $d(t)$ and z^* denotes the conjugate of the complex z. We also define the score function of the process Y by:

$$\psi_Y(u) = -\frac{d}{du}\log p_Y(u)$$

Then, $\Psi_Y(\nu)$ is the discrete Fourier transform of $\psi_Y(y(\tau)), \tau = 1, \ldots, T$. The score function is estimated with a kernel method developed by Pham (Pham, 2003). The computing time is reduced to $3 \times T$ using the cubic spline as kernel. The blind frequency deconvolution (BFD) 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 ψ_Y
3. computation of the gradient (4)
4. updating of $G(\nu) \leftarrow G(\nu) - \mu\nabla_{G^*(\nu)} J(\mathbf{G})$
5. computation of the deconvolution output $y(t)$
6. normalization step

We iterate from 2 to 6 until convergence. The normalization step is required for taking into account scale indeterminacy in $G(\nu)$.

Examples

Fig.2 shows deconvolution results with simulated signals. The source $r(t)$ is a 400-sample Bernouilli-Gaussian process with 50 reflectors. We use a seven order autoregressive direct filter, whose coefficients are [2 0.8 1 0.9 0.8 0.2 0.1], for providing $d(t)$ (Fig.2(b)). We see that the deconvolution output on Fig.2(c) is similar to the source of Fig.2(a). The inverse filter modulus and phase are compared to the theoretical inverse filter on Fig.2(d) and Fig.2(e). The estimation errors are due to the large number parameters optimized, and to the standard deviation of the score function estimator.

Fig.3 shows in (a) the real data of an underwater explosion recorded in a swimming pool. The recorded signal is composed of the direct wave, a reflection on the surface and on the bottom of the swimming pool. While applying the model of Fig.1, the source signature $r(t)$ containing information about the reflection coefficients, and the direct filter w represents the wave generated by the explosion characterized to the "bubble effect". A quite good model of this wave is a non causal moving average (MA) filter. So, we compare on Fig.3 the deconvolution results obtained with spectral egalization, Durbin method and our BFD algorithm. We note that the three methods provide a good positioning of the three events. Spectral egalization and our method realize a better trade-off between deconvolution and noise amplification than the Durbin method in which secondary peaks appear just after reflector. We note that our algorithm gives the nearest to zero output between each reflector, and the magnitude of reflectors seems to be better preserved.

Fig.4 and 5 deal with seismogram recorded on colombian volcano Purace and Galeras. The aim is to separate the resonance effects to the less energetic effects such as excitation and propagation. The recording (Fig.4(a)) is characterized to the long period events (Lesage et al., 2002), with an important resonance. Using the model of Fig.1, the residual signal obtained by deconvolution contains information about the excitation of the volcano, and the resonant filter w gives information about the geometry of the volcano. On Fig.4, we compare the spectral egalization, the Yule Walker method dedicated to autoregressive filter and our frequency algorithm for the Purace recording. We note that our algorithm separate better the excitation to the noise. In Fig.5, we show the observation and the deconvolution output with our method for Galeras.

Conclusion

We purpose a new algorithm of blind deconvolution based on the mutual information rate of the deconvolution output. We write the criterion in frequency domain and we add a compromise between the deconvolution quality and the noise amplification like in Wiener filtering. Seismic deconvolution will be present. The first test about real data give some improvement with respect to spectral egalization or the second order blind methods as Yule Walker and Durbin methods.

Acknowledgments

We would like to thank P. Lesage of Laboratoire de Geophysique Interne et Tectonophysique from Grenoble for his analysis about our results on volcano datasets. We acknowledge "LACROIX" society for the underwater data.

References

Boumahdi, M., 1996, Blind identification using the kurtosis with applications to field data: Signal Processing, **48**, no. 3, 205–216.

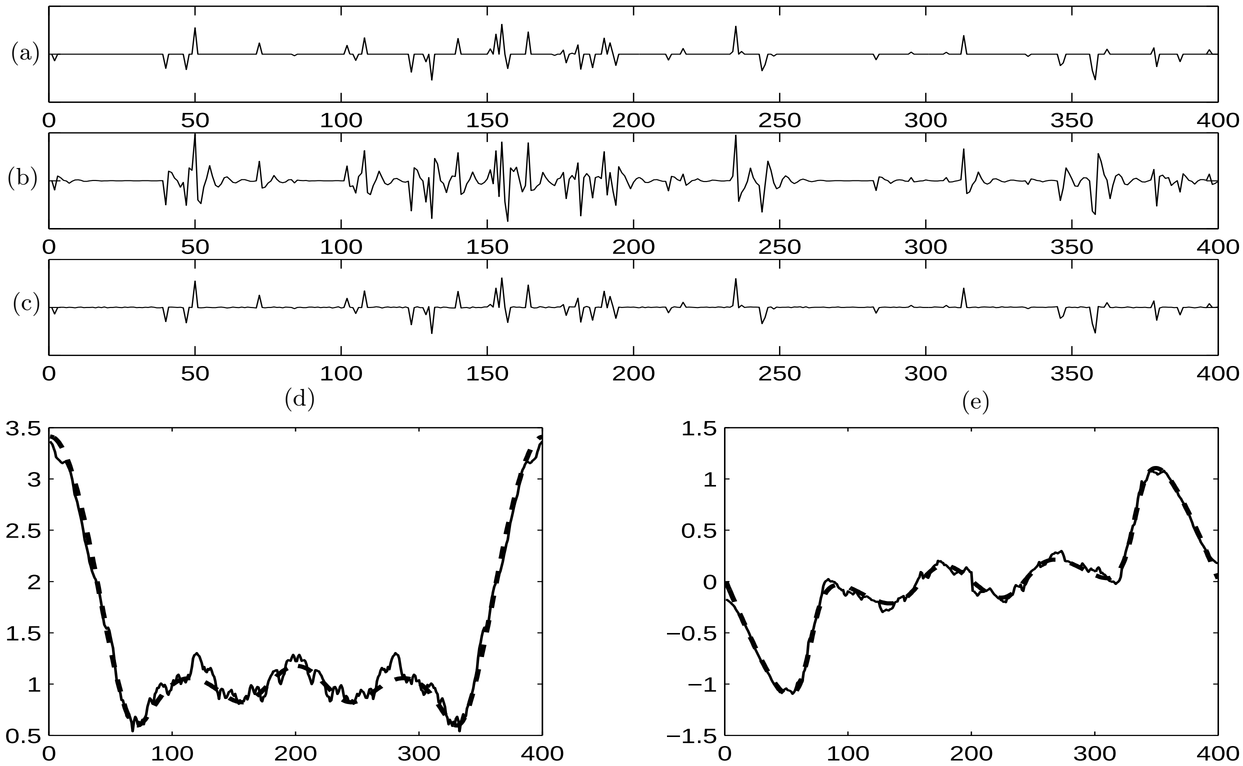

Fig. 2: (a) Reflectivity simulated with a Bernouilli-Gaussian process with 50 reflectors, (b) Observation with a seventh order autoregressive wavelet, (c) Deconvolution output signal, (d) (respectively (e)) In solid line the modulus (respectively phase) of estimated inverse filter, in dashed line the theoretical modulus (respectively phase)

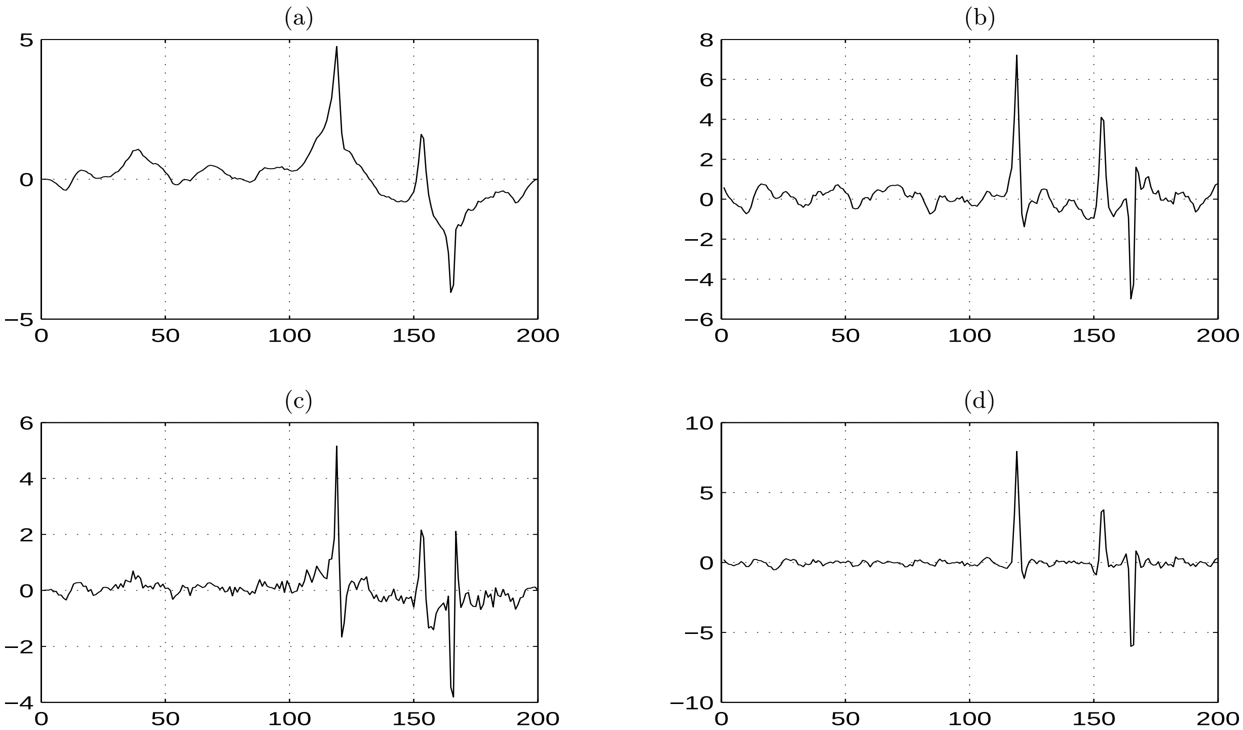

Fig. 3: (a) Observation of underwater explosion, (b) After spectral egalization, (c) After deconvolution by Durbin method, (d) Deconvolution with BFD algorithm with $\mu = 0.01$, $\lambda_1 = 0.1$, $\lambda_2 = 0.1$ and $p = 3$

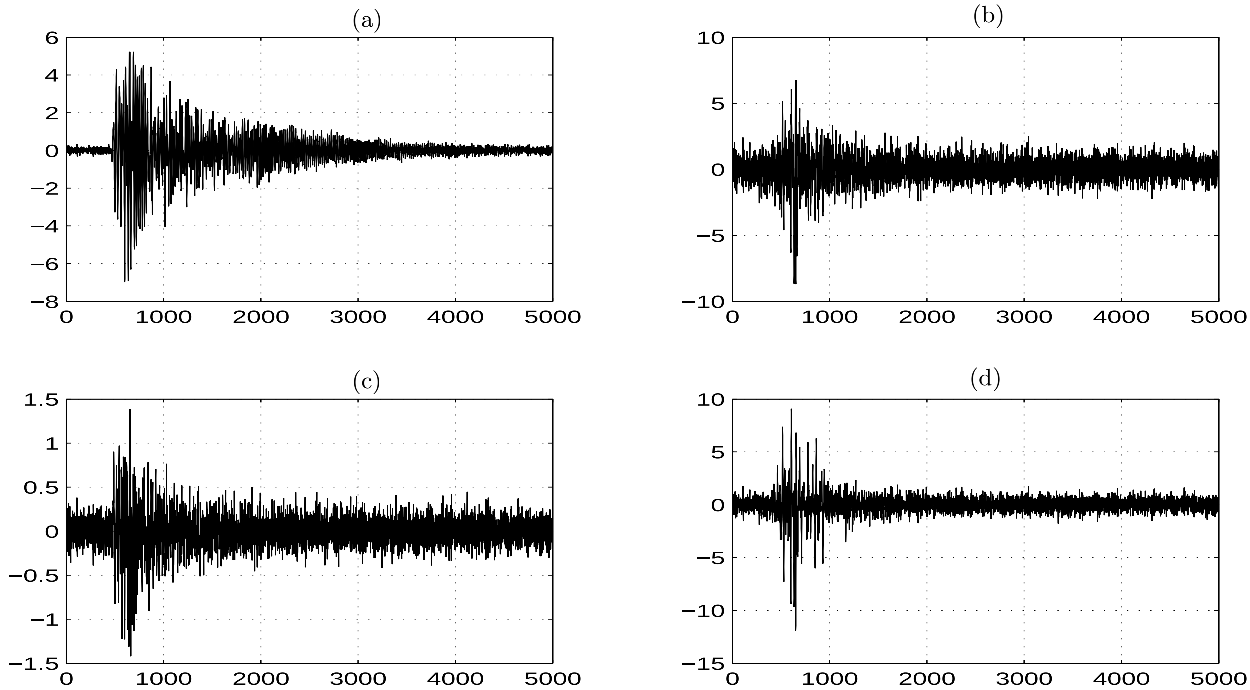

Fig. 4: (a) Purace volcano recording, (b) After spectral egalization, (c) After deconvolution by Yule Walker method, (d) Deconvolution by BFD algorithm with $\mu = 0.01$, $\lambda_1 = 0.1$, $\lambda_2 = 0.1$ and $p = 3$

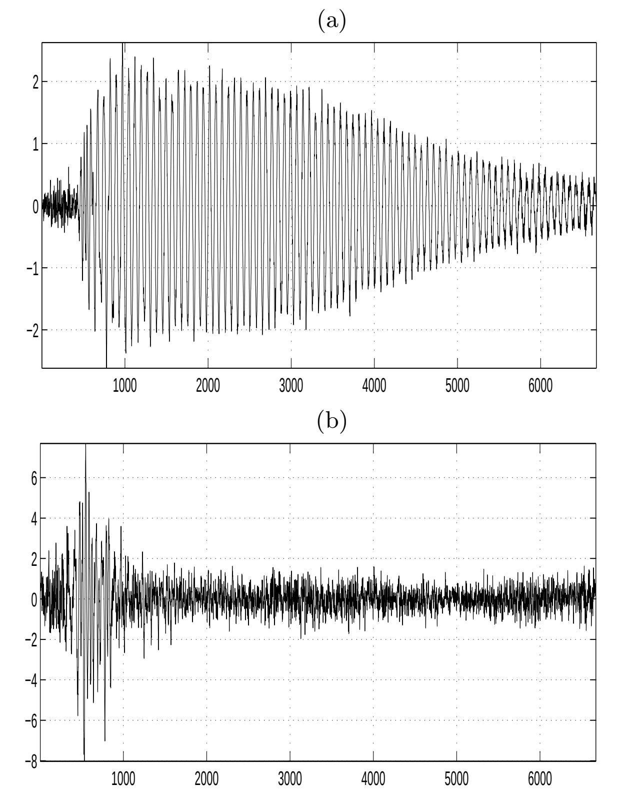

Fig. 5: (a) Galeras volcano recording, (b) Deconvolution by BFD algorithm with $\mu = 0.01$, $\lambda_1 = 0.1$, $\lambda_2 = 0.1$ and $p = 3$

Brandwood, D. H., 1983, A complex gradient operator and its application in adaptative array theory: IEE Proc., **130**, no. 1, 11–16.

Champagnat, F., Goussard, Y., and Idier, J., December 1996, Unsupervised deconvolution of sparse spike trains using stochastic approximation: IEEE Transactions on Geoscience and Remote Sensing, **44**, no. 12, 2988–2998.

Cover, T., and Thomas, J., 1991, Elements of information theory: Wiley Series in Telecommunications.

Lavielle, M., 1993, Bayesian deconvolution of Bernouilli-Gaussian processes: Signal Processing, **33**, 67–79.

Lesage, P., Glangeaud, F., and Mars, J., 2002, 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, **114**, 391–417.

Pham, D.-T., 2003, Fast algorithm for estimating mutual information, entropies and scores functions: Independent Component Analysis, 17–22.

Taleb, A., Solé i Casals, J., and Jutten, C., May 2001, Quasi-nonparametric blind inversion of Wiener systems: IEEE Transactions on Signal Processing, **49**, no. 5, 917–924.

Van Den Bos, A., 1994, Complex gradient operator and hessian: IEE Proceedings on Vision, Image and Signal Processing, **141**, 380–382.