RECOVERY OF A HIGH SHOCK PROBABILITY PROCESS USING BLIND DECONVOLUTION

Francois Combet, Nadine Martin, Pierre Jaussaud
Laboratoire des Images et des Signaux, BP 46,
38402 Saint Martin d'Hères Cedex, France
Tel: 33(0)4 76 82 64 20; fax: 33(0)4 76 82 63 84
e-mail: 1stname.2ndname@lis.inpg.fr

ABSTRACT

We intend in this paper to recover shocks occurring in the situation where they are highly probable using blind deconvolution methods : order 2 (Yule-Walker), higher order (normalized cumulants), and mutual information. We define a *shock probability process* derived from a Bernoulli process and we show that, in opposite to the other two methods, the normalized cumulants are *highly dependent* on the shock probability P. This has consequences on performance versus P, which are studied applying a Kurtosis maximization algorithm on simulations. We finally apply the three blind deconvolution methods to a real signal recorded on a rope transportation line. Results comparison with an *a priori* shock model of the mechanical system confirms that normalized cumulants are less adapted for recovering a high shock probability process.

1 INTRODUCTION

In many practical applications, signals $x(n)$ can be modeled as the filtering of an excitation signal $s(n)$ by a linear and time-invariant system. In case of an unknown system's response f_n, recovering the excitation signal is called a Blind Deconvolution (BD) problem. The inverse filter g_n has to be identified so that the cascade filter impulse response $h_n = f_n * g_n$ (fig.1) has all of its elements zero except one :

$$h(n) = A\delta(n-r) \quad (1)$$

This ideal BD can be theoretically achieved assuming that the excitation time series $s(n)$ is composed of *independent and identically distributed* (iid) samples generated from an underlying random variable S with a non-Gaussian probability density function $p_S(s)$. But the scaled and delay factors A and r are not identifiable, so we can only recover an excitation $y(n)$ which is a scaled and time-shifted image of $s(n)$. A lot of solutions have been proposed in the literature. In this paper, we intend to compare some BD methods for recovering an excitation signal composed of a *finite number* of random localized impulses. For instance, these impulses could represent reflections in non-destructive testing applications [1] or shocks submitted to a mechanical system [2]. We model this excitation signal as an impulse type random process S derived from a Bernoulli process, but with a centered probability density function :

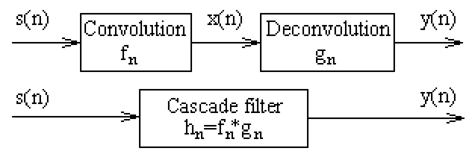
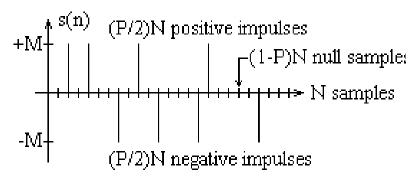


Figure 1: Convolution-deconvolution operation and equivalent cascade system.

Figure 2: N samples realization of the impulse process S for a shock probability P.



$$p_S(s) = \frac{P}{2}[\delta(s-M) + \delta(s+M)] + (1-P)\delta(s) \quad (2)$$

in which M is an undetermined constant and P denotes the probability that the random variable S generates an impulse with same probability to get a positive or a negative sign. This means that in each S realization of N samples $s(1), ..., s(N)$, we find one impulse every $1/P$ samples on average, and we expect to have a total of NP impulses (fig.2). P can be interpreted as a *"shock probability"*. We intend to study BD performance when recovering widely spaced impulses (low shock probability) or close impulses (high shock probability).

In section 2, we present a rapid non exhaustive overview of BD methods that we split in three categories : order 2, higher orders and mutual information. In section 3, we show that higher order methods based on normalized cumulants are highly dependent on the shock probability P. In section 4 we study the P dependence of a Kurtosis maximization method on simulations. BD methods are then tested on a real acceleration signal recorded on a rope transportation line, and supposed to be the mechanical response of close impulses (relatively high P). We finally conclude by comparing the BD results with an *a priori* excitation model of the system.

2 BLIND DECONVOLUTION METHODS

2.1 Order 2 Methods

Order 2 methods [3] do not identify the inverse filter g_n but only a whitening filter which tends to minimize the output's power $E[Y^2]$. The filter's phase is unidentifiable by using order 2 methods, so we have to assume an *a priori* phase (minimum phase, zero phase...). The iid hypothesis implies that the excitation signal is decorrelated (order 2 white process). But order 2 methods are not sufficient for a complete system identification.

2.2 Higher Order Methods

Phase indeterminacy of order 2 methods can be raised by using higher order statistics. In [4] is given an interesting theoretical result based on the *normalized cumulants* defined for a process S by : $K_q(S) = Cum_q(S)/\sigma_S^q$, where $Cum_q(S)$ denotes the q order cumulant of S and σ_S denotes the standard deviation of S. Assuming an iid process S, all $q > 2$ order normalized cumulants are closer to zero for a S filtering than for S. This applied to the cascade filter output Y leads to :

$$|K_q(S)| \geq |K_q(Y)| \tag{3}$$

Furthermore, for any $K_q(S) \neq 0$:
$|K_q(S)| = |K_q(Y)| \Leftrightarrow h_n = A\delta(n-r)$
which corresponds to an ideal BD (1). So the BD problem can be solved by selecting the inverse filter g_n so as to maximize an estimate of $|K_q(Y)|$. This can be achieved by using a *stochastic gradient algorithm* [4]. Assuming a finite impulse response g_n, vector $\mathbf{g}$ of the g_n coefficients is perturbed in the following additive fashion at iteration i :

$$\mathbf{g}_{(i)} = \mathbf{g}_{(i-1)} + \mu \nabla_{\mathbf{g}}[O] \tag{4}$$

where μ is the algorithm step and $\nabla_{\mathbf{g}}[O]$ is the gradient vector of the objective function O being maximized, here $O_q = |K_q(Y)|$. Another method has been proposed in [1] which happens to be faster convergent and no dependent on the step value. We present it in section 4.

2.3 Mutual Information Minimization (MIM)

Another approach of the BD problem, recently proposed in [5], is based on the *mutual information*, defined as follows for a N dimension vector $\mathbf{y}$:
$I(\mathbf{y}) = \sum_{n=1}^{N} H(y(n)) - H(y(1), \ldots, y(N))$
where $H(y(n)) = \int p_Y(u) \log p_Y(u) du$ is the entropy of $y(n)$. Since $I(\mathbf{y})$ is always positive for a dependent samples vector and zero for an iid samples vector, the BD problem can be solved by minimizing $I(\mathbf{y})$. This is achieved by also using a stochastic gradient algorithm, with $O_I = I(\mathbf{y})$. Reader will find all algorithm details in [5]. This method is theoretically the optimal method because of recovering an iid process. However, the output distribution law has to be estimated at each iteration and thus the used algorithm is highly computationally demanding.

3 DEPENDENCE ON THE SHOCK PROBABILITY P

Using the excitation model generated from the process S defined in (2), we aim to study the previously presented methods dependence on the shock probability P. The MIM method is not P dependent as the mutual information is always zero for an iid process, whatever be the distribution law, and so the value of P. Order 2 methods are also P independent as an iid process is also a decorrelated process which has a white power spectral density. However, higher order methods are highly P dependent as shown subsequently. Since $p_S(s)$ is here a symmetrical probability density, all odd order normalized cumulants of S are zero and so unusable. The fourth order normalized cumulant (called Kurtosis) of the S process defined in (2) is readily calculable :

$$Kurt(S) \triangleq K_4(S) = \frac{E[S^4]}{E[S^2]^2} - 3 = \frac{1}{P} - 3 \tag{5}$$

This result shows that $Kurt(S)$ *is non-linearly dependent on the shock probability* P. This will have consequences on BD performances using Kurtosis (section 4). We also calculate the sixth normalized cumulant of S :

$$K_6(S) = \frac{E[S^6]}{E[S^2]^3} - 15\frac{E[S^4]}{E[S^2]^2} + 30 = \frac{1}{P^2} - \frac{15}{P} + 30 \tag{6}$$

which is even more P dependent. Higher q order normalized cumulants of S depend on a term $P^{1-q/2}$, and so are more and more P dependent. This is why we limit our study to the fourth order (Kurtosis). For $P = 1/3$, $Kurt(S)$ is zero and so unusable. Applications here concern a relatively low P value, so we limit our study to $P < 1/3$, which implies $Kurt(S) > 0$ (super-Gaussian process). BD is completed by maximizing $Kurt(Y)$ value so as to approach the equality in (3). In fact, *as we use finite N dimension signals which are not ideal iid processes*, we expect bad performances for P near 1/3 ($Kurt(S)$ near zero) and better performances for lower P values (higher $Kurt(S)$ values). This will be shown in the next section on simulations.

4 KURTOSIS MAXIMIZATION PERFORMANCE VERSUS SHOCK PROBABILITY

4.1 Fast Kurtosis Maximization algorithm

To perform the inverse filter output Kurtosis Maximization (KM), [1] proposed to estimate the inverse filter so that the following objective function is maximized :

$$O_4 = \sum_{n=1}^{N} y^4(n) / [\sum_{n=1}^{N} y^2(n)]^2 \tag{7}$$

Optimizing O_4 with respect to L inverse filter coefficients g_l, i.e. $\partial O_4/\partial g_l = 0$ for $l = 1, \dots, L$ leads to a system of L non-linearly equations that can be written in matrix notation : $\mathbf{b} = \mathbf{R_{xx}g}$, where $\mathbf{R_{xx}}$ is a modified autocorrelation matrix of the observed signal $x(n)$ which terms are :
$R_{xx}(l,j) = \sum_{n=1}^{N} x(n-l)x(n-j)$ for $l, j = 1, \dots, L$
$\mathbf{g} = [g_1 \dots g_L]^T$, and $\mathbf{b}$ is a column vector that contains L normalized intercorrelation terms between $y^3(n)$ and $x(n)$:
$b(l) = [\sum_{n=1}^{N} y^2(n)][\sum_{n=1}^{N} y^3(n)x(n-l)]/[\sum_{n=1}^{N} y^4(n)]$
As this system is highly nonlinear, it is solved iteratively by the following summarized algorithm :
$\mathbf{y_{(i-1)}} \rightarrow \mathbf{b_{(i)}} \qquad \mathbf{R_{xx}^{-1}b_{(i)}} \rightarrow \mathbf{g_{(i)}} \qquad \mathbf{x} * \mathbf{g_{(i)}} \rightarrow \mathbf{y_{(i)}}$
Due to the scale indeterminacy (1), we normalize at each iteration the inverse filter coefficients in order to have an output signal power $E[Y^2] = 1$. Algorithm runs until the output Kurtosis stabilization. We choose the following stop condition : $|O_4^{(i+1)} - O_4^{(i)}|/O_4^{(i+1)} < 10^{-4}$.

4.2 Performance study versus P on simulations

In order to study the KM algorithm performance, we simulate a filter corresponding to a single resonant frequency damped system. The synthesized filter is an ARMA filter which z-transform is :

$$F(z) = \frac{0.79 - 1.63z^{-1} + 0.81z^{-2}}{1 - 1.47z^{-1} + 0.81z^{-2}}$$

This filter has one zero outside the unit circle, and is so a *non minimum phase filter*. We so expect a non causal inverse filter $\mathbf{g}$ that contains an infinity of decreasing coefficients. We choose a sufficient length L so that the truncation effects can be ignored, $L = 25$, with same number of causal and non-causal coefficients : $g_l \neq 0$ for $l = -12, \dots, 12$. We also add a Gaussian white noise to the observed signal $x(n)$, so as to have a Signal to Noise ratio $S/N = 17$ dB. As a measure of performance we use the *intersymbol interference* (ISI) criterion defined in [6] by : $ISI = [\sum_n h_n^2 - max(h_n^2)]/\sum_n h_n^2$

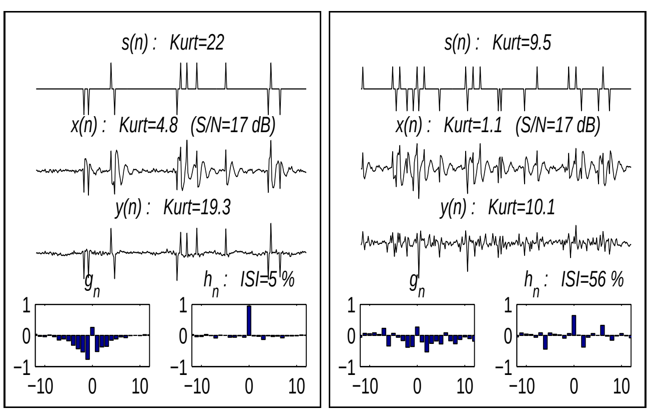


Figure 3: Results of the Kurtosis Maximization algorithm from two S realizations with $N = 300$ samples, left one is for $P = 1/25$ and right one for $P = 2/25$.

Small ISI indicates the proximity to an ideal deconvolution (1). Fig.3 shows two deconvolution results obtained from two different $s(n)$ realizations with the same samples number $N = 300$ but two different shock probability P (1/25 and 2/25). The second simulation has not performed and, *surprisingly*, $y(n)$ Kurtosis is *higher* than $s(n)$ one. This can be interpreted with respect to (5) : *the Kurtosis Maximization algorithm tends to minimize the shock probability P, and, due to estimation errors, risks to converge to a spurious higher local maximum corresponding to less recovered shocks.*

We tested the KM algorithm using 100 independent S realizations for different P values, and for two signal length N (fig.4 top). As expected, the ISI criterion increases versus P on average. Performance is obviously better for a higher N samples number ($N = 600$), but the ISI variances are still high for high P values. This moderate performance is due to interferences between close shocks on the length L of the inverse filter (here $L \simeq 1/P$). In order to reduce shocks interferences, we propose an improvement by applying a *Hamming window* to the g_n coefficients at each algorithm iteration. Results may be biased for a low g_n decrease, but the ISI are much reduced and quasi not P dependent as shown in fig.4 bottom.

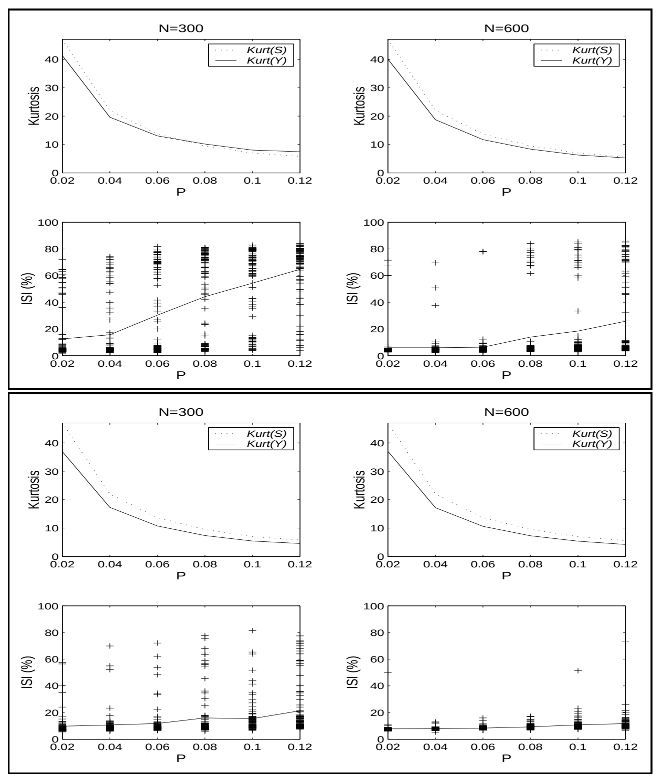


Figure 4: KM method without (top) and with a Hamming window applied to the inverse filter g_n (below), Kurtosis means of S $(\cdots)$ and Y $(-)$, ISI means $(-)$ and ISI realizations $(+)$ of 100 independent simulations for several P values and 2 signal length N.

5 COMPARISON ON A REAL SIGNAL

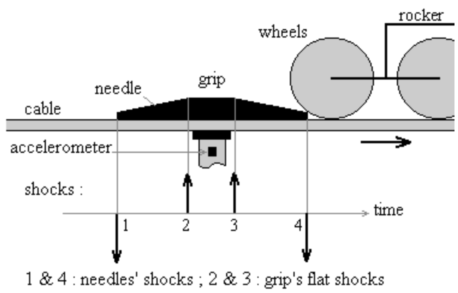


Figure 5: Shock model of the grip.

We present in that section BD results obtained from a real acceleration signal recorded on a chairlift of a rope transportation line [7]. When running a compression tower, the chairlift's grip must insert between the cable and the wheels, which causes shocks (fig.5). The mechanical system and its resonance frequencies are thus excited : shocks causes oscillations of the rocker around its equilibrium position, which are damped by the cable. Recorded vertical acceleration signal is plotted fig.6 (b) as well as an *a priori* excitation model (a) that we built using geometry of the grip and the rocker. But amplitudes and positions of the shocks are not exactly known, nor is the system response. So we use BD methods for recovering the excitation signal. Although shocks positions are deterministic, the excitation model could be seen as one of the ocurrences of the process S defined in (2). We expect 32 shocks among $N = 400$ samples, so a shock probability $P = 32/400 = 0.08$. This value is relatively high for the KM algorithm (fig.4). In fig.6 (bottom) are plotted the different BD results from the KM algorithm without window (c) and with a Hamming window (d), an order 2 method (Yule-Walker) with a causal inverse MA2 filter (e), and the MIM method (f). We also indicate the estimated Kurtosis values of the recovered excitations. We only plot the impulses greater than a threshold fixed experimentally from only noise record. As expected, the KM method is here the less efficient referred to the shock model, despite its higher Kurtosis value. This is due to the high shock probability of the considered application. Applying a Hamming window gives a result better in agreement with the other two methods. The Yule-Walker's result is well in accordance with the model and similar to the MIM result. We so conclude that the mechanical system could be modeled by a minimum phase filter with one resonant frequency (AR filter with 2 conjugate poles).

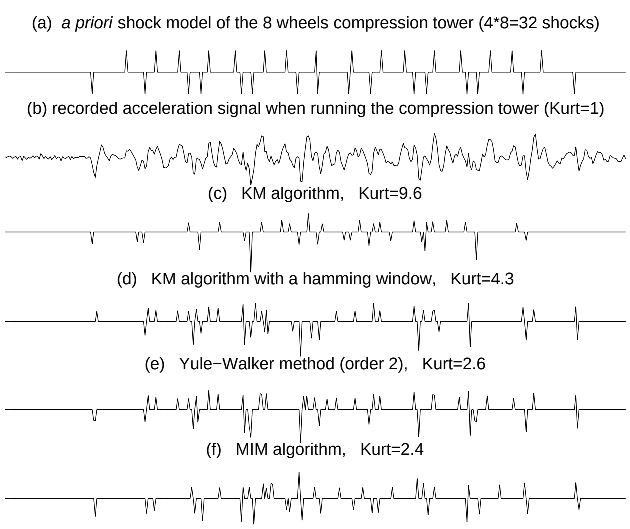


Figure 6: (a) excitation model, (b) real signal, (c-f) results of the Blind Deconvolution methods.

6 CONCLUSION

In this paper, we intended to recover a high shock probability process using Blind Deconvolution methods. We showed that, in opposite to the other methods, the well-known $q > 2$ order methods based on normalized cumulants extrema are *highly dependent on the shock probability* P of the process to be recovered. This has consequences on performance, which was confirmed on simulations and on a real signal. Furthermore, as the P dependence increases versus order q, we expect worse performance using high q orders. This was observed in [1] and [2] *but not interpreted.* We also proposed an improvement of the Kurtosis Maximization algorithm. Comparison of these results with a non-stationary method based on Prony's model [8-9] is under investigation.

References

[1] A.K. Nandi, D. Mämpel, B. Rosher, "Blind Deconvolution of Ultrasonic signals in Nondestructive Testing Applications", *IEEE Transactions on Signal Processing*, VOL 45, NO 5, May 1997.

[2] J.Y. Lee, A.K. Nandi, "Blind Deconvolution of Impacting Signals using Higher-Order Statistics", *Mechanical Systems and Signal Processing*, 12(2), 357-371, 1998.

[3] S. Haykin, "Adaptative filter theory", Prentice Hall 1991.

[4] J.A. Cadzow, "Blind Deconvolution via Cumulant Extrema", *IEEE SP Magazine*, VOL 13, 24-42, May 1996.

[5] A. Taleb, J. Solé, C. Jutten, "Quasi-Nonparametric Blind Inversion of Wiener Systems", *IEEE Transactions on Signal Processing*, VOL 49, NO 5, May 2001.

[6] O. Shalvi, E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)", *IEEE Transactions on Information Theory*, VOL 36, NO 2, 1990.

[7] P. Jaussaud, N. Martin, "Bob Movement and Vehicle Acceleration Running a Support or Compression Tower", *OITAF 8th international congress*, San-Francisco, May 1999.

[8] P. Jaussaud, N. Martin, V. Pellisier, "Detection and modelling of a vehicle acceleration on a compression tower by a Triple Prony Processing", *PSIP*, Paris, 1999.

[9] F. Combet, N. Martin, P.Jaussaud, "Shocks detection and modelling of multi-pulses signals by a Prony Time-Frequency processing", submitted to *ICASSP 2002*.