KURTOSIS MAXIMIZATION FOR BLIND IDENTIFICATION OF NONLINEAR COMMUNICATION CHANNELS

J.-P. Costa, J. LeRoux

Laboratoire I3S, CNRS-UNSA
2000 route des Lucioles, 06410 Biot, France
costa@i3s.unice.fr, leroux@essi.fr

ABSTRACT

This paper presents an original approach for blind deconvolution of a nonlinear communication channel using a criterion based on the second- and fourth-order moments of the input sequence. This approach is a simple extension of kurtosis maximization a method well known in a linear blind identification. We illustrate through a simple example that kurtosis maximization may also be used in a generalized Wiener-Hammerstein nonlinear systems identification process. The only constaint on probability distribution of the unobserved input process is non Gaussianity.

1. INTRODUCTION

The problem of blind deconvolution arises when an unknown system output is observed, but whose input is unobserved. The final purpose is the recovering (deconvolution) of the input process. This is a problem of considerable interest in several fields including channel equalization [1].

Kurtosis maximization has been used for solving blind deconvolution of linear systems [2] [3], [4]. This method is universal in the sense that no restrictions, except non Gaussianity, on the probability distribution of the unobserved input process is imposed. In digital communication channels nonlinear distorsions often arise at high transmission rates, due to saturations generated by amplifiers [5], [6]. Consequently, linear techniques require modifications in order to treat this kind of nonlinear problems.

We propose to identify the nonlinear communication channel using a criterion based on the recovered sequence kurtosis maximization. The communication channel is represented by a generalized Wiener-Hammerstein nonlinear system composed by a cascade of a linear dynamic, a static nonlinearity and a linear dynamic systems ($L-N-L$) [7].

The problem is illustrated in figure 1, where H_1, H_2 are unkown linear dynamic filters and $\varphi(.)$ corresponds to the static nonlinear function. The nonlinear parametric model form is known but the values of the parameters are unknown. We observe the ouput $y(t)$ of H_2 when the input $x(t)$ is an unknown zero mean i.i.d. random variable with a prespecified non gaussian distribution. We intend to recover (deconvolve) the input sequence or equivalently to identify the inverse $H_2^{-1} - \varphi^{-1} - H_1^{-1}$, supposing that H_1, H_2 and $\varphi(.)$ are invertible.

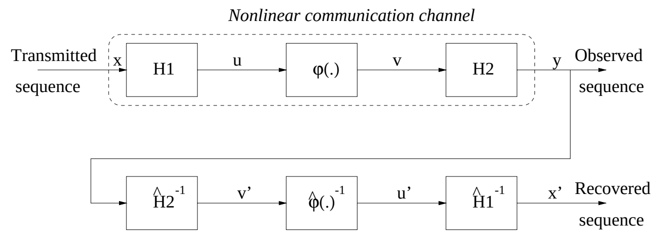

Figure 1: Blind deconvolution model

2. NONLINEAR COMMUNICATION CHANNEL MODEL

Consider the system of Fig.1. In order to facilitate the system inversion, the linear dynamic system is choosen as:

$$H(z) = \frac{1}{1 - a_1 z^{-1} - \ldots - a_m z^{-m}} \qquad (1)$$

where m is the memory length and $a_1, \ldots, a_m$ are the unknown parameters of the filter.

The static nonlinearity can take, for instance, a polynomial form for modeling an optional nonlinear encoder in the V-34 in a telephone channel [8, 6] or the Saleh's analytical model for modeling the nonlinearities of a travelling wave tubes in satellite communication [9, 5].
Here we shall restrict the form of the invertible static nonlinear function: a commonly used function for modeling saturation is a sigmoid function (see figure 2 for $\alpha = 1$, 2 and

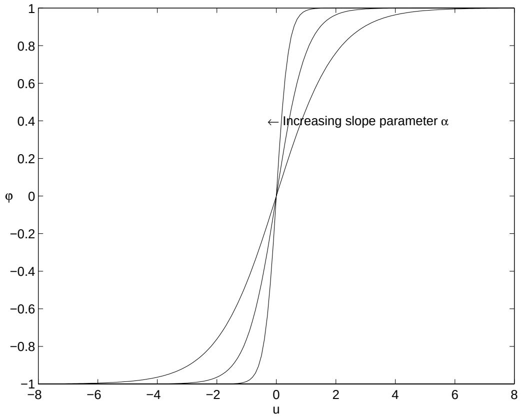

Figure 2: Sigmoid function

5) well known in the neural networks community [10] and defined as follow :

$$\varphi(u) = (\frac{2}{1+\exp(-\alpha u)} - 1)\beta\,. \qquad (2)$$

Note that $\varphi(u) = -\varphi(-u)$.

3. PROBLEM FORMULATION

3.1. Kurtosis maximization in the linear case

It is known that linear prediction methods based on second order statistics are not sufficient for solving the problem of blind deconvolution of scalar linear systems. For that reason, the problem cannot be solved when the input process is Gaussian. Consequently, approaches for blind deconvolution based on high order moments have been proposed [4], [11], [12], [13], [14], [15].

In these methods, the linear filter is identified through the maximization of the kurtosis of the recovered sequence x' under the constant energy constraint $\sigma^2_{x'} = \sigma^2_x$ [4]. Then the probability distribution of the recovered sequence is equal to the probability distribution of the input sequence.

We assume the existence of the moments of $x(t)$ up to the fourth order and we suppose that $K_x \neq 0$, where K_x is the kurtosis associated with $x(t)$ defined by :

$$K_x \triangleq \frac{E[x^4]}{(\sigma_x^2)^2} - 3 \qquad (3)$$

where $E[.]$ stands for the expectation operation.

Blind deconvolution can be perform if

1. $\sigma^2_x = \sigma^2_{x'}$ (energy constraint),
2. $K_x = K_{x'}$, or equivalently $|K_{x'}|$ maximum,
3. there is no measurement noise.

In [4] it is shown that $|K_{x'}|$ has no spurious local maxima under the energy constraint. Therefore, a gradient-search algorithm is expected to converge to the desired response.

3.2. Nonlinear case

This approach based on kurtosis maximization developed by [3] and [4] is efficient in the case of linear systems (channels). We intend to illustrate that it can also be applied to nonlinear communication channel at least in simple cases. We show, with an illustrative example (with and without noise measurement), that the kurtosis of the recovered sequence presents a maximum when the system parameters are correctly estimated. Consequently, kurtosis maximization can be used for this kind of nonlinear system identification.

To the best of our knowledge, this approach has never been used for this kind of nonlinear system analysis.

In order to perform the recovery of the input sequence, the following procedure is used (presently this procedure is not yet automatic):
$K_{x'}$ depends on the parameters of the linear filters H_1 and H_2 and also on α and β. In a first step we choose arbitrarily the parameters of the filter H_2. Then we deduce $\hat{\beta}$ in order to ensure that the static sigmoid function (2) is invertible and we estimate $\alpha\ u(t)$ (it is not necessary to estimate the gain α). Finally the paramaters of H_1 are estimated by maximization of the kurtosis $|K_{x'}|$. This procedure is repeated for a new set of parameters of H_2 until $|K_{x'}|$ reaches its maximum. We do not guarantee that this procedure yields an absolute maximum.

4. SIMULATIONS

4.1. Kurtosis maximization

An illustrative example is proposed as follow :

The input sequence is an i.i.d. uniforme sequence (± 1) with $L = 1000$ symbols . In this case the kurtosis ($K_x = -2$) is negative and we maximize $-K_{x'}$. This maximization is performed under constant energy constraint.

The two linear filters are second order transfer function

$$H_i = \frac{1}{1 - \tau_{i1}\, z^{-1} - \tau_{i2}\, z^{-2}}\,, \qquad (4)$$

with $\tau_{11} = 0.4$, $\tau_{12} = 0.2$ and $\tau_{21} = 0.6$, $\tau_{22} = 0.3$. The nonlinear parameters are $\alpha = 5$ and $\beta = 1$ (see figure 1).

We shall quantify the results of our method with a common measure of identification error which is the normalized mean square error Er defined by

$$Er = \frac{\sum_{i=1}^{L}(x'_i - x_i)^2}{\sum_{i=1}^{L} x_i^2}. \tag{5}$$

The procedure of section 3.2 is used to estimated the parameters of the structure. Results are reported in Table 1. Each row in Table 1 gives the results of the different steps of our kurtosis maximization procedure. The parameters $\hat{\tau}_{21}$ and $\hat{\tau}_{22}$ are chosen, then we estimate $\hat{\beta}$, $\hat{\tau}_{11}$ and $\hat{\tau}_{12}$ minimizing $K(\hat{x})$. The value of $K(\hat{x})$ and of the corresponding $Er\ (dB)$ are given column 6 and 7.

Fixed		Estimated			Evaluated	
$\hat{\tau}_{21}$	$\hat{\tau}_{22}$	$\hat{\beta}$	$\hat{\tau}_{11}$	$\hat{\tau}_{12}$	$K(\hat{x})$	$Er\ (dB)$
0.55	0.25	0.57	0.26	0.28	-1.7119	-9.67
0.58	0.28	0.76	0.21	0.21	-1.7889	-10.64
0.6	0.3	1	0.4	0.2	-2	-246.58
0.62	0.32	0.81	0.14	0.06	-1.6609	-9.20
0.65	0.35	0.6	0.23	0.15	-1.3292	-7.87

Table 1: Estimation of τ_{11} and τ_{12}.

We see that the kurtosis is minimum when $\hat{\tau}_{21}$ and $\hat{\tau}_{22}$ are chosen properly.

Figure 3 represents $K_{x'}(\hat{\tau}_{21}, \hat{\tau}_{22}, \hat{\tau}_{11}, \hat{\tau}_{12}, \beta)$ when $\hat{\tau}_{21} = 0.6$, $\hat{\tau}_{22} = 0.3$, $\hat{\tau}_{12} = 0.2$, $b\hat{e}ta = 0.9$, 0.99 or 1. We note that the kurtosis is minimum when $b\hat{e}ta = 1$ and $\hat{\tau}_{11} = 0.4$.

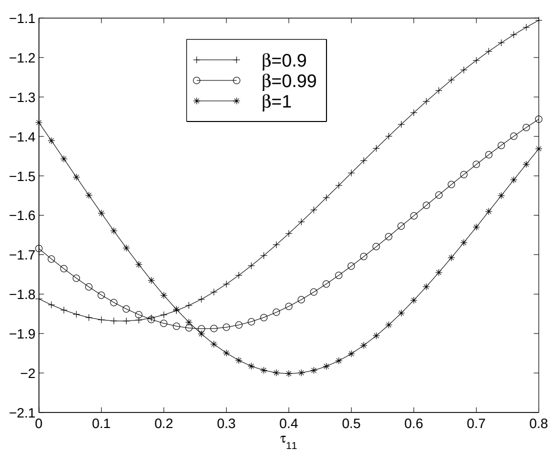

Figure 3: Kurtosis minimization : $\hat{\tau}_{11}$

Figure 4 represents $K_{x'}(\hat{\tau}_{21}, \hat{\tau}_{22}, \hat{\tau}_{11}, \hat{\tau}_{12}, \beta)$ when $\hat{\tau}_{21} = 0.6$, $\hat{\tau}_{22} = 0.3$, $\beta = 1$ and $\hat{\tau}_{11}$ varies from 0.3 to 0.5 with step 0.02. The maximum of $-K_{x'}$ is obtained when the estimation of the parameters yields the actual values of the system parameters.

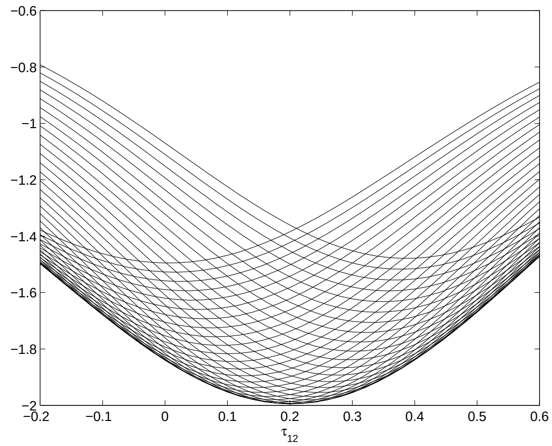

Figure 4: Kurtosis minimization : $\hat{\tau}_{12}$

4.2. Effect of additive noise

So far, we have ignored the presence of additive noise in the system. In practice, this assumption is unrealistic. We propose to study the effect of an additive noise in the structure as shown in figure 5, where the noise b is assume to be a white Gaussian noise $\mathcal{N}(0, \sigma_b^2)$.

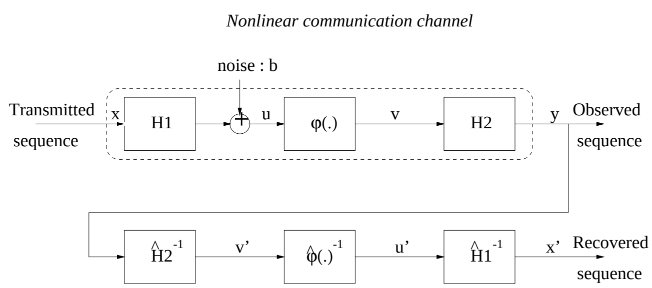

Figure 5: Blind deconvolution model with noise perturbations.

The influence of the additive noise, for different values of σ_b on the performances of the method are presented in Table 2, when $\hat{\tau}_{21} = 0.6$, $\hat{\tau}_{22} = 0.3$ and $\hat{\beta} = 1$. The signal to noise ratio (SNR) is defined by:

$$SNR_{dB} = 10\ \log(\frac{\sigma_x^2}{\sigma_b^2})\,, \tag{6}$$

The performances of this approach increase linearly with the SNR. For a high $SNR \geq 20$ the additive noise has no significant influence on the quality of the recovered sequence.

Fixed		Estimated		Evaluated	
SNR	σ_b	$\hat{\tau}_{11}$	$\hat{\tau}_{12}$	$K(\hat{x})$	$Er\ (dB)$
30	0.031	0.4	0.2	-1.9968	-29.05
25	0.056	0.4	0.2	-1.9860	-24.05
20	0.100	0.4	0.2	-1.9524	-19.06
15	0.178	0.39	0.2	-1.8509	-14.13
10	0.316	0.37	0.2	-1.5748	-9.36
5	0.562	0.3	0.2	-1.0191	-5.04

Table 2: Estimation of τ_{11} and τ_{12} with noise perturbations.

5. TOWARDS A THEORETICAL JUSTIFICATION OF THE PROPOSED METHOD

The proposed method is not yet based on a sound theoretical development, merely on a simulation. However, we may propose the following interpretation.
We suppose that the function φ is close to a linear function

$$\varphi(u) = u - \epsilon(u)\,.$$

If the parameters of the linear filters are accurately estimated while the parameters of φ^{-1} are represented by a linear term

$$\varphi^{-1}(v') = v'\,,$$

the recovered signal $x'(t)$ will be close to the true signal $x(t)$. $x'(t)$ can be expressed as a sum $x'(t) = x(t) + e(t)$ where $e(t)$ is not linearly dependant of $x(t)$ only. The term $e(t)$ in the sum is due to the presence of the filter H_1: $e(t)$ depends on the past or the future of $x(t)$ and not only on $x(t)$. The presence of H_1 (Wiener model) is essential in the identification process. As $x(t)$ is a white noise, the kurtosis of $x'(t)$ will be lower (in absolute value) than the kurtosis of $x(t)$, since $x'(t)$ will appear as a mixture of a white noise and another signal that does not depend only on it: the kurtosis of the mixture is necessary lower than the kurtosis of one of its components [16]. The appropriate choice of the parameters of the nonlinear function will force to zero the amplitude of $e(t)$. Consequently, this procedure will maximise the kurtosis of the estimated output.

6. CONCLUSION

We have proposed to use kurtosis maximization of the recovered sequence for blind identification of a nonlinear communication channel. The model of the structure is composed by a linear filter, a nonlinear static function and a linear filter, which is an extention of the linear case. We briefly address the problem of additive noise in the system and that when the this noise is Gaussian the proposed approach remains usable for a $SNR \geq 20\ dB$.

7. REFERENCES

[1] O. Shalvi and E. Weinstein. ”Super-exponential methods for blind deconvolution”. *SPIE Adaptive signal Processing*, 1565:143–152, 1991.

[2] R. A. Wiggins. ”Minimum entropy deconvolution”. *Geoexploration*, 16:21–35, 1978.

[3] D. L. Donoho. *”On minimum entropy deconvolution”*. in D. F. Findley, Ed. Applied Time Series Analysis II., N.Y. Academic press, pp. 564-608, 1981.

[4] O. Shalvi and E. Weinstein. ”New criteria for blind deconvolution of nonminimum phase systems (channels)”. *IEEE Transactions on Information Theory*, 36(2):312–321, March 1990.

[5] W. Franck and U. Appel. ”Efficient equalization of nonlinear communication channels”. *Proceedings of Icassp’97, Munich, Germany*, 3:2501–2504, 1997.

[6] M. Ibnkahla, J. Sombrin, F. Castanié, and N. J. Bershad. ”Neural networks for modeling nonlinear memoryless communication channels”. *TRANSCOM*, 45(7), 1997.

[7] H. Chen, N. Ishii, and N. Suzumura. ”Structural classification of nonlinear systems by input and output measurements”. *Int. J. Syst. Sci.*, 17(5):741–774, 1986.

[8] A. Saleh. ”Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers”. *IEEE Transactions on Communications*, Com-29(11), 1981.

[9] ”Data communication over the telephone network”. *ITU-T Recommendation V.34*, 1994.

[10] S. Haykin. *”Adaptive Filter Theory”*. Third Edition, Prentice Hall, Inc, A Simon & Schuster Company, New Jersey, 1996.

[11] A. Benveniste, M. Goursat, and G. Ruget. ”Robust identification of a non-minimum phase system: blind adjustment of a linear equalizer in data communication”. *IEEE Transactions on Automatic Control*, 25(3):385–399, 1980.

[12] C. Huet and J. LeRoux. ”Blind system identification using fourth order spectral analysis of complex signals”. *4th IEEE Signal Processing Workshop on HOS, Banff, Canada*, pages 189–193, July 1997.

[13] G. Giannakis and J.M. Mendel. ”Identification of nonminimum phase systems using higher order statistics”. *IEEE Trans. on ASSP*, 37(3):360–377, March 1989.

[14] J. M. Mendel. ”Tutorial on higher order statistics (spectra) in signal processing and system theory: theoretical results and some applications”. *Proceedings of the IEEE*, 79(3):278–305, March 1991.

[15] R. Pan and C. L. Nikias. “The complex cepstrum of higher order cumulants and non-minimum phase system identification”. *IEEE trans. on Signal Processing*, ASSP-36(2), Feb. 1988.

[16] J. Tugnait. ”Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria”. *IEEE Trans. on Sig. Proc.*, 45(3):658–672, March 1997.