

PARAMETRIC APPROACH TO BLIND DECONVOLUTION OF NONLINEAR CHANNELS

Jordi Solé i Casals (*), Anisse Taleb (**), Christian Jutten (**)

(*) University of Vic, Sagrada Família 7, 08500, Vic (Catalunya, Spain)

(**) INPG, LIS, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France

Abstract

A parametric procedure for the blind inversion of nonlinear channels is proposed, based on a recent method of blind source separation in nonlinear mixtures. Experiments show that the proposed algorithms perform efficiently, even in the presence of hard distortion. The method, based on the minimisation of the output mutual information, needs the knowledge of log-derivative of input distribution (the so-called *score function*). Each algorithm consists of three adaptive blocks: one devoted to adaptive estimation of the score function, and two other blocks estimating the inverses of the linear and nonlinear parts of the channel, (quasi-)optimally adapted using the estimated score functions. This paper is mainly concerned by the nonlinear part, for which we propose two parametric models, the first based on a polynomial model and the second on a neural network, while [12, 13] proposed nonparametric approaches.

1 Introduction

When linear models fail, nonlinear models appear to be powerful tools for modelling practical situations. Many researches have been done in the identification and/or the inversion of nonlinear systems. These assume that both the input and the output of the distortion are available, and are based on higher-order input/output cross-correlation [2] or on the application of the Bussgang and Princes theorems [3,8] for nonlinear systems with Gaussian inputs. However, in a real world situation, one often does not have access to the distortion input. In this case, the blind identification of the nonlinearity becomes the only way to solve the problem. This paper is concerned by a particular class of nonlinear systems, composed by a linear subsystem followed by a memoryless nonlinear distortion (Figure 1). This class of nonlinear systems, also known as Wiener systems, is a nice and mathematically attracting model, but also a actual model used in various areas, such as biology [6], industry [1], sociology and psychology (see also [7] and the references therein). Despite its interest, today, there does not exist completely blind procedure for inverting such systems.

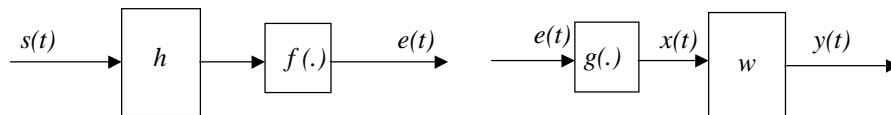


Figure 1: The unknown nonlinear convolution system (left) and the proposed inversion structure (right).

2 Model and assumptions

We suppose that the input of the system $S=\{s(t)\}$ is an unknown non-Gaussian independent and identically distributed (i.i.d.) process, and that subsystems h, f are respectively a linear filter and a memoryless nonlinear function, both unknown and invertible. We would like to estimate $s(t)$ by only observing the system's output. This implies the blind estimation of the inverse structure (Figure 1), composed of similar subsystems: a memoryless nonlinear function g and a linear filter w . Such a system is known as a Hammerstein system. Let s and e be the vectors of infinite dimension defined from the processes $S=\{s(t)\}$ and $E=\{e(t)\}$ respectively, whose t -th entry is $s(t)$ or $e(t)$. The unknown input-output transfer can be written as:

$$e = f(Hs) \quad (1)$$

where:

$$H = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & h(t+1) & h(t) & h(t-1) & \dots \\ \dots & h(t+2) & h(t+1) & h(t) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (2)$$

is a Toeplitz matrix of infinite dimension which represents the action of the filter h to the signal $s(t)$. The matrix H is non-singular provided that the filter h is invertible, i.e. satisfies $h^{-1} * h = h * h^{-1} = \delta_0$, where δ_0 is the Dirac impulse at $t = 0$.

Equation (1) corresponds to a postnonlinear (pnl) model [12]. This model has been recently studied in nonlinear source separation, but only for a finite dimensional case. In fact, with the above parametrization, the i.i.d. nature of $s(t)$ implies the spatial independence of the components of the infinite vector s . Similarly, the output of the inversion structure can be written $y = Wx$ with $x(t) = g(e(t))$. Following [12,13] the inverse system (g, w) can be estimated by minimizing the output mutual information, i.e. spatial independence of y which is equivalent to the i.i.d. nature of $y(t)$.

3 Cost function

The mutual information of a random vector of dimension n , defined by:

$$I(z) = \sum_{i=1}^n H(z_i) - H(z_1, z_2, \dots, z_n) \quad (3)$$

is extended to a vector of infinite dimension, using the notion of *entropy rates* of stationary stochastic process [4]:

$$I(Z) = \lim_{T \rightarrow \infty} \frac{1}{2T+1} \left\{ \sum_{t=-T}^T H(z(t)) - H(z(-T), \dots, z(T)) \right\} = H(z(\tau)) - H(Z) \quad (4)$$

where τ is arbitrary due to the stationarity assumption. We shall notice that $I(Z)$ is always positive and vanishes iff Z is i.i.d. Now, since S is stationary, and since h and w are time-invariant filters, then Y is also stationary, and $I(Y)$ is defined by:

$$I(Y) = H(y(\tau)) - H(Y) \quad (5)$$

Using the Lemma 1 of [13], equation (5) becomes:

$$H(Y) = H(X) + \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} w(t) e^{-j\theta t} \right| d\theta \quad (6)$$

Moreover, since $x(t) = g(e(t))$ and $E = e(t)$ is stationary:

$$H(X) = \lim_{T \rightarrow \infty} \frac{1}{2T+1} \left\{ H(e(-T), \dots, e(T)) + \sum_{t=-T}^T E[\log g'(e(t))] \right\} = H[E] + E[\log g'(e(\tau))] \quad (7)$$

Combining (6) and (7) in (5) leads finally to:

$$I(Y) = H(y(\tau)) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{+\infty} w(t) e^{-jt\theta} \right| d\theta - E[\log g'(e(\tau))] - H[E] \quad (8)$$

4 Theoretical derivation of the inversion algorithm

To derive the optimization algorithm, we need the derivatives of $I(Y)$ (8) with respect to the parameters of the linear part w and of the nonlinear function g .

4.1 Linear subsystem

The linear subsystem is well parameterised by the coefficients of the filter w . The derivative of $I(Y)$ with respect to the coefficient $w(t)$, corresponding to the t -th lag, is easily computed:

$$\frac{\partial I(Y)}{\partial w(t)} = -E[x(\tau-t) \psi_{y(\tau)}(y(\tau))] - \bar{w}(-t) \quad (9)$$

It leads to the following gradient descent algorithm (see [13] for details):

$$w \leftarrow w + \mu \left\{ \psi_{y, \psi_y(y)} + \delta \right\}^* w \quad (10)$$

4.2 Nonlinear subsystem

For this subsystem, we propose two different models for the function g . Both methods are parametric, the first is based on a polynomial model of g , and the second on a multilayer perceptrons.

4.2.1 Polynomial parameterization

Modelling g with a N -degree polynomial leads to:

$$g(t) = \sum_{n=1}^{N+1} a_n t^{n-1} \quad (11)$$

The gradient descent algorithm for g requires the derivatives of $I(Y)$ (8) with respect to the coefficients a_n of the polynomial. The derivatives of the right-side terms of (5) are:

$$\frac{\partial}{\partial a_p} H(y(\tau)) = -E \left[\psi_y(y(\tau)) \sum_{t=-\infty}^{+\infty} g(t) e(\tau-t)^{p-1} \right] \quad (12)$$

and

$$\frac{\partial H(Y)}{\partial a_p} = E \left[\left(\sum_{n=1}^{N+1} a_n (n-1) e(\tau)^{n-2} \right)^{-1} (p-1) e(\tau)^{p-2} \right] \quad (13)$$

Combining (12) and (13) leads to:

$$\frac{\partial I(Y)}{\partial a_p} = -E \left[\psi_y(y(\tau)) \sum_{t=-\infty}^{+\infty} w(t) e(\tau-t)^{p-1} \right] - E \left[\left(\sum_{n=1}^{N+1} (n-1) a_n e(\tau)^{n-2} \right)^{-1} (p-1) e(\tau)^{p-2} \right] \quad (14)$$

Equation (14) is the gradient of $I(Y)$ with respect to polynomial coefficients a_n , and will be used to provide a gradient descent algorithm for estimating g .

4.2.2 Neural network parametrization

In this subsection, we model g using a multilayer perceptron with one hidden layer:

$$g[\mathbf{m}, \mathbf{n}, \boldsymbol{\theta}, u] = \sum_{i=1}^N m_i \sigma(n_i u - \theta_i) \quad (15)$$

The gradient descent algorithm of g requires the derivatives of $I(Y)$ (8) with respect to the network coefficients, \mathbf{m} , \mathbf{n} and $\boldsymbol{\theta}$:

$$\frac{\partial I(Y)}{\partial m_p} = -E \left[\frac{n_p \sigma'(n_p u(t) - \theta_p)}{\sum_{i=1}^N m_i n_i \sigma'(n_i u(t) - \theta_i)} \right] - E \left[\psi_y(y(\tau)) \sum_k w_k \sigma(n_p u(t-k) - \theta_p) \right] \quad (16)$$

$$\begin{aligned} \frac{\partial I(Y)}{\partial n_p} = & -E \left[\frac{m_p (u(t) n_p \sigma''(n_p u(t) - \theta_p) + \sigma'(n_p u(t) - \theta_p))}{\sum_{i=1}^N m_i n_i \sigma'(n_i u(t) - \theta_i)} \right] \\ & - E \left[\psi_y(y(\tau)) \sum_k w_k m_p u(t-k) \sigma'(n_p u(t-k) - \theta_p) \right] \end{aligned} \quad (17)$$

$$\frac{\partial I(Y)}{\partial \theta_p} = E \left[\frac{m_p n_p \sigma''(n_p u(t) - \theta_p)}{\sum_{i=1}^N m_i n_i \sigma'(n_i u(t) - \theta_i)} \right] + E \left[\psi_y(y(\tau)) \sum_k w_k m_p \sigma'(n_p u(t-k) - \theta_p) \right] \quad (18)$$

Equations (16, 17, 18) are the gradients of $I(Y)$ with respect to network coefficients and will be used for deriving a gradient descent algorithm of g .

5 Experimental results

In order to prove the efficiency of the previous algorithms, we simulate a difficult situation, where the input sequence $s(t)$ is an i.i.d. random sequence, filtered by a non-minimum phase FIR filter $h = [-0.082, 0, -0.1793, 0, 0.6579, 0, 0.1793, 0, -0.082]$ and followed by the nonlinear distortion $f(u) = 0.1u + \tanh(5u)$ (see figure 2). Note especially on Fig. 2-right, the saturation due to the function $\tanh(\cdot)$. On the frequency response of h , shown in figure 3, we can see that h a non-minimum phase filter.

The algorithm was trained with a sample size of $T = 1000$. The length of the impulse response of w is set to 21 with equal length for the causal and anti-causal parts. Estimation results, shown in figures 4 and 5, prove the efficiency of the two algorithms, with parametric, either polynomial or neural, models for estimating the inverse of nonlinear function f .

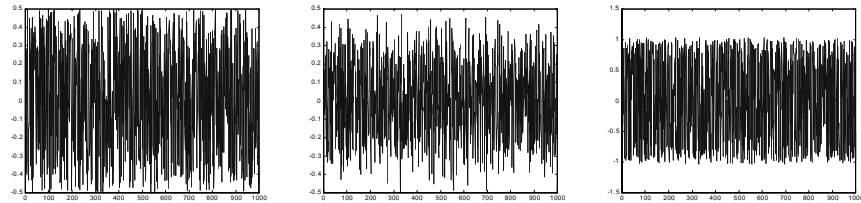


Figure 2. From left to right, input signal $s(t)$, filtered signal and distorted signal $e(t)$

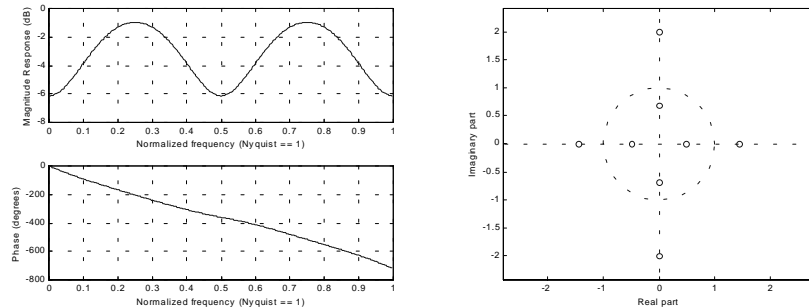


Figure 3. Frequency domain response and zeros of h .

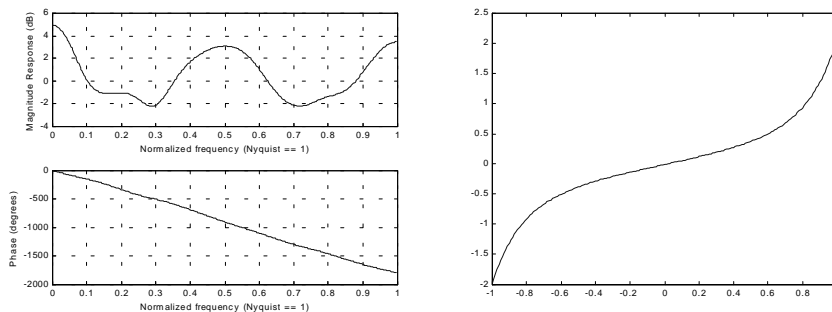


Figure 4. Frequency domain response of w , estimation of the inverse of h (left). Estimated inverse of f based on a 10-degree polynomial model (right).

The performance can be directly measured with $S/N = E[y^2(t)]/E[(s(t) - y(t))^2]$, the output S/N , where N is the error power and S is the estimated signal power. After adequate processing (delaying and re-scaling of $y(t)$), one obtains $S/N \approx +18dB$ with both polynomial and neural models.

6 Conclusions

In this paper, two blind parametric procedures for the inversion of a nonlinear channel (Wiener system) were proposed. Contrary to other blind identification procedures, the system input is not assumed to be Gaussian. If the input is not iid, but is a linear filtering of an iid noise (the so-called innovation), the output provides the innovation. The restitution of the input then requires more information, for instance the distribution or covariance matrix. Moreover, the nonlinear subsystem is unknown and cannot be directly estimated because its input is not available. The inversion procedure, in both cases, is based on the minimization of the mutual information rate of the inverse system output. The estimation of g is done according to a parametric model, using either a polynomial model or a neural network. Both models leads to good results, even in difficult situations (hard nonlinearity and non minimum phase filter).

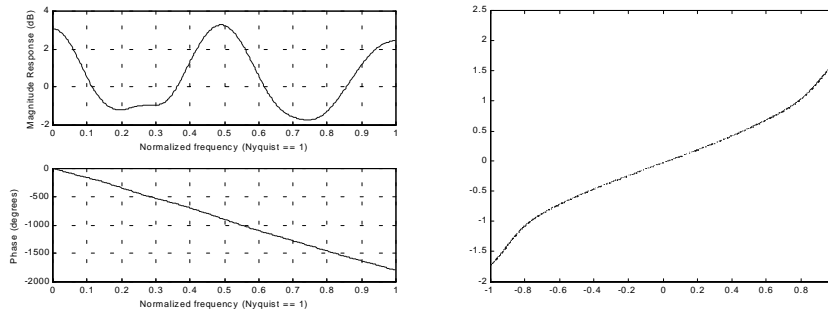


Figure 5. On the left, estimated inverse of h : w frequency domain response. On the right, estimated inverse of f using a multilayer perceptron with 6 neurons in the hidden layer: $x(t)$ vs. $e(t)$

Acknowledgment: This work has been in part supported by the Direcció General de Recerca de la Generalitat de Catalunya.

References

1. R. Bars, I. Bèzi, B. Pilipar, B. Ojhelyi. Nonlinear and long range control of a distillation pilot plant. In *Identification and Syst. Parameter Estimation. 9th IFAC/IFORS Symp.*, pp. 848-853, Budapest (1990)
2. S.A. Bellings, S.Y. Fakhouri. Identification of a class of nonlinear systems using correlation analysis. *Proc. IEEE*, 66 pp. 691-697 (1978)
3. E.D. Boer. Cross-correlation function of a bandpass nonlinear network. *Proc. IEEE*, 64 pp. 1443-1444 (1976)
4. T.M. Cover, J.A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications (1991)
5. W. Härdle. *Smoothing Techniques with implementation in S*. Springer-Verlag (1990)
6. I.W. Hunter. Frog muscle fiber dynamic stiffness determined using nonlinear system identification techniques. *Biophys. J.*, pp. 49-81a (1985)
7. I.W. Hunter, M.J. Korenberg. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol Cybern.*, 55, pp. 135-144 (1985)
8. G. Jacoviti, A. Neri, R. Cusani. Methods for estimating the autocorrelation function of complex stationary process. *IEEE Trans. ASSP*, 35, pp. 1126-1138 (1987)
9. C.L. Nikias, A.P. Petropulu. *Higher-Order Spectra Analysis – A Nonlinear Signal processing Framework*. Englewood Cliffs, NJ: Prentice-Hall (1993)
10. C.L. Nikias, M.R. Raghuveer. Bispectrum estimation: A digital signal processing framework. *Proc. IEEE*, 75 pp. 869-890 (1987)
11. S. Prakriya, D. Hatzinakos. Blind identification of LTI-ZMNL-LTI nonlinear channel models. *Biol. Cybern.*, 55 pp. 135-144 (1985)
12. A. Taleb, C. Jutten. Source separation in postnonlinear mixtures. *IEEE Trans. S.P.*, Vol. 47, n°10, pp.2807-20 (1999).
13. A. Taleb, J. Solé, C. Jutten. Quasy-Nonparametric Blind Inversion of Wiener Systems. *Submitted to IEEE Trans. S.P.*