

Post-Nonlinear Mixtures and Beyond

Jordi Solé-Casals^{1,2} and Christian Jutten²

¹ Signal Processing Group, University of Vic, Sagrada Família 7, 08500, Vic, Spain
jordi.sole@uvic.es

<http://www.uvic.es/eps/recerca/ca/processament/inici.html>

² LIS-(CNRS UMR n°5083), INPG, 46 Av. F.Viallet-38031 Grenoble Cedex, France
Christian.Jutten@inpg.fr

<http://www.lis.inpg.fr/pages.perso/jutten/jutten.htm>

Abstract. Although sources in general nonlinear mixtures are not separable using only statistical independence, a special and realistic case of nonlinear mixtures, the post nonlinear (PNL) mixture is separable choosing a suited separating system. Then, a natural approach is based on the estimation of the separating system parameters by minimizing an independence criterion, like estimated source mutual information. This class of methods requires higher (than 2) order statistics, and cannot separate Gaussian sources. However, use of (weak) prior, like source temporal correlation or nonstationarity, leads to other source separation algorithms, which are able to separate Gaussian sources, and can even, for a few of them, works with second-order statistics. Recently, modeling time correlated sources by Markov models, we propose very efficient algorithms based on minimization of the conditional mutual information. Currently, using the prior of temporally correlated sources, we investigate the feasibility of inverting PNL mixtures with non-bijective non-linearities, like quadratic functions. In this paper, we review the main ICA and BSS results for nonlinear mixtures, present PNL models and algorithms, and finish with advanced results using temporally correlated sources.

1 The nonlinear ICA and BSS problems

1.1 Model and problem

Consider N samples of the m -dimension observed random vector \mathbf{x} , modeled by

$$\mathbf{x} = \mathcal{F}(\mathbf{s}) + \mathbf{n} \quad (1)$$

where \mathcal{F} is an unknown mixing mapping assumed invertible, \mathbf{s} is an unknown n -dimensional source vector containing the source signals s_1, s_2, \dots, s_n , which are assumed to be statistically independent, and \mathbf{n} is an additive noise, independent of the sources.

Such a model is usual in multidimensional signal processing, where each sensor receives an unknown superimposition of unknown source signals at time instants $t = 1, \dots, N$. Then, the goal is to recover the n unknown actual source signals $s_j(t)$ which have given rise to the observed mixtures. This is referred to as the blind source separation (BSS) problem, *blind* since no or very little prior information about the sources is required. Since the only assumption is the independence of sources, the basic idea in blind source separation consists in estimating a mapping \mathcal{G} , only from the observed data \mathbf{x} , such that $\mathbf{y} = \mathcal{G}(\mathbf{x})$ are statistically independent. The method, based on statistical independence, constitutes a generic approach called independent component analysis (ICA).

In the following, we assume that there are as many mixtures as sources ($m = n$) and that noise is zero. In that case (as well as if $m > n$), it is clear that estimating the inverse mapping \mathcal{G} directly provides the sources. On the contrary, if $m < n$, identification of \mathcal{G} and source estimation are unrelated tasks, and extra priors are required for separating the sources.

1.2 Nonlinear mixtures

The general nonlinear ICA problem then consists in estimating a mapping $\mathcal{G} : (R)^n \rightarrow (R)^n$ that yields components

$$\mathbf{y} = \mathcal{G}(\mathbf{x}) \quad (2)$$

which are statistically independent, only using the observations \mathbf{x} .

For general nonlinear mappings, $\mathcal{G} \circ \mathcal{F}$ can lead to independent signals \mathbf{y} which are still mixtures of the independent sources \mathbf{x} . In other words, ICA and BSS are not equivalent: one can easily design a nonlinear mapping which mixes the sources and provides statistically independent variables y_i

$$y_i = h_i(\mathbf{s}) \neq h_i(s_{\sigma(i)}) \quad (3)$$

where $\sigma(i)$ is a permutation over $\{1, 2, \dots, n\}$

Moreover, if the separation would be achieved, each estimated output y_i would only depend on a unique source $y_i = h_i(s_{\sigma(i)})$. Then, strong distortions could still occur, due to the mapping $h_i(\cdot)$. One reason for this is that if u and v are two independent random variables, any of their functions $f(u)$ and $g(v)$ (where f and g are invertible functions) are independent too.

In the nonlinear BSS problem, one then expect to find signals $y_i = h_i(s_{\sigma(i)})$, $i = 1, \dots, n$. Of course, it would be nice to restrict h_i to identity function: then, sources would be recovered up to weaker indeterminacies, *e.g.* scale and permutation indeterminacies as in linear mixtures. But is it possible? In other words, can priors on the sources and/or the mixing mapping be sufficient for this? Generally, using ICA for solving the nonlinear BSS problem requires additional prior informations or suitable regularizing constraints.

1.3 Outline

This paper is organized as follows. After having presented the nonlinear ICA and BSS problems in section 1, we consider two ways for regularizing the problem of source separation in nonlinear mixtures. The first way, based on structural constraints, is explained in Section 2. The second way, based on prior information on the sources, is presented in Section 3. Finally, new ideas for inverting non-bijective functions are investigated in Section 4.

2 Structural constraints

2.1 Linear models

In the case of regular linear models, the mapping \mathcal{F} is linear and can be represented by $\mathbf{x} = \mathbf{A}\mathbf{s}$ where \mathbf{A} is a square invertible matrix. In this case it suffices to constrain the separating model \mathcal{G} to lie in the subspace of invertible square matrices, and one has to estimate a matrix \mathbf{B} such that $\mathbf{y} = \mathbf{B}\mathbf{x} = \mathbf{H}\mathbf{s}$ has independent components. First theoretical results on separability of linear mixtures have been proposed by Comon in 1994 [6].

2.2 PNL mixtures

PNL mixing and separating models In the post-nonlinear (PNL) model, the nonlinear observations have the following specific form (Figure 1):

$$x_i(t) = f_i\left(\sum_{j=1}^n a_{ij}s_j(t)\right), \quad i = 1, \dots, n \quad (4)$$

One can see that the PNL model consists of a linear mixture followed by a component-wise nonlinearity f_i acting on each output independently from the others. The nonlinear functions (distortions) f_i are assumed to be invertible.

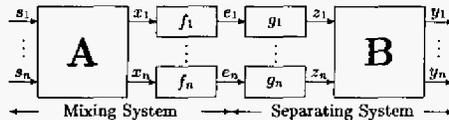


Fig. 1. The mixing-separating system for PNL mixtures.

Besides its theoretical interest, this model belonging to the L-ZMNL³ family suits perfectly for a lot of real-world applications. For instance, such models appear in sensors array processing [14], satellite and microwave communications [17], and in many biological systems [10].

The simplest choice for the separating system \mathcal{G} is the mirror structure of the mixing system \mathcal{F} (see Figure 1). In [22], it has been shown that these mixtures are separable for distributions having at most one Gaussian source, with the same indeterminacies as linear mixtures.

PNL algorithms In [22], the mutual information

$$I(\mathbf{y}) = \int p_{\mathbf{y}}(\mathbf{y}(t)) \log \frac{p_{\mathbf{y}}(\mathbf{y}(t))}{\prod_{i=1}^n p_{y_i}(y_i(t))} d\mathbf{y} \quad (5)$$

between the components y_1, \dots, y_n of the output vector is used as the independence criterion in both linear and non-linear stages. For the linear part, minimization of the mutual information leads to the same estimation equations as for linear mixtures [8, 5]

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{B}} = -\mathbb{E}\{\boldsymbol{\psi}\mathbf{x}^T\} - (\mathbf{B}^T)^{-1} \quad (6)$$

where components ψ_i of the vector $\boldsymbol{\psi}$ are score functions of the components y_i of the vector \mathbf{y} :

$$\psi_i(u) = \frac{d}{du} \log p_i(u) = \frac{p_i'(u)}{p_i(u)} \quad (7)$$

Here $p_i(u)$ is the pdf of y_i and $p_i'(u)$ its derivative.

For the nonlinear stage, one can derive from the estimating equations the gradient rule [22]

$$\frac{\partial I(\mathbf{y})}{\partial \boldsymbol{\theta}_k} = -\mathbb{E} \left\{ \frac{\partial \log |g'_k(\boldsymbol{\theta}_k, x_k)|}{\partial \boldsymbol{\theta}_k} \right\} - \mathbb{E} \left\{ \sum_{i=1}^n \psi_i(y_i) b_{ik} \frac{\partial g_k(\boldsymbol{\theta}_k, x_k)}{\partial \boldsymbol{\theta}_k} \right\} \quad (8)$$

Here x_k is the k th component of the observation vector, b_{ik} is the element ik of the separating matrix \mathbf{B} , and g'_k is the derivative of the k -th nonlinear function g_k . The exact computation algorithm depends on the specific parametric form of the nonlinear mapping $g_k(\boldsymbol{\theta}_k, x_k)$. In [22], a multilayer perceptron network is used for modeling the functions $g_k(\boldsymbol{\theta}_k, x_k)$, $k = 1, \dots, n$.

Contrary to BSS of linear mixtures, separation performance for nonlinear mixtures is strongly related to the estimation accuracy of the score functions (7) [22]. The score functions (7) must be estimated adaptively from the output vector \mathbf{y} . Several alternative ways to do this are considered in [22]. The first approach is to estimate the pdf, and then compute using differentiation the score function. Pdf estimation based on the Gram-Charlier expansion [6, 8] fails except for mild post-nonlinear distortions. For hard nonlinearities, a pdf estimation based on kernel methods is preferable. The second method estimates the score functions directly, and provides very good results for hard

³ L stands for Linear and ZMNL stands for Zero-Memory NonLinearity: it is a separable model with a linear stage followed by a nonlinear (static) distortion.

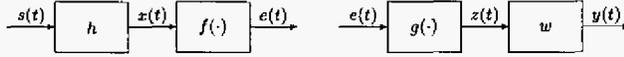


Fig. 2. A Wiener system consists of a filter followed by a distortion (left). For inverting the Wiener system, one uses a Hammerstein system (right), which consists of a distortion followed by a linear filter.

nonlinearities, too. A well performing batch type method for estimating the score functions has been introduced in a later paper [21]. Several other authors have studied methods for blind separation of post-nonlinear mixtures starting from different viewpoints : other parameterizations of non-linear mappings [13, 1, 15, 19], temporal decorrelation of sources [24], geometrical approaches [18, 3], and some tricks to improve the algorithms [20].

2.3 CPNL mixtures

Separability of PNL mixtures can be generalized to convolutive PNL (CPNL) mixtures, in which the instantaneous mixture (matrix \mathbf{A}) is replaced by linear filters (matrix of filters $\mathbf{A}(z)$), where each source is independent and identically distributed (iid) [2]. In fact, denoting $\mathbf{A}(z) = \sum_k \mathbf{A}_k z^{-k}$, and defining $\mathbf{s} \triangleq (\dots, s^T(k-1), s^T(k), s^T(k+1), \dots)^T$ and $\mathbf{e} \triangleq (\dots, e^T(k-1), e^T(k), e^T(k+1), \dots)^T$, we have:

$$\mathbf{e} = \mathbf{f}(\bar{\mathbf{A}}\mathbf{s}) \quad (9)$$

where \mathbf{f} acts componentwise, and:

$$\bar{\mathbf{A}} = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & \mathbf{A}_{k+1} & \mathbf{A}_k & \mathbf{A}_{k-1} & \dots \\ \dots & \mathbf{A}_{k+2} & \mathbf{A}_{k+1} & \mathbf{A}_k & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (10)$$

The iid nature of the source samples, *i.e.* the temporal independence of $s_i(k)$, $i = 1, \dots, n$, insures the spatial independence of \mathbf{s} . Then, the CPNL mixtures can be viewed as a particular PNL mixtures. For FIR mixing matrix $\mathbf{A}(z)$, (9) corresponds to a finite dimension PNL mixture and the separability holds. For more general filter (IIR) matrix, (9) is an infinite dimension PNL mixture, and the separability can be conjectured.

Algorithms for separating sources in CPNL mixtures are based on random processes (instead of random variables) independence, which leads to very tricky criteria. Practically, one can use simplified criterion like $J = \sum_{i,j,k} I(y_i(n)y_j(n-k))$, which demands high computation cost, but can still be simplified.

2.4 Wiener systems

With a suitable parameterization, it can be easily shown that the problem of blind inversion of Wiener systems (Fig. 2) is equivalent to the source separation problem in PNL mixtures [23]. Its output writes as

$$e(t) = f\left(\sum_k h(k)s(t-k)\right) \quad (11)$$

where $s(t)$ is the independent and identically distributed (iid) input, $e(t)$ is the observation, $h(k)$ denotes the entries of the unknown filter H , assumed invertible, and f is the unknown nonlinear mapping, assumed invertible and memoryless.

3 Time correlated sources in nonlinear mixtures

In this section we show that prior information on the sources can simplify the indeterminacies. Specifically, we exploit the temporal correlation of the sources. Each source $s_i(t)$, $i = 1, \dots, n$ is assumed to be temporally correlated, and is modeled by a q -order Markov model, i.e. :

$$p_{s_i}(s_i(t)|s_i(t-1), \dots, s_i(1)) = p_{s_i}(s_i(t)|s_i(t-1), \dots, s_i(t-q)) \quad (12)$$

where p_{s_i} denotes the pdf of the random variable s_i .

Since output independence leads to source separation, a possible approach for separating source is to consider a criterion measuring the independence of the output \mathbf{y} . Following [16, 7], one can use the conditional mutual information of \mathbf{y} , denoted by I :

$$I = \int p_{\mathbf{y}}(\mathbf{y}(t)|\mathbf{y}(t-1), \dots, \mathbf{y}(t-q)) \log \frac{p_{\mathbf{y}}(\mathbf{y}(t)|\mathbf{y}(t-1), \dots, \mathbf{y}(t-q))}{\prod_{i=1}^n p_{y_i}(y_i(t)|y_i(t-1), \dots, y_i(t-q))} d\mathbf{y} \quad (13)$$

which is always nonnegative, and zero if and only if the variables $w_i(t) = y_i(t)|y_i(t-1), \dots, y_i(t-q)$ are statistically independent for $i = 1, \dots, n$, i.e. the signals $y_i(t)$, $i = 1, \dots, n$ are independent Markovian process.

Considering the separation structure (Fig. 1), where $\mathbf{y}(t) = \mathbf{B}\mathbf{z}(t)$ and $z_i(t) = g_i(\theta_i, x_i(t))$,⁴ the outputs can be estimated by minimizing:

$$J(\mathbf{B}, \Theta) = - \sum_{i=1}^n E \left[\log \left| \frac{\partial g_i(\theta_i, x_i(t))}{\partial x_i(t)} \right| \right] - \log |\det(\mathbf{B})| - \sum_{i=1}^n E[\log p_{y_i}(y_i(t)|y_i(t-1), \dots, y_i(t-q))] \quad (14)$$

In practice, under the ergodicity conditions, the mathematical expectation (14) can be estimated by a time averaging, denoted $\hat{J}(\mathbf{B}, \Theta)$, which requires the estimation of the conditional densities of the estimated sources. Asymptotically, extending the results for linear mixtures of Markovian sources [7], the equivalence of the mutual information minimization method with the Maximum Likelihood method still holds for PNL mixtures of Markovian sources. As for linear mixtures, this method based on Markov model is very efficient for post-nonlinear mixture. For an computation cost equal to 3^q , where q is the order of the Markov model, the performance (in term of SNR) is increased and it becomes possible to separate Gaussian sources [11].

4 Compensation of non-invertible nonlinearities

For PNL mixing system with non-bijective functions, the previous algorithms cannot inverse the mixtures, because of the indeterminacy of the inverse functions. For example, when the unknown non-linear functions are $f_i(\cdot) = (\cdot)^2$, the inverse is as $\pm\sqrt{\cdot}$. In this case, assuming the inverse is known, we have to solve the sign indeterminacy. We investigated if this is possible using the time correlation of the signals.

Assuming the non-linear inverse function is known and equal to $\sqrt{\cdot}$, it remains to estimate its sign ϵ and the linear part of the separating structure. Thus, at each time n , the output sample $z_i(n+1)$ is predicted by linear prediction (LP) with the N previous samples:

$$\hat{z}_i(n+1) = \sum_{k=0}^{N-1} c_k z_i(n-k), \quad (15)$$

and we select the sign ϵ which minimizes the square error $(\hat{z}_i(n+1) - \epsilon \sqrt{e_i(n+1)})^2$, where $\epsilon = +1$ or $\epsilon = -1$. In fact, if $|e_i|$ is large, the prediction is easy. On the contrary, the sign determination

⁴ $g_i(\theta_i, x_i(t))$ is a parametric model of $g_i(\cdot)$, where θ_i can represent a set of parameters

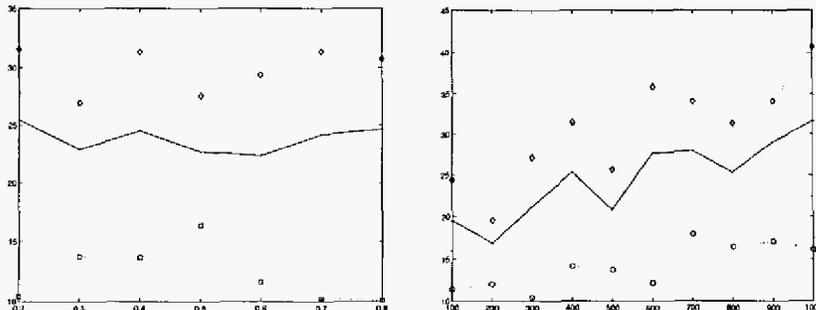


Fig. 3. Minimum (dashed circles), maximum (dashed diamonds) and average (plain) performance over 10 experiments, versus α (left) and versus the number of samples, for $\alpha = 0.5$ (right).

is difficult when $|e_i|$ is close to zero. For avoiding noisy estimation of the sign, it is necessary to chose the LPC order N greater than 2 or 3. Practically, since sign changes occur when $|e_i|$ is close to zero, for decreasing the computation cost, we can estimate the sign only for $|e_i| < \theta$.

The linear part (the matrix \mathbf{B}) is estimated according to estimation equation (6).

4.1 Experimental results

In the first experiment, we consider two random colored sources mixed with a linear matrix

$$\mathbf{A} = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix} \quad (16)$$

with $\alpha \in [0.2, 0.8]$. The non-linear mappings are $f_i(\cdot) = (\cdot)^2$. The sources are generated by filtering a Gaussian noise (500 samples) with the AR filter $1/AR(z)$, where $AR(z) = [1 - 4.6z^{-1} + 8.5z^{-2} - 8.0z^{-3} + 3.8z^{-4} - 0.75z^{-5}]$. Figure 3 plots the minimum, the maximum and the average performance over 10 experiments, measured as the residual error in dB ($E_i = \frac{E|e_i|^2}{E[(y_i - \hat{y}_i)^2]} [dB]$), versus α . We see that the performance does not depends on α , *i.e.* on the mixture hardness.

In the second experiment, we studied how the performance varies with the sample number. In figure 3, using the same mixing system as in the previous experiment, with $\alpha = 0.5$, we plot the maximum, minimum and average values of SNR, versus the sample number of samples. Each experiment is still repeated 10 times. Due to the strong correlation between successive samples, it is necessary to use a large number of samples for achieving a good estimation of the linear part as well as a good decision on the sign of the non-linear part.

Figure 4 is depicted one typical example of recovered source (plain line) and the true source (dashed line). We can observe that sometimes the true sign is lost (due to a wrong estimation) but it mainly occurs for $|e_i|$ close to zero, and the true sign is quickly recovered.

Although this work is promising, the next step, which consists in blindly estimating the (unknown) non-linear mapping $\pm g_i$ and \mathbf{B} , is much more tricky. However, this simple example shows how weak priors can be used for solving ill-posed problems.

5 Concluding remarks

In this paper, we have considered ICA and BSS problems for nonlinear mixture models. It appears clearly BSS and ICA are difficult and ill-posed problems, and regularization is necessary for actually achieving ICA solutions which coincide to BSS.

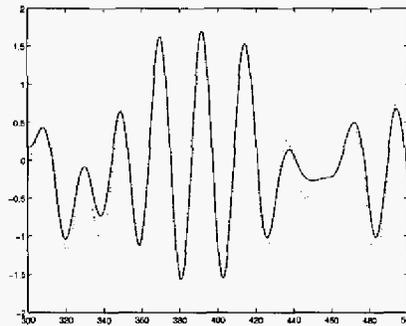


Fig. 4. Recovered source (plain) and the true source (dashed) for one experiment

In this purpose, two main ways can be used. First, solving the nonlinear BSS problem appropriately using only the independence assumption is possible only if mixtures as well as separation structure are structurally constrained: for example post-nonlinear mixtures. Second, prior information on sources, for example temporally correlated sources, can simplify the algorithms or reduce the indeterminacies in the solutions.

A lot of work remains to be done in studying the nonlinear ICA and BSS problems. Especially, the solution of more general non-linear problems, adding a few priors: temporal correlation, more observations than sources [9, 12]. Finally, up to now, the research has addressed mainly theoretical problems. The results will become more widely interesting only if they can be validated on realistic problems using real-world data [4].

6 Acknowledgments

This work has been partly supported by the European Commission project BLISS (IST-1999-14190), and by the University of Vic under the grant R0912.

References

1. S. Achard, D. T. Pham, and C. Jutten. Blind source separation in post nonlinear mixtures. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2001)*, pages 295–300, San Diego, CA, USA, 2001.
2. M. Babaie-Zadeh, C. Jutten, and K. Nayebi. Separating convolutive post non-linear mixtures. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2001)*, pages 138–143, San Diego (California, USA), 2001.
3. M. Babaie-Zadeh, C. Jutten, and K. Nayebi. A geometric approach for separating post nonlinear mixtures. In *Proc. of the XI European Signal Processing Conf. (EUSIPCO 2002)*, volume II, pages 11–14, Toulouse, France, 2002.
4. G. Bedoya, C. Jutten, S. Bermejo, and J. Cabestany. Improving semiconductor-based chemical sensor arrays using advanced algorithms for blind source separation. In *Proceedings of Sicon 04*, January 27–29, New Orleans (USA), 2004.
5. A. Cichocki and S.-I. Amari. *Adaptive Blind Signal and Image Processing - Learning Algorithms and Applications*. J. Wiley, 2002.
6. P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.
7. S. Hosseini, C. Jutten, and D.-T. Pham. Markovian source separation. *IEEE Trans. on Signal Processing*, 51:3009–3019, 2003.

8. A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. J. Wiley, 2001.
9. A. Ilin, S. Achard, and C. Jutten. Bayesian versus constrained structure approaches for source separation in post-nonlinear mixtures. In *Proceedings IJCNN 2004 (to appear)*, 2004.
10. M. Korenberg and I. Hunter. The identification of nonlinear biological systems: LNL cascade models. *Biol. Cybernetics*, 43(12):125–134, December 1995.
11. A. Larue, C. Jutten, and S. Hosseini. Markovian source separation in non-linear mixtures. Granada, Spain, September 2004. Submitted to ICA 2004.
12. J. Lee, C. Jutten, and M. Verleysen. Non-linear ICA using isometric dimensionality reduction. Granada, Spain, September 2004. Submitted to ICA 2004.
13. T.-W. Lee, B. Koehler, and R. Orglmeister. Blind source separation of nonlinear mixing models. In *Neural Networks for Signal Processing VII*, pages 406–415. IEEE Press, 1997.
14. A. Paraschiv-Ionescu, C. Jutten, and G. Bouvier. Source separation based processing for smart sensor arrays. *IEEE Sensors Journal*, 2(6):663–673, December 2002.
15. H. Peng, Z. Chi, and W. Siu. A semi-parametric hybrid neural model for nonlinear blind signal separation. *Int. J. of Neural Systems*, 10(2):79–94, 2000.
16. D. T. Pham. Fast algorithm for estimating mutual information, entropies and score functions. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2003)*, pages 17–22, Nara, Japan, 2003.
17. S. Prakriya and D. Hatzinakos. Blind identification of LTI-ZMNL-LTI nonlinear channel models. *IEEE Trans. on Signal Processing*, 43(12):3007–3013, December 1995.
18. C. Punttonet, M. Alvarez, A. Prieto, and B. Prieto. Separation of sources in a class of post-nonlinear mixtures. In *Proc. 6th European Symp. on Artificial Neural Networks (ESANN'98)*, pages 321–326, Bruges, Belgium, April 1998.
19. M. Solazzi, R. Parisi, and A. Uncini. Blind source separation in nonlinear mixtures by adaptive spline neural networks. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2001)*, pages 254–259, San Diego, CA, USA, 2001.
20. J. Solé-Casals, M. Babaie-Zadeh, C. Jutten, and D.-T. Pham. Improving algorithm speed in PNL mixture separation and Wiener system inversion. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2003)*, Nara (Japan), April 2003.
21. A. Taleb and C. Jutten. Batch algorithm for source separation in post-nonlinear mixtures. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 1999)*, pages 155–160, Aussois, France, 1999.
22. A. Taleb and C. Jutten. Source separation in post-nonlinear mixtures. *IEEE Trans. on Signal Processing*, 47(10):2807–2820, 1999.
23. A. Taleb, J. Solé-Casals, and C. Jutten. Quasi-nonparametric blind inversion of Wiener systems. *IEEE Trans. on Signal Processing*, 49(5):917–924, 2001.
24. A. Ziehe, M. Kawanabe, S. Harmeling, and K.-R. Müller. Separation of post-nonlinear mixtures using ACE and temporal decorrelation. In *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA 2001)*, pages 433–438, San Diego, USA, 2001.