# Flexible ICA solution for nonlinear blind source separation problem

D. Vigliano and A. Uncini

Presented is a new architecture and a new learning algorithm that are exploited to resolve the blind source separation problem under stricter constraints than those considered to date. The mixing model that is assumed is an evolution of the well-known post-nonlinear (PNL) one: the PNL mixing block is followed by a convolutive mixing channel. The flexibility of the algorithm originates from the spline-SG neurons performing an on-line estimation of the score functions.

*Introduction:* The scientific community's interest in blind signal processing and in particular for blind source separation (BSS), performed through independent component analysis (ICA), is growing. The large number of its possible applications imposes the move from a simple static mixing environment (the simple cocktail party mixing model) to the more complex nonlinear one. While the BSS separation in a convolutive environment has been achieved with good results in a large number of ways, the nonlinear one has not been thoughtfully analysed yet. Existing literature assumes the existence of a solution to some particular nonlinear problem and under particular conditions gives a proof of the uniqueness of such a solution. To ensure this uniqueness it is necessary to simplify, in some way, the general nonlinear mixing environment.

Uniqueness only occurs granted in conformal nonlinear static mapping when there are two sources [1] and in the so-called post-nonlinear (PNL) case both with a static and convolutive mixing block [2–4]. Now the trend is to find the strictest nonlinear convolutive mixing environment for which is possible to recover the original independent sources [3, 5]. Within this framework, in this Letter we aim to show that a new kind of mixing problem can be resolved with ICA techniques based on information theory and the adaptive neural network. This new mixing structure consists of a cascade of a PNL block and a convolutive mixing channel. The algorithm exposed here is completely blind; output independence is achieved by optimising a cost function based on information theory.

*Mixing model:* The general formulation for the nonlinear hidden mixing model can be expressed as $\mathbf{x}(n) = \mathscr{F}\{\mathbf{s}(n), \ldots, \mathbf{s}(n-L)\}$, in which $\Re\{\cdot\}$ is a nonlinear mixing function and $\mathbf{s}(n)$ is the original source vector. Several papers show how independence conservation constraint is not strong enough to determine output independence [1, 2] if the transformation $\mathscr{F}$ has no particular form and no other assumption is made about the mixing (and demixing) model. Here a nonlinear convolutive mixing model stricter than the one just explored is considered. The mixed accessible signal $\mathbf{x}[n]$ can be written in a close form as:

$$\mathbf{x}[n] = \mathscr{F}[\mathbf{s}] = \sum_{k=0}^{L-1} \mathbf{Z}[k]\mathbf{F}[\mathbf{A}\mathbf{s}[n-k]] \qquad (1)$$

in which $\mathbf{A}$ is an $N \times N$ matrix (where $N$ represents the sources number), $\mathbf{F}[\mathbf{r}(n)] = [f_1[r_1(n)], \ldots, f_N[r_N(n)]]^T$ are the $N$ nonlinear monotone distorting functions (one per channel) and $\mathbf{Z}[k]$ is a matrix of FIR filters. Considering some specified nonlinear convolutive mixing transform $\mathscr{F}\{\cdot\}$ applied to a vector of independent components $\mathbf{s}$, let $\mathscr{Y}$ be the set of all $\mathbf{y}$ for which the transform $\mathscr{H}\{\cdot\}$ has a diagonal Jacobian; $\mathscr{Y} = \{\mathbf{y} | \mathbf{y} = \mathscr{G} \circ \mathbf{x} = \mathscr{G} \circ \mathscr{F} \circ \mathbf{s} = \mathscr{H} \circ \mathbf{s}\}$. Each element of $\mathscr{Y}$ is a vector of independent components; there are an infinite number of possible functions $\mathscr{G}$ (see note 1). Therefore in $\mathscr{Y}$ there are an infinite number of elements that do not differ from each other for trivial relations. The set $\mathscr{Y}$ collects solutions most of which are of no interest. Thus the output independence is a weak approach to the problem. The solution of the nonlinear convolutive ICA question requires *a priori* assumptions concerning, for example, the mixing–demixing model. Introducing a precise kind of recovery model is sufficient to avoid this strict non-uniqueness of the solution; such a model reduces the weakness of the output independence condition diminishing the cardinality of all possible solutions.

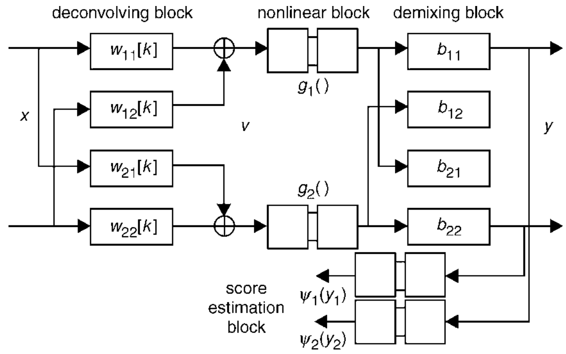


**Fig. 1** *Recovering neural architecture*

*Demixing model:* The best structure to recover signals after the mixture expressed in (1) is the one proposed in Fig. 1. The output vector $\mathbf{y}[n]$ can be expressed as:

$$\mathbf{y}[n] = \mathscr{G}[\mathbf{x}] = \mathbf{BG}\left[\sum_{k=0}^{L-1} \mathbf{W}[k]\mathbf{x}[n-k]\right] \qquad (2)$$

in which $\mathbf{B}$ is an $N \times N$ matrix, $\mathbf{W}[k]$ is a matrix of FIR, and $\mathbf{G}$ is the vector of nonlinear compensating functions. Determining an estimation of the original sources except for some trivial indeterminacy (see note 2) gives the desired solution, according to the theory, for the blind source separation problem. This algorithm measures the output independence by the Kullback-Leibler divergence between $p_Y[\mathbf{y}]$ and $\tilde{p}_Y[\mathbf{y}] = \prod_i p_{y_i}[y_i]$; the highest independency of the output signals is reached where the cost function is at its minimum. To perform the minimisation, stochastic gradient is used. The use of KL divergence as cost function is justified by its suitable mathematical properties. The so-called score functions $\psi_i(y_i) = \dot{p}_i(y_i)/p_i(y_i)$ are a critical issue of this kind of algorithm, those that perform on-line estimation have a better and faster convergence. The neural network of Fig. 1 uses two spline-SG neurons to quickly obtain a good on-line estimation of the score functions and another one to estimate the inverse nonlinear distortion. A direct method that minimises the mean square error is used to obtain the learning rule for the SG-score estimation-neuron only. The stochastic gradient grants the blindness [5]. The use of spline-SG neurons is justified above all by the local on-line learning that characterise it with respect to the other algorithm only widely diffused.

*Experimental results:* The validation of the suggested solution required a challenging environment. In these two experiments the mixtures were sampled at 8 kHz. In both cases the nonlinear distorting functions applied to signals is:

$$\mathbf{F}[f_1(r_1),\ f_2(r_2)]^T = [f_1 = r_1 + 2r_1^3,\ f_2 = 0.5r_2 + \tanh(5r_2)]^T \qquad (3)$$

Every SG-neuron had 53 control points equally spaced on the $x$-axis. The training was achieved with 150 epochs of 19 000 sample signals; the learning rates were $\eta_1 = 10^{-6}$ for the SG neurons and for the FIR taps but $\eta_W = 10^{-5}$ for the $\mathbf{W}$ FIR matrix. We then evaluated the achieved separation by mean of ISI coefficient.

*Experiment 1:* mix of white noise and female voice. The convolving channel was:

$$\mathbf{Z}[k] = \begin{bmatrix} 0.05 - 0.025z^{10} + 0.125z^{20} & -0.045 - 0.035z^{20} \\ 0.025 - 0.0125z^{10} + 0.00645z^{20} & 0.05 + 0.0125z^{20} \end{bmatrix}$$

The mixing matrix was [0.7, −0.2; 0.3, 0.6]. The FIR matrix $\mathbf{W}$ has filters with 70 taps.

*Experiment 2:* mix of male voice and female voice. The convolving channel was:

$$\mathbf{Z}[k] = \begin{bmatrix} 0.08e^{-(k-1)/4} + 0.045e^{-(k-10)/4} & -0.03e^{-(k-1)/2} - 0.015e^{-(k-15)/4} \\ 0.04e^{-(k-1)/6} + 0.02e^{-(k-8)/6} & 0.06e^{-(k-1)/4} + 0.01e^{-(k-15)/4} \end{bmatrix} \quad k = [1, \ldots, 40]$$

The mixing matrix was [0.7, 0.36; 0.4, 0.6]. The FIR matrix $\mathbf{W}$ has filters with 90 taps.

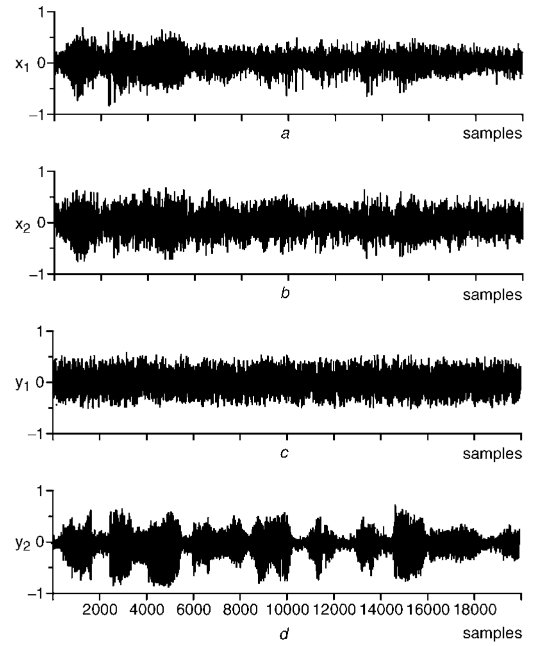


**Fig. 2** *Experiment 1*

*a* Mixed input signal channel 1 *b* Mixed input signal channel 2
*c* Demixed output signal channel 1 *d* Demixed output signal channel 2

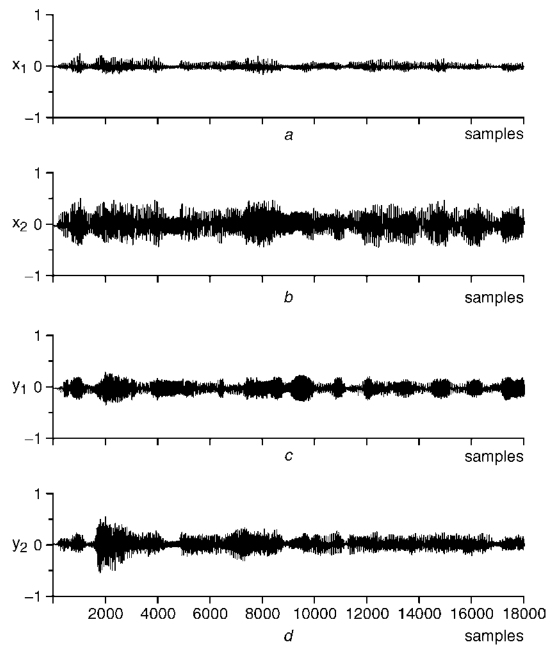


**Fig. 3** *Experiment 2*

*a* Mixed signal channel 1 *b* Mixed signal channel 2
*c* Demixed signal channel 1 *d* Demixed signal channel 2

Fig. 2 shows the mixed and the demixed signals of experiment 1 and Fig. 3 those of experiment 2. These results show that the algorithm performs quite well regardless of the complexity of the mixing environment. The complexity of the mixing structure requires small learning rates that can produce very slow convergence. We have discovered a monotone trend of the ISI coefficient during the learning, it should grant better separation as result of longer training. The final values for ISI coefficient were for experiment 1: $\text{ISI}_B = 0.05$, $\text{ISI}_W = 1.12$, and for experiment 2: $\text{ISI}_B = 0.06$, $\text{ISI}_W = 1.02$.

*Conclusions:* We have presented a new blind algorithm to recover original sources from a novel type of mixing environment composed by a PNL state followed by a convolutive mixing channel. The algorithm uses flexible splines to perform a local on-line estimation of the unknown functions (both the score functions and the nonlinear compensating functions). From the results the good quality of the demixing signals is evident. The algorithm in both cases provided a good separation, deconvolution and nonlinear compensation of the mixed sources. The performances relating to the estimation of score functions were also good, this being an important issue for the learning. The results presented show how the aim of this work has been widely achieved. Better results, in terms of both quality and of computational effort, can be obtained with different demixing structures. For future work, main studies will relate to recovering architectures and the possible types of mixing environment that can be successfully approached by BSS.

*Notes:*

1. Note for example that every single channel distortion preserves the independence of independent variables.
2. The trivial indeterminacy can be a permutation, a time delaying and a scaling of the recovered sources.

 *1 July 2003*

*Electronics Letters Online No: 20031033*
*DOI: 10.1049/el:20031033*

D. Vigliano and A. Uncini (*Dipartimento INFOCOM, Universitá di Roma 'La Sapienza'—Italy, Via Eudossiana, 18, 00184 Roma, Italy*)

**References**

1 HYVARINEN, A., and PAJUNEN, P.: 'Non linear independent component analysis: existence and uniqueness results', *Neural Netw.*, 1999, **12**, (3), pp. 429–439
2 TALEB, A.: 'A generic framework for blind source separation in structured nonlinear models', *IEEE Trans. Signal Process.*, 2002, **50**, (8)
3 MILANI, F., SOLAZZI, M., and UNCINI, A.: 'Blind source separation of convolutive non linear mixtures by flexible spline nonlinear functions'. Proc. IEEE Int. Conf. Acoustic Speech and Signal Processing, ICASSP'02, Orlando, FL, USA, May 2002
4 TAN, Y., WANG, J., and ZURADA, J.M.: 'Non linear blind source separation using radial basis function', *IEEE Trans. Neural Netw.*, 2001, **12**, (1), pp. 124–134
5 SOLAZZI, M., PIAZZA, F., and UNCINI, A.: 'Non linear blind source separation by spline neural network'. ICASSP 2001, Salt Lake City, UT, USA, May 2001