Blind Source Separation of Post-Nonlinear Mixtures Using Evolutionary Computation and a Local Search Procedure

Tiago Macedo Dias, Ricardo Suyama, Leonardo Tomazeli Duarte, Romis Ribeiro de Faissol Attux, and João Marcos Travassos Romano

Abstract— In this work, we propose a new method for source separation of post-nonlinear mixtures that brings together evolutionary-based global search, entropy estimation via order-statistics and a local search step based on the FastICA algorithm. The rationale of the proposal is to attempt to obtain efficient and precise solutions using with parsimony the available computational resources. The new proposal was tested in different scenarios, and, in all cases, we attempted to establish grounds for comparison with an alternative approach whose optimization step does not include the local (memetic) search stage. Simulation results indicate that a good tradeoff between performance and computational cost was indeed reached.

Keywords— Post-Nonlinear, evolutionary algorithms, Blind Source separation, high-order statistics.

Resumo— Neste trabalho, propomos um novo método para separação de fontes que une uma ferramenta de busca global baseada em computação evolutiva, o método de estatísticas de ordem para estimação de entropia e uma etapa de busca local conduzida pelo algoritmo FastICA. A idéia subjacente à proposta é procurar obter soluções precisas e eficientes usando de maneira parcimoniosa os recursos computacionais disponíveis. A nova proposta foi testada em diferentes cenários, e, em todos os casos, tentamos estabelecer bases para comparação com uma abordagem alternativa cujo passo de otimização não inclui o estágio de busca local (ou "memética"). Os resultados obtidos por meio de simulações indicam que um bom compromisso entre desempenho e custo computacional foi, de fato, atingido.

Palavras-Chave— Post-Nonlinear, algoritmos evolutivos, separação cega de fontes, high order statistics.

I. INTRODUCTION

The problem of blind source separation (BSS) is related to the idea of recovering a set of sources from samples that are mixtures thereof. Until the end of the last decade, the majority of the proposed techniques [1] were designed to solve the standard linear and instantaneous mixture problem. However, some practical applications [4] may exhibit a strong nonlinear character, which renders the use of this classical framework impractical: it is thus of paramount relevance to seek a broader systemic view.

Tiago Macedo Dias, Ricardo Suyama, Leonardo Tomazeli Duarte, and João Marcos Travassos Romano from DSPCOM, School of Electrical and Computer Engineering, University of Campinas (UNICAMP) and Romis Ribeiro de Faissol Attux from DCA - School of Electrical and Computer Engineering, University of Campinas (UNICAMP) - C.P. 6101, CEP 13083-970, Campinas - SP, Brazil; {tdias, rsuyama, romano}@decom.fee.unicamp.br, romisri@dca.fee.unicamp.br

In the classical linear scenario, it is fairly usual to associate the solution of a blind source separation problem with the notion of performing ICA (Independent Component Analysis), a methodology derived from the field of data analysis that can be seen as an extension of the classical Principal Component Analysis (PCA). The existence of efficient ICA algorithms - such as the well-known FastICA - is indicative of the degree of maturity reached insofar as the linear case is concerned; on the other hand, a direct extension of these ideas to the nonlinear context must be considered with care, since, in such a case, there is no guarantee that it will be always possible to rely solely on the assumption of mutual independence between sources, which is the very essence of the ICA idea, to carry out the separation task. Fortunately, there are some classes of nonlinear mixture models in which the ICA framework still holds [4], such as that of the so-called Post-Nonlinear (PNL) models.

Post-Nonlinear models are characterized by the existence of memoryless monotonic nonlinearities that further distort linearly-mixed signals. In [4], Taleb and Jutten proposed a solid paradigm for inverting the action of a PNL mixture system that was based on the minimization of the mutual information between the source estimates. Despite its theoretical soundness, this approach suffers from two major practical drawbacks. The first one comes from the fact that the evaluation of the mutual information demands estimation of the marginal entropies, which may be a rather complex task. The second one is related to the presence of local minima in the mutual information-based cost function [9], which poses a problem to the adaptation of the separating system via gradient-based algorithms. Therefore, one may conclude that an effective approach to the PNL problem should be founded on an optimization technique that contains mechanisms to avoid convergence to local minima and also on an efficient entropy estimator, since the cost function to be optimized is built with the aid of the marginal entropies of the recovered signals.

In accordance with these ideas, our approach is founded on a twofold basis. In order to estimate the marginal entropy, one can use an approach based on the so-called order statistics [14]. On the other hand, in order to build an effective search method, it is possible to combine the explorative power of evolutionary algorithms (in particular, of an artificial immune system) with the remarkable local search potential derived from the use of the FastICA algorithm. This new

method, which can be understood as an extension of the work reported in [6], is tested under a number of different scenarios.

The work is structured as follows. In Section II, the fundamentals of the problem of separating PNL mixtures are discussed. Section III presents the algorithms and exposes the proposal for the PNL problem. Simulation results are discussed in Section IV. And finally, Section V presents the concluding remarks.

II. FUNDAMENTALS: BLIND SOURCE SEPARATION OF PNL MIXTURES

Let $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_N(t)]^T$ denote N mutually independent sources and $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T$ be the N mixtures of the source signals, i.e., $\mathbf{x}(t) = \phi(\mathbf{s}(t))$. The aim of a BSS technique is to recover the source signals based solely on the observed samples of the mixtures and on a minimal amount of statistical information about the signals.

The function ϕ can be either a linear or a non-linear function. In the former case, it is a linear BSS problem, in which the following model holds: $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$, being $\mathbf{A}$ the mixing matrix. Separation can be achieved by multiplying the mixture vector by a separating matrix $\mathbf{W}$, such that $\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t)$ and the elements of $\mathbf{y}(t)$ are statistically independent. Ideally, the separating matrix $\mathbf{W}$ will be, up to scaling factors and permutations, inverse to the mixing matrix $\mathbf{A}$. This approach is known as *Independent Component Analysis* (ICA) [1].

On the other hand, the standard BSS problem can be extended by considering ϕ as a nonlinear mixture. In such case, the independence hypothesis, which is the crux of ICA, may no longer be enough to obtain the original sources: this makes the idea of restricting the nonlinear mixing model to a class of separable models very attractive.

The most representative example of a nonlinear separable mixture model is the PNL (Post-Nonlinear) system (Fig. 1). The PNL model divides the problem into its linear and nonlinear parts, wherefore it can be understood as being composed of two separate but interrelated problems. Mathematically, the mixture process is given by $\mathbf{x}(t) = \mathbf{f}(\mathbf{A}\mathbf{s}(t))$, where $\mathbf{f}(\cdot) = [f_1(\cdot), f_2(\cdot), \ldots, f_n(\cdot)]^T$ denotes the nonlinearities applied to each output of the linear mixing stage.
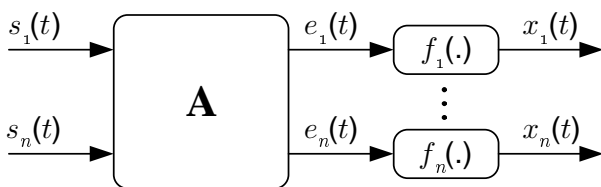

Fig. 1.   The PNL problem structure

In contrast to the linear model, the PNL one may be very suitable to describe the situation wherein the sensors are working in a saturated region [4]. This capability permit us to envisage some applications. For instance, the PNL model could be employed to describe MIMO communication systems with nonlinear amplifier stages.

Source separation of PNL mixtures can be achieved by considering the separating system $\mathbf{y}(t) = \mathbf{W}\mathbf{g}(\mathbf{x}(t))$ [4],

where $\mathbf{g}(\cdot) = [g_1(\cdot), g_2(\cdot), \ldots, g_n(\cdot)]^T$ is a set of nonlinear functions that must be precisely adjusted in order to invert the action of the nonlinearity $\mathbf{f}(\cdot)$. Interestingly, it has been shown [4] that, under some mild conditions over $\mathbf{A}$, $\mathbf{W}$, $\mathbf{f}(\cdot)$ and $\mathbf{g}(\cdot)$, it is possible to separate the sources in this scenario relying exclusively on the well known ICA framework. The relevance of this fact, which cannot be underestimated, is decisive insofar as our option for this model is concerned.

A. Nonlinearities: Source Separation Based on the Minimization of Mutual Information

With Fig. 1 in mind, let us mathematically divide the problem, as proposed before, in its linear and nonlinear parts, by calling $\mathbf{z}(\mathbf{t}) = \mathbf{g}(\mathbf{x}(t))$:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{g}(\mathbf{x}(t)) = \begin{cases} \mathbf{y}(t) = \mathbf{W}\mathbf{z}(t) \\ \mathbf{z}(\mathbf{t}) = \mathbf{g}(\mathbf{x}(t)) \end{cases} \quad (1)$$

As stated in previous sections, source separation is achieved by seeking a condition of independence between the components of the estimated vector $\mathbf{y}$, which means that, to minimize the mutual information between them, the recovery of the source signals is a straightforward procedure.

The mutual information between the components of $\mathbf{y}$ is given by:

$$I(\mathbf{y}) = \sum_i H(y_i) - H(\mathbf{y}), \quad (2)$$

where $H(\mathbf{y})$ represents the joint entropy of $\mathbf{y}$ and $H(y_i)$ is the entropy of each one of its components. Again, having in mind Fig. 1 and the main idea of equation (1), one can develop (2) by applying the entropy transformation law [1]. Therefore, the following expression can be written down

$$I(\mathbf{y}) = \sum_i H(y_i) - H(\mathbf{x}) - \log|\det \mathbf{W}| - E\left\{\log \prod_i |g_i'(x_i)|\right\}, \quad (3)$$

with $g_i'(\cdot)$ denoting the first derivative of the nonlinearity $g_i(\cdot)$. Also note that this equation (3) holds *only if the functions $g_i(\cdot)$ are invertible*. So, when developing the learning algorithm, this restriction must be taken into account.

By observing Eq. (3) one can see that estimation of the mutual information relies basically on an accurate estimation of both $H(y_i)$ and $H(\mathbf{x})$. Taking a closer look at these terms, and considering that our main goal is to minimize Eq.(3) with respect to the parameters of the separating system, it is important to note that $H(\mathbf{x})$ does not depend on the parameters of the separating system and is constant for static mixing systems. This means that it can be ignored in the learning process.

This fact restricts our optimization efforts to the term $H(y_i)$, the marginal entropy of the outputs, which is strongly dependent of the separating system, since it is related to the output $\mathbf{y}$. More will soon be said about the process of estimation of these terms.

III. OPT-AINET AND FASTICA

As discussed in the previous section, in the nonlinear BSS problem, there are two main issues to be addressed: estimation of mutual information by estimating the marginal entropies

of each output y_i and avoidance of local minima. Mutual information is the cost function to be adopted in this work, allied with a method based on the so-called *order statistics* for estimation of the marginal output entropies. The optimization process, on the other hand, is something that, in contrast with the idea presented in [6], is carried out by an evolutionary algorithm together with a local search technique, the FastICA.

As it was mentioned earlier, it is possible to think of the PNL mixing process as being composed of two parts. Analogously, it is natural to apply the same division to the separating system, by using the evolutionary algorithm to adapt the parameters of the nonlinearities $g_i(\cdot)$ given a certain matrix $\mathbf{W}$. Moreover FastICA is used to adapt the linear part of the model, a task to which it is particularly suited. This strategy of combining a general-purpose search tool with a remarkable global search potential and a specialized local search algorithm is in consonance with the notion of memetic algorithm, a name coined having in mind Richard Dawkins' concept of meme - a unit of cultural evolution that can exhibit local refinement [13]

In simple terms, this puts together three elements: an evolutionary technique - the artificial immune network called opt-aiNet -, the FastICA algorithm and a method for entropy estimation based on order statistics - the estimator, it is important to say, leads to a good tradeoff between accuracy and complexity. These three elements give rise to a learning process with two main features:

1) From the standpoint of the opt-aiNet, only the parameters of the nonlinear functions are optimized.
2) Whenever it is necessary to evaluate the fitness of a given set of parameters, the FastICA algorithm is used to calculate the matrix $\mathbf{W}$ that, together with these parameters, will be associated with a certain value of mutual information - the basis of the fitness measure.

Since the ability to perform global search is inherent to the opt-aiNet and the efficiency in performing local search is something typical of the FastICA, satisfactory results are expected.

A. The FastICA algorithm

The FastICA [7] is a widespread algorithm applied, as a rule, in linear BSS problems. The method optimizes a nonlinear contrast function tailored with the aid of a measure of the non-Gaussianity of the signals.

This work employs the algorithm described in [8], which implements an approximate Newton iteration for minimizing the contrast function. This means that, contrarily to gradient-based algorithms, there are no step size parameters to choose.

The steps of the algorithm can be summarized as follows:

1) Center the data to make its mean zero
2) Whiten the data to obtain $\bar{\mathbf{z}}$, i.e., transform the centered data to obtain $\bar{\mathbf{z}}$ such that $R_{\bar{\mathbf{z}}\bar{\mathbf{z}}} = E\{\bar{\mathbf{z}}\bar{\mathbf{z}}^T\} = I$. This can be done using the eigenvalue decomposition of the original autocorrelation matrix;
3) Since the data have been whitened, the separating matrix $\mathbf{W}$ will be orthogonal. Therefore, choose initial values

for $\mathbf{w}_i$ (columns of $\mathbf{W}$), each of unit norm. Orthogonalize $\mathbf{W}$ as in step 4 below
4) For every i, let

$$\mathbf{w}_i \leftarrow E\left\{\mathbf{z}g\left(\mathbf{w}_i^T\mathbf{z}\right)\right\} - E\left\{g'\left(\mathbf{w}_i^T\mathbf{z}\right)\right\}\mathbf{w}_i$$

where $g(x) = \tanh(x)$
5) Do a symmetric orthogonalization of $\mathbf{W}$

$$\mathbf{W} \leftarrow \left(\mathbf{W}\mathbf{W}^T\right)^{-1/2}\mathbf{W}$$

6) If the algorithm has not converged, go back to step 3

B. Evolutionary Optimization Technique - Opt-ainet

Evolutionary algorithms have been applied, in the last decade, to a myriad of engineering problems. In this work, they will be useful due to their remarkable global search potential, which can be decisive in a multimodal context like the one that emerges from the use of an information-theoretical cost function in the Post-Nonlinear scenario. Among a number of possibilities, the artificial immune network was chosen. This is the opt-aiNet (Optimization version of Artificial Immune Network) [12], which is founded on the combination of notions like clonal selection and *affinity maturation* and the concept of *immune network*.

Under the conceptual framework of clonal selection and affinity maturation, the immune system is understood as being composed of cells and molecules that carry receptors for antigens (disease-causing agents). In simple terms, when these receptors recognize a given antigen, they are stimulated to proliferate. During the process, controlled mutation takes place, and, thereafter, the individuals are subjected to a natural selection mechanism that tends to preserve the most adapted.

The immune network theory also states that it is possible for these immune cells and molecules to interact with each other generating "eigenbehavior" even in the absence of antigens. This means that invasions could be seen as perturbations of the status-quo, evolving the immune cells into a new level of adaptation.

In order for this theory to be efficiently applied, there are some important points that need to be taken into account: 1) The fitness function, which is the one being optimized, is, in fact, a measure of affinity between antibody and antigen; 2) Each solution corresponds to the information contained in a given receptor (network cell); 3) The affinity between cells is measured by a simple Euclidean distance.

These are the key points when structuring an opt-aiNet algorithm, which can be seen in more details below:

1) Initialization: randomly create initial network cells;
2) Local search: while the stopping criterion is not met, do:
   a) Clonal expansion: for each network cell, determine its fitness (an objective function to be optimized). Generate a set of N_c antibodies, named clones, which are the exact copies of their parent cell;
   b) Affinity maturation: mutate each clone with a rate that is inversely proportional to the fitness of its parent antibody, which itself is kept unmutated. The mutation follows

$$c' = c + \alpha N(0,1), \text{ with } \alpha = \beta^{-1}\exp\left(-f^*\right) \quad (4)$$

where c' and c represent the mutated and the original individual, respectively; β is a free parameter that controls the decay of the inverse exponential function, and f^* is the fitness of an individual. For each mutated clone, select the one with highest fitness and calculate the average fitness of the selected cells;

   c) Local convergence: if the average fitness of the population does not vary significantly from one iteration to the other, go to the next step; else, return to Step 2;

3) Network interactions: determine the similarity between each pair of network antibodies;

4) Network suppression: eliminate all but one of the network cells whose affinity with each other is lower than a pre-specified threshold σ_s, and determine the number of remaining cells in the network;

5) Diversity: introduce a number of new randomly generated cells into the network and return to Step 2.

Step 1 accounts for initialization, then Step 2 implements a local search applying clonal expansion and affinity maturation based on a fitness-dependent mutation operator; Steps 3 and 4 give the algorithm its main characteristic: an "immune network character" by evolving the antibodies and selecting the "most adapted" ones, thus restricting the size of the population; and at last, Step 5 augments the diversity of the population by adding new features into it. The combination of these stages produces an algorithm that allies a good balance between exploration and exploitation with the notion of seeking a parsimonious use of the available resources.

C. Entropy Estimation using Order Statistics

Order statistics has an historical of successful implementation in BSS problems throughout the literature [14]. As mentioned before, this is an attractive method due to its low computational complexity, when compared to other methods that use density estimation, for example. However, the use of order statistics in entropy estimation does not easily render gradient-based algorithms for blind source separation [14], motivating the use of different optimization tools, like the opt-aiNet algorithm presented in section III-B.

Aiming to introduce the notion of order statistics, one may consider a set of N samples of the random variable Y organized as

$$y_{(1:T)} \leq y_{(2:T)} \leq \cdots \leq y_{(T:T)}. \tag{5}$$

The so-called *kth order statistic*, denoted by $y_{(r:T)}$, is nothing more than the rth value, in ascending order, among the T available samples [14].

In order to clarify the applicability of order statistics to the problem of entropy estimation, let us rewrite the entropy of a random variable Y in terms of its quantile function $Q_Y(u) = \inf\{y \in \Re : P(Y \leq y) \geq u\}$, which is, in fact, the complement of the cumulative distribution function $F_Y(y) = P(Y \leq y)$. Using this definition, it is possible to show

that [14]

$$H(y) = \int_{-\infty}^{\infty} f_Y(\tau) \log Q_Y'[F_Y(\tau)] d\tau = \int_0^1 \log Q_Y'(u)du, \tag{6}$$

where $f_Y(y)$ and $Q_Y'(y)$ denote the probability density function and the derivative of the quantile function of y, respectively.

For practical reasons, in order to evaluate the entropy of a given signal y_i, it is necessary to obtain a discretized form of (6), which is given by

$$H(y_i) \approx \sum_{k=2}^{L} \log\left[\frac{Q_{Y_i}(u_k) - Q_{Y_i}(u_{k-1})}{u_k - u_{k-1}}\right]\frac{u_k - u_{k-1}}{u_L - u_1} \tag{7}$$

with $\{u_1, u_2, \ldots, u_L\}$ denoting a set of increasing number in the interval $[0, 1]$.

The link between entropy estimation and order statistics lies in the close relationship between order statistics and the quantile function. In fact, an estimate of the value of $Q_Y(\frac{k}{T+1})$, called *empirical quantile function*, is given by kth order statistic $y_{(k:T)}$ [14]. Therefore, one can approximate the value of $Q_Y(\cdot)$ in Eq. (7) by $Q_{Y_i}(u) \approx y_{(k:T)}$, for k such that $\frac{k}{T+1}$ is the closest point to u. This simplification results in a fast algorithm for entropy estimation, which is a desirable characteristic when dealing with optimization using evolutionary algorithms.

IV. SIMULATION RESULTS

In order to illustrate the effectiveness of the proposed methodology, this section provides the simulation results from two different scenarios, with 2 and 3 sources, respectively.

A. First scenario

The first PNL-based scenario is composed of two uniformly distributed sources (between $[-1, 1]$) mixed in accordance with the following model:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.6 \\ 0.5 & 1 \end{bmatrix} \text{ and } \begin{array}{rcl} f_1(e_1) & = & \tanh(2e_1) \\ f_2(e_2) & = & 2\sqrt[5]{e_2} \end{array}. \tag{8}$$

From the expression of the PNL model output: $\mathbf{y}(t) = \mathbf{W}\mathbf{g}(\mathbf{x}(t))$, it is observable that $\mathbf{W}$ is the linear square matrix that needs to be optimized. Furthermore, the nonlinear function $\mathbf{g}(\cdot)$ is ideally the inverse of $\mathbf{f}(\cdot)$, and is modeled as a polynomial of order 5 with odd powers: $g_i(x_i) = ax_i^5 + bx_i^3 + cx_i$. Since a crucial requirement of the PNL model is that each function be monotonic, the coefficients of each polynomial are restricted to be positive.

Since the main novelty of this paper is the introduction of a local search term in the optimization process (wherefore our proposal will receive the epithet "memetic"), our results shall be compared with those obtained with a method based solely on an evolutionary algorithm [6] (this method will be referred to as "standard opt-aiNet method"). The tests concerning the standard method involve 10 experiments run with the following parameters: $N_0 = 5$, $N_c = 5$, $\beta = 50$, $\sigma_s = 3$. For the memetic algorithm, the same number of

TABLE I

RESULTS RELATED TO CONVERGENCE AND SIMULATION TIME IN THE
FIRST SCENARIO.

|	opt-aiNet	memetic
Convergence speed (iterations)*	10500	1550
Time spent per iteration (ms)	80	194
Average time to converge (min)	14	3.5

* Average number of iterations spent to reach the algorithm
convergence.

TABLE II

AVERAGE MSE RESULTS FOR THE FIRST SCENARIO.

MSE ($\times 10^{-3}$)	$\hat{s}_1$	$\hat{s}_2$
opt-aiNet	4.7781	2.0373
memetic	0.5753	0.1149

experiments is carried out, but now using the parameters: $N_0 = 10$, $N_c = 7$, $\beta = 60$, $\sigma_s = 3$. Considering that the search space for the opt-aiNet is much larger than that associated with the memetic algorithm, the maximum number of interactions for the first algorithm was set to 15000, and to 3000 interactions for the second.
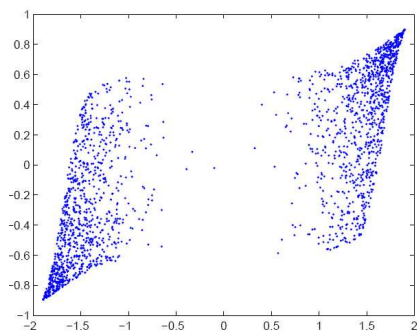
(a) Mixed signals.
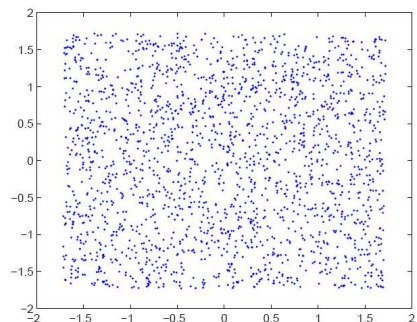
(b) Recovered signals FastICA.

Fig. 2.   Observed and recovered signals in the first scenario.

The average speed, the average time spent per iteration[1] and the average time to converge for both algorithms are depicted in Tab. I. In addition to that, the resulting mean-square errors (MSE) between each source and their estimates are shown in Tab.II.

Clearly, the use of the FastICA algorithm as a local search tool greatly improved the results obtained with the standard

[1]All the simulations were performed in the same machine, an Athlon64 3000+ with 1GB RAM

opt-aiNet method. It can be noticed that the memetic algorithm was able to reduce the number of iterations to converge, but each iteration time was increased due to the FastICA algorithm. Nevertheless, the total time spent by the new algorithm is 4 times smaller than the original proposal.

As far as the estimation error is concerned, it is observed that the memetic approach led to an improvement in the quality of the solution. Due to the local search capability introduced by the FastICA, the new method is able to refine the solution obtained solely with the opt-aiNet algorithm. Fig. 2 shows the mixed and recovered signals using the memetic approach. From this figure, it becomes clear that the separating system was able to restore the uniformly distributed sources in this situation.

B. Second scenario

$$\mathbf{A} = \begin{bmatrix} 1 & 0.6 & 0.5 \\ 0.5 & 1 & 0.4 \\ 0.4 & 0.6 & 1 \end{bmatrix} \text{ and } \begin{array}{rcl} f_1(e_1) &=& 2\sqrt[3]{e_1} \\ f_2(e_2) &=& 2\sqrt[3]{e_2} \\ f_3(e_3) &=& 2\sqrt[3]{e_3} \end{array}.$$

Using the same approach as in the first scenario, the opt-aiNet parameters were set to $N_0 = 10$, $N_c = 5$, $\beta = 50$, $\sigma_s = 3$. As in the previous case, the maximum number of iterations for the opt-aiNet and the memetic algorithms were set to 15000 and 3000, respectively.

Once again, one can see in Tab. III that the memetic technique provided a significant decrease in the convergence time, showing that the approach with a local search tool is very effective. In this case, the reduction in the convergence time is even more pronounced than in the first scenario. Tab. IV shows the estimation error obtained with both methods, also indicating a good performance of the new proposal.

TABLE III

RESULTS RELATED TO CONVERGENCE AND SIMULATION TIME IN THE
FIRST SCENARIO.

	opt-aiNet	memetic
Convergence speed (iterations)*	13750	950
Time spent per iteration (ms)	21	28
Average time to converge (min)	48	8

* Average number of iterations spent to reach the algorithm
convergence.

TABLE IV

AVERAGE MSE RESULTS FOR THE SECOND SCENARIO.

MSE ($\times 10^{-3}$)	$\hat{s}_1$	$\hat{s}_2$	$\hat{s}_3$
opt-aiNet	7.8529	0.4387	0.2187
memetic	1.0332	2.3466	6.0278

V. FINAL COMMENTS AND CONCLUSIONS

When dealing with nonlinear BSS problems, researchers usually face two problems:1) avoidance of local convergence; and 2) the need for an efficient method for evaluating the fitness function. This work proposes a method that is capable of dealing well with these requirements (with the aid of an

artificial immune network and of the order-statistics estimation paradigm), and that, furthermore, contains a FastICA-based local search step that significantly enhances the convergence performance. This method suits nicely the PNL - Post-NonLinear - systemic context, because it is possible to think of the separating system as being formed by stages that may be individually adapted using the global and local search algorithms. The simulations were analyzed having in mind the idea of establishing satisfactory grounds for comparison between this technique and the standard opt-aiNet algorithm proposed in [6]. The results show that the novel (memetic) solution is able to reach a fine tradeoff between performance, speed of convergence and computational effort, which leads us to the conclusion that the idea of allying an evolutionary algorithm to a FastICA-based refinement stage may be decisive in the process of implementing practical and efficient nonlinear separating systems.

REFERENCES

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, 2001.

[2] J. Hérault, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *Proceedings of 10th GRETSI Symposium on Signal and Image Processing*, Nice, May 1985, pp. 1017–1022.

[3] A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing*, John Wiley & Sons, 2002.

[4] A. Taleb and C. Jutten, "Source separation in postnonlinear mixtures," *IEEE Trans. Signal Processing*, vol. 47, no. 10, pp. 2807–2820, October 1999.

[5] S. Achard and C. Jutten, "Identifiability of Post-Nonlinear Mixtures," *IEEE Signal Processing Letters*, vol. 12, no. 5, pp. 423–426, May 2005.

[6] L. T. Duarte, R. Suyama, R. R. F. Attux, F. J. Von Zuben, and J. M. T. Romano, "Blind Source Separation of Post-nonlinear Mixtures Using Evolutionary Computation and Order Statistics," *LNCS 3889*, vol. 6, no. 3889, pp. 66–73, 2006.

[7] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[8] A. Hyvärinen and E. Oja, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[9] F. Rojas, I. Rojas, R. M. Clemente, and C. G. Puntonet, "Nonlinear Blind Source Separation Using Genetic Algorithms," in *Proc. of the 3rd Int. Conf. on Independent Component Analysis and Blind Signal Separation (ICA2001)*, San Diego, California, USA, December 2001, pp. 400–405.

[10] Y. Tan and J. Wang, "Nonlinear blind source separation using higher-order statistics and a genetic algorithm," *IEEE Trans. on Evolutionary Computation*, vol. 5, no. 6, pp. 600–612, December 2001.

[11] R. R. F. Attux, M. B. Loiola, R. Suyama, L. N. de Castro, F. J. Von Zuben, and J. M. T. Romano, "Blind Search for Optimal Wiener Equalizers Using an Artificial Immune Network Model," *EURASIP Journal on Applied Signal Processing – Special Issue on Genetic and Evolutionary Computation for Signal Processing and Image Analysis*, vol. 2003, no. 8, pp. 740–747, 2003.

[12] L. N. de Castro and F. J. Von Zuben, "Learning and Optimization Using the Clonal Selection Principle," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 3, pp. 239–251, June 2002.

[13] N. Krasnogor and J. Smith, "A Tutorial for Competent Memetic Algorithms: Model, Taxonomy, and Design Issues," *IEEE Trans. on Evolutionary Computation*, vol. 9, no. 5, pp. 474–488, October 2005.

[14] D.-T. Pham, "Blind Separation of Instantaneous Mixtures of Sources Based on Order Statistics," *IEEE Trans. Signal Processing*, vol. 48, no. 2, pp. 363–375, February 2000.