An iterative-recursive SOS-based method for separation of Post-Nonlinear Mixtures

Caroline P. A. Moraes, Aline Neves, Denis G. Fantinato

***Abstract*— Independent Component Analysis (ICA) methods are widely used in the linear Blind Source Separation (BSS) problem. Nevertheless, for some practical cases, the linear assumption is not valid, requiring nonlinear mixing models. The Post-Nonlinear is one of the few nonlinear models in which ICA is able to perform source separation. In previous works, an iterative SOS-based algorithm was proposed, combining elements of AMUSE and SOBI. However, instantaneous estimations of the correlation matrix could lead to a loss of performance. In that sense, in this work, we modify the previous algorithm to use an iterative-recursive estimation of the correlation matrices. Due to the SOS-based approach, sources should present a temporal dependency and certain constraints are required on the separation structure. Results show a good performance of the proposed algorithm.**

***Keywords*— Blind Source Separation, Post-Nonlinear, Second-Order Statistics.**

I. INTRODUCTION

Blind Source Separation (BSS) is an important problem within the signal processing area. Essentially, it aims at retrieving the original signals from a set of observed mixed signals [1], [2]. In BSS, both sources and the mixing process are unknown, which leads to an unsupervised approach based only on samples of the mixed signals and statistical characteristics of the sources. At first, research was mainly focused on the linear case, presenting a solid theoretical framework and counting with applications in several areas, such as biomedical signal processing [3], communication systems [1] and geophysical signal analysis [4]. Without any prior information, the problem is not solvable [1], [3]. In that sense, it is usually assumed that sources are mutually independent, which gave rise to the Independent Component Analysis (ICA), a set of techniques widely used for solving the BSS problem [2], such as Infomax, FastICA, Bell-Sejnowski and JADE [1], [2]. Since measuring independence often involves working with Mutual Information, most of the algorithms have an inherent dependence on Higher-Order Statistics (HOS), which leads to certain difficulties, such as high complexity or the necessity of a large number of samples [1]. On the other hand, simpler algorithms were proposed based on Second-Order Statistics (SOS) for solving the linear BSS problem, exploring the temporal dependency among samples, like AMUSE [1], SOBI [5] and TDSEP [2].

For some applications, however, the linear mixing assumption is not enough for retrieving the sources, since the mixing process encompasses nonlinear transformations, as occurs in hyperspectral imaging [6] and chemical sensor arrays [7]. Under a generic nonlinear BSS perspective, ICA methods are intended to fail. In that sense, the studies in this area are focused on constrained nonlinear models where the ICA techniques are still valid, such as linear quadratic, Bilinear and Post-Nonlinear (PNL) [8], [9] models.

Great efforts have been dedicated to the nonlinear BSS problem, including the recent extensions of the SOS-based new methods, assuming temporally dependent sources [9], [10]. In [11] and [12], the authors present some algorithms developed from SOS-based criteria applied to the PNL model, showing that it is possible to use these methods, under certain constraints. However, this approach still lacks an algorithm simpler and robust. The A-SOBIPNL was proposed in [13], consisting of an SOS-gradient-based algorithm for separation of PNL mixtures, using elements of the classical AMUSE and SOBI algorithms. The initial tests showed very good results in terms of SIR ($85\ dB$), but considering reasonably simple simulations scenarios. For a deeper analysis, in [14] the A-SOBIPNL was tested in scenarios with more stringent nonlinear functions, showing promising results. Nevertheless, the instantaneous estimations of the correlation matrices could lead to a loss of performance. In this work, we propose a modification in the A-SOBIPNL, estimating the correlation matrices recursively. As we will show, such a new approach improves the performance of the algorithm and also makes it more efficient in terms of computational cost, since working with a small number of samples is sufficient to obtain a good result.

In Section II, we describe the BSS problem and define the PNL model. In Section III, the recent results involving SOS-based algorithms that motivates this work are presented. The modified SOS-based algorithm is proposed in Section IV. The simulation results are presented and analyzed in Section V. Finally, Section VI concludes the work and presents future perspectives.

II. NONLINEAR BLIND SOURCE SEPARATION

In the following, we describe the general concept of the nonlinear BSS problem.

A. Post-Nonlinear Model

The Post-Nonlinear model assumes that, in the mixing process, there exists a nonlinear component-wise function $\mathbf{f}(\cdot) = [f_1(\cdot)\ \ldots\ f_M(\cdot)]^T$ that is invertible and monotonic. The mixing process can be modeled by:

$$\mathbf{x}(n) = \mathbf{f}(\mathbf{A}\mathbf{s}(n)), \tag{1}$$

Caroline P. A. Moraes, e-mail: caroline.moraes@ufabc.edu.br;Denis G. Fantinato, e-mail: denis.fantinato@ufabc.edu.br; Aline Neves, e-mail: aline.neves@ufabc.edu.br. UFABC - Federal University of ABC, Santo André - SP. This work was partially financial supported by UFABC.

where n is the time index, $\mathbf{x}(n)$ is the M-dimensional observation vector $\mathbf{x}(n) = [x_1(n), x_2(n), \ldots, x_M(n)]^T$, $\mathbf{s}(n)$ is the vector of unknown sources $\mathbf{s}(n) = [s_1(n), s_2(n), \ldots, s_N(n)]^T$, where N is the number of sources and $\mathbf{A}$ is the mixing matrix with dimension $M \times N$.

The aim is to find the matrix $\mathbf{W}$ and the nonlinear function $g(\cdot)$ that could retrieve the sources in a new vector $\mathbf{y}(n) = [y_1(n), y_2(n), \ldots, y_N(n)]^T$, so that:

$$\mathbf{y}(n) = \mathbf{W}\mathbf{g}\left(\mathbf{x}(n)\right), \tag{2}$$

where $\mathbf{W}$, with dimension $M \times N$, ideally is equal to the inverse of $\mathbf{A}$ and $\mathbf{g}(\cdot)$ is a vector with column-wise nonlinear functions, ideally the inverse of $\mathbf{f}(\cdot)$ (hence, strictly monotonic) [1], both admitting scale invariance and/or permutation between the signals. The system model is illustrated in Fig. 1.

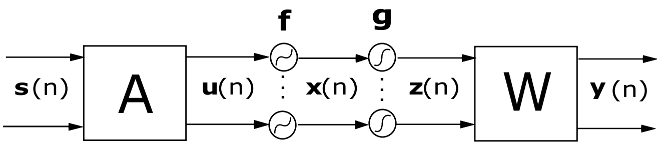

Fig. 1. Post-Nonlinear Mixture Model.

By applying an ICA method, it is expected that $\mathbf{W}$ and $g(\cdot)$ be adjusted so that signals in $\mathbf{y}(n)$ be mutually independent.

III. Second-Order Statistics in the PNL Model

In the linear BSS problem, there are several SOS-based methods that could be applied to adjust $\mathbf{W}$, for instance, TDSEP or WASOBI [1], [3]. However, for the nonlinear case, the use of SOS is still incipient. For PNL mixtures, initial approaches considered SOS along with HOS, adapting matrix $\mathbf{W}$ and nonlinear function $g(\cdot)$ separately. In these cases, the coefficients of the nonlinear function $g(\cdot)$ are adjusted using HOS by exploiting the "Gaussian" effect [15] or by using a priori information [9], [16], while $\mathbf{W}$ is adapted using SOS-based approaches.

Recently, new methods were proposed using the SOS to adjust both linear and nonlinear parts of the PNL model – $\mathbf{W}$ and $g(\cdot)$, respectively [12], [15], [13], [14]. In this case, source separability can be ensured if, besides classical linear BSS conditions are met [2], [3], a linear component in $\mathbf{z}(n) = \mathbf{g}(\mathbf{x}(n))$ do exist. In other words, for a given (and usually unknown) $f(\cdot)$, $g(\cdot)$ must be constrained so that the output be

$$\mathbf{y}(n) = \mathbf{W}\mathbf{g} \circ \mathbf{f}\left(\mathbf{W}\mathbf{A}\mathbf{s}(n)\right) = \mathbf{W}\mathbf{A}\mathbf{s}(n) + \mathbf{W}\tilde{\mathbf{f}}\left(\mathbf{A}\mathbf{s}(n)\right), \tag{3}$$

i.e., the composite function $\mathbf{g} \circ \mathbf{f}(\cdot)$ results in a linear component plus a nonlinear residual $\tilde{\mathbf{f}}(\cdot)$ [11]. In this case, the residue $\tilde{\mathbf{f}}(\cdot)$ can be eliminated by the joint diagonalization of correlation matrices of the recovered sources, for sufficient different time delays. However, a necessary condition is that the linear component in (3) be not suppressed [12].

IV. The MA-SOBIPNL Algorithm

The approach applied in [12] is based on the use of metaheuristics for parameters optimization and is a relevant step towards the use of the SOS in the nonlinear BSS problem. However, this method is computationally costly, mainly due to the populational-based approach.

The work developed in [13] presents a first result in the sense of searching for a simpler algorithm based only on SOS for the PNL model. The method combines two classical SOS-based approaches: AMUSE is used to separate the linear stage of the mixture and SOBI is used to separate the nonlinear stage (being named A-SOBIPNL algorithm). First results shown in [13], [14] suggests that the algorithm can perform well, as long as Eq. (3) is satisfied. However, the algorithm used instantaneous estimations of the correlation matrices, taking into account only the current sample window. In this paper, the algorithm presented in [13] is modified, considering a recursive estimation of the correlation matrices with the objective of increasing the algorithms performance. In this section, we will present MA-SOBIPNL, with its proposed modification.

A. Linear Stage

The AMUSE algorithm is a technique for diagonalization of autocorrelation matrices, which is based on the properties of eigenvalues and eigenvectors [2]. This algorithm requires an initial stage of data whitening, which is usually not used in nonlinear models. However, in A-SOBIPNL, $\mathbf{z}(n)$ is firstly whitened – $\mathbf{z}(n)$ is ideally composed only of linearly mixed signals [13]. The whitening process is based on the correlation matrix of $\mathbf{z}(n)$, whose estimation can be done as

$$\tilde{\mathbf{R}}_Z(d, n) = \frac{1}{L}\sum_{i=0}^{L-1} \mathbf{z}(n-i)\mathbf{z}^T(n-i-d), \tag{4}$$

where L is the number of samples available in a sliding time window and d is a given delay.

Initially, the whitening matrix $\mathbf{V}$ is obtained by estimation of $\tilde{\mathbf{R}}_Z(0, n)$ and then computing $\mathbf{V} = \mathbf{E}\ \mathbf{D}^{-1/2}\ \mathbf{E}^T$, where $\mathbf{D}$ is a diagonal matrix with the eigenvalues and $\mathbf{E}$ is a matrix with the eigenvectors of $\tilde{\mathbf{R}}_Z(0, n)$. Once $\mathbf{V}$ is obtained, the whitening process results in $\mathbf{z}'(n) = \mathbf{V}z(n)$.

In order to improve the quality of the correlation matrix estimation, we use a recursive approach along with the sliding time window and apply a forgetting factor (λ):

$$\begin{aligned}\tilde{\mathbf{R}}_{Z'}(d, n) &= \lambda\tilde{\mathbf{R}}_{Z'}(d, n-1) \\ &\quad + \frac{(1-\lambda)}{(L-d)}\sum_{i=0}^{L-1} \mathbf{z}(n-i)\mathbf{z}^T(n-i-d).\end{aligned} \tag{5}$$

After this process, AMUSE uses the eigenvectors of $\tilde{\mathbf{R}}_{Z'}(d, n)$, with $d \neq 0$ to obtain a separating matrix $\mathbf{W}$, such that $\mathbf{y}(n) = \mathbf{W}\mathbf{z}'(n)$ is mutually uncorrelated.

B. Nonlinear Stage

The nonlinear function $g(\cdot)$ may be written as a linear combination of powers of $\mathbf{x}(n)$ [11]:

$$g_k(x_k(n)) = \mathbf{c}_k^T \, \boldsymbol{\xi}_k(n), \qquad k \in \{1, \ldots, M\} \tag{6}$$

where $\mathbf{c}_k = [c_{k1} \; c_{k2} \; \ldots \; c_{kb}]^T$ are the coefficients of the k-th function $g_k(\cdot)$ with $k = 1, \ldots, M$ and $\boldsymbol{\xi}_k(n) = [x_k^1(n) \; x_k^2(n) \; \ldots \; x_k^b(n)]^T$ is the vector with the k-th mixture to the power of 1 to b, where b is the highest admitted degree.

To adjust the nonlinear part, i.e. the coefficients of $g_k(\cdot)$, SOBI criterion is used [5], [13], whose cost function can be expressed by:

$$J(\mathbf{W}, \mathbf{c}) = \sum_{d \in D} off(\mathbf{R}_Y(d)), \tag{7}$$

where D is the set of delays to be considered, $\mathbf{R}_Y(d) = E\left[\mathbf{y}(n)\, \mathbf{y}^T(n-d)\right]$ and $off(\mathbf{R}_Y(d))$ is

$$off(\mathbf{R}_Y(d)) = \sum_{i \neq j} r_{ij}^2, \tag{8}$$

where i and j represent the row ($i \in \{1, \ldots, N\}$) and the column ($j \in \{1, \ldots, M\}$) of $\mathbf{R}_Y(d)$, respectively. Here, differently from [13], $\mathbf{R}_Y(d)$ will also be estimated recursively:

$$\begin{aligned}\tilde{\mathbf{R}}_{Y'}(d,n) &= \lambda \tilde{\mathbf{R}}_{Y'}(d,n-1) \\ &\quad + \frac{(1-\lambda)}{(L-d)} \sum_{i=0}^{L-1} \mathbf{y}(n-i)\mathbf{y}^T(n-i-d).\end{aligned} \tag{9}$$

According to Eq. (7), ideally, the greater the number of delays used, the more information can be considered by the algorithm, contributing for a better separation.

As adaptation rule, the gradient descent method was used to adjust the coefficients $\mathbf{c}_k$ [13]:

$$\mathbf{c}_k(n+1) = \mathbf{c}_k(n) - \mu \, \frac{\partial J(\mathbf{W}, \mathbf{c})}{\partial \mathbf{c}_k}, \tag{10}$$

where μ is the adaptation step and

$$\frac{\partial J(\mathbf{W}, \mathbf{c})}{\partial \mathbf{c}_k} = \sum_{d \in D} \sum_{i \neq j} \frac{\partial r_{ij}^2(d)}{\partial \mathbf{c}_k}, \tag{11}$$

with

$$\begin{aligned}\frac{\partial r_{ij}^2(d)}{\partial \mathbf{c}_k} &= 2 r_{ij}(d) \left(E\left[\boldsymbol{\xi}_k(n) w_{ik} y_j(n-d)\right.\right. \\ &\qquad \left.\left. + \boldsymbol{\xi}_k(n-d) w_{jk} y_i(n)\right]\right),\end{aligned} \tag{12}$$

where w_{ij} are the coefficients of the matrix $\mathbf{W}$ at line i and column j, respectively.

In order to ensure the validity of Eq. (3), a constraint must be applied to the $\mathbf{c}_k$ coefficients after (10). Following what was done in [13], we will keep the coefficients of c_{k1} non-null and positive, for all k, to preserve the linear term in (3), so that they are not suppressed during the adaptation. This approach requires some kind of partial knowledge about the non-linearity $\mathbf{f}(\cdot)$ and, in this case, $c_{k1} > 0$ implies that $g_k(\cdot)$ will be an increasing monotonic function.

The steps of this modified algorithm, named MA-SOBIPNL, are summarized in Alg. 1. In the first iteration (i.e. $n = 0$), we assume $\lambda = 0$, since we have no previous information about the correlation matrix. Note that, if $\lambda = 0$, we have the previous A-SOBIPNL.

Algorithm 1 MA-SOBIPNL Algorithm

Initialization of MA-SOBIPNL parameters:
d, μ, L; $\mathbf{c} \leftarrow \mathbf{1}$ ($\mathbf{g}(\cdot)$ function); $\mathbf{W} \leftarrow \mathbf{0}$;
for each sliding window **do**
  $\mathbf{z}(n) \leftarrow \mathbf{g}(\mathbf{x}(n))$;
  Linear Stage:
  Whitening of $\mathbf{z}(n)$:
  $\mathbf{z}'(n) \leftarrow \mathbf{V}\mathbf{z}(n)$;
  Recursive estimation of the correlation matrix:

$$\begin{aligned}\tilde{\mathbf{R}}_{Z'}(d,n) &= \lambda \tilde{\mathbf{R}}_{Z'}(d,n-1) \\ &\quad + \frac{(1-\lambda)}{(L-d)} \sum_{i=0}^{L-1} \mathbf{z}(n-i)\mathbf{z}^T(n-i-d);\end{aligned}$$

  Update of $\mathbf{W}$ according to AMUSE;
  $\mathbf{y}(n)$ update : $\mathbf{y}(n) \leftarrow \mathbf{g}(\mathbf{W}\mathbf{z}'(n))$;
  Nonlinear Stage:
  Estimation of $\tilde{\mathbf{R}}_Y(d,n)$ with L samples window:
  for each delay $d \in D$ **do**

$$\begin{aligned}\tilde{\mathbf{R}}_{Y'}(d,n) &= \lambda \tilde{\mathbf{R}}_{Y'}(d,n-1) \\ &\quad + \frac{(1-\lambda)}{(L-d)} \sum_{i=0}^{L-1} \mathbf{y}(n-i)\mathbf{y}^T(n-i-d);\end{aligned}$$

    Gradient estimation by (11) and (12);
  end for
  Coefficients adaptation of $\mathbf{g}(\cdot)$ by Eq. (10);
  Constraints on $\mathbf{g}(\cdot)$ coefficients and normalization;
end for

V. Performance Analysis

A. Simulation Scenario

In this paper we focus on analysing the properties of the MA-SOBIPNL algorithm. Thus, we will consider the cases in which the constraint upon $\mathbf{c}_k$ is or is not applied, situation that could violate (3).

In order to study the performance and the effects of the MA-SOBIPNL algorithm with and without the constraints on the nonlinear function $\mathbf{g}(\cdot)$, in this section we present some simulations results. We consider two independent sources, one with Gaussian distribution (zero mean and unit variance) and the second has an uniform distribution with a range from -1 to 1. The temporal dependence between the samples were inserted in each one of the independent sources through the use of FIR (Finite Impulse Response) filters, with impulse responses: $h_1(z) = 1 + 0.6z^{-1} - 0.2z^{-2} + 0.6z^{-3}$ and $h_2(z) = 1 - 0.4z^{-1} - 0.3z^{-2} + 0.2z^{-3}$. The output of $h_1(z)$ is the source $s_1(n)$ while the output of $h_2(z)$ is $s_2(n)$. Both were normalized.

In the linear mixing stage, the sources were mixed by the matrix $\mathbf{A} = [0.65 \; 0.23; 0.35 \; 0.76]$ and, in the nonlinear stage, $\mathbf{f}(\cdot)$ is

$$\mathbf{f}(\mathbf{u}(n)) = \operatorname{arctanh}(\mathbf{u}(n)) = \operatorname{arctanh}(\mathbf{A}\mathbf{s}(n)), \tag{13}$$

i.e., an inverse hyperbolic tangent function. The chosen separation structure assumes that $\mathbf{W}$ is a 2×2 matrix and $\mathbf{g}(\cdot)$ is, component wise, given by:

$$g_k(x_k(n)) = c_{k1}x_k(n) \; + \; c_{k2}x_k^3(n) \; + c_{k3}x_k^5(n). \qquad (14)$$

The application of the constraint on the coefficients $\mathbf{c}_k$, is possible if there is some information about the function $\mathbf{f}(\cdot)$. In this case, the function has features of an inverse hyperbolic tangent function. In that sense, to keep the validity of the Eq. (3), after each adaptation, we apply $c_{k1} \leftarrow 1$, for $k = \{1, 2\}$. In addition, the coefficients $\mathbf{c}_k$ were normalized to keep the energy of $\mathbf{z}(n)$ constant.

In all cases, an average of 100 independent simulations is considered, being the algorithms performance measured in terms of SIR (*Signal-to-Interference Ratio*), defined as:

$$SIR_{dB} = 10 \log \left(\frac{E[y_i(n)^2]}{E[(s_i(n) - y_i(n))^2]} \right). \qquad (15)$$

B. Performance Analysis - Forgetting factor

The forgetting factor is a parameter that influences the performance of the MA-SOBIPNL algorithm. Firstly, we varied the value of λ from 0.5 to 0.99 and fixed the delays in $d = 3$, using a sliding window of $L = 5$ samples and $\mu = 0.05$.

Firstly analysing λ, Fig. 2 shows the case with constraint and Fig. 3 the case without constraint – moving average was applied in SIR values for curve smoothing. It is noticeable that the algorithm converges to higher values of SIR using $\lambda = 0.99$, reaching $47,3\ dB$ and $44,7\ dB$, respectively. In addition, Fig. 2 shows an interesting relation between λ and SIR: SIR increases as the value of λ increases, showing how the use of a memory on the estimation of the correlation matrices, given by the term $R(d, n-1)$ in (5) and (9), improves the quality of the estimation and the performance of the algorithm. The relation between these parameters in Fig. 3 is not so regular as in the previous case, but for values above 0.8 the SIR values also increase with λ.

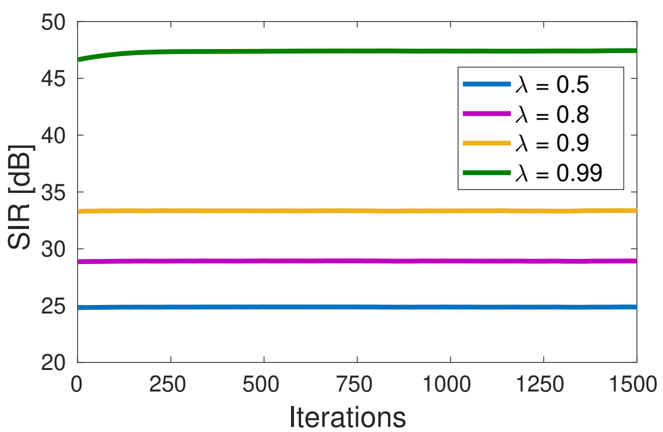

Fig. 2. The convergence analysis with different forgetting factor - Case with constraint

Comparing both figures based on the best results of each one, it is possible to note that the algorithm with constrains converges faster than that without constraints. In Fig. 2 the algorithm only requires 200 iterations to converge while for the case without constraint, the algorithm demands around 2850 iterations to converge, as is shown in Fig. 3. These figures show how unpredictable and occasionally slower the results can be when the restrictions are not applied for each scenario. For the case without constraints in Fig. 3, the solutions did not violate the separation conditions given by (3), resulting in a good performance.

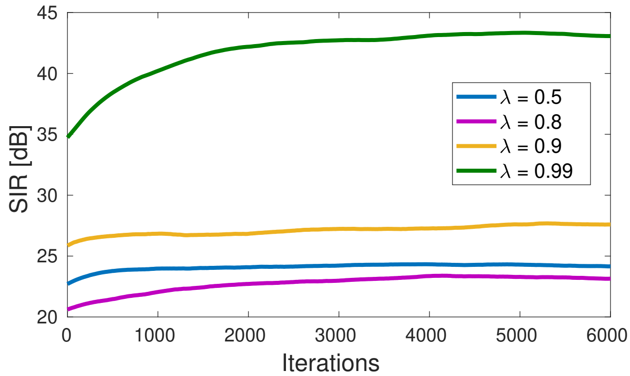

Fig. 3. The convergence analysis with different forgetting factor - Case without constraint

C. The MA-SOBIPNL × A-SOBIPNL

In this second analysis, we compare the performances of the previous A-SOBIPNL[13] with those of the MA-SOBIPNL algorithm proposed in this paper, with and without the constraint of $c_{k1} \leftarrow 1$.

Fig. 4 presents the results with constraints where the A-SOBIPNL is represented by $\lambda = 0$ and based on the previous analyses, in the MA-SOBIPNL we used $\lambda = 0.99$. The number of delays were fixed in $d = 3$, using a sliding window of $L = 5$ samples and $\mu = 0.05$. We verified that the MA-SOBIPNL converges to greater values of SIR than the A-SOBIPNL, i.e., $47,3\ dB$ and $23,2\ dB$, respectively. Additionally, we may observe that even having small values for d and L the MA-SOBIPNL performs very well while the A-SOBIPNL is not able to achieve the same performance with such low values of d and L.

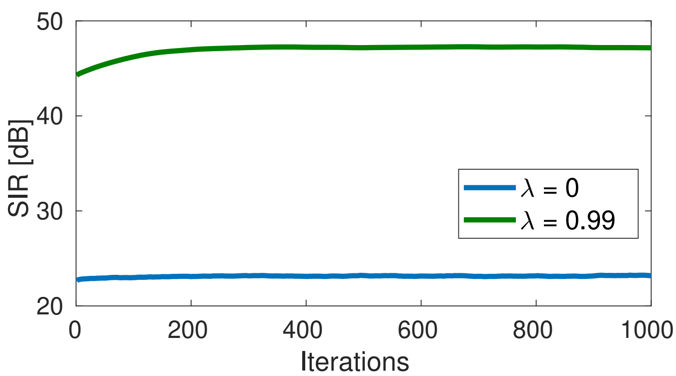

Fig. 4. The convergence in terms of SIR - Case with constraint

For the case without constraint, the influence of the number of delays and window length were very relevant. Fig. 5 shows the results based on three different arrangements of these parameters. The step-size considered was $\mu = 0.05$ and we also compare the performance of the two algorithms. In the first arrangement, we fixed $d = 1$ and $L = 5$. In second we increased the window length for $L = 10$ and the third considered $d = 3$ and $L = 5$. As shown in Fig. 4, the MA-SOBIPNL presented very good values of SIR ($44,7\ dB$) using a small d and L while the A-SOBIPNL needs a larger sliding window to improve its performance ($27,4\ dB$).

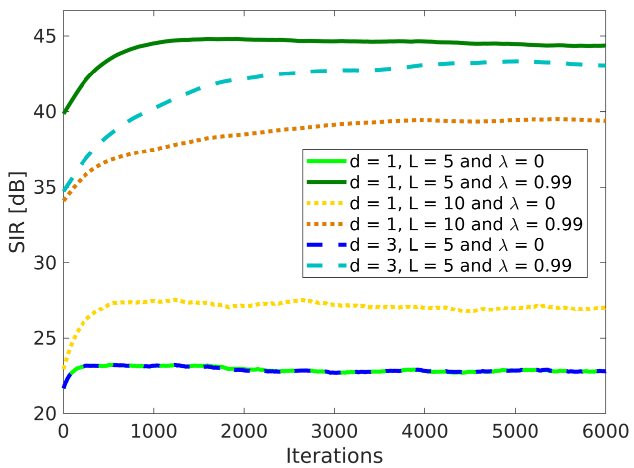

Fig. 5. The convergence in terms of SIR - Case without constraint

Although the algorithm requires more iterations to converge, even in the worst parameters combination (with $d = 1$ and $L = 10$) the MA-SOBIPNL obtained a SIR of $39,5\ dB$ which is better than A-SOBIPNL with its higher SIR value. After several different arrangements, all combinations lead the MA-SOBIPNL to achieve higher SIR levels than those of the previous A-SOBIPNL.

VI. Conclusions

In this work, we propose a modification of the A-SOBIPNL algorithm, which is a method for nonlinear BSS based on Second-Order Statistics and the PNL model. This approach explores the temporal structure of the signals, using two classical SOS algorithms, AMUSE and SOBI. This leads to a very simple gradient-descent method. In the MA-SOBIPNL algorithm proposed in this paper, the correlation matrix estimation is done through a recursive approach, exploiting previous estimations, which leads to an improve in performance.

Simulations results demonstrated the relevance of including some constraints in the adaptation of the nonlinear function parameters. In the evaluated scenarios using the constraints, the algorithm reaches a robust value of SIR ($47,3\ dB$), while in the case without restrictions the value of SIR is unstable, convergence is slower and reaches lower SIR levels. Moreover, the results showed the influence of the forgetting factor in the MA-SOBIPNL. The right choice of this parameter can improve source retrieving compared to the original A-SOBIPNL algorithm. In the future, we consider extending the work using more complex scenarios and applying this method in practical cases.

Acknowledgements

The authors would like to thank University of ABC (UFABC) by the financial support to the work.

References

[1] Comon P. and Jutten C., "Handbook of Blind Source Separation: Independent component analysis and applications," *Academic press*, 2010.

[2] Hyvärinen A., Karhunen J. and Oja E., "Independent component analysis," vol.46, *John Wiley & Sons*, 2004.

[3] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, "Unsupervised signal processing: channel equalization and source separation," in *CRC Press*, 2018.

[4] A. K. Takahata, E. Z. Nadalin, R. Ferrari, L. T. Duarte, R. Suyama, R. R. Lopes, J. M. Romano, and M. Tygel, "Unsupervised processing of geophysical signals: A review of some key aspects of blind deconvolution and blind source separation," in *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 27–35, 2012.

[5] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," in *IEEE Transactions on signal processing*, vol. 45, no. 2, pp. 434–444, 1997.

[6] I. Meganem, Y. Deville, S. Hosseini, P. Deliot, and X. Briottet, "Linearquadratic blind source separation using nmf to unmix urban hyperspectral images," in *IEEE Transactions on signal processing*, vol. 62, no. 7, pp. 1822–1833, 2014.

[7] L. T. Duarte, C. Jutten, and S. Moussaoui, "A bayesian nonlinear source separation method for smart ion-selective electrode arrays," in *IEEE Sensors Journal*, vol. 9, no. 12, pp. 1763–1771, 2009.

[8] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures," in *IEEE Transactions on signal Processing*, vol. 47, no. 10, pp. 2807–2820, 1999.

[9] Y. Deville and L. T. Duarte, "An overview of blind source separation methods for linear-quadratic and post-nonlinear mixtures," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 155–167, Springer, 2015.

[10] L. Jarboui, Y. Deville, S. Hosseini, R. Guidara, A. B. Hamida, and L. T. Duarte, "A second-order blind source separation method for bilinear mixtures," in *Multidimensional Systems and Signal Processing*, vol. 29, no. 3, pp. 1153–1172, 2018.

[11] D. G. Fantinato, L. T. Duarte, Y. Deville, C. Jutten, R. Attux, and A. Neves, "Using taylor series expansions and second-order statistics for blind source separation in post-nonlinear mixtures," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 193–203, Springer, 2018.

[12] D. G. Fantinato, L. T. Duarte, Y. Deville, R. Attux, C. Jutten, and A. Neves, "A second-order statistics method for blind source separation in post-nonlinear mixture," in *Signal Processing*, vol. 155, pp. 63–72, 2019.

[13] J. Saldanha, C. P. A. Moraes, A. Neves, and D. G. Fantinato, "Algoritmo baseado em estatísticas de segunda ordem para separação cega de fontes com misturas não lineares," in *2019 Brazilian Technology Symposium*, pp. 1–5, 2019.

[14] Moraes, C. P. A., Saldanha, J., Attux, R., Duart, L. T., Neves, A. and Fantinato, D. G., "An SOS-Based Algorithm for Source Separation in Nonlinear Mixtures," to be published in *IEEE Statistical Signal Processing Workshop 2021*, pp. 1–5, 2021.

[15] A. Ziehe, M. Kawanabe, S. Harmeling, and K.-R. MÃžller, "Blind separation of post-nonlinear mixtures using linearizing transformations and temporal decorrelation," in *Journal of Machine Learning Research*, vol. 4, no. Dec, pp. 1319–1338, 2003.

[16] D. G. Fantinato, L. T. Duarte, P. Zanini, B. Rivet, R. Attux, and C. Jutten, "A joint second-order statistics and density matching-based approach for separation of post-nonlinear mixtures," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 499–508, Springer, 2017.