INFORMED SEPARATION OF DEPENDENT SOURCES USING JOINT MATRIX DECOMPOSITION

A. Boudjellal[1] *K. Abed-Meraim*[1] *A. Belouchrani*[2] *Ph. Ravier*[1]

[1] Univ. of Orléans, PRISME Laboratory, France; [2] Ecole Nationale Polytechnique, Algiers, Algeria.

ABSTRACT

This paper deals with the separation problem of dependent sources. The separation is made possible thanks to side information on the dependence nature of the considered sources. In this work, we first show how this side information can be used to achieve desired source separation using joint matrix decomposition techniques. Indeed, in the case of statistically independent sources, many BSS methods are based on joint matrix diagonalization. In our case, we replace the target diagonal structure by appropriate non diagonal one which reflects the dependence nature of the sources. This new concept is illustrated with two simple 2×2 source separation examples where second-order-statistics and high-order-statistics are used respectively.

Index Terms— Informed Source Separation, Dependent Source Separation, Matrix Joint Decomposition, Alternating Least Squares, Second-order and High-order-Statistics.

1. INTRODUCTION

Blind Source Separation (BSS) is a class of methods that have been widely used and largely studied for more than thirty years. The key assumption for most of the developed BSS algorithms is to suppose that the sources are mutually independent. The independence information is proven to be sufficient for separating sources blindly [4]. Other BSS algorithms exploit, in addition to the independence information, one or more inherent properties characterizing the original sources (i.e., time coherence, cyclostationarity, sparsity, constant modulus, etc.) [2, 3, 12].

Developed for separating independent sources, most of the existing BSS algorithms present degraded performance when applied on dependent sources. Informed Source Separation (ISS) is a relatively new concept based on introducing side information on the considered sources in order to facilitate separating them later [5].

In this paper, we do not 'add' any additional information to the sources as done in [5]. We simply make use of the information on the nature of sources statistical dependencies, e.g. [11], in order to separate them. Indeed, the separation of dependent sources can be made possible under strong assumptions on the sources structures, e.g. [13, 14, 15] or by using some extra information on the sources, or their statistical dependency model as we do in this work.

Our objective is to focus on ISS using information about the form of dependencies between the sources. In particular, our aim is to generalize the matrix joint diagonalization tool, widely used in BSS [6], to be used in Matrix Joint Decomposition (MJD) where the targeted structure is a (known or partially known) non-diagonal matrix that reflects the nature of considered dependency.

2. DATA MODEL

The instantaneous mixture $\mathbf{x}_t$ of d transmitted sources $\mathbf{s}_t$ received through an m-antenna array is modeled as follows:

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t + \mathbf{n}_t. \tag{1}$$

were $\mathbf{A}$ is the $m \times d$ mixing matrix and $\mathbf{n}_t$ is an additive noise of covariance $\sigma_n^2 \mathbf{I}_m$. By stacking T samples of the received data in one matrix $\mathbf{X} = [\mathbf{x}_1 : \mathbf{x}_T]$, the model (1) becomes:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \tag{2}$$

Standard hypotheses consist of assuming that: (i) The mixing matrix $\mathbf{A}$ is a tall full column rank matrix ($m \geq d$), (ii) The additive noise is white, Gaussian, and independent from the source signals, (iii) The original sources are mutually independent. In our context, the last assumption is no longer valid. The considered sources are assumed to be statistically dependent with known dependency structure.

Informed Source Separation aims to recover the unknown dependent sources from observed mixtures, relying on some assumptions on the statistical properties of the original sources and using additional information about the nature of the dependency between the considered sources. Based on this information, we have to find a $d \times m$ separation matrix $\mathbf{W}$ which output is the estimated source vector (up to scaling and permutation ambiguities [9]):

$$\mathbf{Z} = \mathbf{W}\mathbf{X} = \widehat{\mathbf{S}}. \tag{3}$$

Next, the new concept of separating dependent sources based on the nature of the dependency will be illustrated in two contexts. In the first one, the separation of a mixture of time coherent auto-regressive sources is addressed. The separation algorithm makes use of the Second Order Statistics (SoS). In the second context, a parametric dependency model between the original i.i.d sources similar to the one in [1] is considered. High Order Statistics (HoS) are used for separating the considered sources.

3. SOS-BASED SEPARATION OF AUTO-REGRESSIVE DEPENDENT SOURCES

3.1. General Case: AR(p) Dependent Sources

We consider herein the case of temporally coherent sources with mutual dependencies of known parametric model. To better illustrate our concept, let's consider d auto-regressive (AR) sources of dependent innovation process, i.e.

$$s_i(n) = \sum_{k=1}^{p} b_i(k) s_i(n-k) + \xi_i(n), \qquad i = 1, ..., d \tag{4}$$

where p is the AR order, $\{b_i(k)\}_{1 \le k \le p}$ are the AR coefficients of the i-th source and ξ_i is its associated i.i.d innovation process.

Now, contrary to standard assumption [8], we consider here the situation where the source innovation processes are correlated, i.e. $\mathbf{R}_\xi = E\left(\xi(n)\xi(n)^H\right)$ is a non diagonal matrix, with $\xi(n) = [\xi_1(n), ..., \xi_d(n)]^T$.

Under these data model assumptions, the power spectral density (psd) function of the observed source mixtures can be written is the noiseless case as follows:

$$\mathbf{S}_x(f) = \mathbf{A}\ \mathbf{D}(f, \theta)\ \mathbf{A}^H, \qquad \forall f \tag{5}$$

where the (i, j)-th entry of $\mathbf{D}(f, \theta)$ is given by:

$$\mathbf{D}_{ij}(f, \theta) = \frac{\mathbf{R}_\xi(i, j)}{\mathbf{B}_i(f)\mathbf{B}_j(f)^*} \tag{6}$$

with $\mathbf{B}_i(f) = 1 - \sum_{k=1}^{p} b_i(k)\ e^{-\jmath 2\pi k f}$ and $\mathbf{R}_\xi(i, j)$ is the (i, j)-th entry of the innovation covariance matrix $\mathbf{R}_\xi$. θ represents here the set of parameters modeling the sources psd matrix $\mathbf{D}(f, \theta)$.

If the parameter θ is known (or a prior estimated) one can estimate the mixing matrix $\mathbf{A}$ by joint decomposition of a set of psd matrices $\mathbf{S}_x(f_1), ..., \mathbf{S}_x(f_K)$, $f_1, ..., f_K$ being some chosen frequencies. The matrix $\mathbf{A}$ is estimated as the minimizer of the following Direct Least-Squares (DLS) criterion:

$$\mathbf{A} = \underbrace{argmin}_{\mathbf{A}} \quad \sum_{k=1}^{K} \|\mathbf{S}_x(f_k) - \mathbf{A}\mathbf{D}(f_k, \theta)\ \mathbf{A}^H\|_F^2 \tag{7}$$

In that particular case, one can prove the following identifiability result (proof is omitted due to space limitation):

Theorem 1 *Under the previous data model, the mixing matrix is identifiable from the observed signal SoS if and only if vectors* $\mathbf{b}_i = [b_i(1), ..., b_i(p)]\,, i = 1, ..., d$ *verify* $\mathbf{b}_i \neq \mathbf{b}_j\ \forall i \neq j$.

Remarks:
1) In the standard BSS problem of statistically independent sources, matrices $\mathbf{D}(f, \theta)$ become diagonal, in which case the joint decomposition in (7) reduces to the well known joint diagonalization problem [6].
2) In the more general case where the parameter vector θ is unknown, the source separation can be achieved by solving an extended optimization problem w.r.t unknowns θ and $\mathbf{A}$, i.e. $\mathbf{A} = \underbrace{argmin}_{\mathbf{A},\theta} \quad \sum_{k=1}^{K} \|\mathbf{S}_x(f_k) - \mathbf{A}\mathbf{D}(f_k, \theta)\ \mathbf{A}^H\|_F^2$.

3.2. Illustration Example: AR(1) Dependent Sources

To illustrate the previous approach for separating dependent sources using SoS, a 2-dimensional Gaussian AR(1) sources model is considered, i.e.

$$\mathbf{s}(t) = \mathbf{B}\mathbf{s}(t-1) + \xi(t) \tag{8}$$

where $\mathbf{B} = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$ and $\mathbf{R}_\xi = \begin{bmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}$ is the covariance matrix of the innovation process $\xi(t)$.

The covariance matrix of the mixture signals $\mathbf{X}$ is function of the channel matrix $\mathbf{A}$ and the covariance matrix of the source signals $\mathbf{R}_s$ and is expressed as follows[1]:

$$\begin{aligned} \mathbf{R}_x(\tau) &= \mathbf{A}\mathbf{R}_s(\tau)\mathbf{A}^T, \text{ for } \tau \neq 0 \\ \mathbf{R}_s(\tau) &= \begin{bmatrix} \frac{b_1^\tau \sigma_1^2}{1-b_1^2} & \frac{b_1^\tau \rho}{1-b_1 b_2} \\ \frac{b_2^\tau \rho}{1-b_1 b_2} & \frac{b_2^\tau \sigma_2^2}{1-b_2^2} \end{bmatrix} \end{aligned} \tag{9}$$

Consider a set of K covariance matrices computed at different time lags $\mathcal{S}_x \overset{def}{\equiv} \{\mathbf{R}_x(\tau_k), k = 1, ..., K\}$. Our objective is to find the matrix $\mathbf{A}$ that associates all the couples of matrices $(\mathbf{R}_s(\tau_k), \mathbf{R}_x(\tau_k))\,, k = 1, ..., K$ according to the decomposition in (9).

The same idea has been used by Belouchrani et al. [2] to separate mutually independent and time coherent sources based on their SoS. The main difference with the algorithm SOBI presented in [2] is that the targeted matrices in our case have non-diagonal structure.

As presented in (7), the source separation problem can be solved using the joint matrix decomposition technique which is a generalization of the joint matrix digonalization tool (see [6] and references therein). Matrix $\mathbf{A}$ is estimated as the joint decomposer of the set $\mathcal{S}_x$ such that the following DLS cost

[1]For this simple example, we use the covariance matrices in the time domain, instead of the psd (i.e. their Fourier transform).

function:

$$\mathcal{J}(\mathbf{A}) = \sum_{k=1}^{K} \|\mathbf{R}_x(\tau_k) - \mathbf{A}\mathbf{R}_s(\tau_k)\mathbf{A}^T\|_F^2 \quad (10)$$

is minimized. This minimization problem can be solved by many approaches including the Alternating Least Squares (ALS) method in [6, 7]. The latter has been adapted to our context and used in our simulation experiments.

4. HOS-BASED SEPARATION OF DEPENDENT IID SOURCES

4.1. General Case

Now we consider i.i.d (non coherent), real-valued non Gaussian sources which are mutually dependent according to a known parametric model. For a set of matrices $\mathbf{M}_k, k = 1, 2, ..., K$, let define the contracted cumulant matrices $\mathbf{C}_k = \mathbf{C}_x(\mathbf{M}_k), k = 1, ..., K$ which ij-th entries are defined by [3]:

$$c_{k,ij} = \sum_{p,q=1}^{m} Cum(x_i, x_j, x_p, x_q) m_{k,pq} \quad (11)$$

$m_{k,pq}$ is the pq-th element of $\mathbf{M}_k$. Considering the data model in (2), it can be shown that the previous matrices share the following matrix structure:

$$\mathbf{C}_k = \mathbf{A}\ \mathbf{D}_k\ \mathbf{A}^T \quad (12)$$

where $\mathbf{D}_k = \mathbf{D}(\mathbf{A}, \mathbf{M}_k, \theta)$ are non diagonal matrices depending on $\mathbf{M}_k$, the mixing matrix $\mathbf{A}$, and the sources pdf parameters represented by θ assumed here to be known or a prior estimated.

Similarly, to the SOS-based approach, we propose to estimate the mixing matrix $\mathbf{A}$ using joint decomposition of the matrices $\mathbf{C}_1, \mathbf{C}_2, ...\mathbf{C}_K$, according to:

$$\mathbf{A} = \underbrace{argmin}_{\mathbf{A}}\ \mathcal{J}(\mathbf{A}) = \sum_{k=1}^{K} \|\mathbf{C}_k - \mathbf{A}\ \mathbf{D}_k\ \mathbf{A}^T\|_F^2 \quad (13)$$

This is a specific non linear criterion which deserves the derivation of dedicated optimization algorithms (as for the joint diagonalization problem [6]) which will be the focus of future works.

Next we present a simple 2×2 example to illustrate the concept of separating dependent sources based on their HoS.

4.2. Illustration Example: FGM dependency model

Consider two i.i.d. sources statistically dependent according to the Fairlie-Gumbel-Morgenstern (FGM)-copula with parameter θ [1]. The joint distribution of the two sources is given as follows [10]:

$$F(s_1, s_2) = F_1(s_1)F_2(s_2)\left\{1 + \theta\left(1 - F_1(s_1)\right)\left(1 - F_2(s_2)\right)\right\} \quad (14)$$

if in addition we consider that the sources are uniform $s_i \sim U_{[-\frac{1}{2},\frac{1}{2}]}, i = 1, 2$, the joint probability density function will be:

$$f(s_1, s_2) = 1 + 4\theta s_1 s_2 \quad (15)$$

After some straightforward developments, the entries of the symmetric matrix $\mathbf{D}(\mathbf{A}, \mathbf{M}, \theta)$ are given as follows:

$$\begin{aligned}
d_{11} &= k_1\mathbf{a}_1^T\mathbf{M}\mathbf{a}_1 + k_2\mathbf{a}_2^T\mathbf{M}\mathbf{a}_2 + k_3(\mathbf{a}_2^T\mathbf{M}\mathbf{a}_1 + \mathbf{a}_1^T\mathbf{M}\mathbf{a}_2)\\
d_{22} &= k_1\mathbf{a}_2^T\mathbf{M}\mathbf{a}_2 + k_2\mathbf{a}_1^T\mathbf{M}\mathbf{a}_1 + k_3(\mathbf{a}_2^T\mathbf{M}\mathbf{a}_1 + \mathbf{a}_1^T\mathbf{M}\mathbf{a}_2)\\
d_{21} &= k_2(\mathbf{a}_2^T\mathbf{M}\mathbf{a}_1 + \mathbf{a}_1^T\mathbf{M}\mathbf{a}_2) + k_3(\mathbf{a}_1^T\mathbf{M}\mathbf{a}_1 + \mathbf{a}_2^T\mathbf{M}\mathbf{a}_2)
\end{aligned}$$

with $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2]$ and

$$\begin{aligned}
k_1 &= Cum(s_1, s_1, s_1, s_1) &= -\tfrac{1}{120}\\
k_2 &= Cum(s_1, s_1, s_2, s_2) &= -\tfrac{\theta^2}{648}\\
k_3 &= Cum(s_1, s_1, s_1, s_2) &= -\tfrac{\theta}{360}
\end{aligned} \quad (16)$$

The minimization of the criterion in (13) is done, in our simulation experiment, via a gradient technique under the assumption that θ is known. Matrix $\mathbf{A}$ is updated iteratively according to:

$$\mathbf{A}^{(k)} = \mathbf{A}^{(k-1)} + d\mathbf{A}^{(k)} \quad (17)$$

The 'gradient' matrix $d\mathbf{A}^{(k)}$ is given by

$$d\mathbf{A}^{(k)} = -\lambda \sum_{k=1}^{K} \left(\mathbf{H}_k + \tilde{\mathbf{G}}_k\right) \quad (18)$$

where

$$\begin{aligned}
\tilde{\mathbf{G}}_k &= [\sum_{i,j=1,2} g_{k,ij}\mathbf{V}_{ji} \quad \sum_{i,j=1,2} g_{k,ij}\mathbf{U}_{ji}]\\
\mathbf{G}_k &= \mathbf{A}^T(\mathbf{C}_k - \mathbf{A}\mathbf{D}_k\mathbf{A}^T)\mathbf{A} = \begin{bmatrix} g_{k,11} & g_{k,12}\\ g_{k,21} & g_{k,22}\end{bmatrix}\\
\mathbf{H}_k &= 2(\mathbf{C}_k - \mathbf{A}\mathbf{D}_k\mathbf{A}^T)\mathbf{A}\mathbf{D}_k\\
\mathbf{V}_{11} &= 2\mathbf{M}(k_1\mathbf{a}_1 + k_3\mathbf{a}_2)\\
\mathbf{V}_{22} &= 2\mathbf{M}(k_2\mathbf{a}_1 + k_3\mathbf{a}_2)\\
\mathbf{V}_{21} &= \mathbf{V}_{12} = 2\mathbf{M}(k_3\mathbf{a}_1 + k_2\mathbf{a}_2)\\
\mathbf{U}_{11} &= 2\mathbf{M}(k_3\mathbf{a}_1 + k_2\mathbf{a}_2)\\
\mathbf{U}_{21} &= \mathbf{U}_{12} = 2\mathbf{M}(k_2\mathbf{a}_1 + k_3\mathbf{a}_2)\\
\mathbf{U}_{22} &= 2\mathbf{M}(k_3\mathbf{a}_1 + k_1\mathbf{a}_2)
\end{aligned}$$

and $\lambda > 0$ being a chosen step size parameter.

5. SIMULATION RESULTS

In this section, the new concept of separating mixtures of dependent sources will be validated though the two simple 2×2 scenarios presented below.

In the first scenario, a 2-dimensional auto-regressive signal is generated according to the model in (8) with $\sigma_1 = 5$,

$\sigma_2 = 3$, $\rho = 4$, $b_1 = 0.95$ and $b_2 = 0.85$. The mixing matrix $\mathbf{A}$ is set to $\begin{bmatrix} 1 & 0.85 \\ 0.65 & 1 \end{bmatrix}$.

Figures 1 and 2 show the performance of the ALS algorithm applied for the joint decomposition of a set of SoS based matrices. The separation quality is evaluated through the Mean Rejection Level (MRL) [2] (figure 1) and the SINR criterion defined in [1] (figure 2). One can see that the ALS algorithm converges to the desired solution through the joint decomposition of a set of $K = 5$ covariance matrices.

For the second scenario, the data model is generated according to (12) where the mixing matrix $\mathbf{A}$ and the symmetric matrices $\mathbf{M}_k$, $k = 1, ..., 10$ are generated randomly[2] . The dependency parameter θ in (15) is set to 0.5.

The performance of the gradient based algorithm used for the joint decomposition of a set of matrices based on the HoS of the mixture signals are presented in figure 3 (the decrease of the MRL through the iterations) and figure 4 (the decrease of the cost function through the iterations). One can see through the MRL plot in figure 3 and the DLS criterion evolution in figures 4 that the proposed algorithm converges to the desired solution.

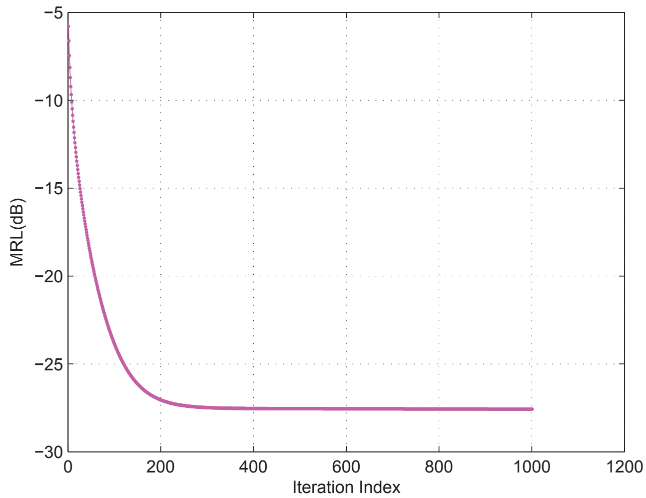

Fig. 1. MRL vs. Iterations: SoS

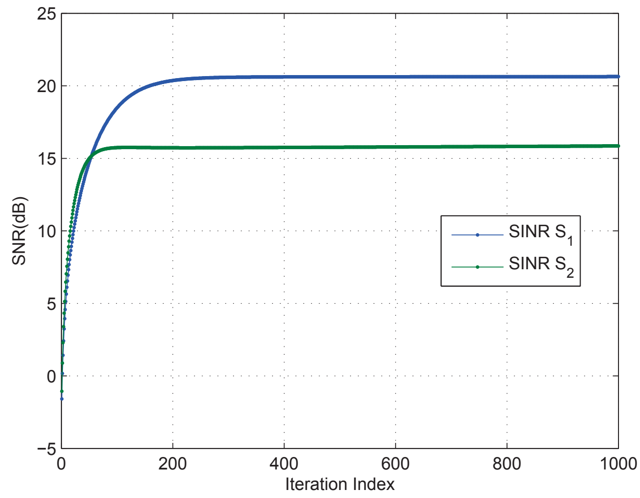

Fig. 2. SNR vs. Iterations: SoS

6. CONCLUSION

In this paper the separation problem of dependent sources has been addressed. It has been shown that the separation is made possible thanks to side information on the dependence nature of the considered sources. For this new concept of separating dependent sources, the joint matrix decomposition technique has been used instead of the joint matrix diagonalization tool where the target matrices have diagonal structure. The effectiveness of the new concept has been illustrated with two simple 2×2 source separation examples based on second-order-statistics and high-order-statistics, respectively. For these two examples, standard algorithms have been used to validate the concept. Better results can be got through more dedicated methods for the considered data model and the nature of the dependency between source signals. In addition, an identifiability result has been established for the case of auto-regressive sources with correlated innovation processes. General identifiability results, especially for HoS based methods are to the best of our knowledge open problem to be considered in future works.

[2]We used in this simulation example the exact stastistics. This is to better illustrate the proposed concept while avoiding the problems of ill convergence of the used gradient algorithm due to noise and finite sample size effects.

7. REFERENCES

[1] A. Keziou, H. Fenniri, K. Messou, E. Moreau, "Blind Source Separation of Independent / Dependent Signals Using a Measure on Copulas", *in Proc. EUSIPCO*, Sep. 2013.

[2] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics", *IEEE Tr. on SP.*, Feb 1997.

[3] J-F. Cardoso, A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals", *IEE Proceedings-F* 1993.

[4] P. Comon, "Independent Component Analysis - A New Concept", *Signal Processing*, April 1994.

[5] S. Gorlow and S. Marchand, "Informed Audio Source Separation Using Linearly Constrained Spatial Filters", *IEEE Trans. on Audio, Speech & Language Processing*, 2013.

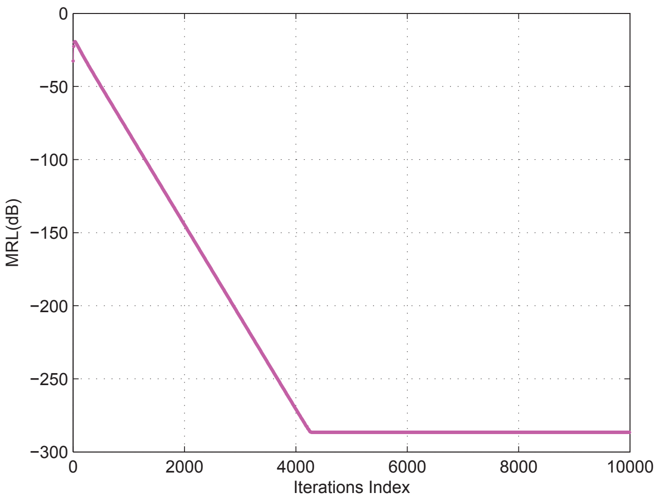

Fig. 3. MRL vs. Iterations: HoS

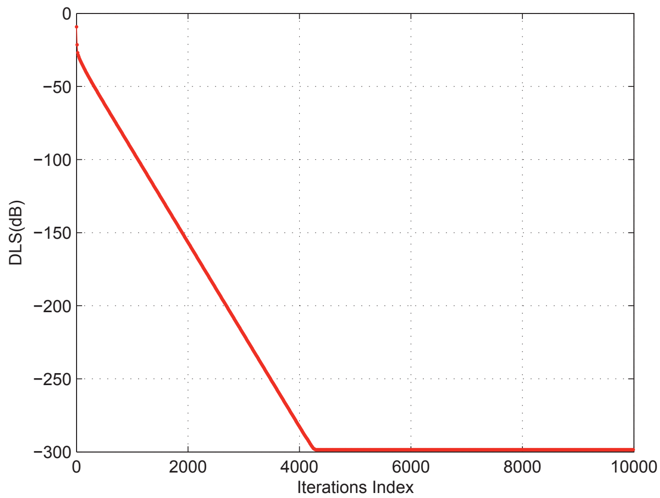

Fig. 4. DLS vs. Iterations: HoS

[6] A. Mesloub, K. Abed-Meraim, A. Belouchrani, "A New Algorithm for Complex Non-Orthogonal Joint Diagonalization Based on Shear and Givens Rotations", *arXiv/1306.0331, to appear in IEEE Tr. on Signal Processing*, 2014.

[7] T. Trainini and E. Moreau, "A Least Squares Algorithm for Global Joint Decomposition of Complex Matrix Sets", *in Proc. CAMSAP*, Dec. 2011.

[8] P. Tichavsk, E. Doron, A. Yeredor, and J. Nielsen, "A Computationally Affordable Implementation of an Asymptotically Optimal BSS Algorithm for AR Sources", *in Proc. EUSIPCO'2006*, Florence, Italy, Sep. 2006.

[9] P. Comon, C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press Inc, 2009.

[10] Pravin K. Trivedi and David M. Zimmer, "Copula Modeling: An Introduction for Practitioners", *Foundations and Trends in Econometrics*, pp. 1111, Vol. 1, No 1, 2005.

[11] S.M. Fouladi and H. Amindavar, "Blind separation of dependent sources using multiwavelet based on M-GARCH model", *International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, 2012.

[12] K. Abed-Meraim, Y. Xiang, J. Manton, and Y. Hua, "Blind source separation using second order cyclostationary statistics", *IEEE Tr. on Signal Processing*, vol. 49, n 4, pp. 694-701, Apr. 2001.

[13] Marc Castella, Selwa Rafi, Pierre Comon and Wojciech Pieczynski, "Separation of instantaneous mixtures of a particular set of dependent sources using classical ICA methods", *EURASIP Journal on Advances in Signal Processing*, 2013.

[14] P. Aguilera, S. Cruces, I. Duran-Diaz, A. Sarmiento, "Blind Separation of Dependent Sources With a Bounded Component Analysis Deflationary Algorithm", *IEEE Signal Processing Letters*, Vol:20 , Issue: 7, July 2013.

[15] Alper T. Erdogan, "A Class of Bounded Component Analysis Algorithms for the Separation of Both Independent and Dependent Sources", *IEEE Transactions on Signal Processing*, Vol:61 , Issue: 22, Nov 2013.