AN ADAPTIVE BLIND SIGNAL SEPARATION BASED ON THE JOINT OPTIMIZATION OF GIVENS ROTATIONS

N.Bienati, U.Spagnolini, M. Zecca

Dip. Elettronica e Informazione, Politecnico di Milano
P.zza L. da Vinci, 32. I-20133 Milano (Italy)
e-mail: bienati@fusberta.elet.polimi.it, spagnoli@elet.polimi.it

ABSTRACT

Blind signal separation (BSS) is a recurrent problem in many multi-sensors applications where observations can be modelled as mixtures of N statistical independent source signals. In this paper we propose the estimation of the orthonormal transformation matrix $\mathbf{Q}$ in the case of whitened observations and a cost function based on the fourth-order moments. $\mathbf{Q}$ is described as combination of elementary Givens rotations and the optimization is carried out jointly for all the rotations. When sub-sets of angles are optimized separately the method reduces to the deflation approach which has been proved to be globally convergent [1]. The joint estimation of Givens rotation matrices has a computational complexity $O(7N^2)$ and, compared to other adaptive BSS, simulations demonstrate that it converges faster and achieves a better attenuation of cross-talks.

1. INTRODUCTION

Applications call for a blind signal separation (BSS) method for memoryless interference systems that is computationally affordable, fast convergent, stable and reasonably accurate. All these properties are equally important and, eventually, can be moderately relaxed. BSS methods rely mainly on the hypothesis of statistical independence of the unknown sources [2]. BSS can be decomposed into two-steps: the observations are first prewhitened by a whitening matrix, and then an orthonormal matrix $\mathbf{Q}$ can be separately calculated by constraining the source separation with the minimization of appropriate cost-functions (orthogonal contrast functions). In this paper we limit ourselves to the case where the prewhitening is known or estimated separately (e.g., by using a LMS or RLS approach) and the rotation matrix $\mathbf{Q}$ has to be estimated. We will assume that the orthogonal contrast-function is based on the fourth-order moments and we concentrate on the way to estimate a matrix $\mathbf{Q}$ by constraining to be an orthonormal transformation.

In iterative methods $\mathbf{Q}$ can be estimated by constraining (at the first order) the updating to result in a matrix $\mathbf{Q}$ that is orthonormal [5]. The constraint can be explicitly added as Lagrange multipliers [6] but a global optimization is not practical as computationally too expensive [3]. In BSS matrix $\mathbf{Q}$ can be conveniently written as a combinations of elementary Givens rotations [4]. In this paper we propose the joint optimization of the Givens rotations (JOG) by exploiting a computationally efficient updating which results in a cost $O(7N^2)$. The approach is similar to the deflation approach proposed in [1] which has been proved to be globally convergent, the main difference here lies in the way the updating is carried out. Compared to the deflation approach, the BSS-JOG shows a faster convergence and lower mutual interference from residual mixing. Here this latter property is conveniently evaluated in term of signal to interference ratio (SIR). The BSS-JOG is compared to EASI [5] as it is a BSS benchmark based on the same contrast-function.

The paper is organized as follows. In the next section we recall the model definitions and the optimization problem under study. Then in Section 3 we describe the BSS-JOG algorithm; the deflation approach [1] is shortly recalled. In addition, it is shown that the sources separation is a stationary point of JOG. The performance analysis with respect to adaptive algorithms are in Section 4. In addition, motivated by the need to reduce the cost for the computations of trigonometric functions (e.g., by addressing a look-up table), the performances are evaluated when the angles-updating is quantized or fixed.

2. PRELIMINARIES

As a general model for BSS let the L observed signals be related to the N independent source signals $\{x_i(k)\}_{i=1}^N$ by the $L \times N$ memoryless channel matrix $\mathbf{A}$:

$$\mathbf{w}(k) = \mathbf{A}\mathbf{x}(k) + \mathbf{n}(k), \quad k \in \mathbb{Z} \tag{1}$$

where $\mathbf{x}(k) = [x_1(k), ..., x_N(k)]^T$ and $\mathbf{n}(t)$ is the additive Gaussian noise. $\mathbf{A}$ is full-column rank, sources are zero mean non-Gaussian with $E[|x_i(k)|^2] = 1$ and $\mathbf{n}(t) \sim \mathcal{N}(0, \sigma^2\mathbf{I}_N)$. Without any loss of generality here we assume in the derivation that signals are real-valued, $L = N$ and $\sigma^2 = 0$ (no-noise). The BSS can operate into two steps [5]. The first step prewhitens the observations according to a whitening matrix $\mathbf{B}$, it results in a set of uncorrelated and normalized signals $\mathbf{z}(k) = \mathbf{B}\mathbf{w}(k)$: $E[\mathbf{z}(k)\mathbf{z}(k)^T] = \mathbf{I}_N$. Since $E[\mathbf{w}(k)\mathbf{w}(k)^T] = \mathbf{A}\mathbf{A}^T$ the matrix $\mathbf{B} = (\mathbf{A}\mathbf{A}^T)^{-1/2}$ or equivalently $\mathbf{A} = \mathbf{B}^{-1}\mathbf{H}$ where $\mathbf{H}$ is an orthonormal transformation. Therefore, the BSS reduces in finding the unknown $\mathbf{Q}$ from the whitened observations $\mathbf{z}(k)$

$$\mathbf{t} = \mathbf{t}(\mathbf{Q}) = \mathbf{Q}\mathbf{z}, \tag{2}$$

An appropriate cost function based on high-order statistics can separate the sources by forcing their independence, thus $\mathbf{Q} \to \mathbf{R}\mathbf{G}$ or equivalently $\mathbf{t} \to \mathbf{R}\mathbf{x}$ apart from a permutation matrix $\mathbf{R}$ and a sign reversal for the sources.

Pre-whitening can be carried out in any of the known methods and it is not dealt with here. We use the RLS approach as it does not delay the BSS because of the ill-conditioning of $E[\mathbf{w}(k)\mathbf{w}(k)^T]$, in addition the cost $O(N^2)$ is comparable with the estimation of $\mathbf{Q}$ proposed below (Section 3).

3. ADAPTIVE ESTIMATION OF THE MATRIX Q

Let us suppose that all the sources have negative kurtosis (sub-Gaussian). The source signals $\mathbf{x}$ can be recovered from $\mathbf{z}$ by estimating the N^2 elements $q_{i,j}(\boldsymbol{\theta}) = [\mathbf{Q}(\boldsymbol{\theta})]_{ij}$ through the minimization of the following *contrast function* [5]:

$$\Psi(\boldsymbol{\theta}) = E\left[\sum_{i=1}^{N} t_i^4(\mathbf{Q}(\boldsymbol{\theta}))\right] = E\left[\sum_{i=1}^{N}\left(\sum_{j=1}^{N} q_{i,j}(\boldsymbol{\theta})\, z_j\right)^4\right], \tag{3}$$

where $\boldsymbol{\theta}$ denotes the vector of $M = N(N-1)/2$ rotation angles. The minimization of $\Psi(\boldsymbol{\theta})$ will be attained by applying the stochastic gradient algorithm, the actual function to be iteratively minimized is

$$\psi(\boldsymbol{\theta}) = \sum_{i=1}^{N} t_i^4(\mathbf{Q}(\boldsymbol{\theta})). \tag{4}$$

The iterative minimization of (4) does not guarantee that $\mathbf{Q}$ converges to a unitary matrix unless constraints are added (see e.g., [6], [7]). Here the unitary matrix $\mathbf{Q}$ is decomposed into the product of M Givens rotations $\mathbf{P}_k$:

$$\mathbf{Q}(\boldsymbol{\theta}) = \prod_{m=1}^{N-1}\prod_{n=1}^{m} \mathbf{P}_{k(m,n)}\left(\theta_{k(m,n)}\right) \tag{5}$$

where $\boldsymbol{\theta} = [\theta_1 ... \theta_M]^T$, $k(m,n) = n + \sum_{i=1}^{m-1} i$ and

$$\mathbf{P}_k(\theta_k) = \begin{bmatrix} \mathbf{I}_{n-1} & & & & \\ & \cos\theta_k & & \sin\theta_k & \\ & & \mathbf{I}_{m-n-1} & & \\ & -\sin\theta_k & & \cos\theta_k & \\ & & & & \mathbf{I}_{M-m} \end{bmatrix}$$

Provided that $\mathbf{Q}$ is decomposed into M Givens rotations that rotate each coordinate plane in the $\mathbb{R}^M$ space, the rotations ordering is not relevant due to the signal permutation ambiguity of BSS. The ordering preserves the similarities with the deflation approach [1]. The rotations (5) are arranged as

$$\mathbf{Q}(\boldsymbol{\theta}) = \prod_{m=1}^{N-1} \mathbf{G}_m \tag{6}$$

where $\mathbf{G}_m = \prod_{n=1}^{m} \mathbf{P}_{k(m,n)}\left(\theta_{k(m,n)}\right)$. It is easy to prove that each matrix $\mathbf{G}_m$ has the following structure

$$\mathbf{G}_m = \begin{bmatrix} \tilde{\mathbf{G}}_m & \mathbf{0}_{m-1,M-m} \\ \mathbf{g}_m^T & \mathbf{0}_{1,M-m} \\ \mathbf{0}_{M-m,m} & \mathbf{I}_{M-m} \end{bmatrix}, \tag{7}$$

where $\mathbf{0}_{p,q}$ is a $p \times q$ matrix whose elements are all zeros, $\tilde{\mathbf{G}}_m$ is a $(m-1) \times m$ matrix, and $\mathbf{g}_m^T$ is a row vector such that $\tilde{\mathbf{G}}_m \mathbf{g}_m = \mathbf{0}_{m-1,1}$. It follows then

$$t_m = \mathbf{g}_m^T \left(\prod_{p=m+1}^{N-1} \mathbf{G}_p\right) \mathbf{z} = \mathbf{g}_m^T \mathbf{y}_m \tag{8}$$

This structure is reminiscent of the one at the basis of the deflation approach proposed in [1] (each matrix $\mathbf{G}_m$ corresponds to one step of the deflative procedure), the main difference here lies in the way the angles $\boldsymbol{\theta}$ are updated as shown below.

The gradient of ψ with respect to the parameters θ_k ($k = 1...M$) is

$$\frac{\partial\psi}{\partial\theta_k} = 4\sum_i t_i^3 \sum_l \frac{\partial q_{i,l}(\theta_k)}{\partial\theta_k} z_l = \mathbf{r}^T \frac{\partial \mathbf{Q}}{\partial\theta_k}\mathbf{z} \tag{9}$$

where $\mathbf{r} = \left[t_1^3 ... t_N^3\right]^T$, $\mathbf{z} = [z_1 ... z_N]^T$, and

$$\frac{\partial\mathbf{Q}}{\partial\theta_k} = \prod_{l=1}^{l=k-1} \mathbf{P}_l(\theta_l) \frac{\partial \mathbf{P}(\theta_k)}{\partial\theta_k} \prod_{l=k+1}^{l=M} \mathbf{P}_l(\theta_l) = \mathbf{L}_k \frac{\partial \mathbf{P}(\theta_k)}{\partial\theta_k} \mathbf{R}_k$$

$$\frac{\partial \mathbf{P}(\theta_k)}{\partial\theta_k} = \begin{bmatrix} \mathbf{0}_{n-1} & & & & \\ & -\sin\theta_k & & \cos\theta_k & \\ & & \mathbf{0}_{m-n-1} & & \\ & -\cos\theta_k & & -\sin\theta_k & \\ & & & & \mathbf{0}_{M-m} \end{bmatrix}$$

here $\mathbf{0}_n = \mathbf{0}_{n,n}$ denotes a $n \times n$ matrix of zeros. The partition of $\frac{\partial\mathbf{Q}}{\partial\theta_k}$ into the product of three matrices ($\mathbf{L}_k$, $\frac{\partial\mathbf{P}(\theta_k)}{\partial\theta_k}$, $\mathbf{R}_k$) is the key for the efficient implementation of the algorithm whose complexity is $O(7N^2)$ for each adaptation step (here we neglect the cost for the evaluations of the trigonometric functions needed to build $\mathbf{Q}(\boldsymbol{\theta}_{n+1})$ as they can be pre-calculated and simply addressed). Indeed, eq. (9) can be rewritten as

$$\frac{\partial\psi}{\partial\theta_k} = \mathbf{r}^T \frac{\partial\mathbf{Q}}{\partial\theta_k}\mathbf{z} = \mathbf{r}^T \mathbf{L}_k \frac{\partial\mathbf{P}(\theta_k)}{\partial\theta_k}\mathbf{R}_k \mathbf{z} = \mathbf{u}_k^T \frac{\partial\mathbf{P}(\theta_k)}{\partial\theta_k}\mathbf{v}_k \tag{10}$$

and the following recursive formulas hold:

$$\begin{cases} \mathbf{u}_1 = \mathbf{r} \\ \mathbf{u}_k^T = \mathbf{u}_{k-1}^T \mathbf{P}_{k-1} \quad 1 < k < M \end{cases} \tag{11a}$$

$$\begin{cases} \mathbf{v}_M = \mathbf{z} \\ \mathbf{v}_{k-1} = \mathbf{P}_k \mathbf{v}_k \quad 1 < k < M \end{cases} \tag{11b}$$

For $k = M, ..., 1$ we first build the vectors $\mathbf{v}_k$; this step requires $M-1$ matrix by vector products, where the matrices are Givens rotation $\mathbf{P}_k$. The second step consists of computing the vectors $\mathbf{s}_k = \frac{\partial\mathbf{P}(\theta_k)}{\partial\theta_k}\mathbf{v}_k$ ($k = 1, ..., M$), notice that every vector $\mathbf{s}_k$ has only 2 non-zero elements. In the third step vectors $\mathbf{u}_k$ are computed and a total of $M-2$ vector by $\mathbf{P}_k$ matrix products are required. A final step is needed in order to compute $\mathbf{u}_k^T \mathbf{s}_k$ ($k = 1, ..., M$). The total number of multiplications needed at each step is listed the table below.

Step	Number of multiplications
1. $\mathbf{v}_k$	$4(M-1)$
2. $\mathbf{s}_k = \frac{\partial\mathbf{P}(\theta_k)}{\partial\theta_k}\mathbf{v}_k$	$4M$
3. $\mathbf{u}_k$	$4(M-2)$
4. $\psi'(\theta_k) = \mathbf{u}_k^T \mathbf{s}_k$	$2M$

The stochastic update algorithm for $\boldsymbol{\theta}$ reduces to

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \delta\dot{\psi}(\boldsymbol{\theta}_n) \tag{12}$$

where $\dot{\psi}(\boldsymbol{\theta}_n) = \left[\mathbf{r}_n^T \frac{\partial\mathbf{Q}(\boldsymbol{\theta}_n)}{\partial\theta_1}\mathbf{z}_n, ..., \mathbf{r}_n^T \frac{\partial\mathbf{Q}(\boldsymbol{\theta}_n)}{\partial\theta_M}\mathbf{z}_n\right]^T$ is derived from (10) for $\boldsymbol{\theta} = \boldsymbol{\theta}_n$ and δ is the step-size. Notice that here the update of the angle $\theta_{k(m,n)}$ depends on all the outputs $\mathbf{t}$, whilst in the deflation approach it depends only on t_m.

Remark 1: Within the interval $[0, 2\pi)$ a stationary point for the updating (12) is any vector $\bar{\boldsymbol{\theta}}$ such that (see e.g., [1]):

$$E\left[\dot{\psi}(\bar{\boldsymbol{\theta}})\right] = 0. \tag{13}$$

Since

$$\frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta_k} = E\left[\sum_{i=1}^{N} 4t_i^3 \frac{\partial t_{i,}(\boldsymbol{\theta})}{\partial\theta_k}\right] = E\left[\mathbf{r}^T \frac{\partial\mathbf{Q}}{\partial\theta_k}\mathbf{z}\right] = E\left[\frac{\partial\psi(\boldsymbol{\theta})}{\partial\theta_k}\right]$$

we conclude that the unique absolute minimum of $\Psi(\boldsymbol{\theta})$ does correspond to a stable stationary point for the updating (12). This proof does not ensure the global convergence of the iterative updating since $\Psi(\boldsymbol{\theta})$ might have local minima (as it can be analytically proved for $N = 2$). It should be recalled that when the optimization of the contrast function (3) can be decoupled into the optimization of each component as for [1] and [6], the steps in [8] can be followed to show that for each of the components the stable stationary points allows the recovery of one source. Since in this case the optimization of the contrast function (12) is carried out globally and not sequentially the proof of global convergence cannot be easily extended in this case. Extensive numerical simulations for $N > 2$ have never given evidences that the BSS-JOG can be trapped in a local minimum implying any degradation of performance with respect to other globally convergent methods (e.g., the deflation approach as shown in Section 4). Moreover, for certain pdf of source signals (for instance, for binary signals) the attractor to the global minimum seems to be so strong that, in absence of additive noise, algorithm (12) permits a complete separation of the sources within less than one hundred adaptation steps.

Remark 2: Each of the 2×2 elements of the rotation matrix $\mathbf{P}_k(\theta_k)$ can be re-written as $\mathbf{P}_k(\rho_k) = \frac{1}{(1+\rho_k^2)^{1/2}}\begin{bmatrix} 1 & \rho_k \\ -\rho_k & 1 \end{bmatrix}$, optimization is now carried out with respect to $\boldsymbol{\rho} = [\rho_1, ..., \rho_M]^T$. This results in a slightly different updating equation (i.e., the only term changed is $\partial\mathbf{P}(\rho_k)/\partial\rho_k$) that could be suitably exploited when the complex valued signals are considered.

4. PERFORMANCE

The performance of the BSS-JOG are evaluated through numerical simulations. Since converge rate and residual interference after the demixing can be considered as design parameters in applications, performances are evaluated by comparing experimentally these two parameters almost independently of the step-size δ. Prewhitening is based on RLS algorithm as the convergence rate is almost independent of the condition number ($CN_{\mathbf{A}}$) of the covariance matrix $E[\mathbf{w}(k)\mathbf{w}(k)^T] = \mathbf{A}\mathbf{A}^T$, performance is mainly due to the estimation of the orthogonal transformation $\mathbf{Q}(k)$. Let $\mathbf{C}(k) = \mathbf{Q}(k)\mathbf{B}(k)\mathbf{A}$ be the global mixing/demixing matrix evaluated during the k-th sample (or iteration as adaptive BSS is performed by one iteration per time sample), the signal to interference ratio (SIR) for the estimated source is:

$$SIR_\ell(k) = 10\log_{10}\left(\frac{\max_n\{|[\mathbf{C}(k)]_{\ell n}|^2\}}{\sum_{n=1}^{N}|[\mathbf{C}(k)]_{\ell n}|^2 - \max_n\{|[\mathbf{C}(k)]_{\ell n}|^2\}}\right). \tag{14}$$

The SIR is a random variable that reflects the performance of a specific BSS algorithm, the number of samples required for convergence T_ℓ (or the convergence speed $1/T_\ell$) and the asymptotic accuracy $SIR_\ell = \lim_{k\to\infty} E[SIR_\ell(k)]$ can be evaluated experimentally for varying step-size δ and $CN_{\mathbf{A}}$. Here T_ℓ and SIR_ℓ are estimated by fitting on the $SIR_\ell(k)$ a piece-line function, in practice:

$$SIR_\ell(k) \approx \begin{cases} k \times SIR_\ell/T_\ell, \text{ for } k < T_\ell \\ SIR_\ell, \text{ for } k \geq T_\ell \end{cases} \tag{15}$$

where $\approx$ stands for the least square fitting. Recall that $SIR_\ell = SIR_\ell(\delta)$ depends on the step δ and $T_\ell = T_\ell(\delta, CN_{\mathbf{A}})$ depends also on the $CN_{\mathbf{A}}$ mostly for the prewhitening stage.

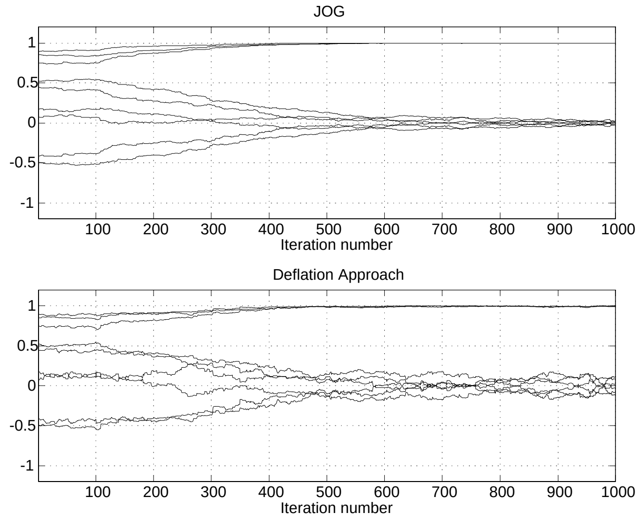

Fig. 1. Evolution of the elements of the 3×3 global matrix $\mathbf{C}$ vs. iterations for BSS-JOG (upper figure) and deflation approach (lower figure). The sources are $x_i(k) \sim \mathcal{U}(-\sqrt{3}, \sqrt{3})$ for $i = 1, 2, 3$, the mixing matrix is orthonormal (BSS after exact prewhitening) and the adaptation step is $\delta = 0.075$.

Performance comparison with Deflation Approach [1]: Since BSS-JOG is conceptually similar to the deflation approach in the updating of the matrix $\mathbf{Q}$ except on the way the block-rotations are updated, the two algorithm are compared with the same step δ. Figure 1 shows the evolution of the elements of the global matrix $\mathbf{C}(k)$ for $N = 3$ uniformly distributed sources and $CN_{\mathbf{A}} = 1$ (i.e., the mixing matrix is unitary). In this case the BSS-JOG outperforms the deflation approach both in term of convergence speed and asymptotic performance. This conclusion is confirmed from Figure 2 where the average values of convergence $T(\delta, CN_{\mathbf{A}})$ is represented with respect to $SIR(\delta)$ for 50 trials of $N = 4$ source signals, each is a four equiprobable levels of pulse amplitude modulation (4PAM): $x_i(k) \in \{\pm 1/\sqrt{5}, \pm 3/\sqrt{5}\}$. Since the use of the same step δ for the two algorithm could be misleading, this plot has the advantage of making the comparison of the performance almost independent on the step δ. For any value of SIR the convergence speed of the deflation approach is lower than the BSS-JOG, for any value of T the SIR of BSS-JOG outperforms by more than 10dB in SIR. This is not surprising, the deflation approach is intrinsically sequential and a loss of performance in term of convergence speed is expected with respect to JOG. However, the advantage of the deflation scheme lies in the proof of conver-

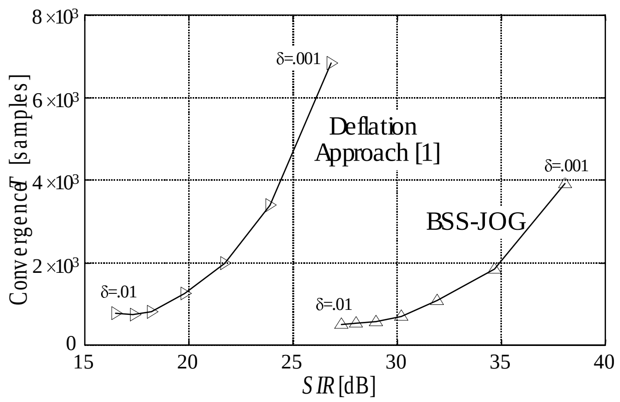

Fig. 2. Converge rate T vs. SIR for BSS-JOG and Deflation Approach [1] for $N = 4$ 4PAM sources.

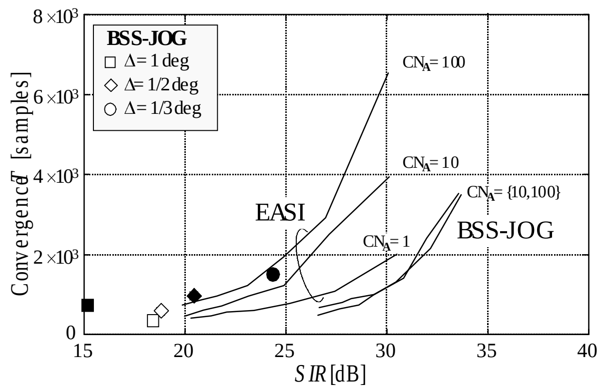

Fig. 3. Converge rate T vs. SIR for BSS-JOG and EASI for four sources of 4PAM signals. BSS-JOG with angle-quantization: quantization A for $\delta = 10^{-2}$ (empty symbols) and quantization B (filled symbols).

gence of the iterative updating which still cannot be proved for BSS-JOG.

Performance comparison with EASI [5]: The adaptive EASI is shown here as it represents a good reference in term of convergence-SIR-cost trade-offs even if it is understood that the RLS prewhitening of BSS-JOG tends to favor JOG algorithm. The comparison is for $N = 4$ source signals with 4PAM as in the example above, the condition numbers are $CN_{\mathbf{A}} = \{1, 10, 100\}$ and the range of steps is $10^{-3} \leq \delta \leq 10^{-2}$. Here $CN_{\mathbf{A}} = 1$ for the mixing matrix is intentionally chosen so as to reduce the EASI to a rotation-only algorithm and thus maximize the convergence rate. Figure 3 shows the comparison of converge rate T vs. asymptotic SIR. The fast prewhitening makes the BSS-JOG be independent of $CN_{\mathbf{A}}$ even if a loss of approx. 4dB arises when compared with Figure 2. For any value of the SIR the convergence of the BSS-JOG is faster than EASI regardless of $CN_{\mathbf{A}}$. A similar conclusion can be driven when considering the SIR for a given convergence rate, the advantage can be quantified (for $CN_{\mathbf{A}} = 1$) as approx. 2-3dB in SIR. Recall that the complexity of EASI is $O(N^2)$.

In real-time implementations trigonometric functions are precomputed and simply accessed (look-up table). The impact in Givens rotations of the trigonometric function is shown in Figure 3 when the angles are quantized with angle-interval Δ. The angles updating (12) is modified as follows:

$$\begin{array}{l}\text{quantization A: } \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \Delta \times round[\,\delta\dot{\psi}(\boldsymbol{\theta}_n)/\Delta] \\ \text{quantization B: } \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \Delta \times sign[\,\dot{\psi}(\boldsymbol{\theta}_n)]\end{array} \tag{16}$$

The quantization A represents the quantized angle-updating still based on the step δ ($round[.]$ denotes the round to the nearest integer) while in the quantization B the angle-updating is bounded to be $\pm\Delta$ and $O(N^2/2)$ multiplications by the step δ are no more necessary. Figure 3 confirms that the quantization B is affected by low convergence for low Δ, and small SIR when Δ is large. In any case, for $\Delta = 1$ deg the convergence is reached into approx. 1000 samples. A suitable strategy to optimize the switching from large Δ (at the first iterations) to small Δ (at final iterations) is not discussed here but can be derived with the help of the performance in Figure 3.

5. CONCLUSIONS

In this paper we have demonstrated that in BSS the decomposition of the rotation matrix into elementary Givens can be conveniently exploited. When the rotation angles are optimized jointly as for the proposed BSS-JOG the convergence rate and the residual interference outperforms with respect to the iterative optimization of the angles as in the deflation approach [1]. The BSS-JOG can be efficiently implemented with a cost $O(7N^2)$. The sensitivity analysis with respect to angle quantization shows that appropriate optimization strategies can be tuned for real-time applications, e.g., in mobile communication systems.

MATLAB codes for BSS-JOG are available at *http://www-dsp.elet.polimi.it* in *TLC* section.

6. REFERENCES

[1] N. Delfosse, P. Loubaton, “Adaptive blind separation of independent sources: A deflation approach”, *Signal Process.*, Vol.45, pp.59-83, July 1995.

[2] P. Comon, “Independent component analysis: a new concept?”, *Signal Process.*, Vol.36, pp.287–314, Apr. 1994.

[3] N. Bienati, “Blind equalization and source signal separation”, Politecnico di Milano, Internal Report No.98.95, 1998.

[4] E. Moreau, O. Macchi, ”A one stage self-adaptive algorithm for source separation”, *Proc. ICASSP*, vol.3, pp.49-52, 1994.

[5] J.F. Cardoso, B.H. Laheld, “Equivariant Adaptive Source Separation”, *IEEE Trans. Signal Processing*, Vol.44, Dec. 1996.

[6] C.B. Papadias, “A Multi-User Kurtosis Algorithm for Blind Source Separation”, *Proc. ICASSP*, vol.5, pp.3144-3147, 2000

[7] S.C. Douglas, “Prewhitened blind source separation with orthogonality constraints”, *Proc. of the 33rd Asilomar Conference*, Vol.1, pp.591-595, 1999.

[8] O. Shalvi, E. Weinstein, “New criteria for blind deconvolution of nonminimum phase systems (channels)”, *IEEE Trans. Information Theory*, Vol.36, pp.312-321, March 1990.