4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), April 2003, Nara, Japan

INITIALIZED JACOBI OPTIMIZATION IN INDEPENDENT COMPONENT ANALYSIS

† *Juan J. Murillo-Fuentes and* † *Rafael Boloix-Tortosa* ‡ *Francisco J. González-Serrano,*

† ATSC. ESI. Universidad de Sevilla. Paseo de los Descubrimientos sn, 41092 Sevilla, Spain.
‡ DTC-EPS, Universidad Carlos III de Madrid. Butarque 15, 28911 Leganés, Madrid, Spain.
E-mail: murillo@esi.us.es

ABSTRACT

The authors propose a new solution to the minimization of marginal entropies (ME) in multidimensional independent component analysis (ICA). Starting from the 'Jacobi optimization' (JO), we focus on a novel method based on initialization. In this method, we first compute the moment matrix for the prewhitened inputs. Then, the moments at each iteration of this Initialized Jacobi Optimization (IJO) are computed as rotations of this matrix. We include a computational comparison between the JO and IJO to design the Optimized Jacobi Optimization (OJO). This new method is available for a wide set of fourth order based contrasts. Experiments have been included to show that this algorithm have an excellent performance at a low computational cost and memory requirements in comparison to other well-known algorithms.

1. INTRODUCTION

Blind separation of sources (BSS) involve the task of obtaining a non-observable set of signals, the so-called sources, from another set of observable signals regarded as mixtures. Here, the adjective "blind" stands for the fact that neither the original sources nor the mixture itself are known. Usually, and in the context of this paper, the assumption of spatial statistical independence is the key to achieve separation. In this sense, the related and more general problem of independent component analysis (ICA) consist of obtaining [1], from a set of component (mixtures in BSS), another set as statistically independent as possible. Contrast functions are cost functions whose minimization yields the solution to the BSS/ICA. These contrast functions may use higher order moments to compute a unitary transformation to diagonalize the associated cumulant tensor of the whitened outputs. They may cancel a set of cumulants out of the diagonal (mutual information MI, JADE [2]), or maximize diagonal entries (minimization of marginal entropies, ME [1]). The diagonalization of this tensor matrix is usually carried out by using the Jacobi Optimization [1], i.e, they operate pairwise minimizing the associated 2-dimensional contrast for every whitened-signal pair in turn over several sweeps until convergence. The ICA and BSS have been applied (see [3] and references therein) to communications [4], [5], biomedical signals such as ECG or EEG [6], monitoring, image [7] or financial data processing, encoding or compression,... These methods share the need for accurate solutions at a low complexity. In this sense, we focus this paper on the development of an algorithm with a low computational burden and similar performance than previous ones.

The paper is organized as follows. We end this section with the matrix model, main assumptions and definitions. Then, a family of fourth order based contrasts, GWE, is introduced in Section 2. The minimization of these sinusoidal function leads to the same solution than the ME and MI approaches [8] at a lower complexity and similar performance than those presented in [1], [9]. In Section 3 we address the n-dimensional case and the 'Jacobi-Optimization' in [1]. We propose to introduce an initialization stage and prove this new approach to minimize the computational burden for low dimensional ($n \leq 15$) problems. Section 4 includes experiments. The last section is devoted to conclusions.

1.1. Matrix model in the BSS/ICA

In its simplest form, the BSS/ICA model reduces to the following matrix model. The entries of the $m \times 1$ mixture vector $\boldsymbol{x}_t$ at time t are instantaneous linear combinations of n statistically independent sources (components) $\boldsymbol{s}_t$, i.e., $\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{s}_t$. If $\boldsymbol{x}_t$ is a stationary ergodic random sequence and the mixing matrix $\boldsymbol{A}$ is non-singular, it is possible to estimate a separation matrix $\boldsymbol{B}$ to obtain the sources. This separating matrix $\boldsymbol{B}$ can be decomposed into the product of a whitening $\boldsymbol{W}$ and a rotation $\boldsymbol{V}$ matrix. The whole process yields

$$\boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{x}_t = \boldsymbol{V}\boldsymbol{W}\boldsymbol{A}\boldsymbol{s}_t = \boldsymbol{V}\boldsymbol{z}_t \quad t = 1, 2, \ldots \tag{1}$$

If we approximate the possible distributions for the sources $\boldsymbol{s}$ by and Edgeworth expansion, rewrite the marginal entrophy contrast function in terms of second-order and fourth-order cumulants, and then minimize it for all possible dis-

tributions [1], it follows that

$$\phi^{ME}(\boldsymbol{y}) \approx \frac{1}{48}\phi_{24}^{ME}(\boldsymbol{y}) = -\frac{1}{48}\sum_i (C_{iiii}^{\boldsymbol{y}})^2 \qquad (2)$$

where, for zero-mean signals, $C_{iiii}^{\boldsymbol{y}} = \mathrm{E}[y_i^4] - 3\mathrm{E}[y_i^2]$ are the marginal cumulants or autocumulants and $\mathrm{E}[\cdot]$ denotes mathematical expectation. Notice that contrast ϕ_{24}^{ME} assumes the outputs are decorrelated. Thus, the problem reduces to the computation of matrix $\boldsymbol{V}$.

2. APPROXIMATIONS TO THE ME

In the two dimensional case, the pair of normalized sources $\bar{\boldsymbol{s}}_t = [\bar{s}_p(t)\ \ \bar{s}_q(t)]^{\mathrm{T}}$ in polar coordinates may be written as $(r(t), \alpha(t))$ so that the outputs yield

$$\begin{bmatrix} y_p(t) \\ y_q(t) \end{bmatrix} = \boldsymbol{V}(\theta) \begin{bmatrix} r(t)\cos(\beta(t)) \\ r(t)\sin(\beta(t)) \end{bmatrix} = \boldsymbol{V}(\theta)\boldsymbol{z}_t \qquad (3)$$

where $\boldsymbol{z}_t = [z_p(t)\ z_q(t)]^{\mathrm{T}}$ are the whitened mixtures, and matrix $\boldsymbol{V}$ performs a rotation of θ so that $\rho(t) = \theta + \beta(t)$ is the angle of vector $\boldsymbol{y}$. Notice that ideally, at separation $\theta + \beta(t) = \alpha(t)$. The contrasts in polar form and the associated estimations of the rotation angle θ may be easily expressed as a closed function of the following complex-valued linear combinations (*centroids*) [10] of the statistics of the outputs

$$\xi_\gamma = \mathrm{E}[r^4(t)e^{j4\beta(t)}] \qquad (4)$$

$$\xi_\eta = \mathrm{E}^2[r^4(t)e^{j2\beta(t)}] \qquad (5)$$

$$\gamma = \mathrm{E}[r^4(t)] - 8 \qquad (6)$$

where $j = \sqrt{-1}$.

Some estimators in this section accept a general expression, the so called 'weighted estimators' (WE) [11], [12]. In order to extend the WE estimator to the ML [13] or MK [14] case we propose the generalized weighted estimator (GWE) as

$$\hat{\theta}_{GWE}(\omega_\gamma, \omega_\xi) = \frac{1}{4}\angle(\omega_\xi\omega_\gamma\xi_\gamma + (1-\omega_\xi)\xi_\eta) \qquad (7)$$
$$0 < \omega_\xi < 1, \quad \omega_\gamma = \pm 1, \gamma$$

where $\angle(\cdot)$ supplies the principal value of its argument. As described in [11], particular cases for this contrast function has been proposed as approximations to the ME contrast function: $\hat{\theta}_{EML} = \hat{\theta}_{GWE}(\gamma, 1)$, $\hat{\theta}_{AML} = \hat{\theta}_{GWE}(\gamma, 1/3)$, $\hat{\theta}_{AEML} = \hat{\theta}_{GWE}(\gamma, 0)$ and $\hat{\theta}_{MaSSFOC} = \hat{\theta}_{GWE}(\gamma, 1/2)$. With the GWE in (7) we may rewrite the estimators in [15], MK [14], [2], SKSE or ML [13] as $\hat{\theta}_{GWE}(\pm 1, 1)$. In [8] the authors show that for a large enough number of observations T, the minimization of $\phi_{24}^{ME}(\theta)$ yields the angle

$$\hat{\theta}_{SICA} = \hat{\theta}_{GWE}(\gamma, 3/7) \qquad (8)$$

3. N-DIMENSIONAL CASE

3.1. Not initizalized Jacobi optimization

The GWE contrast developed in the last sections may be applied directly to the 2-dimensional ICA problem. Comon introduced in [1] the 'Jacobi optimization' to extend the solution of a contrast $\phi(\theta)$ to the n-dimensional problem. Such an algorithm can be summarized as follows.

Algorithm 1 *Non Initialized n-dimensional GWE using Jacobi Optimization: GWE-JO.*

1. *Whitening. Compute a whitening matrix $\boldsymbol{W}$ and the output vector $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$.*
2. *One sweep. For all $g = n(n-1)/2$ pairs, i.e., for $1 \le p < q \le n$, do*
   (a) *Compute the Givens angle $\theta_{pq} = \theta_{GWE}$ in (7) (with $[z_p\ z_q]^{\mathrm{T}} = [y_p\ y_q]^{\mathrm{T}}$).*
   (b) *if $\theta_{pq} > \theta_{min}$, do rotate the pair (y_p, y_q) by θ_{pq} according to (3).*
3. *End? If the number of iterations n_{it} satisfies $n_{it} \ge K = 1 + \sqrt{n}$ or no angle θ_{ij} has been updated, stop. Otherwise go to step 2 for another sweep.*

Thus, the Jacobi approach considers a sequence of 2-dimensional ICA problems. In [9] the algorithms only stops when the set of g Givens rotations have been updated by a value under a threshold θ_{min}. But there is no limit on the number of iterations n_{it}. The value θ_{min} is selected in such a way that rotations by a smaller angle are not 'statistically significant'. Typically $\theta_{min} = 10^{-2}/\sqrt{T}$ where T is the number of samples. In [1] if the algorithm goes through step 2 more than k times with $k \le 1 + \sqrt{n}$ it stops. We use this limit in our approach as it was designed for the ME contrast. However, for a large number of components it may also used a fixed $\theta_{min} = \pi/360$ just to avoid useless computations.

3.2. Initilized Jacobi Optimization

In the previous section the 'Jacobi optimization' was introduced to extend the problem to n dimensions. In the step 2.a of the Algorithm 1, the Givens angle θ_{pq} is computed by using equation (7) with $[z_p, z_q]$ the $[y_p, y_q]$ computed at the previous iteration. Simple calculus and trigonometrics show that θ_{pq} may be written as a function of the moments of the output $\mathrm{E}[y_p^2 y_q^2], \mathrm{E}[y_p^4], \mathrm{E}[y_q^4], \mathrm{E}[y_p y_q^3]$, and $\mathrm{E}[y_p^3 y_q]$. Thus, these five moments are to be computed $n_{it} \cdot g$ times, where $n_{it} \le K = 1 + \sqrt{n}$ is the total number of iterations and g the number of Givens rotations. Bearing this in mind, we will face next the computation of the whole set of moments

just one time at an initial stage and then rotate them at each step of the algorithm.

Proposition 1 *Given the model $\boldsymbol{y}_t = \boldsymbol{V}\boldsymbol{z}_t$ in (1), there exist a symmetric $l \times l$, $l = n(n+1)/2$, matrix*

$$\boldsymbol{M}^z(a(k,l), b(i,j)) = \mu^z_{ijkl} = E[z_i z_j z_k z_l], \tag{9}$$

a diagonal constant matrix $\boldsymbol{S}$ and vectors $\boldsymbol{v}_{pp}, \boldsymbol{v}_{pq}$ and $\boldsymbol{v}_{qq}$ such that the fourth order moments of the outputs, y_p and y_q, yields

$$E[y_p^2 y_q^2] = \boldsymbol{v}_{pp}\boldsymbol{S}\boldsymbol{M}^z\boldsymbol{S}\boldsymbol{v}_{qq}^{\mathrm{T}} \tag{10}$$
$$E[y_p^4] = \boldsymbol{v}_{pp}\boldsymbol{S}\boldsymbol{M}^z\boldsymbol{S}\boldsymbol{v}_{pp}^{\mathrm{T}} \tag{11}$$
$$E[y_q^4] = \boldsymbol{v}_{qq}\boldsymbol{S}\boldsymbol{M}^z\boldsymbol{S}\boldsymbol{v}_{qq}^{\mathrm{T}} \tag{12}$$
$$E[y_p y_q^3] = \boldsymbol{v}_{pq}\boldsymbol{S}\boldsymbol{M}^z\boldsymbol{S}\boldsymbol{v}_{qq}^{\mathrm{T}} \tag{13}$$
$$E[y_p^3 y_q] = \boldsymbol{v}_{pp}\boldsymbol{S}\boldsymbol{M}^z\boldsymbol{S}\boldsymbol{v}_{pq}^{\mathrm{T}}, \tag{14}$$

Proof: See Appendix A.

The formulation introduced above allows an easy computation of the output statistics for a given rotation matrix, as the entries $\boldsymbol{V}(p,q)$ involved are easily arranged in a pair of rotations vectors. Besides, notice that only the subset $1 \leq i \leq j \leq k \leq l \leq n$ is needed to obtain the entries of matrix $\boldsymbol{M}^z$. The number of computed moments reduces from n^4 to $(n+3)!/(n-1)!4!$. Finally, the constant diagonal matrix $\boldsymbol{S}$ may be easily included in the computation of the rotation vectors. Thus, no extra memory is needed and we reduce the number of operations in (10)-(14).

The ICA algorithm by using this algebraic structure yields

Algorithm 2 *n-dimensional GWE using Initialized Jacobi Optimization: GWE-IJO.*

1. *Whitening. Compute a whitening matrix $\boldsymbol{W}$ and set $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$.*
2. *Moments Initialization. Compute matrix $\boldsymbol{M}^z$ in (9).*
3. *One sweep. For all $g = n(n-1)/2$ pairs, i.e., for $1 \leq p < q \leq n$, do*
   (a) *Compute the Givens angle θ_{pq} in $\theta_{GWE}(\boldsymbol{y})$ in (7) for the pair (y_p, y_q) as a function of the moments in (10)-(14).*
   (b) *if $\theta_{pq} > \theta_{min}$, do update the rotation matrix $\boldsymbol{V}$ with rotation angle θ_{pq}.*
4. *End? If the number of iterations n_{it} satisfies $n_{it} \geq K = 1 + \sqrt{n}$ or no angle θ_{pq} has been updated, stop. Otherwise go to step 3 for another sweep.*

In [1] and [16] the set of moments in (10)-(14) is computed at each step (Givens rotation) of the 'Jacobi Optimization' from the outputs $\boldsymbol{y} = \boldsymbol{V}\boldsymbol{z}$. This idea was applied to the GWE-JO in the previous subsection. The main advantage of the formulation presented in this section is that matrix $\boldsymbol{M}$ is computed once. The moments for each pair of outputs needed at each step of Algorithm 2, are computed by a simple product of vectors and matrices. However, at a large number of components, the number of entries of the moment matrix is of order $\mathcal{O}(n^4)$. Thus, the method becomes unavailable. Similarly to the I-SICA method, the JADE algorithm [2], [9], [17] computes a cumulant matrix, $\boldsymbol{\mathcal{C}}$, at the beginning of the algorithm. But this matrix is updated at each step of the Givens angles computation. Another difference between both methods is that $\boldsymbol{\mathcal{C}}$ has $2 \cdot 2n/(n+1)$ times more elements than $\boldsymbol{M}^z$ in (9). This just emphasizes the memory problems. We devote the next paragraphs to provide a solution to this problem.

3.3. Optimized SICA

Both of the non-initialized and initialized algorithms provide the same solution. Thus, we must decide on the number of samples and components to reduce the computational burden. The number of multiplications and accumulations (MACs) is proposed as an index to compare both of the methods. As the computation of a fourth-order moment is approximately $4T$, the number of MACs operations for the non-initialized Algorithm 1 may be approximated by

$$MAC_{JO} = K \cdot g \cdot 5 \cdot 4T + K \cdot g \cdot 4T \tag{15}$$

where $g = n(n-1)/2$ is the number of Givens angles and K is the maximum number of iterations in the 'Jacobi Optimization'. The first term is the computational burden related to the calculation of the moments. The second one must be included to take into account the rotation of the data performed at step 2.b of the proposed GWE-JO Algorithm 1. Regarding the GWE-IJO Algorithm 2, an approximation to the number of MAC operations yields

$$MAC_{IJO} = \frac{(n+3)!}{(n-1)!4!} \cdot 4T + K \cdot g \cdot (l^3 + l) \tag{16}$$

where $l = n(n+1)/2$ is the dimension of the moments matrix $\boldsymbol{M}^z$ in (9). The first term is the number of operations involved in the calculation of this matrix. The second one, the operations in computing (10)-(14) at each Givens angle. The relation between them is as follows

$$\eta(n,T) = \frac{MAC_{JO}}{MAC_{IJO}} = \frac{K \cdot g \cdot 5 \cdot 4T + K \cdot g \cdot 4T}{\frac{(n+3)!}{(n-1)!4!} \cdot 4T + K \cdot g \cdot 5 \cdot (l^2 + l)} \tag{17}$$

In Fig. we 1 have depicted the function $\eta^T(n,T) = 1$. As $n \to \infty$ the number of moments μ^z_{ijkl} becomes of the order

$\mathcal{O}(n^4)$ and makes $\eta(n,T)=1$ non sensitive to the number of samples. This takes place approximately at $n=15$. This allow us to draw the following important conclusions

- For a low number of sources, $n \leq 5$, the GWE-IJO is to be used for any number of samples, $T > 10^2$.
- For $5 < n \leq 15$, the number of samples should be taken into account to decide the method to use.
- For $n > 15$ the GWE-JO should be selected for any number of samples.

Notice that as the GWE-IJO is not to be used for large numbers of components, we avoid memory problems associated to the storage of matrix $\boldsymbol{M}^z$.

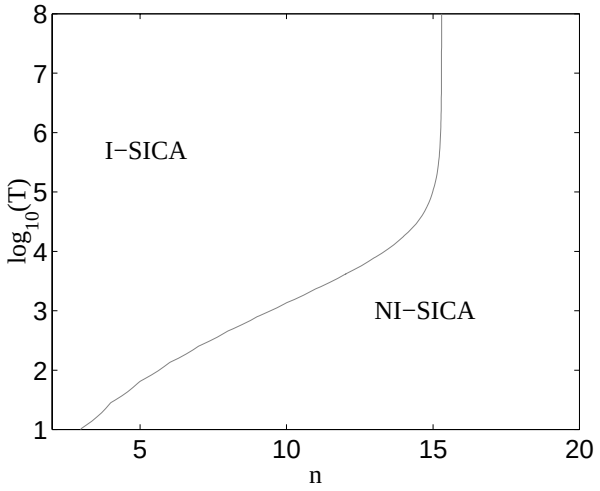

Fig. 1. Computational burden ratio for the GWE-JO and the GWE-IJO methods.

As a result, the following algorithm is proposed,

Algorithm 3 *Optimized Sinusoidal Independent Component Analysis: GWE-OJO.*

1. *Threshold. Compute the condition $\eta(n,T)$ in (3.3).*
2. *Decision on $\eta(n,T)$.*

   If $\eta(n,T) < 1$

   Then go to Algorithm GWE-JO in 1.

   Else go to Algorithm GWE-IJO in 2.

4. EXPERIMENTAL RESULTS

To illustrate the main results of this paper, the interferers to signal power ratio (ISR) is used as an objective separation index [18]. We computed the mean of this index for the set of output components y_i. As GWE we will use $\hat{\theta}_{SICA}$ in (8) [8]. In the following, this contrast solved with OJO will be referred as SICA. The performance of the multidimensional SICA will be compared to the JADE [2], [17], the fourth order based ME method by Comon [1], [16], and the Fast-ICA [19], [20] algorithm. The same whitening method was introduced in all of the algorithms as we focus on the computation of the unitary matrix $\boldsymbol{V}$ in (1). Besides, a few changes were introduced in the code by Comon [16] to save up some operations and the Fast-ICA [20] was executed with the parameters by default including stabilization. In the SICA, JADE and ME by Comon, the Jacobi optimization stops whenever no angle has been updated more than $\pi/360$ rad or it has iterated more than $K = 1 + \sqrt{n}$ times. In the experiments bellow, both of the floating operations (flops) and CPU time will be included. Notice that some flops may be avoided by using intensive storage of data and data access, i.e., by increasing CPU time. Besides, some CPU time may be saved by increasing the number of flops. The simulations have been performed using MATLAB on an Intel Pentium IV 1.4GHz processor with 256 Mb. In the following, random mixing matrices entries a_{ij} are random numbers in the range [-1,+1].

We propose 2 different experiments for different dimensions. We first study the separation of mixtures of 6 sources with different distributions. We then separate 50 mixtures of uniformly distributed signals.

4.1. Performace in the 6-dimensional case

In Fig. 2 $n=6$ zero mean unit variance signals with different distributions were mixed: uniform, laplacian ($\mu = .1$), rayleigh (B=1), exponential ($\mu = 1$), gaussian, and lognormal ($\sigma = .1$). Each point corresponds to the average of 1000 experiments in which the mixing matrix is randomly chosen. Fig. 2.a shows that SICA and ME by comon have a similar peformance, as expected. Besides, JADE and Fast-ICA have good performance close to that of the SICA. Regarding computational features, in Fig. 2.b and Fig. 2.c we study the number of floating point operations and the CPU time. The SICA method presented in this paper clearly outperforms the other ones. Notice that although running the JADE algorithm takes a larger number of flops than the Fast-ICA and the ME by comon, the CPU time is lower than for these two methods.

4.2. Mixture of 50 uniformly distributed sources

As an example of a mixture of a large number of sources we propose the separation of 100 random mixtures of $n=50$ uniformly distributed sources. In this case, we got the better ISR with the SICA algorithm at the lowest CPU time and number of floating operations compared to the ME by Comon and JADE algorithms. The Fast-ICA algorithm took

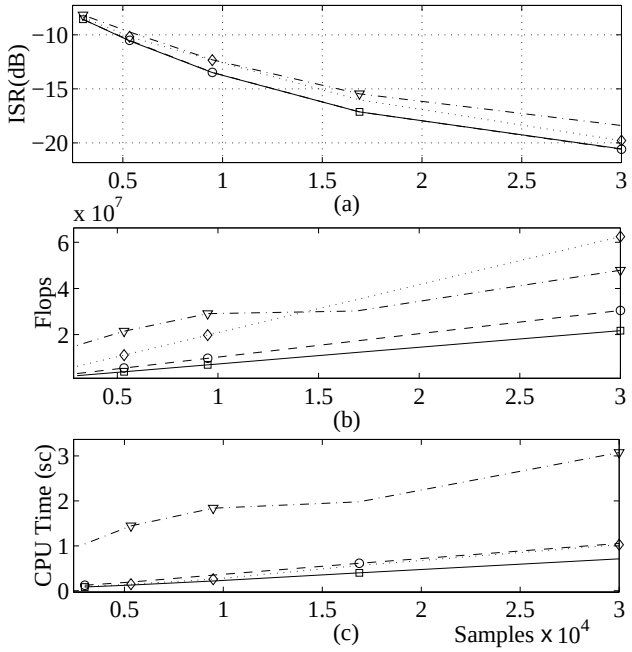

Fig. 2. (a) Mean ISR, (b) Flops and (c) CPU Time in the n=6 dimensional case for the SICA (□), ME by Comon (○), JADE (◇) and Fast-ICA (▽).

more floating operations than the SICA but reduced the CPU time in a 15%. However, it presented a quite poor performance.

Table 1. Comparison between the SICA, JADE, ME by Comon and Fast-ICA algorithms: flops, CPU time, and ISR. Mixture of 50 uniformly distributed sources.

Method	FLOPS	CPU time (sc)	ISR (dB)
SICA	$0.16 \cdot 10^{10}$	60.4	-26.29
ME-Comon	$0.27 \cdot 10^{10}$	74.5	-25.18
JADE	$6.86 \cdot 10^{10}$	499.3	-24.50
FastICA	$0.20 \cdot 10^{10}$	53.5	-6.54

5. CONCLUSIONS

In this paper we present a new approach to independent component analysis based on fourth-order moments under whitening constraint. The starting point is the marginal entropy contrast given by Comon in [1]. This contrast reduces, in the 2-dimensional case, to a sinusoidal function. This strategy results in a simple method with accurate results at a low computational cost. Then, we exploit the algebraic structure of the problem to computationally optimize the method. This new method, Initialized-SICA (I-SICA), computes the set of fourth-order moments at the beginning of the 'Jacobi Optimization'. Then we relate the non-initialized and the initialized versions by studying the number of operations. We conclude that the decision is not sensitive to the number of samples but to the number of components. In the last section devoted to experiments, this new method has been compared to the ICA by Comon, the JADE and the Fast-ICA algorithms. We achieved a similar or lower interference to signal ratio at the lowest number of floating operations or CPU time.

A. PROOF OF PROPOSITION 1

After some simple calculus, (10) yields

$$\mu^y_{ppqq} = \sum_{ij} V_{pi} V_{pj} \sum_{kl} V_{qk} V_{ql} \mu^z_{ijkl} \tag{18}$$

Notice that for $p = q$ we have an analogous result for (11) and (12). Besides,

$$\mu^y_{pppq} = \sum_{ij} V_{pi} V_{pj} \sum_{kl} V_{pk} V_{ql} \mu^z_{ijkl} \tag{19}$$

Let's $\boldsymbol{M}^z$ be a $l \times l$, $l = n(n+1)/2$, symmetric matrix whose entries are the fourth-order moments $\mu^z_{ijkl} : 1 \le i \le j \le n, 1 \le k \le l \le n$. The moment μ^z_{ijkl} is stored in the entry $\boldsymbol{M}^z(a,b)$, where

$$b = \sum_{h=n-i+2}^{n} h + (j-i) : \ 1 \le i \le j \le n \tag{20}$$

$$a = \sum_{h=n-k+2}^{n} h + (l-k) : \ 1 \le k \le l \le n \tag{21}$$

Notice that $\mu_{ijkl} = \mu_{jikl} = \mu_{kjil} = \mu_{ikjl} = \ldots$. Thus, given the column b it is possible to define $\tilde{j} = \sum_{h=n-j+2}^{n} h$ such that only the last $m - \tilde{j}$ entries (moments) are computed. The first $\tilde{j}$ ones were already calculated in previous columns. Hence, we only estimate the subset of different moments μ_{ijkl} $1 \le i \le j \le k \le l \le n$.

Besides, the computation of (18) and (19) may be rewritten by introducing a pair of 'rotation vectors' to left-right multiply matrix $\boldsymbol{M}^z$. The entries a of these vectors written as a function of the entries of the unitary matrix $\boldsymbol{V}$ in (1) yields

$$\begin{aligned} \boldsymbol{v}_{pp}(a) &= 2\boldsymbol{V}(p,k)\boldsymbol{V}(p,l) \\ \boldsymbol{v}_{pq}(a) &= \boldsymbol{V}(p,k)\boldsymbol{V}(q,l) + \boldsymbol{V}(p,k)\boldsymbol{V}(q,l) \\ \boldsymbol{v}_{qq}(a) &= 2\boldsymbol{V}(q,k)\boldsymbol{V}(q,l) \end{aligned} \tag{22}$$

where the indexes k, l and a are related through (21). Finally, as there are different number of repetitions of moments μ_{iiii}, μ_{iijj} and μ_{ijkl} for $1 < i,j,k,l \le n$, some

entries of the rotation vectors should be multiplied by $1/2$. This may be written as a multiplication by a diagonal matrix whose entries $\boldsymbol{S}(a,a)$, $a(k,l)$ as in (21), yield

$$S(a,a) = 1 \;\; l \neq 1 \quad (23)$$

$$S(a,a) = 1/2; \;\; l = 1 \quad (24)$$

B. REFERENCES

[1] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[2] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals," *Proceedings IEE F*, vol. 140, no. 6, pp. 362–370, Dec 1993.

[3] J. F. Cardoso, "Blind signal separation: Statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct 1998.

[4] Antonio Caama no, Juan J. Murillo-Fuentes, Francisco J. González-Serrano, and Javier Ramos, "Natural gradient based multiuser detection," in *PIRMC'02*, Lisbon, Sep 2002.

[5] J.J. Murillo-Fuentes and F.J. González-Serrano, "Median equivariant adaptive separation via independence: application to communications," *Neurocomputing*, vol. (Accepted for publication), 2001.

[6] S. Makeig, T-P. Jung, D. Ghahremani, A.J. Bell, and T.J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proc. Natl. Acad. Sci. USA*, pp. 10979–10984, 1997.

[7] J.J. Murillo-Fuentes, H. Molina-Bulla, and F. González-Serrano, "Independent component analysis applied to digital image watermarking," in *International Conference on Audio, Speech and Signal Processing*, Salt Lake City, USA, May 2001, vol. III, pp. 1997–2000.

[8] J.J. Murillo-Fuentes and F. González-Serrano, "Independent component analysis with sinusoidal fouth-order contrast," in *International Conference on Audio, Speech and Signal Processing*, Salt Lake City, USA, May 2001, vol. V, pp. 2785–2788.

[9] J. F. Cardoso, "High-order contrasts for independent component analysis," *Neural Computation*, vol. 11, no. 1, pp. 157–192, Jan 1999.

[10] Vicente Zarzoso and Asoke K. Nandi, "Blind separation of independent sources for virtually any source probability density function," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2419–2432, September 1999.

[11] Vicente Zarzoso, Frank Herrmann, and K. Nandi, "Weighted closed-form estimators for blind source separation," in *11th International Workshop on Statistical Signal Processing*, Singapore, August 2001.

[12] M. Ghogho, A. Swami, and T. Durrani, "Approximate maximum likelihood blind source separation with arbitrary source pdfs," in *IEEE Workshop on Statistical Signal and Array Processing (SSAP'00)*, Pocono Manor Inn, Pennsylvania, Aug 2000.

[13] F. Harroy and J.-L Lacoume, "Maximum likelihood estimators and cramer-rao bounds in source separation," *Signal Processing*, vol. 55, no. 2, pp. 167–177, 1996.

[14] E. Moreau and O. Macchi, "Higher order contrast for self-adaptive source separation," *International Journal of Adaptive Control and Signal Processing*, vol. 10, no. 1, pp. 19–46, Jan 1996.

[15] P. Comon and E. Moreau, "Improved contrast dedicated to blind separation in communications," in *IEEE International Conference on Acoustics, Speach and Signal Processing*, Munich, Germany, 1997, vol. V, pp. 3453–3456.

[16] P. Comon, "Web page: Matlab source codes," WWW, http://www.i3s.unice.fr/∼comon/matlab.html.

[17] J. F. Cardoso, "Web page: Jade for real-valued data," WWW, ftp://sig.enst.fr/pub/jfc/Algo/Jade/jadeR.m.

[18] V. Zarzoso and A.K. Nandi, *Blind Source Separation*, pp. 167–252, Kluwer Academics Publishers, Boston, 1999.

[19] A. Hyvrinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[20] Jarmo Hurri, Hugo Gävert, Jaakko Särelä, and Aapo Hyvrinen, "Web page: Matlab source codes," 1998, http://www.cis.hut.fi/projects/ica/fastica/.