A NEW ALGORITHM FOR INDEPENDENT COMPONENT ANALYSIS WITH OR WITHOUT CONSTRAINTS

Xuejun Liao and Lawrence Carin

Department of Electrical and Computer Engineering
Duke University,
Durham, NC 27708-0291

ABSTRACT

A new algorithm is developed for independent component analysis (ICA) with or without constraints on the mixing matrix or sources. The algorithm is based on the criterion of Joint Approximate Diagonalization of Eigen-matrices (JADE). We propose a column-wise processing approach to perform joint diagonalization of the cumulant (eigen-) matrices. We utilize the unitary property of diagonalizing matrix **U** and achieve decoupling of its columns via orthogonal projections. We propose a method called Alternating Eigen-search (AE) to maximize the JADE criterion with respect to one column of **U** at a time. The method is extended to the case in which there are application-dependent quadratic constraints imposed on the mixing matrix or sources, resulting in the so-called constrained ICA. Example results are provided to demonstrate the effectiveness and applicability of the algorithm.

1. INTRODUCTION

Independent Component Analysis (ICA) [1-2] is an important technique to solve the problem of blind source separation assuming that the sources are mutually independent. The standard ICA is formulated as

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) = \sum_{k=1}^{N} s_k(t)\mathbf{a}_k \qquad (1)$$

where $\mathbf{y}(t) = [y_1(t)\ \ y_2(t)\ \cdots y_{N_0}(t)]^{\mathrm{T}}$ are N_0 dimensional observed mixed-signals, $\mathbf{s}(t) = [s_1(t)\ \ s_2(t)\ \cdots s_N(t)]^{\mathrm{T}}$ are N dimensional statistically independent source signals with at most one of its components Gaussian, $\mathbf{A} = [\mathbf{a}_1\ \ \mathbf{a}_2\ \ \cdots\ \mathbf{a}_N]$, with $\mathbf{a}_k = [a_{1k}\ \ a_{2k}\ \ \cdots\ a_{N_0 k}]^{\mathrm{T}}$, is the mixing matrix with full column rank, and t is an index for time or other relevant variables.

Cardoso and Souloumiac [2] have proposed an elegant method, called Joint Approximate Diagonalization of Eigen-matrices (JADE), to solve (1). In their original paper [2], the joint diagonalization is achieved via an extended Jacobi technique, which consists of applying successive planar Givens rotations, each time to a pair of columns of the unitary diagonalizing matrix **U** such that the associated 2-by-2 sub-matrices are as diagonal as possible. The unitary **U** together with the whitening matrix determines the mixing matrix **A**. In this paper, we propose a column-wise processing approach to perform joint diagonalization. Instead of successively selecting paired columns of **U** to which to apply planar rotations, we sequentially process the columns of **U** and maximize the JADE criterion with respect to each individual column separately. We utilize the unitary property of **U** and achieve decoupling of its columns by the idea of orthogonal projections [3]. We propose a method called Alternating Eigen-search (AE) to maximize the JADE criterion one column of **U** at a time. An advantage of the new algorithm is that it is easily extended to the case of constrained ICA, where there are application-dependent quadratic constraints imposed on columns of the mixing matrix **A** or the sources $\mathbf{s}(t)$.

2. REVIEW OF THE JADE CRITERION

The JADE-based ICA starts with the pre-whitening of $\mathbf{y}(t)$. Assume the singular value decomposition (SVD) of **A** is $\mathbf{A}=\mathbf{V}_0\mathbf{\Sigma}_0\mathbf{U}$, where $\mathbf{V}_0$ and **U** are unitary and $\mathbf{\Sigma}_0$ is a diagonal matrix of full column rank (we assume **A** has full column rank). If $\mathbf{\Sigma}_0$ is not square, it can always be made so by pruning the rows containing zero singular values (or small singular values when there exist additive noises in $\mathbf{y}(t)$). Denote the pruned version of $\mathbf{\Sigma}_0$ as $\mathbf{\Sigma}$. Accordingly, the corresponding columns of $\mathbf{V}_0$ should be pruned. Denote by **V** the pruned version of $\mathbf{V}_0$. This results in a pruned version of the original SVD, i.e., $\mathbf{A}=\mathbf{V\Sigma U}$, where $\mathbf{\Sigma}$ is N-by-N, **V** is N_0-by-N (full column rank), and $N \le N_0$ is the reduced dimension. Substituting $\mathbf{A}=\mathbf{V\Sigma U}$ into (1) gives

$$\mathbf{y}(t) = \mathbf{V\Sigma U}\mathbf{s}(t) = \mathbf{V\Sigma}\mathbf{z}(t) \qquad (2)$$

where

$$\mathbf{z}(t) = \mathbf{U}\mathbf{s}(t) \qquad (3)$$

are the whitened mixed-signals and $\mathbf{\Sigma}^{-1}\mathbf{V}^H$ defines the whitening matrix. Here we have assumed unit variances for the components of s(t), which can always be made true by absorbing the variances of s(t) into the magnitudes of columns in **A**. This does not affect the results because ICA can only recover the sources and columns of **A** up to a permutation and scaling factor [1-2].

After the whitening, the ICA problem becomes blindly solving (3), for which the JADE criterion is defined as

$$\begin{aligned} JADE(\mathbf{U}) &= \sum_{k,m,n=1}^{N} | cum[s_k(t), s_k^*(t), s_n(t), s_m^*(t)] |^2 \\ &= \sum_{r=1}^{N^2} \| diag(\mathbf{U}^H \mathbf{Q}_z(\mathbf{B}_r)\mathbf{U}) \|^2 = \sum_{k=1}^{N}\sum_{r=1}^{N^2} | \mathbf{u}_k^H \mathbf{Q}_z(\mathbf{B}_r)\mathbf{u}_k |^2 \end{aligned} \quad (4)$$

where $\mathbf{u}_k$ is the k-th column of **U**, $\mathbf{B}_r$ with r=1, 2, ..., N^2 constitute a set of orthonormal bases for the space of N×N matrices, $\mathbf{Q}_z(\mathbf{B}_r)$ is the cumulant matrix defined element-wise as [2]

$$[\mathbf{Q}_z(\mathbf{B}_r)]_{ij} = \sum_{p,q=1}^{N} cum[z_i(t), z_j^*(t), z_q(t), z_p^*(t)] b_{pq}^{(r)} \quad (5)$$

$cum(\cdot)$ denotes the cumulant, * denotes complex conjugate, and $b_{pq}^{(r)}$ is the (p,q)-th element of $\mathbf{B}_r$.

3. THE NEW ALGORITHM FOR ICA

In this section, we develop a new ICA algorithm based on joint diagonalization of $\{\mathbf{Q}_z(\mathbf{B}_r)\}_{r=1}^{N^2}$. For ease of notation, we abbreviate $\mathbf{Q}_z(\mathbf{B}_r)$ as $\mathbf{Q}_r$ in the subsequent sections. Our objective is to find a unitary matrix $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_N]$ such that (4) is maximized, i.e.,

$$\mathbf{U} = \arg\max \sum_{k=1}^{N}\sum_{r}^{N^2} | \mathbf{u}_k^H \mathbf{Q}_r \mathbf{u}_k |^2$$

3.1. Column-wise processing

Our algorithm solves for each column of **U** separately. Each time we find a column, say $\mathbf{u}_k$, that maximizes $\sum_r^{N^2} | \mathbf{u}_k^H \mathbf{Q}_r \mathbf{u}_k |^2$ under the constraint that $\mathbf{u}_k$ has unit norm and is orthogonal to all columns found previously. Specifically, we have a constrained optimization problem

$$\mathbf{u}_k = \arg\max \sum_r^{N^2} | \mathbf{u}_k^H \mathbf{Q}_r \mathbf{u}_k |^2 \quad \text{(6-A)}$$

$$\text{s.t.} \quad \mathbf{u}_k^H \mathbf{u}_k = 1 \quad \text{(6-B)}$$

$$\mathbf{u}_k^H \mathbf{u}_l = 0, \quad l = 1, 2, \cdots, k-1 \quad \text{(6-C)}$$

We use the orthogonal-projection-based idea [3] to eliminate the constraints in (6-C). Denote the N-by-N identity matrix as $\mathbf{I}_N$. Let $\mathbf{U}_{k-1} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_{k-1}]$. Then $\mathbf{U}_{k-1}^H \mathbf{U}_{k-1} = \mathbf{I}_{k-1}$, which is a result of $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_{k-1}$ satisfying (6-B,C). Denote $\Omega_{k-1} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{k-1}\}$, and $\Omega_{k-1}^{\perp}$ the orthogonal complement of Ω_{k-1}. We construct

$$\tilde{\mathbf{P}}_k = \mathbf{I}_N - \mathbf{U}_{k-1}(\mathbf{U}_{k-1}^H \mathbf{U}_{k-1})^{-1}\mathbf{U}_{k-1}^H = \mathbf{I}_N - \mathbf{U}_{k-1}\mathbf{U}_{k-1}^H \quad (7)$$

which is the matrix of the orthogonal projection onto $\Omega_{k-1}^{\perp}$. The orthogonalization of columns of $\tilde{\mathbf{P}}_k$ gives $\mathbf{P}_k$, which satisfies $\mathbf{P}_k^H \mathbf{P}_k = \mathbf{I}_{N-k+1}$. Let $\mathbf{u}_k = \mathbf{P}_k \mathbf{w}_k$. Clearly $\mathbf{u}_k \in \Omega_{k-1}^{\perp}$ and therefore satisfies the constraints in (6-C). Moreover, the constraint (6-B) implies that $(\mathbf{P}_k \mathbf{w}_k)^H \mathbf{P}_k \mathbf{w}_k = 1$, which, using $\mathbf{P}_k^H \mathbf{P}_k = \mathbf{I}_{N-k+1}$, is reduced to $\mathbf{w}_k^H \mathbf{w}_k = 1$. Thus, we have a dual problem to (6) in $\Omega_{k-1}^{\perp}$ as

$$\mathbf{u}_k = \mathbf{P}_k \mathbf{w}_k \quad \text{(8-A)}$$

$$\mathbf{w}_k = \arg\max \sum_r^{N^2} | \mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \mathbf{w}_k |^2 \quad \text{(8-B)}$$

$$\text{s.t.} \quad \mathbf{w}_k^H \mathbf{w}_k = 1 \quad \text{(8-C)}$$

3.2. Symmetric Alternating Eigen-search

We propose a method to solve the constrained optimization problem in (8). The objective is to solve for $\mathbf{w}_k$. The solution $\mathbf{u}_k$ to (6) follows immediately as $\mathbf{u}_k = \mathbf{P}_k \mathbf{w}_k$. First we find the Karush-Kuhn-Tucker (KKT) optimality conditions for (8-B,C). The Lagrangian function for (8-B,C) is

$$L(\mathbf{w}_k, \lambda) = \sum_{r=1}^{N^2} | \mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \mathbf{w}_k |^2 - \lambda(\mathbf{w}_k^H \mathbf{w}_k - 1) \quad (9)$$

Following the definition in [4] for differentiation with respect to complex vectors, we differentiate (9) with respect to $\mathbf{w}_k^*$ and set the result to zero, thus obtaining

$$\mathbf{\Gamma}_k(\mathbf{w}_k)\mathbf{w}_k = \lambda \mathbf{w}_k \quad (10)$$

with the matrix function $\mathbf{\Gamma}_k(.)$ defined as

$$\begin{aligned} \mathbf{\Gamma}_k(\boldsymbol{\beta}) = \frac{1}{2}\sum_{r=1}^{N^2} [& (\boldsymbol{\beta}^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \boldsymbol{\beta}) \mathbf{P}_k^H \mathbf{Q}_r^H \mathbf{P}_k \\ & + (\boldsymbol{\beta}^H \mathbf{P}_k^H \mathbf{Q}_r^H \mathbf{P}_k \boldsymbol{\beta}) \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k] \end{aligned} \quad (11)$$

for any $\boldsymbol{\beta} \in C^{N-k+1}$. It is easily checked that $\mathbf{\Gamma}_k(\boldsymbol{\beta})$ is a Hermitian matrix. (10) together with (8-C) constitute the KKT conditions that the optimal $\mathbf{w}_k$ of (8-B,C) must satisfy. We aim to solve (10) subject to (8-C), thus finding the optimal $\mathbf{w}_k$.

Equations (10) and (8-C) imply that at the stationary point $\mathbf{w}_k$ is an eigenvector of $\mathbf{\Gamma}_k(\mathbf{w}_k)$, with the associated eigenvalue

$$\lambda = \mathbf{w}_k^H \mathbf{\Gamma}_k(\mathbf{w}_k)\mathbf{w}_k = \sum_{r=1}^{N^2} | \mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \mathbf{w}_k |^2 \qquad (12)$$

which is equal to the objective function in (8-B). Therefore the optimal eigenvector $\mathbf{w}_k$ should be associated with the largest eigenvalue of $\mathbf{\Gamma}_k(\mathbf{w}_k)$, to maximize the objective function. However, $\mathbf{\Gamma}_k(\mathbf{w}_k)$ is a matrix function of $\mathbf{w}_k$, and (10) cannot be solved by a simple eigenvalue decomposition.

We propose an iterative algorithm to solve $\mathbf{w}_k$ from (10) subject to (8-C).

Algorithm 1: Symmetric Alternating Eigen-search (Symmetric AE)

Step 1: Initialize $\mathbf{w}_k^{(0)} \in C^{N-k+1}$ and $\mathbf{w}_k^{(0)} \neq \mathbf{0}$, and normalize $\mathbf{w}_k^{(0)}$ to unit L_2 norm, i.e., $[\mathbf{w}_k^{(0)}]^H \mathbf{w}_k^{(0)} = 1$. Define the convergence parameter $\varepsilon \geq 0$. Let $j = 1$.

Step 2: Find the maximum eigenvalue, denoted $\lambda^{(j)}$, of $\mathbf{\Gamma}_k(\mathbf{w}_k^{(j-1)})$, and denote the associated eigenvector with unit L_2 norm as $\mathbf{w}_k^{(j)}$.

Step 3: Check convergence. If $|\lambda^{(j)} - \lambda^{(j-1)}| \leq \varepsilon$, stop; otherwise, let $j = j+1$ and go to step 2.

We have established the following proposition for the convergence of Algorithm 1, the proof of which is omitted here because of limited space.

Proposition 2: For any nonzero unit-norm $\mathbf{w}_k^{(0)}$, the eigenvalue sequence $\{\lambda^{(j)}\}$ produced by the Symmetric AE algorithm converges, i.e., $\lim_{j\to\infty} \lambda^{(j)} = \lambda$. Furthermore, there exists a unit-norm vector $\mathbf{w}_k$ such that λ is the largest eigenvalue of $\mathbf{\Gamma}_k(\mathbf{w}_k)$ and $\mathbf{w}_k$ is the associated eigenvector.

4. EXTENSION TO CONSTRAINED ICA

4.1. Formulation of the Constrained ICA Problem

In this section we extend (1) to the case when there are quadratic constraints on the columns of **A** or on the sources **s**(*t*), and we refer to this extension as constrained ICA. Specifically, our constrained ICA is formulated as

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) = \sum_{k=1}^{N} s_k(t)\mathbf{a}_k \qquad (13\text{-A})$$

$$\text{s.t. } \mathbf{a}_k^H \mathbf{F}_k \mathbf{a}_k \leq 0, \ k = 1, 2, \cdots, N \qquad (13\text{-B})$$

where $\mathbf{F}_k$, $k = 1, 2, \cdots, N$, are assumed to be $N \times N$ Hermitian matrices, or

$$\text{s.t. } \mathbf{s}_k \mathbf{G}_k \mathbf{s}_k^H \leq 0, \ k = 1, 2, \cdots, N \qquad (13\text{-C})$$

where $\mathbf{s}_k = [s_k(1)\ s_k(2)\ \cdots\ s_k(T)]$ with T the number of samples in *t*, and $\mathbf{G}_k$ are $T \times T$ Hermitian matrices.

We now transform the quadratic constraints (13-B) and (13-C) into their "whitened" forms, corresponding to (3), and their forms in $\Omega_{k-1}^{\perp}$, corresponding to (8-B,C). First, we have from the pruned version of SVD of **A** in Section 2

$$\mathbf{a}_k = \mathbf{V}\mathbf{\Sigma}\mathbf{u}_k, \ k = 1, 2, \cdots, N \qquad (14)$$

and from (2) and (3)

$$\mathbf{s}_k = \mathbf{u}_k^H \mathbf{\Sigma}^{-1} \mathbf{V}^H \mathbf{Y} = \mathbf{u}_k^H \mathbf{Z} \qquad (15)$$

where **Y**=[**y**(1) **y**(2) ... **y**(T)] and **Z**=[**z**(1) **z**(2) ... **z**(T)]. Substitution of (14) into (13-B) gives

$$\mathbf{u}_k^H \mathbf{\Sigma}^H \mathbf{V}^H \mathbf{F}_k \mathbf{V}\mathbf{\Sigma}\mathbf{u}_k \leq 0, \ k = 1, 2, \cdots, N \qquad (16)$$

and inserting (15) into (13-C) gives

$$\mathbf{u}_k^H \mathbf{Z}\mathbf{G}_k \mathbf{Z}^H \mathbf{u}_k \leq 0, \ k = 1, 2, \cdots, N \qquad (17)$$

Next, Recalling from Section 3.1 that $\mathbf{u}_k \in \Omega_{k-1}^{\perp}$ and $\mathbf{u}_k = \mathbf{P}_k \mathbf{w}_k$, the constraints (16) are expressed in $\Omega_{k-1}^{\perp}$ as

$$\mathbf{w}_k^H \mathbf{P}_k^H \mathbf{\Sigma}^H \mathbf{V}^H \mathbf{F}_k \mathbf{V}\mathbf{\Sigma}\mathbf{P}_k \mathbf{w}_k \leq 0, \ k = 1, 2, \cdots, N \qquad (18)$$

and (17) expressed in $\Omega_{k-1}^{\perp}$ as

$$\mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Z}\mathbf{G}_k \mathbf{Z}^H \mathbf{P}_k \mathbf{w}_k \leq 0, \ k = 1, 2, \cdots, N \qquad (19)$$

Finally, combining (8) and (18)-(19), we obtain the optimization problem for constrained ICA (13) as

$$\mathbf{u}_k = \mathbf{P}_k \mathbf{w}_k \qquad (20\text{-A})$$

$$\mathbf{w}_k = \arg\max \sum_{r}^{N^2} | \mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \mathbf{w}_k |^2 \qquad (20\text{-B})$$

$$\text{s.t.} \quad \mathbf{w}_k^H \mathbf{w}_k = 1 \qquad (20\text{-C})$$

$$\mathbf{w}_k^H \mathbf{C}_k \mathbf{w}_k \leq 0 \qquad (20\text{-D})$$

where

$$\mathbf{C}_k = \mathbf{P}_k^H \mathbf{\Sigma}^H \mathbf{V}^H \mathbf{F}_k \mathbf{V}\mathbf{\Sigma}\mathbf{P}_k \qquad (21\text{-A})$$

in the case of constraining the mixing matrix **A**, or

$$\mathbf{C}_k = \mathbf{P}_k^H \mathbf{Z}\mathbf{G}_k \mathbf{Z}^H \mathbf{P}_k \qquad (21\text{-B})$$

in the case of constraining the sources **s**(*t*).

Recall that $\mathbf{F}_k$'s and $\mathbf{G}_k$'s are Hermitian matrices, therefore it follows from (21) that $\mathbf{C}_k$'s are Hermitian, too.

4.2. Complementary Alternating Eigen-search

To solve (20-B,C,D), we first derive their KKT optimality conditions. The Lagrangian function is constructed as

$$L(\mathbf{w}_k, \lambda) = \sum_{r=1}^{N^2} [| \mathbf{w}_k^H \mathbf{P}_k^H \mathbf{Q}_r \mathbf{P}_k \mathbf{w}_k |^2 - \eta(\mathbf{w}_k^H \mathbf{w}_k - 1) - \mu(\mathbf{w}_k^H \mathbf{C}_k \mathbf{w}_k)] \qquad (22)$$

with

$$\mu \geq 0 \qquad (23)$$

Differentiate both sides of (22) with respect to $\mathbf{w}_k^*$ and equating the result to zero, we obtain

$$[\mathbf{\Gamma}_k(\mathbf{w}_k) - \mu\,\mathbf{C}_k]\mathbf{w}_k = \eta\mathbf{w}_k \qquad (24)$$

where $\mathbf{\Gamma}_k(.)$ is as defined in (11). (24) together with (23), (20-C,D), and

$$\mu(\mathbf{w}_k^H\mathbf{C}_k\mathbf{w}_k) = 0 \qquad (25)$$

constitute the KKT conditions of (20-B,C,D). We aim to jointly solve these equalities and inequalities, to find the optimal $\mathbf{w}_k$.

Our solution is based on breaking the problem into two sub-problems corresponding, respectively, to $\mu = 0$, and $\mu > 0$. Considering first $\mu = 0$, this reduces to the problem in Section 3.2, that is, (10) subject to (8-C), and Algorithm 1 can be employed to solve it. If the solution $\mathbf{w}_k$ thus obtained satisfies $\mathbf{w}_k^H\mathbf{C}_k\mathbf{w}_k \le 0$, then it is the solution of (20-B,C,D); if not, it is discarded, and we switch to the case of $\mu > 0$.

The second sub-problem, which corresponds to $\mu > 0$, is re-formulated as a set of simultaneous equalities and inequalities,

$$[\mathbf{\Gamma}_k(\mathbf{w}_k) - \mu\,\mathbf{C}_k]\mathbf{w}_k = \eta\mathbf{w}_k \qquad (26\text{-A})$$

$$\mathbf{w}_k^H\mathbf{w}_k = 1 \qquad (26\text{-B})$$

$$\mathbf{w}_k^H\mathbf{C}_k\mathbf{w}_k = 0 \qquad (26\text{-C})$$

$$\mu > 0 \qquad (26\text{-D})$$

Multiplying both sides of (26-A) by $\mathbf{w}_k^H$, and using (26-B), we obtain

$$\eta = \lambda - \mu\upsilon \qquad (27)$$

where λ is as defined in (12) and

$$\upsilon = \mathbf{w}_k^H\mathbf{C}_k\mathbf{w}_k \qquad (28)$$

We have the following observations.

A). Because λ is equal to the objective function, λ should be maximized. Because $\mu>0$, maximizing η means maximizing λ and minimizing υ.

B). (29) shows that μ controls the balance between the maximization of λ and the minimization of υ. A smaller μ will emphasize maximizing λ and de-emphasize minimizing υ, and vice versa. Denote by $\bar{\upsilon}$ the value of υ at the convergence of the maximization of η. Then $\bar{\upsilon}$ will decrease with an increasing μ starting from 0. Since the solution for the case $\mu = 0$ is assumed to have failed, $\bar{\upsilon} > 0$ when $\mu = 0$. Thus the equality $\bar{\upsilon} = 0$ will be reached at $\mu = \mu_{opt}$ with μ_{opt} in the interval $(0, \mu_{\max})$, where $\mu_{\max}$ is a sufficiently large positive number.

Based on the above observations, we propose the following algorithm to solve (26). Denote the complementary matrix pair

$$\mathbf{\Delta}_k(\boldsymbol{\beta}, \mu) = \mathbf{\Gamma}_k(\boldsymbol{\beta}) - \mu(\boldsymbol{\beta}^H\,\mathbf{C}_k\boldsymbol{\beta})\mathbf{I}_{N-k+1} \qquad (29)$$

$$\mathbf{\Pi}_k(\boldsymbol{\beta}, \mu) = \mathbf{\Gamma}_k(\boldsymbol{\beta}) - \mu\,\mathbf{C}_k \qquad (30)$$

It is easily checked that $\mathbf{\Delta}_k(\ ,\ )$ and $\mathbf{\Pi}_k(\ ,\ )$ are both Hermitian. We have

Algorithm 3: Complementary Alternating Eigen-search (Complementary AE)

Step 1: Define the convergence parameters $\varepsilon_\eta \ge 0$ and $\varepsilon_c \ge 0$. Let $\mu_+^{(0)} = 0$ and $\mu_-^{(0)}$ be a sufficiently large positive number. Initialize $\mu^{(0)} = (\mu_-^{(0)} + \mu_+^{(0)})/2$. Let $i = 0$.

Step 2: Initialize $\mathbf{w}_k^{(0)} \in C^{N-k+1}$ and $\mathbf{w}_k^{(0)} \ne \mathbf{0}$, and normalize $\mathbf{w}_k^{(0)}$ to unit L_2 norm. Let $j = 1$.

Step 3: Find the maximum eigenvalue, denoted $\eta^{(j)}$, of $\mathbf{\Delta}_k(\mathbf{w}_k^{(j-1)}, \mu^{(i)})$, and denote the associated eigenvector with unit L_2 norm as $\mathbf{w}_k^{(j)}$. Let $j = j+1$. Find the maximum eigenvalue, denoted $\eta^{(j)}$, of $\mathbf{\Pi}_k(\mathbf{w}_k^{(j-1)}, \mu^{(i)})$, and denote the associated eigenvector with unit L_2 norm as $\mathbf{w}_k^{(j)}$.

Step 4: Check convergence of $\eta^{(j)}$. If $|\eta^{(j)} - \eta^{(j-1)}| \le \varepsilon_\eta$, go to step 5; otherwise, let $j = j+1$ and go back to step 3.

Step 5: Check satisfaction of the quadratic equality constraint (26-C). If $|(\mathbf{w}_k^{(j)})^H\,\mathbf{C}_k\mathbf{w}_k^{(j)}| \le \varepsilon_c$, stop; otherwise, go to step 6.

Step 6: Update μ. If $(\mathbf{w}_k^{(j)})^H\,\mathbf{C}_k\mathbf{w}_k^{(j)} < 0$, let $\mu_-^{(i+1)} = \mu^{(i)}$ and $\mu_+^{(i+1)} = \mu_+^{(i)}$, or, if $(\mathbf{w}_k^{(j)})^H\,\mathbf{C}_k\mathbf{w}_k^{(j)} > 0$, let $\mu_+^{(i+1)} = \mu^{(i)}$ and $\mu_-^{(i+1)} = \mu_-^{(i)}$. Let $\mu^{(i+1)} = (\mu_-^{(i+1)} + \mu_+^{(i+1)})/2$ and $i = i+1$, go back to step 2.

We have the following proposition regarding the convergence of Algorithm 3. The proof is omitted for the limited space.

Proposition 4: For a fixed $\mu^{(i)} > 0$, and starting from any nonzero unit L_2 norm $\mathbf{w}_k^{(0)}$, the eigenvalue sequence $\{\eta^{(j)}\}$ produced by the Complementary AE algorithm converges, i.e., $\lim_{j\to\infty}\eta^{(j)} = \eta$. Furthermore, if we assume the maximum eigenvalues of $\mathbf{\Delta}_k(\mathbf{w}_k^{(2m)}, \mu^{(i)})$ and $\mathbf{\Pi}_k(\mathbf{w}_k^{(2m+1)}, \mu^{(i)})$, $m = 0, 1, \cdots$, all have multiplicity of one, then there exist $\mathbf{w}_k^{(odd)}$ and $\mathbf{w}_k^{(even)}$, and

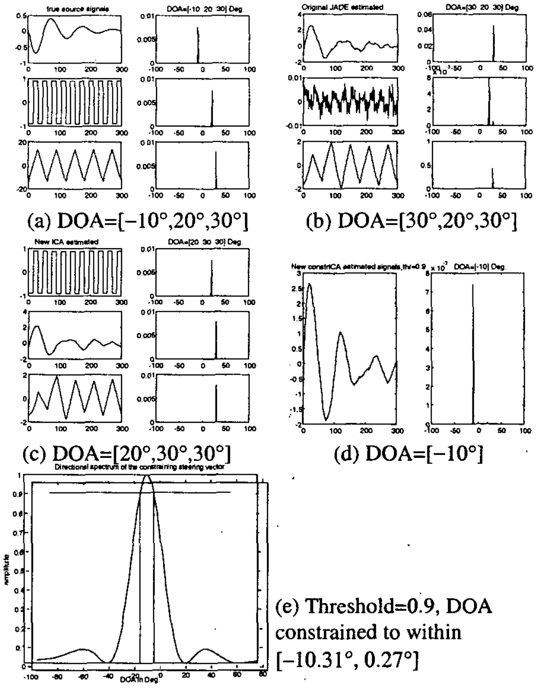
$\gamma_{1,m}, \gamma_{0,m} \in \{-1, 1\}$ such that $\lim_{m\to\infty} \gamma_{1,m} \mathbf{w}_k^{(2m+1)} = \mathbf{w}_k^{(odd)}$ and $\lim_{m\to\infty} \gamma_{0,m} \mathbf{w}_k^{(2m)} = \mathbf{w}_k^{(even)}$.

5. EXAMPLE RESULTS

ICA has been used in such diverse applications as narrowband antenna array beamforming [2], speech [5], biomedicine [6], and wireless communication [7]. Here we take beamforming as an example.

We simulate a linear array with 128 unit-gain omni-directional sensors placed in the far field of the sources. The inter-sensor distance is ¼ the wavelength. The three simulated true source signals with direction of arrival (DOA) of −10°, 20°, and 30° are shown in Fig.1 (a). The estimated source and DOA's by the original JADE and our new ICA, are shown, respectively, in Fig. (b) and (c). In Fig.(d) we show the estimation results from our constrained ICA, by thresholding the squared magnitude of the normalized correlation between the constraining steering vector and the columns of **A** at 0.9, which is equivalent to constraining the DOA of sources between −10.31° and 0.27°, as can be seen from the squared magnitudes of the directional response of the constraining steering vector shown in Fig.1 (e).

It is seen from Fig.1 that both the original JADE and our new ICA recovered the damped-sinusoid and triangle-wave, with tolerable errors. The square-wave is estimated correctly by our new ICA while the original JADE fails for this source. The estimates of DOA from these two algorithms are comparable. Our constrained ICA restricts the search of sources to within [−10.31°, 0.27°], thus pulling out only the damped-sinusoid with the correct DOA estimate of −10°, and the source estimate much improved over the corresponding ones estimated by the original JADE and our new ICA.



(a) DOA=[−10°,20°,30°] (b) DOA=[30°,20°,30°] (c) DOA=[20°,30°,30°] (d) DOA=[−10°] (e) Threshold=0.9, DOA constrained to within [−10.31°, 0.27°]

Fig.1 (a) True source signals (left panel) with DOA's of [−10°, 20°, 30°] (right panel). (b) Source signals (left panel) and DOA's of [30°,20°,30°] (right panel) estimated by original JADE. (c) Source signals (left panel) and DOA's of [20°,30°,30°] (right panel) estimated by our new ICA. (d) Source signals (left panel) and DOA's of [−10°] (right panel) estimated by our constrained ICA. (e) Squared magnitudes of the directional response of the constraining steering vector. Note: Here the order of DOA's shown from left to right is the same as the order of sources shown from top to bottom.

6. CONCLUSIONS

We have presented new algorithms for ICA with or without constraints. The new algorithms are based on maximization of the JADE criterion, subject to application-dependent quadratic constraints in the case of constrained ICA. From the numerous examples we have run, we find that the new ICA algorithm performs at least as well as, and in many cases better than, the original JADE algorithm, and that the constrained ICA not only prune away the unwanted sources correctly but also helps to improve the estimates.

7. REFERENCES

[1] P. Comon, "Independent component analysis, A new concept?", *Signal Processing*, vol. 36, pp. 287–314, 1994.

[2] J.-F. Cardoso and A. Souloumiac, "Blind Beamforming for non-Gaussian Signals," *Proc. Inst. Elect. Eng.* F, vol. 40, pp. 362–370, 1993.

[3] Jian-kang Zhang, Aleksandar Kavcic, Xiao Ma, and Kon Max Wong, " Design of Unitary Precoders for ISI Channels", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002, Vol. 3, p.p. 2265 -2268

[4] Haykin, S., *Adaptive Filter Theory*, Prentice Hall, 3rd edition, 1996

[5] A. Hyvärinen and E. Oja. "Independent Component Analysis: Algorithms and Applications." *Neural Networks*, 13(4-5):411-430, 2000.

[6] Tzyy-Ping Jung, et al. "Imaging brain dynamics using independent component analysis." *Proceedings of the IEEE*, Vol. 89, No. 7, July 2001, pp. 1107 -1122.

[7] T. Ristaniemi and J. Joutsensalo, "On the Performance of Blind Source Separation in CDMA Downlink", *Proc. International Workshop on Independent Component Analysis and Blind Separation of Signals*, Aussois, France, January 11-15, 1999