# A LEAST-SQUARES APPROACH TO JOINT SCHUR DECOMPOSITION

K. Abed-Meraim and Y. Hua

University of Melbourne, Elec. Eng. Dept., Parkville, Victoria 3052 Australia e-mail: abed, yhua@ee.mu.oz.au

#### ABSTRACT

We address the problem of joint Schur decomposition (JSD) of several matrices. This problem is of great importance for many signal processing applications such as sonar, biomedicine, and mobile communications. We first present a least-squares (LS) approach for computing the JSD. The LS approach is shown to coincide with that proposed intuitively by Haardt *et al*, thus establishing the optimality of their criterion in the least-squares sense. Following the LS criterion, we then propose new Jacobi-like algorithms that extend and improve the existing JSD algorithms. An application of the new JSD algorithms to multidimensional harmonic retrieval is also presented.

# 1. INTRODUCTION

Estimating the joint eigenstructure of several matrices is a problem that arises in many multivariate signal processing applications, e.g., joint diagonalization for source separation [3], joint eigendecomposition for parameter pairing [6], joint block-diagonalization for source localization [5], joint Schur decomposition (JSD) for multidimensional harmonic retrieval [1] and blind system identification [2], etc. A detailed review of various techniques for joint eigenstructure estimation is also available in [6]

In this paper, we will focus on the JSD problem [1]. We present a least-squares (LS) approach to joint Schur decomposition. Then, we introduce new iterative JSD algorithms for symmetrical and non-symmetrical cases. These algorithms are Jacobi-like techniques that minimize a squared error cost function iteratively by means of Givens rotations. An advantage of the Jacobi methods is their inherent parallelism, which allows efficient implementations on certain parallel architectures [7]. Another virtue of the Jacobi methods is their favorable rounding-error properties, in the sense that small relative perturbation in the matrices entries cause small relative perturbations in the entries of their engenstructures [8]. In addition to the favorable properties of the Jacobi-like techniques, the proposed algorithms have the advantage to be simpler, more general, and computationally less expensive than those presented in [1, 2].

# 2. PROBLEM FORMULATION

Consider a set of K matrices,  $\mathbf{M}_1, \dots, \mathbf{M}_K, \mathbf{M}_i \in \mathbb{C}^{n \times n}, i = 1, \dots, K$ , that have the following Schur decomposition:

$$\mathbf{M}_i = \mathbf{Q}\mathbf{R}_i\mathbf{Z}^*, \quad i = 1, \cdots, K \tag{1}$$

where  $\mathbf{Q}$  and  $\mathbf{Z}$  are unitary matrices and  $\mathbf{R}_i$ ,  $i = 1, \dots, K$ are upper triangular matrices.  $\mathbf{Z}^*$  denotes the transpose conjugate of  $\mathbf{Z}$ .

The matrices  $\mathbf{M}_i$ ,  $i = 1, \dots, K$  are said to be jointly Schur decomposable under the unitary transforms  $\mathbf{Q}$  and  $\mathbf{Z}$ .

In particular, the JSD problem is said to be *symmetrical* if  $\mathbf{Q} = \mathbf{Z}$ , i.e.,

$$\mathbf{M}_i = \mathbf{Q}\mathbf{R}_i\mathbf{Q}^*, \quad i = 1, \cdots, K$$

The problem of JSD consists in estimating the matrices  $\mathbf{Q}$ ,  $\mathbf{Z}$ , and  $\mathbf{R}_i$ ,  $i = 1, \dots, K$  given the matrices  $\mathbf{M}_i$ ,  $i = 1, \dots, K$ .

A possible approach to the JSD problem would be to apply a conventional Schur decomposition algorithm to  $\mathbf{M}_1$  alone, and then performs the same unitary transform (i.e.,  $\mathbf{Q}$  and  $\mathbf{Z}$ ) on  $\mathbf{M}_2, \dots, \mathbf{M}_K$  to compute  $\mathbf{R}_2, \dots, \mathbf{R}_K$ . Unfortunately, this approach is not consistent since  $\mathbf{M}_1$  may be made upper triangular by a unitary matrix that does not make  $\mathbf{M}_2, \dots, \mathbf{M}_K$  upper triangular.

On the other hand, the matrices  $\mathbf{M}_1, \dots, \mathbf{M}_K$  are given, in practice, by some sample estimated statistics that are corrupted by estimation errors due to noise and finite sample size effects. Thus, they are only "approximately" simultaneously Schur decomposable. This suggests that a viable JSD algorithm must provide a kind of an "average eigenstructure" when it is applied to a set of nearly joint Schur decomposable matrices. An optimal solution based on a least-squares approach is given next<sup>1</sup>.

#### 3. LEAST-SQUARES APPROACH

A least-squares fitting technique consists here in choosing the unitary matrices  $\mathbf{Q}$  and  $\mathbf{Z}$  and the upper triangular matrices  $\mathbf{R}_k$  that minimize the Frobenius norm of the difference between the data matrices  $\mathbf{M}_k$  and the true matrices given by (1)

$$\min_{\mathbf{Q},\mathbf{Z},\{\mathbf{R}_k\}} \sum_{k=1}^{K} \|\mathbf{M}_k - \mathbf{Q}\mathbf{R}_k\mathbf{Z}^*\|^2$$
(2)

To solve the minimization problem we first minimize with respect to the upper triangular matrices  $\{\mathbf{R}_k\}$ , with the matrices  $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_n]$  and  $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_n]$  held fixed.

 $<sup>^{1}</sup>A$  similar approach has been presented in [4] for joint diagonalization.

This latter problem is equivalent to the minimization of each term in (2), i.e.,

$$\min_{\{\mathbf{R}_k\}} \|\mathbf{M}_k - \mathbf{Q}\mathbf{R}_k\mathbf{Z}^*\|^2 \tag{3}$$

Using the vectorizing operator, vec(.) and removing the zero entries of  $\mathbf{R}_k$  (i.e., the strictly lower triangular entries), the minimization turns into

$$\min_{\mathbf{r}_{k}} \|\mathbf{m}_{k} - \mathbf{W}\mathbf{r}_{k}\|^{2} \tag{4}$$

where  $\mathbf{m}_k \stackrel{\text{def}}{=} \operatorname{vec}(\mathbf{M}_k)$  and

$$\mathbf{r}_{k} \stackrel{\text{def}}{=} [\mathbf{R}_{k}(1,1), \mathbf{R}_{k}(1,2), \mathbf{R}_{k}(2,2), \cdots, \mathbf{R}_{k}(1,n), \\ \cdots, \mathbf{R}_{k}(n,n)]^{T} : n(n+1)/2 \times 1$$
$$\mathbf{W} \stackrel{\text{def}}{=} [\overline{\mathbf{z}}_{1} \otimes \mathbf{q}_{1}, \overline{\mathbf{z}}_{2} \otimes \mathbf{q}_{1}, \overline{\mathbf{z}}_{2} \otimes \mathbf{q}_{2}, \cdots, \overline{\mathbf{z}}_{n} \otimes \mathbf{q}_{1}, \\ \cdots, \overline{\mathbf{z}}_{n} \otimes \mathbf{q}_{n}] : n^{2} \times n(n+1)/2$$

where  $\otimes$  denotes the Kronecker product and  $\overline{\mathbf{z}}$  is the conjugate of  $\mathbf{z}$ . The solution to (4) is given by

$$\mathbf{r}_k = \mathbf{W}^{\#} \mathbf{m}_k = (\mathbf{W}^* \mathbf{W})^{-1} \mathbf{W}^* \mathbf{m}_k.$$
 (5)

Since  ${\bf Q}$  and  ${\bf Z}$  are unitary, we can easily check that  ${\bf W}^*{\bf W}={\bf I},$  and thus

$$\mathbf{r}_k = \mathbf{W}^* \mathbf{m}_k \tag{6}$$

Substituting (6) into (2) yields the following estimation criterion:

$$\begin{split} \min_{\mathbf{Q},\mathbf{Z}} \sum_{k=1}^{K} \|(\mathbf{I} - \mathbf{W}\mathbf{W}^*)\mathbf{m}_k\|^2 & \Longleftrightarrow \quad \max_{\mathbf{Q},\mathbf{Z}} \sum_{k=1}^{K} \|\mathbf{W}^*\mathbf{m}_k\|^2 \\ & \longleftrightarrow \quad \max_{\mathbf{Q},\mathbf{Z}} \sum_{k=1}^{K} \mathcal{U}(\mathbf{Q}^*\mathbf{M}_k\mathbf{Z}) \end{split}$$

Using the fact that the Frobenius norm of a matrix is unchanged under unitary transforms, this is finally equivalent to minimizing under unitary transforms  $\mathbf{Q}$  and  $\mathbf{Z}$  the following nonnegative function:

$$\mathcal{C}(\mathbf{Q}, \mathbf{Z}) \stackrel{\text{def}}{=} \sum_{k=1}^{K} \mathcal{L}(\mathbf{Q}^* \mathbf{M}_k \mathbf{Z})$$
(7)

where for any matrix  $\mathbf{M}$ , we define

$$egin{aligned} \mathcal{U}(\mathbf{M}) \stackrel{ ext{def}}{=} \sum_{1 \leq i \leq j \leq n} |\mathbf{M}(i,j)|^2 \ \mathcal{L}(\mathbf{M}) \stackrel{ ext{def}}{=} \sum_{1 \leq j < i \leq n} |\mathbf{M}(i,j)|^2 \end{aligned}$$

# 4. JSD ALGORITHMS

To minimize the JSD criterion (7), we choose here to compute the unitary matrices  $\mathbf{Q}$  and  $\mathbf{Z}$  as products of Givens rotations that we describe next.

<u>Givens rotations</u>: In Jacobi-like algorithms, a unitary matrix  $\mathbf{U}$  is decomposed into a product of elementary Givens rotations, i.e.,

$$\mathbf{U} = \prod_{\# \text{ of sweeps}} \prod_{1 \le p < q \le n} \boldsymbol{\Theta}_{(qp)}$$

where the elementary Givens rotations  $\Theta_{(qp)}$  are defined as unitary matrices where all diagonal elements are 1 except for the two elements equal to c in rows (and columns) pand q. Likewise, all off-diagonal elements of  $\Theta_{(qp)}$  are 0 except for the two elements s and  $-s^*$  at (p,q) and (q,p)respectively. The scalar numbers c and s are given by

$$\begin{cases} c = \cos \theta \\ s = \sin \theta \exp(i\alpha) \end{cases}$$

In the sequel, we describe a procedure to choose the rotation angles  $\theta$  and  $\alpha$  at a particular iteration such that the cost function  $\mathcal{C}(\Theta_{(qp)}, \mathbf{I}), \mathcal{C}(\mathbf{I}, \Theta_{(qp)})$ , or  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$  is decreased to its minimum. To this end, we need to specify the orthogonal transformations

$$\mathbf{M}' = \mathbf{\Theta}^*_{(qp)}\mathbf{M} \tag{8}$$

$$\mathbf{M}'' = \mathbf{M}\mathbf{\Theta}_{(qp)} \tag{9}$$

$$\mathbf{M}^{\prime\prime\prime} = \mathbf{\Theta}^*_{(qp)} \mathbf{M} \mathbf{\Theta}_{(qp)} \tag{10}$$

for any given matrix  $\mathbf{M} \in \mathbb{C}^{n \times n}$ . First notice that these orthogonal transformations change only the rows p and q of  $\mathbf{M}$ , the columns p and q of  $\mathbf{M}$ , and pth and qth rows and pth and qth columns of  $\mathbf{M}$ , respectively. More specifically, the changed entries of  $\mathbf{M}'$ ,  $\mathbf{M}''$ , and  $\mathbf{M}'''$  in their strictly lower triangular part are given by:

$$\begin{split} \mathbf{M}'(p,j) &= c\mathbf{M}(p,j) - s\mathbf{M}(q,j), \quad j p \\ \mathbf{M}''(j,q) &= s\mathbf{M}(j,p) + c\mathbf{M}(j,q), \quad j > q \end{split}$$

and  $\mathbf{M}^{\prime\prime\prime}(k,j) = \mathbf{M}^{\prime}(k,j), k = p, q, j < k$  and  $\mathbf{M}^{\prime\prime\prime}(j,k) = \mathbf{M}^{\prime\prime}(j,k), k = p, q, j > k$  except for the (q,p)-th entry which is given by

$$\mathbf{M}^{\prime\prime\prime}(q,p) = c^{2}\mathbf{M}(q,p) - s^{*}c(M(q,q) - M(p,p)) - \mathbf{M}(p,q)s^{*2}$$
(12)

<u>Non-symmetrical JSD algorithm</u>: The proposed method consists of minimizing iteratively the JSD criterion (7) by successive Givens rotations, starting from  $\mathbf{Q} = \mathbf{I}, \mathbf{Z} = \mathbf{I}$ . At the end of the iterative procedure, we have

$$\mathbf{Q} = \prod_{\text{\# of iterations } 1 \le p < q \le n} \mathbf{\Theta}_{(qp)}$$
$$\mathbf{Z} = \prod_{\text{\# of iterations } 1 \le p < q \le n} \mathbf{\Theta}'_{(qp)}$$

where  $\Theta_{(qp)}$  and  $\Theta'_{(qp)}$  are computed such that  $\mathcal{C}(\Theta_{(qp)}, \mathbf{I})$ and  $\mathcal{C}(\mathbf{I}, \Theta'_{(qp)})$  are minimum<sup>2</sup>, respectively.

The non-symmetrical JSD algorithm can be summarized as follows (using informal notation):

$$\begin{aligned} \mathbf{Q} &= \mathbf{I}, \ \mathbf{Z} = \mathbf{I} \\ \text{for } k = 1, \dots, \# \text{ iterations} \\ \text{for } 1 &\leq p < q \leq n \\ \text{a. } \Theta_{(qp)} = \arg\min_{\theta,\alpha} \mathcal{C}(\Theta_{(qp)}(\theta, \alpha), \mathbf{I}) \\ \mathbf{Q} &:= \mathbf{Q}\Theta_{(qp)} \text{ and } \mathbf{M}_k := \Theta^*_{(qp)}\mathbf{M}_k, \ k = 1, \cdots, K \\ \text{b. } \Theta'_{(qp)} = \arg\min_{\theta,\alpha} \mathcal{C}(\mathbf{I}, \Theta_{(qp)}(\theta, \alpha)) \\ \mathbf{Z} &:= \mathbf{Z}\Theta'_{(qp)} \text{ and } \mathbf{M}_k := \mathbf{M}_k \Theta'_{(qp)}, \ k = 1, \cdots, K \end{aligned}$$

To minimize  $\mathcal{C}(\Theta_{(qp)}, \mathbf{I})$  or  $\mathcal{C}(\mathbf{I}, \Theta'_{(qp)})$ , we force their partial derivatives with respect to parameters  $\theta$  and  $\alpha$  to zero. That leads to the following explicit expressions (for simplicity, we omit here the calculation details):

$$\begin{aligned} \alpha &= \arctan(\frac{\Im m(a)}{\Re e(a)}) \\ \alpha' &= \arctan(\frac{\Im m(a')}{\Re e(a')}) \\ \theta &= \frac{1}{2}\arctan(-2\frac{\Re e(e^{-i\alpha}a)}{b}) \\ \theta' &= \frac{1}{2}\arctan(-2\frac{\Re e(e^{-i\alpha}a')}{b'}) \end{aligned}$$

where

$$a = \sum_{k} \sum_{p \le j < q} \mathbf{M}_{k}(p, j) \mathbf{M}_{k}^{*}(q, j)$$

$$a' = \sum_{k} \sum_{p < j \le q} \mathbf{M}_{k}(j, p)^{*} \mathbf{M}_{k}(j, q)$$

$$b = \sum_{k} \sum_{p \le j < q} |\mathbf{M}_{k}(p, j)|^{2} - |\mathbf{M}_{k}(q, j)|^{2}$$

$$b' = \sum_{k} \sum_{p < j \le q} |\mathbf{M}_{k}(j, p)|^{2} - |\mathbf{M}_{k}(j, q)|^{2}$$

Finally, note that in the real case  $\alpha = 0$  and only the last two equations are used.

Exact Symmetrical JSD algorithm: The symmetrical JSD algorithm has the same iterative structure as the non-symmetrical one, with the exception that the same Givens rotation  $\Theta_{(qp)}$  is applied on both left and right sides of  $\mathbf{M}_1, \dots, \mathbf{M}_K$  (i.e., we use the orthogonal transformation (10) instead of (8) and (9)). To estimate the angle parameters of  $\Theta_{(qp)}$  we have to minimize  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$ . Contrary to the previous situation, this minimization is more complex due to the non-linear term (in terms of c and s)  $\mathbf{M}^{\prime\prime\prime}(q, p)$ .

After some straightforward derivations (that we omit here for sake of simplicity), the minimization of  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$  is shown to be equivalent to the minimization of the quadratic form  $% \left( {{{\left[ {{{{\mathbf{n}}_{{\mathbf{n}}}}} \right]}_{{\mathbf{n}}}}} \right)$ 

$$\min_{\|\mathbf{v}\|=1} (\mathbf{v}^T \mathbf{G} \mathbf{v} + \mathbf{g}^T \mathbf{v}) \tag{13}$$

where  $\mathbf{v} = [\cos(2\theta), \sin(2\theta)\cos(\alpha), \sin(2\theta)\sin(\alpha)]^T$  and **G** (resp. **g**) is a 3 × 3 matrix (resp. a 3 × 1 vector) the expressions of which are given in the appendix.

Using Lagrange multiplier, the minimization of (13) leads to:

$$\mathbf{v} = (\mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{g}$$
(14)

where  $\lambda$  is a real scalar satisfying:

$$\sum_{i=1}^{3} \frac{|\mathbf{u}_i^T \mathbf{g}|^2}{(\lambda_i + \lambda)^2} = 1$$
(15)

 $\{\mathbf{u}_i\}$  and  $\{\lambda_i\}$  being the eigenvectors and eigenvalues of **G**. As we can see from (15), the exact minimization of  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$  involves a 6th order polynomial rooting (or a 4th order polynomial rooting in the real case). In case where (15) has multiple real-valued roots, we select the one corresponding to the minimum value of  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$ . Next, by means of a slight approximation of the JSD criterion, we present an alternative solution to (13) where no polynomial rooting is required.

Approximate Symmetrical JSD algorithm: To simplify the symmetrical JSD algorithm, we choose here to approximate  $|\mathbf{M}^{\prime\prime\prime}(q,p)|^2$  by  $|\mathbf{M}^{\prime\prime\prime}(q,p)|^2 \approx |\mathbf{M}^{\prime\prime}(q,p)|^2 + |\mathbf{M}^{\prime\prime}(q,p)|^2$ . This can be shown to be a first order approximation for  $|\mathbf{M}^{\prime\prime\prime}(q,p)|^2$ . With this approximation, we have

$$\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)}) \approx \mathcal{C}(\Theta_{(qp)}, \mathbf{I}) + \mathcal{C}(\mathbf{I}, \Theta_{(qp)})$$

and thus, the minimization of  $\mathcal{C}(\Theta_{(qp)}, \Theta_{(qp)})$  leads to the explicit expressions:

$$\alpha = \arctan(\frac{\Im m(a+a')}{\Re e(a+a')}), \ \theta = \frac{1}{2}\arctan(-2\frac{\Re e(e^{-i\alpha}(a+a'))}{b+b'})$$

### 5. APPLICATION TO MULTIDIMENSIONAL HARMONIC RETRIEVAL

We consider here the problem of eigenvalues estimation and association using the well known subspace rotation invariance technique [6, 9]. The data model is given by

$$\mathbf{JE} = [\mathbf{E}_0^T, \cdots, \mathbf{E}_{m-1}^T]^T \tag{16}$$

where for  $i = 0, \dots, m-1$ ,  $\mathbf{E}_i = \mathbf{A} \mathbf{\Phi}_i \mathbf{T}$ . The  $n \times d$  matrix  $\mathbf{E}$  has d independent columns and represents some processed data often available as an estimate of "signal subspace" which can be computed from, for example, the eigendecomposition of an array output covariance matrix. The  $ml \times n$  matrix  $\mathbf{J}$  is a selection matrix, each element of which is either zero or one. The  $l \times d$  matrix  $\mathbf{A}$  is unknown and of full column rank. The matrix  $\mathbf{\Phi}_i$  is a full rank diagonal (rotation) matrix. But  $\mathbf{\Phi}_0 = \mathbf{I}$ . The other  $\mathbf{\Phi}_i$ 's often contain the desired information such as frequencies, damping factors, directions and polarizations. The  $d \times d$  matrix  $\mathbf{T}$  is an arbitrary unknown nonsingular matrix. Let  $\mathbf{E}_0^{\#}$  denote the pseudo-inverse of  $\mathbf{E}_0$  and let

$$\mathbf{U}_i = \mathbf{E}_0^{\#} \mathbf{E}_i = \mathbf{T}^{-1} \mathbf{\Phi}_i \mathbf{T}, \quad i = 1, \cdots, m-1$$
(17)

<sup>&</sup>lt;sup>2</sup>Note that the criterion is calculated with the current values of the matrices  $\mathbf{M}_1, \dots, \mathbf{M}_K$  which are updated at each iteration using the orthogonal transformations (8) and (9). For notational simplicity, we keep using the generic notation  $\mathbf{M}_1, \dots, \mathbf{M}_K$  to denote the updated matrices.

The signal parameters (i.e., diagonal matrices  $\Phi_i$ ,  $i = 1, \dots, m-1$ ) can be estimated as the eigenvalues of  $\mathbf{U}_i, i = 1, \dots, m-1$ . This approach however requires a second step of parameter pairing since the eigenvalues are estimated up to a random permutation [6].

By decomposing  $\mathbf{T}^{-1}$  into its QR form, i.e.,  $\mathbf{T}^{-1} = \mathbf{QR}$ , we obtain for  $i = 1, \dots, m-1$ 

$$\mathbf{U}_i = \mathbf{Q} \mathbf{R}_i \mathbf{Q}^*, \quad \mathbf{R}_i = \mathbf{R} \mathbf{\Phi}_i \mathbf{R}^{-1}.$$

As we can see, we have  $\operatorname{diag}(\mathbf{R}_i) = \mathbf{\Phi}_i$ . Therefore, a convenient approach to simultaneously estimate and associate the eigenvalues consists in using a JSD of  $\mathbf{U}_1, \dots, \mathbf{U}_{m-1}$ .

Figure 1 presents a simulation example for 2-D (azimuth/elevation) source localization using the symmetrical JSD algorithm. We consider here a 2-D antennas array with 3 subarrays. The first subarray consists of n = 6 sensors lying uniformly along the x-axis with a sensor to sensor displacement equal to half the wavelength of the signal waves, i.e.,  $\delta = \omega_0/(2c)$ . The two other subarrays are located at  $(\alpha_1, \beta_1) = (\delta, 0)$  and  $(\alpha_2, \beta_2) = (0, \delta)$ , respectively. We assume d = 3 independent sources located at  $(20^\circ, -20^\circ)$ ,  $(30^\circ, 20^\circ)$  and  $(45^\circ, 20^\circ)$ , respectively corrupted by additive Gaussian noise. The sample size is T = 250. The plots show the MSE vs. the SNR (in dB) for azimuth and elevation estimates for the 3 sources. In this context, we observe similar performance accuracy for both exact (in solid line) and approximate (in dashed line) JSD algorithms.



Figure 1. MSE (dB) vs SNR (dB).

### 6. CONCLUSION

In this paper, we have introduced a LS approach for joint Schur decomposition. We have proposed new Jacobi-like algorithms that minimize the squared errors cost function iteratively by means of Givens rotations. The main advantages of the new algorithms are their inherent high parallelism, their robustness to noise and rounding errors, and their computational simplicity (i.e., they are computationally simpler than existing JSD algorithms [1, 2]). We have also presented an application example for multidimensional harmonic retrieval to illustrate the usefulness of the proposed method.

### APPENDIX

Using equation (11) and (12) and the equalities  $c^2 = (\cos(2\theta) + 1)/2$ ,  $|s|^2 = (1 - \cos(2\theta))/2$ , and cs =

 $\sin(2\theta)e^{i\alpha}/2$ , we obtain

$$\begin{aligned} |\mathbf{M}_{k}^{\prime\prime\prime}(q,p)|^{2} &= \mathbf{v}^{T}\mathbf{G}_{k}\mathbf{v} + \mathbf{g}_{k}^{T}\mathbf{v} + cte \\ |\mathbf{M}_{k}^{\prime}(p,j)|^{2} - cte &= -|\mathbf{M}_{k}^{\prime}(q,j)|^{2} + cte = \mathbf{x}_{kj}^{T}\mathbf{v} \\ |\mathbf{M}_{k}^{\prime\prime}(j,p)|^{2} - cte &= -|\mathbf{M}_{k}^{\prime\prime}(j,q)|^{2} + cte = \mathbf{y}_{kj}^{T}\mathbf{v} \end{aligned}$$

where *cte* denotes terms independent from  $(\theta, \alpha)$ . Thus

$$\mathbf{G} = \sum_{k=1}^{K} \mathbf{G}_{k}, \quad \mathbf{g} = \sum_{k=1}^{K} \left( \mathbf{g}_{k} + \sum_{p < j < q} \mathbf{x}_{kj} - \mathbf{y}_{kj} \right)$$

 $\mathbf{G}_k = \frac{1}{4}[g_{ij}]_{1 \le i,j \le 3}$  is a 3  $\times$  3 real symmetric matrix which entries are given by:

$$g_{11} = |\mathbf{M}_k(q,p)|^2 + |\mathbf{M}_k(p,q)|^2 - |\mathbf{M}_k(q,q) - \mathbf{M}_k(p,p)|^2$$

 $g_{12} + ig_{13} = (\mathbf{M}_k^*(q, p) + \mathbf{M}_k^*(p, q))(\mathbf{M}_k(p, p) - \mathbf{M}_k(q, q))$  $g_{22} + ig_{23} = -g_{33} + ig_{32} = -2\mathbf{M}_k^*(q, p)\mathbf{M}_k(p, q)$ 

 $\operatorname{and}$ 

$$\mathbf{g}_{k} = \frac{1}{2} \begin{bmatrix} |\mathbf{M}_{k}(q,p)|^{2} - |\mathbf{M}_{k}(p,q)|^{2} \\ \Re e((\mathbf{M}_{k}(q,p) - \mathbf{M}_{k}(p,q))(\mathbf{M}_{k}^{*}(p,p) - \mathbf{M}_{k}^{*}(q,q)) \\ -\Im m((\mathbf{M}_{k}(q,p) - \mathbf{M}_{k}(p,q))(\mathbf{M}_{k}^{*}(p,p) - \mathbf{M}_{k}^{*}(q,q)) \end{bmatrix}$$
$$\mathbf{x}_{kj} = \begin{bmatrix} |\mathbf{M}_{k}(q,j)|^{2} - |\mathbf{M}_{k}(p,j)|^{2} \\ \Re e(\mathbf{M}_{k}(p,j)\mathbf{M}_{k}^{*}(q,j) \\ \Im m(\mathbf{M}_{k}(p,j)\mathbf{M}_{k}^{*}(q,j)) \end{bmatrix}$$

$$\mathbf{y}_{kj} = \begin{bmatrix} |\mathbf{M}_k(j,q)|^2 - |\mathbf{M}_k(j,p)|^2 \\ \Re e(\mathbf{M}_k^*(j,p)\mathbf{M}_k(j,q) \\ \Im m(\mathbf{M}_k^*(j,p)\mathbf{M}_k(j,q) \end{bmatrix}$$

## REFERENCES

- M. Haardt, K. Hueper, J.B. Moore, and J.A. Nossek, "Simultaneous Schur decomposition of several matrices to achieve automatic pairing in multidimensional harmonic retrieval problem", *Proc. EUSIPCO*, pp. 531–534, 1996.
- [2] Alle-Jan van der Veen, A. Paulraj, "An analytical constant modulus algorithm", *IEEE-SP*, pp. 1136–1155, May 1996.
- [3] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, and E. Moulines, "Blind source separation using second order statistics," *IEEE-SP*, vol. 45, no. 2, pp. 434-444, Feb. 1997.
- [4] M. Wax and J. Scheinvald, "A least-squares approach to joint diagonalization", *IEEE-SP Letters*, vol. 4, pp. 52–53, Feb. 1997.
- [5] A. Belouchrani, A. Moeness, and K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices", *IEEE-SP Letters*, pp. 266–268, Sep. 1997.
- [6] Y. Hua and K. Abed-Meraim, "Techniques of eigenvalues estimation and association", *Digital Signal Processing*, vol. 7, no. 4, pp. 253-259, October 1997.
- [7] P.J. Eberlein, "On one-sided Jacobi methods for parallel computation", SIAM J. Algebraic Discrete Meth. vol. 8, pp. 790-796, 1987.
- [8] G. H. Golub and C. F. Van Loan, "Matrix computations", *The Johns Hopkins University Press*, 1989.
- [9] A. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple invariance ESPRIT", *IEEE-SP*, vol. 40, no. 4, pp. 867-881, April 1992.