Jacobi-Algorithm for Simultaneous Generalized Schur Decomposition in Higher-Order-Only ICA [1]

Lieven De Lathauwer, Joos Vandewalle and Bart De Moor[2]

Published in:

Proc. IEEE Benelux Signal Processing Chapter
Signal Processing Symposium (SPS'98),
Leuven, Belgium,
March 26-27, 1998, pp. 67-70.

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-98-03.ps.Z*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA/COSIC, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 18 05, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista.* E-mail: *Lieven.DeLathauwer@esat.kuleuven.ac.be*, *bart.DeMoor@esat.kuleuven.ac.be*, *Joos.Vandewalle@esat.kuleuven.ac.be*. This research was partially supported by the Flemish Government (the Concerted Research Action GOA-MIPS, the FWO (Fund for Scientific Research - Flanders) projects G.0292.95 and G.0262.97, the FWO Research Communities ICCoS (Identification and Control of Complex Systems) and Advanced Numerical Methods for Mathematical Modelling, the IWT Action Programme on Information Technology (ITA/GBO/T23) - IT-Generic Basic Research), the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs (the Interuniversity Poles of Attraction Programmes IUAP P4-02 and IUAP P4-24) and the European Commission (Human Capital and Mobility Network SIMONET, SC1-CT92-0779). Lieven De Lathauwer is a post-doctoral Research Assistant at the K.U.Leuven. B. De Moor is a senior Research Associate with the F.W.O. and an Associate Professor at the K.U.Leuven. Joos Vandewalle is a Full Professor at the K.U.Leuven. The scientific responsibility is assumed by the authors.

Abstract

In algebraic algorithms for (blind) source separation, (approximate) simultaneous matrix decompositions play an important role. In this paper, we study a Jacobi-iteration for the simultaneous triangularization of a set of matrices under unsymmetric orthogonal transformations of their row and column space.

Jacobi-approach to the Simultaneous Generalized Schur Decomposition

Lieven De Lathauwer, Bart De Moor and Joos Vandewalle

Department of Electrical Engineering ESAT
Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94, 3001 Heverlee - Belgium

`tel. +32/16321805 fax. +32/16321970`
`email : Lieven.DeLathauwer@esat.kuleuven.ac.be`

Abstract

In algebraic algorithms for (blind) source separation, (approximate) simultaneous matrix decompositions play an important role. In this paper, we study a Jacobi-iteration for the simultaneous triangularization of a set of matrices under unsymmetric orthogonal transformations of their row and column space.

1 INTRODUCTION

A *simultaneous generalized Schur decomposition* of a set of real $(R \times R)$-matrices $\{\mathbf{V}_k\}_{(1 \leqslant k \leqslant K)}$ corresponds to a system of equations of the type

$$\begin{aligned} \mathbf{Q} \cdot \mathbf{V}_1 \cdot \mathbf{Z} &= \mathbf{R}_1, & (1)\\ \mathbf{Q} \cdot \mathbf{V}_2 \cdot \mathbf{Z} &= \mathbf{R}_2, & (2)\\ &\vdots & \\ \mathbf{Q} \cdot \mathbf{V}_K \cdot \mathbf{Z} &= \mathbf{R}_K, & (3) \end{aligned}$$

in which $\mathbf{Q}, \mathbf{Z} \in \mathbb{R}^{R \times R}$ are orthogonal and $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K \in \mathbb{R}^{R \times R}$ are upper triangular (a set of two of the equations is called a "generalized Schur decomposition" [1]). Given a set of matrices $\{\mathbf{V}_k\}_{(1 \leqslant k \leqslant K)}$, it is usually impossible to find matrices $\mathbf{Q}$ and $\mathbf{Z}$, and a set $\{\mathbf{R}_k\}_{(1 \leqslant k \leqslant K)}$, such that Eqs. (1-3) are satisfied, but it is natural to solve them in a least-squares sense. The problem then consists of the determination of $\mathbf{Q}$ and $\mathbf{Z}$ such that $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K$ are "as upper triangular as possible"; the function g to be maximized, is given by:

$$g(\mathbf{Q}, \mathbf{Z}) = \sum_k \sum_{q \geqslant p} (\mathbf{Q} \cdot \mathbf{V}_k \cdot \mathbf{Z})^2_{pq} \qquad (4)$$

The decomposition is visualized in Fig. 1.
This simultaneous decomposition plays an important role in signal separation methods, based on the Constant Modulus property [2] or on the mutual statistical independence [3] of the sources. A symmetric version (a "simultaneous Schur decomposition", in which $\mathbf{Q} = \mathbf{Z}$ by definition) appeared in the context of 2-D DOA estimation [4].

For the actual computation of the simultaneous generalized Schur decomposition, an extended QZ-iteration is proposed in [2]. $\mathbf{Q}$ and $\mathbf{Z}$ are updated alternatingly such that, in each step, the optimal $\mathbf{Q}$ (given $\mathbf{Z}$, or vice-versa) is estimated as a product of matrices $\mathbf{H}_1 \mathbf{H}_2 \ldots \mathbf{H}_{R-1}$, that form the equivalent of Householder matrices for the computation of a simple QR-decomposition [1]. E.g. as far as $\mathbf{Q}$ is concerned, $\mathbf{H}_1$ maximally reduces (in least-squares sense) the norm of the below-diagonal part of the first columns of the instantaneous estimates of $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K$. After multiplication with $\mathbf{H}_1$, $\mathbf{H}_2$ minimizes the norm of the below-diagonal part of the second columns, without further affecting the first rows, and so on. In this paper, we investigate a Jacobi-type approach to the problem.

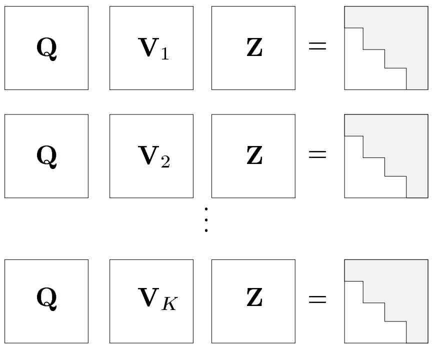

Figure 1: Visualization of a simultaneous generalized Schur decomposition.

2 DERIVATION

We consider an iterative optimization of the function g in Eq. (4), in which at each iteration step the estimate of $\mathbf{Q}$ and the estimate of $\mathbf{Z}$ are both updated by multiplication by a Givens rotation matrix; the purpose is to make the current estimates of $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K$ simultaneously as upper triangular as possible.
First, the estimates of $\mathbf{Q}$ and $\mathbf{Z}$ are initialized as

any orthogonal matrix, e.g. as the identity matrix: $\mathbf{Q}^{(0)} = \mathbf{Z}^{(0)} = \mathbf{I}$. The estimates of $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K$ are initialized accordingly: $\mathbf{R}_1^{(0)} = \mathbf{V}_1$, $\mathbf{R}_2^{(0)} = \mathbf{V}_2, \ldots, \mathbf{R}_K^{(0)} = \mathbf{V}_K$. (As an alternative, initial matrices $\mathbf{Q}^{(0)}$ and $\mathbf{Z}^{(0)}$ could be obtained from the generalized Schur decomposition corresponding to two of the equations in (1-3) [1].)

Assume that at iteration step $l+1$, the estimates $\mathbf{Q}^{(l)}$, $\mathbf{Z}^{(l)}$ and $\mathbf{R}_1^{(l)}, \mathbf{R}_2^{(l)}, \ldots, \mathbf{R}_K^{(l)}$ are available . Let $\mathbf{G}_{ij} \in \mathbb{R}^{R\times R}$ represent an elementary Givens rotation matrix that affects rows i and j, i.e. $\mathbf{G}_{ij}$ equals the identity matrix, except for the entries

$$\begin{aligned}(\mathbf{G}_{ij})_{ii} &= (\mathbf{G}_{ij})_{jj} = \cos\alpha,\\ (\mathbf{G}_{ij})_{ji} &= -(\mathbf{G}_{ij})_{ij} = \sin\alpha,\end{aligned}$$

in which α is the rotation angle (assume e.g. that $j > i$). An update of $\mathbf{Q}^{(l)}$ takes the form of $\mathbf{Q}^{(l+1)} = \mathbf{G}_{ij}\cdot\mathbf{Q}^{(l)}$. Similarly, an update of $\mathbf{Z}^{(l)}$ takes the form of $\mathbf{Z}^{(l+1)} = \mathbf{Z}^{(l)}\cdot{\mathbf{G}'_{ij}}^T$, where the Givens rotation matrix $\mathbf{G}'_{ij}$ is defined in the same way as $\mathbf{G}_{ij}$, in terms of an angle β. At the same time $\mathbf{R}_1^{(l)}, \mathbf{R}_2^{(l)}, \ldots, \mathbf{R}_K^{(l)}$ are updated as $\mathbf{R}_1^{(l+1)} = \mathbf{G}_{ij}\cdot\mathbf{R}_1^{(l)}\cdot{\mathbf{G}'_{ij}}^T$, $\mathbf{R}_2^{(l+1)} = \mathbf{G}_{ij}\cdot\mathbf{R}_2^{(l)}\cdot{\mathbf{G}'_{ij}}^T, \ldots, \mathbf{R}_K^{(l+1)} = \mathbf{G}_{ij}\cdot\mathbf{R_K}^{(l)}\cdot{\mathbf{G}'_{ij}}^T$.

The entries on the strictly lower diagonal part of $\mathbf{R}_k^{(l+1)}$ that are changed by Jacobi-rotation, are given by:

$$\begin{aligned}
(\mathbf{R}_k^{(l+1)})_{ir} &= \cos\alpha(\mathbf{R}_k^{(l)})_{ir} - \sin\alpha(\mathbf{R}_k^{(l)})_{jr},\\
&\qquad 1 \leqslant r < i, \quad (5)\\
(\mathbf{R}_k^{(l+1)})_{jr} &= \sin\alpha(\mathbf{R}_k^{(l)})_{ir} + \cos\alpha(\mathbf{R}_k^{(l)})_{jr},\\
&\qquad 1 \leqslant r < j,\ r \neq i, \quad (6)\\
(\mathbf{R}_k^{(l+1)})_{ji} &= \sin\alpha(\cos\beta(\mathbf{R}_k^{(l)})_{ii} - \sin\beta(\mathbf{R}_k^{(l)})_{ij})\\
&\quad + \cos\alpha(\cos\beta(\mathbf{R}_k^{(l)})_{ji} - \sin\beta(\mathbf{R}_k^{(l)})_{jj}), \quad (7)\\
(\mathbf{R}_k^{(l+1)})_{ri} &= \cos\beta(\mathbf{R}_k^{(l)})_{ri} - \sin\beta(\mathbf{R}_k^{(l)})_{rj},\\
&\qquad i < r \leqslant R,\ r \neq j, \quad (8)\\
(\mathbf{R}_k^{(l+1)})_{rj} &= \sin\beta(\mathbf{R}_k^{(l)})_{ri} + \cos\beta(\mathbf{R}_k^{(l)})_{rj},\\
&\qquad j < r \leqslant R. \quad (9)
\end{aligned}$$

Hence, at iteration step l, the maximization of the function g in Eq. (4) is equivalent to the minimization of

$$h(\alpha,\beta) = \sum_{k=1}^{K}[(\mathbf{R}_k^{(l+1)})_{ji}^2 + \sum_{r=i+1}^{j-1}((\mathbf{R}_k^{(l+1)})_{ri}^2 + (\mathbf{R}_k^{(l+1)})_{jr}^2)] \quad (10)$$

(the other entries do not affect the norm of the below diagonal parts).

Substitution of Eqs. (6-8) in Eq. (10) gives the following explicit expression for h:

$$h(\alpha,\beta) = \sum_{k=1}^{K}\sum_{n=1}^{5} h_{kn}(\alpha,\beta), \quad (11)$$

in which:

$$\begin{aligned}
h_{k1}(\alpha,\beta) &= \sin^2\alpha\,[\cos^2\beta\,(\mathbf{R}_k^{(l)})_{ii}^2 + \sin^2\beta\,(\mathbf{R}_k^{(l)})_{ij}^2\\
&\quad -2\sin\beta\cos\beta\,(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ij}], \quad (12)\\
h_{k2}(\alpha,\beta) &= 2\sin\alpha\cos\alpha\,\{\cos^2\beta\,(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ji}\\
&\quad + \sin^2\beta\,(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{jj} - \sin\beta\cos\beta\\
&\quad [(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{ji} + (\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{jj}]\}, \quad (13)\\
h_{k3}(\alpha,\beta) &= \cos^2\alpha\,[\cos^2\beta\,(\mathbf{R}_k^{(l)})_{ji}^2 + \sin^2\beta\,(\mathbf{R}_k^{(l)})_{jj}^2\\
&\quad -2\sin\beta\cos\beta\,(\mathbf{R}_k^{(l)})_{ji}(\mathbf{R}_k^{(l)})_{jj}], \quad (14)\\
h_{k4}(\alpha,\beta) &= (\sin^2\alpha + \cos^2\alpha)\times\\
&\quad \sum_{r=i+1}^{j-1}[\cos^2\beta\,(\mathbf{R}_k^{(l)})_{ri}^2 + \sin^2\beta\,(\mathbf{R}_k^{(l)})_{rj}^2\\
&\quad -2\sin\beta\cos\beta\,(\mathbf{R}_k^{(l)})_{ri}(\mathbf{R}_k^{(l)})_{rj}], \quad (15)\\
h_{k5}(\alpha,\beta) &= (\sin^2\beta + \cos^2\beta)\times\\
&\quad \sum_{r=i+1}^{j-1}[\cos^2\alpha\,(\mathbf{R}_k^{(l)})_{jr}^2 + \sin^2\alpha\,(\mathbf{R}_k^{(l)})_{ir}^2\\
&\quad +2\sin\alpha\cos\alpha\,(\mathbf{R}_k^{(l)})_{ir}(\mathbf{R}_k^{(l)})_{jr}]. \quad (16)
\end{aligned}$$

Setting the partial derivatives of h, with respect to α and β, equal to zero, leads to a set of biquadratic equations in $\tan\alpha$ and $\tan\beta$:

$$\begin{aligned}
b_1(\beta)\tan^2\alpha + b_2(\beta)\tan\alpha - b_1(\beta) &= 0, \quad (17)\\
b_3(\beta)\tan^2\alpha + b_4(\beta)\tan\alpha + b_5(\beta) &= 0, \quad (18)
\end{aligned}$$

in which $b_n(\beta) = \sum_{k=1}^{K} b_{kn}(\beta)$, with:

$$\begin{aligned}
b_{k1}(\beta) &= \tan^2\beta\,\{(\mathbf{R}_k^{(l)})_{ij}^2 - (\mathbf{R}_k^{(l)})_{jj}^2\\
&\quad + \sum_{r=i+1}^{j-1}[(\mathbf{R}_k^{(l)})_{ir}^2 - (\mathbf{R}_k^{(l)})_{jr}^2]\}\\
&\quad +2\tan\beta\,[(\mathbf{R}_k^{(l)})_{ji}(\mathbf{R}_k^{(l)})_{jj} - (\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ij}]\\
&\quad +\{(\mathbf{R}_k^{(l)})_{ii}^2 - (\mathbf{R}_k^{(l)})_{ji}^2\\
&\quad + \sum_{r=i+1}^{j-1}[(\mathbf{R}_k^{(l)})_{ir}^2 - (\mathbf{R}_k^{(l)})_{jr}^2]\}, \quad (19)\\
b_{k2}(\beta) &= \tan^2\beta\,[(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{jj}\\
&\quad + \sum_{r=i+1}^{j-1}(\mathbf{R}_k^{(l)})_{ir}(\mathbf{R}_k^{(l)})_{jr}]\\
&\quad - \tan\beta\,[(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{ji} + (\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{jj}]\\
&\quad +[(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ji} + \sum_{r=i+1}^{j-1}(\mathbf{R}_k^{(l)})_{ir}(\mathbf{R}_k^{(l)})_{jr}], \quad (20)\\
b_{k3}(\beta) &= (\tan^2\beta - 1)\,[(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ij}\\
&\quad + \sum_{r=i+1}^{j-1}(\mathbf{R}_k^{(l)})_{ri}(\mathbf{R}_k^{(l)})_{jr}]\\
&\quad + \tan\beta\,\{(\mathbf{R}_k^{(l)})_{ij}^2 - (\mathbf{R}_k^{(l)})_{ii}^2
\end{aligned}$$

$$+\sum_{r=i+1}^{j-1}[(\mathbf{R}_k^{(l)})_{rj}^2-(\mathbf{R}_k^{(l)})_{ri}^2]\}, \quad (21)$$

$$\begin{aligned} b_{k4}(\beta) &= (\tan^2\beta-1)\ \times \\ &[(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{ji}+(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{jj}] \\ &+2\tan\beta\,[(\mathbf{R}_k^{(l)})_{ij}(\mathbf{R}_k^{(l)})_{jj}-(\mathbf{R}_k^{(l)})_{ii}(\mathbf{R}_k^{(l)})_{ji}], \quad (22) \end{aligned}$$

$$\begin{aligned} b_{k5}(\beta) &= (\tan^2\beta-1)\ \times \\ &[(\mathbf{R}_k^{(l)})_{ji}(\mathbf{R}_k^{(l)})_{jj}+\sum_{r=i+1}^{j-1}(\mathbf{R}_k^{(l)})_{ri}(\mathbf{R}_k^{(l)})_{rj}] \\ &+\tan\beta\,\{(\mathbf{R}_k^{(l)})_{jj}^2-(\mathbf{R}_k^{(l)})_{ji}^2 \\ &+\sum_{r=i+1}^{j-1}[(\mathbf{R}_k^{(l)})_{rj}^2-(\mathbf{R}_k^{(l)})_{ri}^2]\}. \quad (23) \end{aligned}$$

The global minimum of $h(\alpha,\beta)$ can be determined by computing the various solutions of Eqs. (17-18), and selecting the one corresponding to the smallest value in Eq. (11).

For the solution of the set of biquadratic equations, let us first consider the special case where Eq. (17) is linear in $\tan\alpha$: $b_1(\beta)=0$. The only ways in which a root β_0 of b_1 can lead to a solution of Eqs. (17-18), are (a) $\tan\alpha=0$ and additionally $b_5(\beta_0)=0$, and (b) α is a solution of Eq. (18), for $\beta=\beta_0$, which additionally satisfies $b_2(\beta_0)=0$. Taking into account the definitions of $b_1(\beta)$, $b_2(\beta)$ and $b_5(\beta)$, neither case occurs, generically.

Now let us investigate the general case, i.e. $b_1(\beta)\neq 0$. Substitution of the square roots of Eq. (17) in Eq. (18) (considered as quadratic expressions in the unknown $\tan\alpha$) then leads, after tedious formula manipulations, to the following polynomial of degree 8 in $\tan\beta$:

$$\begin{aligned} &b_1^2(\beta)b_3^2(\beta)+b_1^2(\beta)b_5^2(\beta)-b_1(\beta)b_2(\beta)b_4(\beta)b_5(\beta) \\ &+2b_1^2(\beta)b_3(\beta)b_5(\beta)+b_2^2(\beta)b_3(\beta)b_5(\beta) \\ &-b_1^2(\beta)b_4^2(\beta)+b_1(\beta)b_2(\beta)b_3(\beta)b_4(\beta)=0 \quad (24) \end{aligned}$$

For the roots of this polynomial, the corresponding value of $\tan\alpha$ that gives a solution to Eqs. (17-18), can be found from:

$$\begin{aligned} &(b_2(\beta)b_3(\beta)-b_1(\beta)b_4(\beta))\tan\alpha \\ &-b_1(\beta)(b_3(\beta)+b_5(\beta))=0. \quad (25) \end{aligned}$$

Remark. If the simultaneous generalized Schur decomposition is not exactly satisfied, different orderings of the invariant subspaces may correspond to values of the function g in Eq. (4) that are somewhat different. There is no guarantee that the Jacobi-algorithm will converge to the solution with that specific ordering that leads to the global optimum.

Apart from this re-ordering, there is no formal evidence that the Jacobi-algorithm cannot get stuck in a local optimum; local or global convergence is still an open problem for the computation of other simultaneous matrix decompositions (e.g. [5]) as well. However, our simulations suggest that this does not form a major problem in practice.

3 SIMULATIONS

We have conducted a series of experiments to evaluate the performance of the technique w.r.t. an application in multiway factor analysis; for a detailed description of the simulations, we refer to [6]. Briefly, the simultaneous generalized Schur decomposition can be used to estimate the components of a simultaneous diagonalization by equivalence (congruence) [2, 3], namely

$$\begin{aligned} \mathbf{V}_1 &= \mathbf{U}^{(1)}\cdot\mathbf{D}_1\cdot\mathbf{U}^{(2)^T}, \quad (26) \\ \mathbf{V}_2 &= \mathbf{U}^{(1)}\cdot\mathbf{D}_2\cdot\mathbf{U}^{(2)^T}, \quad (27) \\ &\vdots \\ \mathbf{V}_K &= \mathbf{U}^{(1)}\cdot\mathbf{D}_K\cdot\mathbf{U}^{(2)^T}, \quad (28) \end{aligned}$$

in which $\mathbf{D}_1,\mathbf{D}_2,\ldots,\mathbf{D}_K$ are diagonal and $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$ are regular.

Here we assume 3 real-valued (3×3)-matrices; they are scaled such that $\|\mathbf{V}_1\|^2+\|\mathbf{V}_2\|^2+\|\mathbf{V}_3\|^2=1$; additive zero-mean Gaussian noise is added, having a standard deviation between $1e-3$ and 1. Define a matrix $\mathbf{U}^{(3)}$ by stacking the diagonals of $\mathbf{D}_1,\mathbf{D}_2,\ldots,\mathbf{D}_K$ as rows. We take the condition number of $\mathbf{U}^{(2)}$ and $\mathbf{U}^{(3)}$ equal to 3; for $\mathbf{U}^{(1)}$, three different sets of singular values will be considered: 3,2,1; 30,15,1; 100,50,1.

The following algorithms are applied in 50 Monte Carlo runs: (a) the Jacobi-algorithm, (b) the Jacobi-algorithm, with a restart procedure to check extensively for other invariant subspace orderings, (c) a least-squares matching of both sides of Eqs. (26-28), initialized with the result of (b), and (d) the extended QZ-iteration, referred to in the introduction.

To evaluate the accuracy of the different algorithms, we consider the quality of the estimate $\hat{\mathbf{U}}^{(1)}$ of $\mathbf{U}^{(1)}$; we assume that the columns of $\mathbf{U}^{(1)}$ have unit-length. In Fig. 2 the Root Mean Square Error (RMSE) between $\hat{\mathbf{U}}^{(1)}$ and $\mathbf{U}^{(1)}$ is plotted as a function of the noise level σ_N. For a given noise level and a given algorithm, the error measure has been computed as the average, over the different Monte Carlo simulations, of the Frobenius-norm $\|\hat{\mathbf{U}}^{(1)}-\mathbf{U}^{(1)}\|$; the ordering of the columns of $\hat{\mathbf{U}}^{(1)}$ that corresponds to the ordering of the columns of $\mathbf{U}^{(1)}$, has been determined as the ordering that minimizes the error.

The figure displays the expected performance degradation as the noise level and/or the condition number of $\mathbf{U}^{(1)}$ increases. In none of the simulations, the double-sided Jacobi-iteration technique ended up in a meaningless local optimum: the accuracy obtained

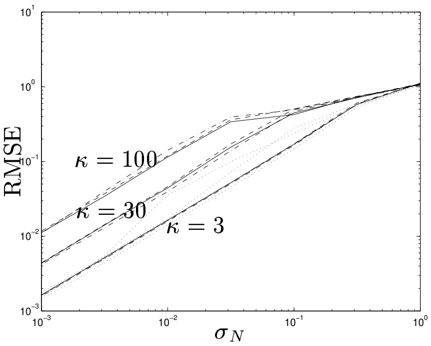

Figure 2: The RMSE for the estimation of $\mathbf{U}^{(1)}$, as a function of the noise level σ_N, for algorithms (a) (solid), (b) (dashed), (c) (dotted) and (d) (dashdot). The upper, middle and lower curves correspond to a condition number κ of $\mathbf{U}^{(1)}$, equal to 100, 30 and 3 respectively.

by algorithms (a) and (b) was always in the same order of magnitude.

In Fig. 3 we have plotted the mean value of $f = \sum_k \sum_{q<p} (\mathbf{Q} \cdot \mathbf{V}_k \cdot \mathbf{Z})^2_{pq}$. First, we notice that method (b) indeed shows a small marginal reduction of the overall Frobenius-norm below the diagonal, with respect to method (a). Next, the figure shows that the extended QZ-iteration does not minimize this norm (the reason for this behaviour is that the way in which $\mathbf{Q}$ and $\mathbf{Z}$ are computed, does not imply monotonic convergence in this sense: e.g. it is possible that the matrix $\mathbf{H}_1$ increases the Frobenius-norm of the part of columns 2 to $R-1$ below the diagonal); this effect is more outspoken as the condition number of $\mathbf{U}^{(1)}$ is larger. On the other hand, it is clear that, in the absence of noise, the theoretical solution is a stationary point of the extended QZ-algorithm; in Fig. 2 we see that the algorithm is still reliable in the presence of noise.

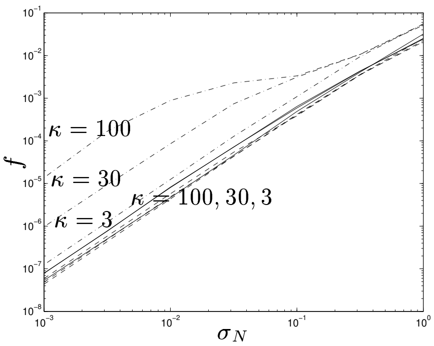

Figure 3: The mean value of f as a function of the noise level σ_N for algorithms (a) (solid), (b) (dashed), (c) (dotted) and (d) (dashdot).

4 CONCLUSION

In this paper, we have presented a Jacobi-algorithm for the computation of the simultaneous generalized Schur decomposition. Further research includes the issue of local vs global convergence, analysis of the convergence speed, and the computation of generalizations that appear in e.g. 2-D DOA estimation.

Acknowledgements

This research was partially supported by the Flemish Government (the Concerted Research Action GOA-MIPS, the FWO (Fund for Scientific Research - Flanders) projects G.0292.95 and G.0262.97, the FWO Research Communities ICCoS (Identification and Control of Complex Systems) and Advanced Numerical Methods for Mathematical Modelling, the IWT Action Programme on Information Technology (ITA/GBO/T23) - IT-Generic Basic Research), the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs (the Interuniversity Poles of Attraction Programmes IUAP P4-02 and IUAP P4-24) and the European Commission (Human Capital and Mobility Network SIMONET, SC1-CT92-0779). B. De Moor is a senior Research Associate with the F.W.O. The scientific responsibility is assumed by the authors.

References

[1] G.H. Golub and C.F. Van Loan. *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, 1991

[2] A.-J. van der Veen and A. Paulraj. *An Analytical Constant Modulus Algorithm*, IEEE Trans. Signal Processing, vol. 44, no. 5, pp. 1136-1155, May 1996

[3] L. De Lathauwer, B. De Moor and J. Vandewalle. *Independent Component Analysis Based on Higher-Order Statistics Only*, Proc. 8th IEEE SP Workshop on Statistical Signal and Array Processing (SSAP-96), Corfu, Greece, pp. 356-359, June 1996

[4] M. Haardt, K. Hüper, J. Moore and J. Nossek. *Simultaneous Schur Decomposition of Several Matrices to Achieve Automatic Pairing in Multidimensional Harmonic Retrieval Problems*, Signal Processing VIII: Theories and Applications, Proc. EUSIPCO-96, Trieste, Italy, vol. 1, pp. 531-534, Sept. 1996

[5] A. Bunse-Gerstner, R. Byers and V. Mehrmann. *Numerical Methods for Simultaneous Diagonalization*, SIAM J. Matrix Anal. Appl., vol. 14, no. 4, pp. 927-949, Oct. 1993

[6] L. De Lathauwer. *Signal Processing based on Multilinear Algebra*, Ph.D. Thesis, K.U.Leuven, E.E. Dept., Sept. 1997