Generalized Givens Rotations Applied to Complex Joint Eigenvalue Decomposition

Ammar MESLOUB

Abstract—This paper shows the different ways of using generalized Givens rotations in complex joint eigenvalue decomposition (JEVD) problem. It presents the different schemes of generalized Givens rotation, justifies the introduced approximations and focuses on the process of extending an algorithm developed for real JEVD to the complex JEVD. Several Joint Diagonalization problem use generalized Givens rotations to achieve the solution, many algorithms developed in the real case exist in the literature and are not generalized to the complex case. Hence, we show herein a simple and not trivial way to get the complex case from the real one. Simulation results are provided to highlight the effectiveness and behaviour of the proposed techniques for different scenarios.

Keywords—Complex Joint EigenValue Decomposition (JEVD), Shear and SHRT, generalized Givens rotations, exact JEVD, approximative JEVD.

I. INTRODUCTION

In this paper, we mainly try to expose the generalized Givens rotation used to solve the Joint EigenValue Decomposition (JEVD) problem of a set of complex square matrices. This problem of JEVD can be found in different applications such as multi-dimensional harmonic retrieval [1], Canonical Polyadic Decomposition (CPD) of tensors [2, 3], Direction of arrival estimation [4], joint angle-delay estimation [5] and Blind Sources Separation (BSS) [6].

The aforementioned problem is widely treated in the literature by using different schemes. The generalized Givens rotations have been used in [7, 8], the LU decomposition is used in [9] and first Taylor approximation is used in [10].

The JEVD is defined for $K$ complex square matrices of dimension $N \times N$ sharing the following structure:

$$M_k = AD_kA^{-1}$$  \hspace{1cm} (1)

Where $k \in \{1, ..., K\}$, $K$ is the number of matrices, $N$ is the matrix dimension. $A$ is unknown square non defective matrix (in the BSS context, this matrix is referred to as mixing matrix) and $D_k$ is the $k^{th}$ diagonal matrix associated to the $k^{th}$ matrix $M_k$.

The last problem consists of looking for $\{A, D_1, \ldots, D_K\}$ by using only the set of the $K$ complex matrices $\{M_1, \ldots, M_K\}$. It can be also defined by finding a matrix $V$ which makes the set of matrices $\{VM_1V^{-1}, \ldots, VM_KV^{-1}\}$ as diagonal as possible specially in the approximate JEVD case (see section V. for more details).

In this paper, we investigate generalized Givens rotations applied to the JEVD problem. The unknown matrix $A$ is decomposed in a product of generalized Givens rotations, according to:

$$A = \prod_{#sweeps} \prod_{1 \leq i < j \leq N} S_{ij}G_{ij}$$  \hspace{1cm} (2)

where $\#sweeps$ represents the number of iterations, $S_{ij}$ and $G_{ij}$ are the elementary Shear and Givens rotations respectively. The problem of JEVD reduces then in the estimation of these elementary rotations.

II. GENERALIZED GIVEN S ROTATION SCHEMES

The elementary rotations of equation (2) can be expressed according to two different schemes. The first one is based on sinus and hyperbolic sinus which needs some complicated developments leading to an efficient estimation.

This first scheme considers elementary rotations that are equal to the identity matrix except for some entries which are the elementary Shear and Givens rotations [7, 11], respectively. The problem of JEVD reduces then in the estimation of these elementary rotations.

$$\begin{bmatrix} G_{ij}(i,i) & G_{ij}(i,j) \\ G_{ij}(j,i) & G_{ij}(j,j) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & e^{-j\varphi} \sin(\theta) \\ -e^{j\varphi} \sin(\theta) & \cos(\theta) \end{bmatrix}$$  \hspace{1cm} (3)

$$\begin{bmatrix} S_{ij}(i,i) & S_{ij}(i,j) \\ S_{ij}(j,i) & S_{ij}(j,j) \end{bmatrix} = \begin{bmatrix} \cosh(y) & e^{j\alpha} \sinh(y) \\ e^{-j\alpha} \sinh(y) & \cosh(y) \end{bmatrix}$$  \hspace{1cm} (4)

$\theta$ is the angle parameter and $\varphi$ is the phase parameter of considered Givens rotation. $y$ is the shear angle and $\alpha$ represents the phase parameter of the elementary Shear rotation. This scheme is the most used one as it can be found in [7, 8, 11–14].

The second scheme allows appropriate approximations that lead to a computationally simplified methods compared to the first scheme. This scheme allows elementary Shear and Givens rotations equal to the identity matrix except for some entries which are given by:

$$\begin{bmatrix} G_{ij}(i,i) & G_{ij}(i,j) \\ G_{ij}(j,i) & G_{ij}(j,j) \end{bmatrix} = \lambda_2 \begin{bmatrix} 1 & g^* \\ -g & 1 \end{bmatrix}$$  \hspace{1cm} (5)

$$\begin{bmatrix} S_{ij}(i,i) & S_{ij}(i,j) \\ S_{ij}(j,i) & S_{ij}(j,j) \end{bmatrix} = \lambda_3 \begin{bmatrix} 1 & s^* \\ s & 1 \end{bmatrix}$$  \hspace{1cm} (6)

Manuscript received May 25, 2021; revised July 18, 2021.

A. MESLOUB is with the Signal Processing Laboratory, Ecole Militaire Polytechnique, Bordj El Bahri, ALGERIA (e-mail: mesloub.a@gmail.com).

Digital Object Identifier (DOI): 10.53907/enpesj.v1i1.49
where $g$ and $s$ are complex numbers associated to Givens and Shear parameters, respectively, $g^*$ is the complex conjugate of $g$. $\lambda_g = \frac{1}{\sqrt{1 + |g|^2}}$ and $\lambda_g = \frac{1}{\sqrt{1 - |s|^2}}$. When this scheme is applied as in [14, 15], both $g$ and $s$ are considered with small norm which simplifies $\lambda_g$ and $\lambda_s$ to the unity ($\lambda_g = \lambda_s \cong 1$).

III. JOINT EIGENVALUE DECOMPOSITION CRITERIA

The JEDV problem uses several criteria as shown in the literature, the difference of these criteria can be seen only in the shear rotation. Before discussing the existing criteria, let us note $M_k'$ the transformed matrices by the shear rotation as

$$M_k' = S_{ij}(y,\alpha)M_kS_{ij}(y,\alpha)$$

(7)

The different criteria can be discussed as

- In [12], the Shear rotation is chosen to minimize the Frobenius norm of transformed matrices $M_k'$, where the index $h$ is chosen by the matrix having the most departure from the normality (see [12] for more details). Then, some approximations and linearisation are introduced to simplify the explicit expression of the optimal value of $'y'$.

$$C_1(y) = ||M_k'||^2$$

(8)

This criterion $C_1(y)$ is used for real JEDV.

- In [7], the criterion of shear transformation takes into account the Frobenius norm of off diagonal matrices, which is the Frobenius norm of matrices $M_k - \text{diag}(M_k)$, $k = 1, \cdots, K$. The Shear parameter $'y'$ is obtained by an analytical solution.

$$C_2(y,\alpha) = \sum_{k=1}^{K} ||M_k' - \text{diag}(M_k')||^2$$

(9)

This criterion $C_2(y,\alpha)$ is used in the complex case.

- In [8], the considered criterion is the sum of square norm of the $(i,j)$th and $(j,i)$th entries of matrices $M_k$, $k = 1, \cdots, K$. The fixed couple $i$ and $j$ satisfies $1 \leq i < j \leq N$. This criterion is a simplified version of the one given in equation (9). However, JDTM (Joint Diagonalization algorithm based on Targeting hyperbolic Matrices) outperforms JUST (Joint Unitary Shear Transformation) in many scenarios (see [8] for details).

$$C_3(y) = \sum_{k=1}^{K} |m_k'(i,j)|^2 + |m_k'(j,i)|^2$$

(10)

- In [13], the considered criterion is the sum of square norm of difference between the $(i,j)$th and $(j,i)$th entries of matrices $M_k$, $k = 1, \cdots, K$. This criterion is introduced to reduce the deviation of matrices $M_k$ from normality. The Shear rotation is explicitly introduced to minimize the deviation from symmetry of matrices $M_k$. Hence, for each elementary rotation $S_{ij}(y)$, the minimized criterion can be expressed as:

$$C_4(y) = \sum_{k=1}^{K} \sum_{1 \leq i < j \leq N} |m_k'(i,j) - m_k'(j,i)|^2$$

(11)

Now, let us generalize the SHRT (SHear RoTation algorithm) method developed in [12] to the complex case.

IV. COMPLEX SHRT METHOD

The objective of this section is to generalize the method proposed in [12] to the complex case. Hence, only the Shear rotation is studied where the Givens rotation is obtained by applying the solution given in [6].

Let us study the transformed matrices by the Shear transform given in (7), where only $i^{th}$, $j^{th}$ rows and columns are affected by the Shear transformation. Modified entries can be written as:

$$m_k'(i,j) = m_k(i,j) \cosh(y) + m_k(j,i) e^{-y} \sinh(y)$$

$$m_k'(j,i) = m_k(i,j) e^{y} \sinh(y) + m_k(j,i) \cosh(y)$$

(12)

and the entries twice affected by the shear rotation can be expressed as:

$$m_k'(i,j) e^{-y} = m_k(i,j) e^{-y} + \xi_k \sinh(2y) + \frac{d_k}{2} \sinh(2y)$$

$$m_k'(j,i) e^{y} = m_k(i,j) e^{y} - \xi_k \sinh(2y) - \frac{d_k}{2} \sinh(2y)$$

(13)

Let us analyse the different parts of $C_{T}(y,\alpha)$ given equation (16), the part $\frac{|m_k'(i,j) + m_k'(j,i)|^2}{2}$ is independent from the shear parameter $y$ and $\alpha$, the part $\frac{|m_k'(i,j) e^{-y} + m_k'(j,i) e^{y}|^2}{2}$ depends only on the phase parameter $\alpha$ and the other parts depend on the two parameters. Then, we have to minimize the part $\frac{|m_k'(i,j) e^{-y} + m_k'(j,i) e^{y}|^2}{2}$ to get optimal phase parameter. The optimal value is

$$\alpha = \frac{1}{2} \arg\{m_k(i,j) m_k'(j,i)\} - \pi$$

(18)

Once this value is computed, we insert it in equation (15) in order to get the shear rotation parameter $y$. After some workouts,
We differentiate equation (19) with respect to $y$ where $V$

The Shear transformation is achieved by computing the shear 

The proposed CSHRT algorithm can be summarized as follows.

Using the solution given in [6], the unitary transformation can 

is achieved by successive unitary and Shear transformations.

A scheme is applied to get the matrix $V$ which has the maximum 

Finally, the developed algorithm, referred to as Complex SHRT (CSHRT), is given in Table I.

| Table I |
| CSHRT ALGORITHM |

| Require | $M_{k}$, $k = 1, \ldots, K$, fixed threshold $\tau$ and maximum sweep $M_{it}$. |
| Initialization | $V = I_{N}$ and $A = I_{N}$. |
| while | max$_{i,j}$ $|y_{i,j}| > \tau$ and (#sweeps $< M_{it}$) |
| for all $1 \leq i < j \leq N$ | Unitary transform |
| Estimate $G_{ij}(\theta, \varphi)$ using solution given in [6]. |
| Updates matrices $V, M_{k}$ using equation (22). |
| Shear transform | Estimate the matrix $M_{h}$ which has the maximum departure to the normality. |
| Compute the phase parameter $\alpha$ using (18). |
| Compute Shear parameter $y$ using equation (21). |
| Matrices updates using equation (23) |
| end for |
| end while. |

The proposed CSHRT algorithm can be summarized as follows. Given the complex matrices $M_{k}, k \in \{1, \ldots, K\}$, an iterative scheme is applied to get the matrix $V$ which is the left inverse of the mixing matrix $A$ given in equation (1). One iteration is achieved by successive unitary and Shear transformations. Using the solution given in [6], the unitary transformation can be applied to get $G(\theta, \varphi)$. The last transformation is used to update the matrices according to

$$
\left\{ \begin{array}{l}
V \leftarrow VG(\theta, \varphi) \\
M_{k} \leftarrow G(\theta, \varphi)^{H}M_{k}G(\theta, \varphi)
\end{array} \right.
$$

where $V$ is the estimated left inverse of the mixing matrix $A$. The Shear transformation is achieved by computing the shear parameters $\alpha$ and $y$. The phase parameter $\alpha$ is obtained by using equation (18) and the angle parameter $y$ is computed by equation (21). Once these parameters are obtained, the first scheme of shear rotation is applied to update the different matrices according:

$$
\begin{align*}
\begin{cases}
V & \leftarrow VS(y, \alpha) \\
M_{k} & \leftarrow S(-y, \alpha)M_{k}S(y, \alpha)
\end{cases}
\end{align*}
$$

The overall proposed algorithm, named complex SHRT (CSHRT), is summarized in Table I. Note that the last algorithm is the first version of CSHRT, other modifications can be introduced to get the other versions. The second version can be developed by taking into account only the entries twice affected given in equation (13) which leads to minimize the criterion $C_{T_{2}}(y, \alpha)$ given in equation (16). The phase parameter $\alpha$ is the same as in (18) but the shear angle $y$ is obtained by:

$$
y = \text{arctanh} \left( \frac{\Re(K_{h}e^{\alpha}) - \xi_{h}^{*}d_{k}}{2(|d_{h}|^{2} + |\xi_{h}|^{2})} + G_{h} \right)
$$

The third version is realized by differentiating equation (19) with respect to $y, \alpha$ and making a linear approximations to the zeros of obtained derivatives. The formula of $y$ is kept unchanged as in (21) but the phase parameter is:

$$
\alpha = \arg \left( \sum_{i=1}^{N} m_{h}(i, l)m_{h}^{*}(j, l) - m_{h}(l, i)m_{h}^{*}(j, l) \right) - \pi
$$

CSHRT$_{1}$ is the version given in Table I, the second version noted CSHRT$_{2}$ is the same as CSHRT$_{1}$ except for $y$ which is computed by (24) instead of (21). The last version CSHRT$_{3}$ is also the same as CSHRT$_{1}$ except for the phase parameter which is obtained by equation (25) instead of equation (18).

V. SIMULATION, RESULTS AND DISCUSSIONS

In this section, we have tested the different proposed algorithm versions and compared with respect to JDTM and SID (Simple Joint Diagonalization algorithm) given in [8, 14] respectively for different cases.

In the first case, the target matrices satisfy exactly the equation (1), this case is named the exact JEVD case. The aim of this part is to show the convergence rate of each algorithm. In the second case, the target matrices did not satisfy the equation (1) which leads to the case of approximative JEVD. The objective of this case is to get the noise robustness of each algorithm.

The used Performance Index (PI) is the same as in [7, 16] evaluated over 100 Monte-Carlo realizations. The definition of PI is

$$
\text{PI}(Q) = \frac{1}{2N(N-1)} \sum_{n=1}^{N} \left( \sum_{m=1}^{N} \frac{|q(n, m)|^{2}}{\max_{k} |q(n, k)|^{2}} - 1 \right) + \frac{1}{2N(N-1)} \sum_{n=1}^{N} \left( \sum_{m=1}^{N} \frac{|q(m, n)|^{2}}{\max_{k} |q(k, n)|^{2}} - 1 \right)
$$

where $Q = VA$ is the global matrix. For a tested algorithm, when the obtained PI value is close to zero, it means that the JEVD quality reached is good.
A. Exact JEVD case

In this case, all matrices $A$ and $\{D_k\}_{k=1,...,K}$ are generated by considering normal distribution, independent and complex for all entries. Hence, the target matrices $\{M_k\}_{k=1,...,K}$ are obtained by using equation (1). All considered algorithms are applied to these matrices in order to realize the JEVD.

We have realized the JEVD of five matrices ($K = 5$) and varying the matrix dimension $N$. $N$ takes these values $\{5, 10, 25, 35\}$. The Figure 1 shows obtained results. Note that as the number of matrices $K$ is less important compared with the matrices dimension $N$, as the JEVD difficulty increases. However, the considered algorithms are differently affected by this difficulty. JDTM is the less affected one and presents the best convergence rate.

The SJD algorithm, as shown in Figures 1(c) and 1(d), diverges completely when the ratio $\frac{K}{N}$ is less than 20% due to the introduced approximations considered by this algorithm as explained in [14]. Note that this algorithm uses the second scheme of generalized Givens rotations by considering small changes ($\alpha_x = \alpha_y \approx 1$). It minimizes the simplified criterion given (9) which introduces this divergence for difficult JEVD.

The performance of the different proposed version of CSHRT are in-between those of SJD and JDTM. This is due to using only one target matrix to estimate shear parameters instead of considering all target matrices. Note that the second version is the most efficient when the CSHRT$_3$ is the worst version which implies that the phase parameter $\alpha$ is better estimated with equation (18) instead of formula (25). By analyzing curves of figure 1, CSHRT$_1$ and CSHRT$_2$ have quadratic convergence where CSHRT$_3$ has a linear convergence.

The second comparison can be done between CSHRT$_1$ and CSHRT$_2$. The difference between these versions is the computation of the shear angle $\gamma$. The first version CSHRT$_1$ minimizes the total criteria given in (15) where the second version CSHRT$_2$ minimizes the simplified criterion given (16). Surprisingly, the second version is better than the first version, it can be seen in curves of figures 1(c) and 1(d).

B. Approximative JEVD case

In this case, the target matrices are generated using the equation given in (27) as:

$$M_k = AD_kA^{-1} + \Xi_k \quad (27)$$

Fig. 1: Mean PI versus sweep number in exact JEVD for different matrix dimensions

Fig. 2: Median PI versus perturbation level in approximate JEVD ($K = N = 5$)

where $\Xi_k$ is a noise matrix. The inverse perturbation level (IPL) can be measured by:

$$PL(dB) = \frac{||AD_kA^{-1}||_F}{||\Xi_k||_F} \quad (28)$$

The noise matrix $\Xi_k$ is generated as $\Xi_k = \varrho_k \Upsilon_k$ where $\Upsilon_k$ is a random matrix (generated at each Monte Carlo run) and $\varrho_k$ is a positive number allowing to tune the perturbation level, it can be seen as kind of matrix SNR.

Algorithms’ performance are evaluated according to the IPL for $K = 5$ and $N = 5$. Obtained results are given in Figure 2 where PI versus IPL curves are shown. Note that all algorithms reach approximatively the same performance as it can be seen in Figures 2 except for CSHRT$_3$ where the phase is less estimated. The difference can be seen in the convergence rate as illustrated in subsection A.

VI. Conclusion

In this paper, the JEVD problem using generalized Givens rotation is considered in the complex case. We have seen the different JEVD criteria and rotation schemes. We have generalized SHRT algorithm to the complex case by considering three versions. Surprisingly, the second version is the most efficient compared to the other ones. By considering the results presented in the simulation section, one can deduce that minimizing the simplified criteria leads to most efficient estimation of shear
parameters (α and y). The presented algorithm is less efficient compared to JDTM, this is due to the introduced simplifications and approximations. These simplifications can be listed as: only one target matrix is considered for shear rotation and the shear parameter y is obtained by using some linear approximations. However, the proposed algorithm seem to be efficient with relatively low computational cost.

REFERENCES


Ammar MESLOUB received the Engineer’s and Master’s degrees in telecommunications from EMP, Algiers, Algeria, in 2006 and 2009 respectively. He received a Ph.D. in signal processing from EMP. He is an associate professor in the signal processing laboratory, EMP. His research interests are in signal processing, time-frequency analysis, and time-frequency array signal processing with applications in radar and communications.