Shortening of Paraunitary Matrices Obtained by Polynomial Eigenvalue Decomposition Algorithms

Jamie Corr*, Keith Thompson*, Stephan Weiss*, Ian K. Proudler†, and John G. McWhirter‡
*Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow, Scotland
†School of Electrical, Electronics and Systems Engineering, Loughborough University, Loughborough, UK
‡School of Engineering, Cardiff University, Cardiff, Wales, UK
{jamie.corr,keith.thompson,stephan.weiss} @strath.ac.uk

Abstract—This paper extends the analysis of the recently introduced row-shift corrected truncation method for paraunitary matrices to those produced by the state-of-the-art sequential matrix diagonalisation (SMD) family of polynomial eigenvalue decomposition (PEVD) algorithms. The row-shift corrected truncation method utilises the ambiguity in the paraunitary matrices to reduce their order. The results presented in this paper compare the effect a simple change in PEVD method can have on the performance of the paraunitary truncation. In the case of the SMD algorithm the benefits of the new approach are reduced compared to what has been seen before however there is still a reduction in both reconstruction error and paraunitary matrix order.

I. INTRODUCTION

Broadband array processing problems are often formulated with a space-time covariance matrix $R\{\tau\}$, which in addition to a spatial components contains an explicit lag $\tau$. Its $z$-transform, $R(z)\mapsto R[\tau]$, yields a polynomial cross-spectral density (CSD) matrix, which extends the symmetric or Hermitian property known from standard matrix algebra to the parahermitian case as $R(z) = R^{\dagger}(z)$, whereby the parahermitian operator $\{\cdot\}$ consists of a Hermitian transposition $\{\cdot\}^H$ and time reversal i.e. $R(z) = R^H(z^{-1})$.

Since the eigenvalue decomposition (EVD) of a covariance matrix provides an optimal factorisation for numerous signal processing problems, its extension from the narrowband to the broadband case has led to polynomial EVD (PEVD) as proposed in [9]. For a parahermitian matrix $R(z)$,

$$R(z) \approx Q(z)D(z)Q(z)^*$$

factorises $R(z)$ into a paraunitary $Q(z)$, such that $Q(z)Q(z)^* = I$, and a diagonal parahermitian $D(z)$,

$$D(z) = \text{diag}\{D_0(z) \ D_1(z) \ldots D_{M-1}(z)\}.$$  

Extension the idea of an ordered EVD [6], the polynomial eigenvalues in $D(z)$ are spectrally majorised, i.e. their power spectral densities $D_m(e^{j\Omega}) = D_m(z)|_{z=e^{j\Omega}}$ satisfy

$$D_{m+1}(e^{j\Omega}) \geq D_m(e^{j\Omega}) \forall \Omega \ m = 0 \ldots (M-1).$$

For FIR paraunitary matrices, the equality in (1) is not guaranteed [9] but likely valid in close approximation for high orders of $Q(z)$ [7].

The PEVD enables a number of applications spanning from filter bank-based channel coding [17], to the design of broadband precoding and equalisation of MIMO systems [12], subband coding [10], broadband angle of arrival estimation [1], and others. A number of these applications are directly influenced by the order of the paraunitary matrix $Q(z)$, such as polynomial subspace decomposition techniques [1], [12], [17], and paraunitary matrices of low order therefore can be crucial.

A number of iterative algorithms have been developed to approximate (1). For example, an approximate PEVD (APEVD) algorithm with fixed order has been reported in [14], but has not been proved to converge. Other algorithms have been proven to converge towards a diagonalised $D(z)$ and can achieve better diagonalisation than APEVD, including the family of second order sequential best rotation (SBR2) algorithms [9], [10] and the family of sequential matrix diagonalisation (SMD) algorithms [2], [11]. Although guaranteed to diagonalise $R(z)$, the SBR2 and SMD algorithms [2], [9]–[11] are unconstrained in their order and therefore the polynomial degrees of both $D(z)$ and $Q(z)$ grow with the number of iterations.

The growing order of the diagonal parahermitian matrix causes difficulties, as its increase is responsible for the rising complexity of iterative PEVD algorithms such as [2], [9]–[11] as iterations go on. To this end, trimming small coefficients at the ends of this matrix has been suggested in [5], [9] in order to reduce the complexity of iterative PEVD algorithm. However, the paraunitary matrix also grows with the number of iterations; while its increase does not impact on the complexity during iterations, the application cost of the finally extracted paraunitary matrix can be high for polynomial subspace-based applications as mentioned above. Therefore, reducing the order of the paraunitary matrix has been suggested in [13], whereby similarly to [5], [9] small outer matrix coefficients are truncated.

In addition to the trimming approach in [13], an enhanced alternative method has been suggested in [3] and applied to SBR2. The aim of this paper therefore is to investigate the so-call shift-corrected version in [3] both SBR2 and SMD families of iterative PEVD algorithms. To accomplish this, Sec. II reviews the PEVD algorithms that will be used to generate paraunitary unitary matrices. Sec. III provides and overview over the two paraunitary truncation approaches to be compared. The results from applying the different truncation methods to the paraunitary matrices produced by the two PEVD
methods are presented in Sec. IV and conclusions are given in Sec. V.

II. PEVD ALGORITHMS

A. General Approach

Starting from the parahermitian matrix $R(z)$, all PEVD algorithms apply a series of elementary paraunitary matrices in order to iterate towards an approximately diagonal polynomial matrix $D(z)$. At the $i$-th iteration, building on a partially diagonalised parahermitian matrix $S^{(i-1)}(z)$, PEVD algorithms build an elementary paraunitary matrix $Q^{(i)}(z)$ from two components: a shift matrix $\Lambda^{(i)}(z)$ which transfers large off-diagonal elements of $S^{(i-1)}(z)$ onto the zero lag, followed by a rotation $Q^{(i)}$ which moves the transferred elements’ energy onto the diagonal. Thus, for the $i$-th iteration we have

$$Q^{(i)}(z) = Q^{(i)}(z) \Lambda^{(i)}(z).$$  

(4)

While various PEVD algorithms share this same general approach, both $\Lambda^{(i)}(z)$ and $Q^{(i)}$ are algorithm dependent. The delay matrix, $\Lambda^{(i)}(z)$, is determined by the search strategy, while the rotation $Q^{(i)}$ is defined by the family of algorithms used. The $i$-th iteration of any iterative PEVD algorithm is then implemented as

$$S^{(i)}(z) = Q^{(i)}(z) S^{(i-1)}(z) \hat{Q}^{(i)}.$$  

(5)

The algorithm is stopped after $I$ iterations if either the off-diagonal energy of $S^{(i)}(z)$ falls below a predefined threshold or $I$ reaches a selected limit. Thereafter $D(z) = S^{(I)}(z)$ and

$$\hat{Q}(z) = \prod_{i=1}^{I} Q^{(i)}(z) \Lambda^{(i)}(z).$$  

(6)

Generally, (6) can be calculated after the algorithm has been executed, as storing the parameters of $Q^{(i)}$ and $\Lambda^{(i)}(z)$ is more efficient than storing and updating the whole paraunitary matrix. Based on the PEVD outline above the following subsections will go into the unique details for both PEVD algorithms utilised in the results section.

B. Second Order Sequential Best Rotation

The second order sequential best rotation (SBR2) algorithm is an extension of the classical Jacobi algorithm for scalar matrices to the polynomial case [9]. Like the Jacobi algorithm, SBR2 starts with a search for the maximum off diagonal element but this now extends to all lags of the polynomial matrix. Starting from $S^{(0)}(z) = R(z)$, at the $i$-th iteration the maximum off diagonal element is found using

$$\{k^{(i)}, \tau^{(i)}\} = \arg \max_{k, \tau} \|s^{(i-1)}(z)\|\infty, \ i = 1 \ldots I.$$  

(7)

With the transform pair $s^{(i-1)}[\tau] \longrightarrow S^{(i-1)}(z)$, the modified $k^{(i)}$th column vector, $s^{(i-1)}[\tau]$, contains only off diagonal elements. The delay matrix $\Lambda^{(i)}(z)$ is then constructed using

$$\Lambda^{(i)}(z) = \text{diag}\{1, \ldots, 1, z^{-\tau^{(i)}}, 1, \ldots, 1\}$$  

(8)

where the parameters $k^{(i)}$ and $\tau$ are used to advance or delay the $k^{(i)}$th column and row, shifting the maximum element by $\tau^{(i)}$ lags onto the zerolag. Finally, the energy from the maximum element is transferred onto the diagonal using the Jacobi transformation

$$Q^{(i)} = \begin{bmatrix} I_1 & \cos \varphi^{(i)} & \ldots & e^{j\varphi^{(i)}} \sin \varphi^{(i)} \\ \vdots & I_2 & \vdots & \vdots \\ -e^{-j\varphi^{(i)}} \sin \varphi^{(i)} & \ldots & \cos \varphi^{(i)} \end{bmatrix},$$  

(9)

where the rotation angles $\varphi^{(i)}$ and $\vartheta^{(i)}$ are determined by the value of the maximum element. The dimensions of the $\Lambda^{(i)}(z)$ are determined by the parahermitian matrix, its sparse nature means that only two rows and columns are modified when proceeding from $S^{(i-1)}(z)$ to $S^{(i)}(z)$.

C. Sequential Matrix Diagonalisation

In addition to the three main PEVD steps described above the sequential matrix diagonalisation (SMD) [11] algorithm includes an additional initialisation step that brings all off diagonal zerolag energy onto the diagonal prior to any shift operations. This results in

$$S^{(0)}(z) = Q^{(0)}(z) R(z) Q^{(0)H}.$$  

(10)

where $Q^{(0)}$ is the modal matrix for the EVD of the zerolag of $R(z)$.

The search step of the SMD is also different in that the $l_{\infty}$ norm in (7) is replaced by an $l_2$ norm, thereby changing the search from maximum element to maximum column norm

$$\{k^{(i)}, \tau^{(i)}\} = \arg \max_{k, \tau} \|s^{(i-1)}[\tau]\|2.$$  

(11)

Using the same modified column vector, $s^{(i-1)}[\tau]$, the parameters $k^{(i)}$ and $\tau^{(i)}$ are again used in (8). For the SMD algorithm, an energy transfer matrix $Q^{(i)}$ that clears all zero lag off-diagonal elements can be found by an EVD at zero lag.

In general, the SMD algorithm transfers more energy per iteration than the SBR2 algorithm. Due to the full EVD being non-sparse and requiring a full matrix multiplication for each lag in the parahermitian matrix, each SMD iteration is more computationally costly than a similar SBR2 iteration. Therefore, SMD overall has a higher complexity than SBR2 to compute an approximate PEVD, but is capable of producing paraunitary matrices of lower order [11], making operations involving $Q(z)$ less costly to apply once the PEVD is calculated.

III. PARAUNITARY MATRIX TRUNCATION METHODS

To reduce the cost of applying the paraunitary matrix $Q(z)$, two different truncation methods have been proposed in [3], [13], which are reviewed below.
A. Lag Based Truncation

The truncation approach specified in [13] reduces the order of the paraunitary matrix by removing the $N_1$ leading and $N_2$ trailing lags, unlike the method for parahermitian truncation in [5, 9] this is done asymmetrically. The trim function for paraunitary matrices can be defined as

$$f_{\text{trim}}(\hat{Q}[n]) = \begin{cases} \hat{Q}[n + N_1] & 0 \leq n < N - N_2 - N_1 \\ 0 & \text{otherwise} \end{cases} .$$

The proportion of energy removed in the $N_1$ leading and $N_2$ trailing lags of $Q[n]$ by the $f_{\text{trim}}(\cdot)$ operation is given by

$$\gamma_{\text{trim}} = 1 - \frac{\sum_n \|f_{\text{trim}}(\hat{Q}[n])\|_F^2}{\sum_n \|Q[n]\|_F^2} = 1 - \frac{1}{M} \sum_n \|f_{\text{trim}}(\hat{Q}[n])\|_F^2 ,$$

where $\|\cdot\|_F$ is the Frobenius norm. To control the impact of the truncation operation on the paraunitary matrix, the parameter $\mu$ is used as an upper bound for the proportion of energy removed, $\gamma_{\text{trim}}$. Here we want to maximise the number of lags removed, $N_1 + N_2$, whilst keeping the energy removed below $\mu$, hence the constrained optimisation problem:

$$\text{maximise} \quad (N_1 + N_2) \quad \text{(12)}$$

$$\text{s.t.} \quad \gamma_{\text{trim}} \leq \mu \quad \text{(13)}$$

The implementation of $f_{\text{trim}}(\cdot)$ is as simple as sequentially removing the outermost matrix coefficients of polynomial $Q(z)$, at either leading or trailing lags, which possess the smallest Frobenius norm whilst ensuring (13) is satisfied.

B. Row-Shift Corrected Truncation

The row-shift corrected truncation method [3] exploits the ambiguity in paraunitary matrices [3, 8]. The ambiguity permits $Q(z)$, from (1), to be replaced by $\hat{Q}(z)$ where $\hat{Q}(z) = \Gamma(z)Q(z)$. As argued in [3], the only viable option for the polynomial matrix $\Gamma(z)$ takes the form

$$\Gamma(z) = \text{diag}\{z^{-\tau_1}, z^{-\tau_2}, \ldots, z^{-\tau_M}\} \quad \text{(14)}$$

which consists of $M$ row shifts by $\tau_m$ samples, where $m = 1 \ldots M$, i.e. for each row of the paraunitary matrix. These row shifts can now be used to align the maximum values in each row so that the overall paraunitary matrix can be truncated further.

We can subdivide the paraunitary matrix, $\hat{Q}(z)$, into its $M$ row vectors $\hat{q}_m(z)$, $m = 1 \ldots M$,

$$\hat{Q}(z) = [\hat{q}_1(z) \ldots \hat{q}_M(z)] . \quad \text{(15)}$$

For the row-shift corrected method each row is truncated individually using

$$f_{\text{shift}}(\hat{q}_m[n]) = \begin{cases} \hat{q}_m[n + N_1,m] & 0 \leq n < T_m \\ 0 & \text{otherwise} \end{cases} , \quad \text{(16)}$$

where the overall length of the truncated vector is $T_m = N - N_{2,m} - N_{1,m}$. The row shifts, $\tau_m$, in (14) are then set equal to $N_{1,m} \forall m = 1 \ldots M$. As each vector has unit energy the proportion of energy to be removed from each row, using the vector-valued truncation, $f_{\text{shift}}(\hat{q}_m[n])$, becomes

$$\gamma_{\text{shift},m} = 1 - \sum_n \|f_{\text{shift}}(\hat{q}_m[n])\|_F^2 . \quad \text{(17)}$$

Similar to the lag-based method, this presents us with the following constrained optimisation problem for $f_{\text{shift}}(\cdot)$:

$$\text{maximise} \quad \min_{m} (N_{1,m} + N_{2,m})$$

$$\text{s.t.} \quad \gamma_{\text{shift},m} \leq \frac{\mu'}{M} \forall m = 1 \ldots M . \quad \text{(19)}$$

The length of resulting paraunitary matrix will be $\max_m T_m$ and the maximum proportion of energy removed will be $\mu'$.

The process outlined above is equivalent to truncating each row of $Q(z)$ with the lag based truncation method in Sec. III-A and [13].

IV. Results

To compare the performance of the different truncation approaches on the two PEVD methods, performance metrics for this comparison are given first, followed by the simulation scenarios over which the comparisons are made.

A. Performance Metrics

Reconstruction Error. When the paraunitary matrix, $Q(z)$, is truncated, the paraunitary property is lost. The paraunitary property states $Q(z)\hat{Q}(z) = I$ therefore the difference from paraunitary is

$$E(z) = I_{M \times M} - \hat{Q}_T(z)\hat{Q}_T(z)^\dagger . \quad \text{(20)}$$

with $\hat{Q}_T(z)$ being the truncated matrix, and the transform $E[\tau] \leftrightarrow E(z)$. When $Q(z)$ is a filter bank, the loss in paraunitarity can be measured as the reconstruction error [16]

$$\xi = \frac{1}{M} \sum_{\tau} \|E[\tau]\|_F^2 . \quad \text{(21)}$$

Diagonalisation. The goal of the PEVD algorithms is to iteratively approximate a diagonal parahermitian matrix. Therefore, a second performance criterion measures the reduction in off-diagonal energy calculated as

$$E_{\text{norm}}^{(i)} = \frac{\sum_m \sum_{k=1}^M \|s_k^{(i)}(\tau)\|_F^2}{\sum_{\tau} \|R[\tau]\|_F^2} , \quad \text{(22)}$$

where $s_k^{(i)}(\tau)$ is the modified column vector from (7) containing only off-diagonal elements.

B. Simulation Scenario

For the following results, the PEVD algorithms are run for 100 iterations recording the performance metrics from Sec. IV-A along with the paraunitary matrix order at each iteration. All results apart from those shown in Sec. IV-E have been averaged over an ensemble of $10^3$ instantiations. The initial parahermitian matrix, $R(z)$, for all of the simulations below was generated using the source model described in [11], which
is randomised to produce a unique $R(z)$ for each instantiation. The paraunitary matrix produced is $R(z) \in \mathbb{C}6 \times 6$ for all of the simulations excluding Sec. IV-E where the dimensions are reduced to $R(z) \in \mathbb{C}4 \times 4$ with the number of lags for all $R(z)$ set at 47.

C. Reconstruction Error

Based on the results in [3], the truncation parameter for the row-shift corrected method is increased by a factor of 5, which leads to the error of the two approaches being approximately equal. For comparison, the truncation parameters across the PEVD methods remain the same with $\mu = 10^{-4}$ and $\mu' = 5\mu$. Fig. 1 shows the reconstruction error for the different PEVD methods, here the SBR2 algorithm using the row-shift corrected method performs the best with an error of $4.5 \times 10^{-4}$ and SBR2 using the original truncation is worst with an error of $4.8 \times 10^{-4}$ after 100 iterations. Initially the error curves start very low but they quickly increase as the outer elements become smaller and the truncation algorithms begin to remove their full quota of energy, be it $\mu$ or $\mu'$. Even with the compensation of $\mu' = 5\mu$, the row-shift correction method still tends to have a slightly lower error for both PEVD methods despite being permitted to remove five times the energy.

D. Truncated Order and Diagonalisation

As previously shown in [3], the row-shift corrected method has a significant effect on reducing the paraunitary order for the SBR2 method, however with the SMD algorithm the same reduction paraunitary order is not apparent for the selected source model. There is still a slight benefit to using the row-shift corrected truncation but due to the nature of the SMD algorithm there tends to be fewer outliers which need to be corrected by the row-shift truncation.

Fig. 3 shows the diagonalisation measure vs. paraunitary order for the different PEVD algorithms and truncation methods. In Fig. 3 there is a similar trend to Fig. 2, with the row-shift corrected method SBR2 outperforming the SMD equivalent.

Crucially for a set level of diagonalisation, SBR2 with the row-corrected truncation generates a paraunitary matrix of lower cost than SMD, thereby negating one of the benefits of the SMD approach identified in Sec. II and in [11] for this particular source model.

E. Examples of Truncated Paraunitary Matrices

Figs. 4 and 5 illustrate the differences in the paraunitary matrices for both the lag-based truncation method with the row-corrected results overlaid. For clarity of the figures, the initial parahermitian matrix has been reduced to $R(z) \in \mathbb{C}4 \times 4$ with all other simulation parameters remaining the same. Clearly in the SBR2 paraunitary matrix the row shift corrected approach has more of an effect than it does for SMD. Whereas the maxima in the rows of the SBR2 paraunitary matrix are delayed and spread out with respect to one another, the SMD paraunitary maxima are clustered and delayed by a similar amount. With SMD, the row-shift correction does not aid in reducing the paraunitary matrix length. For the examples shown in Figs. 4 and 5 the average reduction in order for the row-corrected method is 21 for the SBR2 paraunitary matrix but only 1 for SMD, and the total average row-corrected lengths
are 21 and 29 respectively.

V. CONCLUSION

The recently developed row-shift corrected truncation method has been applied to both the SBR2 and now the SMD PEVD algorithms with varying results. When applied to paraunitary matrices produced by the SMD algorithm the row corrected approach has less of an affect than it does when paraunitary matrices from SBR2 are used. Although not as dramatic as the SBR2 benefits, there are some minor reductions in both reconstruction error and average paraunitary order once the compensation factor for \( \mu' \) is applied. Rather than using the approach of allowing a proportion of energy to be removed from the paraunitary matrix it may be better to specify a maximum acceptable reconstruction error as without \( \mu' \) the row corrected approach would not be as effective.

REFERENCES