Multiple Shift Maximum Element Sequential Matrix Diagonalisation for Parahermitian Matrices

Jamie Corr*, Keith Thompson*, Stephan Weiss*, John G. McWhirter†, Soydan Redif‡, Ian K. Proudler§

*Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow, Scotland
†School of Engineering, Cardiff University, Cardiff, Wales, UK
‡Electrical and Electronic Engineering Department, European University of Lefke, Cyprus
§School of Electrical, Electronics and Systems Engineering, Loughborough University, Loughborough, UK
{jamie.corr,keith.thompson,stephan.weiss}@strath.ac.uk

Abstract—A polynomial eigenvalue decomposition of parahermitian matrices can be calculated approximately using iterative approaches such as the sequential matrix diagonalisation (SMD) algorithm. In this paper, we present an improved SMD algorithm which, compared to existing SMD approaches, eliminates more off-diagonal energy per step. This leads to faster convergence while incurring only a marginal increase in complexity. We motivate the approach, prove its convergence, and demonstrate some results that underline the algorithm's performance.

I. INTRODUCTION

Parahermitian polynomial matrices occur in e.g. broadband sensor array problems: when in calculating correlations of a data vector $\mathbf{x}[n]$ proper time delays rather than phase shifts, as in the narrowband case, must be considered. In this context, a useful quantity is the space-time covariance matrix $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$, where $\mathcal{E}\{\cdot\}$ denotes expectations and $\{\cdot\}^{\mathrm{H}}$ is the Hermitian transpose operator, the z-transform of which is the cross-spectral density (CSD) matrix $\boldsymbol{R}(z) = \sum_\tau \mathbf{R}[\tau]z^{-\tau}$. Define the parahermitian operator $\tilde{\boldsymbol{R}}(z) = \boldsymbol{R}^{\mathrm{H}}(z^{-1})$, i.e Hermitian transposing all matrix-value coefficients and time-reverses the auto- and cross-correlation terms in $\mathbf{R}[\tau]$. Note that the CSD matrix is parahermitian with $\tilde{\boldsymbol{R}}(z) = \boldsymbol{R}(z)$.

An eigenvalue decomposition offers a powerful tool to factorise Hermitian matrices to e.g. reveal subspace decompositions or identify optimal transforms for coding and compression [1], [2]. For the polynomial case a polynomial EVD (PEVD) has been generalised in [4] as

$$\boldsymbol{R}(z) \approx \boldsymbol{H}(z)\boldsymbol{\Gamma}(z)\tilde{\boldsymbol{H}}(z) \quad , \tag{1}$$

where $\boldsymbol{H}(z)$ is paraunitary such that $\boldsymbol{H}(z)\tilde{\boldsymbol{H}}(z) = \tilde{\boldsymbol{H}}(z)\boldsymbol{H}(z) = \mathbf{I}$ and $\boldsymbol{\Gamma}(z)$ is diagonal and spectrally majorised [2]. The approximation in (1) has been suggest to hold very closely if the polynomial order of $\boldsymbol{H}(z)$ is permitted to grow arbitrarily large [3]. This PEVD has found applications in subband coding [5], filter bank-based channel coding [6], design of broadband precoding and equalisation of MIMO systems [7], broadband angle of arrival estimation [8], and other problems.

For the calculation of the PEVD in (1), a number of iterative algorithms have been suggested, including the second order sequential best rotation (SBR2) algorithm [4], an approximate

PEVD [10], a subband coding-optimised version of SBR2 [5] and a sequential matrix decomposition (SMD) algorithm [9]. All these algorithms calculate a sequence of simple paraunitary transformtion with the aim of reducing off-diagonal power in the parahermitian matrix. The SMD algorithms have shown superior convergence due to eliminating an entire column rather than just the maximum off-diagonal element, as in the case of SBR2. The transfer of additional energy comes at the expense of having to perform a matrix multiplication for every lag value of the parahermitian matrix. In this paper, we extend this idea by transferring additional energy from multiple columns, whereby little extra cost over the standard SMD algorithm arises.

In this paper, Sec. II reviews iterative PEVD approximations; Sec. III outlines the proposed algorithm and its convergence; results are shown in Sec. IV and conclusions drawn in Sec. V.

II. ITERATIVE PEVD ALGORITHMS

A. Second Order Sequential Best Rotation Algorithm

The idea of SBR2 is to iteratively diagonalise $\boldsymbol{R}(z)$, whereby at each step the maximum off-diagonal element is identified and its energy transferred onto the diagonal by means of an elementary paraunitary transformation. The latter consists of a delay to bring the element in question onto the lag zero matrix $\mathbf{R}[0]$, where it is eliminated by a Jacobi rotation.

Starting with $\boldsymbol{S}^{(0)}(z) = \boldsymbol{R}(z)$, at the ith iteration we first perform a delay step

$$\boldsymbol{S}^{(i)\prime}(z) = \tilde{\boldsymbol{\Lambda}}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\boldsymbol{\Lambda}^{(i)}(z) , \quad i = 1\ldots I , \tag{2}$$

where

$$\boldsymbol{\Lambda}^{(i)} = \mathrm{diag}\{\underbrace{1 \ldots 1}_{k^{(i)}-1} \, z^{-\tau^{(i)}} \, \underbrace{1 \ldots 1}_{M-k^{(i)}}\} \tag{3}$$

shifts the $k^{(i)}$th column of $\boldsymbol{S}^{(i-1)}(z)$ by $\tau^{(i)}$ samples, and $\tilde{\boldsymbol{\Lambda}}^{(i)}$ shifts the $k^{(i)}$th row in the opposite lag direction.

To find the maximum off-diagonal element, we define a modified column vector $\hat{\mathbf{s}}_k^{(i)}[\tau] \in \mathbb{C}^{M-1}$, which contains all elements in the $k^{(i)}$th column of $\mathbf{S}^{(i)}[\tau]$ except for the diagonal element. Therefore, the optimum parameter set for
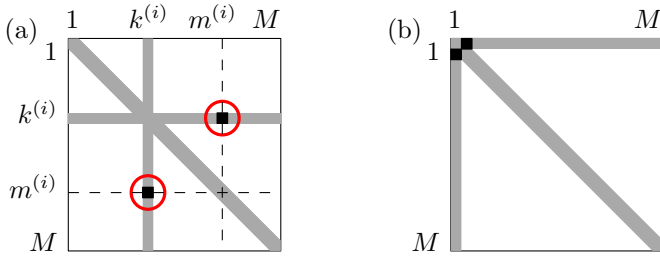
Fig. 1. Sparsity structure of lag zero matrix $\mathbf{S}^{(i)\prime}[0]$ after ith iteration of (a) the SMD algorithm, indicating the maximum off-diagonal element in position $m^{(i)}$ of the $k^{(i)}$th row, and (b) after permutation.

(3) is obtained from

$$\{k^{(i)}, \tau^{(i)}\} = \arg\max_{k,\tau} \|\hat{\mathbf{s}}_k^{(i-1)}[\tau]\|_\infty \quad , \tag{4}$$

such that the optimum off-diagonal element will now lie in the lag zero matrix $\mathbf{S}^{(i),\prime}[0]$.

The elimination of the maximum off-diagonal element is accomplished by a Jacobi rotation, denoted here by a unitary matrix $\mathbf{Q}^{(i)}$,

$$\mathbf{S}^{(i)}(z) = \mathbf{Q}^{(i)\mathrm{H}}\mathbf{S}^{(i)\prime}(z)\mathbf{Q}^{(i)} \quad . \tag{5}$$

This Jacobi rotation is applied to only two rows and columns of $\mathbf{S}^{(i)\prime}(z)$, defined by the column and row indices of the maximum off-diagonal element according to (4). The energy of this maximum element is transferred to the diagonal of $\mathbf{S}^{(i)}[0]$, allocating more energy to the element higher up on the diagonal, which favours but does not guarantee eventual spectral majorisation.

The algorithm has been proven to converge [4], since the paraunitary operations do not alter the total energy in $\mathbf{S}^{(i)\prime}(z)$, while in every step the off-diagonal energy is further minimised. The algorithm stops after I iterations, either when a maximum number of iterations is reached, or if the off-diagonal energy falls below a defined threshold. In this case, the paraunitary matrix

$$\mathbf{H}(z) = \prod_{i=1}^{I} \mathbf{Q}^{(i)}\mathbf{\Lambda}^{(i)}(z) \tag{6}$$

performs the computed decomposition.

B. Sequential Matrix Diagonalisation Algorithm

SMD algorithms differ from SBR2 in that they clear all off-diagonal elements of the zero lag matrix $\mathbf{S}^{(i)}[0]$ at every step. An initialisation step is required to ensure that all instantaneous correlations are removed by means of an EVD,

$$\mathbf{S}^{(0)}[0] = \mathbf{Q}^{(0)\mathrm{H}}\mathbf{R}[0]\mathbf{Q}^{(0)} \quad , \tag{7}$$

such that $\mathbf{S}^{(0)}[0]$ is diagonal, and $\mathbf{S}^{(0)}(z) = \mathbf{Q}^{(0)\mathrm{H}}\mathbf{R}(z)\mathbf{Q}^{(0)}$. Subsequently, at the ith iteration, in a first step the $k^{(i)}$th column is transferred onto the zero lag matrix according to (2), creating a sparsity structure for $\mathbf{S}^{(i)\prime}[0]$ as shown in Fig. 1(a). In the second step, the matrix $\mathbf{S}^{(i)\prime}[0]$ is diagonalised according

to (5), whereby $\mathbf{Q}^{(i)}$ is the modal matrix of an EVD of $\mathbf{S}^{(i)\prime}[0]$ rather than a simple Jacobi rotation.

SMD identifies the $k^{(i)}$th column containing maximum off-diagonal energy by replacing the L_∞-norm in (4) by the L_2-norm. An alternative version, called maximum element SMD (ME-SMD), searches for the column containing the maximum off-diagonal element identical to SBR2 in (4), but performs the SMD-characteristic complete diagonalisation.

The major advantage of SMD with respect to SBR2 is that more energy is transferred onto the diagonal per iteration, hence the algorithm will diagonalise a parahermitian matrix in fewer iterations. However, the matrix $\mathbf{Q}^{(i)}$, although computed only based on $\mathbf{S}^{(i)\prime}[0]$, has to be applied to $\mathbf{S}^{(i)\prime}[\tau]$ for every lag τ. Since $\mathbf{Q}^{(i)}$ no longer has the simple structure of a Jacobi rotation but is non-sparse, SMD has a significantly higher computational complexity than SBR2. However, SMD is capable of achieving levels of diagonalisation that are unobtainable with SBR2, and can generally realise diagonalisation with paraunitary filters of lower order compared to SBR2.

III. MULTIPLE SHIFT ME-SMD ALGORITHM

A. Idea

As discussed in Sec. II-B, the primary advantage of the SMD algorithm over SBR2 is its faster convergence due to eliminating the off-diagonal energy of an entire column in the lag zero matrix. The idea of the proposed algorithm is to move more than one column — and therefore more energy — onto the lag zero matrix and hence reduce even more off-diagonal energy per iterations. This creates a more complex search and EVD per step; however, the subsequent application of the modal matrix to all lags remains the same. The algorithm proposed below is an evolution of the ME-SMD search, but aims to further increase the off-diagonal energy in the zero lag matrix by additional column shifts at every step; we therefore refer to this approach as multiple shift ME-SMD (MSME-SMD) algorithm.

B. Algorithm

The initialisation of the proposed algorithm follows the SMD family with (7). At the ith iteration, we first use (4) to identify the maximum off-diagonal element, and time-shift it with its column onto the lag zero slice, resulting in the sparsity structure shown in Fig. 1(a). By permuting this matrix to the structure in Fig. 1(b), any subsequent operations within the ith iteration will not affect this maximum off-diagonal element as long as the upper 2×2 matrix remains untouched.

Different strategies to identify and time-shift further columns within the ith iteration exist. The strategy employed in MSME-SMD uses a set of reduced search spaces to ensure $(M-1)$ columns are shifted onto the zero lag at each iteration. After the operations shown in Fig. 1 have been completed, the search space shown in Fig. 2(a) is used. Only the highlighted areas in Fig. 2(a) are considered because all their elements can be permuted into the upper 3×3 matrix. If an element outside the search space in Fig. 2(a) is chosen, such as element 2 in Fig. 2(b), applying permutations results in the elements being
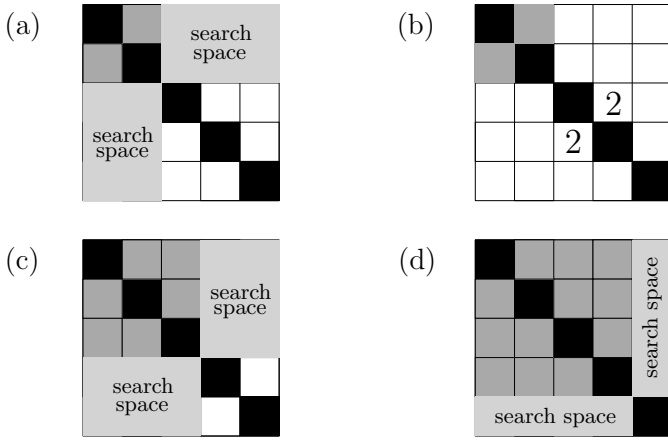
Fig. 2. View of a 5×5 parahermitian matrix during the ith iteration, not showing the lag dimension: starting from the top 2×2 matrix containing the maximum off-diagonal element in (a), (b) shows an example of an element resistant to permutations, the third and fourth stages of the set of reduced search space strategy are shown in (b) and (d).

left in the upper 4×4 rather than 3×3 matrix. Crucially, for the next step when element 2 is chosen, only one further element from the right most column and bottom row, i.e. Fig. 2(d), could be chosen without affecting the previous two maxima. By contrast, the next step when the strategy in Fig. 2(a) is used allows a larger search space. Using another, similar, search space reduction in Fig. 2(c), a third maximum is chosen that can be permuted into the upper 4×4 matrix. Finally a fourth maximum can be chosen from the search space highlighted in Fig. 2(d).

Therefore, a total of $(M-1)$ columns have been shifted during the ith operation, redefining the simple delay matrix in step (2) as a more complex delay & permutation matrix

$$\mathbf{\Lambda}^{(i)} = \mathrm{diag}\{1 \ \ z^{-\tau^{(i,1)}} \ \ \ldots \ \ z^{-\tau^{(i,M-1)}}\} \ \mathbf{P}^{(i)} \qquad (8)$$

whereby the permutation matrix $\mathbf{P}^{(i)}$ accumulates all the column shift operations discussed above. The delays $\tau^{(i,m)}$, $m = 1 \ldots (M-1)$ are the lag values at which the maximum elements for the different columns in Fig. 2 have been found.

Although the sequence of columns and rows has been mixed during the above steps, the ith iteration concludes with applying an ordered EVD [1] as in the ME-SMD algorithms. This ensures that the diagonal elements of the zero lag matrix are ordered in descending energy, thereby encouraging spectral majorisation.

C. Convergence

Theorem 1 (Convergence of the MSME-SMD Algorithm): With a sufficiently large number of iterations I, the multiple-shift ME-SMD algorithm approximately diagonalises $\boldsymbol{R}(z)$ and decreases the power in off-diagonal elements to an arbitrarily low threshold $\epsilon > 0$.

Proof: A number of norms are required to prove Theo-

rem 1. With $s_{m,m}^{(i)}[0]$ the mth diagonal element of $\mathbf{S}^{(i)}[0]$,

$$\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\} \triangleq \sum_{m=1}^{M} |s_{m,m}^{(i)}[0]|^2 \qquad (9)$$

is invariant to shifts and permutations, i.e.

$$\mathcal{N}_1\{\boldsymbol{S}^{(i)\prime}(z)\} = \mathcal{N}_1\{\boldsymbol{\Lambda}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{\Lambda}}^{(i)}(z)\}$$
$$= \mathcal{N}_1\{\boldsymbol{S}^{(i-1)}(z)\} \quad . \qquad (10)$$

The energy of the lag zero matrix

$$\mathcal{N}_2\{\boldsymbol{S}^{(i)}(z)\} \triangleq \|\mathbf{S}^{(i)}[0]\|_{\mathrm{F}}^2 \qquad (11)$$

is invariant under any unitary operation,

$$\mathcal{N}_2\{\boldsymbol{S}^{(i)}(z)\} = \mathcal{N}_2\{\mathbf{Q}^{(i)}\boldsymbol{S}^{(i)\prime}(z)\mathbf{Q}^{(i)\mathrm{H}}\}$$
$$= \mathcal{N}_2\{\boldsymbol{S}^{(i)\prime}(z)\} \quad . \qquad (12)$$

Further,

$$\mathcal{N}_3\{\boldsymbol{S}^{(i)}(z)\} \triangleq \mathcal{N}_2\{\boldsymbol{S}^{(i)}(z)\} - \mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\} \qquad (13)$$
$$\mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\} \triangleq \sum_{\tau} \|\mathbf{S}^{(i)}[\tau]\|_{\mathrm{F}}^2 \qquad (14)$$

where $\| \cdot \|_{\mathrm{F}}$ denotes Frobenius norm and the total energy $\mathcal{N}_4\{\cdot\}$ is invariant under the application of a paraunitary $\boldsymbol{G}^{(i)}(z)$ such that

$$\mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\} = \mathcal{N}_4\{\boldsymbol{G}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{G}}^{(i)}(z)\}$$
$$= \mathcal{N}_4\{\boldsymbol{S}^{(i-1)}(z)\} \quad . \qquad (15)$$

For the off-diagonal norm at the ith iteration,

$$\mathcal{N}_3\{\boldsymbol{S}^{(i)\prime}(z)\} \geq 2\|\hat{\mathbf{s}}_{k^{(i)}}^{(i-1)}[\tau^{(i)}]\|_{\infty}^2 = 2\gamma^{(i)} \quad . \qquad (16)$$

In the following rotation step with $\mathbf{Q}^{(i)}$, this energy is transferred onto the main diagonal such that $\mathcal{N}_3\{\boldsymbol{S}^{(i)}(z)\} = 0$. With

$$\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\} \ > \ \mathcal{N}_1\{\boldsymbol{S}^{(i)\prime}(z)\} + 2\gamma^{(i)}$$
$$= \ \mathcal{N}_1\{\boldsymbol{S}^{(i-1)}(z)\} + 2\gamma^{(i)} \qquad (17)$$

and $\gamma^{(i)} > 0$, $\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ increases monotonically with iteration index i. Since

$$\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\} \leq \mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\} \qquad \forall i \quad , \qquad (18)$$

with the overall energy, $\mathcal{N}_4\{S^{(i)}(z)\}$, remaining constant, $\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ must have a supremum S,

$$S = \sup_i \mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\} \quad . \qquad (19)$$

It follows that for any $\epsilon > 0$ there must be an iteration number I for which $S - \mathcal{N}_1\{\boldsymbol{S}^{(I)}(z)\} < \epsilon$ and so the increase $2\gamma^{(I+i)}$, $i \geq 0$, at any subsequent stage must satisfy

$$2\gamma^{(I+i)} \leq S - \mathcal{N}_1\{\boldsymbol{S}^{(I)}(z)\} < \epsilon \quad . \qquad (20)$$

Hence, for any $\epsilon > 0$, there must be an iteration I by which $\gamma^{(I+i)}$, $i \geq 0$, is bounded by ϵ. ∎
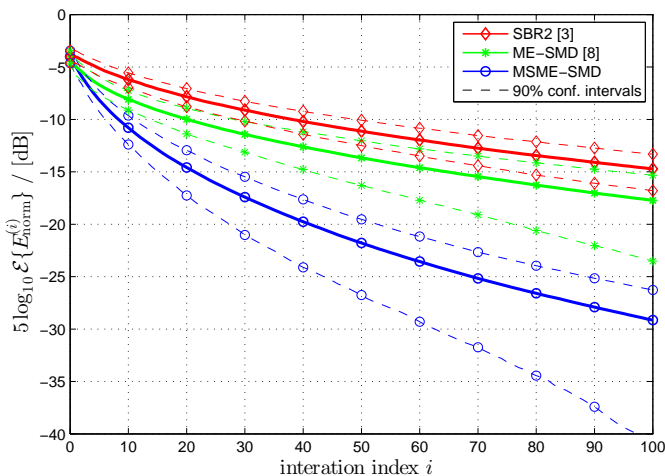
Fig. 3. Comparison of normalised off-diagonalised energy $E_{\text{norm}}^{(i)}$ according to (21) for SBR2, ME-SMD and MSME-SMD, showing ensemble averages with 90% confidence intervals versus iteration index.

Fig. 4. PSDs $S_{m,m}^{(100)}(e^{j\Omega})$ for a) SBR2 [4], b) ME-SMD [9], and c) MSME-SMD demonstrating approximate spectral majorisation.

IV. RESULTS

To demonstrate the proposed algorithm, we assess the reduction of off-diagonal energy at the ith iteration,

$$E_{\text{norm}}^{(i)} = \frac{\sum_\tau \sum_{k=1}^{M} \|\hat{\mathbf{s}}_k^{(i)}[\tau]\|_2^2}{\sum_\tau \|\mathbf{R}[\tau]\|_F^2} \quad , \tag{21}$$

normalised by the total power $\mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\} = \mathcal{N}_4\{\boldsymbol{R}(z)\} = \sum_\tau \|\mathbf{R}[\tau]\|_F^2$. The comparison to the SBR2 [4] and ME-SMD algorithms [5] is calculated over an ensemble of 100 realisations of random parahermitian 5×5 matrices $\boldsymbol{R}(z)$ of order 11. These randomised parahermitian matrices can be generated from matrices $\boldsymbol{A}(z) \in \mathbb{C}^{5\times5}$ of order 6 with independent and identically distributed zero mean unit variance complex Gaussian entries, such that $\boldsymbol{R}(z) = \boldsymbol{A}(z)\tilde{\boldsymbol{A}}(z)$.

The results are depicted in Fig. 3, and confirm the enhanced convergence of ME-SMD over SBR2 due to eliminating an entire off-diagonal column rather than just the maximum off-diagonal element at every iteration step. The proposed MSME-SMD algorithm, by eliminating at least as much as energy as the ME-SMD algorithm per iteration step, provides even faster convergence, and reaches higher levels of diagonalisation as measured by the normalised off-diagonal energy, at the cost of a marginally higher computational complexity.

Fig. 4 shows the power spectral densities along the diagonalised CSD matrix. As hinted earlier, the ordering of energy encourages spectral majorisation — the strict ordering of PSDs at all frequencies — is achieved best by MSME-SMD within the given number of iterations.

V. CONCLUSION

We have presented an enhanced sequential matrix diagonalisation, which iteratively approximates the PEVD of a parahermitian matrix. The algorithm is based on a maximum element SMD version, which at each iteration step brings the maximum off-diagonal element onto the lag zero matrix, where then the entire column is eliminated. Since the main algorithm complexit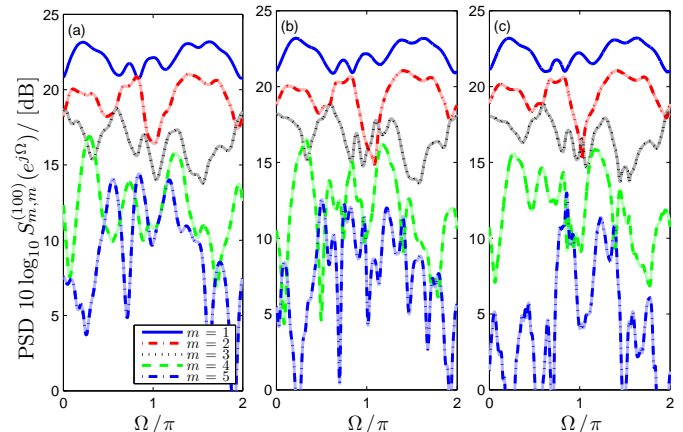y is to apply a unitary matrix of every lag value of the parahermitian matrix, the idea pursued in this paper has been to transfer more off-diagonal energy per iteration step compared to ME-SMD. We have demonstrated the multiple-shift maximum element SMD to be capable of shifting a total of $M - 1$ columns. Due to this additional energy, the algorithm converges significantly faster than both SBR2 and ME-SMD.

REFERENCES

[1] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, 3rd ed., 1996.
[2] P. Vaidyanathan. Theory of optimal orthonormal subband coders. *IEEE Trans SP*, 46(6):1528–1543, June 1998.
[3] S. Icart and P. Comon. Some properties of Laurent polynomial matrices. In *9th IMA Conf. Mathematics in Signal Processing*, Birmingham, UK, Dec. 2012.
[4] J. G. McWhirter, P. D. Baxter, T. Cooper, S. Redif, and J. Foster. An EVD Algorithm for Para-Hermitian Polynomial Matrices. *IEEE Trans SP*, 55(5):2158–2169, May 2007.
[5] S. Redif, J. McWhirter, and S. Weiss. Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition. *IEEE Trans SP*, 59(11):5253–5264, Nov. 2011.
[6] S. Weiss, S. Redif, T. Cooper, C. Liu, P. D. Baxter, and J. G. McWhirter. Paraunitary Oversampled Filter Bank Design for Channel Coding. *EURASIP Journal of Applied Signal Processing*, 2006, 2006.
[7] C. H. Ta and S. Weiss. A Design of Precoding and Equalisation for Broadband MIMO Systems. In *15th Int. Conf. DSP*, pp. 571–574, Cardiff, UK, July 2007.
[8] M. Alrmah, S. Weiss, and S. Lambotharan. An extension of the music algorithm to broadband scenarios using polynomial eigenvalue decomposition. In *19th European Signal Proc. Conf.*, pp. 629–633, Barcelona, Spain, Aug. 2011.
[9] S. Redif, S. Weiss, and J. G. McWhirter. An approximate polynomial matrix eigenvalue decomposition algorithm for para-hermitian matrices. In *11th IEEE Int. Symp. Signal Proc. Inf. Tech.*, pp. 421–425, Bilbao, Spain, Dec. 2011.
[10] A. Tkacenko. Approximate eigenvalue decomposition of para-hermitian systems through successive fir paraunitary transformations. In *IEEE Int. Conf. Acoustics Speech Signal Proc.*, pp. 4074–4077, Mar. 2010.