

Causality-Constrained Multiple Shift Sequential Matrix Diagonalisation for Parahermitian Matrices

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

*Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow, Scotland

†School of Engineering, Cardiff University, Cardiff, Wales, UK

‡School of Electrical, Electronics and Systems Engineering, Loughborough University, Loughborough, UK

{jamie.corr,keith.thompson,stephan.weiss}@strath.ac.uk

Abstract—This paper introduces a causality constrained sequential matrix diagonalisation (SMD) algorithm, which generates a causal paraunitary transformation that approximately diagonalises and spectrally majorises a parahermitian matrix, and can be used to determine a polynomial eigenvalue decomposition. This algorithm builds on a multiple shift technique which speeds up diagonalisation by diagonalisation per iteration step based on a particular search space, which is constrained to permit a maximum number of causal time shifts. The results presented in this paper show the performance in comparison to existing algorithms, in particular an unconstrained multiple shift SMD algorithm, from which our proposed method derives.

I. INTRODUCTION

Polynomial eigenvalue decomposition (PEVD) of parahermitian matrices extends the optimality of the EVD for many narrowband problems to the broadband case, such as for broadband sensor arrays. When calculating covariance matrices for narrowband arrays, the consideration of simple phase shifts suffices; however in the broadband case these must be replaced with actual time delays, leading to covariance matrices that contain polynomials rather than scalar values. Since the narrowband EVD cannot be applied to such matrices, the second order sequential best rotation algorithm (SBR2) [4] has approximated a PEVD, and has found a multitude of applications including broadband direction of arrival estimation [8], precoding and equalisation for broadband MIMO systems [7], and filterbank based channel coding [5], [6]. The SBR2 algorithm is a generalisation of the classical Jacobi algorithm [1] extended to parahermitian polynomial matrices [4].

The idea of Hermitian matrices can be extended to polynomial matrices however in addition to the conjugate symmetry across the diagonal there is also a time reversal i.e. $\mathbf{R}(z) = \mathbf{R}^H(z^{-1})$ where the parahermitian operator $\tilde{\mathbf{R}}(z)$ can be used to signify the Hermitian transpose and time reversal in $\mathbf{R}^H(z^{-1})$. A paraunitary matrix is simply a polynomial matrix whose product with its parahermitian transpose yields the identity, $\mathbf{H}(z)\tilde{\mathbf{H}}(z) = \tilde{\mathbf{H}}(z)\mathbf{H}(z) = \mathbf{I}$ [2]. From [4] the polynomial eigenvalue decomposition (PEVD) of a parahermitian matrix is generalised to

$$\mathbf{R}(z) \approx \mathbf{H}(z)\mathbf{D}(z)\tilde{\mathbf{H}}(z) \quad , \quad (1)$$

where $\mathbf{D}(z)$ is a diagonal polynomial matrix whose diagonals correspond to the approximate polynomial eigenvalues and the

rows of the paraunitary $\mathbf{H}(z)$ are the approximate polynomial eigenvectors of the parahermitian matrix $\mathbf{R}(z)$.

Iterative PEVD algorithms, such as SBR2 [4] or approximate EVD (AEVD) [10], aim to construct the paraunitary matrix, $\mathbf{H}(z)$, through the combination of N simpler paraunitary matrices,

$$\mathbf{H}(z) = \mathbf{G}_N \dots \mathbf{G}_2(z)\mathbf{G}_1(z) \quad , \quad (2)$$

each of which transfers energy from the off-diagonal elements of the parahermitian matrix onto the diagonal. The more energy each of the simpler paraunitary matrices, $\mathbf{G}_n(z)$, transfer to the diagonal, the fewer the number of iterations required to reach a satisfactory paraunitary matrix $\mathbf{H}(z)$.

The sequential matrix diagonalisation algorithm, (SMD) [9], uses a similar approach to construct $\mathbf{H}(z)$ as SBR2 but differences in the techniques used mean each $\mathbf{G}_n(z)$ for SMD transfers more energy per iteration. The algorithm proposed in this paper builds on a recently introduced multiple shift algorithm [3] where the algorithm has been modified to ensure the paraunitary matrix produced is causal (i.e. consists of only delays and no advances) but still transfers more energy per iteration than both SBR2 and SMD.

This paper is arranged as follows: Sec. II reviews existing iterative PEVD algorithms, Sec. III introduces the proposed algorithm, results and conclusions are given in Sec. IV and Sec. V, respectively.

II. EXISTING APPROXIMATE PEVD ALGORITHMS

A. Second Order Sequential Best Rotation Algorithm

The SBR2 algorithm [4] is an iterative approximation of the PEVD of a parahermitian matrix. At every iteration, SBR2 identified the largest off-diagonal element, and, through a series of delay and rotation operations, will eliminate this element and transfer its energy onto the diagonal. The first step of the SBR2 algorithm during the i th iteration is to find the maximum off-diagonal element in the parahermitian matrix and shift it onto the zerolag. A set of modified column vectors, $\hat{\mathbf{s}}_k^{(i)}[\tau] \in \mathbb{C}^{M-1}$, which contain all but the on-diagonal elements, are used to find the column $k^{(i)}$ and lag $\tau^{(i)}$,

$$\{k^{(i)}, \tau^{(i)}\} = \arg \max_{k, \tau} \|\hat{\mathbf{s}}_k^{(i-1)}[\tau]\|_{\infty} \quad , \quad (3)$$

containing the maximum off diagonal element.

Given the lag and column indices, $\tau^{(i)}$ and $k^{(i)}$, the $k^{(i)}$ th column and its complex conjugate row are both shifted in opposite directions by $\tau^{(i)}$ lags using

$$\mathbf{S}^{(i)'}(z) = \tilde{\mathbf{\Lambda}}^{(i)}(z)\mathbf{S}^{(i-1)}(z)\mathbf{\Lambda}^{(i)}(z), \quad i = 1 \dots I, \quad (4)$$

where

$$\mathbf{\Lambda}^{(i)} = \text{diag}\{\underbrace{1 \dots 1}_{k^{(i)}-1} z^{-\tau^{(i)}} \underbrace{1 \dots 1}_{M-k^{(i)}}\} \quad (5)$$

shifts the $k^{(i)}$ th column and the corresponding row is shifted by the same number of lags in the opposite direction using $\tilde{\mathbf{\Lambda}}^{(i)}(z)$.

The energy from the maximum element, now on the zero lag, is transferred onto the diagonal using the Jacobi rotation, $\mathbf{Q}^{(i)}$,

$$\mathbf{S}^{(i)}(z) = \mathbf{Q}^{(i)\text{H}}\mathbf{S}^{(i)'}(z)\mathbf{Q}^{(i)} \quad (6)$$

Due to their sparseness, the Jacobi rotation matrices $\mathbf{Q}^{(i)}$ and $\mathbf{Q}^{(i)\text{H}}$ only affect two rows and columns in the parahermitian matrix. Spectral majorisation of the paraunitary matrix can be encouraged by ordering the zero lag diagonal of the parahermitian matrix after each iteration.

The SBR2 algorithm stops when either a fixed number of iterations, I , have passed or the maximum off-diagonal element falls below a given threshold. The decomposition computed by SBR2 can be performed by a single paraunitary matrix, $\mathbf{H}(z)$,

$$\mathbf{H}(z) = \prod_{i=1}^I \mathbf{Q}^{(i)}\mathbf{\Lambda}^{(i)}(z) \quad , \quad (7)$$

which consists of the product of the I rotation and delay matrices.

B. Sequential Matrix Diagonalisation Algorithm

Rather than just eliminating the maximum element as in an SBR2 iteration, the SMD algorithm clears the entire row and column that is shifted onto the zero lag matrix. Therefore, the simple Jacobi rotation used in the SBR2 algorithm is replaced by a full EVD of the zero lag matrix which transfers all of its off-diagonal energy onto the diagonal.

The SMD algorithm is initialised by calculating the full EVD of the zero lag matrix, $\mathbf{S}^{(0)}[0]$, which is then applied to all lags of the parahermitian matrix; this clears the energy from all off-diagonal elements in the zero lag onto its diagonal. The i th iteration of the SMD algorithm, like SBR2, starts by finding the column to be brought onto the zero lag. For SMD, the L_∞ -norm in (3) is replaced by an L_2 -norm to find the column that contains the maximum off-diagonal energy. Based on the identified lag and column indices, $\tau^{(i)}$ and $k^{(i)}$, (4) is used to bring the respective row and column pair onto the zero lag.

The next step in the i th iteration of SMD is equivalent to (6), but replaces the simple SBR2's Jacobi rotation $\mathbf{Q}^{(i)}$ a non-sparse full EVD of the zero lag matrix, $\mathbf{S}^{(i)'}[0]$. The drawback of using the full EVD is that it is more costly to apply than the simple Jacobi rotation; however rather than

only transferring energy from a single element, the full EVD transfers all off-diagonal energy in the zero lag matrix onto its diagonal. The application of an ordered EVD encourages spectral majorisation akin to the SBR2 algorithm. ordering the zero lag after each iteration. Similar stopping criteria to SBR2 are used but modified slightly to reflect the column norm applied for the parameter search.

The disadvantages of the SMD algorithm with respect to SBR2 are the computational cost of calculating the column norms for the search step, applying a full EVD to the zero lag matrix, and thereafter performing a multiplication with unitary modal matrices at all lags in the parahermitian matrix rather than a simple Jacobi rotation of two rows and columns. The major advantage of the SMD algorithm is its ability to transfer more energy onto the diagonal at each iteration. Transferring more energy gives SMD the ability to diagonalise a parahermitian matrix in far fewer iterations than the SBR2 algorithm.

A maximum element SMD algorithm, ME-SMD, is a lower computational cost version of the SMD algorithm. Rather than using the column norm search described above, ME-SMD uses the maximum element search akin to SBR2 to decide which column should be shifted onto the zero lag matrix.

C. Multiple Shift Maximum Element SMD Algorithm

The distinguishing feature of the MSME-SMD algorithm [3] is in the search and shift operations. In every iteration, MSME-SMD finds and shifts $(M - 1)$ maximum elements onto the zero lag for any $M \times M$ parahermitian matrix. MSME-SMD uses the initialisation step of the SMD algorithm, calculating the full EVD of the zero lag which is applied to all lags. The i th iteration then starts with the same maximum element search as SBR2 (3) followed by the delay step to bring the element identified onto the zero lag. Whereas SBR2 and SMD immediately diagonalise the zero lag, in MSME-SMD a set of reduced search spaces, similar to those described in Sec. III, are used to bring a further $(M - 2)$ elements onto the zero lag. Once the $(M - 1)$ elements are on the zero lag, MSME-SMD then follows the approach of the SMD algorithm where a full EVD of the zero lag is calculated and applied to all lags in the parahermitian matrix.

The major advantage of the MSME-SMD is that during each iteration it shifts more energy than both the SBR2 and SMD onto the diagonal and is able to diagonalise a parahermitian matrix in yet fewer iterations. Compared to the SMD algorithm the search and EVD steps are more costly, although the cost in applying the EVD to the zero lag matrix and subsequently to the modal matrix to all lags of the parahermitian matrix is the same.

III. CAUSALITY CONSTRAINED MSME-SMD ALGORITHM

A. Causality Considerations

The causality of the time shift step in the original SBR2 [4] is not guaranteed, as $\tau^{(i)}$ in (5) can be positive or negative. Note however, that if the maximum element is identified in

column $k^{(i)}$ and row $m^{(i)}$ at lag $\tau^{(i)}$, the parahermitian nature of $\mathbf{S}^{(i-1)}(z) = \tilde{\mathbf{S}}^{(i-1)}(z)$ implies that a corresponding value sits in column $m^{(i)}$ and row $k^{(i)}$ at lag $-\tau^{(i)}$. Therefore, the same maximum element pair shifted by (5) can also be brought onto the zero lag matrix by

$$\mathbf{\Lambda}^{(i)} = \text{diag}\{\underbrace{1 \dots 1}_{m^{(i)}-1} z^{\tau^{(i)}} \underbrace{1 \dots 1}_{M-m^{(i)}}\} . \quad (8)$$

This alternative can be invoked to pick a causal operation from either (5) or (8) at the i th iteration, such that the overall paraunitary matrix in (7) consist of only causal elements. The two operations, although shifting the same two target elements, will however result in different parahermitian matrices $\mathbf{S}^{(i)'}(z)$.

B. Idea

In the MSME-SMD algorithm mentioned in Sec. II-C additional energy compared to the standard SMD is transferred at each iteration by shifting more columns onto the zero lag matrix, whereby the search space in which subsequent column and row shifts are identified plays a crucial role, as will be outlined below. The causality discussed in Sec. III-A will have an impact not just in terms of the causality property of the extracted paraunitary matrix, but also its growth in order with every iteration. By restricting the search space in (3) to positive lags, we below outline a causal multiple shift maximum element SMD (C-MSME-SMD) algorithm and explore some of its properties.

C. Algorithm

The initial step of the C-MSME-SMD algorithm is identical to that of other SMD algorithms, whereby the zero lag matrix is fully diagonalised by an EVD, whose modal matrix is then applied to all lags of the parahermitian matrix. At each iteration, the algorithm shifts $(M-1)$ maximum elements onto the zero lag matrix ensuring that the paraunitary matrix generated is kept causal. C-MSME-SMD then, like the other SMD algorithms, proceeds to carry out a full EVD of the zero lag matrix to complete the iteration.

The search strategy for the C-MSME-SMD algorithm is based on the standard MSME-SMD search [3] with some modifications that lead to a causal paraunitary matrix. To ensure causality in the paraunitary matrix the search is restricted to the positive lag halfspace of the parahermitian matrix so that any elements found are delayed onto the zero lag matrix.

The first element in the i th iteration is found using a maximum element search, similar to (3) but it is restricted to $\tau \geq 0$ for the parahermitian matrix. Once the first element, a , and its conjugate, a^* , from the rear half of the parahermitian matrix have been brought onto the zero lag matrix, for easier understanding of the search strategy we permuted the two maxima into the upper left 2×2 submatrix, we obtain a zero lag matrix similar to Fig. 1(a).

Continuing from Fig. 1(a), the search space for the second element is highlighted in Fig. 1(b), and is to ensure that the previous two elements are not affected. To ensure the extraction of ultimately $(M-1)$ shifts, the complex conjugate

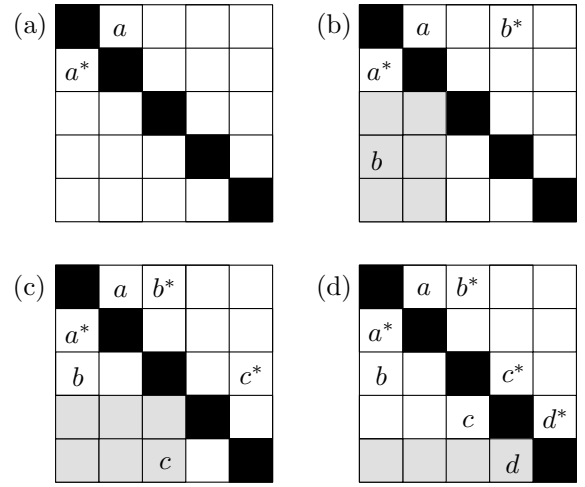


Fig. 1. C-MSME-SMD search strategy for a 5×5 matrix showing a) the initial step after permutations, the reduced search spaces, permutations and choices for b) the second, c) the third, and d) the fourth maximum chosen.

of the next element has to share a row with a previously identified maximum. In the m th step, the search space ensures that only one row is removed from the search space in step $(m+1)$, maximising the number of shifts to $(M-1)$. Choosing element b in Fig. 1(b), its complex conjugate b^* will share row 1 with the first element and only the third row will be removed from the second search space shown in Fig. 1(c) once appropriate permutations are applied. The same search approach is used to find the 3rd element, c , in Fig. 1(c) and 4th element, d , in Fig. 1(d). If we only ensure that previous elements are not affected then we could potentially remove two rows per selection. In the 5×5 case we could only guarantee three or in general $(M/2)$ maximum elements per iteration.

The shift and permutation operations used in the search step can be combined into the delay matrix,

$$\mathbf{\Lambda}^{(i)} = \text{diag}\{1 z^{-\tau^{(i,1)}} \dots z^{-\tau^{(i,M-1)}}\} \mathbf{P}^{(i)} \quad (9)$$

which is the product of the individual delays and permutations used in Fig. 1. The lag values for each of the delays in Fig. 1 are used to generate the delays $\tau^{(i,m)} \geq 0$, $m = 1 \dots (M-1)$. The matrix $\mathbf{P}^{(i)}$ contains the various permutations used to relocate the elements at each of the $(M-1)$ steps in the search.

When the $(M-1)$ maximum elements are all on the zerolag slice, a full ordered EVD of the zero lag matrix, $\mathbf{S}^{(i)}[0]$, is calculated and its modal matrix applied to all lags in the parahermitian matrix. Convergence of the non-causal MSME-SMD algorithm has already been proven [3] which also holds for the this causality constrained C-MSME-SMD algorithm.

The delay matrix $\mathbf{\Lambda}^{(i)}$ as defined in (9) is also applied in the MSME-SMD algorithm, but without the restriction to $\tau^{(i)} \geq 0$. Therefore, the order of $\mathbf{\Lambda}^{(i)}$ in the MSME-SMD could in the worst case be twice as large as in the C-MSME-SMD case, leading to a faster growth in the paraunitary matrix with each iteration. As a drawback, the search space of C-MSME-SMD is only half the size of MSME-SMD, leading to potentially

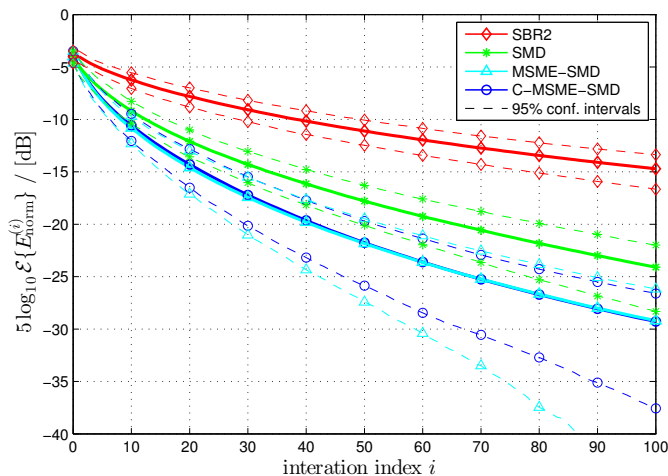


Fig. 2. Convergence plot of mean diagonalisation measure (10) for SBR2, SMD, MSME-SMD and C-MSME-SMD with 95% confidence intervals vs. algorithm iterations for the ensemble simulations

slower diagonalisation. Therefore, the algorithm is likely to offer a trade-off between the diagonalisation performance and the growth in polynomial order compared to the MSME-SMD version.

IV. RESULTS

To compare the convergence of the different PEVD algorithms, the diagonalisation measure of the parahermitian matrix at the i th iteration is calculated as

$$E_{\text{norm}}^{(i)} = \frac{\sum_{\tau} \sum_{k=1}^M \|\hat{\mathbf{s}}_k^{(i)}[\tau]\|_2^2}{\sum_{\tau} \|\mathbf{R}[\tau]\|_F^2}. \quad (10)$$

In (10) the numerator represents the off-diagonal energy, and the denominator is the total energy in the parahermitian matrix, which is invariant under paraunitary operations. Results are averaged over an ensemble of 100 different random 5×5 parahermitian matrices of order 11.

The convergence curves of the SBR2, SMD, MSME-SMD and C-MSME-SMD algorithms in terms of the remaining normalised off-diagonal energy according to (10) are shown in Fig. 2 for comparison. Despite the causality constraint, Fig. 2 shows that C-MSME-SMD converges at an almost identical rate to the non-causal MSME-SMD, both of which converge significantly faster than SMD and SBR2.

The cost to implement a filter bank based on the paraunitary matrix produced by a PEVD algorithm is proportional to its order. The truncation method described in [4] is used to trim any small valued elements at the ends of the paraunitary matrix. Using the same ensemble approach mentioned above, Fig. 3 compares the diagonalisation measure versus the implementation cost for the various PEVD algorithms.

In Fig. 3, both the causal, C-MSME-SMD, and non-causal MSME-SMD algorithms demonstrate very similar implementation costs despite the added constraint of producing a causal paraunitary matrix. Compared with the other causal implementation, SBR2, the major difference is the

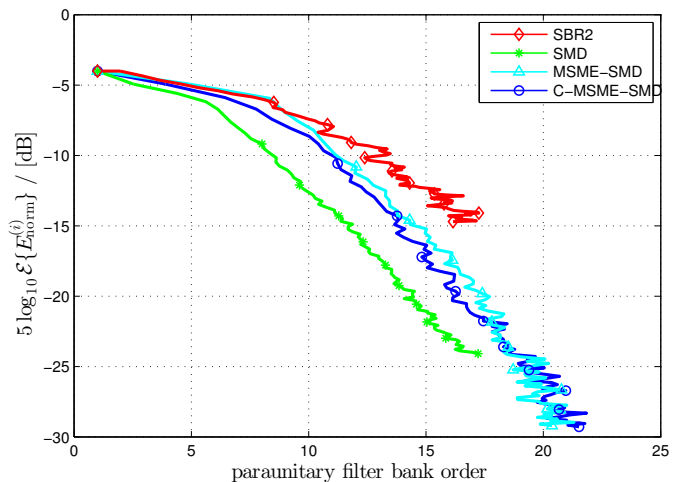


Fig. 3. Mean ensemble diagonalisation measure vs. mean paraunitary filter bank order for SBR2, SMD, MSME-SMD, and C-MSME-SMD.

level of diagonalisation that is achieved. Also C-MSME-SMD outperforms SBR2 in terms of cost for most levels of diagonalisation. The standard SMD algorithm performs best for diagonalisations up to 25 dB w.r.t. the metric in (10), providing paraunitary filter banks with the lowest order. For higher levels of diagonalisation, at least over the number of iterations for which simulations have been run, the MSME-SMD and C-MSME-SMD are the only algorithms that can provide the required level of performance, with no penalty for the causality constraint.

The power spectral densities extracted from the diagonals of the CSD matrix are shown in Figs. 4- 7. As can be clearly seen in Figs. 6 and 7 both of the multiple shift algorithms achieve the best spectral majorisation (i.e. the ordering of the PSDs at all frequencies) after a limited number of 100 iterations. Comparing the causal C-MSME-SMD and non-causal MSME-SMD algorithms, we have found little difference between both algorithms in terms of achieved spectral majorisation over a large number of simulations.

V. CONCLUSION

This paper has presented a causally constrained multiple shift maximum element SMD algorithm for the approximate EVD of a parahermitian matrix. The proposed algorithm is influenced by its non-causal predecessor, the MSME-SMD algorithm, which also brings a total of $(M - 1)$ columns onto the zero lag matrix per iteration. The causality constraint is achieved by limiting the maximum element searches to select portions in the positive lag halfspace of the parahermitian matrix, thus ensuring all elements found have to be delayed rather than advanced onto the zero lag matrix. The results presented here show that the C-MSME-SMD algorithm can achieve similar levels of performance to the unconstrained MSME-SMD algorithm, for the same computational expense, with the added benefit of being causal. The limitation of the search space in the C-MSME-SMD algorithm appears to be compensated by the lower growth in polynomial order.

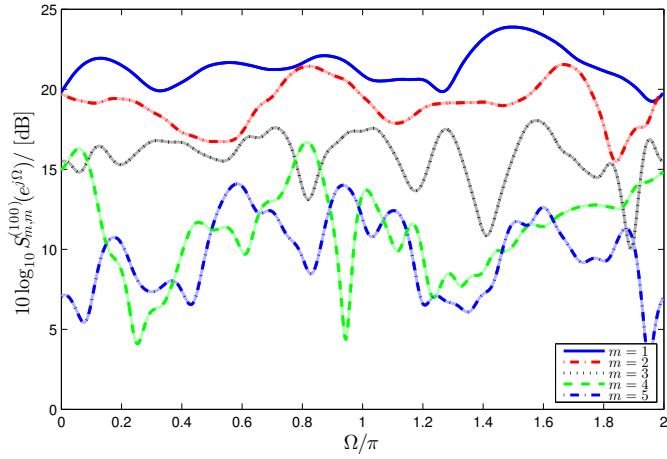


Fig. 4. PSDs $S_{m,m}^{(100)}(e^{j\Omega})$ for SBR2 [4] demonstrating approximate spectral majorisation.

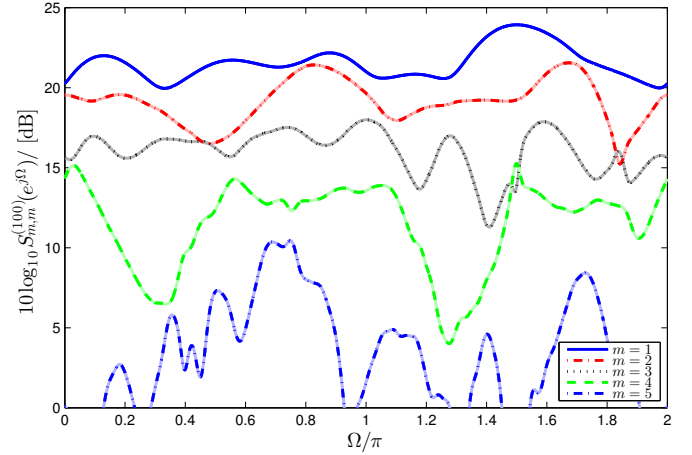


Fig. 7. PSDs $S_{m,m}^{(100)}(e^{j\Omega})$ for C-MSME-SMD demonstrating approximate spectral majorisation.

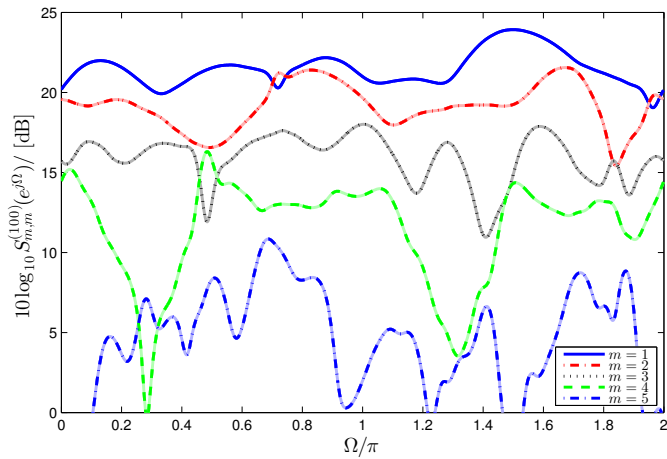


Fig. 5. PSDs $S_{m,m}^{(100)}(e^{j\Omega})$ for SMD [9] demonstrating approximate spectral majorisation.

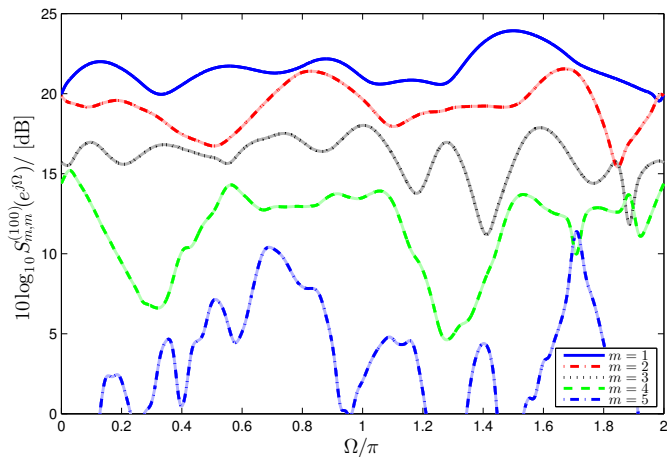


Fig. 6. PSDs $S_{m,m}^{(100)}(e^{j\Omega})$ for MSME-SMD [3] demonstrating approximate spectral majorisation.

Compared to SBR2 and SMD, both the unconstrained and the proposed causality constrained MSME-SMD version perform significantly better in terms of convergence and spectral majorisation.

ACKNOWLEDGEMENT

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/1 and the MOD University Defence Research Collaboration in Signal Processing.

REFERENCES

- [1] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, 3rd ed., 1996.
- [2] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, New Jersey, 1993.
- [3] J. Corr, K. Thompson, S. Weiss, J. G. McWhirter, S. Redif, I. K. Proudler Multiple Shift Maximum Element Sequential Matrix Diagonalisation for Parahermitian Matrices. Submitted to *IEEE Workshop on Statistical Signal Proc.*, Gold Coast, Australia, June, 2014.
- [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.