Sequential Matrix Diagonalization Algorithms for Polynomial EVD of Parahermitian Matrices

Soydan Redif, *Senior Member, IEEE*, Stephan Weiss, *Senior Member, IEEE*, and John G. McWhirter

***Abstract*—For parahermitian polynomial matrices, which can be used, for example, to characterize space-time covariance in broadband array processing, the conventional eigenvalue decomposition (EVD) can be generalized to a polynomial matrix EVD (PEVD). In this paper, a new iterative PEVD algorithm based on sequential matrix diagonalization (SMD) is introduced. At every step the SMD algorithm shifts the dominant column or row of the polynomial matrix to the zero lag position and eliminates the resulting instantaneous correlation. A proof of convergence is provided, and it is demonstrated that SMD establishes diagonalization faster and with lower order operations than existing PEVD algorithms.**

***Index Terms*—MIMO systems, parahermitian matrix, paraunitary matrix, polynomial matrix eigenvalue decomposition.**

I. Introduction

THE eigenvalue decomposition (EVD) of conventional Hermitian matrices plays a central role in DSP, with applications as diverse as principle component analysis, the identification of signal subspaces, and blind signal separation. For some applications, such as MIMO channel decompositions, the singular value decomposition (SVD) is required, but we note that this can always be obtained by means of two EVDs. The EVD is also at the heart of the Karhunen-Loeve transform (KLT) for optimal data compression [1]. These "classical" EVD applications are well suited to narrowband signal processing, where matrices only consist of complex gain factors, or correlations are sufficiently defined by instantaneous covariance matrices.

When addressing broadband signal processing problems, the consideration of only instantaneous correlation is suboptimal if not entirely inappropriate. In the case of a broadband sensor array, for example, information relating to the angle-of-arrival is embedded in the relative time delay of each signal rather than a simple phase shift as in the narrowband case.

Manuscript received July 10, 2014; revised September 30, 2014; accepted October 28, 2014. Date of publication November 04, 2014; date of current version November 27, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ana Perez-Neira. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC)Grant number EP/K014307/1 and the MOD University Defence Research Collaboration in Signal Processing.

S. Redif is with the Electrical and Electronic Engineering Department, European University of Lefke, Gemikonagi 71439, Turkey (e-mail: sredif@eul.edu.tr).

S. Weiss is with the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, Scotland (e-mail: stephan.weiss@strath.ac.uk).

J. G. McWhirter is with Cardiff University, Cardiff CF24 3AA, Wales, U.K. (e-mail: mcwhirterjg@cardiff.ac.uk).

Digital Object Identifier 10.1109/TSP.2014.2367460

A rather obvious approach to decorrelating the broadband sensor signals is to use the independent frequency bin (IFB) method, which splits the broadband spectrum into a number of narrow frequency bands via the discrete Fourier transform (DFT); the narrowband data is then processed using the EVD/SVD. This scheme is also used to achieve spatial multiplexing in wireless communications [2]. However, drawbacks with this method are that correlations and phase-coherence between frequency bands are ignored [3].

If the broadband nature of signals is to be accommodated directly, the relative delays must be carried forward - ideally through space-time covariance matrices, where each entry is not just a single correlation coefficient but an entire auto- or cross-correlation sequence. The corresponding cross-spectral density (CSD) matrix therefore has Laurent polynomial elements and takes the form of a polynomial matrix [4], [5]. The suboptimality of the EVD in broadband situations is then reflected in its inability to diagonalize such a polynomial matrix at more than one time lag.

Generalization of the EVD to polynomial covariance matrices leads to a polynomial EVD (PEVD), which transforms a parahermitian (PH) matrix into a diagonal polynomial matrix by means of paraunitary (PU) matrices or lossless filter banks [6]. The PH and PU properties are generalizations of Hermitian and unitary matrix characteristics to the polynomial case, and will be formally defined later. The existence of such a decomposition is not theoretically guaranteed, although suggestions have been made that any PH matrix can be decomposed with PU matrices of sufficiently high order [7].

Various broadband signal processing tasks can be realized with the help of the PEVD. Its ability to provide a broadband signal subspace decomposition has been exploited for the separation of signals from convolutive mixtures. The use of filterbank based channel coding as a generalization of block coding uses a generator polynomial matrix to span the code subspace, with its complement used as a polynomial parity check matrix to produce a syndrome in the noise-only subspace [8]; here the PEVD's identification of subspaces can yield simple designs [9], [10]. In MIMO communications the PEVD can provide designs for linear [11]–[13] and non-linear [14], [15] precoders and equalizers, extending the EVD's narrowband optimality [16] to the broadband case. The subspace decomposition afforded by the PEVD has also been utilized to generalize the MUSIC algorithm to the case of broadband angle of arrival estimation [17], while the strong decorrelation property has been exploited in pre-processing for broadband beamforming structures [18] and has enabled the design of optimal subband coders [19], [20].

For the calculation of the PEVD, only very limited ideal cases permit an exact decomposition so in general, PEVD algorithms have to rely on iterative approaches. An iterative gradient-based method to diagonalise a PH matrix by means of PU factorization is presented in [21], but is limited to 2×2 PH matrices with a specific structure found in subband coding. Calculation of a PEVD in the DFT domain is performed in [22], whereby the order of the PU filter banks must be strictly limited. This fixed order constraint also applies to a recent DFT-domain approach in [23], but its relaxation of the spectral majorization and paraunitary properties do not necessarily lead to a PEVD as defined in [6], [19]. A fixed order approximate PEVD (AEVD) algorithm operating in the time domain was proposed in [24]. It applies a succession of first-order elementary PU filter stages but does not necessarily lead to a good approximation.

A family of iterative PEVD algorithms based on the second order sequential best rotation (SBR2) approach was proposed previously [6]. In every iteration, SBR2 eliminates the off-diagonal element with maximum magnitude (or the dominant off-diagonal element) of a PH matrix by means of a PU operation. The PU operation is not order-constrained, as in the AEVD, and applies a delay such that the dominant off-diagonal element is transferred onto the zero-lag matrix. A Jacobi rotation then eliminates that element and transfers its energy onto the main diagonal. Because this rotation is applied across all lags, some diagonalization efforts of previous steps will be undone; however, because the dominant off-diagonal element is always targeted, the algorithm has been proven to converge to a good approximate PEVD [6], [20], [25].

In performing a delay operation, SBR2-based algorithms move an entire row and column of the CSD matrix into the zero-lag matrix, where the Jacobi rotation will only eliminate the maximum element. In this paper, we propose and investigate the idea of sequential matrix diagonalisation (SMD) algorithms, which will not only transfer the energy of the maximum element but an entire row and column onto the diagonal, thereby diagonalizing the zero-lag matrix at every iteration. Based on an initial version in [25], below we derive an SMD algorithm which maximizes the energy that is transferred per column in every step, accompanied by a maximum-element SMD (ME-SMD) version, which also diagonalizes the zero-lag matrix at every step but provides a simpler search strategy for a less effective parameter set. The advantages and contributions of the proposed method include:

1) The residual off-diagonal energy is reduced due to targeting the dominant column instead of only the largest off-diagonal element as in the case of SBR2 and applying an ordered EVD on each iteration instead of the Jacobi transformation used by SBR2;
2) The matrix diagonalisation is achieved using lower order PU matrices. This is highly beneficial for a number of applications, including angle-of-arrival estimation and multichannel coding.

The paper is organized as follows. Section II provides notations and definitions used in the remainder of the paper. Iterative PEVD algorithms based on the idea of sequential matrix diagonalisation are introduced in Section III. Based on a mixing model in Section IV, which defines a known ground truth for the PEVD, simulations and results are presented in Section V. Finally, conclusions are drawn in Section VI.

II. Notations and Definitions

Given a vector $\mathbf{x}[n] \in \mathbb{C}^M$ of measurements dependent on discrete time index n and with mean $\mathcal{E}\{\mathbf{x}[n]\} = \underline{0}$, the space-time covariance matrix

$$\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\} \tag{1}$$

measures the correlation corresponding to lag τ, where $\mathcal{E}\{\cdot\}$ represents the expectation operator. Auto-correlation functions of the M measurements in $\mathbf{x}[n]$ reside along the main diagonal of $\mathbf{R}[\tau]$, while cross-correlation terms between the different entries of $\mathbf{x}[n]$ form the off-diagonal terms. Note that due to the definition in (1), $\mathbf{R}[\tau] = \mathbf{R}^{\mathrm{H}}[-\tau]$, where $\{\cdot\}^{\mathrm{H}}$ denotes Hermitian transpose.

The CSD matrix $\boldsymbol{R}(z)$ is obtained by z-transformation of (1),

$$\boldsymbol{R}(z) = \sum_{\tau=-T}^{T} \mathbf{R}[\tau] z^{-\tau}, \tag{2}$$

where the relationship between time domain and transform domain quantities is abbreviated below as $\boldsymbol{R}(z) \bullet\!\!-\!\!-\!\!\circ \mathbf{R}[\tau]$. The support of $\mathbf{R}[\tau]$ is $2T+1$, such that $\mathbf{R}[\tau] = \mathbf{0}\ \forall |t| > T$. Note that dependency on a discrete variable is expressed by square brackets, while dependency on a continuous variable is indicated by round brackets. The quantity $\boldsymbol{R}(z)$ forms a polynomial matrix, or a polynomial with matrix-valued coefficients [4], [5], which is parahermitian (PH), i.e., $\tilde{\boldsymbol{R}}(z) = \boldsymbol{R}^{\mathrm{H}}(z^{-1}) = \boldsymbol{R}(z)$. Polynomial matrices are denoted by their dependency on z, and by their uppercase boldface slanted notation. The superscript $\{\cdot\}^{\mathrm{H}}$ for a polynomial matrix is taken to mean the Hermitian transpose of all polynomial coefficient matrices, while the PH operator $\tilde{\{\cdot\}}$ implies a Hermitian transpose of each coefficient matrix and a replacement of z by z^{-1}, i.e., a Hermitian transposition and time-reversal of the corresponding time domain quantity.

For a PH $\boldsymbol{R}(z)$, the polynomial EVD (PEVD) [6] takes the form

$$\boldsymbol{S}(z) \approx \boldsymbol{H}(z)\boldsymbol{R}(z)\tilde{\boldsymbol{H}}(z), \tag{3}$$

with a diagonal $\boldsymbol{S}(z)$, accomplished by means of decoupling $\boldsymbol{R}(z)$ by a paraunitary (PU) $\boldsymbol{H}(z)$. The diagonalized $\boldsymbol{S}(z) \bullet\!\!-\!\!-\!\!\circ \mathcal{E}\{\mathbf{y}[n]\mathbf{y}^{\mathrm{H}}[n-\tau]\}$ is polynomial, containing on its diagonal the power spectral densities (PSD) of the strongly decorrelated signals

$$\mathbf{y}[n] = \sum_{\nu=0}^{L} \mathbf{H}[\nu]\mathbf{x}[n-\nu] \tag{4}$$

where $\boldsymbol{H}(z) \bullet\!\!-\!\!-\!\!\circ \mathbf{H}[n]$ is of order L. Strong decorrelation [19] implies that the elements $y_m[n], m = 1 \ldots M$, of $\mathbf{y}[n]$ are mutually decorrelated at all lags, such that $\mathcal{E}\{y_i[n]y_j[n-\tau]\} = r_{y_i,y_i}[\tau]\delta(i-j)$ with $r_{y_i,y_i}[\tau]$ arbitrary but fulfilling the necessary properties of an autocorrelation sequence. The approximation sign in (3) indicates that a PEVD decomposition with PU matrices $\boldsymbol{H}(z)$ containing only FIR components does not necessarily exist. However it has been shown that a very close approximation should be possible by letting the filter order grow arbitrarily large [7].

Diagonalisation in (3), or equivalently strong decorrelation in (4), means that

$$\boldsymbol{S}(z) = \text{diag}\{S_{1,1}(z), S_{2,2}(z) \dots S_{M,M}(z)\}. \tag{5}$$

Additionally, akin to an ordered EVD [26] with eigenvalues in descending order, the PSDs $S_{m,m}(e^{j\Omega}) = S_{m,m}(z)|_{z=e^{j\Omega}}, m = 1, 2 \dots M$, should be arranged such that at every normalized angular frequency value Ω

$$S_{m,m}(e^{j\Omega}) \geq S_{m-1,m-1}(e^{j\Omega}), \quad m = 2, \dots M. \tag{6}$$

The property defined by (6) is referred to as spectral majorization [19].

Paraunitarity of a matrix $\boldsymbol{H}(z)$ means that $\boldsymbol{H}(z)\tilde{\boldsymbol{H}}(z) = \tilde{\boldsymbol{H}}(z)\boldsymbol{H}(z) = \mathbf{I}_M$ [5], implementing a lossless filter bank that conserves energy. As a result, for an arbitrary $\mathbf{x}[n]$ with finite energy, in (4), $\sum_n \|\mathbf{x}[n]\|_2^2 = \sum_n \|\mathbf{y}[n]\|_2^2$, where $\|\cdot\|_p$ is the p-norm of a vector. Note that signal powers of elements of $\mathbf{x}[n]$ can be found on the main diagonal of $\mathbf{R}[0]$, which below is referred to as the lag-zero matrix of $\boldsymbol{R}(z)$, such that

$$\text{tr}\{\mathbf{R}[0]\} = \text{tr}\{\mathbf{S}[0]\}, \tag{7}$$

where $\mathbf{S}[\tau] \circ\!\!-\!\!\bullet \boldsymbol{S}(z)$ and $\text{tr}\{\cdot\}$ is the trace operator.

III. PEVD via Polynomial Matrix Diagonalisation

A. Sequential Matrix Diagonalisation Algorithm

To compute the PEVD iteratively, the SMD algorithm at each step eliminates the dominant off-diagonal column (row) entirely, transferring the squared L_2 norm of its off-diagonal elements (off-diagonal energy) onto the main diagonal of the lag-zero coefficient matrix. The dominant off-diagonal column is defined as the one for which this value is greatest. Operating on a CSD matrix $\boldsymbol{R}(z)$, the SMD algorithm starts with a diagonalisation of the lag-zero coefficient matrix $\mathbf{R}[0]$ by means of its modal matrix $\mathbf{Q}^{(0)}$ i.e., $\boldsymbol{S}^{(0)}(z) = \mathbf{Q}^{(0)}\boldsymbol{R}(z)\mathbf{Q}^{(0)\text{H}}$. Note that although the calculation of $\mathbf{Q}^{(0)}$ is only based on the EVD of the lag-zero slice $\mathbf{R}[0]$, it is subsequently applied to the coefficient matrices $\mathbf{R}[\tau]\ \forall\tau$. This initial step corresponds to the instantaneous decorrelation of any underlying time series corresponding to $\boldsymbol{R}(z)$.

In the ith step, $i = 1, 2, \dots L$, the SMD algorithm calculates a transformation of the form

$$\boldsymbol{S}^{(i)}(z) = \boldsymbol{U}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{U}}^{(i)}(z), \tag{8}$$

in which

$$\boldsymbol{U}^{(i)}(z) = \mathbf{Q}^{(i)}\boldsymbol{\Lambda}^{(i)}(z). \tag{9}$$

The product in (9) consists of an elementary PU delay matrix

$$\boldsymbol{\Lambda}^{(i)}(z) = \text{diag}\left\{\underbrace{1 \dots 1}_{k^{(i)}-1} z^{-\tau^{(i)}} \underbrace{1 \dots 1}_{M-k^{(i)}}\right\}, \tag{10}$$

and a unitary matrix $\mathbf{Q}^{(i)}$, with the result that $\boldsymbol{U}^{(i)}(z)$ in (9) is PU by construction. It is convenient for subsequent discussion to define an intermediate variable $\boldsymbol{S}^{(i)\prime}(z)$ where

$$\boldsymbol{S}^{(i)\prime}(z) = \boldsymbol{\Lambda}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{\Lambda}}^{(i)}(z) \tag{11}$$

and

$$\boldsymbol{S}^{(i)}(z) = \mathbf{Q}^{(i)}\boldsymbol{S}^{(i)\prime}(z)\mathbf{Q}^{(i),\text{H}}. \tag{12}$$

The selection of $\boldsymbol{\Lambda}^{(i)}(z)$ and $\mathbf{Q}^{(i)}$ in the ith iteration depends on the position of the dominant off-diagonal column (row) in $\boldsymbol{S}^{(i-1)}(z) \circ\!\!-\!\!\bullet \mathbf{S}^{(i-1)}[\tau]$, as identified by the parameter set

$$\left\{k^{(i)}, \tau^{(i)}\right\} = \underset{k,\tau}{\text{argmax}} \left\|\hat{\mathbf{s}}_k^{(i-1)}[\tau]\right\|_2, \tag{13}$$

where

$$\left\|\hat{\mathbf{s}}_k^{(i-1)}[\tau]\right\|_2 = \sqrt{\sum_{m=1, m\neq k}^{M} \left|s_{m,k}^{(i-1)}[\tau]\right|^2}. \tag{14}$$

$s_{m,k}^{(i-1)}[\tau]$ represents the element in the mth row and kth column of $\mathbf{S}^{(i-1)}[\tau]$ while the hat symbol in (14) signifies that the diagonal element has been omitted from the regular column norm.

The focus on columns does not restrict the generality of the algorithm, since the PH property of $\boldsymbol{S}^{(i-1)}(z)$ ensures that the set identified in (13) also represents the dominant row i.e., the $k^{(i)}$th row at lag $-\tau^{(i)}$. Due to its PH-symmetric form, the shifting process in (11) moves both the dominant off-diagonal row and the dominant off-diagonal column into the zero-lag coefficient matrix and so the modified norm in (14) serves to measure the total energy moved into the zero-lag matrix $\mathbf{S}^{(i)\prime}[0]$.

Since the lag-zero matrix $\mathbf{S}^{(i-1)}[0]$ of $\boldsymbol{S}^{(i-1)}(z) \circ\!\!-\!\!\bullet \mathbf{S}^{(i-1)}[\tau]$ is diagonal, the same property can be imposed on $\mathbf{S}^{(i)}[0]$ by means of the similarity transform in (12) provided $\mathbf{Q}^{(i)}$ is chosen to be the modal matrix obtained from an ordered EVD of $\mathbf{S}^{(i)\prime}[0]$.

The iterative process continues for L steps, say, until $\boldsymbol{S}^{(L)}(z)$ is sufficiently diagonalized with the dominant off-diagonal column (row) norm

$$\max_{k,\tau} \left\|\hat{\mathbf{s}}_k^{(L)}[\tau]\right\|_2 \leq \rho \tag{15}$$

where the value of ρ is chosen to be arbitrarily small. This completes the SMD algorithm and generates an approximate PEVD given by

$$\boldsymbol{S}^{(L)}(z) = \boldsymbol{H}^{(L)}(z)\boldsymbol{R}(z)\tilde{\boldsymbol{H}}^{(L)}(z), \tag{16}$$

with the PU matrix $\boldsymbol{H}^{(L)}(z)$ given by

$$\boldsymbol{H}^{(L)}(z) = \prod_{i=0}^{L-1} \boldsymbol{U}^{(L-i)}(z). \tag{17}$$

To show that the SMD algorithm outlined above performs an approximate PEVD, we state the following theorem:

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

Proof: To prove Theorem 1, a number of norms need to be defined:

$$\mathcal{N}_1\left\{\boldsymbol{S}^{(i)}(z)\right\} \triangleq \sum_{m=1}^{M} \left|s_{m,m}^{(i)}[0]\right|^2 \tag{18}$$

$$\mathcal{N}_2\left\{\boldsymbol{S}^{(i)}(z)\right\} \triangleq \left\|\mathbf{S}^{(i)}[0]\right\|_{\mathrm{F}}^2 \tag{19}$$

$$\mathcal{N}_3\left\{\boldsymbol{S}^{(i)}(z)\right\} \triangleq \mathcal{N}_2\left\{\boldsymbol{S}^{(i)}(z)\right\} - \mathcal{N}_1\left\{\boldsymbol{S}^{(i)}(z)\right\} \tag{20}$$

$$\mathcal{N}_4\left\{\boldsymbol{S}^{(i)}(z)\right\} \triangleq \sum_{\tau}\left\|\mathbf{S}^{(i)}[\tau]\right\|_{\mathrm{F}}^2 \tag{21}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes Frobenius norm and $s_{m,m}^{(i)}[0]$ is the mth diagonal element of $\mathbf{S}^{(i)}[0]$. Note that $\mathcal{N}_1\{\cdot\}$ is invariant under a delay matrix as in (11), i.e.,

$$\begin{aligned}\mathcal{N}_1\left\{\boldsymbol{S}^{(i)\prime}(z)\right\} &= \mathcal{N}_1\left\{\boldsymbol{\Lambda}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{\Lambda}}^{(i)}(z)\right\}\\ &= \mathcal{N}_1\{\boldsymbol{S}^{(i-1)}(z)\},\end{aligned} \tag{22}$$

and that $\mathcal{N}_2\{\cdot\}$ is invariant under a unitary operation, i.e.,

$$\begin{aligned}\mathcal{N}_2\left\{\boldsymbol{S}^{(i)}(z)\right\} &= \mathcal{N}_2\left\{\mathbf{Q}^{(i)}\boldsymbol{S}^{(i)\prime}(z)\mathbf{Q}^{(i)\mathrm{H}}\right\}\\ &= \mathcal{N}_2\left\{\boldsymbol{S}^{(i)\prime}(z)\right\}.\end{aligned} \tag{23}$$

Further, $\mathcal{N}_4\{\cdot\}$ is invariant under the application of a PU $\boldsymbol{U}^{(i)}(z)$ such that

$$\begin{aligned}\mathcal{N}_4\left\{\boldsymbol{S}^{(i)}(z)\right\} &= \mathcal{N}_4\left\{\boldsymbol{U}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{U}}^{(i)}(z)\right\}\\ &= \mathcal{N}_4\left\{\boldsymbol{S}^{(i-1)}(z)\right\}.\end{aligned} \tag{24}$$

The off-diagonal norm of the $k^{(i)}$th column at lag $\tau^{(i)}$ is given by

$$\gamma^{(i)} = \left\|\hat{\mathbf{s}}_{k^{(i)}}^{(i-1)}\left[\tau^{(i)}\right]\right\|_2^2. \tag{25}$$

With step (11), this energy is transferred onto both the $k^{(i)}$th column and $k^{(i)}$th row of the lag zero slice $\mathbf{S}^{(i)\prime}[0]$, such that its total off-diagonal energy is

$$\mathcal{N}_3\left\{\boldsymbol{S}^{(i)\prime}(z)\right\} = 2\gamma^{(i)}. \tag{26}$$

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

$$\begin{aligned}\mathcal{N}_1\left\{\boldsymbol{S}^{(i)}(z)\right\} &= \mathcal{N}_1\left\{\boldsymbol{S}^{(i)\prime}(z)\right\} + 2\gamma^{(i)}\\ &= \mathcal{N}_1\left\{\boldsymbol{S}^{(i-1)}(z)\right\} + 2\gamma^{(i)}\end{aligned} \tag{27}$$

exploiting (22), while the overall energy, $\mathcal{N}_4\{S^{(i)}(z)\}$, remains constant.

Due to (27), $\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ increases monotonically with iteration index i. Because $\mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\}$ is invariant over iterations due to (24) and forms an upper bound

$$\mathcal{N}_1\left\{\boldsymbol{S}^{(i)}(z)\right\} \le \mathcal{N}_4\left\{\boldsymbol{S}^{(i)}(z)\right\} \quad \forall i, \tag{28}$$

$\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ must have a supremum S,

$$S = \sup_i \mathcal{N}_1\left\{\boldsymbol{S}^{(i)}(z)\right\}. \tag{29}$$

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

$$2\gamma^{(L+i)} \le |S - \mathcal{N}_1\{\boldsymbol{S}^{(L)}(z)\}| < \epsilon. \tag{30}$$

Hence, for any $\epsilon > 0$, there must be an iteration L by which $\gamma^{(L)}$ is bounded by ϵ. ■

Note that while $\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ monotonically increases with the iteration index, the value of $\gamma^{(i)}$ in the SMD algorithm does not necessarily decrease monotonically. Each similarity transformation is computed with reference to elements of the lag zero slice $\mathbf{S}^{(i)\prime}[0]$, and is guaranteed to increase $\mathcal{N}_1\{\boldsymbol{S}^{(i)}(z)\}$ by driving the dominant column vector to zero. However, non-zero lag elements of the polynomial matrix $\boldsymbol{S}^{(i)\prime}(z)$, where the same unitary matrix is being applied, can increase the norm of a modified column $|\mathbf{s}_k^{(i)\prime}[\tau]|_2^2$, while reducing the sum of the squares of the diagonal elements. As a result, $\gamma^{(i+1)}$ at the next iteration could be larger than $\gamma^{(i)}$.

The SMD algorithm does not seek to reduce the on-diagonal coefficients for non-zero values of τ, let alone drive them to zero. In the context of strong decorrelation, this would correspond to temporal whitening of the decorrelated signals, which is often highly undesirable and cannot occur as the result of a PU transformation due to the fact that the total PSD is preserved.

It is important to realize that the order of the PU matrices will grow due to the shift operations which are applied. The order necessary to achieve an approximate decomposition cannot be determined prior to applying the SMD algorithm and much of the growth in order which occurs in practice involves coefficient matrices with negligibly small elements. The use of an appropriate truncation procedure, as described in [6] for SBR2, is strongly recommended to curtail unnecessary growth in order.

B. Maximum Element SMD (ME-SMD) Algorithm

This section addresses a lower-cost approximation of the SMD algorithm with respect to the search strategy in each iteration step, compared to the SMD algorithm proposed in Section III.A. The SMD algorithm's search for the maximum norm of modified columns is replaced by the search for the dominant off-diagonal element, hence the term maximum element SMD (ME-SMD) algorithm. While the search is similar to SBR2, the entire shifted column in the lag zero matrix will subsequently be eliminated in SMD fashion.

The modified search strategy can be expressed by replacing the L_2 norm in (14) by the L_∞ norm, such that the search for the optimum parameter set performed at every iteration i becomes

$$\left\{k^{(i)}, \tau^{(i)}\right\} = \underset{k,\tau}{\operatorname{argmax}}\left\|\hat{\mathbf{s}}_k^{(i-1)}[\tau]\right\|_\infty. \tag{31}$$

This parameter may differ from the one determined by the SMD algorithm, since the maximum element targeted by (31) is not necessarily contained in the column with the maximum norm as found by the SMD algorithm according to (13).

Thus, the element search in the ith step is more akin to the SBR2 algorithm, which also picks the largest off-diagonal element. However, in the subsequent rotation step, ME-SMD diagonalizes $\mathbf{S}^{(i)}[0]$, whereas SBR2 only eliminates the dominant off-diagonal element of $\mathbf{S}^{(i)}[0]$. The convergence of the ME-SMD algorithms is covered by the following theorem:

Theorem 2 (Convergence of the ME-SMD Algorithm): With a sufficiently large number of iterations L, the ME-SMD algorithm approximately diagonalizes $\boldsymbol{R}(z)$ and decreases the power in off-diagonal elements to an arbitrarily low threshold $\epsilon > 0$.

Proof: The proof is based on the fact that in each iteration, the ME-SMD algorithm transfers at least as much energy

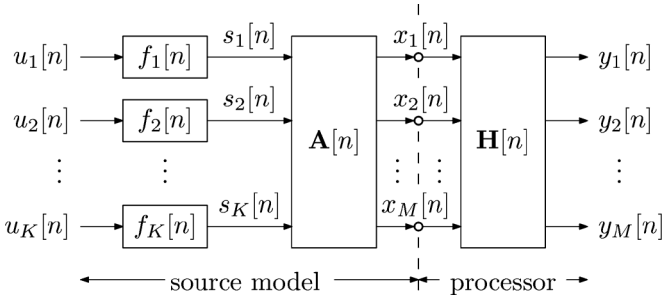


Fig. 1. Source model of K independent Gaussian sources with innovation filters $f_k[n]$ and convolutive mixing matrix $\mathbf{A}[n]$ leading to $M \geq K$ observations $x_m[n]$, followed by a PU processor $\mathbf{H}[n]$ with the aim to generate strongly decorrelated outputs $y_m[n], m = 1 \dots M$.

as SBR2 onto the main diagonal of $\mathbf{S}^{(i)}[0]$. Therefore, the on-diagonal energy grows monotonically and at least as fast as for SBR2. Following on from the proof of convergence for SBR2 in [6], the ME-SMD algorithm also converges. ■

IV. Multichannel Convolutive Mixing Model

A. Source Model

The model for a convolutive mixing system is depicted in Fig. 1. The first stage of this model consists of mutually uncorrelated stochastic processes $s_k[n], k = 1 \dots K$, which emerge from K innovation filters $f_k[n]$ of order N. These innovation filters are excited by uncorrelated, zero mean unit variance complex Gaussian processes $u_k[n] \in \mathcal{N}(0,1)$, such that the cross-correlation

$$\mathcal{E}\{u_l[n]u_k^*[n-\tau]\} = \delta[l-k]\delta[\tau]. \tag{32}$$

The individual PSDs of the signals $s_k[n]$ are therefore $R_{s,k}(e^{j\Omega}) = F_k(e^{j\Omega})F_k^*(e^{j\Omega}) \circ\!\!-\!\!\bullet \mathcal{E}\{s_k[n]s_k^*[n-\tau]\}$ with $F_k(e^{j\Omega}) \bullet\!\!-\!\!\circ f_k[n]$ the Fourier transform of the kth innovation filter. This innovation model is fairly general, but excludes the generation of signals with line spectra [1].

The source vector $\mathbf{s}[n]$, obtained by arranging the K source signals $s_k[n], k = 1 \dots K$, has a diagonal space-time covariance $\mathbf{R}_s[\tau] = \mathcal{E}\{\mathbf{s}[n]\mathbf{s}^{\mathrm{H}}[n-\tau]\}$ and PSD matrix $\boldsymbol{R}_s(z) \bullet\!\!-\!\!\circ \mathbf{R}_s[\tau]$, due to (32). With $\mathbf{s}[n]$ forming the inputs to a convolutive mixing matrix $\mathbf{A}[n] \circ\!\!-\!\!\bullet \boldsymbol{A}(z) \in \mathbb{C}^{M\times K}$ of order P, its outputs $x_m[n], m = 1 \dots M$ are organized in a vector $\mathbf{x}[n]$, with covariance $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$ and CSD matrix

$$\boldsymbol{R}(z) = \boldsymbol{A}(z)\boldsymbol{R}_s(z)\tilde{\boldsymbol{A}}(z), \tag{33}$$

which is entirely based on the innovation filters $F_k(e^{j\Omega}), k = 1 \dots K$, and the convolutive mixing matrix $\boldsymbol{A}(z)$ in Fig. 1.

B. Optimum Decomposition

In order to know the ground truth for the optimum PEVD of $\boldsymbol{R}(z)$ in (33), two conditions are imposed on the realization of the model in Fig. 1 w.r.t. the simulations and results to be presented in Section V:

1) the PSDs of source signals $s_k[n]$, $R_{s_k}(e^{j\Omega})$, are spectrally majorized; and
2) the convolutive mixing matrix $\boldsymbol{A}(z)$ is a PU system.

The spectral majorization helps with some of the metrics to be defined in Section V, and does not restrict our analysis. The paraunitarity of $\boldsymbol{A}(z)$ is an idealizing assumption; however, suggestions in [7] that any PH matrix can be decomposed into a PEVD with PU matrices of sufficient order means that the source model implementation below is restricted only by limiting the order P but not the PU property of the mixing matrix.

With the above selection, $\boldsymbol{R}(z)$ can be decomposed into a PEVD with equality in (3) by using the PU matrix $\boldsymbol{H}(z) = \tilde{\boldsymbol{A}}(z)$. When applied as in Fig. 1 with the M outputs $y_m[n]$ organized in a vector $\mathbf{y}[n]$, this leads to a diagonalized $\boldsymbol{R}_y(z) \bullet\!\!-\!\!\circ \mathbf{R}_y[\tau] = \mathcal{E}\{\mathbf{y}[n]\mathbf{y}^{\mathrm{H}}[n-\tau]\}$, which is spectrally majorized according to (6) such that the diagonal elements

$$R_{y,m}(z) = \begin{cases} F_m(z)\tilde{F}_m(z) & m \leq K \\ 0 & K < m \leq M \end{cases}, \tag{34}$$

i.e., the upper part of $\boldsymbol{R}_y(z)$ matches $\boldsymbol{R}_s(z)$, with the remaining $M-K$ diagonal entries being zero.

C. Spectrally Majorized Innovation Filters

With K moving average (MA) models $f_k[n], k = 1 \dots K$ of order N, spectral majorization requires $|F_k(e^{j\Omega})|^2 \geq |F_{k+1}(e^{j\Omega})|^2 \ \forall \Omega, k = 1 \dots (K-1)$. Starting with unmajorized filters $f_k^{(0)}[n]$ characterized by an arbitrary unit-norm coefficient vector $\mathbf{f}_k^{(0)} \in \mathbb{C}^{L+1}$ and setting $\mathbf{f}_1 = \mathbf{f}_1^{(0)}$, gain factors α_k can be found such that $\mathbf{f}_k = \alpha_k \mathbf{f}_k^{(0)}, k = 2 \dots K$, satisfy spectral majorization. The dynamic PSD range of this basic model can be adjusted with the parameter set $\{K, N\}$.

D. Paraunitary Mixing Matrix

Arbitrary PU matrices $\boldsymbol{A}(z) \in \mathbb{C}^{M\times M}(z)$ of a defined order P can be generated, following a proof in [5], using a concatenation of a unitary $\mathbf{A}_0 \in \mathbb{C}^{M\times M}$ and arbitrary PU first order components

$$\boldsymbol{A}(z) = \mathbf{A}_0 \prod_{p=1}^{P} \boldsymbol{A}_p(z) \tag{35}$$

$$\boldsymbol{A}_p(z) = \mathbf{I} - \mathbf{a}_p\mathbf{a}_p^{\mathrm{H}} + z^{-1}\mathbf{a}_p\mathbf{a}_p^{\mathrm{H}} \tag{36}$$

based on random unit-norm vectors $\mathbf{a}_p \in \mathbb{C}^M, p = 0 \dots P$.

V. Simulations and Results

A. Performance Metrics

Two metrics are defined below, which are normalized such that they can be consistently applied and averaged across ensembles of simulations. It is assumed that the decomposition of $\boldsymbol{R}(z)$ as defined for the source model in (33) is decomposed according to Section III, where $\boldsymbol{S}^{(i)}(z)$ represents the diagonalisation effort at the ith iteration according to (12).

1) Normalized Off-Diagonal Energy: The SMD algorithms minimize the energy residing in the off-diagonal elements of a PH matrix. The off-diagonal energy that remains at the ith iteration is

$$E^{(i)} = \sum_{m=1}^{M} \sum_{\tau} \left\| \hat{\mathbf{s}}_m^{(i)}[\tau] \right\|_2^2, \tag{37}$$

where the modified vector $\hat{\mathbf{s}}_m^{(i)}[\tau]$ and its norm are defined in (14). This can be normalized by the total energy, $\mathcal{N}_4\{\boldsymbol{S}^{(i)}(z)\} = \mathcal{N}_4\{\boldsymbol{R}(z)\}$, recalling the invariance of the $\mathcal{N}_4\{\cdot\}$ norm under PU operations, such that

$$E_{\text{norm}}^{(i)} = \frac{E^{(i)}}{\mathcal{N}_4\{\boldsymbol{R}(z)\}}. \tag{38}$$

The logarithmic metric $5\log_{10} E_{\text{norm}}^{(i)}$ takes into account that covariance matrices already contain quadratic terms.

2) Normalized Coding Gain: Based on the powers of the signals $y_m[n]$ in Fig. 1, the coding gain, which for its maximization requires diagonalisation and spectral majorization of $\boldsymbol{R}_y(z)$ [19], is measured as the ratio of arithmetic and geometric mean of channel variances. At the ith iteration, this variance is $s_{m,m}^{(i)}[0]$ for the mth channel, such that the coding gain is

$$G^{(i)} = \left(\frac{1}{M}\sum_{m=1}^{M} s_{m,m}^{(i)}[0]\right)\left(\prod_{m=1}^{M} s_{m,m}^{(i)}[0]\right)^{-\frac{1}{M}}. \tag{39}$$

Note that the trace

$$\text{tr}\Big\{\mathbf{S}^{(i)}[0]\Big\} = \sum_{m=1}^{M} s_{m,m}^{(i)}[0] = \text{tr}\{\mathbf{R}[0]\} = \text{tr}\{\mathbf{R}_s[0]\}, \tag{40}$$

forming the numerator of (39), is invariant under PU operations.

The optimum coding gain is achieved if $\boldsymbol{R}_y(z)$ fulfils (34), therefore defining the optimum coding gain

$$G_{\text{opt}} = \frac{\frac{1}{M}\sum_{l=k}^{K}\|\mathbf{f}_k\|_2^2}{\prod_{l=k}^{K}\|\mathbf{f}_k\|_2^{2/M}} = \frac{\frac{1}{M}\text{tr}\{\mathbf{R}_s[0]\}}{\det(\mathbf{R}_s[0])^{\frac{1}{M}}}, \tag{41}$$

in dependency of the innovation filter coefficient vectors $\mathbf{f}_k$, whereby $\sum_{l=k}^{K}\|\mathbf{f}_k\|_2^2 = \text{tr}\{\mathbf{R}_s[0]\}$. In (41), it is assumed that $K = M$ to avoid $G_{\text{opt}} \to \infty$. With (41) and (39), the normalized coding gain

$$G_{\text{norm}}^{(i)} = \frac{G^{(i)}}{G_{\text{opt}}} = \left(\frac{\det(\mathbf{R}_s[0])}{\prod_{m=1}^{M} s_{m,m}^{(i)}[0]}\right)^{\frac{1}{M}} \tag{42}$$

arises, such that $0 \le G_{\text{norm}}^{(i)} \le 1$ serves as a measure of how well the approximate PEVD algorithms perform independently of any specific source model.

B. Simulation Scenario

The scenario considered here is for $K = 8$ independent sources and $M = 8$ sensors. The order of the innovation filters and PU mixing matrices are $N = 16$ and $P = 16$, respectively. This leads to an 8×8 CSD matrix of order $2(P+N) = 64$, such that the corresponding space-time covariance matrix $\mathbf{R}[\tau] = \mathbf{0}\forall|\tau| > 32$. Results are averaged across an ensemble of 10^3 randomly generated source models whose dynamic range is limited to realistic values of around 30 dB.

C. Convergence Comparison

The evolution of off-diagonal energy over the iteration steps of the various algorithms is shown in Fig. 2. We see that the ensemble medians $\mathcal{M}\{E_{\text{norm}}^{(i)}\}$ for SMD-type algorithms converge significantly faster than for SBR2 [6] and SBR2C [20]—the convergence curves are separated by several standard deviations of the ensemble, as evidenced by the confidence intervals within which 90% of the ensemble results fall. This gain is due to the enhanced transfer of off-diagonal energy in every step.

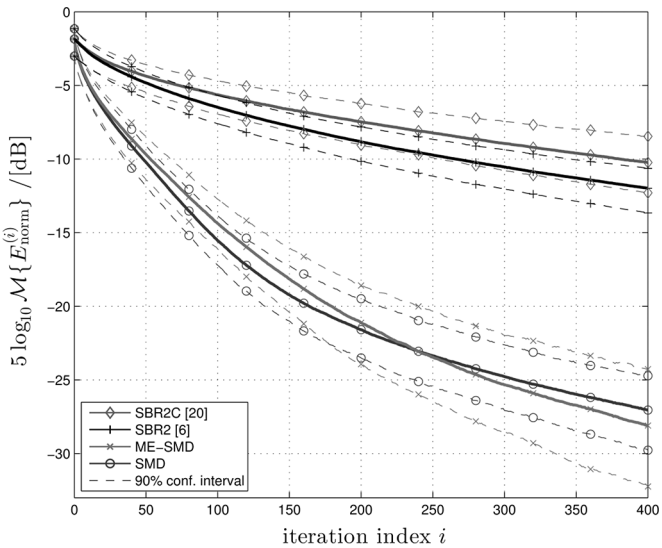


Fig. 2. Ensemble medians for normalized off-diagonal energy $E_{\text{norm}}^{(i)}$ according to (38), and confidence intervals containing 90% of the ensemble probes.

Fig. 2 shows that the two algorithm groups (SMD and SBR2) indeed behave quite differently. Of the SMD algorithms, ME-SMD, with its slightly reduced cost, initially converges slower, but attains a better convergence at higher iteration steps, with very similar ensemble distributions according to quartiles and 5th percentiles. Of the SBR2 algorithms, SBR2 minimizes the dominant off-diagonal element at every step, and so performs better than SBR2C, which optimizes the coding gain instead.

The normalized coding gain for algorithms operating on the ensemble for the scenario are characterized in Fig. 3. Ultimately, all algorithms asymptotically approach a normalized coding gain of unity. Interestingly, the proposed SMD algorithms converge significantly faster than SBR2C [20]. Notice that the performance of SBR2 and SBR2C reverses when considering the ensemble average normalized coding gain, since this metric matches the cost function optimized by SBR2C [20].

D. Spectral Majorization

Spectral majorization, unlike diagonalisation, was not proven for the iterative PEVD algorithms in [6], [20] and Section III, but is targeted by the way off-diagonal energy is transferred at every iteration step. Fig. 4 shows the on-diagonal PSDs of $\boldsymbol{S}^{(50)}(z)$ for the various algorithms after $L = 50$ iteration steps when applied to a single ensemble probe of the scenario. For simplicity of the graphs, only the first four of the eight channels are shown, with the ground truth underlaid in grey. Results from both SMD and ME-SMD are spectrally majorized, with deviations from the ground truth spectrum only for low-power bands.

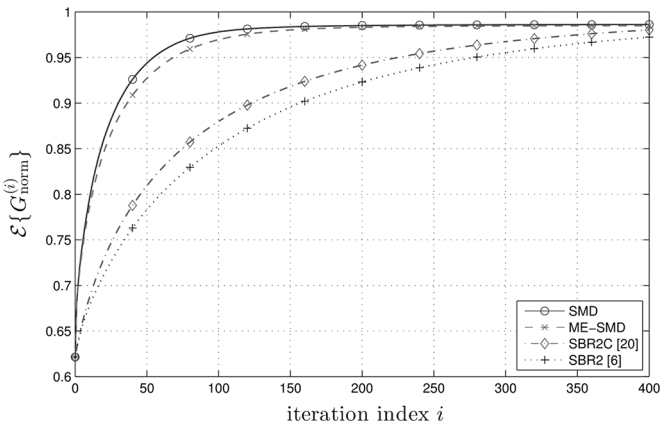


Fig. 3. Ensemble average of the normalized coding gain $G_{\mathrm{norm}}^{(i)}$ according to (42) versus iteration index i for the scenario considered.

In comparison, SBR2 and SBR2C have not fully achieved spectral majorization yet and show some deviations even for higher powered bands.

The PSDs of diagonal elements of $\boldsymbol{S}^{(200)}(z)$ after $L = 200$ iterations are shown in Fig. 5. Most algorithms have converged to a spectrally majorized solution that closely matches the ground truth, shown by the spectrum underlaid in grey. Notice that, for ME-SMD, SBR2 and SBR2C, the lowest subbands are not spectrally majorized in Figs. 5(b), (c) and (d), respectively. In contrast, the SMD algorithm produces PSDs that satisfy the spectral majorization property.

E. Computational Complexity

To assess the complexity of calculating the above decompositions for the given scenario, the ensemble-averaged normalized off-diagonal power $E_{\mathrm{norm}}^{(i)}$ is shown versus the elapsed system time T_i in Fig. 6. The ensemble was simulated on a cluster of desktop personal computers each with Intel Dual-Core 3.20 GHz processor and 4 GB RAM. This graph is obtained by recording both normalized off-diagonal energy $E_{\mathrm{norm}}^{(i)}$ and elapsed system time T_i as a function of the iteration index i, which then permits to relate $E_{\mathrm{norm}}^{(i)}$ to T_i.

It is evident that the SMD type algorithms have a considerably higher computational complexity than the SBR2-family counterparts, which is due to the necessity to apply a full unitary matrix for every lag τ of $\mathbf{S}^{(i)}[\tau]$ rather than just a simple Givens rotation in the case of SBR2, which never involves the processing of more than two rows and columns. The extra cost of the SMD algorithms goes towards unlocking performance regions in terms of reduction of $E_{\mathrm{norm}}^{(i)}$ that are inaccessible to SBR2-type algorithms.

The ME-SMD algorithm has been motivated in Section III.B as an alternative to SMD with a somewhat reduced cost. In Fig. 6(a) this reduction is not directly evident; in parts this may be as the implementations are not optimized w.r.t. the simulation environment or processor platform. Another reason why ME-SMD does not show a consistent reduction in complexity of SMD can be justified from the order of the extracted PU filter banks, which will be discussed next.

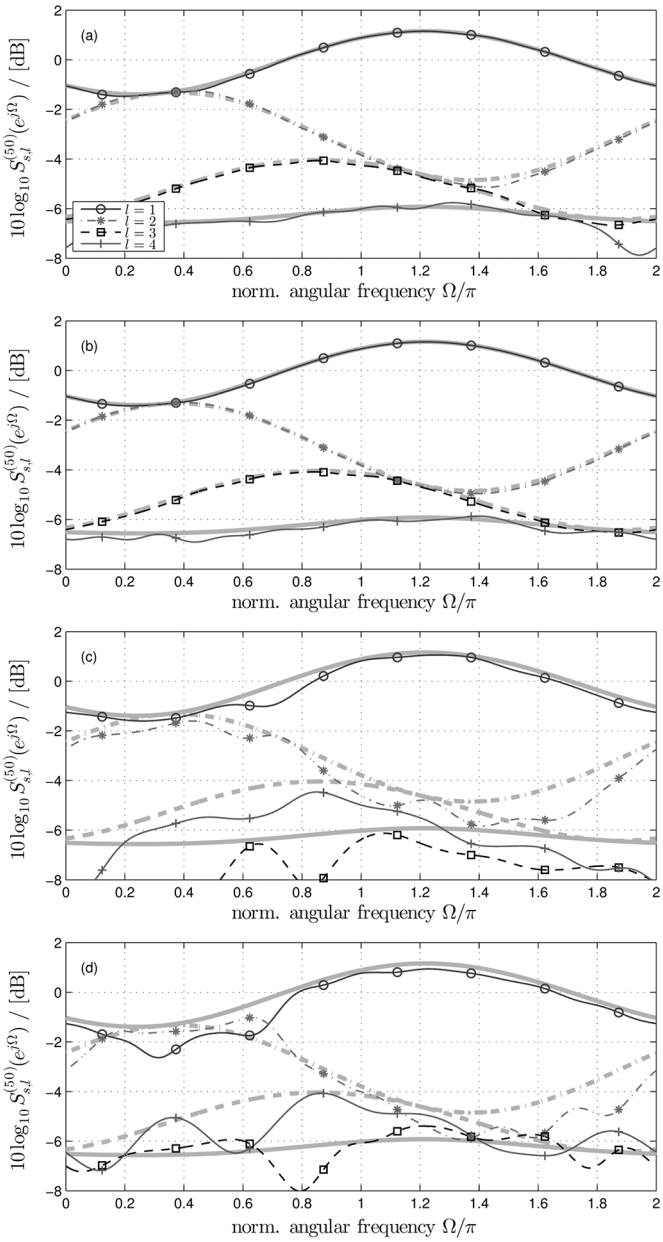


Fig. 4. PSDs of the first four on-diagonal polynomials of $\boldsymbol{S}^{(50)}(z)$ after 50 iterations with (a) SMD, (b) ME-SMD, (c) SBR2 [6] and (d) SBR2C [20], applied to an ensemble probe of the specified scenario, with ideal PSDs underlaid in light shading.

F. Application Cost

Once calculated, the cost of applying the decompositions reached by different algorithms relates directly to the order of the extracted PU matrix $\boldsymbol{H}(z)$ that, for example, can be applied as the processor in Fig. 1. Therefore, ensemble medians for the normalized off-diagonal energy $E_{\mathrm{norm}}^{(i)}$ versus the order of the PU filter banks required to achieve this decomposition are shown in Fig. 6(b). The curves are obtained by recording the PU matrix order at each iteration i, which is then related to the corresponding normalized off-diagonal energy $E_{\mathrm{norm}}^{(i)}$. Note the high orders observed for the PU filter banks, which is due to the high order of the CSD matrix $\boldsymbol{R}(z)$ to be decomposed.

The SMD algorithms offer a consistently lower cost for applying the PU matrix compared to the SBR2 family of

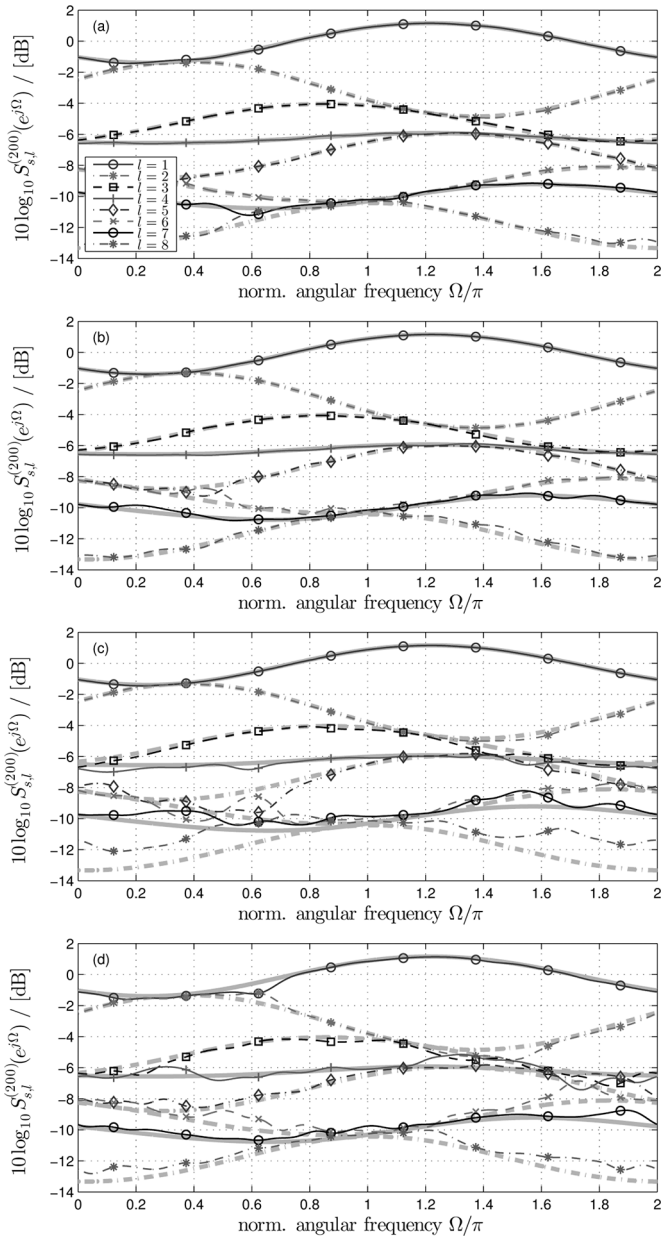


Fig. 5. PSDs of on-diagonal polynomials of $\boldsymbol{S}^{(200)}(z)$ after 200 iterations with (a) SMD, (b) ME-SMD, (c) SBR2 [6] and (d) SBR2C [20], applied to an ensemble probe of the scenario considered, with ideal PSDs underlaid in light shading.

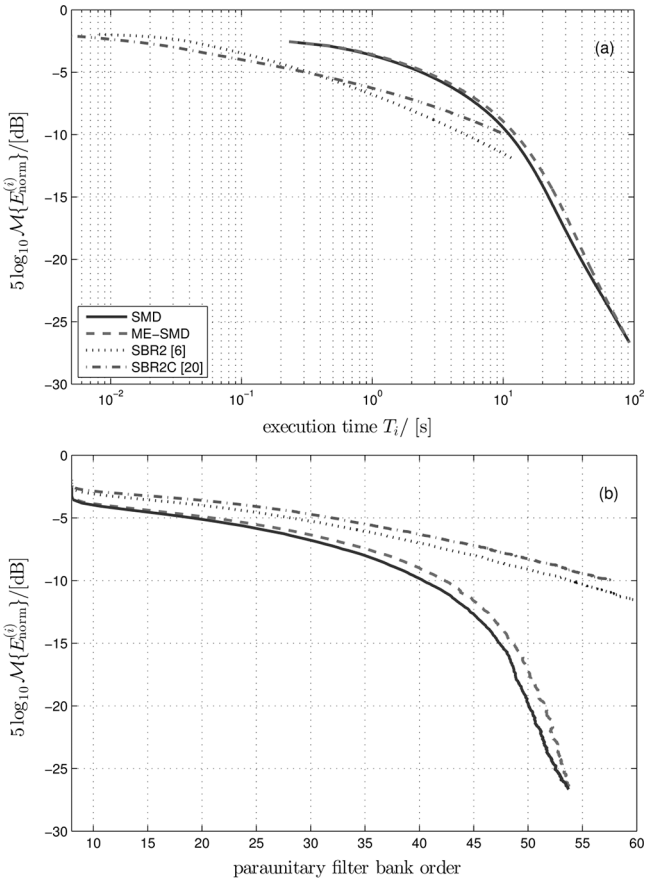


Fig. 6. Ensemble median normalized off-diagonal energy $E^{(i)}_{\text{norm}}$ versus (a) ensemble median elapsed system time and (b) PU filter bank order.

algorithms. Particularly for suppression of off-diagonal energy below -15 dB, SMD algorithms attain this performance with a reasonable order compared to SBR2-type algorithms, which are unable to reach this performance region. Comparing SMD and ME-SMD algorithms, the ME-SMD version requires on average a slightly higher order, which in turn means a higher complexity per iteration step, as a unitary matrix has to be applied at every lag τ of $\mathbf{S}^{(i)}[\tau]$.

VI. Conclusion

Different from previous iterative PEVD algorithms, which only eliminate the maximum (SBR2) or normalized maximum (SBR2C) off-diagonal element at every iteration step, we have proposed a new class of algorithms, termed sequential matrix diagonalisation (SMD), that clears all off-diagonals of the zero-lag matrix. As a result, more energy is transferred onto the main diagonal per iteration, leading to a significantly faster convergence in terms of normalized off-diagonal energy.

However, since the unitary matrix that re-diagonalizes the zero-lag matrix is no longer sparse, its application to matrices at all lags significantly increases complexity. Without any implementation tricks, the algorithm is significantly more complex than the SBR2 family. However, two interesting and important advantages of the SMD algorithms have been demonstrated:

1) SMD algorithms can, within a reasonable number of iteration steps, attain a suppression of off-diagonal energy that previous algorithms were not capable of delivering;
2) as simulations have demonstrated, SMD-based decompositions are achieved with significantly shorter paraunitary filter banks, which are less costly to apply.

In terms of applications, these advantages are expected to bring a significant impact to subspace-based methods such as [17], [18], where enhanced diagonalisation will achieve a better separation of subspaces, while subspace projections, such as for the generator and parity check polynomial matrices in [10], can be performed with lower order paraunitary filters at reduced computational cost.

References

[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York, NY, USA: McGraw-Hill, 1991.

[2] A. Pauraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[3] R. Klemm, "Space-time adaptive processing principles and applications," in *Space-time Adaptive Processing Principles and Applications*, ser. IEE Radar, Sonar, Navigation Avionics, ser. 9. London, U.K.: IEE, 1998.

[4] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1980.

[5] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[6] J. G. McWhirter, P. D. Baxter, T. Cooper, S. Redif, and J. Foster, "An EVD algorithm for para-Hermitian polynomial matrices," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2158–2169, May 2007.

[7] S. Icart and P. Comon, "Some properties of Laurent polynomial matrices," presented at the 9th IMA Conf. Math. Signal Process., Birmingham, U.K., Dec. 2012.

[8] F. Labeau, J.-C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Macq, "Oversampled filter banks as error correcting codes: Theory and impulse noise correction," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4619–4630, Dec. 2005.

[9] S. Weiss, "On the design of oversampled filter banks for channel coding," in *Proc. Eur. Signal Process. Conf.*, Vienna, Austria, Sep. 2004, pp. 885–888.

[10] S. Weiss, S. Redif, T. Cooper, C. Liu, P. D. Baxter, and J. G. McWhirter, "Paraunitary oversampled filter bank design for channel coding," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–10, 2006, Article ID 31346; doi:10.1155/ASP/2006/31346.

[11] R. Brandt and M. Bengtsson, "Wideband MIMO channel diagonalization in the time domain," in *Proc. Int. Symp. Pers., Indoor, Mobile Radio Commun.*, 2011, pp. 1914–1918.

[12] C. H. Ta and S. Weiss, "A design of precoding and equalisation for broadband MIMO systems," in *Proc. 41st Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2007, pp. 1616–1620.

[13] N. Moret, A. Tonello, and S. Weiss, "MIMO precoding for filter bank modulation systems based on PSVD," in *Proc. IEEE 73rd Veh. Technol. Conf.*, Budapest, Hungary, May 2011, pp. 1–5.

[14] C. H. Ta and S. Weiss, "A jointly optimal precoder and block decision feedback equaliser design with low redundancy," in *Proc. 15th Eur. Signal Process. Conf.*, Poznan, Poland, Sep. 2007, pp. 489–492.

[15] J. Foster, J. G. McWhirter, S. Lambotharan, I. Proudler, M. Davies, and J. Chambers, "Polynomial matrix QR decomposition for the decoding of frequency selective multiple-input multiple-output communication channels," *IET Signal Process.*, vol. 6, no. 7, pp. 704–712, Sep. 2012.

[16] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 86–105, Sep. 2007.

[17] S. Weiss, M. Alrmah, S. Lambotharan, J. G. McWhirter, and M. Kaveh, "Broadband angle of arrival estimation methods in a polynomial matrix decomposition framework," presented at the 5th IEEE Int. Workshop Comp. Advances Multi-Sens. Adapt. Process., St. Martin, Dec. 2013.

[18] S. Redif, J. G. McWhirter, P. D. Baxter, and T. Cooper, "Robust broadband adaptive beamforming via polynomial eigenvalues," in *Proc. IEEE/MTS OCEANS*, Boston, MA, Sep. 2006, pp. 1–6.

[19] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1528–1543, Jun. 1998.

[20] S. Redif, J. G. McWhirter, and S. Weiss, "Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5253–5264, Nov. 2011.

[21] P. A. Regalia and D.-Y. Huang, "Attainable error bounds in multirate adaptive lossless FIR filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 1995, vol. 2, pp. 1460–1463.

[22] R. H. Lambert, M. Joho, and H. Mathis, "Polynomial singular values for number of wideband source estimation and principal components analysis," in *Proc. Int. Conf. Independ. Compon. Anal.*, San Diego, CA, Dec. 2001, pp. 379–383.

[23] M. Tohidian, H. Amindavar, and A. M. Reza, "A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices," *EURASIP J. Adv. Signal Process.*, vol. 1, pp. 1–16, 2013.

[24] A. Tkacenko, "Approximate eigenvalue decomposition of para-Hermitian systems through successive FIR paraunitary transformations," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Dallas, TX, USA, Mar. 2010, pp. 4074–4077.

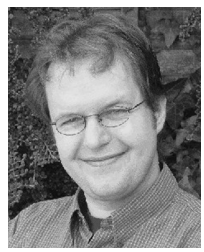
[25] S. Redif, S. Weiss, and J. G. McWhirter, "An approximate polynomial matrix eigenvalue decomposition algorithm for para-Hermitian matrices," in *Proc. 11th IEEE Int. Symp. Signal Process. Inf. Technol.*, Bilbao, Spain, Dec. 2011, pp. 421–425.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.

Soydan Redif (A'05–M'09–SM'12) received the B.Eng. degree (first class honors) in electronic engineering from Middlesex University, London, U.K., in 1998, and the Ph.D. degree in electrical and electronic engineering from the University of Southampton, Southampton, U.K., in 2006.

He joined DERA, working initially on airborne SHF SATCOM systems in 1999. From 2000 to 2007, he was with the Malvern Signal Processing Group at QinetiQ, Malvern, U.K., where his research focused on adaptive signal processing algorithms. From 2008 to 2011, he was an Assistant Professor with Near East University, Nicosia, Cyprus. In 2011, he was a Visiting Lecturer with the European University of Lefke (EUL), Gemikonagi, Cyprus. He has been with EUL since 2011, and is currently the Head of the Department of Electrical and Electronic Engineering. His current research interests include adaptive signal processing, broadband sensor arrays, convolutive blind signal separation, polynomial matrix techniques, and power systems.

Dr. Redif is a U.K. Chartered Engineer, and a member of the IET. He was the recipient of the IEE Award for Outstanding Academic Achievement in 1998.



Stephan Weiss (S'97–A'98–M'99–SM'05) received the Dipl.-Ing. degree from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1995, and the Ph.D. degree from the University of Strathclyde, Glasgow, Scotland, in 1998, both in electronic and electrical engineering.

He is currently a Reader in the Department of EEE at Strathclyde. From 1999 until 2006, he held lecturer and senior lecturer positions within the School of Electronics and Computer Science at the University of Southampton. Prior to this, he was a Visiting Lecturer at the University of Strathclyde during 1998–1999 and a Visiting Scholar at the University of Southern California during 1996–1997. His research interests lie in adaptive, multirate, and array signal processing with applications in communications, audio, and biomedical signal processing, where he has published more than 230 technical papers.

Dr. Weiss is a member of EURASIP. He was the Technical Co-Chair for EUSIPCO 2009, and General Chair for ISPLC 2014 in Glasgow. For his work in biomedical signal processing, he was a co-recipient of the 2001 Research Award of the German society on hearing aids. In 2011, he was a co-recipient of the VTC-Spring Best Paper award in the MIMO systems track.

John G. McWhirter received First Class honors degree in mathematics and the Ph.D. degree in theoretical physics from the Queen's University of Belfast, Belfast, Ireland, in 1970 and 1973, respectively.

He joined the Royal Radar Establishment in Malvern (later to become the Royal Signals and Radar Establishment, and now part of QinetiQ Ltd.) in 1973, where he became a Senior Fellow in the Centre for Signal and Information Processing Group. In 2007, he left QinetiQ to take up his current post as Distinguished Research Professor in the Engineering Department, Cardiff University, Cardiff, U.K. He is also a Visiting Professor in Electrical Engineering at the Queen's University of Belfast. He has been carrying out research on adaptive signal processing since 1980. He has published more than 140 research papers and holds numerous patents. His current research is devoted to broadband sensor arrays, convolutive blind signal separation, and polynomial matrix techniques.

Dr. McWhirter was elected as a Fellow of the Royal Academy of Engineering in 1996 and as a Fellow of the Royal Society in 1999. He is a Fellow of the Institute of Mathematics and its Applications (IMA) and served as President of the IMA in 2002 and 2003. He is also a Fellow of the Institute of Electrical Engineers, a Fellow of the Institute of Physics, and a member of the London Mathematical Society. The signal processing group which he built up in Malvern over many years, received the EURASIP Group Technical Achievement Award for 2003. He was awarded the J. J. Thomson Medal by the Institution of Electrical Engineers in 1994 for his research on systolic arrays.