Impact of Space-Time Covariance Estimation Errors on a Parahermitian Matrix EVD

Connor Delaosa[1], Fraser K. Coutts[1], Jennifer Pestana[2], Stephan Weiss[1]
[1]Dept. of EEE, [2]Dept. of Maths & Stats, University of Strathclyde
Email: {connor.delaosa,fraser.coutts,jennifer.pestana,stephan.weiss}@strath.ac.uk

Abstract—This paper studies the impact of estimation errors in the sample space-time covariance matrix on its parahermitian matrix eigenvalue decomposition. We provide theoretical bounds for the perturbation of the ground-truth eigenvalues and of the subspaces of their corresponding eigenvectors. We show that for the eigenvalues, the perturbation depends on the norm of the estimation error in the space-time covariance matrix, while the perturbation of eigenvector subspaces can additionally be influenced by the distance between the eigenvalues. We confirm these theoretical results by simulations.

Index Terms—broadband array processing; space-time covariance estimation; parahermitian matrix; eigenvalue decomposition.

I. INTRODUCTION

In the field of broadband array processing, polynomial matrix algebra can extend the relative ease of formulating narrowband problems to the broadband case [12], [26]. Optimal solutions to such problems then often depend on extending the utility of eigen- or singular value decompositions to the polynomial case [27]. A number of practiable algorithms such as the 2nd order sequential best rotation (SBR2) [12], [17] or the sequential matrix diagonalisation (SMD) algorithms [4], [18] have been designed to iteratively calculate approximate solutions for these decomposition.

Algorithms such as SBR2 and SMD in turn have enabled a wide field of applications ranging from broadband MIMO systems [22], to beamforming [16], [26], source separation [19] or angle of arrival estimation [25]. The accuracy of e.g. broadband subspace decompositions has been investigated in [1], [6] in order to compare the performance of such algorithms. Known limiting factors for this accuracy are due to the conditioning of the underlying source model [3] as well as algorithm-internal order reductions [4], [5], [7], [23].

In most applications, the cross-spectra density matrix is not readily available but must be estimated from a finite set of data, which results in estimation errors. Therefore, this paper aims to investigate what impact such estimation errors have on the eigenvalues and eigenvectors of a parahermitian matrix. Since we are interested in isolating the errors due to finite sample size from any inaccuracies that arise from iterative algorithms, we here restrict ourselves to the extraction of analytic eigenvalues and vectors from such a parahermitian matrix, which are guaranteed to exist [27], and will restrict our analysis to the unit circle only.

We will first define space-time covariance matrices and their associated CSD matrices in Sec. II. Based in the definition
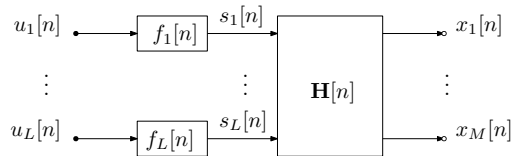
of parahermitian matrix EVD in Sec. III, we then exploit perturbation theory to analyse the problem in Sec. IV. The results are confirmed by simulations in Sec. V.

Fig. 1. Source model with L uncorrelated unit variance Gaussian sources $u_\ell[n]$, $\ell = 1 \dots L$ and a convolutive MIMO mixing matrix $\mathbf{H}[n] \in \mathbb{C}^{M \times L}$ generating M measurements $x_m[n]$, $m = 1 \dots M$.

II. DATA MODEL

A. Space-Time Covariance Matrix

We assume that an M-element array records data into a vector $\mathbf{x}[n] \in \mathbb{C}^M$ with discrete time index n. For narrowband processing, where phase shifts experienced between array elements bear information e.g. on the angle of arrival, it is sufficient to consider the second order statistics contained in the instantaneous covariance matrix $\mathbf{R} = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^H[n]\}$. For broadband processing, where explicit delays τ between array elements rather than just phase shifts must be considered, the second order statistics are captured by a space-time covariance matrix $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^H[n - \tau]\}$. Its z-transform, the cross-spectral density (CSD) $\boldsymbol{R}(z) = \sum_\tau \mathbf{R}[\tau]z^{-\tau}$ — short $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ \mathbf{R}[\tau]$ — contains functions in z as entries.

Since $\mathbf{R}[\tau]$ contains auto- and cross-correlation sequences, it inherits certain symmetries, namely that $\mathbf{R}[\tau] = \mathbf{R}^H[-\tau]$. This implies that the CSD matrix $\boldsymbol{R}(z)$ is a parahermitian matrix such that $\boldsymbol{R}^P(z) = \boldsymbol{R}^H(1/z^*)$, and thus is identical to its time-reversed Hermitian transposition.

B. Source Model

To describe the ground truth of the space-time covariance matrix, we use the source model [18] shown in Fig. 1, which assumes that the data vector $\mathbf{x}[n]$ is generated by convolutively mixing L independent source signals. The individual power spectral densities (PSD) of these source signals $s_\ell[n]$, $\ell = 1 \dots L$, are $S_\ell(z)$, which can be tied to uncorrelated zero mean, unit variance Gaussian signals $u_\ell[n]$ via the innovation filters [15] $f_\ell[n] \circ\!\!-\!\!\bullet F_\ell(z)$, such that $S_\ell(z) = F_\ell(z)F_\ell^P(z)$.

With the convolutive mixing system described by a network of transfer functions $\boldsymbol{H}(z) : \mathbb{C} \to \mathbb{C}^{M \times L}$ •—○ $\mathbf{H}[n]$, the CSD matrix of $\mathbf{x}[n]$ is

$$\boldsymbol{R}(z) = \boldsymbol{H}(z)\boldsymbol{S}(z)\boldsymbol{H}^{\mathrm{P}}(z) \, , \qquad (1)$$

where the diagonal $\boldsymbol{S}(z)$ contains the source PSDs

$$\boldsymbol{S}(z) = \mathrm{diag}\{S_1(z) \ \ S_2(z) \ \ \ldots \ \ S_L(z)\} \, . \qquad (2)$$

C. Sample Space-Time Covariance Matrix and Estimation Errors

We now investigate the effect of determining an estimate $\hat{\boldsymbol{R}}(z)$ of $\boldsymbol{R}(z)$ from N snapshots for data. W.r.t. Fig. 1, these snapshots can be represented as vectors $\mathbf{x}[n] \in \mathbb{C}^M$, $n = 0 \ldots (N - 1)$. The estimation of $\mathbf{R}[\tau]$ is based on these samples,

$$\hat{\mathbf{R}}[\tau] = \frac{1}{N - \tau} \sum_{n=\tau}^{N-1} \mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n - \tau] \, , \ \forall 0 \leq \tau < N \, , \qquad (3)$$

with values for $\tau < 0$ obtainable through $\hat{\mathbf{R}}[\tau] = \hat{\mathbf{R}}^{\mathrm{H}}[-\tau]$. The approach in (3) assumes ergodicity, which implies that $\lim_{N \to \infty} \hat{\boldsymbol{R}}(z) = \boldsymbol{R}(z)$.

We assume that $N > M$, so that there is a chance that the rank of $\hat{\boldsymbol{R}}(z)$ matches that of $\boldsymbol{R}(z)$. Further, we assume that $N \gg \tau_{\mathrm{max}}$, where τ_{max} is the maximum lag for which we want to evaluate $\hat{\mathbf{R}}[\tau]$, to minimise the impact of shorter time series on the estimation as $\tau \to \pm\tau_{\mathrm{max}}$.

Determining the distribution of the sample space-time covariance matrix $\hat{\mathbf{R}}[\tau]$ or $\hat{\boldsymbol{R}}(z)$ is challenging. In the case where $[x][n]$ is only spatially but not temporally correlated, the instantanious sample covariance matrix $\hat{\mathbf{R}}[0]$ is Wishart-distributed [11], [29]. For such a matrix, the distribution of elements depends both on the size of the sample set, N, and also the entries in the ground truth $\mathbf{R}[0]$. If $[x][n]$ is temporally correlated, which is the case here due to the filters $f_\ell[n]$ and the convolutive nature of $\boldsymbol{H}(z)$, then the information contained in the sample set $[x][n], n = 0 \ldots (N - 1)$ is reduced. This has been well-described for the auto-correlation of simple time series that are generated by a first order auto-regressive innovation filters [2], [8], [20].

For the potentially complicated spatial and temporal correlations expressed in $\mathbf{R}[\tau]$, we can qualitatively state that the entries will depend on the values in $\mathbf{R}[\tau]$, as well as some fraction (due to the temporal correlation, and potentially different for each entry) of the sample size N. However, we have been unable to find a description of the exact distribution of $\hat{\mathbf{R}}[\tau]$, and therefore will operate based on the estimation error

$$\mathbf{E}[\tau] = \hat{\mathbf{R}}[\tau] - \mathbf{R}[\tau] \qquad (4)$$

and its norms independent of any particular distribution of $\mathbf{E}[\tau]$. Note that we have

$$\lim_{N \longrightarrow \infty} \mathbf{E}[\tau] = \mathbf{0} \, , \ \ \forall \tau \qquad (5)$$

based on erodicity of $\hat{\mathbf{R}}[\tau]$ and the unbiased estimate in (3).

III. PARAHERMITIAN MATRIX EVD

The parahermitian matrix EVD of the matrix $\mathbf{R}(z)$ is the decomposition

$$\mathbf{R}(z) = \mathbf{U}(z)\boldsymbol{\Lambda}(z)\mathbf{U}^{\mathrm{P}}(z) \, , \qquad (6)$$

where $\mathbf{U}(z)$ is a paraunitary matrix of eigenvectors and $\boldsymbol{\Lambda}(z)$ is a diagonal parahermitian matrix of eigenvalues. We are concerned with how well the matrix EVD of $\mathbf{R}(z)$ matches that of $\hat{\mathbf{R}}(z) = \hat{\mathbf{U}}(z)\hat{\boldsymbol{\Lambda}}(z)\hat{\mathbf{U}}^{\mathrm{P}}(z)$. We note that if the mixing matrix $\boldsymbol{H}(z)$ of the source model in Fig. 1 is restricted to be paraunitary, then the factorisation on the r.h.s. of (1) represents the parahermitian matrix EVD of $\mathbf{R}(z)$ in (6).

If $\mathbf{R}(z)$ is based on stable and causal system components in Fig. 1 and (1), it will be analytic within an annulus containing the unit circle. Since $\hat{\mathbf{R}}(z)$ is based on a finite-length sequence, it will also be analytic. It is shown in [27] that for such $\mathbf{R}(z)$ and $\hat{\mathbf{R}}(z)$, it is possible to ensure that the parahermitian matrix EVD factors $\mathbf{U}(z)$, $\boldsymbol{\Lambda}(z)$, $\hat{\mathbf{U}}(z)$ and $\hat{\boldsymbol{\Lambda}}(z)$ are also analytic. In this case, the eigenvalues in $\boldsymbol{\Lambda}(z)$ and $\hat{\boldsymbol{\Lambda}}(z)$ are unique, while the eigenvectors in the columns of $\mathbf{U}(z)$ and $\hat{\mathbf{U}}(z)$ can be multiplies with arbitrary allpass filters. Note, however, that analyticity of the factors is lost when eigenvalues cross, but are spectrally majorised, a property that is enforced by a number of algorithms [12], [13], [17], [18] that approximate (6).

IV. PERTURBATION OF EIGENVALUES AND EIGENVECTORS

In the following, we explore how an estimation error $\boldsymbol{E}(z)$ in (4) impacts on the parahermitian matrix EVD in (6), i.e. how much the factors $\hat{\mathbf{U}}(z)$ and $\hat{\boldsymbol{\Lambda}}(z)$ of the sample CSD matrix $\hat{\boldsymbol{R}}(z)$ deviate from the ground truth $\mathbf{U}(z)$ and $\boldsymbol{\Lambda}(z)$. For this analysis, we concentrate on the behaviour on the unit circle, i.e. for $z = \mathrm{e}^{\mathrm{j}\Omega}$, and utilise a number of results from matrix perturbation theory [10], [21] to first investigate the eigenvalues in Sec. IV-A, and then the eigenvectors in Sec. IV-B.

A. Impact of Estimation Error on Eigenvalues

At any frequency Ω_0, evaluation of (4) at $z = \mathrm{e}^{\mathrm{j}\Omega_0}$ on unit circle gives $\boldsymbol{E}(\mathrm{e}^{\mathrm{j}\Omega_0}) = \hat{\boldsymbol{R}}(\mathrm{e}^{\mathrm{j}\Omega_0}) - \boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_0})$ so that, by Weyl's theorem [21] we have the following bounds on the perturbation of the mth eigenvalue $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0})$ of $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_0})$,

$$\lambda_{\mathrm{min}}(\mathrm{e}^{\mathrm{j}\Omega_0}) + \lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \leq \hat{\lambda}_m(\mathrm{e}^{\mathrm{j}\Omega_0})$$
$$\hat{\lambda}_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \leq \lambda_{\mathrm{max}}(\mathrm{e}^{\mathrm{j}\Omega_0}) + \lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \, ,$$

where $\hat{\lambda}_m(\mathrm{e}^{\mathrm{j}\Omega})$ is the mth eigenvalue of $\hat{\boldsymbol{R}}(\mathrm{e}^{\mathrm{j}\Omega_0})$, and $\lambda_{\mathrm{min}}(\mathrm{e}^{\mathrm{j}\Omega_0})$ and $\lambda_{\mathrm{max}}(\mathrm{e}^{\mathrm{j}\Omega})$ are the minimum and maximum eigenvalue of $\boldsymbol{E}(\mathrm{e}^{\mathrm{j}\Omega_0})$. After reshuffling,

$$\lambda_{\mathrm{min}}(\mathrm{e}^{\mathrm{j}\Omega_0}) \leq \hat{\lambda}_m(\mathrm{e}^{\mathrm{j}\Omega_0}) - \lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \leq \lambda_{\mathrm{max}}(\mathrm{e}^{\mathrm{j}\Omega_0}) \qquad (7)$$

provides bounds for the change between the ground-truth eigenvalues $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0})$ and the sample estimate-based $\hat{\lambda}_m(\mathrm{e}^{\mathrm{j}\Omega_0})$.

Alternatively, the Hoffman-Wielandt theorem states that for all M eigenvalues

$$\sum_{m=1}^{M} \left(\hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0}) \right)^2 \leq \|\boldsymbol{E}(e^{j\Omega_0})\|_F^2 \qquad (8)$$

holds. Further, the Bauer-Fike theorem [28] guarantees that

$$|\hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0})| \leq \kappa\{\hat{\boldsymbol{U}}(e^{j\Omega_0})\}\|\boldsymbol{E}(e^{j\Omega_0})\|_2, \qquad (9)$$

with $\kappa\{\mathbf{A}\}$ the condition number of the matrix $\mathbf{A}$. Since $\hat{\boldsymbol{U}}(e^{j\Omega_0})$ is unitary by definition, $\kappa\{\hat{\boldsymbol{U}}(e^{j\Omega_0})\} = 1$, and (9) simplifies further.

Overall, both the bounds (7) and (9) relate the deviation between the ground-truth eigenvalues and their sample estimate-based counterparts directly to the estimation error $\boldsymbol{E}(e^{j\Omega})$, but are independent of the absolute size and relative distance between eigenvalues. In particular, because of (5) we have $\lim_{N \to \infty} \hat{\lambda}_m(e^{j\Omega_0}) = \lambda_m(e^{j\Omega_0})$.

B. Impact of Estimation Error on Eigenvectors

Assume that the mth eigenvectors corresponding to the mth eigenvalues $\lambda_m(e^{j\Omega})$ and $\hat{\lambda}_m(e^{j\Omega})$ are $\mathbf{u}_m(e^{j\Omega})$ and $\hat{\mathbf{u}}_m(e^{j\Omega})$, respectively. Assume further that for the sample set N, $\lim_{N \to \infty} \hat{\lambda}_m(e^{j\Omega}) = \lambda_m(e^{j\Omega})$, $m = 1 \ldots M$. Due to the phase ambiguity of eigenvectors, a robust comparison between $\mathbf{u}_m(e^{j\Omega})$ and $\hat{\mathbf{u}}_m(e^{j\Omega})$ must be based on subspace angles or -correlations [9], [14]. This is even more pivotal if eigenvalues have a C-fold algebraic multiplicity, such that e.g. $\lambda_m(e^{j\Omega}) = \ldots \lambda_{m+C-1}(e^{j\Omega})$ at a particular frequency Ω, since in this case the corresponding eigenvectors can form any orthonormal basis within a C-dimensional subspace. In the vicinity of such an algebraic multiplicity, eigenvectors can be sensitive, while the subspace in which they are contained remains invariant. We therefore focus on the subspaces in which eigenvectors of $\boldsymbol{R}(e^{j\Omega})$ and $\hat{\boldsymbol{R}}(e^{j\Omega})$ exist.

To analyse the subspace of eigenvectors belonging to a C-fold algebraic multiplicity of eigenvalues at a frequency Ω_0, we reorder both eigenvalues and eigenvectors as follows. We let $\boldsymbol{\Lambda}(e^{j\Omega_0}) = \text{blockdiag}\{\boldsymbol{\Lambda}_1(e^{j\Omega_0}), \boldsymbol{\Lambda}_2(e^{j\Omega_0})\}$, where $\boldsymbol{\Lambda}_1(e^{j\Omega_0}) \in \mathbb{R}^{C \times C}$ contains the multiple eigenvalues, and $\boldsymbol{\Lambda}_2(e^{j\Omega_0})$ the remaining $M - C$ eigenvalues. We similarly partition $\boldsymbol{U}(e^{j\Omega_0}) = [\boldsymbol{U}_1(e^{j\Omega_0}), \boldsymbol{U}_2(e^{j\Omega_0})]$, such that $\mathcal{U}_1(e^{j\Omega_0}) = \text{range}\{\boldsymbol{U}_1(e^{j\Omega_0})\}$ is the subspace containing the eigenvectors corresponding to the C multiple eigenvalues in $\boldsymbol{\Lambda}_1(e^{j\Omega_0})$. We now want to measure the distance between $\mathcal{U}_1(e^{j\Omega_0})$, and the subspace $\hat{\mathcal{U}}_1(e^{j\Omega_0})$ of the corresponding eigenvectors of $\hat{\boldsymbol{R}}(e^{j\Omega_0})$.

A suitable metric for the distance between two subspaces is based on the projection operators $\boldsymbol{P}_1(e^{j\Omega_0}) = \boldsymbol{U}_1(e^{j\Omega_0})\boldsymbol{U}_1^H(e^{j\Omega_0})$ and the similarly defined $\hat{\boldsymbol{P}}_1(e^{j\Omega_0})$ based on the estimated eigenvectors. In this case [9],

$$\text{dist}\{\mathcal{U}_1(e^{j\Omega_0}), \hat{\mathcal{U}}_1(e^{j\Omega_0})\} = \|\boldsymbol{P}_1(e^{j\Omega_0}) - \hat{\boldsymbol{P}}_1(e^{j\Omega_0})\|_2 .$$

Similar to above, we also partition the perturbation

$$\boldsymbol{E}(e^{j\Omega_0}) = \left[\begin{array}{cc} \underbrace{\boldsymbol{E}_{11}(e^{j\Omega_0})}_{C} & \underbrace{\boldsymbol{E}_{21}^H(e^{j\Omega_0})}_{M-C} \\ \boldsymbol{E}_{21}(e^{j\Omega_0}) & \boldsymbol{E}_{22}(e^{j\Omega_0}) \end{array} \right]. \qquad (10)$$

By defining the spectral distance δ between the cluster of C eigenvalues in $\boldsymbol{\Lambda}_1(e^{j\Omega_0})$ and the next-nearest neighbour in $\boldsymbol{\Lambda}_2(e^{j\Omega_0})$,

$$\delta = \min_{\substack{\lambda_1 \in \boldsymbol{\Lambda}_1(e^{j\Omega_0}) \\ \lambda_2 \in \boldsymbol{\Lambda}_2(e^{j\Omega_0})}} |\lambda_1 - \lambda_2| > 0 , \qquad (11)$$

we find that the subspace distance in (10) is bounded such that

$$\text{dist}\{\mathcal{U}_1(e^{j\Omega_0}), \hat{\mathcal{U}}_1(e^{j\Omega_0})\} \leq \frac{4}{\delta} \|\boldsymbol{E}_{21}(e^{j\Omega_0})\|_2 , \qquad (12)$$

as long as the overall perturbation is limited by $\|\boldsymbol{E}(e^{j\Omega_0})\|_2 < \delta/5$ [9]. This can be satisfied by selecting the sample size N sufficiently large.

The above analysis for assessing the subspace distance between ground truth and estimated eigenvectors can be applied for all $\Omega = \Omega_0$ and in turn for all eigenvectors. Similarly to the perturbation of eigenvalues, the perturbation of eigenvector subspaces depends on the estimation error, measured here by $\|\boldsymbol{E}_{21}(e^{j\Omega_0})\|_2 < \|\boldsymbol{E}(e^{j\Omega_0})\|_2$. However, the mismatch between the ground truth and the estimated subspaces will also depend on the distance between the associated eigenvalues: the closer eigenvalues are located, the more perturbed the subspaces of individual associated eigenvectors can become.

V. SIMULATIONS AND RESULTS

We use two scenarios with reference to the source model in Fig. 1, and a number of metric to underpin the above analysis. *Model 1.* For $M = L = 2$, the innovation filters $F_1(z) = 1 + z^{-1}$ and $F_2 = 1 + jz^{-1}$ are followed by an elementary paraunitary mixing matrix $\boldsymbol{H}(z) = \mathbf{I} - \mathbf{v}\mathbf{v}^H + z^{-1}\mathbf{v}\mathbf{v}^H$ with $\mathbf{v} = [1, \ j]^T/\sqrt{2}$ [24]. Evaluated on the unit circle, the ground truth eigenvalues of $\boldsymbol{R}(z)$ are the source PSDs $\lambda_1(e^{j\Omega}) = 2 + \cos\Omega$ and $\lambda_2(e^{j\Omega}) = 2 + \sin\Omega$, which cross at $\Omega = \frac{\pi}{4}$ and $\Omega = \frac{5\pi}{4}$.
Model 2. A controlled set of innovation filters and a convolutive mixing matrix, each of order 10, creates a system of $L = 3$ sources, whos PSDs define well-separated (i.e. naturally spectrally majorised) eigenvalues on the unit circle.

Each model can be excited by different instantiations of independent and identically distributed complex Gaussian noise, and estimates $\hat{\mathbf{R}}(z)$ can be calculated from N snapshots of data $\mathbf{x}[n]$, $n = 0 \ldots (N-1)$. By performing the EVD on discrete DFTs bins, we measure a number of samples of $\hat{\lambda}_m(e^{j\Omega_k})$ and $\hat{\mathbf{u}}(e^{j\Omega_k})$ along the frequency axis for $\Omega_k = \frac{2\pi k}{K}$, $k = 0 \ldots (K-1)$, which we can compare to the ground truth.

For Model 1, the distribution of the estimation error metric $\|\boldsymbol{E}(e^{j\Omega})\|_2$ for $N = 10^3$ over an ensemble of 10^5 simulations is shown in Fig. 2, which demonstrates a range of error distributions and hence error powers over frequency. The measured distribution of $\hat{\lambda}_m(e^{j\Omega})$ in Fig. 3 suggests that the deviation from the ground truth depends on $\|\boldsymbol{E}(e^{j\Omega})\|_2$ as established in
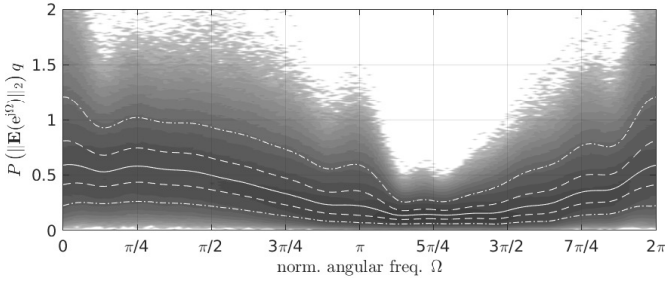
Fig. 2. Distribution of estimation error for Model 1 with 5, 25, 75 and 95th percentiles. The median is shown as a solid line.
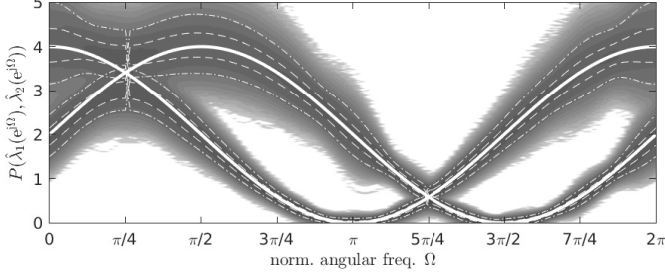

Fig. 3. Distribution of estimated eigenvalues for Model 1, with percentiles as defined in Fig. 2 and the ground truth as solid lines.
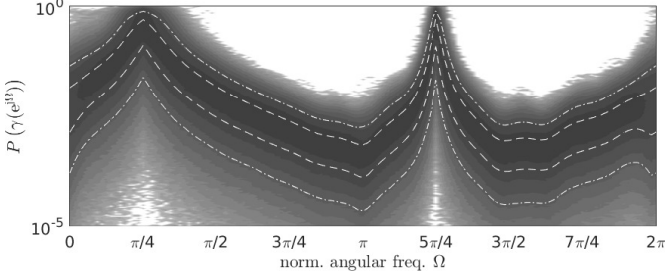

Fig. 4. Distribution of $\gamma_m(\mathrm{e}^{\mathrm{j}\Omega})$ in (13), with percentiles as defined in Fig. 2.

Fig. 5. Distribution of normalised eigenvalue mismatch $\chi_m(\mathrm{e}^{\mathrm{j}\Omega_0})$.
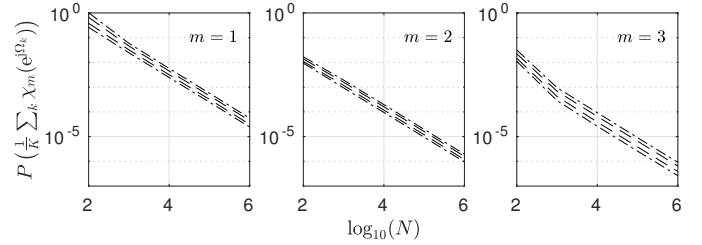
Fig. 6. Measured distribution of total eigenvalue mismatch for Model 2, with percentiles as defined in Fig. 2.

Fig. 7. Measured distribution of total subspace mismatch for Model 2, with percentiles as defined in Fig. 2.

(9), but also on the absolute value of $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$. The assignment of $\hat\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$ to the correct ground truth $\lambda_j(\mathrm{e}^{\mathrm{j}\Omega})$, $m, j = 1, 2$, is difficult in the vicinity of the crossing points of the PSD, leading to some disturbance in the percentiles around these frequencies in Fig. 3.

To measure how close the subspaces spanned by the ground truth and estimated eigenvectors $\boldsymbol{u}_m(\mathrm{e}^{\mathrm{j}\Omega})$ and $\hat{\boldsymbol{u}}_m(\mathrm{e}^{\mathrm{j}\Omega})$ are, we use a modified version of the subspace correlation,

$$\gamma_m(\mathrm{e}^{\mathrm{j}\Omega}) = 1 - |\hat{\boldsymbol{u}}_m^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{u}_m(\mathrm{e}^{\mathrm{j}\Omega})| \ . \tag{13}$$

where the Hermitian angle is insensitive to the eigenvectors' arbitrary phase shifts. Small values of $\gamma_m(\mathrm{e}^{\mathrm{j}\Omega})$ means that subspaces are aligned, while $\gamma_m(\mathrm{e}^{\mathrm{j}\Omega}) = 1$ indicates orthogonality. In Fig. 4, the measured distribution* of $\gamma_m(\mathrm{e}^{\mathrm{j}\Omega})$ shows higher subspace alignment where eigenvalues according to Fig. 3 are clearly separated. Near algebraic multiplicities, a mismatch in subspaces arises as analysed in (12).

For Model 2, the distance between the ground truth eigenvalues reduces the problem of associating estimates $\hat\lambda_m(\mathrm{e}^{\mathrm{j}\Omega_k})$ correctly at every frequency. In order to measure the statistics across the ensemble and the frequency bins relative to the boundary in (9), we employ a normalised metric
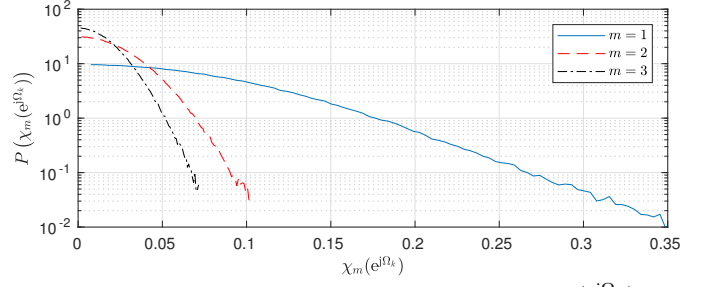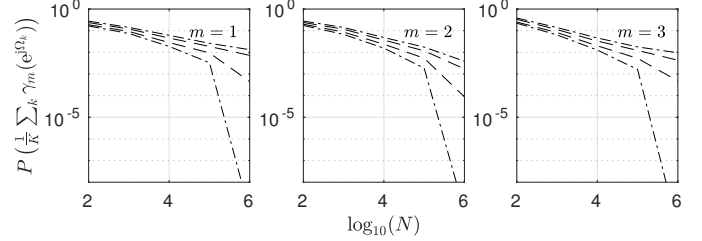
$$\chi_m(\mathrm{e}^{\mathrm{j}\Omega_0}) = \frac{|\hat\lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0}) - \lambda_m(\mathrm{e}^{\mathrm{j}\Omega_0})|}{\max\{|\lambda_{\max}(\mathrm{e}^{\mathrm{j}\Omega_0})\}|, |\lambda_{\min}(\mathrm{e}^{\mathrm{j}\Omega_0})|\}} \ , \tag{14}$$

such that $0 \le \chi_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \le 1$. The estimated probability density function of $\chi_m(\mathrm{e}^{\mathrm{j}\Omega_0})$, measured over an ensemble of 5000 runs and across 512 frequency bins, is shown in Fig. 5, demonstrating that the normalised bound $\chi_m(\mathrm{e}^{\mathrm{j}\Omega_0}) \le 1$ is satisfied. The distribution of total metrics — i.e. averaged over frequency — for the subspace and eigenvalue mismatches $\chi_m(\mathrm{e}^{\mathrm{j}\Omega})$ and $\gamma_m(\mathrm{e}^{\mathrm{j}\Omega})$ are depicted in Figs. 6 and 7, demonstrating the improved estimation of the parahermitian matrix EVD factors as the sample size N increases.

VI. CONCLUSIONS

When estimating a cross-spectral density matrix from a finite data set, the arising estimation error results in perturbations of the ground truth eigenvalues and eigenvectors of its parahermitian matrix eigenvalue decomposition. Evaluated on the unit circle, bounds for the perturbation of both quantities can be stated, which in case of the eigenvalues depends on the estimation error, while the eigenvector perturbation additionally depends on the distance between the ground truth eigenvalues. These findings have been demonstrated in and underpined by a number of simulations.

REFERENCES

[1] M. Alrmah, J. Corr, A. Alzin, K. Thompson, and S. Weiss. Polynomial subspace decomposition for broadband angle of arrival estimation. In *Sensor Signal Processing for Defence*, pages 1–5, Edinburgh, Scotland, Sept 2014.

[2] R. L. Anderson. Distribution of the serial correlation coefficient. *Ann. Math. Statist.*, 13(1):1–13, 03 1942.

[3] J. Corr, K. Thompson, S. Weiss, I. Proudler, and J. McWhirter. Impact of source model matrix conditioning on PEVD algorithms. In *IET/EURASIP Intelligent Signal Processing*, London, UK, December 2015.

[4] J. Corr, K. Thompson, S. Weiss, I. Proudler, and J. McWhirter. Row-shift corrected truncation of paraunitary matrices for PEVD algorithms. In *23rd European Signal Processing Conference*, pages 849–853, Nice, France, August/September 2015.

[5] J. Corr, K. Thompson, S. Weiss, I. Proudler, and J. McWhirter. Shortening of paraunitary matrices obtained by polynomial eigenvalue decomposition algorithms. In *Sensor Signal Processing for Defence*, Edinburgh, Scotland, September 2015.

[6] F. Coutts, K. Thompson, S. Weiss, and I. Proudler. Impact of fast-converging pevd algorithms on broadband aoa estimation. In *Sensor Signal Processing for Defence Conference*, pages 1–5, London, UK, December 2017.

[7] J. Foster, J. G. McWhirter, and J. Chambers. Limiting the order of polynomial matrices within the SBR2 algorithm. In *IMA International Conference on Mathematics in Signal Processing*, Cirencester, UK, December 2006.

[8] H. Goldsmith. The exact distributions of the serial correlation coefficients and an evaluation on some approximate distributions. *Journal of Statistical Computation and Simulation*, 5(2):115–134, 1977.

[9] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[10] T. Kato. *Perturbation Theory for Linear Operators*. Springer, 1980.

[11] R. K. Mallik. The pseudo-Wishart distribution and its application to MIMO systems. *IEEE Transactions on Information Theory*, 49(10):2761–2769, October 2003.

[12] J. G. McWhirter, P. D. Baxter, T. Cooper, S. Redif, and J. Foster. An EVD Algorithm for Para-Hermitian Polynomial Matrices. *IEEE Transactions on Signal Processing*, 55(5):2158–2169, May 2007.

[13] J. G. McWhirter and Z. Wang. Insight into the SBR2 algorithm. In *11th IMA Conference on Mathematics in Signal Processing*, Birmingham, UK, December 2016.

[14] J. Mosher and R. Leahy. Source localization using recursively applied and projected (rap) music. *IEEE Transactions on Signal Processing*, 47(2):332–340, February 1999.

[15] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 3rd edition, 1991.

[16] S. Redif. *Polynomial Matrix Decompositions and Paraunitary Filter Banks*. PhD thesis, School of Electronics & Computer Science, University of Southampton, July 2006.

[17] S. Redif, J. McWhirter, and S. Weiss. Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition. *IEEE Transactions on Signal Processing*, 59(11):5253–5264, November 2011.

[18] S. Redif, S. Weiss, and J. McWhirter. Sequential matrix diagonalization algorithms for polynomial EVD of parahermitian matrices. *IEEE Transactions on Signal Processing*, 63(1):81–89, January 2015.

[19] S. Redif, S. Weiss, and J. McWhirter. Relevance of polynomial matrix decompositions to broadband blind signal separation. *Signal Processing*, 134:76–86, May 2017.

[20] A. Steinhardt and J. Makhoul. On the autocorrelation of finite-length sequences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(6):1516–1520, Dec 1985.

[21] G. W. Stewart and J.-g. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

[22] C. H. Ta and S. Weiss. A Jointly Optimal Precoder and Block Decision Feedback Equaliser Design With Low Redundancy. In *15th European Signal Processing Conference*, pages 489–492, Poznan, Poland, September 2007.

[23] C. H. Ta and S. Weiss. Shortening the order of paraunitary matrices in SBR2 algorithm. In *6th International Conference on Information, Communications & Signal Processing*, pages 1–5, Singapore, Dec. 2007.

[24] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, Englewood Cliffs, 1993.

[25] S. Weiss, M. Alrmah, S. Lambotharan, J. McWhirter, and M. Kaveh. Broadband angle of arrival estimation methods in a polynomial matrix decomposition framework. In *IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pages 109–112, Dec 2013.

[26] S. Weiss, S. Bendoukha, A. Alzin, F. Coutts, I. Proudler, and J. Chambers. MVDR broadband beamforming using polynomial matrix techniques. In *23rd European Signal Processing Conference*, pages 839–843, Nice, France, September 2015.

[27] S. Weiss, J. Pestana, and I. Proudler. On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix. *IEEE Transactions on Signal Processing*, 2018. to appear.

[28] J. Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford University Press, 1965.

[29] L. Zhuang and A. T. Walden. Sample mean versus sample Fréchet mean for combining complex Wishart matrices: A statistical study. *IEEE Transactions on Signal Processing*, 65(17):4551–4561, September 2017.