



Impact of space–time covariance matrix estimation on bin-wise eigenvalue and eigenspace perturbations

Connor Delaosa ^a, Jennifer Pestana ^b, Ian K. Proudler ^a, Stephan Weiss ^a ^{*}

^a Department of Electronic & Electrical Engineering, University of Strathclyde, 204 George Street, Glasgow, G1 1XW, UK

^b Dept. of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow, G1 1XH, UK

ARTICLE INFO

Keywords:

Covariance
Space–time processing
Estimation errors
Eigenvalue decomposition
Eigenvalue perturbation
Eigenspace perturbation

ABSTRACT

In the context of broadband multichannel signal processing, problems can often be formulated using a space–time covariance matrix, and solved using a diagonalisation of this quantity via a polynomial or analytic eigenvalue decomposition (EVD). In this paper, we address the impact that an estimation of the space–time covariance has on the factors of such a decomposition. In order to address this, we consider a linear unbiased estimator based on Gaussian distributed data, and characterise the variance of this estimate, as well as the variance of the error between the estimate and the ground truth. These quantities in turn enable to find expressions for the bin-wise perturbation of the eigenvalues, which depends on the error variance of the estimate, and for the bin-wise perturbation of the eigenspaces, which depends on both the error variance but also on the eigenvalue distance. We adapt a number of known bounds for ordinary matrices and demonstrate the fit of these bounds in simulations. In order to minimise the error variance of the estimate, and hence the perturbation of the EVD factors, we discuss a way to optimise the lag support of the space–time covariance estimate without access to the ground truth on which the estimate is based.

1. Introduction

The second order statistics of a measurement vector $\mathbf{x}[n] \in \mathbb{C}^M$ acquired by an M -element array over discrete time, indexed by $n \in \mathbb{Z}$, are contained in the space–time covariance matrix $\mathbf{R}[\tau] = \mathcal{E} \{ \mathbf{x}[n] \mathbf{x}^H[n - \tau] \}$, where $\mathcal{E} \{ \cdot \}$ is the expectation operator and $\{ \cdot \}^H$ denotes Hermitian transposition. In narrowband signal processing, time delays with which e.g. sources illuminate the different array elements are entirely captured by phase shifts. In this narrowband case it therefore suffices to consider the instantaneous covariance matrix $\mathbf{R}[0]$ only. If broadband signals are to be addressed, explicit time delay information must be preserved, and this requires the inclusion of the lag component τ in the space–time covariance matrix $\mathbf{R}[\tau]$, which thus includes auto- and cross-correlation sequences as entries.

Many narrowband array processing applications such as beamforming [1–3] or angle of arrival estimation [4] rely on the eigenvalue decomposition (EVD) of the instantaneous covariance matrix $\mathbf{R}[0]$ for optimal narrowband solutions. For the broadband case, the narrowband solutions can be generalised using a similar factorisation of the space–time covariance matrix $\mathbf{R}[\tau]$. This however requires the diagonalisation of $\mathbf{R}[\tau]$ for every lag τ , which is equivalent to strong decorrelation of the underlying data $\mathbf{x}[n]$ [5]. Because such decompositions are generally formulated for the cross-spectral density (CSD) matrix $\mathbf{R}(z) =$

$\sum_{\tau} \mathbf{R}[\tau] z^{-\tau}$ — a Laurent polynomial matrix [6,7] — they are termed polynomial matrix eigenvalue decompositions (PEVD) [8–10]. Various PEVD algorithms have been investigated over the two decades, see [8,11–21], and form the basis of algorithms for various applications including broadband beamforming [22–25], broadband signal compaction and coding [13,26,27], broadband source separation [27,28], scene discovery [29], or broadband angle of arrival estimation [30–32].

For most applications, the space–time covariance matrix must be estimated from measurements $\mathbf{x}[n]$ over a limited number of, say N , snapshots i.e. only a data set $\{ \mathbf{x}[n] \in \mathbb{C}^M, n = 0 \dots (N - 1) \}$ is available to obtain an estimate $\hat{\mathbf{R}}[\tau]$ of $\mathbf{R}[\tau]$. This may be either due to the limited availability of data, or the need to restrict the estimation to an interval over which the data can be assumed to be stationary. While various investigations have been undertaken into the accuracy of the above decompositions [33,34], and limiting factors due to algorithm-internal order reductions [35–38] or the conditioning of the underlying source model [39], it has only been relatively recently that the estimation errors of $\mathbf{R}[\tau]$ and their impact on the factorisation of the PEVD of $\mathbf{R}(z)$ [40] have been investigated.

Since the estimated quantity $\hat{\mathbf{R}}[\tau]$ will likely differ from the ground truth space–time covariance matrix $\mathbf{R}[\tau]$, the EVD factors of $\hat{\mathbf{R}}[\tau]$ will

* Corresponding author.

E-mail address: stephan.weiss@strath.ac.uk (S. Weiss).

generally also differ from those of $\mathbf{R}[\tau]$. Assessing the impact of any discrepancies of an ordinary matrix on its eigenvalues and eigenvectors has been well studied under the topic of perturbation theory [41–46]. Rellich [47] showed that for a Hermitian matrix \mathbf{A}_0 and an analytic perturbation term $\mathbf{A}(\epsilon)$, the eigenvalues and eigenspaces of $\mathbf{A}_0 + \mathbf{A}(\epsilon)$ are also analytic in $\epsilon \in \mathbb{R}$. While his work has been useful in order to establish the existence of an eigenvalue decomposition of $\mathbf{R}(z)$ [9], here the perturbation term is random and not analytic.

To characterise the perturbation of the covariance matrix, we need to consider the statistics that are involved in estimating its constituent auto- and cross-correlation sequences from a finite number of samples, N . To date, results have been derived for the broadband single channel case, i.e. for the sample auto-correlation sequence. For this case, various attempts have been undertaken for random signals that can be modelled as first order auto-regressive processes [48,49], or by more general data models [50–52]. For the multichannel case, analysis has been restricted to narrowband signals and the case where snapshots $\mathbf{x}[n]$ are temporally independent; this leads to the spatial covariance matrix being Wishart distributed [53,54]. This is insufficient for the distribution of the sample space–time covariance matrix, $\hat{\mathbf{R}}[\tau]$, which is constructed from multichannel broadband data, where the correlation between subsequent snapshots $\mathbf{x}[n]$ and $\mathbf{x}[n+1]$ is vital and has explicitly motivated the inclusion of the lag parameter τ [7–9,55].

Therefore, in this paper we want to characterise space–time covariance estimation and its impact on the perturbation of PEVD factors. Based on initial work in [56], we expand by analysing potential estimators, whereby an incorrect choice can lead to a rank-one space–time covariance irrespective of the number of contributing sources. For a linear unbiased estimator, we state a closed-form solution for the expected variance of the error between $\hat{\mathbf{R}}[\tau]$ and $\mathbf{R}[\tau]$, which depends both the sample size N and the ground truth $\mathbf{R}[\tau]$. Particularisation of our results agree with [50,53,54] and with results from spectral estimation such as [52]. This error variance can in turn be linked to the impact on the PEVD of $\hat{\mathbf{R}}[\tau]$. Initial work in [40] is expanded by the derivation and demonstration of bounds for eigenvalues and eigenspaces. In order to minimise this impact, we also investigate over which range of lags $\hat{\mathbf{R}}[\tau]$ should be evaluated, for which initial ideas had been reported in [57].

Below, we commence with a definition of the space–time covariance matrix, and review its properties and matrix factorisation in Section 2. In Section 3 we analyse the sample cross-correlation sequence, which is used for the sample space–time covariance in Section 4 and its perturbation effects in Section 5, followed by experimental verification in Section 6. Section 7 considers how the lag-support of a space–time covariance estimate may be optimised in order to minimise its deviation from the ground truth $\mathbf{R}[\tau]$, and to best limit the perturbation of its bin-wise EVD factors. A summary and conclusions can be found in Section 8.

2. Space–time covariance matrix and analytic EVD

2.1. Data model and space–time covariance

Given M sensor measurements $x_m[n]$, $m = 1 \dots M$, organised in a column vector $\mathbf{x}[n] = [x_1[n] \dots x_M[n]]^T$, the space–time covariance matrix of the data was defined in Section 1 as $\mathbf{R}[\tau] = \mathcal{E} \{ \mathbf{x}[n] \mathbf{x}^H[n - \tau] \}$. The source model or innovation filter [58] in Fig. 1 ties this data vector $\mathbf{x}[n]$ to L zero-mean unit-variance mutually independent complex circularly symmetric Gaussian sources u_ℓ , $\ell = 1 \dots L$, such that $\mathcal{E} \{ u_\ell[n] u_\nu[n - \tau] \} = \delta[\tau] \delta[\ell - \nu]$ for $\nu = 1 \dots L$ [59]. As a result, the space–time covariance matrix can be expressed as

$$\mathbf{R}[\tau] = \sum_n \mathbf{H}[n] \mathbf{H}^H[n - \tau]$$

where $\mathbf{H}[n] \in \mathbb{C}^{M \times L}$ is a matrix of deterministic filters. If the entry in the m th row and ℓ th column of $\mathbf{H}[n]$ represents the impulse response

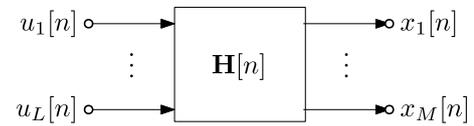


Fig. 1. Source model for M convolutively mixed signals arising from L independent unit-variance zero-mean sources.

$h_{m\ell}[n]$ between the ℓ th source and the m th sensor, then

$$r_{m\mu}[\tau] = \sum_n \sum_{\ell=1}^L h_{m\ell}[n] h_{\mu\ell}^*[n - \tau] \quad (1)$$

is a cross-correlation sequence that occupies the m th row and μ th column of $\mathbf{R}[\tau]$, with $\{\cdot\}^*$ denoting complex conjugation.

2.2. Cross-spectral density matrix

Since the space–time covariance matrix comprises auto- and cross-correlation sequences, it satisfies the symmetry $\mathbf{R}[\tau] = \mathbf{R}^H[-\tau]$. Its z -transform, the cross-spectral density (CSD) matrix $\mathbf{R}(z) = \sum_{\tau} \mathbf{R}[\tau] z^{-\tau}$ – or for short $\mathbf{R}(z) \longleftrightarrow \mathbf{R}[\tau]$ to denote a transform pair – therefore is a parahermitian matrix, such that its parahermitian transpose, denoted by the operator $\{\cdot\}^P$, is equal to itself: $\mathbf{R}^P(z) = \{\mathbf{R}(1/z^*)\}^H = \mathbf{R}(z)$ [55].

2.3. Analytic eigenvalue decomposition

A parahermitian, analytic $\mathbf{R}(z)$ admits an analytic EVD [9]

$$\mathbf{R}(z) = \mathbf{Q}(z) \mathbf{A}(z) \mathbf{Q}^P(z), \quad (2)$$

where $\mathbf{Q}(z)$ is a paraunitary matrix of eigenvectors and $\mathbf{A}(z)$ is a diagonal parahermitian matrix of eigenvalues. In most standard cases, these factors can be selected to be analytic [10].¹ If so, then the factors $\mathbf{Q}(z)$ and $\mathbf{A}(z)$ on the r.h.s. of (2) are generally algebraic or transcendental. While for an analytic $\mathbf{R}(z)$, an analytic solution for $\mathbf{A}(z)$ is unique, the eigenvectors can contain arbitrary allpass filters, and only the eigenspaces within which these vectors reside are unique.

If the space–time covariance matrix is estimated from a finite set of samples, the obtained matrix $\hat{\mathbf{R}}[\tau] \longleftrightarrow \hat{\mathbf{R}}(z)$ will differ from $\mathbf{R}(z)$. Thus also the eigenvalues and eigenspaces can be expected to be perturbed. Since the estimate $\hat{\mathbf{R}}(z)$ will be a Laurent polynomial and therefore analytic due to its finite nature, we would ideally be interested in the perturbation of the analytic factorisation $\hat{\mathbf{R}}(z) = \hat{\mathbf{Q}}(z) \hat{\mathbf{A}}(z) \hat{\mathbf{Q}}^P(z)$. However, while such an analytic factorisation is guaranteed by the theorems in [9,10,60,61], Section 5 will argue that we can currently only state the perturbation at individual frequencies $\Omega_0 \in \mathbb{R}$ for an evaluation of $\hat{\mathbf{R}}(z)$ on the unit circle.

3. Cross-correlation estimation

Practically, the space–time covariance matrix must be estimated from data. If only a set of N snapshots of $\mathbf{x}[n]$, $n = 0 \dots (N - 1)$, is available, then generally the estimate for the space–time covariance matrix, $\hat{\mathbf{R}}[\tau]$, will experience estimation errors. In this section, we focus on the estimation of the cross-correlation sequence since (1) is the most general component of the space–time covariance matrix.

¹ Many current algorithms aim to provide the McWhirter decomposition [8], which is not necessarily analytic, and approximates $\mathbf{Q}(z)$ and a spectrally majorised $\mathbf{A}(z)$ by Laurent polynomials.

3.1. Biased estimation

The cross-correlation sequence between two signals $x_m[n]$ and $x_\mu[n]$, $m, \mu \in \{1 \dots M\}$, is defined as

$$r_{m\mu}[\tau] = \mathcal{E} \left\{ x_m[n] x_\mu^*[n - \tau] \right\}. \quad (3)$$

Assuming strong ergodicity and therefore by implication stationarity for the involved signals [62], the estimation of $r_{m\mu}[\tau]$ over a set of N time snapshots can be performed in different ways. For later use in Section 4.1, one specific cross-correlation estimate is

$$\hat{r}_{m\mu}^{(\text{biased})}[\tau] = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-|\tau|-1} x_m[n + \tau] x_\mu^*[n], & \tau \geq 0; \\ \frac{1}{N} \sum_{n=0}^{N-|\tau|-1} x_m[n] x_\mu^*[n - \tau], & \tau < 0. \end{cases} \quad (4)$$

The superscript will differentiate this estimate from an unbiased one discussed in Section 3.2.

We first calculate the mean of the estimator in (4) for $\tau \geq 0$,

$$\begin{aligned} \text{mean}\{\hat{r}_{m\mu}^{(\text{biased})}[\tau]\} &= \mathcal{E} \left\{ \hat{r}_{m\mu}^{(\text{biased})}[\tau] \right\} \\ &= \frac{1}{N} \sum_{n=0}^{N-|\tau|-1} \mathcal{E} \left\{ x_m[n + \tau] x_\mu^*[n] \right\} \\ &= \frac{1}{N} \sum_{n=0}^{N-|\tau|-1} r_{m\mu}[\tau] = \frac{N - |\tau|}{N} r_{m\mu}[\tau]. \end{aligned}$$

Similar analysis can be performed for $\tau < 0$, such that for $|\tau| < N$

$$\text{mean}\{\hat{r}_{m\mu}^{(\text{biased})}[\tau]\} = \frac{N - |\tau|}{N} r_{m\mu}[\tau]. \quad (5)$$

This shows that the estimator in (4) is biased for all lag values τ except $\tau = 0$.

Due to the finite length of the signals over the interval $0 \leq n < N$, it is possible to state their z -transforms $X_m(z) = \sum_{n=0}^{N-1} x_m[n] z^{-n}$. In this case the cross-spectral density estimate $\hat{R}_{m\mu}^{(\text{biased})}(z) \leftarrow \hat{r}_{m\mu}^{(\text{biased})}[\tau]$ is

$$\hat{R}_{m\mu}^{(\text{biased})}(z) = \frac{1}{N} X_m(z) X_\mu^*(z), \quad (6)$$

which will be further examined in Section 4.1. The triangular data window in (7) is responsible for the bias of the estimator. Note that for the case $m = \mu$ and $z = e^{j\Omega}$, (6) represents the periodogram, which is known to be an inconsistent, biased estimator for the power spectral density [63].

3.2. Unbiased estimation

With a lag-dependent normalisation compared to (4), an estimate of $r_{m\mu}[\tau]$ over a sample size N is defined as

$$\hat{r}_{m\mu}[\tau] = \begin{cases} \frac{1}{N - |\tau|} \sum_{n=0}^{N-|\tau|-1} x_m[n + \tau] x_\mu^*[n], & \tau \geq 0; \\ \frac{1}{N - |\tau|} \sum_{n=0}^{N-|\tau|-1} x_m[n] x_\mu^*[n - \tau], & \tau < 0 \end{cases} \quad (7)$$

and this can be shown to be unbiased. For example for $\tau \geq 0$,

$$\begin{aligned} \text{mean}\{\hat{r}_{m\mu}[\tau]\} &= \mathcal{E} \left\{ \hat{r}_{m\mu}[\tau] \right\} \\ &= \frac{1}{N - |\tau|} \sum_{n=0}^{N-|\tau|-1} \mathcal{E} \left\{ x_m[n + \tau] x_\mu^*[n] \right\} \\ &= \frac{1}{N - |\tau|} \sum_{n=0}^{N-|\tau|-1} r_{m\mu}[\tau] = r_{m\mu}[\tau], \end{aligned}$$

i.e. the mean of the quantity estimated via (7) is indeed the cross-correlation sequence defined in (3).

3.3. Variance of estimator

The variance of the unbiased cross-correlation sequence estimator in (7) is defined as

$$\begin{aligned} \text{var}\{\hat{r}_{m\mu}[\tau]\} &= \mathcal{E} \left\{ (\hat{r}_{m\mu}[\tau] - r_{m\mu}[\tau]) (\hat{r}_{m\mu}[\tau] - r_{m\mu}[\tau])^* \right\} \\ &= \mathcal{E} \left\{ \hat{r}_{m\mu}[\tau] \hat{r}_{m\mu}^*[\tau] \right\} - \mathcal{E} \left\{ \hat{r}_{m\mu}[\tau] \right\} r_{m\mu}^*[\tau] \\ &\quad - r_{m\mu}[\tau] \mathcal{E} \left\{ \hat{r}_{m\mu}^*[\tau] \right\} + r_{m\mu}[\tau] r_{m\mu}^*[\tau] \\ &= \mathcal{E} \left\{ \hat{r}_{m\mu}[\tau] \hat{r}_{m\mu}^*[\tau] \right\} - r_{m\mu}[\tau] r_{m\mu}^*[\tau]. \end{aligned} \quad (8)$$

Inserting the estimation in (7) into $\mathcal{E} \left\{ \hat{r}_{m\mu}[\tau] \hat{r}_{m\mu}^*[\tau] \right\}$ results in (8) containing a fourth-order term involving the signals $x_m[n]$ and $x_\mu[n]$.

The above fourth-order term can be expressed in terms of cumulants, but for Gaussian signals cumulants of order three and above are zero [64,65]. This property also holds for the complex-valued case [66], such that this fourth-order term simplifies as follows²

$$\begin{aligned} \mathcal{E} \left\{ x_m[n] x_\mu^*[n - \tau] x_m^*[n] x_\mu[n - \tau] \right\} &= \\ &= \mathcal{E} \left\{ x_m[n] x_\mu^*[n - \tau] \right\} \cdot \mathcal{E} \left\{ x_m^*[n] x_\mu[n - \tau] \right\} \\ &\quad + \mathcal{E} \left\{ x_m[n] x_m^*[n] \right\} \cdot \mathcal{E} \left\{ x_\mu^*[n - \tau] x_\mu[n - \tau] \right\} \\ &\quad + \mathcal{E} \left\{ x_m[n] x_\mu[n - \tau] \right\} \cdot \mathcal{E} \left\{ x_\mu^*[n - \tau] x_m^*[n] \right\}. \end{aligned}$$

Therefore, for $\tau \geq 0$, the variance of the estimator in (7) becomes

$$\begin{aligned} \text{var}\{\hat{r}_{m\mu}[\tau]\} &= \frac{1}{(N - |\tau|)^2} \sum_{n,v=0}^{N-|\tau|-1} \left(\mathcal{E} \left\{ x_m[n + \tau] x_\mu^*[n] \right\} \cdot \right. \\ &\quad \cdot \mathcal{E} \left\{ x_m^*[v + \tau] x_\mu[v] \right\} + \\ &\quad + \mathcal{E} \left\{ x_m[n + \tau] x_m^*[v + \tau] \right\} \mathcal{E} \left\{ x_\mu^*[n] x_\mu[v] \right\} \\ &\quad + \mathcal{E} \left\{ x_m[n + \tau] x_\mu[v] \right\} \mathcal{E} \left\{ x_\mu^*[n] x_m^*[v + \tau] \right\} \left. \right) \\ &\quad - r_{m\mu}[\tau] r_{m\mu}^*[\tau] \\ &= \frac{1}{(N - |\tau|)^2} \sum_{n,v=0}^{N-|\tau|-1} \left(\mathcal{E} \left\{ x_m[n] x_m^*[v] \right\} \cdot \right. \\ &\quad \cdot \mathcal{E} \left\{ x_\mu^*[n] x_\mu[v] \right\} + \\ &\quad + \mathcal{E} \left\{ x_m[n] x_\mu[v - \tau] \right\} \mathcal{E} \left\{ x_m^*[v] x_\mu^*[n - \tau] \right\} \left. \right). \end{aligned} \quad (9)$$

The same result can be obtained for $\tau < 0$, and matches results reached in [52]. Note that the first term in (9) can be simplified as

$$\begin{aligned} &\sum_{n,v=0}^{N-|\tau|-1} \left(\mathcal{E} \left\{ x_m[n] x_m^*[v] \right\} \mathcal{E} \left\{ x_\mu^*[n] x_\mu[v] \right\} \right) \\ &= \sum_{n,v=0}^{N-|\tau|-1} \left(\mathcal{E} \left\{ x_m[n] x_m^*[n - (n - v)] \right\} \cdot \right. \\ &\quad \cdot \mathcal{E} \left\{ x_\mu^*[n] x_\mu[n - (n - v)] \right\} \left. \right) \\ &= \sum_{n,v=0}^{N-|\tau|-1} r_{mm}[n - v] r_{\mu\mu}^*[n - v] \\ &= \sum_{t=-N+|\tau|+1}^{N-|\tau|-1} (N - |\tau| - |t|) r_{mm}[t] r_{\mu\mu}^*[t]. \end{aligned}$$

The second term in (9) can be addressed similarly. With $\bar{r}_{m\mu}[\tau] = \mathcal{E} \left\{ x_m[n] x_\mu[n - \tau] \right\}$ denoting the complementary cross-correlation sequence, the variance of the sample cross-correlation sequence becomes

² We are grateful to one of our anonymous reviewers for pointing out that this is the Isserlis formula for complex-valued processes. The formula holds more generally than for Gaussian signals as long as the assumption of the summability of fourth-order moments is satisfied [63].

$$\text{var}\{\hat{r}_{m\mu}[\tau]\} = \frac{1}{(N-|\tau|)^2} \sum_{t=-N+|\tau|+1}^{N-|\tau|-1} (N-|\tau|-|t|) \cdot \left(r_{mm}[t]r_{\mu\mu}^*[t] + \bar{r}_{m\mu}[\tau+t]\bar{r}_{m\mu}^*[\tau-t] \right). \quad (10)$$

As special cases, if in the source model in Fig. 1, $u[n]$ is complex valued with a circularly-symmetric distribution, then $\bar{r}_{m\mu}[\tau] = 0 \forall \tau \in \mathbb{Z}$. If all signals are real valued, then $\bar{r}_{m\mu}[\tau] = r_{m\mu}[\tau]$.

3.4. Comparison with known results

The result in (10) generalises a number of solutions reported in the literature. In the real-valued, single channel case, i.e. $\mathbf{x}[n] \in \mathbb{R}^L$ and $m = \mu$, then (10) simplifies to

$$\text{var}\{\hat{r}_{mm}[\tau]\} = \frac{1}{(N-|\tau|)^2} \sum_{t=-N+|\tau|+1}^{N-|\tau|-1} (N-|\tau|-|t|) \cdot \left(|r_{mm}[t]|^2 + r_{mm}[\tau+t]r_{mm}[\tau-t] \right).$$

This matches with the result reported in [50].

Next, we consider the narrowband case without any temporal correlation of signals. In this case the transfer function $\mathbf{H}(z) \circ \rightarrow \mathbf{H}[n]$ is an ordinary matrix, $\mathbf{H}(z) = \mathbf{H}_0$, and the signals $x_m[n]$ and $x_\mu[n]$ only have non-zero correlation for the case $\tau = 0$. If further $\mathbf{u}[n] \in \mathbb{R}^L$ and $\mathbf{H}_0 \in \mathbb{R}^{M \times L}$, then the space-time covariance $\mathbf{R}[\tau] = \mathbf{H}_0 \mathbf{H}_0^T \delta[\tau]$ is Wishart-distributed. For the instantaneous and real case, (10) simplifies to

$$\text{var}\{\hat{r}_{m\mu}[0]\} = \frac{1}{N} \left(r_{mm}[0]r_{\mu\mu}[0] + |r_{m\mu}[0]|^2 \right),$$

which indeed matches the variance of a Wishart distribution.

4. Sample space-time covariance matrix estimation

The cross-correlation estimation of Section 3 has a profound effect on the construction of a space-time covariance matrix. We first explore the impact of a biased estimation of the type in (4) before considering the estimation and modelling error when employing an unbiased estimator.

4.1. Biased estimator

Due to the r.h.s. of (6) being an outer product of a single vector, $\hat{\mathbf{R}}^{(\text{biased})}(z)$ by its very structure is a rank one matrix, and therefore only has a single non-zero eigenvalue. The corresponding principal eigenvector of (6) can be obtained by normalising $\mathbf{x}(z) = \sum_{n=0}^{N-1} \mathbf{x}[n]z^{-n}$, $\hat{\mathbf{q}}_1^{(\text{biased})}(z) = \frac{\Phi(z)}{\sqrt{\mathbf{x}^P(z)\mathbf{x}(z)}} \mathbf{x}(z)$, (11)

where $\Phi(z)$ is an arbitrary allpass filter. Note that the normalisation involves a $\sqrt{\cdot}$ operation on a polynomial so $\hat{\mathbf{q}}_1(z)$ contains potentially algebraic or transcendental functions in z and these operations can be approximated using a Maclaurin or Taylor series [9]. Thus the principal eigenvalue is

$$\hat{\Lambda}_1^{(\text{biased})}(z) = \frac{1}{N} \mathbf{x}^P(z)\mathbf{x}(z), \quad (12)$$

with all remaining eigenvalues $\hat{\Lambda}_m^{(\text{biased})}(z) = 0$ for $m = 2 \dots M$.

While this is not obvious in the time domain, the z -domain analysis of the sample CSD matrix directly shows that the biased cross-correlation sequence estimation via (4) leads to a catastrophic collapse in the rank of the estimated matrix.

4.2. Unbiased estimator

Since we know that for the cross-correlation sequence the estimator in (7) is unbiased, building up a sample space-time covariance matrix $\hat{\mathbf{R}}[\tau]$ from such components also leads to an unbiased estimate, such that $\mathcal{E}\{\hat{\mathbf{R}}[\tau]\} = \mathbf{R}[\tau]$.

Assume that the space-time covariance matrix has support of length $2\tau_{\max} + 1$, i.e. $\mathbf{R}[\tau] = \mathbf{0} \forall |\tau| > \tau_{\max}$. Further assume that $\hat{\mathbf{R}}[\tau]$ is estimated over a support length of $2T + 1$. We define the mismatch as

$$\mathbf{E}[\tau] = \mathbf{R}[\tau] - \hat{\mathbf{R}}[\tau], \quad (13)$$

and can state a mean square modelling error as

$$\begin{aligned} \xi &= \sum_{\tau=-\infty}^{\infty} \mathcal{E}\{\|\mathbf{E}[\tau]\|_{\text{F}}^2\} \\ &= \underbrace{\sum_{\tau=-T}^T \mathcal{E}\{\|\mathbf{E}[\tau]\|_{\text{F}}^2\}}_{\xi_1} + 2 \underbrace{\sum_{\tau=T+1}^{\tau_{\max}} \|\mathbf{R}[\tau]\|_{\text{F}}^2}_{\xi_2}, \end{aligned} \quad (14)$$

where the first term, ξ_1 , is an estimation error due to (10), while the second term, ξ_2 , represents a truncation error. Note that $\xi_2 = 0$ if $T \geq \tau_{\max}$.

For the estimation error, using (10) leads to

$$\begin{aligned} \xi_1 &= \sum_{\tau=-T}^T \sum_{m,\mu=1}^M \text{var}\{\hat{r}_{m\mu}[\tau]\} \\ &= \sum_{\tau=-T}^T \sum_{t=-N+|\tau|+1}^{N-|\tau|-1} \frac{N-|\tau|-|t|}{(N-|\tau|)^2} \left(\text{tr}\{|\mathbf{R}[t]|^2\} + \text{vec}\{\bar{\mathbf{R}}[\tau-t]\}^H \text{vec}\{\bar{\mathbf{R}}[\tau+t]\} \right), \end{aligned} \quad (15)$$

where the operator $\text{vec}\{\cdot\}$ vectorises its argument, T is the support of the estimate, and $\bar{\mathbf{R}}[\tau]$ is the complementary space-time covariance matrix holding the complementary cross-correlation sequences defined in Section 3.3. Therefore, the modelling error ξ depends only on the space-time covariance matrix itself, the sample size N , and the support of the estimate, T .

4.3. Optimum support for unbiased estimator

With the mean square modelling error ξ depending on the ground truth $\mathbf{R}[\tau]$, the sample size N , and the chosen maximum lag T , the only parameter typically under design control is the support $|\tau| \leq T$ over which $\mathbf{R}[\tau]$ is evaluated. The optimum value T_{opt} for T in terms of minimising the mean square modelling error therefore is

$$T_{\text{opt}} = \arg \min_T \xi. \quad (16)$$

In general, this will be a trade-off between the terms ξ_1 and ξ_2 . Since $T > \tau_{\max}$ leads to $\xi_2 = 0$, and ξ_1 generally grows with increasing T , we find $T_{\text{opt}} < \tau_{\max}$, i.e. it appears better to underestimate than to overestimate the support of $\mathbf{R}[\tau]$ in practice. A heuristic sample-set based approximate optimisation of the support window has been investigated in [56].

Example 1. A space-time covariance matrix $\mathbf{R}(z) : \mathbb{C} \rightarrow \mathbb{C}^{2 \times 2}$ of order 120 is obtained from the source model in [16]. If the estimate based on $N = 10^3$ samples is calculated over a window $|\tau| \leq T$, then the estimation and truncation error terms are shown in Fig. 2. Note that here $T_{\text{opt}} = 9$, which is substantially smaller than the support of the ground truth. Δ

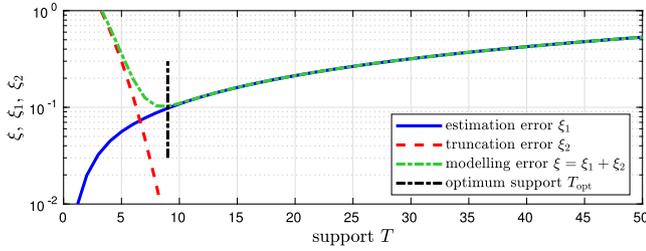


Fig. 2. Errors incurred when estimating $\mathbf{R}[\tau]$ from $N = 10^3$ samples over the support $|\tau| \leq T$.

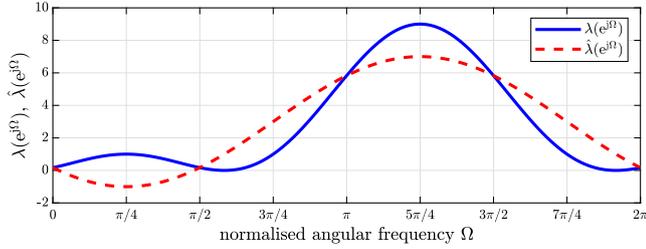


Fig. 3. Power spectral density $\lambda(e^{j\Omega})$ corresponding to $\mathbf{R}[\tau] \in \mathbb{C}^{1 \times 1}$, and $\hat{\lambda}(e^{j\Omega})$ obtained from a support-limited $\hat{\mathbf{R}}[\tau]$.

4.4. Truncation and loss of positive semi-definiteness

If $\mathbf{R}(z) : \mathbb{C} \rightarrow \mathbb{C}^{M \times M}$ is a space–time covariance matrix or emerges from a product $\mathbf{R}(z) = \mathbf{A}(z)\mathbf{A}^H(z)$, with $\mathbf{A}(z) : \mathbb{C} \rightarrow \mathbb{C}^{M \times L}$ with L and $\mathbf{A}(z)$ arbitrary, then its evaluation on the unit circle is positive semi-definite, i.e. $\mathbf{R}(e^{j\Omega}) \succeq 0 \forall \Omega \in \mathbb{R}$. The introduction of estimation or truncation errors may destroy this property, such that $\hat{\mathbf{R}}(z) \not\succeq 0$. This loss has been observed in [50], and been addressed recently in e.g. [67,68]. We highlight this loss of positive semi-definiteness by way of an example.

Example 2. For $M = 1$ and $\mathbf{A}(z) = 1 - e^{j\pi/4}z^{-1} + jz^{-2}$, $\mathbf{R}(z) = \mathbf{A}(z)\mathbf{A}^H(z)$ reduces to a power spectral density, and is equal to its only eigenvalue $\lambda(z)$. In the case of truncating it to a support length of 3 instead of 5, the evaluation of $\hat{\lambda}(z)$ on the unit circle, $\hat{\lambda}(e^{j\Omega})$, is shown in Fig. 3. Over the interval $0 < \Omega < \frac{\pi}{2}$, $\hat{\lambda}(e^{j\Omega}) < 0$, i.e. the estimated PSD now takes on some negative values. Δ

Thus, for the biased estimator in (4), while truncation removes the problem of structural rank-deficiency in (6), the estimation with finite data removes the guarantee of positive semi-definiteness that comes with (6), and eigenvalues may take on negative values when evaluated on the unit circle.

5. Perturbation of eigenvalues and eigenspaces

This section explores how an estimation error $E(z) \circ \rightarrow \mathbf{E}[\tau]$ in (13) impacts on the parahermitian matrix EVD in (2), i.e. how much the quantities $\hat{\mathbf{Q}}(z)$ and $\hat{\boldsymbol{\Lambda}}(z)$ of the sample CSD matrix $\hat{\mathbf{R}}(z)$ deviate from the ground truth. For the analysis, we evaluate on the unit circle, i.e. for $z = e^{j\Omega}$, and utilise a number of results from matrix perturbation theory [43,46] to first investigate the eigenvalues in Section 5.1. To remove the ambiguity of eigenvectors from the analysis, we will secondly analyse the perturbation of eigenspaces in Section 5.2.

5.1. Impact of modelling error on eigenvalues

At any specific sample point Ω_0 , the evaluation of the z -transform of (13) at $z = e^{j\Omega_0}$ on the unit circle gives $E(e^{j\Omega_0}) = \mathbf{R}(e^{j\Omega_0}) - \hat{\mathbf{R}}(e^{j\Omega_0})$ so that, by Weyl's theorem [43], we have the following bounds on the

perturbation of the m th eigenvalue $\lambda_m(e^{j\Omega_0})$ of $\mathbf{R}(e^{j\Omega_0})$,

$$\begin{aligned} \lambda_{\min}^E(e^{j\Omega_0}) + \lambda_m(e^{j\Omega_0}) &\leq \hat{\lambda}_m(e^{j\Omega_0}) \\ \hat{\lambda}_m(e^{j\Omega_0}) &\leq \lambda_{\max}^E(e^{j\Omega_0}) + \lambda_m(e^{j\Omega_0}), \end{aligned}$$

where $\hat{\lambda}_m(e^{j\Omega_0})$ is the m th eigenvalue of $\hat{\mathbf{R}}(e^{j\Omega_0})$, and $\lambda_{\min}^E(e^{j\Omega_0})$ and $\lambda_{\max}^E(e^{j\Omega_0})$ are the minimum and maximum eigenvalue of $E(e^{j\Omega_0})$ respectively. After reshuffling,

$$\lambda_{\min}^E(e^{j\Omega_0}) \leq \hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0}) \leq \lambda_{\max}^E(e^{j\Omega_0}) \quad (17)$$

provides bounds for the discrepancy between eigenvalues $\lambda_m(e^{j\Omega_0})$ and $\hat{\lambda}_m(e^{j\Omega_0})$.

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

$$\sum_{m=1}^M (\hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0}))^2 \leq \|E(e^{j\Omega_0})\|_F^2 \quad (18)$$

holds. Further, the Bauer–Fike theorem [45] guarantees that

$$|\hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0})| \leq \kappa\{\hat{\mathbf{U}}(e^{j\Omega_0})\} \|E(e^{j\Omega_0})\|_2, \quad (19)$$

with $\kappa\{\mathbf{A}\}$ the condition number of the matrix \mathbf{A} . Since $\hat{\mathbf{U}}(e^{j\Omega_0})$ is unitary by definition, $\kappa\{\hat{\mathbf{U}}(e^{j\Omega_0})\} = 1$, and (19) simplifies further. It can also be noted that the bound on (19) is always an upper bound for (17).

Overall, both the bounds (17) and (19) relate the deviation between the ground truth and estimated eigenvalues directly to the estimation error $E(e^{j\Omega_0})$, but are not tied to the absolute size and relative distance between eigenvalues. Relative bounds [69] can also help to explore the effect of perturbation but assume a rank one perturbation and well-separated eigenvalues, which is not guaranteed here. For such relative bounds, it has been shown in e.g. [70,71] that bounds for a specific eigenvalue are scaled by that eigenvalue, such that the eigenvalue perturbation can become multiplicative [72].

5.2. Impact of modelling error on eigenspaces

Assume that the m th eigenvectors associated with the m th eigenvalues $\lambda_m(e^{j\Omega})$ and $\hat{\lambda}_m(e^{j\Omega})$ are $\mathbf{q}_m(e^{j\Omega})$ and $\hat{\mathbf{q}}_m(e^{j\Omega})$, respectively. Assume further that as the sample size $N \rightarrow \infty$ the estimate-based eigenvalue $\hat{\lambda}_m(e^{j\Omega}) \rightarrow \lambda_m(e^{j\Omega})$. Due to the phase ambiguity of the eigenvectors, subspace angles or -correlations [73,74] must be used in order to compare $\mathbf{q}_m(e^{j\Omega})$ and $\hat{\mathbf{q}}_m(e^{j\Omega})$. These metrics are particularly important if eigenvalues have a C -fold algebraic multiplicity, such that e.g. $\lambda_m(e^{j\Omega_0}) = \dots = \lambda_{m+C-1}(e^{j\Omega_0})$ at a particular frequency Ω_0 , 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 ill-defined, while the subspace in which they are contained remains invariant [46]. We therefore focus on the subspaces in which eigenvectors of $\mathbf{R}(e^{j\Omega})$ and $\hat{\mathbf{R}}(e^{j\Omega})$ exist.

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

The subspace distance is an appropriate metric for the distance between two subspaces. It is defined using $\mathbf{P}_1(e^{j\Omega_0}) = \mathbf{Q}_1(e^{j\Omega_0})\mathbf{Q}_1^H(e^{j\Omega_0})$, and similarly $\hat{\mathbf{P}}_1(e^{j\Omega_0})$, for the estimated eigenvectors. In this case [73], the subspace distance is

$$\text{dist}\{\mathcal{U}_1(e^{j\Omega_0}), \hat{\mathcal{U}}_1(e^{j\Omega_0})\} = \|\mathbf{P}_1(e^{j\Omega_0}) - \hat{\mathbf{P}}_1(e^{j\Omega_0})\|_2. \quad (20)$$

Similar to above, we also partition the perturbation

$$E(e^{j\Omega_0}) = \begin{bmatrix} \underbrace{E_{11}(e^{j\Omega_0})}_{C} & \underbrace{E_{21}^H(e^{j\Omega_0})}_{M-C} \\ \underbrace{E_{21}(e^{j\Omega_0})}_{C} & \underbrace{E_{22}(e^{j\Omega_0})}_{M-C} \end{bmatrix}. \quad (21)$$

By defining the spectral distance $\delta(\Omega_0)$ between the group of C repeated eigenvalues in $A_1(e^{j\Omega_0})$ and the next-nearest neighbour in $A_2(e^{j\Omega_0})$ at a specific angular frequency Ω_0 as

$$\delta(\Omega_0) = \min_{\substack{\lambda_1(\Omega_0) \in A_1(e^{j\Omega_0}) \\ \lambda_2(\Omega_0) \in A_2(e^{j\Omega_0})}} |\lambda_1(e^{j\Omega_0}) - \lambda_2(e^{j\Omega_0})| > 0, \quad (22)$$

we find that the subspace distance in (20) is bounded such that

$$\text{dist}\{\mathcal{U}_1(e^{j\Omega_0}), \hat{\mathcal{U}}_1(e^{j\Omega_0})\} \leq \frac{4}{\delta(\Omega_0)} \|E_{21}(e^{j\Omega_0})\|_2, \quad (23)$$

as long as the overall perturbation is limited by the condition [73]

$$\|E(e^{j\Omega_0})\|_2 \leq \frac{\delta(\Omega_0)}{5}. \quad (24)$$

This can be satisfied by selecting the sample size N sufficiently large.

The above analysis for assessing the subspace distance between ground truth and sample-based eigenvectors can be applied for any $\Omega = \Omega_0$ and in turn for eigenvectors at any frequency. Similarly to the perturbation of eigenvalues, the perturbation of eigenspaces depends on the estimation error, measured by $\|E_{21}(e^{j\Omega_0})\|_2 < \|E(e^{j\Omega_0})\|_2$. 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.

6. Perturbation-related experiments

6.1. Scenarios

Model 1. To underpin the above analysis by simulations, we define a scenario as in Fig. 1 with $M = L = 2$. The ground truth eigenvalues of $R(z)$ are $\lambda_1(z) = z + 2 + z^{-1}$ and $\lambda_2(z) = -jz + 2 + jz^{-1}$, corresponding to PSDs $\lambda_1(e^{j\Omega}) = 2 + 2\cos\Omega$ and $\lambda_2(e^{j\Omega}) = 2 + 2\sin\Omega$ on the unit circle which cross at $\Omega = \frac{\pi}{4}$ and $\Omega = \frac{5\pi}{4}$ [9,75]. The eigenvectors are given by the elementary paraunitary mixing matrix $Q(z) = \mathbf{I} - \mathbf{v}\mathbf{v}^H + z^{-1}\mathbf{v}\mathbf{v}^H$ with $\mathbf{v} = [1, j]^T/\sqrt{2}$ [55].

Model 2. W.r.t. Fig. 1, we use a source model $H(z) = Q(z)F(z)$ as introduced in [16], such that $F(z)F^P(z)$ contains spectrally majorised eigenvalues, and the paraunitary $Q(z)$. The polynomial orders of both $Q(z)$ and $F(z)$ can be controlled, and $Q(z)$ can be obtained as a concatenation of elementary paraunitary matrices [13,55].

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 $\mathbf{x}[n]$, $n = 0 \dots (N-1)$ according to (7) and using a support $|\tau| \leq T_{\text{opt}}$ that minimises the overall error in (16). By performing the EVD on discrete DFT bins, we calculate a number of samples of $\hat{\lambda}_m(e^{j\Omega_k})$ and $\hat{\mathbf{q}}(e^{j\Omega_k})$ along the frequency axis for $\Omega_k = \frac{2\pi k}{K}$, $k = 0 \dots (K-1)$, which we can compare to the ground truth.

6.2. Unbiased estimator

We first demonstrate the accuracy of the variance of a sample cross-correlation sequence in (10). For a cross-correlation created by means of an innovation filter model of order 5 with a single source ($L = 1$) through the model given in Fig. 1 and (1), for $N = 10^2$, (10) is compared to the mean variance over an ensemble of size 10^5 in Figs. 4 and 5 for real- and complex-valued cases.

Generally, mean ensemble results in Figs. 4(a) and 5(a) and (b) underline that the estimator is unbiased. The theoretical variance of the estimator in (1) also matches the ensemble results very closely; since the lag support is almost the same as the sample size, the effective

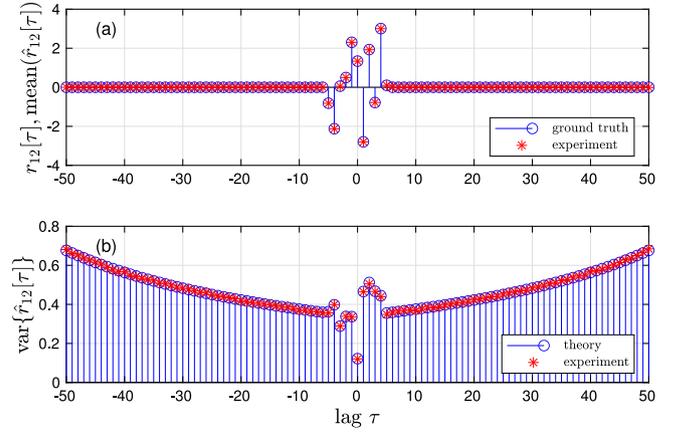


Fig. 4. (top) ground truth and mean sample cross-correlation sequence, and (bottom) its variance, both calculated according to (10) and estimated from real valued data.

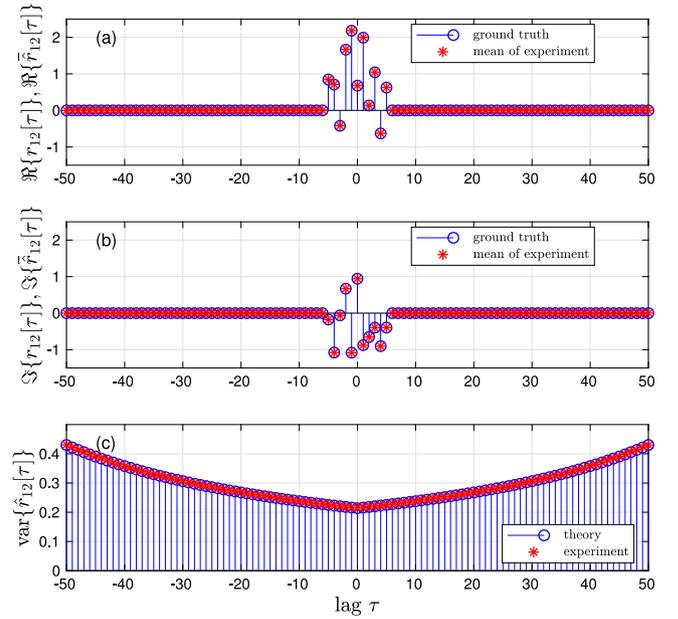


Fig. 5. Complex-valued equivalent to Fig. 4, with (top) real part, (middle) imaginary part, and (bottom) variance.

sample size of $N - |\tau|$ increases the variance for larger values of $|\tau|$ in Figs. 4(b) and 5(c). These two figures exhibit a notable difference though: the absence of the complementary cross-correlation terms $\bar{r}_{mu}[\tau]$ from (10) in the case of circularly-symmetric complex Gaussian excitation means that the specific cross-correlation-related structure over the lag range $|\tau| \leq 6$ that was present in the real-valued case in Fig. 4(b) has now disappeared in Fig. 5(c).

6.3. Modelling error

To check the accuracy of the expected estimation error (15), an $\mathbf{R}[\tau]$ of order 10^2 is generated via Model 2 in Section 6.1. For $N = 500$ and a range of supports T , Fig. 6 compares results for the theoretical values ξ_1 and ξ_2 to the distribution $p(\hat{\xi})$ of the experimental quantity $\hat{\xi}$ obtained over an ensemble of 10^4 instantiations. These match well, and also demonstrate that in this case the optimum estimated support $T_{\text{opt}} = 10$ is significantly shorter than the ground truth support $\tau_{\text{max}} = 50$. The shading in Fig. 6 and in subsequent figures is a coarse 3-D visualisation of the distributions, in addition to the more precise bounds within which 50% and 95% of the experiments fall.

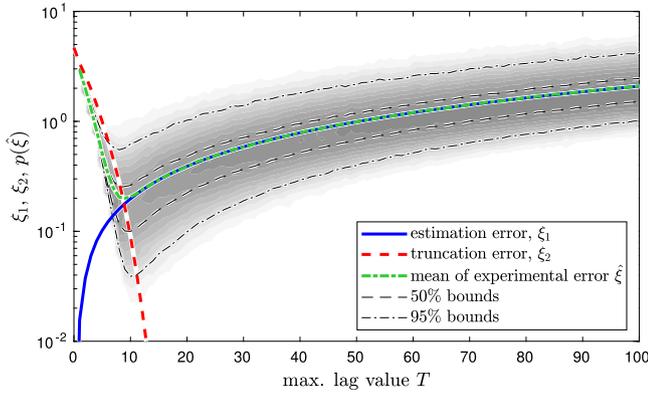


Fig. 6. Comparison of the experimentally measured distribution of the modelling error due to the truncation and expected estimation errors.

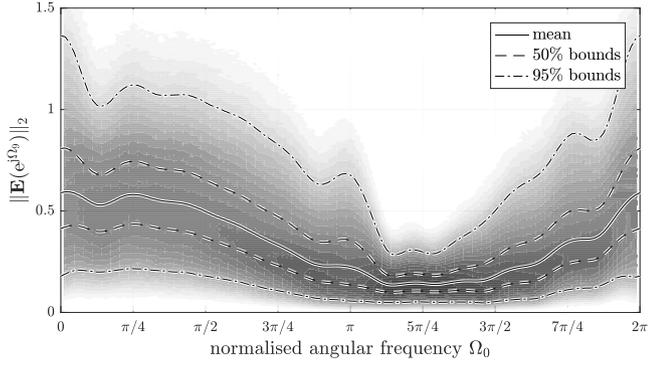


Fig. 7. Distribution of estimation error $\|E(e^{j\Omega_0})\|_2$ for Model 1.

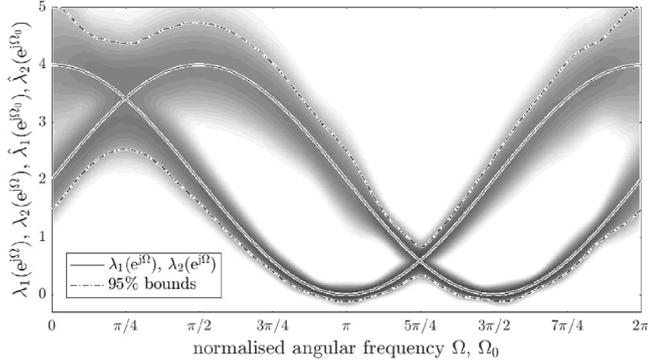


Fig. 8. Distribution of estimated eigenvalues for Model 1, with the ground truth as solid lines.

6.4. Perturbation experiments

For Model 1 in Section 6.1, the distribution of the estimation error metric $\|E(e^{j\Omega_0})\|_2$ for $N = 10^3$ over an ensemble of 10^5 simulations is shown in Fig. 7, which demonstrates the frequency-dependency of (21).

The measured distribution of $\hat{\lambda}_m(e^{j\Omega_0})$ in Fig. 8 suggests that the deviation from the ground truth does indeed depend on $\|E(e^{j\Omega_0})\|_2$ as established in (19). The shading that indicates the spread of the distribution also highlights that larger eigenvalues are more perturbed compared to smaller ones, as explored by the comments on relative bounds at the end of Section 5.1. Here, the perturbation of the eigenvalues depends on the absolute value of $\lambda_m(e^{j\Omega_0})$, with the estimation error having an approximately multiplicative effect.

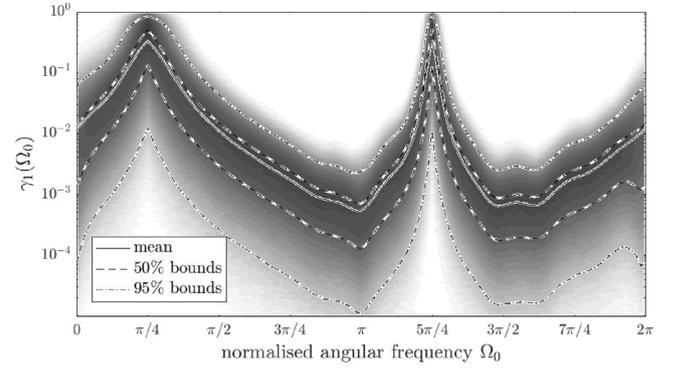


Fig. 9. Distribution of $\gamma_1(\Omega_0)$ in (25) for Model 1.

To measure how close the subspaces spanned by the ground truth and estimated eigenvectors $q_m(e^{j\Omega_0})$ and $\hat{q}_m(e^{j\Omega_0})$ are, we use a modified version of the subspace correlation,

$$\gamma_m(\Omega_0) = 1 - |\hat{q}_m^H(e^{j\Omega_0})q_m(e^{j\Omega_0})|, \quad (25)$$

where the Hermitian angle is insensitive to the eigenvectors' arbitrary phase shifts. Small values of $\gamma_m(\Omega_0)$ mean that subspaces are aligned, while $\gamma_m(\Omega_0) = 1$ indicates orthogonality. In Fig. 9, the measured distribution of $\gamma_1(\Omega_0)$ shows higher subspace alignment where eigenvalues, according to Fig. 8, are clearly separated. Near algebraic multiplicities, a mismatch in subspaces arises as analysed in (23).

6.5. Bounds on eigenvalue and eigenspace perturbations

Normalised Weyl Bound. Assuming that the upper and lower limits in (17) are separated, i.e. that $\lambda_{\max}^E(e^{j\Omega_0}) \neq \lambda_{\min}^E(e^{j\Omega_0})$, we can define a modified eigenvalue error

$$\epsilon_{\text{Weyl},m}(\Omega_0) = \frac{\Delta\lambda_m(e^{j\Omega_0}) - \lambda_{\min}^E(e^{j\Omega_0})}{\lambda_{\max}^E(e^{j\Omega_0}) - \lambda_{\min}^E(e^{j\Omega_0})}, \quad (26)$$

where $\Delta\lambda_m(e^{j\Omega_0}) = \hat{\lambda}_m(e^{j\Omega_0}) - \lambda_m(e^{j\Omega_0})$. The modified Weyl bound is normalised to $0 \leq \epsilon_{\text{Weyl},m}(\Omega_0) \leq 1$.

Normalised Hoffman-Wielandt Bound. Based on (18), we define

$$\epsilon_{\text{HW}}(\Omega_0) = \frac{\sum_m (\Delta\lambda_m(e^{j\Omega_0}))^2}{\|E(e^{j\Omega_0})\|_F^2}. \quad (27)$$

Therefore the Hoffman-Wielandt bound for this normalised quantity becomes $0 \leq \epsilon_{\text{HW}}(\Omega_0) \leq 1$.

Normalised Bauer–Fike Bound. Using (19), we define

$$\epsilon_{\text{BF},m}(\Omega_0) = \frac{|\Delta\lambda_m(e^{j\Omega_0})|}{\|E(e^{j\Omega_0})\|_2}. \quad (28)$$

Similarly to the previous two bounds, the normalised Bauer–Fike quantity satisfies $0 \leq \epsilon_{\text{BF},m}(\Omega_0) \leq 1$.

Model 1 of Section 6.1 generates the distributions for each of the above quantities in Figs. 10–12. For the normalised Weyl bounds $\epsilon_{\text{Weyl},m}(\Omega_0)$ in (26), Fig. 10(a) shows the case of $m = 1$ and Fig. 10(b) the case of $m = 2$. Interestingly, the ensemble values of $\epsilon_{\text{Weyl},m}(\Omega_0)$ cover the ranges $\frac{1}{2} \leq \epsilon_{\text{Weyl},1}(\Omega_0) \leq 1$ and $0 \leq \epsilon_{\text{Weyl},2}(\Omega_0) \leq \frac{1}{2}$. When the two eigenvalues are maximally separated, the distribution of $\epsilon_{\text{Weyl},m}(\Omega_0)$ across these ranges appears nearly uniform, while at sample points $\Omega_0 = \{\frac{\pi}{4}, \frac{5\pi}{4}\}$ where the eigenvalues possess an algebraic multiplicities of two, $\epsilon_{\text{Weyl},2}(\Omega_0)$ satisfies the lower bound while $\epsilon_{\text{Weyl},1}(\Omega_0)$ takes on the upper bound. The bounds are therefore satisfied; the usefulness of any bound increases the tighter it is, which here is the case at algebraic multiplicities.

The normalised Hoffman-Wielandt bound (27) in Fig. 11 behaves similarly to the normalised Weyl bound, and appears to be uniformly distributed across its expected range when the eigenvalues are well

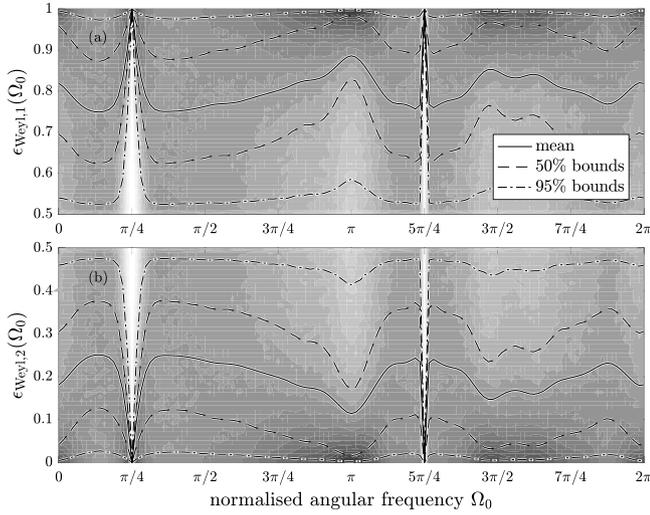


Fig. 10. Distribution of the normalised Weyl bound (26) for $m = 1, 2$ for the example given Model 1 in Section 6.1.

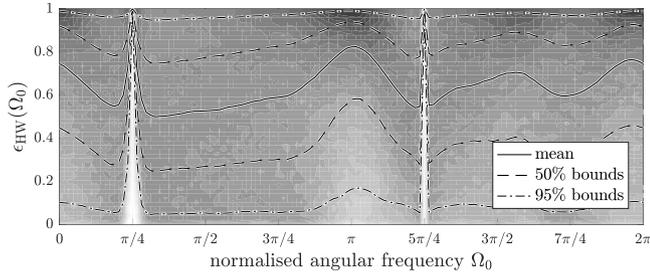


Fig. 11. Distribution of normalised Hoffman-Wielandt bound (27) for the scenario in Section 6.1.

separated, and takes on values close to the upper bound when eigenvalues match for $\Omega_0 = \{\frac{\pi}{4}, \frac{5\pi}{4}\}$. Fig. 12(a) and (b) show the normalised Bauer-Fike bounds for $m = 1$ and $m = 2$, respectively. Similar to the other bounds, at frequencies where eigenvalues are well separated, the distribution of the values $\epsilon_{BF,m}(\Omega_0)$ is approximately uniform and the bounds are loose. However for frequencies where the eigenvalues possess non-trivial algebraic multiplicities, the relative bound $\epsilon_{BF,m}(\Omega_0)$ is tight and therefore particularly useful.

The bounds on the eigenspaces described in Section 5.2 can be demonstrated using similar metrics as for the eigenvalues. Using a normalised version of (23),

$$\psi(\Omega_0) = \frac{\delta(\Omega_0) \text{dist}\{\mathcal{U}'_1(e^{j\Omega_0}), \hat{\mathcal{U}}'_1(e^{j\Omega_0})\}}{4 \|\mathbf{E}_{21}(e^{j\Omega_0})\|_2}, \quad (29)$$

such that $0 \leq \psi(\Omega_0) \leq 1$ as long as (24) is satisfied. Fig. 13(a) shows that condition (24) is violated near $\Omega_0 = \{\frac{\pi}{4}, \frac{5\pi}{4}\}$. As a result, the distribution of the normalised eigenspace perturbation does not satisfy $\psi(\Omega_0) \leq 1$ near these frequency points as evident from Fig. 13(b). Thus, the subspace perturbation bound is only useful where eigenvalues are sufficiently well separated, and breaks down near frequencies where the eigenvalues possess non-trivial algebraic multiplicities. Where eigenvalues are sufficiently separated, the bounds may be translated to performance bounds of signal processing algorithms, such as demonstrated for a broadband angle of arrival estimation application using a polynomial multiple signal classification algorithm in [76].

7. Support estimation

Previous sections have explored the impact that modelling errors due to (i) truncation and (ii) estimation errors have on the bin-wise

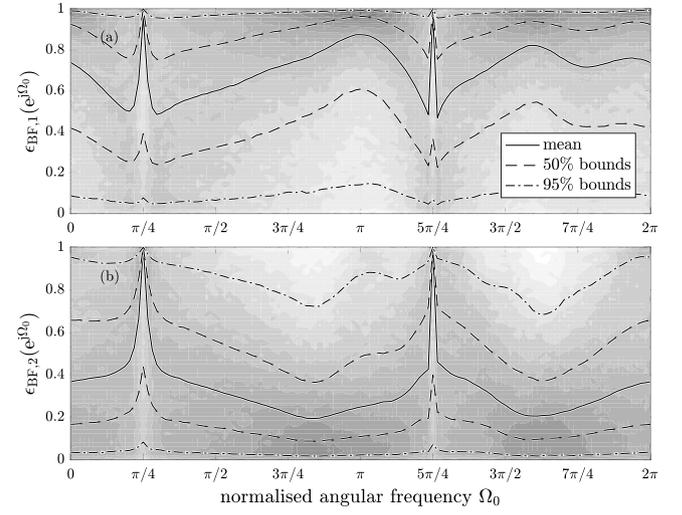


Fig. 12. Distribution of the normalised Bauer-Fike bound (28) for (a) $m = 1$ and (b) $m = 2$ for Model 1 in Section 6.1.

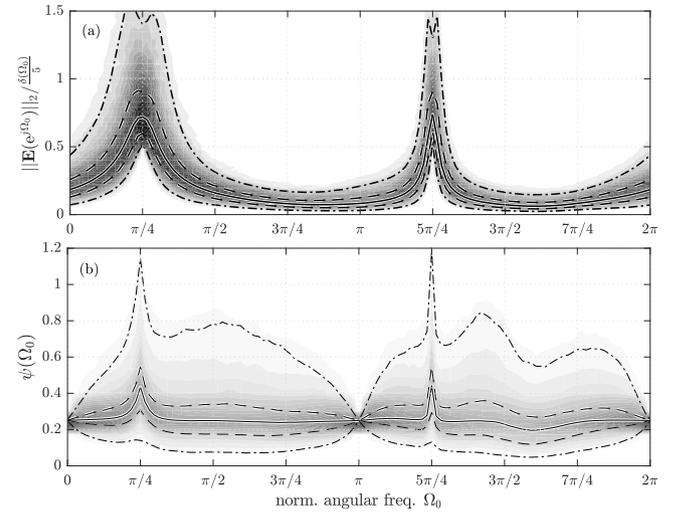


Fig. 13. (a) Distribution of the overall normalised perturbation $\|\mathbf{E}(e^{j\Omega_0})\|_2 / \frac{\delta(\Omega_0)}{5}$, where values greater than one indicate that the condition (24) has been violated, and (b) the eigenspace perturbation bound (23) for the scenario in Section 6.1.

perturbation – with some bounds – of the eigenvalues and eigenspaces of a space–time covariance matrix. Therefore, in the section below, we want to concentrate on a suitable estimation of the support of the space–time covariance, which aims to minimise the modelling error and thereby the above perturbations, based on an initial idea presented in [57].

7.1. Preliminary considerations

The ensemble-optimum support defined in (16) and illustrated in Fig. 2 reflects the optimum adjustment across the entire ensemble under the assumption that the ground truth space–time covariance $\mathbf{R}[\tau]$ is known. Firstly, in practice $\mathbf{R}[\tau]$ is unknown, and the computation cannot be performed. Secondly for individual estimates the best support may differ from the ensemble-optimum value, which the following example briefly explores.

Example 3. For the scenario of Example 1, Fig. 6 highlights that for different percentiles of the distribution, both the minimum modelling error $\hat{\xi}$ as well as the optimum support for a particular estimate can

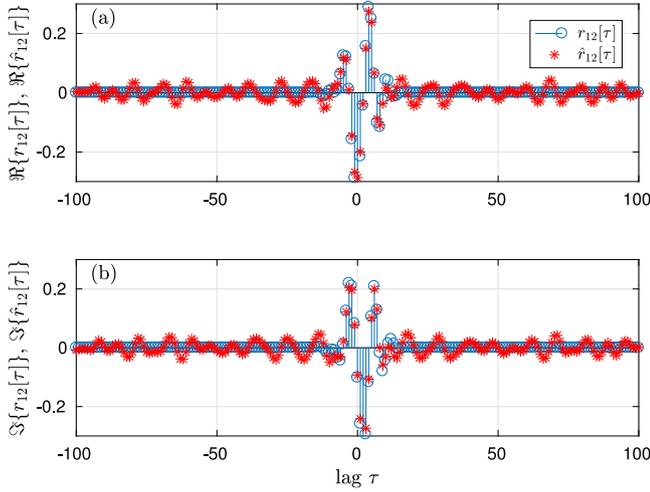


Fig. 14. (a) Real and (b) imaginary part of a cross-correlation sequence $r_{12}[\tau]$ and its estimate $\hat{r}_{12}[\tau]$ based on $N = 10^3$ snapshots of data.

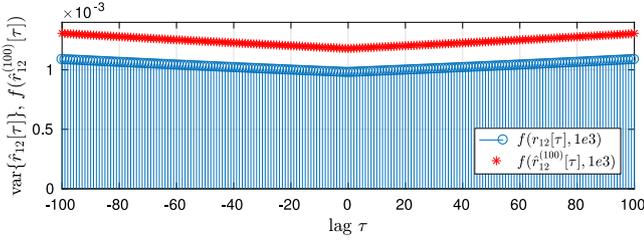


Fig. 15. Variance of the sample cross-correlation estimate shown in Fig. 14, and an approximation based on a sample estimate $\hat{r}_{12}[\tau]$ for $|\tau| \leq 100$.

deviate from the ensemble-optimum values. Specifically considering the 95% range of experiments, for the distribution shown Fig. 6 the support can vary between 8 and 10 lags, where $\min\{\hat{\xi}\}$ takes on values between approximately $4 \cdot 10^{-2}$ and $6 \cdot 10^{-1}$. Δ

Therefore, in addition to the ensemble-optimum values T_{opt} and $\min \xi$ which we have discussed in previous sections, we are therefore also interested in values $T_{\text{opt,p}}$ and its estimate $\hat{T}_{\text{opt,p}}$ which are based on a single measurement or ‘probe’ (motivating the subscript ‘p’) that contains N snapshots of data. Note that $T_{\text{opt,p}}$ is based on the knowledge of both $\mathbf{R}[\tau]$ and the specific estimate $\hat{\mathbf{R}}[\tau]$. Hence our aim is to find an estimate $\hat{T}_{\text{opt,p}}$ that minimises $\hat{\xi}$ as best as possible and is solely based on $\hat{\mathbf{R}}[\tau]$. Based on (10), the variance of a cross-correlation estimate $\text{var}\{\hat{r}_{m\mu}[\tau]\}$ is a function of both the ground truth cross-correlation $r_{m\mu}[\tau]$ and the sample size N , i.e.

$$\text{var}\{\hat{r}_{m\mu}[\tau]\} = f(r_{m\mu}[\tau], N). \quad (30)$$

Let $\hat{r}_{m\mu}^{(T)}[\tau]$ be a sample cross-correlation estimate with support $2T + 1$, such that $\hat{r}_{m\mu}^{(T)}[\tau] = 0 \forall |\tau| > T$. For a sufficiently large values of N and N/T , (30) can be very roughly approximated by

$$f(r_{m\mu}[\tau], N) \sim f(\hat{r}_{m\mu}^{(T)}[\tau], N),$$

i.e. instead of the inaccessible ground truth, the variance of the estimation will be based on the sample estimate itself.

Example 4. Fig. 14 provides an example for $r_{12}[\tau] \in \mathbb{C}$ and a sample estimate $\hat{r}_{12}^{(100)}[\tau]$. In addition to $\text{var}\{\hat{r}_{12}[\tau]\} = f(r_{12}[\tau], N)$, Fig. 15 shows $f(\hat{r}_{12}^{(100)}[\tau], N)$, which behaves very similarly, and therefore may provide an approximation to $\text{var}\{\hat{r}_{12}[\tau]\}$. Δ

7.2. Support estimation approach

To estimate a suitable support, recall from e.g. (14) that the mean square modelling error consists of an estimation error and a truncation error term. The estimation error ξ_1 , in the case of a limitation to T lags, can now be approximated based on $f(\hat{r}_{m\mu}^{(T)}[\tau], N)$. To exclude estimation error behaviour for small lags, particularly in case of real-valued Gaussian data as in Fig. 4, we focus on a portion of ξ_1 in the tail section of $\text{var}\{\hat{r}_{m\mu}[\tau]\}$ of length T_0 via

$$\chi_1[T] = \sum_{\tau=T+1}^{T+T_0} \sum_{m,\mu} f(\hat{r}_{m\mu}^{(T)}[\tau], N). \quad (31)$$

Note that $\chi_1[T]$ roughly approximates an estimation error portion for an estimate of larger support, $\hat{r}_{m\mu}^{(T+T_0)}[\tau]$, measured over the range $T \leq |\tau| \leq T + T_0$.

If we consider the energy in the tail section of an estimate $\hat{r}_{m\mu}^{(T+T_0)}[\tau]$ of increased support, then

$$\chi_2[T] = \sum_{\tau=T+1}^{T+T_0} \sum_{m,\mu} f(\hat{r}_{m\mu}^{(T+T_0)}[\tau], N). \quad (32)$$

will contain approximately the same portion of the estimation error, but also include a term that would otherwise have been truncated in $\hat{r}_{m\mu}^{(T)}[\tau]$ if T was too short as support. Therefore the comparison of $\chi_1[T]$ and $\chi_2[T]$ enables us to detect if a truncation error has been incurred, in the case that the support T of $\hat{r}_{m\mu}^{(T)}[\tau]$ was too restrictive. Thus, we utilise the variable $\gamma[T] = \chi_2[T]/\chi_1[T]$ in order to estimate $\hat{T}_{\text{opt,p}}$ for the optimum support T_{opt} : a value $\gamma[T] \approx 1$ would indicate that an estimation error has been incurred, while $\gamma[T] \gg 1$ implies an additional truncation error.

We now check for which value T the behaviour of $\gamma[T]$ changes from the truncation-dominated part in e.g. Fig. 4, to a portion where only an estimation error is incurred. For this, assume that we inspect $\gamma[T]$ for sufficiently high values of T , say beyond some value T_1 , which is bound to exceed the support of $\mathbf{R}[\tau]$, or at least ensure that any elements of $\mathbf{R}[\tau]$ will be small compared to an estimation error. Over a range of values, say over $T_1 \leq \tau \leq 2T_1$, we determine (i) a linear least squares fit $\gamma_{\text{LF}}[T]$ to $\gamma[T]$, such that $\sum_{T=T_1}^{2T_1} |\gamma[T] - \gamma_{\text{LF}}[T]|^2$ is minimised, and (ii) the standard deviation σ of $\gamma[T]$ w.r.t. this linear fit $\gamma_{\text{LF}}[T]$. We then check whether for smaller values of T , i.e. $T < T_1$, values of $\gamma[T]$ fall outside a pipe of width $K\sigma$ around this linear fit, and estimate the optimum support as

$$\hat{T}_{\text{opt,p}} = \arg \max_T T \quad \text{s.t. } \gamma[T] > \gamma_{\text{LF}}[T] + K\sigma, \quad (33)$$

with K yet to be determined.

Example 5. The approach is exemplified for the scenario of Example 1 in Fig. 16(a), where $T_1 = 50$, and $\gamma[T]$ as well as the linear fit with a pipe of width $K\sigma$, here for $K = 5$, are shown. In this case, $\hat{T}_{\text{opt,p}} = 14$. With exact knowledge of $\mathbf{R}[\tau]$, and given the sample estimate $\hat{\mathbf{R}}[\tau]$, we can exactly determine the modelling and truncation errors $\hat{\xi}$ and ξ_2 , and therefore the estimation error $\hat{\xi}_1$, displayed in Fig. 16(b). The search for a minimum of $\hat{\xi}$ for this example yields $T_{\text{opt,p}} = 10$. A more accurate value for $\hat{T}_{\text{opt,p}}$ could have been obtained for a wider pipe, i.e. for a larger value of K . Δ

While there are statistical approaches for the testing of ratios of variances such as the F -test [77], their application is not straightforward, as e.g. successive ratios $\gamma[T]$ in Fig. 16(a) are not independent. Therefore, K is maintained as a variable, and we will explore the effect that different values have on the support estimation below.

7.3. Support estimation ensemble simulation

To briefly demonstrate the accuracy and sensitivity of the proposed support estimation, we consider a numerical example of a randomised matrix $\mathbf{R}[\tau]$ of spatial dimension $M = 6$ and $\mathbf{R}[\tau] = \mathbf{0}$ for $|\tau| >$

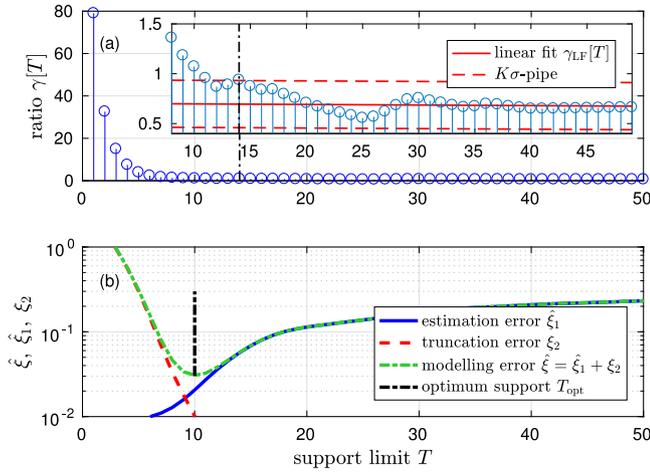


Fig. 16. (a) Ratio $\gamma[T]$ with threshold selection by a $K\sigma$ -pipe based on a single sample set, and (b) optimum support based on the ground truth $\mathbf{R}[\tau]$ and estimate $\hat{\mathbf{R}}[\tau]$ with support $2T + 1$.

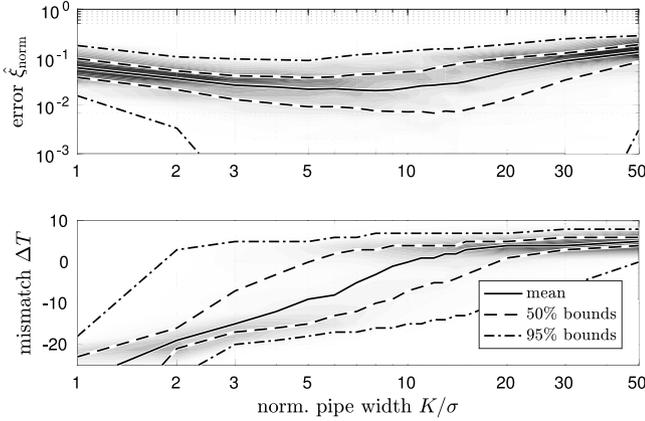


Fig. 17. Distribution of metrics as a function of the pipe width K when estimating the support of $\hat{\mathbf{R}}[\tau]$.

T_0 with $T_0 = 30$. As performance metrics over an ensemble of 10^3 randomisations $\mathbf{R}[\tau]$, constructed using the source model in Fig. 1 using parameters from [16], we utilise two performance metrics. Firstly, we assess the support mismatch ΔT

$$\Delta T = T_{\text{Opt},p} - \hat{T}_{\text{Opt},p}. \quad (34)$$

A positive value for ΔT means underestimation, $\Delta T < 0$ overestimation of T_{opt} . The effect of over- vs underestimation on the modelling error is unequal, see e.g. Fig. 6. Secondly, a normalised modelling error

$$\hat{\xi}_{\text{norm}} = \frac{\sum_{m,\mu,\tau} f(\hat{r}_{m,\mu}^{(T_{\text{Opt},p})}, N) - f(\hat{r}_{m,\mu}^{(\hat{T}_{\text{Opt},p)}}, N)}{\sum_{m,\mu,\tau} f(r_{m\mu}[\tau], N)} \quad (35)$$

creates a metric that is somewhat independent of a specific realisation $\mathbf{R}[\tau]$, such that results become comparable across the ensemble.

The distribution for the performance metrics are shown in Fig. 17 as function of different pipe widths K . The normalised modelling error $\hat{\xi}_{\text{norm}}$ in Fig. 17(a) indicates that the result is relatively insensitive to the selection of K , although the metric shows a minimum at around $K = 8$. For this value, the mean of the distribution of ΔT is approximately zero, i.e. the estimate $\hat{T}_{\text{Opt},p}$ coincides with the optimum support $T_{\text{Opt},p}$.

8. Conclusions

This paper has addressed the estimation of the space–time covariance matrix, and the perturbation of this estimate itself as well as its

parahermitian matrix eigenvalue decomposition due to finite sample size effects. Through analysis of biased and unbiased estimators, this paper has shown the problem of rank-deficiency when estimating a space–time covariance matrix using a biased estimator and the effect of truncating these covariance matrices. For unbiased estimates, a link between the sample size and perturbations of the space–time covariance matrix has been presented, where the eigenvalue perturbation is affected by the norm of the discrepancy between ground truth and estimated space–time covariance matrices and the eigenspace perturbation is additionally affected by the distance between ground truth and estimated eigenvalues.

We have further presented bounds for these perturbations to both the eigenvalues and eigenspaces. In the case of the eigenvalues, the bounds are particularly useful near non-trivial algebraic multiplicities, where experiments indicate that they are tight. The opposite is the case for the eigenspaces — here the near algebraic multiplicities, the condition under which bounds can be derived are violated, and the bounds may offer less applicability, for example when assessing subspace leakage for subspace-based detection [78–80] or applications such that the multiple-signal classification (MUSIC) algorithm [4,81]. In this context, the analysis and bounds presented in this paper has been used to assess the impact of estimation errors on a polynomial extension of the MUSIC algorithm due to subspace perturbations in [76]. It is also key to assess subspace perturbations for broadband subspace methods to detect weak transient broadband signals [82–84], and the basis of studying effects of perturbed eigenvalues across frequency bins [85].

In order to minimise perturbation effects on the EVD factors, we have suggested a method to optimise the support when estimating the space–time covariances. While the variance of the estimate depends on the ground truth covariance, we have replaced this term by the estimated space–time covariance, and demonstrated that a useful support estimation is possible, with impact across the applications that we have summarised in the introduction to this paper.

CRedit authorship contribution statement

Connor Delaosa: Writing – review & editing, Writing – original draft, Validation, Software, Formal analysis. **Jennifer Pestana:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Ian K. Proudler:** Writing – review & editing, Formal analysis. **Stephan Weiss:** Writing – review & editing, Writing – original draft, Software, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Stephan Weiss reports financial support was provided by Engineering and Physical Sciences Research Council. Connor Delaosa reports financial support was provided by Defence Science and Technology Laboratory. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We very gratefully acknowledge the insightful comments and corrections by the anonymous reviewers. We further would like to thank Nicolas J. Goddard of Dstl and Dr Samuel Somasundaram of Thales Underwater Systems for triggering this research with their questions. C. Delaosa was supported by Dstl and a John Anderson Research Award while working at the University of Strathclyde. The work of S. Weiss was supported in parts by the Engineering and Physical Sciences Research Council (EPSRC), United Kingdom Grant number EP/S000631/1 and the MOD University Defence Research Collaboration in Signal Processing.

Data availability

Data will be made available on request.

References

- [1] B.D. Van Veen, K.M. Buckley, Beamforming: a versatile approach to spatial filtering, *IEEE ASSP Mag.* 5 (2) (1988) 4–24.
- [2] H.L. Van Trees, *Detection, Estimation and Modulation Theory: Optimum Array Processing*, Wiley, New York, 2002.
- [3] S. Haykin, *Adaptive Filter Theory*, second ed., Prentice Hall, Englewood Cliffs, 1991.
- [4] R.O. Schmidt, Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Prop.* 34 (3) (1986) 276–280.
- [5] P. Vaidyanathan, Theory of optimal orthonormal subband coders, *IEEE Trans. Signal Process.* 46 (6) (1998) 1528–1543.
- [6] I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.
- [7] V. Neo, S. Redif, J. McWhirter, J. Pestana, I. Proudler, S. Weiss, P. Naylor, Polynomial eigenvalue decomposition for multichannel broadband signal processing, *IEEE Signal Process. Mag.* 40 (7) (2023) 18–37.
- [8] J.G. McWhirter, P.D. Baxter, T. Cooper, S. Redif, J. Foster, An EVD algorithm for para-Hermitian polynomial matrices, *IEEE Trans. Signal Process.* 55 (5) (2007) 2158–2169.
- [9] S. Weiss, J. Pestana, I.K. Proudler, On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix, *IEEE Trans. Signal Process.* 66 (10) (2018) 2659–2672.
- [10] S. Weiss, J. Pestana, I.K. Proudler, F.K. Coutts, Corrections to on the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix, *IEEE Trans. Signal Process.* 66 (23) (2018) 6325–6327.
- [11] A. Tkacenko, P. Vaidyanathan, Iterative greedy algorithm for solving the fir paraunitary approximation problem, *IEEE Trans. Signal Process.* 54 (1) (2006) 146–160.
- [12] A. Tkacenko, Approximate eigenvalue decomposition of para-Hermitian systems through successive FIR paraunitary transformations, in: *IEEE ICASSP*, Dallas, TX, 2010, pp. 4074–4077.
- [13] S. Redif, J. McWhirter, S. Weiss, Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition, *IEEE Trans. Signal Process.* 59 (11) (2011) 5253–5264.
- [14] M. Tohidian, H. Amindavar, A.M. Reza, A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices, *EURASIP J. Adv. Signal Process.* 2013 (1) (2013) 1–16.
- [15] J. Corr, K. Thompson, S. Weiss, J. McWhirter, S. Redif, I. Proudler, Multiple shift maximum element sequential matrix diagonalisation for parahermitian matrices, in: *IEEE SSP*, Gold Coast, Australia, 2014, pp. 312–315.
- [16] S. Redif, S. Weiss, J. McWhirter, Sequential matrix diagonalization algorithms for polynomial EVD of parahermitian matrices, *IEEE Trans. Signal Process.* 63 (1) (2015) 81–89.
- [17] Z. Wang, J.G. McWhirter, J. Corr, S. Weiss, Multiple shift second order sequential best rotation algorithm for polynomial matrix EVD, in: *23rd EUSIPCO*, Nice, France, 2015, pp. 844–848.
- [18] V.W. Neo, P.A. Naylor, Second order sequential best rotation algorithm with householder reduction for polynomial matrix eigenvalue decomposition, in: *IEEE ICASSP*, Brighton, UK, 2019, pp. 8043–8047.
- [19] S. Weiss, I.K. Proudler, F.K. Coutts, Eigenvalue decomposition of a parahermitian matrix: extraction of analytic eigenvalues, *IEEE Trans. Signal Process.* 69 (2021) 722–737.
- [20] S. Weiss, I. Proudler, F. Coutts, F. Khattak, Eigenvalue decomposition of a parahermitian matrix: extraction of analytic eigenvectors, *IEEE Trans. Signal Process.* 71 (2023) 1642–1656.
- [21] F.A. Khattak, I.K. Proudler, S. Weiss, Scalable analytic eigenvalue extraction algorithm, *IEEE Access* 12 (2024) 166652–166659.
- [22] S. Redif, J. McWhirter, P. Baxter, T. Cooper, Robust broadband adaptive beamforming via polynomial eigenvalues, in: *OCEANS*, Boston, MA, 2006, pp. 1–6.
- [23] S. Weiss, S. Bendoukha, A. Alzin, F. Coutts, I. Proudler, J. Chambers, MVDR broadband beamforming using polynomial matrix techniques, in: *23rd EUSIPCO*, Nice, France, 2015, pp. 839–843.
- [24] A. Alzin, F. Coutts, J. Corr, S. Weiss, I. Proudler, J. Chambers, Polynomial matrix formulation-based capon beamformer, in: *IMA Conf. Math. in Signal Process.* Birmingham, UK, 2016.
- [25] V.W. Neo, E. d’Olne, A.H. Moore, P.A. Naylor, Fixed beamformer design using polynomial eigenvalue decomposition, in: *IWAENC*, 2022, pp. 1–5.
- [26] S. Weiss, C.H. Ta, C. Liu, A wiener filter approach to the design of filter bank based single-carrier precoding and equalisation, in: *IEEE ISPLC*, Pisa, Italy, 2007, pp. 493–498.
- [27] V. Neo, C. Evers, S. Weiss, P. Naylor, Signal compaction using polynomial EVD for spherical array processing with applications, *IEEE/ ACM Trans. ASPL* 31 (2023) 3537–3549.
- [28] S. Redif, S. Weiss, J. McWhirter, Relevance of polynomial matrix decompositions to broadband blind signal separation, *Signal Process.* 134 (2017) 76–86.
- [29] S. Weiss, N.J. Goddard, S. Somasundaram, I.K. Proudler, P.A. Naylor, Identification of broadband source-array responses from sensor second order statistics, in: *SSPD*, London, UK, 2017, pp. 1–5.
- [30] M. Alrmah, S. Weiss, S. Lambbotharan, An extension of the MUSIC algorithm to broadband scenarios using polynomial eigenvalue decomposition, in: *EUSIPCO*, Barcelona, Spain, 2011, pp. 629–633.
- [31] S. Weiss, M. Alrmah, S. Lambbotharan, J. McWhirter, M. Kaveh, Broadband angle of arrival estimation methods in a polynomial matrix decomposition framework, in: *IEEE CAMSAP*, 2013, pp. 109–112.
- [32] A. Hogg, V. Neo, S. Weiss, C. Evers, P. Naylor, A polynomial eigenvalue decomposition MUSIC approach for broadband sound source localization, in: *IEEE WASPAA*, New Paltz, NY, 2021.
- [33] M. Alrmah, J. Corr, A. Alzin, K. Thompson, S. Weiss, Polynomial subspace decomposition for broadband angle of arrival estimation, in: *Sensor Signal Processing for Defence*, Edinburgh, Scotland, 2014, pp. 1–5.
- [34] F. Coutts, K. Thompson, S. Weiss, I. Proudler, Impact of fast-converging PEVD algorithms on broadband AoA estimation, in: *SSPD*, London, UK, 2017, pp. 1–5.
- [35] J. Foster, J.G. McWhirter, J. Chambers, Limiting the order of polynomial matrices within the SBR2 algorithm, in: *IMA Conf. Math. in Signal Process.* Cirencester, UK, 2006.
- [36] C.H. Ta, S. Weiss, Shortening the order of paraunitary matrices in SBR2 algorithm, in: *ICICSP*, Singapore, 2007, pp. 1–5.
- [37] J. Corr, K. Thompson, S. Weiss, I. Proudler, J. McWhirter, Row-shift corrected truncation of paraunitary matrices for PEVD algorithms, in: *EUSIPCO*, Nice, France, 2015, pp. 849–853.
- [38] J. Corr, K. Thompson, S. Weiss, I. Proudler, J. McWhirter, Shortening of paraunitary matrices obtained by polynomial eigenvalue decomposition algorithms, in: *Sensor Signal Processing for Defence*, Edinburgh, Scotland, 2015.
- [39] J. Corr, K. Thompson, S. Weiss, I. Proudler, J. McWhirter, Impact of source model matrix conditioning on PEVD algorithms, in: *IET/EURASIP Intelligent Signal Processing*, London, UK, 2015.
- [40] C. Delaosa, F.K. Coutts, J. Pestana, S. Weiss, Impact of space–time covariance estimation errors on a parahermitian matrix EVD, in: *10th IEEE Workshop on Sensor Array and Multichannel Signal Processing*, 2018, pp. 1–5.
- [41] F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Courant Institute of Mathematical Sciences, New York University, New York, 1954.
- [42] T. Kato, *Perturbation Theory for Linear Operators*, in: *Die Grundlehren Der Mathematischen Wissenschaften in Einzeldarstellungen*, 132, Springer, Berlin, 1966.
- [43] G.W. Stewart, J.-g. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.
- [44] G.H. Golub, C.F. Van Loan, *Matrix Computations*, fourth ed., John Hopkins University Press, Baltimore, Maryland, 2013.
- [45] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.
- [46] T. Kato, *Perturbation Theory for Linear Operators*, Springer, 1980.
- [47] F. Rellich, Störungstheorie der Spektralzerlegung. I. Mitteilung. Analytische Störung der isolierten Punkteigenwerte eines beschränkten Operators, *Math. Ann.* 113 (1937) DC–DCXIX.
- [48] R.L. Anderson, Distribution of the serial correlation coefficient, *Ann. Math. Stat.* 13 (1) (1942) 1–13.
- [49] H. Goldsmith, The exact distributions of the serial correlation coefficients and an evaluation on some approximate distributions, *J. Stat. Comput. Simul.* 5 (2) (1977) 115–134.
- [50] S. Kay, The effect of sampling rate on autocorrelation estimation, *IEEE Trans. ASSP* 29 (4) (1981) 859–867.
- [51] A. Steinhart, J. Makhoul, On the autocorrelation of finite-length sequences, *IEEE Trans. ASSP* 33 (6) (1985) 1516–1520.
- [52] D.D. Ariananda, G. Leus, Compressive wideband power spectrum estimation, *IEEE Trans. Signal Process.* 60 (9) (2012) 4775–4789.
- [53] R.K. Mallik, The pseudo-Wishart distribution and its application to MIMO systems, *IEEE Trans. Inform. Theory* 49 (10) (2003) 2761–2769.
- [54] L. Zhuang, A.T. Walden, Sample mean versus sample Fréchet mean for combining complex Wishart matrices: A statistical study, *IEEE Trans. Signal Process.* 65 (17) (2017) 4551–4561.
- [55] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.
- [56] C. Delaosa, J. Pestana, N.J. Goddard, S. Somasundaram, S. Weiss, Sample space–time covariance matrix estimation, in: *IEEE ICASSP*, Brighton, UK, 2019, pp. 8033–8037.
- [57] C. Delaosa, J. Pestana, N.J. Goddard, S.D. Somasundaram, S. Weiss, Support estimation of a sample space–time covariance matrix, in: *SSPD*, Brighton, UK, 2019.
- [58] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, third ed., McGraw-Hill, New York, 1991.
- [59] B. Picinbono, On circularity, *IEEE Trans. Signal Process.* 42 (12) (1994) 3473–3482.
- [60] G. Barbarino, V. Noferini, On the Rellich eigendecomposition of para-Hermitian matrices and the sign characteristics of $*$ -palindromic matrix polynomials, *Linear Algebra Appl.* 672 (2023) 1–27.

- [61] S. Weiss, I.K. Proudler, G. Barbarino, J. Pestana, J.G. McWhirter, On properties and structure of the analytic singular value decomposition, *IEEE Trans. Signal Process.* 72 (2024) 2260–2275.
- [62] W.A. Gardner, *Introduction to Random Process: with Application to Signals and Systems*, second ed., McGraw-Hill, 1990.
- [63] D.R. Brillinger, *Time Series: Data Analysis and Theory*, SIAM, 2001.
- [64] W. Bär, F. Dittrich, Useful formula for moment computation of normal random variables with nonzero means, *IEEE Trans. Autom. Control* 16 (3) (1971) 263–265.
- [65] J.M. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications, *Proc. IEEE* 79 (3) (1991) 278–305.
- [66] P.J. Schreier, L.L. Scharf, *Statistical Signal Processing of Complex-Valued Data. The Theory of Improper and Non-Circular Signals*, Cambridge University Press, 2010.
- [67] N.J. Higham, N. Strabić, V. Šego, Restoring definiteness via shrinking, with an application to correlation matrices with a fixed block, *SIAM Rev.* 58 (2) (2016) 245–263.
- [68] F. Wen, L. Chu, R. Ying, P. Liu, Fast and positive definite estimation of large covariance matrix for high-dimensional data analysis, *IEEE Trans. Big Data* (2019) 1.
- [69] I.C.F. Ipsen, B. Nadler, Refined perturbation bounds for eigenvalues of Hermitian and non-Hermitian matrices, *SIAM J. Matrix Anal. Appl.* 31 (1) (2009) 40–53.
- [70] R.-C. Li, Relative perturbation theory III: More bounds on eigenvalue variation, *Linear Algebra Appl.* 266 (1997) 337–345.
- [71] S.C. Eisenstat, I.C.F. Ipsen, Three absolute perturbation bounds for matrix eigenvalues imply relative bounds, *SIAM J. Matrix Anal. Appl.* 20 (1998).
- [72] W.L. Jianxin Chen, S. Pang, Multiplicative relative perturbation bounds of eigenvalues for diagonalizable matrices, *Int. J. Comput. Math.* 88 (14) (2011) 3061–3068.
- [73] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., John Hopkins University Press, Baltimore, Maryland, 1996.
- [74] J. Moshier, R. Leahy, Source localization using Recursively Applied and Projected (RAP) MUSIC, *IEEE Trans. Signal Process.* 47 (2) (1999) 332–340.
- [75] S. Icart, P. Comon, Some properties of laurent polynomial matrices, in: 9th IMA Conf. Math. in Signal Process. Birmingham, UK, 2012.
- [76] C. Delaosa, J. Pestana, S. Weiss, I.K. Proudler, Subspace perturbation bounds with an application to angle of arrival estimation using the MUSIC algorithm, in: International Conference on Sensor Signal Processing for Defence, Edinburgh, UK, 2020.
- [77] I. Bronshtein, K. Semendyayew, *Handbook of Mathematics*, Springer, Heidelberg, 2015.
- [78] J. Thomas, L. Scharf, D. Tufts, The probability of a subspace swap in the SVD, *IEEE Trans. Signal Process.* 43 (3) (1995) 730–736.
- [79] M. Shaghghi, S.A. Vorobyov, Subspace leakage analysis and improved doa estimation with small sample size, *IEEE Trans. Signal Process.* 63 (12) (2015) 3251–3265.
- [80] P. Vallet, X. Mestre, P. Loubaton, Performance analysis of an improved music DoA estimator, *IEEE Trans. Signal Process.* 63 (23) (2015) 6407–6422.
- [81] B.A. Johnson, Y.I. Abramovich, X. Mestre, Music, g-music, and maximum-likelihood performance breakdown, *IEEE Trans. Signal Process.* 56 (8) (2008) 3944–3958.
- [82] S. Weiss, C. Delaosa, J. Matthews, I.K. Proudler, B. Jackson, Detection of weak transient signals using a broadband subspace approach, in: International Conference on Sensor Signal Processing for Defence, Edinburgh, Scotland, 2021, pp. 65–69.
- [83] V.W. Neo, S. Weiss, P.A. Naylor, A polynomial subspace projection approach for the detection of weak voice activity, in: Sensor Signal Processing for Defence Conference, London, UK, 2022, pp. 1–5.
- [84] C.A.D. Pahalsan, L.H. Crockett, S. Weiss, Detection of weak transient broadband signals using a polynomial subspace and likelihood ratio test approach, in: 32nd European Signal Processing Conference, Lyon, France, 2024, pp. 1312–1316.
- [85] F.A. Khattak, S. Weiss, I.K. Proudler, J.G. McWhirter, Space-time covariance matrix estimation: Loss of algebraic multiplicities of eigenvalues, in: 56th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 2022.