Eigenvalue Decomposition of a Parahermitian Matrix: Extraction of Analytic Eigenvectors

Stephan Weiss, *Senior Member, IEEE*, Ian K. Proudler, Fraser K. Coutts, *Member, IEEE*, and Faizan A. Khattak

***Abstract*—An analytic parahermitian matrix admits in almost all cases an eigenvalue decomposition (EVD) with analytic eigenvalues and eigenvectors. We have previously defined a discrete Fourier transform (DFT) domain algorithm which has been proven to extract the analytic eigenvalues. The selection of the eigenvalues as analytic functions guarantees in turn the existence of unique one-dimensional eigenspaces in which analytic eigenvectors can exist. Determining such eigenvectors is not straightforward, and requires three challenges to be addressed. Firstly, one-dimensional subspaces for eigenvectors have to be woven smoothly across DFT bins where a non-trivial algebraic multiplicity causes ambiguity. Secondly, with the one-dimensional eigenspaces defined, a phase smoothing across DFT bins aims to extract analytic eigenvectors with minimum time domain support. Thirdly, we need to check whether the DFT length, and thus the approximation order, is sufficient. We propose an iterative algorithm for the extraction of analytic eigenvectors and prove that this algorithm converges to the best of a set of stationary points. We provide a number of numerical examples and simulation results, in which the algorithm is demonstrated to extract the ground truth analytic eigenvectors arbitrarily closely.**

***Index Terms*—eigenvalue decomposition, parahermitian matrix factorisation, analytic functions, phase retrieval, quadratic programming with quadratic constraints.**

I. Introduction

If M sensors record a multi-channel broadband signal $\mathbf{x}[n] \in \mathbb{C}^M$ with time index $n \in \mathbb{Z}$, any signal processing needs to deal with time delays between elements of $\mathbf{x}[n]$, as opposed to the narrowband case where mere phase shifts suffice. Therefore, the space-time covariance matrix $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$, with $\mathcal{E}\{\cdot\}$ the expectation operator and $\{\cdot\}^{\mathrm{H}}$ the Hermitian transposition, has to include a lag parameter $\tau \in \mathbb{Z}$. The space-time covariance $\mathbf{R}[\tau]$ inherits the symmetries of its constituent auto- and cross-correlation sequence entries, and so $\mathbf{R}[\tau] = \mathbf{R}^{\mathrm{H}}[-\tau]$. Thus, its z-transform, the cross-spectral density (CSD) matrix $\boldsymbol{R}(z) = \sum_\tau \mathbf{R}[\tau] z^{-\tau}$, satisfies the parahermitian property $\boldsymbol{R}(z) = \boldsymbol{R}^{\mathrm{P}}(z)$ [1], whereby $\boldsymbol{R}^{\mathrm{P}}(z) \equiv \{\boldsymbol{R}(1/z^*)\}^{\mathrm{H}}$ requires a Hermitian transposition and time reversal. This CSD matrix can assist in formulating and solving broadband array problems, including e.g. beamforming [2]–[4], blind source separation [5], multichannel coding [6], [7], speech enhancement [8]–[10], or MIMO system design [11]–[13].

Optimum narrowband solutions at a fixed frequency Ω_0 are typically based on the diagonalisation of the narrowband covariance matrix $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_0}) = \boldsymbol{R}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega_0}}$. To formulate equivalent optimum broadband solutions requires the diagonalisation of $\mathbf{R}[\tau]$ for all lags τ, or equivalently the diagonalisation of the CSD matrix $\boldsymbol{R}(z)$ for every value of z. It is shown in [14]–[16] that unless $\mathbf{x}[n]$ is time-multiplexed, a parahermitian $\boldsymbol{R}(z)$ admits a parahermitian matrix EVD (PhEVD)

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z) \tag{1}$$

with an analytic paraunitary $\boldsymbol{Q}(z)$, such that $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}$, and an analytic diagonal $\boldsymbol{\Lambda}(z) = \mathrm{diag}\{\lambda_1(z), \ldots, \lambda_M(z)\}$. Both $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{Q}(z)$ are absolutely convergent but generally transcendental functions in z, i.e. (infinite) Laurent series. Due to this analyticity, Laurent polynomials can approximate these functions arbitrarily well by truncation.

A problem related to (1) is the McWhirter decomposition or polynomial EVD (PEVD) [17] $\boldsymbol{R}(z) \approx \boldsymbol{U}(z)\boldsymbol{D}(z)\boldsymbol{U}^{\mathrm{P}}(z)$. This is an approximate factorisation, using Laurent polynomial matrices $\boldsymbol{U}(z)$ and $\boldsymbol{D}(z)$. The diagonal parahermitian matrix $\boldsymbol{D}(z)$ contains the M approximate polynomial eigenvalues $d_\mu(z)$, $\mu = 1, \ldots, M$. These eigenvalues are spectrally majorised [18] such that, on the unit circle, $d_\mu(\mathrm{e}^{\mathrm{j}\Omega}) \geq d_{\mu+1}(\mathrm{e}^{\mathrm{j}\Omega})$ $\forall\Omega$ and $\mu = 1, \ldots, (M-1)$. As a result, in the case of overlapping eigenvalues $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$ in (1) as shown in Fig. 1(a), the functions $d_\mu(\mathrm{e}^{\mathrm{j}\Omega})$ are piecewise analytic, permuted versions of $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$ as depicted in Fig. 1(b), which are no longer differentiable and so are not analytic functions. Further, the corresponding eigenvectors will also be permuted, piece-wise analytic functions, but because of their orthonormality will include discontinuities at the points of permutation. The approximation of non-differentiable and discontinuous functions requires infinite Laurent series for $\boldsymbol{D}(z)$ and $\boldsymbol{U}(z)$, which are no longer absolutely convergent; their approximations therefore lead to much higher order functions than those for the analytic functions $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{Q}(z)$ in (1).

A large number of algorithms have targeted solutions to the McWhirter decomposition or PEVD over the last two decades. This includes the second order sequential best rotation (SBR2) family of algorithms [7], [17], [20] and the sequential matrix diagonalisation algorithms (SMD, [21], [22]). These algorithms generally encourage spectral majorisation, while SBR2

This work was supported in parts by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/S000631/1 and the MOD University Defence Research Collaboration in Signal Processing. F. Khattak is the recipient of a Commonwealth PhD Scholarship.

S. Weiss, I.K. Proudler, and F.A. Khattak are with the Centre for Signal & Image Processing (CeSIP), Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, Scotland (e-mail {ian.proudler,stephan.weiss,faizan.khattak}@strath.ac.uk).

F.K. Coutts is with the Institute for Digital Communications (IDCOM), University of Edinburgh, Scotland (e-mail fraser.coutts@ed.ac.uk).

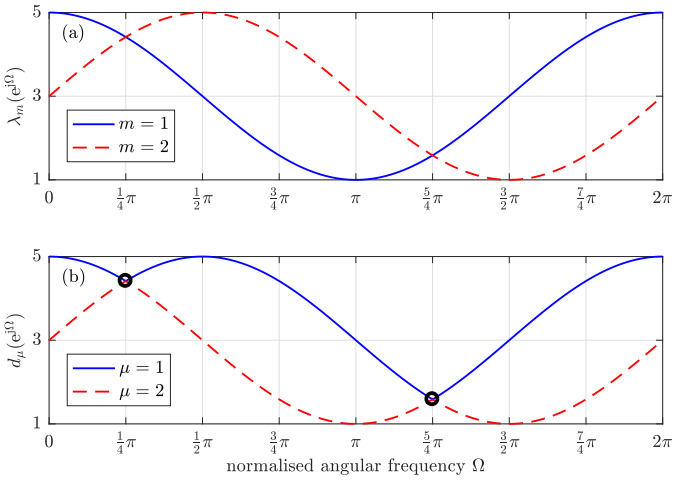

Fig. 1. Example of (a) analytic eigenvalues $\lambda_m(e^{j\Omega})$ [14], [19] and (b) spectrally majorised eigenvalues $d_\mu(e^{j\Omega})$, $m, \mu = 1, 2$, with non-differentiable points circled.

has been proven to converge to a spectral majorised solution [23]. Even though no convergence proof is provided, the fixed-order algorithms in [24]–[26] also target the spectrally majorised solution. Spectral majorisation can be useful, as it, e.g., provides optimality for subband coding [7], [18]. However, for subspace decompositions, targeting a spectrally majorised decomposition can lead to subspace perturbations at the permutation points that distinguish analytic from spectrally majorised solutions (e.g., see later Fig. 7); these perturbations inherently demand high approximation orders, resulting in an unnecessarily high implementation cost for applications in e.g. [2], [3], [5], [6], [11].

We are therefore interested in algorithms that can approximate the analytic solution in (1). Due to analyticity of both $\boldsymbol{R}(z)$ and its desired factors, it suffices to consider such functions on the unit circle, $\boldsymbol{R}(e^{j\Omega}) = \boldsymbol{R}(z)|_{z=e^{j\Omega}}$, where a solution in z may be obtained by re-parameterisation [27]. Related efforts on EVD, singular value (SVD) and QR decompositions have been undertaken for general matrices $\boldsymbol{A}(t)$ that depend on a real parameter $t \in \mathbb{R}$ on some interval [28]–[35]. For a detailed review, please refer to [36]; note however that an analytic decomposition of $\boldsymbol{A}(t)$ does not necessarily imply its existence for a matrix $\boldsymbol{A}(e^{j\Omega})$ that is periodic in $\Omega \in \mathbb{R}$ [15]. If it exists, then this periodicity can be exploited via efficient algorithms based on the discrete Fourier transform, unlike the solutions in [32]–[35].

DFT-based approaches somewhat negate the advantage of spectral coherence that is guaranteed by the time-domain PEVD algorithms in [7], [17], [20]–[22], [37]. Instead, similarly to [32]–[35] spectral coherence needs to be re-introduced across bins. The smooth polynomial EVD in [38] accomplishes this by monitoring the orthogonality of eigenvectors across bins. This can be misleading [36], [39], since eigenvectors are phase-ambiguous [14] and, at a C-fold algebraic multiplicity of eigenvalues, can form an arbitrary C-fold basis. Therefore, as a first part, in [36], we have based the extraction of analytic eigenvalues for (1) only on their analyticity, and have provided an algorithm with proven convergence, which iteratively grows the DFT order until a desired approximation error can be reached. In each iteration step, spectral coherence is re-established in principle by considering all possible associations of eigenvalues across frequency bins, which are evaluated via a smoothness metric based on differentiability [40]. In practice, the method in [36] follows a maximum likelihood sequence estimation approach that starts from the first DFT bins, and by adding further bins progressively only retains associations across these frequency bins that are sufficiently smooth. A metric that assesses truncation or time domain aliasing is then used in [36] to terminate the iteration and thus the algorithm at a desired threshold.

In this paper, we target the extraction of analytic eigenvectors based on the assumption that the analytic eigenvalues have been obtained using e.g. [36]. We characterise the ambiguities that exist for DFT-based binwise EVDs, and first resolve the problem of creating one-dimensional subspaces across DFT bins that contain non-trivial algebraic multiplicities; different from an initial investigation in [41], we here propose a closed-form solution. Thereafter, we perform a phase smoothing based on the analyticity of each of the eigenvectors separately. Beyond the discussion in [41], we show that this problem is generally NP hard, and provide some insight into a smoothness-based cost function. This analysis permits a guaranteed convergence to the smoothest of a set of possible functions, and hence the lowest-order approximation. We finally show that by iterating the subspace extraction and phase smoothing, convergence to an analytic solution can be guaranteed.

In the following, Sec. II defines important properties of the PhEVD in (1). The general DFT-based approach is outlined in Sec. III, followed by the extraction of 1-d eigenspaces and phase smoothing in Secs. IV and V. The growth of the approximation order and overall convergence is discussed in Sec. VI, with simulations in Sec. VII. Sec. VIII provides conclusions.

II. Parahermitian Matrix EVD

Via the space-time covariance matrix in Sec. II-A, we discuss relevant properties of a parahermitian matrix. Based on this, we review its PhEVD in Sec. II-B before we explore polynomial approximations of its factors in Sec. II-C.

A. Space-Time Covariance Matrix

A space-time covariance $\mathbf{R}[\tau]$ can be tied to the model in Fig. 2, where L_s mutually independent source signals $s_\ell[n]$, $\ell = 1, \ldots, L_\mathrm{s}$ contribute to the measurement vector $\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^\mathrm{T}$ via a convolutive mixing system $\mathbf{H}[n] \in \mathbb{C}^{M \times L_\mathrm{s}}$. The z-transform leads to a matrix $\boldsymbol{H}(z) = \sum_n \mathbf{H}[n] z^{-n} \colon \mathbb{C} \to \mathbb{C}^{M \times L_\mathrm{s}}$ of transfer functions. In the following we use the short hand notation $\mathbf{H}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{H}(z)$ for a transform pair [27]. In turn, the L_s source signals $s_\ell[n]$ in Fig. 2 can each be modelled as the output of an innovation filter $g_\ell[n] \circ\!\!-\!\!\bullet\, G_\ell(z)$ that is excited by a zero mean unit variance uncorrelated signal $u_\ell[n]$ [42]. With $\boldsymbol{G}(z) = \mathrm{diag}\{G_1(z) \ldots G_{L_\mathrm{s}}(z)\}$, the CSD matrix $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ\, \mathbf{R}[\tau]$,

$$\boldsymbol{R}(z) = \boldsymbol{H}(z)\boldsymbol{G}(z)\boldsymbol{G}^\mathrm{P}(z)\boldsymbol{H}^\mathrm{P}(z)\,, \tag{2}$$

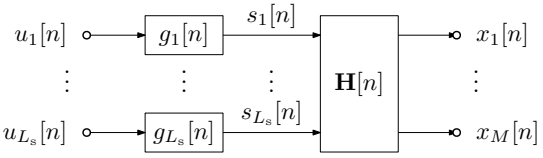

Fig. 2. Source model for measurement vector $\mathbf{x}[n]$ [36].

is thus determined by the source model components in Fig. 2.

If the systems $\boldsymbol{G}(z)$ and $\boldsymbol{H}(z)$ in (2) are causal and stable, and therefore analytic, then the CSD matrix $\boldsymbol{R}(z)$ will be analytic. Because of its structure, $\boldsymbol{R}(z) = \boldsymbol{A}(z)\boldsymbol{A}^{\mathrm{P}}(z)$, where $\boldsymbol{A}(z) = \boldsymbol{H}(z)\boldsymbol{G}(z)$, and so $\boldsymbol{R}(z)$ is guaranteed to be a parahermitian and positive semi-definite matrix. Further, when evaluated on the unit circle, $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{R}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega}}$ is a self-adjoint matrix such that $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{R}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})$ [43], i.e. for any particular normalised angular frequency $\Omega_0 \in \mathbb{R}$, $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_0})$ satisfies the Hermitian property.

B. Parahermitian Matrix EVD

If a parahermitian matrix $\boldsymbol{R}(z)$ is analytic and the measurement data $\mathbf{x}[n]$ does not originate from some multiplexing operation [15], then it is possible to state the following for its eigenvalues in (1):

Theorem 1 (Eigenvalues of a parahermitian matrix): Any analytic, parahermitian, and non-multiplexed $\boldsymbol{R}(z) : \mathbb{C} \to \mathbb{C}^{M\times M}$ possesses M unique analytic eigenvalues.

Proof. See [15] based on analysis in [14], which extends results for analytic self-adjoint matrices in [28] to the parahermitian case. ■

We assume that there are no identical eigenvalues, i.e. $\lambda_m(z) = \lambda_\mu(z)\ \forall z$, with $m, \mu = 1, \ldots, M$, is only true for $m = \mu$. This implies that for the model in Fig. 2, at the very least $L_s \geq (M-1)$.[1] In this case, the uniqueness theorem for analytic functions [44] states that there may only be a finite number of intersections of the functions $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$, $m = 1, \ldots, M$, i.e. frequency points where $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ possesses eigenvalues with an algebraic multiplicity greater than one. Under this premise, we can state the following theorem for the eigenvectors of $\boldsymbol{R}(z)$:

Theorem 2 (Eigenvectors of a parahermitian matrix): For the above analytic $\boldsymbol{R}(z)$ with no identical eigenvalues, there exists a unique one-dimensional eigenspace for every eigenvalue. The corresponding eigenvectors lying in these eigenspaces can be analytic, but are ambiguous up to an arbitrary allpass function.

Proof. See [15] and the analysis in [14] for the existence of analytic eigenvectors, which again follows as a generalisation of the self-adjoint matrix case in [28]. Because of its importance, in the following we explicitly consider the phase ambiguity. For this, we define a diagonal and paraunitary matrix $\boldsymbol{\Psi}(z)$,

$$\boldsymbol{\Psi}(z) = \mathrm{diag}\{\psi_1(z),\ \ldots,\ \psi_M(z)\}\ . \tag{3}$$

[1]We exclude identical eigenvalues, as the identification of 1-d eigenspaces as performed in Sec. IV otherwise is difficult and beyond the scope of this work.

Since $\boldsymbol{\Psi}(z)\boldsymbol{\Psi}^{\mathrm{P}}(z) = \mathbf{I}$ and diagonal matrices permute, we have $\boldsymbol{\Lambda}(z) = \boldsymbol{\Psi}(z)\boldsymbol{\Lambda}(z)\boldsymbol{\Psi}^{\mathrm{P}}(z)$ and so $\boldsymbol{R}(z)$ in (1) can be written as

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Psi}(z)\boldsymbol{\Lambda}(z)\boldsymbol{\Psi}^{\mathrm{P}}(z)\boldsymbol{Q}^{\mathrm{P}}(z)\ .$$

The product $\boldsymbol{Q}(z)\boldsymbol{\Psi}(z)$ now defines the phase ambiguity of the eigenvector $\boldsymbol{q}_m(z)$, i.e. the mth column of $\boldsymbol{Q}(z)$: if $\boldsymbol{q}_m(z)$ is an eigenvector corresponding to the eigenvalue $\lambda_m(z)$, then $\psi_m(z)\boldsymbol{q}_m(z)$ is also a valid eigenvector. If the remaining $(M-1)$ eigenvectors are selected to be analytic, then the mth eigenvector can only reside in a uniquely defined orthogonal subspace, refered to as the 1-d eigenspace. Many different functions lie with this eigenspace, differing by phase functions $\psi_m(z)$, with only some being analytic. ■

Example 1: Consider $\boldsymbol{R}(z) : \mathbb{C} \to \mathbb{C}^{2\times 2}$ from [14], [19],

$$\boldsymbol{R}(z) = \begin{bmatrix} \frac{1-\mathrm{j}}{2}z + 3 + \frac{1+\mathrm{j}}{2}z^{-1} & \frac{1+\mathrm{j}}{2}z^{2} + \frac{1-\mathrm{j}}{2} \\ \frac{1+\mathrm{j}}{2} + \frac{1-\mathrm{j}}{2}z^{-2} & \frac{1-\mathrm{j}}{2}z + 3 + \frac{1+\mathrm{j}}{2}z^{-1} \end{bmatrix}, \tag{4}$$

with eigenvalues $\boldsymbol{\Lambda}(z) = \mathrm{diag}\{z{+}3{+}z^{-1};\ -\mathrm{j}z{+}3{+}\mathrm{j}z^{-1}\}$. The eigenvectors are $\boldsymbol{q}_m(z) = \psi_m(z)[1, \pm z^{-1}]^{\mathrm{T}}/\sqrt{2}$, where

$$\psi_m(z) = \mathrm{e}^{\mathrm{j}\varphi}\prod_n \frac{a^*_{m,n} - z^{-1}}{1 - a_{m,n}z^{-1}} \tag{5}$$

is an arbitrary allpass filter with $|a_{m,n}| < 1$ and $a_{m,n} \in \mathbb{C}\ \forall m, n$ [45]. Note that unless the allpass $\psi_m(z)$ takes on a simple delay with $a_{m,n} = 0\ \forall m, n$, the support of $\mathbf{q}_m[n] \circ\!\!-\!\!\bullet\ \boldsymbol{q}_m(z)$ is infinite. Therefore, even though the choice of the allpass filter does not affect analyticity, it nonetheless influences the support and potential complexity if the eigenvectors are to be approximated by polynomials. △

C. Polynomial Approximation of PhEVD Factors

Even if $\boldsymbol{R}(z)$ is a Laurent polynomial, i.e. a function of finite order, the factors $\boldsymbol{Q}(z)$ and $\boldsymbol{\Lambda}(z)$ in (1) will in many cases be Laurent series, i.e. of infinite order, and represent algebraic or even transcendental functions [14], [15]. We therefore would like to approximate $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{Q}(z)$ by finite order functions, i.e. generally by Laurent polynomials. For the approximation of analytic eigenvalues, we can state the following theorem:

Theorem 3 (Laurent polynomial approximation of eigenvalues): The best Nth-order approximation $\hat{\boldsymbol{\Lambda}}^{(N)}(z)$ of an analytic $\boldsymbol{\Lambda}(z)$ in the least squares sense is obtained by truncating $\boldsymbol{\Lambda}(z)$ symmetrically to the required order.

Proof. See Theorem 2 in [36]. ■

The approximation of the paraunitary $\boldsymbol{Q}(z)$ should ideally lead to lossless filter bank implementations that process data to e.g. achieve strong decorrelation [18] or to generate subspace decompositions. We are therefore interested in approximate eigenvectors $\hat{\mathbf{q}}_m^{(N)}[n] \circ\!\!-\!\!\bullet\ \hat{\boldsymbol{q}}_m^{(N)}(z)$ that are causal and finite, i.e. for $\hat{\boldsymbol{q}}_m^{(N)}(z)$ to be a polynomial rather than a Laurent polynomial. With a rectangular window

$$p_N[n] = \begin{cases} 1\ , & 0 \leq n \leq N, \\ 0\ , & \text{otherwise}, \end{cases} \tag{6}$$

we can state for the polynomial approximation $\hat{\boldsymbol{q}}_m^{(N)}(z)$ of $\boldsymbol{q}_m(z)$:

Theorem 4 (Polynomial approximation of eigenvectors): The best Nth order polynomial approximation $\hat{\boldsymbol{q}}_m^{(N)}(z)$ of an analytic eigenvector $\boldsymbol{q}_m(z)$ in the least squares sense is obtained by masking

$$\hat{\mathbf{q}}_m^{(N)}[n] = p_N[n]\mathbf{q}_m[n-\Delta_m] \,, \tag{7}$$

whereby the shift Δ_m must minimise the overall truncation error in the least squares sense.

Proof. We want to initially find the Nth order Laurent polynomial approximation $\hat{\mathbf{w}}_m^{(N)}[n] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{w}}_m^{(N)}(\mathrm{e}^{\mathrm{j}\Omega})$ of a single eigenvector $\mathbf{q}_m[n] \circ\!\!-\!\!\bullet\, \boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})$. For the approximation error, we defined

$$\zeta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \|\boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega}) - \hat{\boldsymbol{w}}_m^{(N)}(\mathrm{e}^{\mathrm{j}\Omega})\|_2^2 \mathrm{d}\Omega \,. \tag{8}$$

Due to Parseval, this is equivalent to

$$\zeta = \sum_n \|\mathbf{q}_m[n] - \hat{\mathbf{w}}_m^{(N)}[n]\|_2^2 = \epsilon_- + \epsilon_0 + \epsilon_+ \tag{9}$$

with

$$\epsilon_- = \sum_{n=-\infty}^{-\Delta_m-1} \|\mathbf{q}_m[n]\|_2^2 \,, \qquad \epsilon_+ = \sum_{n=N-\Delta_m+1}^{\infty} \|\mathbf{q}_m[n]\|_2^2 \,,$$
$$\epsilon_0 = \sum_{n=-\Delta_m}^{N-\Delta_m} \|\mathbf{q}_m[n] - \hat{\mathbf{w}}_m^{(N)}[n]\|_2^2 \,.$$

All terms are non-negative; to minimise ζ, we require $\epsilon_0 = 0$, and set $\hat{\mathbf{w}}_m^{(N)}[n] = \mathbf{q}_m[n]$ for $-\Delta_m \leq n \leq N-\Delta_m$. There will be at least one Δ_m such that the overall truncation error $\epsilon_- + \epsilon_+$ is minimised. To ensure that the approximation is causal, we shift $\hat{\mathbf{w}}_m^{(N)}[n]$ by Δ_m samples, such that $\hat{\mathbf{q}}_m^{(N)}[n] = \hat{\mathbf{w}}_m^{(N)}[n-\Delta_m]$. ■

Example 2: Since $\boldsymbol{q}_m(z)$ is analytic, $\mathbf{q}_m[n]$ is absolutely convergent and decays at least as fast as an exponential function, say

$$\|\mathbf{q}_m[n]\|_2 \leq \begin{cases} a^{-n} & n < 0 \,, \\ b^n & n \geq 0 \,. \end{cases} \tag{10}$$

For the purpose of this example, we assume equality in (10), such that

$$\epsilon_- = \frac{a^{2\Delta_m+2}}{1-a^2} \,, \qquad \epsilon_+ = \frac{b^{2N-2\Delta_m+2}}{1-b^2} \,. \tag{11}$$

Therefore, for the minimum truncation error, we have that

$$\Delta_m = \frac{1}{2}\log_{ab}\frac{a^2(1-b^2)}{b^{2N+2}(1-a^2)} \,. \tag{12}$$

Specifically, for a symmetric decay with $a = b$ we obtain $\Delta_m = \frac{N}{2}$. △

The adjustment of the delay Δ_m is equivalent to the shift-compensated truncation method discussed in [46].

Note that the Fourier transform $\hat{\boldsymbol{q}}_m^{(N)}(\mathrm{e}^{\mathrm{j}\Omega}) \bullet\!\!-\!\!\circ\, \hat{\mathbf{q}}_m^{(N)}[\tau]$ converges uniformly to $\mathrm{e}^{-\mathrm{j}\Omega\Delta_m}\boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})$ as N increases [47]. Therefore by sufficiently adjusting the approximation order N, we can decrease the approximation error to an arbitrarily small value at every frequency $\Omega \in \mathbb{R}$. Further, while we used the variable N for both $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{Q}(z)$, generally different approximation orders are required for these two quantities if similar approximation errors are to be obtained.

Example 3: To demonstrate the potential mismatch in order, consider $\boldsymbol{R}(z) = [3,\ z;\ z^{-1},\ 3]$. This matrix possesses constant eigenvalues with $\boldsymbol{\Lambda}(z) = \mathrm{diag}\{4,\ 2\}$ of order zero, while $\boldsymbol{Q}(z) = [1,\ 1;\ z^{-1},\ -z^{-1}]/\sqrt{2}$ has order one. △

III. Bin-Wise EVD and General Approach

Since we operate in the DFT domain, we review the properties of an EVD at one sample point of $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ in Sec. III-A, and see how this relates to the PhEVD in Sec. III-B. This will expose quantities that are already known assuming previously determined analytic eigenvalues [36], and quantities that still need to be determined in order to extract analytic eigenvectors. A general algorithm of the latter is outlined in Sec. III-C.

A. EVD of a Hermitian Matrix

For a Hermitian matrix $\mathbf{R} \in \mathbb{C}^{M\times M}$, the eigenvalues λ_m, $m = 1, \ldots, M$ are uniquely determined up to an arbitrary ordering. The corresponding eigenvectors $\mathbf{q}_m$, such that $\mathbf{R}\mathbf{q}_m = \lambda_m\mathbf{q}_m$ are ambiguous. Assuming a C_i-fold algebraic multiplicity of eigenvalues, such that $\lambda_i = \lambda_{i+1} = \ldots = \lambda_{i+C_i-1}$, with corresponding eigenvectors $\mathbf{q}_i$, $\mathbf{q}_{i+1}$, …, $\mathbf{q}_{i+C_i-1}$, then for any arbitrary unitary $\mathbf{V}_i{}' \in \mathbb{C}^{C_i\times C_i}$,

$$\left[\mathbf{q}_i', \ldots, \mathbf{q}_{i+C_i-1}'\right] = \left[\mathbf{q}_i, \ldots, \mathbf{q}_{i+C_i-1}\right]\mathbf{V}_i' \tag{13}$$

are also valid eigenvectors of $\mathbf{R}$. This means that the eigenvectors can be selected as an arbitrary basis within a C_i-dimensional subspace. In case of a distinct eigenvalue with $C_i = 1$, this unitary matrix reduces to a 1×1 quantity $\mathbf{V}_i' = \mathrm{e}^{\mathrm{j}\psi_i}$ that still imposes an arbitrary phase shift $\psi_i \in \mathbb{R}$ [48].

For the factorisation $\mathbf{R} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}$, assume that the eigenvalues in $\boldsymbol{\Lambda}$ are majorised, i.e. $\lambda_1 \geq \ldots \geq \lambda_M$. A reordering in $\boldsymbol{\Lambda}'$ can be accomplished via a permutation matrix $\mathbf{P} \in \mathbb{R}^{M\times M}$ s.t. $\boldsymbol{\Lambda}' = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^{\mathrm{T}}$. If there are M' distinct eigenvalues and the ith one has an algebraic multiplicity C_i, $\boldsymbol{\Lambda}$ is block-diagonal,

$$\boldsymbol{\Lambda} = \mathrm{blockdiag}\{\boldsymbol{\Lambda}_1, \ldots, \boldsymbol{\Lambda}_{M'}\} \,,$$

where $\boldsymbol{\Lambda}_i = \lambda_i\mathbf{I}_{C_i}$ with $\mathbf{I}_{C_i}$ is $C_i \times C_i$, $i = 1, \ldots, M'$. Hence for any paraunitary matrix $\mathbf{V}_i'$ we have $\mathbf{V}_i'\boldsymbol{\Lambda}_i\mathbf{V}_i'^{\mathrm{H}} = \boldsymbol{\Lambda}_i$. Thus we can write $\boldsymbol{\Lambda}' = \mathbf{P}\mathbf{V}'\boldsymbol{\Lambda}\mathbf{V}'^{\mathrm{H}}\mathbf{P}^{\mathrm{T}} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^{\mathrm{T}}$ where the unitary matrix $\mathbf{V}' \in \mathbb{C}^{M\times M}$ is:

$$\mathbf{V}' = \mathrm{blockdiag}\{\mathbf{V}_1', \ldots, \mathbf{V}_{M'}'\} \,,$$

and $\mathbf{V}_i{}' \in \mathbb{C}^{C_i\times C_i}$ with $\sum_{i=1}^{M'} C_i = M$. Hence we have

$$\begin{aligned}\mathbf{R} &= \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}} = \mathbf{Q}\mathbf{V}'^{\mathrm{H}}\mathbf{P}^{\mathrm{T}}\mathbf{P}\mathbf{V}'\boldsymbol{\Lambda}\mathbf{V}'^{\mathrm{H}}\mathbf{P}^{\mathrm{T}}\mathbf{P}\mathbf{V}'\mathbf{Q}^{\mathrm{H}} \\ &= \left(\mathbf{Q}\mathbf{V}'^{\mathrm{H}}\mathbf{P}^{\mathrm{T}}\right)\boldsymbol{\Lambda}'\left(\mathbf{P}\mathbf{V}'\mathbf{Q}^{\mathrm{H}}\right) = \mathbf{Q}'\boldsymbol{\Lambda}'\mathbf{Q}'^{\mathrm{H}} \,,\end{aligned} \tag{14}$$

where $\boldsymbol{\Lambda}' = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^{\mathrm{T}}$ holds the permuted eigenvalues, and $\mathbf{Q}' = \mathbf{Q}\mathbf{V}'^{\mathrm{H}}\mathbf{P}^{\mathrm{T}}$ contains eigenvectors that are (i) reordered and (ii) form a new C_i-dimensional basis for the eigenspace corresponding to an eigenvalue with algebraic multiplicity C_i.

B. EVD at Sample Points of $\boldsymbol{R}(z)$

We evaluate $\boldsymbol{R}(z) : \mathbb{C} \longrightarrow \mathbb{C}^{M\times M}$ for K sample points along the unit circle at equispaced normalised angular frequencies $\Omega_k = \frac{2\pi}{K}k$, $k = 0, \ldots, (K-1)$. Evaluating the PhEVD in (1) at one such sample point Ω_k, we obtain

$$\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k})\boldsymbol{Q}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_k}) \;, \tag{15}$$

where $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \mathrm{diag}\{\lambda_1(\mathrm{e}^{\mathrm{j}\Omega_k}), \ldots, \lambda_M(\mathrm{e}^{\mathrm{j}\Omega_k})\}$ contains the appropriate sample points of the analytic eigenvalues $\lambda_m(z)$, $m = 1, \ldots, M$, and the mth column of $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})$ holds the sample points of the corresponding mth analytic eigenvector.

Note that if, as is conventional, the eigenvalues are ordered there is no reason to assume this ordering is the same in adjacent bins. Indeed, ordering the eigenvalues separately in each bin will produce spectral majorisation and, in general, dissociate the 'spectral coherence' of $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega})$ across bins. Therefore, in the EVD

$$\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \mathbf{Q}_k\boldsymbol{\Lambda}_k\mathbf{Q}_k^{\mathrm{H}} \;, \tag{16}$$

the r.h.s. factors are not directly equivalent to the terms on the r.h.s. of (15); in particular, the eigenvalues in $\boldsymbol{\Lambda}_k = \mathrm{diag}\{\lambda_{1,k}, \cdots, \lambda_{M,k}\}$ may be permuted w.r.t. to those in $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k})$.

The relation between the sample points of the analytic functions $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k})$ and $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})$ in (15), and the factors $\boldsymbol{\Lambda}_k$ and $\mathbf{Q}_k$ of the EVD in (16) is described by the ambiguity in (14), i.e.

$$\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \mathbf{P}_k\boldsymbol{\Lambda}_k\mathbf{P}_k^{\mathrm{T}} \tag{17}$$

$$\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \mathbf{Q}_k\mathbf{V}_k^{\mathrm{H}}\mathbf{P}_k^{\mathrm{T}} \;. \tag{18}$$

Akin to Sec. III-A, (17) and (18) contain permutations $\mathbf{P}_k$ and block-diagonal unitary matrices $\mathbf{V}_k$ that are specific to the kth bin. Specifically, $\mathbf{P}_k$ adjusts the ordering of eigenvalues in the kth bin, while

$$\mathbf{V}_k = \mathrm{blockdiag}\{\mathbf{V}_{k,1}, \ldots, \mathbf{V}_{k,M_k}\} \;, \tag{19}$$

where $M_k \leq M$ is the number of distinct eigenvalues in the k bin, and the dimension of the otherwise arbitrary unitary matrices $\mathbf{V}_{k,\mu} \in \mathbb{C}^{C_{k,\mu}\times C_{k,\mu}}$, $\mu = 1, \ldots, M_k$, is that of the algebraic multiplicity $C_{k,\mu}$ of the corresponding eigenvalues [14], [36]. Note that $\sum_{\mu=1}^{M_k} C_{k,\mu} = M$.

C. Analytic Eigenvector Extraction Approach

With the permutation matrices $\mathbf{P}_k$ and the unitary block-diagonal $\mathbf{V}_k$, $k = 0, \ldots, (K-1)$, as defined above, (18) reflects the challenges that the retrieval of analytic eigenvectors from the bin-wise EVDs of $\boldsymbol{R}(z)$ poses. We briefly discuss why we can assume $\mathbf{P}_k$ is known, and therefore only have to focus on $\mathbf{V}_k$ thereafter.

The permutation $\mathbf{P}_k$ needs to create a smooth association of both eigenspaces and eigenvalues across DFT bins. Although it is possible to estimate $\mathbf{P}_k$ based on both eigenvalues and eigenvectors [38], it is arguably better to only rely on the eigenvalues—see Sec. I. The extraction of analytic eigenvalues has been addressed in [36], and an algorithm with proven convergence to the analytic eigenvalues exists, which is capable of retrieving both the permutations $\mathbf{P}_k$, $k = 0, \cdots, (K-1)$, as well as the minimum approximation order $K_{\boldsymbol{\Lambda}}$ for a set approximation error for $\boldsymbol{\Lambda}(z)$. Therefore, for any DFT size $K \geq K_{\boldsymbol{\Lambda}}$, we assume $\mathbf{P}_k$ to be given via [36].

Algorithm 1: Extraction of Analytic Eigenvectors

1: determine analytic eigenvalues $\boldsymbol{\Lambda}(z)$ of $\boldsymbol{R}(z)$ [36];
2: initialise K to exceed the support of $\boldsymbol{R}(z)$ and $\boldsymbol{\Lambda}(z)$;
3: **repeat**
4: obtain $\mathbf{P}_k$, $k = 0, \ldots (K-1)$ from $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega_k})$;
5: calculate $\mathbf{Q}_k$ via EVD;
6: determine $\mathbf{A}_k$ for any $C_{k,\mu} > 1$;
7: determine $\boldsymbol{\Psi}_k$, $k = 0, \ldots, (K-1)$;
8: interpolate $\hat{\boldsymbol{Q}}(z)$ from $\hat{\mathbf{Q}}_k^{(K)} = \mathbf{Q}_k\boldsymbol{\Psi}_k^{\mathrm{H}}\mathbf{A}_k^{\mathrm{H}}\mathbf{P}_k^{\mathrm{T}}$;
9: calculate approximation error ζ
10: $K \leftarrow 2K$;
11: **until** $(\zeta < \epsilon) \vee (K > 2K_{\max})$.

To approximate the analytic eigenvectors in $\boldsymbol{Q}(z)$ from (18), it therefore remains to determine suitable unitary block-diagonal matrices $\mathbf{V}_k$ and an approximation order $K_{\mathbf{Q}}$ for $\boldsymbol{Q}(z)$ to guarantee a sufficiently small approximation error w.r.t. the analytic solution. As discussed below, the unitary matrix $\mathbf{V}_k$ can be determined as the product of two matrices,

$$\mathbf{V}_k = \mathbf{A}_k\boldsymbol{\Psi}_k \;, \tag{20}$$

subject to the DFT order being sufficiently large. The procedure then consists of three steps:

1) The alignment matrix $\mathbf{A}_k$ aligns 1-d eigenspaces smoothly across algebraic multiplicities; it contains along its diagonal unitary blocks of size $C_{k,\mu}$, where an eigenvalue $\lambda_{k,\mu}$ has an algebraic multiplicity $C_{k,\mu} > 1$, and diagonal values of one elsewhere.
2) The 1-d smooth eigenspaces can contain many different functions since only their magnitude is constrained. Generally such functions will not be analytic unless their phases are selected appropriately. Therefore, the diagonal matrix $\boldsymbol{\Psi}_k$,
$$\boldsymbol{\Psi}_k = \mathrm{diag}\{\mathrm{e}^{\mathrm{j}\psi_{k,1}}, \ldots, \mathrm{e}^{\mathrm{j}\psi_{k,M}}\} \;, \tag{21}$$
adjusts the phases in the kth bin to create analytic eigenvectors within each of the smooth 1-d eigenspaces. Note that $\boldsymbol{\Psi}_k$ closely relates to the arbitrary allpass filters that can modify the analytic eigenvectors in (3).
3) Unless $\mathbf{V}_k$ admits a sufficiently small approximation error, we return to 1) with an increased DFT length K.

These steps are outlined in Algorithm 1. The algorithm terminates when a sufficient accuracy or a maximum DFT length, $K_{\max}$, has been reached. We will later in Sec. VI-C see that a finite approximation error for the extracted analytic eigenvalues according to [36] does not impact on the correct association of the eigenvalues and eigenvectors, and hence does not affect steps 1 and 2. It may however impact on the achievable minimum approximation error ζ in step 3. We will further elaborate on the above steps 1), 2), and 3) in Secs. IV, V, and VI, respectively.

IV. Smooth One-Dimensional Eigenspaces

The purpose of this section is to construct the alignment matrices $\mathbf{A}_k$, $k = 0, \ldots, (K-1)$ for (20), which we address on a bin-by-bin basis. If in the kth bin all eigenvalues in (17) are distinct, then only a simple ordering of the eigenvectors via $\mathbf{P}_k$ is required, and $\mathbf{A}_k = \mathbf{I}_M$. Otherwise, if the kth bin contains any eigenvalues with an algebraic multiplicity greater than one, Sec. IV-A outlines the rationale for constructing $\mathbf{A}_k$ via a phase alignment across algebraic multiplicities discussed in Sec. IV-B, followed by the selection of smooth one-dimensional eigenspaces in Sec. IV-C.

A. Rationale

Recall that in case of a $C_{k,\mu}$-fold algebraic multiplicity of the μth eigenvalue in the kth bin, with $C_{k,\mu} > 1$, the corresponding eigenvectors can form an arbitrary basis within a $C_{k,\mu}$-dimensional subspace. However, to approximate analytic eigenvectors, we first need to weave smooth 1-d eigenspaces through such a manifold. Based on the assumption of non-identical eigenvalues and supported by the uniqueness theorem of analytic functions [44], we know that non-trivial algebraic multiplicities can only occur at a finite and isolated number of frequency points. We therefore compare the eigenvectors at such points against those in their direct vicinity, typically within a small fraction of the bin width.

Let Ω_k be a frequency bin where the eigenvalues $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega_k}) = \ldots = \lambda_{m+C_{k,m}-1}(\mathrm{e}^{\mathrm{j}\Omega_k})$ share a $C_{k,m}$-fold non-trivial algebraic multiplicity. Since we know the analytic eigenvalues to a predefined accuracy from [36], we can find frequency points $\Omega_k - \Delta$ and $\Omega_k + \Delta$ where $\Delta \ll 2\pi/K$ and the $C_{k,m}$ eigenvalues are sufficiently distinct. Let $\mathbf{Q}_{k-}$ and $\mathbf{Q}_{k+}$ be the matrices of eigenvectors obtained by EVDs from $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}(\Omega_k-\Delta)})$ and $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}(\Omega_k+\Delta)})$, and appropriately ordered using $\mathbf{P}_{k-}$ and $\mathbf{P}_{k+}$ based on the knowledge of the analytic eigenvalues [36]. Since the phases of $\mathbf{Q}_{k-}$ and $\mathbf{Q}_{k+}$ are incoherent — recall the difference between (15) and (16) — we first align phases across $\mathbf{Q}_{k-}$ and $\mathbf{Q}_{k+}$ in Sec. IV-B as a prerequisite for interpolating through the appropriate $C_{k,m}$-dimensional subspace of $\mathbf{Q}_k$ in Sec. IV-C.

B. Phase-Alignment Across Algebraic Multiplicities

Let $\mathbf{q}_{\mu,k-}$ and $\mathbf{q}_{\mu,k+}$ be the μth eigenvectors in the columns of $\mathbf{Q}_{k-}$ and $\mathbf{Q}_{k+}$. W.l.o.g., we retain $\mathbf{Q}_{k-}$ as it is, but change $\mathbf{q}_{\mu,k+}$ by a phase term $\vartheta_{m,k}$ that satisfies the optimisation problem

$$\vartheta_{\mu,k}^{\text{opt}} = \arg\min_{\vartheta_{\mu,k}} \|\mathbf{q}_{\mu,k-} - \mathrm{e}^{-\mathrm{j}\vartheta_{\mu,k}}\mathbf{q}_{\mu,k+}\|_2^2 \,, \quad \mu = m, \cdots, m + C_{k,m} - 1 \,. \tag{22}$$

Noting that $\|\mathbf{q}_{\mu,k-}\|_2 = \|\mathbf{q}_{\mu,k+}\|_2 = 1$, (22) simplifies to

$$\vartheta_{\mu,k}^{\text{opt}} = \arg\min_{\vartheta_{\mu,k}} \left\{ 1 - \mathrm{Re}\left\{ \mathrm{e}^{-\mathrm{j}\vartheta_{\mu,k}} \mathbf{q}_{\mu,k-}^{\mathrm{H}} \mathbf{q}_{\mu,k+} \right\} \right\} , \tag{23}$$

such that (22) is minimised by

$$\vartheta_{\mu,k}^{\text{opt}} = \angle\{\mathbf{q}_{\mu,k-}^{\mathrm{H}} \mathbf{q}_{\mu,k+}\} \,. \tag{24}$$

With the phase shifts $\vartheta_{\mu,k}^{\text{opt}}$, $\mu = m, \ldots, m + C_{k,m} - 1$, the eigenvectors $\mathbf{q}_{\mu,k-}$ and $\mathbf{q}_{\mu,k+}$ become aligned as closely as possible in the least squares sense.

C. Smooth Eigenspace Selection in Algebraic Multiplicities

Similar to the block-diagonal partitioning of $\mathbf{V}_k$ in (19), we now form the block diagonal $\mathbf{A}_k$ of (20), where every unitary subblock $\mathbf{A}_{k,\mu} \in \mathbb{C}^{C_{k,\mu} \times C_{k,\mu}}$ belongs to a $C_{k,\mu}$-fold algebraic multiplicity in the kth bin. For each such $C_{k,\mu}$ identical eigenvalues let $\mathbf{U}_{k,\mu} \in \mathbb{C}^{M \times C_{k,\mu}}$ contain the corresponding $C_{k,\mu}$ eigenvectors. Recalling (18) and (20), there is an ambiguity in the eigenvectors such that if $\mathbf{U}_{k,\mu}$ are valid eigenvectors then, for any arbitrary unitary matrix $\mathbf{A}_{k,\mu}$, so are $\mathbf{U}_{k,\mu}\mathbf{A}_{k,\mu}^{\mathrm{H}}$.

Since we assume that there are no eigenvalues that are identical for all frequencies, non-trivial algebraic multiplicities can only occur at isolated frequency points such as Ω_k [44]. This means that the algebraic multiplicities resolve immediately to the left and right of Ω_k, where we define $\mathbf{U}_{k-,\mu} \in \mathbb{C}^{M \times C_{k,\mu}}$ and $\mathbf{U}_{k+,\mu} \in \mathbb{C}^{M \times C_{k,\mu}}$ as the corresponding eigenvectors in $\mathbf{Q}_{k-,\mu}$ and $\mathbf{Q}_{k+,\mu}$, evaluated at $\Omega = \Omega_k \pm \Delta$. In order to smoothly continue eigenspaces, $\mathbf{A}_{k,\mu}$ must satisfy two conditions,

$$\mathbf{U}_{k,\mu}\mathbf{A}_{k,\mu}^{\mathrm{H}} = \lim_{\Delta \to 0} \mathbf{U}_{k-,\mu} \,, \tag{25}$$

$$\mathbf{U}_{k,\mu}\mathbf{A}_{k,\mu}^{\mathrm{H}} = \lim_{\Delta \to 0} \mathbf{U}_{k+,\mu}\boldsymbol{\Theta}_{k,\mu} \,, \tag{26}$$

such that the columns of $\mathbf{U}_{k,\mu}\mathbf{A}_{k,\mu}^{\mathrm{H}}$ align with the corresponding 1-d eigenspaces to the immediate left and right of the kth bin. The matrix $\boldsymbol{\Theta}_{k,\mu} = \mathrm{diag}\{\mathrm{e}^{\mathrm{j}\vartheta_{\mu,k}}, \ldots, \mathrm{e}^{\mathrm{j}\vartheta_{\mu+C_{k,\mu}-1,k}}\}$ performs the phase alignment described in Sec. IV-B between $\mathbf{U}_{k-,\mu}$ and $\mathbf{U}_{k+,\mu}$.

Replacing the limits $\Delta \to 0$ in (25) and (26) by a small deviation $\Delta \ll 2\pi/K$, we formulate the following constrained problem for $\mathbf{A}_{k,\mu}$:

$$\min_{\mathbf{A}_{k,\mu}} \chi_{k,\mu} \qquad \text{s.t.} \qquad \mathbf{A}_{k,\mu}^{\mathrm{H}}\mathbf{A}_{k,\mu} = \mathbf{I} \,. \tag{27}$$

with

$$\chi_{k,\mu} = \|\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k-,\mu} - \mathbf{I}\|_{\mathrm{F}}^2 + \|\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k+,\mu}\boldsymbol{\Theta}_{k\mu} - \mathbf{I}\|_{\mathrm{F}}^2 \,,$$

and $\|\cdot\|_{\mathrm{F}}$ the Frobenius norm. Expanding this cost term using the trace operator $\mathrm{tr}\{\cdot\}$

$$\begin{aligned}\chi_{k,\mu} = \mathrm{tr}\big\{&(\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k-,\mu} - \mathbf{I})^{\mathrm{H}}(\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k-,\mu} - \mathbf{I})\big\} \\ &+ \mathrm{tr}\left\{(\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k+,\mu}\boldsymbol{\Theta}_{k,\mu} - \mathbf{I})^{\mathrm{H}} \cdot \right. \\ &\left. \cdot(\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\mathbf{U}_{k+,\mu}\boldsymbol{\Theta}_{k,\mu} - \mathbf{I})\right\}\end{aligned} \tag{28}$$

reveals a quadratic expression in $\mathbf{A}_{k,\mu}$. However, the unitary constraint generally causes this problem to be non-convex [49]. Nevertheless, a gradient approach using matrix-valued differentiation [50], Wirtinger calculus [51], and a suitable initialisation leads to a workable iterative solution [41].

Interestingly, a closed-form solution to the optimisation problem in (27) exists[2]. Rewriting $\chi_{k,\mu}$ as

$$\chi_{k,\mu} = \|\mathbf{A}_{k,\mu}\mathbf{U}_{k,\mu}^{\mathrm{H}}\left[\mathbf{U}_{k-,\mu} \,,\ \ \mathbf{U}_{k+,\mu}\boldsymbol{\Theta}_{k,\mu}\right] - \left[\mathbf{I} \,,\ \ \mathbf{I}\right]\|_{\mathrm{F}}^2$$

[2]The closed-form solution via a Procustes problem was kindly suggested by one of the reviewers.

reveals an orthogonal Procrustes problem [48], which is equivalent to

$$\chi_{k,\mu} = \|\mathbf{A}_{k,\mu} - (\mathbf{U}_{k-,\mu} + \mathbf{U}_{k+,\mu}\mathbf{\Theta}_{k,\mu})^{\mathrm{H}}\,\mathbf{U}_{k,\mu}\|_{\mathrm{F}}^2\,, \quad (29)$$

i.e. we want to find a unitary matrix $\mathbf{A}_{k,\mu}$ that matches $(\mathbf{U}_{k-,\mu} + \mathbf{U}_{k+,\mu}\mathbf{\Theta}_{k,\mu})^{\mathrm{H}}\mathbf{U}_{k,\mu}$ in the least squares sense.

Let the operation $\mathbf{Y} = \mathbf{\Pi}\{\mathbf{X}\}$ find the unitary matrix $\mathbf{Y}$ closest to $\mathbf{X} \in \mathbb{C}^{C_{k,\mu}\times C_{k,\mu}}$ in the least squares sense; based on the SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}$, we have $\mathbf{\Pi}\{\mathbf{X}\} = \mathbf{U}\mathbf{V}^{\mathrm{H}}$ [48]. Therefore,

$$\mathbf{A}_{k,\mu} = \mathbf{\Pi}\{(\mathbf{U}_{k-,\mu} + \mathbf{U}_{k+,\mu}\mathbf{\Theta}_{k,\mu})^{\mathrm{H}}\mathbf{U}_{k,\mu}\} \quad (30)$$

is the solution to (27).

Thus, to solve the subspace alignment problem across algebraic multiplicities we first obtain $\mathbf{\Theta}_{k,\mu}$ as defined in (26), and thereafter evaluate (30). Once all algebraic multiplicities within a bin have been addressed, we can assemble

$$\mathbf{A}_k = \mathrm{blockdiag}\{\mathbf{A}_{k,1}, \mathbf{A}_{k,2}, \ldots, \mathbf{A}_{k,M_k}\}$$

as the alignment matrix for the kth bin. In practice, we define an algebraic multiplicity of order C if within a bin we encounter C eigenvalues that are within a small limit ϵ_{AM} of each other. We then find a small Δ which can be iteratively increased if necessary, such that we find new bins $\Omega_k \pm \Delta$ where those C eigenvalues separate by more than several ϵ_{AM}.

Example 4: For a parahermitian matrix with known analytic eigenvalues and eigenvectors, Fig. 3(a) shows the analytic eigenvalues that can be extracted, e.g., by the method in [36]. The black circles in Fig. 3(a) indicate which of the $K = 2^{10}$ bins contain non-trivial algebraic multiplicities. It can be seen that there are a further four non-trivial algebraic multiplicities in the continuous eigenvalues but they do not coincide with the $K = 2^{10}$ frequency bins. In terms of the algorithm under discussion, it is the multiplicities at $\Omega = \{0, \pi, \frac{3}{2}\pi\}$ that are of interest, where the eigenvectors are not clearly defined. To measure the smoothness of extracted 1-d eigenspaces, the Hermitian angle of eigenvectors evolving with Ω are inspected. We calculate the Hermitian angle w.r.t. a reference point, here $\mathbf{r} = \mathbf{q}_1(\mathrm{e}^{\mathrm{j}\Omega})|_{\Omega=0}$, such that

$$\cos\alpha_m(\mathrm{e}^{\mathrm{j}\Omega_k}) = |\mathbf{q}_m^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_k})\cdot\mathbf{r}|\,. \quad (31)$$

The subspace angle $\alpha_m(\mathrm{e}^{\mathrm{j}\Omega_k})$ in (31) for $\mathbf{q}_m^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_k})$ before and after performing the above smoothing of 1-d eigenspaces are shown in Fig. 3(b) and (c), respectively. Note that the eigenvectors corresponding to the eigenvalues with non-trivial algebraic multiplicities result in discontinuities in the subspace angle, while the application of the above method correctly aligns the 1-d eigenspaces across algebraic multiplicities, providing a smooth evolution of $\alpha_m(\mathrm{e}^{\mathrm{j}\Omega_k})$. △

Note that the extraction of smooth 1-d eigenspaces is indicated by a smooth evolution of the subspace angle in, e.g., Fig. 3(c). Since (31) takes the absolute value on its r.h.s., the subspace angle is blind to phase shifts. Thus with respect to (20), we have identified $\mathbf{A}_k$, but now need to determine $\mathbf{\Psi}_k$ in each bin in order to determine an analytic eigenvector within each smooth 1-d eigenspace. This is the purpose of the following section.

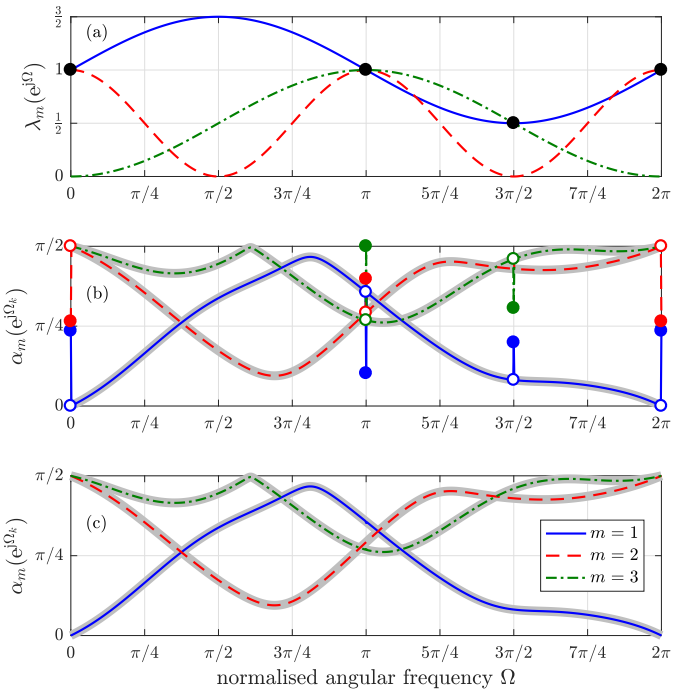

Fig. 3. For an example of a parahermitian matrix $\boldsymbol{R}(z) : \mathbb{C} \to \mathbb{C}^{3\times 3}$, (a) analytic eigenvalues evaluated in $K = 2^{10}$ DFT bins resulting in 3 algebraic multiplicities indicated by black markers, and the evolution of subspace angle in (31) for the corresponding eigenvectors (b) before and (c) after creation of smooth 1-d eigenspaces, with the ground truth under-laid in grey.

V. Extraction of Analytic Eigenvectors from Smooth Eigenspaces

A. Overview of Approach

To determine the sample points $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})$ of the desired analytic eigenvectors based on the evaluated $\mathbf{Q}_k$ in (18) requires $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k}) = \mathbf{Q}_k\mathbf{P}_k\mathbf{A}_k\,\mathbf{\Psi}_k$ via (20). With the permutation matrix $\mathbf{P}_k$ given by the analytic eigenvalue extraction [36] and the alignment matrix $\mathbf{A}_k$ calculated in Sec. IV, each column of $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})$ matches the corresponding column of $\mathbf{Q}_k\mathbf{P}_k\mathbf{A}_k$ save for a phase shift contained in $\mathbf{\Psi}_k$. Specifically, the kth sample point of the mth analytic eigenvector $\mathbf{q}_{k,m}$, i.e. $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega_k})$, and the corresponding column $\mathbf{u}_{k,m}$ of $\mathbf{Q}_k\mathbf{P}_k\mathbf{A}_k$ are related as $\mathbf{q}_{k,m} = \mathrm{e}^{\mathrm{j}\psi_{k,m}}\mathbf{u}_{k,m}$, whereby $\mathbf{\Psi}_k = \mathrm{diag}\{\mathrm{e}^{\mathrm{j}\psi_{k,1}}, \ldots, \mathrm{e}^{\mathrm{j}\psi_{k,M}}\}$. Finding $\mathbf{\Psi}_k$ therefore consists of M decoupled problems. In this section, we representatively focus on a single such problem, and by dropping subscripts for the mth eigenvector, want to determine sample points $\mathbf{q}_k$ of a vector of analytic functions by phase-aligning given vectors $\mathbf{u}_k$, $k = 0, \ldots, (K-1)$, across all K bins. Thus, within each of the one-dimensional eigenspaces determined in Sec. IV, we now aim to establish an analytic eigenvector.

The aim is to select the phase values ψ_k, $k = 0, \ldots, (K-1)$ such that the sample points $\mathbf{q}_k$ admit a maximally smooth interpolation $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) \bullet\!\!-\!\!\circ\, \hat{\mathbf{q}}^{(K)}[n]$ as per (7), with

$$\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) = \left[\hat{Q}_1^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}), \ldots, \hat{Q}_M^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})\right]^{\mathrm{T}}\,. \quad (32)$$

If $q_{\mu,k}$ and $u_{\mu,k}$ are the μth elements of the vectors $\mathbf{q}_k \in \mathbb{C}^M$ and $\mathbf{u}_k \in \mathbb{C}^M$, respectively, then each function $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ is

obtained by interpolation from the set of points contained in the vectors $\mathbf{q}'_\mu \in \mathbb{C}^K$ and $\mathbf{u}'_\mu \in \mathbb{C}^K$,

$$\mathbf{q}'_\mu = [q_{\mu,0},\ \ldots,\ q_{\mu,K-1}]^{\mathrm{T}} \tag{33}$$

$$\mathbf{u}'_\mu = [u_{\mu,0},\ \ldots,\ u_{\mu,K-1}]^{\mathrm{T}}\ . \tag{34}$$

These two vectors of samples are related via

$$\mathbf{q}'_\mu = \mathrm{diag}\{\mathbf{a}\}\,\mathbf{u}'_\mu = \mathrm{diag}\{\mathbf{u}'_\mu\}\,\mathbf{a} \tag{35}$$

based on a vector of phase shifts

$$\mathbf{a} = \left[\mathrm{e}^{\mathrm{j}\psi_0},\ \ldots,\ \mathrm{e}^{\mathrm{j}\psi_{K-1}}\right]^{\mathrm{T}}\ . \tag{36}$$

Below, we elaborate on maximally smooth interpolations $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ in Sec. V-B, define the power of derivatives of the interpolant as a measure of differentiability, and utilise it as a cost function for optimising the phase shifts in $\mathbf{a}$ in Sec. V-C.

B. Maximally Smooth Interpolation

The smoothest possible interpolation for $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ is given by the Dirichlet interpolation, in the sense that it provides the shortest possible support in the time domain [36]. In the Fourier domain, the Dirichlet function $\mathcal{P}_K(\mathrm{e}^{\mathrm{j}\Omega}) = \sum_{n=0}^{K-1} \mathrm{e}^{-\mathrm{j}\Omega n} \bullet\!\!-\!\!\circ\, p_K[n]$ given by (6) leads to

$$\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) = \frac{1}{K}\sum_{k=0}^{K-1} q_{\mu,k}\mathcal{P}_K\left(\mathrm{e}^{\mathrm{j}(\Omega-2\pi k/K)}\right) \tag{37}$$

for the smoothest possible interpolant based on the K sample points contained in $\underline{q}'_\mu$. Note that

$$\begin{aligned}\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) &= \frac{1}{K}\sum_{k=0}^{K-1} q_{\mu,k}\sum_{n=0}^{K-1}\mathrm{e}^{-\mathrm{j}(\Omega-2\pi k/K)n}\\ &= \frac{1}{K}\sum_{n=0}^{K-1}\mathrm{e}^{-\mathrm{j}\Omega n}\sum_{k=0}^{K-1} q_{\mu,k}\mathrm{e}^{\mathrm{j}2\pi nk/K}\\ &= \frac{1}{\sqrt{K}}\mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\cdot\mathbf{W}_K^{\mathrm{H}}\cdot\mathbf{q}'_\mu\ ,\end{aligned}$$

where

$$\mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) = [1,\ \mathrm{e}^{-\mathrm{j}\Omega},\ \ldots,\ \mathrm{e}^{-\mathrm{j}\Omega(K-1)}]\ ,$$

and $\mathbf{W}_K$ is a K-point DFT matrix scaled by $1/\sqrt{K}$ to be unitary.

C. Measuring Smoothness of an Eigenvector

Since we are looking for an analytic solution, the interpolation $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$, $\mu = 1,\ldots,M$, must be infinitely differentiable. Similar to finding analytic eigenvalues in [36], therefore the power in a pth derivative of $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ provides a suitable metric, $\xi_p^{(K)}$,

$$\xi_p^{(K)} = \sum_{\mu=1}^{M}\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\partial^p\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})}{\partial\Omega^p}\right|^2\mathrm{d}\Omega\ . \tag{38}$$

For the pth derivative of $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$, we have that

$$\frac{\partial^p\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})}{\partial\Omega^p} = \frac{1}{\sqrt{K}}\mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\mathbf{D}_K^p\cdot\mathbf{W}_K^{\mathrm{H}}\cdot\mathbf{q}'_\mu\ ,$$

where

$$\mathbf{D}_K = \mathrm{diag}\{0,\ -\mathrm{j},\ ,-2\mathrm{j},\ \ldots,\ -(K-1)\mathrm{j}\}\ . \tag{39}$$

Due to Parseval's theorem [27], [39], [52], for any $\mathbf{x}\in\mathbb{C}^K$,

$$\frac{1}{2\pi}\int_0^{2\pi}\left|\mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\mathbf{x}\right|^2\mathrm{d}\Omega = \mathbf{x}^{\mathrm{H}}\mathbf{x}\ .$$

Therefore, for (38), we obtain

$$\xi_p^{(K)} = \frac{1}{K}\sum_{\mu=1}^{M}\left\|\mathbf{D}_K^p\cdot\mathbf{W}_K^H\cdot\mathbf{q}'_\mu\right\|_2^2\ .$$

Using (35), this cost can be tied to the phase vector $\mathbf{a}$ in (36),

$$\xi_p^{(K)} = \mathbf{a}^{\mathrm{H}}\mathbf{C}_{K,p}\mathbf{a}\ , \tag{40}$$

where

$$\mathbf{C}_{K,p} = \sum_{\mu=1}^{M}\mathrm{diag}\{\mathbf{u}'^{\mathrm{H}}_\mu\}\,\mathbf{W}_K\mathbf{D}_K^{2p}\mathbf{W}_K^{\mathrm{H}}\mathrm{diag}\{\mathbf{u}'_\mu\}\ .$$

Thus the pth derivative, and therefore the smoothness, of the vector of the smoothest possible interpolants in (32) can be expressed in terms of the vector $\mathbf{a}$ of phase shifts that we would like to optimise.

D. Optimisation Problem

Based on the above analysis, we need to solve the following optimisation problem

$$\begin{aligned}&\min_{\mathbf{a}}\ \mathbf{a}^{\mathrm{H}}\mathbf{C}_{K,p}\mathbf{a}\\ &\text{s.t.}\quad |a_k| = 1,\ \forall k = 0,\ldots,(K-1)\ ,\end{aligned} \tag{41}$$

where a_k is the kth element of $\mathbf{a}$. Even though the cost term is quadratic in $\mathbf{a}$, the condition is awkward.

To understand the type of optimisation problem, note that, using (35), (40) can be reformulated as $\xi_p^{(K)} = \mathbf{x}^{\mathrm{H}}\mathbf{C}\mathbf{x}$, where $\mathbf{x}$ is the concatenation of the M vectors $\mathbf{q}'_\mu$, $\mu = 1,\ldots,M$, and $\mathbf{C} = \mathbf{C}_{K,p}\otimes\mathbf{I}_M$, with $\otimes$ denoting the Kronecker product. We need to include a unit norm constraint on the eigenvector $\mathbf{q}_k$ via some matrix $\mathbf{F}_k$ as $\mathbf{x}^{\mathrm{H}}\mathbf{F}_k\mathbf{x} = 1$. Additionally, $\mathbf{q}_k$ must be orthogonal to the remaining $(M-1)$ eigenvectors in the kth bin. This can be achieved by appropriately embedding the latter in a matrix $\mathbf{G}$ such that $\mathbf{G}\mathbf{x} = \underline{0}$. The resulting formulation

$$\begin{aligned}&\min_{\mathbf{x}}\ \mathbf{x}^{\mathrm{H}}\mathbf{C}\mathbf{x}\\ &\quad\text{s.t.}\quad \mathbf{x}^{\mathrm{H}}\mathbf{F}_k\mathbf{x} - 1 = 0, \qquad k = 0,\ldots,(K-1)\\ &\qquad\quad \mathbf{x}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\mathbf{G}\mathbf{x} = 0\end{aligned}$$

represents a quadratically constrained quadratic programming (QCQP) problem [53], [54], which is non-convex and even for semi-definite matrices $\mathbf{C}$, $\mathbf{F}_k$, and $\mathbf{G}^{\mathrm{H}}\mathbf{G}$ NP hard.

QCQP problems can be solved via a linearisation in an inflated parameter space — typically squaring the number of parameters — subject to a rank one constraint. A popular solution approach is semidefinite relaxation [54], which suppresses the rank one constraint, thus resulting in a convex optimisation problem for which various solvers exist [55]. Since here the number of frequency bins could be rather large,

operating with K^2 optimisation parameters is not feasible. We therefore explore a different approach below, which exploits some particular properties of our cost function.

E. Cost Function Properties

To incorporate the constraint into the cost function of (41), we define a vector of phases

$$\boldsymbol{\psi}^{\mathrm{T}} = [\psi_0,\ \psi_1,\ \ldots,\ \psi_{K-1}] \,.$$

Using (36), this allows to reformulate (41) as

$$\min_{\boldsymbol{\psi}} \mathbf{a}^{\mathrm{H}} \mathbf{C}_{K,p} \mathbf{a} \,. \tag{42}$$

The gradient of (42) can be determined via the operator

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\psi}} &= \frac{\partial \mathbf{a}^{\mathrm{H}}}{\partial \boldsymbol{\psi}} \frac{\partial}{\partial \mathbf{a}^*} + \frac{\partial \mathbf{a}^{\mathrm{T}}}{\partial \boldsymbol{\psi}} \frac{\partial}{\partial \mathbf{a}} \\ &= -\mathrm{j}\,\mathrm{diag}\{\mathbf{a}^*\} \frac{\partial}{\partial \mathbf{a}^*} + \mathrm{j}\,\mathrm{diag}\{\mathbf{a}\} \frac{\partial}{\partial \mathbf{a}} \,, \end{aligned} \tag{43}$$

where $\mathrm{diag}\{\mathbf{a}\}$ is a diagonal matrix containing the elements of the vector $\mathbf{a}$ along its diagonal. Then, noting that $\mathbf{C}_{K,p} = \mathbf{C}_{K,p}^{\mathrm{H}}$,

$$\begin{aligned} \frac{\partial \xi_p^{(K)}}{\partial \boldsymbol{\psi}} &= -\mathrm{j}\,\mathrm{diag}\{\mathbf{a}^*\}\, \mathbf{C}_{K,p} \mathbf{a} \ + \ \mathrm{j}\,\mathrm{diag}\{\mathbf{a}\}\, \mathbf{C}_{K,p}^{\mathrm{H}} \mathbf{a}^* \\ &= 2\mathrm{Im}\{\mathrm{diag}\{\mathbf{a}^*\}\, \mathbf{C}_{K,p} \mathbf{a}\} \ . \end{aligned} \tag{44}$$

With the gradient defined, we can make the following statement about some of the stationary points of this cost function:

Theorem 5 (Stationary points): If the time-domain support of an eigenvector is limited, then some stationary points of the cost function to be minimised in (42) are closely related by a small shift $\kappa \in \mathbb{Z}$ applied to the time domain eigenvector.

Proof. We assume that for a phase vector $\boldsymbol{\psi}_{\min}$ with corresponding $\mathbf{a}_{\min}$, according to (44), we have

$$\frac{\partial \xi_p^{(K)}(\boldsymbol{\psi}_{\min})}{\partial \boldsymbol{\psi}} = 2\mathrm{Im}\{\mathrm{diag}\{\mathbf{a}_{\min}^*\}\, \mathbf{C}_{K,p} \mathbf{a}_{\min}\} = \underline{0} \,.$$

A shift by $\kappa \in \mathbb{Z}$ samples in the time domain is equivalent to a phase shift applied to $\mathbf{a}_{\min}$, such that

$$\begin{aligned} \mathbf{a}_\kappa &= \mathrm{diag}\Big\{ [1,\ \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}\kappa},\ \ldots,\ \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}(K-1)\kappa}] \Big\}\, \mathbf{a}_{\min} \\ &= \mathrm{diag}\{\mathbf{w}_\kappa\}\, \mathbf{a}_{\min} \,. \end{aligned} \tag{45}$$

For the gradient, we obtain

$$\begin{aligned} \frac{\partial \xi_p^{(K)}(\boldsymbol{\psi}_\kappa)}{\partial \boldsymbol{\psi}} &= 2\mathrm{Im}\{\mathrm{diag}\{\mathbf{a}_\kappa^*\}\, \mathbf{C}_{K,p} \mathbf{a}_\kappa\} \\ &= 2\mathrm{Im}\Big\{\mathrm{diag}\{\mathbf{a}_{\min}^*\}\, \mathbf{C}_{K,p}^{(\kappa)} \mathbf{a}_{\min}\Big\} \ , \end{aligned}$$

where the modified term $\mathbf{C}_{K,p}^{(\kappa)}$ is given by

$$\begin{aligned} \mathbf{C}_{K,p}^{(\kappa)} &= \mathrm{diag}\{\mathbf{w}_\kappa^*\} \sum_\mu \mathrm{diag}\big\{\mathbf{u}_\mu'^*\big\}\, \mathbf{W}_K \mathbf{D}_K^{2p} \mathbf{W}_K^{\mathrm{H}} \\ &\qquad \cdot \mathrm{diag}\big\{\mathbf{u}_\mu'\big\}\, \mathrm{diag}\{\mathbf{w}_\kappa\} \\ &= \sum_\mu \mathrm{diag}\big\{\mathbf{u}_\mu'^*\big\}\, \mathrm{diag}\{\mathbf{w}_\kappa^*\}\, \mathbf{W}_K \mathbf{D}_K^{2p} \mathbf{W}_K^{\mathrm{H}} \\ &\qquad \cdot \mathrm{diag}\{\mathbf{w}_\kappa\}\, \mathrm{diag}\big\{\mathbf{u}_\mu'\big\} \ . \end{aligned} \tag{46}$$

Note that

$$\mathrm{diag}\{\mathbf{w}_\kappa^*\}\, \mathbf{W}_K = \mathbf{W}_K \begin{bmatrix} \mathbf{0} & \mathbf{I}_{K-\kappa} \\ \mathbf{I}_\kappa & \mathbf{0} \end{bmatrix} \,.$$

Therefore, (46) simplifies to

$$\mathbf{C}_{K,p}^{(\kappa)} = \sum_\mu \mathrm{diag}\big\{\mathbf{u}_\mu'^*\big\}\, \mathbf{W}_K \tilde{\mathbf{D}}_K^{2p} \mathbf{W}_K^{\mathrm{H}} \mathrm{diag}\big\{\mathbf{u}_\mu'\big\}$$

with

$$\tilde{\mathbf{D}}_K^{2p} = \mathrm{diag}\{\kappa,\ \ldots,\ (K-1),\ 0,\ \ldots,\ (\kappa-1)\}^{2p} \ . \tag{47}$$

Note that $\mathbf{W}_K^{\mathrm{H}} \mathrm{diag}\big\{\mathbf{u}_\mu'\big\}$ contains the time domain coefficients equivalent to the sample points in $\mathbf{u}_\mu'$ along its diagonal. If the time domain support is limited to the first L coefficients with $L < K - \kappa$, then comparing (39) with (47) and for a small value κ we have

$$\lim_{k\to\infty} \frac{k^{2p}}{(k+\kappa)^{2p}} = 1 \,. \tag{48}$$

Thus, the weighting applied by $\tilde{\mathbf{D}}_K^{2p}$ is approximately proportionate to that applied by $\mathbf{D}_K^{2p}$, and therefore for the gradient we obtain

$$\frac{\partial \xi_p^{(K)}(\boldsymbol{\psi}_\kappa)}{\partial \boldsymbol{\psi}} \approx 2\mathrm{Im}\{\mathrm{diag}\{\mathbf{a}_{\min}^*\}\, \mathbf{C}_{K,p} \mathbf{a}_{\min}\} = \underline{0} \,.$$

The deviation from proportionality for small k in (48) means that a phase shift by a small $\kappa \in \mathbb{Z}$ only leads approximately to a new stationary point. ■

Example 5: In Example 4, the ground truth analytic $\boldsymbol{Q}(z)$ has a length of $L = 6$. We here investigate the cost function for $K = 16$ bins of its first eigenvector based on the power in its $p = 2, 3, 4, 5$th derivative. If the $M = 3$ components $\mathbf{u}_\mu'$, $\mu = 1, 2, 3$ in (34) match the analytic solution, then ideally $\psi_k = 0$, $k = 0, \ldots, (K-1)$. We now select $\psi_k = 2\pi\kappa k/K$, $k = 0, \ldots, (K-1)$, with $\kappa \in \mathbb{R}$. Fig. 4(a) shows a cut through the cost function $\xi_2^{(16)}$ for $\kappa \in [0,\ 16]$. Note that for small integer values $\kappa \in \mathbb{Z}$, the cost function exhibits deep notches. The gradient of the cost function in Fig. 4(b) shows that for small $\kappa \in \mathbb{Z}$, we approximately encounter stationary points, which together with the evidence of Fig. 4(a) indicates that these stationary points are minima of the cost function. Note that as p increases, the notches in the cost function and the gradient norm deepen, but do not change location.

Two examples for interpolations $\hat{Q}_\mu^{(16)}(\mathrm{e}^{\mathrm{j}\Omega})$ according to (37) based on $\kappa = \{0, 1\}$ are shown in Fig. 5. For $\kappa = 0$ in Fig. 5(a), we obtain the lowest order eigenvector with maximum smoothness. In comparison, Fig. 5(b) illustrates the case $\kappa = 1$, which is modulated w.r.t. the interpolation for $\kappa = 0$, and hence oscillates faster and is less smooth. However, it is also an analytic solution (with the time domain shifted and now of length 7) and hence represents a local minimum for the cost function. △

F. Phase Smoothing Algorithm

Based on Theorem 5, we combine an iterative gradient search with a phase shift approach. For the iterative gradient search, in the ith iteration the phase vector $\boldsymbol{\psi}[i]$ is updated as

$$\boldsymbol{\psi}[i+1] = \boldsymbol{\psi}[i] - \rho \mathbf{H}^{-1} \frac{\partial \xi_p^{(K)}(\boldsymbol{\psi}[i])}{\partial \boldsymbol{\psi}} \ , \tag{49}$$

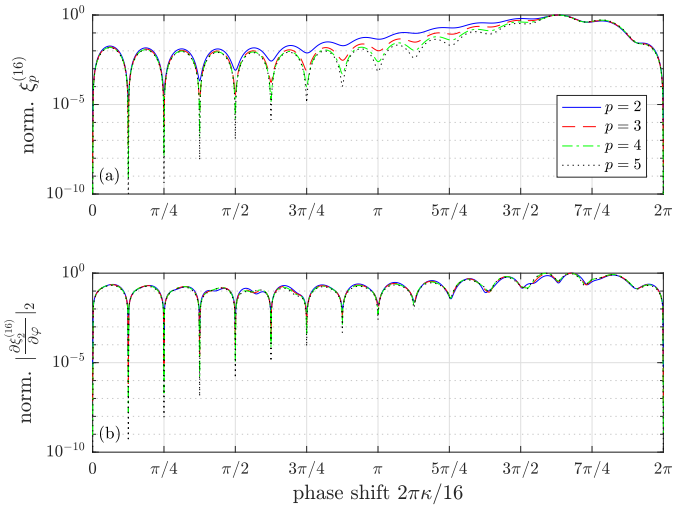

Fig. 4. (a) Cost function $\xi_p^{(16)}$, and (b) norm of its gradient for a single eigenvector with an applied phase shift $\psi_k = 2\pi\kappa k/16$ with $\kappa \in [0,\ 16]$. The curves for each $p = 2, 3, 4, 5$ are normalised with unity as maximum.

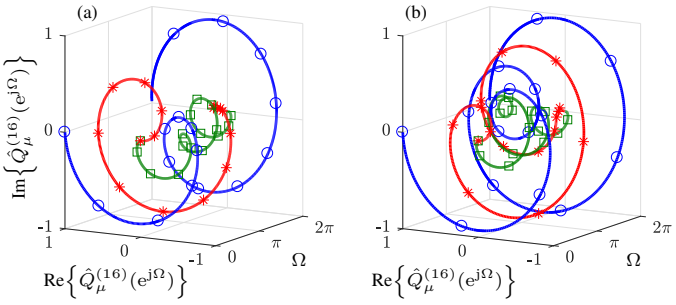

Fig. 5. Interpolations $\hat{Q}_\mu^{(16)}(e^{j\Omega})$ for $\mu = 1, 2, 3$ in blue, red, and green, respectively, for (a) $\kappa = 0$ and (b) $\kappa = 1$. The markers show the values of $\mathbf{q}'_\mu$ in (35).

with the Hessian matrix [56]

$$\mathbf{H} = 2\mathrm{Re}\{\mathrm{diag}\{\mathbf{a}^*\}\,\mathbf{C}_{K,p}\mathrm{diag}\{\mathbf{a}\} - \mathrm{diag}\{\mathrm{diag}\{\mathbf{a}^*\}\,\mathbf{C}_{K,p}\mathbf{a}\}\}\,.$$

This Gauss-Newton algorithm will converge to a stationary point of the cost function. Once sufficiently converged, a number of phase shifts are applied, and for each of these the above gradient search is performed to move to a related stationary point. For the stationary point with the smallest cost function value, again phase shifted points can be used to identify potential related stationary points, and the procedure is repeated until the cost function cannot be further minimised by gradient searches and shifts. This algorithm is summarised in Algorithm 2, with SP abbreviating 'stationary point'.

Theorem 6 (Convergence of phase smoothing algorithm): With a suitable step size ρ and a DFT length K that exceeds the support of the analytic solution, the phase smoothing Algorithm 2 will converge to the stationary point, within a set $\mathcal{S}$, with the smallest cost function $\xi_p^{(K)}$ in (41). The set $\mathcal{S}$ consists of a collection of stationary points near phase shifted versions of each other.

Proof. For a suitable step size ρ, the Gauss-Newton scheme will converge to a stationary point — this could be a (local) minimum or a saddle point. If a saddle point is encountered, the fact that phase shifted versions only approximately co-

Algorithm 2: Phase Smoothing

1: For some initialisation, find an SP $\mathbf{a}_{\mathrm{SP}}$ using (49);
2: **repeat**
3: set $\mathbf{a}'_{\mathrm{SP}} = \mathbf{a}_{\mathrm{SP}}$;
4: from $\mathbf{a}_{\mathrm{SP}}$, generate $\mathbf{a}_\kappa$, $\kappa = 0, \ldots, (K-1)$ via (45);
5: for every $\mathbf{a}_\kappa$ find an SP $\mathbf{a}_{\mathrm{SP},\kappa}$ using (49);
6: find $\mathbf{a}_{\mathrm{SP}} = \arg\min_\kappa \xi_p^{(K)}(\mathbf{a}_{\mathrm{SP},\kappa})$;
7: **until** $\|\mathbf{a}_{\mathrm{SP}} - \mathbf{a}'_{\mathrm{SP}}\|_2 < \epsilon$.

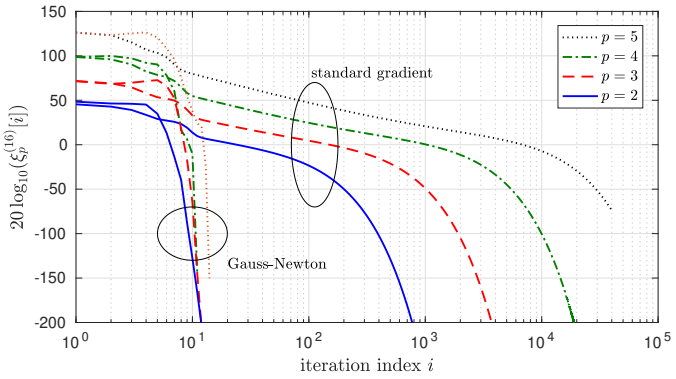

Fig. 6. Cost function evolution in dependency of the iteration index i and the derivative order p of Algorithm 2.

incide with new stationary points means that the algorithm cannot get repeatedly trapped in a saddle point. For a DFT length greater than the support of the analytic solution, Theorem 5 applies to any stationary point. This guarantees that through phase shifts we can re-initialise the gradient search to reach a solution with a smaller cost function value, until the minimum within the set of phase-related stationary points is found. ■

Example 6: Learning curves for (49) for determining the phase smoothing of the eigenvectors in Example 3 are shown in Fig. 6. The conditioning of the cost function worsens with an increase in the derivative order p, and significantly benefits from the inclusion of the inverse Hessian in (49) compared to a standard gradient search with $\mathbf{H} = \mathbf{I}$. △

VI. Adjustment of Approximation Order

We now aim to determine criteria to identify whether a phase-smoothed solution in K bins yields a sufficient approximation. For this, we investigate two criteria — (i) a loss in paraunitarity, and (ii) a reconstruction error — before commenting on the overall convergence.

A. Loss of Paraunitarity

Based on an EVD in each of K bins, the matrices $\mathbf{Q}_k$ in (16) are unitary. After the extraction of 1d eigenspaces and phase smoothing, the resulting approximation $\hat{\boldsymbol{Q}}^{(K)}(z)$ will therefore also be unitary for $z = e^{j2\pi k/K}$, $k = 0, \ldots, (K-1)$. However, generally $\boldsymbol{E}^{(K)}(e^{j\Omega}) = \hat{\boldsymbol{Q}}^{(K)}(e^{j\Omega})\hat{\boldsymbol{Q}}^{(K),\mathrm{H}}(e^{j\Omega}) - \mathbf{I} = \mathbf{0}$ $\forall\Omega \in \mathbb{R}$ cannot be assumed unless the approximation order K is sufficiently large. We therefore assess

$$\zeta_{\mathrm{PU}} = \frac{1}{2\pi}\int_0^{2\pi} \left\|\boldsymbol{E}^{(K)}(e^{j\Omega})\right\|_{\mathrm{F}}^2 \mathrm{d}\Omega \qquad (50)$$

to measure how well paraunitarity of an interpolation is satisfied. For the evaluation of (50), Parseval's theorem permits to assess the energy in $\mathbf{E}^{(K)}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{E}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$. In practice, instead of determining $\mathbf{E}^{(K)}[n]$ via a K-point IDFT from the sample points $\mathbf{Q}_k$, we perform a zero-padded $2K$-IDFT. This effectively performs a two-fold Dirichlet interpolation of $\boldsymbol{E}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ and ensures that the product $\hat{\boldsymbol{Q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})\hat{\boldsymbol{Q}}^{(K),\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})$ — equivalent to a cyclic convolution in the time domain — does not cause any time-domain aliasing due to cyclic wrap-around. Thus the approximation according to (7) in Theorem 4 applies.

Lemma 1 (Sufficiency of Loss of Paraunitarity): If for a small threshold ϵ_{PU}, $\zeta_{\mathrm{PU}} > \epsilon_{\mathrm{PU}}$, then the approximation order K is too small.

Proof. We know that an analytic paraunitary solution exists. For (50), a two-fold Dirichlet interpolation of the phase-smoothed result is evaluated. If for some frequency, a deviation of $\hat{\boldsymbol{Q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ from orthonormality is found, then the interpolation is incorrect, and the sample points on which interpolation is based are not sufficiently closely spaced. Therefore, $\zeta_{\mathrm{PU}} > \epsilon_{\mathrm{PU}}$ is sufficient to indicate that the DFT size K, which also represents the approximation order, is too small. ■

Note that $\zeta_{\mathrm{PU}} \leq \epsilon_{PU}$ does not guarantee that a suitable K has been found. Hence a second criterion is required.

B. Reconstruction Error

With the approximation $\hat{\boldsymbol{Q}}^{(K)}(z)$, as well as the extracted analytic eigenvalues $\hat{\boldsymbol{\Lambda}}(z)$ using [36], we can reconstruct $\hat{\mathbf{R}}^{(K)}[\tau] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{Q}}^{(K)}(z)\hat{\boldsymbol{\Lambda}}(z)\hat{\boldsymbol{Q}}^{(K),\mathrm{P}}(z)$, and determine

$$\zeta_{\mathrm{R}} = \sum_{\tau} \left\| \mathbf{R}[\tau] - \hat{\mathbf{R}}^{(K)}[\tau] \right\|_{\mathrm{F}}^2 \tag{51}$$

as a reconstruction error. Provided that the eigenvalue from [36] is sufficiently accurate, we can state:

Lemma 2 (Reconstruction Error): If and only if for a small threshold ϵ_{R}, $\zeta_{\mathrm{R}} < \epsilon_{\mathrm{R}}$, then the approximation order K is sufficient.

Proof. If the approximation is sufficiently good and $\hat{\boldsymbol{Q}}^{(K)}(z)$ is close to $\boldsymbol{Q}(z)$, then ζ_{R} must be small. For the 'only if' part, if ζ_{R} is small, then $\hat{\boldsymbol{Q}}^{(K)}(z)$ behaves almost indistinguishably from $\boldsymbol{Q}(z)$. Due to the uniqueness of the parahermitian matrix EVD [14], [15], and the uniqueness of analytic functions [44], in fact $\hat{\boldsymbol{Q}}^{(K)}(z)$ must closely approximate the analytic eigenvectors save of their phase ambiguity. ■

C. Order Increase and Overall Convergence

The quantities defined in Secs. VI-A and VI-B now enable a criterion for the loop in Algorithm 1. There, we will iterate as long as $\zeta_{\mathrm{PU}} > \epsilon_{\mathrm{PU}}$ for some threshold ϵ_{PU}. Once $\zeta_{\mathrm{PU}} \leq \epsilon_{\mathrm{PU}}$, because loss of paraunitarity is only sufficient, we still iterate as long as $\zeta_{\mathrm{R}} > \epsilon_{\mathrm{R}}$, with some threshold ϵ_{R}.

Theorem 7 (Convergence to Analytic Eigenvectors): Algorithm 1 with the prescribed loop criterion and suitable thresholds ϵ_{PU} and ϵ_{R} can converge close to analytic eigenvectors.

Proof. If an approximation order K is too short to extract analytic eigenvalues, likely both ζ_{PU} and ζ_{R}, but at the very least the latter, will not fall below their thresholds. Thus, Algorithm 1 will increase K. Since analytic eigenvalues exist and decay at least exponentially [14], [15], at some point the truncation error will be sufficient small, such that the support of the analytic eigenvalues is shorter than K. Thus, Theorem 6 then guarantees to yield smooth eigenvectors, save of some arbitrary phase. Note that criterion ζ_{PU} is independent of any previously extracted eigenvalues via [36], and hence does not depend on the accuracy of those eigenvalues. The accuracy of the eigenvalues will impact on the necessary and sufficient criterion ζ_{R}, and thus potentially limit the achievable approximation error for the eigenvectors. Since it is possible to extract analytic eigenvalues with arbitrarily small approximation errors [36], the only limitation for approximation errors on the eigenvectors are the approximation order K, and Theorem 6 potentially returning a stationary point with an insufficient smoothness. ■

In practice, for all simulations run to date with known ground truth, Algorithm 2 has in fact found the global minimum, hence allowing Algorithm 1 to converge arbitrarily closely to the analytic eigenvectors.

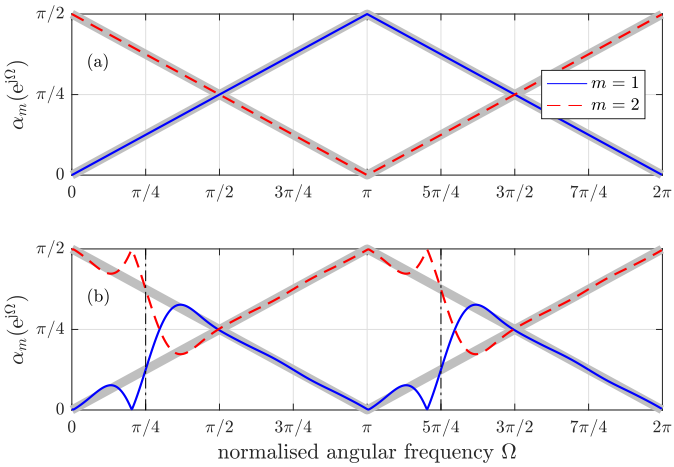

Fig. 7. Hermitian angles for the extracted eigenvectors using (a) the proposed algorithm and (b) the SBR2 algorithm [7], [17]. The ground truth angles are under-laid in grey. The lower plot highlights the frequencies $\Omega = \{\frac{\pi}{4}, \frac{5\pi}{4}\}$, where algebraic multiplicities enforce a permutation as dash-dotted lines.

VII. Simulations and Results

A. Numerical Example

As an initial numerical example, we utilise $\boldsymbol{R}(z) : \mathbb{C} \to \mathbb{C}^{2\times 2}$ from Example 1. The ground truth eigenvectors have a minimum order of 1 for a phase term $\psi_m(z) = 1$ in (5). The Hermitian angles $\alpha_m(\mathrm{e}^{\mathrm{j}\Omega})$ for the eigenvectors $\boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})$ relative to $\boldsymbol{q}_1(\mathrm{e}^{\mathrm{j}0})$ as defined in (31) are shown in Fig. 7(a), together with the Hermitian angles for the extracted eigenvectors using the proposed algorithm, which are also of order 1 and match the ground truth.

When using the SBR2 algorithm [7], [17], the approximated eigenvalues are guaranteed to be spectrally majorised [23], and as a result, the extracted eigenvectors need to approximate discontinuities at $\Omega = \frac{\pi}{4}$ and $\Omega = \frac{5\pi}{4}$. Fig. 7(b) shows the

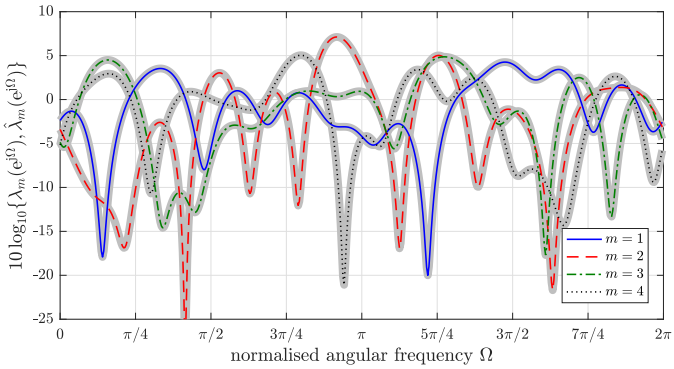

Fig. 8. Analytic eigenvalues for one ensemble probe with $L = 12$, with the ground truth $\lambda_m(e^{j\Omega})$ underlaid in grey and the extracted eigenvalues $\hat{\lambda}_m(e^{j\Omega})$ according to [36] in colour.

progression of the Hermitian angles of these estimated eigenvectors, which require an approximation order of 87 — almost two orders of magnitude above the order of the extracted analytic eigenvectors, and with a significant reconstruction error compared to the results using the proposed algorithm.

B. Ensemble Test

In order to form a more exhaustive assessment of the proposed algorithm against benchmarks with proven convergence such as SBR2 [7], [17] and SMD [21], we use the model outlined in (2) and Fig. 2, to generate an ensemble of CSD matrices $\boldsymbol{R}(z) : \mathbb{C} \to \mathbb{C}^{4\times4}$, where the innovation filters $G_\ell(z) \bullet\!\!-\!\!\circ\, g_\ell[n]$, $\ell = 1,\ldots,4$ are of order $L \in \{1,2,\ldots,12\}$. The $L+1$ coefficients of $g_\ell[n]$ for every ensemble probe are uncorrelated and drawn from a zero mean unit variance complex Gaussian distribution, and normalised such that $\sum_n |g_\ell[n]|^2 = 1$. The convolutive mixing matrices $\boldsymbol{H}(z)$, also of order L, are generated from a concatenation of L elementary paraunitary matrices [1],

$$\boldsymbol{H}(z) = \prod_{i=1}^{L} \left(\mathbf{I} + (z^{-1}-1)\mathbf{v}_i\mathbf{v}_i^{\mathrm{H}}\right) \ , \tag{52}$$

whereby each vector $\mathbf{v}_i$ has its uncorrelated coefficients drawn from a zero mean unit variance complex Gaussian distribution, but is normalised such that $\|\mathbf{v}_i\|_2^2 = 1$. Because of the resulting paraunitarity of $\boldsymbol{H}(z)$, (2) directly reflects the analytic parahermitian EVD with the lowest order $\boldsymbol{Q}(z) = \boldsymbol{H}(z)$, and $\boldsymbol{\Lambda}(z) = \boldsymbol{G}(z)\boldsymbol{G}^{\mathrm{P}}(z)$. Using this model, for each value of L, an ensemble of 500 randomised CSD matrices $\boldsymbol{R}(z)$ of order $4L$ is generated for every value of L. As an example for one ensemble probe, Fig. 8 characterises the eigenvalues on the unit circle, $\boldsymbol{\Lambda}(e^{j\Omega})$, for $L = 12$.

When factorising $\boldsymbol{R}(z)$ with the various methods, in the case of SBR2 and SMD, internal trimming in the algorithm curbs the growth in polynomial order as iterations progress; additionally, the order of the extracted paraunitary matrices is optimised using a shift-corrected truncation described in [46], which removes the potential time shift encountered in (5). The algorithms terminate after 300 or 600 iterations for SMD and SBR2, respectively, or if the off-diagonal energy has

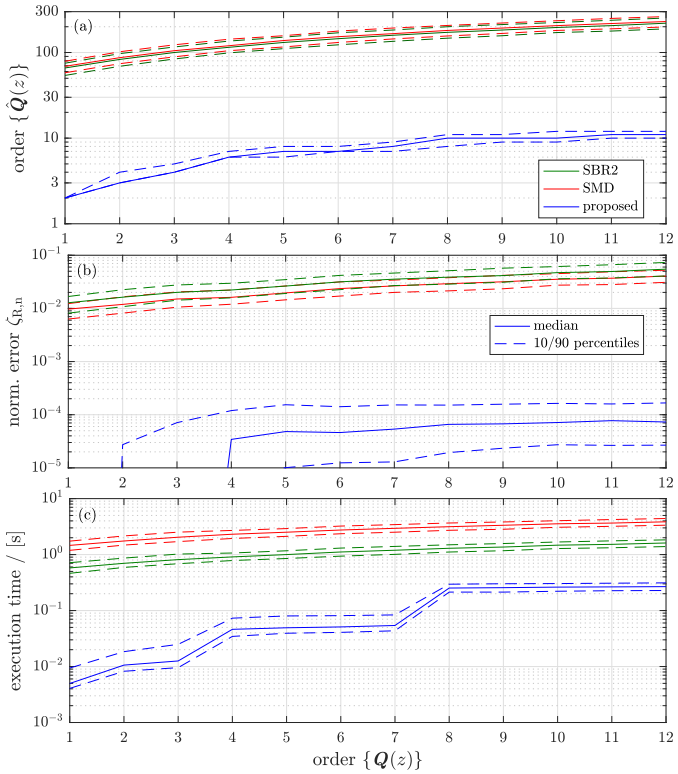

Fig. 9. Ensemble results showing (a) the achievable order, (b) reconstruction error, and (c) execution time for the proposed eigenvector extraction vs the performance of SBR2 [7], [17] and SMD [21] algorithms.

dropped below 10^{-4} of the total energy in $\boldsymbol{R}(z)$. The proposed algorithm operates with a paraunitarity error threshold $\epsilon_{\mathrm{PU}} = 10^{-5}$; although (50) could not be shown to represent a necessary convergence measure, in practice across 6000 ensemble simulations, it was sufficient to check $\zeta_{\mathrm{PU}} < \epsilon_{\mathrm{PU}}$ alone in order to achieve a reconstruction $\zeta_{\mathrm{R}} < 10^{-4}$.

Fig. 9(a) shows the order of the estimated eigenvectors, in $\hat{\boldsymbol{Q}}(z)$, versus that of the ground truth eigenvalues generated via (52). Additionally, Fig. 9(b) and (c) show the normalised reconstruction error $\zeta_{\mathrm{R,n}} = \zeta_{\mathrm{R}} / \sum_\tau \|\mathbf{R}[\tau]\|_{\mathrm{F}}^2$, whereby ζ_{R} is normalised for comparison across an ensemble of different CSD matrices, and the execution time of the algorithms.

It is evident from Fig. 9 that the proposed analytic eigenvalue extraction achieves the task with an approximation order close to the ground truth, and with a small reconstruction error. In contrast, because in many cases the ground truth eigenvalues are not spectrally majorised, the spectrally majorised solution targeted by the SBR2 and SMD algorithms requires significantly higher approximation orders in Fig. 9(a). Subsequently, the generally insufficient approximation of a discontinuity by a polynomial leads to poorer reconstruction errors in Fig. 9(b). The low order achieved by the proposed algorithm therefore also means a significant reduction in execution time as shown in Fig. 9(c). The latter clearly shows the execution time is affected more by the required DFT length K — a power of 2 — than the order of the ground truth $\boldsymbol{Q}(z)$.

For the results in Fig. 9, the analytic eigenvalue extraction was based on the exact eigenvalues. We have also estimated the analytic eigenvalues using the algorithm in [36] with

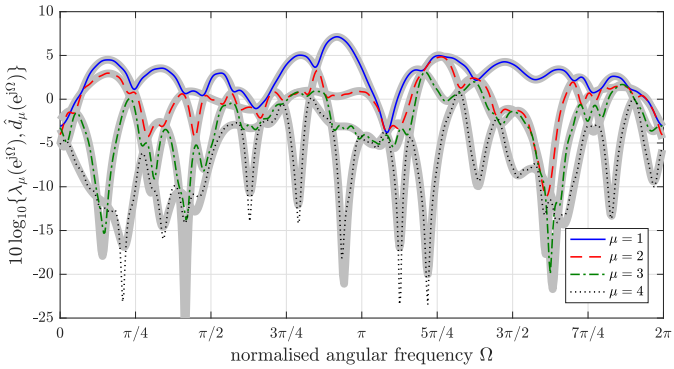

Fig. 10. Eigenvalues for one ensemble probe with $L = 12$, with the analytic ground truth $\lambda_m(e^{j\Omega})$ underlaid in grey and the approximately spectrally majorised solution $\hat{d}_m(e^{j\Omega})$ obtained by the SMD algorithm [21] in colour.

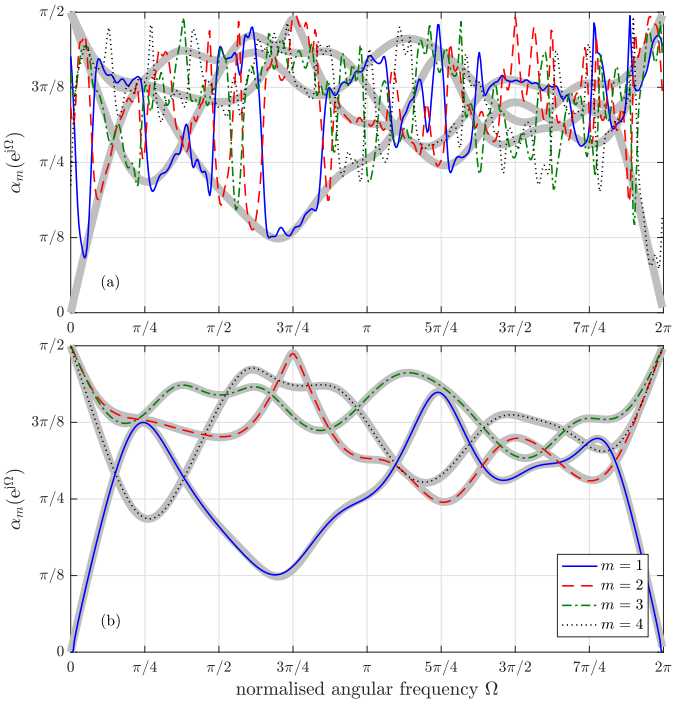

Fig. 11. Evolution of the Hermitian angles (31) of 1-d subspaces for one ensemble probe for $L = 12$, with the ground truth for $\boldsymbol{Q}(z)$ underlaid in grey; the coloured curves show the Hermitian angles for (a) the eigenvectors $\hat{\boldsymbol{U}}(z)$ obtained by the SMD algorithm [21], and (b) the extracted analytic eigenvectors $\hat{\boldsymbol{Q}}(z)$ using the proposed algorithm.

a predefined precision of 10^{-5}, followed by the proposed analytic eigenvector extraction. Because of the high accuracy of the extracted eigenvalues, there is no significant difference to the results shown in Fig. 9, with the orders remaining the same, and achievable values for the normalised reconstruction error $\zeta_{\mathrm{R},n}$ within 1.7% of those with perfect knowledge of the analytic eigenvalues.

To exemplify the performance difference between the proposed analytic eigenvector extraction and the SBR2/SMD algorithms, for the ensemble probe characteristed in Fig. 8, Fig. 10 shows the eigenvalues $\hat{d}_\mu(e^{j\Omega})$ extracted by the SMD algorithm, which generally satisfy spectral majorisation. The resulting approximation of the permuted ground truth analytic eigenvalues leads to the estimated eigenvectors having to approximate piecewise analytic, discontinuous functions, as evident from the evolution of the Hermitian angles of the estimated eigenvectors in Fig. 11(a). There, at the algebraic multiplicities of the eigenvalues shown in Fig. 10, the estimated eigenvectors switch between ground truth eigenvectors, underlaid in grey occur, resulting in large associated time domain support. In contrast, the estimated analytic eigenvectors using the proposed algorithm in Fig. 11(b) follow the ground truth closely, resulting in smooth eigenspaces that permit much lower order approximations than those obtainable by the SMD-estimated functions in Fig. 11(a).

VIII. Conclusions

We have proposed an analytic eigenvector extraction algorithm for parahermitian matrices. The method relies on the previous extraction of analytic eigenvalues via [36], and then iteratively increases the DFT length and hence approximation order until a suitably close approximation is found. In each iteration, the method extracts smooth 1-d eigenspaces across frequency bins, which need to be suitably continued across algebraic multiplicites. This is followed by a phase smoothing operation, which identifies eigenvectors of short order within every 1-d eigenspace. The approach demonstrates significantly enhanced performance over state-of-the-art algorithms, such as SBR2 and SMD. While we have not proven its necessity, the error in paraunitarity seems to be sufficient, across all our simulations, as a metric to decide when to terminate the algorithm, and has the benefit of being independent of the accuracy with which analytic eigenvalues have been extracted.

While some applications such as subband coding [7], [18] rely on spectral majorisation of the eigenvalues, subspace-based applications such as those required for broadband beamforming [3] or source separation [5] benefit from both the accurate extraction of subspaces and the computational savings that result from the significantly lower polynomial order of the analytic eigenvectors obtainable with [36] and the proposed approach.

Acknowledgments

We would like to thank our anonymous reviewers for their helpful and constructive input. We are particularly grateful to one of them for suggesting the closed-form solution in Sec. IV-C using the Procrustes method.

References

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

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

[3] S. Weiss, S. Bendoukha, A. Alzin, F. Coutts, I. Proudler, and J. Chambers, "MVDR broadband beamforming using polynomial matrix techniques," in *Proc. 23rd Eur. Signal Process. Conf.*, Nice, France, pp. 839–843, Sep. 2015.

[4] V. W. Neo, E. d'Olne, A. H. Moore, and P. A. Naylor, "Fixed beamformer design using polynomial eigenvalue decomposition," in *2022 Int. Workshop on Acoustic Signal Enhancement*, Bamberg, Germany, pp. 1–5, Sep. 2022.

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

[6] S. Weiss, S. Redif, T. Cooper, C. Liu, P. Baxter, and J. McWhirter, "Paraunitary oversampled filter bank design for channel coding," *EURASIP J. Adv. Signal Processing*, Vol. 2006, pp. 1–10, 2006.

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

[8] V. W. Neo, C. Evers, and P.A. Naylor, "Speech dereverberation performance of a polynomial-EVD subspace approach," in *Proc. 29th Eur. Signal Process. Conf.*, pp. 221–225, Dublin, Ireland, Aug. 2021

[9] ——, "Enhancement of Noisy Reverberant Speech Using Polynomial Matrix Eigenvalue Decomposition,"" in IEEE/ACM Trans. Acoustics, Speech and Language Process., vol. 29, pp. 3255-3266, Oct. 2021.

[10] ——, "Polynomial matrix eigenvalue decomposition of spherical harmonics for speech enhancement," in *ICASSP*, pp. 786–790, May 2021.

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

[12] N. Moret, A. Tonello, and S. Weiss, "MIMO precoding for filter bank modulation systems based on PSVD," in *Proc. IEEE Vehicular Technology Conf. — Spring*, Budapest, Hungary, pp. 1–5, May 2011.

[13] D. Hassan, S. Redif, J. G. McWhirter, and S. Lambotharan, "Polynomial GSVD beamforming for two-user frequency-selective MIMO channels," *IEEE Trans. Signal Process.*, vol. 69, pp. 948–959, Jan. 2021.

[14] S. Weiss, J. Pestana, and I.K. Proudler, "On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2659–2672, May 2018.

[15] S. Weiss, J. Pestana, I. Proudler, and F. Coutts, "Corrections to "on the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix"," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6325–6327, Dec. 2018.

[16] G. Barbarino and V. Noferini, "On the Rellich eigendecomposition of para-Hermitian matrices and the sign characteristics of $*$-palindromic matrix polynomials," *arXiv:2211.15539*, 2022.

[17] J.G. McWhirter, P.D. Baxter, T. Cooper, S. Redif, and J. Foster, "An EVD Algorithm for Para-Hermitian Polynomial Matrices," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2158–2169, May 2007.

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

[19] S. Icart and P. Comon, "Some properties of Laurent polynomial matrices," in *9th IMA Conf. Mathematics in Signal Process.*, pp. 1-5, Birmingham, UK, Dec. 2012.

[20] Z. Wang, J.G. McWhirter, J. Corr, and S. Weiss, "Multiple shift second order sequential best rotation algorithm for polynomial matrix EVD," in *Proc. 23rd Eur. Signal Process. Conf.*, Nice, France, pp. 844–848, Sep. 2015.

[21] S. Redif, S. Weiss, and J. McWhirter, "Sequential matrix diagonalization algorithms for polynomial EVD of parahermitian matrices," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 81–89, Jan. 2015.

[22] J. Corr, K. Thompson, S. Weiss, J. McWhirter, S. Redif, and I. Proudler, "Multiple shift maximum element sequential matrix diagonalisation for parahermitian matrices," in *Proc. IEEE Statistical Signal Process. Conf.*, Gold Coast, Australia, p. 312–315, June 2014.

[23] J. G. McWhirter and Z. Wang, "A novel insight to the SBR2 algorithm for diagonalising para-Hermitian matrices," in *11th IMA Conf. Mathematics in Signal Process.*, pp. 1-5, Birmingham, UK, Dec. 2016.

[24] A. Tkacenko and P. Vaidyanathan, "On the spectral factor ambiguity of FIR energy compaction filter banks," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 380–385, Jan. 2006.

[25] ——, "Iterative greedy algorithm for solving the FIR paraunitary approximation problem," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 146–160, Jan. 2006.

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

[27] B. Girod, R. Rabenstein, and A. Stenger, *Signals and Systems*. Chichester: J. Wiley & Sons, 2001.

[28] F. Rellich, "Störungstheorie der Spektralzerlegung. I. Mitteilung. Analytische Störung der isolierten Punkteigenwerte eines beschränkten Operators," *Mathematische Annalen*, **113**:DC–DCXIX, 1937.

[29] L. Dieci and T. Eirola, "On smooth decompositions of matrices," *SIAM J. Matrix Analysis Appl.*, vol. 20, no. 3, pp. 800–819, 1999.

[30] E.S. Van Vleck, *Numerical algebra, matrix theory, differential-algebraic equations and control theory: Festschrift in honor of Volker Mehrmann*. Springer, ch. Continuous Matrix Factorizations, pp. 299–318, 2015.

[31] B. De Moor and S. Boyd, "Analytic properties of singular values and vectors," KU Leuven, Tech. Rep., 1989.

[32] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nicols, "Numerical computation of an analytic singular value decomposition of a matrix valued function," *Numer. Math*, vol. 60. pp. 1–40, 1991.

[33] K. Wright, "Differential equations for the analytic singular value decomposition of a matrix," *Num. Mathematik*, vol. 63, no. 1, pp. 283–295, Dec. 1992.

[34] D. Janovská, V. Janovský, and K. Tanabe, "An algorithm for computing the analytic singular value decomposition," *Int. J. Math. & Comp.Sciences*, vol. 2, no. 11, pp. 765–770, 2008.

[35] Y. Nakatsukasa, V. Noferini, and N. Trefethen, "Computing the analytic SVD," Chebfun – numerical computing with functions, 2016.

[36] S. Weiss, I.K. Proudler, and F.K. Coutts, "Eigenvalue decomposition of a parahermitian matrix: Extraction of analytic eigenvalues," *IEEE Trans. Signal Process.*, vol. 69, pp. 722–737, Jan. 2021.

[37] V.W. Neo and P.A. Naylor, "Second order sequential best rotation algorithm with Householder reduction for polynomial matrix eigenvalue decomposition," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, Brighton, UK, pp. 8043–8047, May 2019.

[38] M. Tohidian, H. Amindavar, and A.M. Reza, "A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices," *EURASIP J. Advanced Signal Process.*, vol. 2013, no. 1, pp. 1–16, 2013.

[39] S. Weiss, I.K. Proudler, F.K. Coutts, and J. Pestana, "Iterative approximation of analytic eigenvalues of a parahermitian matrix EVD," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, Brighton, UK, pp. 8038-8042, May 2019.

[40] S. Weiss, J. Selva, and M.D. Macleod, "Measuring smoothness of trigonometric interpolation through incomplete sample points," in *28th Eur. Signal Process. Conf.*, Amsterdam, Netherlands, pp. 2319–2323, Jan. 2021.

[41] S. Weiss, I. Proudler, F. Coutts, and J. Deeks, "Extraction of analytic eigenvectors from a parahermitian matrix," in *Int. Conf. Sensor Signal Process. Defence*, pp. 1-5, Edinburgh, UK, Jan. 2020.

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

[43] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*. New York: Academic Press, 1982.

[44] D. Burdick and F.D. Lesley, "Some uniqueness theorems for analytic functions," *The American Mathematical Monthly*, vol. 82, no. 2, pp. 152–155, 1975.

[45] P. A. Regalia, S. K. Mitra, and P. P. Vaidyanathan, "The digital all-pass filter: a versatile signal processing building block," *Proc. IEEE*, vol. 76, no. 1, pp. 19–37, Jan. 1988.

[46] J. Corr, K. Thompson, S. Weiss, I. Proudler, and J. McWhirter, "Row-shift corrected truncation of paraunitary matrices for PEVD algorithms," in *Proc. 23rd Eur. Signal Process. Conf.*, Nice, France, pp. 849–853, Aug. 2015,

[47] L. V. Ahlfors, *Complex analysis: An introduction to the theory of analytic functions of one complex variable*. NY: McGraw-Hill, 1953.

[48] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, ML: John Hopkins Univ. Press, 1996.

[49] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[50] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, June 2007.

[51] W. Wirtinger, "Zur formalen Theorie der Funktionen von mehreren komplexen Veränderlichen," *Math. Annalen*, vol. 97, pp. 357 – 376, 1926.

[52] A.V. Oppenheim, R.W. Schafer, and J.R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Pearson, 1999.

[53] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 87–109, May 2015.

[54] Z. Luo, W. Ma, A. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[55] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[56] F. K. Coutts, K. Thompson, J. Pestana, I. Proudler, and S. Weiss, "Enforcing eigenvector smoothness for a compact DFT-based polynomial eigenvalue decomposition," in *Proc. IEEE 10th Sensor Array*

Multichannel Signal Process. Workshop, Sheffield, UK, pp. 159-163, July 2018.

Stephan Weiss received a Dipl.-Ing. degree from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1995, and a Ph.D. degree from the University of Strathclyde, Glasgow, Scotland, in 1998, both in electronic and electrical engineering.

He is professor for signal processing at the University of Strathclyde, Glasgow, following previous academic appointments at both the Universities of Strathclyde and Southampton. His research interests lie in adaptive, multirate, and array signal processing with applications in acoustics, communications, audio, and biomedical signal processing, where he has published more than 300 technical papers. For his work in biomedical signal processing, he was a co-recipient of the 2001 research award of the German society on hearing aids. In 2011 he was a co-recipient of the VTC-Spring best paper award in the MIMO systems track.

Dr Weiss is a member of EURASIP and a senior member of the IEEE. He was the technical co-chair for EUSIPCO 2009 and general chair of IEEE ISPLC 2014, both organised in Glasgow, and special session co-chair for ICASSP 2019.

Ian K. Proudler graduated from Oxford University in 1978 having read Physics. He spent two years doing R&D work in the electronics industry before obtaining a Ph.D. in Digital Signal Processing from Cambridge University in 1984.

He is currently a Visiting Professor of Signal Processing at the University of Strahclyde. From 1986 until 2011 he worked in the Defence sector looking into various adaptive digital signal processing issues such as: numerical stability and efficient computation; antenna algorithm for HF communications; signal separation for ESM purposes; magnetic detection for maritime surveillance; and GPS anti-jam systems. He has published some 100+ research papers, contributed to four textbooks and holds a patent on an adaptive filtering architecture.

He was awarded the John Benjamin Memorial Prize, in 1992 and 2001, and the IEE J.J. Thomson Medal, in 2002, for his work on signal processing algorithms. He was an Honorary Editor for IEE Proceedings: Radar, Sonar and Navigation for ten years. He has been on the organising committee of several international conferences.

Fraser K. Coutts graduated with an MEng (distinction) in Electronic & Electrical Engineering from the University of Strathclyde in 2015, and received his PhD in 2019 from the same institution.

Dr Coutts is a research associate at the Institute for Digital Communications (IDCOM) at the University of Edinburgh, where he is working on secondary inference from compacted data. He previously was a PhD researcher within the Centre for Signal & Image Processing (CeSIP) at the University of Strathclyde, holding a prestigious Caledonian Scholarship from the Carnegie Trust. He has collaborated widely, and his interests span linear algebraic methods for signal processing, hyperspectral imaging, Doppler radar, as well as general algorithm development and implementations. He has published more than 30 peer-reviewed conference and journal papers in these areas, and received a best paper award at IEEE SAM'18.

Faizan A. Khattak graduated with a BEng in Electronic & Electrical Engineering/ from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan, in 2016, and received an MSc in Electronic & Electrical Engineering with distinction from the University of Strathclyde, Gkasgow, Scotland, in 2019.

The recipient of a prestigious Commonwealth Scholarship, he is currently a PhD candidate at the University of Strathclyde pursuing theory, algorithms, and applications of polynomial matrix algebra. His interests generally are in signal processing and communications, with an emphasis on algorithm design and implementation.