
Article:
Khattak, F. A., Proudler, I.K., Weiss, S. et al. (Accepted: 2026) Analytic Gram-Schmidt Orthogonalisation and QR Decomposition for Polynomial Matrices. IEEE Transactions on Signal Processing. ISSN: 1053-587X (In Press)

Analytic Gram-Schmidt Orthogonalisation and QR Decomposition for Polynomial Matrices

Faizan A. Khattak, Ian K. Proudler, Stephan Weiss, and Sebastian J. Schlecht

***Abstract*—We extend the concept of Gram-Schmidt orthogonalisation from a set of ordinary vectors to vectors of analytic functions. Firstly, we address the normalisation of such a vector. Even if its Euclidean norm possesses spectral zeros, we show that it is always possible to find an analytic normalised vector that has unit norm everywhere on the unit circle. Secondly, a sequence of such normalisation steps can be utilised for a Gram-Schmidt procedure applied to analytic matrices of full spatial rank. Analogous to the case of ordinary matrices, where the Gram-Schmidt procedure leads to a QR factorisation, we prove that an analytic QR decomposition exists with an analytic paraunitary matrix and an upper-right triangular matrix of analytic functions as factors. Thirdly, we present algorithms with proven convergence for both the analytic vector normalisation and the analytic QR decomposition. This type of decomposition can find applications in broadband multiple-input multiple-output systems, and we compare our algorithmic realisation to existing solutions in the literature, which do not necessarily converge towards an analytic decomposition.**

I. Introduction

The QR decomposition (QRD) is a key component of the basic toolkit of linear algebraic decompositions — the 'big six' [1]. With these decompositions, ordinary matrices can be factorised; some of these decompositions have been extended to matrices of analytic functions in the complex variable z, such as matrices of transfer functions of causal stable systems, including the eigenvalue [2]–[4] and the singular value decompositions [5], [6]. While standard algebraic tools would only factorise such matrices for one specific parameter value, analytic decompositions perform a decomposition for all z within a region of convergence. These analytic decompositions can be used to solve broadband multichannel problems [7] such as MIMO precoding and equalisation [8], [9], subband coding [10], angle of arrival estimation [11], broadband adaptive beamforming [12], or speech enhancement [13]–[15].

The QRD for polynomial matrices has previously been considered in the literature [16]–[23]. However, these algorithms were developed without an understanding whether an analytic QRD exists or not. Algorithms approximating the QRD of a polynomial matrix include an iterative time domain approach in [17], [20], which uses elementary paraunitary operations to clear the lower left-triangular part of a matrix column by column. An enhanced updating scheme which removes more power from the lower-left triangular matrix in every step is presented in [21], [22]. A discrete Fourier transform (DFT) domain method in [23] performs a bin-wise QRD and aims to establish phase-coherence across frequency bins. The approach in [16], [18], [19] does not address existence either, but notes that infinite-length solutions may arise because of the operations that are performed on polynomial terms. To address this, the approach in [18], [19] operates in the domain of rational polynomials but retains the denominator polynomials separately, until a normalisation step is performed at the end. Ultimately, such algorithms have been applied in communications receivers such as [17], [20] in a Bell Labs layered space-time encoding architecture [27] or similarly in [16], [18], [19], whereby instead of many ordinary per-subcarrier QRDs, a polynomial QRD is determined from a reduced number of pilot subcarriers, from which the QR factors in the remaining subcarriers can be evaluated at low cost. This requires that such a QRD of a polynomial must possess sufficiently smooth factors.

In this paper, we address the extension of the Gram-Schmidt orthogonalisation procedure [28] to find an orthogonal basis for the columns of a matrix $\boldsymbol{A}(z)$ that is analytic in $z \in \mathbb{C}$, resulting in a paraunitary matrix $\boldsymbol{Q}(z)$ and, as a byproduct, in an upper triangular matrix $\boldsymbol{R}(z)$ that also serve as factors of an analytic extension of the QR decomposition. A paraunitary matrix is the extension of an orthogonal or unitary matrix to analytic matrices. Defining the parahermitian transpose $\boldsymbol{Q}^{\mathrm{P}}(z) = \{\boldsymbol{Q}(1/z^*)\}^{\mathrm{H}}$ we have $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}_M$ [29], where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Crucial, in this context, is that on the unit circle, the extension requires a frequency-dependent normalisation of a vector; we show that even though its norm may be zero at isolated frequency points, the normalisation can always be performed. As long as the determinant of $\boldsymbol{A}(z)$ only possesses isolated zeros on the unit circle, an extended Gram-Schmidt procedure will always lead to a paraunitary analytic $\boldsymbol{Q}(z)$. As a result, we show that a QRD of the type $\boldsymbol{A}(z) = \boldsymbol{Q}(z)\boldsymbol{R}(z)$ with an upper triangular analytic matrix $\boldsymbol{R}(z)$ exists for any such matrix $\boldsymbol{A}(z)$. This provides the basis for the decompositions to which the approximate algorithms in [16]–[19], [21], [23] aspire, including smoothness of the factors due to the infinite differentiability of analytic functions [26]. Further, in this paper, we provide algorithms for normalisation of a vector of analytic functions, as well as for the analytic Gram-Schmidt procedure and for an analytic QRD.

The work of F.A. Khattak was supported by the Commonwealth Scholarship Commission while studying towards a PhD at the University of Strathclyde.

F.A. Khattak is with the School of Computer Science, University of Leeds, Leeds LS2 9JT, UK (e-mail f.a.khattak@leeds.ac.uk).

I.K. Proudler, and S. Weiss are with the Dept. of Electronic & Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, Scotland (e-mail {ian.proudler;stephan.weiss}@strath.ac.uk).

S.J. Schlecht is with the Chair for Multimedia Communications and Signal Processing, Friedrich-Alexander Universität Erlangen-Nürnberg, Cauerstraße 7, 91058 Erlangen, Germany (e-mail sebastian.schlecht@fau.de).

Below, Sec. II reviews the Gram-Schmidt procedure, its relation to the QRD, and some background on linear algebra for matrices of analytic functions. The Gram-Schmidt procedure is extended to an analytic version in Sec. III. In order to numerically extract solutions from a given $\boldsymbol{A}(z)$, Sec. IV first addresses the normalisation of a vector of analytic functions. This allows to construct algorithms for an analytic Gram-Schmidt procedure and an analytic QRD in Sec. V. Sec. VI explores and benchmarks these against existing algorithms in a number of simulations. Conclusions are drawn in Sec. VII. Code for all algorithms and for reproducing the figures is available on GitHub[1].

II. Preliminaries

We first address relevant background material — this includes the (modified) Gram-Schmidt procedure for ordinary matrices in Sec. II-A. We further summarise some background on the algebra of analytic functions in Sec. II-B, in preparation for the extension of the Gram-Schmidt orthogonalisation and the QR decomposition in subsequent sections.

A. Gram-Schmidt Orthogonalisation

Let us assume that a set of linearly independent vectors $\mathbf{a}_m \in \mathbb{C}^M$, $m = 1, \ldots, M$ forms the columns of a thus full rank matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_M] \in \mathbb{C}^{M \times M}$. From this set of vectors $\mathbf{a}_m$, we aim to find a set of orthogonal vectors $\mathbf{q}_m$, $m = 1, \ldots, M$ using Gram-Schmidt orthogonalisation such that $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_M]$ is unitary. Gram-Schmidt in its standard form maintains the strict ordering of the columns of $\mathbf{A}$. It starts by normalising $\mathbf{a}_1$,

$$\mathbf{q}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|_2 \ , \tag{1}$$

with respect to its ℓ_2-norm $\|\cdot\|_2$. A subsequent mth basis vector $\mathbf{q}_m$, $m = 2, \ldots, M$, is iteratively calculated by projecting away from the already established $(m-1)$ orthonormal basis vectors, and then normalising it to unit norm, such that

$$\mathbf{q}_m = \bar{\mathbf{a}}_m / \|\bar{\mathbf{a}}_m\|_2 \ , \tag{2}$$

where

$$\bar{\mathbf{a}}_m = \left(\mathbf{I}_M - \sum_{\mu=1}^{m-1} \mathbf{q}_\mu \mathbf{q}_\mu^{\mathrm{H}} \right) \mathbf{a}_m = \mathbf{a}_m - \sum_{\mu=1}^{m-1} \mathbf{q}_\mu (\mathbf{q}_\mu^{\mathrm{H}} \mathbf{a}_m) \ , \tag{3}$$

and $\{\cdot\}^{\mathrm{H}}$ is the Hermitian transposition. Note that if $\mathbf{A}$ is not full rank, the algorithm stops at some iteration m with $\|\bar{\mathbf{a}}_m\|_2 = 0$. To overcome this one can modify the Gram-Schmidt orthogonalisation to include column pivoting (see e.g. Algorithm 5.4.1 in [28]), which moves columns with larger norms to the left. Ultimately we end up with a number of columns of vanishing norm on the right hand side. For simplicity we assume that the columns $\mathbf{a}_m$ are linearly independent and therefore that $\mathbf{A}$ has full rank.

Gram-Schmidt orthogonalisation does not only find an orthogonal basis, but also links to the QR decomposition. With appropriate definitions for $r_{m,m}$ and $r_{m,\mu}$ as elements of an upper right triangular matrix $\mathbf{R}$, (2) and (3) can be formulated as

$$r_{m,m} \mathbf{q}_m = \mathbf{a}_m - \sum_{\mu=1}^{m-1} \mathbf{q}_\mu r_{m,\mu} \ . \tag{4}$$

Thus the Gram-Schmidt procedure can be written as $\mathbf{QR} = \mathbf{A}$, i.e. the QR decomposition of $\mathbf{A}$ [28]. In the QR decomposition of a full rank matrix $\mathbf{A}$, it is possible to introduce a diagonal matrix $\mathbf{\Phi} = \mathrm{diag}\{\mathrm{e}^{\mathrm{j}\phi_1}, \ldots, \mathrm{e}^{\mathrm{j}\phi_M}\}$, such that

$$\mathbf{A} = \mathbf{QR} = \underbrace{\mathbf{Q}\mathbf{\Phi}}_{\mathbf{Q}'} \underbrace{\mathbf{\Phi}^{\mathrm{H}}\mathbf{R}}_{\mathbf{R}'} \ . \tag{5}$$

This illustrates the non-uniqueness of the QR decomposition: the columns of $\mathbf{Q}$ and the rows of $\mathbf{R}$ can be multiplied by arbitrary phase shifts to yield the equally valid QR factors $\mathbf{Q}'$ and $\mathbf{R}'$. This ambiguity is sometimes removed by selecting $\mathbf{\Phi}$ such that the diagonal components of $\mathbf{R}'$ are positive real.

A modification to (2) can be beneficial to remove some sensitivity of Gram-Schmidt to rounding and error propagation. This modified Gram-Schmidt procedure performs the normalisation step prior to the projection step, whereby $\bar{\mathbf{a}}_{m,\mu}$ is the modified mth column of $\mathbf{A}$ at iteration step $\mu = 1, \ldots, M$. With the initialisations $\bar{\mathbf{a}}_{m,1} = \mathbf{a}_m$, $m = 1, \ldots, M$, the modified Gram-Schmidt procedure calculates at the μth iteration, $\mu = 1, \ldots, M$, the normalised μth column

$$\mathbf{q}_\mu = \frac{\bar{\mathbf{a}}_{\mu,\mu}}{\|\bar{\mathbf{a}}_{\mu,\mu}\|_2} \ , \tag{6}$$

and updates the remaining columns for $i = \mu + 1, \ldots, M$,

$$\bar{\mathbf{a}}_{i,\mu+1} = \bar{\mathbf{a}}_{i,\mu} - \mathbf{q}_\mu (\mathbf{q}_\mu^{\mathrm{H}} \bar{\mathbf{a}}_{i,\mu}) \ . \tag{7}$$

This simple modification significantly reduces the orthogonalisation error compared to the classical procedure in (2) when evaluated numerically with finite machine precision [24], [25]. Note that we have

$$\bar{\mathbf{a}}_{m,m} = \prod_{\mu=1}^{m-1} (\mathbf{I}_M - \mathbf{q}_\mu \mathbf{q}_\mu^{\mathrm{H}}) \mathbf{a}_m = \bar{\mathbf{a}}_m \ , \tag{8}$$

i.e. the end results of (3) and (7) are identical because the projections

$$\prod_{\mu=1}^{m-1} (\mathbf{I}_M - \mathbf{q}_\mu \mathbf{q}_\mu^{\mathrm{H}}) = \mathbf{I}_M - \sum_{\mu=1}^{m-1} \mathbf{q}_\mu \mathbf{q}_\mu^{\mathrm{H}} \tag{9}$$

can be shown to be equivalent by multiplying out the l.h.s. expression.

B. Vectors and Matrices of Analytic Functions

In this paper we are interested in extending the Gram-Schmidt procedure and the QR decomposition to, respectively, vectors and matrices of analytic functions. A vector $\boldsymbol{x}(z)$ is analytic in z if all its elements are functions that possess a region of convergence extending beyond the unit circle. Equivalently, its time domain correspondence $\mathbf{x}[n]$, such that $\boldsymbol{x}(z) = \sum_n \mathbf{x}[n] z^{-n}$ or for short $\mathbf{x}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{x}(z)$, must contain

[1] https://github.com/SebastianJiroSchlecht/PolynomialGramSchmidt

elements that are absolutely convergent [26]. While such functions may have infinite support and generally can be algebraic or even transcendental, via delay and truncation operations these can be approximated arbitrarily closely by polynomials, i.e. by finite impulse response filters if elements of $\boldsymbol{x}(z)$ represent transfer functions [7].

A crucial tool when working with analytic functions is the uniqueness theorem [26]. If two analytic functions are identical within some part of their region of convergence, then they will be identical within their entire region of convergence. It therefore often suffices that when constructing analytic functions, to do so over a finite region of the z plane. In this paper, we will often focus on the unit circle.

The analyticity of functions is generally preserved under standard arithmetic operations except for division. While for two analytic functions $a(z)$ and $b(z)$, $a(z)\pm b(z)$ and $a(z)b(z)$ are also analytic, the analyticity of $a(z)/b(z)$ is restricted to those functions $b(z)$ that do not include any spectral zeros, i.e. zeros that lie on the unit circle. Thus, mathematically, analytic functions form a ring rather than a field [7], and division has to be treated with caution. This will require particular care when extending normalisations such as in (2) or (6) to vectors of analytic functions. For this reason, we often use a primed notation such as $a'(\Omega)$ if it is uncertain whether a function can be analytically extended into the complex plane[2], as opposed to the notation $a(\mathrm{e}^{\mathrm{j}\Omega})$ which implies that a reparameterisation with $z = \mathrm{e}^{\mathrm{j}\Omega}$ yields an analytic function.

Some properties and operations generalise easily from ordinary vectors and matrices [29]. For example, the parahermitian transposition $\{\cdot\}^{\mathrm{P}}$ applied to a matrix $\boldsymbol{A}(z)$ involves a Hermitian transposition but also a time-reversal, such that $\boldsymbol{A}^{\mathrm{P}}(z) = \{\boldsymbol{A}(1/z^*)\}^{\mathrm{H}}$. A parahermitian matrix $\boldsymbol{\Psi}(z)$ extends Hermitian symmetry such that $\boldsymbol{\Psi}(z) = \boldsymbol{\Psi}^{\mathrm{P}}(z)$.

C. Analytic Vector Inner Product and Magnitude Spectrum

Given the comments in Sec. II-B, the inner product $\boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{b}(z)$ between two analytic vectors $\boldsymbol{a}(z) : \mathbb{C} \to \mathbb{C}^M$ and $\boldsymbol{b}(z) : \mathbb{C} \to \mathbb{C}^M$ yields an analytic expression. We speak of $\boldsymbol{a}(z)$ and $\boldsymbol{b}(z)$ being orthogonal, if $\boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{b}(z) = 0$ $\forall z$ (assumed within the region of convergence). For $\boldsymbol{a}(z) = [a_1(z), \ldots, a_M(z)]^{\mathrm{T}}$, the scalar function

$$r(z) = \boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{a}(z) = \sum_{m=1}^{M} a_m(z)a_m^{\mathrm{P}}(z) \tag{10}$$

is parahermitian, such that $r(z) = r^{\mathrm{P}}(z)$. By construction $r(z)$ satisfies the properties of a power spectral density: when evaluated on the unit circle, $r(\mathrm{e}^{\mathrm{j}\Omega}) = \sum_m |a_m(\mathrm{e}^{\mathrm{j}\Omega})|^2$ is non-negative real, i.e. $r(\mathrm{e}^{\mathrm{j}\Omega}) \in \mathbb{R}$ and $r(\mathrm{e}^{\mathrm{j}\Omega}) \geq 0$ $\forall\Omega$. Therefore with $r(z)$, we can define a magnitude spectrum:

Definition 1 (Squared magnitude spectrum of an analytic vector): Let the squared magnitude spectrum of an analytic vector $\boldsymbol{a}(z)$ be $r(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{a}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega}}$.

[2]E.g. for the analytic EVD [2]–[4] and the analytic SVD [4], [5], there does not always exist a solution that is analytic in $z \in \mathbb{C}$.

Note that for specific values of Ω, the magnitude spectrum $\sqrt{r(\mathrm{e}^{\mathrm{j}\Omega})} = \|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ is just the Euclidean vector norm. Further, if $\boldsymbol{a}(z)$ is modified by some arbitrary analytic allpass function $\varphi(z)$, then $\boldsymbol{a}(z)$ and $\tilde{\boldsymbol{a}}(z) = \varphi(z)\boldsymbol{a}(z)$ have identical squared magnitude spectra, since

$$r(z) = \tilde{\boldsymbol{a}}^{\mathrm{P}}(z)\tilde{\boldsymbol{a}}(z) = \boldsymbol{a}^{\mathrm{P}}(z)\varphi^{\mathrm{P}}(z)\varphi(z)\boldsymbol{a}(z) = \boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{a}(z)\,, \tag{11}$$

with an allpass function satisfying $\varphi(z)\varphi^{\mathrm{P}}(z) = 1$. We say a vector $\boldsymbol{q}(z) : \mathbb{C} \to \mathbb{C}^M$ possesses a unit magnitude spectrum if $\boldsymbol{q}^{\mathrm{P}}(z)\boldsymbol{q}(z) = 1$ $\forall z$, which on the unit circle implies $\boldsymbol{q}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{q}(\mathrm{e}^{\mathrm{j}\Omega}) = 1$. A matrix $\boldsymbol{Q}(z) = [\boldsymbol{q}_1(z), \ldots, \boldsymbol{q}_M(z)]$ comprising of mutually orthogonal vectors $\boldsymbol{q}_m(z)$, $m = 1, \ldots, M$ with unit magnitude spectra is called a paraunitary matrix and satisfies $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}_M$.

Example 1: The polyphase analysis filter bank for a Daubechies 4 wavelet decomposition [29], [30] is given by

$$\boldsymbol{Q}(z) = \frac{1}{4\sqrt{2}} \begin{bmatrix} c_0 + c_2 z^{-1} & c_3 + c_1 z^{-1} \\ c_1 + c_3 z^{-1} & -c_2 - c_0 z^{-1} \end{bmatrix}, \tag{12}$$

with $c_0 = (1+\sqrt{3})$, $c_1 = (3+\sqrt{3})$, $c_2 = (3-\sqrt{3})$, and $c_3 = (1-\sqrt{3})$. It is easy to verify that $\boldsymbol{Q}(z) = [\boldsymbol{q}_1(z),\ \boldsymbol{q}_2(z)]$ is a paraunitary matrix whose columns indeed satisfy $\boldsymbol{q}^{\mathrm{P}}{}_m(z)\boldsymbol{q}_\mu(z) = \delta[m-\mu]$, with $\delta[n]$ the Kronecker function [31]. △

Returning to the squared magnitude spectrum, because $r(\mathrm{e}^{\mathrm{j}\Omega})$ is non-negative real and $r(z)$ is analytic thus satisfying the Wiener-Paley condition, the Wiener-Hopf or spectral factorisation [33]

$$r(z) = \beta_+(z)\beta_+^{\mathrm{P}}(z) \tag{13}$$

exists where $\beta_+(z)$ is stable, causal, and minimum phase in the non-strict sense, i.e. it does not possess any zeros for $|z| > 1$. Therefore, $\beta_+(z)$ is analytic, and the Weierstrass factorisation theorem (see e.g. Theorem 7 in [26]) allows to write it as an infinite product of zeros,

$$\beta_+(z) = C \underbrace{\prod_{\ell}(1 - \mathrm{e}^{\mathrm{j}\Theta_\ell} z^{-1})}_{\beta_0(z)} \cdot \underbrace{\prod_{n}(1 - \gamma_n z^{-1})}_{\beta_{\min}(z)}\,, \tag{14}$$

with $C \in \mathbb{C}$, where $\beta_0(z)$ contains the spectral zeros of $\beta_+(z)$ at normalised angular frequencies $\Omega = \Theta_\ell$, and $\beta_{\min}(z)$ is strictly minimum phase with all zeros γ_n inside the unit circle, i.e. $|\gamma_n| < 1$.

III. Analytic Modified Gram-Schmidt and Existence of an Orthonormal Basis

Before extending Gram-Schmidt orthogonalisation to a set of vectors of analytic functions, we discuss issues that arise from the normalisation of such vectors. This impacts on the existence of the analytic Gram-Schmidt procedure and the analytic QR decomposition.

A. Normalisation of a Vector of Functions

For an analytic vector $\boldsymbol{a}(z)$, we first generalise (1) for $|z| = 1$, before extending the result beyond the unit circle. On the unit circle, we define

$$\boldsymbol{q}'(\Omega) = \frac{\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})}{\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2} = \frac{\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})}{\sqrt{\boldsymbol{a}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})}} \,. \tag{15}$$

Since it is initially unclear if the division by the magnitude spectrum — i.e. a normalisation to unit length at every frequency Ω— in (15) results in an analytic function in $\Omega \in \mathbb{R}$ and whether it can be extended to $z \in \mathbb{C}$ with a non-vanishing region of convergence, we denote the result for now as $\boldsymbol{q}'(\Omega)$. Two challenges arise for the evaluation of (15) and the analyticity of its result: how to address (i) the square root in $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2$, and (ii) a division by zero in the case of $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2 = 0$. We first exclude case (ii) and address case (i) by the following lemma.

Lemma 1 (Normalisation in the absence of spectral zeros): If an analytic vector $\boldsymbol{a}(z)$ has a squared magnitude spectrum that is free of spectral zeros, i.e. $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2^2 > 0$, then its normalisation to a unit magnitude spectrum for $|z| = 1$ can be selected to be analytic in z.

Proof. Consider the term $\sqrt{\boldsymbol{a}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})}$ in (15). Using (11), (13), and (14), and with $r(z)$ lacking spectral zeros, we have

$$r(z) = \boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{a}(z) = C\beta_{\min}(z)\beta_{\min}^{\mathrm{P}}(z) \,, \tag{16}$$

where, compared to (14), $\beta_0(z) = 1$. Therefore, in the z plane for the divisor in (15), we obtain

$$\frac{1}{\sqrt{r(z)}} = \frac{1}{\sqrt{C}} \prod_n \frac{1}{\sqrt{1-\gamma_n z^{-1}}} \frac{1}{\sqrt{1-\gamma_n^* z}} \tag{17}$$

and investigate its factors in turn.

Firstly, in (17) any term $1/\sqrt{1-\gamma_n z^{-1}}$ can expand into a Maclaurin series [34],

$$\frac{1}{\sqrt{1-\gamma_n z^{-1}}} = \sum_{\nu=0}^{\infty} \underbrace{\frac{1}{\nu!} \left\{ \prod_{i=0}^{\nu-1} (\frac{1}{2}+i) \right\}}_{\chi_\nu} \gamma_n^\nu z^{-\nu} \,, \tag{18}$$

with monotonically decreasing Maclaurin coefficients $|\chi_\nu| < |\chi_{\nu-1}| \leq 1 \;\forall\; \nu > 1$. Since γ_n^ν decays exponentially, (18) is an absolutely convergent power series in z^{-1} [2]. Secondly in (17) for the remaining conjugate factor $(1-\gamma_n^* z)$ in (16), the inversion and square root similarly leads to a Maclaurin expansion that represents an absolutely convergent power series in z, i.e. it is anti-causal. Therefore, the square root operations in (15) result in an analytic Laurent series. Note that even if all elements of $\boldsymbol{a}(z)$ are polynomial, the expression $1/\sqrt{\boldsymbol{a}^{\mathrm{P}}(z)\boldsymbol{a}(z)}$ may likely no longer be of finite order; in fact, it can become an algebraic or even transcendental function [2], [18]. Nonetheless, since these series are absolutely convergent, the inverse square root term is analytic and, as $\boldsymbol{a}(z)$ is analytic, the normalised vector $\boldsymbol{q}(z)$ is analytic. By implication on the unit circle we can write $\boldsymbol{q}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{q}'(\Omega)$. ■

Note that there is an allpass ambiguity to the normalisation of a polynomial vector. If $\boldsymbol{q}(z)$ is an analytic normalised version of $\boldsymbol{a}(z)$, and $\phi(z)$ is an arbitrary allpass function then $\phi(z)\boldsymbol{q}(z)$ is also has a unit magnitude spectrum due to the fact that $|\phi(\mathrm{e}^{\mathrm{j}\Omega})| = 1$ on the unit circle[3]. All that remains is to address case (ii) of spectral zeros of $r(z)$:

Theorem 1 (Analyticity of a normalised vector): The result of normalising a vector of analytic functions as in (15) can be selected to be analytic in z.

Proof. The case of no spectral zeros is covered by Lemma 1. We now need to investigate the case of the magnitude spectrum of a vector $\boldsymbol{a}(z)$ becoming zero for some values of Ω. The normalised vector is the ratio of two terms. When the denominator is zero so is the numerator. So heuristically, according to de l'Hôpital's rule [34], the result may be well defined. Recall the squared magnitude spectrum of $\boldsymbol{a}(z)$ is $r(z)$ as in (10). It is a sum of M non-negative power spectral density terms. Assume that $r(\mathrm{e}^{\mathrm{j}\Omega}) = 0$ at N isolated frequency points $\Omega = \Theta_\ell$, $\ell = 1, \ldots, N$. This requires a spectral zero term $(1 - \mathrm{e}^{\mathrm{j}\Theta_\ell} z^{-1})$ in (14) to be common to all elements $a_m(z)$, $m = 1, \ldots, M$. We define

$$\boldsymbol{a}(z) = \beta_0(z)\tilde{\boldsymbol{a}}(z) \,, \tag{19}$$

where $\beta_0(z)$ contains all N common spectral zeros at frequencies Θ_ℓ of $\boldsymbol{a}(z)$ according to (14). Then consider[4]

$$\begin{aligned} \boldsymbol{q}'(\Omega) &= \frac{\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})}{\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2} = \frac{\beta_0(\mathrm{e}^{\mathrm{j}\Omega})}{\sqrt{\beta_0^*(\mathrm{e}^{\mathrm{j}\Omega})\beta_0(\mathrm{e}^{\mathrm{j}\Omega})}} \frac{\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})}{\|\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})\|_2} \\ &= \sqrt{\frac{\beta_0(\mathrm{e}^{\mathrm{j}\Omega})}{\beta_0^*(\mathrm{e}^{\mathrm{j}\Omega})}} \frac{\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})}{\|\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})\|_2} = \beta'(\Omega)\tilde{\boldsymbol{q}}(\mathrm{e}^{\mathrm{j}\Omega}) \,. \end{aligned} \tag{20}$$

Since $\tilde{\boldsymbol{a}}(z)$ has no common spectral zeros and therefore $\|\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})\|_2 > 0 \;\forall\Omega$, according to Lemma 1 the term

$$\tilde{\boldsymbol{q}}(z) = \frac{\tilde{\boldsymbol{a}}(z)}{\sqrt{\tilde{\boldsymbol{a}}^{\mathrm{P}}(z)\tilde{\boldsymbol{a}}(z)}} \tag{21}$$

is analytic. For the N common spectral zero terms at normalised angular frequencies Θ_ℓ, $\ell = 1, \ldots, N$, of $\boldsymbol{a}(z)$ gathered in $\beta_0(z)$, we obtain

$$\beta'(\Omega) = \sqrt{\frac{\beta_0(\mathrm{e}^{\mathrm{j}\Omega})}{\beta_0^*(\mathrm{e}^{\mathrm{j}\Omega})}} = \prod_{\ell=1}^{N} \sqrt{\frac{1-\mathrm{e}^{\mathrm{j}(\Theta_\ell-\Omega)}}{1-\mathrm{e}^{-\mathrm{j}(\Theta_\ell-\Omega)}}}. \tag{22}$$

For one such square route term, with $\theta \in \mathbb{R}$ note that

$$\frac{1-\mathrm{e}^{\mathrm{j}\theta}}{1-\mathrm{e}^{-\mathrm{j}\theta}} = \frac{\mathrm{e}^{\mathrm{j}\theta}(\mathrm{e}^{-\mathrm{j}\theta}-1)}{1-\mathrm{e}^{-\mathrm{j}\theta}} = -\mathrm{e}^{\mathrm{j}\theta} \tag{23}$$

represents a key step in our argument, and we obtain for (22)

$$\beta'(\Omega) = \prod_{\ell=1}^{N} \sqrt{-\mathrm{e}^{\mathrm{j}(\Theta_\ell-\Omega)}} = \pm \left\{ \prod_{\ell=1}^{N} \mathrm{e}^{\mathrm{j}(\Theta_\ell+\pi)/2} \right\} \mathrm{e}^{-\mathrm{j}\frac{N}{2}\Omega} \,, \tag{24}$$

[3]Note that this allpass ambiguity is analogous to that of analytic eigenvectors [2], [4] or analytic left- or right-singular vectors [5], which have to satisfy unit norm on the unit circle and are non-unique w.r.t. arbitrary allpass functions

[4]Recall that since it is initially unclear if the division by the norm in (15) results in an analytic function, we denote the result for now as $\boldsymbol{q}'(\Omega)$.

where $\pm$ addresses the ambiguity of the square root of a complex-valued number. Hence $\beta'(\Omega)$ is an allpass function, but it is not necessarily analytic due to a potentially fractionally spaced delay term $z^{-\frac{N}{2}}$ after reparameterisation with $z = \mathrm{e}^{\mathrm{j}\Omega}$. Note $\boldsymbol{q}'(\Omega) = \beta'(\Omega)\tilde{\boldsymbol{q}}(\mathrm{e}^{\mathrm{j}\Omega})$, and thus $\boldsymbol{q}'(\Omega)$ and $\tilde{\boldsymbol{q}}(\mathrm{e}^{\mathrm{j}\Omega})$ are both valid normalised versions of $\boldsymbol{a}(\Omega)$ but we choose the latter as it can always be analytically extended. Hence the above $\tilde{\boldsymbol{q}}(z)$ is an analytic normalisation of $\boldsymbol{a}(z)$. ■

Two observations arise from Theorem 1: (i) the analysis can be generalised to show that any common factor — not just spectral zeros— across the functions $a_m(z)$, $m = 1, \ldots, M$, is irrelevant for the normalisation of $\boldsymbol{a}(z)$, and (ii) in line with (11), $\tilde{\boldsymbol{q}}(z)$ can be modified by an analytic allpass filter and remain a valid analytic normalisation of $\boldsymbol{a}(z)$.

Example 2: The vector

$$\boldsymbol{a}(z) = (1 + z^{-2}) \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \tag{25}$$

possesses $L = 2$ common spectral zeros at $z = \pm\mathrm{j}$, i.e. for $\Theta_{1,2} = \pm\frac{\pi}{2}$. Removing $\beta_0(z) = (1 + z^{-2})$ and normalising according to (20) leads to

$$\tilde{\boldsymbol{q}}(z) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \tag{26}$$

as the analytic normalisation of (25). △

B. Impact of Undetected Common Zeros

In theory, Sec. III-A provides a clear route to the normalisation of an analytic vector, provided that spectral zeros common to all elements $a_m(z)$, $m = 1, \ldots, M$, of $\boldsymbol{a}(z)$ can be determined and removed. In practice, a search for common zeros (including spectral ones), requires the determination of the greatest common divisor (GCD) across all functions $a_m(z)$, $m = 1, \ldots, M$. This however is computationally costly at best, and numerically ill-conditioned at worst. Despite recent efforts and refinements [35]–[37], the calculation of the GCD is still mostly reliant on Euclid's algorithm from his 7th book of elements, published in 300BC [38]. This algorithm becomes increasingly sensitive and computationally costly for moderate to large polynomials, hence using it to find all common zeros is not easy. Hence we briefly explore the effect if none or not all common spectral zeros of $\boldsymbol{a}(z)$ are eliminated. In fact, we will later not even try to find them and instead mitigate the resulting effect.

Assume that L common spectral zeros at $\Omega = \Theta_\ell$, $\ell = 1, \ldots, L \leq N$, are missed by a GCD algorithm. Let

$$\boldsymbol{q}'(\Omega) = \pm \left\{ \prod_{\ell=1}^{L} \mathrm{e}^{\mathrm{j}(\Theta_\ell + \pi)/2} \right\} \mathrm{e}^{-\mathrm{j}\frac{L}{2}\Omega} \tilde{\boldsymbol{q}}(\mathrm{e}^{\mathrm{j}\Omega}) \,, \tag{27}$$

where $\boldsymbol{q}'(\Omega)$ differs from $\tilde{\boldsymbol{q}}(\mathrm{e}^{\mathrm{j}\Omega})$ in (20) by the inclusion of the missed factor $\beta'(\Omega)$. For L an even number, (27) is 2π-periodic. Then, using a reparameterisation $z = \mathrm{e}^{\mathrm{j}\Omega}$, an analytic solution arises with

$$\boldsymbol{q}(z) = \pm z^{-L/2} \prod_{\ell=1}^{L} \mathrm{e}^{\mathrm{j}(\Theta_\ell + \pi)/2} \tilde{\boldsymbol{q}}(z) \,, \tag{28}$$

where $\tilde{\boldsymbol{q}}(z)$ is given by (21). The leading term $\pm z^{-L/2} \prod_{\ell=1}^{L} \mathrm{e}^{\mathrm{j}(\Theta_\ell + \pi)/2}$ represents a delay and phase change, i.e. an allpass function and can be ignored according to (11).

Example 3: If for the normalisation of the vector $\boldsymbol{a}(z)$ in Example 2 the GCD is ignored, (28) yields $\pm z^{-1}(\mathrm{e}^{\mathrm{j}(\frac{\pi}{2} - \pi)/2})(\mathrm{e}^{\mathrm{j}(-\frac{\pi}{2} - \pi)/2})\tilde{\boldsymbol{q}}(z) = \mp z^{-1}\tilde{\boldsymbol{q}}(z)$ as the solution. This differs from (26) via the allpass $\phi(z) = \mp z^{-1}$. △

Note that for odd values of L, the term $\mathrm{e}^{-\mathrm{j}\frac{L}{2}\Omega}$ in (27) is 4π-periodic, and while analytic in $\Omega \in \mathbb{R}$, the expansion into the complex plane via the substitution $z = \mathrm{e}^{\mathrm{j}\Omega}$ is not analytic in $z \in \mathbb{C}$ and represents a fractional delay [39]. This is exemplified below.

Example 4: For $\boldsymbol{a}(z) = (1 + z^{-1})[1;\ z^{-1}]$, removing the common factor term $(1 + z^{-1})$ with a single ($L = 1$) spectral zero at $\Omega = \pi$ yields the same normalised vector $\tilde{\boldsymbol{q}}(z)$ as for Example 2 in (26). If the common factor is not removed, the normalisation yields $\pm z^{-\frac{1}{2}}\ \tilde{\boldsymbol{q}}(z)$, which means that the solution contains a fractional delay $z^{-\frac{1}{2}}$. Its time domain equivalent, a sinc function [39], decays with $1/n$. Since this is slower than the exponential convergence that is required of an analytic function, such a solution is not analytic. △

Hence, the case of L being odd leads to a non-analytic solution with a fractional delay, which we will need to address when seeking a normalisation algorithm — see Sec. IV-C.

C. Extended Modified Gram-Schmidt Orthogonalisation

We now extend the modified Gram-Schmidt procedure to a set of linearly independent analytic vectors $\boldsymbol{a}_m(z) : \mathbb{C} \to \mathbb{C}^M$, $m = 1, \ldots, M$, that form the columns of an analytic matrix $\boldsymbol{A}(z)$. Analogous to Sec. II-A, $\bar{\boldsymbol{a}}_{m,\mu}(z)$ denotes the modified mth column of $\boldsymbol{A}(z)$ at iteration μ, with $m, \mu = 1, \ldots, M$. Based on the initialisation $\bar{\boldsymbol{a}}_{m,1}(z) = \boldsymbol{a}_m(z)$, $m = 1, \ldots, M$, at the μth iteration, $\mu = 1, \ldots, M$, we calculate

$$\boldsymbol{q}_\mu(z) = \frac{\bar{\boldsymbol{a}}_{\mu,\mu}(z)}{\sqrt{\bar{\boldsymbol{a}}^{\mathrm{P}}{}_{\mu,\mu}(z)\bar{\boldsymbol{a}}_{\mu,\mu}(z)}} \,, \tag{29}$$

followed by

$$\bar{\boldsymbol{a}}_{i,\mu+1}(z) = \bar{\boldsymbol{a}}_{i,\mu}(z) - \boldsymbol{q}_\mu(z) \left(\boldsymbol{q}^{\mathrm{P}}{}_\mu(z)\bar{\boldsymbol{a}}_{i,\mu}(z) \right) \tag{30}$$

for $i = \mu + 1, \ldots, M$. This is a simple extension of standard linear algebra to the domain of analytic functions [15]; this is permitted, since multiplications and additions of analytic functions in (30) yield an analytic result, and the analyticity of (29) is covered by Theorem 1.

As for the ordinary Gram-Schmidt procedure in Sec. II-A, (29) and (30) can be repeated for $1 \leq \mu \leq M$ unless we find that, at iteration μ say, $\boldsymbol{a}_\mu(z)$ and therefore due to (8) $\bar{\boldsymbol{a}}_{\mu,\mu}(z)$ is zero. So to complete the orthogonalisation of $\boldsymbol{A}(z)$, we need to define the equivalent of the full rank property of an ordinary matrix $\mathbf{A}$ in Sec. II-A. To extend the notion of rank to matrices of functions, we define:

Definition 2 (Full spatial rank): A square matrix $\boldsymbol{A}(z)$ has full spatial rank if there exists at least one Ω_0 such that its determinant satisfies $\det\{\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})\} \neq 0$.

As a sum of products of analytic functions, the determinant is an analytic function. Note that as explained in Theorem 1, columns $\boldsymbol{a}(z)$ can have isolated common spectral zeros, which are therefore also spectral zeros of the determinant of $\boldsymbol{A}(z)$. According to the uniqueness theorem of analytic functions [26], an analytic function can only be zero over some interval if it is zero everywhere within its region of convergence. Thus the determinant is not the zero function if it is non-zero in at least one point on the unit circle.

As an alternative to Definition 2, full spatial rank can be defined via the analytic singular values of $\boldsymbol{A}(z)$ (see Sec. III-B of [41]), which also may only possess isolated spectral zeros. Thus, an $M \times M$ matrix $\boldsymbol{A}(z)$ of full spatial rank will possess M linearly independent columns in the sense that no column can be formed as a sum of the remaining columns weighted by analytic functions. But different from the full rank definition in the ordinary matrix case, full spatial rank does not imply that $\boldsymbol{A}(z)$ is invertible. This would require that $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})$ is invertible everywhere on the unit circle, but this is not possible at frequencies Ω where $\det\{\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})\} = 0$.

With the extended modified Gram-Schmidt procedure above and based on Definition 2, we can formally establish the existence of an analytic orthonormal basis:

Lemma 2 (Existence of an analytic orthonormal basis): Based on the columns $\boldsymbol{a}_m(z)$, $m = 1, \ldots, M$, of an analytic matrix $\boldsymbol{A}(z)$ of full spatial rank, the procedure defined by (29) and (30) can find analytic orthonormal basis vectors $\boldsymbol{q}_m(z)$ that form a paraunitary matrix $\boldsymbol{Q}(z) = [\boldsymbol{q}_1(z), \ldots, \boldsymbol{q}_M(z)]$.

Proof. For $\boldsymbol{A}(z)$ with full spatial rank, M iterations of the procedure in (29) and (30) are possible, and analytic and orthogonal basis vectors $\boldsymbol{q}_m(z)$ can be obtained from the columns $\boldsymbol{a}_m(z)$, $m = 1, \ldots, M$ of $\boldsymbol{A}(z)$. Because there are M mutually orthogonal vectors $\boldsymbol{q}_m(z)$ with unit magnitude spectrum, we have $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}_M$, i.e. $\boldsymbol{Q}(z)$ is paraunitary. ■

D. Existence of an Analytic QR Decomposition

In the ordinary matrix case, the Gram-Schmidt procedure applied to a matrix $\boldsymbol{A}$ directly links to the QR decomposition. Now with (29) and (30), and based on Lemma 2, we can state:

Theorem 2 (Existence of an analytic QR decomposition): An analytic matrix $\boldsymbol{A}(z)$ of full spatial rank admits an analytic QR decomposition,

$$\boldsymbol{A}(z) = \boldsymbol{Q}(z)\boldsymbol{R}(z) \;, \tag{31}$$

with an analytic paraunitary $\boldsymbol{Q}(z)$ and an upper-right triangular analytic $\boldsymbol{R}(z)$.

Proof. If $\boldsymbol{A}(z) : \mathbb{C} \to \mathbb{C}^{M\times M}$ has full spatial rank, then the extended modified Gram-Schmidt orthogonalisation in (29) and (30) can complete M iterations. Lemma 2 guarantees that an analytic $\boldsymbol{Q}(z)$ exists as a result of this Gram-Schmidt orthogonalisation. Analogous to the ordinary matrix case (see e.g. Algorithm 5.2.5 in [28]), the normalisation and projection steps in (29) and (30) directly imply (31) with an upper right triangular matrix $\boldsymbol{R}(z)$. Because (29) can always yield an analytic result and the inner products in (30) are guaranteed to be analytic, $\boldsymbol{R}(z)$ can always be analytic, and may we written as

$$\boldsymbol{R}(z) = \begin{bmatrix} \boldsymbol{q}_1^{\mathrm{P}}(z)\boldsymbol{a}_1(z) & \boldsymbol{q}_1^{\mathrm{P}}(z)\boldsymbol{a}_2(z) & \ldots & \boldsymbol{q}_1^{\mathrm{P}}(z)\boldsymbol{a}_M(z) \\ 0 & \boldsymbol{q}_2^{\mathrm{P}}(z)\boldsymbol{a}_2(z) & \ldots & \boldsymbol{q}_2^{\mathrm{P}}(z)\boldsymbol{a}_M(z) \\ \vdots & \ddots & & \vdots \\ 0 & \ldots & 0 & \boldsymbol{q}_M^{\mathrm{P}}(z)\boldsymbol{a}_M(z) \end{bmatrix}, \tag{32}$$

with $\boldsymbol{q}_m(z)$ and $\boldsymbol{a}_m(z)$, $m = 1, \ldots, M$ the columns of the analytic matrices $\boldsymbol{A}(z)$ and $\boldsymbol{Q}(z)$. ■

Theorem 2 provides the foundation for algorithms that have been introduced to approximate (31), such as in [17]–[19], [23]. As a generalisation of the phase ambiguity of the standard QR factors in (5), with (31) and the allpass ambiguity of $\boldsymbol{Q}(z)$—see comment (ii) after Theorem 1 in Sec. III-A—, we can state that with a diagonal matrix $\boldsymbol{\Phi}(z) = \mathrm{diag}\{\phi_1(z), \ldots .\phi_M(z)\}$ of analytic allpass filters $\phi_m(z)$, $m = 1, \ldots, M$, the modified matrices $\boldsymbol{Q}(z)\boldsymbol{\Phi}(z)$ and $\boldsymbol{\Phi}^{\mathrm{P}}(z)\boldsymbol{R}(z)$ are also valid factors of the analytic QR decomposition.

Ambiguity can be suppressed in the ordinary QR decomposition by mandating the diagonal components of $\mathbf{R}$ to be positive real. The extension to the case of analytic functions would be for the diagonal components $r_{m,m}(\mathrm{e}^{\mathrm{j}\Omega})$ of $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ to be non-negative real. But this is not always possible: as a counterexample for a degenerate $M \times 1$ matrix, i.e. a vector, $\boldsymbol{a}(z)$, the analytic QR decomposition $\boldsymbol{a}(z) = \boldsymbol{q}(z)r(z)$ where $r(z)$ is a simple polynomial. The analytic singular value decomposition $\boldsymbol{a}(z) = \boldsymbol{u}(z)\left(\sigma(z)v^{\mathrm{P}}(z)\right)$ is then equivalent with $\boldsymbol{u}(z) = \boldsymbol{q}(z)$ and $\sigma(z)v^{\mathrm{P}}(z) = r(z)$. However, for the analytic SVD to exist, $\sigma(\mathrm{e}^{\mathrm{j}\Omega})$ must be permitted to have sign changes if $\boldsymbol{a}(z)$ has zero crossings [5], i.e. may not be non-negative, and cannot be restricted to be real if the number of those zero crossings is odd [32]. Alternative ways[5] to remove degrees of freedom include allpass modifications such that either (i) the diagonal components are positive real for $z = 1$, i.e. for DC, or (ii) to force diagonal components to be minimum phase, with the leading coefficient positive real.

IV. Analytic Vector Normalisation Algorithm

As a component of an expanded modified Gram-Schmidt procedure, in the following we first concentrate on an algorithm that is capable of normalising a vector $\boldsymbol{a}(z)$ of analytic functions, yielding an analytic result $\boldsymbol{q}(z)$ whose existence is guaranteed by Theorem 1.

A. Overall Approach

Sec. III-B has highlighted that the detection of common spectral zeros within the elements of a vector $\boldsymbol{a}(z)$ is difficult in practice. We therefore do not attempt to find a GCD. This has two consequences: (i) there may be an odd number of common spectral zeros, in which case we have to compensate for a fractional delay in order to attain an analytic result; (ii) the normalisation result may have a larger support than

[5]As kindly suggested by one of the anonymous reviewers.

necessary, because common factors have not been removed from $\boldsymbol{a}(z)$ and because of the presence of an arbitrary allpass filter. On the latter, note that even if $\boldsymbol{q}(z)$ has finite support, with an arbitrary allpass $\varphi(z)$ the support of $\varphi(z)\boldsymbol{q}(z)$ may be significantly larger.

The proposed normalisation algorithm will operate in the DFT domain. We are evaluating $\boldsymbol{a}(z)$ on the unit circle $z = \mathrm{e}^{\mathrm{j}\Omega}$ and in K frequency bins with frequencies Ω_k, $k = 0, \ldots, (K-1)$, such that

$$\mathbf{a}_k = \boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega_k}) \; . \tag{33}$$

Recalling that $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ can become zero, in order to avoid the case $\|\mathbf{a}_k\|_2 = 0$ and hence normalisation issues, the bin frequencies Ω_k are not necessarily selected to be uniform. This is addressed in Sec. IV-B. The problem of detecting and counting common spectral zeros between bins, addressing potential sign changes resulting from zeros crossings, and compensating for the potentially resulting fractional delay will be outlined in Sec. IV-C. Sec. IV-D identifies a sufficient DFT size, which relies on a time domain aliasing criterion and on finding an allpass function to shorten the support of $\mathbf{q}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{q}(z)$. Sec. IV-E summarises these algorithmic steps.

B. Avoiding DFT Bins with Zero Norm

Since we will be performing a bin-wise normalisation of $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$ followed by a reconstruction to obtain $\mathbf{q}[n]$ via an inverse Fourier transform, we need to avoid bins where $\|\mathbf{a}_k\|_2 = 0$. In related DFT domain methods where particular bins can be problematic, interpolation across [40] or omission [41] of such bins are an option. We here follow a simple scheme from [42] of shifting problematic bins: if for a particular Ω'_k, on a uniform grid $\Omega'_k = 2\pi k/K$, we find that $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega'_k})\|_2 \leq \epsilon$ with a small constant ϵ, then — recalling that zeros can only occur at isolated frequencies — we shift this bin frequency by an amount $\delta_k < 2\pi/K$ such that for $\Omega_k = \Omega'_k + \delta_k$ we have $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega_k})\|_2 > \epsilon$. Thus, the bin frequencies Ω_k form a non-uniform grid, and the sample points $\{\mathbf{a}_k\}$ are obtained from a non-uniform DFT [43], [44] of $\mathbf{a}[n] \circ\!\!-\!\!\bullet\, \mathbf{a}(z)$.

In order to reconstruct $\mathbf{a}[n]$ from $\{\mathbf{a}_k\}$, a non-uniform inverse DFT [43], [44] can be invoked. This reconstruction will be accurate for a general $\mathbf{a}[n]$ as long as the K bin frequencies Ω_k, $k = 0, \ldots, K$ are sufficiently distinct and thus the non-uniform inverse DFT matrix will be well conditioned. For such a non-uniform inverse DFT, efficient schemes akin to a fast Fourier transform (FFT) exist [43], [44]; if only very few bins are shifted off a uniform grid, then a numerically efficient scheme is to perform a simple interpolation of the missing uniform grid points followed by a standard inverse FFT [42].

C. Detecting and Compensating for Zero Crossings

We now need to detect where and how often the norm satisfies $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|^2 = 0$, i.e. where we have $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega}) = \underline{0}$. Since according to Sec. IV-B spectral zeros do not coincide with bin frequencies, these can only occur between bins. We therefore investigate what happens in the vicinity $\Omega \in [\Theta_\ell - \epsilon;\ \Theta_\ell + \epsilon]$ of a spectral zero at frequency Θ_l for some small ϵ, $0 < \epsilon \ll 1$. Assume that such a spectral zero has multiplicity P, i.e. that $\Theta_\ell = \Theta_{\ell+1} = \ldots = \Theta_{\ell+P-1}$. Let $\boldsymbol{a}(z) = (1 - \mathrm{e}^{\mathrm{j}\Theta_\ell} z^{-1})^P \tilde{\boldsymbol{a}}(z)$, where $\tilde{\boldsymbol{a}}(z)$ is without a spectral zero at $\Omega = \Theta_\ell$. With a Taylor series expansion [34] of $\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega})$ around this point, we have $\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Omega}) = \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\Omega - \Theta_\ell)$. In the vicinity of $\Omega = \Theta_\ell$, this expansion yields

$$\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}(\Theta_\ell \pm \epsilon)}) = \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\epsilon) \; ,$$

i.e. for a small $\epsilon \to 0$, the terms captured by $\mathcal{O}(\epsilon) \to \underline{0}$, while $\tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}(\Theta_\ell)}) \neq \underline{0}$ by definition. With the series expansion $\mathrm{e}^x = \sum_{i=0}^{\infty} \frac{1}{i!}(x)^i$ [34], for $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$ at $\Omega = \Theta_\ell - \epsilon$ we then have

$$\begin{aligned} \boldsymbol{a}(\mathrm{e}^{\mathrm{j}(\Theta_\ell - \epsilon)}) &= (1 - \mathrm{e}^{\mathrm{j}\epsilon})^P \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}(\Theta_\ell - \epsilon)}) \\ &= [\mathrm{j}\epsilon + \mathcal{O}(\epsilon^2)]^P \left\{ \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\epsilon) \right\} \\ &= (\mathrm{j}\epsilon)^P \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\epsilon^{P+1}) \; . \end{aligned} \tag{34}$$

Likewise, at $\Omega = \Theta_\ell + \epsilon$, we obtain

$$\begin{aligned} \boldsymbol{a}(\mathrm{e}^{\mathrm{j}(\Theta_\ell + \epsilon)}) &= (-\mathrm{j}\epsilon)^P \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\epsilon^{P+1}) \\ &= (-1)^P (\mathrm{j}\epsilon)^P \tilde{\boldsymbol{a}}(\mathrm{e}^{\mathrm{j}\Theta_\ell}) + \mathcal{O}(\epsilon^{P+1}) \; , \end{aligned} \tag{35}$$

Comparing (34) and (35) for very small ϵ shows that $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$ changes sign at a spectral zero of odd multiplicity P. Below, we refer to the case P being odd as a zero crossing, whereas for an even value of P the elements of $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$ only tangentially touch the origin. While an even multiplicity P is of little consequence overall, for zero crossings we must (i) correct a sign change for the normalisation, and (ii) count towards whether overall we have an odd number of zeros on a 2π interval of Ω.

In order to detect a zero crossing, assume that K is sufficiently large such that the coherence bandwidth [45] of $\mathbf{a}(\mathrm{e}^{\mathrm{j}\Omega})$ significantly exceeds the regular bin distance $2\pi/K$. Under such conditions, the rate of change of $\mathbf{a}(\mathrm{e}^{\mathrm{j}\Omega})$ from bin to bin is small. Thus, if no zero crossing is enclosed between bins Ω_{k-1} and Ω_k, we find that $\mathbf{a}_k \approx \mathbf{a}_{k-1}$. However, if a zero crossing of $\mathbf{a}(\mathrm{e}^{\mathrm{j}\Omega})$ occurs between these two bins[6], then $\mathbf{a}_k \approx -\mathbf{a}_{k-1}$. The bin-wise normalised vector $\mathbf{u}_k = \mathbf{a}_k / \|\mathbf{a}_k\|_2$ shows the same behaviour with $\mathbf{u}_k \approx \pm \mathbf{u}_{k-1}$, where a positive sign indicates that no zero crossing is enclosed, and a sign change means that a zero crossing has occured. To obtain an indicator for zero crossings, note therefore that $\mathbf{u}_k^{\mathrm{H}} \mathbf{u}_{k-1} \approx \pm 1$. In order to force a decision from the generally complex valued term $\mathbf{u}_k^{\mathrm{H}} \mathbf{u}_{k-1}$, we eliminate the potentially small imaginary part using the real operator $\Re\{\cdot\}$ and employ the modified signum function $\mathrm{sign}(x) = 1$ for $x \geq 0$ and $\mathrm{sign}(x) = -1$ otherwise, such that $\mathrm{sign}(\Re\{\mathbf{u}_k^{\mathrm{H}} \mathbf{u}_{k-1}\}) = \pm 1$. Setting $d_0 = 1$, we define for $k = 1, \ldots, (K-1)$

$$d_k = \mathrm{sign}(\Re\{\mathbf{u}_k^{\mathrm{H}} \mathbf{u}_{k-1}\}) d_{k-1} \; . \tag{36}$$

[6]The case of two adjacent bins enclosing more than one zero crossing is addressed in Sec. IV-D; it will be avoided if the DFT size K is sufficiently large.

Algorithm 1: Bin-wise vector normalisation

Input: $\{\mathbf{a}_k\}$, $\{\Omega_k\}$, K as non-uniform DFT parameters of $\mathbf{a}[n]$;
$\mathbf{u}_0 = \mathbf{a}_0/\|\mathbf{a}_0\|_2$, $d_0 = 1$, $\mathbf{a}_K = \mathbf{a}_0$;
for $k = 1 \to K$ **do**
  $\mathbf{u}_k = \mathbf{a}_k/\|\mathbf{a}_k\|_2$;
  if $\Re\{\mathbf{u}_{k-1}^{\mathrm{H}}\mathbf{u}_k\} < 0$ **then** $d_k = -d_{k-1}$;
  else $d_k = d_{k-1}$;
end
if $d_K \neq d_0$ **then** $\Theta = \frac{1}{2}$ **else** $\Theta = 0$;
for $k = 0 \to (K-1)$ **do**
  $\mathbf{q}_k = \mathrm{e}^{-\mathrm{j}\Theta\Omega_k} d_k \mathbf{u}_k$;
end
Output: $\{\mathbf{q}_k\}$

Thus, $d_k = \pm 1$: starting with a positive sign for the first bin $k = 0$, a zero crossing between subsequent bins $k-1$ and k results in a sign change of d_k.

We can use the indicator d_k for two purposes. Firstly, at a zero crossing of $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$, a sign change occurs for all its elements. Hence to obtain a smooth normalisation, we require to change the sign of $\mathbf{u}_k$ at every such transition. Secondly, we want to detect if there is an odd number of zero crossings over a 2π interval. This can be accomplished by defining $d_K = \mathrm{sign}(\Re\{\mathbf{u}_K^{\mathrm{H}}\mathbf{u}_{K-1}\}) = \mathrm{sign}(\Re\{\mathbf{u}_0^{\mathrm{H}}\mathbf{u}_{K-1}\})$. Then, if $d_K = d_0$, the number of zero crossings is even, otherwise for $d_K = -d_0$ it is odd and a half-sample fractional delay has occured and must be compensated. In the DFT domain, this compensation can be implemented using a phase shift equivalent to a half-sample delay via the Kronecker function $\delta[\cdot]$. Then

$$\mathbf{q}_k = \mathrm{e}^{-\mathrm{j}\frac{1}{2}\Omega_k \delta[d_0 + d_K]} d_k \mathbf{u}_k \tag{37}$$

represents sample points of 2π-periodic functions whose non-uniform IDFT will yield absolutely convergent time domain functions, and thus can be analytically extended via reparameterisation with $z = \mathrm{e}^{\mathrm{j}\Omega}$. This procedure is summarised in Algorithm 1. A simple case with a zero crossing, and hence a sign change and the need for a fractional delay compensation, is given in the following example.

Example 5: Using the vector $\boldsymbol{a}(z)$ from Example 2, its norm is zero for $\Omega = \pi$, and we find that for any even-length DFT the appropriate bin has to be shifted according to Sec. IV-B. Using $K = 64$, Fig. 1 shows the norm of $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ and the shifted 33rd bin. With this frequency-shift implemented, the bin-normalised components of $\mathbf{u}_k = [u_{1,k}; u_{2,k}]^{\mathrm{T}}$ for a DFT length of $K = 64$ are all well-defined and displayed in Fig. 2(a); the zero-crossing of the norm at $\Omega = \pi$ leads to discontinuity in the vector elements at that frequency, which is eliminated by the sign correction via d_k in Fig. 2(b). This elimination comes at the cost of $d_k\mathbf{u}_k$ no longer being 2π-periodic. Note the opposing signs of the vector components for $\Omega_0 = 0$ and $\Omega_K = 2\pi$. Applying a phase shift equivalent to a half sample delay as in (37) yields a 2π-periodic functions in Fig. 2(c), which belong to an analytic solution $z^{-1}\boldsymbol{q}(z)$,

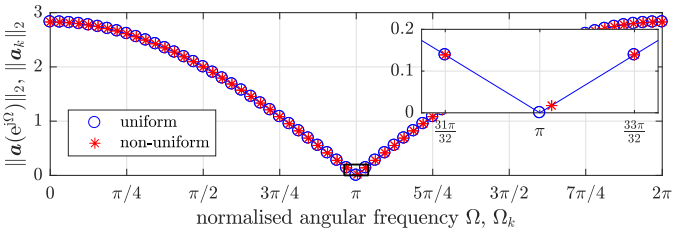

Fig. 1. Norm $\|\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ for Example 4, the evaluation with a uniform $K = 64$ point DFT, and the shift of the bin at $\Omega = \pi$ for the non-uniform DFT avoiding $\|\mathbf{a}_k\|_2 = 0$ for any value of $k = 0, \ldots, (K-1)$.

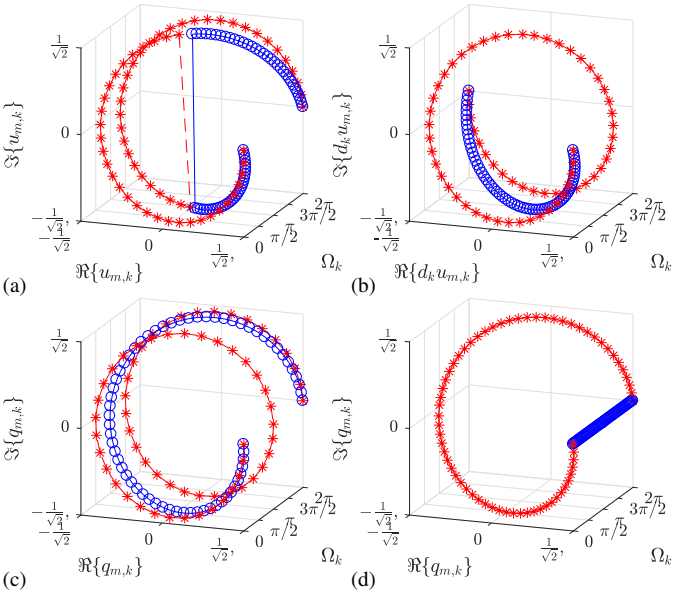

Fig. 2. Components for example of a normalised vector (a) $\mathbf{u}_k$, (b) including a sign correction $d_k\mathbf{u}_k$, and (c) a phase shift $\mathrm{e}^{-\mathrm{j}\frac{1}{2}\Omega_k} d_k\mathbf{u}_k$ equivalent to a half sample delay and (d) a half sample advance $\mathrm{e}^{\mathrm{j}\frac{1}{2}\Omega_k} d_k\mathbf{u}_k$. The two vector components are colour-coded, with $m = 1$ in blue circles ($\circ$) and $m = 2$ in red asterisks ($*$).

with $\boldsymbol{q}(z)$ from (26). With an allpass filter $\varphi(z) = z$ applied, the solution $\boldsymbol{q}(z)$ shown in Fig. 2(d) is also analytic, but has a lower order than the solution in Fig. 2(c) as can be seen from the reduction in the number of 2π turns. △

D. Phase Smoothing and Sufficient DFT Size

For the bin-wise normalisation of Sec. IV-C we have not yet stated what the DFT size K should be. Too low a value of K will lead to insufficient resolution, and the bin width may not be small enough compared to the coherence bandwidth of $\boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega})$ for the premise $\mathbf{a}_k \approx \pm\mathbf{a}_{k-1}$ of Sec. IV-C to hold. The consequence will also be that adjacent bins may enclose more than one zero-crossing; thus, with respect to Sec. IV-C, an even number of such zero crossings would remain undetected, while an odd number of zero crossings would be detected as a single zero crossing between bins. Insufficient resolution in the DFT domain is equivalent to the time-domain reconstruction suffering from time-domain aliasing. Such a challenge will be eventually resolved if K

is selected sufficiently large, which will be addressed further below in Lemmas 3 and 4: the former will present a necessary criterion that is computationally inexpensive to evaluate, while the later covers a both necessary and sufficient criterion, but the sufficiency comes at an increased computational burden.

Additionally the allpass ambiguity for the normalised vector as mentioned in a comment after the proof to Theorem 1 in Sec. III-A directly relates to the DFT size. Let $\hat{\mathbf{q}}^{(K)}[n] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{q}}^{(K)}(z)$ be the non-uniform IDFT of the sample points $\{\mathbf{q}_k\}$ returned by Algorithm 1 when operating with K bins. Note that while the bin-wise normalisation of an analytic function preserves phase coherence between bins, an allpass function will affect the required order of a solution $\hat{\boldsymbol{q}}^{(K)}(z)$. Hence, in the context of finding a suitable DFT size K, it will be important to select the arbitrary allpass such that $\hat{\boldsymbol{q}}^{(K)}(z)$ attains an order that is as small as possible. This is equivalent to $\hat{\mathbf{q}}^{(K)}[n]$ possessing minimum support, and $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ to contain maximally smooth functions. This is evident in Example 5, where the solution in Fig. 2(d) is smoother compared to that in Fig. 2(c).

Hence before attempting any reconstruction from $\{\mathbf{q}_k\}$, phase smoothing should be applied (see Algorithm 2 in [46]). For this, we evaluate the smoothness of a Dirichlet interpolation $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ through the sample points $\{\mathbf{q}_k\}$. The μth element $\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ of $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ can be obtained from a vector $\bar{\mathbf{q}}_\mu \in \mathbb{C}^K$ containing the K sample points of the μth element of $\{\mathbf{q}_k\}$ via

$$\hat{Q}_\mu^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) = \frac{1}{\sqrt{K}} \mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) \bar{\mathbf{W}}_K^{\mathrm{H}} \bar{\mathbf{q}}_\mu \,, \tag{38}$$

where $\mathbf{e}_K^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) = [1, \mathrm{e}^{-\mathrm{j}\Omega}, \ldots, \mathrm{e}^{-\mathrm{j}\Omega(K-1)}]$. Further, we have $\bar{\mathbf{W}}_K = \mathrm{diag}\{\mathrm{e}^{-\mathrm{j}2\pi\delta_0/K}, \ldots, \mathrm{e}^{-\mathrm{j}2\pi\delta_{K-1}/K}\}\, \mathbf{W}_K$, with $\mathbf{W}_K$ a K-point DFT matrix and δ_k, $k = 0, \ldots, (K-1)$ the shift applied to the kth frequency bin as outlined in Sec. IV-B. As such, (38) represents a non-uniform inverse DFT, followed by a discrete time (but continuous frequency, i.e. $\Omega \in \mathbb{R}$) Fourier transform. The infinite differentiability associated with analyticity [26] implies smoothness, which can be assessed via a cost function that measures the power in derivatives of $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ with respect to frequency [46]. By incorporating non-uniform sampling via (38), we can use this cost function for phase smoothing in the same fashion as applied for eigenvectors in [40]. Maximising the smoothness of $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ minimises the time domain support of $\hat{\mathbf{q}}^{(K)}[n]$. This ensures that for a sufficiently large DFT size K, the effect of time domain aliasing in $\hat{\mathbf{q}}^{(K)}[n]$ is as small as possible.

If the phase smoothing procedure is not applied, the resulting solution remains valid but may exhibit a larger time-domain support than necessary. Therefore, in applications where the polynomial order of $\hat{\boldsymbol{Q}}(z)$ is not a critical metric, the computationally expensive phase smoothing procedure can be omitted. However, when phase smoothing is applied, the resulting time-domain support is guaranteed to be minimal [40].

We can employ the Dirichlet interpolation in (38) as a first criterion to establish whether K is sufficiently large. While the normalisation on the unit circle is bound to be correct in the K sample points Ω_k, $k = 0, \ldots, (K-1)$, for any value of K, it can deviate elsewhere. We therefore measure

$$\zeta_{\mathrm{U}} = \frac{1}{2\pi} \int_0^{2\pi} \left| \hat{\boldsymbol{q}}^{(K),\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) \hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) - 1 \right|^2 \mathrm{d}\Omega \,. \tag{39}$$

This general deviation from unity measured by (39) can equivalently be expressed via Parseval's theorem as

$$\zeta_{\mathrm{U}} = \sum_n \left| \hat{\mathbf{q}}^{(K),\mathrm{H}}[-n] * \hat{\mathbf{q}}^{(K)}[n] - \delta[n] \right|^2 \,, \tag{40}$$

which assesses time domain aliasing using the convolution operator $*$. Based on discrete parameters $\hat{\mathbf{q}}^{(K)}[n]$, (40) can be evaluated numerically. The computational effort for this is relatively low. We can state:

Lemma 3 (Loss of unit magnitude): With ζ_{U} as defined in (39) and some suitably low threshold ϵ_{U}, for $\zeta_{\mathrm{U}} > \epsilon_{\mathrm{U}}$ the DFT size K is insufficient.

Proof. The existence of an analytic solution for $\boldsymbol{q}(z)$ is guaranteed by Theorem 1. This solution must satisfy $\|\boldsymbol{q}(\mathrm{e}^{\mathrm{j}\Omega})\|_2 = 1$ $\forall\Omega$ which implies from (39) that $\zeta_{\mathrm{U}} = 0$. Therefore if the total (i.e. integrated) error in unity ζ_{U} exceeds a threshold ϵ_{U}, then K is not large enough. A value $\zeta_{\mathrm{U}} > \epsilon_{\mathrm{U}}$ therefore indicates that the sample points Ω_k are not sufficiently closely spaced to allow a suitable interpolation. ■

Due to analyticity, $\mathbf{q}[n]$ must be absolutely convergent. Since K also relates to the approximation order, an insufficient value of K may — but is not guaranteed to — lead to time domain aliasing. Hence as long as $\zeta_{\mathrm{U}} > \epsilon_{\mathrm{U}}$, we know that K is insufficient, but from $\zeta_{\mathrm{U}} < \epsilon_{\mathrm{U}}$ we cannot deduce that a large enough value of K has been reached. We therefore require a second, sufficient criterion.

Since $\boldsymbol{q}(z)$ must align with $\boldsymbol{a}(z)$ but can carry an arbitrary allpass factor $\varphi(z)$, we can assess the projection of $\boldsymbol{a}(z)$ onto the orthogonal complement of $\boldsymbol{q}(z)$, i.e. $(\mathbf{I}_M - \boldsymbol{q}(z)\boldsymbol{q}^{\mathrm{P}}(z))$. Note that the allpass ambiguity cancels out in the term $\boldsymbol{q}(z)\boldsymbol{q}^{\mathrm{P}}(z)$, see e.g. (11), and we have that, for the correct values,

$$\left(\mathbf{I}_M - \boldsymbol{q}(z)\boldsymbol{q}^{\mathrm{P}}(z)\right) \boldsymbol{a}(z) = \underline{0} \,. \tag{41}$$

We therefore define the projection error ζ_{P},

$$\zeta_{\mathrm{P}} = \frac{1}{2\pi} \int_0^{2\pi} \left\| \left(\mathbf{I}_M - \hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega}) \hat{\boldsymbol{q}}^{(K),\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) \right) \boldsymbol{a}(\mathrm{e}^{\mathrm{j}\Omega}) \right\|_2^2 \mathrm{d}\Omega \,, \tag{42}$$

based on the same Dirichlet interpolation $\hat{\boldsymbol{q}}^{(K)}(\mathrm{e}^{\mathrm{j}\Omega})$ as used in (39). Exploiting Parseval, the criterion (42) can be equivalently formulated in the time domain as

$$\zeta_{\mathrm{P}} = \sum_n \left\| \left(\mathbf{I}_M \delta[n] - \hat{\mathbf{q}}^{(K)}[n] * \hat{\mathbf{q}}^{(K),\mathrm{H}}[-n] \right) * \mathbf{a}[n] \right\|_2^2 \,. \tag{43}$$

Analogous to ζ_{U}, (43) can be evaluated from discrete coefficients $\hat{\mathbf{q}}^{(K)}[n]$ obtained from a non-uniform IDFT of $\{\mathbf{q}_k\}$, but requires a matrix evaluation rather than an inner product, and is therefore computationally more expensive to evaluate than (40).

Lemma 4 (Projection error): If and only if the projection error ζ_{P} in (42) and the error in unity ζ_{U} in (39) fall below some small thresholds ϵ_{P} and ϵ_{U}, then the DFT size K will be sufficiently large.

Proof. If the approximation $\hat{\boldsymbol{q}}^{(K)}(z)$ for sufficiently large K approaches $\boldsymbol{q}(z)$, due to (41) we have $\zeta_{\mathrm{P}} \longrightarrow 0$. In turn for the 'only if' part, with $\zeta_{\mathrm{P}} = 0$, $\hat{\boldsymbol{q}}^{(K)}(z)$ sits in the same 1-d subspace as $\boldsymbol{a}(z)$ and $\boldsymbol{q}(z)$. The vector $\boldsymbol{q}(z)$ has unit norm by construction and, with $\zeta_{\mathrm{U}} < \epsilon_{\mathrm{U}}$, $\hat{\boldsymbol{q}}(z)$ can be considered to have unit norm for all intents and purposes. Similar to analytic eigenvectors in 1-d subspaces —see Sec. VI of [40] — the only variability of unit norm analytic vectors can be an arbitrary analytic allpass function. Thus, $\hat{\boldsymbol{q}}^{(K)}(z)$ must match $\boldsymbol{q}(z)$ except for an arbitrary allpass function as seen in (11). ■

With Lemmas 3 and 4 we now have two criteria to check convergence of a normalisation algorithm. The former, based on the error in unit magnitude, is a necessary but not proven to be sufficient criterion with reasonably low computational cost. The latter, which additionally assesses a projection error, offers a both necessary and sufficient condition, but is computationally more costly than the former.

E. Overall Normalisation Algorithm

We now have all components in place to normalise an analytic vector $\boldsymbol{a}(z)$, as summarised in Algorithm 2. Note that although a small projection error ζ_{P} is both necessary and sufficient, evaluating it is computationally expensive, so it is prudent to check for a small ζ_{U} first. An inner loop iteratively increases the DFT length until the unit magnitude criterion of Lemma 3 is satisfied. The DFT sizes are computed as powers of two starting from an order that just exceeds the order of $\boldsymbol{a}(z)$. Selecting powers of two enables the use of FFT algorithms despite potentially having to interpolate a limited number of non-uniformly spaced bins [42]. An outer loop additionally evaluates the slightly more complicated but thus sufficient projection error criterion of Lemma 4. The algorithm terminates if the latter metric falls below the threshold ϵ_{P}, or if the DFT length exceeds a preset limit $K_{\max}$.

Theorem 3 (Convergence of normalisation algorithm): With suitable thresholds ϵ_{U} and ϵ_{P}, for any given analytic vector Algorithm 2 will converge towards an analytic normalised vector.

Proof. For any given analytic vector $\boldsymbol{a}(z)$, Theorem 1 has established that a normalised, analytic $\boldsymbol{q}(z)$ exists. Irrespective of the allpass ambiguity of $\boldsymbol{q}(z)$, analyticity implies absolute convergence of $\mathbf{q}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{q}(z)$. Since the DFT size K increases with every iterations and $\mathbf{q}[n]$ decays at least exponentially, K will eventually exceed the support of $\boldsymbol{q}(z)$, which can be detected by the necessary and sufficient criteria of Lemma 4. ■

The normalised vector extracted by Algorithm 2 is written as $\hat{\mathbf{q}}^{(K)}[n]$, with the hat notation to indicate that at DFT length and therefore approximation order K, this may indeed be a approximation, depending on how the thresholds ϵ_{U} and ϵ_{P} are selected.

Algorithm 2: Analytic vector normalisation

Input: $\mathbf{a}[n]$; ϵ_{U}; ϵ_{P}; $K_{\max}$;
$K = 2^{\lceil 1+\log_2 K_0 \rceil}$ with K_0 the support of $\mathbf{a}[n]$;
repeat
  repeat
    determine $\{\mathbf{a}_k | k = 0, \ldots, (K-1);\ \|\mathbf{a}_k\|_2 > \epsilon\}$ as non-uniform DFT of $\mathbf{a}[n]$;
    execute Algorithm 1 to obtain $\{\mathbf{q}_k\}$;
    apply phase smoothing to $\{\mathbf{q}_k\}$ via Algorithm 2 from [40];
    evaluate ζ_{U};
    $K \leftarrow 2K$;
  until $(\zeta_{\mathrm{U}} < \epsilon_{\mathrm{U}}) \vee (K > K_{\max})$;
  evaluate ζ_{P};
until $(\zeta_{\mathrm{P}} < \epsilon_{\mathrm{P}}) \vee (K > K_{\max})$;
Output: $\hat{\mathbf{q}}^{(K)}[n]$ as non-uniform IDFT of $\{\mathbf{q}_k\}$

Algorithm 3: Analytic modified Gram-Schmidt

Input: $\boldsymbol{A}(z)$;
$\boldsymbol{A}(z) = [\boldsymbol{a}_{1,1}(z),\ \ldots,\ \boldsymbol{a}_{M,1}(z)]$;
for $m = 1 : M$ **do**
  $\hat{\boldsymbol{q}}_m(z)$ via (29) using Algorithm 2;
  for $\mu = m+1 : M$ **do**
    calculate $\bar{\boldsymbol{a}}_{\mu,m+1}(z)$ via (30)
  end
end
Output: $\hat{\boldsymbol{Q}}(z) = [\hat{\boldsymbol{q}}_1(z),\ \ldots,\ \hat{\boldsymbol{q}}_M(z)]$;

V. Gram-Schmidt-Based Analytic QR Decomposition

A. Analytic Modified Gram-Schmidt Procedure

To extend the modified Gram-Schmidt procedure to analytic matrices, we rely on the results for the normalisation in Theorem 1 and the convergence of the associated Algorithm 2 guaranteed by Theorem 3. We here use this to extend the modified Gram-Schmidt procedure based on the initialisation

$$\boldsymbol{A}(z) = [\bar{\boldsymbol{a}}_{1,1}(z),\ \ldots,\ \bar{\boldsymbol{a}}_{M,1}(z)]\ . \tag{44}$$

Then at the mth iteration, $m = 1, \ldots, M$, we perform one normalisation according to (29) and $M - m - 1$ projections using (30) for $\mu = m+1, \ldots, M$. If phase smoothing is applied to $\boldsymbol{q}_m(z)$ after the normalisation step as part of Algorithm 2, then this will help to shorten the support of $\boldsymbol{q}_m(z)$ and directly that of the extracted $\boldsymbol{Q}(z)$. It may also simplify the normalisation step in (30). The procedure is summarised in Algorithm 3.

B. Analytic QR Decomposition Algorithm

Based on the existence of the analytic QRD in Sec. III-D, we determine the upper triangular matrix $\hat{\boldsymbol{R}}(z)$ via (32) using $\boldsymbol{A}(z)$ and the estimate $\hat{\boldsymbol{Q}}(z)$ obtained from Algorithm 3. It is also possible to determine entries of $\hat{\boldsymbol{R}}(z)$ as part of the evaluation of Algorithm 3 akin to the method in Algorithm

5.4.1 of [28], such that e.g. the diagonal entries of $\hat{\boldsymbol{R}}(z)$ are square magnitude spectra that have already been explicitly calculated. The convergence of Algorithm 3 and thus the above analytic QR decomposition are covered by Theorem 3 for the normalisation and the general Gram-Schmidt procedure.

Note that even if $\boldsymbol{A}(z)$ has full spatial rank, it may not be invertible. Thus the matrix $\hat{\boldsymbol{R}}(z)$ may inherit non-invertibility. Closer inspection of the normalisation process reveals that

$$\boldsymbol{A}(z) = \hat{\boldsymbol{Q}}(z)\boldsymbol{\Gamma}(z)\tilde{\boldsymbol{R}}(z)\,, \tag{45}$$

where the diagonal matrix $\boldsymbol{\Gamma}(z)$ contains all spectral zeros of the normalisation process of Sec. IV, such that the analytic right triangular matrix $\tilde{\boldsymbol{R}}(z)$ has a non-zero determinant everywhere on the unit circle and is now invertible[7].

VI. Simulations and Results

A. Performance Metrics

To assess the performance of the QR decomposition $\boldsymbol{A}(z) = \hat{\boldsymbol{Q}}(z)\hat{\boldsymbol{R}}(z)$ produced by the proposed algorithm and by different benchmark methods, we formulate various metrics for the factors $\hat{\boldsymbol{Q}}(z)$ and $\hat{\boldsymbol{R}}(z)$. In the ideal case, $\hat{\boldsymbol{Q}}(z)$ is a paraunitary matrix that satisfies $\hat{\boldsymbol{Q}}(z)\hat{\boldsymbol{Q}}^{\mathrm{P}}(z) = \mathbf{I}_M$. To assess how close $\hat{\boldsymbol{Q}}(z)$ is to a paraunitary matrix, we therefore measure, akin to (40),

$$\eta_{\mathrm{PU}} = \sum_n \|\hat{\mathbf{Q}}[n] * \hat{\mathbf{Q}}^{\mathrm{H}}[-n] - \mathbf{I}\delta[n]\|_{\mathrm{F}}^2\,, \tag{46}$$

which assesses the loss in paraunitarity in the time domain.

As a second criterion, we check how close $\hat{\boldsymbol{R}}(z)$ is to an upper triangular matrix. This is not sensible for the proposed analytic algorithm if $\hat{\boldsymbol{R}}(z)$ is evaluated as an upper triangular matrix in Sec. V-B. However, if an approximation of $\boldsymbol{R}(z)$ is evaluated via $\hat{\boldsymbol{R}}(z) = \hat{\boldsymbol{Q}}^{\mathrm{P}}(z)\boldsymbol{A}(z)$, such a metric provides another measure for the precision of $\hat{\boldsymbol{Q}}(z)$. This metric is also important for benchmarking algorithms such as [17], [20], [21], which yield only approximately upper right triangular matrices. By separating $\hat{\boldsymbol{R}}(z)$ into

$$\hat{\boldsymbol{R}}(z) = \bar{\boldsymbol{R}}(z) + \boldsymbol{L}(z)\,, \tag{47}$$

where $\bar{\boldsymbol{R}}(z)$ contains the upper-right triangular part of $\hat{\boldsymbol{R}}(z)$ and $\boldsymbol{L}(z)$ $\bullet\!\!-\!\!\circ$ $\mathbf{L}[n]$ the components of $\hat{\boldsymbol{R}}(z)$ below its diagonal, we define

$$\eta_{\mathrm{TR}} = \frac{\sum_n \|\mathbf{L}[n]\|_{\mathrm{F}}^2}{\sum_n \|\mathbf{R}[n]\|_{\mathrm{F}}^2} = \frac{\sum_n \|\mathbf{L}[n]\|_{\mathrm{F}}^2}{\sum_n \|\mathbf{A}[n]\|_{\mathrm{F}}^2}\,, \tag{48}$$

which therefore measures the normalised energy below the diagonal of $\hat{\boldsymbol{R}}(z)$. Since in (48), the ground truth factor $\boldsymbol{R}(z)$ in the denominator may not be known, we exploit the fact that it is only a paraunitary operation away from the given matrix $\boldsymbol{A}(z)$ i.e. $\sum_n \|\mathbf{R}[n]\|_{\mathrm{F}}^2 = \sum_n \|\mathbf{A}[n]\|_{\mathrm{F}}^2$. If $\hat{\boldsymbol{R}}(z)$ was a perfectly upper triangular matrix, then $\eta_{\mathrm{TR}} = 0$.

[7]This factorisation is in fact at the heart of the modified Gram-Schmidt orthogonalisation procedure [49]. Effectively one is using a modified QR decomposition i.e. instead of $\mathbf{A} = \mathbf{QR}$, one writes $\mathbf{A} = \mathbf{QD}\tilde{\mathbf{R}}$. Here $\mathbf{D}$ is diagonal and $\tilde{\mathbf{R}}$ is upper-triangular with ones on the diagonal.

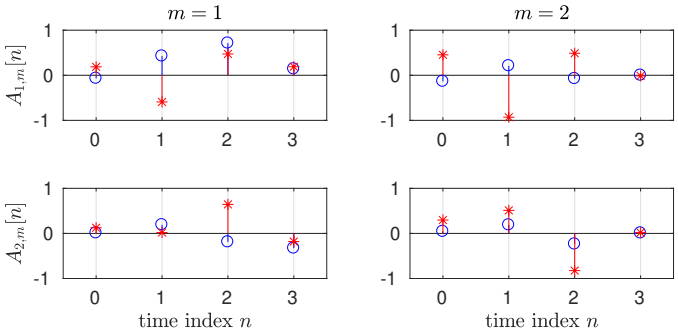

Fig. 3. Matrix $\mathbf{A}[n]$ $\circ\!\!-\!\!\bullet$ $\boldsymbol{A}(z)$ of Example 6, showing the real ($\circ$) and imaginary parts ($*$) of the elements $A_{\mu,m}[n]$ in the μth row and mth column.

As a further criterion, we consider the reconstruction error $\boldsymbol{E}(z) = \boldsymbol{A}(z) - \hat{\boldsymbol{Q}}(z)\bar{\boldsymbol{R}}(z)$ and its associated metric

$$\eta_{\mathrm{REC}} = \sum_n \|\mathbf{E}[n]\|_{\mathrm{F}}^2 \tag{49}$$

with $\mathbf{E}[n]$ $\circ\!\!-\!\!\bullet$ $\boldsymbol{E}(z)$. Note that we have

$$\begin{aligned}\boldsymbol{E}(z) &= \boldsymbol{A}(z) - \hat{\boldsymbol{Q}}(z)\big(\hat{\boldsymbol{Q}}^{\mathrm{P}}(z)\boldsymbol{A}(z) - \boldsymbol{L}(z)\big) &(50)\\ &= \big(\mathbf{I} - \hat{\boldsymbol{Q}}(z)\hat{\boldsymbol{Q}}^{\mathrm{P}}(z)\big)\mathbf{A}(z) \;-\; \hat{\boldsymbol{Q}}(z)\boldsymbol{L}(z)\,, &(51)\end{aligned}$$

where the first term is related to the loss in paraunitarity, and the second term to the triangularisation. Hence, η_{REC} captures aspects of both η_{PU} and η_{TR}.

B. Worked Examples

Example 6: The matrix $\boldsymbol{A}(z)$ shown in Fig. 3 is constructed as $\boldsymbol{A}(z) = \boldsymbol{Q}(z)\boldsymbol{R}(z)$, with a 2nd order random elementary paraunitary matrix $\boldsymbol{Q}(z)$ [29] and the upper triangular matrix

$$\boldsymbol{R}(z) = \begin{bmatrix} 1 + \mathrm{j}z^{-1} & \frac{1}{2} - \frac{1}{2}z^{-1} \\ 0 & 1 - z^{-1} \end{bmatrix}\,. \tag{52}$$

With $\boldsymbol{R}(z) = [\boldsymbol{r}_1(z),\ \boldsymbol{r}_2(z)]$, on the unit circle we find for its column norms

$$\begin{aligned}\|\boldsymbol{r}_1(\mathrm{e}^{\mathrm{j}\Omega})\|_2^2 &= 2 + 2\sin\Omega\,,\\ \|\boldsymbol{r}_2(\mathrm{e}^{\mathrm{j}\Omega})\|_2^2 &= \frac{5}{2} - \frac{5}{2}\cos\Omega\,.\end{aligned}$$

Since $\boldsymbol{Q}(z)$ is paraunitary, this implies $\|\boldsymbol{a}_m(\mathrm{e}^{\mathrm{j}\Omega})\|_2 = \|\boldsymbol{r}_m(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ for $m = 1, 2$. These norms are depicted in Fig. 4 on the interval $[0;\ 2\pi)$, where $\|\boldsymbol{a}_1(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ has a zero at $\Omega = \frac{3}{2}\pi$, while $\|\boldsymbol{a}_2(\mathrm{e}^{\mathrm{j}\Omega})\|_2$ has a zero at $\Omega = 0$. The determinant $\det\{\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})\}$, as shown in Fig. 4, also reflects these spectral zeros.

Algorithm 3 with thresholds $\epsilon_{\mathrm{U}} = \epsilon_{\mathrm{P}} = 10^{-12}$ converges in a single iteration, with a DFT size of $K = 8$, and a projection error ζ_{P} below -289.1 dB. The error in paraunitarity η_{PU} is -288.8 dB, with an error in triangularisation η_{TR} of -297.1 dB and a reconstruction error η_{REC} of -287.1 dB. Here, the small finite order of the ground truth matrices $\boldsymbol{Q}(z)$ and $\boldsymbol{R}(z)$ leads to compact factors $\hat{\boldsymbol{Q}}(z)$ and $\hat{\boldsymbol{R}}(z)$ of the

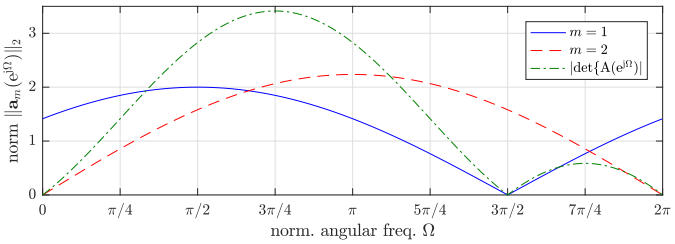

Fig. 4. Column norms and modulus of the determinant of $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})$ for Example 6, with spectral zeros at $\Omega = 0$ and $\Omega = \frac{3}{2}\pi$.

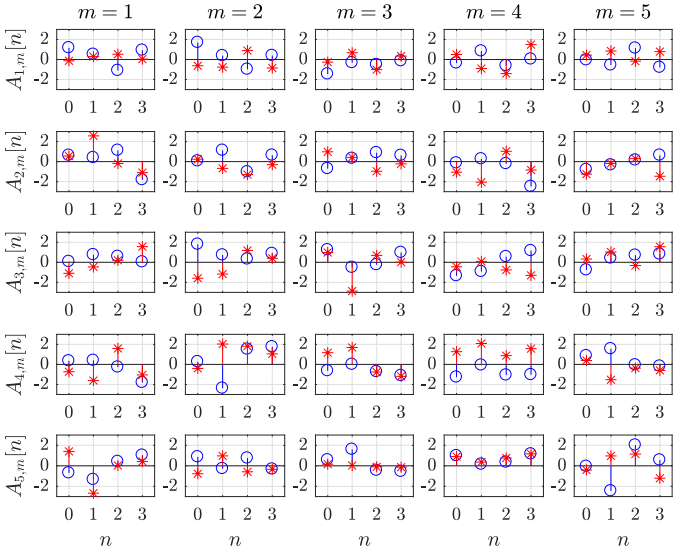

Fig. 5. Elements of matrix $\mathbf{A}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{A}(z)$ for Example 7, with real ($\circ$) and imaginary parts ($*$).

same order, such that $\boldsymbol{\Phi}$ is constant rather than a frequency-dependent matrix of allpass filters. In contrast, the column-wise PQRD from [17], [20], operating in the time domain, completes after 340 iterations and reaches factors of orders 61 and 59 due to the lack of phase smoothing, with metrics η_{PU} of -30.1 dB, η_{TR} of -55.4 dB, and a reconstruction error η_{REC} of -27.3 dB. The DFT domain method in [23] fails to return a result, as it cannot cope with zero column norms or singularities of the determinant occurring at frequencies that coincide with DFT bins. This is the case for $\Omega = 0$ and $\Omega = \frac{3}{2}\pi$, as seen in Fig. 4. △

In Example 6, there is a finite order ground truth, that is — within the allpass/phase ambiguity — recovered by the proposed solution. Generally, the QR factors can be Laurent series that due to their analyticity are absolutely convergent but yet can be infinite in order. The next example addresses such a more general case.

Example 7: A 5×5 matrix $\boldsymbol{A}(z)$ of order 3 is characterised by the impulse responses of its elements, with independent coefficients $A_{m,\mu}[n] \sim \mathcal{CN}(0,1)$, in Fig. 5. Applying Algorithm 3 with thresholds $\epsilon_{\mathrm{U}} = \epsilon_{\mathrm{P}} = 10^{-12}$ in Algorithm 2 requires iterations up to a DFT order of 2^{10} to reach factors $\hat{\boldsymbol{Q}}(z)$ with a loss of paraunitarity η_{PU} of -49.2 dB and a

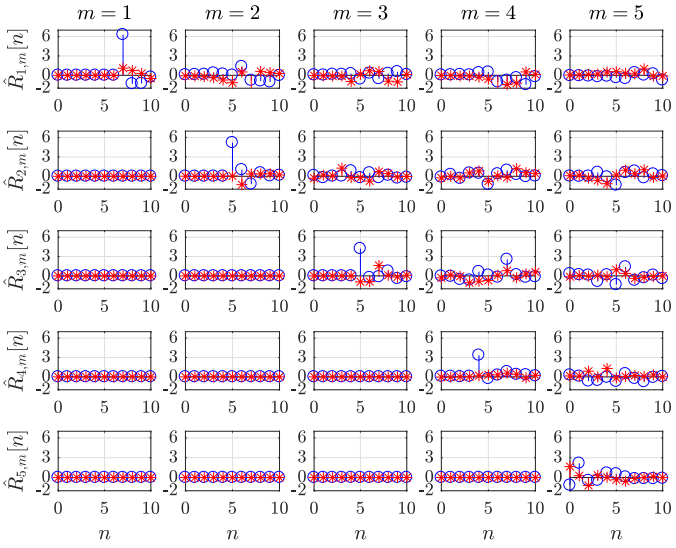

Fig. 6. Elements of matrix $\hat{\mathbf{R}}[n] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{R}}(z)$ for Example 7, with real ($\circ$) and imaginary parts ($*$).

$\hat{\boldsymbol{R}}(z)$ with an error in triangularisation η_{TR} of -58.5 dB. The reconstruction error η_{REC} is -41.3 dB. If these values are used to motivate trimming $\hat{\boldsymbol{R}}(z)$ to a precision of -50 dB, i.e. to remove leading and trailing Laurent polynomial coefficients that fall below this threshold as common in polynomial matrix decomposition algorithms [47], [50]–[52], then its order is approximately 45. The central 15 coefficients of $\hat{\mathbf{R}}[n] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{R}}(z)$ are shown in Fig. 6, demonstrating its upper triangular nature. △

C. Ensemble Simulation

For a variable spatial dimension $M \in \{2, 4, 8, 16, 32\}$, we generate an ensemble of matrices $\boldsymbol{A}(z) = \boldsymbol{Q}(z)\boldsymbol{R}(z)$ from randomised factors $\boldsymbol{Q}(z)$ and $\boldsymbol{R}(z)$. Based on random vectors $\mathbf{v} \in \mathbb{C}^M$ with zero mean unit norm complex Gaussian entries, we generate random elementary paraunitary matrices $\boldsymbol{Q}(z) = \mathbf{I}_M - \mathbf{v}\mathbf{v}^{\mathrm{H}}/\|\mathbf{v}\|^2(1-z^{-1})$ according to [29], while for $\boldsymbol{R}(z) = \mathbf{R}_0 + \mathbf{R}_1 z^{-1}$, the coefficients of the upper right triangular matrices $\mathbf{R}_n$, $n = 0, 1$, are also zero mean unit norm complex Gaussian distributed. Some of these matrices contain spectral zeros for the norm, or zeros that are very close to the unit circle. For the viability of the method in [23], ensemble probes $\boldsymbol{A}(z)$ whose determinant $\det \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})$ falls below $\epsilon = 10^{-12}$ in any bin for a $K_{\max} = 2^{12}$-point DFT are excluded from simulations.

In terms of algorithms, we compare the time domain elimination by column (PQRD-BC) method in [17], [20], a multiple-shift version (PQRD-MS) from [21] which is similar to [22], and the DFT-domain approach (PQRD-DFT) from [23] with the proposed method. For the computationally intensive time domain methods PQRD-BF and PQRD-MS, we simulate an ensemble of 10^3 random matrices $\boldsymbol{A}(z)$ for every value for M. For PQRD-DFT and the proposed approach, the ensemble size is 10^4.

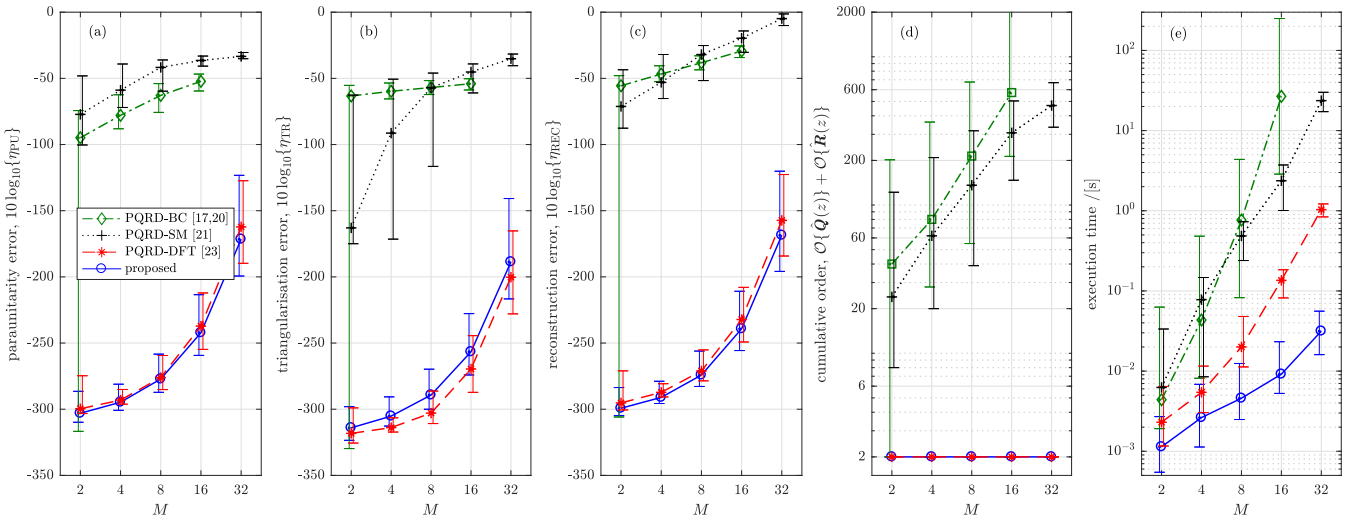

Fig. 7. Comparison of the proposed method with PQRD-BC [17], [20], PQRD-SM [21], and PQRD-DFT [23], showing the (a) paraunitarity error η_{PU}, (b) triangularisation error η_{TR}, (c) reconstruction error η_{REC}, (d) order to the extracted factors $\hat{\boldsymbol{Q}}(z)$ and $\hat{\boldsymbol{R}}(z)$, and (e) execution time of the algorithms versus the spatial dimension M. The curves show the median, while the error bars indicate 5- and 95-percentiles of the ensemble distributions. For better distinction, the error bars are slightly offset horizontally w.r.t. each other.

The ensemble results for the proposed method and for three benchmark algorithms are displayed in Fig. 7. The time domain methods PQRD-BD and PQRD-MS are allowed a maximum of 500 iterations; since their order tends to grow with every iteration, an internal truncation of leading and trailing coefficients below 10^{-8} is performed [17], [47], [48] to limit the order of the factors. This has the effect of introducing a lower bound on the paraunitarity error η_{PU}, see Fig. 7(a). Here the DFT domain methods perform significantly better. Since the triangularisation of the time domain methods is only approximated and the factorisation does not necessarily target the analytic solution, the same difference in performance can be observed for the triangularisation error η_{TR} defined in (48) in Fig. 7(b), and for the reconstruction error η_{REC} of (49) in Fig. 7(c). For all algorithms, the performance degrades as M increases.

The cumulative order of the extracted factors $\hat{\boldsymbol{Q}}(z)$ and $\hat{\boldsymbol{R}}(z)$, $\mathcal{O}\{\hat{\boldsymbol{Q}}(z)\}+\mathcal{O}\{\hat{\boldsymbol{R}}(z)\}=\mathcal{O}\{\hat{\boldsymbol{Q}}(z)\hat{\boldsymbol{R}}(z)\}$, is shown in Fig. 7(d). For the time domain methods PQRD-BC ad PQRD-SM, which as in Example 5 do not necessarily target the analytic solution, the number of iterations grows with M, as does therefore the order to the approximated factors. Note that PQRD-SM, which shifts more energy per iteration step and therefore requires fewer iterations than PQRD-BC, leads to a slightly lower order. Nonetheless, the orders are several magnitudes above what can be achieved with the DFT-domain methods, which in all cases extract orders equivalent to those of the ground truth factors.

Finally, Fig. 7(e) illustrates the execution times of the ensemble simulations. For small values of M, the time-domain methods are economic, but their complexity grows quickly with M. Particularly PQRD-BC, which converges slower per iteration and tends to yield higher order factors compared to the other algorithms as evident in Fig. 7(d), becomes increasingly costly with M, such that for $M=32$, it became infeasible to simulate an ensemble; it is for this reason that the PQRD-BC results for $M=32$ are missing from the plots in Fig. 7. The complexity of the DFT domain methods also grows with M. The proposed method has a cost that is slightly lower than that of the bin-wise QR approach PQRD-DFT from [23]. Note however that PQRD-DFT cannot cope with spectral zeros that coincide with bin frequencies, as has been the case in Example 5. In contrast, the proposed method based on Algorithm 3 exhibits execution times that are at least as fast or lower[8] but has proven convergence and no exclusions due to spectral zeros.

VII. Conclusions

In order to extend Gram-Schmidt orthogonalisation to vectors of analytic functions, this paper has first investigated the normalisation of such a vector. We have shown that on the unit circle, the magnitude spectrum — the Euclidean norm for specific frequency values — of an analytic vector may possess an isolated number of spectral zeros, but yet it is always possible to find an analytic solution to the normalisation process. This includes the case of an odd number of spectral zeros on the interval $0\leq\Omega<2\pi$, which requires a phase correction in order to avoid fractional delays for the solution. An algorithm with proven convergence has been presented.

The normalisation can then lead to an analytic extension of the Gram-Schmidt procedure when applied to the columns of an analytic matrix $\boldsymbol{A}(z)$. Provided that this matrix has full spatial rank, i.e. the determinant $\det\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})$ only has isolated zeros, the analytic Gram-Schmidt procedure leads to analytic QR decompositions, whose existence we have therefore estab-

[8]This includes some simple computational tricks, such as replacing time domain convolutions by multiplications in the DFT domain [53].

lished. An algorithm for this analytic QR decomposition has also been presented and shown to converge.

In examples and simulations, we demonstrate aspects of the approach and benchmarked it against existing polynomial QR methods. Time domain methods do not produce results with compact support, and have a computational requirement that does not scale well with the spatial dimension M. We have improved on an existing DFT domain algorithm by covering all possibilities, particularly when spectral zeros of the matrix determinant coincide with DFT bins. The approach can be applied to a number of problems, ranging e.g. from broadband MIMO communications [16], [18], [19] to the orthogonalisation or completion of polynomial matrices for the blocking matrix design in broadband beamformers [7].

ACKNOWLEDGEMENTS

We gratefully acknowledge the kind assistance by John G. McWhirter and Fraser K. Coutts in providing the code for [17], [20], [21].

REFERENCES

[1] G.W. Stewart, "The decompositional approach to matrix computation," *Computing in Science Engineering*, vol. 2, no. 1, pp. 50–59, Jan 2000.

[2] 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.

[3] S. Weiss, J. Pestana, I.K. Proudler, and F.K. 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.

[4] G. Barbarino and V. Noferini, "On the Rellich eigendecomposition of para-Hermitian matrices and the sign characteristics of $*$-palindromic matrix polynomials," *Linear Algebra Appl.*, vol. 672, pp. 1–27, Sep. 2023.

[5] S. Weiss, I.K. Proudler, G. Barbarino, J. Pestana, and J.G. McWhirter, "On properties and structure of the analytic singular value decomposition," *IEEE Trans. Signal Process.*, vol. 72, pp. 2260–2275, Apr. 2024.

[6] G. Barbarino, "On the periodicity of singular vectors and the holomorphic block-circulant SVD on the unit circumference," *Linear Algebra Appl.*, vol. 721, pp. 465–483, Sep. 2025.

[7] V.W. Neo, S. Redif, J.G. McWhirter, J. Pestana, I.K Proudler, S. Weiss, and P.A. Naylor, "Polynomial eigenvalue decomposition for multichannel broadband signal processing: A mathematical technique offering new insights and solutions," *IEEE Signal Process. Mag.*, vol. 40, no. 7, pp. 18–37, Nov. 2023.

[8] N. Moret, A. Tonello, and S. Weiss, "MIMO precoding for filter bank modulation systems based on PSVD," in *IEEE 73rd Vehicular Technology Conf.*, May 2011, pp. 1–5.

[9] X. Mestre and D. Gregoratti, "A parallel processing approach to filterbank multicarrier MIMO transmission under strong frequency selectivity," in *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, May 2014, pp. 8078–8082.

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

[11] M. Alrmah, S. Weiss, and S. Lambotharan, "An extension of the MUSIC algorithm to broadband scenarios using polynomial eigenvalue decomposition," in *19th European Signal Process. Conf.*, Barcelona, Spain, Aug. 2011, pp. 629–633.

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

[13] V.W. Neo, C. Evers, and P.A. Naylor, "Enhancement of noisy reverberant speech using polynomial matrix eigenvalue decomposition," *IEEE/ACM Trans. Audio, Speech, and Language Process.*, vol. 29, pp. 3255–3266, 2021.

[14] A.O.T. Hogg, V.W. Neo, S. Weiss, C. Evers, and P.A. Naylor, "A polynomial eigenvalue decomposition MUSIC approach for broadband sound source localization," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, Oct. 2021.

[15] V.W. Neo, C. Evers, S. Weiss, and P.A. Naylor, "Signal compaction using polynomial EVD for spherical array processing with applications," *IEEE/ACM Trans. Audio, Speech, and Language Process.*, vol. 31, pp. 3537–3549, 2023.

[16] D. Cescato, M. Borgmann, H. Bölcskei, J. Hansen, and A. Burg, "Interpolation-based QR decomposition in MIMO-OFDM systems," in *IEEE Workshop on Signal Processing Advances in Wireless Communications*, Jun. 2005, pp. 945–949.

[17] J. Foster, J.G. McWhirter, M. Davies, and J.A. Chambers, "An algorithm for calculating the QR and singular value decompositions of polynomial matrices," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1263–1274, Mar. 2010.

[18] D. Cescato and H. Bölcskei, "QR decomposition of Laurent polynomial matrices sampled on the unit circle," *IEEE Trans. Information Theory*, vol. 56, no. 9, pp. 4754–4761, Sep. 2010.

[19] ——, "Algorithms for interpolation-based QR decomposition in MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1719–1733, Apr. 2011.

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

[21] F.K. Coutts, J. Corr, K. Thompson, S. Weiss, I.K. Proudler, and J.G. McWhirter, "Multiple shift QR decomposition for polynomial matrices," in *IMA Int. Conf. Mathematics in Signal Process.*, Birmingham, UK, Dec. 2016.

[22] D. Hassan, S. Redif, S. Lambotharan, and I.K. Proudler, "Sequential polynomial QR decomposition and decoding of frequency selective MIMO channels," in *26th European Signal Process. Conf.*, Rome, Italy, Sep. 2018, pp. 460–464.

[23] F.A. Khattak, M.A. Bakhit, I.K. Proudler, and S. Weiss, "Smooth QR decomposition of polynomial matrices," in *9th IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Process.*, Los Sueños, Costa Rica, Dec. 2023.

[24] F. Ling and J. Proakis, "A generalized multichannel least squares lattice algorithm based on sequential processing stages," *IEEE Trans. Acoustics, Speech, and Signal Process.*, vol. 32, no. 2, pp. 381–389, 1984.

[25] I.K. Proudler, T. Shepherd, and J.G. McWhirter, "Computationally efficient QRD-based wide-band beamforming," in *IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, 1990, pp. 1799–1802 vol.3.

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

[27] G.D. Golden, C.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Lett.*, vol. 35, pp. 14–16, Jan. 1999.

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

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

[30] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.

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

[32] S. Weiss and G. Barbarino, "Relations between analytic spectral and singularvalue decompositions," in *33rd European Signal Process. Conf.*, Isola delle Femmine, Italy, Sep. 2025.

[33] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes, ii: The linear predictor," *Acta Mathematica*, vol. 99, pp. 93–137, 1958.

[34] I. Bronshtein and K. Semendyayew, *Handbook of Mathematics*. Heidelberg: Springer, 2015.

[35] S. Barnett, "Greatest common divisor of several polynomials," *Mathematical Proc. Cambridge Philosophical Society*, vol. 70, pp. 263–268, 9 1971.

[36] P. van Dooren, "Rational and polynomial matrix factorizations via recursive pole-zero cancellation," *Linear Algebra and its Applications*, vol. 137-–138, no. 0, pp. 663–697, 1990.

[37] A.-G. Wu, G.-R. Duan, and Y. Xue, "Kronecker maps and sylvester-polynomial matrix equations," *IEEE Trans. Automatic Control*, vol. 52, no. 5, pp. 905–910, 2007.
[38] Euclid., *Euclide's elements; the whole fifteen books compendiously demonstrated. [. . .].* London: Isaac Barrow, 1751.
[39] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the Unit Delay," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.
[40] S. Weiss, I.K Proudler, F.K. Coutts, and F.A Khattak, "Eigenvalue decomposition of a parahermitian matrix: extraction of analytic eigenvectors," *IEEE Trans. Signal Process.*, vol. 71, pp. 1642–1656, Apr. 2023.
[41] S. Weiss, S. J. Schlecht, and M. Moonen, "Best least squares paraunitary approximation: Analytic Procrustes problem," *IEEE Trans. Signal Process.*, vol. 74, pp. 904–919, Feb. 2026.
[42] F. A. Khattak, I. K. Proudler, and S. Weiss, "Scalable analytic eigenvalue extraction algorithm," *IEEE Access*, vol. 12, pp. 166 652–166 659, Dec. 2024.
[43] D. Potts, G. Steidl, and M. Tasche, *Fast Fourier Transforms for Nonequispaced Data: A Tutorial*. Boston, MA: Birkhäuser Boston, 2001, pp. 247–270.
[44] G. Plonka, D. Potts, G. Steidl, and M. Tasche, *Numerical Fourier Analysis*. Birkhäuser, 2019.
[45] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.
[46] S. Weiss, J. Selva, and M. D. Macleod, "Measuring smoothness of trigonometric interpolation through incomplete sample points," in *28th European Signal Process. Conf.*, Amsterdam, Netherlands, Jan. 2021, pp. 2319–2323.
[47] 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.
[48] 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.
[49] G. Proakis, C.M. Rader, F. Ling, C.L. Nikias, M. Moonen, and I.K. Proudler, *Algorithms for statistical signal processing*. Prentice Hall, NJ, 2002.
[50] J. Foster, J. G. McWhirter, and J. Chambers, "Limiting the order of polynomial matrices within the SBR2 algorithm," in *IMA International Conference on Mathematics in Signal Processing*, Cirencester, UK, December 2006.
[51] C. H. Ta and S. Weiss, "Shortening the order of paraunitary matrices in SBR2 algorithm," in *6th International Conference on Information, Communications & Signal Processing*, Singapore, Dec. 2007, pp. 1–5.
[52] J. Corr, K. Thompson, S. Weiss, I. Proudler, and J. McWhirter, "Row-shift corrected truncation of paraunitary matrices for PEVD algorithms," in *23rd European Signal Processing Conference*, Nice, France, August/September 2015, pp. 849–853.
[53] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1983.

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, Glasgow, Scotland, in 2019. He is currently a Postdoctoral Research Fellow in the Department of Computer Science at the University of Leeds, UK.

As a recipient of a prestigious Commonwealth Scholarship, he recently completed his PhD 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.

Ian K. Proudler received an M.A. degree in physics from Oxford University,U.K., in 1978, and a Ph.D. degree in digital signal processing from Cambridge University in 1984. He is currently a Visiting Professor of Signal Processing at the University of Strahclyde, Glasgow, U.K. Between degrees he spent two years doing R&D work in the U.K. electronics industry.

From 1986 until 2011 he worked in the U.K. 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 authored or co-authored more than 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.

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. He is a professor of signal processing at the University of Strathclyde, following previous academic appointments with the University of Southampton and the University of Strathclyde. His research interests include adaptive, multi-rate, and array signal processing, with applications in acoustics, communications, audio, and biomedical signal processing. He was the Technical Co-Chair of EUSIPCO 2009, General Chair of IEEE ISPLC 2014, and part of the organising committees for ICASSP 2019 and SSP 2025. He is a Subject Editor for Elsevier Signal Processing.

Sebastian J. Schlecht is an Associate Professor of Signal Processing at the Chair of Multimedia Communications and Signal Processing at the University of Erlangen-Nuremberg, Germany. He earned a Diploma in Applied Mathematics from the University of Trier, Germany, in 2010, followed by an M.Sc. in Digital Music Processing at Queen Mary University of London, U.K., in 2011. He completed his PhD at the International Audio Laboratories Erlangen, Germany, in 2017. Between 2019 and 2024, he held the position of Professor of Practice for Sound in Virtual Reality at Aalto University, Finland. He has received multiple best paper awards, including those from JAES, WASPAA, SSP, DAFX (2018, 2021, 2022, 2023, 2025), and AES AVAR.