Impact of Estimation Errors of a Matrix of Transfer Functions onto Its Analytic Singular Values and Their Potential Algorithmic Extraction

Mohammed A. Bakhit, Faizan A. Khattak, Ian K. Proudler, and Stephan Weiss
Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, Scotland
{mohammed.bakhit,faizan.khattak,ian.proudler,stephan.weiss}@strath.ac.uk

***Abstract*—A matrix of analytic functions $\boldsymbol{A}(z)$, such as the matrix of transfer functions in a multiple-input multiple-output (MIMO) system, generally admits an analytic singular value decomposition (SVD), where the singular values themselves are functions. When evaluated on the unit circle, for the sake of analyticity, these singular values must be permitted of become negative. In this paper, we address how the estimation of such a matrix, causing a stochastic perturbation of $\boldsymbol{A}(z)$, results in fundamental changes to the analytic singular values: for the perturbed system, we show that their analytic singular values lose any algebraic multiplicities and are strictly non-negative with probability one. We present examples and highlight the impact that this has on algorithmic solutions to extracting an analytic or approximate analytic SVD.**

I. Introduction

The singular value decomposition (SVD) is a well known and widely used mathematical technique within the toolkit of linear algebra. It admits a factorisation of any given matrix $\mathbf{A} \in \mathbb{C}^{M\times L}$ into a diagonalised form via unitary matrices [1, 2]. In signal processing, this factorization has found extensive applications summarised in e.g. [3, 4]. Specifically in signal processing for multiple-input multiple-output communications, the SVD enables to determine precoding and equalisation matrices that decouple a narrowband communications channel, i.e. a matrix comprising of complex gain factors, in order to achieve system optimality in various senses [5]. The desire to extent this utility to the broadband case where the channel matrix contains impulse responses — or, in the z-domain, transfer functions — has given rise to the investigation and application of a polynomial SVD [6–19].

In a polynomial matrix SVD [6], the SVD factors, i.e. the singular values and left- and right-singular vectors, themselves become frequency-dependent. The left- and right-singular vectors, which form so-called paraunitary matrices and implement lossless filter banks [20], can be approximated by matrices of finite impulse response filters [21–24]. The order of these filters then contributes to the implementation cost of any application; hence it is advantageous to find polynomial SVD factorisations that are as compact in order as possible. SVD algorithms in [6, 10, 25] are based on a set of algorithms that either favour or even guarantee [26] spectrally majorised singular values; these converge towards piecewise analytic functions and are therefore difficult to approximate. More recently, the existence of an analytic SVD has been established [27, 28], which postulates significantly smoother and therefore lower order SVD factors than may be achievable with current techniques. The extraction of analytic SVD factors therefore seems attractive, with initial attempts reported in [16–19] which operate analogously to the principles behind analytic EVD algorithms in [29–36].

In practice, for example in a communications system where the channel needs to be estimated from a finite amount of data, the system to which we want to apply an analytic SVD will be perturbed by a random component. In related work for the analytic eigenvalue decomposition of a space-time covariance matrix [37], a profound challenge has occurred: even if the space-time covariance matrix is subjected to only very small estimation errors [38], their perturbation effect on the analytic eigenvalues cannot be neglected. According to [38], unless the transition to an infinite sample size and a zero estimation error is made, the eigenvalues of the perturbed matrix will be spectrally majorised even if the unperturbed eigenvalues are not. Missing the smoothness of the ground truth solution means that any identified solution will require a significantly higher approximation order for the EVD factors, resulting in potentially high implementation costs.

Therefore, in this paper we want to investigate how approximation errors in a matrix of analytic functions will impact on its analytic SVD, and specifically its analytic singular values. Below, in Sec. II we review the analytic SVD, followed by a discussion on how the estimation perturbs a system matrix in Sec. III. Its impact of the perturbation of the singular values is laid out in Sec. IV, which contains the main findings of this paper: the singular values of a random perturbed matrix are non-negative and spectrally majorised with probability one. For this, we provide some examples in Sec. V and discuss consequences Sec. VI.

II. Analytic Singular Value Decomposition

A. Standard Singular Value Decomposition

For any given matrix $\mathbf{A} \in \mathbb{C}^{M\times L}$, without loss of generality[1] assuming $M \leq L$, there is a singular value decomposition

The work of Mohammed Bakhit is funded by Mathworks and the University of Strathclyde. Faizan Khattak is the recipient of a scholarship of the Commonwealth Scholarship Commission.

[1] Otherwise we address $\mathbf{A}^{\mathrm{H}}$ instead of $\mathbf{A}$.

[2] such that

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\rm H} \; , \tag{1}$$

where $\mathbf{U} \in \mathbb{C}^{M\times M}$ and $\mathbf{V} \in \mathbb{C}^{L\times L}$ are the left- and right-singular vectors. The quantity $\boldsymbol{\Sigma} \in \mathbb{R}^{M\times L}$ in (1) is a diagonal matrix containing the unique singular values σ_m, $m = 1, \ldots, M$, such that

$$\sigma_m \geq \sigma_{m+1} \geq 0, \qquad \forall m = 1, \ldots, (M-1) \; . \tag{2}$$

As per the standard definition of the SVD, the singular values are constrained to be non-negative. If there are singular values $\sigma_m = \sigma_{m+1} = \cdots = \sigma_{m+C-1}$, we speak of these singular values as possessing a C-fold algebraic multiplicity [2, 39].

B. Extension to Matrices of Analytic Functions

Extending (1) to polynomial matrices, or generally to matrices of functions $\boldsymbol{A}(z) : \mathbb{C} \rightarrow \mathbb{C}^{M\times L}$ that are analytic in z, requires an analytic SVD [27, 28]. Unless the system $\boldsymbol{A}(z)$ contains any multiplexing operation, such as implicit block filtering [40–42], the analytic SVD takes the form

$$\boldsymbol{A}(z) = \boldsymbol{U}(z)\boldsymbol{\Sigma}(z)\boldsymbol{V}^{\rm P}(z) \; . \tag{3}$$

In (3), $\boldsymbol{U}(z) \in \mathbb{C}^{M\times M}$ and $\boldsymbol{V}(z) \in \mathbb{C}^{L\times L}$ are the left- and right-singular vectors, respectively. Note that $\boldsymbol{U}(z)$ and $\boldsymbol{V}(z)$ are paraunitary matrices i.e. $\boldsymbol{U}(z)\;\boldsymbol{U}(z)^{\rm P} = \mathbf{I}$, $\boldsymbol{V}(z)\boldsymbol{V}^{\rm P}(z) = \mathbf{I}$, whereby the parahermitian operator $\{\cdot\}^{\rm P}$ implies a Hermitian transposition and time reversal such that $\boldsymbol{U}^{\rm P}(z) = \{\boldsymbol{U}^{\rm H}(1/z^*)\}^{\rm H}$. The diagonal matrix $\boldsymbol{\Sigma}(z) = \mathrm{diag}\{\sigma_1(z), \ldots, \sigma_M(z)\} \in \mathbb{C}^{M\times L}$ contains the singular values $\sigma_m(z)$, $m = 1, \ldots, M$. In order to admit analyticity of the SVD factors in (3), the singular values must be permitted to become negative on the unit circle, i.e. the restriction of non-negativity as known from the standard SVD in Sec. II-A must be dropped. This is known for the case where a matrix $\boldsymbol{A}(t)$ is a function in a real parameter $t \in \mathbb{R}$ [43, 44] as well as for the case of a dependency of $\boldsymbol{A}(z)$ on a complex-valued variable $z \in \mathbb{C}$ [27, 28].

Example 1: Consider the matrix

$$\boldsymbol{A}(z) = \frac{1}{2}\left[\begin{array}{cc} (\frac{1}{4}-\mathrm{j})z + 1 + (\frac{1}{4}+\mathrm{j})z^{-1} & -(\frac{1}{4}+\mathrm{j})z - 1 - (\frac{1}{4}-\mathrm{j})z^{-1} \\ (\frac{1}{4}+\mathrm{j})z + 1 + (\frac{1}{4}-\mathrm{j})z^{-1} & -(\frac{1}{4}-\mathrm{j})z - 1 - (\frac{1}{4}+\mathrm{j})z^{-1} \end{array}\right] , \tag{4}$$

which has the analytic singular values $\sigma_1(z) = \frac{1}{4}z + 1 + \frac{1}{4}z^{-1}$ and $\sigma_2(z) = -\mathrm{j}z + \mathrm{j}z^{-1}$. One option for the analytic left- and right-singular vectors is $\boldsymbol{U}(z) = [1,\;1;\;1\;\;-1]/\sqrt{2}$ and $\boldsymbol{V}(z) = z^{-1}[1,\;1; -1,\;1]/\sqrt{2}$. The evaluation of the analytic singular values on the unit circle, $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$, $m = 1, 2$, is shown in Fig. 1. △

Analyticity is an important property, since analytic functions can be approximated arbitrarily closely by shift and truncation operations [33]. If analyticity is denied, the singular values may only form piece-wise analytic functions, whereby non-differentibilities arise which are difficult to approximate even by polynomial factors of very high order [32]; this has consequences for the implementation of such factors for e.g. precoding and equalisation of MIMO communications channels [7, 11], where high approximation orders for $\boldsymbol{U}(z)$ and $\boldsymbol{V}(z)$ result in their realisation via filter banks [20] requiring very long filters. .

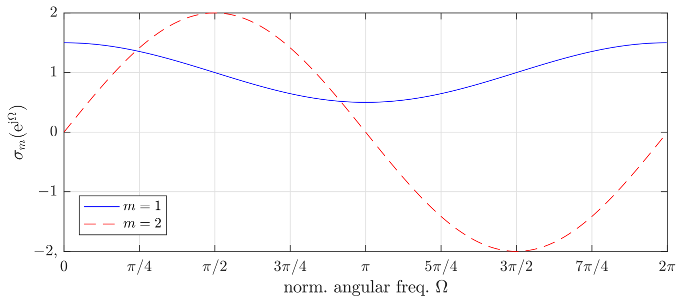

Fig. 1. Example of analytic singular values of matrix $\boldsymbol{A}(z)$ in (4).

C. Ambiguities of the Analytic SVD

To explore the ambiguities of the analytic SVD in (3), consider its equivalent representation using a sum of orthogonal rank-one terms,

$$\boldsymbol{A}(z) = \sum_{m=1}^{M} \mathbf{u}_m(z)\sigma_m(z)\mathbf{v}_m^{\rm P}(z) \; . \tag{5}$$

With the analytic SVD maintaining $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega}) \in \mathbb{R}$ but dropping $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega}) \geq 0$, if $\sigma_m(z)$ is a valid mth analytic singular vector, then so is $-\sigma_m(z)$. We can therefore obtain a different but equally valid mth analytic singular value $\sigma'_m(z)$ and corresponding analytic left- and right-singular vectors $\boldsymbol{u}'_m(z)$ and as $\boldsymbol{v}'_m(z)$ as

$$\sigma'_m(z) = \mathrm{e}^{\mathrm{j}\pi\zeta_m}\sigma_m(z) \tag{6}$$

$$\boldsymbol{u}'_m(z) = \varphi_m(z)\boldsymbol{u}_m(z) \tag{7}$$

$$\boldsymbol{v}'_m(z) = \varphi_m(z)\mathrm{e}^{\mathrm{j}\pi\zeta_m}\boldsymbol{v}_m(z) \; , \tag{8}$$

where $\zeta_m \in \{0, 1\}$ and an allpass $\varphi_m(z)$ are arbitrary. If we exclude the case of identical singular values, such that $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega}) = \sigma_\mu(\mathrm{e}^{\mathrm{j}\Omega}) \forall \Omega$ for some $\mu \neq m$, where further choices arise [27], (6) – (8) expresses the ambiguity of the analytic SVD.

III. Random Perturbation of an Analytic Matrix

In this section, we briefly highlight how a matrix $\mathbf{A}(z)$ may be randomly perturbed e.g. through the process of estimation. We provide an example, where $\mathbf{A}(z)$ is obtained by system identification via the Wiener solution. This estimate $\hat{\boldsymbol{A}}(z)$ will appear randomly perturbed with respect to the $\boldsymbol{A}(z)$, and we state factors that influence the variance of this estimate in case of a system identification process involving Gaussian distributed signals.

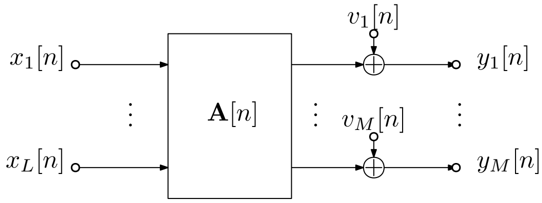

Fig. 2. System matrix $\mathbf{A}[n] \circ\!\!-\!\!\bullet \boldsymbol{A}(z)$ excited by inputs $x_\ell[n]$, $\ell = 1,\ldots,L$ source contributions and providing M outputs $y_m[n]$ corrupted by additive noise $v_m[n]$, $m = 1,\ldots,M$.

A. System Identification Setup

Instead of having direct access to a matrix of transfer functions $\boldsymbol{A}(z)$ with L inputs and M outputs, consider the scenario where we can only measure the input-output behaviour of $\boldsymbol{A}(z)$ as illustrated in Fig. 2. Assume that we have L zero-mean unit-variance uncorrelated sources x_ℓ, $\ell = 1,\ldots,L$, contribute to the measurements at M sensors via a matrix $\mathbf{A}[n] \in \mathbb{C}^{M\times L}$ of impulse responses. For $\mathbf{A}[n]$, the element $a_{m,\ell}[n]$ in the mth row and ℓth column,

$$\mathbf{A}[n] = \begin{bmatrix} a_{1,1}[n] & \cdots & a_{1,L}[n] \\ \vdots & \ddots & \vdots \\ a_{M,1}[n & \cdots & a_{M,L}[n] \end{bmatrix}$$

is the impulse response connecting the ℓth source to the mth sensor. Thus the contribution of the mth sensor from all L sources can be written as

$$y_m[n] = \sum_{\ell=1}^{L} a_{m,\ell}[n] * x_\ell[n] + \mathbf{v}_m[n] \ , \tag{9}$$

where $*$ denotes the convolution operator and $v_m[n]$ is additive spatially and temporally uncorrelated noise of variance σ_v^2. Overall, using vectors $\mathbf{x}[n] = [x_1[n],\ldots,x_L[n]]^{\mathrm{T}}$, $\mathbf{v}[n] = [v_1[n],\ldots,v_M[n]]^{\mathrm{T}}$, and $\mathbf{y}[n] = [y_1[n],\ldots,y_M[n]]^{\mathrm{T}}$, we can write

$$\mathbf{y}[n] = \mathbf{A}[n] * \mathbf{x}[n] + \mathbf{v}[n] \ . \tag{10}$$

The aim is now to identify $\mathbf{A}[n]$ from a finite number of measurements $\mathbf{x}[n]$ and $\mathbf{y}[n]$, $n = 0,\ldots,(N-1)$.

B. Wiener Solution

For general adaptive system identification system of $\mathbf{A}[n]$, we adjust a matrix $\hat{\mathbf{A}}[n]$, excited by the same input $\mathbf{x}[n]$ and generating a separate output $\hat{\mathbf{y}}[n]$. The aim is to adjust $\hat{\mathbf{A}}[n]$ such that the mean square error

$$\begin{aligned} \xi_{\mathrm{MSE}} &= \mathcal{E}\big\{\|\hat{\mathbf{y}}[n] - \mathbf{y}[n]\|_2^2\big\} \\ &= \mathcal{E}\Big\{\|\hat{\mathbf{A}}[n] * \mathbf{x}[n] - \mathbf{y}[n]\|_2^2\Big\} \end{aligned} \tag{11}$$

is minimised, where $\|\cdot\|_2$ is the ℓ_2-norm and $\mathcal{E}\{\cdot\}$ the expectation operator. Thus, we solve the problem

$$\hat{\mathbf{A}}_*[n] = \arg\min_{\hat{\mathbf{A}}[n]} \xi_{\mathrm{MSE}} \ , \tag{12}$$

where $\hat{\mathbf{A}}_*[n]$ is the minimum mean square error or Wiener solution [45, 46].

The Wiener solution involves a data covariance matrix derived from $\mathbf{x}[n]$, and the cross-correlation between $\mathbf{x}[n]$ and $\mathbf{y}[n]$. These correlation terms themselves must be estimated from a finite number N of samples of $\mathbf{x}[n]$ and $\mathbf{y}[n]$, $0 \le n < N$. For an unbiased estimator and Gaussian data, the variance of these estimates depends on both the size N and the true correlation of the signals [47].

C. Variance of Estimation Error

For the mean square error, note that inserting (10) into (11) yields

$$\xi_{\mathrm{MSE}} = \mathcal{E}\Big\{\|(\hat{\mathbf{A}}[n] - \mathbf{A}[n]) * \mathbf{x}[n] + \mathbf{v}[n]\|_2^2\Big\} \ . \tag{13}$$

The identification error

$$\mathbf{E}[n] = \hat{\mathbf{A}}[n] - \mathbf{A}[n] \tag{14}$$

is the perturbation of $\mathbf{A}[n]$ in which we are interested. Further, if $\mathbf{E}[n]$ has a temporal order J, we have

$$\mathbf{E}[n] * \mathbf{x}[n] = [\mathbf{E}[0],\ldots,\mathbf{E}[J]] \cdot \begin{bmatrix} \mathbf{x}[n] \\ \vdots \\ \mathbf{x}[n-J] \end{bmatrix} = \mathbf{E}\mathbf{x}_n \tag{15}$$

Assuming that the input signal, the noise, and with a wider stretch the identification error $\mathbf{E}[n]$ are mutually independent, (13) can be written as

$$\begin{aligned} \xi_{\mathrm{MSE}} &= \mathcal{E}\big\{\mathrm{tr}\big\{\mathbf{E}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{E}^{\mathrm{H}}\big\}\big\} + \mathcal{E}\big\{\mathrm{tr}\big\{\mathbf{v}[n]\mathbf{v}^{\mathrm{H}}[n]\big\}\big\} \\ &= \mathrm{tr}\big\{\mathcal{E}\big\{\mathbf{E}^{\mathrm{H}}\mathbf{E}\big\}\,\mathcal{E}\big\{\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\big\}\big\} + \mathrm{tr}\big\{\mathcal{E}\big\{\mathbf{v}[n]\mathbf{v}^{\mathrm{H}}[n]\big\}\big\} \ , \end{aligned}$$

with $\mathrm{tr}\{\cdot\}$ the trace operator and exploiting $\mathrm{tr}\{\mathbf{ABC}\} = \mathrm{tr}\{\mathbf{CAB}\}$. With $\mathcal{E}\{\mathbf{x}_n\mathbf{x}_n\} = \mathbf{I}_{JM}$ and $\mathcal{E}\big\{\mathbf{v}[n]\mathbf{v}^{\mathrm{H}}[n]\big\} = \sigma_v^2\mathbf{I}_M$, the MSE simplifies to

$$\begin{aligned} \xi_{\mathrm{MSE}} &= \mathcal{E}\big\{\mathrm{tr}\big\{\mathbf{E}^{\mathrm{H}}\mathbf{E}\big\}\big\} + M\sigma_v^2 \\ &= \mathcal{E}\Bigg\{\sum_n \|\mathbf{E}[n]\|_{\mathrm{F}}^2\Bigg\} + M\sigma_v^2 \ . \end{aligned} \tag{16}$$

In (16), the MSE consists of the minimum mean square error components $M\sigma_v^2$ assuming that the estimate $\hat{\mathbf{A}}[n]$ is not curtailed in length to incur truncation errors. Any estimation errors due to a finite sample size N as discussed in Sec. III-B will therefore contribute to an offset of the MSE with respect to the MMSE by the variance of the estimation error of $\mathbf{A}[n]$, i.e. by $\mathcal{E}\{\sum_n \|\mathbf{E}[n]\|_{\mathrm{F}}^2\}$. In the following, we assume that $\mathbf{E}[n]$ itself is random with its elements being Gaussian distributed.

IV. Perturbation of Singular Values

We are now interested in how a random perturbation of $\boldsymbol{A}(z) = \sum_n \mathbf{A}[n]z^{-n}$, or short $\boldsymbol{A}(z) \bullet\!\!-\!\!\circ \mathbf{A}[n]$, by a Gaussian term $\boldsymbol{E}(z) \bullet\!\!-\!\!\circ \mathbf{E}[n]$ impacts on the analytic singular values of the perturbed matrix. We first assess the perturbation in isolated frequency bins, before expanding our reasoning to the entire frequency axis.

A. Bin-Wise Perturbation

Based on the perturbation of $\mathbf{A}[n]$ in (14), we now define the z-domain equivalent as

$$\hat{\boldsymbol{A}}(z) = \boldsymbol{A}(z) + \boldsymbol{E}(z) \; , \tag{17}$$

where $\boldsymbol{E}(z) \bullet\!\!-\!\!\circ \mathbf{E}[n]$ is the estimation error whose variance has been assessed in Sec. III-C. With a bin-wise view, the singular values ς_m of $\hat{\boldsymbol{A}}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega_0}}$ evaluated at a normalised angular frequency Ω_0, i.e. via an SVD of $\hat{\boldsymbol{A}}(\mathrm{e}^{\mathrm{j}\Omega_0})$, are now stochastic quantities that obey some probability distribution. Therefore, in the case that the singular values $|\sigma_m(\mathrm{e}^{\mathrm{j}\Omega_0})|$ of $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})$ have a C-fold algebraic multiplicity, C singular values ς_μ, $\mu = m, \ldots, m+C-1$, are sampled from this distribution. As a result, the singular values ς_m will be distinct with probability one. This is analogous to the case of eigenvalues of a randomly perturbed parahermitian matrices [37]. Thus, the perturbed bin-wise singular values will, with probability one, not possess any non-trivial algebraic multiplicities.

For small singular values, G.W. Stewart in [48] states that they "tend to increase under perturbation, and the increment is proportional to $\sqrt{M}$". If $|\sigma_m(\mathrm{e}^{\mathrm{j}\Omega_0})|$ is the modulus of a small singular value of $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})$, then the square of the corresponding singular value of $\hat{\boldsymbol{A}}(\mathrm{e}^{\mathrm{j}\Omega_0})$ can be expanded as

$$\varsigma_m^2 = (|\sigma_m(\mathrm{e}^{\mathrm{j}\Omega_0})| + \gamma)_m^2 + \eta_m^2 \; , \tag{18}$$

where the terms γ and η are bounded by [49, 50] such that

$$|\gamma_m| \leq \|\mathbf{P}\boldsymbol{E}(\mathrm{e}^{\mathrm{j}\Omega_0})\|_2 \tag{19}$$

$$\inf{}_2\{\mathbf{P}_\perp\boldsymbol{E}(\mathrm{e}^{\mathrm{j}\Omega_0})\} \leq \eta_m \leq \|\mathbf{P}_\perp\boldsymbol{E}(\mathrm{e}^{\mathrm{j}\Omega_0})\|_2 \; . \tag{20}$$

The matrix $\boldsymbol{P}$ is the orthogonal projection into the column space of $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})$, $\boldsymbol{P}_\perp$ its complement, and $\inf_2\{\mathbf{X}\}$ the smallest singular value of $\mathbf{X}$ [48].

While particularly η^2 creates a bias for small $|\sigma_m(\mathrm{e}^{\mathrm{j}\Omega_0})|$, for $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega_0}) = 0$, the lower bound in (20) will be zero, and theoretically $\varsigma_m = 0$ is possible. Note that ς_m^2 is also an eigenvalue of the Hermitian matrix $\mathbf{R} = \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})\boldsymbol{A}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_0})$. In [51–53] the distributions of small eigenvalues and the limits for the condition number of a $\mathbf{R}$ for a Gaussian $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega_0})$ are derived. Since the condition number of $\mathbf{R}$ in e.g. [51] has an upper bound, we must have $\lambda_{\min}(\mathbf{R}) > 0$, and therefore by implication $|\varsigma_m| > 0$. Thus, a randomly perturbed matrix will, with probability one, possess no zero singular values.

The probability distribution of singular values can be difficult to express, see e.g. [54], but we provide two simple examples.

Example 2: Consider the degenerate case of $\mathbf{A}$, $\mathbf{E} \in \mathbb{C}^{1\times 1}$ but $\mathbf{A} = 0$, and $\mathbf{E}$ complex Gaussian distributed. Then we have $\sigma = 0$ but $\varsigma = |\mathbf{E}|$. Thus, note that for any finite perturbation $\mathbf{E}$, indeed $\varsigma > 0$. Further, given the complex Gaussian distribution of $\mathbf{E}$, ς will be Rician distributed. △

Example 3: Recall the matrix $\boldsymbol{A}(z)$ from Example 1. With a bin-wise evaluation at $\Omega_0 = \pi$, Fig. 3 shows the histogram of the two singular values over 10^4 different random perturbations . While $\sigma_2(\mathrm{e}^{\mathrm{j}\Omega_0}) = 0$, it is evident from Fig. 1(a) that the distribution of ς_2 does not include zero. The histograms

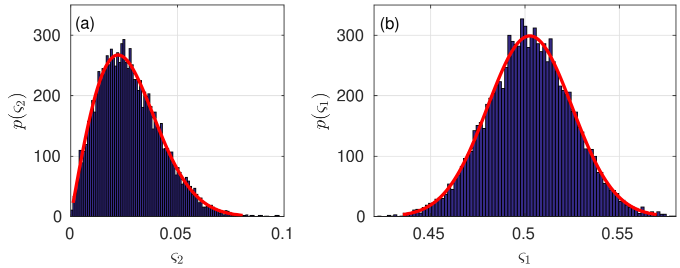

Fig. 3. Histograms of perturbed zero singular values and fits by a Rician distribution (in red) for (a) ς_2 and (b) ς_1.

in Fig. 3 also include Rician fits, which suggest that ς_m, $m = 1, 2$, even for $\varsigma_m \gg 0$ still adhere to this distribution in good approximation. △

Thus, overall the perturbed singular values will, with probability one, have only trivial algebraic multiplicities and satisfy $\varsigma > 0$.

B. Perturbed Analytic Singular Values

In order to assess the singular values of the perturbed system $\hat{\mathbf{A}}(z)$, note that $\hat{\mathbf{A}}(z)$ remains analytic in $z \in \mathbb{C}$ and therefore admits an analytic singular value decomposition [27]. Assessing these analytic singular values $\hat{\sigma}_m(z)$ on the unit circle, at every frequency the result from Sec. IV-A has to hold. As a consequence, the analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ of $\hat{\boldsymbol{A}}(z)$ will

1) neither intersect each other
2) nor cross the zero line

with probability one.

As an interpretation, on the unit circle the analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ will approximate functions that arise from extracting spectrally majorised versions of $\pm\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$. Where they intersect — and this includes the intersections of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ with $-\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ —, a permutation occurs. We illustrate this via the following simple example.

Example 4: The system $\boldsymbol{A}(z)$ from Example 1 is now perturbed by a random term $\boldsymbol{E}(z) \bullet\!\!-\!\!\circ \mathbf{E}[n]$ of the same order as $\boldsymbol{A}(z)$. The perturbation is such that for each element $e_{m,\mu}[n]$, $m, \mu = 1, 2$ and $0 \leq n \leq 2$, of $\mathbf{E}[n]$, $e_{m,\mu}[n] \sim \mathcal{CN}(0, \sigma_e^2)$ with $\sigma_e^2 = 10^{-4}$. The singular values of $\hat{\boldsymbol{A}}(z) = \boldsymbol{A}(z) + \boldsymbol{E}(z)$ are illustrated in Fig. 4, where the analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ are linearly interpolated from a very high resolution bin-wise SVD evaluation. Fig. 4(b) highlights the loss of an algebraic multiplicity of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$, while the previous zero-crossing of $\sigma_2(\mathrm{e}^{\mathrm{j}\Omega})$ at $\Omega = \pi$ no longer occurs for $\hat{\sigma}_2(\mathrm{e}^{\mathrm{j}\Omega})$. △

V. Examples and Simulation

A. System Simulation

In addition to the previous examples, we want to highlight the perturbation of a more complicated matrix $\boldsymbol{A}(z)$ by estimation errors with different variance terms. For this, we determine $\boldsymbol{A}(z)$ from a given ground truth factorisation in (1). The matrices holding the left- and right-singular values

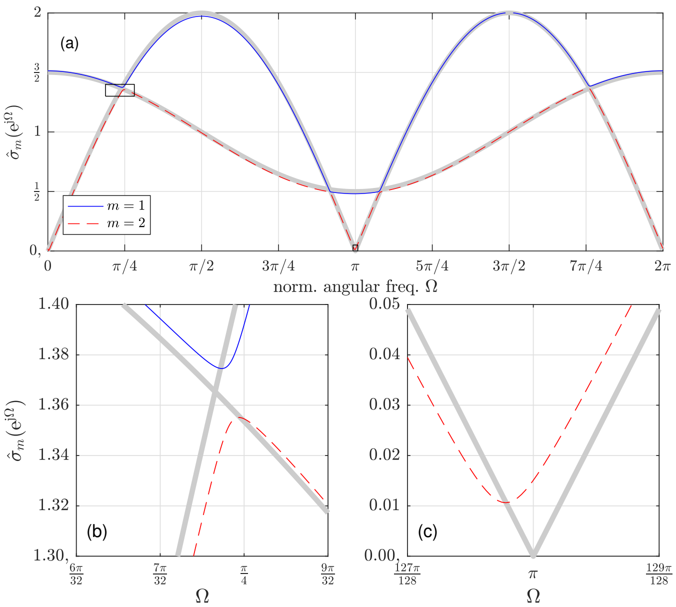

Fig. 4. (a) analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ of the randomly perturbed matrix $\hat{\mathbf{A}}(z)$, with the moduli of the ground truth analytic singular values $|\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})|$ underlaid in grey; zoomed version of this graph near (b) an algebraic multiplicity and (c) a zero crossing of the ground truth singular values.

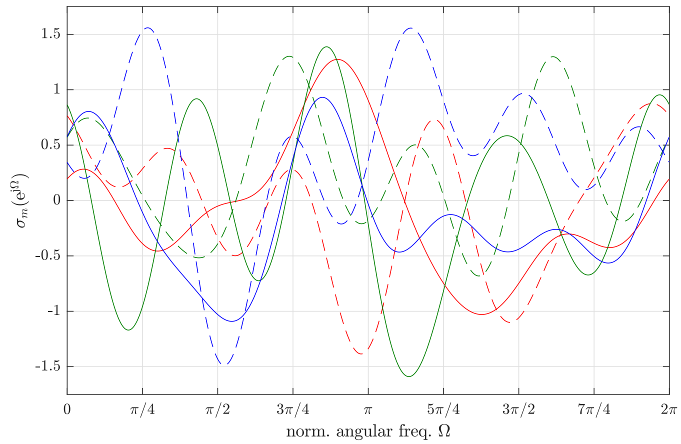

Fig. 5. Ground truth singular values $\sigma_m(z)$, $m = 1, \ldots, 6$ of system $\boldsymbol{A}(z)$.

are obtained form a concatenation of random elementary paraunitary matrices [20]

$$\boldsymbol{Q}(z) = (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{H}}) + \mathbf{w}\mathbf{w}^{\mathrm{H}} z^{-1} \; , \tag{21}$$

where $\mathbf{w}$ is a random unit-norm vector such that $\|\mathbf{w}\|_2 = 1$. With a concatenation of first-order terms drawn from random instantiations of (21), we create $\boldsymbol{U}(z) : \mathbb{C} \rightarrow \mathbb{C}^{6\times 6}$ of order 10, and likewise $\boldsymbol{V}(z)$. The ground truth analytic singular values $\sigma_m(z)$, which for real-valuedness of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ must be parahermitian, are created from random finite impulse responses $s_m[n]$ of length 6 via

$$\sigma_m(z) = s_m(z) + s_m^{\mathrm{P}}(z) \; . \tag{22}$$

It is easy to check that with an arbitrary $s_m[n]$, with (22) we have indeed $\sigma_m(z) = \sigma_m^{\mathrm{P}}(z)$. This generates a matrix $\boldsymbol{A}(z) : \mathbb{C} \rightarrow \mathbb{C}^{6\times 6}$ of order 26. For a particular instance, the resulting singular values $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$, $m = 1, \ldots, 6$, are illustrated in Fig. 5.

B. Variable Perturbation

We now simulate an estimated matrix $\hat{\boldsymbol{A}}(z)$ by randomly perturbing $\boldsymbol{A}(z)$ by a term $\boldsymbol{E}(z)$. The size and order of $\boldsymbol{E}(z)$ are selected to match the parameters of $\boldsymbol{A}(z)$, and similar to Example 4, its coefficients are picked from a complex Gaussian distribution $\mathcal{CN}(0, \sigma_e^2)$, but we will change the variance σ_e^2, which is meant to represent the variance of the estimation error. In order to relate this estimation error to the energy in $\boldsymbol{A}(z)$, we define the normalised variance

$$\sigma^2_{e,\mathrm{norm.}} = \frac{\sum_n \|\mathbf{E}[n]\|_{\mathrm{F}}^2}{\sum_n \|\mathbf{A}[n]\|_{\mathrm{F}}^2} \; . \tag{23}$$

For values of $\sigma^2_{e,\mathrm{norm}} = \{0.3; 10^{-2}; 10^{-4}\}$, the singular values of the perturbed matrices $\hat{\boldsymbol{A}}(z) = \boldsymbol{A}(z) + \boldsymbol{E}(z)$ are provided in Figs. 6, 7, and 8.

It is evident in Figs. 6–8 that with decreasing perturbation, i.e. a smaller $\sigma^2_{e,\mathrm{norm}}$, the singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ on a bin-wise view approach the ground truth. They remain strictly spectrally majorised, i.e. do not intersect, and closer and closer approach zero where $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ for $\sigma^2_{e,\mathrm{norm}} \rightarrow 0$. Yet, the singular valyes will only reach and intersect zero in the limit for $\sigma^2_{e,\mathrm{norm}} = 0$.

VI. Summary and Impact

We have investigated how a random perturbation affects a matrix $\boldsymbol{A}(z)$, which may arise for example when estimating a matrix of transfer functions, and how this perturbation impacts on its analytic singular values. On the unit circle, where the analytic singular values of $\boldsymbol{A}(z)$ can have algebraic multiplicities and zero-crossings, the analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$, $m = 1, \ldots, M$, of $\hat{\boldsymbol{A}}(z)$ will have lost both the no-trivial algebraic multiplicities and the zero-crossings with probability one. Thus, the analytic singular values of the randomly perturbed system with be strictly spectrally majorised and non-negative, even if the ground truth analytic singular values are not. While for a decreasing perturbation, the bin-wise singular values will converge to the ground truth, it is only in the transition to a zero perturbation that the analytic singular values will match the ground truth, and potentially be able to intersect and change sign.

The loss of algebraic multiplicities and zero-crossings has a profound impact on potential algorithms to extract the analytic SVD factors from a perturbed matrix. If we pursue the extraction of $\hat{\sigma}_m(z)$ directly, then this task will become the more difficult the smaller the estimation error is. This is somewhat counter-intuitive, as one would expect an estimate to be more useful the more confident it is. However, since for a decreasing perturbation the functions $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ converge stronger towards non-differentibilities where the ground truth $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ possesses intersections and zero-crossings, the approximation

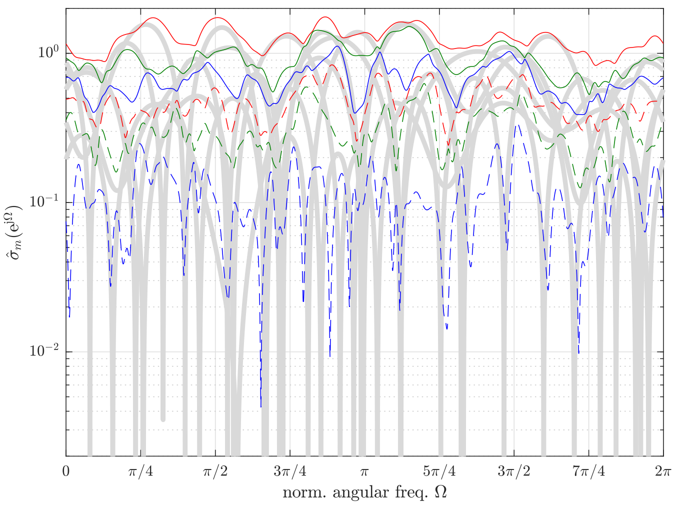

Fig. 6. Analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ of the perturbed system $\hat{\boldsymbol{A}}(z)$, with $\sigma^2_{e,\mathrm{norm}} = 0.3$; the moduli of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ are underlaid in grey.

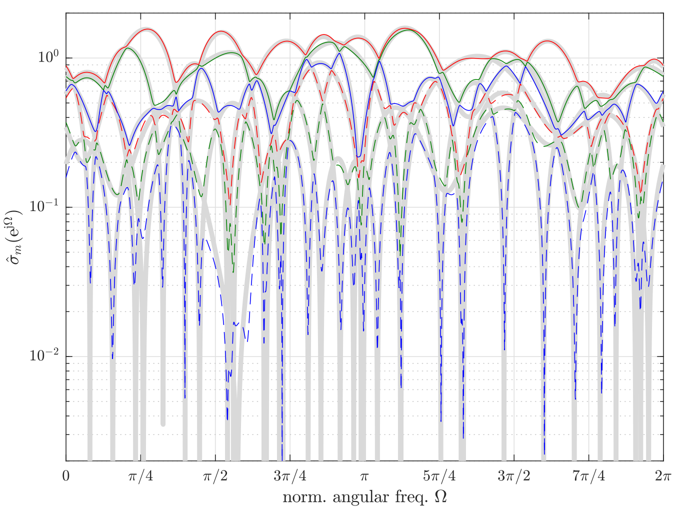

Fig. 7. Analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ of the perturbed system $\hat{\boldsymbol{A}}(z)$, with $\sigma^2_{e,\mathrm{norm}} = 0.01$; the moduli of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ are underlaid in grey.

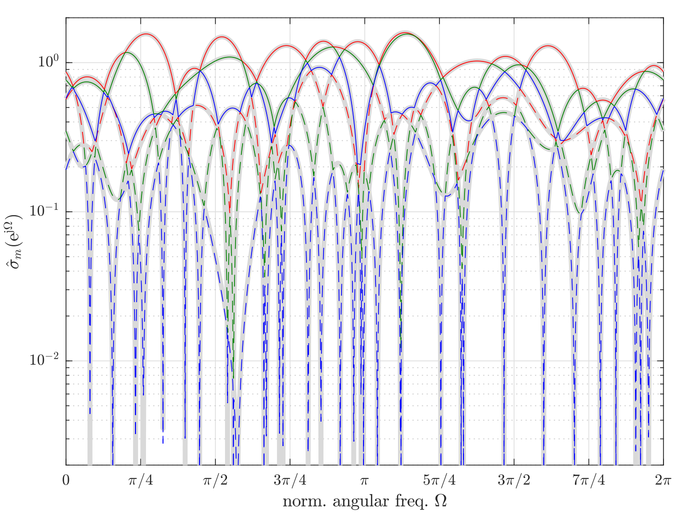

Fig. 8. Analytic singular values $\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega})$ of the perturbed system $\hat{\boldsymbol{A}}(z)$, with $\sigma^2_{e,\mathrm{norm}} = 10^{-4}$; the moduli of $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ are underlaid in grey.

of $\hat{\sigma}_m(z)$ generally requires higher and higher polynomial orders for the SVD factors. Hence, direct approaches such as for the analytic eigenvalue decomposition in [32, 33] will become very costly, and a new methods akin to [55] need to be found that can target the less complex ground truth rather than the exact SVD solution of the perturbed system.

References

[1] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. New York: Academic Press, 1980.

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

[3] E. F. Deprettere, Ed., *SVD and Signal Processing: Algorithms, Applications and Architectures*. NLD: North-Holland Publishing Co., 1989.

[4] R. J. Vaccaro, Ed., *SVD and Signal Processing II: Algorithms, Analysis and Applications*. Elsevier Science Pub. Co., 1991.

[5] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86–105, Sep. 2007.

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

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

[8] W. Al-Hanafy, A. P. Millar, C. H. Ta, and S. Weiss, "Broadband SVD and non-linear precoding applied to broadband MIMO channels," in *42nd Asilomar Conf. SSC*, Pacific Grove, CA, USA, Oct. 2008, pp. 2053–2057.

[9] H. Zamiri-Jafarian and M. Rajabzadeh, "A polynomial matrix svd approach for time domain broadband beamforming in mimo-ofdm systems," in *IEEE Vehicular Technology Conference*, spring 2008, pp. 802–806.

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

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

[12] X. Mestre and D. Gregoratti, "A parallel processing approach to filterbank multicarrier MIMO transmission under strong frequency selectivity," in *IEEE ICASSP*, May 2014, pp. 8078–8082.

[13] A. Ahrens, A. Sandmann, S. Lochmann, and Z. Wang, "Decomposition of optical mimo systems using polynomial matrix factorization," in *2nd IET Int. Conf. Intelligent Signal Processing*, London, UK, 2015, pp. 1–6.

[14] A. I. Pérez-Neira, M. Caus, R. Zakaria, D. L. Ruyet, E. Kofidis, M. Haardt, X. Mestre, and Y. Cheng, "MIMO signal processing in offset-QAM based filter bank multicarrier systems," *IEEE Trans. Signal Processing*, vol. 64, no. 21, pp. 5733–5762, Nov. 2016.

[15] S. Weiss, S. J. Schlecht, O. Das, and E. De Sena, "Polynomial Procrustes problem: Paraunitary approximation of matrices of analytic functions," in *31st European Signal Processing Conference*, Helsinki, Finland, Sep. 2023, pp. 1629–1633.

[16] F. Khattak, I. Proudler, and S. Weiss, "Generalized polynomial power method," in *International Conference on Sensor Signal Processing for Defence*, Edinburgh, Scotland, Sep. 2023, accepted.

[17] F. Khattak, I. Proudler, J. McWhirter, and S. Weiss, "Generalised sequential matrix diagonalisation for the SVD of polynomial matrices," in *International Conference on Defence for Signal Processing*, Edinburgh, Scotland, Sep. 2023, submitted.

[18] M.A. Bakhit, F.A. Khattak, I.K. Proudler, S. Weiss, and G.W. Rice, "Compact order polynomial singular value decomposition of a matrix of analytic functions," in *9th IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Los Sueños, Costa Rica, Dec. 2023.

[19] F. A. Khattak, M. Bakhit, I. K. Proudler, and S. Weiss, "Extraction of analytic singular values of a polynomial matrix," in *32nd European Signal Processing Conference*, Lyon, France, Aug. 2024, pp. 1297–1301.

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

[21] J. Foster, J.G. McWhirter, and J.A. Chambers, "Limiting the order of polynomial matrices within the SBR2 algorithm," in *IMA Int. Conf. on Mathematics in Signal Processing*, Cirencester, UK, Dec. 2006.

[22] C.H. Ta and S. Weiss, "Shortening the order of paraunitary matrices in SBR2 algorithm," in *6th Int. Conf. Information, Communications & Signal Processing*, Singapore, pp. 1–5, Dec. 2007.

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

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

[25] J. G. McWhirter, "An algorithm for polynomial matrix SVD based on generalised Kogbetliantz transformations," in *18th EUSIPCO*, Aalborg, Denmark, Aug. 2010, pp. 457–461.

[26] J. G. McWhirter and Z. Wang, "A novel insight to the SBR2 algorithm for diagonalising para-hermitian matrices," in *11th IMA Conf. Maths in Signal Processing*, Birmingham, UK, December 2016.

[27] 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 Processing*, vol. 72, pp. 2260–2275, 2024.

[28] 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.

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

[30] S. Weiss, I.K. Proudler, F.K. Coutts, and J. Pestana, "Iterative approximation of analytic eigenvalues of a parahermitian matrix EVD," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, May 2019.

[31] S. Weiss, I.K. Proudler, F.K. Coutts, and J. Deeks, "Extraction of analytic eigenvectors from a parahermitian matrix," in *Int. Conf. Sensor Signal Processing or Defence*, Edinburgh, UK, Sep. 2020.

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

[33] S. Weiss, I.K. Proudler, F.K. Coutts, and F. Khattak, "Eigenvalue decomposition of a parahermitian matrix: extraction of analytic eigenvectors," *IEEE Trans. Signal Process.*, vol. 71, pp. 1642–1656, Apr. 2023.

[34] F. A. Khattak, I. K. Proudler, and S. Weiss, "Extension of power method to para-Hermitian matrices: polynomial power method," in *31st European Signal Processing Conference*, Helsinki, Finland, pp. 1564–1568, Sep. 2023.

[35] F.A. Khattak, I.K. Proudler, and S. Weiss, "Scalable extraction of analytic eigenvalues from a parahermitian matrix," in *32nd European Signal Processing Conference*, Lyon, France, Aug. 2024, pp. 1317–1321.

[36] F.A. Khattak, I.K. Proudler, and S. Weiss, "Scalable analytic eigenvalue extraction algorithm," *IEEE Access*, 2024, submitted.

[37] F. A. Khattak, S. Weiss, I. K. Proudler, and J. G. McWhirter, "Space-time covariance matrix estimation: Loss of algebraic multiplicities of eigenvalues," in *56th Asilomar Conf. SSC*, Pacific Grove, CA, Oct. 2022.

[38] F. A. Khattak, I. K. Proudler, and S. Weiss, "Enhanced space-time covariance estimation based on a system identification approach," in *SSPD*, London, UK, Sep. 2022.

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

[40] H. Scheuermann and H. Göckler, "A comprehensive survey of digital transmultiplexing methods," *Proc. IEEE*, vol. 69, no. 11, pp. 1419–1450, 1981.

[41] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE*, vol. 78, no. 1, pp. 56–93, Jan. 1990.

[42] A. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal Transmultiplexers in Communication: a Review," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 979–995, Apr. 1998.

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

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

[45] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, New York: Prentice Hall, 1985.

[46] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs: Prentice Hall, 1991.

[47] C. Delaosa, J. Pestana, N. J. Goddard, S. Somasundaram, and S. Weiss, "Sample space-time covariance matrix estimation," in *IEEE ICASSP*, Brighton, UK, May 2019, pp. 8033–8037.

[48] G. Stewart, "Perturbation theory for the singular value decomposition," Institute for Advanced Computer Studies, Univ. of Maryland, Tech. Rep. UMIACS-TR-90-124, Oct. 1990.

[49] ——, "A note on the perturbation of singular values," *Linear Algebra Appl.*, vol. 28, pp. 213–216, 1979.

[50] ——, "A second order perturbation expansion for small singular values," *Linear Algebra Appl.*, vol. 56, pp. 231–235, 1984.

[51] J.-M. Azaïs and M. Wschebor, "Upper and lower bounds for the tails of the distribution of the condition number of a gaussian matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 426–440, 2004.

[52] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Eigenvalues and condition numbers of complex random matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 441–456, 2004.

[53] W. Anderson and M. T. Wells, "The exact distribution of the condition number of a gaussian matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1125–1130, 2010.

[54] J. Shen, "On the singular values of Gaussian random matrices," *Linear Algebra Appl.*, vol. 326, no. 1, pp. 1–14, 2001.

[55] S.J. Schlecht and S. Weiss, "Reconstructing analytic dinosaurs: Polynomial eigenvalue decomposition for eigenvalues with unmajorised ground truth," in *32nd European Signal Processing Conference*, Lyon, France, Aug. 2024, pp. 1287–1291.