Generalized Polynomial Power Method

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

***Abstract*—The polynomial power method repeatedly multiplies a polynomial vector by a para-Hermitian matrix containing spectrally majorised eigenvalue to estimate the dominant eigenvector corresponding the dominant eigenvalue. To limit the order of the resulting vector, truncation is performed in each iteration. This paper extends the polynomial power method from para-Hermitian matrices to a general polynomial matrix for determining its dominant left- and right-singular vectors and the corresponding singular value. The proposed extension assumes that the dominant singular is positive on the unit circle. The resulting algorithm is compared with a state-of-the-art PSVD algorithm and provides better accuracy with reduced computation time and lower approximation orders for the decomposition.**

I. Introduction

The algebra of polynomial matrices has proven to be useful in solving a variety of problems related to broadband sensor arrays, particularly through the application of two main operations: the polynomial eigenvalue decomposition (PEVD) [1–8] and the polynomial singular value decomposition (PSVD) [9–11]. The PEVD is restricted to para-Hermitian matrices only, where the matrix is equal to its transpose-conjugate time-reversed version [12]. In contrast, the PSVD can be applied to any polynomial matrix. Thus, PSVD algorithms find applications in a variety of problems such as MIMO design [13, 14], paraunitary filter design [7], or beamforming [15].

Typically, the PSVD is calculated using two PEVDs [10] or via a polynomial QR decomposition (PQRD) [9], both of which are computationally expensive. However, a dedicated PSVD algorithm exists which is the generalization of the second-order sequential best rotation (SBR2)[4], and it exploits the Kogbetliantz transformation[16]. This method iteratively transfers the energy onto the diagonal. While [16] performs only an approximate diagonalization, its performance is still better than that achieved via PEVD or PQRD approaches. An SVD with analytic factors exists [17, 18], such that there are unique singular values that are real on the unit circle, and left- and right-singular vectors that share a common ambiguity w.r.t. arbitrary allpass functions. The above algorithm ignore this coupled ambiguity, and hence typically yield complex-valued approximations of the singular values.

In order to overcome the deficiencies of the above PSVD algorithms, in this document, we extend the polynomial power method [19] from a para-Hermitian matrix to a general polynomial matrix for the computation of the dominant left- and right-singular vector and the singular value. The polynomial power method is an extension of the ordinary power method [20] to para-Hermitian matrices where a polynomial vector is repeatedly multiplied by a para-Hermitian matrix, and the resulting vector converges to the dominant eigenvector provided that the matrix is spectrally majorised. Similar to the power method, the generalised polynomial power method can be coupled with deflation in order to compute an entire PSVD.

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

II. Polynomial Singular Value Decomposition

For an analytic, non-multiplexed $\boldsymbol{A}(z) \in \mathbb{C}^{M\times N}, M \geq N$, the analytic SVD exists [18]

$$\boldsymbol{A}(z) = \boldsymbol{U}(z)\boldsymbol{\Sigma}(z)\boldsymbol{V}^{\mathrm{P}}(z)\,, \tag{1}$$

such that $\boldsymbol{\Sigma}(z) = \mathrm{diag}\{\sigma_1(z), \ldots, \sigma_N(z)\} \in \mathbb{C}^{M\times N}$ contains the analytic singular values and the matrices $\boldsymbol{U}(z) \in \mathbb{C}^{M\times M}$, $\boldsymbol{V}(z) \in \mathbb{C}^{N\times N}$ are paraunitary i.e. $\boldsymbol{U}(z)\boldsymbol{U}^{\mathrm{P}}(z) = \mathbf{I}, \boldsymbol{V}(z)\boldsymbol{V}^{\mathrm{P}}(z) = \mathbf{I}$, and contains the left- and right analytic singular vectors, respectively. Note that the parahermitian operation, $\{\cdot\}^{\mathrm{P}}$, involves a Hermitian transposition and time reversal of its arguments, such that e.g. $\boldsymbol{U}^{\mathrm{P}}(z) = (\boldsymbol{U}(1/z^*)^{\mathrm{H}}$. Unlike singular values of constant matrices, which must be real and positive semi-definite [20], the analytic singular values evaluated on unit circle for $z = \mathrm{e}^{\mathrm{j}\Omega}$ must be permitted to take on negative values. This is similarly known for matrices that depend analytically on a continuous, real parameter on some interval [21, 22].

Generally, the analytic singular values of a matrix $\boldsymbol{A}(z)$ may intersect. However, if $\boldsymbol{A}(z)$ is estimated from finite data e.g. via system identification [23], it will have spectrally majorised singular values

$$\sigma_i(\mathrm{e}^{\mathrm{j}\Omega}) \geq \sigma_{i+1}(\mathrm{e}^{\mathrm{j}\Omega}) \;\forall\; \Omega,\; i = 1, \ldots, (N-1) \tag{2}$$

with probability one [23]. We therefore assume the property (2) to hold for the remainder of this paper.

III. Polynomial Power Method for Parahermitian Matrix

The underlying idea is to extend the polynomial power method, proposed in [19] for para-Hermitian matrices, to general polynomial matrices for the extraction of the dominant singular vectors and singular value. Thus this section provides a brief summary of the polynomial power method. We denote a para-Hermitian matrix by $\boldsymbol{R}(z)$, which must therefore must satisfy $\boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}(z)$ [12].

This is a peer reviewed, accepted author manuscript of the following research contribution: Khattak, F. A., Proudler, I. K., & Weiss, S. (2023). Generalised polynomial power method. Paper presented at *12th International Conference on Sensor Signal Processing for Defence*, Edinburgh, United Kingdom.

A. Overall Rationale

This method is an extension of the ordinary power iteration [20] from Hermitian matrices to para-Hermitian matrices where an arbitrary polynomial vector $\boldsymbol{x}^{(0)}(z)$ is repeatedly multiplied with a para-Hermitian matrix $\boldsymbol{R}(z)$ to obtain a sequence of polynomial vectors. After k iterations, we have

$$\boldsymbol{x}^{(k)}(z) = \boldsymbol{R}(z)\boldsymbol{x}^{(k-1)}(z) = \boldsymbol{R}^k(z)\boldsymbol{x}^{(0)}(z)\,. \tag{3}$$

In each iteration, $\boldsymbol{x}^{(k)}(z)$ has to be normalized, such that $\boldsymbol{x}^{(k)}_{\text{norm}}(z)$ satisfies $\boldsymbol{x}^{(k),\text{P}}_{\text{norm}}(z)\boldsymbol{x}^{(k)}_{\text{norm}}(z) = 1$. This normalization can be performed on the unit-circle, due to analyticity, with ease. To limit the order growth of $\boldsymbol{x}^{(k)}(z)$, truncation can be applied. This can be achieved by either limiting the order to the estimated support of the eigenvector obtained from [24] by shifted-truncation [25], or removing any trailing coefficients below a small threshold. The iterations are stopped once a suitable defined difference between consecutive polynomial vectors falls below some threshold. The overall analysis that connects the polynomial power method with the ordinary power method is given next.

B. Polynomial Power Method Analysis

The initial $\boldsymbol{x}^{(0)}(z)$ can be represented as a linear combination of the eigenvectors of $\boldsymbol{R}(z)$

$$\begin{aligned}\boldsymbol{x}^{(0)}(z) &= \boldsymbol{Q}(z)\boldsymbol{c}(z)\\ &= c_1(z)\boldsymbol{q}_1(z) + \cdots + c_M(z)\boldsymbol{q}_M(z)\,,\end{aligned} \tag{4}$$

where $\boldsymbol{q}_m(z)$, $m = 1\ldots, M$, is the mth analytic eigenvector of $\boldsymbol{R}(z)$ and $\boldsymbol{c}(z) = [c_1(z), \ldots, c_M]^{\text{T}} \in \mathbb{C}^M$ is a vector of analytic weighting factors. Due to analyticity, we can restrict the analysis to the unit circle. Therefore, z can be substituted with $\text{e}^{\text{j}\Omega}$ to evaluate and iterate on the unit circle. Combining (4) and (3) with the fact that $\boldsymbol{R}(z)$ has spectrally-majorised eigenvalues due to being estimated from finite data, we have

$$\begin{aligned}\boldsymbol{x}^{(k)}(\text{e}^{\text{j}\Omega}) = \lambda_1^k(\text{e}^{\text{j}\Omega})\Bigg[&c_1(\text{e}^{\text{j}\Omega})\boldsymbol{q}_1(\text{e}^{\text{j}\Omega})\\ &+ \sum_{m=2}^{M} c_m(\text{e}^{\text{j}\Omega})\left(\frac{\lambda_m(\text{e}^{\text{j}\Omega})}{\lambda_1(\text{e}^{\text{j}\Omega})}\right)^k \boldsymbol{q}_m(\text{e}^{\text{j}\Omega})\Bigg]\,.\end{aligned} \tag{5}$$

The summation term will converge towards zero for $k \to \infty$. This permits us to re-write (5) as

$$\boldsymbol{x}^{(k)}(\text{e}^{\text{j}\Omega}) = \lim_{k\to\infty} \lambda_1^k(\text{e}^{\text{j}\Omega})c_1(\text{e}^{\text{j}\Omega})\boldsymbol{q}_1(\text{e}^{\text{j}\Omega})\,. \tag{6}$$

The term $\boldsymbol{x}^{(k)}(\text{e}^{\text{j}\Omega})$ in (6) is normalized in each iteration such that it has unit norm on the unit circle. This normalisation is carried out in the DFT domain. If $c_1(\text{e}^{\text{j}\Omega})$ possesses any spectral nulls for some Ω, the resulting division by zero in the normalisation process can be avoided by regularization [19]. For a sufficiently large k, the normalized vectors become

$$\boldsymbol{x}^{(k)}_{\text{norm}}(\text{e}^{\text{j}\Omega}) = \hat{\boldsymbol{q}}_1(z) = g_1(\text{e}^{\text{j}\Omega})\boldsymbol{q}_1(\text{e}^{\text{j}\Omega})\,, \tag{7}$$

where $g_1(\text{e}^{\text{j}\Omega}) = c_1(\text{e}^{\text{j}\Omega})/|c_1(\text{e}^{\text{j}\Omega})|$ is an allpass filter. This allpass filter generalises the phase ambiguity of the eigenvectors of a standard matrix.

Algorithm 1: PPM Algorithm [19]

Input: $\boldsymbol{R}(z), \epsilon, k_{\max}$
Output: $\hat{\mathbf{q}}_1(z), \hat{\lambda}_1(z)$
$\boldsymbol{x}^{(0)}(z) \in \mathbb{C}^M$, $k \leftarrow 0$, $\gamma = \infty$;
$\tilde{\boldsymbol{x}}^{(0)}_{\text{norm}}(z) \leftarrow$ normalise & order limit $\boldsymbol{x}^{(0)}(z)$;
while $\gamma > \epsilon$ & $k < k_{max}$ **do**
  $k \leftarrow k + 1$;
  $\boldsymbol{x}^{(k)}(z) \leftarrow \boldsymbol{R}(z)\tilde{\boldsymbol{x}}^{(k-1)}_{\text{norm}}(z)$;
  $\boldsymbol{x}^{(k)}_{\text{norm}}(z) \leftarrow$ normalisation $\boldsymbol{x}^{(k)}(z)$;
  $\tilde{\boldsymbol{x}}^{(k)}_{\text{norm}}(z) \leftarrow$ order limitation of $\boldsymbol{x}^{(k)}_{\text{norm}}(z)$;
  update γ
end
$\hat{\boldsymbol{q}}_1(z) = \tilde{\boldsymbol{x}}^{(k)}_{\text{norm}}(z)$;
$\hat{\lambda}_1(z) = \tilde{\boldsymbol{x}}^{(k),\text{P}}_{\text{norm}}(z)\boldsymbol{R}(z)\tilde{\boldsymbol{x}}^{(k)}_{\text{norm}}(z)$;

The stopping criterion for the polynomial power method is to measure the overall deviation

$$\gamma = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\alpha(\Omega)|^2 \text{d}\Omega \tag{8}$$

of the Hermitan angle $\alpha(\Omega)$, defined as

$$\begin{aligned}\alpha(\Omega) = \angle\{\boldsymbol{x}^{(k)}_{\text{norm}}(\text{e}^{\text{j}\Omega}), \boldsymbol{x}^{(k-1)}_{\text{norm}}(\text{e}^{\text{j}\Omega})\} =\\ \text{acos}\left(\frac{|\boldsymbol{x}^{(k),\text{H}}_{\text{norm}}(\text{e}^{\text{j}\Omega})\boldsymbol{x}^{(k-1)}_{\text{norm}}(\text{e}^{\text{j}\Omega})|}{\|\boldsymbol{x}^{(k)}_{\text{norm}}(\text{e}^{\text{j}\Omega})\|_2 \cdot \|\boldsymbol{x}^{(k-1)}_{\text{norm}}(\text{e}^{\text{j}\Omega})\|_2}\right)\,.\end{aligned}$$

Once γ falls below a threshold ϵ, the process can be terminated. The corresponding dominant eigenvalue can be extracted as

$$\hat{\lambda}_1(z) = \tilde{\boldsymbol{x}}^{(k),\text{P}}_{\text{norm}}(z)\boldsymbol{R}(z)\tilde{\boldsymbol{x}}^{(k)}_{\text{norm}}(z)\,. \tag{9}$$

This constitutes the polynomial version of the power method for spectrally majorised matrices, with the overall procedure summarised in Algorithm 1. For further details, please refer to [19].

IV. Ordinary Generalised Power Method

We know that the conventional reduced SVD of $\mathbf{A} \in \mathbb{C}^{M\times N}$ with $M \geq N$, given as $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\text{H}}$ with $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N] \in \mathbb{C}^{M\times N}$, $\mathbf{\Sigma} \in \mathbb{R}^{N\times N}$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_N] \in \mathbb{C}^{N\times N}$, can be obtained through the ordinary power method [20]. In order to determine the right-singular vectors, the power method can be applied to $\mathbf{A}^{\text{H}}\mathbf{A} \in \mathbb{C}^{N\times N}$ as its eigenvectors are in fact the right-singular vectors of $\mathbf{A}$. After $\hat{\mathbf{v}}_i = \text{e}^{\text{j}\phi}\mathbf{v}_i$, $i = 1, \ldots, N$ has been found, the singular values and left-singular vectors can be obtained via

$$\sigma_i = ||\mathbf{A}\hat{\mathbf{v}}_i||_2, \Rightarrow \hat{\mathbf{u}}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i} = \text{e}^{\text{j}\phi}\mathbf{u}_i,\ i = 1, \ldots, M\,, \tag{10}$$

where $\text{e}^{\text{j}\phi}$ is an arbitrary phase shift. Note that the phase ambiguity of the left- and right-singular vectors is coupled.

Alternatively, the left-singular vector can be determined by applying the power method to $\mathbf{A}\mathbf{A}^{\text{H}}$ and then the singular value can be computed later. Since left- and right-singular

vetors are determined independently, their phase ambiguities are no longer coupled. Hence, if $\hat{\mathbf{u}}_1 = \mathrm{e}^{\mathrm{j}\alpha}\mathbf{u}_1$ and $\hat{\mathbf{v}}_1 = \mathrm{e}^{\mathrm{j}\phi}\mathbf{v}_1$, the resulting estimated singular value will be $\hat{\sigma}_1 = \hat{\mathbf{u}}_i^{\mathrm{H}}\mathbf{A}\hat{\mathbf{v}}_1 = \mathrm{e}^{-\mathrm{j}\alpha}\sigma_1\mathrm{e}^{\mathrm{j}\phi}$ i.e. it will not in general be real-valued. However, real-valuedness, and therefore phase coupling of the left- and right-singular vectors, can be achieved buy adjusting the phase of $\hat{\sigma}_i$. Such a procedure may not be possible in the case of polynomial matrices as will become clear in the following section.

This shows that the power iteration is not restricted to Hermitian matrices, but can indeed be applied to any matrix. This motivates us to combine the above concept with the already established polynomial power iterationsw, however. The aim is to drop the restrictition to para-Hermitian matrices, such that the dominant singular vectors may be computed in first instance, with the option of later performing a full SVD of a polynomial matrix through deflation.

V. Generalized Polynomial Power Method

We now aim to extend the polynomial power method reviewed in Sec. III to generalise the SVD approach summarised in Sec. IV to the case of a polynomial matrix $\boldsymbol{A}(z) : \mathbb{C} \rightarrow \mathbb{C}^{M\times N}$.

A. Polynomial Iterations Analysis

For an initial $\boldsymbol{x}^{(0)}(z) = \boldsymbol{V}(z)\boldsymbol{c}(z)$, the polynomial iteration can be applied to a para-Hermitian matrix $\boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z)$ where after k iterations, with z substituted by $\mathrm{e}^{\mathrm{j}\Omega}$, we obtain

$$\begin{aligned}\boldsymbol{x}^{(k)}(\mathrm{e}^{\mathrm{j}\Omega}) &= \boldsymbol{A}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{x}^{(k-1)}(\mathrm{e}^{\mathrm{j}\Omega}) \\ &= \sum_{n=1}^{N} \boldsymbol{v}_n(\mathrm{e}^{\mathrm{j}\Omega})\sigma_n^{2k}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{v}_m^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{V}(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{c}(\mathrm{e}^{\mathrm{j}\Omega}) \,,\end{aligned} \tag{11}$$

which can be re-written as

$$\begin{aligned}\boldsymbol{x}^{(k)}(\mathrm{e}^{\mathrm{j}\Omega}) = \sigma_1^{2k}(\mathrm{e}^{\mathrm{j}\Omega})\Bigg[&c_1(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{v}_1(\mathrm{e}^{\mathrm{j}\Omega}) \\ &+ \sum_{m=2}^{M} c_m(\mathrm{e}^{\mathrm{j}\Omega})\left(\frac{\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})}{\sigma_1(\mathrm{e}^{\mathrm{j}\Omega})}\right)^{2k}\boldsymbol{v}_m(\mathrm{e}^{\mathrm{j}\Omega})\Bigg] \,.\end{aligned} \tag{12}$$

Since the singular values of $\boldsymbol{A}(z)$ are spectrally-majorised, so are the eigenvalues of the para-Hermitian matrix $\boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z)$ i.e. $\sigma_n^2(\mathrm{e}^{\mathrm{j}\Omega}) \geq \sigma_{n+1}^2(\mathrm{e}^{\mathrm{j}\Omega})$ $n = 1, \ldots, N-1$. Hence $\boldsymbol{x}^{(k)}(\mathrm{e}^{\mathrm{j}\Omega})$ converges to a scaled version of $\boldsymbol{v}_1(\mathrm{e}^{\mathrm{j}\Omega})$ for sufficiently large k similar to (6). Similarly, after normalization, we have

$$\boldsymbol{x}_{\mathrm{norm}}^{(k)}(\mathrm{e}^{\mathrm{j}\Omega}) = \hat{\boldsymbol{v}}_1(\mathrm{e}^{\mathrm{j}\Omega}) = g_1(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{v}_1(\mathrm{e}^{\mathrm{j}\Omega}) \;\forall\; \Omega \,, \tag{13}$$

where $g_1(z)$ is an allpass filter. Similar to the power method for para-Hermitian matrix, this generalised approach also includes truncation and normalization in each iteration which can be found in the original algorithm reported in [19]. Similarly, the problem of singularities in $c_1(\mathrm{e}^{\mathrm{j}\Omega})$ can be handled either through regularization a or modification to the initialization if a spectral zero is encountered.

The estimation of the dominant singular value and the corresponding left-singular vector is not straightforward and needs careful consideration. Assuming, we follow the first method described in Sec. IV in (10), which is to determine the singular value and then the left-singular vector. For this method, the frequency dependent version for extracting the dominant singular value will be

$$\hat{\sigma}_m(\mathrm{e}^{\mathrm{j}\Omega}) = ||\boldsymbol{A}_m(\mathrm{e}^{\mathrm{j}\Omega})\hat{\boldsymbol{v}}_m(\mathrm{e}^{\mathrm{j}\Omega})||_2, \; m = 1, \ldots, N \,, \tag{14}$$

which forces $\sigma_m(\mathrm{e}^{\mathrm{j}\Omega})$ to be positive $\forall\; \Omega$ due to the norm operator whereas the theory behind the analytic decomposition existence shows that the singular value can be negative on the unit-circle [21, 22]. Forcing the singular values to be positive violates this condition, thus with this method, the obtained decomposition might differ from the decomposition given in [18]. Alternatively, if the matrix $\boldsymbol{A}(z)$ is known to be positive semi-definite, the singular values are guaranteed to be real and positive and so this method gives the correct decomposition. Once the singular value is obtained with an acceptable accuracy, which we discuss further below, via (14) the dominant left-singular vector can be obtained as

$$\hat{\boldsymbol{u}}_1(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\Omega})\hat{\boldsymbol{v}}_1(\mathrm{e}^{\mathrm{j}\Omega})/\hat{\sigma}_1(\mathrm{e}^{\mathrm{j}\Omega}) = g_1(\mathrm{e}^{\mathrm{j}\Omega})\boldsymbol{u}_1(\mathrm{e}^{\mathrm{j}\Omega}) \,. \tag{15}$$

The allpass factor $g_1(\mathrm{e}^{\mathrm{j}\Omega})$ is the same as that of the right-singular vector, such that their ambiguities are coupled. This coupling results in the singular value being real-valued on the unit-circle. Both (14) and (15) can be implemented in DFT domain. Adjusting the size of this DFT is discussed further below.

The second method, described in Sec. IV, determines the left-singular vector by applying the polynomial power method to $\boldsymbol{A}(z)\boldsymbol{A}^{\mathrm{P}}(z)$ and then computes the singular value as $\hat{\sigma}_1(z) = \hat{\boldsymbol{u}}_1^{\mathrm{P}}(z)\boldsymbol{A}(z)\hat{\boldsymbol{v}}_1(z)$. This method does not impose the condition of singular value being positive on the unit circle, and so it can allow the analytic decomposition given in (1) to be achievable for any $\boldsymbol{A}(z)$. However, to retain real-valuedness for the singular values on the unit circle, the left- and right-singular vector have to have a common allpass factor. Thus if both the left- and right-singular vectors are independently extracted by applying the polynomial power method to $\boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z)$ and $\boldsymbol{A}(z)\boldsymbol{A}^{\mathrm{P}}(z)$, respectively, the allpass factor in $\hat{\boldsymbol{u}}_1(z)$ and $\hat{\boldsymbol{v}}_1(z)$ will, in general, not be be coupled. Hence, the second method may not be desireable to be used unless a common phase shift can be found. Another reason for avoiding this method is invoking two polynomial power methods is computationally more expensive compared to the first approach.

To optimize the computational efficiency of the first method, we estimate the right singular vector using the polynomial power method in case of $M \geq N$. For the case $M \leq N$, all of the above approaches can used instead to factorise $\boldsymbol{A}^{\mathrm{P}}(z)$.

B. Sufficient DFT Size

1) Dominant Singular Value: Once $\boldsymbol{v}_1(z)$ is determined with Algorithm 1, the singular value can be determined via (14) in the DFT domain. To determine a sufficient DFT size,

time-domain aliasing can be utilized [2]. Thus (14) can be evaluated at increasing DFT sizes until

$$\xi_{\hat{\sigma}} = \sum_{\tau} \frac{|\hat{\sigma}_1^{(K)}[\tau] - \hat{\sigma}_1^{(K/2)}[\tau]|^2}{|\hat{\sigma}_1^{(K)}[\tau]|^2} \,, \tag{16}$$

where $\hat{\sigma}_1^{(K)}[\tau]$ represents the time-domain equivalent of (14) obtained with a $K-$point inverse DFT, until $\xi_{\hat{\sigma}}$ falls below a certain low threshold ε_1. A small value of $\xi_{\hat{\sigma}}$ indicates that $K/2$ can be considered sufficient for approximate the dominant singular. If $\hat{\boldsymbol{v}}_i(z)$ is extracted with satisfactory accuracy, a DFT size of $K = \mathcal{O}\{\boldsymbol{A}(z)\hat{\boldsymbol{v}}_i(z)\} + 1$ should generally suffice, where $\mathcal{O}\{\cdot\}$ measures the polynomial order of its argument.

2) Left Singular Vector: Similarly, to determine a sufficient DFT size for (15), time-domain aliasing may be captured via the error w.r.t. normality in the time-domain as

$$\xi_u = \sum_{\tau} |\hat{\mathbf{u}}_1^{\mathrm{H}}[-\tau] * \hat{\mathbf{u}}_1[\tau] - \delta[\tau]|_2^2, \quad \tau \in \mathbb{Z} \,. \tag{17}$$

A similar criterion has been utilised for the DFT size in [3]. There, it is shown as a necessary criterion; while sufficiency has not been proven, in practise is has generally been shown to suffice in all simulations.

It follows that for a sufficient DFT size ξ_u will be small since $\hat{\mathbf{u}}_1[\tau]$ should be normal. Thus (15) is implemented at increasing DFT size until ξ_u falls below a some given threshold ε_u.

VI. Simulations and Results

A. Numerical Example

To demonstrate the potential of the generalized polynomial power method, we assume a simple case of $\boldsymbol{A}(z)$ where we know the ground truth factorisation according to (1). For these factors, $\boldsymbol{\Sigma}(z) \in \mathbb{C}^{3\times 2}$ contains

$$\sigma_1(z) = \frac{1}{2}z + 4 + \frac{1}{2}z^{-1}, \sigma_2(z) = \frac{1}{4}z + 1 + \frac{1}{4}z^{-1} \,, \tag{18}$$

which are spectrally majorised. The left-singular vector are constructed via elementary paraunitary operation given as [12]

$$\boldsymbol{U}(z) = \prod_{i=1}^{2} \{\mathbf{I} - (1 - z^{-1})\mathbf{e}_i\mathbf{e}_i^{\mathrm{H}})\} \,, \tag{19}$$

where $\mathbf{e}_{i=1,2} = [1, 1, \mp 1]^{\mathrm{T}}/\sqrt{(3)} \in \mathbb{C}^3$ are unit-norm vectors. The right-singular vectors in $\boldsymbol{V}(z) \in \mathbb{C}^{2\times 2}$ of order 2 are generated by the same approach with $\mathbf{e}_1 = [1, -1]^{\mathrm{T}}/\sqrt{2}$ and $\mathbf{e}_2 = [-1, 0]^{\mathrm{T}}$. The polynomial matrix $\boldsymbol{A}(z)$ is then defined as $\boldsymbol{U}(z)\boldsymbol{\Sigma}(z)\boldsymbol{V}^{\mathrm{P}}(z)$.

Algorithm 1 is executed with $\epsilon = 10^{-12}$, $k_{\max} = 10^3$, $\boldsymbol{R}(z) = \boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z)$ and $\boldsymbol{x}^{(0)}(z) = 1$. The truncation method employed is the order limitation [19] where the order of $\boldsymbol{x}^{(k)}(z)$ post-normalisation is limited to the estimated support obtained from [24]. Algorithm 1 converges in 44 iterations resulting in $\xi_v = 1.4 \times 10^{-11}$. Once the left-singular vector is esimated, the corresponding singular value is estimated via (14). With $K = 16$, the time-domain aliasing

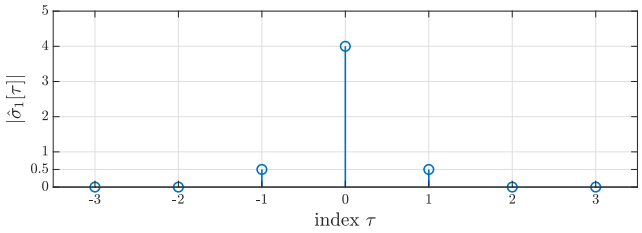

Fig. 1. Polynomial Power method based estimated dominant singular value coefficients for the numerical example.

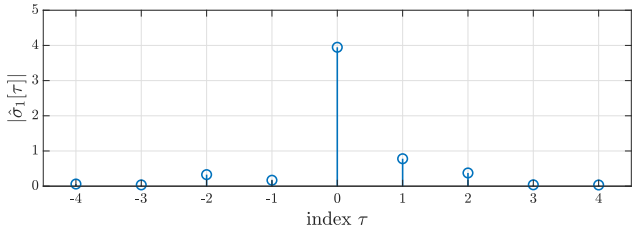

Fig. 2. GSBR2 based estimated dominant singular value coefficients for the numerical example.

$\xi_\lambda = 8\times 10^{-28}$. The trailing coefficients of $\hat{\sigma}_1[\tau]$ are truncated on either side of $\tau = 0$ via a threshold of 10^{-10}. This results in order of 6 whereas the ground-truth singular value has an order of 2. The coefficients are illustrated in Fig. 1 where the coefficients at $\tau = 0, \pm 1$ exactly match the ground-truth coefficients in (18), whereas the coefficients at $\tau = \pm 2, \pm 3$ are smaller than 10^{-5}. The normalized squared difference between the estimated and the ground-truth singular value, which can be defined similar to (16) as

$$\xi_\sigma = \sum_{\tau} \frac{|\sigma_1[\tau] - \hat{\sigma}_1[\tau]|^2}{|\sigma_1[\tau]|^2} \,, \tag{20}$$

is 3.5×10^{-12}. The corresponding left-singular vector is then obtained from (15) with a DFT of size $K = 16$. Thereafter, the order is limited by a shifted-truncation to 3, which achieves a metric of $\xi_v = 9 \times 10^{-12}$.

The GSBR2 is executed with $\mu_{\mathrm{PU}} = 10^{-5}$, $\epsilon = 10^{-5}$ and $\mu_{\mathrm{PH}} = 10^{-5}$ for 1000 iterations and results in $\xi_v = 9.6\times 10^{-5}$, $\xi_u = 1.5 \times 10^{-5}$ and $\xi_\lambda = 2.7 \times 10^{-2}$. The dominant singular value estimated with GSBR2 has order 8 whose coefficients are illustrated in Fig. 2. It is evident that the estimated singular value is neither conjugate symmetric and nor the coefficients match the ground truth coefficients except at $\tau = 0$ where $\hat{\sigma}[0] = 3.95 \approx 4$. This loss of conjugate symmetry may be the cause of large value of ξ_λ.

B. Ensemble Test

In a more extensive test, we evaluate the proposed method against the Kogbetliantz transformation-based method [11] which we refer to as the generalised second order sequential best rotation (GSBR2) algorithm via an ensemble consisting of 500 randomised instantiations of $\boldsymbol{A}(z) \in \mathbb{C}^{3\times 2}$ such that each instance has $\mathcal{O}\{\boldsymbol{U}(z)\} = \mathcal{O}\{\boldsymbol{V}(z)\} = 10$ and $\mathcal{O}\{\boldsymbol{\Sigma}(z)\} = 20$. All the instantiations have spectrally majorised singular values.

TABLE I
PERFORMANCE COMPARISON OF GSBR2 AND GPPM

Metrics	GSBR2	GPPM
$\mathcal{O}\{\hat{\boldsymbol{u}}_1(z)\}$	966 ± 185	10
$\mathcal{O}\{\hat{\boldsymbol{v}}_1(z)\}$	422 ± 126	10
$\mathcal{O}\{\hat{\sigma}_1(z)\}$	96 ± 38	57 ± 4
ξ_v	$(1.2 \pm 0.8) \times 10^{-3}$	$(5.5 \pm 4.5) \times 10^{-5}$
ξ_u	$(1.6 \pm 0.85) \times 10^{-3}$	$(5.5 \pm 4.5) \times 10^{-5}$
ξ_σ	0.09 ± 0.07	$(1.5 \pm 1.3) \times 10^{-5}$
time(s)	0.67 ± 0.15	0.44 ± 0.19

For the proposed method, Algorithm 1 is simulated with $\epsilon = 10^{-10}$, $k_{\max} = 10^3$, $\boldsymbol{R}(z) = \boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z)$ and $\boldsymbol{x}^{(0)}(z) = 1$. The order of the product vector is limited to 10, with its order estimated through the method in [24], followed by shifted-truncation [19, 25]. The corresponding singular value and the left-singular vector are extracted at $K = 2^{\lceil \log_2(\mathcal{O}\{\boldsymbol{A}(z)\hat{\boldsymbol{u}}_1(z)\}) \rceil}$ where $\lceil . \rceil$ denotes ceiling operation. The trailing coefficients of the estimated left-singular are truncated below a threshold of 10^{-10} while the right-singular vector is similarly order-limited to its estimated support.

GSBR2 is simulated with $\mu_{\mathrm{PU}} = 10^{-4}$, $\mu_{\mathrm{PH}} = 2 \times 10^{-10}$ employing the original truncation method of SBR2/SMD [4, 5]. The algorithm is allowed to perform a maximum of 200 iteration; however, the execution is terminated if the off-diagonal terms fall below 10^{-6}.

The ensemble average for all the metrics is shown in Table I. It is evident that the proposed method provides a more compact order approximation for both the left- and right-singular vectors and the singular value compared to the GSBR2. Moreover, the errormetrics ξ_u and ξ_v of the proposed method's extracted singular vectors reach orders of magnitude below those obtained with GSBR2. Likewise, the normalized squared difference between the estimated and ground-truth singular value is orders of magnitude lower for the polynomial method than GSBR2. The potential reason for the large deviation of the GSBR2's estimated singular value is likely the imperfect conjugate symmetry due to the uncoupled allpass ambiguity of the estimated left- and right-singular vectors. Also, SBR2-type algorithms are known to only achieves a relatively poor diagonalisation compared to their DFT-domain counterparts in e.g. [2, 3].

VII. CONCLUSION

The polynomial power method, which was initially proposed form para-Hermitian matrices, has been extended into the generalized polynomial power method for computing the dominant left- and right-singular vectors and their corresponding singular value of a polynomial matrix. The proposed extension provides better estimation of the singular vectors with lower order approximation as compared to the only direct PSVD algorithm based on the Kogbetliantz method. The proposed method promises better results and can be further utilized to compute the PSVD of a polynomial matrix through the polynomial matrix deflation analogous to ordinary matrix deflation.

REFERENCES

[1] S. Weiss, J. Pestana, and I. K. Proudler, "On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix," *IEEE Transactions on Signal Processing*, **66**(10):2659–2672, 2018.

[2] S. Weiss, I. K. Proudler, and F. K. Coutts, "Eigenvalue decomposition of a parahermitian matrix: Extraction of analytic eigenvalues," *IEEE Transactions on Signal Processing*, **69**:722–737, 2021.

[3] S. Weiss, I. Proudler, F. Coutts, and F. Khattak, "Eigenvalue decomposition of a parahermitian matrix: Extraction of analytic eigenvectors," *IEEE Transactions on Signal Processing*, to appear, 2023.

[4] J. G. McWhirter, P. D. Baxter, T. Cooper, S. Redif, and J. Foster, "An EVD algorithm for para-Hermitian polynomial matrices," *IEEE Transactions on Signal Processing*, **55**(5):2158–2169, 2007.

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

[6] V. W. Neo, C. Evers, and P. A. Naylor, "Speech enhancement using polynomial eigenvalue decomposition," in *IEEE WASPAA*, 2019, pp. 125–129.

[7] S. Redif, J. G. McWhirter, P. D. Baxter, and T. Cooper, "Robust broadband adaptive beamforming via polynomial eigenvalues," in *OCEANS*, 2006, pp. 1–6.

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

[9] J.A. Foster, J. G. McWhirter, M. R. Davies, and J. A. Chambers, "An algorithm for calculating the QR and singular value decompositions of polynomial matrices," *IEEE Transactions on Signal Processing*, **58**(3):1263–1274, 2010.

[10] Z. Wang, A. Sandmann, J. G. McWhirter, and A. Ahrens, "Multiple shift SBR2 algorithm for calculating the SVD of broadband optical MIMO systems," in *Int. Conf. TSP*, 2016, pp. 433–436.

[11] J.G. McWhirter, "An algorithm for polynomial matrix SVD based on generalised kogbetliantz transformations," in *EUSIPCO*, 2010, pp. 457–461.

[12] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*. USA: Prentice-Hall, Inc., 1993.

[13] N. Moret, A. Tonello, and S. Weiss, "MIMO precoding for filter bank modulation systems based on PSVD," in *IEEE VTC*, 2011, pp. 1–5.

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

[15] S. Weiss and I. Proudler, "Comparing efficient broadband beamforming architectures and their performance trade-offs," in *DSP*, **1**:417–423, 2002.

[16] G. E. Forsythe and P. Henrici, "The cyclic jacobi method for computing the principal values of a complex matrix," *Transactions of the American Mathematical Society*, **94**(1):1–23, 1960.

[17] G. Barbarino and V. Noferini, "On the Rellich eigendecomposition of para-hermitian matrices and the sign characteristics of $*$-palindromic matrix polynomials," *Linear Algebra & Its Appl.*, **672**:1–27, 2023.

[18] S. Weiss, I. K. Proudler, J. P. G. Barbarino, and J. G. McWhirter, "On properties and structure of the analytic singular value decomposition," submitted 2023.

[19] F. A. Khattak, I. K. Proudler, and S. Weiss, "Extension of power method to para-hermitian matrices: Polynomial power method," in *EUSIPCO*, submitted 2023.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.

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

[22] 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.

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

[24] F. A. Khattak, I. K. Proudler, and S. Weiss, "Support estimation of analytic eigenvectors of parahermitian matrices," in *2022 Int. Symp. RAEE & CS*, 2022.

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