INVERSION OF PARAHERMITIAN MATRICES

Stephan Weiss, Andrew P. Millar and Robert W. Stewart

Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow G1 1XW, Scotland, UK
stephan.weiss@eee.strath.ac.uk

ABSTRACT
Parahermitian matrices arise in broadband multiple-input multiple-output (MIMO) systems or array processing, and require inversion in some instances. In this paper, we apply a polynomial eigenvalue decomposition obtained by the sequential best rotation algorithm to decompose a parahermitian matrix into a product of two parahermitian matrices. Its Fourier pair can be defined. Its Fourier pair can be defined. Its Fourier pair can be defined. Its Fourier pair of the form

\[ R(x) = \delta \{ x[n] \} \cdot x^H(n-\tau) \]

(1)
can be defined. Its Fourier pair \( R(z) \) is defined as

\[ R(z) = \sum_{\tau=-\infty}^{\infty} R(\tau)z^{-\tau} \]

(2)
is a power spectral matrix, which takes the form of a matrix polynomial in \( z \) [1]. This power spectral matrix is parahermitian, fulfilling

\[ R(z) = R(z^*) \in \mathbb{C}^{M \times M}(z) \]

(3)
with the parahermitian operator \( \tilde{\cdot} \) that implies complex conjugate transposition and reversal of the polynomials, i.e. \( \tilde{R(z)} = R(z^{-1}) \). An example of a 3 \times 3 parahermitian matrix of order 4 is given in Fig. 1.

The inversion of such a parahermitian matrix is required e.g. for the generalised Wiener filter sought in [2] and [3].

The inversion of parahermitian matrices also arises as part of the pseudo-inverse of an arbitrary rectangular polynomial matrix \( C(z) \in \mathbb{C}^{M \times N}(z) \), which can represent the transfer function of a broadband MIMO system. Here, the transfer path between each pair of transmit and receive antennas requires to be modelled by an FIR filter, instead of the simpler complex gain factor that can be used in the narrowband case. To compute a zero-forcing linear precoder or equaliser, the pseudo-inverse of \( C(z) \) is given by

\[ C^{-1}(z) = \begin{cases} \tilde{C}(z)(\tilde{C}(z))^+ & M \geq N \\ \tilde{C}(z)(C(z)\tilde{C}(z))^{-1} & M \leq N \end{cases} \]

(3)
where the products \( \tilde{C}(z)C(z) \) and \( C(z)\tilde{C}(z) \) are parahermitian. Such MIMO systems can be found in multichannel deconvolution problems in audio and acoustics [4, 5, 6, 7] as well as in communications [2, 3].

This paper is organised as follows. Sec. 2 will review existing time- and frequency-domain methods for the inversion of polynomial matrices. The proposed approach will be outlined in Sec. 3 and require the inversion a auto-correlation sequences, which will be addressed in Sec. 4. Finally, simulation results will be presented in Sec. 5 and conclusions be drawn in Sec. 6.

2. STATE OF THE ART
To the best of our knowledge, polynomial space-time covariance matrices have not been inverted previously. However, inversion approaches exist for dispersive MIMO systems in audio and acoustics, as well as communications. In the following, a number of the approaches mentioned in Sec. 1 will be outlined.

2.1 Time-Domain / MMSE Inversion
An early reference to the inversion of MIMO systems can be found in [4], where a linear system with convolutional matrices is setup that allows to solved for the inverse system using standard linear algebraic techniques. However, this requires to select the order of the inverse system a priori, on which the accuracy of the solution will depend.

2.2 Frequency Domain Inversion
For a similar problem in acoustics, [5] uses a DFT approach to reduce the broadband problem into narrowband problems that can be independently solved using standard matrix inversion techniques in each frequency bin. Evaluating on a finite grid of frequency points \( \Omega \), the inversion of a MIMO system matrix \( C(z) \) is based on the DFT representation \( C(z)\mid_{z=e^{i\Omega}} = C(e^{i\Omega}) \), such that

\[ \tilde{C}^{-1}(e^{i\Omega}) = C^t(e^{i\Omega})A(e^{i\Omega}) \]

(4)
whereby the matrix \( A(e^{i\Omega}) \) is a reference control system, which can e.g. be utilised to permit a delay for the overall system. The results

![Figure 1: Example of a 3 x 3 parahermitian matrix R(z).](image-url)
is transformed back into the time domain by means of an Inverse DFT, but if the number of frequency bins is selected too low, and the inverse is longer than anticipated, wrap-around due to the DFT implementing a cyclic convolution will occur.

Similar to the solution in [5], OFDM approaches to broadband MIMO inversion will be based on a solution in the DFT domain. In the following, an approach based on polynomial matrices will be proposed and evaluated.

3. POLYNOMIAL EVD-BASED INVERSION

The inversion technique for parahermitian matrices proposed in this paper is based on the polynomial eigenvalue decomposition (PEVD) by McWhirter et al. [8]. First, the PEVD is characterised, before the inversion methods and a practical algorithm to implement a PEVD are presented.

3.1 Polynomial EVD

A polynomial eigenvalue decomposition of a parahermitian matrix $\mathbf{R}(z) \in \mathbb{C}^{M \times M}(z)$ is defined as

$$\mathbf{R}(z) = \mathbf{Q}(z)\Lambda(z)\mathbf{Q}(z)$$

whereby $\mathbf{Q}(z) \in \mathbb{C}^{M \times M}(z)$ is paraunitary, i.e.

$$\mathbf{Q}(z)\mathbf{Q}(z) = \mathbf{Q}(z)\mathbf{Q}(z)^H = \mathbf{I}$$

and $\Lambda(z) \in \mathbb{C}^{M \times M}(z)$ is parahermitian and diagonal with diagonal elements $\Lambda_i(z)$ ordered such that the power spectral densities $\Lambda_i(e^{j\Omega})$ fulfill

$$\Lambda_i(e^{j\Omega}) \geq \Lambda_{i+1}(e^{j\Omega}), \quad \forall \Omega, \quad i = 0 \ldots (M-2) .$$

The property (7) is called spectral majorisation. While a practical decomposition algorithm developed in [8] will be discussed later, an example for the spectrally majorised $\Lambda(z)$ arising from the decomposition of the matrix in Fig. 1 is given in Fig. 2.

3.2 Polynomial Inverse

Based on the PEVD, the inverse can be formulated as

$$\mathbf{R}^{-1}(z) = \mathbf{Q}(z)\Lambda^{-1}(z)\mathbf{Q}(z).$$

It is straightforward to show that

$$\mathbf{R}^{-1}(z)\mathbf{R}(z) = \mathbf{R}(z)\mathbf{R}^{-1}(z) = \mathbf{I}. $$

The paraunitarity of $\mathbf{Q}(z)$ plays a vital role in the simplicity of this inverse. It remains to invert the diagonal polynomial matrix $\Lambda(z)$, which can be achieved by inverting all elements along on the main diagonal,

$$\Lambda^{-1}(z) = \begin{bmatrix} \hat{\Lambda}_0(z) & \vdots & \vdots & \hat{\Lambda}_{M-1}(z) \end{bmatrix},$$

whereby $\Lambda_i(z)\hat{\Lambda}_i(1) = 1$. Next, a practical decomposition to determine $\mathbf{Q}(z)$ will be reviewed, before methods to invert the off-diagonal elements $\Lambda_i(z)$ are discussed in Sec. 4.

3.3 Sequential Best Rotation Algorithm

SBR2 is an iterative broadband eigenvalue decomposition technique based on second order statistics only and can be seen as a generalisation of the Jacobi algorithm. The decomposition after $L$ iterations is based on a paraunitary matrix $\mathbf{U}_L(z)$,

$$\mathbf{U}_L(z) = \prod_{i=0}^{L} \mathbf{Q}_i(z) \mathbf{P}(z)$$

whereby $\mathbf{Q}_i(z)$ is a Jacobi rotation and the matrix $\mathbf{P}(z)$ a paraunitary matrix of the form

$$\mathbf{P}(z) = \mathbf{I} - \nu v_i v_i^H + z^{-\Lambda_i} v_i v_i^H$$

with $v_i = [0 \cdots 0 1 0 \cdots 0]^H$ containing zeros except for a unit element in the $\delta$th position. Thus $\mathbf{P}(z)$ is an identity matrix with the $\delta$th diagonal element replaced by a delay $z^{-\Lambda_i}$.

At the $i$th step, SBR2 will eliminate the largest off-diagonal element of the matrix $\mathbf{U}_{L-1}(z)\mathbf{R}(z)\mathbf{U}_{L-1}(z)$, which is defined by the two corresponding sub-channels and by a specific lag index. By delaying the two contributing sub-channels appropriately with respect to each other by selecting the position $\delta$ and the delay $\Delta_i$, the lag value is compensated. Thereafter a Jacobi rotation $\mathbf{Q}_i(z)$ can eliminate the targeted element such that the resulting two terms on the main diagonal are ordered in size, leading to a diagonalisation and at the same time accomplishing a spectral majorisation.

SBR2 only achieves an approximate diagonalisation after a finite number of iteration steps when off-diagonal elements are smaller than a threshold $\delta$,

$$\mathbf{R}(z) = \mathbf{Q}(z)(\Lambda(z) + \mathbf{E}(z))\mathbf{Q}(z)^H$$

with $\Lambda(z)$ diagonal and $\mathbf{E}(z)$ a non-sparse error matrix with $\| \mathbf{E}(z) \|_\infty \leq \delta$. Here, the infinity norm $\| \mathbf{R}(z) \|_\infty$ is defined as returning the largest element across all matrix-valued coefficients of the polynomial $\mathbf{R}(z)$,

$$\| \mathbf{R}(z) \|_\infty = \max \| \mathbf{R}(z) \|_\infty .$$

An alternative stopping criterion is to define a maximum number of iterations for SBR2 [11, 13].

3.4 Error Gain

SBR2 will reach a decomposition where the matrix $\mathbf{Q}(z)$ is perfectly paraunitary by definition. However, SBR2 only achieves an approximate diagonalisation, and as a result the inversion process has an inherent error as off-diagonal terms will be ignored when computing (10) later.

The decomposition with SBR2 can be characterised as follows

$$\mathbf{R}(z) = \mathbf{Q}(z)\Lambda(z)\mathbf{Q}(z) + \mathbf{Q}(z)\mathbf{E}(z)\mathbf{Q}(z)$$

$$= \mathbf{Q}(z)\Lambda(z)\mathbf{Q}(z) - \mathbf{E}(z) .$$
Theorem 2. An important measure in the inversion process is the amplification of elements in the resulting sequence due to the finite number of iterations of SBR2. Ignoring off-diagonal elements in the inversion process, we exclude spectral zeros from the inverse of $E_i(z)$ due to the inversion process, resulting in $E_2(z)$. An interesting performance measure of inversion is therefore the error gain

$$Y = \frac{\|E_2(z)\|_F}{\|E_1(z)\|_F},$$

where the Frobenius norm of a polynomial matrix $R(z) = \sum_{\nu = -\infty}^{\infty} R_{\nu} z^{-\nu}$ is defined as

$$\|R(z)\|_F = \left( \sum_{\nu = -\infty}^{\infty} \|R_{\nu}\|_2^2 \right)^{\frac{1}{2}}. \tag{19}$$

Next, we will concentrate on the inversion of the on-diagonal elements of $A(z)$.

### 4. Inversion of Autocorrelation Sequences

This section addresses the inversion of on-diagonal elements of $A(z)$. These elements have the properties of auto-correlation sequences, i.e.

$$r_\mu[\tau] = r_\mu^*[\tau] \Leftrightarrow R_\mu(z) = R_\mu^*(z^{-1}).$$

This symmetry can be exploited in the inversion process, since the inverse of a linear phase single-input single-output (SISO) system must also be a linear phase system and therefore have the same symmetry properties [10, 9, 12]. From $R_\mu(z)R_\mu^{-1}(z) = 1$ we deduce $r_\mu[\tau] * s_\mu[\tau] = \delta[\tau]$ where $s_\mu[\tau] = R_\mu(z)$ is the inverse of the auto-correlation sequence. We hence use $S_\mu(z)$ to describe the inverse of $R(z)$ due to potential truncation errors in the methods described below.

#### 4.1 Spectral Factorisation

Due to its minimum phase property, each auto-correlation function can be factored into

$$A(z) = \Lambda_{\min}(z)\Lambda_{\max}(z)$$

with $\Lambda_{\min}(z)$ minimum and $\Lambda_{\max}(z)$ maximum phase. In the inversion process, we exclude spectral zeros from the $A_i(z)$, as this would lead to a non-invertible sequence. Do to symmetry, $\Lambda_{\max}(z) = \Lambda_{\min}(z^{-1})$.

We calculate $\Lambda_{\min}(z)$ and have

$$\Lambda^{-1}_\min(z) = \Lambda^{-1}_{\min}(z)\Lambda^{-1}_{\max}(z^{-1}) \tag{21}$$

First order sections of $\Lambda_{\min}(z)$ are inverted using geometric series expansions of appropriate lengths in the time domain. Thereafter, convolution yields $\hat{A}_{\min}(\tau)$, and the estimate of the inverse is computed according to

$$s_i[\tau] = \hat{A}_{\min}[\tau] * s_{\min}[-\tau]$$

and truncated to the range $-T \leq \tau \leq T$.

#### 4.2 Time Domain / MMSE Inversion

The time domain inversion is based on a convolutional matrix description of the convolution of an auto-correlation sequence $r[n]$ and its inverse $s[n]$.

$$\begin{bmatrix} r[N] \\ \vdots \\ r[-N] \end{bmatrix} \begin{bmatrix} s[-T] \\ \vdots \\ s[T] \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$As = d$$

with $A \in \mathbb{C}^{(2T+2N+1) \times (2T+1)}$, $s \in \mathbb{C}^{(2T+1)}$ and $d \in \mathbb{Z}^{(2T+2N+1)}$. A solution can be obtained via the left pseudo-inverse,

$$s = (A^H A)^{-1} A^H d \tag{22}$$

This solution should have the same symmetry properties as $r[n]$, and any deviation from symmetry must be due to numerical problems in the inversion process. The symmetry error

$$\varepsilon = \|s - Js^*\|_2$$

should be as small as possible.

A minimum mean square error solution to (22) can be obtained by including the noise-to-signal ratio for regularisation purposes.

### 4.3 Inversion with Explicit Symmetry Constraint

An ill-conditioned $A$ can lead to an asymmetric solution in (22). Hence it is advantageous to enforce symmetry in the setup.

This can be performed by a Lagrangian approach, which solves the constrained optimisation problem

$$\text{find } \min \|As - d\|_2^2$$

subject to $s = Js^*$. \tag{24}

Instead of solving this Lagrangian problem, the next section discusses a direct approach of embedding the constraint into the formulation.

### 4.4 Inversion with Implicit Symmetry Constraint

The symmetry condition can be incorporated into the system equation by formulating

$$\begin{bmatrix} \Re(A) & -\Im(A) \\ \Im(A) & \Re(A) \end{bmatrix} \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} = \begin{bmatrix} d \\ 0 \end{bmatrix} \tag{25}$$

In this, the inverse is implicitly constrained by only defining half the response as

$$w = \begin{bmatrix} s[-T] \\ \vdots \\ s[1] \end{bmatrix}_{\frac{N}{2}}$$

with

$$\Re(s) = \begin{bmatrix} I_F \\ 0^T \end{bmatrix} \Re(w) = M_1 \Re(w) \tag{26}$$

$$\Im(s) = \begin{bmatrix} I_F \\ 0^T \end{bmatrix} \Im(w) = M_2 \Im(w) \tag{27}$$
to reconstruct the real and imaginary part of the true solution. Therefore the problem formulation becomes

\[
\begin{bmatrix}
M_1 \Re(A) & -M_2 \Im(A) \\
M_2 \Re(A) & M_2 \Re(A)
\end{bmatrix}
\begin{bmatrix}
\Re(w) \\
\Im(w)
\end{bmatrix} = \begin{bmatrix} d \\ \tilde{d}_c \end{bmatrix}
\]

and the solution is reached via the pseudo-inverse

\[
s = [M_1 \quad iM_2](A^\top A_c)^{-1} A^\top \tilde{d}_c.
\]

5. RESULTS

5.1 Comparison of ACS Inversion Approaches

The two approaches of constrained time domain inversion and spectral factorisation are compared in a number of simulations, with Figs. 3 and 4 showing the inversion of an autocorrelation sequence for two different lengths of the inverse. Additionally, Fig. 5 addresses the inversion error — measured Euclidean distance of the convolved autor-correlation sequence and its inverse from a Dirac impulse —, and Fig. 6 the computational complexity of the proposed schemes.

For the same length, spectral factorisation is worse in terms of inversion error but lower in terms of complexity. The MMSE approaches both achieve the same accuracy for inversion, but the implicitly constrained method offers a lower complexity due to real valued arithmetic even for complex valued problems, and the resulting response is perfectly symmetric. Unconstrained optimisation requires complex valued arithmetic, and while the error in symmetry is not significant, it grows with the order of the inverse, as evident from Fig. 7.

5.2 Parahermitian Matrix Inversion

As an example, the inversion of the autocorrelation sequences for the matrix in Fig. 2 is shown in Fig. 8 with the convolution of the original parahermitian matrix and its inverse shown in Fig. 9.

6. CONCLUSIONS

This paper has presented an inversion approach for parahermitian polynomial matrices. A number of issues, such as the error gain applied to approximation errors and the possibility of regularisation that may be linked to the size of off-diagonal terms, still need to be investigated, and will be included in the full paper. Initial results indicate that the proposed approach can work well with reasonable complexity, and presents an attractive approach to existing inversion methods for arbitrary broadband MIMO systems.

REFERENCES

Figure 5: Ensemble MSE in inverting ACS for various methods and orders. Dashed lines show one standard deviation.

Figure 6: Computational complexity of inversion methods measured in terms of CPU time on a very old laptop.

Figure 7: Symmetry error $\varepsilon$ of unconstrained inversion. Note that spectral factorisation and constrained inversion yield $\varepsilon = 0$.

Figure 8: Inverted diagonal matrix $\Lambda^{-1}(e^{i\Omega})$ for the matrix $R(z)$ in Fig. 1. Inverse responses are generated by the constrained MMSE method with $T = 100$.

Figure 9: System $D(z) = R(z)S(z)$, with $R(z)$ given in 1.