# Introduction to Polynomial Matrix Algebra and Applications

Stephan Weiss and Ian K. Proudler

Centre for Signal & Image Processing
Department of Electronic & Electrical Engineering
University of Strathclyde, Glasgow, Scotland, UK

UDRC-EURASIP Summer School, Edinburgh, 30 June 2022

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/S000631/1 and the MOD University Defence Research Collaboration in Signal Processing.

# Presentation Overview

1. Overview

Part I: Polynomial Matrices and Decompositions

2. Polynomial matrices and basic operations
3. Parahermitian matrix / polynomial eigenvalue decomposition (PhEVD /PEVD)
4. Iterative PEVD algorithms
5. PEVD Matlab toolbox

Part II: Beamforming & Source Separation Applications

6. Broadband MIMO decoupling
7. Broadband angle of arrival estimation
8. Broadband beamforming
9. Source-sensor transfer function extraction
10. Weak transient signal detection

# What is a Polynomial Matrix?

- A polynomial matrix is a polynomial with matrix-valued coefficients, e.g.:

$$\boldsymbol{A}(z) = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} z^{-1} + \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} z^{-2} \; ; \qquad (1)$$

- a polynomial matrix can equivalently be understood a matrix with polynomial entries, i.e.

$$\boldsymbol{A}(z) = \begin{bmatrix} 1 + z^{-1} - z^{-2} & -1 + z^{-1} + 2z^{-2} \\ -1 + z^{-1} + z^{-2} & 2 - z^{-1} - z^{-2} \end{bmatrix} \; ; \qquad (2)$$

- we may also encounter matrix-valued power series, Laurent polynomials, and Laurent series.

# Matrix-Valued Polynomials and Power Series

- A power series $a(z)$ arises as the $z$-transform

$$a(z) = \sum_n a[n] z^{-n} \qquad \text{or} \quad \text{short} \qquad a(z) \bullet\!\!-\!\!\circ\, a[n] \;, \tag{3}$$

- for $a(z)$ to exist as a power series, $a[n]$ must be
  causal: $a[n] = 0 \;\forall n < 0$;
  absolutely convergent: $\sum_n |a[n]| < \infty$
- absolute convergence implies that $a[n]$ decays at least as fast as an exponential function;
- a polynomial is a power series, but of finite length;
- polynomials or power series can form the entries of a matrix $\boldsymbol{A}(z)$.

# Example of a Power Series

- For the geometric series

$$a[n] = \begin{cases} 0, & n < 0 \\ (\frac{1}{2})^n, & n \geq 0 \end{cases} \tag{4}$$

  we have

$$\sum_n |a[n]| = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots \quad = 2 < \infty \; ; \tag{5}$$

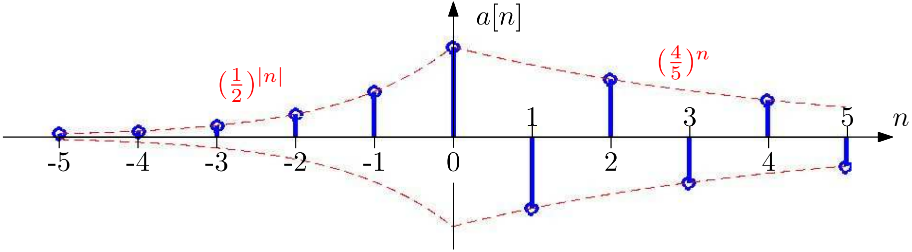
- therefore $a[n]$ is an absolutely convergent power series, and $a(z)$ exists as an analytic function;
- here, for $a(z)$:

$$a(z) = 1 + \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} + \frac{1}{8}z^{-3} + \ldots \quad = \frac{1}{1 - \frac{1}{2}z^{-1}} \; ; \tag{6}$$

- this looks like the transfer function of a causal infinite impulse response (IIR) filter.

# Laurent Series and Laurent Polynomials

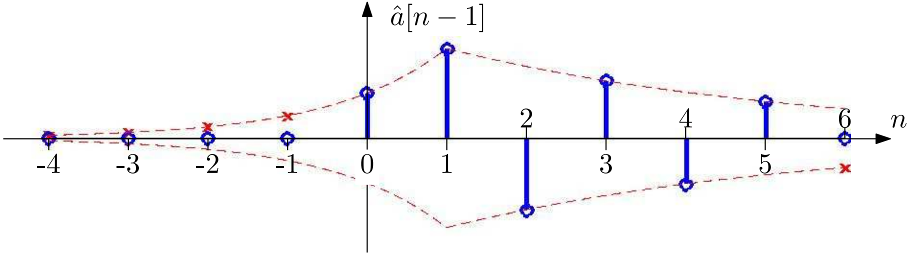
- A Laurent series $a[n]$ is potentially infinite, but can include non-negative terms for both $n \geq 0$ and $n < 0$;
- for $a(z) \bullet\!\!-\!\!\circ\; a[n]$ to exist, $a[n]$ needs to decay at least exponentially in both positive and negative time direction;




- if it possesses finite support, $a(z)$ is a Laurent polynomial.

# Analyticity and Polynomial Approximation

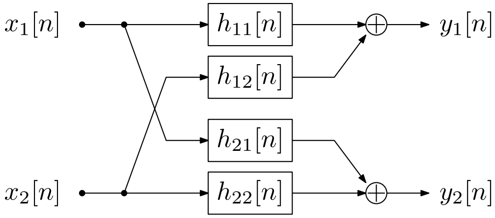
- Absolute convergence of $a[n]$ implies analyticity of $a(z)$ $\bullet\!\!-\!\!\circ$ $a[n]$;
- the best approximation of an infinite order, analytic $a(z)$ in the least squares sense is by truncation (power series $\longrightarrow$ polynomial);
- likewise, a Laurent series can be approximated by a polynomial through truncation ($\longrightarrow$ Laurent polynomial) and an appropriate delay ($\longrightarrow$ polymomial);




- hence polynomials can typically approximate any general analytic function well, and arbitrarily closely.

# Where Do Polynomial Matrices Arise?

- A multiple-input multiple-output (MIMO) system could be made up of a number of finite impulse response (FIR) channels:

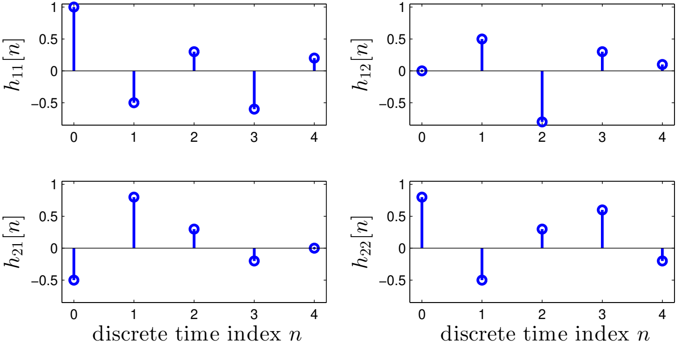



- writing this as a matrix of impulse responses:

$$\mathbf{H}[n] = \begin{bmatrix} h_{11}[n] & h_{12}[n] \\ h_{21}[n] & h_{22}[n] \end{bmatrix} . \tag{7}$$

# Transfer Function of a MIMO System

- Example for MIMO matrix $\mathbf{H}[n]$ of impulse responses:

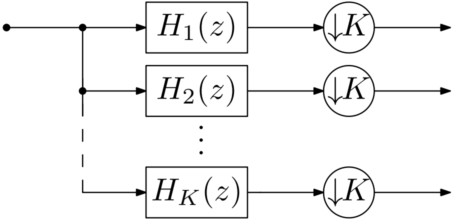



- the transfer function of this MIMO system is a polynomial matrix:

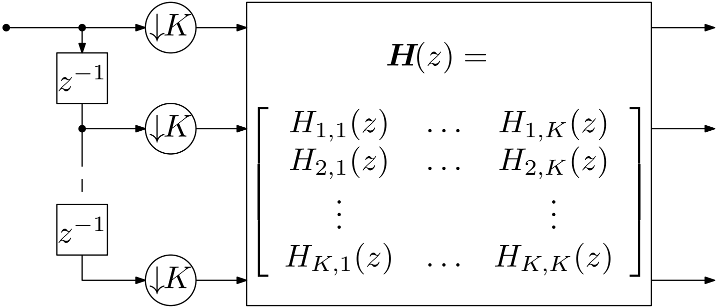
$$\boldsymbol{H}(z) = \sum_{n=-\infty}^{\infty} \mathbf{H}[n] z^{-1} \qquad \text{or} \qquad \boldsymbol{H}(z) \;\bullet\!\!-\!\!\circ\; \mathbf{H}[n] \tag{8}$$

# Analysis Filter Bank

- Critically decimated $K$-channel analysis filter bank:



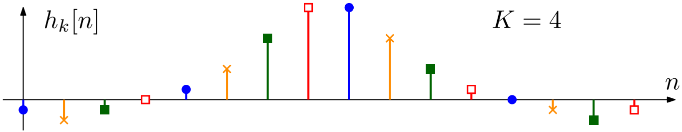

- equivalent polyphase representation:



# Polyphase Analysis Matrix

- With the $K$-fold polyphase decomposition of the analysis filters

$$H_k(z) = \sum_{n=1}^{K} H_{k,n}(z^K) z^{-n+1} \qquad (9)$$

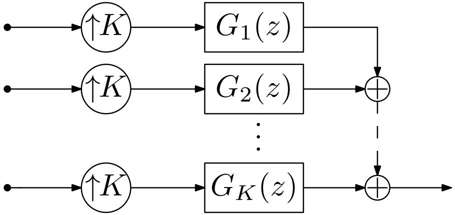



- the polyphase analysis matrix is a polynomial matrix:

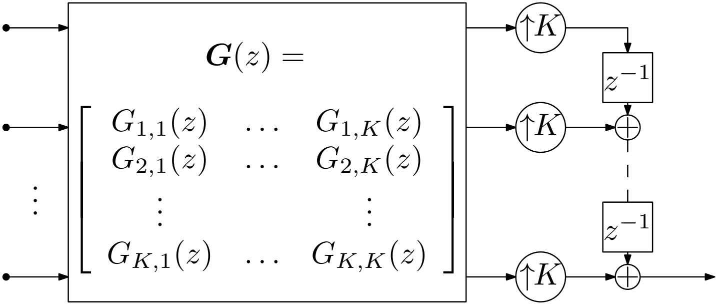
$$\boldsymbol{H}(z) = \begin{bmatrix} H_{1,1}(z) & H_{1,2}(z) & \dots & H_{1,K}(z) \\ H_{2,1}(z) & H_{2,2}(z) & \dots & H_{2,K}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{K,1}(z) & H_{K,2}(z) & \dots & H_{K,K}(z) \end{bmatrix} \qquad (10)$$

# Synthesis Filter Bank

- Critically decimated $K$-channel synthesis filter bank:




- equivalent polyphase representation:



# Polyphase Synthesis Matrix

- Analoguous to analysis filter bank, the synthesis filters $G_k(z)$ can be split into $K$ polyphase components, creating a polyphse synthesis matrix

$$\boldsymbol{G}(z) = \begin{bmatrix} G_{1,1}(z) & G_{1,2}(z) & \dots & G_{1,K}(z) \\ G_{2,1}(z) & G_{2,2}(z) & \dots & G_{2,K}(z) \\ \vdots & \vdots & \ddots & \vdots \\ G_{K,1}(z) & G_{K,2}(z) & \dots & G_{K,K}(z) \end{bmatrix} \qquad (11)$$

- operating analysis and synthesis back-to-back, perfect reconstruction is achieved if

$$\boldsymbol{G}(z)\boldsymbol{H}(z) = \mathbf{I} \; ; \qquad (12)$$

- i.e. for perfect reconstruction, the polyphase analysis matrix must be invertible: $\boldsymbol{G}(z) = \boldsymbol{H}^{-1}(z)$.

# Space-Time Covariance Matrix

- Measurements obtained from $M$ sensors are collected in a vector $\mathbf{x}[n] \in \mathbb{C}^M$:

$$\mathbf{x}^{\mathrm{T}}[n] = [x_1[n] \;\; x_2[n] \;\; \ldots \;\; x_M[n]] \;; \tag{13}$$

- with the expectation operator $\mathcal{E}\{\cdot\}$, the spatial correlation is captured by $\mathbf{R} = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\}$;
- for spatial and temporal correlation, we require a space-time covariance matrix

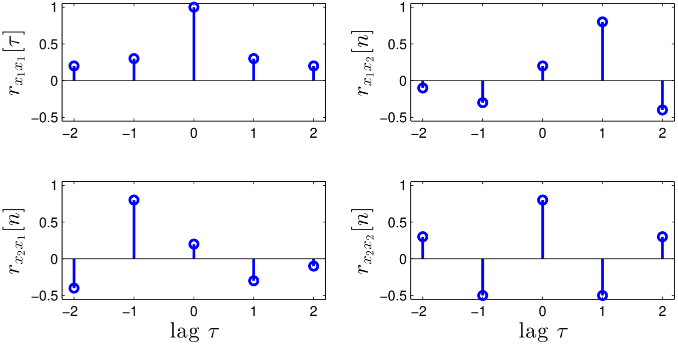
$$\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\} \tag{14}$$

- this space-time covariance matrix contains auto- and cross-correlation terms, e.g. for $M = 2$

$$\mathbf{R}[\tau] = \begin{bmatrix} \mathcal{E}\{x_1[n]x_1^*[n-\tau]\} & \mathcal{E}\{x_1[n]x_2^*[n-\tau]\} \\ \mathcal{E}\{x_2[n]x_1^*[n-\tau]\} & \mathcal{E}\{x_2[n]x_2^*[n-\tau]\} \end{bmatrix} \tag{15}$$

# Cross-Spectral Density Matrix

- example for a space-time covariance matrix $\mathbf{R}[\tau] \in \mathbb{R}^{2\times 2}$:




- the cross-spectral density (CSD) matrix

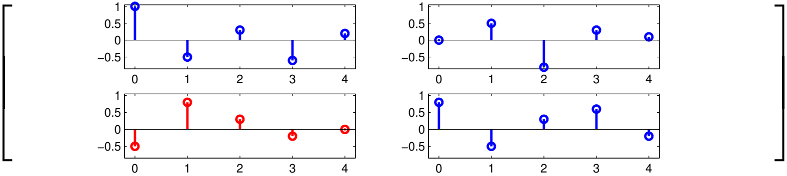
$$\boldsymbol{R}(z) \circ\!\!-\!\!\bullet \mathbf{R}[\tau] \tag{16}$$

is a polynomial matrix.

# Parahermitian Operator

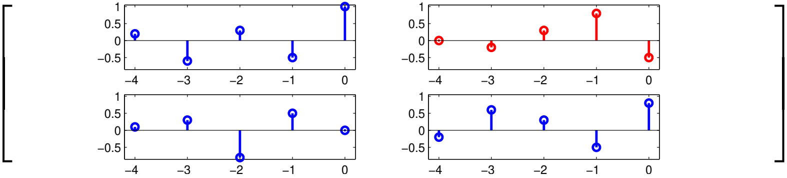
- A parahermitian operation is indicated by $\{\cdot\}^{\mathrm{P}}$, and compared to the Hermitian transposition of a matrix additionally performs a time-reversal;
- example:

$\boldsymbol{A}(z) =$




- parahermitian $\boldsymbol{A}^{\mathrm{P}}(z) = \boldsymbol{A}^{\mathrm{H}}(1/z^*)$:

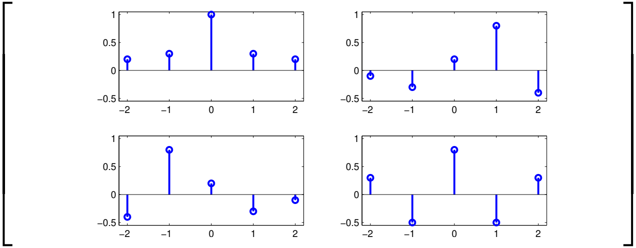
$\tilde{\boldsymbol{A}}(z) =$



# Parahermitian Property

- A polynomial matrix $\boldsymbol{R}(z)$ is parahermitian if $\boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}^{\mathrm{H}}(1/z^*) = \boldsymbol{R}(z)$;
- this is an extension of the symmetric (if $\mathbf{R} \in \mathbb{R}$) or or Hermitian (if $\mathbf{R} \in \mathbb{C}$) property to the polynomial case:
  transposition, complex conjugation and time reversal (in any order) do not alter a parahermitian $\boldsymbol{R}(z)$;
- any CSD matrix is parahermitian;
- example:

$\boldsymbol{R}(z) =$




$= \boldsymbol{R}^{\mathrm{P}}(z)$

# Paraunitary Matrices

- Recall that $\mathbf{A} \in \mathbb{C}$ (or $\mathbf{A} \in \mathbb{R}$) is a unitary (or orthonormal) matrix if $\mathbf{A}\mathbf{A}^{\mathrm{H}} = \mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}$;
- in the polynomial case, $\mathbf{A}(z)$ is paraunitary if

$$\boldsymbol{A}(z)\boldsymbol{A}^{\mathrm{P}}(z) = \boldsymbol{A}^{\mathrm{P}}(z)\boldsymbol{A}(z) = \mathbf{I} \tag{17}$$

- therefore, if $\boldsymbol{A}(z)$ is paraunitary, then the polynomial matrix inverse is simple:

$$\boldsymbol{A}^{-1}(z) = \boldsymbol{A}^{\mathrm{P}}(z) \tag{18}$$

- example: polyphase analysis or synthesis matrices of perfectly reconstructing (or lossless) filter banks are usually paraunitary.

# Attempt of Gaussian Elimination

- System of polynomial equations:

$$\begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z) \end{bmatrix} \cdot \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} = \begin{bmatrix} B_1(z) \\ B_2(z) \end{bmatrix} \qquad (19)$$

- modification of 2nd row:

$$\begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{11}(z) & \frac{A_{11}(z)}{A_{21}(z)} A_{22}(z) \end{bmatrix} \cdot \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} = \begin{bmatrix} B_1(z) \\ \frac{A_{11}(z)}{A_{21}(z)} B_2(z) \end{bmatrix} \qquad (20)$$

- upper triangular form by subtracting 1st row from 2nd:

$$\begin{bmatrix} A_{11}(z) & A_{12}(z) \\ 0 & \frac{A_{11}(z)A_{22}(z) - A_{12}(z)A_{21}(z)}{A_{21}(z)} \end{bmatrix} \cdot \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} = \begin{bmatrix} B_1(z) \\ \bar{B}_2(z) \end{bmatrix} \qquad (21)$$

- penalty: we end up with rational functions rather than polynomials.

# Parahermitian Matrix Eigenvalue Decomposition I

- For a Hermitian matrix $\mathbf{R} = \mathbf{R}^{\mathrm{H}}$, we know that an eigenvalue decomposition (EVD) $\mathbf{R} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}$ exists [18, 22];
- for eigenvalues $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_M\}$ and eigenvectors $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_M]$:

$$\mathbf{R}\mathbf{q}_m = \lambda_m \mathbf{q}_m$$

- eigenvalues $\lambda \in \mathbb{R}$;
- eigenvectors can be chosen as orthonormal, but may have an arbitary phase shift: $\mathbf{q}'_m = e^{j\varphi}\mathbf{q}_m$ is also an eigenvector;
- in case of an algebraic multiplicity $C$: $\lambda_m = \lambda_{m+1} = \cdots = \lambda_{m+C-1}$, only a $C$-dimensional subspace is defined, within which the eigenvectors can form an arbitrary orthonormal basis, with any unitary $\mathbf{V}$:

$$[\mathbf{q}'_m, \ \ldots \ \mathbf{q}'_{m+C-1}] = [\mathbf{q}_m, \ \ldots \ \mathbf{q}_{m+C-1}] \, \mathbf{V} \ , \tag{22}$$

# Parahermitian Matrix Eigenvalue Decomposition II

University of Strathclyde Engineering

- A standard EVD can diagonalise $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ\, \mathbf{R}[\tau]$ only for one specific value of $z$ or of $\tau$, respectively;
- we are interested in the EVD of a parahermitian matrix $\boldsymbol{R}(z)$ such that

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z)\; \boldsymbol{\Lambda}(z)\; \boldsymbol{Q}^{\mathrm{P}}(z)\;, \qquad (23)$$

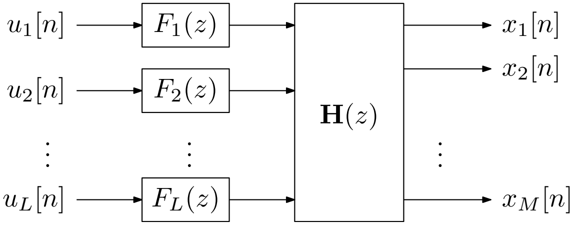
- $\boldsymbol{Q}(z) = [\boldsymbol{q}_1(z), \ldots, \boldsymbol{q}_M(z)]$ must be paraunitary, such that

$$\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}\;; \qquad (24)$$

- $\boldsymbol{\Lambda}(z) = \mathrm{diag}\{\lambda_1(z), \ldots, \lambda_M\}$ must be diagonal and parahermitian;
- the parahermitian property implies that on the unit circle, $\lambda(\mathrm{e}^{\mathrm{j}\Omega}) = \lambda(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega}} \in \mathbb{R}$;
- we call (23) a parahermitian matrix EVD.

# Analyticity of $\boldsymbol{R}(z)$

- The analyticity of $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$ can be tied to a source model [26, 40]




- the innovation filters $F_\ell(z)$, $\ell = 1, \ldots, L$ describe the spectral shape of the $L$ contributing source signals;
- a convolutive mixing system $\boldsymbol{H}(z) : \mathbb{C} \rightarrow \mathbb{C}^{M \times N}$ models the transfer paths between the $L$ sources and $M$ sensors;
- if $F_\ell(z)$ and $\boldsymbol{H}(z)$ are stable and causal, then $\boldsymbol{R}(z) = \boldsymbol{H}(z)\boldsymbol{F}(z)\boldsymbol{F}^{\mathrm{P}}(z)\boldsymbol{H}^{\mathrm{P}}(z)$ is analytic.

# Analytic EVD

- Franz Rellich (1939,[28]) for a self-adjoint, analytic $\boldsymbol{R}(t) = \boldsymbol{R}^{\mathrm{H}}(t)$, $t \in \mathbb{R}$:

$$\boldsymbol{R}(t) = \boldsymbol{Q}(t)\boldsymbol{\Lambda}(t)\boldsymbol{Q}^{\mathrm{H}}(t) \; ;$$

- $\boldsymbol{Q}(t)$ and $\boldsymbol{\Lambda}(t)$ *can* be chosen analytic;

- similarly for an arbitrary (i.e. not necessarily Hermitian or square) analytic matrix, de Moor & Boyd (1989, [14]) and Bunse-Gerstner (1991,[10]) established an analytic SVD.

# EVD on the Unit Circle

- Analyticity: $\boldsymbol{R}(z)$ is uniquely definited by its representation on the unit circle, $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{R}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega}}$;
- $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ is self-adjoint: $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{R}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega})$, i.e. Hermitian for every $\Omega$;
- EVD on the unit circle:

$$\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{Q}(\Omega) \cdot \boldsymbol{\Lambda}(\Omega) \cdot \boldsymbol{Q}^{\mathrm{H}}(\Omega) \ . \tag{25}$$

- for every $\Omega$, $\boldsymbol{Q}(\Omega)$ and $\boldsymbol{\Lambda}(\Omega)$ fulfill the properties of the EVD;
- (25) is covered by Rellich [28];
- $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ is $2\pi$-periodic, but the same periodicity cannot be guaranteed for $\boldsymbol{Q}(\Omega)$ and $\boldsymbol{\Lambda}(\Omega)$ [41].

# Matrix Perturbation Theory

- Intuitive explanation of Rellich [28]: if we know that $\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})$ varies smoothly, what can be say about $\boldsymbol{Q}(\Omega)$ and $\boldsymbol{\Lambda}(\Omega)$?
- eigenvalues (Hoffman-Wielandt, 1953,[22]):

$$\sum_i |\lambda_i(\Omega) - \lambda_i(\Omega + \Delta\Omega)| \leq \|\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) - \boldsymbol{R}(\mathrm{e}^{\mathrm{j}(\Omega+\Delta\Omega)})\|_{\mathrm{F}} \ , \tag{26}$$

- subspace distance for eigenvectors / eigenspaces (Golub & van Loan,[18]):

$$\boldsymbol{Q}^{\mathrm{H}}(\Omega)\left(\boldsymbol{R}(\mathrm{e}^{\mathrm{j}(\Omega+\Delta\Omega)}) - \boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega})\right)\boldsymbol{Q}(\Omega) = \left[\begin{array}{cc} \underbrace{\begin{array}{c}\mathbf{E}_{11}(\mathrm{e}^{\mathrm{j}\Omega},\Delta\Omega) \\ \mathbf{E}_{21}(\mathrm{e}^{\mathrm{j}\Omega},\Delta\Omega)\end{array}}_{C} & \underbrace{\begin{array}{c}\mathbf{E}_{21}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega},\Delta\Omega) \\ \mathbf{E}_{22}(\mathrm{e}^{\mathrm{j}\Omega},\Delta\Omega)\end{array}}_{M-C} \end{array}\right] . \tag{27}$$

$$\mathrm{dist}\{\mathcal{Q}_1(\Omega), \mathcal{Q}_1(\Omega + \Delta\Omega)\} \leq \frac{4}{\delta}\|\mathbf{E}_{21}(\mathrm{e}^{\mathrm{j}\Omega}, \Delta\Omega)\|_{\mathrm{F}} \ . \tag{28}$$

# Existence and Uniqueness of an Analytic PhEVD

- If $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$ is analytic, and the data $\mathbf{x}[n]$ does not originate from a multiplexing operation, then we have

$$\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega}) \cdot \boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega}) \cdot \boldsymbol{Q}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega}) \; ; \tag{29}$$

- the factors $\boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega})$ and $\boldsymbol{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega})$ are analytic in $\mathrm{e}^{\mathrm{j}\Omega}$;
- therefore, $\mathbf{Q}[n] \circ\!\!-\!\!\bullet \boldsymbol{Q}(\mathrm{e}^{\mathrm{j}\Omega})$ and $\boldsymbol{\Lambda}[\tau] \circ\!\!-\!\!\bullet \boldsymbol{\Lambda}[\tau]$ are absolutely convergent;
- we can reparameterise (29) as [40]

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z) \cdot \boldsymbol{\Lambda}(z) \cdot \boldsymbol{Q}^{\mathrm{P}}(z) \; ; \tag{30}$$

- the eigenvalues in $\boldsymbol{\Lambda}(z)$ are unique up to a permutation;
- if eigenvalues are distinct, then eigenvectors are unique up to an allpass filter $\Psi_\ell(z)$;
- with $\boldsymbol{\Psi}(z) = \mathrm{diag}\{\Psi_1(z), \ldots, \Psi_M(z)\}$,

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

# Numerical Example for a 2x2 Matrix

- Consider the parahermitian matrix $\boldsymbol{R}(z) = \boldsymbol{U}(z)\boldsymbol{\Gamma}(z)\boldsymbol{U}^{\mathrm{P}}(z)$:

$$\boldsymbol{R}(z) = \begin{bmatrix} \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} & \frac{1+j}{2}z^{2} + \frac{1-j}{2} \\ \frac{1+j}{2} + \frac{1-j}{2}z^{-2} & \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} \end{bmatrix}; \tag{31}$$

- it can be shown that for the eigenvalues,

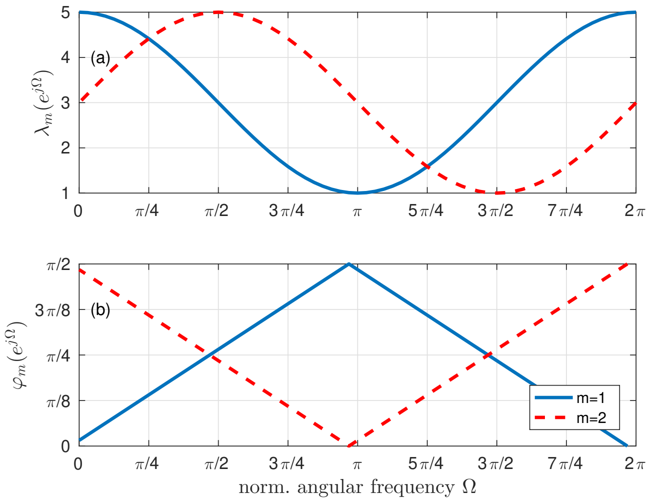
$$\boldsymbol{\Lambda}(z) = \begin{bmatrix} z + 3 + z^{-1} & \\ & -jz + 3 + jz^{-1} \end{bmatrix}; \tag{32}$$

- for the eigenvectors, one possible solution is

$$\boldsymbol{U}(z) = [\boldsymbol{u}_1(z),\, \boldsymbol{u}_2(z)] \qquad \text{with} \qquad \boldsymbol{u}_{1,2}(z) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ \pm z^{-1} \end{bmatrix}; \tag{33}$$

- we'll evaluate on the unit circle, and for the eigenvectors inspect the Hermitian angle $\cos\varphi_m = |\boldsymbol{q}_1^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}0}) \cdot \boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})|$.

# Numerical Example for a 2x2 Matrix cont'd



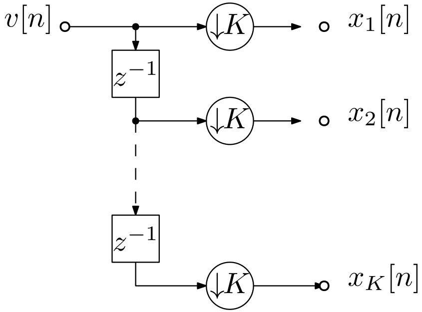

- eigenvalues $\mathbf{\Lambda}(\mathrm{e}^{\mathrm{j}\Omega}) = \mathrm{diag}\{\lambda_1(\mathrm{e}^{\mathrm{j}\Omega})\ \lambda_M(\mathrm{e}^{\mathrm{j}\Omega})\}$;
- Hermitian angles $\cos\varphi_m = |\boldsymbol{q}_1^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}0}) \cdot \boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})|$.

# Non-Existence of an Analytic PhEVD

- Recall due to Rellich [28]

$$\boldsymbol{R}(\mathrm{e}^{\mathrm{j}\Omega}) = \boldsymbol{Q}(\Omega) \cdot \boldsymbol{\Lambda}(\Omega) \cdot \boldsymbol{Q}^{\mathrm{H}}(\Omega) \; ; \tag{34}$$

- if $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ \; \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$ is analytic, but the data $\mathbf{x}[n]$ is $K$-fold multiplexed, then $\boldsymbol{Q}(\Omega)$ and $\boldsymbol{\Lambda}(\Omega)$ will be $2K\pi$ periodic;




- as such, we can only find an analytic EVD if $\boldsymbol{R}(z)$ is $K$-fold oversampled [41]:

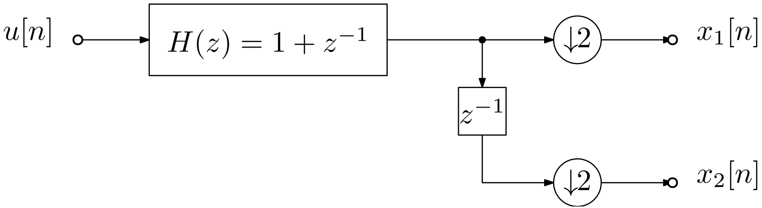
$$\boldsymbol{R}(z^K) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z) \; . \tag{35}$$

# Numerical Example

- Consider the analytic CSD matrix [33, 11]

$$\boldsymbol{R}(z) = \begin{bmatrix} 2 & 1+z^{-1} \\ z+1 & 2 \end{bmatrix} ; \qquad (36)$$

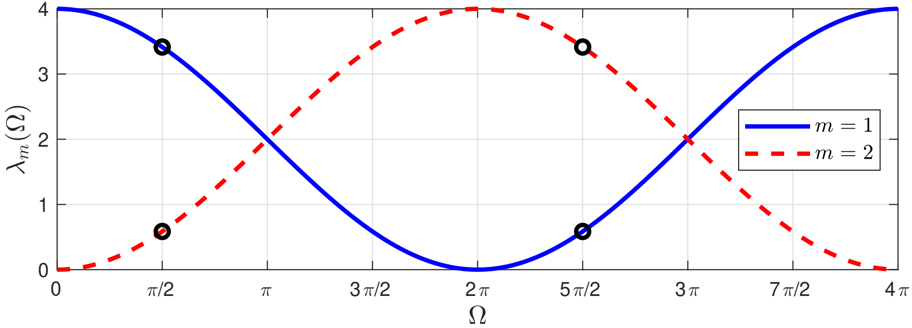
- this is a pseudo-circulant system [34] that can be created by the following multiplexing operation with uncorrelated $u[n] \in \mathcal{N}(0,1)$:



# Numerical Example cont'd

- We can find

$$\boldsymbol{R}(z) = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ z^{-\frac{1}{2}} & z^{-\frac{1}{2}} \end{bmatrix}\begin{bmatrix} z^{\frac{1}{2}}+2+z^{-\frac{1}{2}} & \\ & -z^{\frac{1}{2}}+2-z^{-\frac{1}{2}} \end{bmatrix}\begin{bmatrix} 1 & z^{\frac{1}{2}} \\ -1 & z^{\frac{1}{2}} \end{bmatrix} \; ; \quad (37)$$




- note that the eigenvalues are modulated versions of each other.
- fractional powers of $z$ are not analytic — we need to oversample by two.

# Exact Calculation for a $2 \times 2$ Matrix

- Given an arbitrary parahermitian $\boldsymbol{R}(z) \in \mathbb{C}^{2\times 2}$;
- eigenvalues $\gamma_{1,2}(z)$ can be directly computed in the $z$-domain as the roots of

$$\det\{\gamma(z)\mathbf{I} - \boldsymbol{R}(z)\} = \gamma^2(z) - T(z)\gamma(z) + D(z) = 0$$

- determinant $D(z) = \det\{\boldsymbol{R}(z)\}$ and trace $T(z) = \mathrm{trace}\{\boldsymbol{R}(z)\}$;
- this leads to

$$\gamma_{1,2}(z) = \frac{1}{2}T(z) \pm \frac{1}{2}\sqrt{T(z)T^{\mathrm{P}}(z) - 4D(z)}\;; \tag{38}$$

- awkward: $T(z)T^{\mathrm{P}}(z) - 4D(z) = S(z)S^{\mathrm{P}}(z)$ is parahermitian, but so must be the result of the square root.

# Exact Calculation cont'd

- Maclaurin series: for every root of $S(z)$,

$$\sqrt{1-\beta z^{-1}} = \sum_{n=0}^{\infty} \xi_n \beta^n z^{-n} \tag{39}$$

$$\frac{1}{\sqrt{1-\alpha z^{-1}}} = \left(\sum_{n=0}^{\infty} \xi_n \alpha^n z^{-n}\right)^{-1} = \sum_{n=0}^{\infty} \chi_n \alpha^n z^{-n} \tag{40}$$

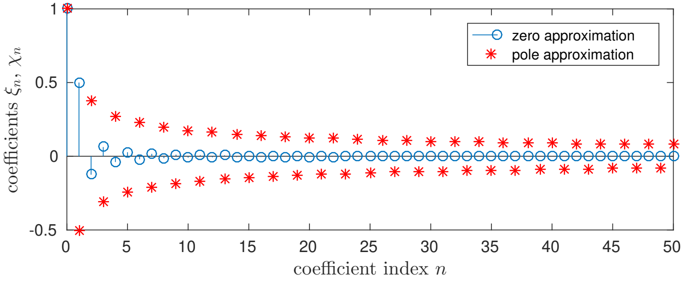
- with coefficients

$$\xi_n = (-1)^n \binom{\frac{1}{2}}{n} = \frac{(-1)^n}{n!} \prod_{i=0}^{n-1} \left(\frac{1}{2} - i\right) \quad , \tag{41}$$

$$\chi_n = (-1)^n \binom{-\frac{1}{2}}{n} = \frac{(-1)^{n-1}}{n!} \prod_{i=0}^{n-1} \left(\frac{1}{2} + i\right) \quad . \tag{42}$$

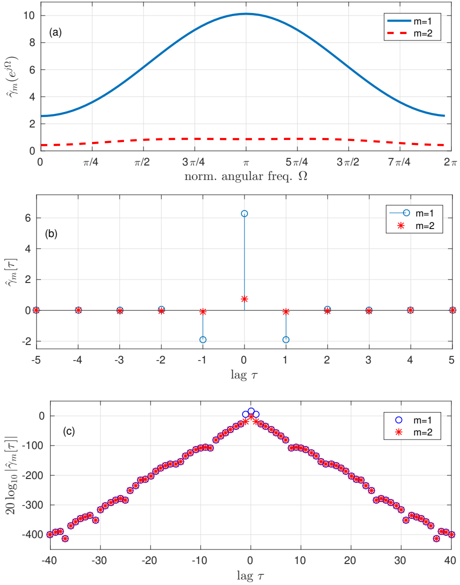
# Maclaurin Series

- Coefficients $\xi_n$ and $\chi_n$ for $n = 0 \ldots 50$:




- these coefficients additionally dampen a geometric series;
- only if $S(z)$ has double zeros (and double poles) is a polynomial (rational) solution possible;
- in general, the result are transcendental eigenvalues.

# Numerical Example




- Example from Icart & Comon (2012,[21]):

$$\boldsymbol{R}(z) = \begin{bmatrix} 1 & 1 \\ 1 & -2z + 6 - 2z^{-1} \end{bmatrix}$$

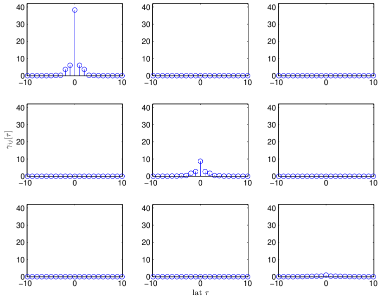
- (a) solution on unit circle;
- (b) coefficients of analytic eigenvalues;
- (c) decay of coefficients.
- solution generally can be transcendental, i.e. neither finite nor rational!

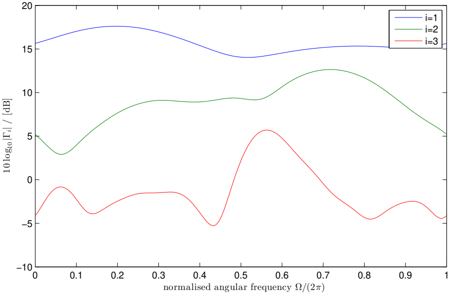
# Polynomial Eigenvalue Decomposition

- Polynomial EVD or McWhirter decomposition [24] of the CSD matrix

$$\boldsymbol{R}(z) \approx \boldsymbol{U}(z)\,\boldsymbol{\Gamma}(z)\,\boldsymbol{U}^{\mathrm{P}}(z) \tag{43}$$

- with paraunitary, polynomial $\boldsymbol{U}(z)$, s.t. $\boldsymbol{U}(z)\boldsymbol{U}^{\mathrm{P}}(z) = \mathbf{I}$;
- diagonalised and spectrally majorised Laurent polynomial $\boldsymbol{\Gamma}(z)$:



# Numerical Example

- We return to the previous example of a parahermitian matrix:

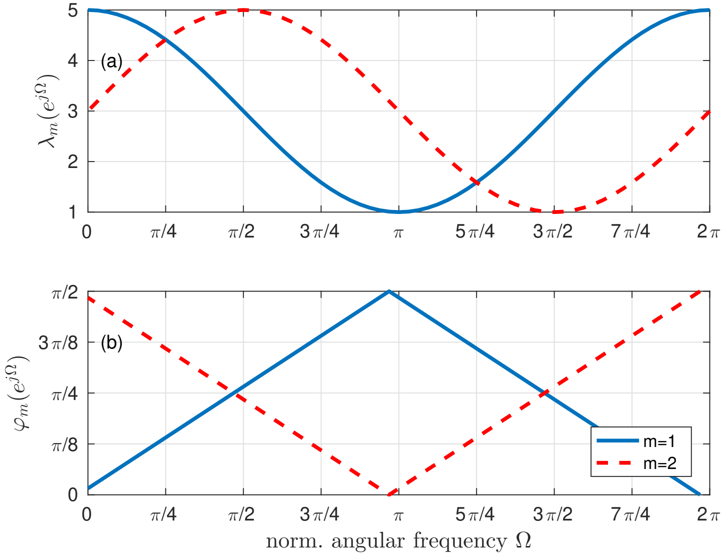
$$\boldsymbol{\Lambda}(z) = \begin{bmatrix} z + 3 + z^{-1} & \\ & -jz + 3 + jz^{-1} \end{bmatrix}$$

$$\boldsymbol{Q}(z) = [\boldsymbol{q}_1(z),\ \boldsymbol{q}_2(z)] \qquad \text{with} \qquad \boldsymbol{q}_{1,2}(z) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \pm z^{-1} \end{bmatrix} ;$$

- parahermitian matrix $\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z)$:

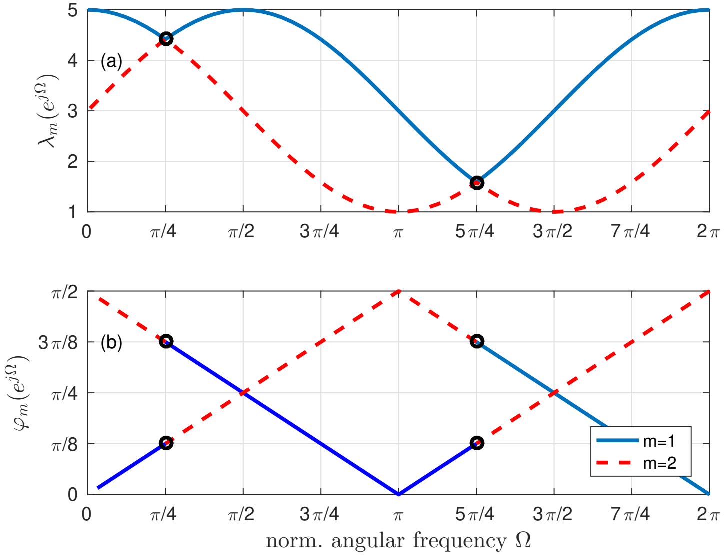
$$\boldsymbol{R}(z) = \begin{bmatrix} \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} & \frac{1+j}{2}z^{2} + \frac{1-j}{2} \\ \frac{1+j}{2} + \frac{1-j}{2}z^{-2} & \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} \end{bmatrix} .$$

# Numerical Example — Analytic Solution



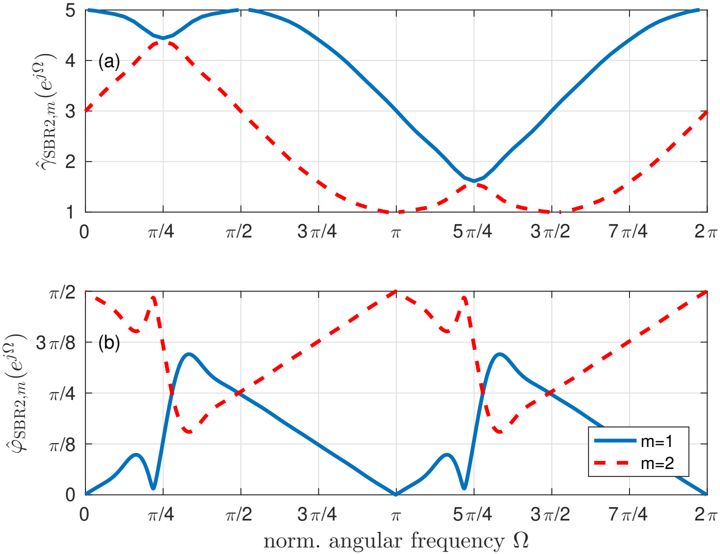

- Recall from earlier:
- analytic (and therefore infinitely differentiable) eigenvalues $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$;
- smooth Hermitian angles $\cos\varphi_m = |\boldsymbol{q}_1^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}0}) \cdot \boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})|$.

# Numerical Example — Ideal Spectral Majorisation




- Analytic eigenvalues are permuted where they intersect;
- resulting spectrally majorised eigenvalues are piecewise analytic but not differentiable;
- corresponding eigenvectors are piecewise analytic but not continuous.

# Numerical Example — PEVD Algorithmic Solution




- Using the SBR2 algorithm in [24] to approximate the McWhirter factorisation;
- spectrally majorised eigenvalues $\mathbf{\Gamma}(z)$ of order 24;
- corresponding eigenvectors in $\boldsymbol{U}(z)$ of order 84.

# Iterative PEVD Algorithms

- Second order sequential best rotation (SBR2, McWhirter 2007, [24]);
- iterative approach based on an elementary paraunitary operation:

$$\begin{aligned} \boldsymbol{S}^{(0)}(z) &= \boldsymbol{R}(z) \\ &\vdots \\ \boldsymbol{S}^{(i+1)}(z) &= \tilde{\boldsymbol{H}}^{(i+1)}(z)\boldsymbol{S}^{(i)}(z)\boldsymbol{H}^{(i+1)}(z) \end{aligned}$$

- $\boldsymbol{H}^{(i)}(z)$ is an elementary paraunitary operation, which at the $i$th step eliminates the largest off-diagonal element in $\boldsymbol{s}^{(i-1)}(z)$;
- stop after $L$ iterations:

$$\hat{\boldsymbol{\Lambda}}(z) = \boldsymbol{S}^{(L)}(z) \qquad , \qquad \boldsymbol{Q}(z) = \prod_{i=1}^{L} \boldsymbol{H}^{(i)}(z)$$

- sequential matrix diagonalisation (SMD) and
- multiple-shift SMD (MS-SMD) will follow the same scheme ...

# Elementary Paraunitary Operation

- An elementary paraunitary matrix [34] is defined as

$$\boldsymbol{H}^{(i)}(z) = \mathbf{I} - \mathbf{v}^{(i)}\mathbf{v}^{(i),\mathrm{H}} + z^{-1}\mathbf{v}^{(i)}\mathbf{v}^{(i),\mathrm{H}} \quad , \quad \|\mathbf{v}^{(i)}\|_2 = 1$$

- we utilise a different definition:

$$\boldsymbol{H}^{(i)}(z) = \boldsymbol{D}^{(i)}(z)\mathbf{Q}^{(i)}$$

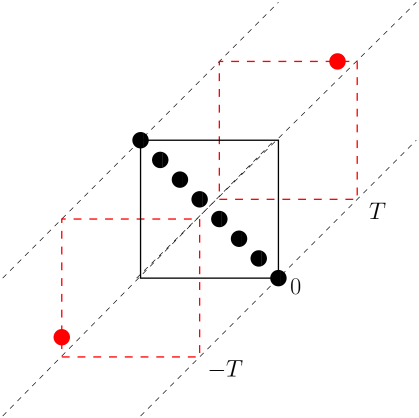
- $\boldsymbol{D}^{(i)}(z)$ is a delay matrix:

$$\boldsymbol{D}^{(i)}(z) = \mathrm{diag}\{1 \; \ldots \; 1 \; z^{-\tau} \; 1 \; \ldots \; 1\}$$

- $\mathbf{Q}^{(i)}(z)$ is a Givens rotation.

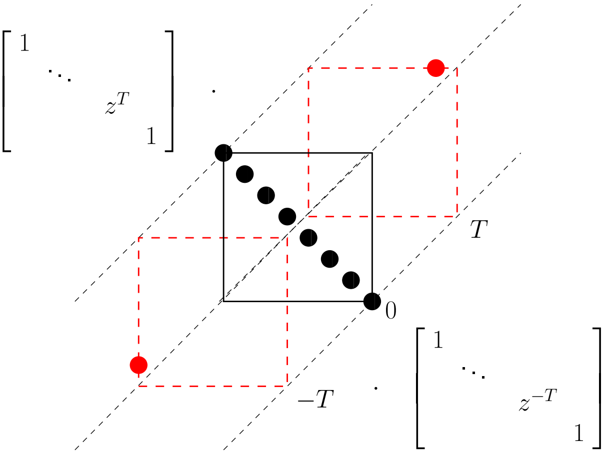
# Sequential Best Rotation Algorithm (McWhirter [45])

- At iteration $i$, consider $\boldsymbol{S}^{(i-1)}(z) \circ\!\!-\!\!\bullet\ \mathbf{S}^{(i-1)}[\tau]$



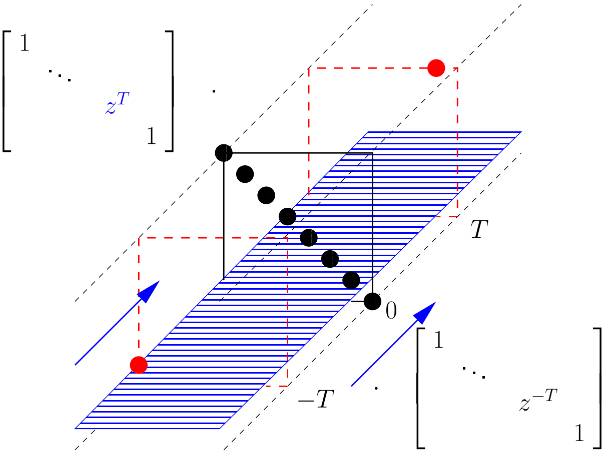
# Sequential Best Rotation Algorithm (McWhirter [45])

- $\tilde{\boldsymbol{D}}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\boldsymbol{D}^{(i)}(z)$



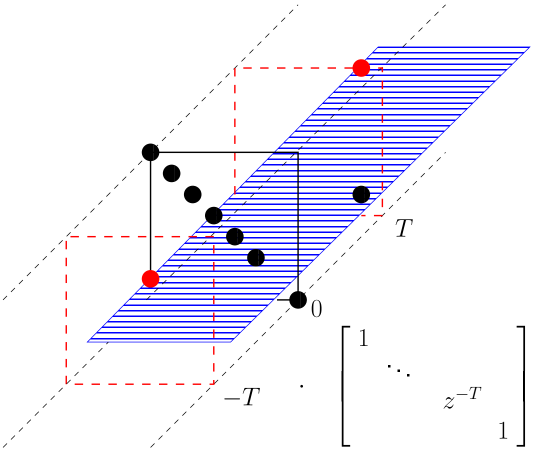
# Sequential Best Rotation Algorithm (McWhirter [45])

- $\tilde{\boldsymbol{D}}^{(i)}(z)$ advances a row-slice of $\boldsymbol{S}^{(i-1)}(z)$ by $T$



# Sequential Best Rotation Algorithm (McWhirter [45])

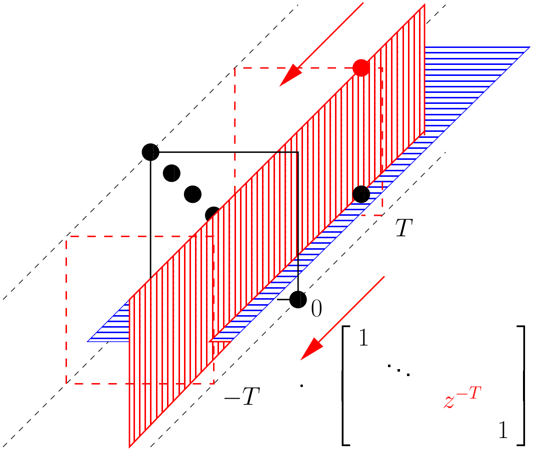
- the off-diagonal element at $-T$ has now been translated to lag zero




$$\cdot \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & z^{-T} & \\ & & & 1 \end{bmatrix}$$

# Sequential Best Rotation Algorithm (McWhirter [45])

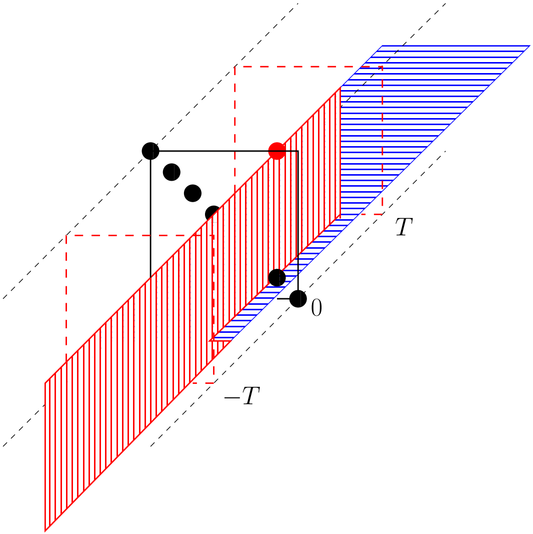
- $\mathbf{D}^{(i)}(z)$ delays a column-slice of $\boldsymbol{S}^{(i-1)}(z)$ by $T$




$$\cdot \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & z^{-T} & \\ & & & 1 \end{bmatrix}$$

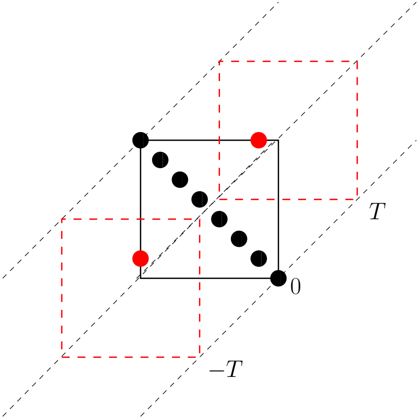
# Sequential Best Rotation Algorithm (McWhirter [45])

- the off-diagonal element at $-T$ has now been translated to lag zero



# Sequential Best Rotation Algorithm (McWhirter [45])

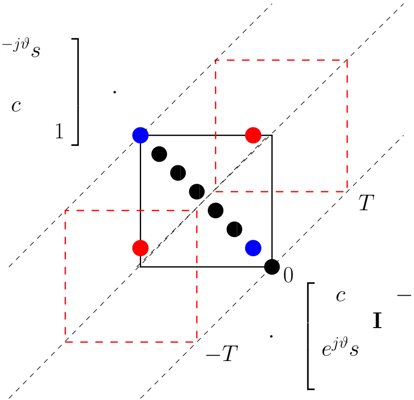
- the step $\tilde{\boldsymbol{D}}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\boldsymbol{D}_{(i)}(z)$ has brought the largest off-diagonal elements to lag 0.



# Sequential Best Rotation Algorithm (McWhirter [45])

- Jacobi step to eliminate largest off-diagonal elements by $\mathbf{Q}^{(i)}$

$$\begin{bmatrix} c & & e^{-j\vartheta}s & \\ & \mathbf{I} & & \\ -e^{j\vartheta}s & & c & \\ & & & 1 \end{bmatrix} \cdot$$

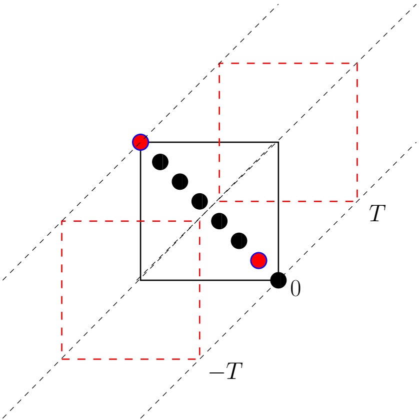



$$\cdot \begin{bmatrix} c & & -e^{-j\vartheta}s & \\ & \mathbf{I} & & \\ e^{j\vartheta}s & & c & \\ & & & 1 \end{bmatrix}$$

# Sequential Best Rotation Algorithm (McWhirter [45])

- iteration $i$ is completed, having performed

$$\boldsymbol{S}^{(i)}(z) = \mathbf{Q}^{(i)}\boldsymbol{D}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\tilde{\boldsymbol{D}}^{(i)}(z)\tilde{\mathbf{Q}}^{(i)}(z)$$



# SBR2 Outcome

- At the $i$th iteration, the zeroing of off-diagonal elements achieved during previous steps may be partially undone;
- however, the algorithm has been shown to converge, transfering energy onto the main diagonal at every step (McWhirter 2007);
- after $L$ iterations, we reach an approximate diagonalisation

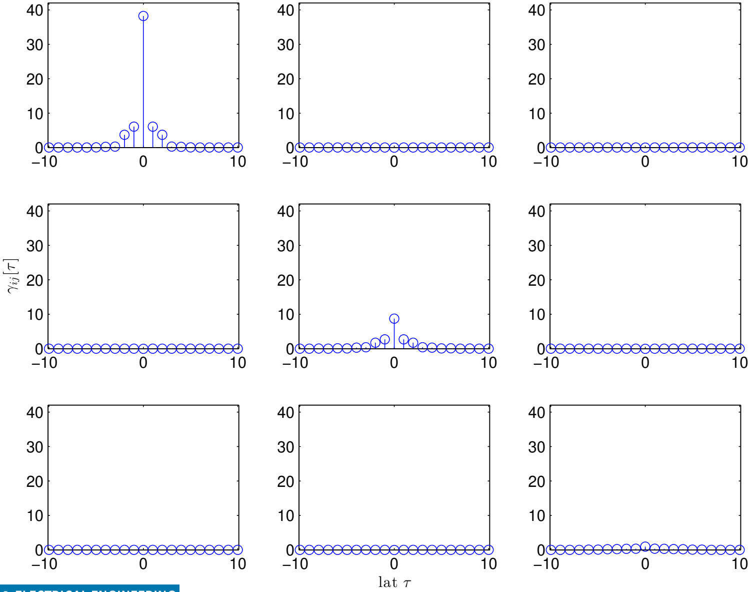
$$\hat{\mathbf{\Lambda}}(z) = \boldsymbol{S}^{(L)}(z) = \tilde{\boldsymbol{Q}}(z)\boldsymbol{R}(z)\boldsymbol{Q}(z)$$

  with

$$\boldsymbol{Q}(z) = \prod_{i=1}^{L} \boldsymbol{D}^{(i)}(z)\mathbf{Q}^{(i)}$$

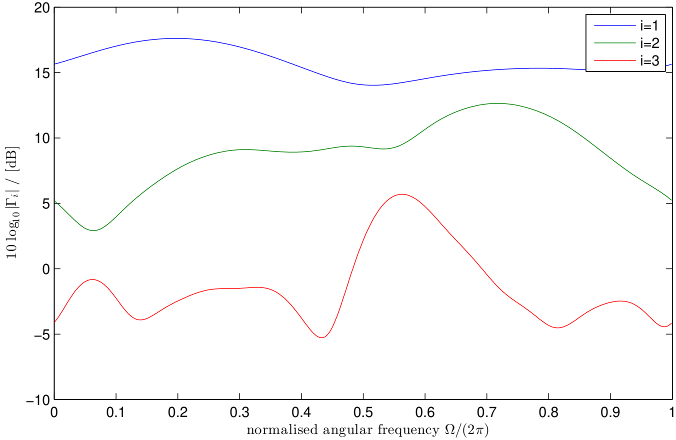
- diagonalisation of the previous $3 \times 3$ polynomial matrix ...

# SBR2 Example — Diagonalisation



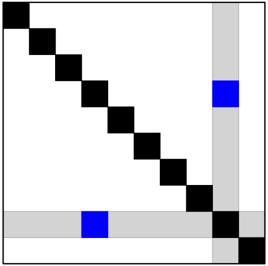
# SBR2 Example — Spectral Majorisation

- The on-diagonal elements are spectrally majorised



# SBR2 — Givens Rotation

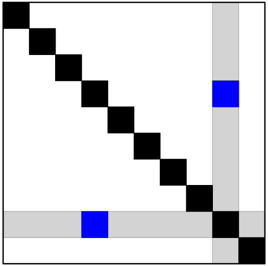
- A Givens rotation eliminates the maximum off-diagonal element once brought onto the lag-zero matrix;
- note I: in the lag-zero matrix, one column and one row are modified by the shift:

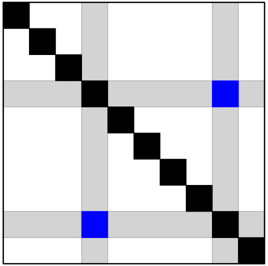


- note II: a Givens rotation only affects two columns and two rows in every matrix;
- Givens rotation is relatively low in computational cost!

# SBR2 — Givens Rotation

- A Givens rotation eliminates the maximum off-diagonal element once brought onto the lag-zero matrix;
- note I: in the lag-zero matrix, one column and one row are modified by the shift:

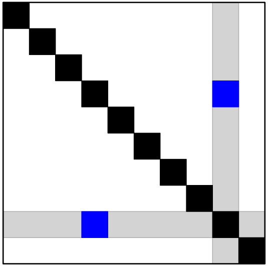




- note II: a Givens rotation only affects two columns and two rows in every matrix;
- Givens rotation is relatively low in computational cost!

# Sequential Matrix Diagonalisation (SMD, [27])

- Main idea — the zero-lag matrix is diagonalised in every step;
- initialisation: diagonalise $\mathbf{R}[0]$ by EVD and apply modal matrix to all matrix coefficients $\longrightarrow \boldsymbol{S}^{(0)}$;
- at the $i$th step as in SBR2, the maximum element (or column with max. norm) is shifted to the lag-zero matrix:



- an EVD is used to re-diagonalise the zero-lag matrix;
- a full modal matrix is applied at all lags — more costly than SBR2.

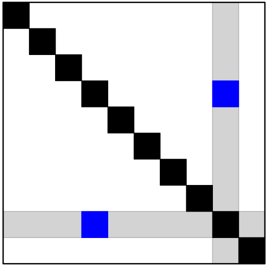
# Sequential Matrix Diagonalisation (SMD, [27])

- Main idea — the zero-lag matrix is diagonalised in every step;
- initialisation: diagonalise $\mathbf{R}[0]$ by EVD and apply modal matrix to all matrix coefficients $\longrightarrow \boldsymbol{S}^{(0)}$;
- at the $i$th step as in SBR2, the maximum element (or column with max. norm) is shifted to the lag-zero matrix:

- an EVD is used to re-diagonalise the zero-lag matrix;
- a full modal matrix is applied at all lags — more costly than SBR2.

# Multiple Shift SMD (SMD)

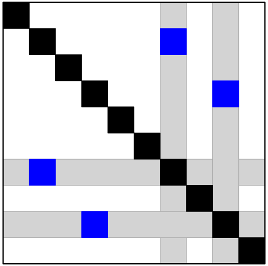
- SMD converges faster than SBR2 — more energy is transfered per iteration step;
- SMD is more expensive than SBR2 — full matrix multiplication at every lag;
- this cost will not increase further if more columns / rows are shifted into the lag-zero matrix at every iteration



- MS-SMD will transfer yet more off-diagonal energy per iteration;
- because the total energy must remain constant under paraunitary operations, SBR2, SMD and MS-SMD can be proven to converge.

# Multiple Shift SMD (SMD)

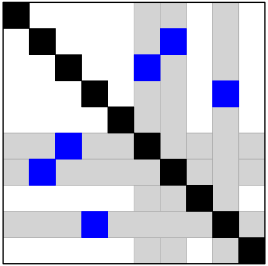
- SMD converges faster than SBR2 — more energy is transfered per iteration step;
- SMD is more expensive than SBR2 — full matrix multiplication at every lag;
- this cost will not increase further if more columns / rows are shifted into the lag-zero matrix at every iteration



- MS-SMD will transfer yet more off-diagonal energy per iteration;
- because the total energy must remain constant under paraunitary operations, SBR2, SMD and MS-SMD can be proven to converge.

# Multiple Shift SMD (SMD)

- SMD converges faster than SBR2 — more energy is transfered per iteration step;
- SMD is more expensive than SBR2 — full matrix multiplication at every lag;
- this cost will not increase further if more columns / rows are shifted into the lag-zero matrix at every iteration



- MS-SMD will transfer yet more off-diagonal energy per iteration;
- because the total energy must remain constant under paraunitary operations, SBR2, SMD and MS-SMD can be proven to converge.

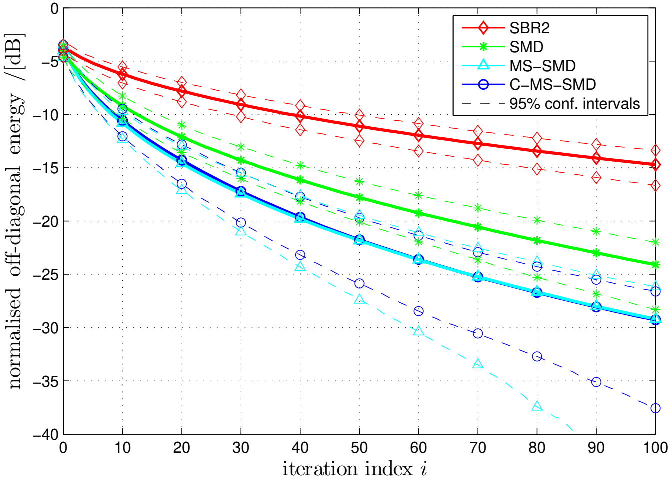
# Multiple Shift SMD (SMD)

- SMD converges faster than SBR2 — more energy is transfered per iteration step;
- SMD is more expensive than SBR2 — full matrix multiplication at every lag;
- this cost will not increase further if more columns / rows are shifted into the lag-zero matrix at every iteration

- MS-SMD will transfer yet more off-diagonal energy per iteration;
- because the total energy must remain constant under paraunitary operations, SBR2, SMD and MS-SMD can be proven to converge.

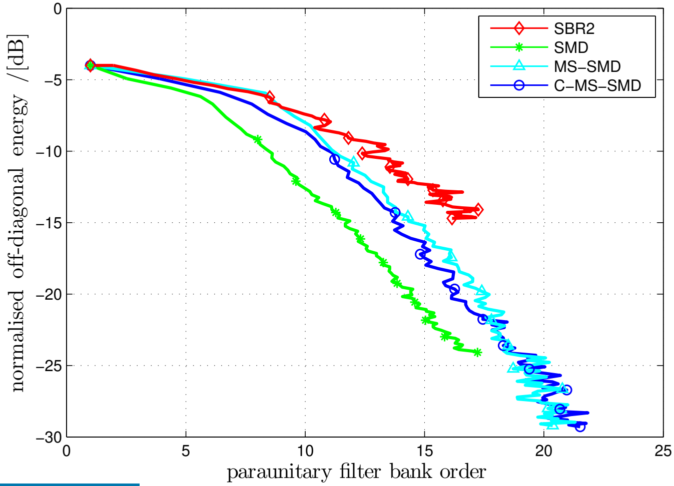
# SBR2/SMD/MS-SMD Convergence

- Measuring the remaining normalised off-diagonal energy over an ensemble of space-time covariance matrices:



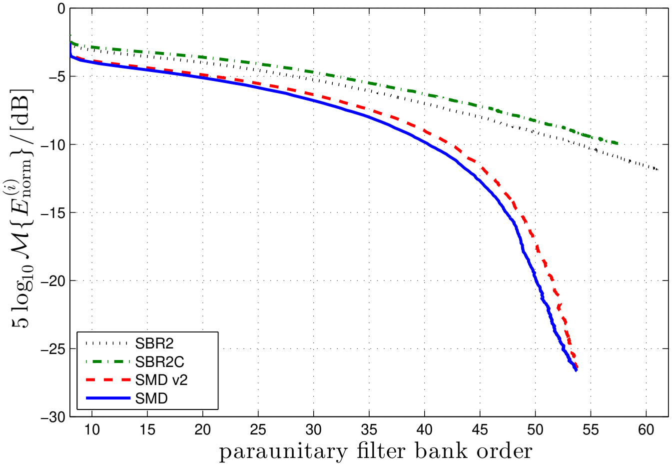
# SBR2/SMD/MS-SMD Application Cost 1

- Ensemble average of remaining off-diagonal energy vs. order of paraunitary filter banks to decompose 4x4x16 matrices:



# SBR2/SMD/MS-SMD Application Cost 2

- Ensemble average of remaining off-diagonal energy vs. order of paraunitary filter banks to decompose 8x8x64 matrices:

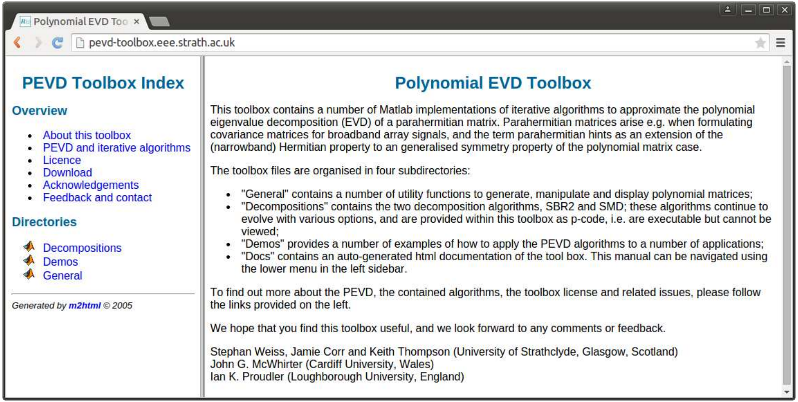


# MATLAB Polynomial EVD Toolbox

University of Strathclyde Engineering

- The MATLAB polynomial EVD toolbox can be downloaded from

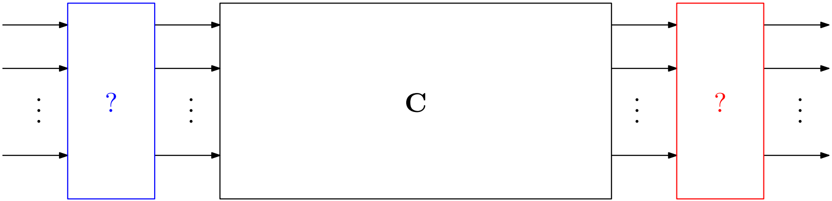
  `pevd-toolbox.eee.strath.ac.uk`




- the toolbox contains a number of iterative algorithms to calculate an approximate PEVD, related functions, and demos.

# Narrowband MIMO Communications

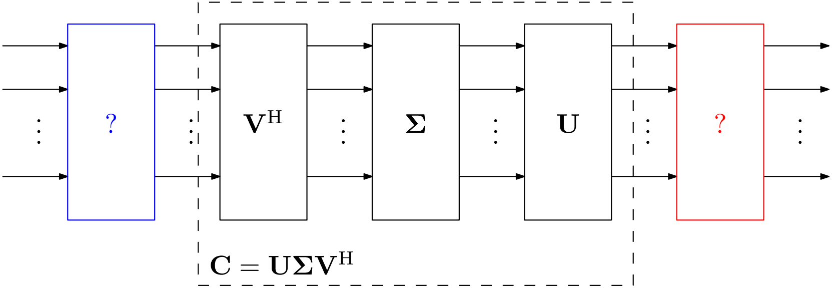
- a narrowband channel is characterised by a matrix $\mathbf{C}$ containing complex gain factors;
- problem: how to select the precoder and equaliser?




- overall system;

# Narrowband MIMO Communications

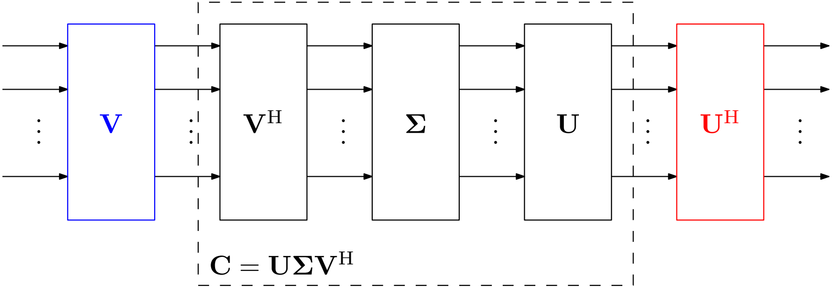
- a narrowband channel is characterised by a matrix $\mathbf{C}$ containing complex gain factors;
- problem: how to select the precoder and equaliser?




- the singular value decomposition (SVD) factorises $\mathbf{C}$ into two unitary matrices $\mathrm{U}$ and $\mathbf{V}^{\mathrm{H}}$ and a diagonal matrix $\mathbf{\Sigma}$;

# Narrowband MIMO Communications

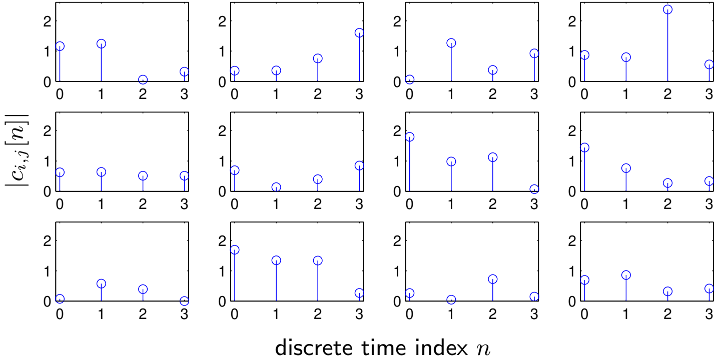
- a narrowband channel is characterised by a matrix $\mathbf{C}$ containing complex gain factors;
- problem: how to select the precoder and equaliser?




- we select the precoder and the equaliser from the unitary matrices provided by the channel's SVD;
- the overall system is diagonalised, decoupling the channel into independent single-input single-output systems by means of unitary matrices.

# Broadband MIMO Channel

- The channel is a matrix of FIR filters; example for a $3 \times 4$ system $\mathbf{C}[n]$:

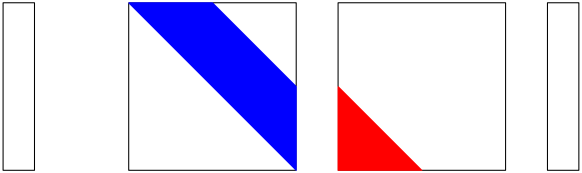



- the transfer function $\boldsymbol{C}(z)$ $\bullet\!\!-\!\!\circ$ $\mathbf{C}[n]$ is a polynomial matrix;
- an SVD can only diagonalise $\mathbf{C}[n]$ for one particular lag $n$.

# Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known [20]):
  1. divide spectrum into narrowband channels;
  2. address each narrowband channel independently using narrowband-optimal techniques;

  drawback: ignores spectral coherence across frequency bins;
- optimum filter bank transceiver (if channel itself is known [30, 31, 29]):
  1. block processing;
  2. inter-block interference is eliminated by guard intervals;
  3. resulting matrix can be diagonalised by SVD;
- both techniques invest DOFs into the guard intervals, which are generally not balanced against other error sources.

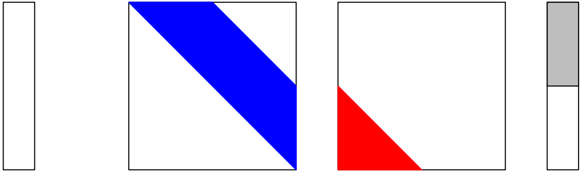


$$\mathbf{y}(z) = ( \mathbf{C}_0 + \mathbf{C}_1 z^{-1} ) \cdot \mathbf{x}(z)$$

# Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known [20]):
  1. divide spectrum into narrowband channels;
  2. address each narrowband channel independently using narrowband-optimal techniques;

  drawback: ignores spectral coherence across frequency bins;
- optimum filter bank transceiver (if channel itself is known [30, 31, 29]):
  1. block processing;
  2. inter-block interference is eliminated by guard intervals;
  3. resulting matrix can be diagonalised by SVD;
- both techniques invest DOFs into the guard intervals, which are generally not balanced against other error sources.

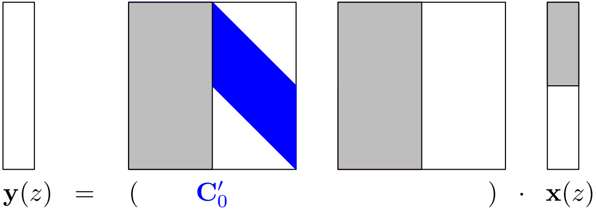


$$\mathbf{y}(z) = \left( \mathbf{C}_0 + \mathbf{C}_1 z^{-1} \right) \cdot \mathbf{x}(z)$$

# Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known [20]):
  1. divide spectrum into narrowband channels;
  2. address each narrowband channel independently using narrowband-optimal techniques;

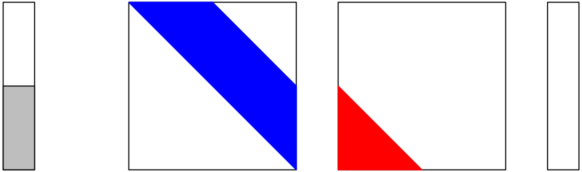
  drawback: ignores spectral coherence across frequency bins;
- optimum filter bank transceiver (if channel itself is known [30, 31, 29]):
  1. block processing;
  2. inter-block interference is eliminated by guard intervals;
  3. resulting matrix can be diagonalised by SVD;
- both techniques invest DOFs into the guard intervals, which are generally not balanced against other error sources.



# Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known [20]):
  1. divide spectrum into narrowband channels;
  2. address each narrowband channel independently using narrowband-optimal techniques;

  drawback: ignores spectral coherence across frequency bins;
- optimum filter bank transceiver (if channel itself is known [30, 31, 29]):
  1. block processing;
  2. inter-block interference is eliminated by guard intervals;
  3. resulting matrix can be diagonalised by SVD;
- both techniques invest DOFs into the guard intervals, which are generally not balanced against other error sources.

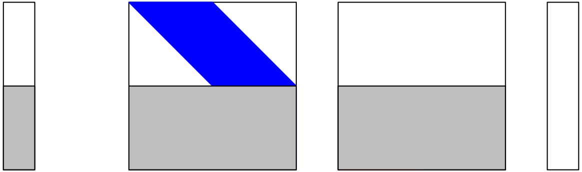


$$\mathbf{y}(z) \;=\; \left( \quad \mathbf{C}_0 \quad + \quad \mathbf{C}_1 z^{-1} \quad \right) \cdot \mathbf{x}(z)$$

# Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known [20]):
  1. divide spectrum into narrowband channels;
  2. address each narrowband channel independently using narrowband-optimal techniques;

  drawback: ignores spectral coherence across frequency bins;
- optimum filter bank transceiver (if channel itself is known [30, 31, 29]):
  1. block processing;
  2. inter-block interference is eliminated by guard intervals;
  3. resulting matrix can be diagonalised by SVD;
- both techniques invest DOFs into the guard intervals, which are generally not balanced against other error sources.



$$\mathbf{y}(z) = ( \quad \mathbf{C}'_0 \quad ) \cdot \mathbf{x}(z)$$

# Polynomial Singular Value Decompositions

- Iterative algorithms have been developed to determine a polynomial eigenvalue decomposition (EVD) for a parahermitian matrix $\boldsymbol{R}(z) = \boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}^{\mathrm{H}}(z^{-1})$:

$$\boldsymbol{R}(z) \approx \boldsymbol{H}(z)\boldsymbol{\Gamma}(z)\boldsymbol{H}^{\mathrm{P}}(z)$$

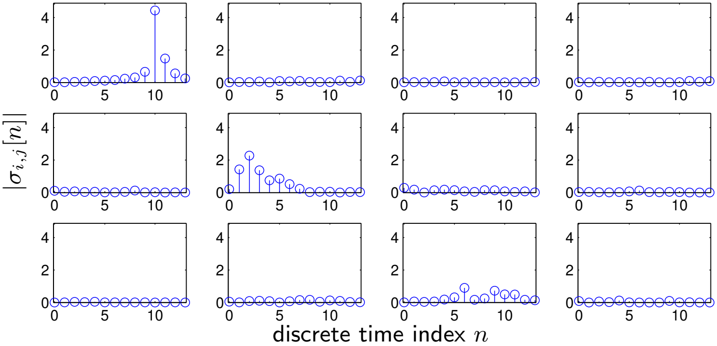
- paraunitary $\boldsymbol{H}(z)\boldsymbol{H}^{\mathrm{P}}(z) = \mathbf{I}$, diagonal and spectrally majorised $\boldsymbol{\Gamma}(z)$;
- polynomial SVD of channel $\boldsymbol{C}(z)$ can be obtained via two EVDs:

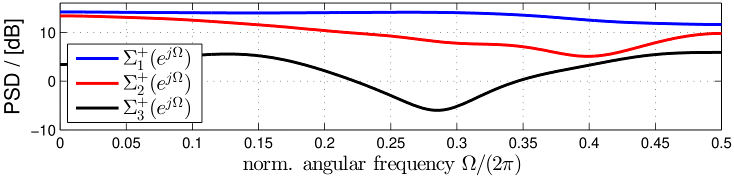
$$\boldsymbol{C}(z)\boldsymbol{C}^{\mathrm{P}}(z) = \boldsymbol{U}(z)\boldsymbol{\Sigma}^{+}(z)\boldsymbol{\Sigma}^{-}(z)\boldsymbol{U}^{\mathrm{P}}(z)$$
$$\boldsymbol{C}^{\mathrm{P}}(z)\boldsymbol{C}(z) = \boldsymbol{V}(z)\boldsymbol{\Sigma}^{-}(z)\boldsymbol{\Sigma}^{+}(z)\boldsymbol{V}^{\mathrm{P}}(z)$$

finally:

$$\boldsymbol{C}(z) = \boldsymbol{U}(z)\boldsymbol{\Sigma}^{+}(z)\boldsymbol{V}^{\mathrm{P}}(z)\ .$$

# MIMO Application Example

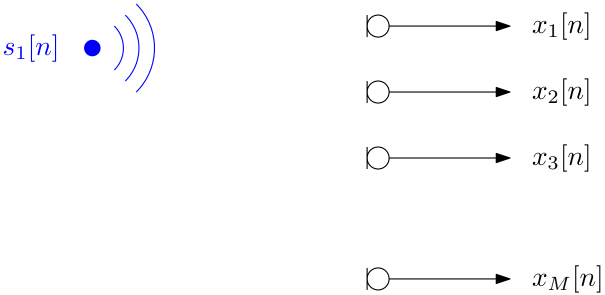



- Polynomial SVD of the previous $\boldsymbol{C}(z) : \mathbb{C} \rightarrow \mathbb{C}^{3\times 4}$;
- the singular value spectra are majorised.

# Narrowband Source Model

- Scenario with sensor array and far-field sources:

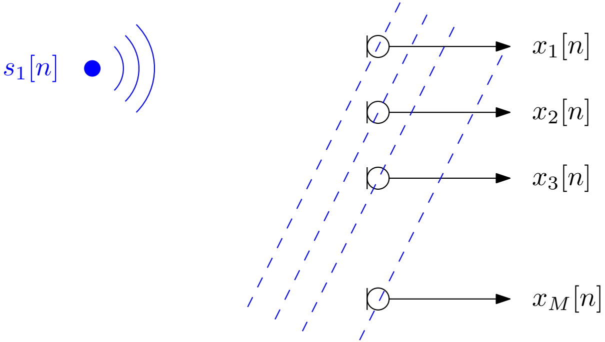



- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector
- data model:

$$\mathbf{x}[n] =$$

# Narrowband Source Model

- Scenario with sensor array and far-field sources:

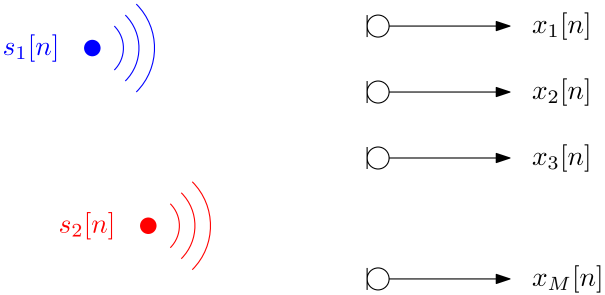



- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector $\mathbf{s}_1$
- data model:

$$\mathbf{x}[n] = s_1[n] \cdot \mathbf{s}_1$$

# Narrowband Source Model

- Scenario with sensor array and far-field sources:

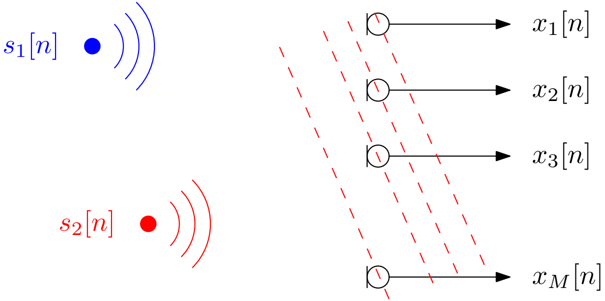



- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector $\mathbf{s}_1$
- data model:

$$\mathbf{x}[n] = s_1[n] \cdot \mathbf{s}_1$$

# Narrowband Source Model

- Scenario with sensor array and far-field sources:

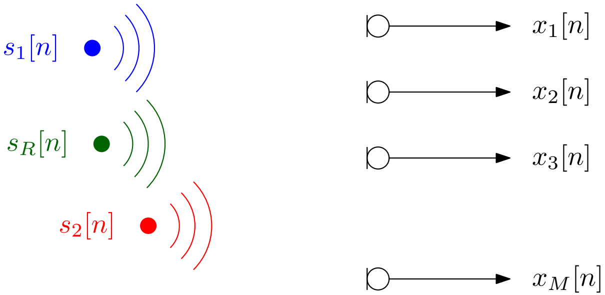



- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector $\mathbf{s}_1$, $\mathbf{s}_2$
- data model:

$$\mathbf{x}[n] = s_1[n] \cdot \mathbf{s}_1 + s_1[n] \cdot \mathbf{s}_2$$

# Narrowband Source Model

- Scenario with sensor array and far-field sources:

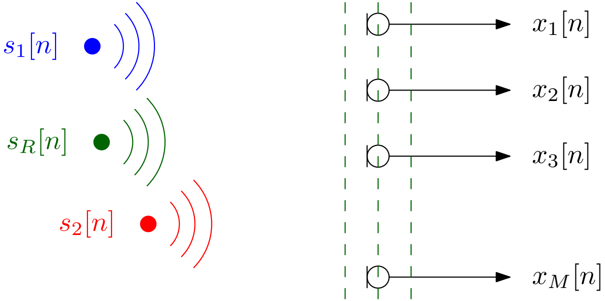



- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector $\mathbf{s}_1$, $\mathbf{s}_2$
- data model:

$$\mathbf{x}[n] = s_1[n] \cdot \mathbf{s}_1 + s_1[n] \cdot \mathbf{s_2}$$

# Narrowband Source Model

- Scenario with sensor array and far-field sources:




- for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector $\mathbf{s}_1$, $\mathbf{s}_2$, $\ldots \mathbf{s}_R$;
- data model:

$$\mathbf{x}[n] = s_1[n] \cdot \mathbf{s}_1 + s_1[n] \cdot \mathbf{s_2} + \cdots + s_R[n] \cdot \mathbf{s}_R = \sum_{r=1}^{R} s_r[n] \cdot \mathbf{s}_r$$

# Steering Vector

- A signal $s[n]$ arriving at the array can be characterised by the delays of its wavefront (neglecting attenuation):

$$\begin{bmatrix} x_0[n] \\ x_1[n] \\ \vdots \\ x_{M-1}[n] \end{bmatrix} = \begin{bmatrix} s[n-\tau_0] \\ s[n-\tau_1] \\ \vdots \\ s[n-\tau_{M-1}] \end{bmatrix} = \begin{bmatrix} \delta[n-\tau_0] \\ \delta[n-\tau_1] \\ \vdots \\ \delta[n-\tau_{M-1}] \end{bmatrix} * s[n] \circ\!\!-\!\!\bullet\ \mathbf{a}_\vartheta(z)S(z)$$

- if evaluated at a narrowband normalised angular frequency $\Omega_i$, the time delays $\tau_m$ in the broadband steering vector $\mathbf{a}_\vartheta(z)$ collapse to phase shifts in the narrowband steering vector $\mathbf{a}_{\vartheta,\Omega_i}$,

$$\mathbf{a}_{\vartheta,\Omega_i} = \mathbf{a}_\vartheta(z)|_{z=e^{j\Omega_i}} = \begin{bmatrix} e^{-j\tau_0\Omega_i} \\ e^{-j\tau_1\Omega_i} \\ \vdots \\ e^{-j\tau_{M-1}\Omega_i} \end{bmatrix} .$$

# Data and Covariance Matrices

- A data matrix $\mathbf{X} \in \mathbb{C}^{M \times L}$ can be formed from $L$ measurements:

$$\mathbf{X} = \left[ \begin{array}{cccc} \mathbf{x}[n] & \mathbf{x}[n+1] & \ldots & \mathbf{x}[n+L-1] \end{array} \right]$$

- assuming that all $x_m[n]$, $m = 1, 2, \ldots M$ are zero mean, the (instantaneous) data covariance matrix is

$$\mathbf{R} = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\} \approx \frac{1}{L}\mathbf{X}\mathbf{X}^{\mathrm{H}}$$

  where the approximation assumes ergodicity and a sufficiently large $L$;
- Problem: can we tell from $\mathbf{X}$ or $\mathbf{R}$ (i) the number of sources and (ii) their orgin / time series?
- w.r.t. Jonathon Chamber's introduction, we here only consider the underdetermined case of more sensors than sources, $M \geq K$, and generally $L \gg M$.

# SVD of Data Matrix

University of Strathclyde Engineering

- Singular value decomposition of $\mathbf{X}$:

$$\boxed{\mathbf{X}} = \boxed{\mathbf{U}}\;\boxed{\boldsymbol{\Sigma}}\;\boxed{\mathbf{V}^{\mathrm{H}}}$$

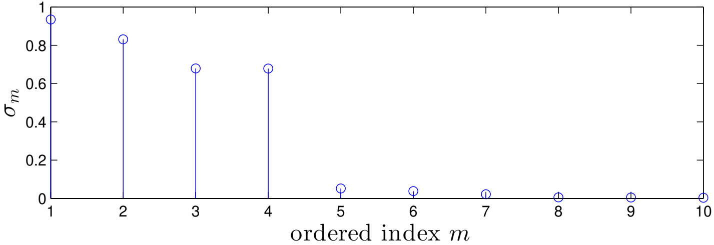
- unitary matrices $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_M]$ and $\mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_L]$;
- diagonal $\boldsymbol{\Sigma}$ contains the real, positive semidefinite singular values of $\mathbf{X}$ in descending order:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & & 0 & \sigma_M & 0 & \ldots & 0 \end{bmatrix}$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_M \geq 0$.

# Singular Values

- If the array is illuminated by $R \leq M$ linearly independent sources, the rank of the data matrix is
$$\mathrm{rank}\{\mathbf{X}\} = R$$
- only the first $R$ singular values of $\mathbf{X}$ will be non-zero;
- in practice, noise often will ensure that $\mathrm{rank}\{\mathbf{X}\} = M$, with $M - R$ trailing singular values that define the noise floor:




- therefore, by thresholding singular values, it is possible to estimate the number of linearly independent sources $R$.

# Subspace Decomposition

- If $\operatorname{rank}\{\mathbf{X}\} = R$, the SVD can be split:

$$\mathbf{X} = [\mathbf{U}_s \;\; \mathbf{U}_n] \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_n \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^{\mathrm{H}} \\ \mathbf{V}_n^{\mathrm{H}} \end{bmatrix}$$

- with $\mathbf{U}_s \in \mathbb{C}^{M\times R}$ and $\mathbf{V}_s^{\mathrm{H}} \in \mathbb{C}^{R\times L}$ corresponding to the $R$ largest singular values;
- $\mathbf{U}_s$ and $\mathbf{V}_s^{\mathrm{H}}$ define the signal-plus-noise subspace of $\mathbf{X}$:

$$\mathbf{X} = \sum_{m=1}^{M} \sigma_m \mathbf{u}_m \mathbf{v}_m^{\mathrm{H}} \approx \sum_{m=1}^{R} \sigma_m \mathbf{u}_m \mathbf{v}_m^{\mathrm{H}}$$

- the complements $\mathbf{U}_n$ and $\mathbf{V}_n^{\mathrm{H}}$,

$$\mathbf{U}_s^{\mathrm{H}} \mathbf{U}_n = \mathbf{0} \qquad , \qquad \mathbf{V}_s \mathbf{V}_n^{\mathrm{H}} = \mathbf{0}$$

define the noise-only subspace of $\mathbf{X}$.

# SVD via Two EVDs

- Any Hermitian matrix $\mathbf{A} = \mathbf{A}^{\mathrm{H}}$ allows an eigenvalue decomposition

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}}$$

with $\mathbf{Q}$ unitary and the eigenvalues in $\mathbf{\Lambda}$ real valued and positive semi-definite;

- postulating $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}$, therefore:

$$\begin{aligned} \mathbf{X}\mathbf{X}^{\mathrm{H}} &= (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}})(\mathbf{V}\mathbf{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}) = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}} && (44) \\ \mathbf{X}^{\mathrm{H}}\mathbf{X} &= (\mathbf{V}\mathbf{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}})(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}) = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{H}} && (45) \end{aligned}$$

- (ordered) eigenvalues relate to the singular values: $\lambda_m = \sigma_m^2$;
- the covariance matrix $\mathbf{R} = \frac{1}{L}\mathbf{X}\mathbf{X}$ has the same rank as the data matrix $\mathbf{X}$, and with $\mathbf{U}$ provides access to the same spatial subspace decomposition.

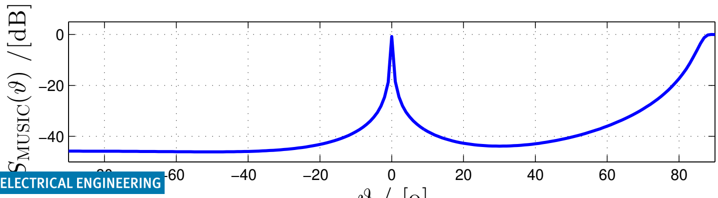
# Narrowband MUSIC Algorithm

- EVD of the narrowband covariance matrix identifies signal-plus-noise and noise-only subspaces

$$\mathbf{R} = [\mathbf{U}_s \;\; \mathbf{U}_n] \begin{bmatrix} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^{\mathrm{H}} \\ \mathbf{U}_n^{\mathrm{H}} \end{bmatrix}$$

- scanning the signal-plus-noise subspace could only help to retrieve sources with orthogonal steering vectors;
- therefore, the multiple signal classification (MUSIC) algorithm scans the noise-only subspace for minima, or maxima of its reciprocal

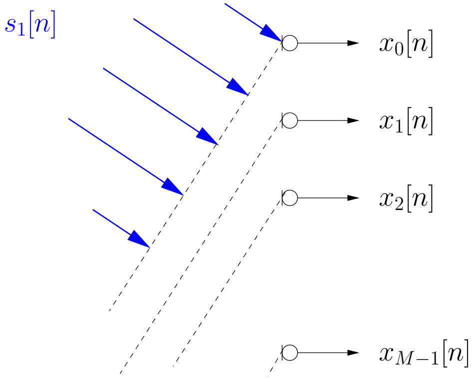
$$S_{\mathrm{MUSIC}}(\vartheta) = \frac{1}{\|\mathbf{U}_n \mathbf{a}_{\vartheta,\Omega_i}\|_2^2}$$



# Narrowband Source Separation

- Via SVD of the data matrix $\mathbf{X}$ or EVD of the covariance matrix $\mathbf{R}$, we can determine the number of linearly independent sources $R$;
- using the subspace decompositions offered by EVD/SVD, the directions of arrival can be estimated using e.g. MUSIC;
- based on knowledge of the angle of arrival, beamforming could be applied to $\mathbf{X}$ to extract specific sources;
- overall: EVD (and SVD) can play a vital part in narrowband source separation;
- what about broadband source separation?

# Broadband Array Scenario




- Compared to the narrowband case, time delays rather than phase shifts bear information on the direction of a source.

# Broadband Steering Vector

- A signal $s[n]$ arriving at the array can be characterised by the delays of its wavefront (neglecting attenuation):

$$\begin{bmatrix} x_0[n] \\ x_1[n] \\ \vdots \\ x_{M-1}[n] \end{bmatrix} = \begin{bmatrix} s[n-\tau_0] \\ s[n-\tau_1] \\ \vdots \\ s[n-\tau_{M-1}] \end{bmatrix} = \begin{bmatrix} \delta[n-\tau_0] \\ \delta[n-\tau_1] \\ \vdots \\ \delta[n-\tau_{M-1}] \end{bmatrix} * s[n] \circ\!\!-\!\!\bullet\; \mathbf{a}_\vartheta(z) S(z)$$

- if evaluated at a narrowband normalised angular frequency $\Omega_i$, the time delays $\tau_m$ in the broadband steering vector $\mathbf{a}_\vartheta(z)$ collapse to phase shifts in the narrowband steering vector $\mathbf{a}_{\vartheta,\Omega_i}$,

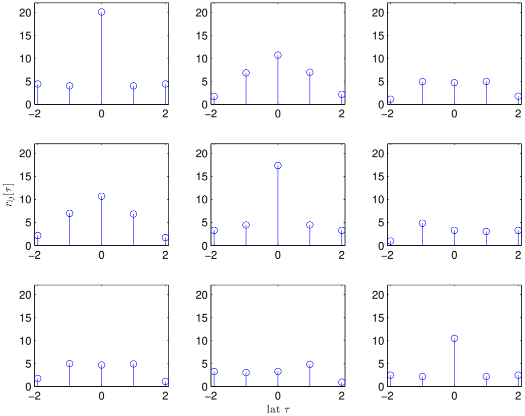
$$\mathbf{a}_{\vartheta,\Omega_i} = \mathbf{a}_\vartheta(z)|_{z=e^{j\Omega_i}} = \begin{bmatrix} e^{-j\tau_0\Omega_i} \\ e^{-j\tau_1\Omega_i} \\ \vdots \\ e^{-j\tau_{M-1}\Omega_i} \end{bmatrix} \; .$$

# Space-Time Covariance Matrix

- If delays must be considered, the (space-time) covariance matrix must capture the lag $\tau$:

$$\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n] \cdot \mathbf{x}^{\mathrm{H}}[n-\tau]\}$$

- $\mathbf{R}[\tau]$ contains auto- and cross-correlation sequences:



# Cross Spectral Density Matrix

- $z$-transform of the space-time covariance matrix is given by

$$\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}_n \mathbf{x}^{\mathrm{H}}_{n-\tau}\} \quad \circ\!\!-\!\!\bullet \quad \boldsymbol{R}(z) = \sum_l S_l(z)\mathbf{a}_{\vartheta_l}(z)\tilde{\mathbf{a}}_{\vartheta_l}(z) + \sigma_N^2 \mathbf{I}$$

  with $\vartheta_l$ the direction of arrival and $S_l(z)$ the PSD of the $l$th source;
- $\boldsymbol{R}(z)$ is the cross spectral density (CSD) matrix;
- the instantaneous covariance matrix (no lag parameter $\tau$)

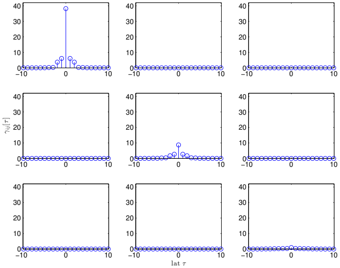
$$\mathbf{R} = \mathcal{E}\{\mathbf{x}_n \mathbf{x}^{\mathrm{H}}_n\} = \mathbf{R}[0] \ .$$

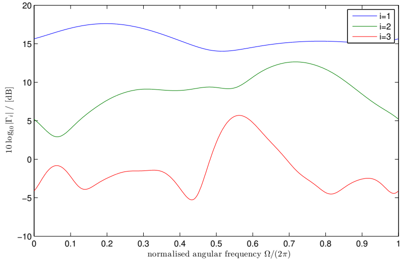
# Polynomial MUSIC (PMUSIC, [3])

- Based on the polynomial EVD of the broadband covariance matrix

$$\mathbf{R}(z) \approx \underbrace{[\mathbf{Q}_s(z) \;\; \mathbf{Q}_n(z)]}_{\mathbf{Q}(z)} \underbrace{\begin{bmatrix} \mathbf{\Lambda}_s(z) & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n(z) \end{bmatrix}}_{\mathbf{\Lambda}(z)} \begin{bmatrix} \tilde{\mathbf{Q}}_s(z) \\ \tilde{\mathbf{Q}}_n(z) \end{bmatrix}$$

- paraunitary $\mathbf{Q}(z)$, s.t. $\mathbf{Q}(z)\tilde{\mathbf{Q}}(z) = \mathbf{I}$;
- diagonalised and spectrally majorised $\mathbf{\Lambda}(z)$:



# PMUSIC cont'd

- Idea —- scan the polynomial noise-only subspace $Q_n(z)$ with broadband steering vectors

$$\Gamma(z,\vartheta) = \tilde{\mathbf{a}}_\vartheta(z)\tilde{\mathbf{Q}}_n(z)\mathbf{Q}_n(z)\mathbf{a}_\vartheta(z)$$

- looking for minima leads to a spatio-spectral PMUSIC

$$S_{\mathrm{PSS-MUSIC}}(\vartheta,\Omega) = \left(\Gamma(z,\vartheta)|_{z=e^{j\Omega}}\right)^{-1}$$

- and a spatial-only PMUSIC

$$S_{\mathrm{PS-MUSIC}}(\vartheta) = \left(2\pi \oint \Gamma(z,\vartheta)|_{z=e^{j\Omega}} d\Omega\right)^{-1} = \Gamma_\vartheta^{-1}[0]$$

with $\Gamma_\vartheta[\tau] \circ\!\!-\!\!\bullet\, \Gamma(z,\vartheta)$.

# Simulation I — Toy Problem

- Linear uniform array with critical spatial and temporal sampling;
- broadband steering vector for end-fire position:

$$\mathbf{a}_{\pi/2}(z) = [1 \;\; z^{-1} \;\; \cdots \;\; z^{-M+1}]^{\mathrm{T}}$$

- covariance matrix

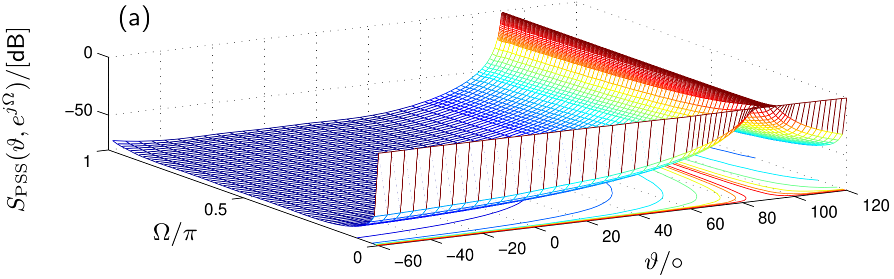
$$\mathbf{R}(z) = \mathbf{a}_{\pi/2}(z)\tilde{\mathbf{a}}_{\pi/2}(z) = \begin{bmatrix} 1 & z^{1} & \dots & z^{M-1} \\ z^{-1} & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ z^{-M+1} & \dots & \dots & 1 \end{bmatrix} \quad .$$

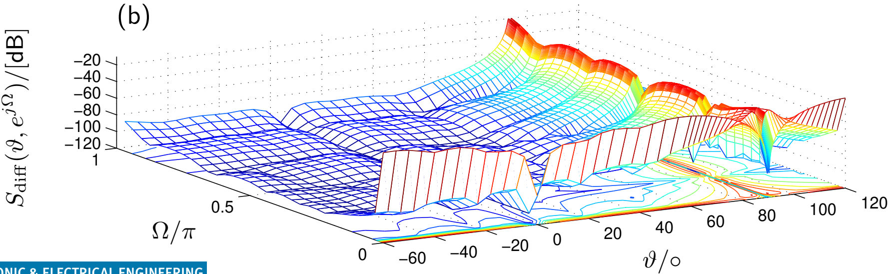
- PEVD (by inspection)

$$\mathbf{Q}(z) = \mathbf{T}_{\mathrm{DFT}}\mathrm{diag}\{1 \;\; z^{-1} \;\; \cdots \;\; z^{-M+1}\} \quad ; \qquad \mathbf{\Lambda}(z) = \mathrm{diag}\{1 \; 0 \; \cdots \; 0\}$$

- simulations with $M = 4$ ...

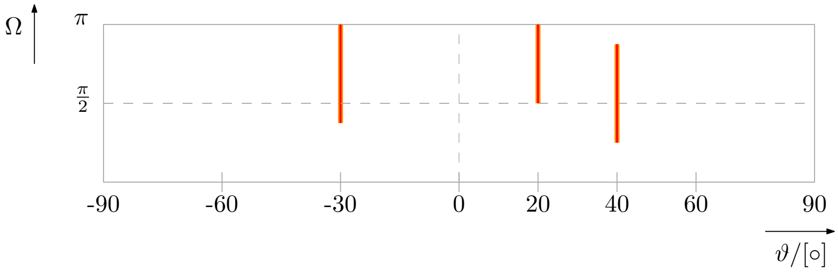
# Simulation I — PSS-MUSIC



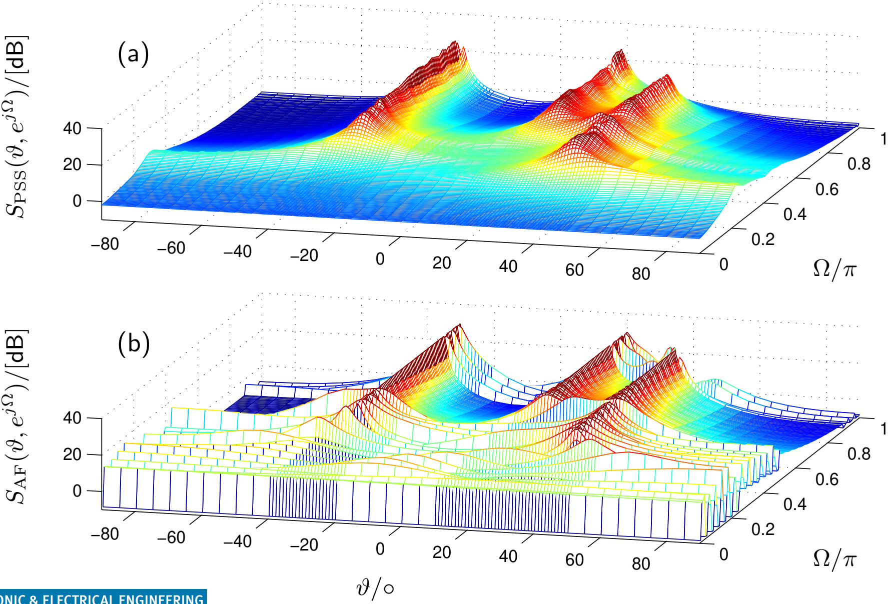
# Simulation II

- $M = 8$ element sensor array illuminated by three sources;
- source 1: $\vartheta_1 = -30°$, active over range $\Omega \in [\frac{3\pi}{8};\ \pi]$;
- source 2: $\vartheta_2 = 20°$, active over range $\Omega \in [\frac{\pi}{2};\ \pi]$;
- source 3: $\vartheta_3 = 40°$, active over range $\Omega \in [\frac{2\pi}{8};\ \frac{7\pi}{8}]$; and



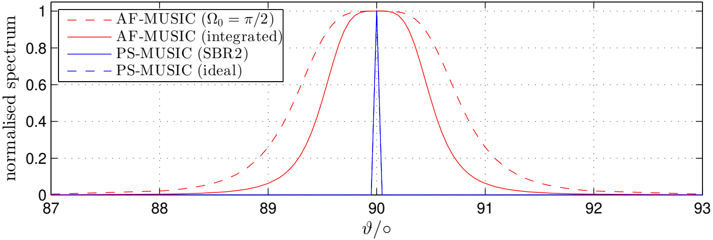

- filter banks as innovation filters, and broadband steering vectors to simulate AoA;
- space-time covariance matrix is estimated from $10^4$ samples.

# Simulation II — PSS-MUSIC

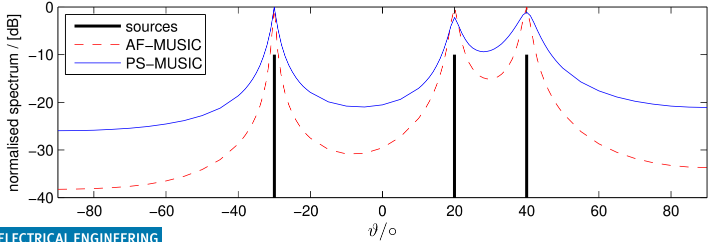


# PS-MUSIC Comparison

- Simulation I (toy problem): peaks normalised to unity:




- Simulation II: inaccuracies on PEVD and broadband steering vector

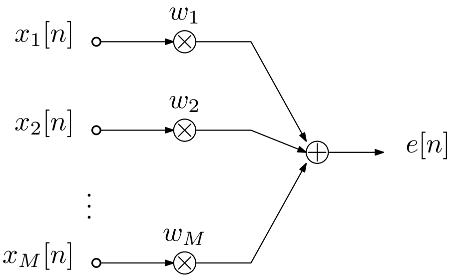


# AoA Estimation — Conclusions

- We have considered the importance of SVD and EVD for narrowband source separation;
- narrowband matrix decomposition real the matrix rank and offer subspace decompositions on which angle-of-arrival estimation alhorithms such as MUSIC can be based;
- broadband problems lead to a space-time covariance or CSD matrix;
- such polynomial matrices cannot be decomposed by standard EVD and SVD;
- a polynomial EVD has been defined;
- iterative algorithms such as SBR2 can be used to approximate the PEVD;
- this permits a number of applications, such as broadband angle of arrival estimation;
- broadband beamforming could then be used to separate broadband sources.

# Narrowband Minimum Variance Distortionless Response Beamformer

- Scenario: an array of $M$ sensors receives data $\mathbf{x}[n]$, containing a desired signal with frequency $\Omega_{\mathrm{s}}$ and angle of arrival $\vartheta_{\mathrm{s}}$, corrupted by interferers;
- a narrowband beamformer applies a single coefficient to every of the $M$ sensor signals:



# Narrowband MVDR Problem

- Recall the space-time covariance matrix:

$$\mathbf{R}[\tau] = \mathcal{E}\left\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\right\}$$

- the MVDR beamformer minimises the output power of the beamformer:

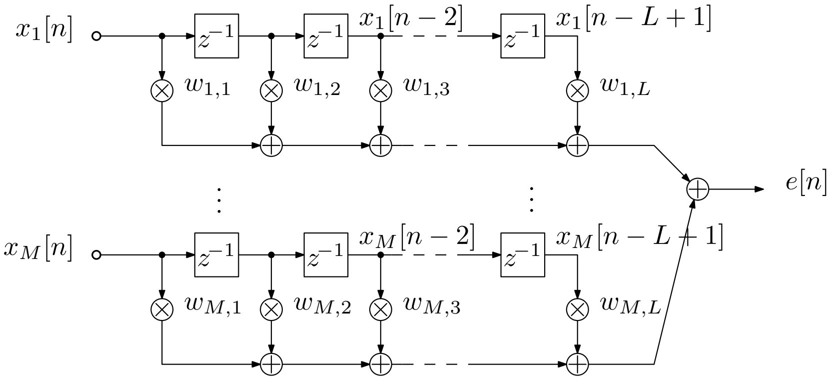
$$\min_{\mathbf{w}} \mathcal{E}\left\{|e[n]|^2\right\} = \min_{\mathbf{w}} \mathbf{w}^{\mathrm{H}}\mathbf{R}[0]\mathbf{w} \tag{46}$$

$$\text{s.t.} \quad \mathbf{a}^{\mathrm{H}}(\vartheta_{\mathrm{s}}, \Omega_{\mathrm{s}})\mathbf{w} = 1 \; , \tag{47}$$

- this is subject to protecting the signal of interest by a constraint in look direction $\vartheta_{\mathrm{s}}$;
- the steering vector $\mathbf{a}_{\vartheta_{\mathrm{s}},\Omega_{\mathrm{s}}}$ defines the signal of interest's parameters.

# Broadband MVDR Beamformer

- Each sensor is followed by a tap delay line of dimension $L$, giving a total of $ML$ coefficients in a vector $\mathbf{v} \in \mathbb{C}^{ML}$ [9, 8, 35]



# Broadband MVDR Beamformer

- A larger input vector $\mathbf{x}_n \in \mathbb{C}^{ML}$ is generated; also including lags;
- the general approach is similar to the narrowband system, minimising the power of $e[n] = \mathbf{v}^{\mathrm{H}}\mathbf{x}_n$;
- however, we require several constraint equations to protect the signal of interest, e.g.

$$\mathbf{C} = [\mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_0), \ \ \mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_1) \ \ldots \ \mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_{L-1})] \tag{48}$$

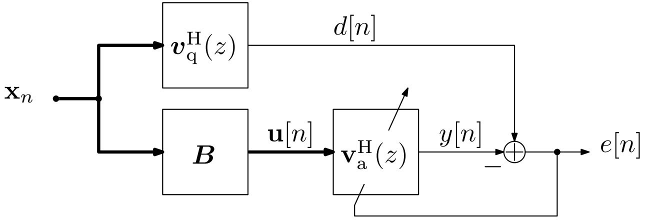
- these $L$ constraints pin down the response to unit gain at $L$ separate points in frequency:

$$\mathbf{C}^{\mathrm{H}}\mathbf{v} = \mathbf{1} \ ; \tag{49}$$

- generally $\mathbf{C} \in \mathbb{C}^{ML \times L}$, but simplifications can be applied if the look direction is towards broadside.

# Generalised Sidelobe Canceller

- A quiescent beamformer $\mathbf{v}_{\mathrm{q}} = \left(\mathbf{C}^{\mathrm{H}}\right)^{\dagger}\mathbf{1} \in \mathbb{C}^{ML}$ picks the signal of interest;
- the quiescent beamformer is optimal for AWGN but generally passes structured interference;
- the output of the blocking matrix $\mathbf{B}$ contains interference only, which requires $[\mathbf{BC}]$ to be unitary; hence $\mathbf{B} \in \mathbb{C}^{ML \times (M-1)L}$;
- an adaptive noise canceller $\mathbf{v}_{\mathrm{a}} \in \mathbb{C}^{(M-1)L}$ aims to remove the residual interference:




- note: all dimensions are determined by $\{M, L\}$.

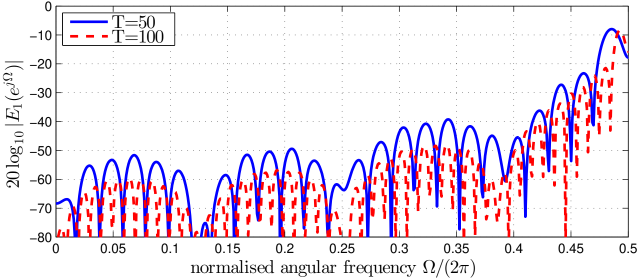
# Polynomial Matrix MVDR Formulation

- Power spectral density of beamformer output: $R_e(z) = \tilde{\boldsymbol{w}}(z)\boldsymbol{R}(z)\boldsymbol{w}(z)$
- proposed broadband MVDR beamformer formulation:

$$\min_{\mathbf{w}(z)} \oint_{|z|=1} R_e(z)\frac{dz}{z} \tag{50}$$

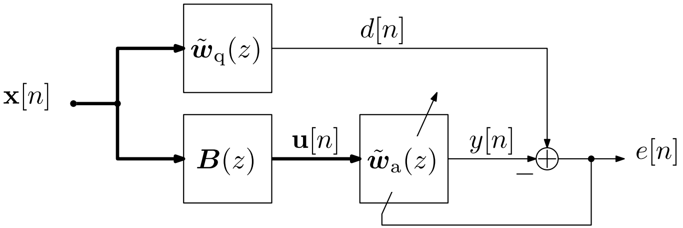
$$\text{s.t.} \quad \tilde{\boldsymbol{a}}(\vartheta_{\mathrm{s}}, z)\boldsymbol{w}(z) = F(z) \; . \tag{51}$$

- precision of broadband steering vector, $|\tilde{\boldsymbol{a}}(\vartheta_{\mathrm{s}}, z)\boldsymbol{a}(\vartheta_{\mathrm{s}}, z) - 1|$, depends on the length $T$ of the fractional delay filter:



# Generalised Sidelobe Canceller

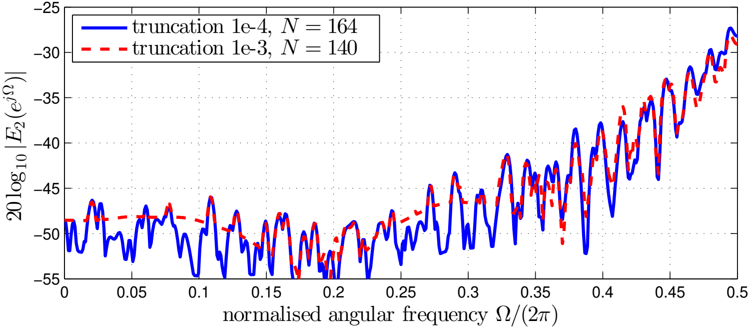
- Instead of performing constrained optimisation, the GSC projects the data and performs adaptive noise cancellation:




- the quiescent vector $\mathbf{w}_{\mathrm{q}}(z)$ is generated from the constraints and passes signal plus interference;
- the blocking matrix $\boldsymbol{B}(z)$ has to be orthonormal to $\mathbf{w}_{\mathrm{q}}(z)$ and only pass interference.

# Design Considerations

- The blocking matrix can be obtained by completing a paraunitary matrix from $\mathbf{w}_{\rm q}(z)$;
- this can be achieved by calculating a PEVD of the rank one matrix $\mathbf{w}_{\rm q}(z)\tilde{\mathbf{w}}_{\rm q}(z)$;
- this leads to a block matrix of order $N$ that is typically greater than $L$;
- maximum leakage of the signal of interest through the blocking matrix:



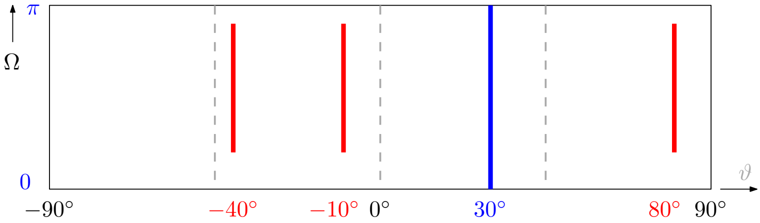
# Computational Cost

- With $M$ sensors and a TDL length of $L$, the complexity of a standard beamformer is dominated by the blocking matrix;
- in the proposed design, $\mathbf{w}_{\mathrm{a}} \in \mathbb{C}^{M-1}$ has degree $L$;
- the quiescent vector $\mathbf{w}_{\mathrm{q}}(z) \in \mathbb{C}^{M}$ has degree $T$;
- the blocking matrix $\boldsymbol{B}(z) \in \mathbb{C}^{(M-1)\times M}$ has degree $N$;
- cost comparison in multiply-accumulates (MACs):

| component | GSC cost | |
|---|---|---|
| | polynomial | standard |
| quiescent beamformer | $MT$ | $ML$ |
| blocking matrix | $M(M-1)N$ | $M(M-1)L^2$ |
| adaptive filter (NLMS) | $2(M-1)L$ | $2(M-1)L$ |

# Example

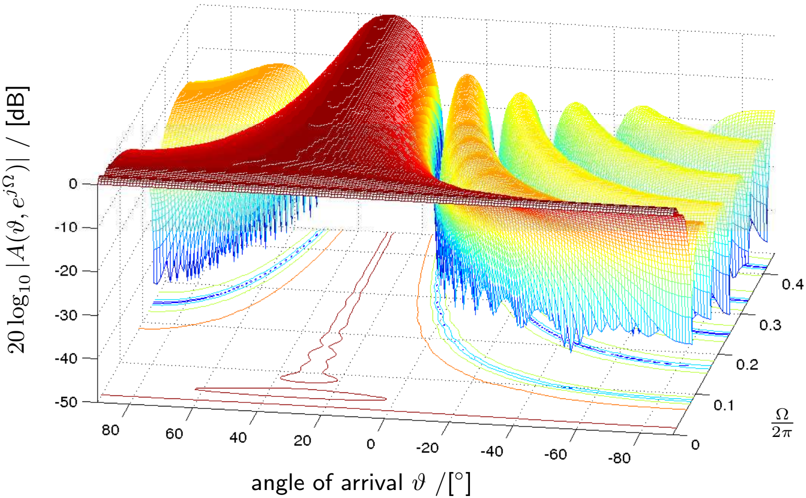
- We assume a signal of interest from $\vartheta = 30^\circ$;
- three interferers with angles $\vartheta_i \in \{-40^\circ, -10^\circ, 80^\circ\}$ active over the frequency range $\Omega = 2\pi \cdot [0.1;\ 0.45]$ at signal to interference ratio of -40 dB;




- $M = 8$ element linear uniform array is also corrupted by spatially and temporally white additive Gaussian noise at 20 dB SNR;
- parameters: $L = 175$, $T = 50$, and $N = 140$;
- cost per iteration: 10.7 kMACs (proposed) versus 1.72 MMACs (standard).

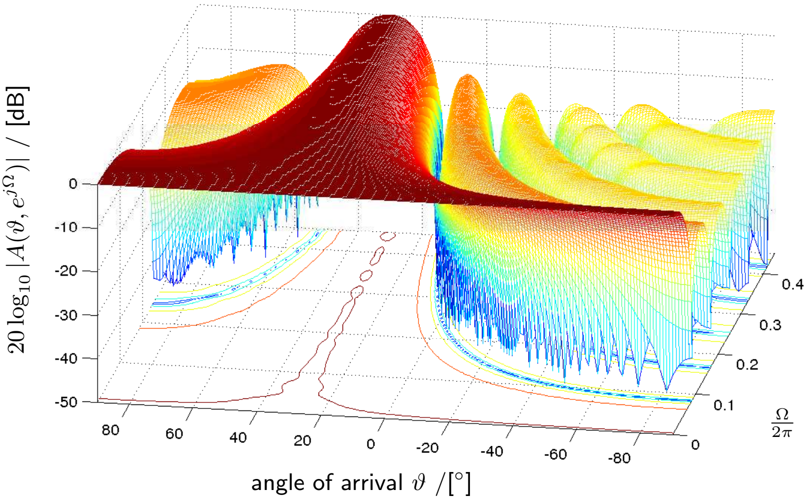
# Quiescent Beamformer

- Directivity pattern of quiescent standard broadband beamformer:



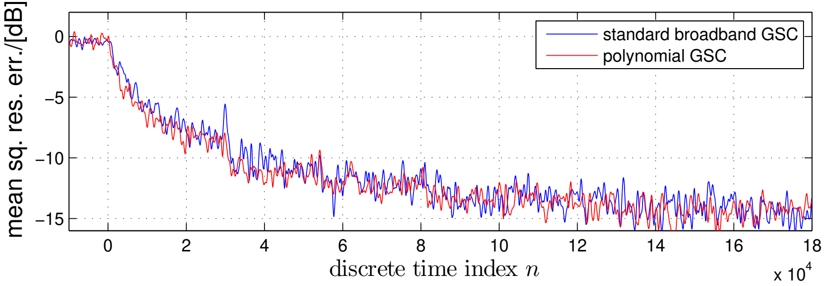
# Quiescent Beamformer

- Directivity pattern of quiescent proposed broadband beamformer:



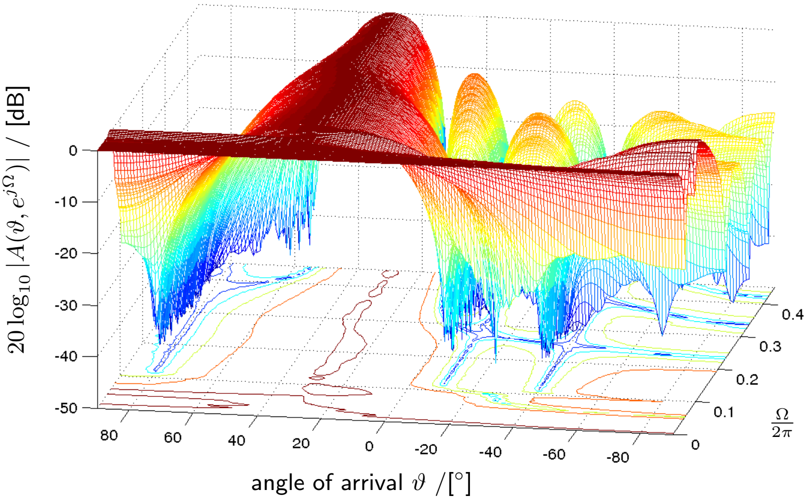
# Adaptation

- Convergence curves of the two broadband beamformers, showing the residual mean squared error (i.e. beamformer output minus signal of interest):



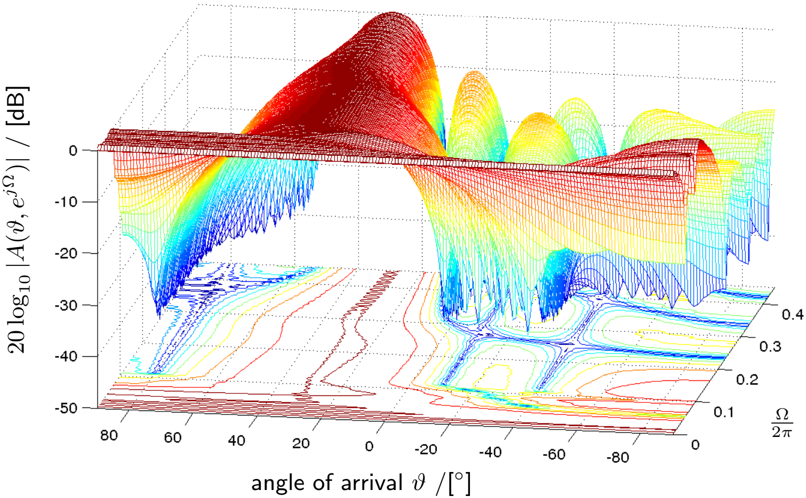
# Adapted Beamformer

- Directivity pattern of adapted proposed broadband beamformer:



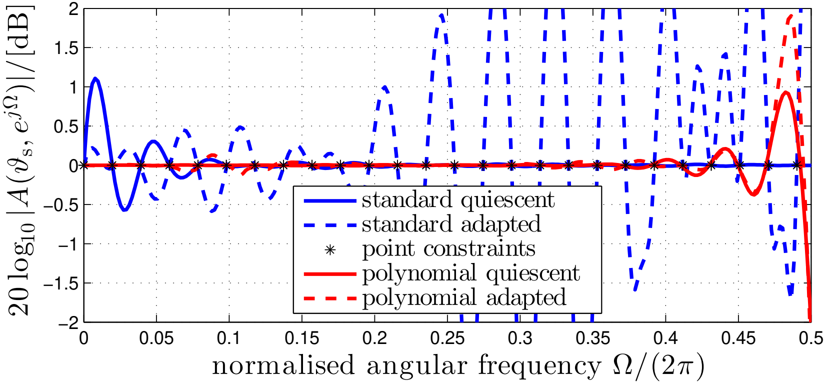
# Adapted Beamformer

- Directivity pattern of adapted standard broadband beamformer:



# Gain in Look Direction

- Gain in look direction $\vartheta_{\mathrm{s}} = 30^\circ$ before and after adaptation:



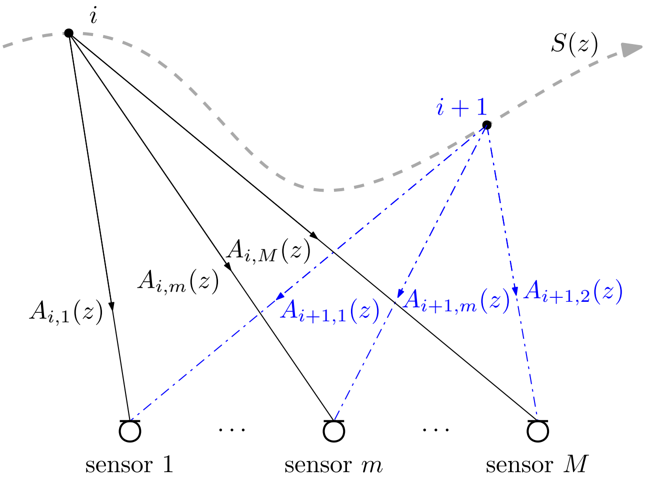

- due to signal leakage, the standard broadband beamformer after adaptation only maintains the point constraints but deviates elsewhere.

# Broadband Beamforming Conclusions

- Based on the previous AoA estimation, beamforming can help to extract source signals and thus perform "source separation";
- broadband beamformers usually assume pre-steering such that the signal of interest lies at broadside;
- this is not always given, and difficult for arbitary array geometries;
- the proposed beamformer using a polynomial matrix formulation can implement abitrary constraints;
- the performance for such constraints is better in terms of the accuracy of the directivity pattern;
- because the proposed design decouples the complexities of the coefficient vector, the quiescent vector and block matrix, and the adaptive process, the cost is significantly lower than for a standard broadband adaptive beamformer.

# Source Extraction Application

- We take $M$-array measurements of a single source:




- 2nd order statistics: $\boldsymbol{R}_i(z) = S(z)\boldsymbol{a}_i(z)\boldsymbol{a}_i^{\mathrm{P}}(z) = \gamma_{i,m}(z)\boldsymbol{u}_i(z)\boldsymbol{u}_i^{\mathrm{P}}$.

# Application Example

- 2nd order stats: $\boldsymbol{R}_i(z) = S(z)\boldsymbol{a}_i(z)\boldsymbol{a}_i^{\mathrm{P}}(z) = \gamma_{i,m}(z)\boldsymbol{u}_i(z)\boldsymbol{u}_i^{\mathrm{P}}$;
- difference: $\boldsymbol{u}_i(z)$ is normal, $\boldsymbol{u}_i^{\mathrm{P}}(z)\boldsymbol{u}_i(z) = 1$, while $\boldsymbol{a}_i(z)$ is not:

$$\boldsymbol{a}_i^{\mathrm{P}}(z)\boldsymbol{a}_i(z) = A_{i,(-)}(z)A_{i,(+)}(z) = A_{i,(+)}^{\mathrm{P}}(z)A_{i,(+)}(z)$$

  with minimum-phase $A_{(+)}(z)$;
- therefore:

$$\boldsymbol{u}_i(z) = \frac{\boldsymbol{a}_i(z)}{A_{i,(+)}(z)}$$
$$\gamma_i(z) = A_{i,(+)}(z)S(z)A_{i,(+)}^{\mathrm{P}}(z) \; ,$$

- for a single measurement, we can say nothing about $\boldsymbol{a}_i^{\mathrm{P}}(z)$ or $S(z)$.

# Application — Multiple Measurements

- If we have several measurements $i = 1 \ldots I$:

$$\boldsymbol{u}_i(z) = \frac{\boldsymbol{a}_i(z)}{A_{i,(+)}(z)}$$

$$\gamma_i(z) = A_{i,(+)}(z) S(z) A^{\mathrm{P}}_{i,(+)}(z) \; ,$$

- we can extract $S(z)$ as the greatest common divisor

$$\hat{S}(z) = \mathsf{GCD}\{\lambda_1(z) \; \ldots \; \lambda_I(z)\} \; ; \qquad (52)$$

- we can also extract the $A_{i,(+)}(z)$, and hence determine the vectors $\boldsymbol{a}_i(z)$ save of an arbitrary phase response.

# Application — Frequency Domain Attempt

- As an alternative, we take measurements in independent frequency bins:

$$\begin{aligned} \mathbf{R}_{i,k} &= \boldsymbol{R}_i(e^{j\Omega_k}) = \boldsymbol{a}_i(e^{j\Omega_k})S(e^{j\Omega_k})\boldsymbol{a}_i^{\mathrm{H}}(e^{j\Omega_k}) + \sigma_n^2\mathbf{I} \quad (53) \\ &= \mathbf{q}_{i,k}\lambda_{i,k}\mathbf{q}_{i,k}^{\mathrm{H}} \; . \quad (54) \end{aligned}$$

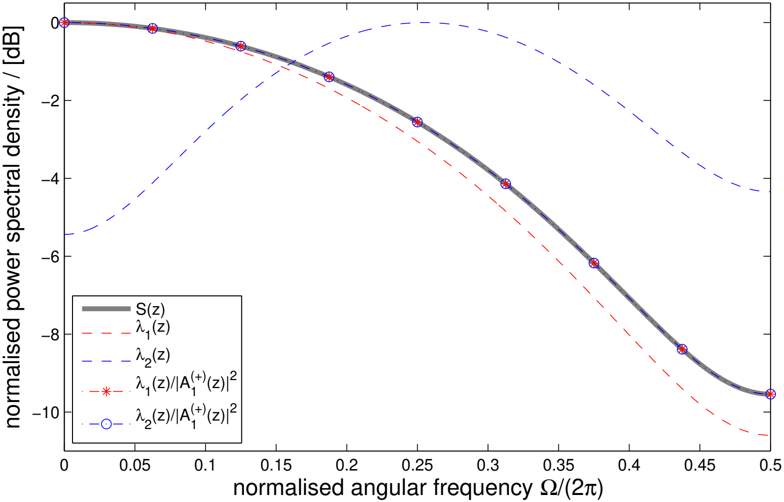
- principal eigenvectors and eigenvalues for the measurement campaigns are

$$\begin{aligned} \mathbf{q}_{i,k} &= \frac{\boldsymbol{a}_i(e^{j\Omega_k})}{|\boldsymbol{a}_i(e^{j\Omega_k})|} \; , \quad (55) \\ \lambda_{i,k} &= S(e^{j\Omega_k})|\boldsymbol{a}_i(e^{j\Omega_k})|^2 \; . \quad (56) \end{aligned}$$

- because of the normalisation, nothing can be extracted about the source or the transfer functions.

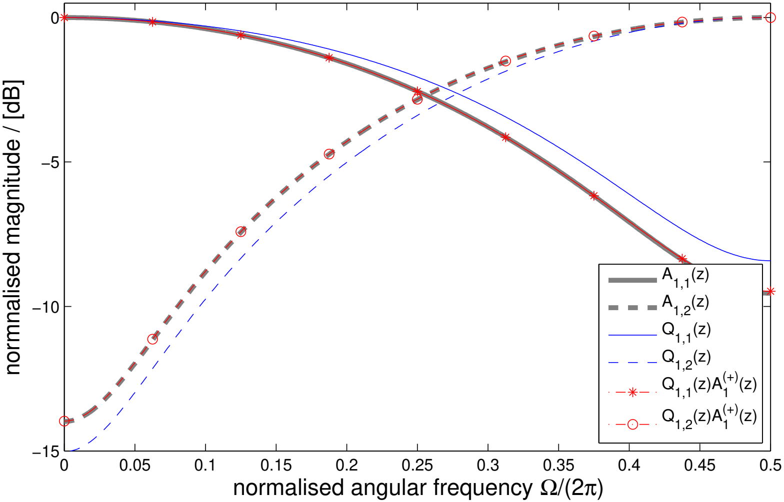
# Application — Results I

- Eigenvalues / source PSD for two measurements $i = \{0, 1\}$:



# Application — Results II

- Eigenvectors / magnitude response for measurement $i = \{0\}$:

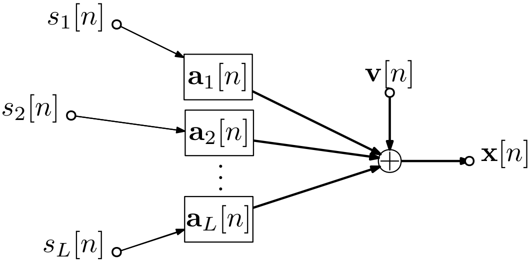


# Application — Summary

- we can extract the source PSD and the magnitude responses once we have at least two measurements;
- an independent frequency bin approach does not yield anything;
- the polynomial approach rests on an accurate parahermitian EVD, and an accurate root finding / GCD algorithm;
- root finding is numerically challenging;
- nevertheless the example gives a glimpse of the type of advantages that a "holistic" broadband approach offers.

# Problem & Model

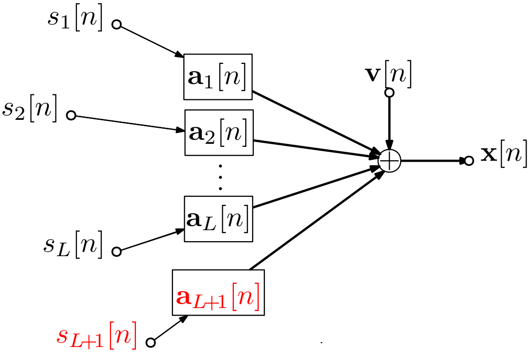
- A number of broadband stationary sources $s_\ell[n]$, $\ell = 1, \ldots, L$, illuminate an $M$-element sensor array;
- each transfer path is modelled by a vector of impulse responses $\mathbf{a}_\ell[n] \in \mathbb{C}^M$;
- stationary additive, spatially and temporally uncorrelated noise $\mathbf{v}[n] \in \mathbb{C}^M$;




$$\mathbf{x}[n] = \sum_{\ell=1}^{L} \mathbf{a}_\ell[n] * s_\ell[n] + \mathbf{v}[n]$$

# Problem & Model

- A number of broadband stationary sources $s_\ell[n]$, $\ell = 1, \dots, L$, illuminate an $M$-element sensor array;
- each transfer path is modelled by a vector of impulse responses $\mathbf{a}_\ell[n] \in \mathbb{C}^M$;
- stationary additive, spatially and temporally uncorrelated noise $\mathbf{v}[n] \in \mathbb{C}^M$;
- a broadband transient signal $s_{L+1}[n]$ enters the scene at some point in time;
- aim: we want to detect the onset of this transient signal, which may be weak in power [38];
- assumption: $M > L$.

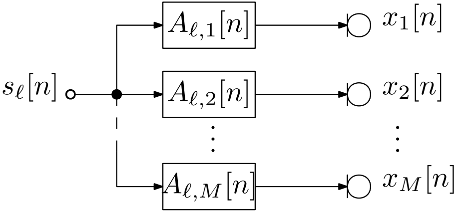



$$\mathbf{x}[n] = \sum_{\ell=1}^{L+1} \mathbf{a}_\ell[n] * s_\ell[n] + \mathbf{v}[n]$$

# Model

- Each source, $s_\ell[n]$, contributes to the data vector $\mathbf{x}[n] = [x_1[n], \ \ldots \ , x_M[n]]^\mathrm{T}$ via a steering vector $\mathbf{a}_\ell[n] = [A_{\ell,1}[n], \ldots A_{\ell,M}[n]]^\mathrm{T}$;
- compact with $\mathbf{A}[n] = [\mathbf{a}_1[n] \ldots \mathbf{a}_L[n]]$ and $\mathbf{s}[n] = [s_1[n], \ldots, s_L[n]]^\mathrm{T}$:

$$\mathbf{x}[n] = \mathbf{A}[n] * \mathbf{s}[n] + \mathbf{v}[n] \ ;$$




- space-time covariance: $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^\mathrm{H}[n-\tau]\}$:

$$\begin{aligned} \mathbf{R}[\tau] &= \mathbf{A}[\tau] * \mathcal{E}\{\mathbf{s}[n]\mathbf{s}^\mathrm{H}[n-\tau]\} * \mathbf{A}^\mathrm{H}[-\tau] + \mathcal{E}\{\mathbf{v}[n]\mathbf{v}^\mathrm{H}[n-\tau]\} && (57) \\ &= \mathbf{A}[\tau] * \mathbf{\Gamma}[\tau] * \mathbf{A}^\mathrm{H}[-\tau] + \sigma_v^2 \mathbf{I}_M \delta[\tau] \ . && (58) \end{aligned}$$

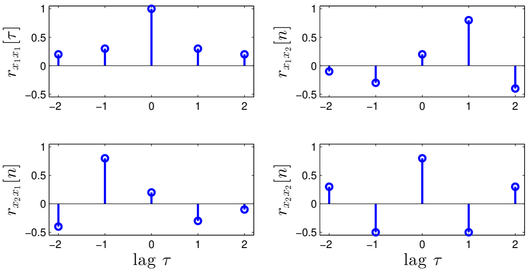
# Cross-Spectral Density Matrix

- Transfer function matrix $\boldsymbol{A}(z) = \sum_n \mathbf{A}[n]z^{-n}$ (short $\boldsymbol{A}(z) \bullet\!\!-\!\!\circ\, \mathbf{A}[n]$) is a polynomial in $z \in \mathbb{C}$;
- cross-spectral density $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ\, \mathbf{R}[\tau]$:

$$\boldsymbol{R}(z) = \boldsymbol{A}(z)\boldsymbol{\Gamma}(z)\boldsymbol{A}^{\mathrm{P}}(z) + \sigma_v^2\mathbf{I}_M \;;$$

- parahermitian property:

$$\boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}^{\mathrm{H}}(1/z^*) = \boldsymbol{R}(z) \;;$$




- when evaluated for a specific normalised angular frequency $\Omega_0$: $\mathbf{R}_0 = \boldsymbol{R}(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega_0}}$;
- $\mathbf{R}_0$ is a constant matrix that describes a *narrowband* problem;
- $\mathbf{R}_0$ is Hermitian $\longrightarrow$ eigenvalue decomposition (EVD) $\mathbf{R}_0 = \mathbf{Q}_0\boldsymbol{\Lambda}_0\mathbf{Q}_0^{\mathrm{H}}$.

# Narrowband EVD and Subspace Decomposition

- We assume an ordered EVD $\mathbf{R}_0 = \mathbf{Q}_0 \mathbf{\Lambda}_0 \mathbf{Q}_0^{\mathrm{H}}$, where $\mathbf{\Lambda}_0 = \mathrm{diag}\{\lambda_1, \ldots, \lambda_M\}$ with $\lambda_\ell \geq \lambda_{\ell+1}$, $\ell = 1, \ldots, (M-1)$;
- partitioning enables a subspace decomposition:

$$\mathbf{R}_0 = \left[\begin{array}{c:c} \mathbf{Q}_{\mathrm{s}} & \mathbf{Q}_{\mathrm{n}} \end{array}\right] \left[\begin{array}{c:c} \mathbf{\Lambda}_s + \sigma_v^2 \mathbf{I}_L & \\ \hdashline & \sigma_v^2 \mathbf{I}_{M-L} \end{array}\right] \left[\begin{array}{c} \mathbf{Q}_{\mathrm{s}}^{\mathrm{H}} \\ \hdashline \mathbf{Q}_{\mathrm{n}}^{\mathrm{H}} \end{array}\right]$$

- source enumeration: eigenvalues above noise floor = number of uncorrelated sources;
- $\mathbf{y}[n] = \mathbf{Q}_{\mathrm{n}}^{\mathrm{H}} \mathbf{x}[n] \in \mathbb{C}^{M-L}$ only contains noise;
- power in $\mathbf{y}[n]$: $\mathcal{E}\{\|\mathbf{y}[n]\|_2^2\} = (M-L)\sigma_v^2$ because of orthonormality of $\mathbf{Q}$.

# Broadband EVD

- Space-time covariance $\mathbf{R}[\tau]$ or equivalently the CSD matrix $\boldsymbol{R}(z)$ are only diagonalised by the EVD for a specific value $\tau$ or $z$;
- for an analytic $\boldsymbol{R}(z)$ that is not derived from multiplexed data, there exists a parahermitian matrix EVD [40, 41]

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z) \; ; \tag{59}$$

- $\boldsymbol{\Lambda}(z)$ is diagonal, parahermitian, analytic, and unique;
- eigenvectors in $\boldsymbol{Q}(z)$ are paraunitary, analytic, and unique up to an arbitrary allpass function;
- paraunitarity $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \boldsymbol{Q}^{\mathrm{P}}(z)\boldsymbol{Q}(z) = \mathbf{I}$ implies losslessness;
- a number of algorithms can approximate (59) [24, 26, 27, 44, 42, 43].

# Broadband Subspace Decomposition

- The parahermitian matrix EVD $\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z)$ enables a broadband subspace decomposition:

$$\boldsymbol{R}(z) = \left[\begin{array}{c:c} \boldsymbol{Q}_{\mathrm{s}}(z) & \boldsymbol{Q}_{\mathrm{n}}(z) \end{array}\right] \left[\begin{array}{c:c} \boldsymbol{\Lambda}_s(z) + \sigma_v^2 \mathbf{I}_L & \\ \hdashline & \sigma_v^2 \mathbf{I}_{M-L} \end{array}\right] \left[\begin{array}{c} \boldsymbol{Q}_{\mathrm{s}}^{\mathrm{P}}(z) \\ \hdashline \boldsymbol{Q}_{\mathrm{n}}^{\mathrm{P}}(z) \end{array}\right]$$

- $\mathbf{Q}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{Q}(z)$ describes a lossless filter bank;
- data vector component in the noise-only subspace: $\mathbf{y}[n] = \mathbf{Q}_{\mathrm{n}}^{\mathrm{H}}[-n] * \mathbf{x}[n]$;
- again, it can be shown that ideally $\mathcal{E}\{\|\mathbf{y}[n]\|_2^2\} = (M-L)\sigma_v^2$.

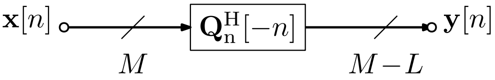
# ‘Syndrome’ Idea

University of Strathclyde Engineering

- We estimate $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ \mathbf{R}[\tau]$ over a window of data, with $L < M$ stationary sources present;
- compute parahermitian matrix EVD, perform source enumeration, and determine the eigenvectors spanning the noise-only subspace, $\boldsymbol{Q}_{\mathrm{n}}(z)$;
- if an additional source $s_{L+1}[n]$ enters the scene, it will likely protrude into the noise-only subspace;
- we therefore monitor the syndrome vector

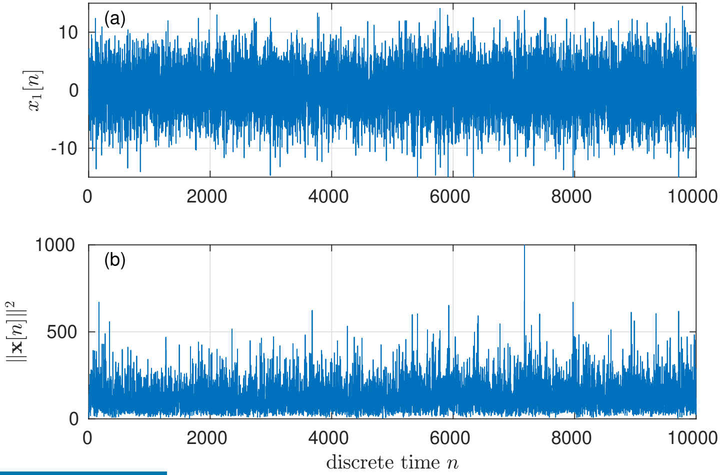
$$\mathbf{y}[n] = \mathbf{Q}_{\mathrm{n}}^{\mathrm{H}}[-n] * \mathbf{x}[n] \tag{60}$$

for a change in power, or for any structured / correlated components.



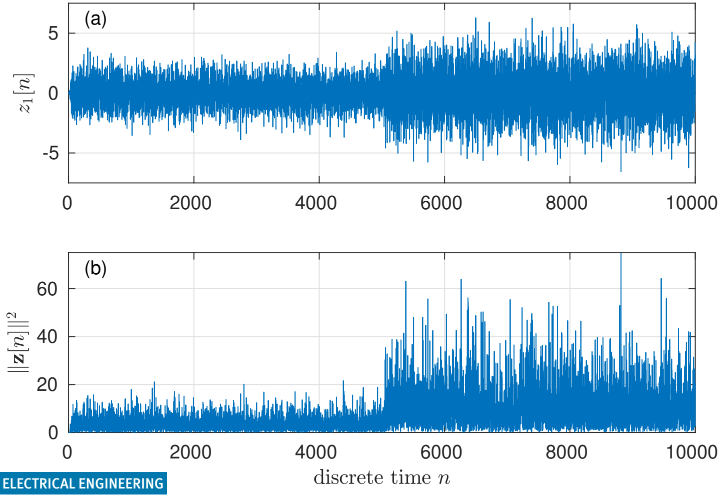
# Intuitive Example I

- $M = 6$ sensors, $L = 3$ stationary sources; weak transient source at $n = 5000$;
- monitoring a sensor output $x_1[n]$:



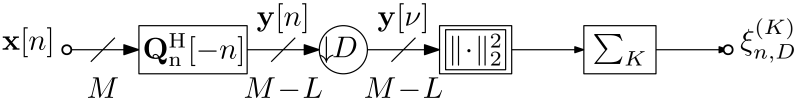
# Intuitive Example II

- $M = 6$ sensors, $L = 3$ stationary sources; weak transient source at $n = 5000$;
- monitoring a syndrome element $y_1[n]$:



# Proposed Approach

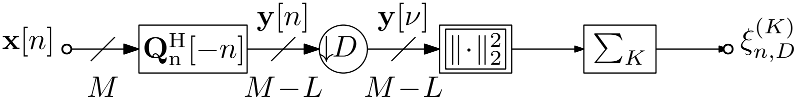
- We use the statistics and evaluated parahermitian matrix EVD of a previous time window, and utilise the broadband noise-only subspace spanned by the columns of $\boldsymbol{Q}_{\mathrm{n}}(z)$;
- being analytic, $\boldsymbol{Q}_{\mathrm{n}}(z)$ can typically be approximated well by low-order polyomials, and is relatively inexpensive to implement;




- because of the processing, elements of the syndrome vector $\mathbf{y}[n]$ are spatially and temporally correlated;
- decimation by $D$ can break temporal correlation and further reduces complexity;
- we can average over consecutive syndrome vectors to increase discrimination;
- $\xi_{n,D}^{(K)}$ is generalised $\chi^2$ distributed if temporal correlation is suppressed [32, 13].

# Decimated Processor

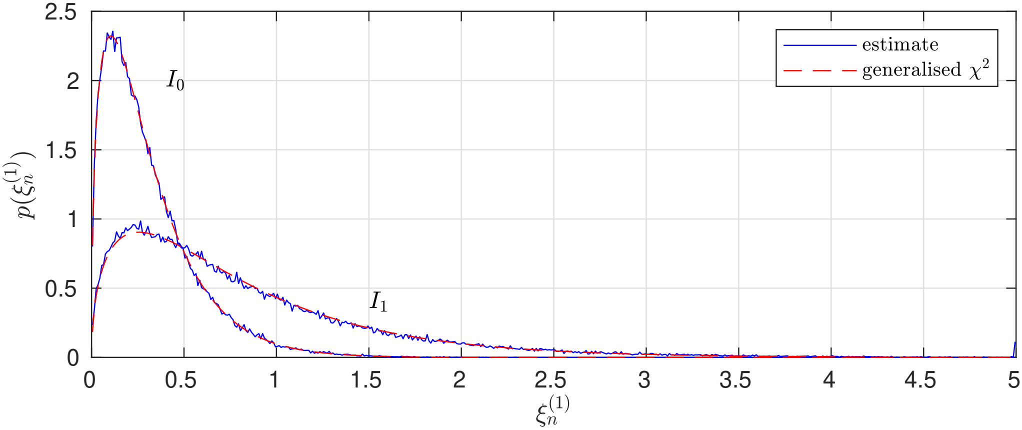
- The proposed subspace projection is followed by a decimation by $D$:




- cost advantage: a polyphase implementation integrates the decimation with the processor, reducing operations by a factor of $D$;
- temporal decorrelation: if the temporal correlation does not exceed $D$ lags, the decimation will temporally decorrelate susequent snapshots of the syndrome vector $\mathbf{y}[\nu]$.

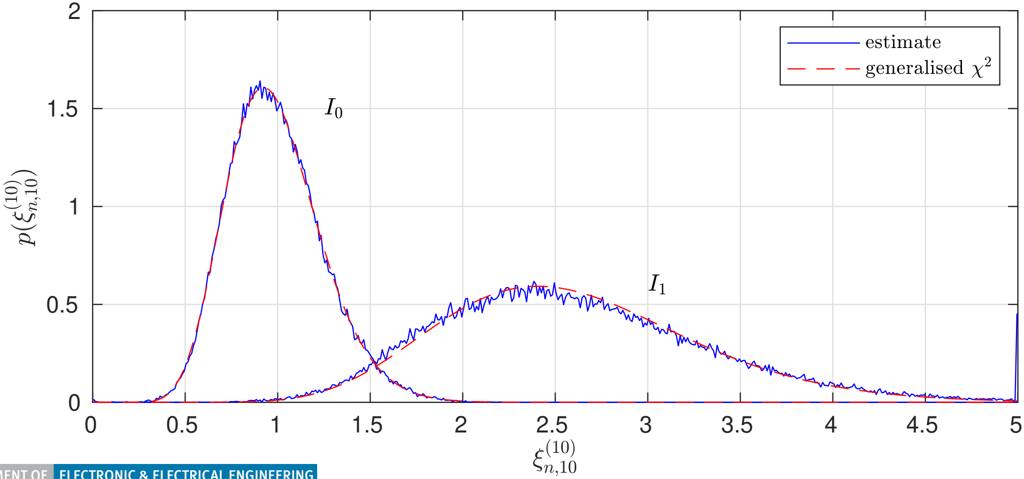
# Results I — Statistics

- $M = 6$ sensors, $L = 2$ stationary sources, transfer functions determined by radio propagation model for dense urban environment (polynomial order $\approx 40$);
- statistics of output for $I_0$: no transient versus $I_1$: transient present; $K = 1$;



# Results I — Statistics

- $M = 6$ sensors, $L = 2$ stationary sources, transfer functions determined by radio propagation model for dense urban environment (polynomial order $\approx 40$);
- statistics of output for $I_0$: no transient versus $I_1$: transient present; $K = 10$;



# Results II — Sources and Propagation Environment

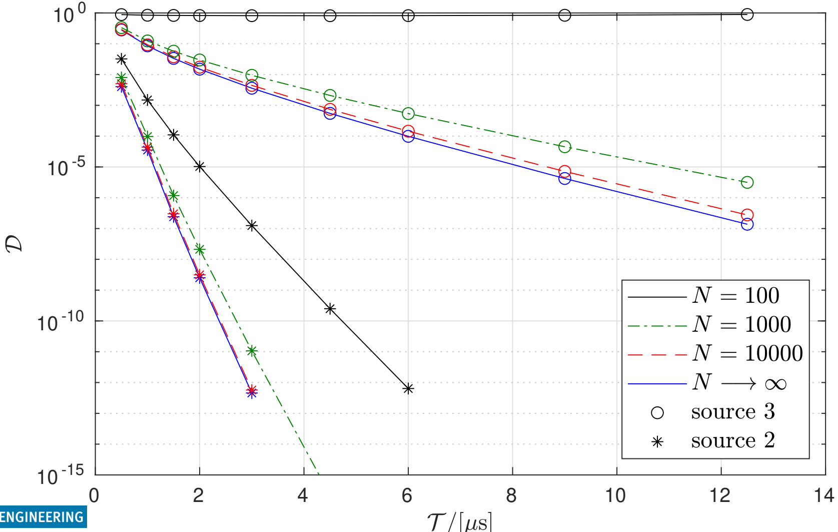
- Realistic 20MHz urban scenario with dispersive impulse responses;
- $M = 6$ sensors;
- total power of contributions of three different sources:

| signal | power |
|---|---|
| source 1 | 0.0000 dB |
| source 2 | -4.3494 dB |
| source 3 | -13.2865 dB |
| noise | -16.2865 dB |

- we use either source 2 or 3 as transient signal; the two remaining sources are stationary.

# Results III — Discrimination vs Decision Time

- Averaging increasingly separates the distributions for $I_0$ and $I_1$ — measured as discrimination $\mathcal{D}$: derived from the ROC [19];
- averaging also increases the time to compute $\xi_{n,D}^{(T)} \longrightarrow$ decision time $\mathcal{T}$ (for a 20MHz channel);
- $N$ here is the window over which the space-time covariance is estimated [15, 16, 17];



## Summary

- We have discussed a broadband subspace approach to detect the presence of weak transient signals;
- this is based on second order statistics of sensor array data — the space-time covariance matrix — and a polynomial matrix EVD;
- this covariance matrix and its decomposition can be computed off-line; for low-cost implementations, see e.g. [12, 23]
- a subspace decomposition for the noise-only subspace determines a syndrome vector;
- in the absence of a transient signal, this syndrome only contains noise;
- a transient signal is likely to protrude into the noise-only subspace, and a change in energy can be detected even if the signal is weak;
- discrimination can be traded off against decision time;
- in audio, the approach is utilised to detect the onset of weak speakers;
- in future, we may investigate time-varying channels and subspace leakage.

# References I

[1] M. Alrmah, M. Hussin, S. Weiss, and S. Lambotharan.
Comparison of broadband direction of arrival estimation algorithms.
In *9th IMA Mathematics in Signal Processing Conference*, Birmingham, UK, December 2012.

[2] M. Alrmah and S. Weiss.
Filter bank based fractional delay filter implementation for widely accurate broadband steering vectors.
In *5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Saint Martin, December 2013.

[3] M. Alrmah, S. Weiss, and S. Lambotharan.
An extension of the MUSIC algorithm to broadband scenarios using polynomial eigenvalue decomposition.
In *19th European Signal Processing Conference*, pages 629–633, Barcelona, Spain, August 2011.

[4] M. Alrmah, S. Weiss, and J. McWhirter.
Implementation of accurate broadband steering vectors for broadband angle of arrival estimation.
In *IET Intelligent Signal Processing*, London, UK, December 2013.

[5] M. Alrmah, S. Weiss, S. Redif, S. Lambotharan, and J. McWhirter.
Angle of arrival estimation for broadband signals: A comparison.
In *IET Intelligent Signal Processing*, London, UK, December 2013.

[6] A. Alzin, F. Coutts, J. Corr, S. Weiss, I. Proudler, and J. Chambers.
Polynomial matrix formulation-based Capon beamformer.
In *IMA International Conference on Signal Processing in Mathematics*, Birmingham, UK, December 2016.

[7] A. Alzin, F. Coutts, J. Corr, S. Weiss, I. K. Proudler, and J. A. Chambers.
Adaptive broadband beamforming with arbitrary array geometry.
In *IET/EURASIP Intelligent Signal Processing*, London, UK, December 2015.

[8] K. M. Buckley.
Spatial/Spectral Filtering with Linearly Constrained Minimum Variance Beamformers.
*IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35(3):249–266, March 1987.

# References II

[9] K. M. Buckley and L. J. Griffith.
An Adaptive Generalized Sidelobe Canceller with Derivative Constraints.
*IEEE Transactions on Antennas and Propagation*, 34(3):311–319, March 1986.

[10] 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*, 60:1–40, 1991.

[11] F. Coutts, K. Thompson, S. Weiss, and I. Proudler.
A comparison of iterative and DFT-based polynomial matrix eigenvalue decompositions.
In *IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Curacao, December 2017.

[12] F. K. Coutts, I. K. Proudler, and S. Weiss.
Efficient implementation of iterative polynomial matrix evd algorithms exploiting structural redundancy and parallelisation.
*IEEE Transactions on Circuits and Systems I: Regular Papers*, 66(12):4753–4766, Dec. 2019.

[13] A. Das and W. S. Geisler.
A method to integrate and classify normal distributions.
*arXiv e-prints*, page arXiv:2012.14331, Dec. 2020.

[14] B. De Moor and S. Boyd.
Analytic properties of singular values and vectors.
Technical report, KU Leuven, 1989.

[15] C. Delaosa, F. K. Coutts, J. Pestana, and S. Weiss.
Impact of space-time covariance estimation errors on a parahermitian matrix EVD.
In *10th IEEE Workshop on Sensor Array and Multichannel Signal Processing*, pages 1–5, July 2018.

# References III


[16] C. Delaosa, J. Pestana, N. J. Goddard, S. Somasundaram, and S. Weiss.
Sample space-time covariance matrix estimation.
In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8033–8037, May 2019.

[17] C. Delaosa, J. Pestana, N. J. Goddard, S. D. Somasundaram, and S. Weiss.
Support estimation of a sample space-time covariance matrix.
In *Sensor Signal Processing for Defence*, Brighton, UK, March 2019.

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

[19] J. A. Hanley and B. J. McNeil.
The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve.
*Radiology*, 143:26–36, 1982.

[20] L. Hanzo, M. Münster, B. J. Choi, and T. Keller.
*OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs, and Broadcasting*.
IEEE and John Wiley & Sons, Chichester, UK, 2003.

[21] S. Icart and P. Comon.
Some properties of Laurent polynomial matrices.
In *9th IMA Conference on Mathematics in Signal Processing*, Birmingham, UK, December 2012.

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

# References IV


[23] F. A. Khattak, S. Weiss, and I. K. Proudler.
Fast givens rotation approach to second order sequential best rotation algorithms.
In *International Conference in Sensor Signal Processing for Defence*, pages 40–44, Edinburgh, Scotland, September 2021.

[24] 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, May 2007.

[25] V. Neo, C. Evers, S. Weiss, and P. A. Naylor.
Polynomial matrix eigenvalue decompositionexploiting spherical microphone array processing.
*IEEE Transactions on Signal Processing*, 2022.

[26] S. Redif, J. McWhirter, and S. Weiss.
Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition.
*IEEE Transactions on Signal Processing*, 59(11):5253–5264, November 2011.

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

[28] F. Rellich.
Störungstheorie der Spektralzerlegung. III. Mitteilung. Analytische, nicht notwendig beschränkte Störung.
*Mathematische Annalen*, 116:555–570, 1939.

[29] A. Scaglione, G. B. Giannakis, and S. Barbarossa.
Filterbank Transceivers Optimizing Information Rate in Block Transmission over Dispersive Channels.
*IEEE Transactions on Information Theory*, 45(4):1019–1032, April 1999.

# References V


[30] A. Scaglione, G. B. Giannakis, and S. Barbarossa.
Redundant Filterbank Precoders and Equalizers. I. Unification and Optimal Designs.
*IEEE Transactions on Signal Processing*, 47(7):1988–2006, July 1999.

[31] A. Scaglione, G. B. Giannakis, and S. Barbarossa.
Redundant Filterbank Precoders and Equalizers. II. Blind Channel Estimation, Synchronization, and Direct Equalization.
*IEEE Transactions on Signal Processing*, 47(7):2007–2022, July 1999.

[32] H. Solomon and M. A. Stephens.
Distribution of a sum of weighted chi-square variables.
*Journal of the American Statistical Association*, 72(360):881–885, 1977.

[33] M. Tohidian, H. Amindavar, and A. M. Reza.
A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices.
*EURASIP Journal on Advances in Signal Processing*, 2013(1):1–16, 2013.

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

[35] B. D. Van Veen and K. M. Buckley.
Beamforming: A Versatile Approach to Spatial Filtering.
*IEEE Acoustics, Speech, and Signal Processing Magazine*, 5(2):4–24, April 1988.

[36] S. Weiss, M. Alrmah, S. Lambotharan, J. McWhirter, and M. Kaveh.
Broadband angle of arrival estimation methods in a polynomial matrix decomposition framework.
In *IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pages 109–112, Dec 2013.

# References VI


[37] S. Weiss, S. Bendoukha, A. Alzin, F. Coutts, I. Proudler, and J. Chambers.
MVDR broadband beamforming using polynomial matrix techniques.
In *23rd European Signal Processing Conference*, pages 839–843, Nice, France, September 2015.

[38] S. Weiss, C. Delaosa, J. Matthews, I. Proudler, and B. Jackson.
Detection of weak transient signals using a broadband subspace approach.
In *International Conference on Sensor Signal Processing for Defence*, pages 65–69, Edinburgh, Scotland, Sept. 2021.

[39] S. Weiss, N. J. Goddard, S. Somasundaram, I. K. Proudler, and P. A. Naylor.
Identification of broadband source-array responses from sensor second order statistics.
In *Sensor Signal Processing for Defence Conference*, pages 1–5, London, UK, December 2017.

[40] 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, May 2018.

[41] S. Weiss, J. Pestana, I. K. Proudler, and F. K. Coutts.
Corrections to "on the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix".
*IEEE Transactions on Signal Processing*, 66(23):6325–6327, Dec 2018.

[42] S. Weiss, I. Proudler, F. Coutts, and J. Deeks.
Extraction of analytic eigenvectors from a parahermitian matrix.
In *International Conference on Sensor Signal Processing or Defence*, Edinburgh, UK, 2020.

[43] 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.

# References VII


[44] 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.

[45] S. Weiss, C. H. Ta, and C. Liu.
A wiener filter approach to the design of filter bank based single-carrier precoding and equalisation.
In *IEEE International Symposium on Power Line Communications and Its Applications*, pages 493–498, Pisa, Italy, March 26-28 2007.