Source Separation and Beamforming

Stephan Weiss and Ian K. Proudler

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

UDRC Summer School, 1 July 2021

With many thanks to:
J.G. McWhirter, I.K. Proudler, J. Pestana, J. Corr, Z. Wang,
F.K. Coutts, C. Delaosa, and F. Khattak

Polynomial Matrix Co-Enthusiasts

Presentation Overview

1. Overview

Part I: Polynomial Matrices and Decompositions

2. Polynomial matrices and basic operations
   2.1 occurence: MIMO systems, filter banks, space-time covariance
   2.2 basic properties and operations
3. Polynomial eigenvalue decomposition (PEVD)
4. Iterative PEVD algorithms
   4.1 sequential best rotation (SBR2)
   4.2 sequential matrix diagonalisation (SMD)
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 from 2nd order statistics
10. Summary and materials

What is a Polynomial Matrix?

University of
Strathclyde
Engineering

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

$$\boldsymbol{A}(z) = \left[\begin{array}{rr} 1 & -1 \\ -1 & 2 \end{array} \right] + \left[\begin{array}{rr} 1 & 1 \\ 1 & -1 \end{array} \right] z^{-1} + \left[\begin{array}{rr} -1 & 2 \\ 1 & -1 \end{array} \right] z^{-2} \quad (1)$$

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

$$\boldsymbol{A}(z) = \left[\begin{array}{cc} 1 + z^{-1} - z^{-2} & -1 + z^{-1} + 2z^{-2} \\ -1 + z^{-1} + z^{-2} & 2 - z^{-1} - z^{-2} \end{array} \right] \quad (2)$$

▶ polynomial matrices could also contain rational polynomials, but the notation would not be as easily interchangeable as (1) and (2).

Where Do Polynomial Matrices Arise?

University of
Strathclyde
Engineering

▶ A multiple-input multiple-output (MIMO) system could be made up of a number of finite impulse response (FIR) channels:
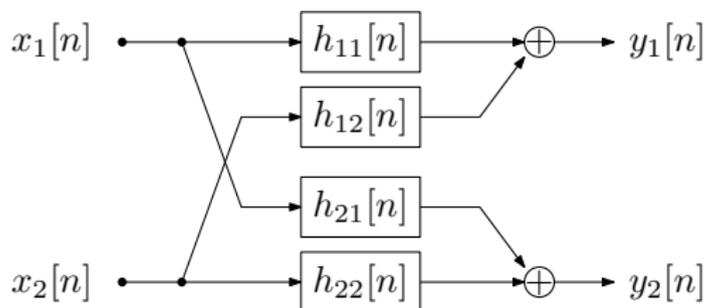
▶ writing this as a matrix of impulse responses:

$$\mathbf{H}[n] = \left[\begin{array}{cc} h_{11}[n] & h_{12}[n] \\ h_{21}[n] & h_{22}[n] \end{array} \right] \tag{3}$$

Transfer Function of a MIMO System

▶ Example for MIMO matrix $\mathbf{H}[n]$ of impulse responses:
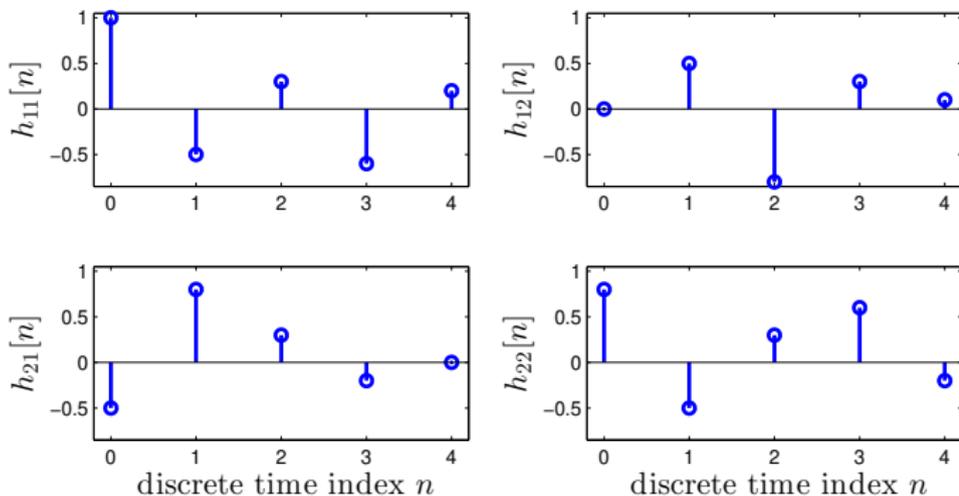
▶ the transfer function of this MIMO system is a polynomial matrix:

$$\boldsymbol{H}(z) = \sum_{n=-\infty}^{\infty} \mathbf{H}[n]z^{-1} \qquad \text{or} \qquad \boldsymbol{H}(z) \;\bullet\!\!-\!\!\circ\; \mathbf{H}[n] \qquad (4)$$

Analysis Filter Bank

▶ Critically decimated K-channel analysis filter bank:
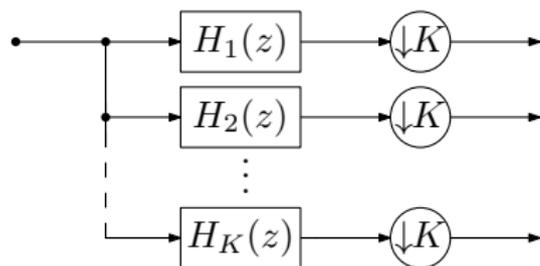
▶ equivalent polyphase representation:

Polyphase Analysis Matrix

University of
Strathclyde
Engineering
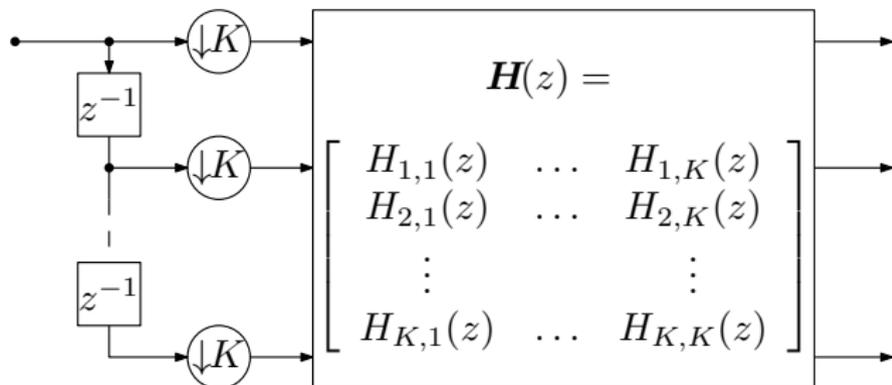
▶ With the K-fold polyphase decomposition of the analysis filters

$$H_k(z) = \sum_{n=1}^{K} H_{k,n}(z^K) z^{-n+1} \tag{5}$$
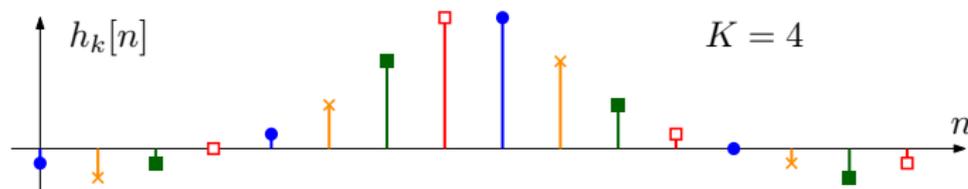
▶ the polyphase analysis matrix is a polynomial matrix:

$$\boldsymbol{H}(z) = \begin{bmatrix} H_{1,1}(z) & H_{1,2}(z) & \dots & H_{1,K}(z) \\ H_{2,1}(z) & H_{2,2}(z) & \dots & H_{2,K}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{K,1}(z) & H_{K,2}(z) & \dots & H_{K,K}(z) \end{bmatrix} \tag{6}$$

Synthesis Filter Bank

- Critically decimated K-channel synthesis filter bank:
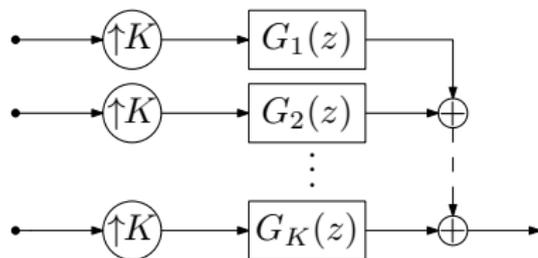
- equivalent polyphase representation:
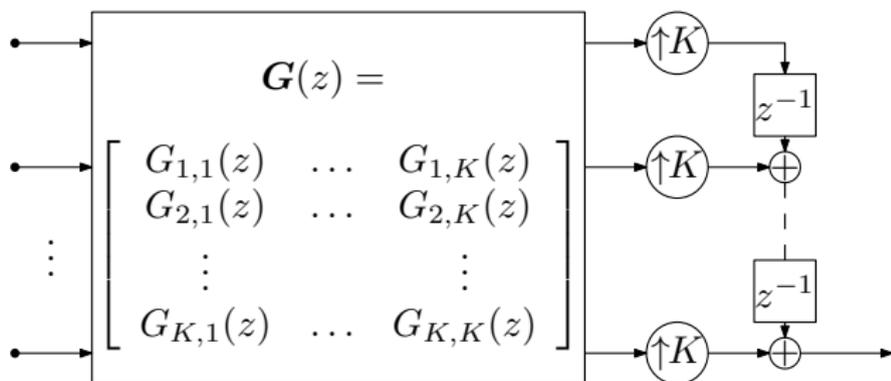
Polyphase Synthesis Matrix

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

$$\boldsymbol{G}(z) = \left[\begin{array}{cccc} G_{1,1}(z) & G_{1,2}(z) & \ldots & G_{1,K}(z) \\ G_{2,1}(z) & G_{2,2}(z) & \ldots & G_{2,K}(z) \\ \vdots & \vdots & \ddots & \vdots \\ G_{K,1}(z) & G_{K,2}(z) & \ldots & G_{K,K}(z) \end{array} \right] \quad (7)$$

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

$$\boldsymbol{G}(z)\boldsymbol{H}(z) = \mathbf{I} \; ; \quad (8)$$

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

Space-Time Covariance Matrix

University of
Strathclyde
Engineering

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

$$\mathbf{x}^{\mathrm{T}}[n] = [x_1[n] \ \ x_2[n] \ \ \ldots \ \ x_M[n]] \ ; \tag{9}$$

▶ 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

$$\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n - \tau]\} \tag{10}$$

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

$$\mathbf{R}[\tau] = \left[\begin{array}{cc} \mathcal{E}\{x_1[n]x_1^*[n - \tau]\} & \mathcal{E}\{x_1[n]x_2^*[n - \tau]\} \\ \mathcal{E}\{x_2[n]x_1^*[n - \tau]\} & \mathcal{E}\{x_2[n]x_2^*[n - \tau]\} \end{array} \right] \tag{11}$$

Cross-Spectral Density Matrix

▶ example for a space-time covariance matrix $\mathbf{R}[\tau] \in \mathbb{R}^{2 \times 2}$:
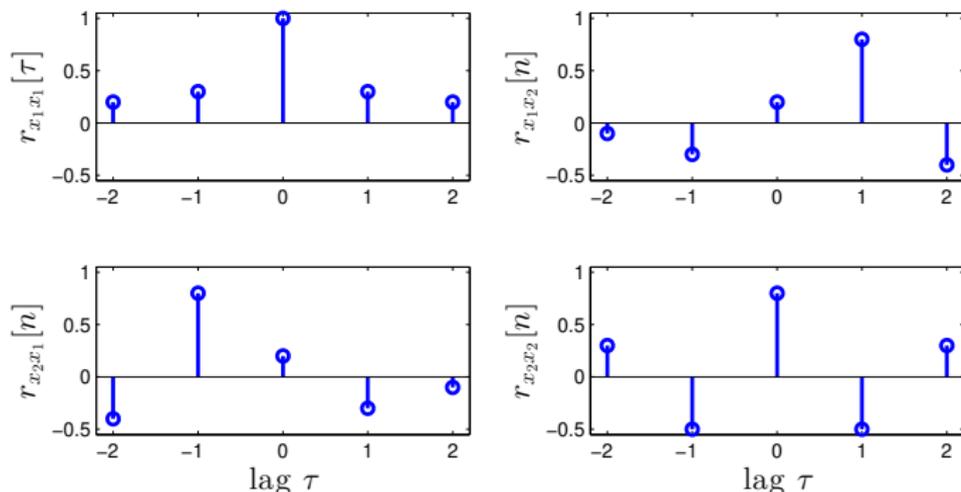
▶ the cross-spectral density (CSD) matrix

$$\boldsymbol{R}(z) \circ\!\!-\!\!\bullet \mathbf{R}[\tau] \qquad (12)$$

is a polynomial matrix.

Parahermitian Operator

- A parahermitian operation is indicated by $\{\cdot\}^{\mathrm{P}}$, and compared to the Hermitian ($=$ complex conjugate transpose) of a matrix additionally performs a time-reversal;

- example:

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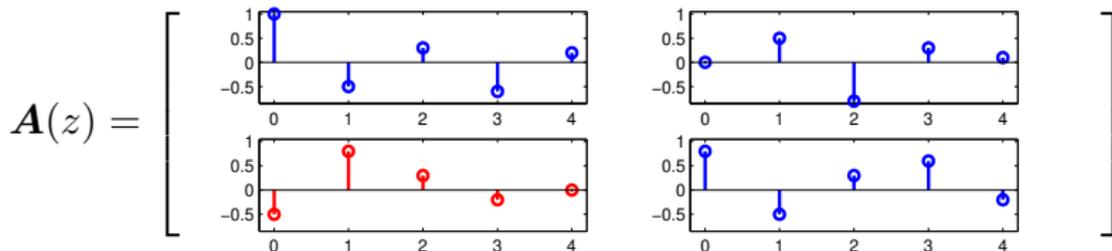
- parahermitian $\boldsymbol{A}^{\mathrm{P}}(z) = \boldsymbol{A}^{\mathrm{H}}(1/z^*)$:

$$\boldsymbol{A}^{\mathrm{P}}(z) = $$
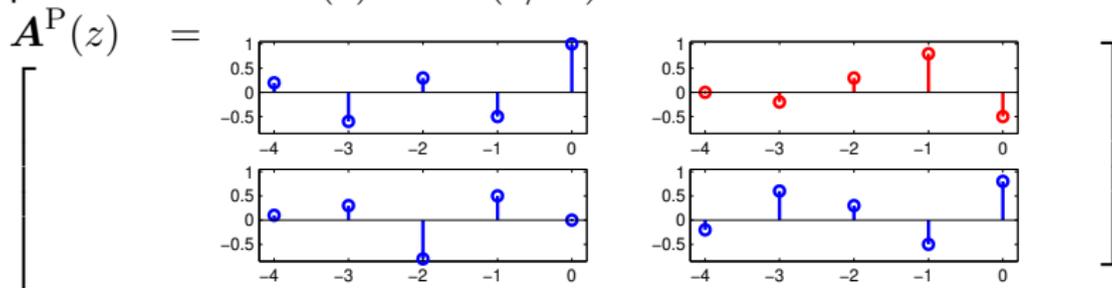
Parahermitian Property

- 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:

$$\boldsymbol{R}(z) = \begin{bmatrix} & \\ & \end{bmatrix} = \boldsymbol{R}^{\mathrm{P}}(z)$$

Paraunitary Matrices

University of
Strathclyde
Engineering
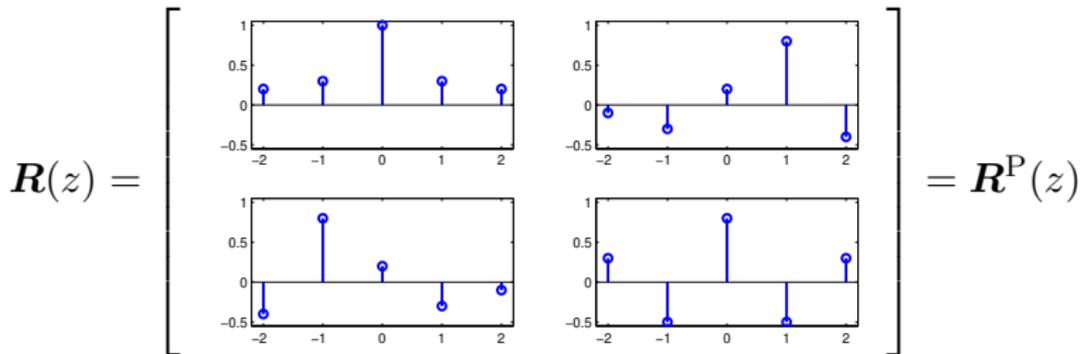
▶ Recall that $\mathbf{A} \in \mathbb{C}$ (or $\mathbf{A} \in \mathbb{R}$) is a unitary (or orthonormal) matrix if $\mathbf{A}\mathbf{A}^{\mathrm{H}} = \mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}$;

▶ in the polynomial case, $\mathbf{A}(z)$ is paraunitary if

$$\mathbf{A}(z)\mathbf{A}^{\mathrm{P}}(z) = \mathbf{A}^{\mathrm{P}}(z)\mathbf{A}(z) = \mathbf{I} \qquad (13)$$

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

$$\mathbf{A}^{-1}(z) = \mathbf{A}^{\mathrm{P}}(z) \qquad (14)$$

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

Attempt of Gaussian Elimination

University of
Strathclyde
Engineering

▶ System of polynomial equations:

$$\left[\begin{array}{cc} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z) \end{array}\right] \cdot \left[\begin{array}{c} X_1(z) \\ X_2(z) \end{array}\right] = \left[\begin{array}{c} B_1(z) \\ B_2(z) \end{array}\right] \qquad (15)$$

▶ modification of 2nd row:

$$\left[\begin{array}{cc} A_{11}(z) & A_{12}(z) \\ A_{11}(z) & \frac{A_{11}(z)}{A_{21}(z)}A_{22}(z) \end{array}\right] \cdot \left[\begin{array}{c} X_1(z) \\ X_2(z) \end{array}\right] = \left[\begin{array}{c} B_1(z) \\ \frac{A_{11}(z)}{A_{21}(z)}B_2(z) \end{array}\right] \quad (16)$$

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

$$\left[\begin{array}{cc} A_{11}(z) & A_{12}(z) \\ 0 & \frac{A_{11}(z)A_{22}(z)-A_{12}(z)A_{21}(z)}{A_{21}(z)} \end{array}\right] \cdot \left[\begin{array}{c} X_1(z) \\ X_2(z) \end{array}\right] = \left[\begin{array}{c} B_1(z) \\ \bar{B}_2(z) \end{array}\right]$$

$$(17)$$

▶ penalty: we end up with rational polynomials.

Polynomial Eigenvalue Decomposition

[McWhirter Decomposition, 2007]

▶ Polynomial EVD of the CSD matrix

$$\boldsymbol{R}(z) \approx \hat{\boldsymbol{U}}(z) \; \hat{\boldsymbol{\Gamma}}(z) \; \hat{\boldsymbol{U}}^{\mathrm{P}}(z) \qquad (18)$$

▶ with paraunitary $\hat{\boldsymbol{U}}(z)$, s.t. $\hat{\boldsymbol{U}}(z)\hat{\boldsymbol{U}}^{\mathrm{P}}(z) = \mathbf{I}$;

▶ diagonalised and spectrally majorised $\hat{\boldsymbol{\Gamma}}(z)$:
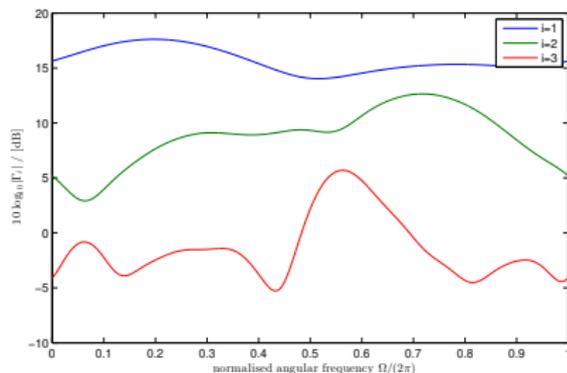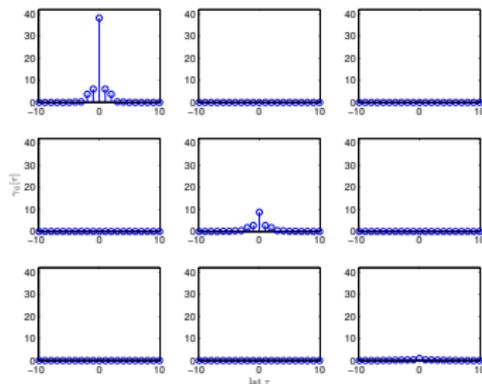
▶ algorithms exist, with proven convergence; **but we do not know to which values they converge**.

Source Model for CSD Matrix

University of
Strathclyde
Engineering

▶ General source model for a CSD matrix with L unit variance, uncorrelated and mutually independent inputs $u_\ell[n]$, $\ell = 1 \ldots L$, and M measurements $x_m[n]$, $m = 1 \ldots M$:
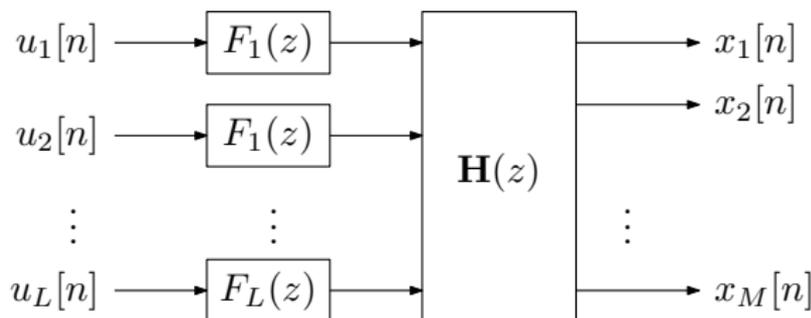
▶ source PSDs via innovation filter: $S_\ell(z) = F_\ell(z) F_\ell^{\mathrm{P}}(z)$;

▶ for the CSD matrix:

$$\boldsymbol{R}(z) = \boldsymbol{H}(z) \boldsymbol{F}(z) \boldsymbol{F}^{\mathrm{P}}(z) \boldsymbol{H}^{\mathrm{P}}(z) \ . \tag{19}$$

Definitions around the CSD Matrix

► The innovation filters $F_\ell(z)$ and $\boldsymbol{H}(z)$ can be polynomials or rational functions in z, but are assumed to be causal and stable;

► they can be represented as a power series (or matrix of power series), $n \in \mathbb{N}$,

$$X(z) = \sum_n x_n z^{-n}$$

► we have a Laurent series, if $n \in \mathbb{Z}$;

► power and Laurent series do not need to converge (=absolutely summable, $\sum_n |x_n| < \infty$);

► they are analytic if they converge;

► if the sums are finite, we speak of polynomials ($n \in \mathbb{N}$) and Laurent polynomials ($n \in \mathbb{Z}$);

► $\boldsymbol{R}(z)$ is a Laurent series, and analytic in an annulus $\mathcal{D}$;

► analyticity: $\boldsymbol{R}(z)$ is infinitely differentiable.

Eigenvalue Decomposition (EVD)

University of
Strathclyde
Engineering

- A Hermitian matrix $\mathbf{R} = \mathbf{R}^{\mathrm{H}}$ be positive semidefinite (given for any covariance matrix, or anything arising from $\mathbf{R} = \mathbf{A}\mathbf{A}^{\mathrm{H}}$, with $\mathbf{A}$ arbitrary);

- eigenvalues and eigenvectors:

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

- eigenvalues $\lambda \in \mathbb{R}$, $\lambda \geq 0$;

- eigenvectors are chosen to be orthonormal, but 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 $\mathbf{V}$:

$$[\mathbf{q}'_m, \ \ldots \ \mathbf{q}'_{m+C-1}] = [\mathbf{q}_m, \ \ldots \ \mathbf{q}_{m+C-1}] \, \mathbf{V} \ , \qquad (20)$$

Problem Statement

University of
Strathclyde
Engineering

▶ We want to investigate the existence and uniquess of the EVD of a positive semidefinite parahermitian matrix $\boldsymbol{R}(z)$:

$$\boldsymbol{R}(z) = \boldsymbol{U}(z)\,\boldsymbol{\Gamma}(z)\,\boldsymbol{U}^{\mathrm{P}}(z)\,, \qquad (21)$$

▶ $\boldsymbol{U}(z)$ must be paraunitary;
▶ $\boldsymbol{\Gamma}(z)$ must be diagonal and parahermitian.
▶ we call this a parahermitian matrix EVD;
▶ why do we care about the z-domain representation? If we can write $\boldsymbol{U}(z)$ and $\boldsymbol{\Gamma}(z)$ as convergent power/Laurent series, we can immediately derive time domain realisations $\mathbf{U}[n]$ and $\boldsymbol{\Gamma}[n]$.

EVD on the Unit Circle

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

$$\boldsymbol{R}(e^{j\Omega}) = \boldsymbol{Q}(e^{j\Omega})\boldsymbol{\Lambda}(e^{j\Omega})\boldsymbol{Q}^{\mathrm{H}}(e^{j\Omega}) . \tag{22}$$

- for every Ω, $\boldsymbol{Q}(e^{j\Omega})$ and $\boldsymbol{\Lambda}(e^{j\Omega})$ fulfill the properties of the EVD;
- can we find $\mathbf{U}(z)$ and $\boldsymbol{\Gamma}(z)$ such that on the unit circle

$$\mathbf{U}(e^{j\Omega}) = \boldsymbol{Q}(e^{j\Omega}) \tag{23}$$

$$\boldsymbol{\Gamma}(e^{j\Omega}) = \boldsymbol{\Lambda}(e^{j\Omega}) \tag{24}$$

- all now depends on the continuity/smoothness of $\boldsymbol{Q}(e^{j\Omega})$ and $\boldsymbol{\Lambda}(e^{j\Omega})$;

Matrix Perturbation Theory

University of
Strathclyde
Engineering

▶ If we know that $\boldsymbol{R}(e^{j\Omega})$ varies smoothly, what can be say about $\boldsymbol{Q}(e^{j\Omega})$ and $\boldsymbol{\Lambda}(e^{j\Omega})$?

▶ eigenvalues (Hoffman-Wielandt, 1953):

$$\sum_i |\lambda_i(e^{j\Omega}) - \lambda_i(e^{j(\Omega + \Delta\Omega)})| \leq \|\boldsymbol{R}(e^{j\Omega}) - \boldsymbol{R}(e^{j(\Omega+\Delta\Omega)})\|_{\mathrm{F}}, \quad (25)$$

▶ eigenvector subspace (Golub & van Loan):

$$\boldsymbol{Q}^{\mathrm{H}}(e^{j\Omega}) \left(\boldsymbol{R}(e^{j(\Omega+\Delta\Omega)}) - \boldsymbol{R}(e^{j\Omega}) \right) \boldsymbol{Q}(e^{j\Omega}) = \\ \left[\begin{array}{cc} \underbrace{\mathbf{E}_{11}(e^{j\Omega}, \Delta\Omega)}_{C} & \underbrace{\mathbf{E}_{21}^{\mathrm{H}}(e^{j\Omega}, \Delta\Omega)}_{} \\ \mathbf{E}_{21}(e^{j\Omega}, \Delta\Omega) & \mathbf{E}_{22}(e^{j\Omega}, \Delta\Omega) \end{array} \right]. \quad (26)$$

$$\underbrace{\phantom{\mathbf{E}_{11}(e^{j\Omega}, \Delta\Omega)}}_{C} \quad \underbrace{\phantom{\mathbf{E}_{22}(e^{j\Omega}, \Delta\Omega)}}_{M-C}$$

$$\mathrm{dist}\{\mathcal{Q}_1(e^{j\Omega}), \mathcal{Q}_1(e^{j(\Omega+\Delta\Omega)})\} \leq \frac{4}{\delta} \|\mathbf{E}_{21}(e^{j\Omega}, \Delta\Omega)\|_{\mathrm{F}}. \quad (27)$$

Weierstrass (1885)

University of
Strathclyde
Engineering

- ▶ As long as $\boldsymbol{Q}(e^{j\Omega})$ is continuous, it can be arbitrarily closely approximated by a converging series of polynomials;

- ▶ the functions may change as the approximation order changes;

- ▶ hence there is not just one function in the limit;

- ▶ we would like something more precise, and ideally obtain an approximant by truncating the limit function.

Analytic EVD

- Franz Rellich (1937) for a self-adjoint, analytic $\boldsymbol{R}(e^{j\Omega})$:

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

- $\boldsymbol{Q}(e^{j\Omega})$ and $\boldsymbol{\Lambda}(e^{j\Omega})$ *can* be chosen analytic;

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

Eigenvalue Considerations
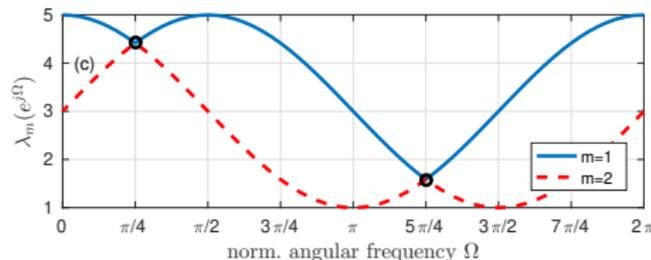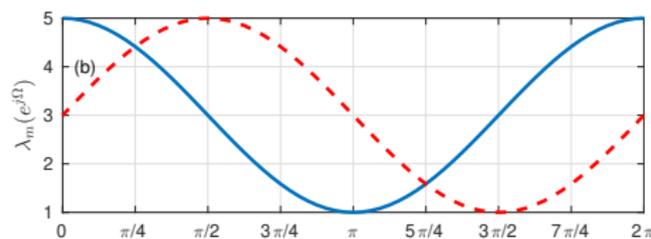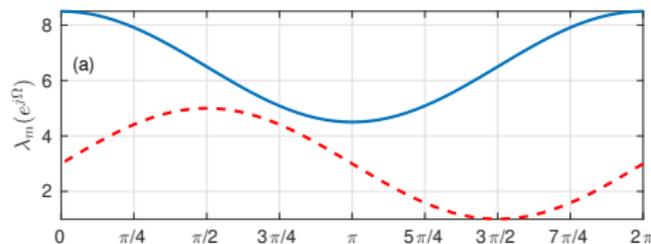
We distinguish three cases:

- non-overlapping eigenvalues $\lambda_m(e^{j\Omega})$, where all eigenvalues have algebraic multiplicity one for all frequencies Ω;

- overlapping, maximally smooth eigenvalues;

- overlapping, spectrally majorised PSDs.

Distinct Eigenvalues — Eigenvalues

University of
Strathclyde
Engineering

- ▶ Since $\lambda_m(e^{j\Omega}) \in \mathbb{R}$ we can write

$$\hat{\lambda}_m^{(N)}(e^{j\Omega}) = \sum_{\ell=0}^{N} c_\ell e^{j\ell\Omega} + c_\ell^* e^{-j\ell\Omega}, \quad c_\ell \in \mathbb{C} \; ; \qquad (28)$$

- ▶ with $\hat{\boldsymbol{\Lambda}}^{(N)}(e^{j\Omega}) = \text{diag}\left\{\hat{\lambda}_1^{(N)}(e^{j\Omega}), \, \dots \, , \hat{\lambda}_M^{(N)}(e^{j\Omega})\right\}$, uniform convergence implies that for every $\epsilon_\Lambda > 0$ there exists $N_0 > 0$ such that (Stone-Weierstrass)

$$\sup_{\Omega \in [0,2\pi)} \left\| \hat{\boldsymbol{\Lambda}}^{(N)}(e^{j\Omega}) - \boldsymbol{\Lambda}(e^{j\Omega}) \right\| < \epsilon_\Lambda, \qquad (29)$$

- ▶ this requires Hölder continuity for $\boldsymbol{\Lambda}(e^{j\Omega})$:

$$|X(e^{j\Omega_1}) - X(e^{j\Omega_2})| \le C|e^{j\Omega_1} - e^{j\Omega_2}|^\alpha \qquad (30)$$

Analytic Eigenvalues

University of
Strathclyde
Engineering

▶ We had

$$\hat{\lambda}_m^{(N)}(e^{j\Omega}) = \sum_{\ell=0}^{N} c_\ell e^{j\ell\Omega} + c_\ell^* e^{-j\ell\Omega}, \quad c_\ell \in \mathbb{C} ;$$

▶ for absolutely summability $\sum_\ell |c_\ell| < \infty$, we require a slightly stricter form of Hölder with $\alpha > \frac{1}{2}$;

▶ then there exist unique eigenvalues of a parahermitian matrix:

$$\gamma_m(z) = \sum_{\ell=-\infty}^{\infty} c_\ell z^{-\ell}$$

▶ these are generally infinite sums or transcendental functions in z (Weierstrass);

▶ an arbitrarily close approximation by Laurent polynomials is possible via truncation:

$$\hat{\gamma}_m(z) = \sum_{\ell=-N}^{N} c_\ell z^{-\ell}$$

Example of a 2×2 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 $T(z) = \text{trace}\{\boldsymbol{R}(z)\}$;
- ▶ this leads to

$$\gamma_{1,2}(z) = \frac{1}{2}T(z) \pm \frac{1}{2}\sqrt{T(z)T^{\mathrm{P}}(z) - 4D(z)} \; ; \qquad (31)$$

- ▶ 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.

Example cont'd

- Colin Maclaurin to the rescue: for every root of $S(z)$,

$$\sqrt{1 - \beta z^{-1}} = \sum_{n=0}^{\infty} \xi_n \beta^n z^{-n} \tag{32}$$

$$\frac{1}{\sqrt{1 - \alpha z^{-1}}} = \left(\sum_{n=0}^{\infty} \xi_n \alpha^n z^{-n}\right)^{-1} \tag{33}$$

$$= \sum_{n=0}^{\infty} \chi_n \alpha^n z^{-n} \tag{34}$$

- with coefficients

$$\xi_n = (-1)^n \binom{\frac{1}{2}}{n} = \frac{(-1)^n}{n!} \prod_{i=0}^{n-1} \left(\frac{1}{2} - i\right) , \tag{35}$$

$$\chi_n = (-1)^n \binom{-\frac{1}{2}}{n} = \frac{(-1)^{n-1}}{n!} \prod_{i=0}^{n-1} \left(\frac{1}{2} + i\right) . \tag{36}$$

Example cont'd

- Coefficients ξ_n and χ_n for $n = 0 \ldots 50$:
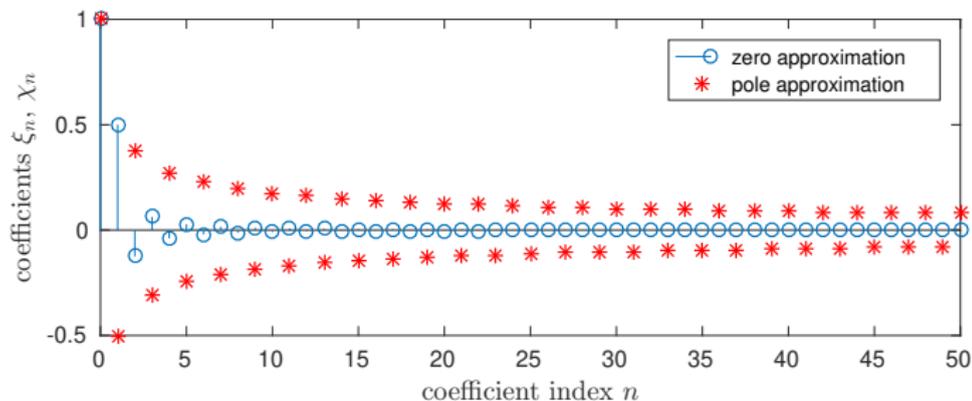
- these coefficients additionally dampen a geometric series;
- only if $S(z)$ has double zeros (and double poles) is a polynomial (rational) solution possible;
- in general, the result are transcendental eigenvalues.

Numerical Example
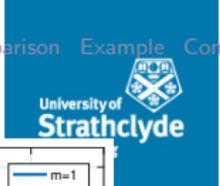
University of
Strathclyde

Example from Icart & Comon (2012):

$$\boldsymbol{R}(z) = \left[\begin{array}{cc} 1 & 1 \\ 1 & -2z + 6 - 2z^{-1} \end{array} \right]$$

- solution on unit circle;
- coefficients of analytic eigenvalues;
- decay of coefficients.
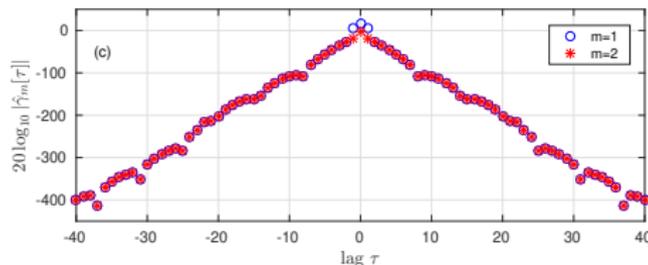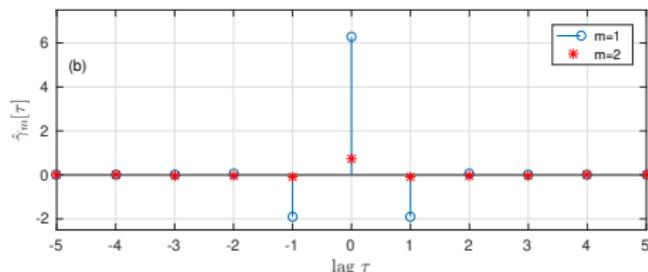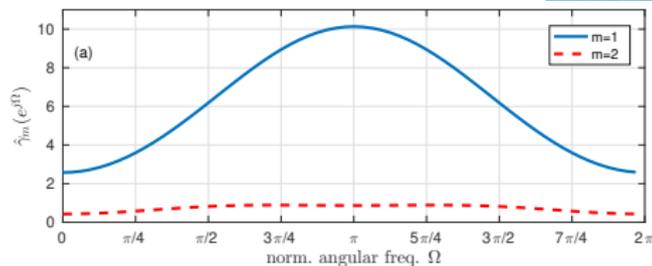
Eigenvectors

- ▶ Recall: eigenvectors $\boldsymbol{q}_m(e^{j\Omega})$ have an arbitrary phase response;
- ▶ if this phase response is Hölder continuous, analytic eigenvectors exists; some phase ambiguity remains;
- ▶ if the phase response is discontinuous at Ω_0,

$$\lim_{N\to\infty} \hat{\boldsymbol{q}}_m^{(N)}(e^{j\Omega_0}) = \frac{1}{2}\lim_{\Omega\to 0}\Big(\boldsymbol{q}_m(e^{j(\Omega_0-\Omega)}) + \\ + \boldsymbol{q}_m(e^{j(\Omega_0+\Omega)})\Big) ; \qquad (37)$$

  but this Fourier series does converge neither uniformly nor absolutely; analytic eigenvalues do not exist;

- ▶ there is a grey area (continuous but not Hölder) where we cannot make any statement;
- ▶ analytic approximations are possible in both cases.

Overlapping Eigenvalues

▶ We now inspect the case of overlapping eigenvalues:
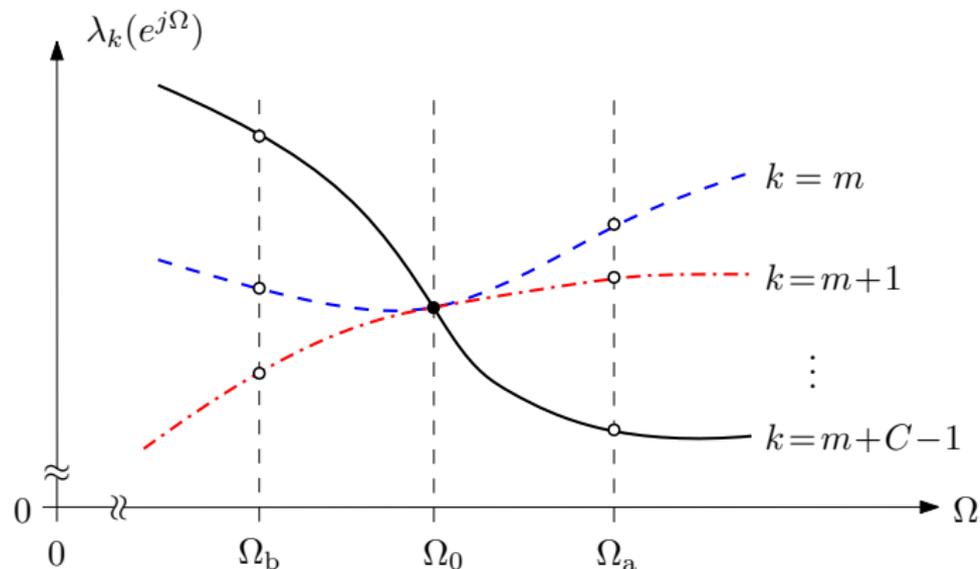
▶ maximally smooth continuation: analytic case ... same as before;
▶ spectrally majorised case: eigenvalues $\lambda_m(e^{j\Omega})$ can be shown to be Lipschitz (stronger than Hölder) and analytic eigenvalues $\gamma_m(z)$ exis s.t. $\gamma_m(z)|_{z=e^{j\Omega}} = \lambda_m(e^{j\Omega})$.

Polynomial Approximation of Eigenvalues
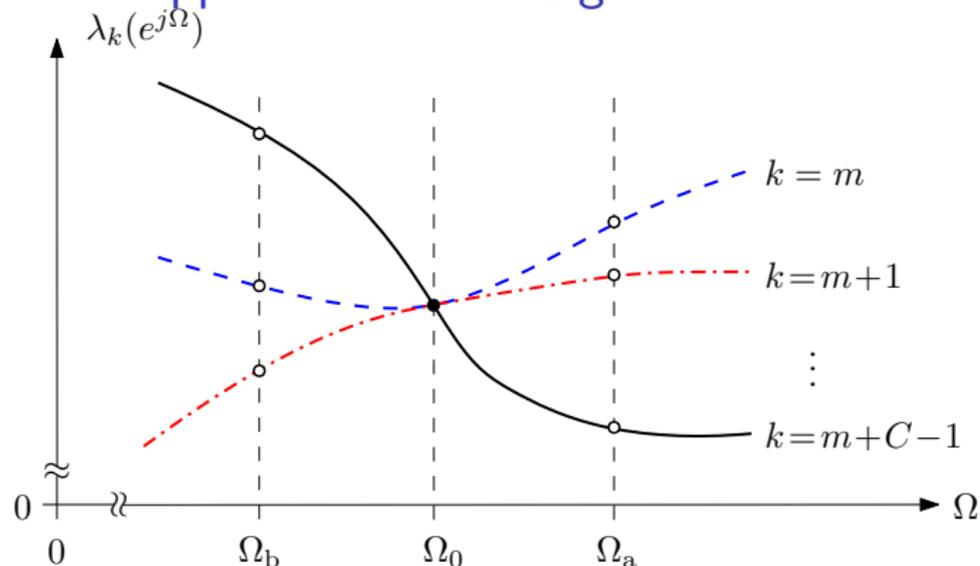
- ▶ Maximally smooth / analytic case will generally require a transcendental solution;
- ▶ for the spectrally majorised case, the convergence of the Laurent series will be slower;
- ▶ we expect to need a higher approximation order in case of spectral majorisation.

Eigenvectors for Analytic Eigenvalues
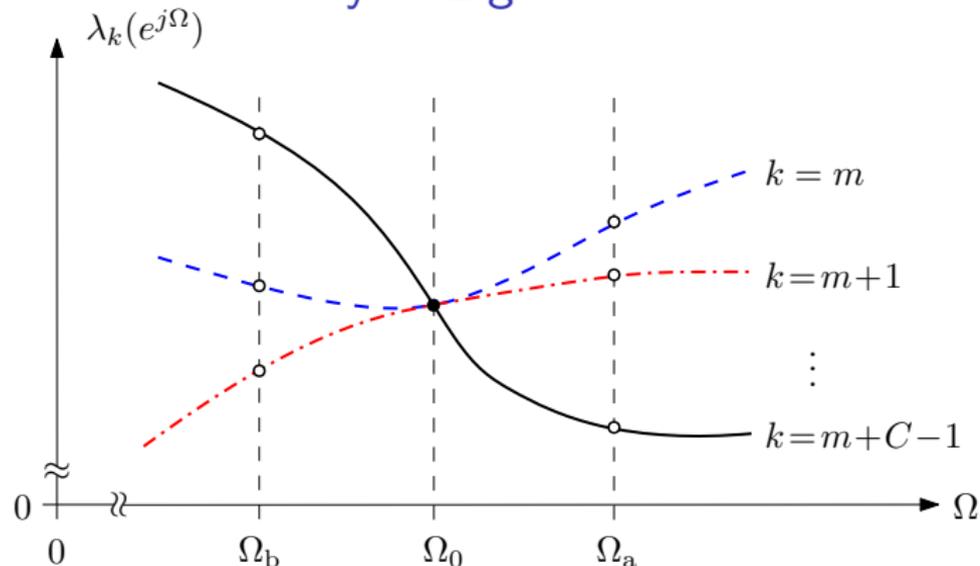
- For analytic eigenvalues, unique 1-d subspaces for the eigenvectors can be shown to continuous blend in and out of algebraic multiplicities;
- the existence of analytic eigenvectors $\mathbf{u}_m(z)$ again depends on the selection of the phase response within each subspace.

Eigenvectors for Spectrally Majorised Eigenvalues

University of
Strathclyde
Engineering
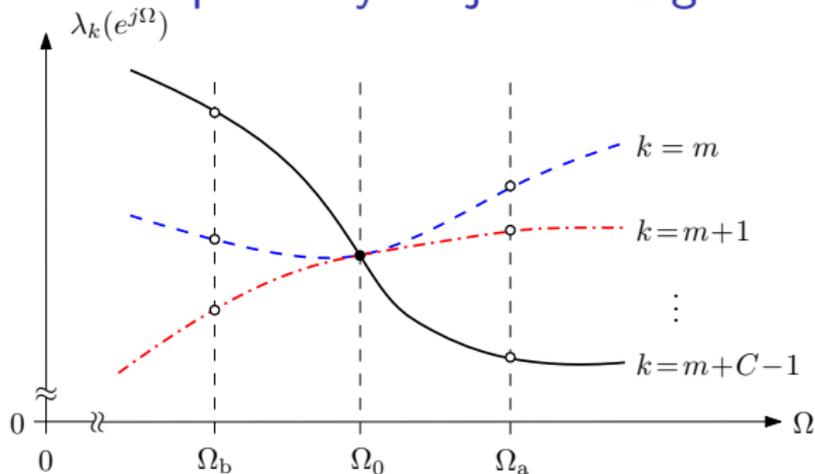
- For spectrally majorised eigenvalues, eigenvectors are discontinuous in the algebraic multiplicity;
- hence no exact analytic eigenvectors $\boldsymbol{u}_m(z)$ exist in this case, irrespective of the phase response;
- a Fourier series representation can be found if $\boldsymbol{q}_m(e^{j\Omega})$ has only a finite number of discontinuities;
- the series is not analytic, but a truncation can be.

Numerical Example

University of
Strathclyde
Engineering

▶ We take a parahermitian matrix with known factors:

$$\mathbf{\Gamma}(z) = \left[\begin{array}{cc} z + 3 + z^{-1} & \\ & -jz + 3 + jz^{-1} \end{array} \right]$$

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

▶ parahermitian matrix $\boldsymbol{R}(z) = \boldsymbol{U}(z)\mathbf{\Gamma}(z)\boldsymbol{U}^{\mathrm{P}}(z)$:

$$\boldsymbol{R}(z) = \left[\begin{array}{cc} \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} & \frac{1+j}{2}z^2 + \frac{1-j}{2} \\ \frac{1+j}{2} + \frac{1-j}{2}z^{-2} & \frac{1-j}{2}z + 3 + \frac{1+j}{2}z^{-1} \end{array} \right].$$

Example — Maximally Smooth Case

▶ The factors represent the solution for the analytic eigenvalues and eigenvectors;
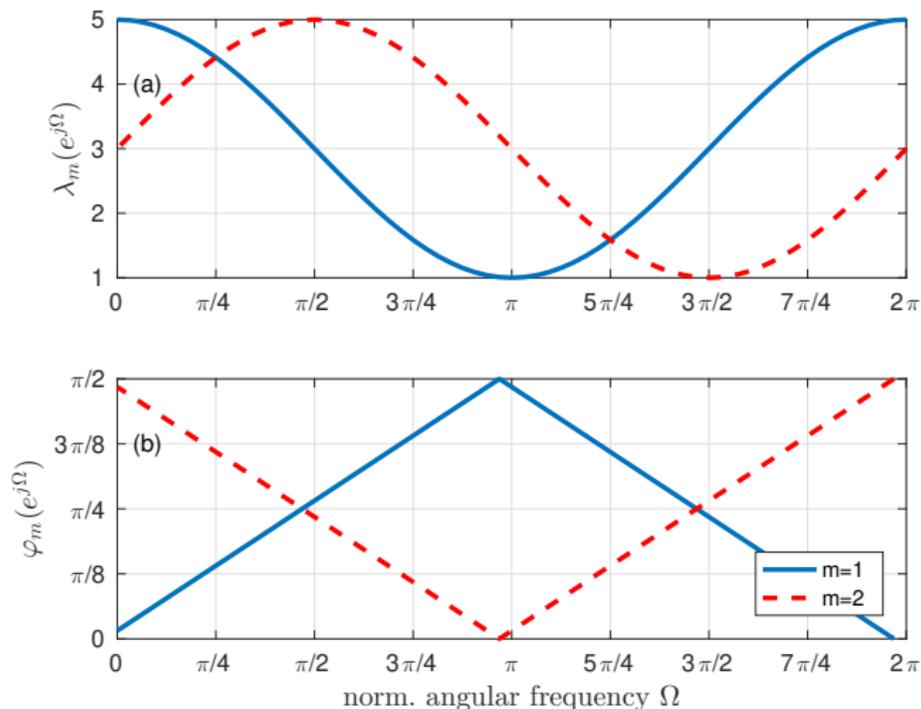
Example — Spectral Majorisation

▶ On the unit circle, we end up with piecewise smooth eigenvalues and piecewith continuous eigenvectors — analytic solution only for the former!
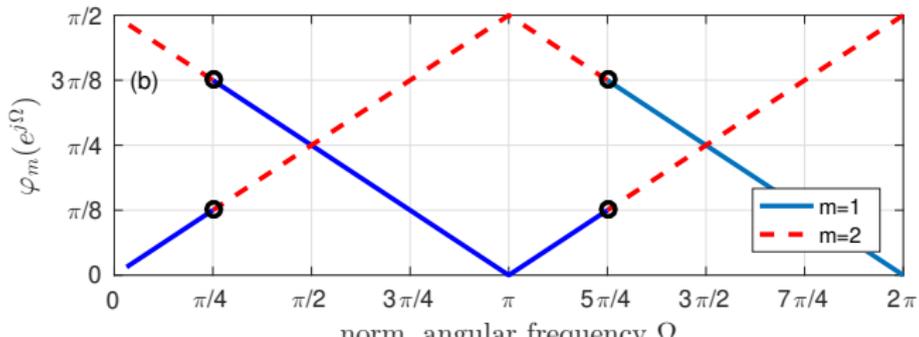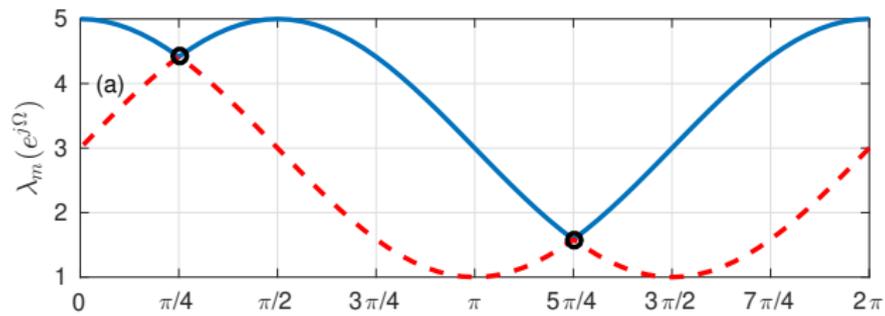
Example — Algorithmic Solution

- Using an algorithm to approximate the McWhirter factorisation (order 24 for $\hat{\mathbf{\Gamma}}(z)$ and order 84 for $\hat{\mathbf{U}}(z)$):

McWhirter Decomposition

University of
Strathclyde
Engineering
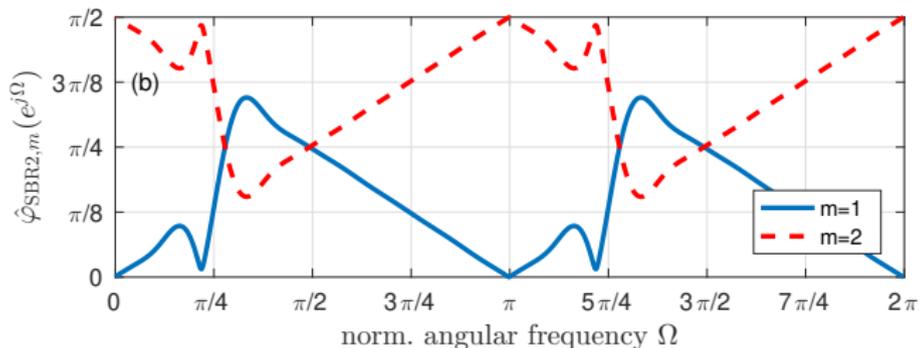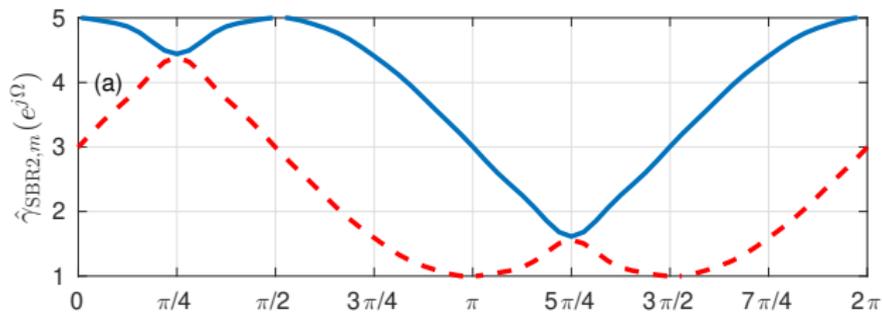
► John McWhirter et al. (2007): polynomial eigenvalue decomposition of a parahermitian matrix:

$$\boldsymbol{R}(z) \approx \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z)$$

► paraunitary (i.e. lossless) matrix $\boldsymbol{Q}(z)$, s.t. $\boldsymbol{Q}(z)\boldsymbol{Q}^{\mathrm{P}}(z) = \mathbf{I}$;

► diagonal and spectrally majorised $\boldsymbol{\Lambda}(z)$:

Parahermitian Matrix EVD (PEVD)

University of
Strathclyde
Engineering
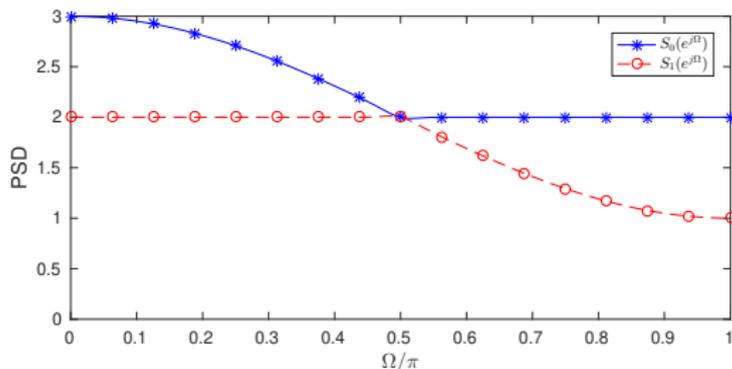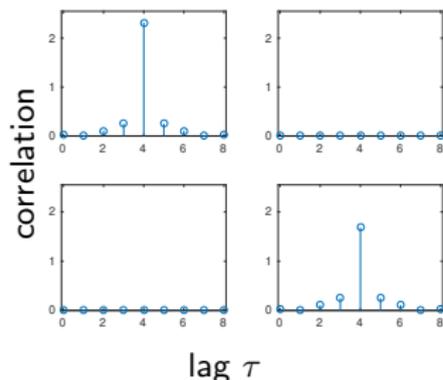
▶ Franz Rellich (1937): for $\boldsymbol{R}(e^{j\Omega})$ analytic, there exist analytic eigenvectors $\boldsymbol{\Gamma}(e^{j\Omega})$ and analytic eigenvalues $\boldsymbol{U}(e^{j\Omega})$;

▶ can be generalised to

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

lag τ
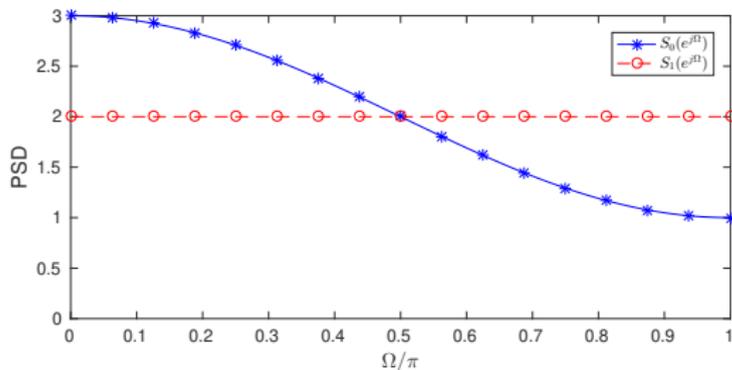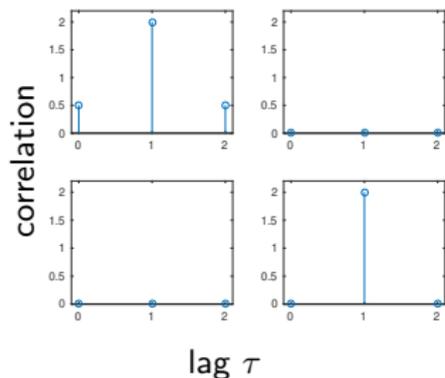
▶ eigenvalues are unique, eigenvectors can be modified by arbitrary allpass filters $H(z)$ (s.t. $H(z)H^{\mathrm{P}}(z) = 1$),

$$\boldsymbol{R}(z)\boldsymbol{u}(z)H(z) = \gamma(z)\boldsymbol{u}(z)H(z) \; .$$

Iterative PEVD Algorithms

- ▶ Second order sequential best rotation (SBR2, McWhirter 2007);
- ▶ 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 ith 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

University of
Strathclyde
Engineering

▶ An elementary paraunitary matrix [Vaidyanathan] is defined as

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

▶ we utilise a different definition:

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

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

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

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

Sequential Best Rotation Algorithm (McWhirter)

- At iteration i, consider $\boldsymbol{S}^{(i-1)}(z)$ ○—● $\boldsymbol{S}^{(i-1)}[\tau]$
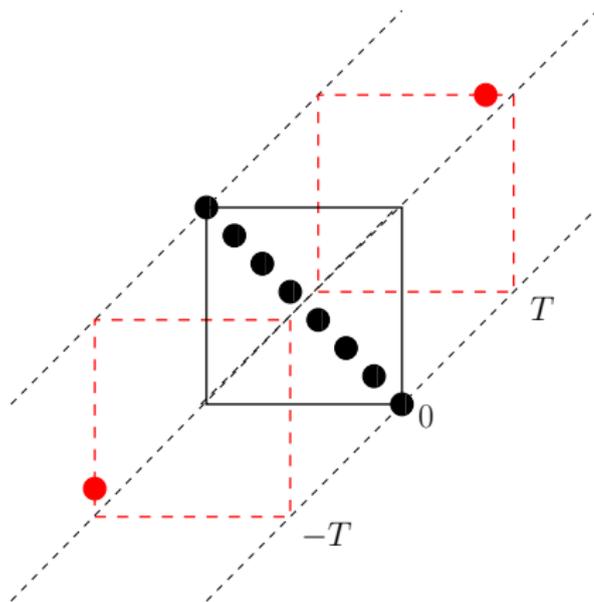
Sequential Best Rotation Algorithm (McWhirter)

- $\tilde{\boldsymbol{D}}^{(i)}(z)\boldsymbol{S}^{(i-1)}(z)\boldsymbol{D}^{(i)}(z)$
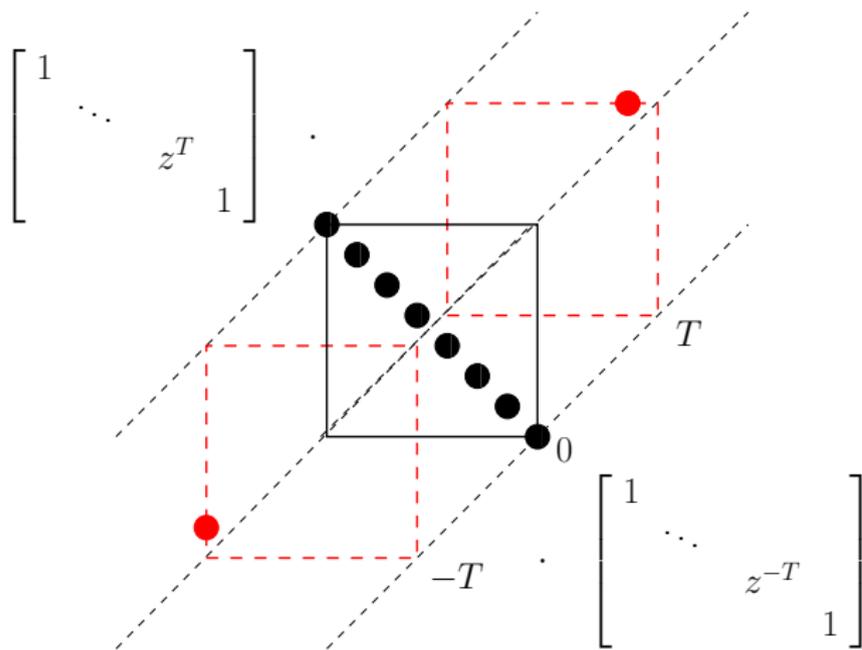
Sequential Best Rotation Algorithm (McWhirter)

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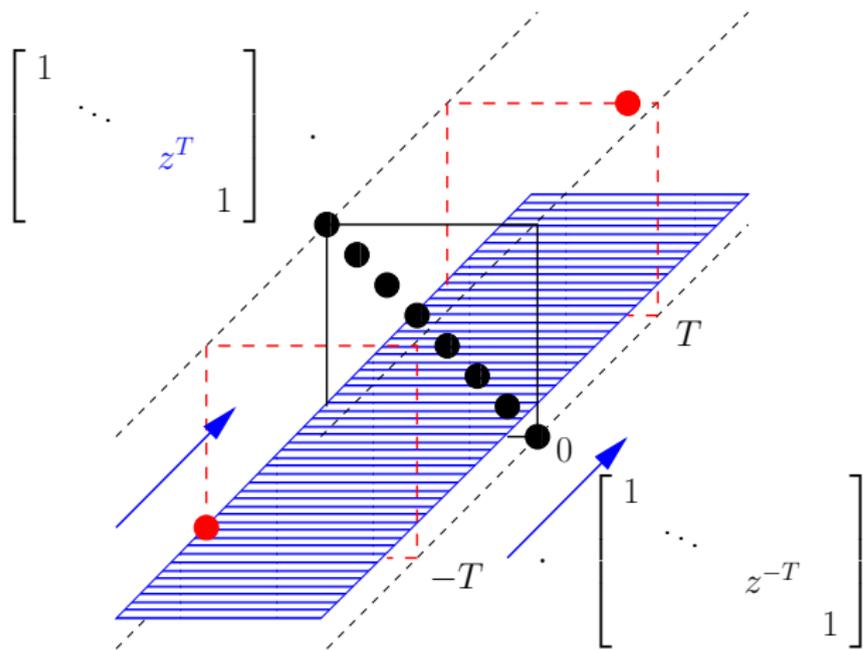
Sequential Best Rotation Algorithm (McWhirter)

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

$$\cdot \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & z^{-T} & \\ & & & 1 \end{bmatrix}$$
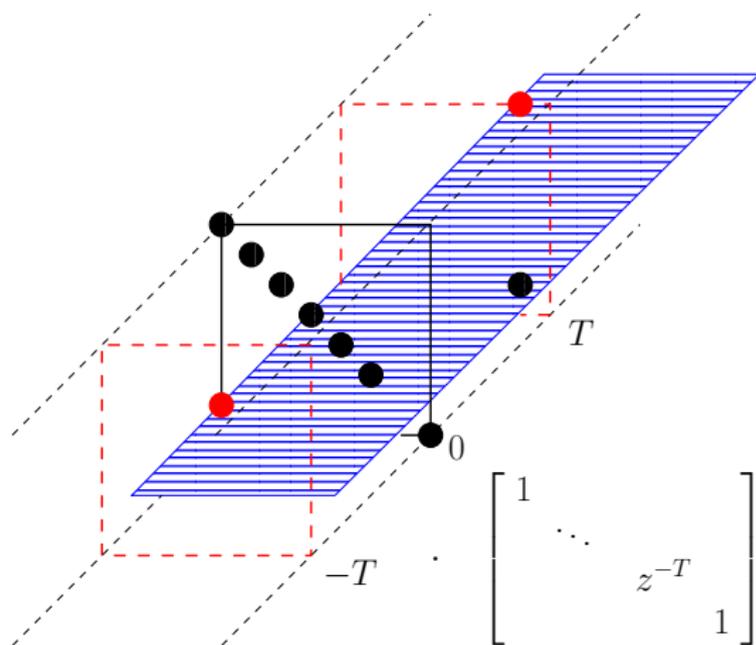
Sequential Best Rotation Algorithm (McWhirter)

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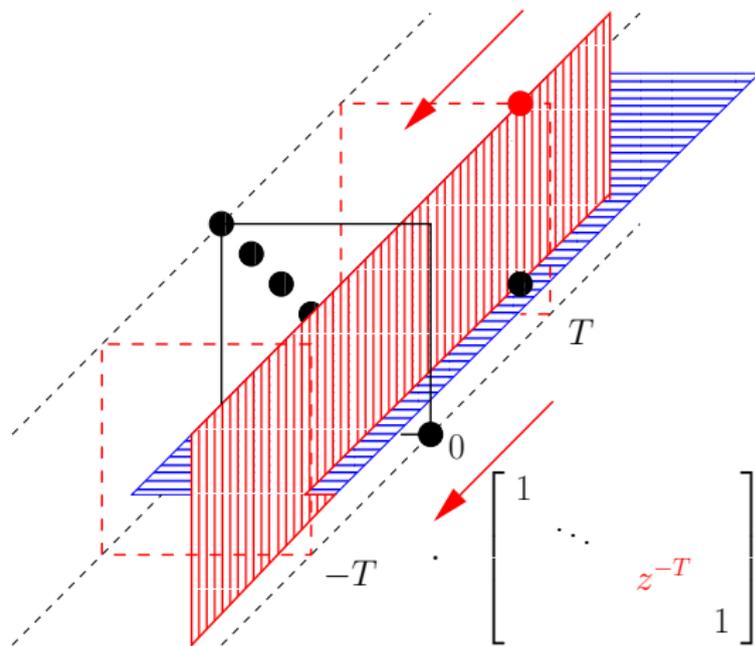
Sequential Best Rotation Algorithm (McWhirter)

- the off-diagonal element at $-T$ has now been translated to lag zero
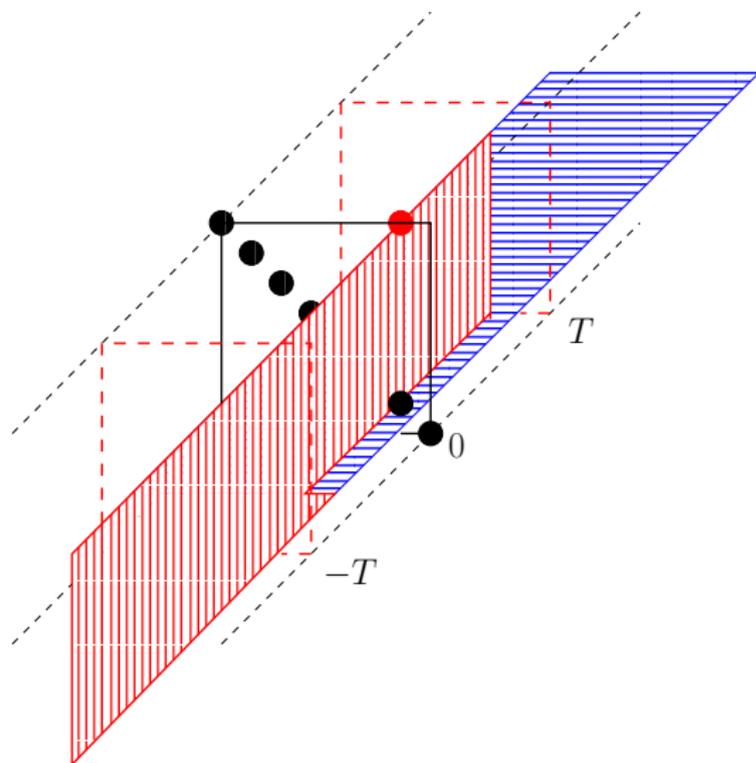
Sequential Best Rotation Algorithm (McWhirter)

- 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.
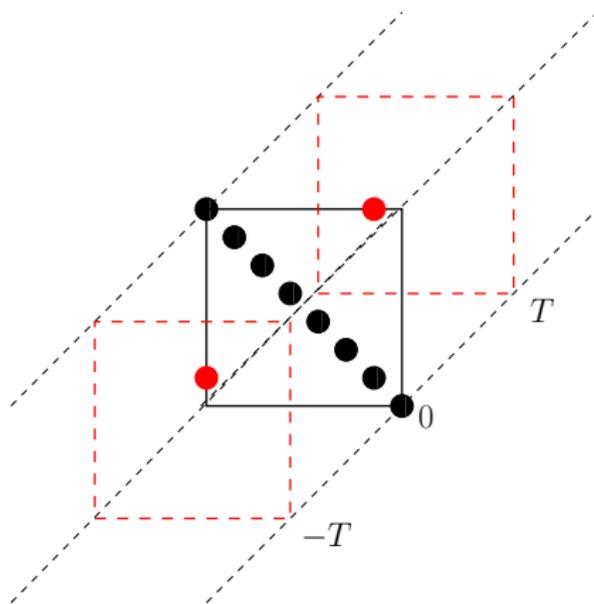
Sequential Best Rotation Algorithm (McWhirter)

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

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

$$\cdot \begin{bmatrix} c & & -e^{-j\vartheta}s \\ & \mathbf{I} & \\ e^{j\vartheta}s & & c \\ & & & 1 \end{bmatrix}$$
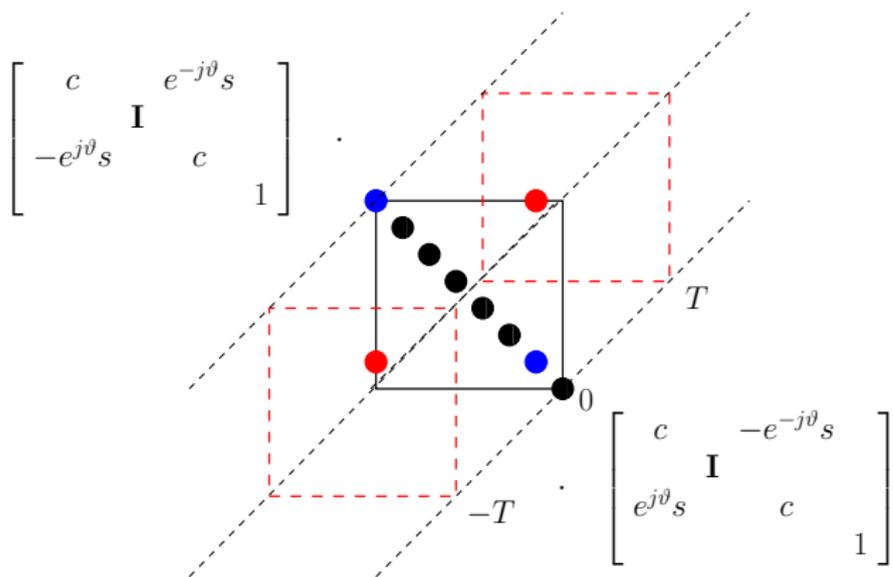
Sequential Best Rotation Algorithm (McWhirter)

▶ iteration i is completed, having performed

$$\boldsymbol{S}^{(i)}(z) = \mathbf{Q}^{(i)} \boldsymbol{D}^{(i)}(z) \boldsymbol{S}^{(i-1)}(z) \tilde{\boldsymbol{D}}^{(i)}(z) \tilde{\mathbf{Q}}^{(i)}(z)$$
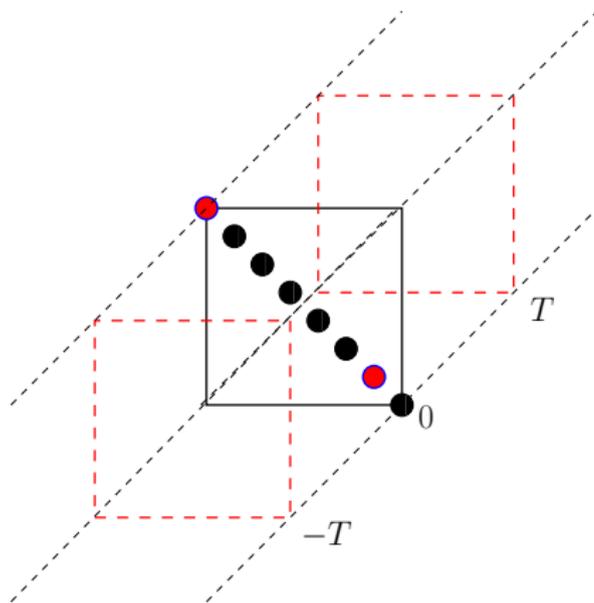
SBR2 Outcome

▶ At the ith 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

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

with

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

▶ diagonalisation of the previous 3×3 polynomial matrix ...

SBR2 Example — Diagonalisation

SBR2 Example — Spectral Majorisation

University of
Strathclyde
Engineering
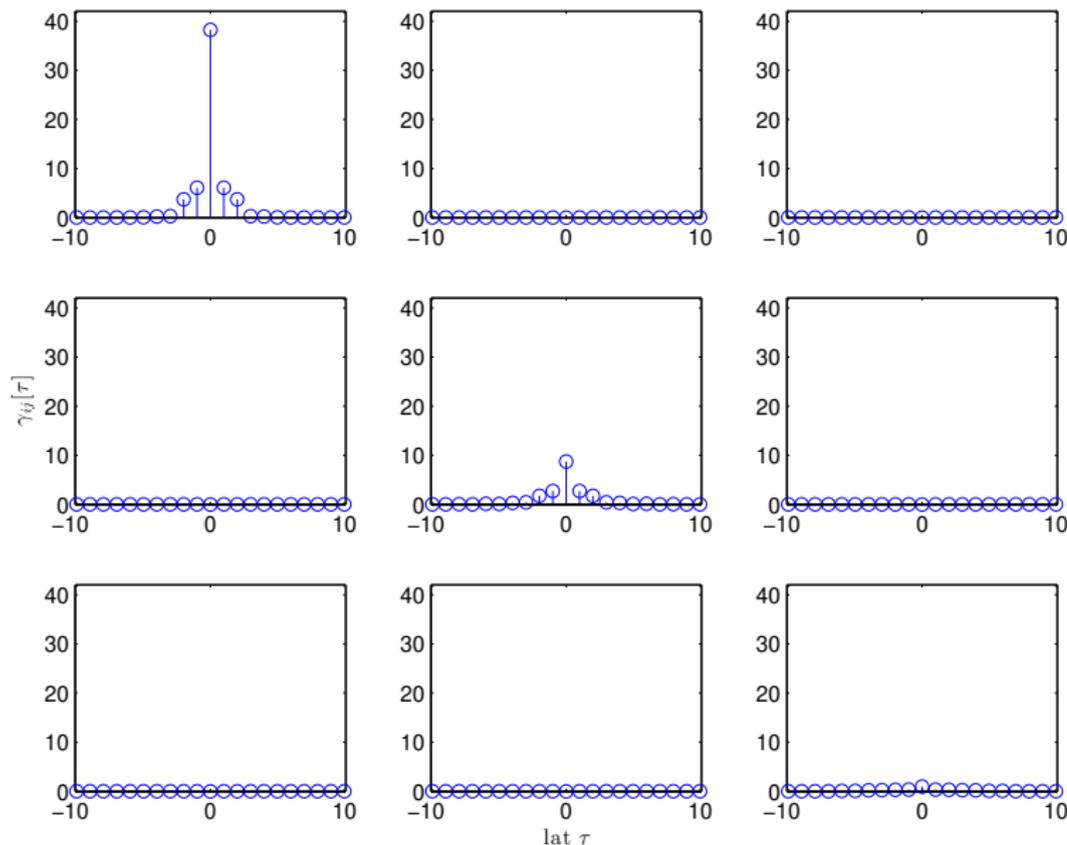
▶ The on-diagonal elements are spectrally majorised
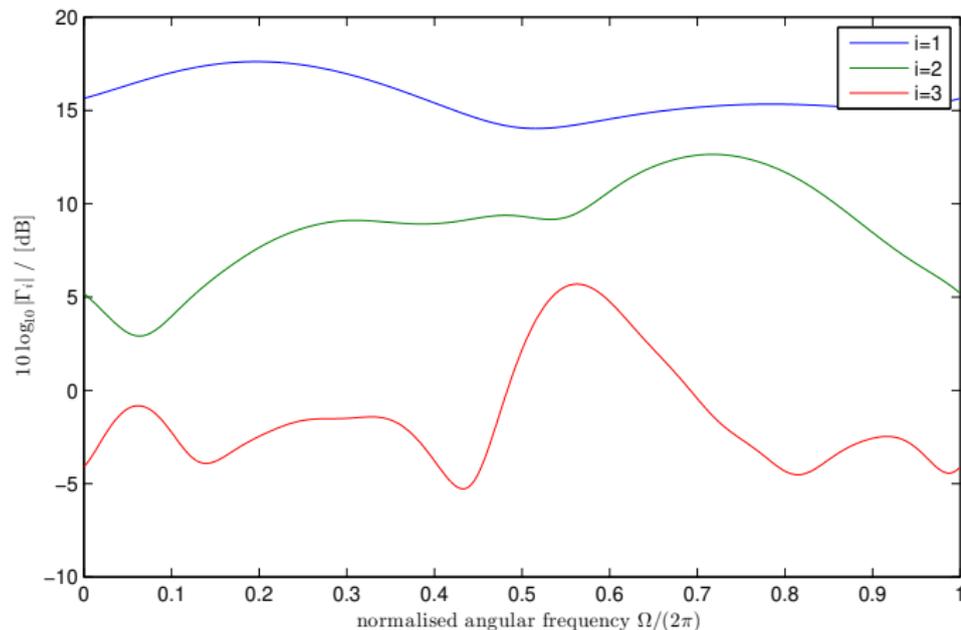
SBR2 — Givens Rotation

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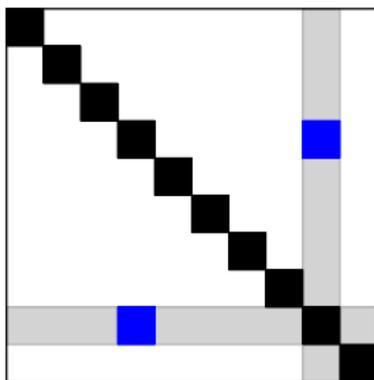
- ▶ note II: a Givens rotation only affects two columns and two rows in every matrix;
- ▶ Givens rotation is relatively low in computational cost!

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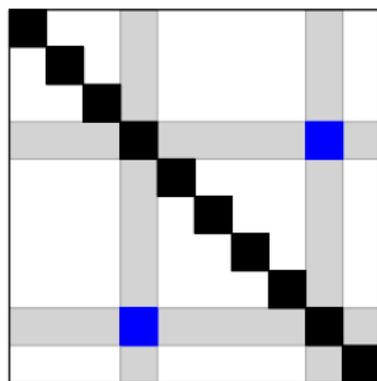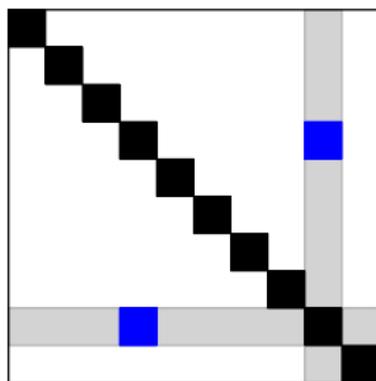
- ▶ note II: a Givens rotation only affects two columns and two rows in every matrix;
- ▶ Givens rotation is relatively low in computational cost!

Sequential Matrix Diagonalisation (SMD)

University of
Strathclyde
Engineering

[Redif *et al.*, IEEE Trans SP 2015]

- ▶ 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 ith step as in SBR2, the maximum element (or column with max. norm) is shifted to the lag-zero matrix:
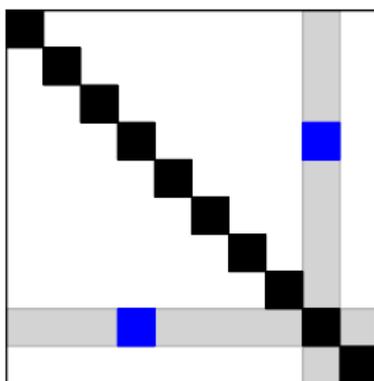
- ▶ an EVD is used to re-diagonalise the zero-lag matrix;
- ▶ a full modal matrix is applied at all lags — more costly than SBR2.

Sequential Matrix Diagonalisation (SMD)

[Redif *et al.*, IEEE Trans SP 2015]

University of
Strathclyde
Engineering

- ▶ 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 $\longrightarrow S^{(0)}$;
- ▶ at the ith step as in SBR2, the maximum element (or column with max. norm) is shifted to the lag-zero matrix:

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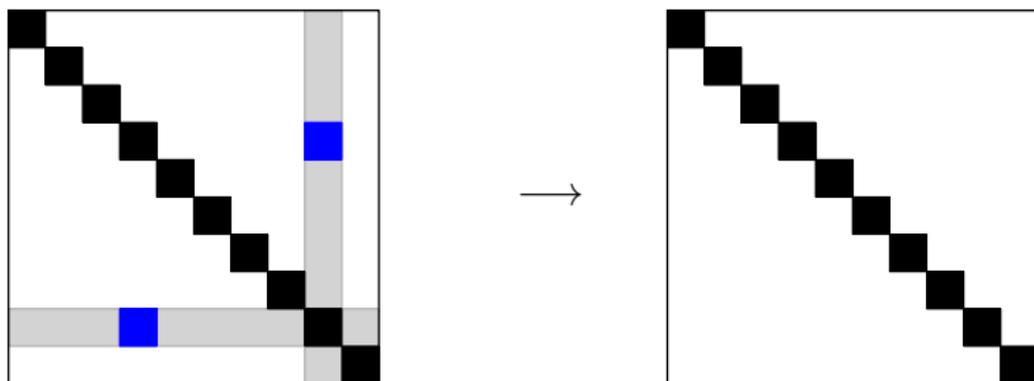
- ▶ an EVD is used to re-diagonalise the zero-lag matrix;
- ▶ a full modal matrix is applied at all lags — more costly than SBR2.

Multiple Shift SMD (SMD)

University of
Strathclyde
Engineering

▶ 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

$\longrightarrow$

- 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

University of
Strathclyde
Engineering

- ▶ Measuring the remaining normalised off-diagonal energy over an ensemble of space-time covariance matrices:
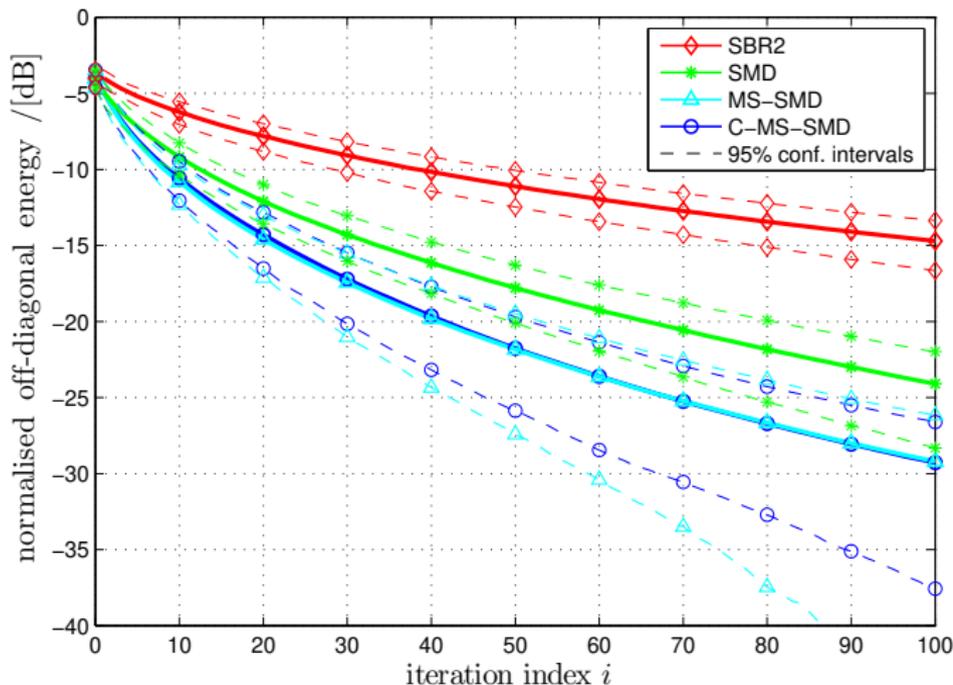
SBR2/SMD/MS-SMD Application Cost 1

▶ Ensemble average of remaining off-diagonal energy vs. order of paraunitary filter banks to decompose 4x4x16 matrices:
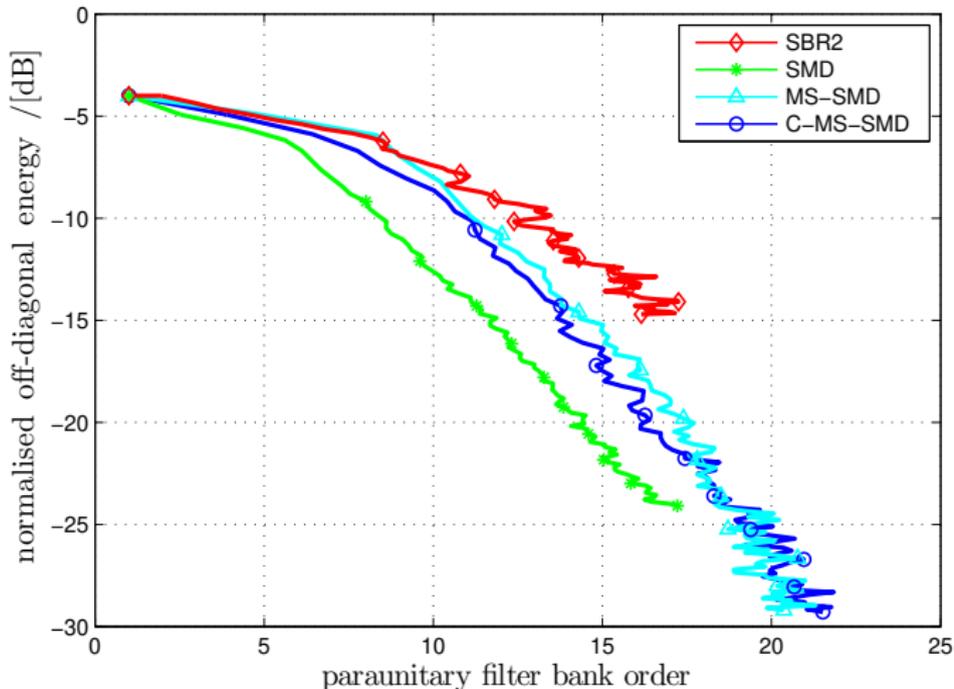
SBR2/SMD/MS-SMD Application Cost 2

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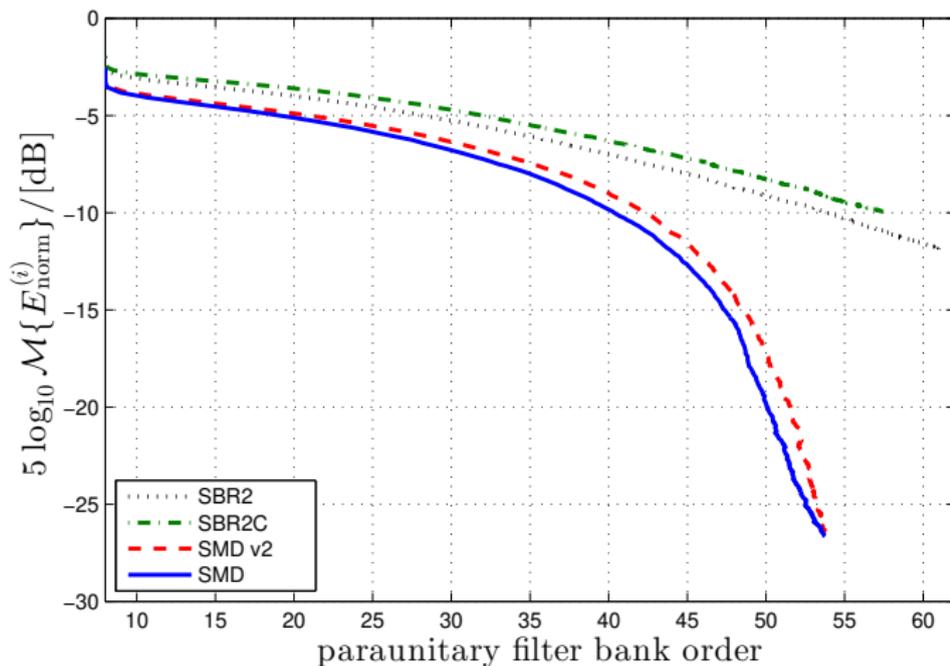
MATLAB Polynomial EVD Toolbox

▶ The MATLAB polynomial EVD toolbox can be downloaded from
`pevd-toolbox.eee.strath.ac.uk`
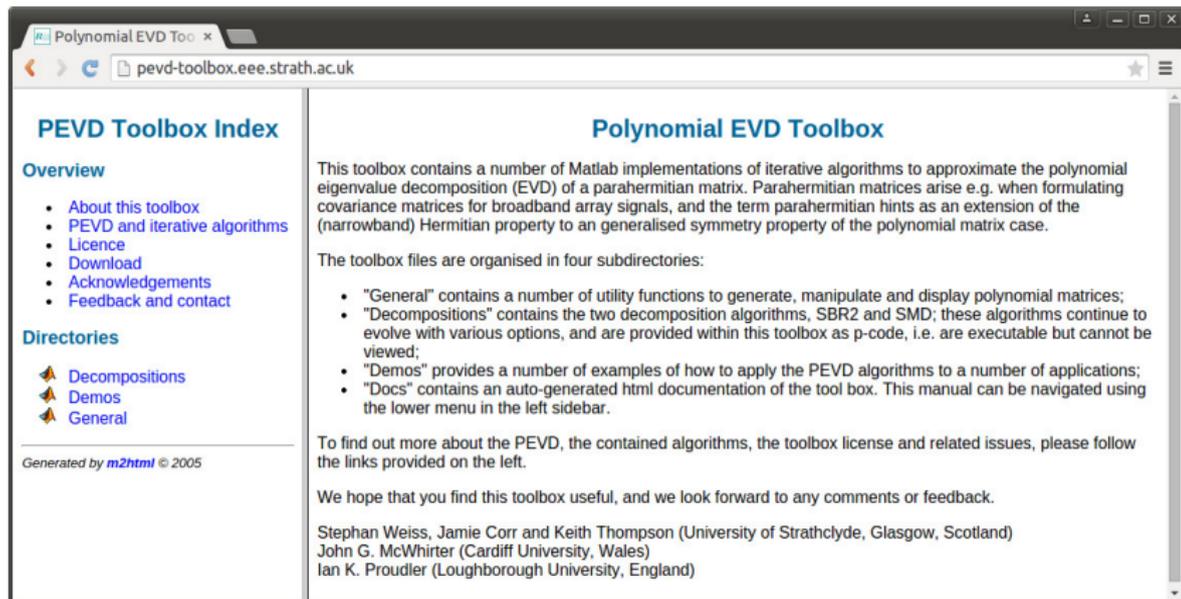
▶ the toolbox contains a number of iterative algorithms to calculate an approximate PEVD, related functions, and demos.

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?
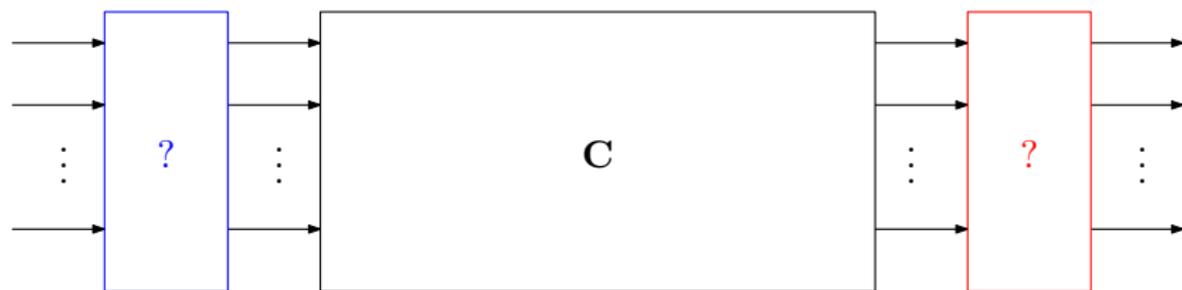
- overall system;

Narrowband MIMO Communications

University of
Strathclyde
Engineering

▶ a narrowband channel is characterised by a matrix $\mathbf{C}$ containing complex gain factors;

▶ problem: how to select the precoder and equaliser?
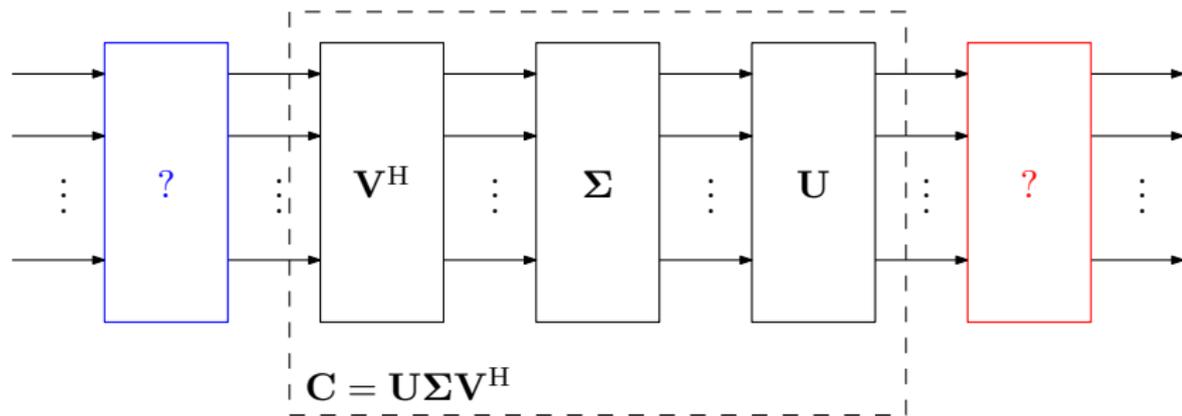
▶ the singular value decomposition (SVD) factorises $\mathbf{C}$ into two unitary matrices U and $\mathbf{V}^{\mathrm{H}}$ and a diagonal matrix $\mathbf{\Sigma}$;

Narrowband MIMO Communications

University of
Strathclyde
Engineering

▶ a narrowband channel is characterised by a matrix $\mathbf{C}$ containing complex gain factors;

▶ problem: how to select the precoder and equaliser?

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}$$

▶ 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 now a matrix of FIR filters; example for a 3×4 MIMO system $\mathbf{C}[n]$:

discrete time index n
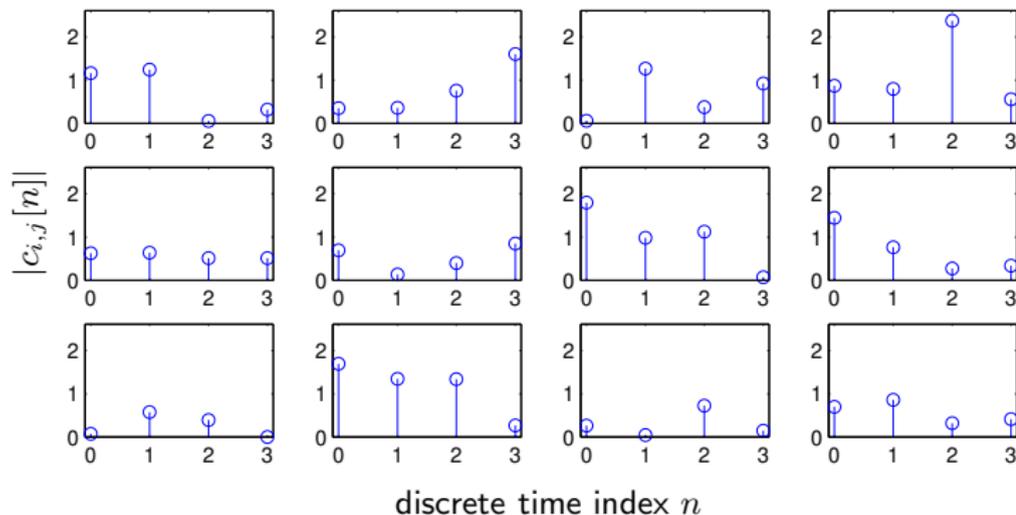
▶ the transfer function $C(z)$ •——○ $\mathbf{C}[n]$ is a polynomial matrix;

▶ an SVD can only diagonalise $\mathbf{C}[n]$ for one particular lag n.

Standard Broadband MIMO Approaches

University of
Strathclyde
Engineering

- ▶ OFDM (if approximate channel length is known):
  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):
  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.

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

Standard Broadband MIMO Approaches

- OFDM (if approximate channel length is known):
  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):
  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.

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

Standard Broadband MIMO Approaches

University of
Strathclyde
Engineering

- ▶ OFDM (if approximate channel length is known):
    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):
    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.

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

Standard Broadband MIMO Approaches

University of
Strathclyde
Engineering

- ▶ OFDM (if approximate channel length is known):
    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):
    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.

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

Standard Broadband MIMO Approaches

University of
Strathclyde
Engineering

- ▶ OFDM (if approximate channel length is known):
    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):
    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.

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

Polynomial Singular Value Decompositions

University of
Strathclyde
Engineering

- ▶ 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:

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

finally:

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

MIMO Application Example

University of
Strathclyde
Engineering

▶ Polynomial SVD of the previous $C(z) \in \mathbb{C}^{3 \times 4}$ channel matrix:

discrete time index n
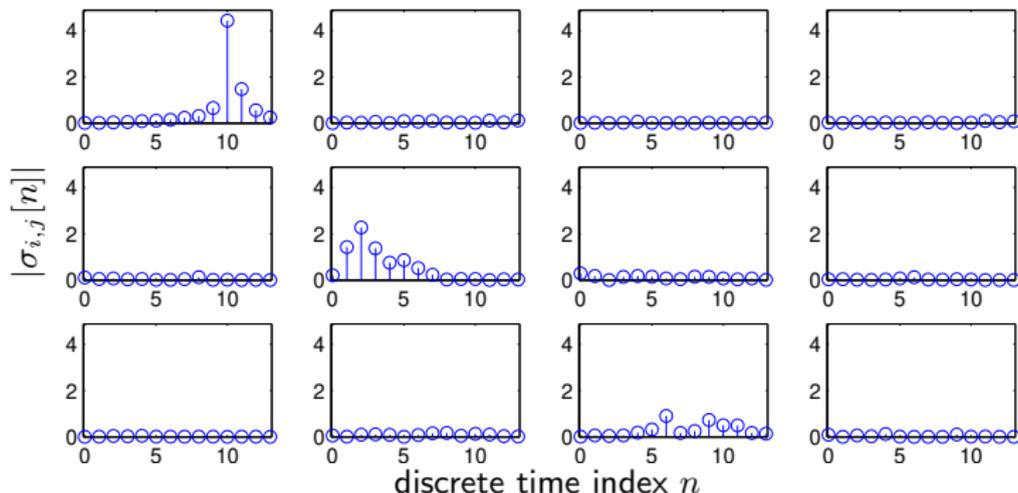
▶ the singular value spectra are majorised:
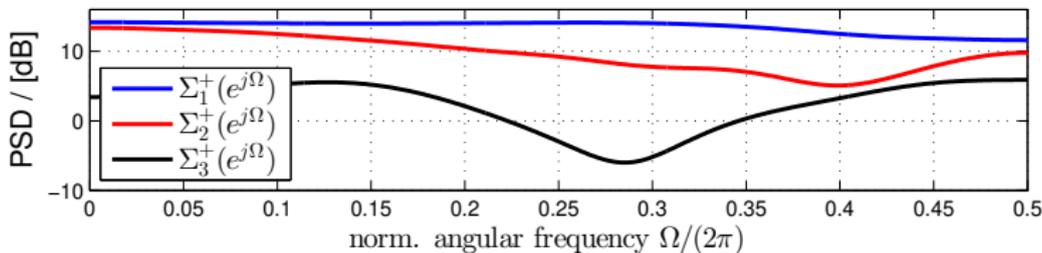
norm. angular frequency $\Omega/(2\pi)$

Narrowband Source Model

▶ Scenario with sensor array and far-field sources:
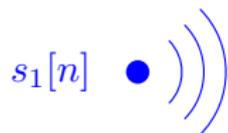
▶ for the narrowband case, the source signals arrive with delays, expressed by phase shifts in a steering vector

▶ 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 $\mathbf{s}_1$, $\mathbf{s}_2$, ... $\mathbf{s}_R$;

▶ 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

University of
Strathclyde
Engineering

▶ 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 Ω_i, the
time delays τ_m in the broadband steering vector $\mathbf{a}_\vartheta(z)$ collapse to
phase shifts in the narrowband steering vector $\mathbf{a}_{\vartheta,\Omega_i}$,

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

Data and Covariance Matrices

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

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}[n] & \mathbf{x}[n+1] & \dots & \mathbf{x}[n+L-1] \end{bmatrix}$$

▶ assuming that all $x_m[n]$, $m = 1, 2, \dots M$ 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

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

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

▶ unitary matrices $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_M]$ and $\mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_L]$;

▶ diagonal $\mathbf{\Sigma}$ contains the real, positive semidefinite singular values of $\mathbf{X}$ in descending order:

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

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

$$\text{rank}\{\mathbf{X}\} = R$$

▶ only the first R singular values of $\mathbf{X}$ will be non-zero;

▶ in practice, noise often will ensure that $\text{rank}\{\mathbf{X}\} = M$, with
  $M - R$ trailing singular values that define the noise floor:
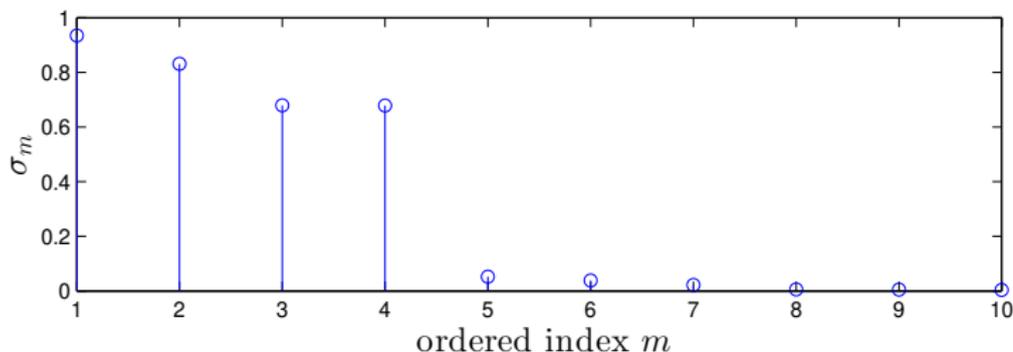
▶ therefore, by thresholding singular values, it is possible to estimate
  the number of linearly independent sources R.

Subspace Decomposition

University of
Strathclyde
Engineering

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

$$\mathbf{X} = [\mathbf{U}_s \ \mathbf{U}_n] \left[\begin{array}{cc} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_n \end{array} \right] \left[\begin{array}{c} \mathbf{V}_s^{\mathrm{H}} \\ \mathbf{V}_n^{\mathrm{H}} \end{array} \right]$$

- 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

University of
Strathclyde
Engineering

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

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

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

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

$$\mathbf{X}\mathbf{X}^{\mathrm{H}} = (\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}})(\mathbf{V}\boldsymbol{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}} \quad (38)$$
$$\mathbf{X}^{\mathrm{H}}\mathbf{X} = (\mathbf{V}\boldsymbol{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}})(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}) = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{H}} \quad (39)$$

▶ (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

University of
Strathclyde
Engineering

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

$$\mathbf{R} = [\mathbf{U}_s \ \mathbf{U}_n] \left[\begin{array}{cc} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n \end{array} \right] \left[\begin{array}{c} \mathbf{U}_s^{\mathrm{H}} \\ \mathbf{U}_n^{\mathrm{H}} \end{array} \right]$$

▶ 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

University of
Strathclyde
Engineering
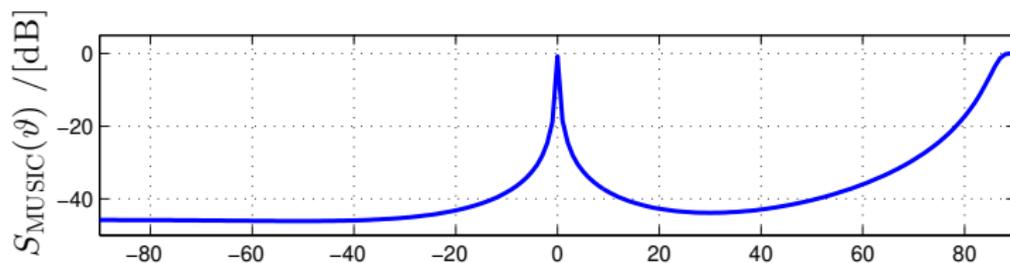
- ▶ Via SVD of the data matrix $\mathbf{X}$ or EVD of the covariance matrix $\mathbf{R}$, we can determine the number of linearly independent sources R;
- ▶ using the subspace decompositions offered by EVD/SVD, the directions of arrival can be estimated using e.g. MUSIC;
- ▶ based on knowledge of the 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
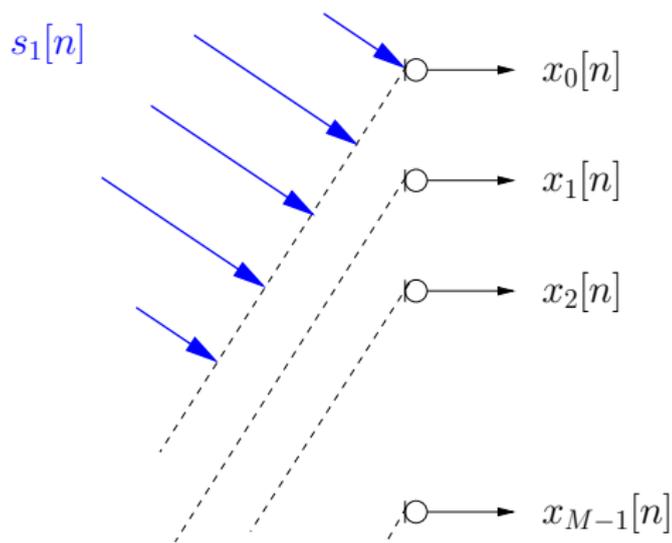
- ▶ Compared to the narrowband case, time delays rather than phase shifts bear information on the direction of a source.

Broadband Steering Vector

University of
Strathclyde
Engineering

▶ 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 Ω_i, the time delays τ_m in the broadband steering vector $\mathbf{a}_\vartheta(z)$ collapse to phase shifts in the narrowband steering vector $\mathbf{a}_{\vartheta,\Omega_i}$,

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

Space-Time Covariance Matrix

▶ If delays must be considered, the (space-time) covariance matrix must capture the lag τ:

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

▶ $\mathbf{R}[\tau]$ contains auto- and cross-correlation sequences:
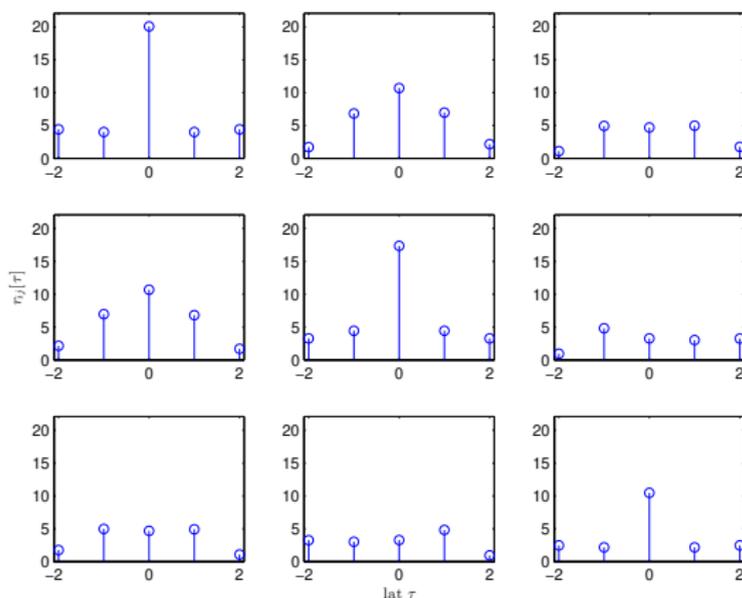
Cross Spectral Density Matrix

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

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

  with ϑ_l the direction of arrival and $S_l(z)$ the PSD of the lth source;

- $\boldsymbol{R}(z)$ is the cross spectral density (CSD) matrix;

- the instantaneous covariance matrix (no lag parameter τ)

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

Polynomial MUSIC (PMUSIC)

[Alrmah, Weiss, Lambotharan,*EUSIPCO* (2011)]

University of
Strathclyde
Engineering

- ▶ Based on the polynomial EVD of the broadband covariance matrix

$$\mathbf{R}(z) \approx \underbrace{[\mathbf{Q}_s(z) \ \mathbf{Q}_n(z)]}_{\mathbf{Q}(z)} \underbrace{\left[\begin{array}{cc} \mathbf{\Lambda}_s(z) & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n(z) \end{array}\right]}_{\mathbf{\Lambda}(z)} \left[\begin{array}{c} \mathbf{Q}_s^{\mathrm{P}}(z) \\ \mathbf{Q}_n^{\mathrm{P}}(z) \end{array}\right]$$

- ▶ paraunitary $\mathbf{Q}(z)$, s.t. $\mathbf{Q}(z)\mathbf{Q}^{\mathrm{P}}(z) = \mathbf{I}$;
- ▶ diagonalised and spectrally majorised $\mathbf{\Lambda}(z)$:
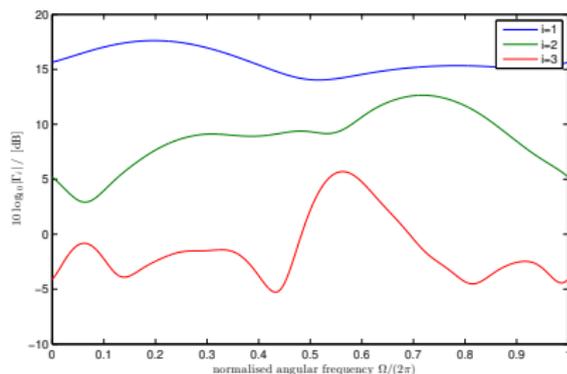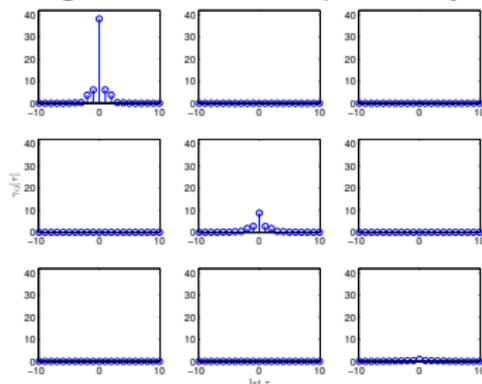
PMUSIC cont'd

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

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

- looking for minima leads to a spatio-spectral PMUSIC

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

- 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

University of
Strathclyde
Engineering

- ▶ 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

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

- ▶ 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

(a)

(b)
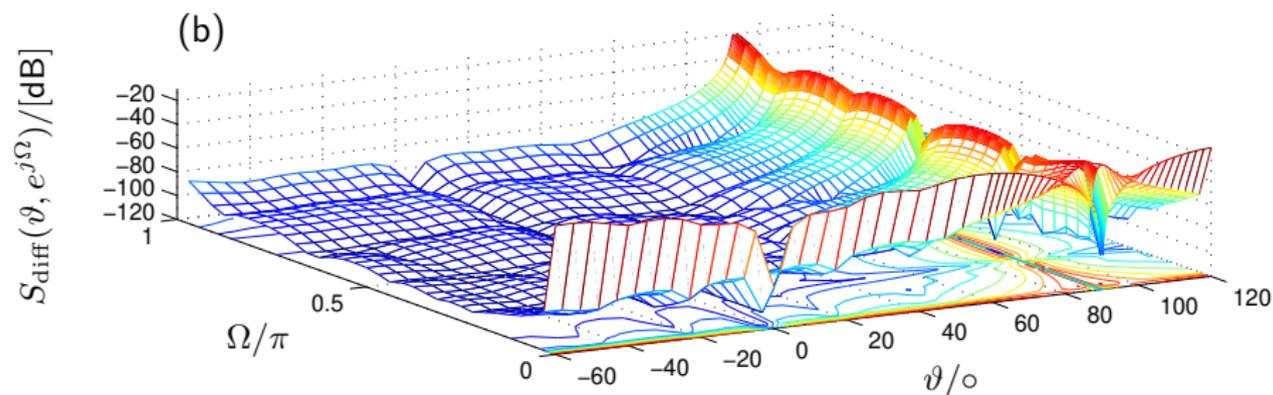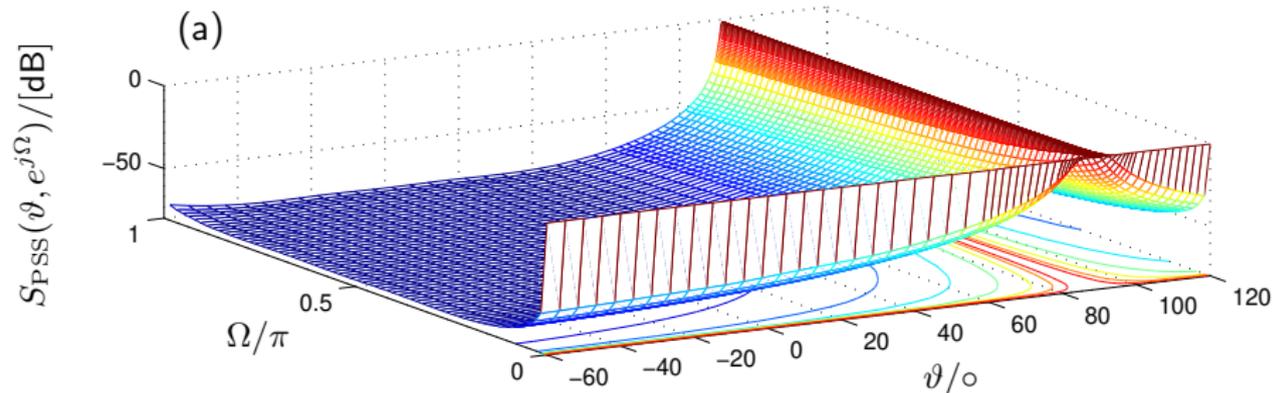
Simulation II

- $M = 8$ 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
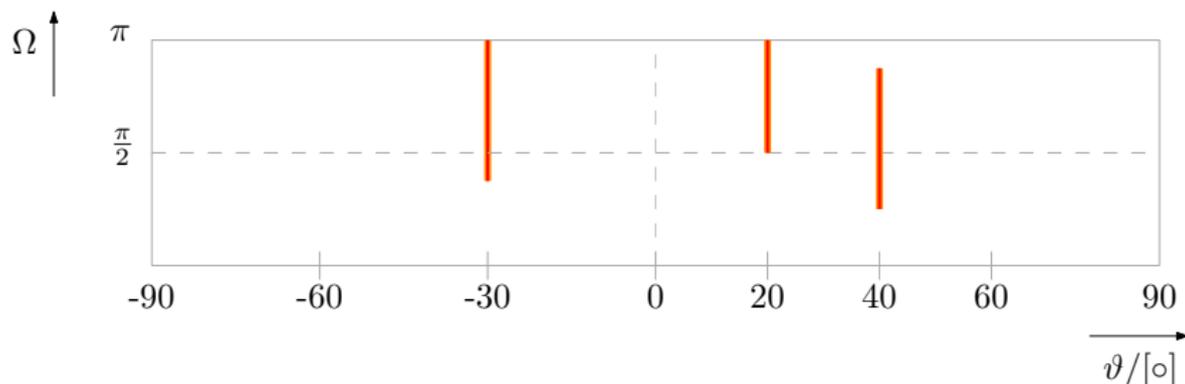
- filter banks as innovation filters, and broadband steering vectors to simulate AoA;
- space-time covariance matrix is estimated from 10^4 samples.

Simulation II — PSS-MUSIC
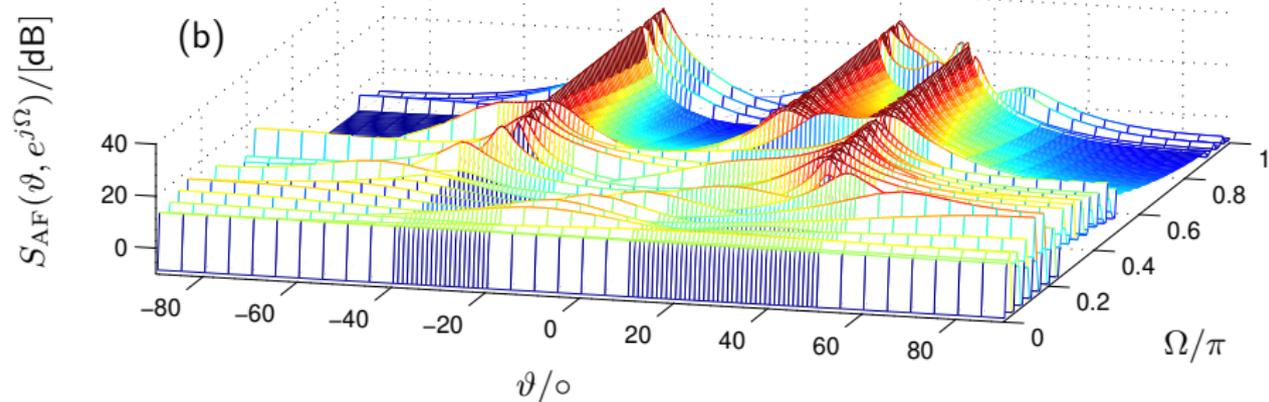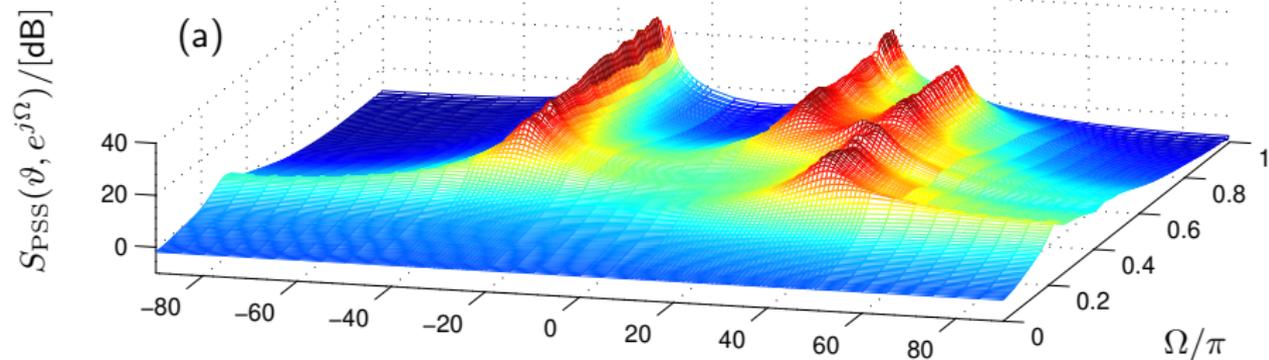
PS-MUSIC Comparison

▶ Simulation I (toy problem): peaks normalised to unity:
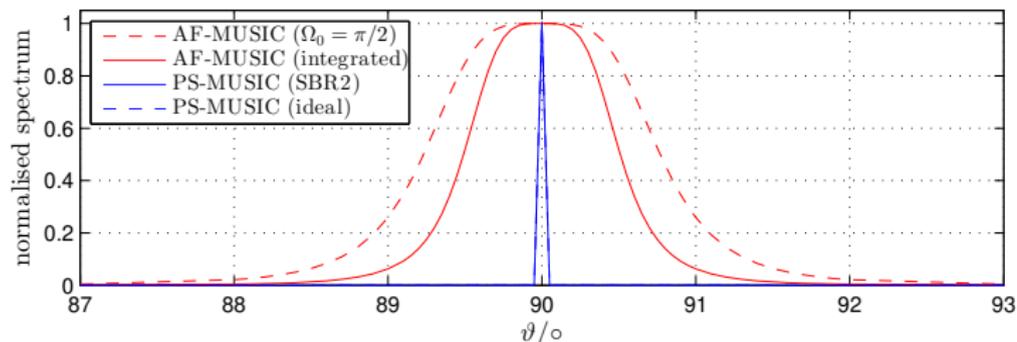
▶ Simulation II: inaccuracies on PEVD and broadband steering vector
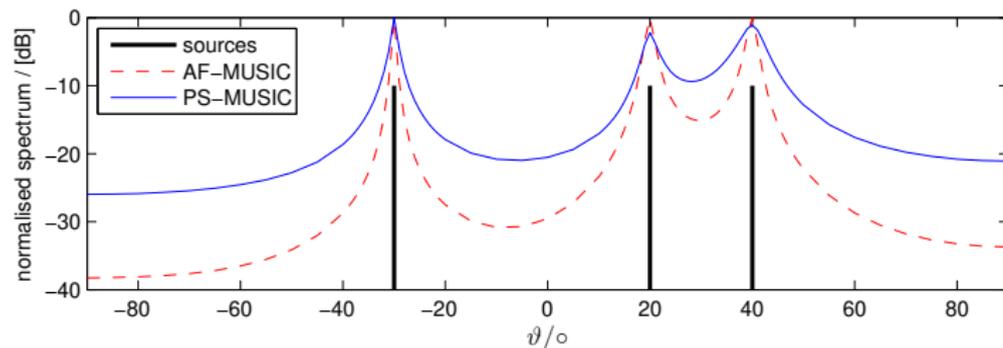
AoA Estimation — Conclusions

University of
Strathclyde
Engineering

- ▶ 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 Ω_s and angle of arrival ϑ_s, corrupted by interferers;

▶ a narrowband beamformer applies a single coefficient to every of the M sensor signals:
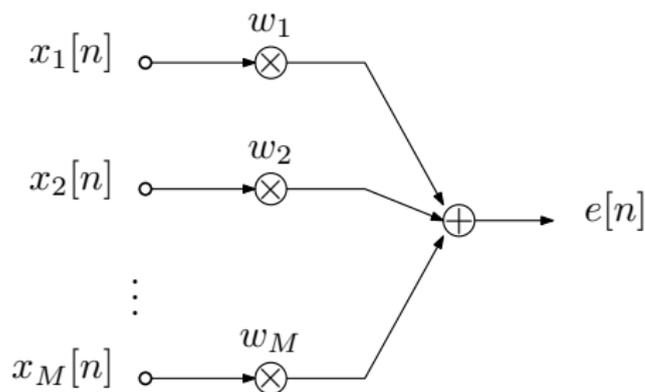
Narrowband MVDR Problem

University of
Strathclyde
Engineering

▶ Recall the space-time covariance matrix:

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

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

$$\min_{\mathbf{w}} \mathcal{E}\big\{|e[n]|^2\big\} = \min_{\mathbf{w}} \mathbf{w}^{\mathrm{H}}\mathbf{R}[0]\mathbf{w} \tag{40}$$

$$\text{s.t.} \quad \mathbf{a}^{\mathrm{H}}(\vartheta_{\mathrm{s}}, \Omega_{\mathrm{s}})\mathbf{w} = 1 \,, \tag{41}$$

▶ this is subject to protecting the signal of interest by a constraint in look direction ϑ_{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 $\mathbf{v} \in \mathbb{C}^{ML}$
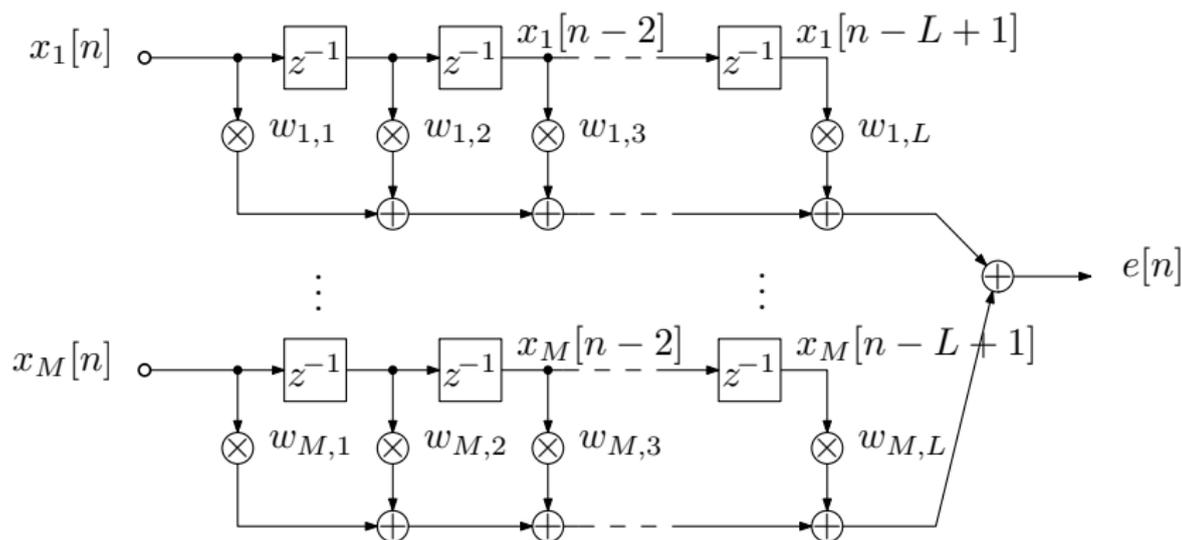
Broadband MVDR Beamformer

- ▶ 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.

$$\mathbf{C} = [\mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_0), \ \mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_1) \ \ldots \ \mathbf{s}(\vartheta_{\mathrm{s}}, \Omega_{L-1})] \qquad (42)$$

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

$$\mathbf{C}^{\mathrm{H}} \mathbf{v} = \mathbf{1} ; \qquad (43)$$

- ▶ 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 $[\mathbf{B}\mathbf{C}]$ 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:
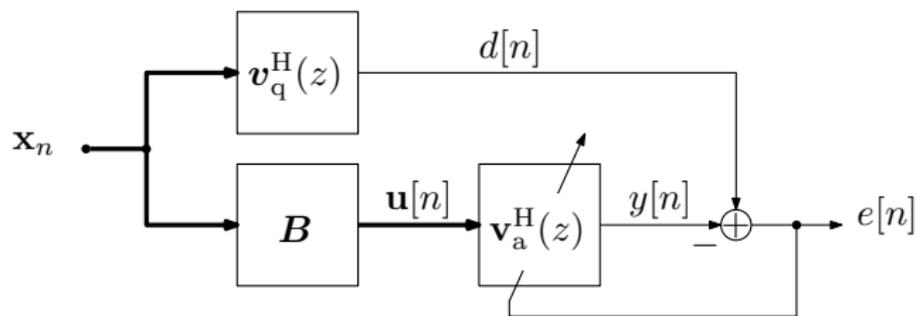
- ► note: all dimensions are determined by $\{M, L\}$.

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:

$$\min_{\mathbf{w}(z)} \oint_{|z|=1} R_e(z)\frac{dz}{z} \qquad (44)$$

$$\text{s.t.} \quad \tilde{\boldsymbol{a}}(\vartheta_\mathrm{s}, z)\boldsymbol{w}(z) = F(z) . \qquad (45)$$

- 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:
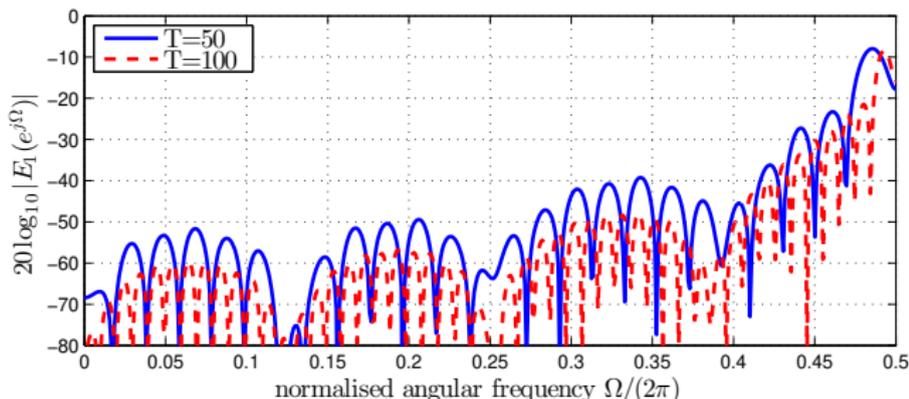
Generalised Sidelobe Canceller

▶ Instead of performing constrained optimisation, the GSC projects the data and performs adaptive noise cancellation:
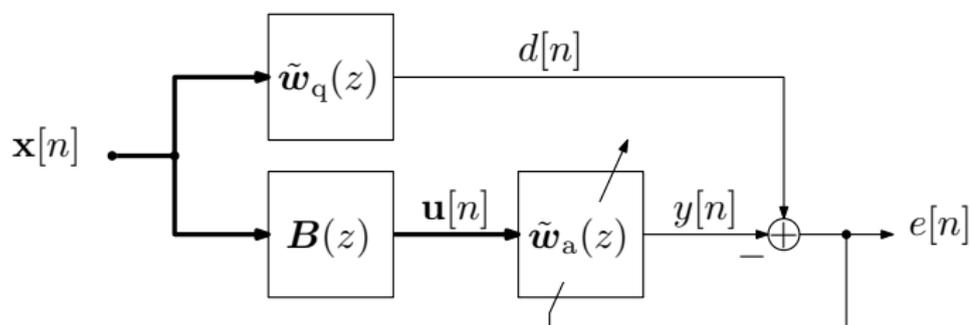
▶ the quiescent vector $\mathbf{w}_{\mathrm{q}}(z)$ is generated from the constraints and passes signal plus interference;

▶ the blocking matrix $\boldsymbol{B}(z)$ has to be orthonormal to $\mathbf{w}_{\mathrm{q}}(z)$ and only pass interference.

Design Considerations

University of
Strathclyde
Engineering

- ▶ The blocking matrix can be obtained by completing a paraunitary matrix from $\mathbf{w}_{\mathrm{q}}(z)$;
- ▶ this can be achieved by calculating a PEVD of the rank one matrix $\mathbf{w}_{\mathrm{q}}(z)\tilde{\mathbf{w}}_{\mathrm{q}}(z)$;
- ▶ 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

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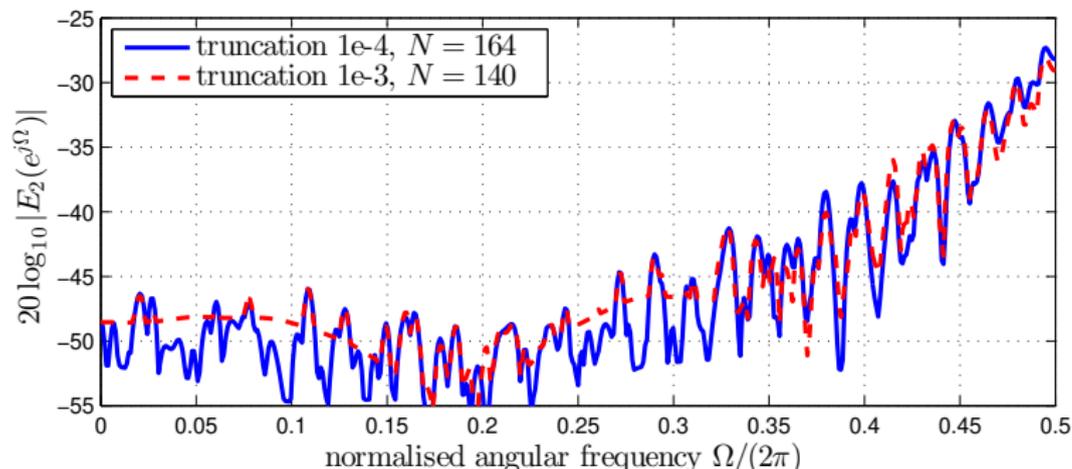
- With M sensors and a TDL length of L, the complexity of a standard beamformer is dominated by the blocking matrix;
- in the proposed design, $\mathbf{w}_{\mathrm{a}} \in \mathbb{C}^{M-1}$ has degree L;
- the quiescent vector $\mathbf{w}_{\mathrm{q}}(z) \in \mathbb{C}^{M}$ has degree T;
- the blocking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

	GSC cost	
component	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 $\vartheta = 30°$;
- three interferers with angles $\vartheta_i \in \{-40°, -10°, 80°\}$ active over the frequency range $\Omega = 2\pi \cdot [0.1; \ 0.45]$ at signal to interference ratio of -40 dB;
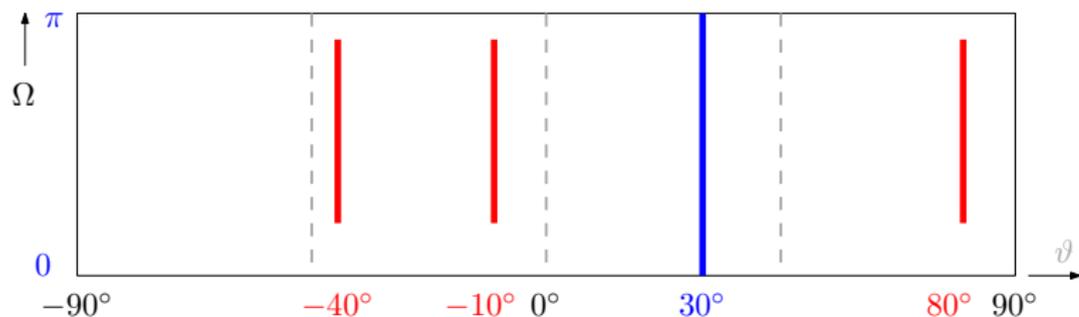
- $M = 8$ element linear uniform array is also corrupted by spatially and temporally white additive Gaussian noise at 20 dB SNR;
- parameters: $L = 175$, $T = 50$, and $N = 140$;
- cost per iteration: 10.7 kMACs (proposed) versus 1.72 MMACs (standard).

Quiescent Beamformer

▶ Directivity pattern of quiescent standard broadband beamformer:

Quiescent Beamformer

University of
Strathclyde
Engineering
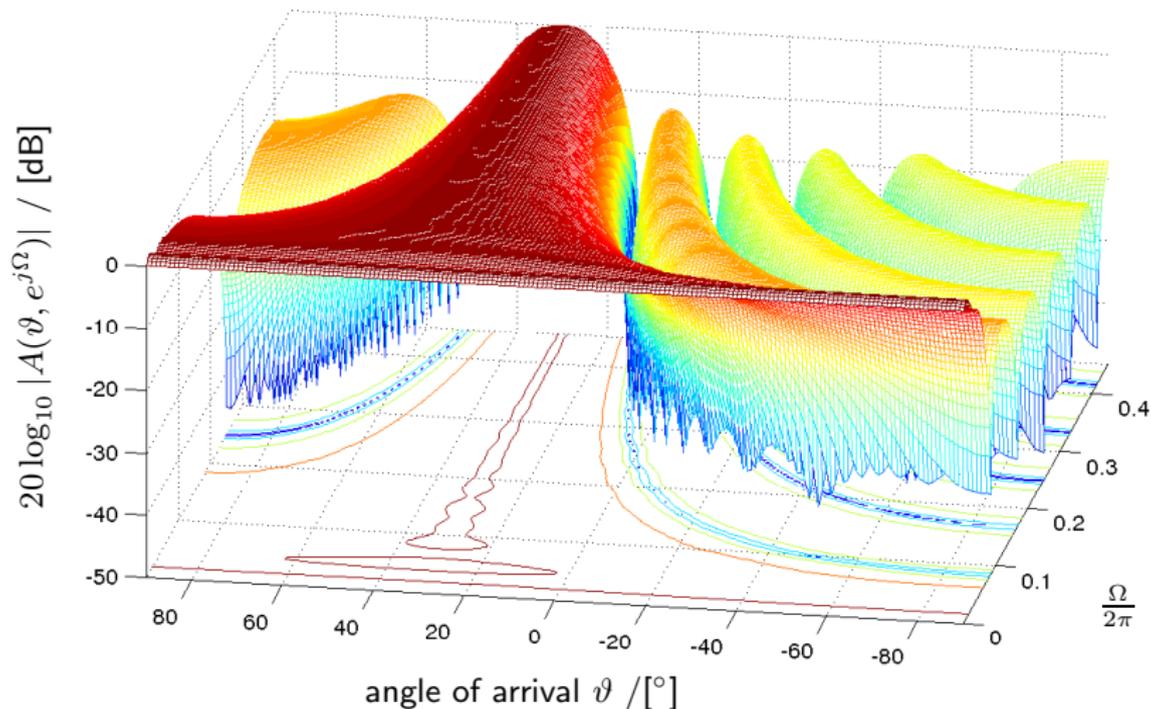
- ▶ Directivity pattern of quiescent proposed broadband beamformer:
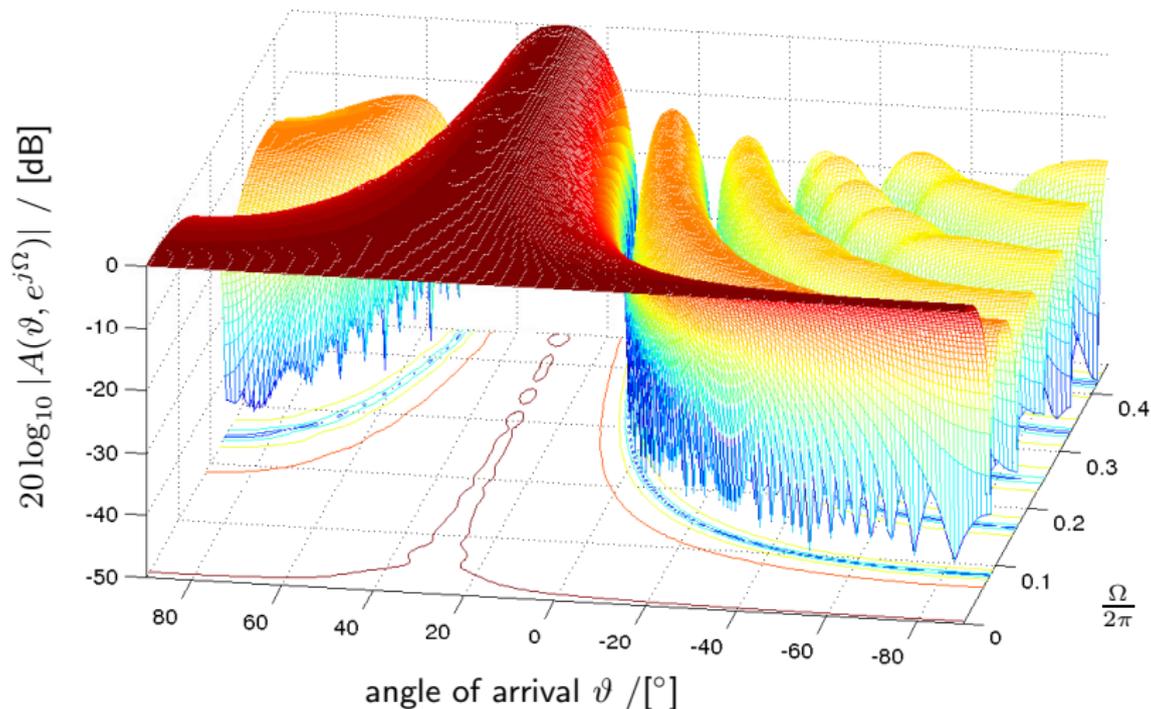
Adaptation

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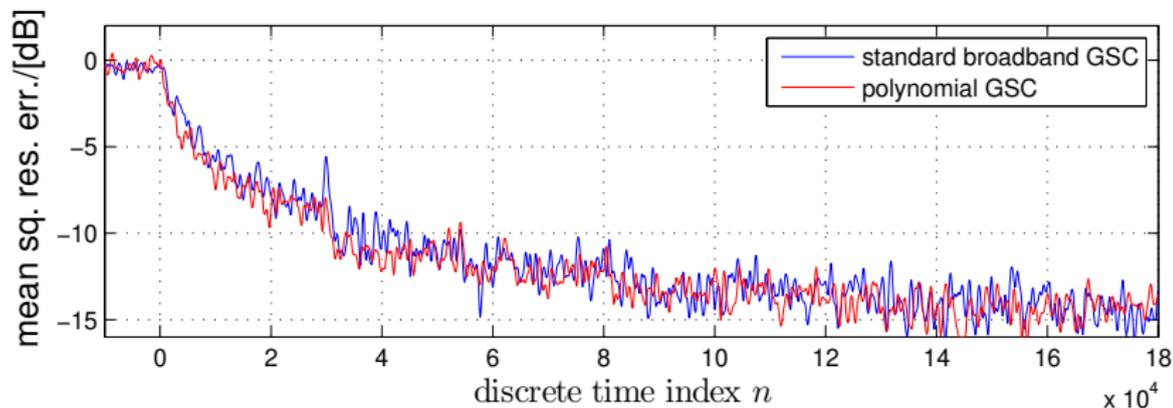
Adapted Beamformer

▶ Directivity pattern of adapted proposed broadband beamformer:
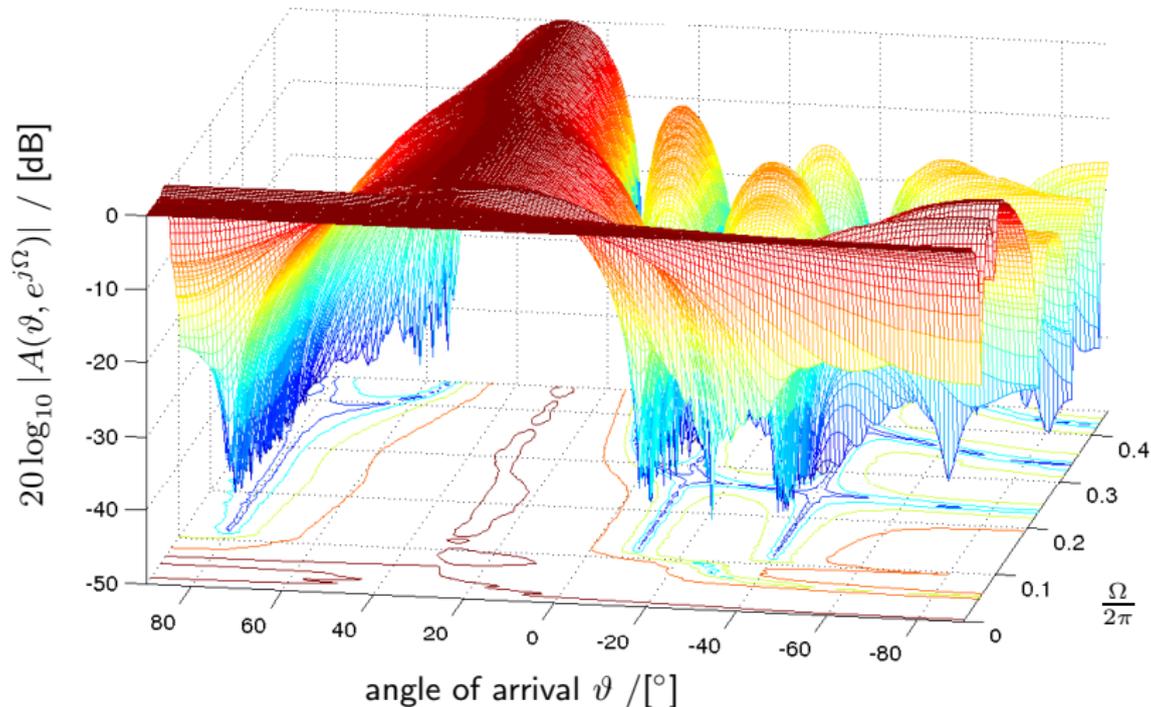
Adapted Beamformer

▶ Directivity pattern of adapted standard broadband beamformer:

Gain in Look Direction

University of
Strathclyde
Engineering
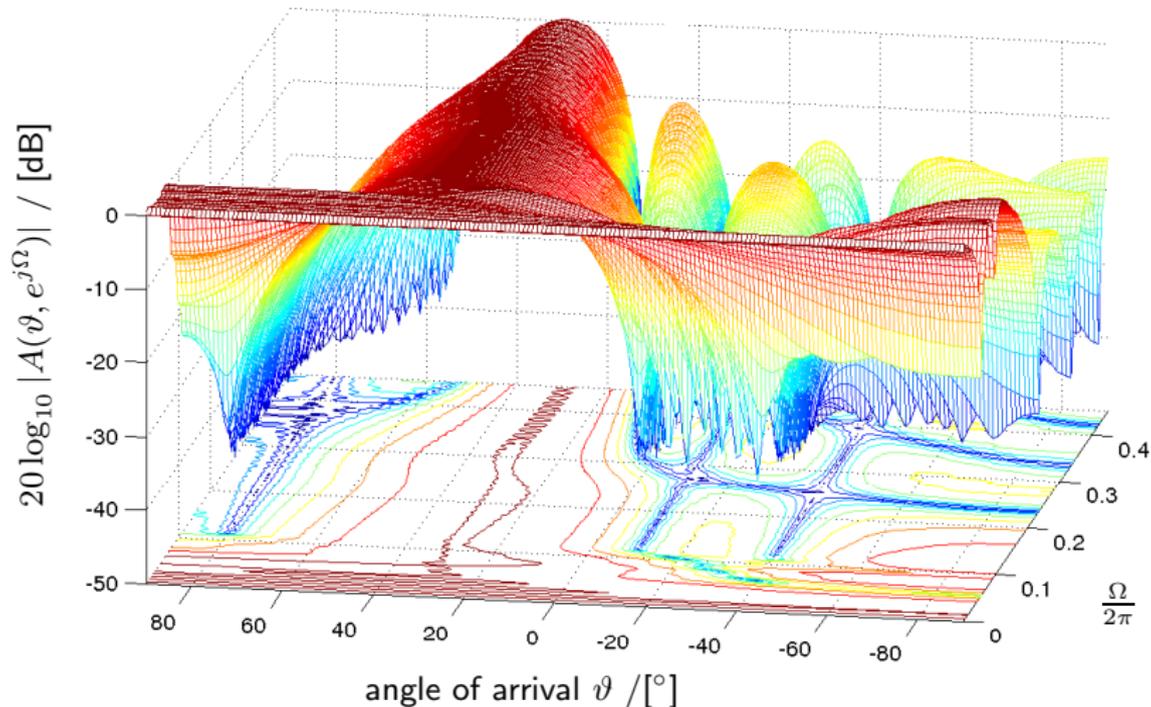
▶ Gain in look direction $\vartheta_s = 30°$ before and after adaptation:
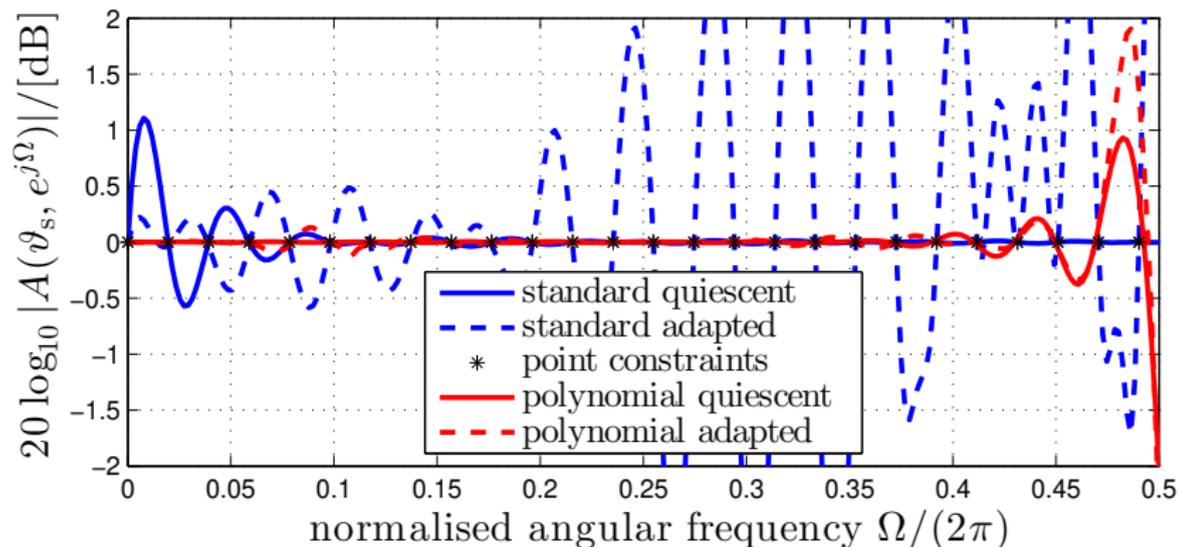
▶ due to signal leakage, the standard broadband beamformer after adaptation only maintains the point constraints but deviates elsewhere.

Broadband Beamforming Conclusions

University of
Strathclyde
Engineering

▶ 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-Sensor Transfer Functions

University of
Strathclyde
Engineering

▶ We take M-array measurements of a single source:
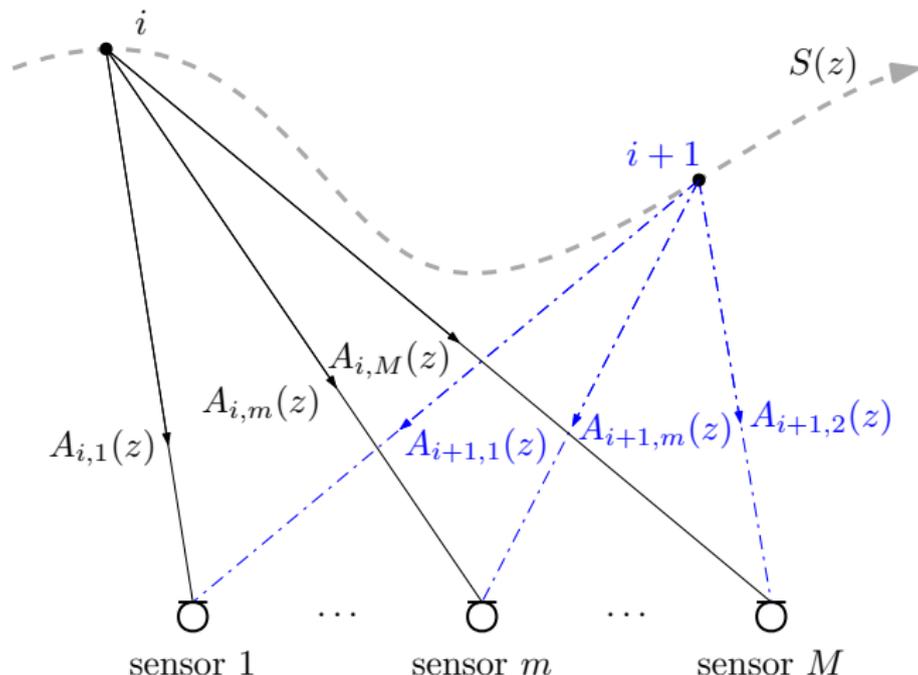
▶ 2nd order stats: $\boldsymbol{R}_i(z) = S(z)\boldsymbol{a}_i(z)\boldsymbol{a}_i^{\mathrm{P}}(z) = \gamma_i(z)\boldsymbol{u}_i(z)\boldsymbol{u}_i^{\mathrm{P}}$.

Transfer Functions and PEVD

University of
Strathclyde
Engineering

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

- 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:

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

- from a single measurement $\boldsymbol{R}_i(z)$, we cannot say anything about $\boldsymbol{a}_i(z)$ or $S(z)$.

Multiple Measurements

University of
Strathclyde
Engineering

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

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

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

$$\hat{S}(z) = \mathsf{GCD}\{\gamma_1(z) \; \ldots \; \gamma_I(z)\} \;;$$

▶ we can then extract the terms $A_{i,(+)}(z)$, and hence determine the vectors $\boldsymbol{a}_i(z)$ save of an arbitrary phase response due to the allpass $H_i(z)$:

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

Alternative DFT Domain Attempt

University of
Strathclyde
Engineering

- As an alternative, we take measurements in independent frequency bins $\Omega_k = \frac{2\pi k}{K}$:

$$\mathbf{R}_{i,k} = \boldsymbol{R}_i(e^{j\Omega_k}) = \boldsymbol{a}_i(e^{j\Omega_k})S(e^{j\Omega_k})\boldsymbol{a}_i^{\mathrm{H}}(e^{j\Omega_k})$$
$$= \mathbf{q}_{i,k}\lambda_{i,k}\mathbf{q}_{i,k}^{\mathrm{H}} .$$

- principal eigenvectors and eigenvalues for the measurement campaigns are

$$\mathbf{q}_{i,k} = \frac{\boldsymbol{a}_i(e^{j\Omega_k})}{|\boldsymbol{a}_i(e^{j\Omega_k})|} ,$$
$$\lambda_{i,k} = S(e^{j\Omega_k})|\boldsymbol{a}_i(e^{j\Omega_k})|^2 .$$

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

Numerical Example

University of
Strathclyde
Engineering

▶ Source with power spectral density

$$S(z) = \frac{1}{2}z + \frac{5}{4} + \frac{1}{2}z^{-1}$$

▶ vector of transfer functions during campaign $i = 1$:

$$\boldsymbol{a}_1(z) = \left[\begin{array}{ccc} 1 & + & \frac{1}{2}z^{-1} \\ \frac{3}{4} & - & \frac{1}{2}z^{-1} \end{array} \right]$$

▶ vector of transfer functions during campaign $i = 2$:

$$\boldsymbol{a}_2(z) = \left[\begin{array}{ccc} \frac{4}{5} & - & \frac{1}{2}z^{-1} \\ -\frac{1}{2} & + & z^{-1} \end{array} \right] ;$$

▶ based on these: PEVD computations for $\boldsymbol{R}_1(z)$ and $\boldsymbol{R}_2(z)$, and GCD calculation based on eigenvalues.

Numerical Results — Source PSD

▶ Eigenvalues / source PSD for both measurements $i = \{1, 2\}$:
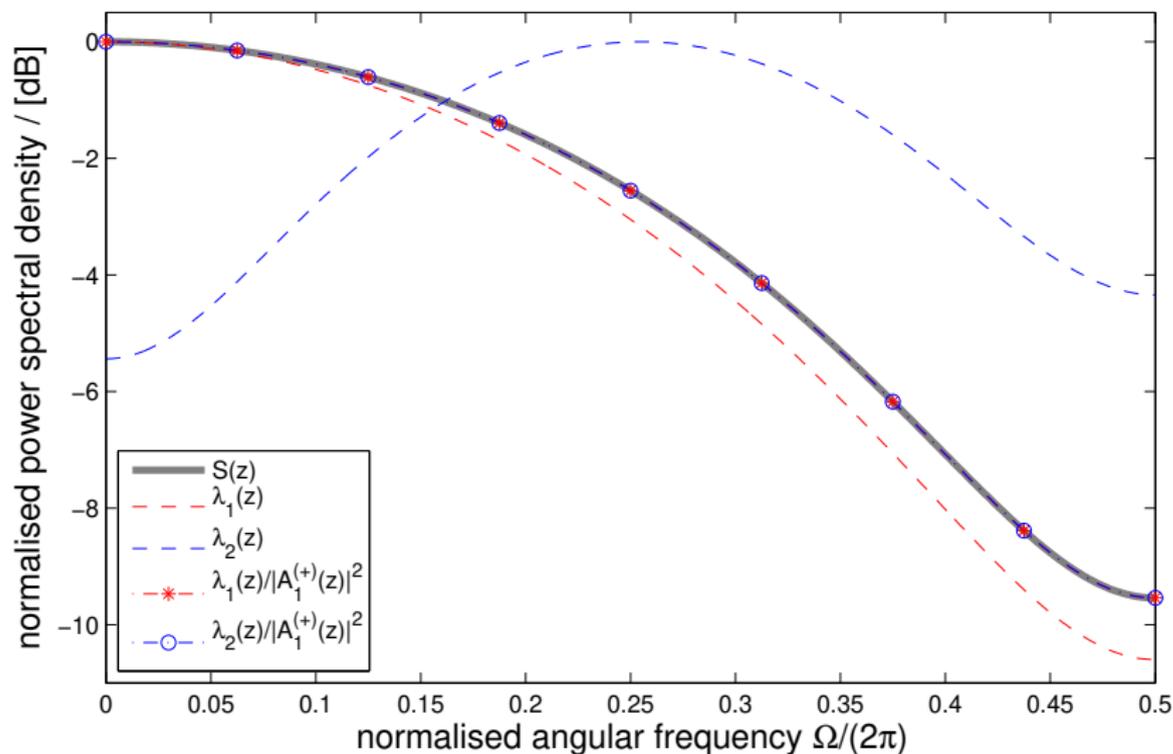
Numerical Result — Magnitude Responses I

University of
Strathclyde
Engineering

▶ Eigenvectors / magnitude response for measurement $i = \{1\}$:
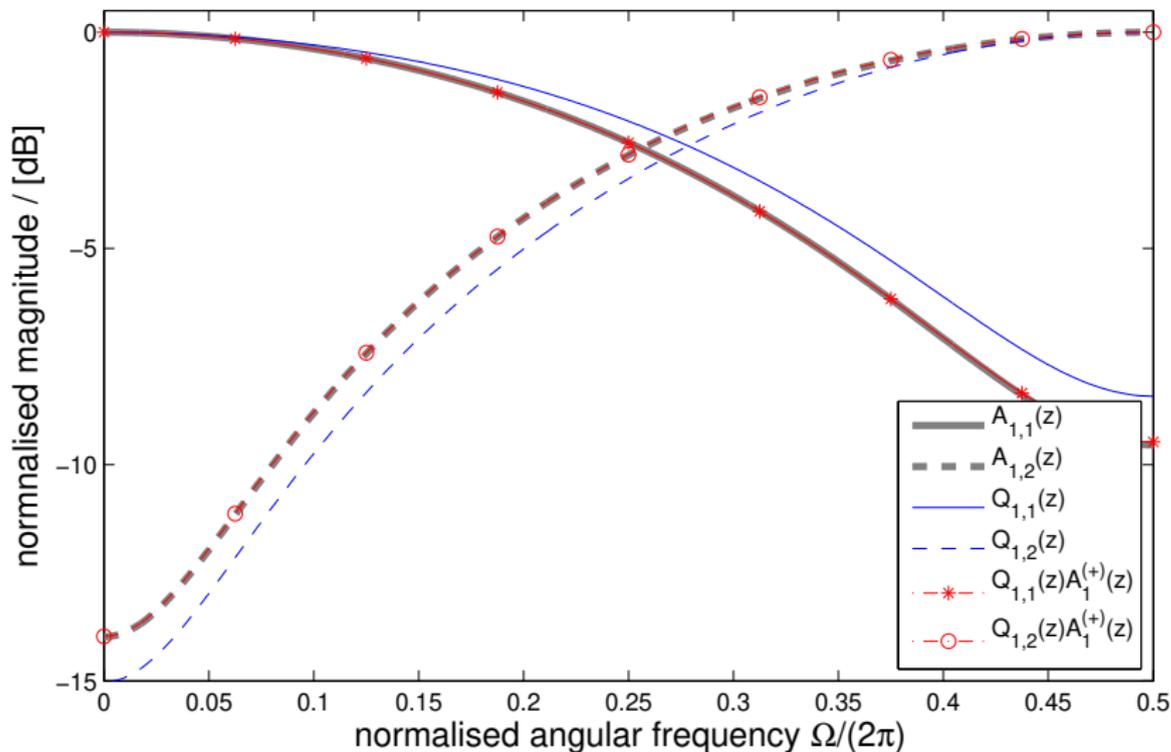
Numerical Result — Magnitude Responses II

▶ Eigenvectors / magnitude response for measurement $i = \{2\}$:
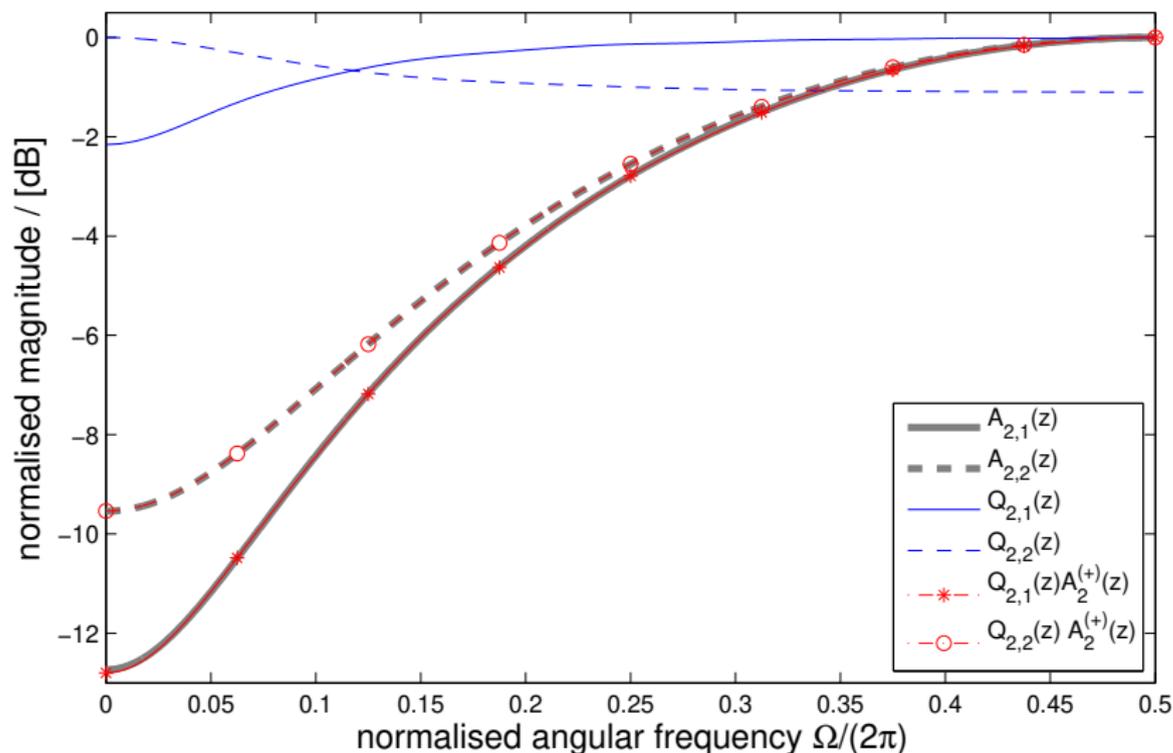
Application Summary and Critique

▶ 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: research since Euclid (300BC), with robust root-finding methods still on-going (Gröbner bases, algebraic geometry);

▶ nevertheless the approach gives a glimpse of the type of advantages that a coherent broadband approach can offer.

Additional Material

University of
Strathclyde
Engineering

- ▶ Key papers:
  1 J.G. McWhirter, P.D. Baxter, T. Cooper, S. Redif, and J. Foster: "An EVD Algorithm for Para-Hermitian Polynomial Matrices," IEEE Trans SP, **55**(5): 2158-2169, May 2007.
  2 S. Redif, J.G. McWhirter, and S. Weiss: "Design of FIR Paraunitary Filter Banks for Subband Coding Using a Polynomial Eigenvalue Decomposition," IEEE Trans SP, **59**(11): 5253-5264, Nov. 2011.
  3 S. Redif, S. Weiss, and J.G. McWhirter: "Sequential matrix diagonalisation algorithms for polynomial EVD of parahermitian matrices," IEEE Trans SP, **63**(1): 81–89, Jan. 2015.

- ▶ If interested in the discussed methods and algorithms, please download the free Matlab PEVD toolbox from

  pevd-toolbox.eee.strath.ac.uk

- ▶ for questions, please feel free to ask:
  - Stephan Weiss (stephan.weiss@strath.ac.uk) or
  - Fraser Coutts (fraser.coutts@strath.ac.uk)
  - Connor Deloasa (connor.delaosa@strath.ac.uk)