Polynomial Eigenvalue Decomposition for Multichannel Broadband Signal Processing

Vincent W. Neo, *Member, IEEE*, Soydan Redif, *Senior Member, IEEE*, John G. McWhirter, Jennifer Pestana, Ian K. Proudler, Stephan Weiss, *Senior Member, IEEE*, and Patrick A. Naylor, *Fellow, IEEE*

***Index Terms*—Polynomial matrix, polynomial eigenvalue decomposition, multichannel broadband processing, space-time covariance matrix, lossless filter b anks, b roadband beamforming, subband coding, speech enhancement**

I. Introduction

This article is devoted to the polynomial eigenvalue decomposition (PEVD) and its applications in broadband multichannel signal processing, motivated by the optimum solutions provided by the eigenvalue decomposition (EVD) for the narrowband case [1], [2]. In general, the successful techniques from narrowband problems can also be applied to broadband ones, leading to improved solutions. Multichannel broadband signals arise at the core of many essential commercial applications such as telecommunications, speech processing, healthcare monitoring, astronomy and seismic surveillance, and military technologies like radar, sonar and communications [3]. The success of these applications often depends on the performance of signal processing tasks, including data compression [4], source localization [5], channel coding [6], signal enhancement [7], beamforming [8], and source separation [9]. In most cases and for narrowband signals, performing an EVD is the key to the signal processing algorithm. Therefore, this paper aims to introduce PEVD as a novel mathematical technique suitable for many broadband signal processing applications.

A. Motivations and Significance

In many narrowband signal processing applications, such as beamforming [8], signal enhancement [7], subband coding [6] and source separation [9], the processing is performed based on the covariance matrix. The instantaneous spatial covariance matrix, computed using the outer product of the multichannel data vector, can capture the phase shifts between narrowband signals arriving at different sensors. In the narrowband case, diagonalization of the spatial covariance matrix often leads to optimum solutions. For example, the multiple signal classification (MUSIC) algorithm uses an EVD of the instantaneous spatial covariance matrix to perform super-resolution direction finding [5], [10].

The defining feature of a narrowband problem is the fact that a time delayed version of a signal can be approximated by the undelayed signal multiplied by a phase shift. The success of narrowband processing therefore depends on the accuracy of this approximation which varies from problem to problem. It is well known that as this approximation degrades, various issues start to occur when using narrowband algorithms. In array processing problems, this is often because some quantity in the algorithm that is related to direction of arrival starts to depend on the frequency of the signal. For example, in direction of arrival (DOA) algorithms, a wideband source can appear to be spatially distributed. Another issue is that of multipath. Reflections can cause problems as the different multipath signals are derived from a single source but arrive at the sensors at different times. This leads to various issues which can be advantageous or disadvantageous depending on one's point of view. In beamforming and direction of arrival estimation this causes a problem as the bearing to the source is clearly not well defined. However in signal recovery problems like speech enhancement or communication systems, multipath is advantageous as the signals can be combined to improve the signal-to-noise ratio. This is however only possible if the multipath signals are coherent. With narrowband processing, multipath signals appear to decorrelate as the delay increases. Multipath signals can also cause frequency dependent fading whereas narrowband processing can only deal with flat fading. Hence for some problems it is desirable to depart from narrowband processing and introduce some form of frequency dependent processing.

For the broadband case, one common approach is to divide each broadband signal into multiple narrowband signals. While these narrowband signals are often processed independently by well-established and optimal narrowband techniques that are typically based on the EVD, splitting the broadband signal into independent frequency bins neglects spectral coherence and thus ignores correlations between different discrete Fourier transform (DFT) bins [11], [12]. As a result, optimal narrowband solutions applied in independent DFT bins give rise to suboptimal approaches to the overall broadband problem [13]. Broadband optimal solutions in the DFT domain need to consider the cross-coupling between DFT bins via cross-terms, but the number of terms depends on the signal-to-

V. W. Neo and P. A. Naylor are with Imperial College London, UK (email: {vincent.neo09, p.naylor}@imperial.ac.uk). S. Redif is with American University of the Middle East, Kuwait (email: Soydan.Redif@aum.edu.kw). J. G. McWhirter is with Cardiff University, UK (email: mcwhirterjg@cardiff.ac.uk). J. Pestana, I. K. Proudler and S. Weiss are with University of Strathclyde, UK (email: {jennifer.pestana, ian.proudler, stephan.weiss}@strath.ac.uk).

The work of S. Weiss was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant no. EP/S000631/1 and the UK MOD University Defence Research Collaboration in Signal Processing. The work of P. A. Naylor was funded through UK EPSRC grant no. EP/S035842/1 and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 956369.

noise ratio (SNR) and cannot be determined in advance [14], [15]. Another approach uses tapped delay line (TDL) processing [16]–[18], but the performance depends on the filter length, which is challenging to determine in practice. These approaches highlight the lack of generic tools to solve broadband problems directly.

Polynomial matrices are widely used in control theory and signal processing. In the control domain, these matrices are used to describe multivariable transfer functions for multiple-input multiple-output (MIMO) systems [19]. Control systems are usually designed for continuous-time systems and are analyzed in the Laplace domain. There, factorizations, such as the Smith or Smith–McMillan decomposition, of matrices in the Laplace variable s, target unimodularity which is critical in the control context for invertibility, or spectral factorizations with minimum phase components in order to minimize time delays [20]. More recently, within digital signal processing (DSP), multirate DSP exploits polynomial matrices to describe lossless filter bank systems using polyphase notation [6], [20]. In multichannel broadband arrays or convolutively mixed signals, the array signals are generally correlated in time across different sensors. Therefore, the time delays for broadband signals cannot be represented by only phase shifts but need to be explicitly modelled. The relative time shifts are captured using the space-time covariance polynomial matrix, where decorrelation over a range of time shifts can be achieved using a PEVD [21].

While the initial work on the PEVD was a numerical algorithm [21], the existence of the decomposition of an analytic, positive semi-definite parahermitian matrix like the space-time covariance matrix has only recently been proved [22]–[24]. In most cases, unique parahermitian eigenvalues and paraunitary eigenvectors for a parahermitian matrix EVD exist but are of infinite length. However, being analytic, they permit good approximations by finite-length factors, which are still helpful in many practical applications such as beamformers [25], MIMO communications [26], source coding [27], signal enhancement [28], [29], source separation [30], source identification [31] and DOA estimation [32].

B. Outline of the Article

This article is organized as follows. Section II provides a primer on the relevant mathematical concepts. Section III introduces the notations and gives a background on multichannel array processing, including the use of spatial and space-time covariance matrices and the inadequacies of two common approaches. Section IV first introduces the PEVD, whose analytic eigenvalues and eigenvectors are first described before their approximations by numerical algorithms are presented. Section V will demonstrate the use of PEVD for some multichannel broadband applications, namely adaptive beamforming, subband coding, and speech enhancement. Concluding remarks and future perspectives are provided in Section VI.

II. Mathematical Background

Analytic Functions. In the time-domain, the key to describing the propagation of a broadband signal through a linear time-invariant system (LTI) is the difference equation, where the system output $y[n]$ depends on a weighted average of the input $x[n]$ and past values of both $y[n]$ and $x[n]$. This difference equation

$$y[n] = \sum_{\nu \geq 0} b[\nu] x[n-\nu] + \sum_{\mu > 0} a[\mu] y[n-\mu] \tag{1}$$

is straightforward to implement but does not lend itself to simple algebraic manipulations. For example, the difference equation for the concatenation of two LTI systems is not easily expressed in terms of the difference equations for the two component systems. For this reason, the z-transform $x(z) = \sum_n x[n] z^{-n}$ with $z \in \mathbb{C}$, or for short $x(z) \bullet\!\!-\!\!\circ\, x[n]$, can be used to turn the time-domain convolution into the multiplicative expression $y(z) = \mathit{h}(z) \cdot x(z)$, which is easy to manipulate [33], [34].

The z-transform exists as long as the time-domain quantities are absolutely summable, i.e. for $x(z)$ we require $\sum_n |x[n]| < \infty$. Values of z for which the z-transform is finite define the region of convergence (RoC), which therefore must include at least the unit circle since $\sum_n x[n] \mathrm{e}^{-\mathrm{j}\Omega n} \leq \sum_n |x[n]|$. For values of z within this RoC, the function $x(z)$ is complex analytic, which has profound consequences. Analytic functions mathematically belong to a *ring*, such that any addition, subtraction, and multiplication will produce an analytic result. These operations potentially reduce the RoC. Dividing by an analytic function also results in an analytic function, as long as the divisor does not have spectral zeros; again this operation may shrink the RoC. For example with $\mathit{b}(z)$ and $\mathit{a}(z)$ analytic and the latter without any zeros on the unit circle then $\mathit{h}(z) = \mathit{b}(z)/\mathit{a}(z)$ is also guaranteed to be analytic. Note that the same cannot be said for non-analytic functions. This is important since non-analytic functions can be difficult to approximate optimally in practice – see below. (For more on the algebra of analytic functions see the "Algebra of Functions" siderbar.)

Laurent Series, Power Series, and Polynomials. Throughout this article, we often represent z-transforms by series, i.e., by expressions of the form

$$\mathit{h}(z) = \sum_{n=N_1}^{N_2} h[n] z^{-n} \,. \tag{2}$$

This is motivated by the fact that analytic functions can be represented by a Taylor (or, equivalently, power) series within the region of convergence. More generally, we are interested in Laurent series, power series, Laurent polynomials and polynomials, which we distinguish below.

For finite N_1 and N_2 in (2), $\mathit{h}(z)$ is a *Laurent polynomial* if N_1 and N_2 have opposing signs. If N_1 and N_2 share the same sign, i.e. if $\mathit{h}(z)$ is purely an expression in powers of either z^{-1} or z, it is a *polynomial*. Typically, by a polynomial, we refer to an expression that contains powers in z^{-1}. If interpreted as a transfer function, a polynomial $\mathit{h}(z)$ in z^{-1} refers to a causal finite impulse response filter. If it possesses finite coefficients, then a polynomial or Laurent polynomial $\mathit{h}(z)$ will always be absolutely summable and hence be analytic.

A *Laurent series* is characterized by $N_1 \to -\infty$ and $N_2 \to \infty$, while for a *power series*, $\mathit{h}(z)$ only contains powers in

strictly either z^{-1} (for $N_1 \geq 0$ and $N_2 \to \infty$) or z (for $N_1 \to -\infty$ and $N_2 \leq 0$). Both Laurent and power series possess a generally infinite coefficient sequence $\{h[n]\}$. Such sequences can be used to represent rational functions, where $\hbar(z) = b(z)/a(z)$ is a ratio of two polynomials; with respect to (1), such a power series can describe an infinite impulse response filter. Further and more generally, Laurent and power series can also represent transcendental functions, which are absolutely convergent but may not be representable by a finite number of algebraic operations such as a ratio.

In signal processing, polynomials and convergent power series can represent quantities such as finite or infinite impulse responses of either causal or anti-causal stable systems. In contrast, Laurent series or Laurent polynomials appear as a result of correlation operations. First, assume that a zero-mean unit-variance uncorrelated random signal $x[n]$ excites a system with impulse response $h[n]$. With the input autocorrelation sequence $r_x[\tau] = \mathbb{E}\{x[n]x^*[n-\tau]\} = \delta[\tau]$, the output autocorrelation is $r_y[\tau] = \sum_n h^*[-n]h[\tau - n]$ [35], where $\mathbb{E}\{\cdot\}$ and $[\cdot]^*$ are the expectation and complex conjugate operators, respectively. Then its z-transform, the power spectral density $\mathfrak{r}_y(z) \bullet\!\!-\!\!\circ\, r_y[\tau]$, $\mathfrak{r}_y(z) = \hbar(z)\hbar^*(1/z^*)$, will be a Laurent series if $\hbar(z)$ is a power series, and a Laurent polynomial if $\hbar(z)$ is a polynomial.

Polynomial Approximation and Polynomial Arithmetic. By the Weierstrass Theorem, any continuous function can be arbitrarily well approximated by a polynomial of sufficient degree, but in general it can be nontrivial to construct the approximating polynomials. However, for analytic functions, such as Laurent or power series $\hbar(z)$, this approximation can be easily obtained by truncating $h[n] \circ\!\!-\!\!\bullet\, \hbar(z)$ to the required order. If the result is a Laurent polynomial as in (2), describing, e.g., the impulse response of a non-causal system, then a polynomial (or causal system) can be obtained by a delay by N_1 sampling periods. Thus, by delay and truncation, all of the above expressions describing analytic functions — Laurent series, power series, and Laurent polynomials — can be arbitrarily closely approximated by polynomials.

Key Statement

While operations on analytic functions tend to yield analytic functions, the same is not true for polynomials: e.g. the ratio of polynomials generally yields a rational function but not a polynomial. Nonetheless, since the resulting function is analytic, it can be approximated arbitrarily closely by a polynomial via appropriate delay and truncation operations.

Matrices of Analytic Functions and Polynomial Matrices. In this article, we consider matrices whose entries are analytic functions in general and their close approximation by polynomial matrices, in particular. The mathematical theory of polynomial matrices that depend on a real parameter has been studied in e.g. [36]. This has found application, for example, in the control domain [37]. Within signal processing, polynomial matrices have been used in filter bank theory. Specifically, polyphase notation [38] has been utilized to allow efficient implementation. Here, polynomial matrices in the form of polyphase analysis and synthesis matrices describe networks of filters operating on demultiplexed single-channel data. More generally, polynomial matrices have been used to define space-time covariance matrices on the demultiplexed data streams [20] or directly for multichannel data [21].

Algebra of Functions

We are interested in matrices whose entries are more general than complex numbers. Specifically, we are interested in entries that are analytic functions. Matrices whose entries are analytic functions rather than real or complex numbers, and the algebraic manipulation of these matrices may at first seem a little exotic, but many operations for real or complex numbers carry over to this setting.

There are several different classes of function depending on what properties they have. For example there are discontinuous functions, continuous but non-differentiable functions, functions that are continuous and differentiable up to a certain order, and functions that are continuous and differentiable for all orders. The class of analytic functions, by definition, have locally convergent power series. Consequently, they are infinitely differentiable and are easier to work with than other types of function. These series might have a finite number of terms but in general there are infinitely many. The truncation of these series results in polynomial approximations of the underlying analytic function.

Analytic functions can be algebraic or transcendental. An algebraic function $f(x)$ is a function that is a root of a polynomial equation. More specifically, f is algebraic if it satisfies $p(x, f(x)) = 0$ for some irreducible polynomial $p(x, y)$ with coefficients in some field. Examples of algebraic functions include rational functions and nth roots of polynomials. Note that the inverse function of an algebraic function (if it exists) is also algebraic. An analytic function that is not algebraic is called a transcendental function. Examples include e^x, $\sin(x)$ and $\cos(x)$. Such functions have power series representations with an infinite number of terms.

Let us first consider analytic functions on their own. A function $f(z)$ is (complex) analytic in a domain (an open set) $\mathcal{D} \subset \mathbb{C}$ if at each point, it can be written as a locally convergent Taylor series. (Note that this means that such a function is infinitely differentiable.) The set $\mathcal{D}$ is known as the domain of analyticity of $f(z)$. We note that two different analytic functions $f(z)$ and $g(z)$, may have different domains of analyticity, say $\mathcal{D}_f$ and $\mathcal{D}_g$. When we operate on these functions, we will assume that $\mathcal{D}_f$ and $\mathcal{D}_g$ overlap, i.e., that they have a nontrivial intersection, and will restrict $f(z)$ and $g(z)$ to this common domain $\mathcal{D} = \mathcal{D}_f \cap \mathcal{D}_g$.

Then, we can perform certain fundamental operations

on analytic functions, and the result will also be an analytic function with the same domain of analyticity $\mathcal{D}$. In particular, if $f(z)$ and $g(z)$ are analytic on a domain $\mathcal{D} \subset \mathbb{C}$ then $f(z) + g(z)$ is analytic, that is, it can be expressed as a locally convergent power series for any $z \in \mathcal{D}$. Similarly, $f(z) - g(z)$ and $f(z) \cdot g(z)$ are analytic. Things become a little more complicated when we consider quotients of the form $f(z)/g(z)$, but the result is analytic everywhere except at zeros of $g(z)$, as might be expected. This "closure" is important since it means that as we manipulate analytic functions we do not need to worry if the result is also analytic. Note as well that if the product $f(z) \cdot g(z) \equiv 0$ on $\mathcal{D}$ then $f(z) \equiv 0$ on $\mathcal{D}$ or $g(z) \equiv 0$ on $\mathcal{D}$.

If we now restrict our attention to polynomials in z, which are analytic everywhere, and Laurent polynomials, which are analytic everywhere except $z = 0$, then we can say something more. Indeed, if $f(z)$ and $g(z)$ are (Laurent) polynomials, then $f(z)+g(z)$, $f(z)-g(z)$ and $f(z)\cdot g(z)$ are not just analytic but are also (Laurent) polynomials. Now, however, we must exercise some care when considering quotients $f(z)/g(z)$, since the result will be analytic in $\mathcal{D}$ (except at the zeros of $g(z)$) but will not be a (Laurent) polynomial in general. However, $f(z)/g(z)$, or indeed any analytic function can be arbitrarily well approximated by polynomials as discussed above.

Let us now consider matrices $\mathcal{R}(z)$ whose entries are analytic functions in $\mathcal{D} \subset C$. We start by noting that for any fixed $z_0 \in \mathbb{C}$, the matrix $\mathcal{R}(z_0)$ is simply a matrix of complex numbers that can be manipulated in the usual ways. For example, we can multiply $\mathcal{R}(z_0)$ by another (conformable) matrix or vector, or compute the EVD of $\mathcal{R}(z_0)$. When we instead allow z to vary, it is still possible to form, say, matrix–matrix or matrix–vector products with $\mathcal{R}(z)$. Indeed, using the arguments in the previous paragraphs, if $\mathcal{R}(z)$ has analytic (polynomial) entries then the resulting matrix or vector will also have analytic (polynomial) entries. However, it is not immediately obvious that we can write down a single, z-dependent eigenvalue decomposition of $\mathcal{R}(z)$ that holds for all values of $z \in \mathcal{D}$. That this is true in certain circumstances is proved in a remarkable result from Rellich [39].

Polynomial Matrix Factorizations. A number of polynomial matrix factorizations have been introduced in the past. Since we are particularly interested in diagonalizations of matrices, these prominently include the Smith and the Smith–Macmillan forms for matrices of polynomials and rational functions, respectively [20]. Popular in the control domain, these allow a decomposition into a diagonal term and two outer factors that are invertible but generally non-orthogonal polynomial matrices. Further, spectral factorizations [37], [40] involve the decomposition of a matrix into a product of a causal, stable matrix and its time-reversed, complex conjugated, and transposed version. These are matrix-valued extensions of Wiener's factorization of a power spectral density into minimum- and maximum phase components, and are supported by numerical tools such as PolyX [41].

In control theory, minimizing the delay of a system is a critical design issue, and hence many of the existing matrix decompositions, such as spectral factorization, emphasize the minimum-phase equivalent of the resulting matrix factors. In signal processing, the delay is often a secondary issue, while e.g. energy preservation (or unitarity) of a transform is crucial. Therefore, in the following, we will explore the diagonalization of an analytic matrix by means of energy-preserving transformations.

III. Preliminaries: Representing broadband signals

A. Signal Model

The received signal at the mth sensor for the discrete-time index n is

$$x_m[n] = \sum_{\ell=1}^{L} \sum_{\tau=0}^{T} a_{\ell,m}[\tau] s_\ell[n-\tau] + v_m[n], \quad m = 1, \ldots, M \ , \tag{3}$$

where $a_{\ell,m}[n]$ models the channel from the ℓth source signal $s_\ell[n]$ to the mth sensor and is an element of $\mathbf{A}[n] \in \mathbb{C}^{L \times M \times T}$, $v_m[n]$ is the additive noise at the mth sensor assumed uncorrelated with the L source signals, and T is the maximum order of any of the channel impulse responses. The received data vector for M sensors is $\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^{\mathrm{T}} \in \mathbb{C}^M$, and each element has a mean of $\mathbb{E}\{x_m[n]\} = 0 \ \forall m$, where $[\cdot]^{\mathrm{T}}$ represents the transpose operator. Similarly, the source and noise data vectors are $\mathbf{s}[n] \in \mathbb{C}^L$ and $\mathbf{v}[n] \in \mathbb{C}^M$, respectively.

Broadband source signals, which naturally arise in, for example, audio, speech, communications, sonar, and radar, are directly reflected by (3). It is also applicable for narrowband systems. Here, the source signal is often described by a complex exponential, $\mathrm{e}^{\mathrm{j}\Omega n}$, where $\mathrm{j} = \sqrt{-1}$ is the imaginary number and Ω is the normalized angular frequency. This means that (3) can be simplified by setting $T = 0$. As an alternative to (3), as shown in Fig. 1, the L source signals, $s_\ell[n] \ \ \forall \ell$, could be generated using spectral-shaped noise obtained by filtering uncorrelated, zero mean unit variance complex Gaussian random variables, $u_\ell[n] \in \mathcal{N}(0, 1)$, through the innovation filters, $f_\ell[n]$ [35].

The channel model in (3) can describe systems in diverse scenarios, for example, instantaneous and convolutive mixtures, near-field and far-field sources, and anechoic and reverberant environments. The signal model in (3) is often simplified by taking the z-transform. However, care is needed as the z-transform of a random signal does not exist. Nonetheless, in the case of deterministic, absolutely summable signals, the z-transform of (3) may be written, in matrix-vector form, as

$$\boldsymbol{x}(z) = \boldsymbol{\mathcal{A}}^P(z)\boldsymbol{\imath}(z) + \boldsymbol{\upsilon}(z) \ , \tag{4}$$

where $\mathbf{A}[n] \ \circ\!\!-\!\!\bullet \ \boldsymbol{\mathcal{A}}(z) \in \mathbb{C}^{L \times M}$, $\mathbf{s}[n] \ \circ\!\!-\!\!\bullet \ \boldsymbol{\imath}(z) \in \mathbb{C}^L$, and $\mathbf{v}[n] \ \circ\!\!-\!\!\bullet \ \boldsymbol{\upsilon}(z) \in \mathbb{C}^M$ are z-transform pairs of the channel matrix, source and noise vectors, respectively. The symbol $[\cdot]^{\mathrm{P}}$ denotes the parahermitian operator, $\boldsymbol{\mathcal{A}}^{\mathrm{P}}(z) = \boldsymbol{\mathcal{A}}^{\mathrm{H}}(1/z^*)$,

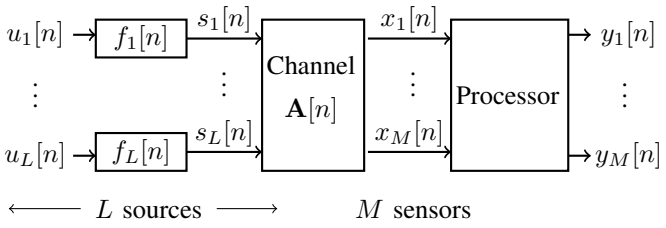


Fig. 1: Multichannel system model for L spectral-shaped source signals and M sensors. Uncorrelated noise signals $\mathbf{v}[n]$, not drawn in the figure, are optionally added to each sensor based on (3).

which involves a hermitian transpose followed by a time-reversal operation [20], where $[.]^{\mathrm{H}}$ denotes the hermitian transpose operator. The well-known equivalence of convolution in the time-domain and multiplication in the z-domain [33] is expressed in (3) and (4).

B. Covariance Matrices

The covariance matrix used in many narrowband subspace-based approaches [5], [8], [10] is described by

$$\mathbf{R} = \mathbb{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\} \,, \tag{5}$$

using the data vector obtained from (3). The (m,ℓ)th element of $\mathbf{R}$ is $r_{m,\ell} = \mathbb{E}\{x_m(n)x_\ell^*(n)\}$, and the expectation operation is performed over n. In practice, the expectation is approximated using the sample mean where the inner product between the received signals at the mth and ℓth sensor is computed before normalizing by the total number of samples N. Because the inner product is calculated sample-wise, the covariance matrix instantaneously captures the spatial relationship between different sensors. This paper will call it the instantaneous (or spatial) covariance matrix.

When the system involves convolutive mixing or broadband signals, time delays between signals at different sensors need to be modelled. This spatio-temporal relationship is explicitly captured by the space-time covariance matrix, parameterized by the discrete-time lag parameter $\tau \in \mathbb{Z}$, defined as [21]

$$\mathbf{R}[\tau] = \mathbb{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\} \,. \tag{6}$$

The (m,ℓ)th element of $\mathbf{R}[\tau]$, arising from sensors with a fixed geometry, is $r_{m,\ell}[\tau] = \mathbb{E}\{x_m[n]x_\ell^*[n-\tau]\}$ and again the expectation operation is performed over n, where wide-sense temporal stationarity is assumed. The auto-correlation and cross-correlation sequences are obtained when $m = \ell$ and $m \neq \ell$, respectively. Furthermore, (5) can be seen as a special case of (6) when only the instantaneous lag is considered, i.e., $\mathbf{R}[0]$ is the coefficient of z^0 when $\tau = 0$, as shown in Fig. 2.

The z-transform of the space-time covariance matrix in (6), $\mathbf{R}[\tau] \circ\!\!-\!\!\bullet\, \mathcal{R}(z) \in \mathbb{C}^{M\times M}$, is known as a cross spectral density (CSD) and is represented by

$$\mathcal{R}(z) = \sum_{\tau=-\infty}^{\infty} \mathbf{R}[\tau] z^{-\tau} \,, \tag{7}$$

is a parahermitian polynomial matrix satisfying the property, $\mathcal{R}^{\mathrm{P}}(z) = \mathcal{R}(z)$. The polynomial matrix can be interpreted as a matrix-valued polynomial (function) or a polynomial with matrix coefficients, i.e., $\mathbf{R}[\tau]$ is the matrix coefficient of $z^{-\tau}$. This is visualized in the middle of Fig. 2 , which describes the temporal evolution of the spatial relationship across the entire array. Equivalently, the same polynomial matrix can also be interpreted as a matrix with polynomial elements, representing the temporal correlation in the z-domain between sensor pairs, for example, element $r_{3,1}(z)$ for sensors 3 and 1 on the right of Fig. 2.

Key Statement

The space-time covariance matrix completely captures the second-order statistics of multichannel broadband signals via auto- and cross-correlation functions. Its z-transform has the useful property of being parahermitian.

C. Comparison With Other Broadband Signal Representations

The multichannel signal model introduced in Section III will be compared against two signal representations commonly encountered in array processing. They are the TDL and short-time Fourier transform (STFT) approaches.

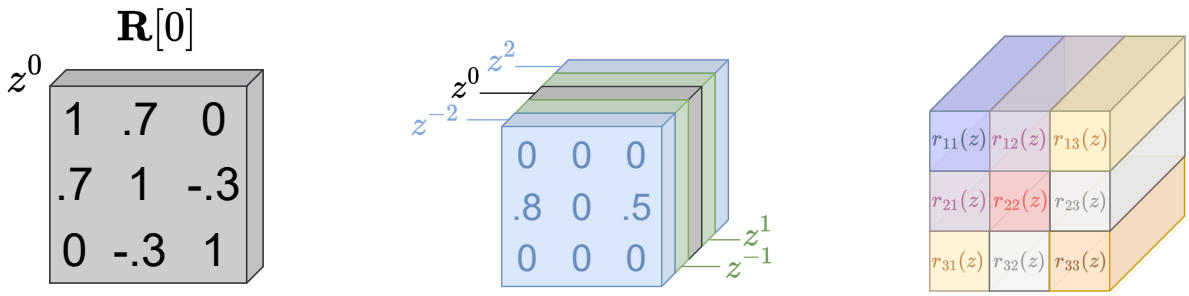
1) Tapped Delay Line (TDL) Processing: The relative delays with which broadband signals arrive at different sensors cannot be sufficiently modelled by phase shifts because they can only be accurate at a single frequency. Therefore, these delays need to be implemented by filters that possess at the very least frequency-dependent phase shifts. Such filters must rely on processing a temporal window of the signals; the access to this window can, in the FIR case, be provided by TDLs that are attached to each of the M array signals. The length T of these TDLs will determine the accuracy with which such delays—often of a fractional nature [18]—are realized.

Based on the array signal vector $\mathbf{x}[n]$, a T-element TDL provides data that can be represented by a concatenated vector $\boldsymbol{\chi}[n] = [\mathbf{x}^{\mathrm{T}}[n], \ldots, \mathbf{x}^{\mathrm{T}}[n-T+1]]^{\mathrm{T}} \in \mathbb{C}^{MT}$, which holds both spatial and temporal samples. For the covariance matrix of $\boldsymbol{\chi}[n]$, $\mathbf{R}_\chi = \mathbb{E}\{\boldsymbol{\chi}[n]\boldsymbol{\chi}^{\mathrm{H}}[n]\}$, we have

$$\mathbf{R}_\chi = \begin{bmatrix} \mathbf{R}[0] & \ldots & \mathbf{R}[T-1] \\ \vdots & \ddots & \vdots \\ \mathbf{R}[-T+1] & \ldots & \mathbf{R}[0] \end{bmatrix} . \tag{8}$$

Although of different dimensions, the covariance $\mathbf{R}_\chi$ thus contains as sub-matrices the same terms that also make up the space-time covariance matrix $\mathbf{R}[\tau]$. However, it is not necessarily clear prior to processing how large or small T should be selected. Apart from its impact on the accuracy of a delay implementation, if T is selected smaller than the coherence time of the signal, then some temporal correlations for lags $|\tau| \geq T$ in the signals are missed, leading to a potentially insufficient characterization of the signals' second order statistics. If T is set too large, then no extra correlation information is included but additional noise is added.

The EVD of the covariance matrix $\mathbf{R}_\chi = \mathbf{Q}_\chi \boldsymbol{\Lambda}_\chi \mathbf{Q}_\chi^{\mathrm{H}}$ gives access to MT eigenvalues in $\boldsymbol{\Lambda}_\chi$. In inspecting these



Example of matrix at zero-lag. Polynomial with matrix coefficients. Matrix with polynomial elements.

Fig. 2: A typical spatial covariance matrix for the zero lag, i.e., $\tau = 0$, is the matrix coefficient of z^0, shown on the left. In general, each matrix slice corresponds to a coefficient of the polynomial, shown in the centre. The same polynomial matrix is also a matrix comprising polynomial elements represented by tubes in the same cube shown on the right.

eigenvalues, there no longer is any separation between space and time, and, for example, a single broadband source that is captured by the array in its data vector $\mathbf{x}[n]$ can generate, depending on its bandwidth, anything between one and $T+\Delta$ non-zero eigenvalues, where Δ is the maximum propagation delay across the array that any source can experience. Hence, tasks such as source enumeration can become challenging.

Furthermore, in narrowband processing, a common procedure is to project the received signals $\mathbf{x}$ onto the so-called signal subspace as this would suppress some of the noise [7]. The signal subspace is defined by partitioning the eigenvalues by magnitude and selecting the subset of eigenvectors corresponding to the larger eigenvalues. Mimicking this in the broadband case would mean partitioning $\mathbf{Q}_\chi = \begin{bmatrix}\mathbf{Q}_s & \mathbf{Q}_n\end{bmatrix}$ where $\mathbf{Q}_s$ correspond to the larger eigenvalues. In the narrowband case, it is well known that, in general, the projected signals $\underline{\mathbf{y}}[n] = \mathbf{Q}_s^{\mathrm{H}}\,\boldsymbol{\chi}[n]$ are not the source signals but merely span the same space. However, if only one source signal is present then the projected signal is the source signal. In the broadband case, not even this is true since, as noted above, we might have more than one eigenvalue per signal.

2) Short-time Fourier Transform (STFT): If we take a T-point DFT $\mathbf{W}$ of each of the TDLs in $\boldsymbol{\chi}[n]$, we evaluate $\boldsymbol{\xi}[n] = (\mathbf{W} \odot \mathbf{I}_M)\boldsymbol{\chi}[n]$, with $\odot$ the Kronecker product. The DFT-domain covariance $\mathbf{R}_\xi = \mathbb{E}\{\boldsymbol{\xi}[n]\boldsymbol{\xi}^{\mathrm{H}}[n]\} = (\mathbf{W} \odot \mathbf{I}_M)\mathbf{R}_\chi(\mathbf{W} \odot \mathbf{I}_M)^{\mathrm{H}}$ is generally non-sparse due to cross-coupling between DFT bins. This cross-coupling does not subside even as T is increased. For bin-wise processing — i.e. processing each of the frequency bins across the array independently of other frequency bins — many of the terms in $\mathbf{R}_\xi$ are neglected, leading to processing that can be very low-cost but generally is only suboptimal. To achieve optimality, time-domain criteria must be embedded in the processing, which generally leads to cross-terms between bins [14], [42]. The generally dense nature of $\mathbf{R}_\xi$ can be relaxed when employing more frequency-selective subband methods over DFT-processing, but cross-terms at least between adjacent subband still remain [43]. Together with the increased computation cost of such filters over the DFT, this thus negates the low-complexity aspiration of this approach.

Key Statement

Broadband processing requires accurately representing fractional time delays. Previous approaches do not lead to proper generalizations of the narrowband algorithms and are often suboptimal.

IV. Polynomial Matrix Eigenvalue Decomposition

As discussed in Section III-C, conventional approaches to processing broadband signals have some shortcomings. Arguably, this is because the incorrect signal representation was used. Specifically, the use of a tapped delay-line, with either time-domain or frequency domain processing, mixes up the spatial and temporal dimensions. This section builds on the signal model in Section III, representing the broadband system using z-transforms or "polynomials". Guided by the successful use of linear algebra, i.e. EVD, in narrowband systems, this section focuses on the decomposition of parahermitian polynomial matrices such as the space-time covariance matrix. That is, given a parahermitian polynomial matrix $\mathcal{R}(z) = \mathcal{R}^{\mathrm{P}}(z)$, does a decomposition $\mathcal{R}(z) = \mathcal{Q}(z)\boldsymbol{\Lambda}(z)\mathcal{Q}^{\mathrm{P}}(z)$ exist, where $\boldsymbol{\Lambda}(z)$ is diagonal and $\mathcal{Q}(z)$ is paraunitary?

Note that the EVD can diagonalize a parahermitian matrix $\mathbf{R}[\tau]$ for only one specific lag $\tau = \tau_0$, or alternatively $\mathcal{R}(z) \bullet\!\!-\!\!\circ\, \mathbf{R}[\tau]$ for one specific value $z = z_0$. The unitary matrices that accomplish the diagonalization at that value are unlikely to diagonalize the matrix at other values $\tau \neq \tau_0$ or $z \neq z_0$. We, therefore, require a decomposition that diagonalizes $\mathcal{R}[\tau]$ for all values of τ, or $\mathcal{R}(z)$ for all values of z within the RoC. We address the existence of such a decomposition via the analytic EVD in Section IV-A, and provide some comments on numerical algorithms in Section IV-B.

A. Analytic EVD

The key to a more general EVD is the work by Franz Rellich [44], who in the context of quantum mechanics, investigated a matrix-valued function $\mathbf{A}(t)$ that is self-adjoint, i.e. $\mathbf{A}(t) = \mathbf{A}^{\mathrm{H}}(t)$, and analytic in t on some real interval. Matrix-valued functions of this type admit a decomposition $\mathbf{A}(t) = \mathbf{U}(t)\boldsymbol{\Gamma}(t)\mathbf{U}^{\mathrm{H}}(t)$ with matrix-valued functions $\mathbf{U}(t)$ and $\boldsymbol{\Gamma}(t)$ that are also analytic in t, and where $\boldsymbol{\Gamma}(t)$ is diagonal

and $\mathbf{U}(t_0)$ is unitary for any specific value $t = t_0$. These results were obtained through perturbation analysis [45], where for the EVD of a matrix $\mathbf{A}(t_0) = \mathbf{U}(t_0)\mathbf{\Gamma}(t_0)\mathbf{U}^{\mathrm{H}}(t_0)$, a change of $\mathbf{A}(t_0)$ by some small Hermitian matrix results in only a limited perturbation of both the eigenvalues and eigenvectors. There is no such guarantee if $\mathbf{A}(t)$ is not analytic in t; even infinite differentiability does not suffice [45].

To decompose a matrix $\mathcal{R}(z)$ that is analytic in the complex-valued parameter z, it suffices to investigate $\mathcal{R}(z)$ on the unit circle for $z = \mathrm{e}^{\mathrm{j}\Omega}$. This is due to the uniqueness theorem for analytic functions, which guarantees that if two functions are identical on some part of their region of convergence (ROC) —here the unit circle, which must always be included — they must be identical across the entire ROC. Although $\Omega \in \mathbb{R}$, Rellich's results do not directly apply as they do not imply a 2π-periodicity. Without such periodicity, it is not possible to re-parameterize the EVD factors by replacing $\mathrm{e}^{\mathrm{j}\Omega}$ with z and hence produce an EVD that is analytic in z. However, it has recently been shown that Rellich's result admits $2\pi N$-periodic eigenvalue functions, and, furthermore, $N = 1$ unless the data generating $\mathcal{R}(z)$ emerges from N-fold multiplexing or block filtering [23], [24]. Analytic eigenvector functions then exist with the same periodicity as the eigenvalues [46]. Therefore, an analytic EVD for this N-fold multiplexed system

$$\mathcal{R}(z^N) = \mathcal{Q}(z)\mathbf{\Lambda}(z)\mathcal{Q}^{\mathrm{P}}(z) \tag{9}$$

exists with analytic factors such that $\mathbf{\Lambda}(z)$ is diagonal. The matrix $\mathcal{Q}(z)$ contains the eigenvector functions and for $z = \mathrm{e}^{\mathrm{j}\Omega_0}$ is unitary. For a general z, $\mathcal{Q}(z)$ is paraunitary, such that $\mathcal{Q}(z)\mathcal{Q}^{\mathrm{P}}(z) = \mathcal{Q}^{\mathrm{P}}(z)\mathcal{Q}(z) = \mathbf{I}$. Paraunitarity is an extension of the orthonormal and unitary properties of matrices from the real- and complex-valued cases to matrices that are functions in a complex variable [20]. For ease of exposition in the following, we will talk of 'analytic matrices $\mathcal{X}(z)$' with the understanding that $\mathcal{X}(z)$ is a matrix-valued analytic function.

In the analytic EVD of (9), the eigenvalue function is $\mathbf{\Lambda}(z) = \mathrm{diag}\{\lambda_1(z), \ldots, \lambda_M(z)\}$, where $\mathrm{diag}(\cdot)$ forms a diagonal matrix from its argument. When evaluated on the unit circle, the eigenvalues $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega})$, $m = 1, \ldots, M$, are real-valued, and unique up to a permutation. If there are M distinct eigenvalues, i.e., $\lambda_m(\mathrm{e}^{\mathrm{j}\Omega}) = \lambda_\mu(\mathrm{e}^{\mathrm{j}\Omega})$ only for $m = \mu$ for any $m, \mu = 1, \ldots, M$, then the corresponding eigenvectors $\boldsymbol{q}_m(z)$ in $\mathcal{Q}(z) = [\boldsymbol{q}_1(z), \ldots, \boldsymbol{q}_M(z)]$ are unique up to an arbitrary allpass function, i.e., $\boldsymbol{q}'_m(z) = \psi_m(z)\,\boldsymbol{q}_m(z)$ is also a valid analytic eigenvector of $\mathcal{R}(z^N)$, where $\psi_m(z)$ is allpass.

As an example, consider the system from [23]

$$\mathcal{R}(z) = \begin{bmatrix} \frac{1-\mathrm{j}}{2}z + 3 + \frac{1+\mathrm{j}}{2}z^{-1} & \frac{1+\mathrm{j}}{2}z^2 + \frac{1-\mathrm{j}}{2} \\ \frac{1+\mathrm{j}}{2} + \frac{1-\mathrm{j}}{2}z^{-2} & \frac{1-\mathrm{j}}{2}z + 3 + \frac{1+\mathrm{j}}{2}z^{-1} \end{bmatrix}, \tag{10}$$

which is constructed from eigenvalues $\mathbf{\Lambda}(z) = \mathrm{diag}\{z + 3 + z^{-1},\ \mathrm{j}z + 3 - \mathrm{j}z^{-1}\}$, and corresponding eigenvectors $\boldsymbol{q}_{1,2}(z) = 1/\sqrt{2}[1,\ \pm z^{-1}]^{\mathrm{T}}$. The evaluation of the eigenvalues on the unit circle, $\lambda_{1,2}(\mathrm{e}^{\mathrm{j}\Omega})$ is shown in Fig. 3(a). For the eigenvectors, the Hermitian angle $\varphi_m(\mathrm{e}^{\mathrm{j}\Omega}) = \arccos(|\boldsymbol{q}_1^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}0})\,\boldsymbol{q}_m(\mathrm{e}^{\mathrm{j}\Omega})|)$, is drawn in Fig. 3(b). Note that due to the analyticity of the EVD factors, all these quantities evolve smoothly with the normalized angular frequency Ω. An allpass modification of the eigenvectors might be as simple as imposing a delay; while this will not affect $\varphi_m(\mathrm{e}^{\mathrm{j}\Omega})$, it can increase support of $\mathbf{Q}[n] \circ\!\!-\!\!\bullet\, \mathcal{Q}(z)$.

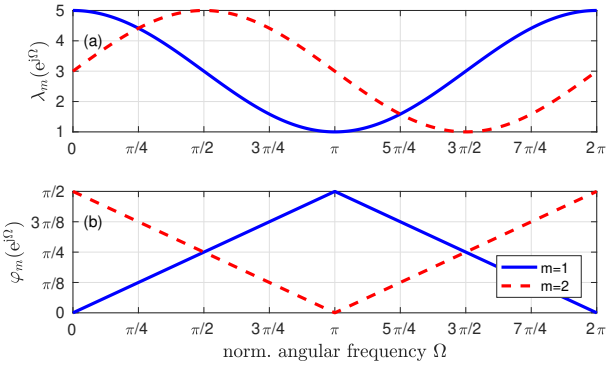


Fig. 3: Example for (a) analytic eigenvalues on the unit circle and (b) Hermitian angles of the corresponding analytic eigenvectors, measured against a reference vector [23].

While in the above example, the factorization yields polynomial factors, this does not have to be the case – they could be Laurent or power series. For example, modifying the previous eigenvectors by arbitrary allpass functions does not invalidate the decomposition but it may change the order of $\boldsymbol{q}_m(z)$ to ∞, i.e, a power series. More generally, Laurent polynomial matrices $\mathcal{R}(z)$ are likely to lead to algebraic or even transcendental functions as EVD factors [23], [24]. Nonetheless, recall from Section I that analyticity implies absolute convergence in the time-domain. Therefore, the best least-squares approximation is achieved by truncation. Further, as the approximation order increases, the approximation error can be made arbitrarily small.

The components of the analytic PEVD also have some useful properties. The matrix of eigenvectors $\mathbf{Q}[n]$ can be viewed as a lossless filter bank. Clearly, it transforms the input times series into another set of time series. However, being paraunitary, the energy in the output signals is the same as that of the input signals. Furthermore, the output signals are strongly decorrelated. That is, any two signals have zero cross-correlation coefficient at all lags. Significantly, the signals are not temporally whitened, i.e., do not have an impulse as their auto-correlation function. Note that the order of a z-transform is connected to the time-domain support of the corresponding time-series. Thus, the computational cost of implementing such a filter bank is related to the order of $\mathcal{Q}(z)$. In general, the eigenvalues of a narrowband covariance matrix have differing magnitudes, with the presence of small values indicating approximate linear dependency between the input signals. Similarly, the eigenvalues on the diagonal of $\mathbf{\Lambda}[n]$ can show linear dependence but in a frequency-dependent manner.

Key Statement

The pioneering work of Rellich showed that an analytic EVD exists for a matrix function. Applying this to a space-time covariance matrix on the unit circle introduces some additional constraints but results in the existence of an analytic PEVD.

B. PEVD Algorithms

The first attempt at producing a PEVD algorithm began with the second-order sequential best rotation (SBR2) [21], which was motivated by Jacobi's method for numerically computing the EVD [2]. The PEVD of $\mathcal{R}(z)$, i.e., (7), as given by (9) for $N = 1$ and established in Section IV-A can be approximated using an iterative algorithm and is expressed as [21], [47]

$$\mathcal{R}(z) \approx \mathcal{U}(z)\mathbf{\Lambda}(z)\mathcal{U}^{\mathrm{P}}(z)\,, \tag{11}$$

where the columns of the polynomial matrix, $\mathcal{U}(z) \in \mathbb{C}^{M\times M}$, correspond to the eigenvectors with their associated eigenvalues on the diagonal polynomial matrix, $\mathbf{\Lambda}(z) \in \mathbb{C}^{M\times M}$. The Laurent polynomial matrix factors $\mathcal{U}(z)$ and $\mathbf{\Lambda}(z)$ are necessarily analytic being of finite order. However, under certain circumstances, the theoretical factors in the PEVD might not be analytic [23], [24]. In which case, the Laurent polynomial matrix factors $\mathcal{U}(z)$ and $\mathbf{\Lambda}(z)$ are only approximations to the true factors. Hence, the approximation in (11).

Rewriting (11) as

$$\mathbf{\Lambda}(z) \approx \mathcal{U}^{\mathrm{P}}(z)\mathcal{R}(z)\mathcal{U}(z)\,, \tag{12}$$

the diagonalization of $\mathcal{R}(z)$ can be achieved by generalized similarity transformations with $\mathcal{U}(z)$ satisfying the paraunitary or lossless condition [20]

$$\mathcal{U}^{\mathrm{P}}(z)\mathcal{U}(z) = \mathcal{U}(z)\mathcal{U}^{\mathrm{P}}(z) = \mathbf{I}\,, \tag{13}$$

where $\mathbf{I}$ is the identity matrix. The similarity transform $\mathcal{U}(z)$ may be calculated via an iterative algorithm such as the SBR2 or sequential matrix diagonalization (SMD) [47]. Here, a sequence of elementary paraunitary transformations $\mathcal{G}_i(z)$ $(i = 1, \dots)$ are applied to $\mathcal{R}(z)$ until the polynomial matrix becomes approximately diagonal i.e. starting from $\tilde{\mathcal{R}}_0(z) = \mathcal{R}(z)$, the following expression is iterated

$$\tilde{\mathcal{R}}_i(z) = \mathcal{G}_i^{\mathrm{P}}(z)\tilde{\mathcal{R}}_{i-1}(z)\mathcal{G}_i(z) \tag{14}$$

until $\tilde{\mathcal{R}}_{N_I}(z)$ is approximately diagonal for some N_I. An elementary paraunitary transformation takes the form of the product of a unitary transformation and a polynomial delay matrix, $\mathrm{diag}\{1, \dots, 1, z^n, 1, \dots, 1\}$.

Figure 4 shows the steps involved during every iteration of SBR2. At each iteration, the algorithm searches for the off-diagonal element with the largest magnitude across all z-planes, as marked in red in Fig. 4. If the magnitude exceeds a predefined threshold, a delay polynomial matrix is applied to bring the element to the principal z^0-plane, as shown in Fig. 4. A unitary matrix, designed to zero out two elements on the zero-lag plane, is applied to the entire polynomial matrix. Note that applying one elementary paraunitary transformation may make some previously small off-diagonal elements larger, but overall the algorithm converges to a diagonal matrix. As observed in Fig. 4, the delay step can increase the polynomial order and make it unnecessarily large. Therefore, a trimming procedure [21] is used to control the growth of polynomial order by discarding negligibly small coefficients in the outer planes, e.g., z^{-4} and z^4 in Fig. 4. Furthermore, the similarity transformations in (12) affect a pair of dominant elements so that the search space can be halved due to the preservation of symmetry. The algorithm terminates when the magnitudes of all off-diagonal elements fall below the pre-set threshold or when a user-defined maximum number of iterations is reached.

This has led to a family of time-domain algorithms based on SBR2 [21] and SMD [47]. The computational complexity of these numerical algorithms is at least $\mathcal{O}(M^3T)$ due to matrix multiplication applied to every lag [48]. The additional complexity incurred over the EVD approach is essential for the temporal decoupling of broadband signals. Furthermore, some promising efforts using parallelizable hardware [49] and numerical tricks [50] have been proposed and the decomposition can be computed in a fraction of a second. These algorithms are also guaranteed to produce polynomial, paraunitary eigenvectors but tend to generate spectrally majorized eigenvalues, which may not be analytic. Two functions $f_1(z)$ and $f_2(z)$ are said to be spectrally majorized if, on the unit circle, one function's magnitude is always greater than the other's. Figure 5 shows the results of using the SMD algorithm to process the matrix used to generate Fig. 3. In Fig. 5, the eigenvalue in blue is always greater than the one in red. In contrast, the (analytic) eigenvalues in Fig. 3 intersect and are not spectrally majorized. As shown in Fig. 5(b), forcing a spectrally majorized solution for the eigenvalues leads to the eigenvectors having discontinuities which are difficult to approximate with polynomials. To get an accurate result, high-order polynomials are required. This, in turn, has consequences for the implementation cost of any signal processing based on the output of these algorithms. Note, however, that spectral majorization can be advantageous in some situations – see Section V-B.

Unlike the SBR2 algorithm [21], there is no proof that the SMD algorithm will always produce spectrally majorized eigenvalues, although evidence from the use of this algorithm strongly supports this conjecture. Given the issues that spectral majorization produces in terms of exploiting $\mathcal{Q}(z)$ as a filter bank or for identifying subspaces, recent work has been directed at designing an algorithm that can produce a PEVD whose components are guaranteed to be analytic. One such approach [51] involves working in the frequency domain and taking steps to ensure that the spectral coherence is not lost.

A number of algorithms have been designed for decompositions of fixed order, and without proven convergence. This includes the approximate EVD (AEVD) algorithm [52], which in the time-domain applies a fixed number of elementary paraunitary operations in an attempt to diagonalise $\mathcal{R}(z)$. In the DFT domain, [53] aims to extracts maximally smooth eigenvalues and eigenvectors, which therefore can target the extraction of the analytic solution.

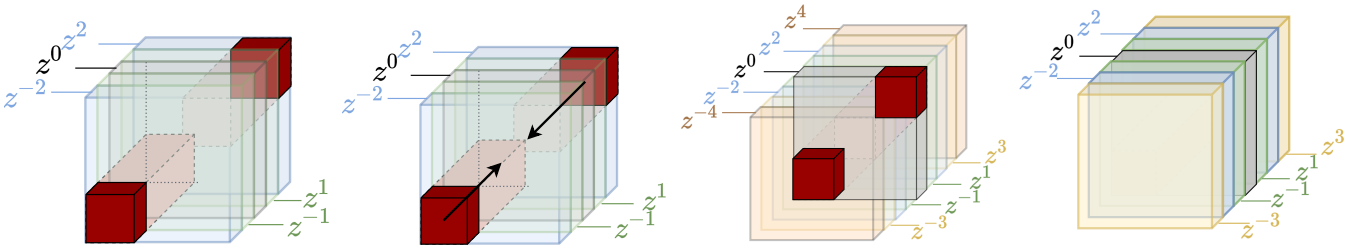


Fig. 4: Each PEVD iteration involves the following four steps. The polynomial matrix is first searched for the maximum off-diagonal across all lags (leftmost). The second delay step brings the largest element to the principal z^0-plane (second from the left). The third is the zeroing step which transfers energy from the off-diagonal elements to the diagonal (second from the right). The final trimming step discards negligibly small coefficients in the outer matrix slices (rightmost).

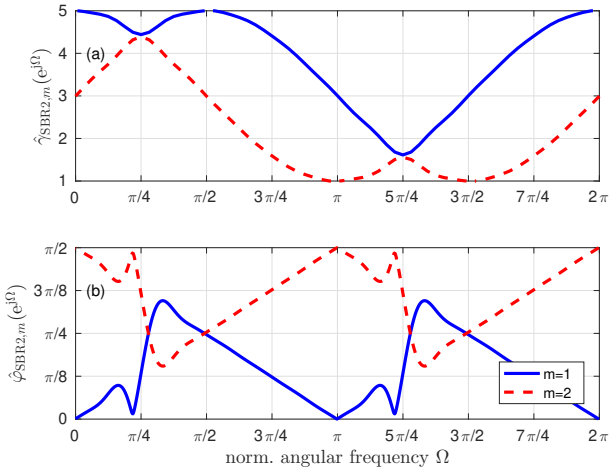


Fig. 5: Results of using SMD to decompose the matrix used in Figure 3: (a) eigenvalues on the unit circle and (b) Hermitian angles of the corresponding analytic eigenvectors, measured against a reference vector [23].

Key Statement

Approximating analytic functions by polynomials allows the development of PEVD algorithms based on an elementary paraunitary operator. The resulting algorithms are guaranteed to produce polynomial, paraunitary eigenvectors but tend to generate spectrally majorized eigenvalues. This property has benefits as well as drawbacks.

V. Example Applications using PEVD

This section highlights three application cases, that demonstrate key examples where PEVD-based approaches can offer advantages over state-of-the-art processing. In Section V-A, we demonstrate how for adaptive beamforming, the computational complexity is decoupled from the tap delay line length that otherwise determines the cost of a broadband adaptive beamformer. Section V-B shows how in subband coding, the PEVD can generate a system with optimized coding gain, and helps to formulate optimum compaction filter banks that previously could only be stated for the two-channel case. Finally, Section V-C addresses how the preservation of spectral coherence can provide perceptually superior results over DFT-based speech enhancement algorithms.

A. PEVD-based Adaptive Beamforming

To explore PEVD-based beamforming, we first recall some aspects of narrowband beamforming before defining a linearly constrained minimum variance (LCMV) beamformer using both tapped delay line (TDL) and PEVD-based formulations. We work with an arbitrary geometry of M sensors, but for simplicity, assume free-space propagation and for the array to be sufficiently far-field to neglect any loss in amplitude across its sensors.

Spatial Filtering and Steering Vector. Spatial filtering uses the fact that wavefronts arriving from different sources have a different delay profile when arriving at the sensors. If there are L spatially separated sources, then for the ℓth source, $\ell = 1, \ldots, L$, let this delay profile be $\{\tau_{\ell,1}, \ldots, \tau_{\ell,M}\}$, where $\tau_{\ell,m}$ is the delay at the mth sensor with respect to some common reference point. We further define a vector of transfer functions $\boldsymbol{a}_\ell(z) = [\delta_{\tau_{\ell,1}}(z), \ldots, \delta_{\tau_{\ell,M}}(z)]^{\mathrm{T}}$ containing fractional delay filters, where $d_\tau[n] \circ\!\!-\!\!\bullet\, \delta_\tau(z)$ implements a delay by $\tau \in \mathbb{R}$ samples [18]. We refer to $\boldsymbol{a}_\ell(z)$ as a broadband steering vector since, when evaluated at a fixed frequency Ω_ℓ, the ℓth source can be regarded as a narrowband signal with centre frequency Ω_ℓ, in which this vector of functions reduces to the well-known steering vector $\boldsymbol{a}_\ell(z)|_{z=\mathrm{e}^{\mathrm{j}\Omega_\ell}} = \boldsymbol{a}_\ell(\mathrm{e}^{\mathrm{j}\Omega_\ell}) \in \mathbb{C}^M$. The latter contains the phase shifts that each sensor experiences with respect to the ℓth source. If at least two sensors satisfy the spatial sampling theorem, and for a particular frequency $\Omega_\ell = \Omega_0$, this steering vector is unique with respect to the direction of arrival of the ℓth source.

We want to process the array data $\mathbf{x}[n] \in \mathbb{C}^M$ by a vector of filters $\mathbf{w}[n] \circ\!\!-\!\!\bullet\, \boldsymbol{w}(z)$, with $\boldsymbol{w}^{\mathrm{P}}(z) = [\mathcal{W}_1(z), \ldots, \mathcal{W}_M(z)]$ and $\mathcal{W}_m(z) \bullet\!\!-\!\!\circ\, w_m[n]$ is the filter processing the mth sensor signal, such that the array output $y[n]$ is sum of the filtering operations, $y[n] = \sum_\nu \mathbf{w}^{\mathrm{H}}[-\nu]\mathbf{x}[n-\nu] = \sum_{m,\nu} w_m[\nu] x_m[n-\nu]$. The definition of the filter vector $\mathbf{w}[n]$ with its time reversed and conjugated weights may seem cumbersome, but it follows similar conventions for complex-valued data [54] and will later simplify the z-transform notation.

Narrowband Beamforming. In the narrowband case, the delay filters can be replaced by complex coefficients in a vector $\mathbf{w}^{\mathrm{H}} = [w_1, \ldots, w_M] \in \mathbb{C}^M$ that implement phase shifts. To generate a different gain f_ℓ, $\ell = 1, \ldots, L$ with respect to each of the L sources, the beamformer defined by $\mathbf{w}$ must satisfy the constraint equation

$$\underbrace{\begin{bmatrix} \boldsymbol{a}_1^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_1}) \\ \vdots \\ \boldsymbol{a}_L^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\Omega_M}) \end{bmatrix}}_{\mathbf{C}} \mathbf{w} = \underbrace{\begin{bmatrix} f_1 \\ \vdots \\ f_L \end{bmatrix}}_{\mathbf{f}}, \tag{15}$$

where $\mathbf{C} \in \mathbb{C}^{L \times M}$ and $\mathbf{f} \in \mathbb{C}^L$ are the constraint matrix and associated gain vector for the L constraints. In the presence of spatially white noise, the minimum mean-square error (MMSE) solution is the quiescent beamformer $\mathbf{w}_{\mathrm{q}} = \mathbf{C}^{\dagger}\mathbf{f}$ [54], where $\mathbf{C}^{\dagger}$ is the pseudo-inverse of $\mathbf{C}$. If the noise is spatially correlated, then the LCMV formulation

$$\min_{\mathbf{w}} \mathbb{E}\{|y[n]|^2\} \qquad \text{s.t.} \quad \mathbf{C}\mathbf{w} = \mathbf{f} \tag{16}$$

provides the MMSE solution, now constrained by (15). Solutions to (16) include, for example, the Capon beamformer, $\mathbf{w}_{\mathrm{opt}} = (\mathbf{R}[0])^{-1}\mathbf{C}^{\mathrm{H}}[\mathbf{C}(\mathbf{R}[0])^{-1}\mathbf{C}^{\mathrm{H}}]^{-1}\mathbf{f}$ or the generalized sidelobe canceller (GSC). For the GSC, a 'quiescent beamformer' $\mathbf{w}_{\mathrm{q}}$ implements the constraints in $\mathbf{C}$, and a 'blocking matrix' $\mathbf{B}$ is constructed such that $\mathbf{C}\mathbf{B} = \mathbf{0}$; thus, when operating on the array data $\mathbf{x}[n]$, its output is free of any components appearing in the quiescent beamformer output. All that remains is to suppress any signal components in the quiescent beamformer output that correlate with the blocking matrix output. This unconstrained optimization problem for the vector $\mathbf{w}_{\mathrm{a}}$ in Fig. 6(a) can be addressed by adaptive filtering algorithms via a noise cancellation architecture [54]. The overall response of the adapted GSC is $\mathbf{w} = \mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}}$.

TDL-based Generalized Sidelobe Canceller. In the broadband case, each sensor is followed by a TDL of length T in order to implement a finite impulse response filter that can resolve explicit time delays [55]. This leads to the concatenated data vector $\boldsymbol{\chi}[n] = [\mathbf{x}^{\mathrm{H}}[n], \ldots, \mathbf{x}^{\mathrm{H}}[n-T+1]]^{\mathrm{H}} \in \mathbb{C}^{MT}$ presented in Section III-C1. The weight vector $\mathbf{v} \in \mathbb{C}^{MT}$ now performs a linear combination across this spatio-temporal window, such that the beamformer output becomes $y[n] = \mathbf{v}^{\mathrm{H}}\boldsymbol{\chi}[n]$. Analogous to (15), a constraint equation $\mathbf{C}_{\mathrm{b}}\mathbf{v} = \mathbf{f}_{\mathrm{b}}$ defines the frequency responses in a number of directions. The constraint formulation for a linear array with a look direction towards broadside is as straightforward as in the narrowband case [56]. For linear arrays with off-broadside constraints, or for arbitrary arrays, the formulation of constraints becomes more tricky, and can be based on stacked narrowband constraints across a number of DFT bins akin to the single frequency formulation that leads to (15). Since it may not be clear how many such constraints should be stacked, robust approaches start with a large number of these, which are then trimmed to a reduced set of linearly independent constraints using e.g., a QR decomposition [57]. Overall, with respect to the narrowband case, the dimensions of the constraint matrix and constraining vector will increase approximately T-fold such that $\mathbf{C}_{\mathrm{b}} \in \mathbb{C}^{TL \times TM}$ and $\mathbf{f}_{\mathrm{b}} \in \mathbb{C}^{TL}$.

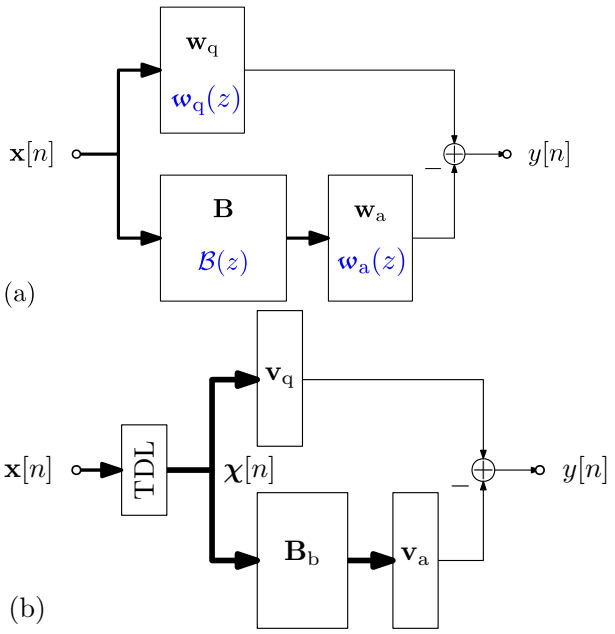


Fig. 6: Generalized sidelobe canceller for (a) the narrowband (black quantities in boxes) and the PEVD-based cases (blue quantities in boxes), and (b) the TDL-based case.

For the broadband GSC [58], a quiescent beamformer $\mathbf{v}_{\mathrm{q}} = \mathbf{C}_{\mathrm{b}}^{\dagger}\mathbf{f}_{\mathrm{b}} \in \mathbb{C}^{TL}$ will generate an output that still contains any structured interference that is not addressed by the constraint equation. A signal vector correlated with this remaining interference is produced by the blocking matrix $\mathbf{B}_{\mathrm{b}} \in \mathbb{C}^{T(M-L) \times TM}$, whose columns, akin to the narrowband case, must span the nullspace of $\mathbf{C}_{\mathrm{b}}$ such that $\mathbf{C}_{\mathrm{b}}\mathbf{B}_{\mathrm{b}} = \mathbf{0}$. Its output is then linearly combined by an adaptive filter $\mathbf{v}_{\mathrm{a}} \in \mathbb{C}^{T(M-L)}$ such that the overall beamformer output in Fig. 6(b) is minimized in the MSE sense. Note that the TDL length determines the dimensions of all GSC components, with the overall adapted response of the beamformer, with respect to the input $\mathbf{x}[n]$ extended to the TDL representation in $\boldsymbol{\chi}[n]$, being $\mathbf{v} = \mathbf{v}_{\mathrm{q}} - \mathbf{B}_{\mathrm{b}}\mathbf{v}_{\mathrm{a}}$.

PEVD-Based Generalized Sidelobe Canceller. In the PEVD-based approach, we replace narrowband quantities in the narrowband formulation by their polynomial equivalents to address the broadband case. This includes substituting the Hermitian transpose $\{\cdot\}^{\mathrm{H}}$ by a parahermitian transposition $\{\cdot\}^{\mathrm{P}}$. Thus, the constraint equation becomes

$$\underbrace{\begin{bmatrix} \boldsymbol{a}_1^{\mathrm{P}}(z) \\ \vdots \\ \boldsymbol{a}_L^{\mathrm{P}}(z) \end{bmatrix}}_{\mathcal{C}(z)} \boldsymbol{w}(z) = \underbrace{\begin{bmatrix} f_1(z) \\ \vdots \\ f_L(z) \end{bmatrix}}_{\boldsymbol{f}(z)}. \tag{17}$$

The constraint matrix $\mathcal{C}(z)$ is therefore made up of broadband steering vectors, and the gain vector $\boldsymbol{f}(z)$ contains the transfer functions $f_\ell(z)$, $\ell = 1, \ldots, L$ that should be imposed on the L sources at the beamformer output. Both quantities are of the same dimensions as in the narrowband case, but are now

functions of the complex variable z. Writing the beamformer output as $y[n] = \sum_\nu \mathbf{w}^{\mathrm{H}}[-\nu]\mathbf{x}[n-\nu]$ allows the broadband LCMV problem to be formulated as [25]

$$\min_{\boldsymbol{w}(z)} \oint_{|z|=1} \boldsymbol{w}^{\mathrm{P}}(z)\boldsymbol{\mathcal{R}}(z)\boldsymbol{w}(z)\frac{\mathrm{d}z}{z} \qquad \text{s.t.} \quad \boldsymbol{\mathcal{C}}(z)\mathbf{w}(z) = \boldsymbol{f}(z)\ , \tag{18}$$

where $\boldsymbol{\mathcal{R}}(z)$ is the CSD matrix of $\mathbf{x}[n]$. The evaluation of (18) at a single frequency Ω_0 leads back to the narrowband formulation via the substitution $z = \mathrm{e}^{\mathrm{j}\Omega_0}$.

The solution to the broadband LCMV problem can be found as the equivalent of the Capon beamformer $\boldsymbol{w}_{\mathrm{opt}} = \boldsymbol{\mathcal{R}}^{-1}(z)\boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)\{\boldsymbol{\mathcal{C}}(z)\boldsymbol{\mathcal{R}}^{-1}(z)\boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)\}^{-1}\boldsymbol{f}(z)$, which is a direct extension of the narrowband formulation. In order to access this solution, the inversion of the parahermitian matrices $\boldsymbol{\mathcal{R}}(z)$ and subsequently $\boldsymbol{\mathcal{C}}(z)\boldsymbol{\mathcal{R}}^{-1}(z)\boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)$ can be accomplished via PEVDs [59]. Once factorized, the resulting paraunitary matrices are straightforward to invert, and it remains to invert the individual eigenvalues; for this, recall the comment on analytic functions as divisors in Section II. Alternatively, to avoid the nested matrix inversions of the Capon beamformer and to exploit iterative schemes for their general numerical robustness, an iterative unconstrained optimization can be performed via a broadband PEVD-based generalized sidelobe canceller (GSC), whereby with respect to Fig. 6(a), the quiescent beamformer is $\boldsymbol{w}_{\mathrm{q}}(z) = \boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)\{\boldsymbol{\mathcal{C}}(z)\boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)\}^{-1}\boldsymbol{f}(z)$. The pseudo-inverse of a polynomial matrix for the quiescent solution can again be obtained via a PEVD of the parahermitian term $\boldsymbol{\mathcal{C}}(z)\boldsymbol{\mathcal{C}}^{\mathrm{P}}(z)$ [59]. Furthermore, its subspace decomposition also reveals the nullspace of $\boldsymbol{\mathcal{C}}(z)$ that can be used to define the columns of the blocking matrix $\boldsymbol{\mathcal{B}}(z)$ such that $\boldsymbol{\mathcal{C}}(z)\boldsymbol{\mathcal{B}}(z) = \mathbf{0}$. It only remains to operate a vector $\boldsymbol{w}_{\mathrm{a}}(z)$ of $(M-L)$ adaptive filters on the output of the blocking matrix to complete the optimization of this PEVD-based GSC. Note that the overall response of the beamformer is $\boldsymbol{w}(z) = \boldsymbol{w}_{\mathrm{q}}(z) - \boldsymbol{\mathcal{B}}(z)\boldsymbol{w}_{\mathrm{a}}(z)$.

Key Statement

Using polynomial matrix notations and the PEVD, narrowband approaches, such as the Capon beamformer or the GSC, can be directly extended to the broadband case.

Note that compared to the narrowband GSC in Fig. 6(a), all quantities have retained their dimensions but are now functions of z. It now remains to set the polynomial orders of the different GSC components for an implementation. The quiescent vector $\boldsymbol{w}_{\mathrm{q}}(z)$ depends on the constraint formulation, and its order J_1 determines the accuracy of the fractional delay filters. The order J_2 of the blocking matrix $\boldsymbol{\mathcal{B}}(z)$ needs to be sufficiently high such that no source signal components covered by the constraint equation leak into the adaptive part $\boldsymbol{w}_{\mathrm{a}}(z)$. The order J_3 of the latter has to be sufficient to minimize the power of the output, $y[n]$. Thus, unlike in the TDL-based broadband beamformer case, the orders of the components are somewhat decoupled.

If the optimization of the adaptive part is addressed by inexpensive least-squares-type algorithms, the computational cost in both the TDL and the PEVD-based approaches is governed by the blocking matrix. In the TDL-based case, it requires $T^2(M^2 - ML)$ multiplications and additions, while the PEVD-based blocking matrix only expends $J_2(M^2 - ML)$ such operations. With typically $J_2 \approx T$, the PEVD-based realization is less expensive by a factor of approximately the length of the TDL, T.

Numerical Example. A linear array with $M = 8$ elements spaced by half the wavelength of the highest frequency component has a look direction towards $\vartheta = 30°$, which is protected by a constraint. Three 'unknown' interferers with directions $\vartheta = \{-40°, -10°, 80°\}$ are active over a frequency range of $0.2 \leq \Omega/\pi \leq 0.9$ at -20 dB SINR, and need to be adaptively suppressed. The data is further corrupted by spatially and temporally white additive noise 50 dB below the signal levels of the interferers. A TDL-based GSC operates with a TDL length of $T = 175$. For a PEVD-based GSC, the adaptive filter uses the same temporal dimension $J_3 = T$, but to match the MSE performance of the TDL-based version, a length of $J_1 = 51$ for the fractional delay filters in the quiescent beamformer and a temporal dimension of $J_2 = 168$ for the blocking matrix suffices. The adaptive filter is adjusted by a normalized least mean squares algorithm [54]. Note that $J_2 < T$. Overall, per iteration, the PEVD-based GSC takes 12.3 kMACs, while the TDL-based GSC requires 3.46 MMACs, which is indeed more than a factor of T higher.

To evaluate the beamformer performance, we determine the gain response or directivity pattern of the beamformer by probing the adapted overall beamformer response by sweeping a broadband steering vector $\boldsymbol{a}_i(z)$ across a set of angles $\{\varphi_i\}$ with a corresponding delay profile. For the directivity pattern, the angle-dependent transfer function $\mathcal{G}(z, \varphi_i) = \boldsymbol{w}^{\mathrm{P}}(z)\boldsymbol{a}_i(z)$ can be evaluated on the unit circle. For the PEVD-based GSC, this directivity pattern is shown in Fig. 7(a); the response (not displayed) for the TDL-based GSC is very similar. A difference can, however, be noted in the look direction, which in the case of the TDL-based GSC is protected by a number of point constraints along the frequency axis, as highlighted in Fig. 7(b). The gain response satisfies these point constraints, but it shows significant deviations from the ideal, flat response between the constrained frequencies. In contrast, the PEVD-based beamformer is based on a single broadband constraint equation, which shows a significantly lower deviation from the desired look direction gain. This is due to the formulation in the time-domain, which preserves spectral coherence. There are downsides, and the gain response will break down closer to $\Omega = \pi$ due to the imperfections that are inherent in fractional delay filters operating close to half the sampling rate.

Key Statement

The PEVD-based generalized sidelobe canceller can implement the constraint equation more easily and more precisely than the TDL-based version, and possesses a significantly lower complexity when addressing non-trivial constraints.

B. PEVD-Based Subband Coding

Data Representation and Task. Although this article addresses techniques for array signals, in many circumstances, mul-

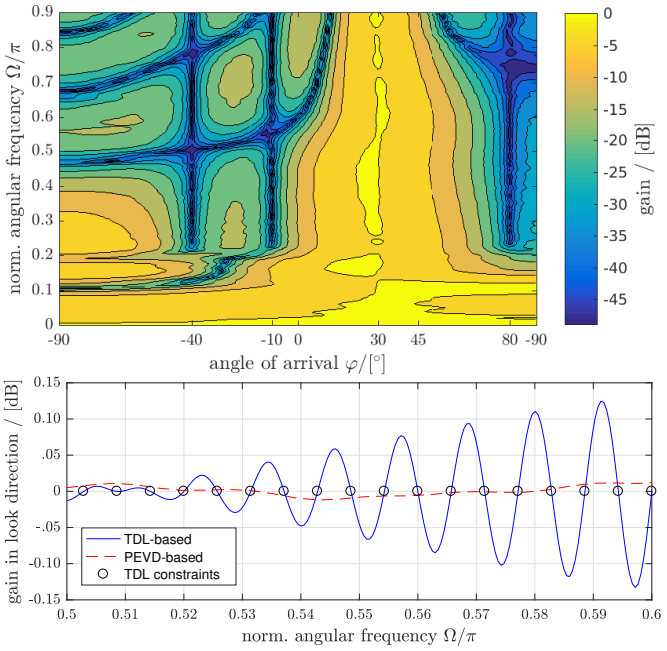


Fig. 7: (a) directivity patterns for adapted PEVD-based generalized sidelobe canceller, plotted above, and (b) the gain response in look direction ($\varphi = 30°$) for both the PEVD- and TDL-based GSC, plotted below.

tichannel signal representations are derived from a single-channel signal by demultiplexing [20], [60]–[62]. Let $x[\nu]$ be such a single channel signal. Demultiplexing by M and an implicit decimation operation by the same factor, or serial-to-parallel conversion, is performed to obtain a data vector $\mathbf{x}[n] = [x[nM],\ x[nM-1]\ ,\ldots,x[nM-M+1]]^{\mathrm{T}}$. This demultiplexed vector $\mathbf{x}[n]$ possesses the same form as the data vectors previously considered in Section III-A. While the amount and type of samples that are held in $\mathbf{x}[n]$ remain unaltered from those in $x[\nu]$, the representation in $\mathbf{x}[n]$ allows clever data reduction and coding schemes through filter bank-based processing, for which we will ultimately exploit the PEVD.

Principal Component Filter Banks and Subband Coding. Generally, we want to process the data $\mathbf{x}[n]$ through a transformation such that $\mathbf{y}[n] = \sum_\nu \mathbf{Q}^{\mathrm{H}}[-\nu]\mathbf{x}[n-\nu]$. Specifically, we wish this transformation to be lossless, i.e. for $\mathcal{Q}(z) \bullet\!\!-\!\!\circ\, \mathbf{Q}[n]$ to be paraunitary, such that a perfect reconstruction via $\mathbf{x}[n] = \sum_\nu \mathbf{Q}[\nu]\mathbf{y}[n-\nu]$ is possible. To the original unmultiplexed single channel signal, the transformation $\mathbf{Q}^{\mathrm{H}}$ represents the analysis filter bank, whereas the transformation $\mathbf{Q}$ implements the synthesis (reconstruction) filter bank [20]. The matrices $\mathcal{Q}^{\mathrm{P}}(z)$ and $\mathcal{Q}(z)$ are known as the analysis and synthesis polyphase matrices, respectively, and the paraunitarity of $\mathcal{Q}(z)$ guarantees perfect reconstruction of the overall filter bank system when operating back-to-back.

The polyphase matrix $\mathcal{Q}(z)$ can be designed to implement a series of lowpass, bandpass, and highpass filters to split the signal $x[\nu]$ into signal components with different spectral content. However, the filter bank $\mathcal{Q}(z)$ can also be signal-dependent. Chiefly amongst such systems are principal component, or optimum compaction, filter banks (PCFBs) which aim to assign as much power of $x[\nu]$ into as few successive components of $\mathbf{y}[n]$ as possible. The purpose of this is to discard some components of $\mathbf{y}[n]$, thus producing a lower-dimensional representation of the data. A closely related task is subband coding, where a quantization is performed on $\mathbf{y}[n]$ rather than on $x[\nu]$. Thus, a higher bit resolution is dedicated to those subbands of $\mathbf{y}[n]$ that possess higher power. By not increasing the overall number of bits w.r.t. $x[\nu]$, the judicious distribution of the coding effort results in an increase in the coding gain measure: the ratio between the arithmetic and geometric means of the variances of the subband signals in $\mathbf{y}[n]$ [60]. A coding gain greater than one can be exploited as an increased signal to quantization noise power under constant word length, or in terms of a reduction in the number of bits required for quantization while retaining the same quality for the quantized signals.

Optimum Coding Gain and PEVD. To maximize the coding gain under the constraint of paraunitarity of $\mathcal{Q}(z)$, two necessary and sufficient conditions of $\mathbf{y}[n]$ have been identified [60]: (i) the subband signals in $\mathbf{y}[n]$ must be strongly decorrelated, such that $\mathcal{R}_y(z)$ is diagonal; and (ii) they must be spectrally majorized, such that for the elements $S_m(z)$ along its diagonal, on the unit circle, we have $S_m(\mathrm{e}^{\mathrm{j}\Omega}) \geq S_{m+1}(\mathrm{e}^{\mathrm{j}\Omega})\ \forall\Omega$ and $m = 1,\ldots,(M-1)$. Due to Parseval's theorem, this implies that the powers of the subband signals are also ordered in a descending fashion. While under the paraunitary constraint, this does not change the arithmetic mean, it minimizes the geometric mean of the subband variances and thus maximizes the coding gain. Optimum subband coders $\mathbf{Q}(z)$ have been derived for the case of a zeroth order filter bank, where they reduce to the Karhunen-Loève transform (KLT), and for the infinite-order filter bank case [60], [62]. Executing the PEVD, described in Section IV-B, on $\mathbf{R}[\tau]$ leads to $\mathbf{R}_y[\tau] = \mathbf{\Lambda}[\tau]$ that directly satisfies the above conditions, and thus provides a solution to a subband coder of finite order [27], whose theoretical evaluation had otherwise eluded the research community except for the case of $M = 2$ [61].

As discussed in Section IV-B, at each SBR2-algorithm iteration, the parameters of the elementary paraunitary operator are selected such that the most dominant cross-covariance term of the input space-time covariance matrix is zeroed. There are two problems with this "greedy" optimization approach: (i) cross-correlation energies spread amongst subbands of weakest power can, in general, be ignored, which limits the extent to which spectral majorization is performed; (ii) there is a stronger tendency to annihilate cross-correlations due to noise in powerful subbands rather than true cross-terms related to weak subbands, which causes a degradation in strong decorrelation performance. The coding-gain variant of SBR2, namely SBR2C [27], alleviates these problems because it uses a cost function based on the coding-gain measure, which is proportionately equally receptive to cross-correlations between any of the subbands.

Numerical Example. Consider a signal $x[\nu]$ described by a fourth-order autoregressive model [27], [35]; its PSD $\mathcal{S}(\mathrm{e}^{\mathrm{j}\Omega})$ is shown in Fig. 8(top). Demultiplexing by $M = 4$ produces a pseudo-circulant matrix $\mathcal{R}(z)$ whose analytic eigenvalues

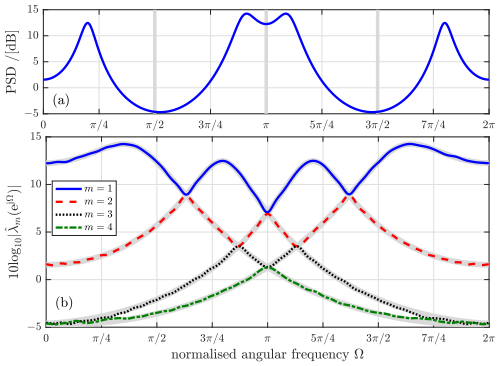


Fig. 8: PSD of input $x[\nu]$ (top) and eigenvalues (bottom) extracted by the SMD algorithm for the subband coding problem; the $M = 4$ times folded PSD of the input signal is underlaid in grey.

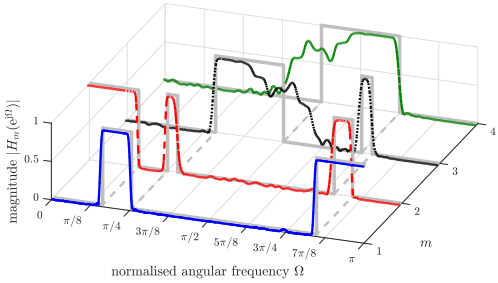


Fig. 9: Magnitude responses $|H_m(e^{j\Omega})|$, $m = 1, \ldots, 4$, of the $M = 4$ channel filter bank equivalent to the polyphase analysis matrix $\mathcal{Q}(z)$, with the theoretical PCFB of infinite order underlaid in grey.

$\mathcal{S}_m(\Omega) = \mathcal{S}(e^{j(\Omega/M - 2\pi(m-1))})$, $m = 1, \ldots, M$, are 8π-periodic, modulated versions of this PSD [20]. For $0 \leq \Omega \leq 2\pi$, they are depicted as grey underlaid curves in Fig. 8(bottom). Although the 8π periodicity of these functions means that $\mathcal{R}(z)$ has no analytic eigenvalues, an estimated CSD matrix $\hat{\mathcal{R}}(z)$, here based on 10^4 samples of $x[\nu]$, does possess an analytic EVD due to the perturbation by the estimation error [24]. Applying the SMD [27] to $\hat{\mathcal{R}}(z)$ will generate a strongly decorrelated signal vector $\mathbf{y}[n]$ via a paraunitary operation $\mathcal{Q}(z)$. The eigenvalues $\hat{\lambda}_m(e^{j\Omega})$ extracted by the SMD algorithms are also shown in Fig. 8(bottom). These closely match the folded PSD of $x[\nu]$ underlaid in grey, but are spectrally majorized.

Interpreting $\mathcal{Q}(z)$ as a polyphase analysis matrix, the associated four-channel filter bank is characterized in Fig. 9. The theoretically optimum infinite-order PCFB [60] is also shown. These are obtained by assigning every demultiplexed frequency component of $x[\nu]$ to one of four filters, in descending magnitude. This yields a binary mask in the Fourier domain, which would require the implementation of infinite sinc functions in the time-domain. In contrast, the finite order filters computed by the SMD algorithm, each derived from an eigenvector in $\mathcal{Q}(z)$ corresponding to the eigenvalues in Fig. 8, very closely approximate the PCFB, except where the input PSD is small and arguably unimportant.

Ensemble Results. To demonstrate the wider benefit of the proposed subband coder design, a randomly generated ensemble of 100 moving average processes of order 14 (MA(14)) produces signals $x[\nu]$ that are demultiplexed by $M = 4$. For each ensemble probe, the space-time covariance matrix is estimated from 2^{11} samples of $x[\nu]$ as a basis for the subband coder design. In order to average subband coding results across this ensemble, we normalize the coding gain obtained for each instance in the ensemble by the maximum coding gain of the infinite-order PCFB; the latter can be derived from each of the MA(14) processes [27]. This ensemble-averaged normalized coding gain verses the order of the polynomial matrix $\mathcal{Q}(z)$ is shown in Fig. 10. The figure shows results for the KLT, the approximate EVD (AEVD) algorithm in [52], as well as for SBR2C and SMD. The KLT is the optimum zeroth order subband coder. The AEVD algorithm is a fixed order technique that aims to generate a polynomial EVD but without proven convergence. Note that, like the AEVD algorithm, the SMD algorithm for zeroth order systems (i.e. length one polynomials) reduces to an ordinary EVD that is equivalent to the KLT and optimum for narrowband source signals — as shown in the figure. Both SBR2C and SMD converge towards the optimum performance of an infinite-order PCFB as the polynomial order of $\mathcal{Q}(z)$ increases. This is indeed what would be expected since the PEVD is effectively the broadband generalization of the KLT. Due its specific targeting of the coding gain and the resulting enhanced spectral majorization, SBR2C here outperforms SMD, and thus provides a highly useful trade-off between polynomial order and coding gain.

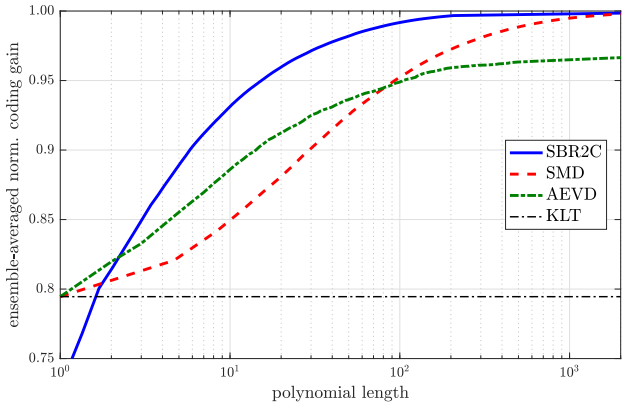


Fig. 10: Averaged normalized coding gain in dependence of the length (order plus one) of $\mathcal{Q}(z)$ for an ensemble of random MA(14) processes and for the case of demultiplexing with $M = 4$.

Key Statement

Hitherto, algorithms for $M > 2$-channel paraunitary filter banks for subband coding were suboptimal. PEVD-designed M-channel filter banks now closely approximate the ideal system.

C. Polynomial Subspace Speech Enhancement

Speech enhancement is important for applications involving human-to-human communications, such as hearing aids

and telecommunications and human-to-machine interactions, like robot audition, voice-controlled systems and automatic speech recognition (ASR). These speech signals are often captured by multiple microphones, commonly found in many devices today, and provide opportunities for spatial processing. Moreover, speech signals captured by different microphones naturally exhibit temporal correlations, especially in reverberant acoustic environments. This section will show that it is advantageous to use PEVD algorithms to capture and process these spatio-temporal correlations, thus preserving spectral coherence. A more comprehensive treatment with listening examples and code is available [28].

Multichannel Reverberant Signal Model. Consider a scenario where there is a single speaker $s[n]$, an array of microphones, and uncorrelated background noise $\mathbf{v}[n]$. The speech propagates from the source to each microphone m through the channels with acoustic impulse responses (AIRs), $a_{\ell,m}[n]$, assumed to be time-invariant. The AIR models the direct path propagation of the speech signal from the speaker to the microphone as well as reverberation due to multipath reflections from objects and walls in enclosed rooms. Background noise is then added to each microphone. The signal models in Section III-A with $\ell = 1$ can describe this situation. Across M microphones, the signal vector $\mathbf{x}[n] \in \mathbb{C}^M$ is used to compute the space-time covariance matrix $\mathbf{R}_{\mathbf{x}}[\tau]$ in (6) and its z-transform $\mathcal{R}_{\mathbf{x}}(z)$ in (7).

Exploiting the reverberation model in [63], the early reflections in the AIR represent closely spaced distinct echoes that perceptually reinforce the direct path component and may improve speech intelligibility in certain conditions. On the other hand, the late reflections in the AIR consist of randomly distributed small amplitude components and the associated late reverberant signal components are commonly assumed to be mutually uncorrelated with the direct path and early signal components [28], [63]. Thus, (7) can be written as

$$\begin{aligned}\mathcal{R}_{\mathbf{x}}(z) &= \tilde{\boldsymbol{a}}(z)\tilde{\boldsymbol{a}}^{\mathrm{P}}(z)\imath_s(z) + \mathcal{R}_{\boldsymbol{\ell}}(z) + \mathcal{R}_{\mathbf{v}}(z) \qquad (19)\\ &= \mathcal{R}_{\tilde{\mathbf{s}}}(z) + \mathcal{R}_{\tilde{\mathbf{v}}}(z)\ , \qquad (20)\end{aligned}$$

where $\imath_s(z) \bullet\!\!-\!\!\circ\ r_s[\tau]$ and $r_s[\tau]$ is the auto-correlation sequence of the source. The space-time covariance matrices of the late reverberation $\mathcal{R}_{\boldsymbol{\ell}}(z)$ is modelled as a spatially diffuse field. This is combined with the space-time covariance of the ambient noise $\mathcal{R}_{\mathbf{v}}(z)$ to form $\mathcal{R}_{\tilde{\mathbf{v}}}(z)$. The channel polynomial vector is $\tilde{\boldsymbol{a}}(z) = [\tilde{a}_1(z), \ldots, \tilde{a}_M(z)] \in \mathbb{C}^M$, where $\tilde{a}_m(z)$ is a polynomial obtained by taking the z-transform of the direct path and early reflections in the AIR from the source to the mth microphone, i.e., $\tilde{a}_m(z) = \sum_{i=0}^{I} \tilde{a}_m[i]z^{-i}$, dropping ℓ for brevity.

PEVD-based Speech Enhancement. The PEVD of (20) decomposes the polynomial matrix into

$$\mathcal{R}_{\mathbf{x}}(z) = \begin{bmatrix}\mathcal{U}_{\tilde{s}}(z) \mid \mathcal{U}_{\tilde{v}}(z)\end{bmatrix} \left[\begin{array}{c|c} \boldsymbol{\Lambda}_{\tilde{s}}(z) & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{\Lambda}_{\tilde{v}}(z) \end{array}\right] \left[\begin{array}{c} \mathcal{U}^{\mathrm{P}}_{\tilde{s}}(z) \\ \hline \mathcal{U}^{\mathrm{P}}_{\tilde{v}}(z) \end{array}\right], \qquad (21)$$

where $\{.\}_{\tilde{s}}$ and $\{.\}_{\tilde{v}}$ are associated with the signal-plus-noise (or simply signal) and noise-only (or simply noise) subspaces, respectively. Unlike some speech enhancement approaches, the proposed method does not use any noise or relative transfer function (RTF) estimation algorithms since the strong decorrelation property of the PEVD implicitly orthogonalizes the subspaces across all time lags in the range of τ. Consequently, speech enhancement can be achieved by combining components in the signal subspace while nulling components residing in the noise subspace.

The paraunitary $\mathcal{U}(z)$ is a lossless filter bank with an all-pass frequency response. This implies that $\mathcal{U}(z)$ can only distribute spectral power among channels and not change the total signal and noise power over all subspaces. The eigenvector filter bank is used to process the microphone signals using

$$\mathbf{y}[n] = \sum_{\nu} \mathbf{U}^{\mathrm{H}}[-\nu]\mathbf{x}[n-\nu]\ , \qquad (22)$$

where $\mathbf{U}[n] \ \circ\!\!-\!\!\bullet\ \mathcal{U}(z) \in \mathbb{C}^{M\times M}$. Since the polynomial eigenvector matrix $\mathcal{U}(z)$ is constructed from a series of delay and unitary matrices, each filter $\mathbf{u}_m[n]$ has a filter-and-sum structure.

For a single source, the signal subspace has a dimension of 1. Therefore, the enhanced signal can be extracted from the first channel of the processed outputs $\mathbf{y}[n]$. The enhanced output $y_1[n]$, associated with the signal subspace, comprises mainly speech components, originally distributed over all microphones but now summed coherently. In contrast, the noise subspace is dominated by ambient noise and the late reverberation in the acoustic channels. The orthogonality between subspaces is a result of strong decorrelation, expressed as $\mathcal{R}_{\mathbf{y}}(z) = \boldsymbol{\Lambda}(z)$, where $\mathcal{R}_{\mathbf{y}}(z) \bullet\!\!-\!\!\circ\ \mathbf{R}_{\mathbf{y}}[\tau]$ is computed from $\mathbf{R}_{\mathbf{y}}[\tau] = \mathbb{E}\{\mathbf{y}[n]\mathbf{y}^{\mathrm{H}}[n-\tau]\}$.

In practice, assuming quasi-stationarity, the speech signals are processed frame by frame such that $\mathbf{R}_{\mathbf{x}}[\tau]$ in (6) can be recursively estimated. Additionally, the two-sided z-transform $\mathcal{R}_{\mathbf{x}}(z)$ in (7) can be approximated by some truncation window W, which determines the extent of the supported temporal correlation of the speech signal. The time-domain PEVD algorithms, such as SBR2 and SMD, are used to compute (21) because they preserve spectral coherence of the speech signals and do not introduce audible artefacts. The proposed algorithm can also cope in noise-only and reverberation-only scenarios, as explored in [28]. Experimental results are next presented to show these general principles applied for a specific case.

Experimental Setup. Anechoic speech signals, which were sampled at 16 kHz, were taken from the TIMIT corpus [64]. AIR measurements and babble noise recordings for the $M = 3$ channel 'mobile' array were taken from the ACE corpus [65]. The ACE Lecture Room 2 has a reverberation time T_{60} of 1.22 s. For each Monte-Carlo simulation, 50 trials were conducted. In each trial, sentences from a randomly selected speaker were concatenated to have 8 to 10 s duration. The anechoic speech signals were then convolved with the AIRs for each microphone before being corrupted by additive noise. The SNRs ranged from -10 dB to 20 dB.

Comparative Algorithms. PEVD-based enhancement can be compared against other algorithms such as the oracle multichannel Wiener filter (OMWF), weighted power minimum distortionless response (WPD) and two subspace approaches,

multichannel subspace (MCSUB) [66] and coloured subspace (COLSUB) [67], which use an EVD and a generalized eigenvalue decomposition (GEVD), respectively. Furthermore, unlike the PEVD approach, noise estimation is required for GEVD. The OMWF is based on the concatenation of a minimum variance distortionless response (MVDR) beamformer followed by a single-channel Wiener filter. The OMWF provides an ideal performance upper-bound since it uses complete prior knowledge of the clean speech signal, based on [68] where the filter length is 80. Practical multichannel Wiener filters, which rely on the relative transfer function and noise estimation algorithms, do not perform as well as OMWF, and comparative results can be found in [28]. WPD is an integrated method for noise reduction and dereverberation [69]. Ground truth DOA is provided to compute the steering vector for WPD to avoid signal direction mismatch errors. PEVD does not use any knowledge of the speech, DOA, and array geometry.

Experiments presented here to illustrate comparative performance use PEVD parameters, chosen following [28], including $\delta = \sqrt{N_1/3} \times 10^{-2}$ denoting the threshold of the dominant off-diagonal column norm, where N_1 is the square of the trace-norm of $\mathbf{R_{xx}}(0)$, trim factor $\mu = 10^{-3}$ and $L = 500$ iterations. In all experiments, the frame size T and window W are set to 1600. With this parameter selection, correlations within 100 ms, which were assumed to include the direct path and early reflection components, were captured and used by the algorithm. The source corresponding to these experiments is available [28].

Evaluation Measures. Frequency-weighted segmental signal-to-noise ratio (FwSegSNR) can be used to evaluate noise reduction and normalized signal-to-reverberant ratio (NSRR) and Bark spectral distortion (BSD) for dereverberation [28]. To measure speech intelligibility and to account for processing artefacts, short-time objective intelligibility (STOI) can be used. These measures are computed for the signals before and after enhancement using the proposed and benchmark algorithms. The improvement Δ is reported. Positive Δ values show improvements in all measures except ΔBSD, for which a negative value indicates a reduction in spectral distortions.

Results and Discussions. An illustrative example based on a clean speech $s[n]$ corrupted by 5 dB babble noise in the reverberant ACE Lecture Room 2 is presented in Fig. 11. The spectrogram of the first microphone signal $x_1[n]$ shows temporal smearing due to reverberation and the addition of babble noise. Comparing the plots for $x_1[n]$ with the processed signals $y_1[n]$, the dotted cyan boxes in Fig. 11 qualitatively show the attenuation and some suppression of the babble noise and reverberation for PEVD and COLSUB. This is supported by Table I, which shows that PEVD significantly improves STOI and NSRR while coming second in FwSegSNR and BSD after COLSUB. Although COLSUB makes the most significant improvement in FwSegSNR, speech structures are destroyed during the processing, for example, the solid white boxes comparing $s[n]$ and $y_1[n]$ between 3.0 to 3.3 s and 4.2 to 4.7 s in Fig. 11, resulting in artefacts in the listening examples and the lowest improvement in STOI. OMWF, which uses complete knowledge of the clean speech signal, is the second best in STOI and slightly improves other metrics, similar to WPD, which uses the ground truth steering vector. MCSUB offers limited improvement. Listening examples also highlight that PEVD does not introduce audible processing artefacts into the enhanced signal [28].

TABLE I: Enhancement of a single reverberant speech in Lecture Room 2 and 5 dB ACE babble noise.

Algorithm	FwSegSNR	STOI	NSRR	BSD
Noisy	-10.9 dB	0.664	-7.57 dB	0.69 dB
OMWF	-11.1 dB	0.747	-7.42 dB	0.60 dB
MCSUB	-11.7 dB	0.711	-11.5 dB	0.93 dB
COLSUB	**-6.60 dB**	0.678	-7.90 dB	**0.35 dB**
PEVD	-8.21 dB	**0.750**	**-6.13 dB**	0.40 dB
WPD	-8.90 dB	0.723	-6.27 dB	0.45 dB

Results for the Monte-Carlo simulation involving 50 speakers in Lecture Room 2 and corrupted by -10 dB to 20 dB babble noise are shown in Fig. 12. For SNR$\leq$10 dB, COLSUB outperforms other algorithms in ΔFwSegSNR but gives the worst ΔSTOI. On the other hand, OMWF, designed to minimize speech distortion using knowledge of clean speech, performs the best in ΔSTOI but not in ΔFwSegSNR. This also reflects the fact that speech intelligibility may not necessarily be affected by noise levels, up to some limit. Despite not being given any information on the target speech, PEVD performs comparably to OMWF and ranks first in ΔNSRR and second in ΔFwSegSNR and ΔSTOI.

At 20 dB SNR, algorithms targeting reverberation like WPD perform better than noise reduction approaches. Similar to generalized weighted prediction error (GWPE) in the reverberation-only case in [28], WPD processes the reverberant signals aggressively by removing most early reflections but not the direct path and late reflections, as observed in the listening examples. Furthermore, WPD uses ground truth DOA to compute the ideal steering vector, leading to the best improvement in ΔBSD and ΔSTOI. Listening examples for PEVD indicate that the direct path and early reflections are retained in the enhanced signal in the first channel. The late reverberations, absent in the enhanced signal, are observed in the second and third channels because of orthogonality [28]. Even without additional information, PEVD performs comparably to WPD and ranks second in ΔNSRR and ΔSTOI.

Despite not being given knowledge of the DOA, target speech and array geometry, PEVD consistently ranks first for ΔNSRR and second in ΔSTOI and ΔFwSegSNR over the range of scenarios. Comprehensive results with listening examples and code for the noise-only, reverberation-only, and more noisy reverberant scenarios are available [28].

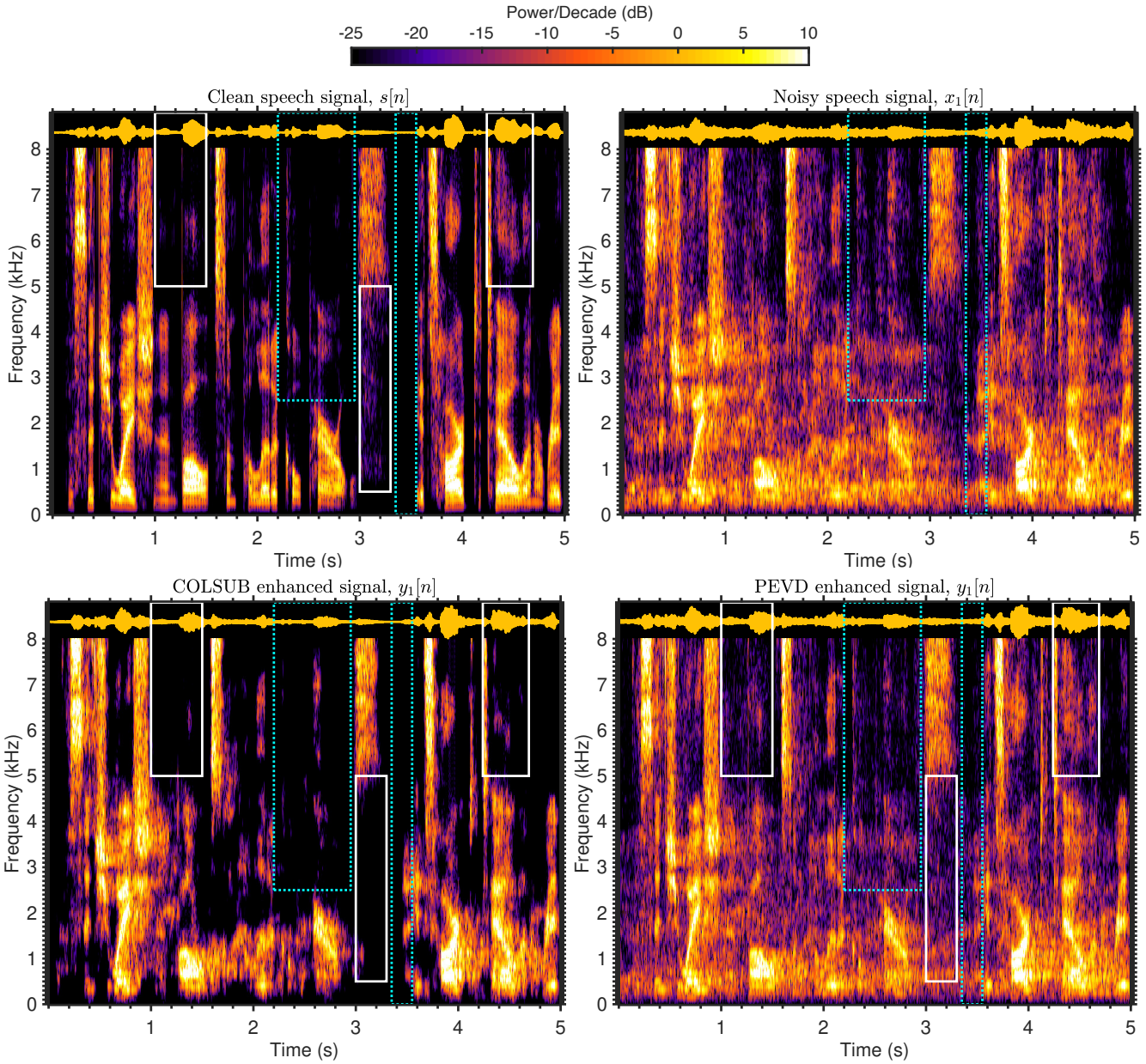


Fig. 11: Spectrograms, with corresponding time-domain signals above each, for the processing of a noisy reverberant speech example in ACE Lecture Room 2 and 5 dB babble noise. Dotted cyan boxes highlight noise and reverberation suppressed regions as a result of processing. Solid white boxes highlight regions where speech structures are lost using COLSUB but not PEVD processing. Listening examples are available [28].

Key Statement

PEVD-based speech enhancement consistently improves noise reduction metrics, speech intelligibility scores, and dereverberation measures over a wide range of acoustic scenarios. This blind and unsupervised algorithm requires no knowledge of the array geometry and does not use any channel or noise estimation algorithms but performs comparably to an oracle algorithm. More notably, due to the preservation of the spectral coherence using time-domain PEVD algorithms, the proposed algorithm does not introduce noticeable processing artefacts into the enhanced signal. Code and listening examples are provided in [28].

VI. Conclusions and Future Perspectives

This article has demonstrated the use of polynomial matrices to model broadband multichannel signals and the use of the PEVD to process them. Previous approaches of using tapped delay lines and short-time Fourier transforms do not lead to proper generalization of narrowband algorithms and are sub-optimal. Instead of only considering the instantaneous covariance matrix, the space-time covariance matrix has been proposed to completely capture the second-order statistics of multichannel broadband signals. Motivated by the optimum processing of narrowband signals using the EVD, i.e. for a single lag, the PEVD has been proposed to process broadband signals across a range of time lags. In most cases, an ana-

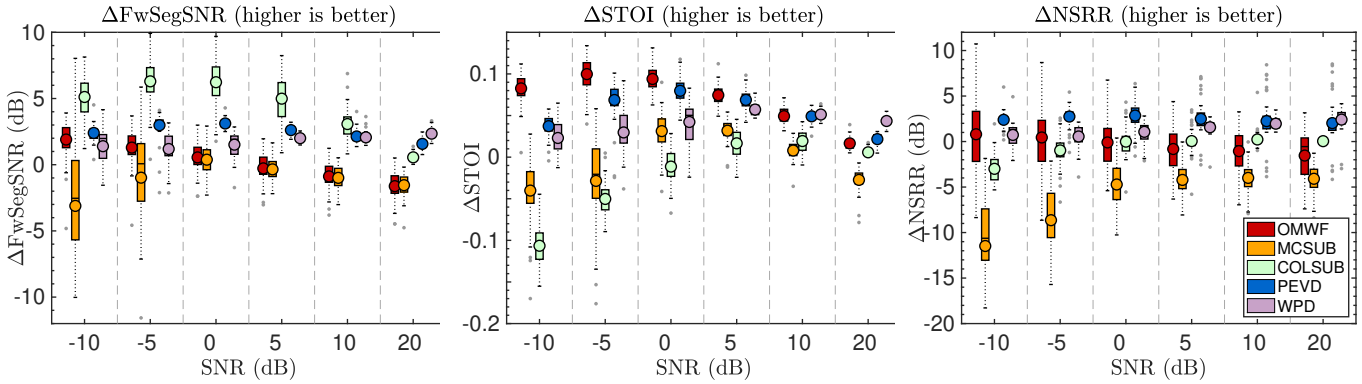


Fig. 12: Comparison of speech enhancement performance for recorded AIR and babble noise in ACE Lecture Room 2 with 1.22 s reverberation time.

lytic PEVD exists and can be approximated by polynomials using numerical algorithms, which tend to generate spectrally majorized eigenvalues and paraunitary eigenvectors.

PEVD-based processing for three example applications has been presented and is advantageous over state-of-the-art processing. The PEVD approach can implement the constraints more easily and precisely for adaptive broadband beamforming while achieving a lower complexity than the TDL-based approach. For multichannel subband coding, the PEVD design approximates the ideal optimal data encoding system and overcomes the previous issues with the more than two-channel case. The PEVD-based algorithm, which only uses the microphone signals, can consistently enhance speech signals without introducing any audible artefacts and performs comparably to an oracle algorithm, as observed in the listening examples. In addition to the applications presented in this article, PEVD is also successfully used for blind source separation [30], MIMO system design [26], source identification [31] and broadband DOA estimation [32].

Future Work. Similar extensions from EVD to an analytic or polynomial EVD can be undertaken for other linear algebraic operations, e.g., non-parahermition EVD, singular value decomposition (SVD), QR decomposition, generalized SVD (GSVD). Algorithms for the SVD and the QR decomposition have appeared but are yet without a theoretical foundation with respect to their existence. Powerful narrowband techniques such as independent component analysis (ICA) may find their polynomial equivalents. While a number of low-cost implementations have already emerged, algorithmic scalability is an area of active investigation. We hope that these theoretical and algorithmic developments will motivate the signal processing community to experiment with polynomial techniques, and take these beyond the successful application areas showcased in this article. Resources including code and demo pages are available [28], [70].

References

[1] G. Strang, *Linear Algebra and Its Application*, 2nd ed. Academic Press, 1980.

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

[3] S. Haykin and K. J. R. Liu, Eds., *Handbook on Array Processing and Sensor Networks*. John Wiley & Sons, Inc., 2010.

[4] N. S. Jayant and P. Noll, *Digital Coding of Waveforms Principles and Applications to Speech and Video*. Prentice Hall, 1984.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[6] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Upper Saddle River, NJ, USA: Prentice Hall, 1995.

[7] M. Moonen and B. De Moor, *SVD and Signal Processing, III: Algorithms, Architectures and Applications*. Elsevier Science Publishers B. V., 1995.

[8] H. L. Van Trees, *Optimal Array Processing. Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 2002.

[9] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, 1st ed. Academic Press, 2010.

[10] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, 1st ed. Prentice Hall, 2000.

[11] R. Klemm, *Space-time Adaptive Processing Principles and Applications*, ser. IEE Radar, Sonar, Navigation Avionics. IEE, London, UK, 1998, no. 9.

[12] A. Rao and R. Kumaresan, "On decomposing speech into modulated components," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 3, pp. 240–254, May 2000.

[13] S. Weiss and I. K. Proudler, "Comparing efficient broadband beamforming architectures and their performance trade-offs," in *Proc. IEEE Int. Conf. Digital Signal Process. (DSP)*, Jul. 2002, pp. 417–424.

[14] W. Kellermann and H. Buchner, "Wideband algorithms versus narrowband algorithms for adaptive filtering in the DFT domain," in *Proc. Asilomar Conf. on Signals, Syst. & Comput.*, 2003, pp. 1–5.

[15] Y. Avargel and I. Cohen, "System identification in the short-time Fourier transform domain with crossband filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 4, pp. 1305–1319, May 2007.

[16] B. Widrow, P. Mantey, L. Griffiths, and B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, 1967.

[17] R. T. Compton, Jr., "The relationship between tapped delay-line and FFT processing in adaptive arrays," *IEEE Trans. Antennas Propag.*, vol. 36, no. 1, pp. 15–26, Jan. 1988.

[18] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay [FIR/all pass filters design]," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.

[19] T. Kailath, *Linear Systems*, 1st ed. Englewood Cliffs, NJ 07632, USA, 1980.

[20] P. P. Vaidyanathan, *Multirate Systems and Filters Banks*, 1st ed. New Jersey, USA: Prentice Hall, 1993.

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

[22] S. Icart and P. Comon, "Some properties of Laurent polynomial matrices," in *IMA Int. Conf. on Math. in Signal Process.*, Dec. 2012, pp. 1–4.

[23] S. Weiss, J. Pestana, and I. K. Proudler, "On the existence and uniqueness of the eigenvalue decomposition of a para-Hermitian matrix," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2659–2672, May 2018.

[24] S. Weiss, J. Pestana, I. K. Proudler, and F. K. Coutts, "Corrections to "On the existence and uniqueness of the eigenvalue decomposition of a para-Hermitian matrix"," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6325–6327, Dec. 2018.

[25] S. Weiss, S. Bendoukha, A. Alzin, F. K. Coutts, I. K. Proudler, and J. Chambers, "MVDR broadband beamforming using polynomial matrix techniques," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2015, pp. 839–843.

[26] R. Brandt and M. Bengtsson, "Wideband MIMO channel diagonalization in the time domain," in *Proc. Int. Symp. on Personal, Indoor and Mobile Radio Commun.*, 2011, pp. 1958–1962.

[27] S. Redif, J. G. McWhirter, and S. Weiss, "Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5253–5264, Nov. 2011.

[28] V. W. Neo, C. Evers, and P. A. Naylor, "Enhancement of noisy reverberant speech using polynomial matrix eigenvalue decomposition," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 3255–3266, Oct. 2021.

[29] J. Corr, J. Pestana, S. Weiss, I. K. Proudler, S. Redif, and M. Moonen, "Investigation of a polynomial matrix generalised EVD for multi-channel Wiener filtering," in *Proc. Asilomar Conf. on Signals, Syst. & Comput.*, 2016, pp. 1354–1358.

[30] S. Redif, S. Weiss, and J. G. McWhirter, "Relevance of polynomial matrix decompositions to broadband blind signal separation," *Signal Process.*, vol. 134, pp. 76–86, May 2017.

[31] 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 Process. for Defence Conf. (SSPD)*, 2017, pp. 1–5.

[32] W. Coventry, C. Clemente, and J. Soraghan, "Enhancing polynomial MUSIC algorithm for coherent broadband sources through spatial smoothing," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2017, pp. 2448–2452.

[33] A. Oppenheim and R. W. Schafer, *Digital Signal Processing*, 2nd ed. Prentice Hall, 1993.

[34] B. Girod, R. Rabebstein, and A. Stenger, *Signals and Systems*. John Wiley & Sons, Inc., 2001.

[35] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw-Hill, 1991.

[36] I. Gohberg, P. Lancaster, and L. Rodamn, *Matrix Polynomials*, 2nd ed. SIAM, 2009.

[37] V. Kučera, *Analysis and Design of Discrete Linear Control Systems*. Prentice Hall, 1991.

[38] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, New Jersey, USA: Prentice Hall, 1983.

[39] F. Rellich, "Störungstheorie der Spektralzerlegung. I. Mitteilung. Analytische Störung der isolierten Punkteigenwerte eines beschränkten Operators," *Mathematische Annalen*, vol. 113, pp. DC–DCXIX, 1937. [Online]. Available: https://eudml.org/doc/159886

[40] M. Davis, "Factoring the spectral matrix," *IEEE Trans. Autom. Control*, vol. 8, no. 4, pp. 296–305, Oct. 1963.

[41] P. Ltd., "The polynomial toobox," 2005. [Online]. Available: http://www.polyx.com

[42] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[43] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 40, no. 8, pp. 1862–1875, Aug. 1992.

[44] F. Rellich, *Perturbation Theory of Eigenvalue Problems*. Gordon and Breach Science Publishers Inc., 1969.

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

[46] G. Barbarino and V. Noferini, "On the Rellich eigendecomposition of para-hermitian matrices and the sign characteristics of $*$-palindromic matrix polynomials," 2022. [Online]. Available: https://arxiv.org/abs/2211.15539

[47] S. Redif, S. Weiss, and J. G. McWhirter, "Sequential matrix diagonalisation algorithms for polynomial EVD of para-Hermitian matrices," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 81–89, Jan. 2015.

[48] F. K. Coutts, J. Corr, K. Thompson, S. Weiss, I. K. Proudler, and J. G. McWhirter, "Memory and complexity reduction in para-Hermitian matrix manipulations of PEVD algorithms," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2016, pp. 1633–1637.

[49] S. Kasap and S. Redif, "Novel field-programmable gate array architecture for computing the eigenvalue decomposition of para-hermitian polynomial matrices," *IEEE Trans. Very Large Scale Integration Syst.*, Mar. 2014.

[50] F. K. Coutts, I. K. Proudler, and S. Weiss, "Efficient implementation of iterative polynomial matrix EVD algorithms exploiting structural redundancy and parallelisation," *IEEE Trans. Circuits Syst. I*, vol. 66, no. 12, pp. 4753–4766, Dec. 2019.

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

[52] A. Tkacenko, "Approximate eigenvalue decomposition of para-Hermitian systems through successive FIR paraunitary transformations," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2010, pp. 4074–4077.

[53] M. Tohidian, H. Amindavar, and A. M. Reza, "A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices," *EURASIP J. on Applied Signal Process.*, vol. 1, no. 93, pp. 1–16, Apr. 2013.

[54] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Prentice Hall, 1991.

[55] K. M. Buckley, "Spatial/spectral filtering with linear constrained minimum variance beamformers," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 3, pp. 249–266, Mar. 1987.

[56] W. Liu and S. Weiss, *Wideband Beamforming — Concepts and Techniques*. John Wiley & Sons, Inc., 2010.

[57] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, 2005.

[58] K. M. Buckley, "Broad-band beamforming and the generalized sidelobe canceller," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 1322–1323, Oct. 1986.

[59] S. Weiss, A. P. Millar, and R. W. Stewart, "Inversion of para-Hermitian matrices," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2010, pp. 447–451.

[60] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1528–1543, Jun. 1998.

[61] B. Xuan and R. Bamberger, "FIR principal component filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 930–940, Apr. 1998.

[62] A. Kirac and P. P. Vaidyanathan, "Theory and design of optimum FIR compaction filters," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 903–919, Apr. 1998.

[63] P. A. Naylor and N. D. Gaubitch, Eds., *Speech Dereverberation*. Springer-Verlag, 2010.

[64] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, and V. Zue, "TIMIT acoustic-phonetic continuous speech corpus," Linguistic Data Consortium (LDC), Philadelphia, USA, Corpus LDC93S1, 1993.

[65] J. Eaton, N. D. Gaubitch, A. H. Moore, and P. A. Naylor, "Estimation of room acoustic parameters: the ACE challenge," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 10, pp. 1681–1693, Oct. 2016.

[66] F. Jabloun and B. Champagne, "A multi-microphone signal subspace approach for speech enhancement," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2001, pp. 205–208.

[67] Y. Hu and P. C. Loizou, "A subspace approach for enhancing speech corrupted by colored noise," *IEEE Signal Process. Lett.*, vol. 9, no. 7, pp. 204–206, Jul. 2002.

[68] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2230–2244, Sep. 2002.

[69] T. Nakatani and K. Kinoshita, "A unified convolutional beamformer for simultaneous denoising and dereverberation," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 903–907, Jun. 2019.

[70] S. Weiss, J. Corr, K. Thompson, J. G. McWhirter, and I. K. Proudler, "Polynomial EVD toolbox," 2014. [Online]. Available: http://pevd-toolbox.eee.strath.ac.uk