FIXED BEAMFORMER DESIGN USING POLYNOMIAL EIGENVALUE DECOMPOSITION

Vincent W. Neo ⦿, *Emilie d'Olne* ⦿, *Alastair H. Moore* ⦿, *Patrick A. Naylor*[†] ⦿

Department of Electrical and Electronic Engineering, Imperial College London, UK
{vincent.neo09, emilie.dolne16, alastair.h.moore, p.naylor}@imperial.ac.uk

ABSTRACT

Array processing is widely used in many speech applications involving multiple microphones. These applications include automatic speech recognition, robot audition, telecommunications, and hearing aids. A spatio-temporal filter for the array allows signals from different microphones to be combined desirably to improve the application performance. This paper will analyze and visually interpret the eigenvector beamformers designed by the polynomial eigenvalue decomposition (PEVD) algorithm, which are suited for arbitrary arrays. The proposed fixed PEVD beamformers are lightweight, with an average filter length of 114 and perform comparably to classical data-dependent minimum variance distortionless response (MVDR) and linearly constrained minimum variance (LCMV) beamformers for the separation of sources closely spaced by 5 degrees.

Index Terms— polynomial eigenvalue decomposition, microphone arrays, fixed beamformers, MVDR, LCMV

1. INTRODUCTION

Array processing is widely used in many speech applications involving multiple microphones. These applications include automatic speech recognition [1], robot audition [2], telecommunications and hearing aids [3]. Beamforming, or the design of a spatio-temporal filter for the array, combines signals from different microphones to extract the desired signal arriving from a specific direction [4]. Processing these extracted signals instead of the microphone signals usually improves the application performance.

The beamformer can be designed to be fixed or adaptive [5,6]. In fixed or data-independent designs, the array can have a specific spatial response by proper selection of the filter weights. Although fixed beamformers are limited in dynamic acoustical environments, they are still in use due to their simplicity, effectiveness and low complexity [1,7], which is particularly important for on-device processing with limited power and computation.

Data-dependent or adaptive beamformers rely on the statistical properties of the signals [8]. They can be designed to adapt to time-changing acoustical environments based on statistically optimal criteria such as maximum signal to noise ratio (SNR), minimum mean-square error (MMSE) and linearly constrained minimum variance (LCMV) [9, 10]. These have led to well-known beamformers such as the multi-channel Wiener filter (MWF) [11, 12], minimum variance distortionless response (MVDR) [5,13,14] and generalized sidelobe canceller (GSC) [15, 16]. In most approaches, the microphone signals are usually processed in the short-time Fourier transform domain. However, this approach divides the broadband speech into multiple narrowband signals, thus ignoring the correlation between different DFT bins and phase coherence across bands [17].

An alternative approach can be formulated in the z-domain using polynomial matrices computed from the microphone signals. Polynomial matrices can simultaneously capture the space, time and frequency correlations and are suitable for modelling multi-channel broadband signals. The polynomial matrices are processed using an iterative polynomial matrix eigenvalue decomposition (PEVD) algorithm such as second-order sequential best rotation (SBR2) [18, 19] and sequential matrix diagonalization (SMD) [20, 21] in the time-domain or [22] in the frequency-domain. PEVD algorithms have been found useful for speech enhancement [23, 24], source separation [25, 26], source localization [27, 28] and channel coding [29].

In [30, 31], polynomial MVDR and GSC have been developed by formulating and solving the well-known optimization problem in the z-domain using polynomial matrix techniques. In this paper, we will investigate fixed beamformers designed by PEVD for arbitrary arrays. During training, the learnt PEVD filterbanks are stored in a look-up table, and the entries are later retrieved and used directly for testing. The novel contributions of this paper are (i) an analysis of the fixed beamformer design using PEVD, (ii) a visual interpretation of the eigenvectors generated by the PEVD, and (iii) a comparison of the proposed approach with the classical MVDR and LCMV beamformers in noisy reverberant environments.

2. PROBLEM FORMULATION

2.1. Signal Model

The received signal at the q-th microphone for sample index n is

$$x_q(n) = \sum_{p=1}^{P} h_{p,q}(n) * s_p(n) + v_q(n) , \qquad (1)$$

where $h_{p,q}(n)$ represents the time-invariant room impulse response (RIR) from the p-th source to the q-th microphone, $s_p(n)$ is the p-th source signal, $v_q(n)$ represents the additive noise at the q-th microphone, and $*$ denotes the linear convolution operator. The noise signals are assumed to be zero-mean, uncorrelated with each other and the source signals. The data vector over Q microphones is $\mathbf{x}(n) = [x_1(n), \ldots, x_Q(n)]^T \in \mathbb{R}^Q$, where $[\cdot]^T$ is the transpose operator.

2.2. Polynomial Matrix Eigenvalue Decomposition

The space-time covariance matrix [18, 32], parameterized by time lag $\tau \in \mathbb{Z}$, is computed using

$$\mathbf{R}(\tau) = \mathbb{E}\{\mathbf{x}(n)\mathbf{x}^T(n - \tau)\} , \qquad (2)$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator over n. Each element, $r_{p,q}(\tau)$, is the correlation sequence between the p-th and q-th microphone signals. This produces auto- and cross-correlation sequences on the diagonals and off-diagonals, respectively.

The z-transform of (2),

$$\mathscr{R}(z) = \sum_{\tau=-\infty}^{\infty} \mathbf{R}(\tau) z^{-\tau} , \qquad (3)$$

denoted by $\mathbf{R}(\tau) \circ\!\!-\!\!\bullet \mathscr{R}(z)$, is a para-Hermitian polynomial matrix satisfying $\mathscr{R}(z) = \mathscr{R}^P(z) = \mathscr{R}^H(1/z^*)$, where $[\cdot]^*, [\cdot]^H, [\cdot]^P$ are the complex conjugate, Hermitian and para-Hermitian operators respectively. The para-Hermitian eigenvalue decomposition (EVD) of $\mathscr{R}(z) \in \mathbb{C}^{Q \times Q}$ in (3) is [32, 33]

$$\mathscr{R}(z) = \mathcal{U}(z)\,\mathbf{\Lambda}(z)\,\mathcal{U}^P(z) , \qquad (4)$$

where the columns of $\mathcal{U}(z) \in \mathbb{C}^{Q \times Q}$ are the polynomial eigenvectors and the elements on the diagonal matrix $\mathbf{\Lambda}(z) \in \mathbb{C}^{Q \times Q}$ are the polynomial eigenvalues. Iterative PEVD algorithms based on the SBR2 [18, 19] and SMD [20, 21] are used to approximate (4) by Laurent polynomial factors. Diagonalization of (4) is achieved using

$$\mathbf{\Lambda}(z) = \mathcal{U}^P(z)\,\mathscr{R}(z)\,\mathcal{U}(z) . \qquad (5)$$

3. BEAMFORMER DESIGNS USING POLYNOMIAL EVD

3.1. Polynomial Eigenvector Beamformer Design

In most applications involving physical signals, there is often coupling in space, time and frequency because of the wave equation [9], motivating the need for a decoupling approach such as the PEVD. The eigenvector $\mathcal{U}(z)$ performs a filter-and-sum operation on the microphone inputs to generate strongly decorrelated signals, i.e., output signals are not spatially correlated over a range of time lags.

The acoustic channel in (1) can be more compactly represented by $\mathbf{H}(n) \circ\!\!-\!\!\bullet \mathcal{H}(z)$, where each element is $h_{p,q}(n)$. As the sources propagate through the acoustic channels, the space-time polynomial matrix of the microphone signals, corrupted by spatially and temporally uncorrelated noise with equal power σ_v^2, is

$$\mathscr{R}_{\mathbf{x}}(z) = \mathcal{H}^P(z)\mathscr{R}_{\mathbf{s}}(z)\mathcal{H}(z) + \sigma_v^2 \mathbf{I} , \qquad (6)$$

where $\mathscr{R}_{\mathbf{x}}(z)$ and $\mathscr{R}_{\mathbf{s}}(z)$ are the space-time covariance matrices of the microphone and source signals, and $\mathbf{I}$ is the identity matrix.

Consider the case when source signals are generated by uncorrelated, zero-mean unit variance Gaussian processes $g_p(n) \in N(0,1)$ with the following cross-correlation sequence

$$r_{m,l}(\tau) = \mathbb{E}\{g_m(n)g_l(n-\tau)\} = \delta(m-l)\delta(\tau) . \qquad (7)$$

This implies that $\mathscr{R}_{\mathbf{s}}(z) = \mathbf{I}$. Further, assume that sources satisfy spectral majorization such that [21, 29]

$$\gamma_{g_{m-1}}(e^{j\Omega}) \geq \gamma_{g_m}(e^{j\Omega}) \quad \forall \Omega, \; m = 2, \dots, P , \qquad (8)$$

where $\gamma_{g_m}(e^{j\Omega})$ is the power spectral density of $g_m(n)$. Applying the PEVD and rearranging, (6) becomes
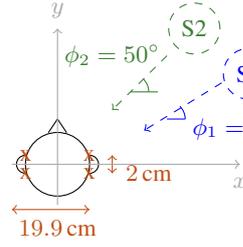
$$\mathbf{\Lambda}(z) - \sigma_v^2 \mathbf{I} = \mathcal{U}^P(z)\mathcal{H}^P(z)\mathcal{H}(z)\mathcal{U}(z) . \qquad (9)$$

This implies that the right hand side term in (9) must also achieve diagonalization, indicating that $\mathcal{U}^P(z)$ might decorrelate $\mathcal{H}(z)$. The eigenvector filterbank generated by the iterative PEVD algorithms satisfy the para-unitary or lossless condition [34] such that

$$\mathcal{U}^P(z)\mathcal{U}(z) = \mathcal{U}(z)\mathcal{U}^P(z) = \mathbf{I} . \qquad (10)$$

The first consequence of (10) implies that $\mathcal{U}(z)$ cannot change the total power over the subspaces but can only redistribute spectral

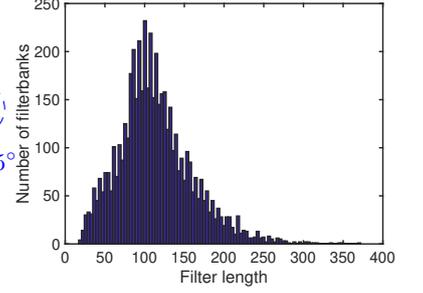(a) Setup for sources S1 and S2.   (b) Histogram of filter lengths.

Fig. 1: Experiment setup and trained filterbank lengths.

power among the microphones [18, 23]. This gives $\mathcal{U}(z)$ a stable and all-pass characteristic [18]. The second is the orthogonality such that the inner product between different columns of $\mathcal{U}(z)$ equals 0.

For this choice of source signals, i.e. white Gaussian, the PEVD filterbank is designed to primarily target only the acoustic channels. With a different selection of source signals, another set of $\mathcal{U}(z)$ will be generated to jointly decorrelate in space, time and frequency.

3.2. Broadband Steering Vectors and Beam Patterns

To understand the spatio-spectral processing of the designed beamformer, the response of the array to a plane wave is computed in Section 4. The broadband steering vector, which implements fractional delays, is defined as [35]

$$\boldsymbol{a}_\phi(z) = [1, a_2(z), \dots, a_Q(z)]^T , \qquad (11)$$

where $a_q(z) = \sum_n a_q(n)z^{-n}$, $a_q(n) = \mathrm{sinc}(nT_s - \Delta\tau_q)$, $\Delta\tau_q$ is the time delay from the source to the q-th microphone relative to the first reference microphone and T_s is the sampling period. The time delays are embedded in the angles of arrival ϕ and can be modelled explicitly using the array geometry [27]. In particular, a narrowband source signal with frequency Ω_0 arriving at the array is $e^{j\Omega_0 nT_s}\mathbf{a}_\phi(\Omega_0)$, where $\mathbf{a}_\phi(\Omega_0) = [1, \dots, e^{-j\Omega_0 \Delta\tau_Q}]^T$ is expressed in terms of phase shifts from the reference microphone.

The broadband array response can be computed using

$$\mathscr{B}(\phi, z) = \mathcal{U}^P(z)\boldsymbol{a}_\phi(z) . \qquad (12)$$

The array response at frequency Ω can be obtained by evaluating (12) on the unit circle such that

$$\mathbf{B}(\phi, \Omega) = \mathscr{B}(\phi, z)|_{z=e^{j\Omega}} . \qquad (13)$$

It is also common to reparameterize ϕ in (13) using wavenumber $\mathbf{k}$ for a more general representation. In this case, (13) is termed as the frequency-wavenumber response function [9]. In this paper, only the azimuth angle ϕ is considered, and (13) will be sufficient but this can straightforwardly be extended to consider the general case.

4. SIMULATION AND RESULTS

4.1. Setup and Evaluation Metrics

Spatially and temporally white Gaussian noise was used as the source signals for training the filterbanks and simulating sensor noise. Pink noise from the Noisex database [36] was also used for training. Anechoic speech signals sampled at 16 kHz were taken from the TIMIT corpus [37]. For each speaker, short utterances were concatenated to generate signals of 8 to 10 s. During both training
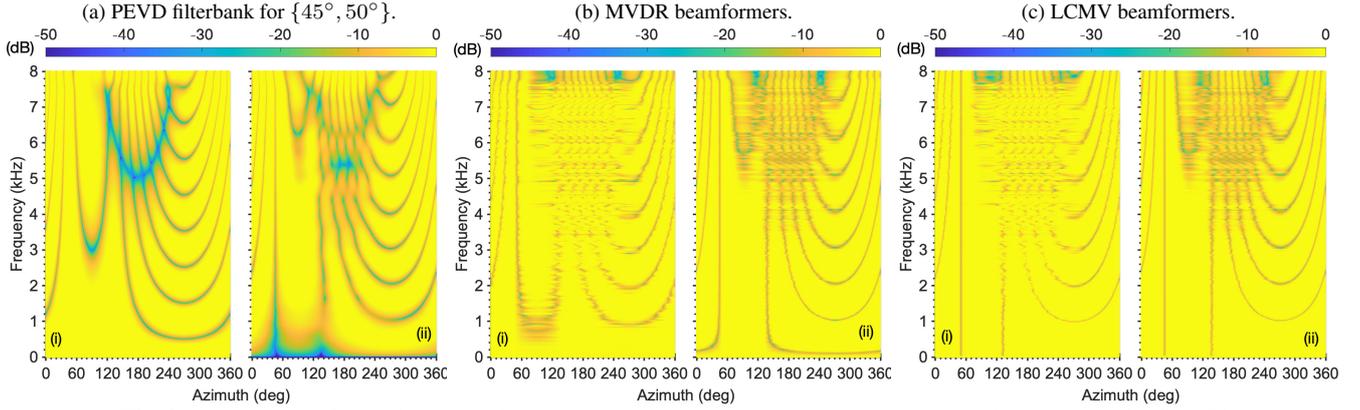
Fig. 2: Comparison of beamformers trained on two white, uncorrelated noise sources positioned at $\phi = \{45°, 50°\}$.

and testing, each source signal was convolved with the RIRs and summed before adding 30 dB SNR sensor noise.

Anechoic RIRs were simulated using [38] for a single source located at different angular positions $0° \leq \phi < 360°$ in steps of $5°$ relative to the x-axis in Fig. 1, which shows a $Q = 4$ free-field hearing aid array. Different combinations of $P = 2$ sources at positions $(\phi_i, \phi_j), i \neq j$, and the SMD algorithm [21] were used to train the fixed beamformers. To guarantee spectral majorization in (8) so that the PEVD outputs are ordered in a desired sequence as explained in Section 3.1, the signal power of source S1 at ϕ_i was 10 dB higher than S2 at ϕ_j. When both sources are white Gaussian, the filterbank has been designed to primarily target the acoustic channel, as explained in Section 3.1. For each azimuth pair, the learnt PEVD filterbank was stored in a look-up table, the entries of which were later retrieved and directly used for testing.

During testing, the setup is shown in Fig. 1(a) where S1 and S2 were positioned at $\phi_1 = 45°$ and $\phi_2 = 50°$ from the x-axis, 2 m from the hearing aid array. The speech sources were adjusted to the same power levels at the source positions using [39] implemented in [40]. RIRs for a 5 m $\times$ 4 m $\times$ 6 m room with a reverberation time of 300 ms were generated using [41]. For performance evaluation, short-time objective intelligibility (STOI) [42] and signal to interference ratio (SIR) [43] are used to measure speech intelligibility and interference rejection. Although each PEVD filterbank consists of 4 beamformers, results for only the first 2 are presented for brevity.

4.2. Experiments and Discussions

4.2.1. Experiment 1: Beam Patterns for White Noise Sources

Two white noise sources are spaced $5°$ apart, and the first source is 10 dB higher in power than the second source. The PEVD beamformers are trained on these signals for different source angles. The minimum, maximum and mean lengths of the filters are 17, 372 and 114, respectively with the histogram shown in Fig. 1(b).

The beam patterns for the PEVD beamformers are shown in Fig. 2(a). The first output in Fig. 2(a)(i) has a response that resembles a delay-and-sum beamformer, with unit gain in the look direction $\phi_1 = 45°$ for all frequencies. This is expected since the SMD algorithm [21] uses a series of delays to maximize the auto-correlation sequences on the diagonal eigenvalue polynomial matrix while enforcing an ordering from the largest to smallest energy. The second output in Fig. 2(a)(ii) shows a deep null steered in the direction of the first source position at $\phi_1 = 45°$ across all frequencies. The first and second beamformers also exhibit orthogonality characteristics in space and z-domain, e.g., unit gain and a null at $\phi_1 = 45°$ for PEVD beamformers 1 and 2, respectively.

For reference, MVDR and LCMV beamformers for $\{45°, 50°\}$ are also presented in Fig. 2(b) and (c). The beamformer weights are calculated using the statistics estimated from the received signals in the array and processed using [44]. The ground truth steering vectors are provided to ensure unit gain in the look directions as shown in Fig. 2(b). For the LCMV, hard constraints, i.e., unit gain in the target direction and zero gain in the interferer direction for all frequencies, are used to generate the beam patterns in Fig. 2(c). Compared to the second LCMV in Fig. 2(c)(ii), the second PEVD beamformer in Fig. 2(a)(ii) has additionally provided some attenuation around the low frequencies near the null direction $\phi_1 = 45°$. Note that this direction has unit gain for the first beamformer in Fig. 2(a)(i).

4.2.2. Experiment 2: Separation of Speech Sources

In this experiment, the source signals are two male speakers of the same signal power at their respective positions. The previously trained PEVD filterbank for white sources located at $\{45°, 50°\}$ shown in Fig. 2(a) is used for this experiment without any modifications. The beamformer weights for MVDR and LCMV are calculated using the statistics estimated from the received signals in the array while the ground truth steering vectors are provided.

The beam patterns for the MVDR and LCMV are provided in Fig. 3(b) and (c), respectively. The PEVD beamformer trained on white noise signals in Fig. 2(a)(ii) has a beam pattern that resembles the second LCMV beamformer in Fig. 3(c)(ii) — both impose null constraints on $\phi_1 = 45°$ except for the low frequency attenuation provided by PEVD in the neighbourhood of the $\phi_1 = 45°$.

The source separation results for an anechoic environment are shown in Table 1. In the PEVD filterbanks designed for $\{\phi_i, \phi_j\}$, the first beamformer behaves like a delay-and-sum steered towards ϕ_i while the second beamformer behaves like a cancellation or spatial filter nulling signals arriving from ϕ_i. Consequently, the PEVD cancellation beamformers performs the best in STOI and even out-

Table 1: Source separation for an anechoic environment. PEVD and PEVD′ are trained on 2 white, and 1 white and 1 pink sources.

Algorithm	S1 ($\phi_1 = 45°$)		S2 ($\phi_2 = 50°$)	
	STOI	SIR (dB)	STOI	SIR (dB)
Received	0.809	0.240	0.640	- 0.358
PEVD $\{45°, 50°\}$	0.811	0.206	0.844	15.394
PEVD $\{50°, 45°\}$	0.932	16.943	0.644	- 0.111
MVDR	0.922	13.727	0.826	12.077
LCMV	0.856	20.226	0.796	23.164
PEVD′ $\{45°, 50°\}$	0.806	0.195	0.369	11.953
PEVD′ $\{50°, 45°\}$	0.366	12.392	0.636	- 0.109

(a) PEVD$'$ $\{45°, 50°\}$ trained on white and pink noise. (b) MVDR beamformers. (c) LCMV beamformers.

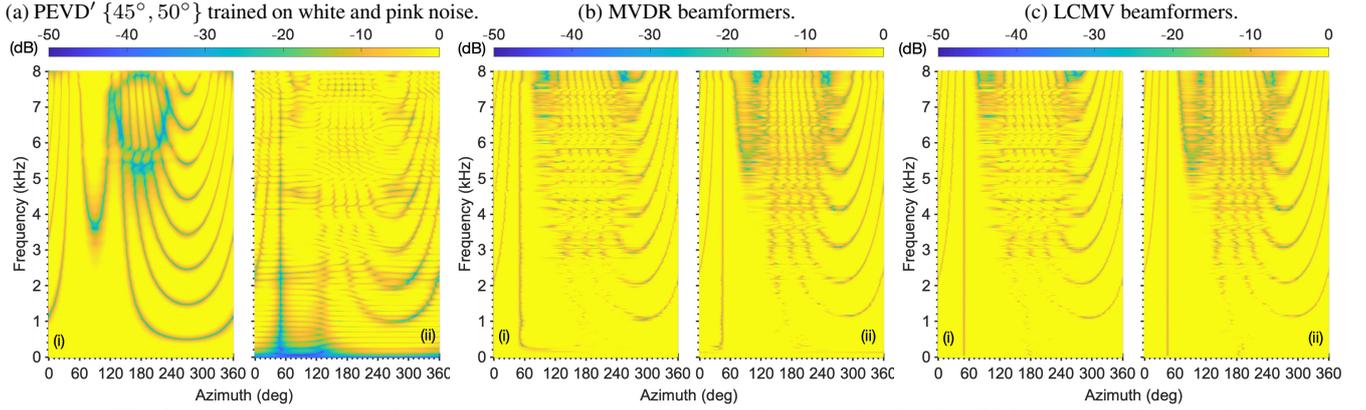**Fig. 3**: Comparison of beamformers applied to the two talkers positioned at $\phi = \{45°, 50°\}$ for an anechoic scenario.

perform MVDR by 0.018 for S2. In terms of SIR, the PEVD cancellation filters performs better than MVDR by more than 3 dB but not as well as LCMV beamformers which places deep notch at the interferer direction. However, PEVD cancellation beamformers performs better than LCMV in STOI by up to 0.15.

The results for a room with a T_{60} of 300 ms is shown in Table 2. In this case, the fixed beamformers designed by PEVD are used directly. The PEVD$\{50°, 45°\}$ cancellation beamformer that extracts S1 at $\phi = 45°$ gives the most significant improvement in STOI and SIR simultaneously. The other PEVD beamformers do not generally worsen STOI and SIR scores as much as MVDR and LCMV. Although the ground truth steering vectors are provided, the estimation of signal statistics using the microphone signals rather than the ground truth noise covariance matrices makes the beamformers perform more like the minimum power distortionless response (MPDR) and linearly constrained minimum power (LCMP), respectively. The beam patterns for the reverberant room are also provided [45].

4.2.3. Experiment 3: Different Source Signals for Training

A pink, instead of the white, noise signal is used as the source S2 for training and this trained PEVD filterbank is denoted by PEVD$'$. Compared to PEVD in Fig. 2(a), the spatio-spectral characteristics of PEVD$'$ in Fig. 3(a) are less smooth. The source separation results in Table 1 show that PEVD beamformers which are trained on white noise signals targeting spatial separation perform better than PEVD$'$. This suggests that using pink noise for training may not be suited for the separation of speakers because (8) may no longer be guaranteed.

5. CONCLUSION

We have proposed a fixed beamformer design using PEVD. Depending on the source signals used for training, the PEVD can design spatially, and possibly spectrally, orthogonal filterbanks for arbitrary arrays. We have utilized the array geometry information and introduced beam pattern analysis to interpret the eigenvector filterbanks generated by PEVD. For the separation of two source signals, analysis of the beam patterns has shown that the first and second beamformers behave like delay-and-sum and cancellation filterbanks, respectively. The fixed PEVD filterbanks are lightweight, with an average filter length of 114 and perform as well as, if not better than, data-dependent MVDR and LCMV beamformers even for closely spaced sources ($5°$). Informal listening examples are available [45].

6. REFERENCES

[1] D. Sharma *et al.*, "Spatial processing front-end for distant ASR exploiting self-attention channel combinator," in *Proc. IEEE*

Table 2: Source separation performance in reverberant room.

Algorithm	S1 ($\phi_1 = 45°$)		S2 ($\phi_2 = 50°$)	
	STOI	SIR (dB)	STOI	SIR (dB)
Received	0.698	0.298	0.566	0.055
PEVD $\{45°, 50°\}$	0.682	0.950	0.557	- 0.133
PEVD $\{50°, 45°\}$	0.715	1.102	0.533	- 0.502
MVDR	0.649	- 0.483	0.582	0.128
LCMV	0.638	4.958	0.491	- 3.940

Int. Conf. on Acoust., Speech and Signal Process. (ICASSP), 2022, pp. 7997–8001.

[2] C. Evers and P. A. Naylor, "Acoustic SLAM," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 9, pp. 1484–1498, Sep. 2018.

[3] A. H. Moore *et al.*, "Binaural mask-informed speech enhancement for hearing aids with head tracking," in *Proc. Int. Workshop on Acoust. Signal Enhancement (IWAENC)*, Tokyo, Japan, Sep. 2018, pp. 461–465.

[4] B. D. van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[5] M. S. Brandstein and D. B. Ward, Eds., *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer-Verlag, 2001.

[6] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of Signal Enhancement and Array Signal Processing*. Wiley-IEEE Press, 2018.

[7] S. Sun *et al.*, "Multiple fixed beamformers with a spacial Wiener-form postfilter for far-field speech recognition," in *Asia-Pacific Signal and Inform. Process. Assoc. Annual Summit and Conf. (APSIPA)*, 2019.

[8] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2002.

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

[10] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction," *Speech Commun.*, vol. 49, no. 7–8, pp. 636–656, Aug. 2007.

[11] 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.

[12] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the noise reduction Wiener filter," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1218–1234, Jul. 2006.

[13] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.

[14] E. d'Olne, A. H. Moore, and P. A. Naylor, "Model-based beamforming for wearable microphone arrays," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2021.

[15] S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

[16] I. Cohen, S. Gannot, and B. Berdugo, "An integrated real-time beamforming and post filtering system for nonstationary noise environments," *EURASIP J. on Applied Signal Process.*, vol. 11, pp. 1064–1073, 2003.

[17] 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.

[18] J. G. McWhirter *et al.*, "An EVD algorithm for para-hermitian polynomial matrices," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2158–2169, May 2007.

[19] Z. Wang, J. G. McWhirter, J. Corr, and S. Weiss, "Multiple shift second order sequential best rotation algorithm for polynomial matrix EVD," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2015, pp. 844–848.

[20] V. W. Neo and P. A. Naylor, "Second order sequential best rotation algorithm with Householder transformation for polynomial matrix eigenvalue decomposition," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2019, pp. 8043–8047.

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

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

[23] 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.

[24] V. W. Neo, C. Evers, and P. A. Naylor, "Polynomial matrix eigenvalue decomposition of spherical harmonics for speech enhancement," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, Jun. 2021, pp. 786–790.

[25] 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.

[26] V. W. Neo, C. Evers, and P. A. Naylor, "Polynomial matrix eigenvalue decomposition-based source separation using informed spherical microphone arrays," in *Proc. IEEE Workshop on Appl. of Signal Process. to Audio and Acoust. (WASPAA)*, New Paltz, NY, Oct. 2021, pp. 201–205.

[27] M. A. Alrmah, S. Weiss, and S. Lambotharan, "An extension of the MUSIC algorithm to broadband scenarios using a polynomial eigenvalue decomposition," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2011, pp. 629–633.

[28] A. O. T. Hogg *et al.*, "A polynomial eigenvalue decomposition MUSIC approach for broadband sound source localization," in *Proc. IEEE Workshop on Appl. of Signal Process. to Audio and Acoust. (WASPAA)*, New Paltz, NY, Oct. 2021.

[29] 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.

[30] S. Weiss *et al.*, "MVDR broadband beamforming using polynomial matrix techniques," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2015, pp. 839–843.

[31] A. Alzin *et al.*, "Adaptive broadband beamforming with arbitrary array geometry," in *Proc. IET Int. Conf. on Intelligent Signal Process.*, Dec. 2015, pp. 1–6.

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

[33] S. Weiss, J. Pestana, I. K. Proudler, and F. K. Coutts, "Corrections to "On the Existence and Uniqueness of the Eigenvalue Decomposition of a Parahermitian Matrix"," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6325–6327, Dec. 2018.

[34] P. P. Vaidyanathan, *Multirate Systems and Filters Banks*. New Jersey, USA: Prentice Hall, 1993.

[35] 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.

[36] A. Varga and H. J. M. Steeneken, "Assessment for automatic speech recognition II: NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Commun.*, vol. 3, no. 3, pp. 247–251, Jul. 1993.

[37] J. S. Garofolo *et al.*, "DARPA timit acoustic-phonetic continuous speech corpus CD-rom," Nat. Inst. of Standards and Tech. (NIST), NIST Interagency/Internal Report (NISTIR), Feb. 1993.

[38] A. H. Moore, "Microphone array classes in the ELOBES project," Jan. 2020. [Online]. Available: https://github.com/ImperialCollegeLondon/sap-elobes-microphone-arrays

[39] "Objective measurement of active speech level," Int. Telecommun. Union (ITU-T), Recommendation, Mar. 1993.

[40] D. M. Brookes, "VOICEBOX: A speech processing toolbox for MATLAB," 1997. [Online]. Available: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html

[41] E. A. P. Habets, "Room impulse response generator," Technische Universiteit Eindhoven (TU/e), Tech. Rep., 2006.

[42] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "An algorithm for intelligibility prediction of time-frequency weighted noisy speech," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 7, pp. 2125–2136, Sep. 2011.

[43] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006.

[44] R. E. Crochiere, "A weighted overlap-add method for short-time Fourier analysis/Synthesis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 1, pp. 99–102, Feb. 1980.

[45] V. W. Neo, "Fixed beamformer design using PEVD," May 2022. [Online]. Available: https://vwn09.github.io/research/pevd-beamformer-iwaenc