Wideband MPDR Beamforming for a Behind-The-Ear Microphone Array

Teddy Herrera
Department of Electrical and Computer Engineering, Naval Postgraduate School
1 University Circle, Monterey, CA 93943, USA
herrera.teddy45@gmail.com

***Abstract*—Modern hearing aids operate in acoustically complex environments where speech signals are inherently wideband and microphone arrays are severely size-constrained. Traditional narrowband beamforming techniques, which rely on frequency-domain phase steering, are therefore ill-suited for behind-the-ear (BTE) hearing aid applications. This work presents a simulation study of a finite impulse response (FIR)–based wideband minimum power distortionless response (MPDR) beamformer formulated in the space-time domain and tailored to a realistic BTE microphone array geometry. By avoiding frequency-domain approximations, the proposed approach preserves inter-frequency coherence essential for speech intelligibility. The MPDR formulation is recast into a linearly constrained minimum variance (LCMV) framework to enable explicit control over mainlobe steering and spatial null placement in both azimuth and elevation. Using a three-microphone BTE configuration, beampattern analyses demonstrate effective wideband steering toward a desired speech source and accurate suppression of interferers at prescribed directions, while illustrating fundamental limitations imposed by array geometry and degrees of freedom. Results confirm that up to two independent nulls can be reliably enforced while maintaining a distortionless target response. This study establishes a static wideband beamforming baseline suitable for compact hearing aid arrays and provides a foundation for future adaptive implementations such as generalized sidelobe canceller architectures in dynamic acoustic environments.**

***Index Terms*—MPDR, Wideband Beamforming, Microphone Array, Hearing aids**

I. Introduction

MODERN hearing aids are increasingly tasked with enhancing speech in complex acoustic environments, filled with multiple competing sound sources and noise. While typical narrowband beamforming techniques – such as phase steering and direction of arrival estimation – perform well in radar and sonar applications where signals are narrowband and array sizes are relatively large, they are ill suited for hearing aid applications.

This is due to the compact nature of in-the-ear (ITE) and behind-the-ear (BTE) hearing aid designs which limit the number of microphone elements to maintain a lightweight, low-power operation. Additionally, speech signals are inherently wideband spanning several kilohertz (kHz), which challenges narrowband assumptions. Speech signals are inherently wideband signals and span the 250 Hz to 8 kHz range [1] [2].

Wideband beamforming offers a solution but deviates fundamentally from the narrowband techniques. To maintain coherence across the frequencies, wideband signals must be processed in the time-domain. Van Trees [3], presents two methods capable of maintaining this coherence: the first approach utilizes a discrete Fourier transform (DFT)-based beamformer, which is ill-suited for processing speech due to its reliance on breaking the signal into discrete frequency bins which fails to capture inter-frequency correlations critical for speech intelligibility and directional. The second approach employs finite impulse response (FIR) at each microphone output to approximate the desired spatial response developed by Frost [4] which is far better suited for speech enhancement tasks.

This paper simulates an FIR-based wideband minimum power distortionless response (MPDR) beamformer tailored to a realistic BTE microphone array based on the configuration from Kayser et al [5]. The beamformer is formulated as a space-time optimization, problem, avoiding frequency domain approximations. Furthermore, the system is recast into the Linearly Constrained Minimum Variance (LCMV) framework to demonstrate precise control of mainlobe direction and null placement, enabling effective interference and noise suppression.

Beampatterns for the array are visualized in azimuth and elevation to demonstrate the spatial selectivity and wideband nature of the design and provides a foundational step towards practical deployment of adaptive wideband beamforming structures, such as the generalized sidelobe canceller (GSC) discussed by Van Trees [3], and Doclo et al. [6].

II. Problem Formulation

A. Signal Model

Following Van Trees [3], we adopt a space-time signal model suitable for wideband beamforming. The observation vector at time t is defined by stacking delayed microphone signals into a space-time vector:

$$\mathbf{x}_t = \begin{bmatrix} x_1(t) \\ x_1(t-1) \\ \vdots \\ x_1(t-(L-1)) \\ x_2(t) \\ \vdots \\ x_M(t-(L-1)) \end{bmatrix} \in R^{ML} \tag{1}$$

where M is the number of microphones in the array, and L is the number of FIR taps per microphone.

The microphone signals contain contributions from the target source $s_d(t)$, multiple interferers $s_i(t)$, and additive noise n_t:

$$\mathbf{x}_t = \sum_{i=1}^{N} \mathbf{H}_i \mathbf{s}_i + \mathbf{n}_t. \tag{2}$$

Here, $\mathbf{H}_i$ represents the space-time channel response from source i to the microphone array, and $\mathbf{s}_i$ is the vector of delayed samples of the source $s_i(t)$. For wideband signals like speech, this space-time model inherently preserves cross-frequency correlations, critical for speech intelligibility.

B. Wideband MPDR Beamforming Formulation

Minimum variance distortionless response (MVDR) beamformers are commonly used in narrowband beamforming applications because they minimize output power at a specific frequency. The minimum power distortionless response (MPDR) formulation generalizes the narrowband approach by directly minimizing output power in the time domain, making it suitable for wideband signals like speech. MPDR minimizes the array output power while maintaining a distortionless response in the direction of the desired source. The optimization problem is defined as:

$$\min_{\mathbf{w}} \quad \mathbf{w}^T \mathbf{R}_{xx} \mathbf{w} \tag{3}$$

subject to:

$$\mathbf{w}^T \mathbf{g} = 1. \tag{4}$$

Where, $\mathbf{w}$ is the concatenated FIR weight vector across all microphones, $\mathbf{R_{xx}}$ is the space-time covariance matrix of the microphone array signals, and $\mathbf{g}$ is the space time steering vector for the target direction, constructed by stacking the delayed impulse responses. In wideband beamforming $\mathbf{g}$ differs from the narrowband steering vector, which applies frequency dependent phase shifts, by "stacking" the temporal responses across the microphones. This structure captures both spatial and temporal information, preserving cross-band correlations critical for speech processing.

The closed-form solution to the MPDR problem is given by:

$$\mathbf{w}_{\text{MPDR}} = \frac{\mathbf{R}_{xx}^{-1}\mathbf{g}}{\mathbf{g}^T \mathbf{R}_{xx}^{-1} \mathbf{g}}. \tag{5}$$

This ensures that the beamformer minimizes output power while preserving the target signal undistorted.

C. LCMV Recasting

While MPDR enforces a distortionless constraint for the target direction, it provides limited control over interference and noise. To enable practical control over both the mainlobe and sidelobes, the problem is extended into the linearly constrained minimum variance (LCMV) framework. LCMV extends MPDR by allowing multiple linear constraints:

$$\min_{\mathbf{w}} \quad \mathbf{w}^T \mathbf{R}_{xx} \mathbf{w} \tag{6}$$

subject to:

$$\mathbf{C}^T \mathbf{w} = \mathbf{f}. \tag{7}$$

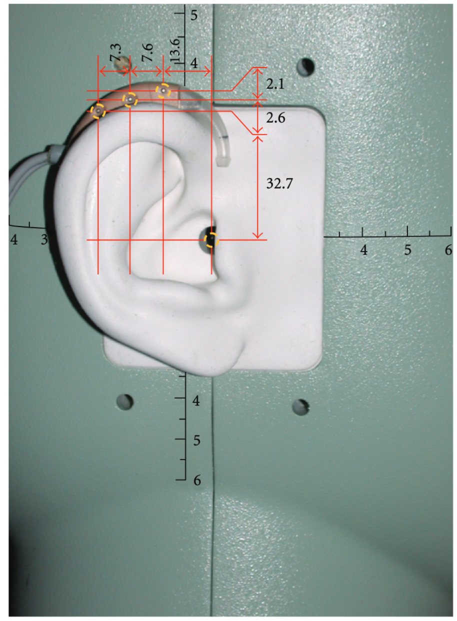

Fig. 1. Behind-the-ear hearing aid microphone array geometry. Source [5]

where $\mathbf{C} \in R^{ML \times K}$ contains K constraint vectors, and $\mathbf{f} \in R^K$ defines the desired gain for each constraint (e.g., 1 for the target, 0 for nulls).

The LCMV solution is given by:

$$\mathbf{w}_{\text{LCMV}} = \mathbf{R}_{xx}^{-1}\mathbf{C}\left(\mathbf{C}^T \mathbf{R}_{xx}^{-1}\mathbf{C}\right)^{-1}\mathbf{f}. \tag{8}$$

The LCMV extends the MPDR formulation and enables precise null placement at known interferer locations while maintaining target signal integrity in static noise environments. This provides a baseline for performance and analysis before implementing adaptive filters for dynamic noise environments.

III. Simulation Framework

A. Array Geometry

The BTE microphone array geometry in the simulation is modeled after the configuration developed by Kayser et al [5]. The array consists of three microphone elements positioned non-uniformly in a housing behind the ear, as shown in Figure 1. The microphone positions are defined relative to the ear canal, and this paper defines the coordinate system origin to be centered at the ear canal and in the same plane as the sensors.

The positive x axis is defined to point toward the face, the positive y axis points laterally outward from the ear canal, and the positive z axis points toward the top of the head. Figure 2 defines the azimuth angle ψ and elevation angle θ within defined coordinate system. This non-uniform and non-linear array geometry reflects realistic BTE microphone spacing and is critical for modeling direction-dependent spatial selectivity and beamformer performance.

B. Signal Model

The target signal is modeled as a broadband speech-like source arriving from a specified direction θ_t, ψ_t. Interfering

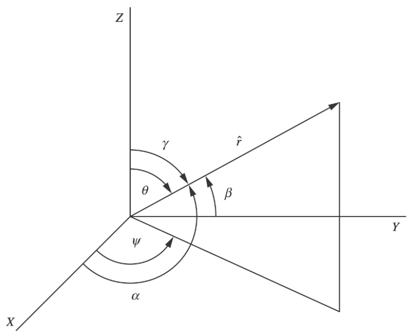

Fig. 2. Definition of azimuth angle ψ and elevation angle θ within the array coordinate system. Source [7]

sources are placed at user-defined angles $\theta_{\text{null}}, \psi_{\text{null}}$ to evaluate null placement capabilities and limitations. Both target and interference signals are modeled using ideal wideband impulse responses simulated as delayed copies of the source signal based on time-of-arrival differences at each microphone. Additive white Gaussian noise is also added at each microphone channel.

While speech bandwidth is from 250 Hz to 8 kHz, to ensure a high quality reconstruction and to account for leakage outside of the 8kHz it is recommended to sample above 16 kHz [8]. This paper uses the same sampling frequency as Kayser et al. [5] used to acquire their data. To model the time delays for the array's spatial aperture and speech signal spread, the FIR filters were defined to have a length of L = 256 taps per microphone channel. This balances temporal resolution with computational efficiency. Given the sampling rate, this filter length provides a frequency resolution of 188 Hz, which is adequate for resolving speech components and facilitating precise beamsteering and null placement [2].

C. Beamformer Configuration and Constraints

The FIR-based wideband LCMV beamformer is configured to maintain a distortionless response in the target direction while placing spatial nulls at known interferer locations. Each micropohone channel applies an FIR filter with L taps, forming a spacetime weight vector $\mathbf{w} \in R^{ML}$.

For each desired direction (θ, ψ), a space-time steering vector $\mathbf{g}(\theta, \psi)$ is generated. This vector accounts for the relative time delays between the target or interferer direction and each microphone, stacked across the L filter taps. The target steering vector $\mathbf{g}_{\text{target}}$ is computed for the desired source direction. Similarly, null steering vectors $\mathbf{g}_{\text{null}_i}$ are generated at each interferer location located at θ and ψ.

The beamformer spatial filtering behavior is defined through the constraint matrix $\mathbf{C}$, which stacks the target and null steering vectors:

$$\mathbf{C} = [\mathbf{g}_{\text{target}}, \mathbf{g}_{\text{null}_1}, \mathbf{g}_{\text{null}_2}, \ldots, \mathbf{g}_{\text{null}_K}] \tag{9}$$

where K is the total number of null constraints.

The desired response vector $\mathbf{f} \in R^{(1+K)}$ is defined as:

$$\mathbf{f} = [1, 0, 0, \ldots, 0]^T,$$

enforcing unity gain in the target direction and nulls in the interferer directions.

For this simulation, the space-time covariance matrix $\mathbf{R}_{xx}$ is assumed to be the identify matrix:

$$\mathbf{R}_{xx} = \mathbf{I}.$$

This assumption isolates the effect of spatial filtering and constraint enforcement without introducing correlation effects from environmental noise or reverberation.

The optimal space-time weight vector $\mathbf{w}_{\text{LCMV}}$ is computed by solving the LCMV optimization problem denoted in Equation 6. This formulation ensures distortionless target preservation while suppressing energy from constrained null directions.

While this work evaluates the beamformer's using a static LCMV formulation, the same space-time framework naturally extends to adaptive implementations discussed within [3], [6].

IV. Results

The configured wideband LCMV beamformer is evaluated through spatial beampattern analysis to demonstrate its ability to steer the mainlobe toward the target direction while placing nulls at specified interferer locations. All beampatterns are visualized in decibels (dB), normalized to the maximum output.

A. Broadside Beampattern

Figure 3 shows the beampattern when steering directly toward broadside, defined as $\theta = 90^\circ$, $\psi = 90^\circ$. As expected, the mainlobe is centered on the target direction with symmetric sidelobes determined by array geometry and FIR filter length. No null constraints are applied in this case.

B. Elevation Null Placement

Next, a spatial null is placed in the elevation plane at $\theta = 45^\circ$, $\psi = 90^\circ$. Figure 4 confirms that the beamformer successfully forms a deep null at the specified elevation while preserving the mainlobe at broadside.

C. Azimuth Null Placement

Figure 5 presents the beampattern with a null placed at $\theta = 90^\circ$, $\psi = 30^\circ$ in the azimuth plane. The beamformer demonstrates the ability to accurately place a spatial null while maintaining gain in the target direction.

D. Null Placement Limitation (Infeasible Null)

To demonstrate the beamformer's spatial limitation, a null is placed at $\theta = 80^\circ$, $\psi = 80^\circ$, within the half-power beamwidth of the mainbeam. Figure 6 shows that the beamformer can suppress this direction at the cost of the mainlobe being severly underpowered in the target direction.

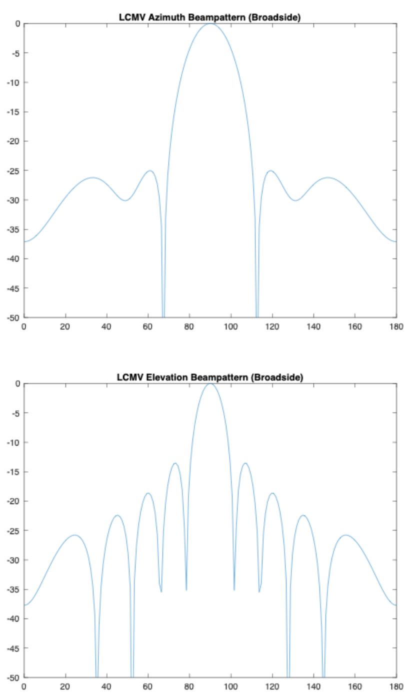

Fig. 3. Azimuth (Top) and Elevation (bottom) Beampatterns for the wideband beamformer steered to braodside with no null constraints.

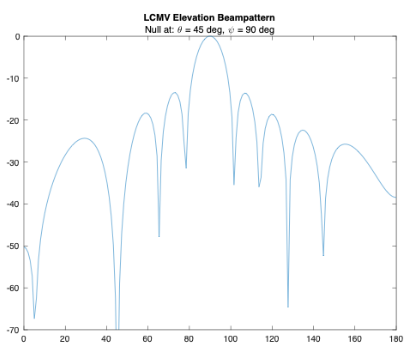

Fig. 4. Beampattern with elevation null at $\theta = 45°$, $\psi = 90°$.

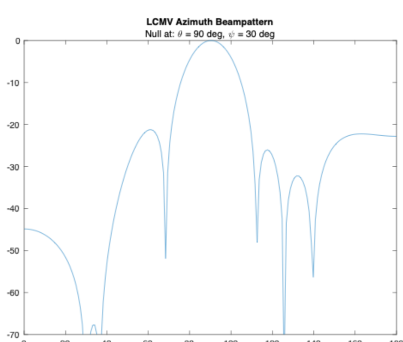

Fig. 5. Beampattern with azimuth null at $\theta = 90°$, $\psi = 30°$.

E. Maximum Number of Nulls

Finally, the theoretical limit of null placement is tested. Since the array contains $M = 3$ microphones, it can support a maximum of $M - 1 = 2$ independent nulls in addition to the target constraint. Figure 7 demonstrates successful null placement when two nulls are imposed. Attempts to exceed this limit degrade the beamformer's ability to enforce all constraints.

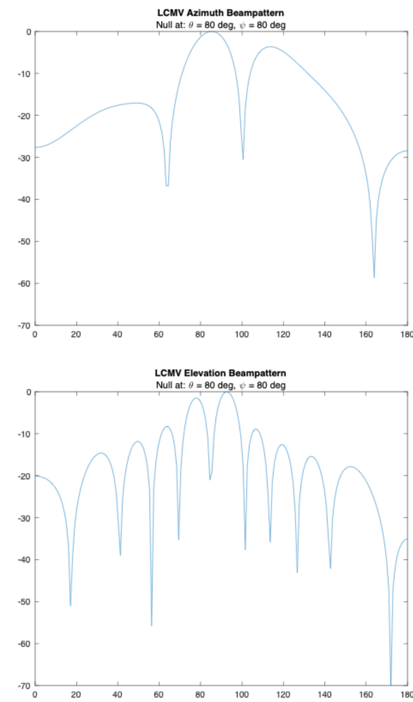

Fig. 6. Azimuth (Top) and Elevation (bottom) showing a failed null attempt at $\theta = 80°$, $\psi = 80°$.

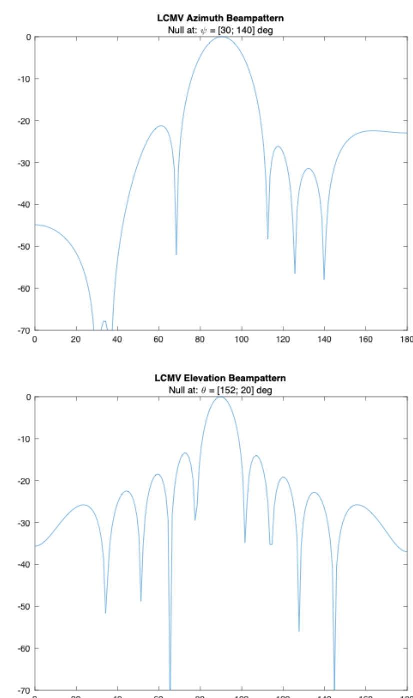

Fig. 7. Azimuth (Top) and Elevation (bottom) beampatterns showing maximum two nulls enforced with $M = 3$ microphones.

F. Discussion

These results validate that the FIR-based wideband LCMV beamformer accurately steers toward the target while placing spatial nulls at both elevation and azimuth directions. As expected, the ability to enforce nulls is bounded by the array's degrees of freedom, limiting the system to $M-1$ independent nulls. These results illustrate the trade-off between array size, FIR length, and spatial resolution in realistic BTE hearing aid configurations.

V. Conclusion

This paper presented a simulation of an FIR-based wideband minimum power distortionless response (MPDR) beamformer designed for a realistic behind-the-ear hearing aid microphone array. By formulating the beamformer in the space-time domain, the approach preserves wideband speech signal coherence, overcoming the limitations narrowband methods. The linearly constrained minimum variance (LCMV) framework was used to enable precise mainlobe steering and spatial null placement.

Simulation results demonstrated that the beamformer effectively steers toward the target direction while placing deep nulls at specified interferer locations in both azimuth and elevation. As expected, the system's directional filtering capability and the number of enforceable nulls are fundamentally constrained by the array's physical geometry and degrees of freedom. With three microphones, the array reliably places up to two nulls while preserving the desired target response.

Future work should recast the beamformer into a generalized sidelobe canceller structure (GSC) to enable dynamic adaptation to moving interferers and time-varying acoustic scenes, building on the static LCMV baseline demonstrated here [3], [6]. Additionally, future simulations should aim to utilize realistic acoustic environments and incorporate data generated from Kayser et al. [5] to evaluate performance in real-world conditions. Extending the model to a binaural configuration could further improve spatial cue preservation, enhancing speech intelligibility and localization for hearing aid users.

Overall, the results provide a strong foundation for practical wideband beamforming in hearing aids and highlight the key areas for future development.

References

[1] ITU-T Recommendation G.722, "7 kHz audio-coding within 64 kbit/s," November 1988.

[2] S. L. Bell, "Filtering to Match Hearing Aid Insertion Gain to Individual Ear Acoustics," Trends in Amplification, vol. 13, no. 3, pp. 181–189, Aug. 2009, doi: https://doi.org/10.1177/1084713809344974.

[3] H. L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory, Wiley-Interscience, 2002.

[4] O. L. Frost, "An Algorithm for Linearly Constrained Adaptive Array Processing," Proceedings of the IEEE, vol. 60, no. 8, pp. 926–935, 1972.

[5] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Evaluation of Binaural Noise Reduction Algorithms for Hearing Aids Using a Virtual Acoustic Environment," The Journal of the Acoustical Society of America, vol. 125, no. 1, pp. 164–170, 2009.

[6] S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, "Multichannel Signal Enhancement Algorithms for Assisted Listening Devices: Exploiting Spatial Diversity Using Multiple Microphones," IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 18–30, Mar. 2015.

[7] L. J. Ziomek, An Introduction to Sonar Systems Engineering. 2022.

[8] H. R. Staff, "Ensuring High-fidelity in Hearing Aid Sound Processing," The Hearing Review, Mar. 05, 2009. https://hearingreview.com/practice-building/practice-management/ensuring-high-fidelity-in-hearing-aid-sound-processing