Sparsity promoting adaptive beamforming for hearing aids

Ching-Hua Lee, Gokce Sarar, Bhaskar D. Rao, and Harinath Garudadri
Department of Electrical and Computer Engineering, University of California, San Diego

Contact Information:
Department of Electrical and Computer Engineering
University of California, San Diego
9500 Gilman Drive #0436, La Jolla, CA 92093

Email: `gsarar@ucsd.edu`

Abstract

Objective:

- To improve noise management in hearing aids by investigating advanced beamforming techniques in Open Speech Platform (OSP), `http://openspeechplatform.ucsd.edu`.

Methods:

- We adopt the generalized sidelobe canceller (GSC) as the baseline system and demonstrate improvements over one microphone case.
- Further improvements are achieved by leveraging the underlying sparsity of the adaptive filter for fast convergence.
- Sparsity is promoted by incorporating the ℓ_p norm as a sparsity penalty term in the least mean square (LMS) optimization problem, resulting in the sparsity promoting LMS (SLMS) algorithm.
- Future revisions of OSP are expected to support multiple microphones. We demonstrate SLMS for multi microphone beamforming use case.

Results

- The conventional GSC beamformer provides **14.2 dB of Signal-to-Interference Ratio (SIR)** and **0.22 of Hearing-Aid Speech Quality Index (HASQI)** improvements over the system with only one microphone. Incorporating SLMS in GSC, with a sparsity level of $p = 1.3$, we observed an improvement of **17.7 dB in SIR** and **0.28 in HASQI**.
- With use of SLMS, in the case of multiple interferences (3 in our experiments), in a 2-microphone setup we can achieve **12.5 dB SIR** and **0.16 HASQI** improvement, and with the use of additional microphones, in an 8-microphone setup, we can achieve **17.6 dB SIR** and **0.26 HASQI** improvement over the 1-microphone setup.

1 The baseline adaptive beamformer

The generalized sidelobe canceller (GSC) [1] is adopted as the baseline system to perform wideband beamforming for speech, as illustrated in Figure 1.

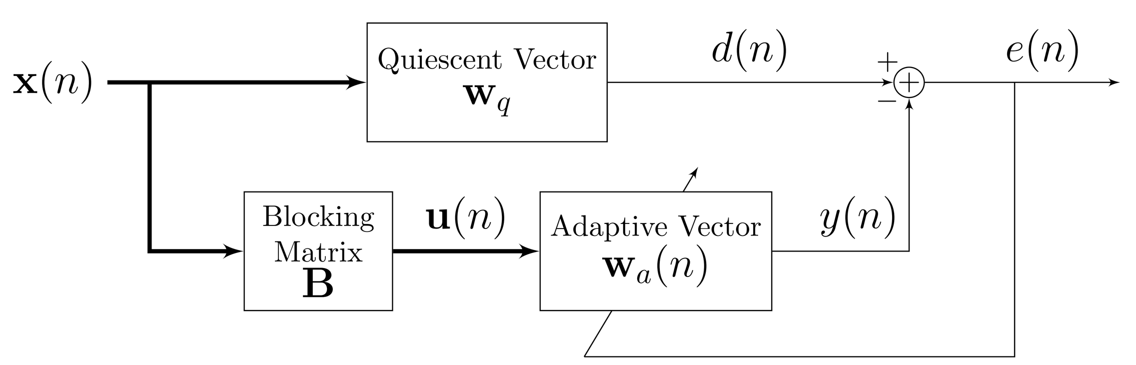

Figure 1: Block diagram of the GSC beamformer.

For a two microphone setting, a basic two-microphone GSC beamformer for hearing aids [2] is shown in Figure 2. Least mean squares (LMS) type algorithms are typically adopted for adjusting coefficients of the adaptive filter. The following "modified LMS" [3] is suggested for speech processing to alleviate the effect of abrupt onset and fluctuation of speech power:

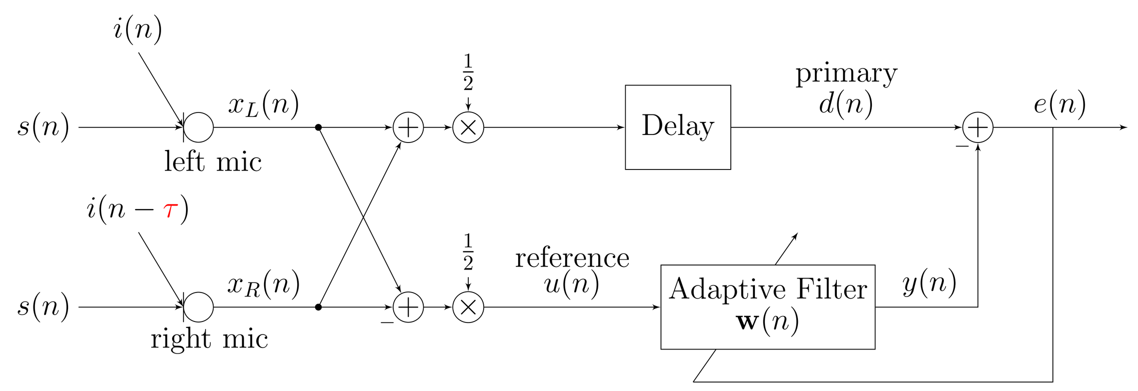

Figure 2: Block diagram of the basic two-mic GSC beamforming system. $d(n)$ corresponds to the enhanced broadside target signal by the delay-and-sum beamformer and $y(n)$ corresponds to the interference component. The goal is to cancel the interference by minimizing the power of $e(n)$ by adapting the coefficients of $\mathbf{w}(n)$.

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{M\hat{\sigma}^2(n) + \delta}\mathbf{u}(n)e(n), \quad (1)$$

where $\mu > 0$ is the step size parameter, M is the filter length, $\delta > 0$ is a small constant to prevent division by zero, and

$$\hat{\sigma}^2(n) = \rho\hat{\sigma}^2(n-1) + (1-\rho)(u^2(n) + e^2(n)) \quad (2)$$

is the running power estimate with a forgetting factor $0 < \rho \leq 1$.

2 Sparsity promoting adaptive beamformer

2.1 Incorporating sparsity

The adaptive filter at convergence shows a sparse nature as can be seen in Figure 3.

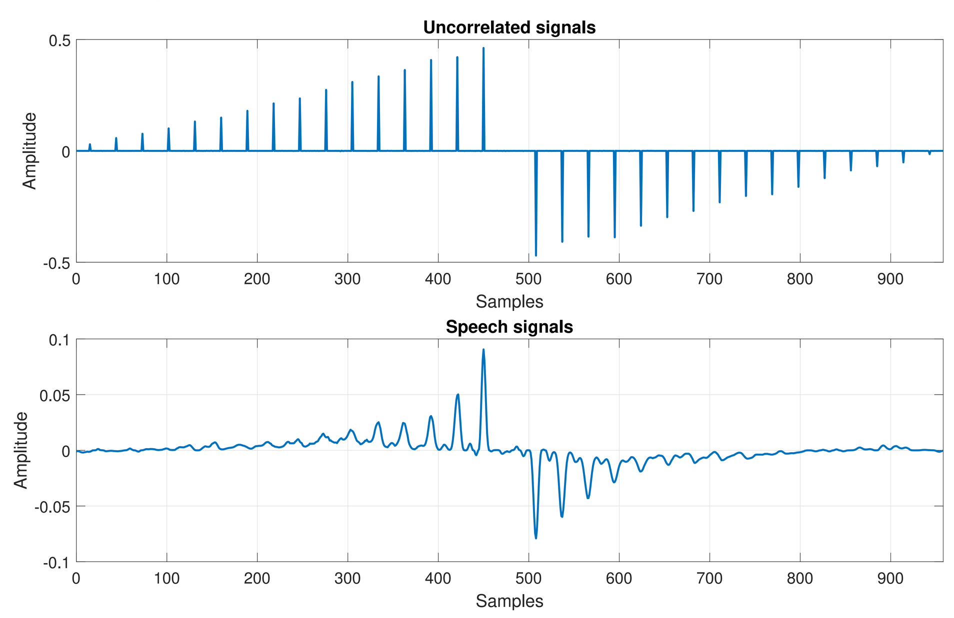

Figure 3: The adaptive filter coefficients of the two-mic GSC at convergence. In order to leverage this sparse nature, so that fast convergence can be achieved, we incorporated SLMS, where we adapt filter taps according to their weights, ie. filter taps with large weights are adapted faster.

The ordinary LMS considers the following mean squared error minimization problem:

$$\min_{\mathbf{w}} \quad J(\mathbf{w}) = E\left[|e(n)|^2\right] = E\left[|d(n) - \mathbf{w}^T\mathbf{u}(n)|^2\right]. \quad (3)$$

To promote sparsity we incorporate the ℓ_p norm as a sparsity penalty term to the LMS objective function:

$$\min_{\mathbf{w}} \quad J(\mathbf{w}) = E\left[|d(n) - \mathbf{w}^T\mathbf{u}(n)|^2\right] + \lambda\|\mathbf{w}\|_p^p, \quad (4)$$

where λ is the regularization coefficient and $\|\mathbf{w}\|_p^p = \sum_{i=0}^{M-1}|w_i|^p$, $p \in (0, 2]$ is the ℓ_p norm diversity measure.

2.2 The proposed algorithm

Setting $\lambda = 0$ in the end of the optimization results in the Sparsity promoting LMS (SLMS) [4] algorithm:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{M\hat{\sigma}^2(n) + \delta}\mathbf{G}(n)\mathbf{u}(n)e(n), \quad (5)$$

where

$$\mathbf{G}(n) = \mathrm{diag}\{g_0(n), g_1(n), ..., g_{M-1}(n)\} \quad (6)$$

is called the "**step-size control matrix**" and

$$g_i(n) = \frac{(|w_i(n)| + c)^{2-p}}{\frac{1}{M}\sum_{j=0}^{M-1}(|w_j(n)| + c)^{2-p}}, \quad p \in (0, 2], \quad (7)$$

where $c > 0$ is a small constant to prevent stagnation.

3 Performance Evaluation

For the anechoic scenario with the desired speech signal at the broadside, with the GSC beamformer operating in the 96 kHz domain, we performed 3 sets of experiments with the setups shown in Fig 4:

- The effect of the ℓ_p norm diversity measure p is investigated when there is one interference at $45°$ as can be seen in Figure 5.
- The performance of the GSC is investigated with and without SLMS for different angle of arrival (AoA) of the single interference as can be seen in Figure 6.
- Our initial experiments in multi-microphone multi-interference setup, are tested for different p values, with 2 and 8 microphones and with 3 interferences as can be seen in Figure 7.

We evaluated them with signal-to-interference ratio (SIR) and hearing-aid speech quality index (HASQI) [5].

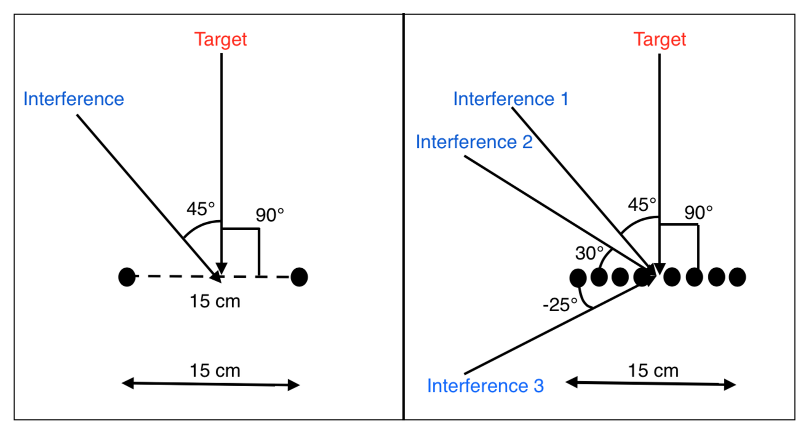

Figure 4: (Left) One interference - 2 microphone setup, (right) multi microphone - multi interference setup with 8 microphones

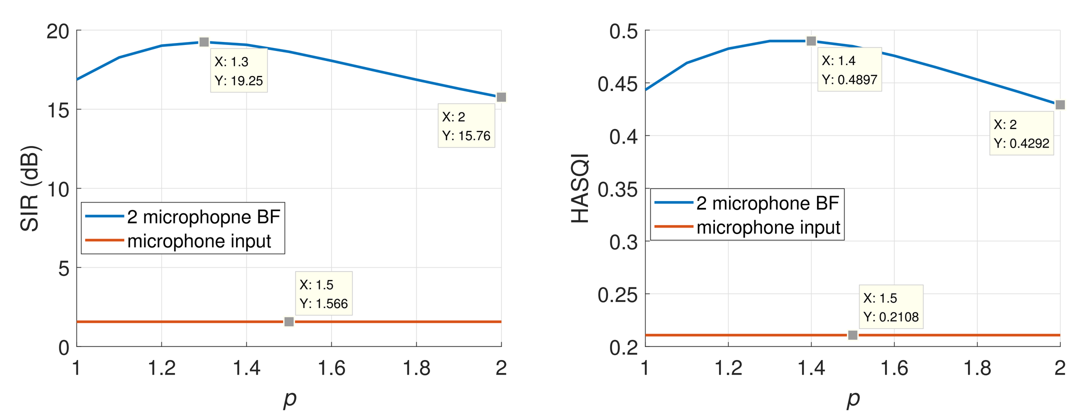

Figure 5: SIR and HASQI performance vs. p value with the single interference at $45°$ in the 2 microphone setup.

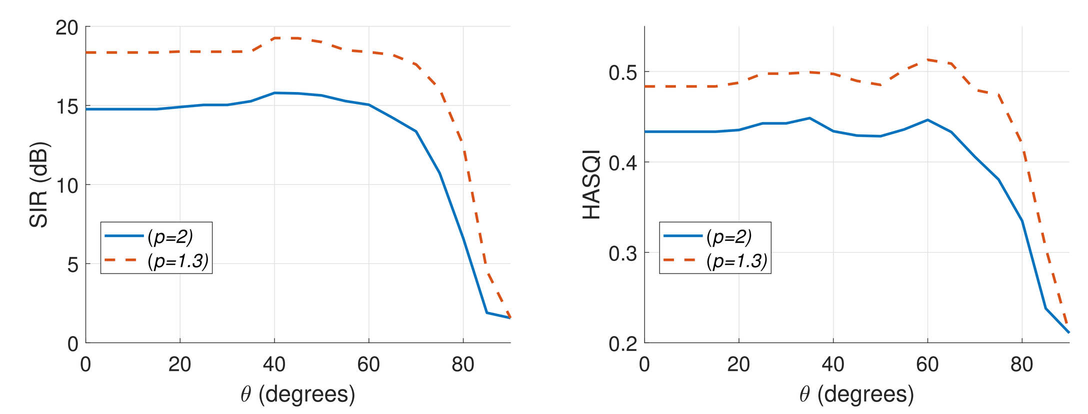

Figure 6: SIR and HASQI performance vs. AoA of the interference in the 2 microphone setup.

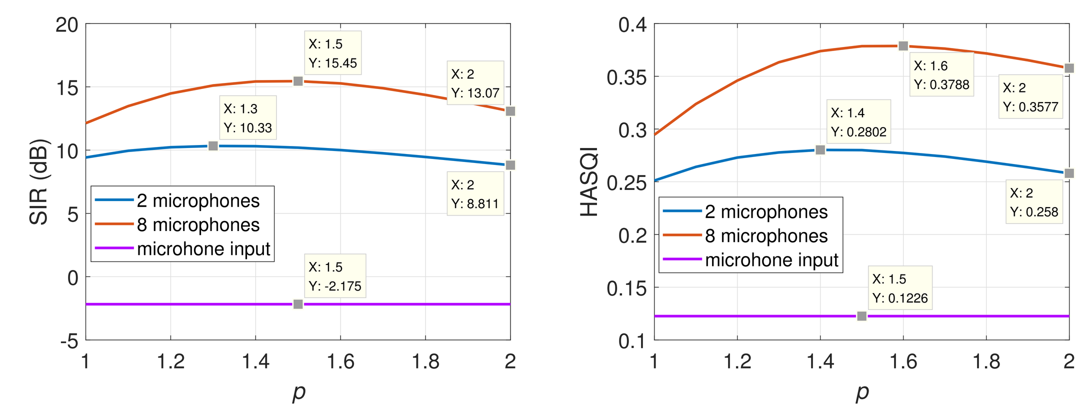

Figure 7: SIR and HASQI performance vs. p value with 3 interference at $30°, 45°$ and $-25°$ in 2 and 8 microphone setup.

4 Conclusion

In this contribution, we present a baseline adaptive beamformer for use in OSP that achieves **14.2 dB of SIR and 0.22 of HASQI** improvement with a 2-mic beamformer, compared with no beamformer in the case of single interference. Incorporating SLMS, a technique to leverage sparsity for faster convergence, we obtain a total improvement of **17.7 dB in SIR** and **0.28 in HASQI**. We show that the improved system can be extended to multiple microphones and shows significant improvement in the case of multiple interferences. In the case of 3 interferences, with 8 microphones, we achieve an improvement of **17.6 dB SIR** and **0.26 HASQI** improvement over the case without beamformer. Our future work would be on integrating beamforming for noise management in the real-time system and developing a robust system for reverberant environments and array misalignment.

5 Acknowledgements

This work was supported by NIH R01-DC015436: A Real-time, Open, Portable, Extensible Speech Lab and NIH R33-DC015046: Self-fitting of Amplification: Methodology and Candidacy.

References

[1] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, vol. 30, no. 1, pp. 27–34, 1982.

[2] J. E. Greenberg and P. M. Zurek, "Evaluation of an adaptive beamforming method for hearing aids," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1662–1676, 1992.

[3] J. E. Greenberg, "Modified LMS algorithms for speech processing with an adaptive noise canceller," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 4, pp. 338–351, 1998.

[4] C.-H. Lee, B. D. Rao, and H. Garudadri, "Sparsity promoting LMS for adaptive feedback cancellation,," in *Proc. Europ. Signal Process. Conf. (EUSIPCO)*, 2017, pp. 226–230.

[5] J. M. Kates and K. H. Arehart, "The hearing-aid speech quality index (HASQI) version 2," *J. Audio Eng. Soc.*, vol. 62, no. 3, pp. 99–117, 2014.