Technical Documentation ver. 1

4/11/2018 Anshuman Ganguly, Yiya Hao

Architecture for Eight microphone DOA for Raspberry Pi

1. **Algorithm:**
   a. 16-channel Super-directive Minimum Variance Distortionless Response (MVDR)[1]-based Fixed Beamforming;
   b. 16-channel MMSE-LSA based Post-filtering (optional);
   c. 16-channel Feature extraction and Single-feature based Unsupervised VAD;
   d. 16-channel Feature Smoothening and Adaptive Thresholding;
   e. Beam-Ranking for Speech Source localization to estimate DOA $\hat{\theta}$;
   f. Beam-selection for Speech Enhancement to obtain Enhanced Speech $\hat{s}(n)$.

2. **Implementation:**
   a. MATLAB 2017b for testing and debugging
   b. Visual Studio 2016 for C/C++ and Java Native Interface (JNI) for Android

3. **Frame length**:
   a. MATLAB Implementation: 100ms@16kHz = 1600. FFT size = 2048.
   b. C/Java Implementation: 40ms @48kHz. FFT size = 2048.

4. **Tunable parameters (for VAD):**
   a. Initial Threshold time (T) for VAD: 10 frames@100ms/frame = 1sec
   b. Duration Threshold (D): 2 frames@100ms/frame = 200ms. (to differentiate between sustained speech and short-duration noise)
   c. $\alpha_{As} = 10^{-6}$ (Fast attack), α_{Ds}= 0.5 (Slow decay); α_{An}= 0.7 (Slow attack), $\alpha_{Dn} = 10^{-6}$ (Fast decay).

5. **Comparison with other Eight Mic DOA estimation:** Cross-correlation(CC) [2] (No VAD and No Post-processing)

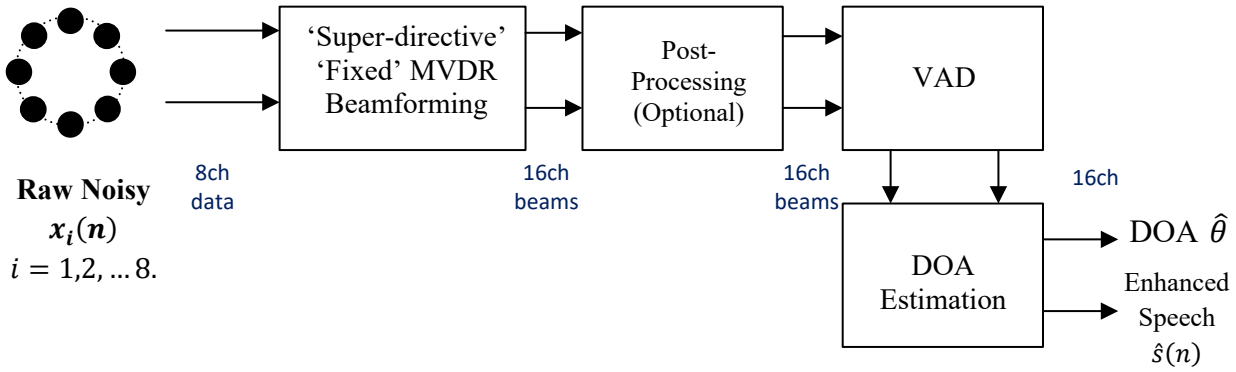
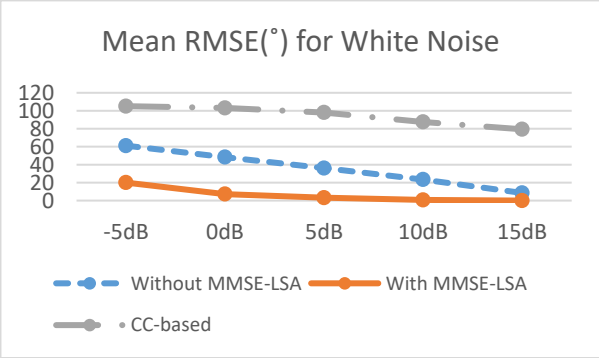
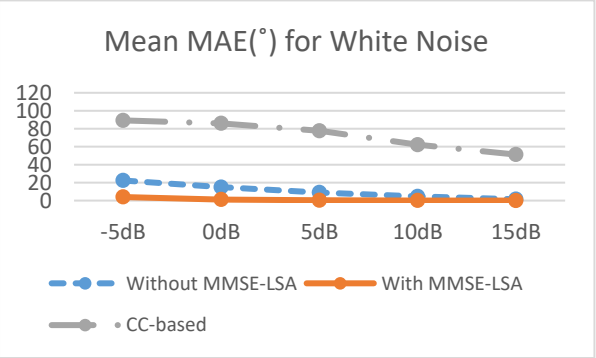


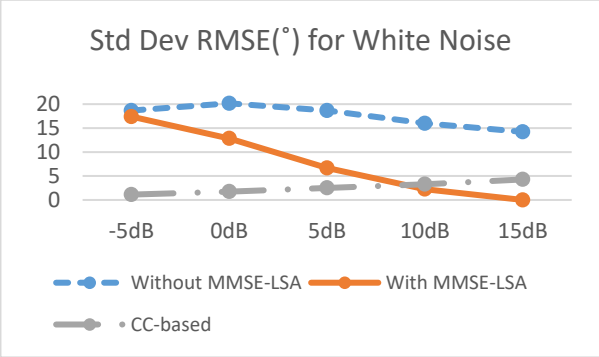
Figure 1. Proposed Architecture for Speech source localization using Uniform circular microphone array with 8 microphones

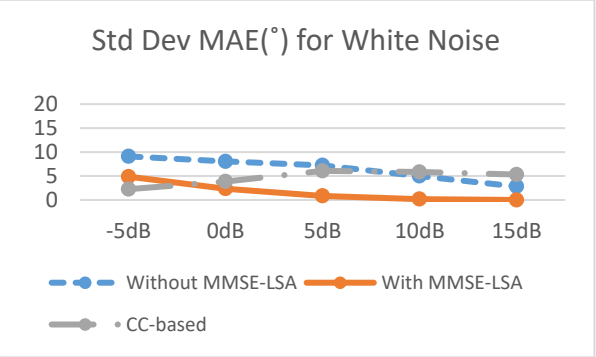
Experimental Evaluation

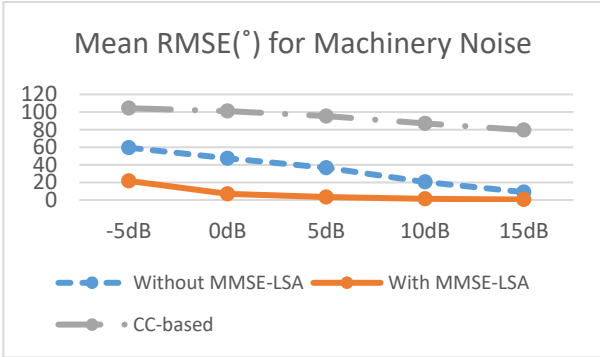
Experiment 1: DOA Estimation

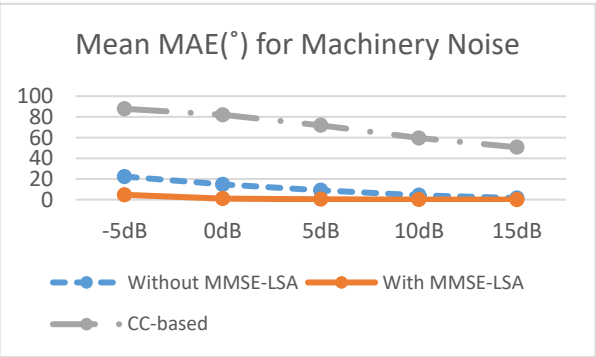


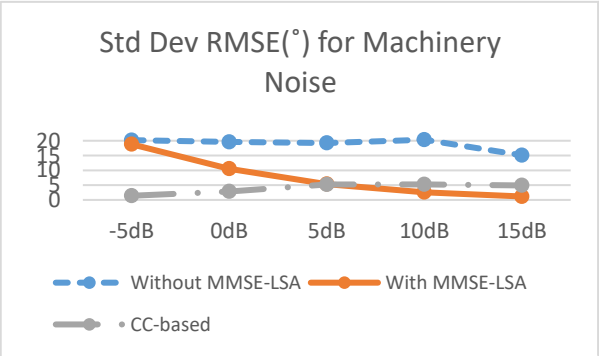


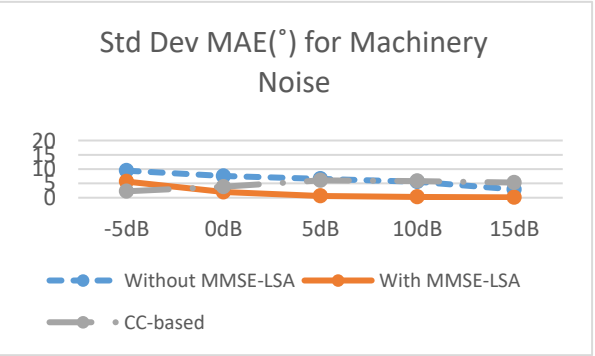












Refer: https://www.utdallas.edu/ssprl/hearing-aid-project/research-papers/

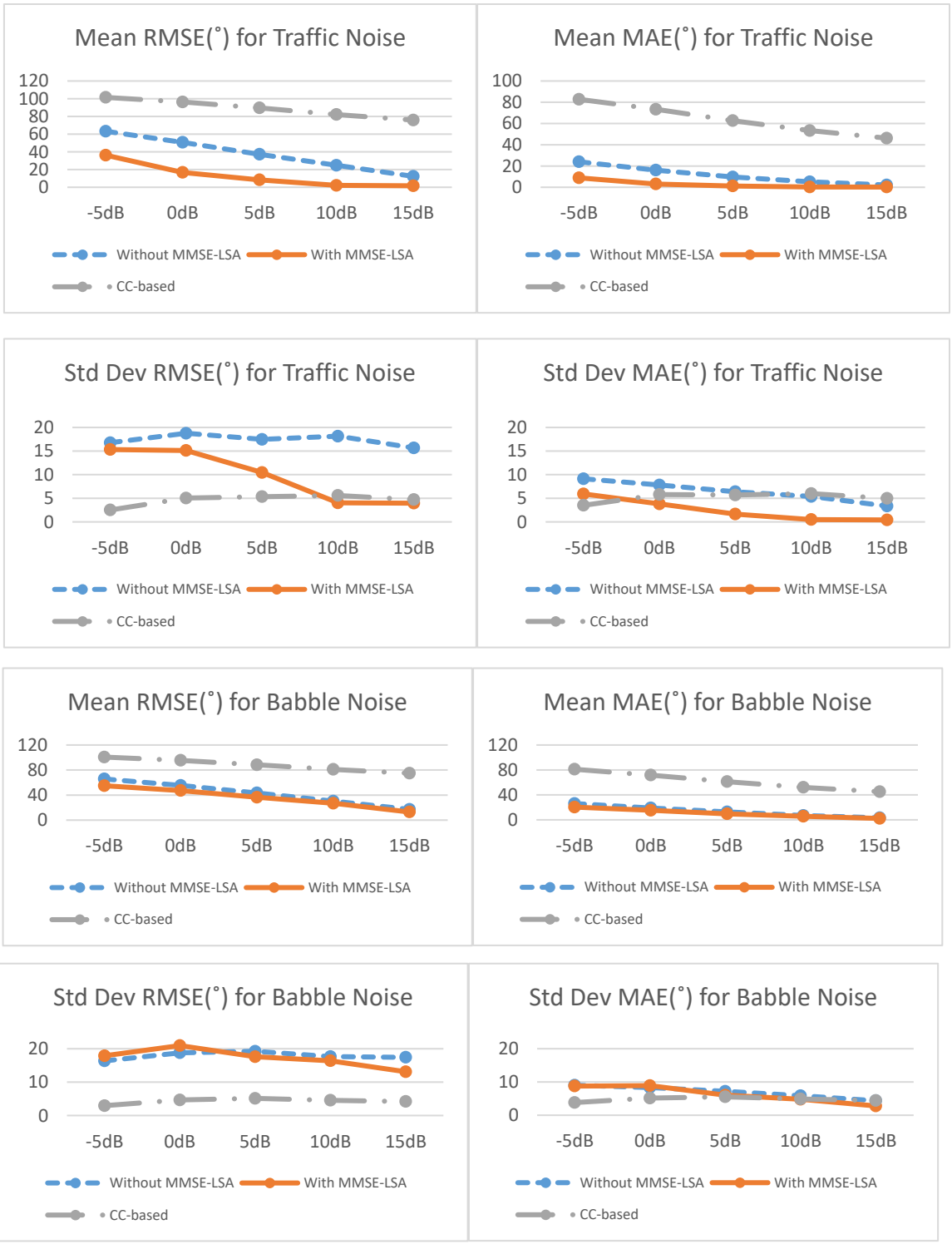
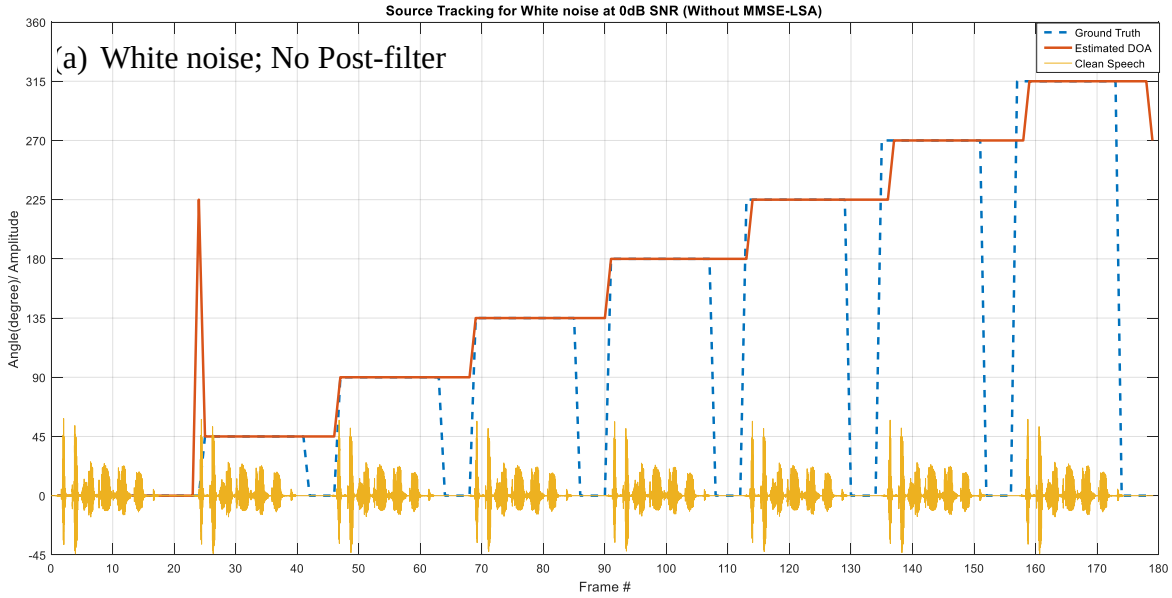
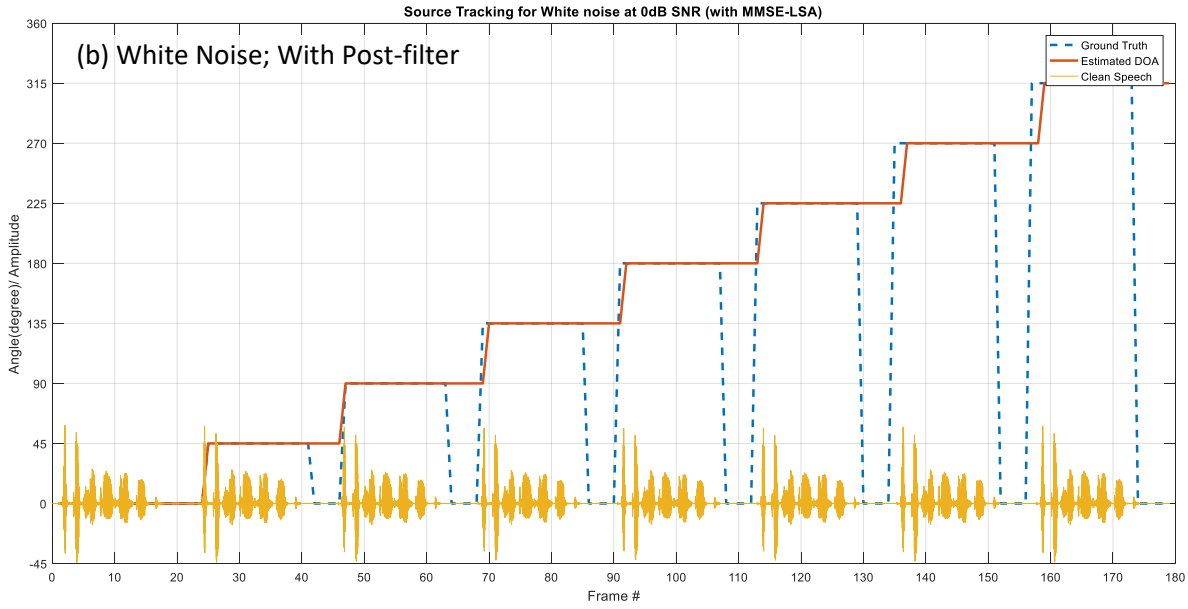


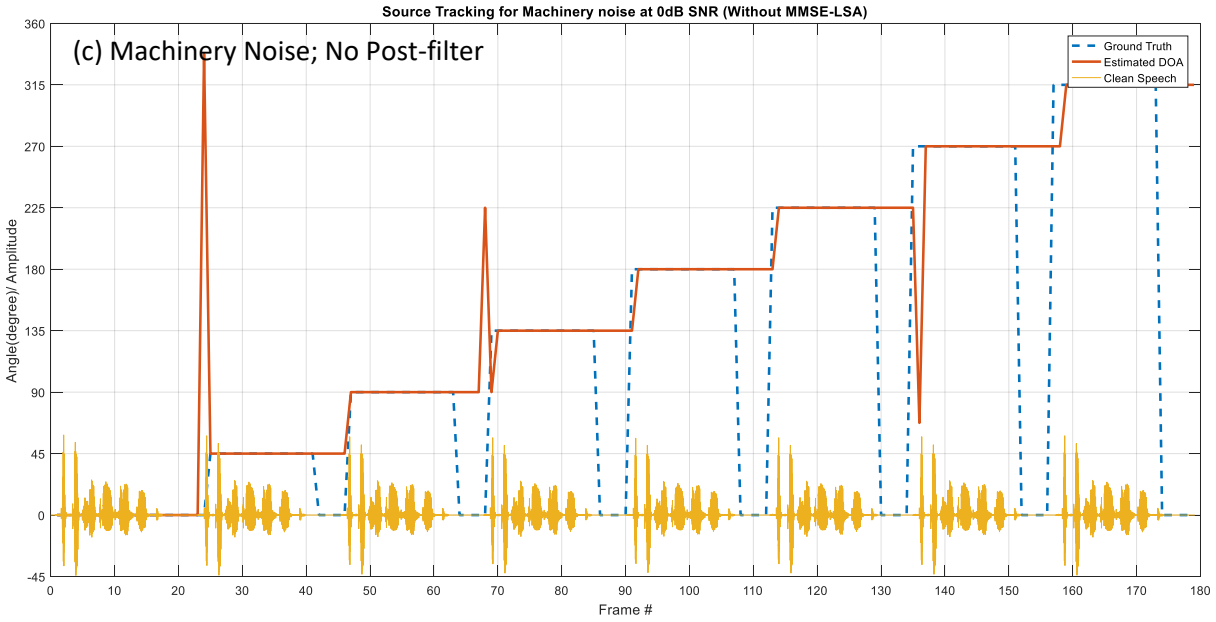
Figure 2. Quantitative analysis of the Proposed DOA estimation method versus CC-based method in [38]. Lower values of RMSE and MAE (both mean and std. dev.) are preferred.

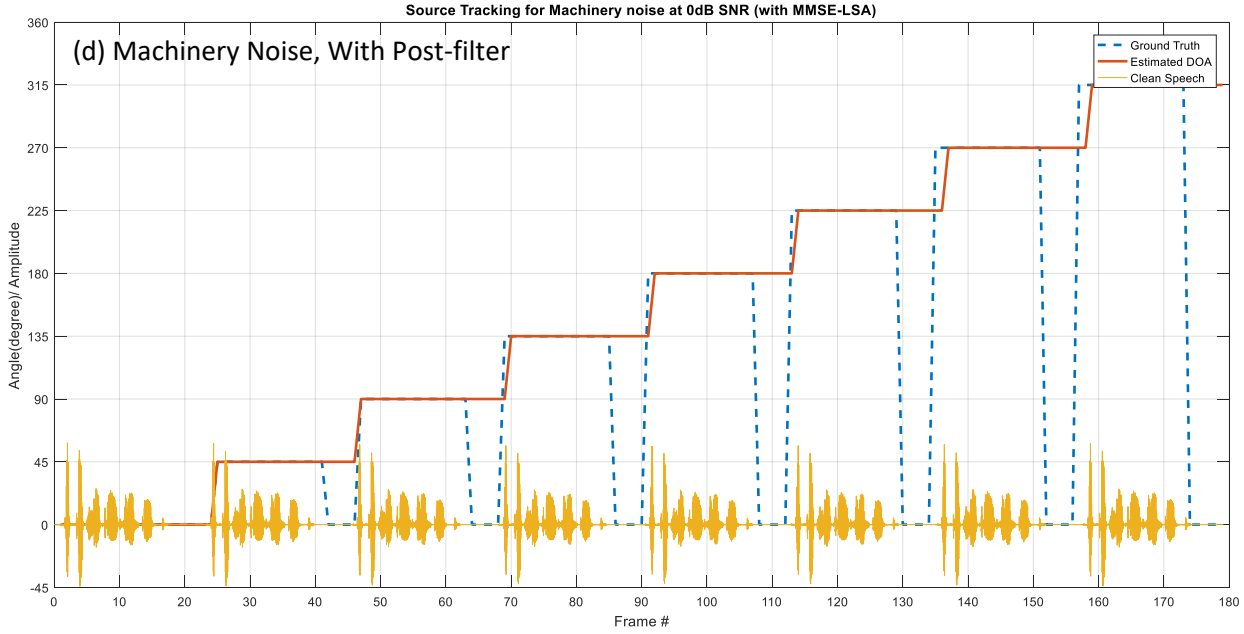
Experiment 2: DOA Tracking



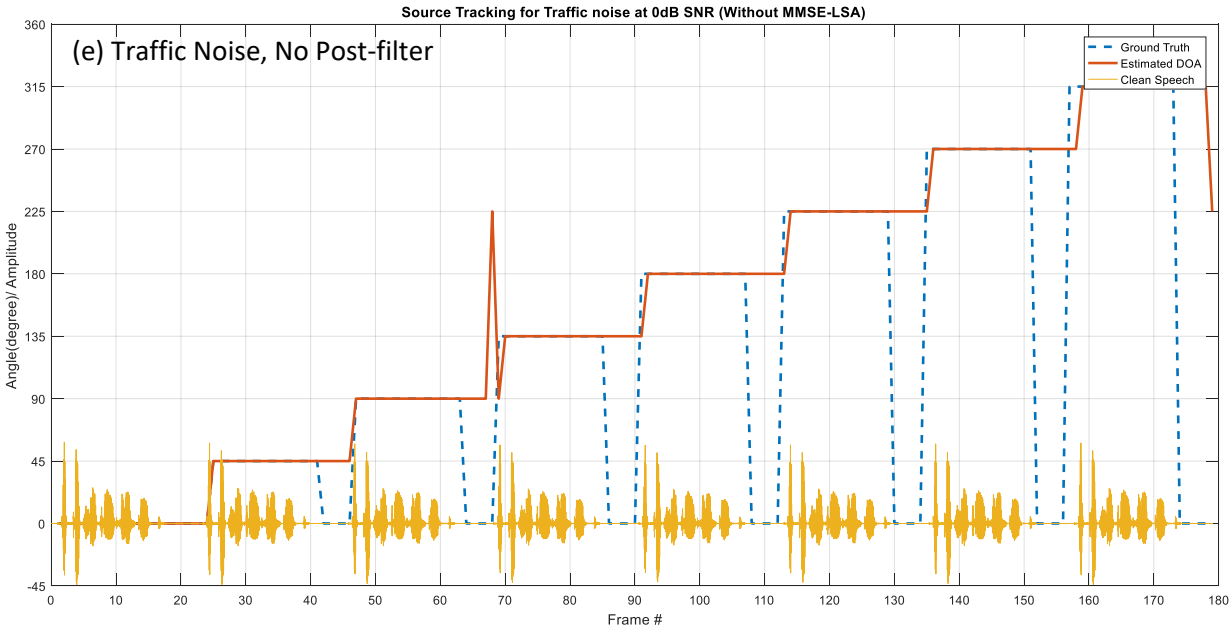


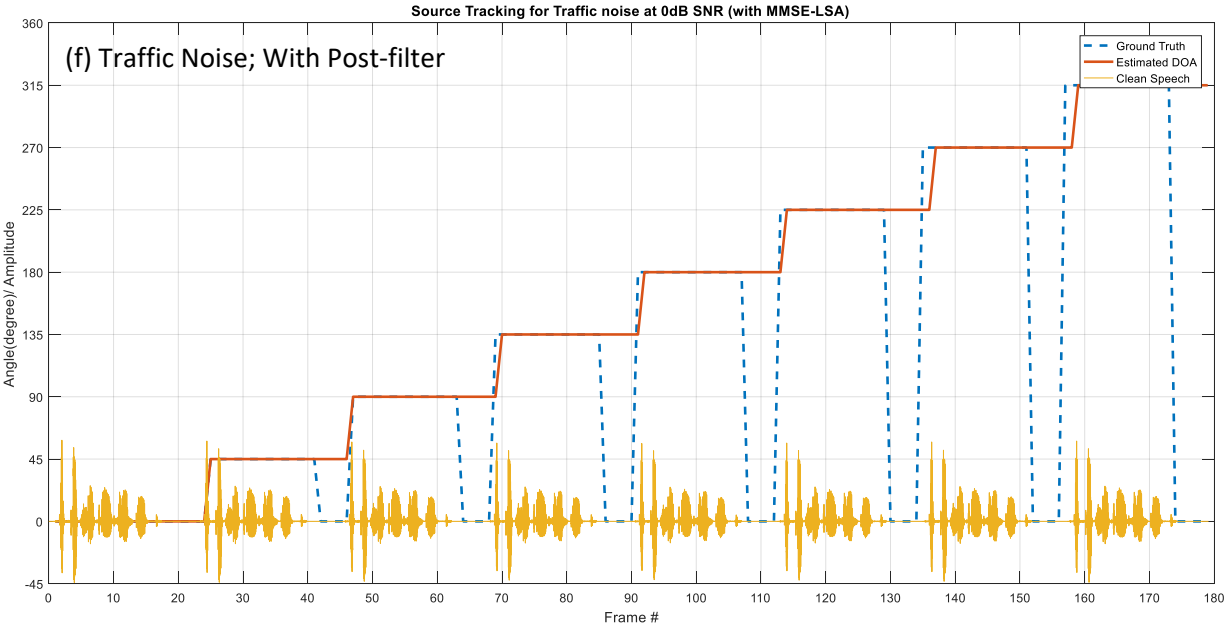
Refer: https://www.utdallas.edu/ssprl/hearing-aid-project/research-papers/





Refer: https://www.utdallas.edu/ssprl/hearing-aid-project/research-papers/





Refer: https://www.utdallas.edu/ssprl/hearing-aid-project/research-papers/

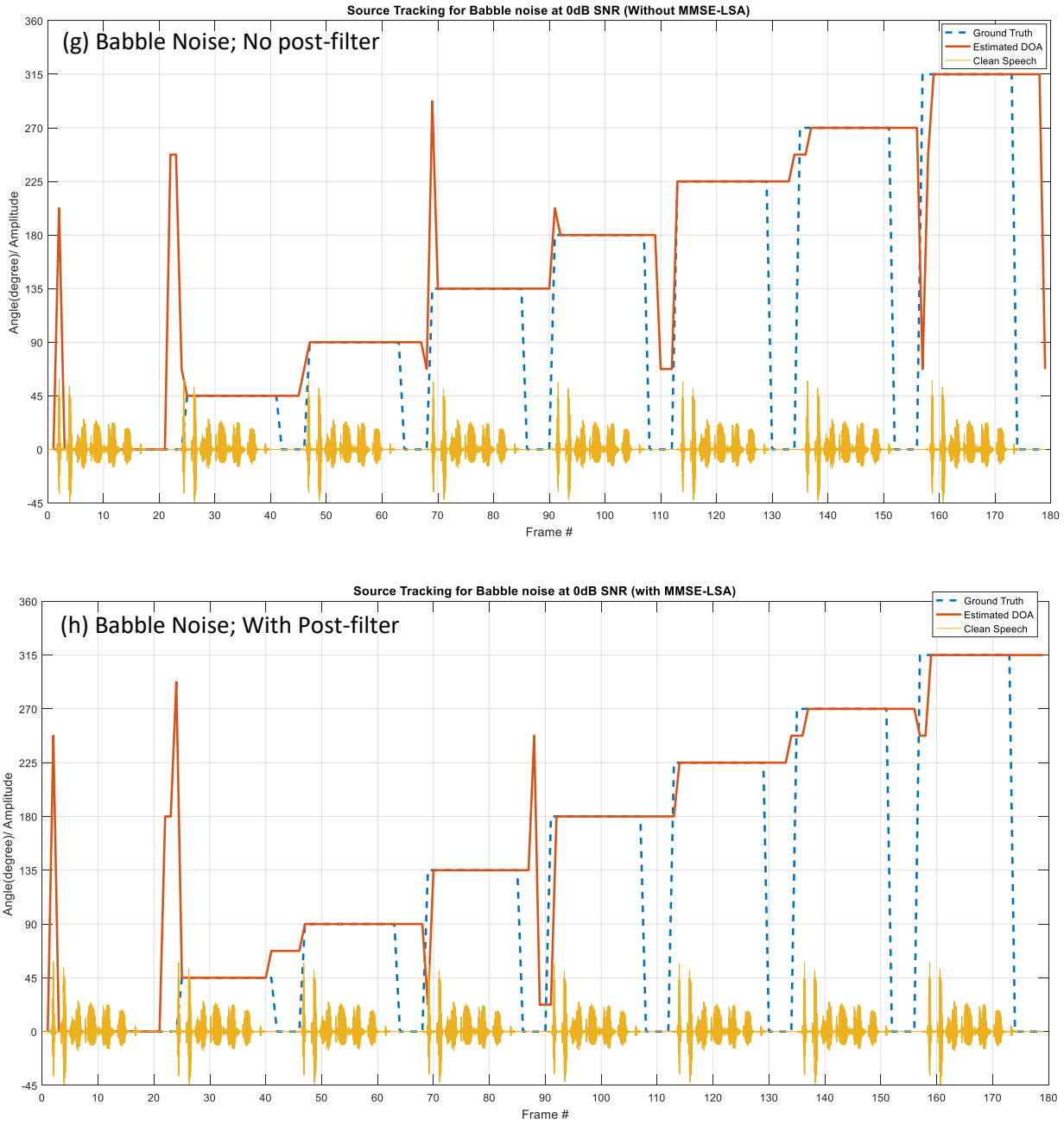


Figure 3. Source tracking results for different noise types at 0dB SNR for (a)White noise,(c) Machinery noise, (e) Traffic noise, (g) Babble noise without Post filter and for (b)White noise,(d) Machinery noise, (f) Traffic noise, (h) Babble noise with Post-filter (MMSE-LSA).

Real Time Implementation on Matrix Creator/Raspberry Pi

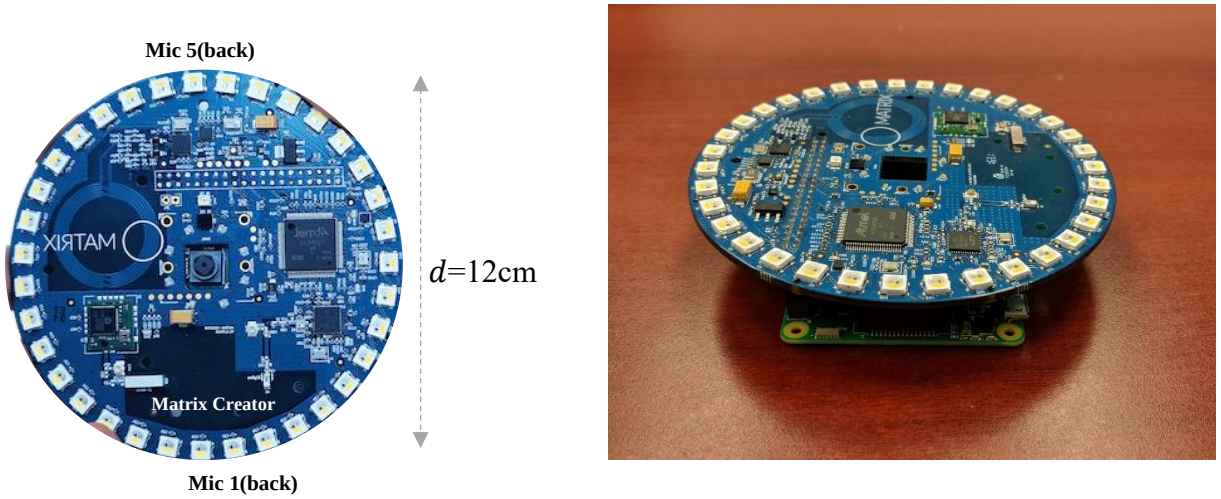


Figure 4. Proposed Hardware –Matrix Creator for Raspberry Pi 3

The proposed method will be implemented on a Matrix Creator (MC) [3] (Figure 4) evaluation board that is driven by a Raspberry Pi3(RPi3) model B. The MC board features an eight omnidirectional MEMS microphone array, controlled by an ARM Cortex M3 microcontroller with a Xilinx Spartan 6 FPGA board. It interfaces to an RPi3 board using the general Purpose Input-Output (GPIO) and the Serial Peripheral Interface (SPI) protocol to transmit commands and audio data in real-time. The MC board collects the raw data from the eight-microphone circular array and streams it to the RPi3 for the signal processing part. The DOA estimate is communicated to the user by a ring of 35 LEDs on the periphery of the MC board (like the Amazon Echo [10]). The enhanced speech can be hard-wired through a 3.5mm stereo output. Alternately, since the RPi3 supports both Bluetooth and Wi-Fi wireless protocols, the DOA estimate and the enhanced speech could be transmitted to either a paired smartphone for additional processing or even a wireless headset. The DOA estimate would be displayed on the smartphone's display, while the enhanced speech could be streamed directly to the user's hearing aids/earphones.

Refer: https://www.utdallas.edu/ssprl/hearing-aid-project/research-papers/

References

[1] Vorobyov, Sergiy A. "Principles of minimum variance robust adaptive beamforming design." *Signal Processing 93.12 (2013): 3264-3277.*

[2] Clénet, Boris, and Harald Romsdorfer. *Circular microphone array based beamforming and source localization on reconfigurable hardware.* Diss. Master's thesis, Graz University of Technology, 2010.

[3] January 2018 [Online] Matrix Creator- The IoT Development Board for Building Incredibly Smart Products. Available: https://www.matrix.one/products/creator.