

ACOUSTIC IMAGING ABOARD THE INTERNATIONAL SPACE STATION (ISS): CHALLENGES AND PRELIMINARY RESULTS

*Luca Bondi^{*1}, Gabriel Chuang^{*2,3}, Christopher Ick^{*1,4}, Adarsh Dave^{*1,3}
Charles Shelton¹, Brian Coltin^{2,5}, Trey Smith², Samarjit Das¹*

¹ Bosch Research ² NASA Ames Research Center ³ Carnegie Mellon University
⁴ New York University ⁵ KBR

^{*}Equal contribution

ABSTRACT

Design and execution of high fidelity acoustic sensing in complex environments poses a number of practical challenges, from accurately measuring the geometry of the setup and estimating the channel response, to time synchronization amongst the sources and receivers. When acoustic experiments are performed on-board the International Space Stations (ISS), the number of constraints and obstacles vastly increases, due to the combination of a highly unpredictable acoustic environment, and restricted availability of crew time. In this paper, we present our preliminary results with a *first-of-a-kind* acoustic imaging experiment performed aboard the ISS, highlighting the difference between simulations, laboratory measurements, and in-space experiments. We hope that these experiments and results will help the research community in realizing high performance acoustic imaging capabilities in complex environments.

Index Terms— autonomous sensing, acoustic imaging, source localization

1. INTRODUCTION

Plans for future human space exploration missions typically include pressurized space habitats that are designed for astronauts, but must also operate uncrewed for long periods. For example, NASA's planned Gateway outpost near the Moon is expected to host astronauts for a few weeks each year [1]. During uncrewed periods, there is a critical need for intra-vehicular robots (IVR) operating within the habitat to provide caretaking functions such as emergency fault recovery, routine maintenance, and supporting on-board science experiments. Autonomous state assessment, including autonomously detecting current anomalies or predicting future anomalies, is a pre-requisite for such functions [2].

Acoustic signals transmitted through structure or air can diagnose many classes of faults. For example, the ISS carries an Ultrasonic Leak Detector (ULD) which is used by astronauts to locate any cabin air leaking into space [3]. The ULD converts ultrasound signals into the astronaut's hearing range.

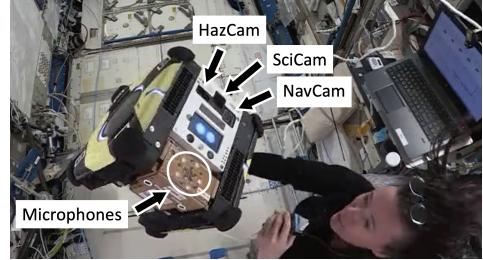


Fig. 1: NASA astronaut Megan McArthur configuring the SoundSee payload for Astrobe on the ISS.

While it has been used successfully for leak isolation, astronauts report that the many background ultrasound sources on the ISS make using the ULD a painstaking process; its single microphone must be scanned very close to any possible leaks to detect faint signals. Our current work aims to automate this leak isolation process with a mobile inspector robot, which could register the acoustic spatial information into a 3D map.

Our target robotic testing platform is NASA's free-flying Astrobe robot, three of which have been operating inside the ISS since 2019 [4]. Their primary function is to provide a platform for micro-gravity robotics research. They also act as free-flying cameras to monitor astronaut activities, and collect sensor surveys. Each Astrobe has three payload bays available to host guest science payloads.

The SoundSee payload, developed by Bosch USA in partnership with Astrobotic Technology Inc, is a microphone array with 20 individual MEMS capsules, capable of capturing audio recordings at sampling rates varying from 44.1kHz to 192kHz [5]. Fig. 1 depicts the SoundSee payload installed on Honey, one of the three Astrobies in orbit, during a recent ISS activity. SoundSee endows Astrobe with the capacity to hear, making acoustic imaging a concrete possibility onboard the ISS.

Acoustic imaging techniques have seen success in several practical industrial and academic [6] domains, such as medical monitoring [7, 8], machine fault detection [9], and bio-acoustics [10]. Arrays of microphones used in specific geometries with classical beamforming techniques can be used

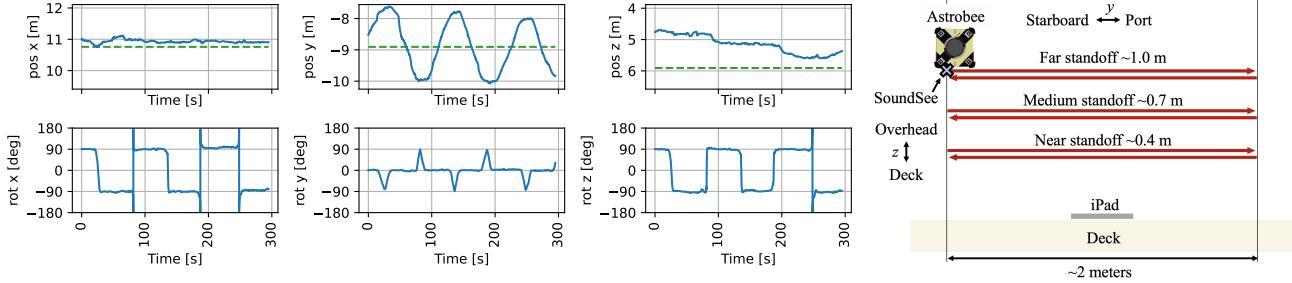


Fig. 2: Manual transects trajectory of SoundSee on the ISS (blue lines). iPad position is represented by dashed green lines. The array is moved forward on the y axis for $\sim 2\text{m}$, rotated around the z axis by 180° then moved backward on the y axis to the starting point, three times at different standoff distances z from the iPad. Absolute position (*top row*). Intrinsic rotation angles (*bottom row*). Nominal transect geometry (*right*). The microphone array is always facing downwards.

to estimate direction-of-arrival from incoming sound sources [11]. Combining these techniques with synthetic aperture methods using mobile recording platforms and specific travel trajectories allows relatively simple microphone array platforms to accurately reconstruct acoustic scenes [12]. Autonomous robotic platforms such as Astrobee, combined with the SoundSee payload have the potential to use these methods to create an automated IVR system for ULD.

In this paper, we present the challenges inherent in conducting in-orbit acoustic imaging experiments, introducing the environment, the constraints, and the time-synchronization difficulties encountered in Sec. 2. Sec. 3 illustrates how signal-processing techniques are used to align data-streams from different clock domains. Sec. 4 describes the simulation and laboratory setup used for comparison to the in-orbit experiment. Sec. 5 shows the results obtained with baseline beamforming techniques for source localization.

2. ENVIRONMENT AND SETUP

Most Astrobee activities use the Japanese Kibo module of the ISS [13], where the Astrobee docking station is located. Kibo is the largest ISS module, with $\sim 11 \times 2.5 \times 2.5\text{ m}$ accessible volume in the cabin. Like most ISS modules, it hosts a variety of science experiment payloads and astronaut activities, and its internal configuration changes substantially over time.

SoundSee mounts in Astrobee’s bottom forward payload bay, with its microphone array on the lower part of Astrobee’s forward face. Three of Astrobee’s cameras are on the upper part of its forward face, with fields of view that overlap with SoundSee: the SciCam (a camera which transmits streaming video of astronaut activities), the NavCam (a wide field of view camera used for navigation), and the HazCam (a LIDAR sensor that returns 3D point clouds useful for hazard avoidance as well as mapping 3D structure) as seen in Fig. 1.

Astrobee navigation relies on a localization system that can operate anywhere in the ISS, using a combination of inertial sensing and NavCam imagery that enables both visual odometry and recognizing visual landmarks from a prior map of the ISS interior [14]. The same position and orientation

information that enables navigation can be combined with acoustic imagery and HazCam LIDAR point clouds to form a 3D acoustic map. ISS position coordinates reported in this paper use the standard Space Station Analysis Coordinate System [15], which is also used for Astrobee navigation.

For our initial data collection on the ISS, we opted for a series of moving passes, or transects, over an unmoving sound source. We used a 10.5 inch 2017 iPad Pro, attached to the deck of the Kibo module, to play a stimulus signal through all the four loudspeakers, while moving the Astrobee back and forth over the iPad at varying target distances. The center position of the iPad is estimated via triangulation of NavCam images to be $(10.76, -8.90, 5.91)\text{m}$.

Rather than flying the desired trajectories autonomously with Astrobee, an astronaut manually moved the robot along the transect, which eliminated propulsion noise from the acoustic experiment. Autonomous flight will be addressed in future experiments. Three transects took place, with the Astrobee offset from the iPad by approximately 0.4m, 0.7m, and 1.0m, respectively. Fig. 2 shows position and orientation of SoundSee during the transects.

The SoundSee payload features a custom-built microphone array in a spiral configuration (Fig. 5). Out of the 20 available microphones, for this experiment we use the two outermost rings, featuring five microphones each, on concentric circles of $\sim 6\text{cm}$ and $\sim 2\text{cm}$ diameter, respectively.

Challenges: Rules around safety and acoustic well-being for the astronauts limit the range of sounds that can be played on the ISS, barring the use of colored noise, chirps, or pure tones. We resorted to a 10s music loop with a heavy harmonic structure. The relatively long periodicity is a challenge for beamforming techniques that rely upon the assumption of a stationary source and microphone. Additionally, the acoustic environment of the ISS is uncharacterized, and is impractical to characterize due to a lack of crew time, necessary equipment, and environmental variability; this also makes the acoustic environment very hard to recreate on the ground. Time-synchronization across data-streams is crucial to consider, as both imaging and synthetic aperture methods rely on

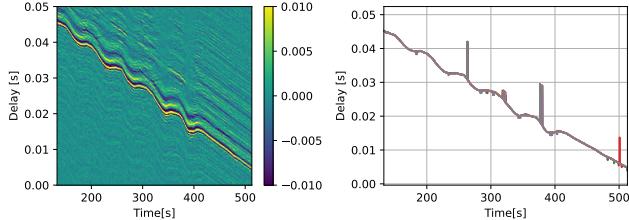


Fig. 3: Time-varying delay between $s(t)$ and $r_i(t)$ before drift compensation. Time-delay matrix $\hat{\Delta}_{sr_i}(t)$, color represent the normalized cross-correlation value (*left*). Time-delay signal $\hat{\delta}_{sr_i}(t)$ for the ten microphones; spikes occur when sounds from other sources are significantly louder than the iPad. (*right*).

time-aligned signal readings and position measurements for accurate results. However, technical issues necessitated special considerations to be made to re-align our recordings using available data.

3. TIME ALIGNMENT OF DATA STREAMS

To accurately estimate the offset between the recorded audio stream and the trajectory of the microphone array, we opt for a geometry-driven approach. Specifically, we can estimate the offset via cross-correlation between: 1) the time-varying delay between the stimulus signal and the microphone array, $\delta_{sr}(t)$; and 2) the time-varying distance between the iPad position and the SoundSee position $\gamma_{sr}(t)$.

While $\gamma_{sr}(t)$ is trivial to compute given the trajectory, extracting $\delta_{sr}(t)$ is challenging. As the stimulus signal $s(t)$ is played back from a device (iPad) that is not clock-synchronous with i^{th} microphone array signals $r_i(t)$, two parameters need to be estimated before computing $\delta_{sr}(t)$: the offset o_{sr} and the clock drift d_{sr} between $s(t)$ and $r_i(t)$.

We first estimate o_{sr} via cross-correlation between $s(t)$ and $r_0(t)$, neglecting the presence of a drift between the two signals, and observing that the offset is not a function of the microphone index. We then compute $\hat{\delta}_{sr}(t)$, the time-varying delay between $s(t)$ and $r_i(t)$ without drift compensation. In particular, given $s(t)$ and $r_i(t)$ we first compute $\hat{\Delta}_{sr_i}(t) \in \mathbb{R}^{D \times T}$, a time-delay matrix where each column is the cross-correlation between a window of $s(t)$ and the corresponding window of $r_i(t)$ - Fig. 3 (*left*). As expected, the harmonic nature of the stimulus signal plays a clear role in $\hat{\Delta}_{sr_i}(t)$, as the auto-correlation signature of $s(t)$ is clearly visible in the columns of $\hat{\Delta}_{sr_i}(t)$. After a 2D Gaussian smoothing step, we track the maximum value along the columns of $\hat{\Delta}_{sr_i}(t)$ to build $\hat{\delta}_{sr_i}(t)$, the time-varying delay signal between $s(t)$ and $r_i(t)$ - Fig. 3 (*right*).

By averaging over the active microphones, we obtain $\hat{\delta}_{sr}(t)$, the average time-varying delay between the source signal and the microphone array, prior to drift compensation. We then compute the clock drift d_{sr} via linear regression on $\hat{\delta}_{sr}(t)$ between seconds 420 and 500. The drift-compensated average time-varying delay $\delta_{sr}(t)$ between $s(t)$ and $r_i(t)$ can

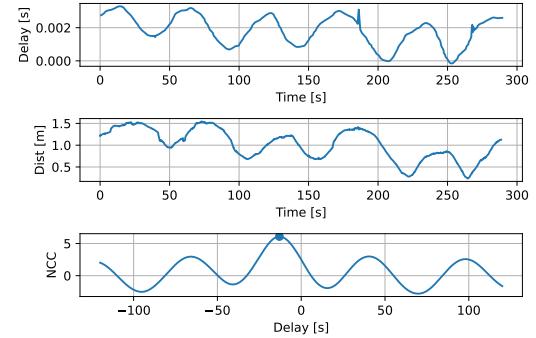


Fig. 4: Drift-compensated time-varying delay $\delta_{sr}(t)$ (*top*); time-varying distance between source and array $\gamma_{sr}(t)$ (*middle*); normalized cross-correlation between the two signals (*bottom*)



Fig. 5: Back-side of microphone array used in SoundSee unit (*left*); marked in green are the outer rings of 10 microphones used to capture data. Image of SoundSee mounted on robotic platform inside sound chamber for experiments (*right*).

be easily derived from $\hat{\delta}_{sr}(t)$ and d_{sr} .

Fig. 4 shows the drift-compensated time-varying delay $\delta_{sr}(t)$, the time-varying distance between source and SoundSee $\gamma_{sr}(t)$, and the normalized cross-correlation between the two signals, used to extract the offset between the recorded audio stream and the trajectory of the microphone array.

4. SIMULATIONS AND LAB EXPERIMENTS

To estimate the quality of imaging results from the ISS experiment, we replicate both in simulation and in a controlled lab environment the same transect acquisition.

Acoustic simulations. Simulated experiments are computed in a virtual acoustic environment using PyRoomAcoustics (PRA)[16], a Python module for acoustic simulation in specific room configurations using the image source method. Using a shoebox-style model of a room, we design an acoustic space mirroring the approximate dimensions of the Kibo module on board the ISS. The iPad is modeled as an omni-directional point sound source, whereas the SoundSee microphone array unit is modeled as a set of omni-directional point microphones, in the same dual ring configuration as on hardware. The image source method used in PRA is designed to compute the room impulse response (RIR) between a given source location and a given microphone location. However, this RIR is computed with the assumption of neither the mi-

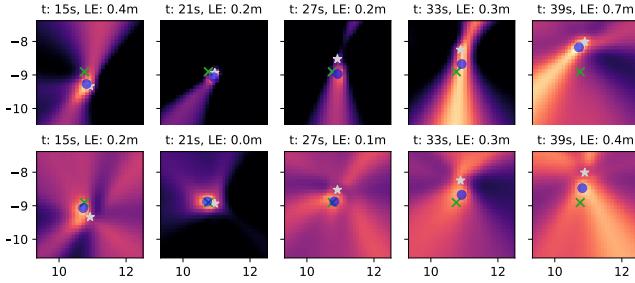


Fig. 6: Acoustic images computed for ISS data (*top row*) and simulated data (*bottom row*). Green cross: source position. White star: array position. Blue dot: estimated source position. Titles report timestamp (t) and localization error (LE).

crophone nor source being in motion. We assume that the microphone position is stationary in each trajectory point for the duration of the analysis window. From here, we use the audio centered at each corresponding time-step and location, along with the computed RIR, to simulate the microphone array response at that time and location.

Experiments in SoundSee lab. Controlled acoustic experiments were conducted in Bosch’s Pittsburgh office, inside a sound isolation chamber (WhisperRoom SE 2000) measuring 2.4m by 4.1m, without any modification to its native acoustic properties. Audio was piped into the chamber with a Genelec 8320A speaker, replacing the iPad on the ISS without further source modelling. Object locations in the chamber were tracked by a Motive OptiTrack motion capture system, featuring four ceiling-corner mounted USB cameras. OptiTrack systems are broadly capable of tracking with sub-10 millisecond latency, within 2 millimeter accuracy. All audio was recorded using a SoundSee unit identical to the one on the ISS. A SuperDroids Mecanum robot was used for motion control in the anechoic chamber during acoustic captures. SoundSee was attached to 4 ft speaker mount on the robot to control height and angle of the microphones.

5. RESULTS

Acoustic images are generated with the Music [17] beamforming algorithm as implemented in Acoular [11]. The imaging plane spans x and y axes for the width and depth of the Kibo module, with a resolution of 10cm. Analysis is focused on a third octave bandwidth, centered at 6kHz, where resolution of the array is maximized and aliasing artifacts are negligible. The position of the source is estimated to be on the grid point with the largest response to the beamforming algorithm. Fig. 6 shows some sample acoustic images from ISS data, overlaid with the array, the real, and the estimated source positions.

Fig. 7 shows the localization error as a function of the displacement between the microphone array and the iPad along the y axis for each analysis point. The blue dots (ISS) result from analyzing the ISS manual transect at 70cm standoff distance. The orange crosses (Lab) result from the lab acqui-

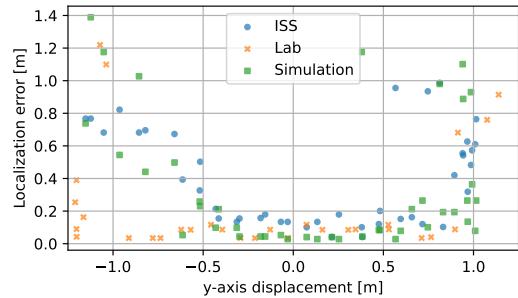


Fig. 7: Localization error as a function of the displacement between the microphone array and the source along y axis.

sition of a transect at the same standoff distance. The green squares (Simulation) are obtained by simulating the ISS trajectory in a room with an RT60 of 2s. All the walls in the simulated room have the same frequency-independent absorption coefficient of 0.13. The same music stimulus used on the ISS is used for both simulations and lab experiments.

As expected, the localization error increases as the distance between the source and the array increases. While the localization error for the simulated scenario and the lab acquisition is smaller than from ISS data, both the position of the SoundSee unit on the ISS and the position of the iPad are estimated via NavCam images and other sensors, and are thus subject to noise. Interestingly, changing the absorption coefficient of the simulated room is enough to get a reasonable approximation of the ISS localization error, despite treating the iPad as a spherical point source, and neglecting the polar pattern of the array capsules. This conclusion would imply that the lab experiment, conducted in a semi-anechoic chamber, may yield lower localization errors, as observed in Fig. 7.

All three experiments show large localization errors at greater than 1m displacement on either side of the speaker, potentially due to the speaker’s directionality. While the ISS and simulation series show linearly increasing errors between 0.5 and 1m to the left of the speaker, this is not shared by the lab series. This behavior will be subject to further investigation.

6. CONCLUSIONS

We presented the challenges and possible solutions to consider before computing acoustic imaging and source localization in a highly unpredictable acoustic environment. Preliminary results show that small changes to a simulated environment can be used to approximate the error distribution that could be expected in the target environment. Future works will leverage this experience to prepare for in-space acquisitions with the aid of simulations and laboratory measurements, more advanced beamforming algorithms [18], by characterizing sources and receiver, setting expectations and validating hypotheses on the uncertain and constantly-changing acoustic landscape on the ISS.

7. ACKNOWLEDGEMENT

We would like to thank the ISS National Laboratory for sponsoring our space mission. We would also like to thank astronauts Megan, Vick, Shannon, and Kayla for performing the on-orbit experiments with SoundSee. Finally, we would like to thank our partners on the ISS Astrobee Facility team, the ISAAC project team, and all the other ISS team members who worked so diligently behind the scenes to make these experiments possible.

8. REFERENCES

- [1] Jason C. Crusan, R. Marshall Smith, Douglas A. Craig, Jose M. Caram, John Guidi, Michele Gates, Jonathan M. Krezel, and Nicole B. Herrmann, “Deep space gateway concept: Extending human presence into cislunar space,” in *2018 IEEE Aerospace Conference*, 2018, pp. 1–10.
- [2] Trey Smith, Maria Bualat, Abiola Akanni, Oleg Alexandrov, Laura Barron, J Benton, Gabriel Chuang, Brian Coltin, Terry Fong, Janette Garcia, Kathryn Hamilton, Lewis Hill, Marina Moreira, Robert Morris, Nicole Ortega, Joseph Pea, Jonathan Rogers, Misha Savchenko, Khaled Sharif, and Ryan Soussan, “ISAAC: An integrated system for autonomous and adaptive caretaking,” in *Proc. Int. Space Station Res. Dev. Conf. (ISS-RDC)*, 2021.
- [3] Willam C. Wilson, Neil C. Coffey, and Eric I. Madaras, “Leak detection and location technology assessment for aerospace applications,” Tech. Rep. NASA/TM-2008-215347, NASA Langley Research Center, 2008.
- [4] Maria G. Bualat, Jonathan S. Barlow, Jose V. Benavides, Brian Coltin, Lorenzo J. Flückiger, Marina Gouveia Moreira, Kathryn Hamilton, and Trey Smith, “Astrobee on-orbit commissioning,” in *Proc. IAF Int. Conf. Space Ops*, 2021.
- [5] Neel V. Patel, “Nasa will use a robot to listen out for danger on the iss,” *MIT Technology Review*, Oct 30, 2019.
- [6] Holger Hewener, Christoph Risser, Selina Barry-Hummel, Heinrich Fonfara, Marc Fournelle, and Steffen Tretbar, “Integrated 1024 channel ultrasound beamformer for ultrasound research,” in *2020 IEEE International Ultrasonics Symposium (IUS)*, 2020, pp. 1–4.
- [7] Xiaoqian Song, Maokun Li, Fan Yang, Shenheng Xu, and Aria Abubakar, “Feasibility study of acoustic imaging for human thorax using an acoustic contrast source inversion algorithm,” *The Journal of the Acoustical Society of America*, vol. 144, no. 5, pp. 2782–2792, 2018.
- [8] Mark Haynes and Mahta Moghaddam, “Large-domain, low-contrast acoustic inverse scattering for ultrasound breast imaging,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 11, pp. 2712–2722, 2010.
- [9] E. Cardenas Cabada, Q. Leclere, J. Antoni, and N. Hamzaoui, “Fault detection in rotating machines with beamforming: Spatial visualization of diagnosis features,” *Mechanical Systems and Signal Processing*, vol. 97, pp. 33–43, 2017, Special Issue on Surveillance.
- [10] Matthew Wijers, Andrew Loveridge, David W. Macdonald, and Andrew Markham, “Caracal: a versatile passive acoustic monitoring tool for wildlife research and conservation,” *Bioacoustics*, vol. 30, no. 1, pp. 41–57, 2021.
- [11] Ennes Sarradj, Gert Herold, Adam Kujawski, Tom Genesch, Simon Jekosch, Mikolaj Czuchaj, and Art Pelling, *Acoular*, 2015.
- [12] Boqiang Fan and Samarjit Das, “Synthetic Aperture Acoustic Imaging with Deep Generative Model Based Source Distribution Prior,” in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, June 2021, pp. 1420–1424, IEEE.
- [13] Human Space Systems and Utilization Program Group, “Kibo handbook,” Tech. Rep., Japanese Aerospace Exploration Agency (JAXA), 2007.
- [14] Brian Coltin, Jesse Fusco, Zack Moratto, Oleg Alexandrov, and Robert Nakamura, “Localization from visual landmarks on a free-flying robot,” in *Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS)*, 2016, pp. 4377–4382.
- [15] Johnson Space Center National Aeronautics and Space Administration, “Space station reference coordinate systems,” Tech. Rep., International Space Station Program, 2001.
- [16] Robin Scheibler, Eric Bezzam, and Ivan Dokmanić, “Pyroomacoustics: A Python package for audio room simulations and array processing algorithms,” *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 351–355, Apr. 2018, arXiv: 1710.04196.
- [17] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, mar 1986.
- [18] Yujie Gu and Amir Leshem, “Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3881–3885, 2012.