PROCEEDINGS of the
24th International Congress on Acoustics

October 24 to 28, 2022 in Gyeongju, Korea

Recalibration of the NASA Exterior Effects Room

E.K. Ellington SCOTT[1]; Andrew CHRISTIAN[2]; Matthew BOUCHER[2]; Stephen RIZZI[3];
Aric AUMANN[4]; Haik MARTIROSYAN[5]

[1] Graduate Program in Architectural Acoustics, Rensselaer Polytechnic Institute (NIFS Intern), USA

[2] Structural Acoustics Branch, NASA Langley Research Center, USA

[3] Aeroacoustics Branch, NASA Langley Research Center, USA

[4]Analytical Services and Materials, USA

[5]Zeiss Industrial Metrology (NIFS Intern), USA

ABSTRACT

The Exterior Effects Room (EER) is a psychoacoustic testing facility at NASA Langley Research Center which primarily focuses on testing human response to aircraft noise. The 39-seat auditorium houses a real-time spatial audio system and flyover simulation environment. The audio server utilizes an implementation of three-dimensional vector base amplitude panning (VBAP), a perceptual spatial audio technique exploiting loudspeaker triplets to place a virtual sound source at an arbitrary spatial position. Due to the irregular room geometry and nonuniform loudspeaker setup, the audio server applies equalization filtering to compensate for spectral coloration attributed to loudspeaker installation, crossover filtering, and delay/gain offsets. These filters can incorporate measurements taken at multiple listening points, allowing a more extensive listening area than filters derived from measurements taken at a single listening point. This work endeavors to update the existing equalization filter generation process by producing a new set of filters for various human subject testing scenarios.

Keywords: Equalization, Vector Base Amplitude Panning, Spatial Audio.

1. INTRODUCTION

NASA has a long history of performing laboratory psychoacoustic research for community noise applications across the aerospace industry: from fixed wing aircraft and rotorcraft of all shapes and sizes (see Rizzi for a recent overview (1)), to current efforts to produce quiet supersonic aircraft (2),

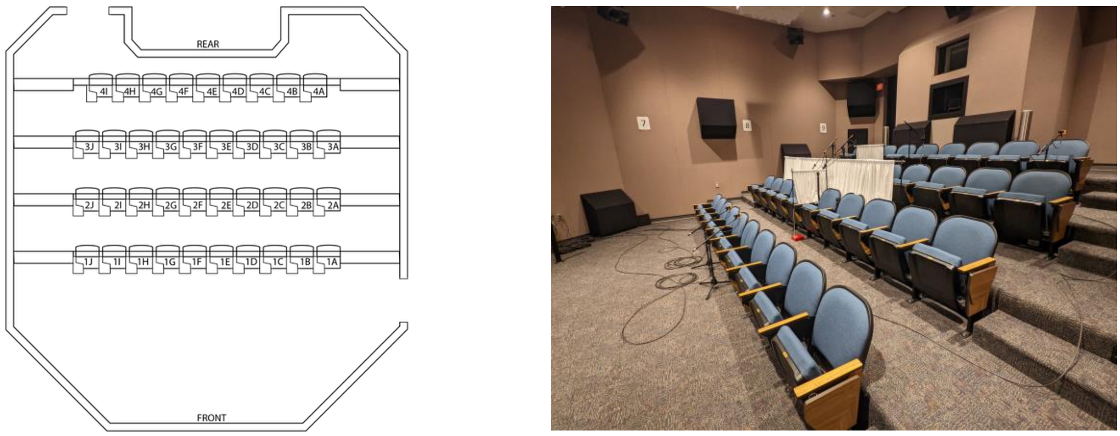

Figure 1 – Exterior Effects Room: (a) Floor plan, (b) Photograph

to early versions of wind turbines (3) and beyond. One of the psychoacoustic test facilities that has seen a considerable amount of use over the past decade is the Exterior Effects Room (EER) at Langley Research Center. Originally designed as a small auditorium, it was retrofit with speakers and acoustical treatments to allow for use as a psychoacoustic testing facility (4). This facility was well suited for presentation of recorded aircraft flyover noises given its seating arrangement. Originally, single-event psychoacoustic testing required large panels of actual subjects sitting in the field (i.e., outdoors, on chairs, responding via pencil and paper, e.g., Beranek (5)). This required extensive coordination between pilots attempting to fly different aircraft consistently in the prevailing (and ever changing) environmental conditions. In a laboratory environment such as the EER, researchers can control various aspects of the acoustic scene and be sure that there is a high degree of reproducibility between presentations (e.g., to different groups of subjects).

Originally, an array of 10 loudspeakers was used to render moving noise sources in the EER, and these speakers were manually equalized to the center of the room using the electroacoustic equipment available at the time. In the late 2000s the EER was completely renovated into its current configuration (6). The setup now utilizes a 31-channel playback system of twenty-seven Klein & Hummel O300 and four Klein & Hummel O900 subwoofer loudspeakers (**Figure 1**). Satellites are arranged roughly in three rings in the EER, with those in the lower two rings baffled to protect the loudspeakers from accidental contact by researchers and subjects. The subwoofers are fully embedded in the corner facets of the EER.

Although it is possible to use the facility with any multichannel audio playback device – patching into the speakers can be accomplished either via a ¼" TRS patchbay or "soft patched" via AVB-networked audio devices – the primary playback device is a Vectsonic AuSIM3D audio rendering server (7). This system implements a Vector Base Amplitude Panning (VBAP) approach to spatialize sounds within the EER. VBAP is an extension of stereophonic panning that can create the impression of a sound coming from anywhere in the EER (6), and the server can apply this technique both to recorded/auralized sources (e.g., rotorcraft flyover noise that was captured in synchrony with GPS data (8)) as well as to real-time inputs. This system is integrated with a three-dimensional visual environment that, while not typically used for psychoacoustic testing purposes, has great utility as a communication tool (e.g., to stakeholders and the public).

In addition to the VBAP processing, the AuSIM3D server can apply compensatory equalization filters for each speaker in the array. Originally, these filters (along with compensation for differing time-of-flight delay and gain) were derived for two multichannel microphone arrays within the EER (6). Since that work, the renovated EER has been used for many psychoacoustic tests (e.g., most recently (8)). Subjects typically sit in 4 seats near the geometric center of the room,[1] however, the original equalization process was undertaken before any of these tests. This fact, combined with recent upgrades to the playback system – from an all-analog system to one built around an AVB audio network, in addition to any effects of aging/replacement of audio equipment – make it a prudent time to resurrect the equalization process and undertake it for the system as it exists today.

This paper details the procedure to recalibrate the EER multichannel audio system for different subject configurations using the current hardware and physical configuration. First, the EER VBAP system is described. Then the method of deriving equalization filters from recordings of pink noise at various locations (including subject seat locations) is developed. Example filter shapes are shown for combinations of these recordings.

2. VECTOR BASE AMPLITUDE PANNING

The common left-right stereophonic panning scheme, as in most contemporary recorded music, applies differential gain between two channels of audio in order to create the illusion that a sound is coming from a particular azimuthal/horizontal direction, as shown in the left of Figure 2 (see e.g., Begault (10)). VBAP generalizes this concept to 3 Cartesian (or 2 polar) dimensions to derive gains for triplets of speakers to generate localization cues in both the azimuthal and elevation/vertical directions.

Given the unit vectors of the loudspeaker positions, $\boldsymbol{l}_{(\cdot)} = [l_x \quad l_y \quad l_z]^T$, a loudspeaker triplet is used to reproduce a virtual source within a triangle formed by the 3 locations (Figure 2R). The gain for each loudspeaker is calculated using a linear combination of the loudspeaker vectors,

[1] A series of recent tests have also used a single subject location near one of the subwoofers. See, most recently, Rafaelof et al. (9).

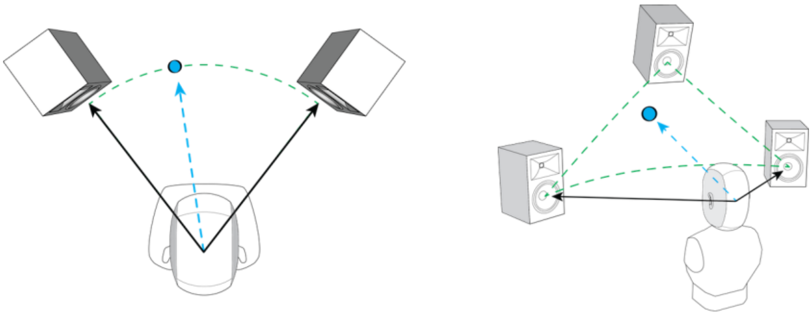

Figure 2 – VBAP loudspeaker arrangements where blue is the virtual source vector and black are the loudspeaker vectors. (a) Stereo pair. (b) Triplet.

$$\begin{gathered}\boldsymbol{p} = g_n \boldsymbol{l}_n + g_m \boldsymbol{l}_m + g_k \boldsymbol{l}_k, \\ \boldsymbol{g} = \boldsymbol{p}^T \boldsymbol{L}_{nmk}^{-1}\end{gathered} \tag{1}$$

where $\boldsymbol{g}$ is the gain vector, $\boldsymbol{p}$ is the virtual source vector, and $\boldsymbol{L}_{nmk} = [\boldsymbol{l}_n \quad \boldsymbol{l}_m \quad \boldsymbol{l}_k]^T$.

The EER spatial audio system augments VBAP with virtual source positions to remedy triangle ambiguity between loudspeaker triplet sets. The arrangement of the ceiling loudspeakers creates indistinguishable triangles – loudspeakers arranged in a rectangle form two geometrically similar triangles. This phenomenon leads to undesirable level changes when panning through the triplets. Consequently, virtual loudspeakers were used to create unique triangle sets (**Figure 3**). Gains are calculated using the virtual loudspeaker vectors and then distributed between the physical loudspeakers within the neighboring loudspeaker triplets. For example, for a sound positioned within

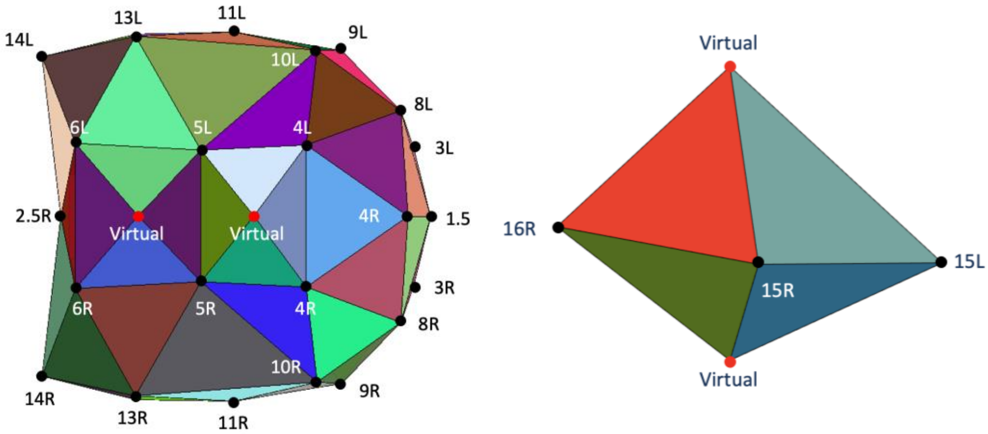

Figure 3 – EER VBAP triangles of loudspeaker triplets. Physical loudspeaker positions are black. Virtual loudspeakers positions are red. (a) Satellite triangle sets. (b) Subwoofer triangle sets.

the triangle 6R, 6L, and Virtual, the gains are calculated using normal VBAP. The gain assigned to the virtual speaker is then shared equally among its neighbors (5L, 5R, 6L, 6R), so that 6L and 6R will have more gain than 5L and 5R. The same technique was also applied to the subwoofers: virtual sources were placed above and below the plane of the 4 subwoofers in order to render sounds coming from any elevation. EER VBAP processing also uses spread angle to minimize gain artifacts that can become apparent when a source location results in playback through a single loudspeaker. This is done by defining a small ring of virtual sources around the current source position (6).

3. LOUDSPEAKER EQUALIZATION

Equalization for each loudspeaker is individually calculated given a measurement at a set of listening locations; one speaker's equalization is not dependent on another's. To implement equalization of each loudspeaker in the EER, three main components are required: equalization filters in second-order sections structure, relative gain compensation, and time delay compensation. The following sections describe the methods utilized to calculate these elements of equalization.

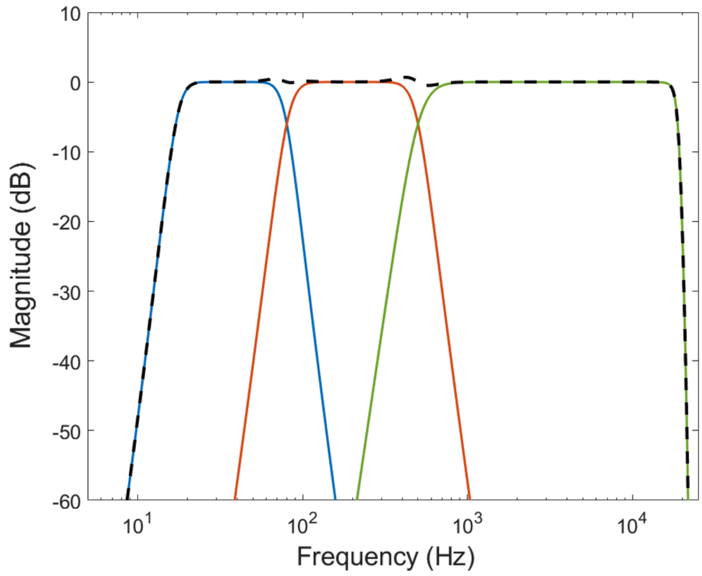

Figure 4 – Desired response of subwoofers (blue), satellites (orange, green), and the entire frequency range of the EER system at a sampling frequency of 48 kHz (dashed).

3.1 Equalization Filter

Implementation of the equalization accounts for both spectral coloration and crossover filtering. Conceptually, this is accomplished by inverting the difference between the measured response and the desired response shown in Figure 4. An eighth-order Linkwitz-Riley filter[2] was used as the desired response due to minimum reaction at the crossover positions.

While an exact inverse of the frequency response may be obtained for a single measurement with zero error, the limitations of this simplistic process may cause the resulting filter to exhibit undesirable artifacts such as excessive gain when compensating for a null caused by a reflection. Spatial averaging is considered in the equalization process by incorporating the frequency response of multiple measurement locations within the listening space. The following sections will discuss the techniques used to generate the equalization filters.

3.1.1 FIR Wiener Filters

Wiener filtering, a linear optimal filtering, allows for the estimation of a desired response assuming a linear-time invariant system. The process minimizes the mean-squared error between the desired and measured system response (11). This process is carried out via Toeplitz matrix operations (12). The number of FIR taps used for the satellite and subwoofers are 8,192 (2^{13}) and 32,768 (2^{15}), respectively. More taps are needed for the subwoofers to generate FIR filters that have the necessary resolution at low frequency. Incorporating multiple responses in the same filter improves the conditioning of the matrix in the Wiener filtering process. Unfortunately, using these high-order FIR filters mean that their implementation is both computationally very expensive and would add a prohibitive amount of delay into the system for tracked sources. Therefore, the direct use of the FIR filters is restricted to offline simulation.

3.1.2 IIR Filters

Lower order IIR filters which approximate the large FIR filters were computed to allow real-time playback within the Vectsonic system. Minimum-phase IIR filters which estimate the magnitude responses of the FIR filters were generated via a Gauss/Newton spectral factorization method (13). To improve frequency resolution of the satellite IIR approximation, the FIR filter approximation divided the passband into two segments: 80 Hz to 500 Hz and 500 Hz to 18.75 kHz.[3] If not split up,

---

[2] That is to say that this has an equivalent response to two cascaded 8th-order Butterworth filters.

[3] The upper frequency limit is determined by the sampling frequency of the generated set. It is $f_s/2.56$, or 18.75 kHz for the 48 kHz filters shown in this paper.

the IIR estimator will tend to assign too many points of control to higher frequencies, as frequencies are linearly weighted in the process. The two segments are then joined to form a single IIR filter. Orders were reduced significantly to 25 second-order sections and 55 second-order sections for subwoofers and satellite loudspeakers, respectively.

3.2 Time Delay Compensation

The system time delay between input and output was calculated using a peak detection algorithm in the time domain for a single measurement point (typically the center of the room). The compensated delay was calculated by subtracting each delay estimate from the maximum measured delay; the delay differences were then added to the uncompensated delays. These delays account for the full path, including the digital signal path and geometric position relative to the selected measurement position. This allows the AuSIM3D engine to process the delay in a single line for each loudspeaker (6).

3.3 Relative Gain Compensation

Gain compensation was estimated using the mean of the one-third octave band level of each loudspeaker's passband. The same compensation technique was utilized with relative gain as mentioned with delay – levels were subtracted from the maximum mean gain, compensating each loudspeaker with the difference from the maximum.

4. RECORDINGS AND RESULTS

During the development of the original equalization filters, frequency response measurements were taken at fourteen microphone positions – twelve spaced throughout the designated listening area and two placed near the room center (6). The two-microphone array (2Mic) was used to obtain a small but accurate sweet spot intended for human-subject experimentation (though before it was known where exactly the subjects would sit). The twelve-microphone array was used to obtain a larger sweet spot that was a compromise between the larger distribution of positions and could be used for facility demonstrations to large groups.

During the current equalization process, 21 measurements were taken with locations strategically selected given experiences with past tests. The locations measured are indicated in Figure 5. Eleven seat location recordings were set roughly to a center-of-head position; consistency of this location was maintained by using a jig that fit over the armrests of the seats. The main four subject locations were included in this set (4Mic), along with the extrema of the seating space. Recordings were also taken at locations "in the pit" (or "on the stage," depending on your proclivities), at the ends of an aisle, and at several locations that matched those from the previous effort. All mics faced upwards,

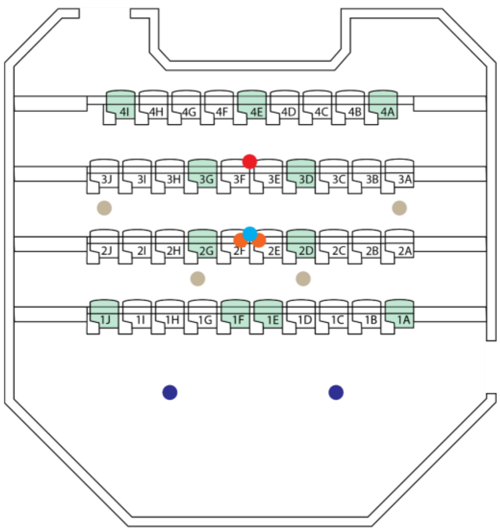

Figure 5 – Frequency response measurements used in updated equalization framework. Positions include center location (cyan), two-microphone configuration (orange), seat locations (green), listening area positions (dark blue, beige, red).

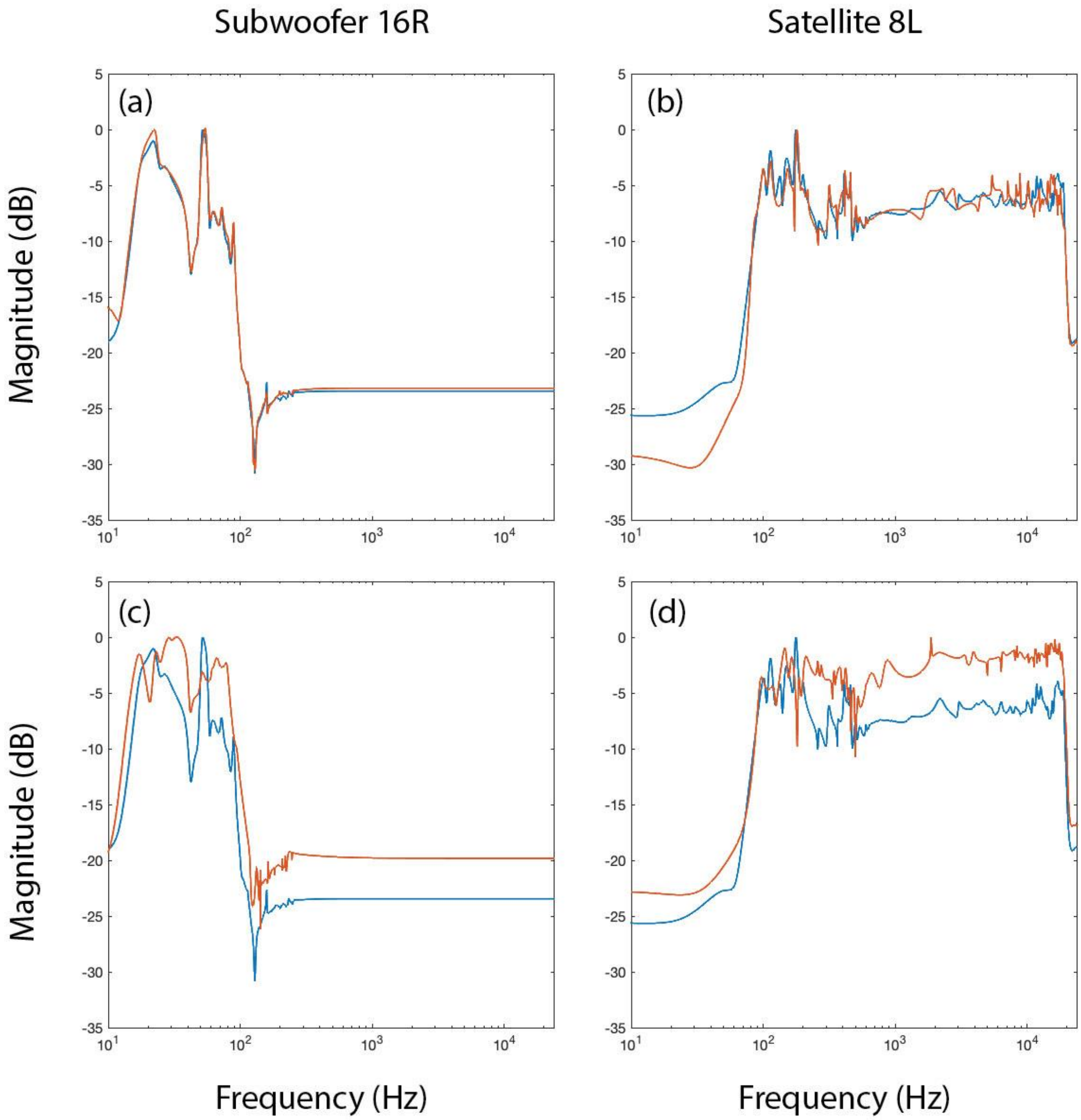

Figure 6 – Comparison between old and new equalization filters: (a, c) Subwoofer 16R, (b, d) Satellite 8L. (a, b) Comparison between old 2Mic filters (blue) and 2Mic recreation with new measurements (orange). (c, d) Comparison between old 2Mic filters (blue) and new 4Mic filters derived from subject seat location measurements (orange).

except for positions that were meant to match those from the previous effort.[4]

The mics used were GRAS 40AQ capsules on 26CA preamplifiers. Two GRAS 12AX ICP power supplies were used to power the mics and apply a +20 dB gain within the EER. The microphone signals were digitized using an RME ADI-8 QS in the nearby EER control room and were transmitted via a MADI connection to a computer running the Reaper DAW. 8 channels were captured at once – 7 microphone "response" signals and a direct reference "stimulus" signal which was a digital copy of the signals being sent to the speakers. As the speakers are self-powered, there are no amplifier settings for each channel that could add complications. Pure tone calibration recordings were captured using a B&K Type 4231 pistonphone.

[4] Prior to the first psychoacoustic test following the latest renovation, a set of white curtains were added to block the view of one subject from another. These curtains, while mainly acoustically transparent, were not installed during the original recording effort. They were included in this set, as they are always now present during human subject testing in the EER.

4.1 Equalization Filters

The recording suite consists of two-minute pink noise excitations per-speaker and per-microphone location. These recordings are processed into incoherently-averaged single-sided power spectral densities via FFTs of 32,768 (2^{15}) samples. At a 48 kHz sampling rate, this results in a binwidth of 1.5 Hz. A 50% overlap is used, to generate 240 averages. This was done for all speaker-microphone combinations – 651 responses. Choosing to include a response (or not) in the equalization process effectively extends the intended "sweet spot" of the listening area to those locations to the greatest extent possible (given the constraints of the other included responses).

Figures 6a and 6b show examples of the original 2Mic equalization filter of a satellite and subwoofer with filters derived from the new measurements (and using the new workflow) that were made to replicate the old 2Mic case. This shows a large amount of agreement between the two efforts. Although there are some minor discrepancies, especially in the higher frequencies (where these sorts of measurements are more sensitive), the trends largely match. These results are typical of the agreement found for all of the speakers between the new and old measurements. Evidence of the linear weighting of the IIR-estimation process is evident by the increasing "jaggedness" of the filter responses towards the top of the partitioned passbands (consider the smoothness of the functions just below and above 500 Hz).

Figures 6c and 6d compare magnitudes of the old 2Mic equalization filters used for human subject testing with the new 4Mic case derived from the measurements of the subject seat locations. As expected, the crossover filter limits are consistent while the bandpass variations show some marked differences in response. Because the typical psychoacoustic test positions the subjects in these four seat locations, the new 4Mic equalization filters are expected to improve the sound reproduction over the 2Mic equalization filters (new or old).

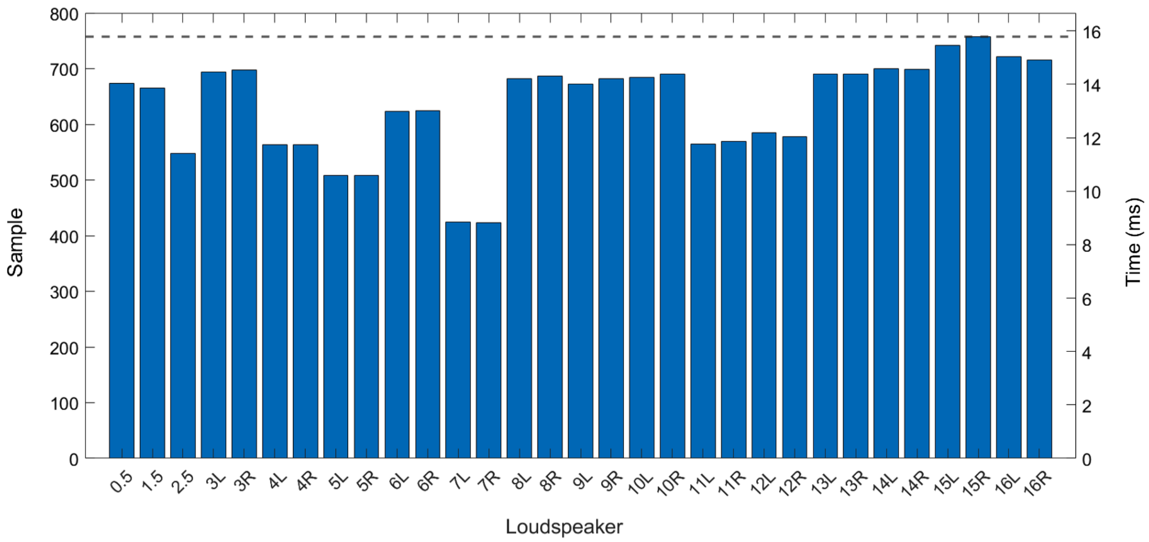

Figure 7 – Full-path time delay compensation for each satellite and subwoofer using a measurement made in the geometric center of the room as the reference.

4.2 Delay and Gain

The same suite of recordings can be processed in order to determine the relative delays and overall gain values for the speakers. In the past, a microphone at the geometric center of the room was used to derive the delays, and this location was included in the current recording suite. When employed as the timing reference point, the maximum delay is approximately 16 ms or 750 samples (see Figure 7). In comparison to the original measurements, the most considerable difference is within 0.5 ms, showing the updated framework is highly consistent with the original results. It is possible to use another microphone location to derive these delays, for example to accommodate a single subject sitting close to a subwoofer, but for typical EER usage going forward, the center mic will be used to derive the timing information.

Finally, the same recordings can also be used to set the relative gain offset between speakers. These results also showed consistency with previously derived gains. The process by which the gains are derived is tied to the microphones selected for the filter derivation, and not that chosen for the delay. That is, it is possible to use the center mic to derive the timing, and the 4 subject seat locations to derive the filter and gains, but unwise to split the filter and gain computations between different mics.

ACKNOWLEDGEMENTS

This project was initiated during Haik Martirosyan's stay at NASA Langley as a NIFS intern and was largely completed during Ellington Scott's NIFS internship. Funding for this work was provided by the Revolutionary Vertical Lift Technology project – thanks to Susan Gorton, Benny Lunsford, Noah Schiller, and others. Thanks to Jaye Moen for building the jig that kept the placement of the microphones consistent between seats in the EER.

REFERENCES

[1] Rizzi, S. A., "Toward Reduced Aircraft Community Noise Impact via a Perception-Influenced Design Approach," InterNosie 2016, Hamburg, Germany, 2016. p 2-26.
[2] Rathsam, J., Klos, J. and A. Loubeau, "Effects of chair vibration on indoor annoyance ratings of sonic booms," The Journal of the Acoustical Society of America, Vol. 143, No. 1, 2018.
[3] Shepherd, K. P., "Detection of Low Frequency Impulsive Noise from Large Wind Turbine Generators," NASA CR-172511, 1985.
[4] Hubbard, H. H. and C. A. Powell, "Acoustic Facilities for Human Factors Research at NASA Langley Research Center," NASA TM-84585, 1981.
[5] Beranek, L., "Riding the Waves: A Life in Sound, Science, and Industry," MIT Press, Cambridge, MA, 2008.
[6] Faller K.J., Rizzi, S.A. and A. A. Aumann, "Acoustic Performance of a Real-Time Three-Dimensional Sound-Reproduction System," NASA TM-2013-218004, 2013.
[7] "AuSIM3D Vectsonic," <http://www.ausim3d.com/products/Vectsonic/index.html> , accessed 8/2022.
[8] Boucher, M. A., Christian, A. W., Krishnamurthy, S. and S. A. Rizzi, "A perceptual evaluation of the efficacy of Sound Exposure Level in the rating of annoyance to helicopter noise," Presented at the VFS International 76th Annual Forum & Technology Display, Virginia Beach, VA, 2020.
[9] Rafaelof, M., Christian, A. W., Shepherd, K. P., Rizzi, S. A. and Stephenson, J. H., "Audibility of Multiple, Low-Frequency Tonal Signals in Noise," NASA TM-2019-220398, 2019.
[10] Begault, D. R., "3D Sound: For Virtual Reality and Multimedia," AP Professional, Boston, MA, 1994.
[11] Haykin, S. "Adaptive Filter Theory (5th Ed.)," Pearson, Upper Saddle River, NJ, 2014.
[12] Roberts, R. A. and C. T. Mullis, "Digital Signal Processing," Addison-Wesley Longman Publishing Co., Inc., 1987.
[13] Dennis Jr., J. E. and R. B. Schnabel, "Numerical methods for unconstrained optimization and nonlinear equations," Society for Industrial and Applied Mathematics, 1996.