Capsule Calibration Approaches for Low-Cost Higher Order Ambisonic Microphone Arrays

Charles J Middlicott[1,2] & Bruce J Wiggins[2]

[1] *Sky Labs Brentwood, Essex UK* [2] *University of Derby, Derby UK*

Correspondence should be addressed to Author (charles@charlesmiddlicott.co.uk)

ABSTRACT

Low-cost electret microphone capsules, such as those accessible to the audio hobbyist or DIY audio electronics community are typically available 'unmatched'. Ideally one would like to ensure these capsules all reside within a very small production / performance tolerance. In a cost-focused, mass production environment it is not feasible to measure and match large batches of capsules hence the acoustic performance will differ from capsule to capsule. This paper summarises work around approaches to capsule calibration in an effort to mitigate/overcome the problems caused by capsule mismatch, with a specific focus on its applications ambisonics and multi-microphone arrays.

1 Introduction

This paper is the second instalment in a multi-part series of publications focusing on the topic of capsule calibration for low-cost higher order microphone arrays, with a specific interest in its applications for ambisonics. The implementation of capsule calibration approaches, and associated array development (simulation & Characterization) will be discussed in brief in the following sections, the author directs the reader to the previous publications [1], [2] for a longer form review of these topics.

The purpose of this paper is to present to a review of the capsule calibration approaches investigated in respect to ambisonic microphone arrays. The key focus lies within resulting performance of these approaches. This is ascertained by robust analysis and evaluation when the approaches are considered at both an individual capsule level in the spherical harmonic domain.

2 Prototype Array Development

The authors PhD study explored the topic capsule calibration for ambisonic arrays, predominantly utilizing electret capsules (ECM). This necessitated prototyping several microphone arrays throughout the course of this wider investigation, due to this a modular approach to hardware design was adopted where Two multi-channel designs were developed, incorporating 4 & 8 channels of capsule conditioning circuitry. This meant that a range of array geometries could be characterized with relative ease (by swapping out capsule holders). The modular designs shown in **Figure 1** provided the TSB140A-T / AO-T capsules with 48V phantom power, resulting in a balanced output [3], [4].

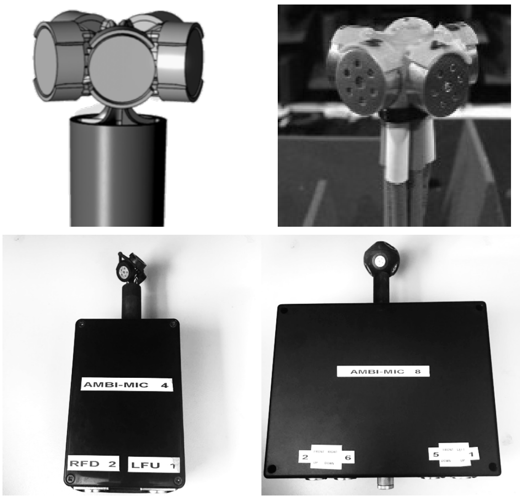

Figure 1 – Array Designs (CAD / Prototypes & Modules Electronics)

Initial investigation utilized a horizontal-only 2nd order prototype (which resides in the **2H1V** category in mixed-order ambisonics case [5]), as such this five channel array is capable of collectively encoding 2nd order horizontal components and 1st order vertical components. The rationale behind this approach was to limit the channel count (and required circuitry) in addition to minimizing measurement complexity.

2.1 Array Simulation

A set of microphone array simulation routines have been developed as a tool to aid in the assessment of theoretical array designs, in addition to physical prototypes, as per [1] and [2]. These tools have been packaged as a freely available set of Windows / Mac applications [6].

The simulation is based around the performance of a theoretically ideal higher-order microphone array under anechoic conditions. Here the theoretical ideal refers to requirement that the elements within an array exhibit identical performance characteristics such as directivity, frequency and phase response.

This routine is used to inform the design process, where parameters such as array radius, number of capsules and differing sampling distribution on the sphere can be investigated before the development of the physical prototype begins. Such simulations are then to be used as a framework upon which the performance of a physical prototype can measured be evaluated against. Comparing these simulated and measured responses not only provides a baseline for the measured array to be compared against, but also exposes any errors and physical limitations inherent in the array design. This further aids the assessment of capsule calibration performance and post-processing where the simulation used as a benchmark.

2.2 Array Characterization

Array characterization was undertaken in a Hemi-Anechoic Chamber, utilizing Farinas Exponential Sine Sweep (ESS) method [7] & [8]. A large matrix of impulse responses (IRs) has been measured to characterize the response of the prototype array. In the interests of time and simplicity, the proposed measurement system utilized a remotely controlled turntable to rotate the DUT [9]. This approach was taken instead of using multiple loudspeakers to both streamline & automate the characterization process. This resulted in 181 IRs per microphone (in 2° increments across the azimuth plane). This resolution mirrors the virtual source positions used in the array simulation.

To further limit the overall measurement duration required the authors opted to use a 15-second logarithmic sine sweep from 20Hz to 20kHz, this included a 3 second pad pre and post resulting in a 21-second sweep. The rationale behind using a 15-second sweep was due to a trade-off between measurement resolution at low frequency and overall measurement duration. The padding has been implemented to mitigate any noise imparted between each turntable rotation. In total 905 sweeps were measured, resulting in a ***M x D*** matrix of IRs, with a total measurement duration of ~5 hours. A graphical representation of the measurement system can be observed below.

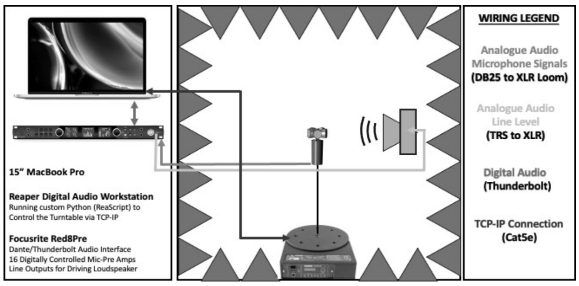

Figure 2 – Diagram of Array Characterization Approach

3 Implementation

Two core methods to capsule calibration have been considered, these are ‘Gain-Matching’, implemented in the time domain and approaches which reside under the umbrella of filtering in the frequency domain. The latter covers a range of approaches utilising Nelson Kirkeby inversion to generate calibration filters. Collectively these two approaches span optimisation over both the axial position and all/or a subset of the measurement positions, this helps consider variations in capsule response over a wider azimuthal angle of incidence.

3.1 Gain Matching

Gain matching is the process of applying gain or attenuation to one microphone signal to level match it to another. This gain/attenuation term is denoted as the correction factor. As gain matching is the application of a scalar correction factor to an input signal, it will merely increase or decrease the overall magnitude of the input signal, such that this has no effect on the frequency response of the microphones within an array (i.e. no filtering in the frequency domain is utilized in this implementation).

Calculating a correction factor to level match an array of microphones can be achieved in several ways, utilizing the measured microphone filters in both the time domain and frequency domain where averaging is utilized (resulting in averaged-gain-matching). In all cases, a single capsule is used as a reference for which other capsules are level matched to. The 'reference' capsule used in these approaches was ascertained programmatically by analyzing the deviation in axial frequency response (opting to use the capsule with the flattest frequency response).

Four contrasting methods for gain-matching have been considered to optimize this general approach, these four methods are….

1). ***Gain Matching in Time Domain***
2). ***Gain Matching at a Single Frequency***
3). ***Averaging from DC to Nyquist***
4). ***1/3rd Octave Averaging***

3.1.1 Gain Matching - Time Domain

When gain matching in the time domain, the absolute magnitude of the ***m***th microphones measured axial filter is compared to the reference capsule. The purpose is to generate/apply a scalar correction factor to increase/attenuate the output level of the measured filters.

Applying the correction factor to the corresponding ***m***th microphone filter is achieved by scalar multiplication, resulting in a matrix of gain matched microphone filters. The response of this approach can be observed pre/post correction in **Figure 3**.

Initial analysis considered the minima & maxima of the magnitude in the frequency domain (pre and post calibration), where a decrease in the difference is expected. This was calculated across octave bands from 31Hz to 16kHz.

3.1.2 Gain Matching – 1kHz

Gain Matching at a single frequency, aims to decrease the deviation between the capsule responses to a greater degree than afforded by the time-domain based approached. To implement this a similar approach is used, yet the Fast-Fourier Transform is utilised to transform each of the ***m***th microphones axial filters into the frequency domain.

The magnitude of each filter at 1kHz is determined and this is used in conjunction with the reference capsule to generate a correction factor. In this case the magnitude of the responses is considered in dB in the frequency domain, logarithmic to linear conversion needs be applied to transform the correction factor in dB into a scalar that can be used as part of a multiplication operation to generate the corrected responses, which can be observed below in **Figure 4.**

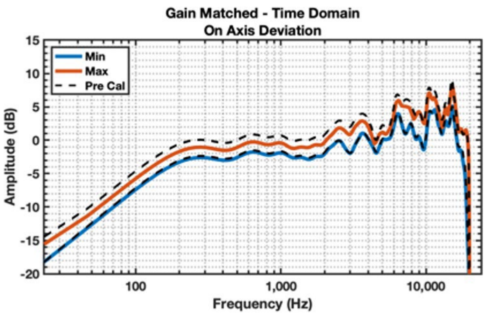

Figure 3 – Gain Matching in the Time Domain

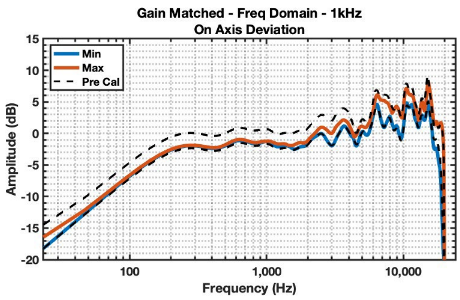

Figure 4 – Gain Matching at 1kHz

3.1.3 Gain Matching – DC to Nyquist

Gain-Matching from DC to Nyquist works in a similar manner to matching at 1kHz, however in this case the mean magnitude is taken across DC/Nyquist rather than relying on a single frequency. The purpose here is to ensure that the resulting performance of the calibration isn't being needlessly restricted. Averaging across DC/Nyquist in the axial case resulted in responses that offered a marginal improvement to the raw responses but a decrease in performance is observed compared to gain-matching in the time domain.

3.1.4 Gain Matching – 1/3rd Oct. Averaging

The author opted to modify the previous approach to average in frequency domain across 1/3rd octaves. This began at the extremities *(50Hz to 20kHz)* and converged about 1kHz *(800Hz to 1.25kHz).* The purpose here was to generate correction factors calculated by averaging between all 13 sets of octave limits, such that an optimal set of averaging limits can be defined. Across the axial measurement position averaging across 160Hz to 6.3Hz proved to offer the best performance increase. However, when one considered the response over all angular measurement positions, 50Hz to 20kHz proved to offer the best performance, this can be observed below in **Figure 5**. The effect over azimuth will be discussed in a later section.

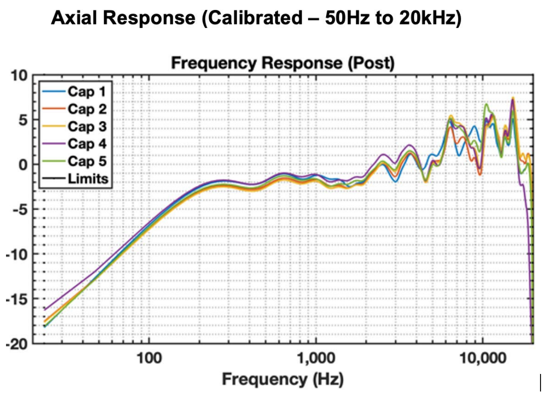

Figure 5 – Gain Matching in the Time Domain

3.2 Filtering in the Frequency Domain

Capsule calibration approaches which utilise filtering in the frequency domain differ to the frequency-domain based gain-matching approaches considered previously. Here, rather than using the frequency-domain data to generate scalar correction factors, (by averaging over specific frequency indices), the frequency domain representation of the filters is used to generate correction filters. This results in a frequency response that has been manipulated such that the amount of correction differs across each frequency bin. These correction filters are derived for each of the $\boldsymbol{m}^{th}$ microphones by utilising Nelson-Kirkeby inversion per the equation below.

$$\boldsymbol{H_{inv}(\omega)} = \frac{\boldsymbol{H_{target}(\omega)} \cdot \boldsymbol{Conj(H(\omega))}}{\boldsymbol{Conj(H(\omega))} \cdot \boldsymbol{H(\omega)} + \boldsymbol{e}} \qquad \textbf{(1)}$$

When the resulting inverse filter is convolved with the matrix of microphone filters the result is a calibrated microphone where the frequency response has been corrected. Careful consideration needs to be taken with respect to the regularization parameters utilised when generating these inverse filters. This will be discussed in more details in the following sections. A graphical representation of the differing stages of each approach can be observed below.

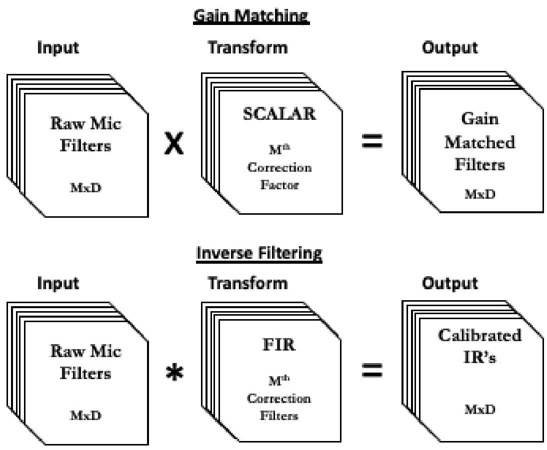

Figure 6 – Gain Matching vs Filtering in the Frequency Domain

The distinct difference here is that when the transform from raw input to calibrated output occurs FIR Correction filters are utilised instead of Scalar correction factors.

3.2.1 Approaches to Regularization

Regarding the regularization parameter ***(e)*** from equation 1, To add additional flexibility to the approach a frequency dependent implementation has been utilised per [1]. Here a varying ***(e)*** parameter is used across the frequency range of interest. The author has opted to use three regularization values between a set of programmatically dervied frequency ranges (Low Limit ,Centre & High Limit).

This results in a gradual decrease in the regularization parameter towards 0 between the centre limits, when the high limit is reached the regularization parameter then increases above 0 such that excess amplification is not imparted across the extremities of the frequency range. To further the robustness of this approach the author opted to use a fixed 10dB amplification limit to programmatically define the required regularization values across frequency.

3.2.2 Calibration to a Measured Capsules Axial Response

The authors have shifted from investigating the use of scalar correction factors to generating capsule calibration filters. To ensure we are not adversely coloring the array response the authors opted to calibrate the array of capsules to one of the measured capsules axial responses. This required determining the flattest of the measured capsules responses in the frequency domain and choosing that capsule as the target/reference. It was expected that this approach would result in a minimal amount of inverse filtering required to 'correct' for the axial position. To implement this approach $\boldsymbol{H_{target}(\omega)}$ in equation 1 becomes the 'reference' capsules axial filter.

3.2.3 Calibration to a Flat Axial Response

In contrast when using Nelson-Kirkeby inversion and calibrating to a flat response the target response $\boldsymbol{H_{target}(\omega)}$ resembles a Kronecker Delta function which exhibits a flat frequency response. The target function is at a minimum, twice the length of the filter to be inverted in all cases to ensure there is ample time to calculate a stable inverse filter.

3.2.4 Calibration by Diffuse Field Equalization

The distinct benefit afforded by diffuse-field equalization is that it is capable of considering the response of a microphone across varying angles of incidence when the user is looking to generate a correction filter. This is instead of optimizing solely across the axial positions as per the previous approaches. This requires many Characterization measurements which are then averaged to generate an average diffuse-field response. The inverse filter of this average is then calculated by means of Nelson-Kirkeby inversion.

4 Reviewing Raw Measured Responses

When compared to the reference capsule response in section **3.1** the absolute error of the remaining capsules can be calculated. This highlights the disparity in capsule response over azimuth. Most approaches considered rely on the axial measurement filters to generate a calibration factor or filter, this inherently optimizes across the axial position over/above the off-axis response. The use of absolute error gives insight into the effects of these methods as any degradation across azimuth will result in errors encoded in the SHD domain.

The absolute error shown in **Figure 7** reveals a significant difference in the angular position of the nulls in the directivity responses from capsule to capsule, with capsule #2 showing the largest disparity here, with capsule #4 generally exhibiting a greater overall absolute difference.

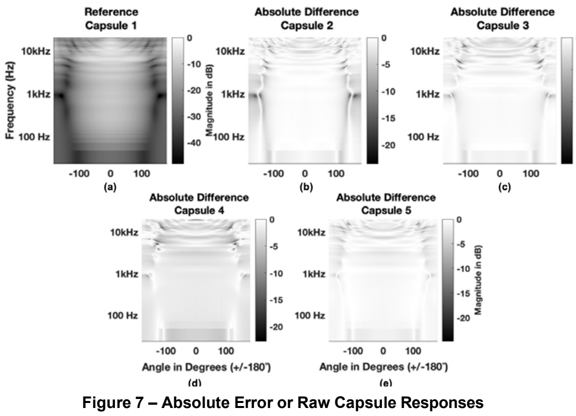

Figure 7 – Absolute Error or Raw Capsule Responses

When the raw capsule signals are encoded into the spherical harmonic domain the uncalibrated capsules cause encoding errors and skew the resulting directivity responses as shown below in **Figure 8.**

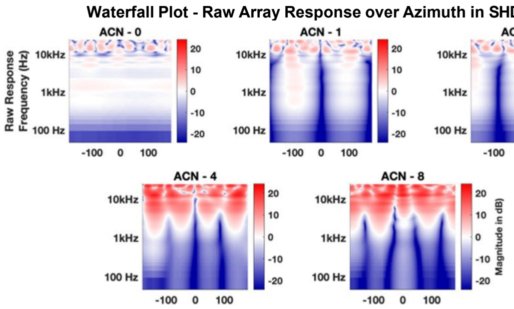

Figure 8 – Raw Array Response in SHD

4.1 Review of Performance Metrics

To adequately assess the performance of these calibration approaches, the authors employed a methodical framework to array analysis which considered the responses at various stages during the signal chain. RMS Error was used (over both frequency and angle) to characterize any improvement in the error between capsules as a direct result of calibration. This was considered at a capsule level and the result was also evaluated in the SHD.

5 Performance at a Capsule Level

Of the approaches utilizing a measured capsule as a reference, (such as gain-matching and calibration to a measured capsule response), analysis has shown that averaging 1/3rd octaves between 50Hz and 20kHz outperforms all the previous gain-matching approaches implemented across the axial position..

However, when frequency-domain based filtering is considered, calibrating to capsule #1's axial response yields a comparable result to averaging over 50Hz to 20kHz, with the distinct benefit that at the axial position there is negligible to 0 error (as this is position the inverse filters generated during the calibration process are optimized for). Interestingly, calibrating to capsule #1 at a capsule level when considered across azimuth has shown to yield a marginally poorer performance when compared to the raw response. Whereas the remaining frequency-domain based filtering approaches have indicated that the use of more drastic inverse filtering is beneficial to correcting the capsules responses. Further to this, as expected calibration to a flat axial response has shown to exhibit closest to the ideal response, this is to be taken with caution as the references in these cases are somewhat approach specific.

Using Diffuse-field equalization results in the largest decrease in error (performance gain) which is highlighted when considering the RMSE over frequency, but the angular RMSE result of **3.595dB** reveals the slight increase in error. For the axial position this is due to each calibrated capsule exhibiting a slightly difference axial response over the audible range.

6 Performance in the SHD

To consider the array designs robustness and the effect of using unmatched capsules, array responses were generated in the SHD. This was achieved by using the raw measured capsule and an array of synthesized capsules (using capsule 1) to highlight the performance of an ideal array using identical capsules. When comparing the responses from **Figure 8** with **Figure 9** one can observe greater uniformity across azimuth when considering the synthesized array of capsules, which alludes to the importance of capsule calibration when signals are be encoded in the SHD (for applications such as ambisonics. The aim of these calibration approaches is to minimize the error (relative/absolute difference) between the raw signals and the theoretical array responses generated using the array simulation.

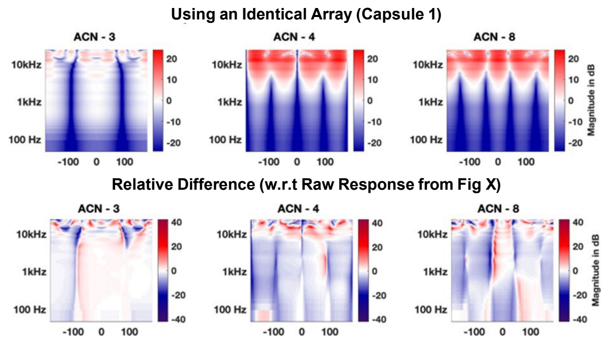

Figure 9 – Raw vs Identical Capsules (2nd Order SH)

In **Figure 10** below the benefit of capsules exhibiting uniformity can be easily observed. It is evident here that in cases where there is no capsule matching within an array or any effort taken to calibrate the capsules, encoding errors are introduced in particular where nulls are expected in the directivity response.

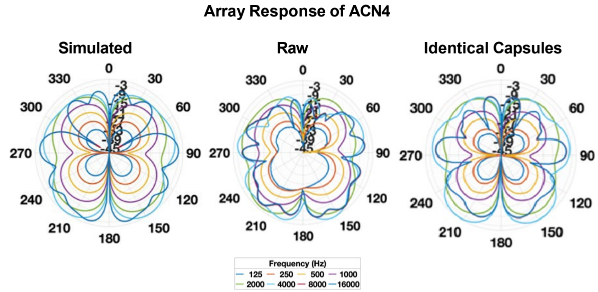

Figure 10 – Array Response (Sim, Raw and Indentical Capsules)

6.1 Gain-Matching Approaches

When considering the effect of approaches to gain-matching in the SHD the time-domain based approached has shown to out-perform all the alternative gain-matching approaches. In comparison averaging from 160Hz to 6.3kHz causes skewing of directivity response in the SHD.

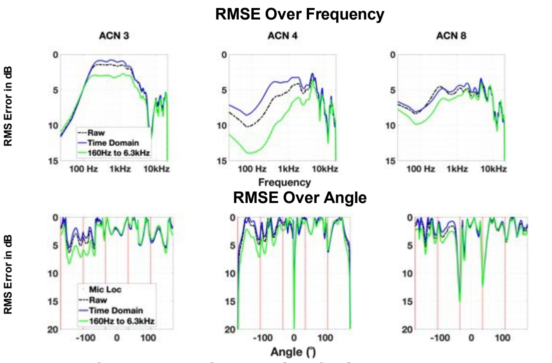

Figure 11 – Comparing RMSE in the SHD for Gain-Matching Approaches

6.2 Diffuse-Field Equalization

Utilizing Diffuse-field equalization has shown to offer a marked improvement above all the prior calibration approaches considered. This is likely due to taking into account each capsules response over multiple angles of incidence to generate calibration filters. It can be observed that employing diffuse field equalization results in a greater accuracy in encoding in the SHD over the nulls in the directivity response. This is evident in **Figures 12 & 13** respectively.

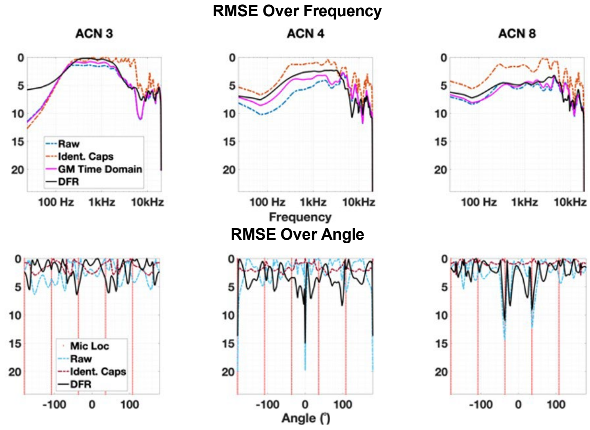

Figure 12 – RMSE in the SHD using Diffuse Field Equalization

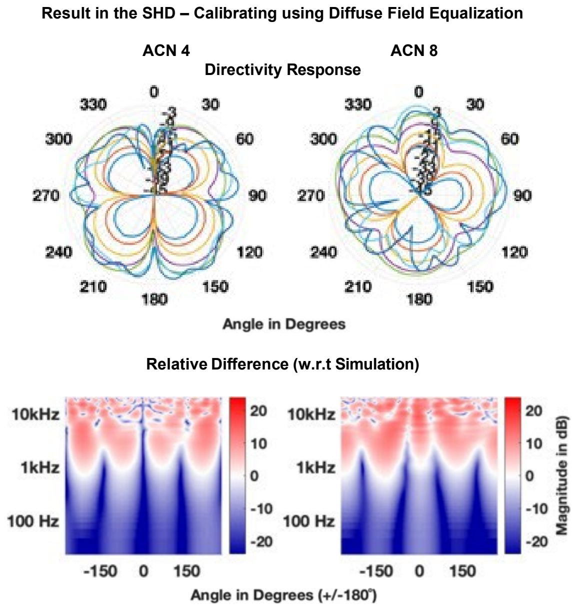

Figure 13 – Response in SHD (Using Diffuse Field Equalization)

7 Conclusions & Future Work

Calibration using gain-matching, (when evaluated in the SHD), has shown a consistent decrease in performance. The contributing factor being a skewing/smearing of the directivity response. This occurs in all scenarios where the magnitude in the frequency domain is used to generate correction factors, at a fixed frequency / or averaged in some manner. However, the simplest gain-matching implementation, in the time domain, has shown a slight improvement upon the raw responses.

If one then considers which approaches have shown the best performance in the SHD, Calibration to Capsule #1's Axial response has shown an improvement above that of Gain-Matching in the Time Domain. This proves that inverse filtering is of great use – (if applied carefully - without employing excessive amplification or attenuation).

Further improvement is yielded when using Diffuse-Field Equalization due to the nature of averaging a large set of angular measured, this correction considers more than just the axial response of the capsules within an array, this results in an inverse filter that although will not correct to a perfectly flat response will improve the average correction over a larger angular range.

Considering the findings from this study that author suggests, if feasible, to calibrate capsules for use within in ambisonics by Diffuse Field Equalization. This work has not only shown that to yield the best improvement in response, but it has also shown that one does not need characterize across the azimuth plane at 2˚. Instead capturing a smaller subset of data corresponding to a 20˚ azimuth resolution will offer a comparable response in the SHD to that of having performed array characterization in 2˚ increments over azimuth

The author has proposed future work which focuses on how the investigated approaches to capsule calibration scale to periphonic array designs such as those shown at the bottom portion of **Figure 1.** In addition to this avenue of research (to further validate these approaches to capsule calibration), the author plans to integrate the approach used for array characterization into a shareable application which combines the measurement approach with turntable control and appropriate post processing. The aim here is to automatically generate filters to implement the calibration approaches considered in this paper and the authors wider PhD thesis [10].

References

[1] C. Middlicott and B. Wiggins, "Calibration Approaches for Higher Order Ambisonic Microphones," Oct. 2019, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=20673.

[2] C. Middlicott and B. Wiggins, "Development of Ambisonic Microphone Design Tools—Part 1," Oct. 2018, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=19753.

[3] JLI-Electronics, "JLI-140A-T Capsule Datasheet," *JLI-140A-T Capsule Datasheet*. JLI Electronics, p. 1, 2019, Accessed: Dec. 12, 2019. [Online]. Available: https://www.jlielectronics.com/content/JLI-140A-T.pdf.

[4] JLI-Electronics, "JLI-140AO-T.pdf." p. 1, 2019.

[5] C. Travis, "A New Mixed Order Scheme for Ambisonic Signals," in *1st Internation Symposium on Ambisonics*, 2009, pp. 1–6.

[6] C. J. Middlicott, "MATLAB Tools for Ambisonic Microphone Array Design," 2018. https://www.charlesmiddlicott.co.uk/GUI (accessed Sep. 12, 2022).

[7] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," *Audio Eng. Soc.*, no. I, pp. 1–15, 2000.

[8] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *AES 122nd Convention*, 2007, p. Paper 7121, 1-21, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14106.

[9] Outline, "Outline ET250-3D Turntable," Brescia, Italy, 2017. [Online]. Available: https://outline.it/download/Documents/Manuals/Strumenti di misura/ET250-3D_Manual.pdf.

[10] C. J. Middlicott, "An Investigation Into The Calibration Of Low-Cost Microphones For Higher-Order Microphone Arrays," University of Derby, 2022.