Measuring Speech Intelligibility Using a Binaural Decode of a First-Order Ambisonics Microphone

Name: Tom Cotter **Degree:** BSc (Hons) Live Sound Technology **Supervisor:** Matthew Hickling

1 INTRODUCTION

Current standardised methods of measuring speech intelligibility are non-inclusive of the human auditory system's binaural advantages; specifying omnidirectional receivers (British Standards Institution, 2011) (Wijngaarden & Drullman, 2008, p. 4514).

Research has been completed into replacing omnidirectional receivers with Head and Torso Simulators (HATS). However, these are not commercially or practically viable. In solution, a binaural decode of a first-order Ambisonics microphone (DUT) was implemented.

2 AIM

To investigate using a binaural decode of a First-Order Ambisonics (FOA) microphone to measure speech intelligibility.

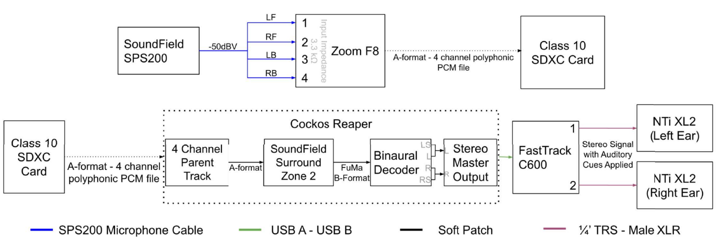
3 OBJECTIVES

1. With a noise source at arbitrary azimuths, gather STIPA results for various decoders and synthesis approaches and compare to HATS
2. Conduct STIPA measurements in accordance with BS 60268-16:2011, setting a noise source at multiple azimuth's
3. Repeat measurements using the HATS and DUT
4. Repeat all measurements in contrasting acoustic environments
5. Measure interaural level differences of the HATS and DUT

4 METHODOLOGY

- STIPA measurements were complete in hemi-anechoic, semi-reverberant and reverberant environments (Wijngaarden & Drullman, 2008) (Tang et. al., 2017)(Andreopoulou et. al., 2015)
- The target source was fixed on-axis to the receiver; with a noise source at NR52 & NR63 masking levels rotated in 15° increments
- Measurements were conducted in accordance with BS 60268-16:2011 to provide standardised data and repeated using the HATS and DUT
- The DUT was calibrated to level match the HATS when no masking was present
- DUT performance was measured through the extent at which STIPA results resembled those obtained with a HATS

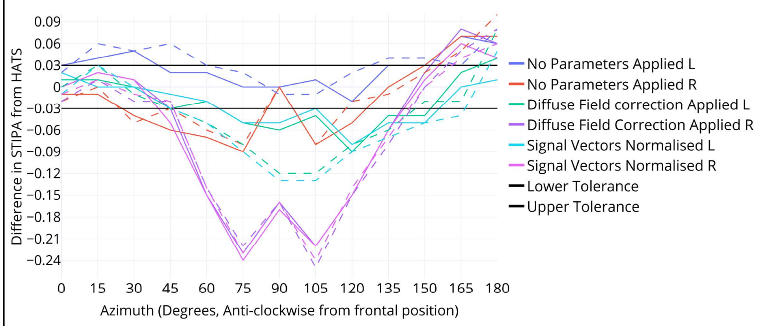
5 SYSTEM DESIGN



6 RESULTS AND DISCUSSION

- Encapsulation of diffuse components in parametric decoders mitigate better-ear listening (McCormack & Delikaris-Manias, 2019, p. 2)
- Magnitude Least-Squares (MLS) decoding and diffuse field correction improves shadowed ear accuracy where partial shadowing occurs
- Diffuse field correction reduces accuracy where opaque shadowing of high frequency occurs and mitigates pinna shadowing
- Normalising signal vectors in the binaural decode decreases accuracy due to altering signal-to-noise ratio's

Hemi-Anechoic Chamber - Parameter Crossover Angle



Dashed Line = Masking NR52. Solid Line = Masking NR63.

- 'W' trend exhibited in the above figure is a result of spreading of the contralateral pressure lobe resulting from diffraction due to FOA low resolution (Politis, et al., 2018)
- Low resolution caused sound field localisation errors, inducing error no matter the synthesis approach (Politis, et al., 2018)
- Accuracy decreases at lower masking levels and longer reverberation times due to localisation error increasing

MLS With Diffuse Correction in All Environments - Masking NR63

7 CONCLUSION

1. Low resolution of FOA incorrectly encodes the sound field, with error increasing for lower masking levels and in diffuse fields
2. For future investigations with a higher-order microphone:
   a) Must use a linear decoder
   b) MLS is most accurate synthesis method
   c) Straight decodes and diffuse field correction improve accuracy for opaque and partial shadowing of high frequencies respectively; with the latter underestimating pinna shadowing

REFERENCES

ANDREOPOULOU, A., BEGAULT, D. R. & KATZ, B. F. G., 2015. Inter-Laboratory Round Robin HRTF Measurement Comparison. s.l.:IEEE Journal of Selected Topics in Signal Processing.

BRITISH STANDARDS INSTITUTION, 2011. Sound System Equipment Part 16: Objective rating of Speech intelligibility by speech transmission index. s.l.:BSI Standards Publication.

McCORMACK, L. & DELIKARIS-MANIAS, S., 2019. Parametric First-Order Ambisonic Decoding for Headphones Utilising the Cross-pattern Coherence Algorithm. 1st EEA Spatial Audio Signal Processing Symposium.

POLITIS, A., TERVO, S. & PULKKI, V., 2018. COMPASS: Coding and Multidirectional Parameterization of Ambisonic Sound Scenes. s.l.:IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

TANG, Y., LIU, Q., WANG, W. & COX, T. J., 2017. A non-intrusive method for estimating binaural speech intelligibility from noise-corrupted signals captured by a pair of microphones. s.l.:Elsevier B.V..

WIJNGAARDEN, J. V. & DRULLMAN, R., 2008. Binaural Intelligibility Prediction based on the speech transmission index. Journal of the Acoustical Society of America, Volume 123, pp. 4514-4523.