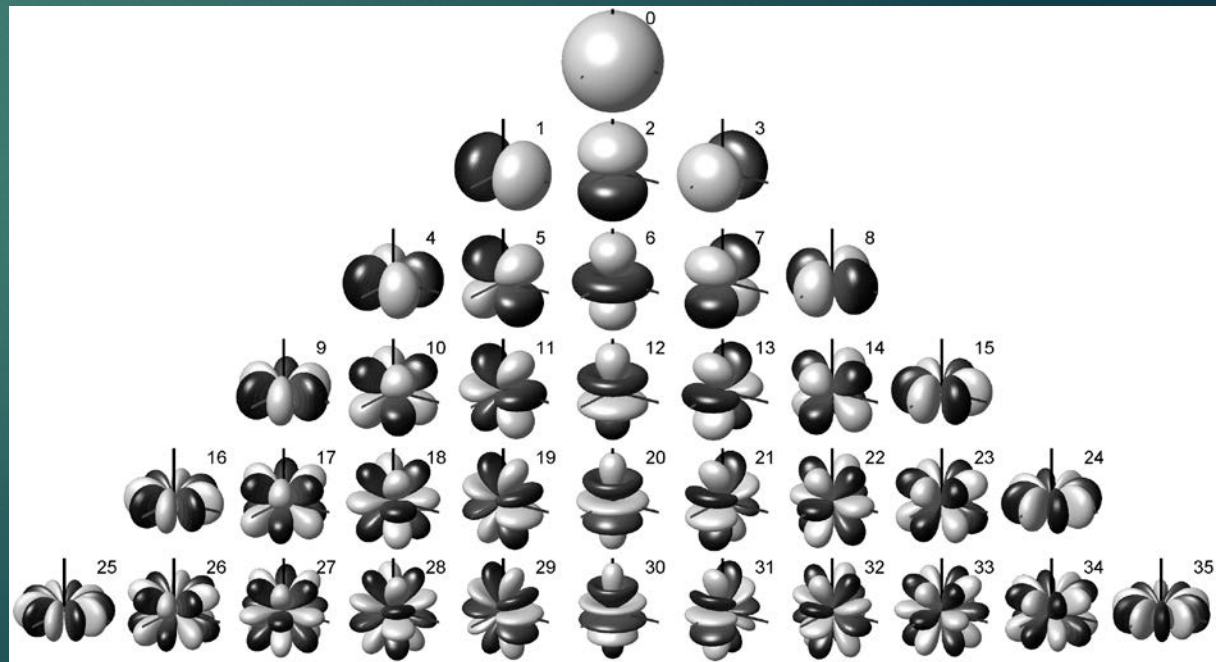
Simulating & Evaluating HOA Planar Arrays

ALEXANDER VILKAITIS, BRUCE WIGGINS, UNIVERSITY OF DERBY

UNIVERSITY of DERBY

What is HOA?

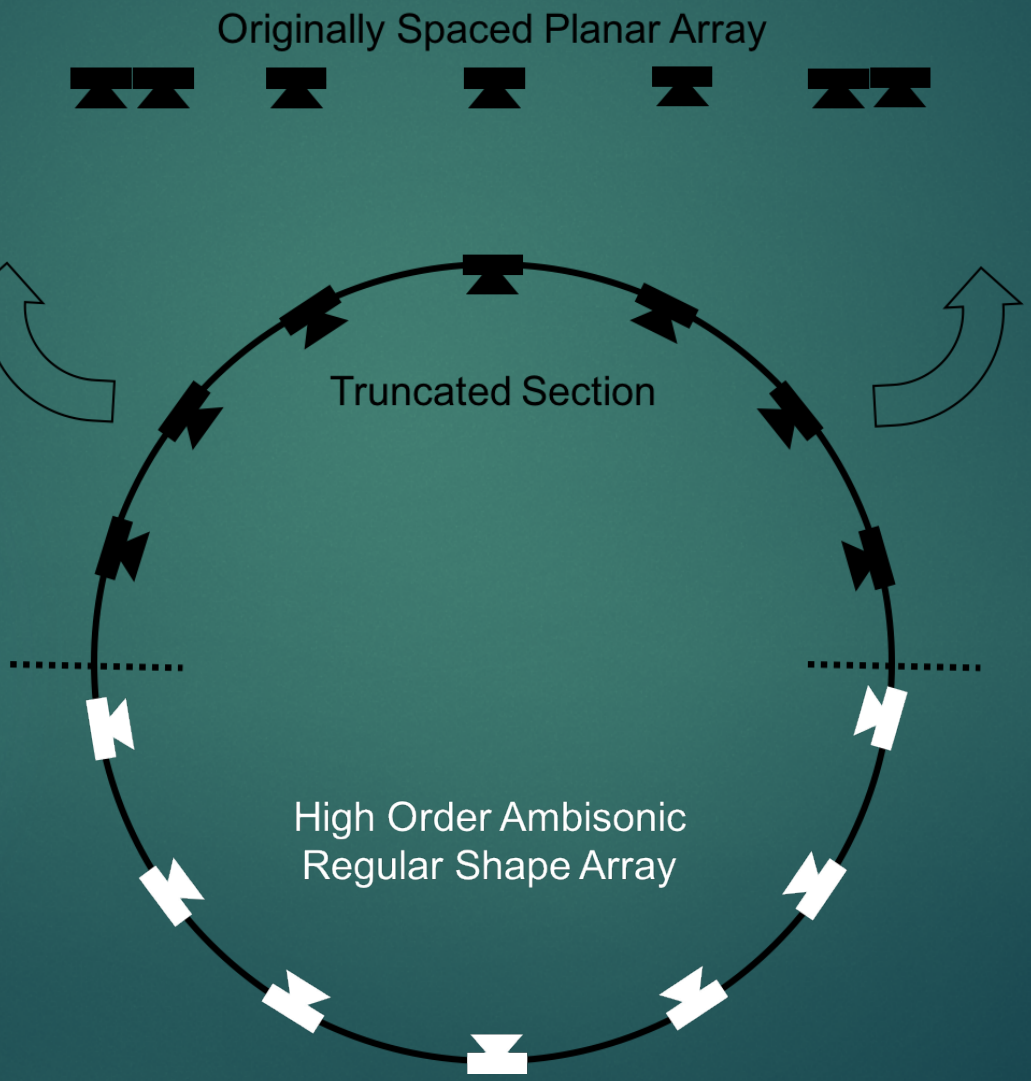
- Higher Order Ambisonics
- Based on spherical decomposition of a sound field
- Driving function for surround speaker arrays
- Array agnostic
- Flexible



UNIVERSITY of DERBY

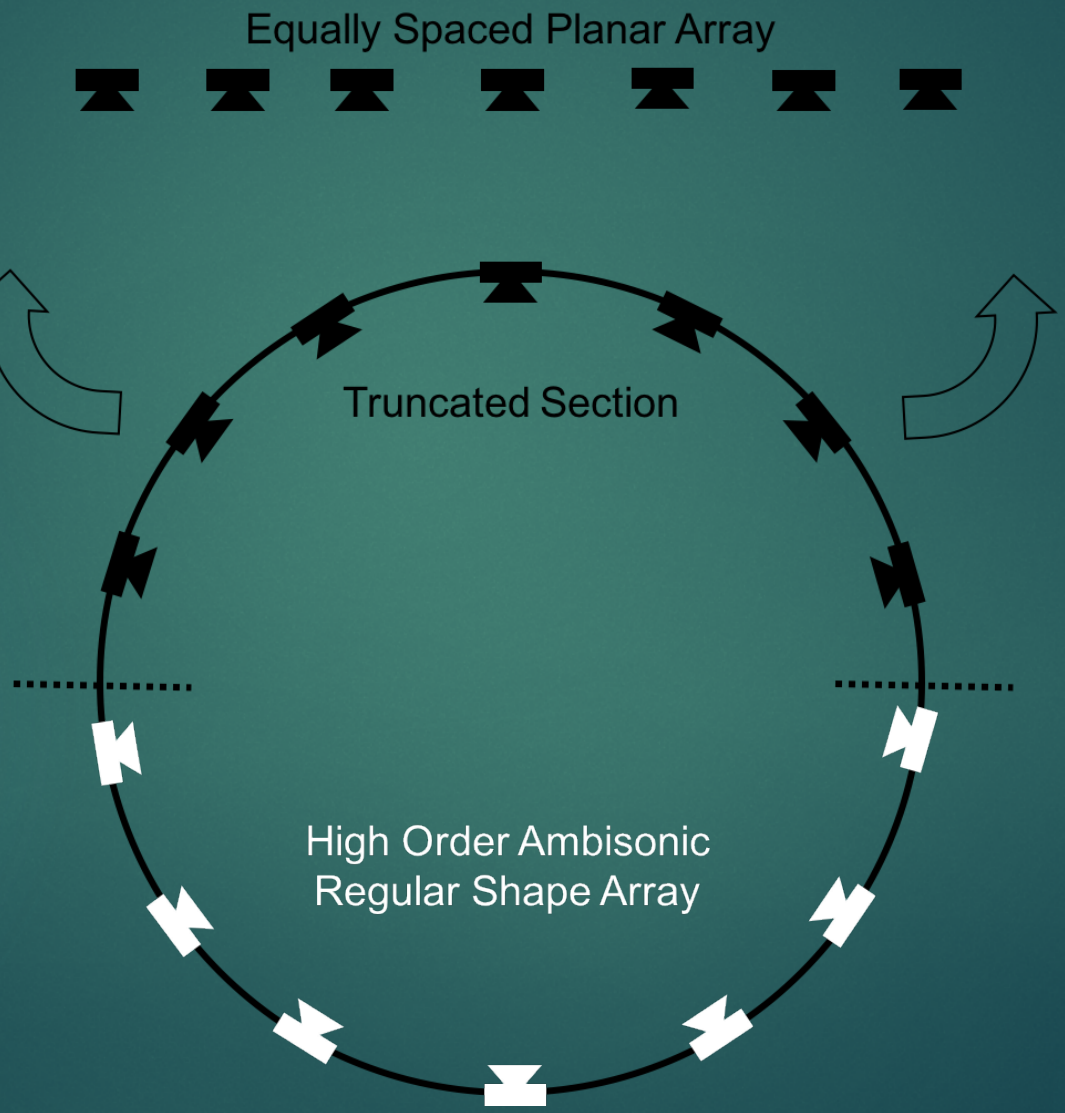
Why Planar?

- From now this array configuration will be referred to as originally spaced as it keeps the x-axis coordinates as if still a circular array



UNIVERSITY of DERBY

- This version will be referred to as equally spaced as they have been spaced equidistant in the 4m array



UNIVERSITY of DERBY

Simulation Aims

- Matlab script simulating 1st, 3rd and 5th order ambisonics
- Evaluate Pressure, Extended Energy Vector (EEV) and localisation error
- HRTF Inter Aural Level Difference (ILD) analysis
  - Shows ILD cues at each synthesised virtual source position compared to a real source in same position
- HRTF Inter Aural Time Difference (ITD) analysis
  - Shows ITD cues at each synthesised virtual source position compared to a real source in same position

Simulation Process

- First calculate number of speakers and their position
- Encode a virtual source at position (*x*,*y*)
- Encode speaker (secondary source positions)
- Use psuedo inverse method to create ambisonic decoder
- Use these gain coefficients in plotting the pressure wave
- Evaluate simulation area with EEV and Localisation Error
- HRTF Analysis

What is the Extended Energy Vector?

- EEE developed by Peter Stitt[1]
- Method of predicting localisation
- EEE improves prediction at off centre listening positions
- Perceptual weight (the extended bit) is assigned to each loudspeaker gain in relation to relative arrival times, levels and direction of the loudspeaker signals

$$\mathbf{E} = \frac{\Sigma_{i=1}^{N} |w_i(\alpha) G_i / r_i|^2 \mathbf{u}_i}{\Sigma_{i=1}^{N} |w_i(\alpha) G_i / r_i|^2} \quad (1)$$

where $w_i(\alpha)$ is the precedence effect weighting, G_i is the loudspeaker gain, r_i is the loudspeaker distance from the listening position, N is the total number of loudspeakers and $\mathbf{u}_i$ is the unit vector in the direction of the i-th loudspeaker.

UNIVERSITY of DERBY

[1] Extended Energy Vector Prediction of Ambisonically Reproduced Image Direction at Off-Centre Listening Positions, Peter Stitt, Stéphanie Beret, and Maarten Van Walstijn, AES Journal 64, No.5, May 2016

Psychoacoustic Optimisation

- MaxRe and In-Phase are psychoacoustic optimisation methods
- Both aim to improve off centre listening and general reproduction of frequencies above 500Hz
- MaxRe optimises the energy vector to improve ILD reproduction at HF whilst maximising the energy in the source direction
- In-phase decoding eliminates all negative gains, meaning no out of phase components are emitted from the speakers

2D max-rE:

$$rE = cos(\pi/2N + 2)$$
$$g_n = cos(n\pi/(2N + 2))$$

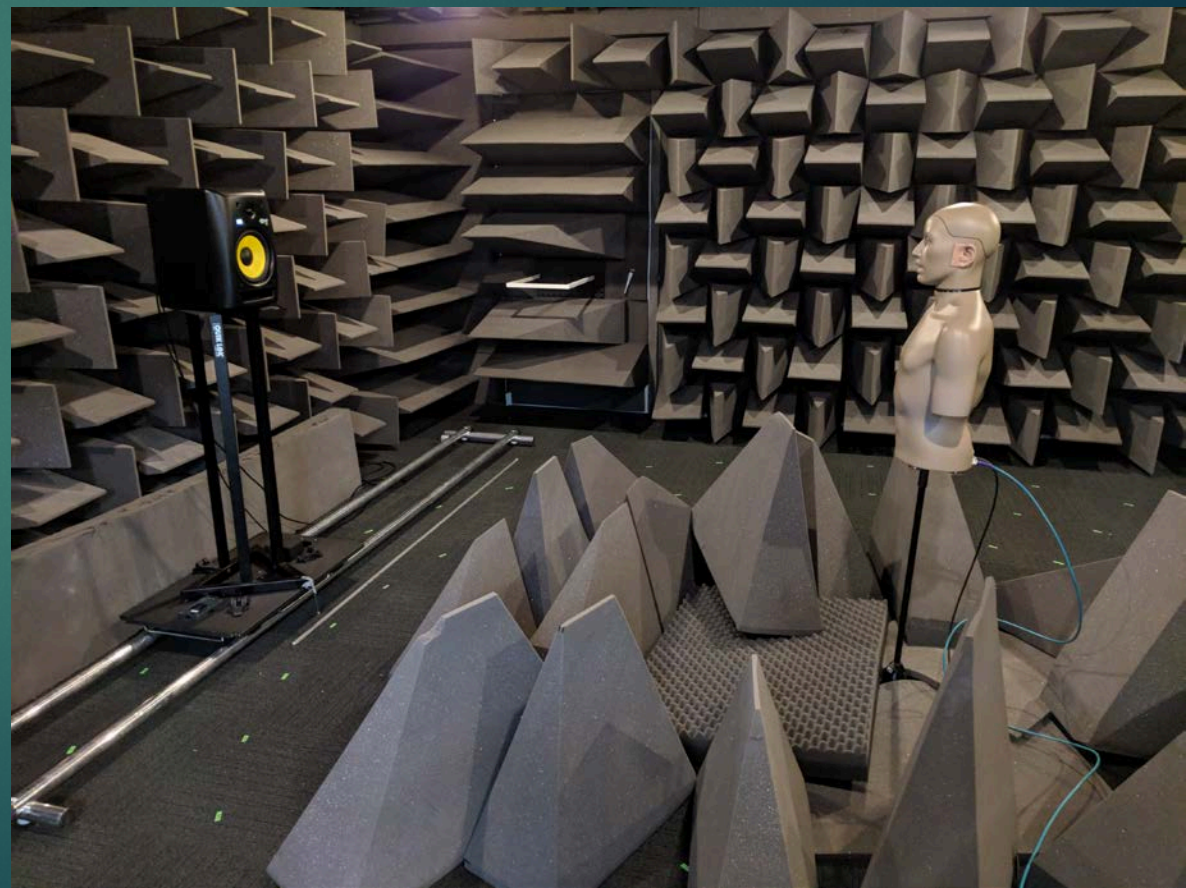
2D in-phase:

$$g_n = \frac{(N!)^2}{(N+n)!\times(N-n)!}$$

UNIVERSITY of DERBY

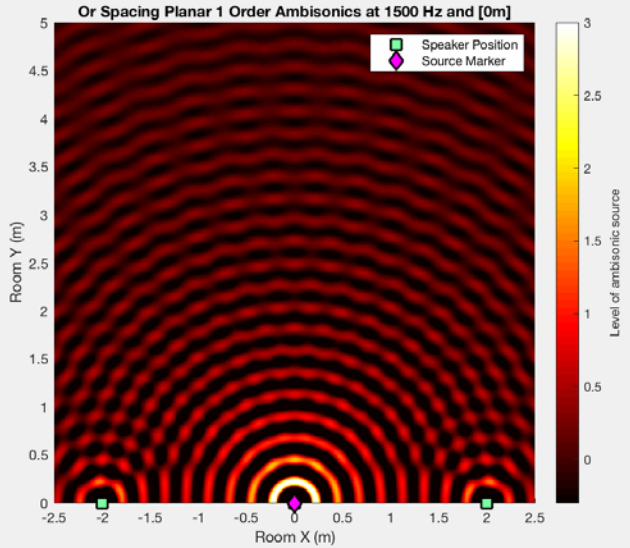
HRTF Capture

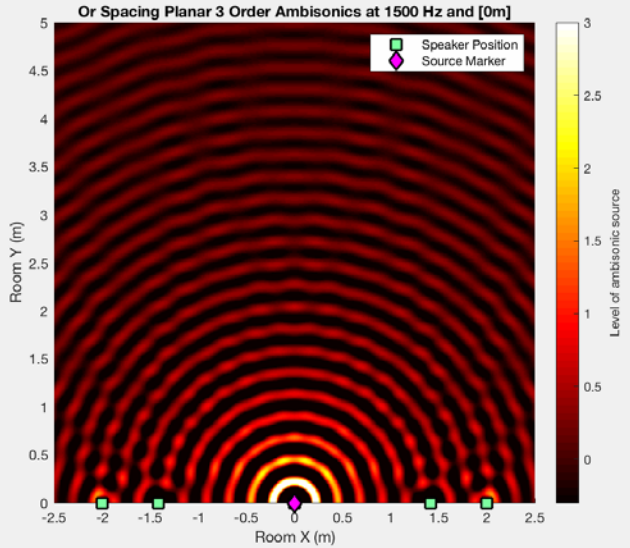
- Assuming symmetry of 4m array – so recording 0-2m
- Camera slider to ensure movement in only one dimension
- KRK ROKIT 8 Source
- Audomatica CLIO FW 10
- KEMAR Head and Torso Simulator
- Record one IR (each ear channel) then move speaker to next position

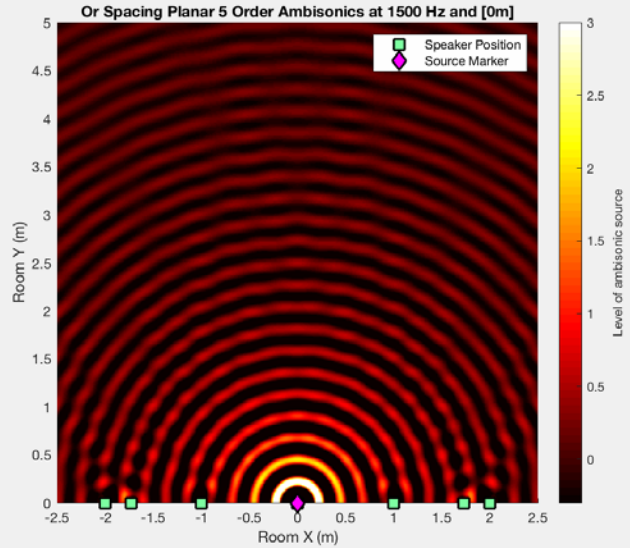


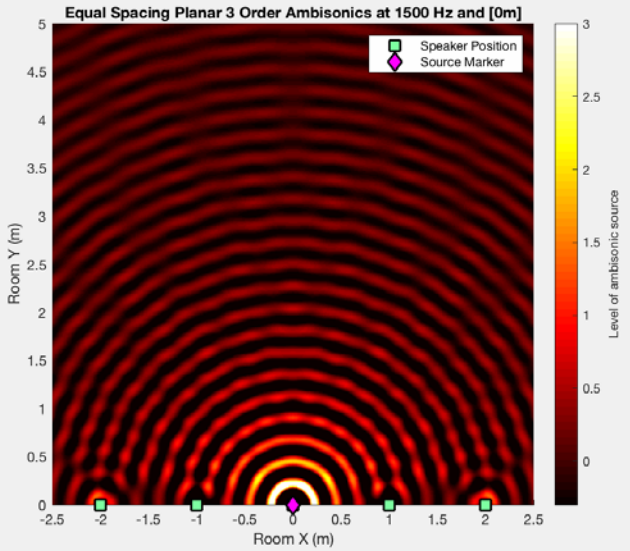
UNIVERSITY of DERBY

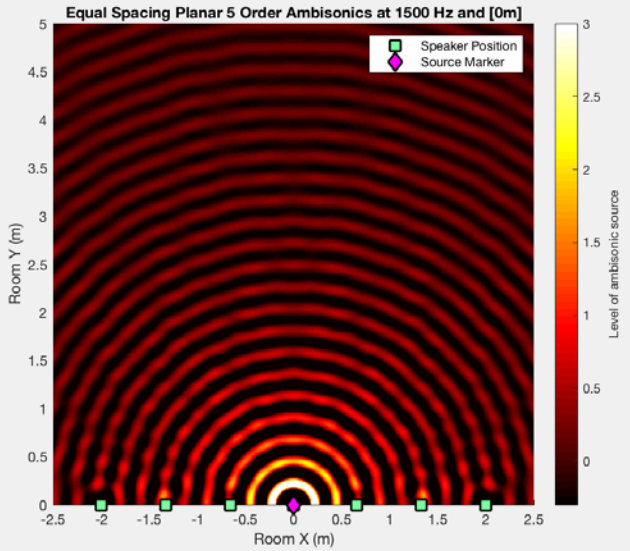
Original Spacing vs Equal Spacing





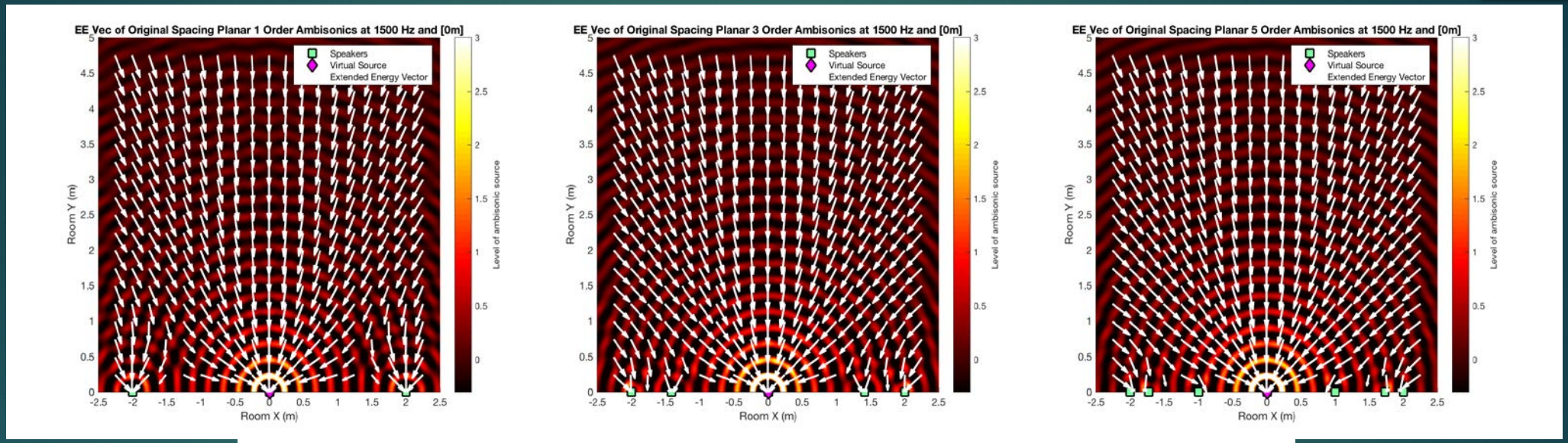


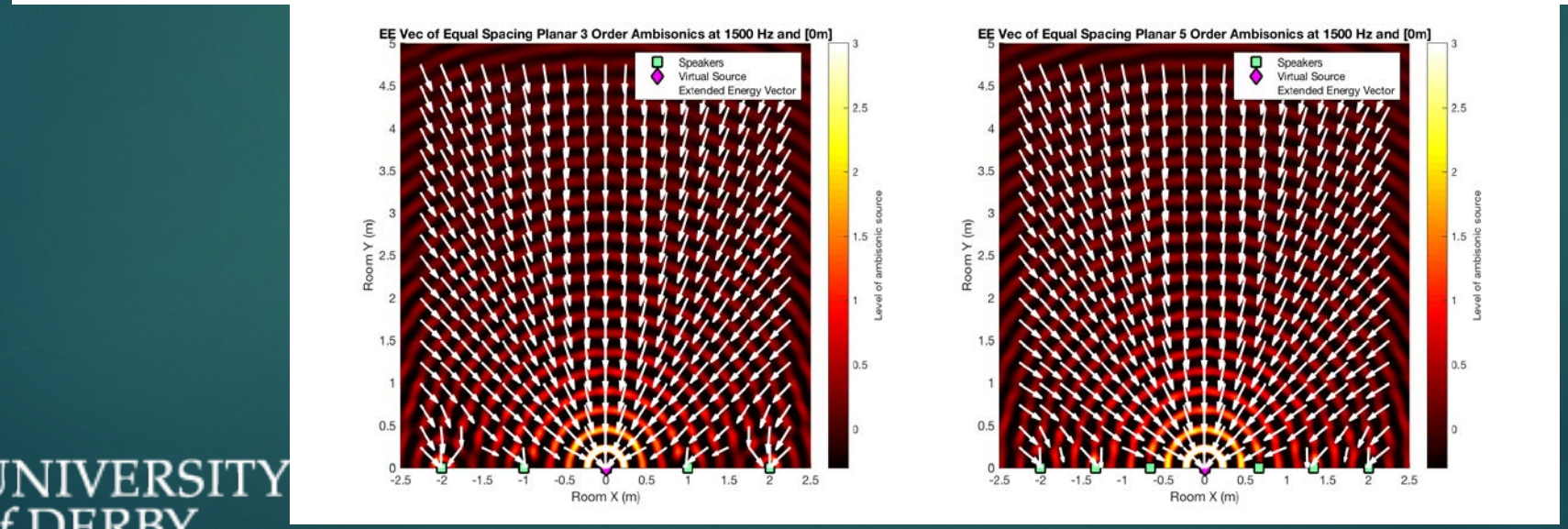




UNIVERSITY of DERBY

Original Spacing vs Equal Spacing Cont.

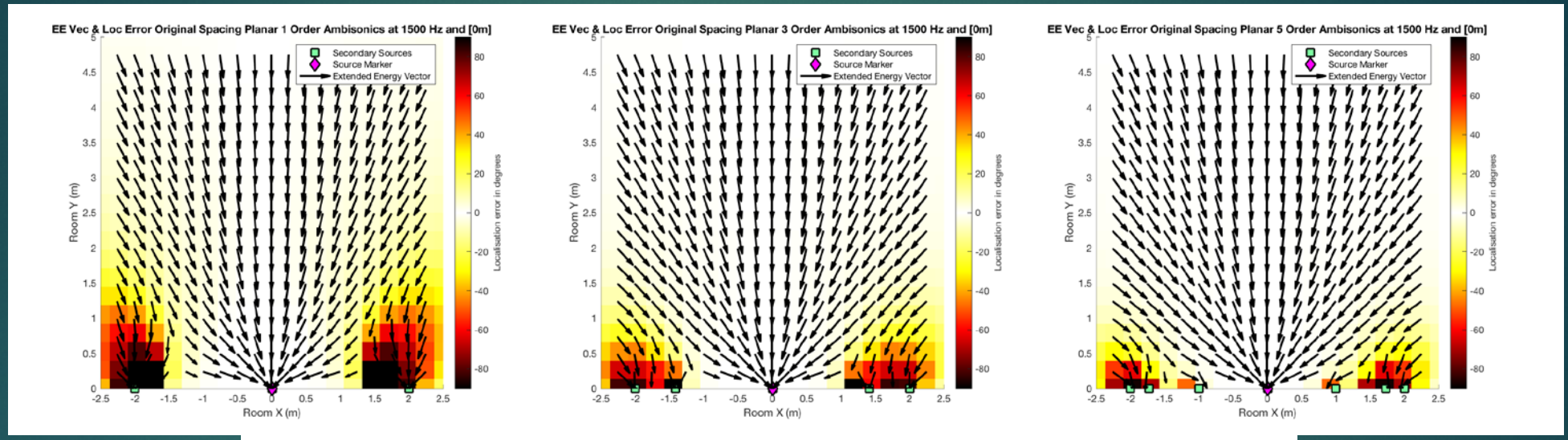


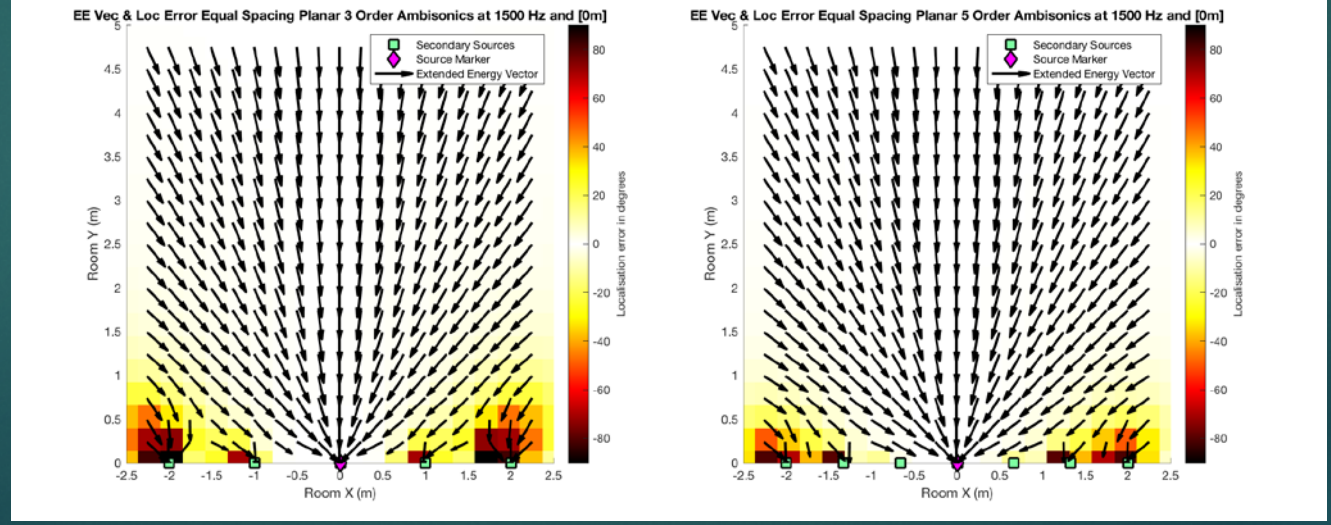


UNIVERSITY of DERBY

Extended Energy Vector & Localisation Error

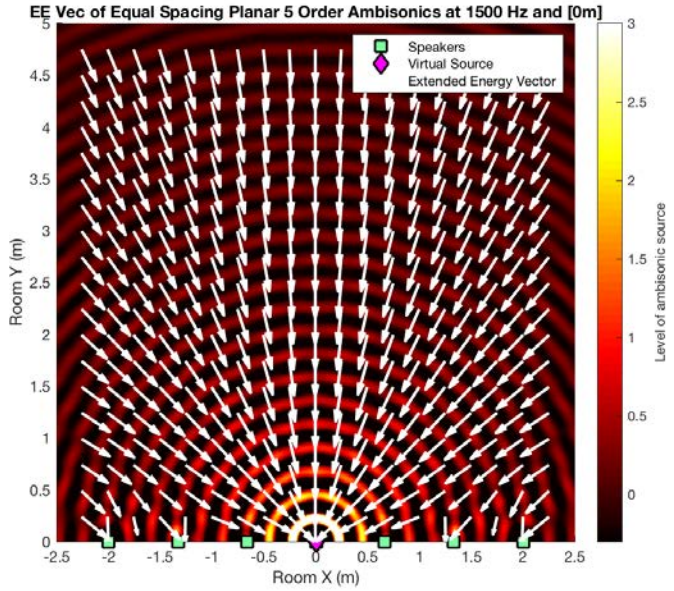
Equal spacing performs better, less localisation error at the edges of the array

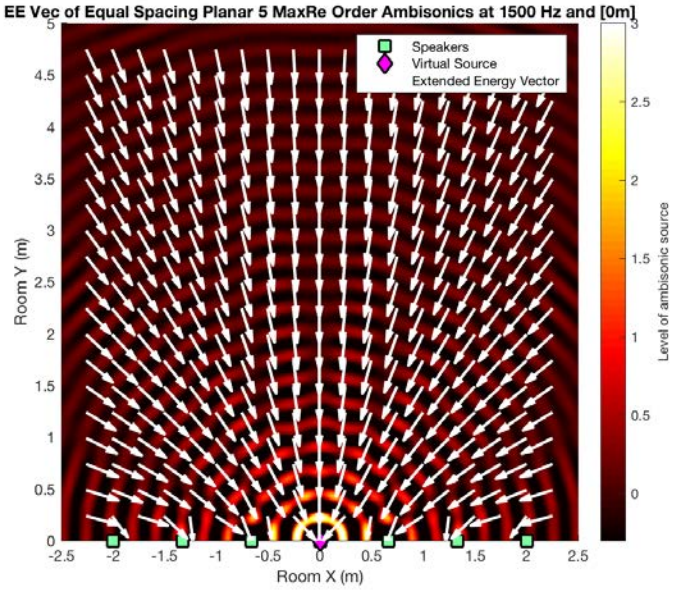


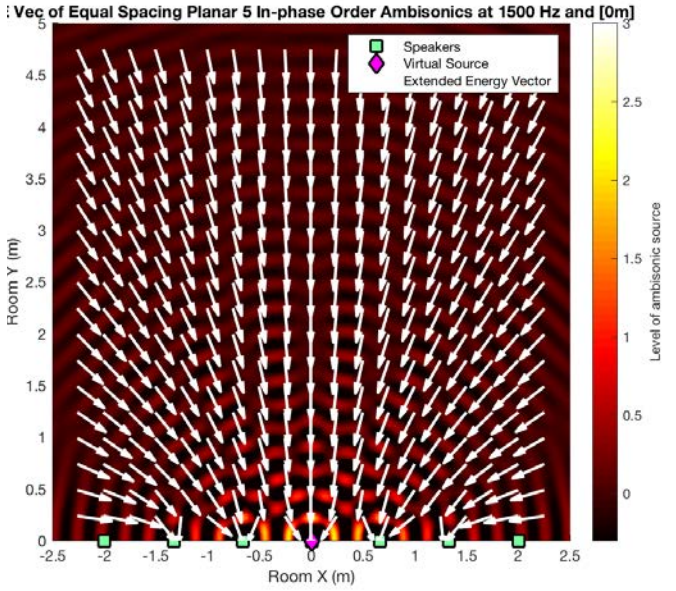


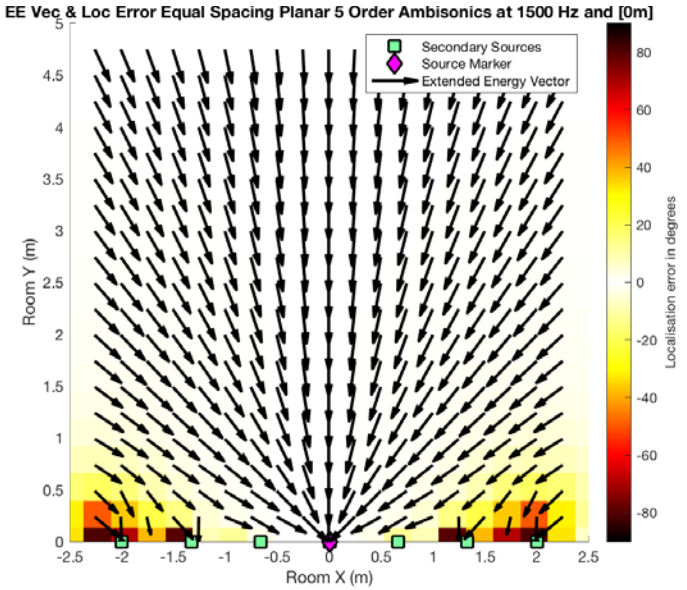
UNIVERSITY of DERBY

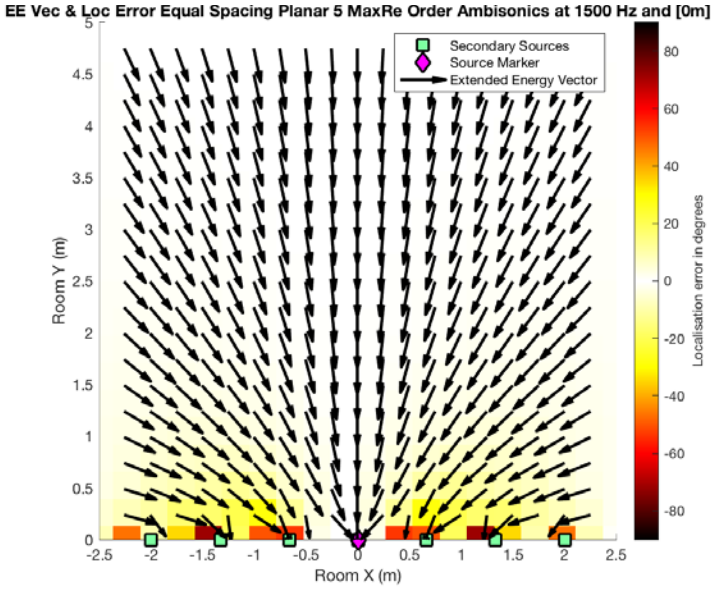
Basic Decoding/MaxRE/In-phase at 5th Order

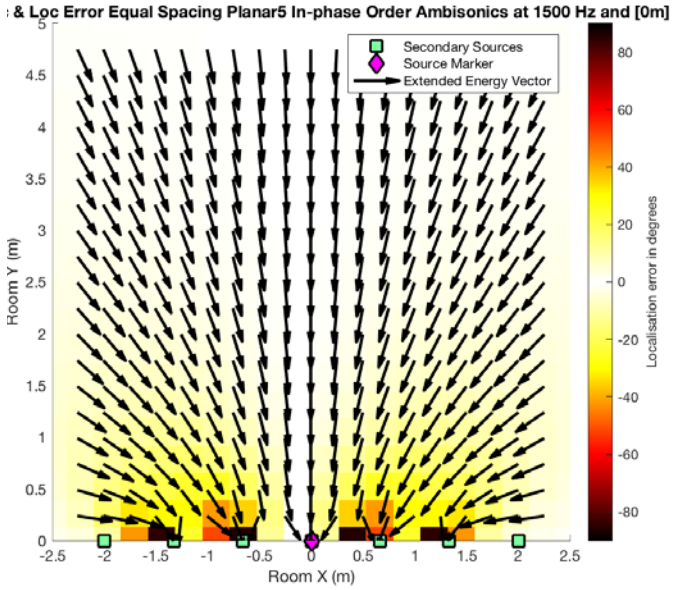




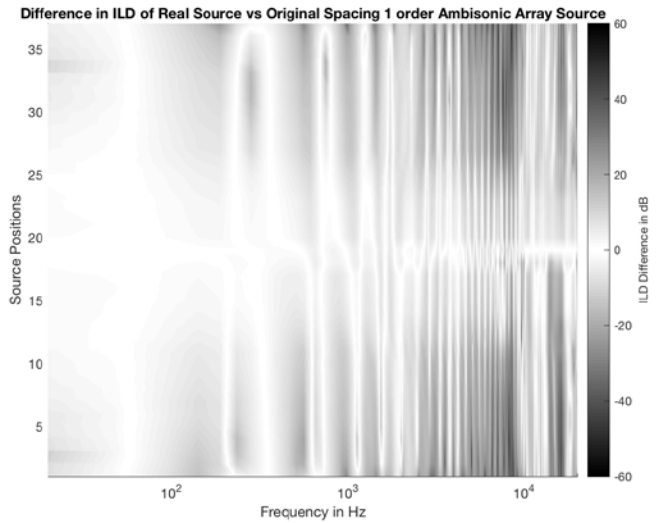


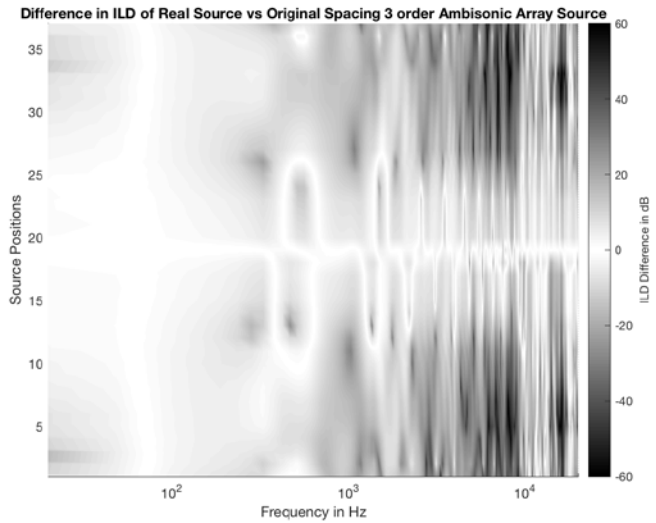


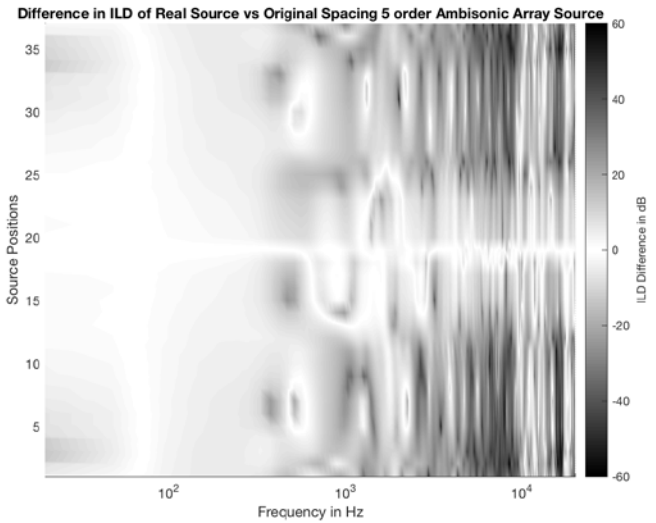


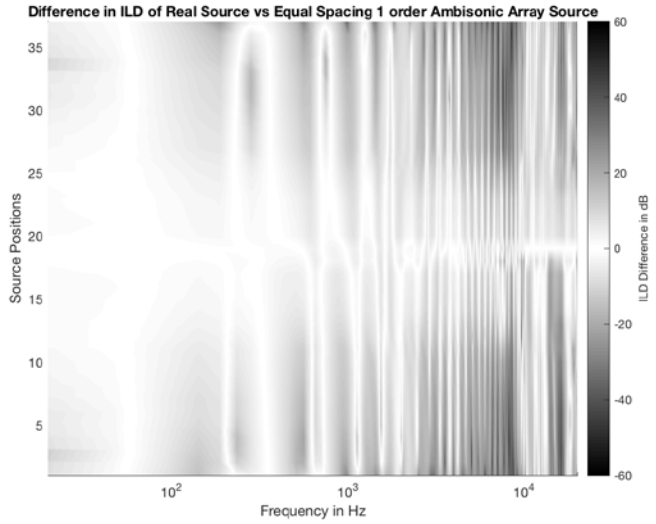


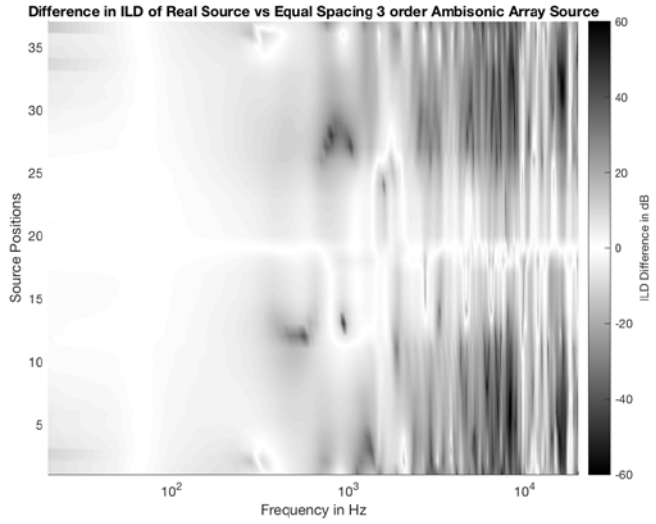
HRTF Analysis - Original VS Equally Spaced

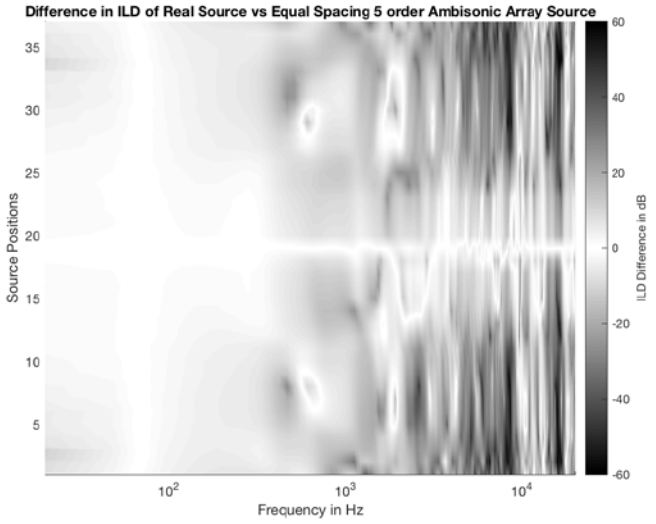






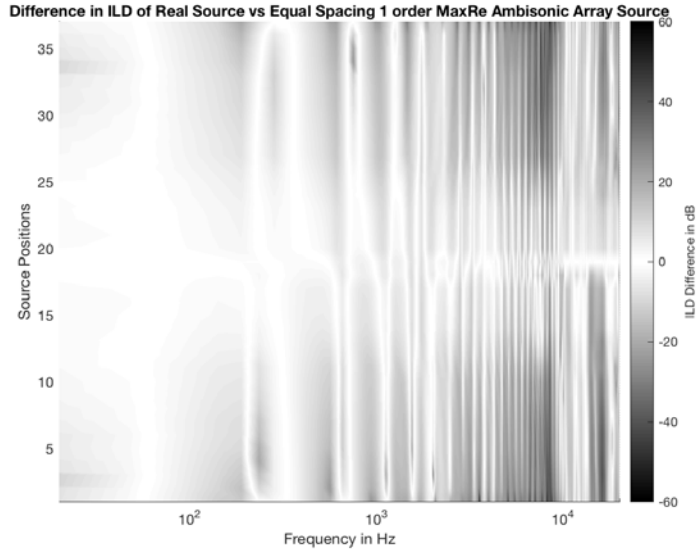


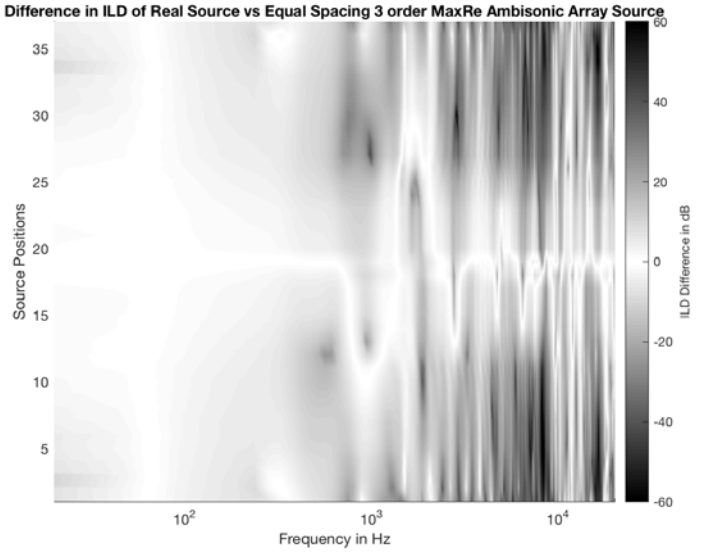


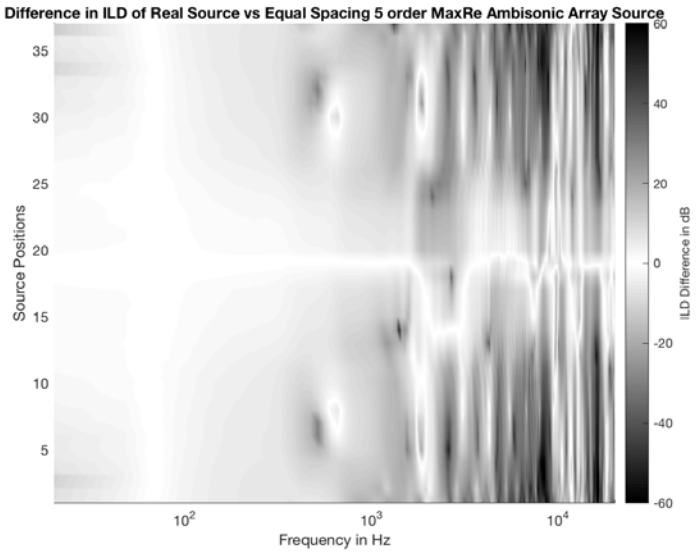


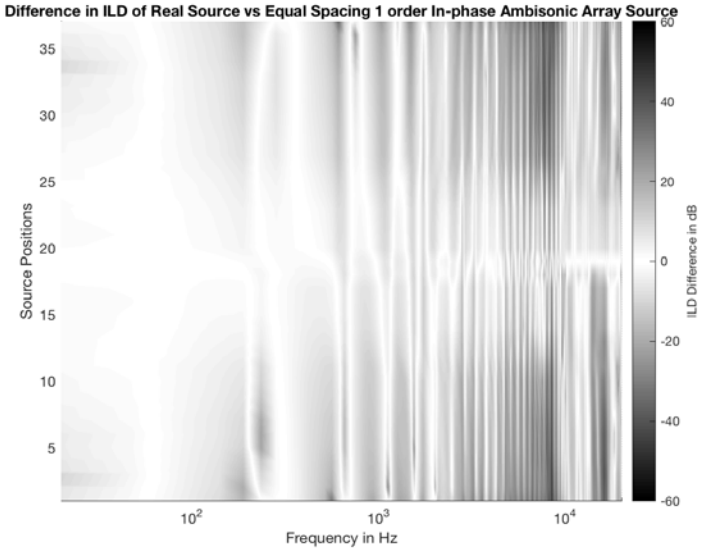
UNIVERSITY of DERBY

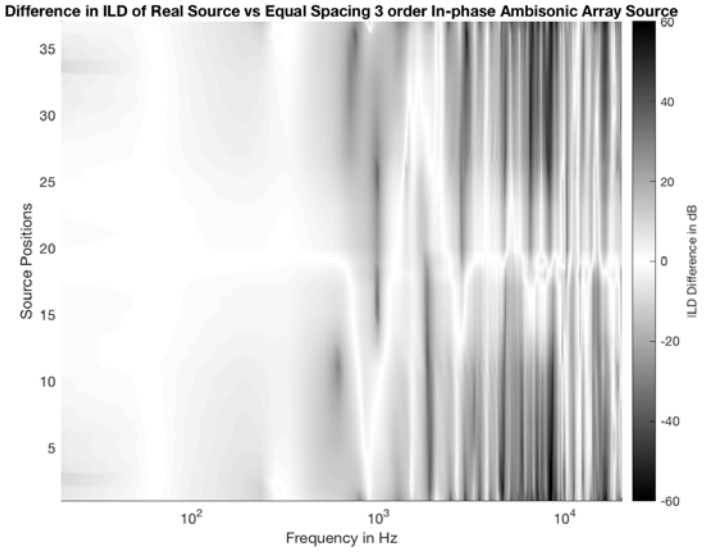
MaxRe vs In-phase

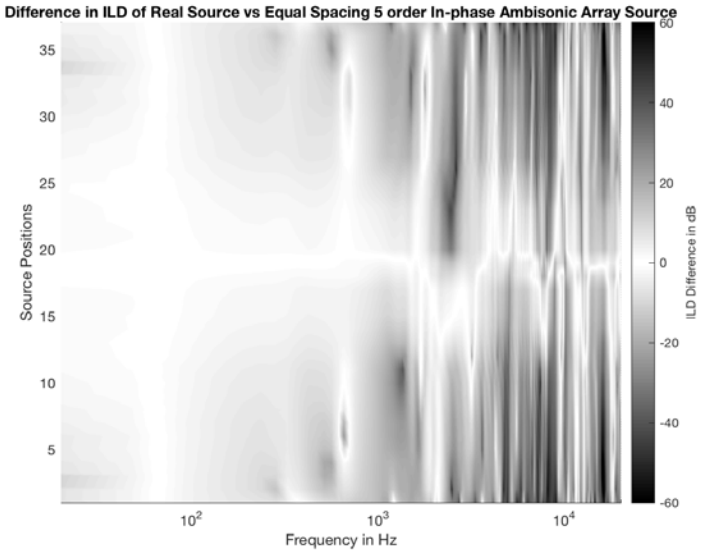






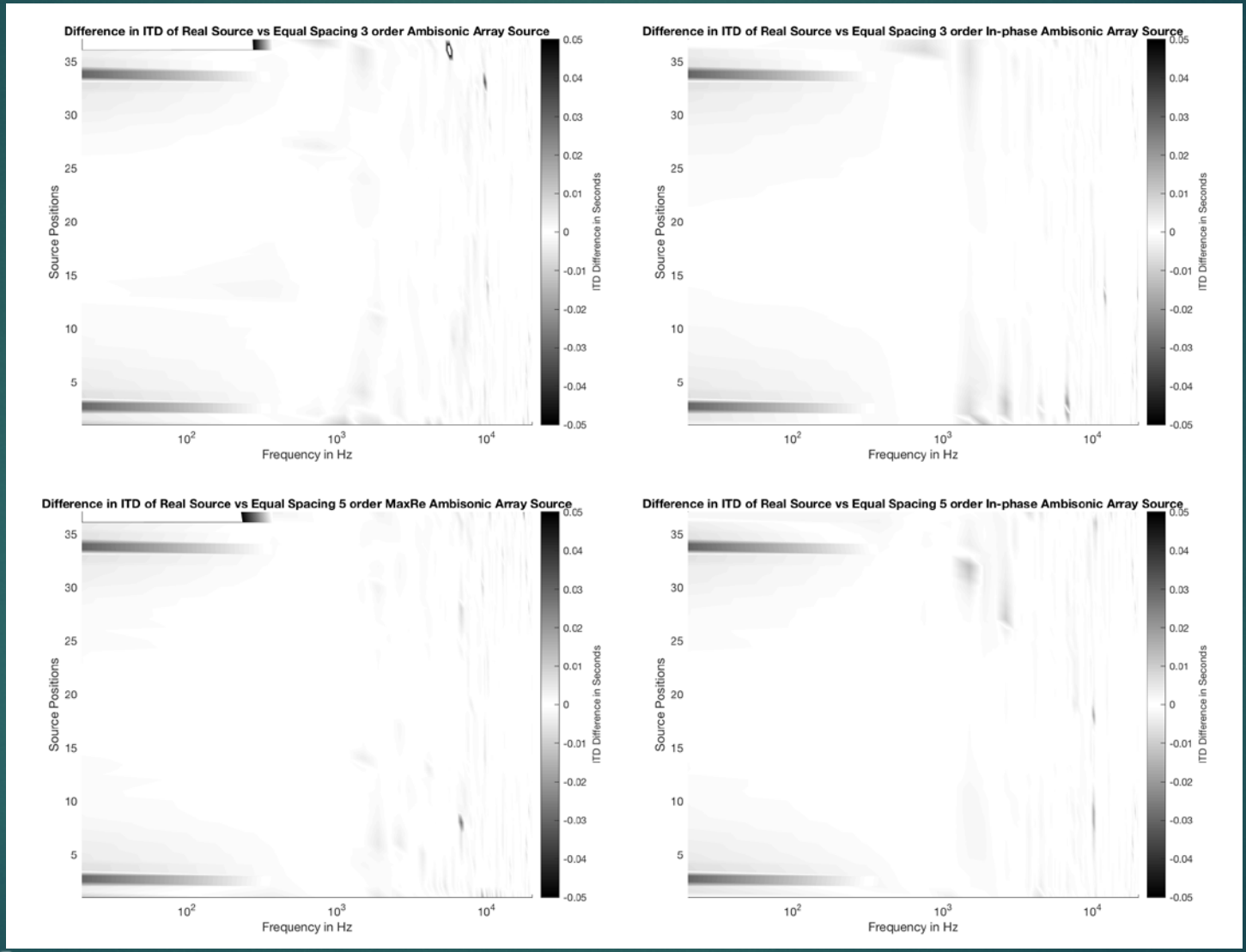






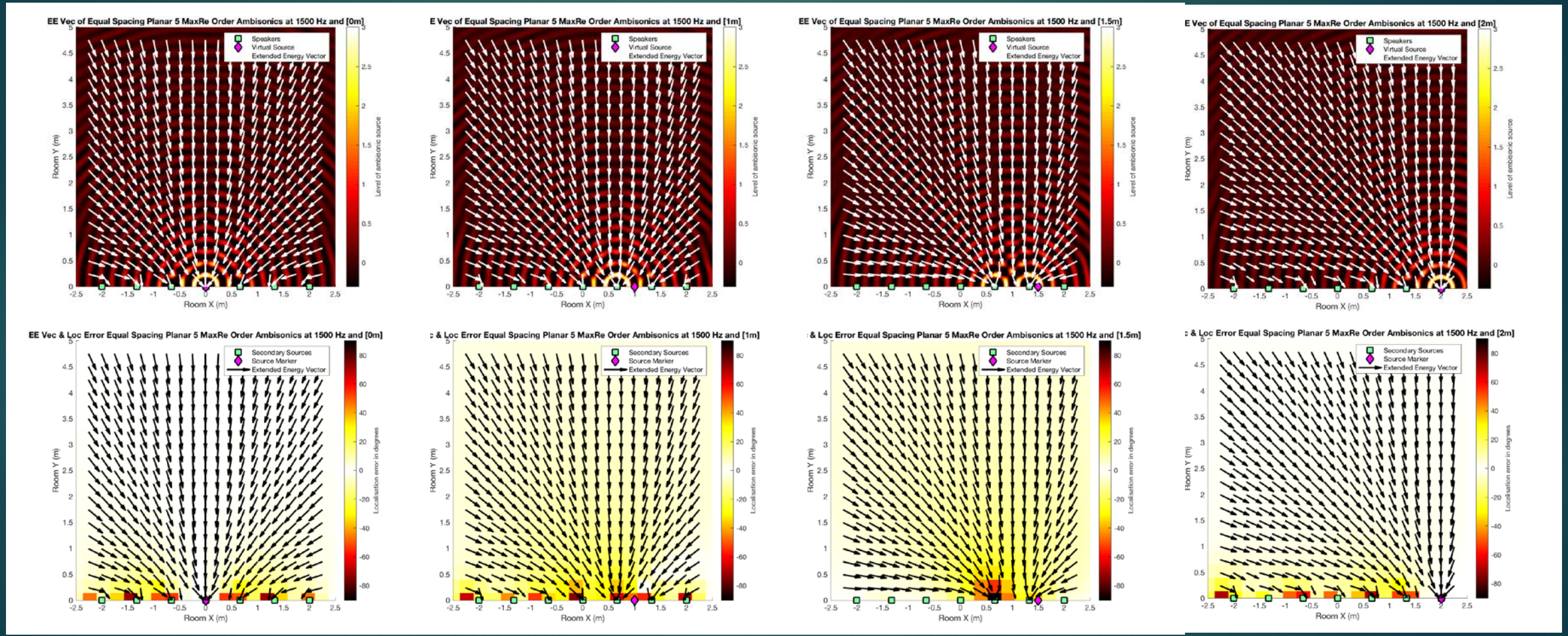
UNIVERSITY of DERBY

ITD Equal Spacing MaxRe vs In-phase



UNIVERSITY of DERBY

Off centre source?



UNIVERSITY of DERBY

Summary

- Equally spaced performs better than originally spaced
- Good localisation cues up to 1500Hz, need denser arrays to increase frequency at which localisation cues are reproduced accurately
- Higher the order the better the reproduction of virtual sources
- MaxRE best compromise for psychoacoustic optimisation, not In-phase which creates smearing of source (which ambisonic equivalent panning developed by Neukom and Schacher uses)
- ITD not a useful metric at this frequency range

Further Work

- Create a percentage error metric to allow quantitative comparison instead of visual
- Subjective listening tests
- Bigger, faster and denser measurement of HRTF's for higher order planar arrays – arduino controlled motor for the camera slider (and publically release)
- Distributed Kemar HRTF receiver positions in order to analyse off-centre performance
- Model cardioid source, currently assumed omnidirectional
- Compare planar HOA to WFS
- Investigate beamforming of sources to create surround reflections
- Distance encoding implementation

UNIVERSITY of DERBY