

# DISTANCE CODING AND PERFORMANCE OF THE MARK 5 AND ST350 SOUNDFIELD MICROPHONES AND THEIR SUITABILITY FOR AMBISONIC REPRODUCTION

T Spenceley    University of Derby, Derby, UK  
B Wiggins     University of Derby, Derby, UK

## 1 INTRODUCTION

Capturing and replaying distance cues for multi-channel audio is currently an under-explored and under-exploited area. Panners that successfully give control of distance do not, currently, exist. However, recordings made with 1<sup>st</sup> order ambisonic, Soundfield, microphones replayed over an ambisonic rig can give realistic results with respect to distance perception (particularly when bringing sound sources inside the speaker array).

Near-field effect, resulting from the wave front curvature of near-field sources, is one cue recorded by the microphone, but not reproduced by software or hardware panners. Papers by Daniel<sup>1 & 2</sup> discuss the encoding and decoding of ambisonic material with particular reference to higher-order ambisonics, and describe 'near-field coding' filters which encode near-field effect while pre-compensating for finite loudspeaker reproduction distance. While existing research concentrates on its simulation, this report documents an investigation into near-field effect in Soundfield ST350 and MKV tetrahedral microphones. It is found that, as a result of calibration for a flat frequency response at a practical source distance, the Soundfield microphone responses bear strong similarity to various near-field coding filters, suggesting the existence of an optimum loudspeaker array radius for positional localisation. On determination of this distance, recordings may be adapted for proper reproduction at any chosen reference distance using the WigWare ambisonic plug-ins created at the University of Derby.

## 2 SOUNDFIELD MICROPHONES

### 2.1 Overview

In order for a microphone to be able to successfully capture the information necessary for ambisonic reproduction, the polar patterns as shown in Figure 1 must be recorded from the same point in space. These signals are collectively known as B-Format, and are sufficient to reproduce with-height 1<sup>st</sup> order ambisonics.

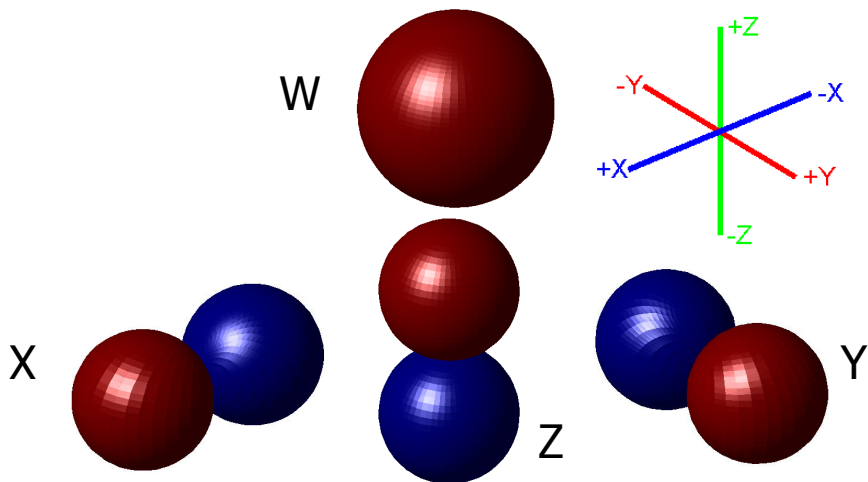


Figure 1 – 1<sup>st</sup> Order B-Format Polar Patterns

However, recording coincidentally in three dimensions proves to be problematic. Coincident microphone techniques in two dimensions (such as Blumlein Stereo or Mid/Side) are possible where the microphones can be made coincident in the X – Y axis but not in the Z axis (although this still causes some mis-alignment problems); however, in three dimensions it is desirable for the polar patterns to be equally accurate in all three dimensions, which among other benefits, allows for steering of the microphone’s patterns to be carried out without compromise. This problem was solved by Gerzon and Craven<sup>3</sup> by the use of four sub-cardioid microphone capsules mounted in a tetrahedral arrangement. This arrangement is shown in Figure 2.



Figure 2 - Tetrahedral Arrangement of Sub-Cardioid Microphone Capsules.

The arrangement shown in Figure 2 is not exactly coincident, but is equally non-coincident in each axis’ direction. This simplifies correcting for the spacing of the capsules, in that each type of B-Format output will require the same filter (i.e. the omni output will require filtering in one way, all three figure of eight responses in another).

The capsules are orientated in the directions shown in Table 1. The signals coming from these mic capsules are collectively known as A-Format.

Capsule	Azimuth	Elevation
A	45 <sup>0</sup>	35.3 <sup>0</sup>
B	135 <sup>0</sup>	-35.3 <sup>0</sup>
C	-45 <sup>0</sup>	-35.3 <sup>0</sup>
D	-135 <sup>0</sup>	35.3 <sup>0</sup>

Table 1 – Orientation of the four capsules in a SoundField Microphone.

Simple manipulations can be performed on these four capsules' outputs so as to construct the four pick-up patterns of B-Format as shown in Equation 1. A graphical representation of the four cardioid capsule responses and the four 1<sup>st</sup> order components derived from these are shown in Figure 3.

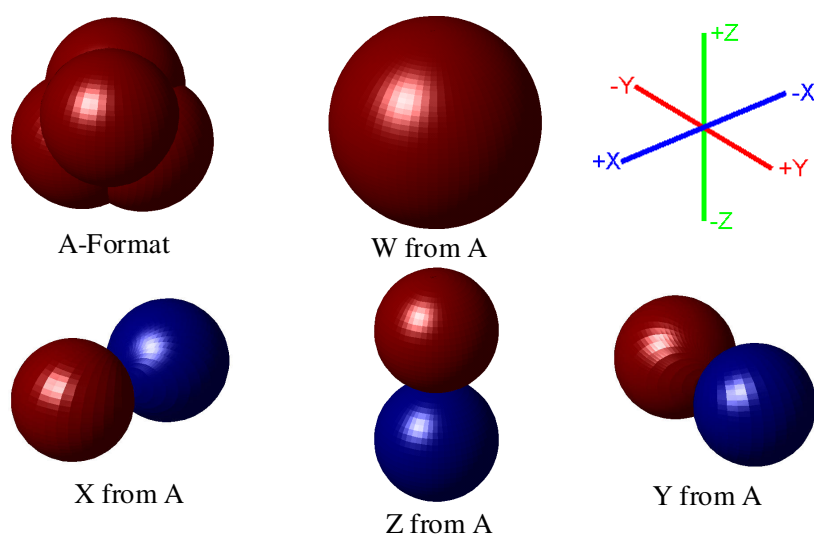
$$W = 0.5 \times (A + B + C + D)$$

$$X = (A + C) - (B + D)$$

$$Y = (A + B) - (C + D)$$

$$Z = (A + D) - (B + C)$$

**Equation 1 – Equations used to convert A-Format signals into B-Format.**



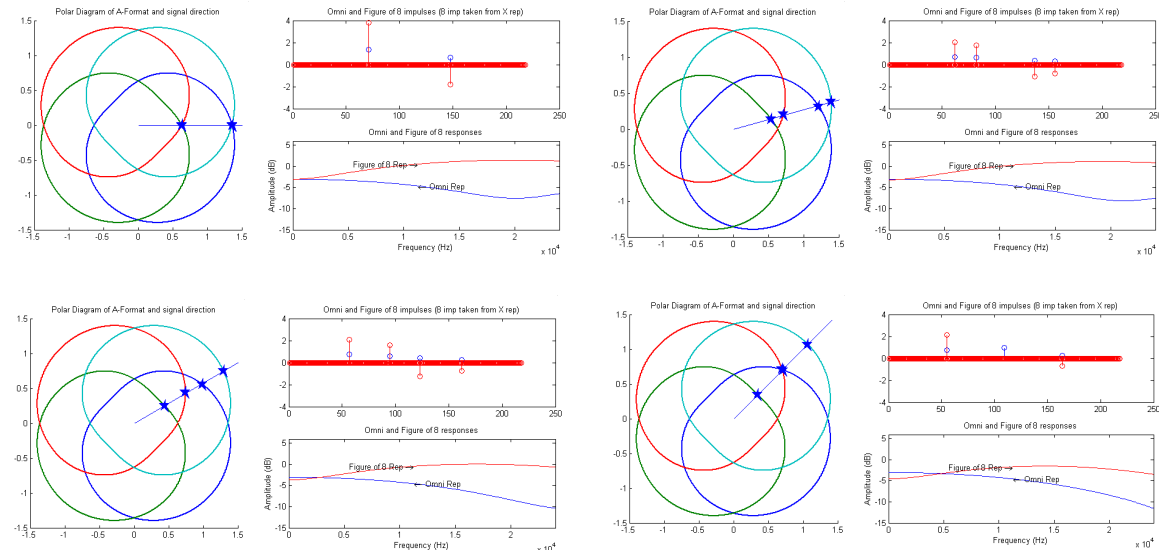
**Figure 3 - B-Format spherical harmonics derived from the four capsules of an A-Format microphone (assuming perfect coincidence). Red represents in-phase and blue represents out-of-phase pickup.**

The four capsules providing the A-Format signals are not perfectly coincident. This has the effect of misaligning the capsules in time/phase (they are so close that they do not significantly affect the amplitude response of the capsules, except at higher frequencies where shadowing is an issue), which results in colouration (filtering) of the resulting B-Format signals. As all of the capsules are equally non-coincident then any colouration will be the same for each order, i.e. the 0<sup>th</sup> order component will be filtered in one way, and the 1<sup>st</sup> order components will be filtered in another way.

To illustrate the frequency response characteristics of a SoundField microphone, it is simpler to assume that the microphone only works horizontally. Each of the four sub-cardioid capsules has no elevation angle, only an azimuth as described earlier. The equations that construct W, X, and Y will still be the same but the Z component will not be constructed. Figure 4 shows a number of representations of a sound being recorded from four different directions, 0°, 15°, 30° and 45° and indicates what amplitude each capsule will record, what timing mismatches will be present, and finally a frequency response for the W and X signals. It can be seen that the two channels not only have different frequency responses, but also these responses change as the source moves around the microphone. It must be remembered that the overall amplitude of the X channel will change due to the fact that the X channel has a figure of eight response.

Looking at Figure 4 shows a clear problem with having the capsules spaced in this way, and that is the fact that the frequency response of the B-Format signals change as the source moves around the microphone. The smaller the spacing, the less of a problem it becomes (as the changes move up in frequency due to the shortening of the wavelengths when compared to the spacing of the

capsules). Figure 4 is based on the approximate spacing of the SoundField MKV microphone capsules<sup>4</sup>.



**Figure 4 - Simulated frequency responses of a two-dimensional, multi-capsule A-Format to B-Format processing using a capsule spacing radius of 1.2cm.**

These responses can be corrected using filtering techniques, but only the average or on-axis response will be correct (depending on the calibration technique used), with the sound changing timbrally as it is moved around the microphone.

Although the frequency response deviations sound like a large problem, they are not generally noticed and are combined with other errors in the signal chain such as microphone capsule imperfections and loudspeaker responses. Also Farrah<sup>5</sup> claims that similar coincident stereo techniques have a far greater error than the SoundField microphone anyway – “*Closeness of the array allows compensations to be applied to produce B-format signal components effectively coincident up to about 10 kHz. This contrasts vividly with conventional stereo microphones where capsule spacing restricts coincident signals up to about 1.5 kHz*”. What is being referred to here is the frequency at which the filtering becomes non-constant. If the graphs in the omni-directional signal response are observed, it can be seen that its frequency response remains constant up to around 15 kHz, and it is the spacing and matching of the capsules that defines this frequency. The closer the capsules in space and response, the higher the frequency until non-uniformity is observed.

## 2.2 Near-Field Effect

The explanation of the workings of the Soundfield microphone above only take into account a sound source at a fixed distance from the microphone, however, the microphone’s response will change in a more complex manner when the distance of the source is taken into account.

Near-field, or proximity, effect refers to the magnitude and phase changes at low frequencies due to the difference in pressure gradient resulting from the spherical wave fronts of near-field sources. The prominence of this distance cue, present in 1<sup>st</sup> order ambisonic components and above, increases with order. While near-field effect naturally affects the 1<sup>st</sup> order components of any recording made with tetrahedral array microphones such as those produced by Soundfield, it must be simulated when encoding ambisonic material. For 1<sup>st</sup> order components, a 6.02dB per octave boost filter should be used<sup>6</sup>. The corner frequency for this filter can be determined by simple analysis of the pressure at a given distance from the source.

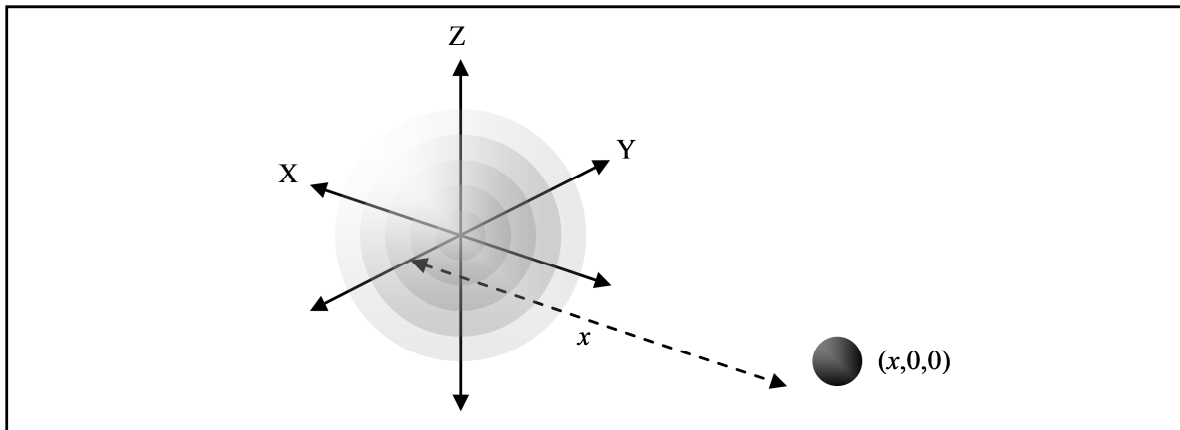


Figure 5 – Estimating pressure at an on-axis point at distance x from a source.

Assuming the pressure amplitude is unity at 1m, the pressure p for an on-axis point x metres from a sine wave source can be found by<sup>7</sup>:

$$p = \frac{1}{x} \sin(\omega(t - \frac{x}{c}))$$

Where  $\omega = 2\pi f$

**Equation 2**

To obtain the first-order component response to this pressure, the time-integral of the gradient of Equation 2 must be found:

$$g = \int \frac{dp}{dx} dt$$

$$\frac{dp}{dx} = -\frac{\omega}{cx} \cos(\omega(t - \frac{x}{c})) - \frac{1}{x^2} \sin(\omega(t - \frac{x}{c}))$$

$$g = -\frac{1}{cx} \sin(\omega(t - \frac{x}{c})) + \frac{1}{\omega x^2} \cos(\omega(t - \frac{x}{c})) + k$$

**Equation 3**

As the two terms of Equation 3 are 90° out-of-phase, it is clear that the 3dB cut-off frequency for a filter simulating near-field effect is located where the magnitudes of the two terms are equal:

$$\frac{1}{cx} = \frac{1}{\omega x^2}$$

$$\omega x = c$$

$$f = \frac{c}{2\pi x}$$

**Equation 4**

Figure 6 shows g (in red), and its two component terms, for distances of 1m and 0.5m. The first term of Equation 3 has been used as a reference in order to show clearly the relative amplification resulting from near-field effect. It should be noted that the method used thus far to determine near-field effect is only valid for 1<sup>st</sup> order components.

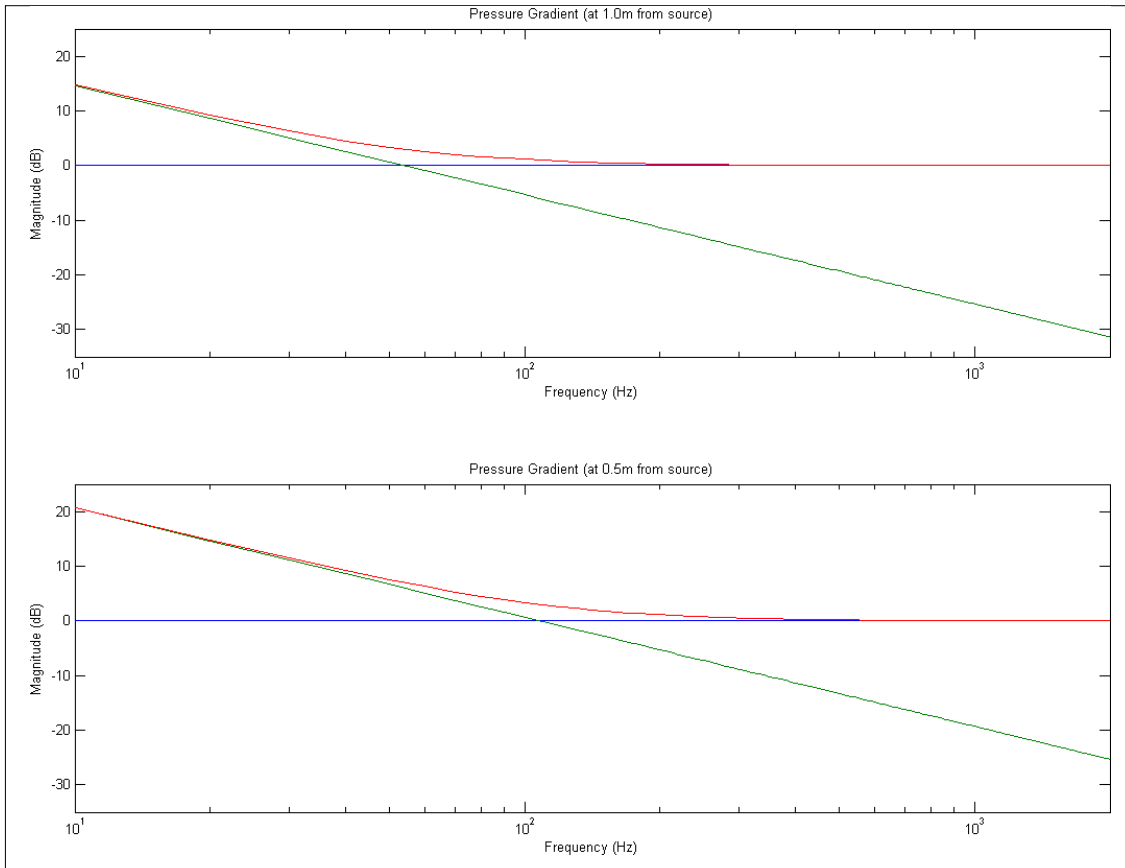


Figure 6 – Near-field effect on 1<sup>st</sup> order components at 1m and 0.5m.

Figure 7 illustrates the low-frequency boost resulting from near-field effect on ambisonic components up to 8<sup>th</sup> order. Filters matching these curves for the encoding of near-field effect would be unstable due to the infinite amplification at low frequencies<sup>1</sup>.

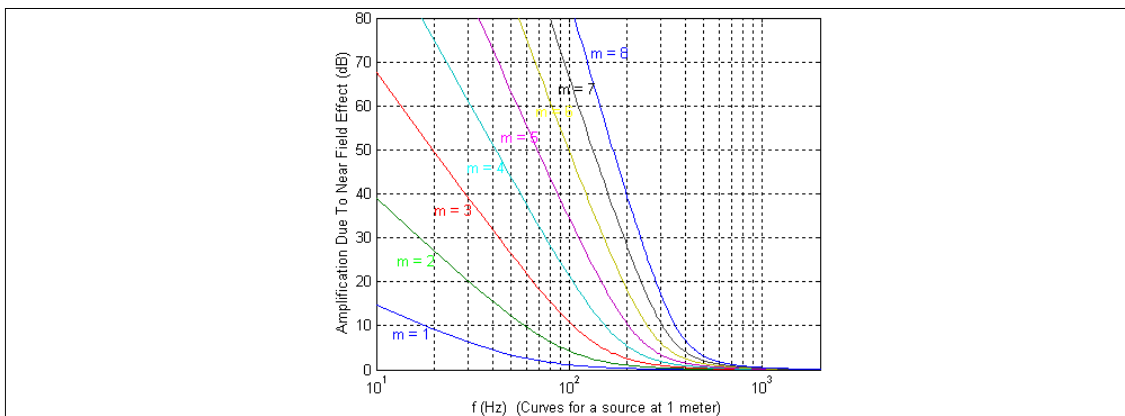


Figure 7 – Near-field effect on 1<sup>st</sup> to 8<sup>th</sup> order components at 1m taken from Daniel<sup>1</sup>.

### 2.3 Near-Field Compensation and NFC Filters

So far, only the near-field effect resulting from near-field sources has been considered. Also of importance is the near-field effect of loudspeakers used to reproduce ambisonic material. In order

to accurately reproduce both far- and near-field sources, finite loudspeaker distance must be compensated for (see Gerzon<sup>6</sup>).

Daniel<sup>1</sup> proposes 'near-field coding' (NFC) filters, which encode near-field effect while pre-compensating for loudspeaker distance. Crucially, these NFC filters give finite amplification, removing the instability issues seen in Figure 7. Example 1<sup>st</sup> to 4<sup>th</sup> order NFC filters, from a realisation by Adriaensen<sup>7</sup>, are shown in Figure 8. The final amplification of an m<sup>th</sup> order NFC shelving filter is given by:

$$m \times 20 \log_{10} \left( \frac{R}{\rho} \right) \text{ dB}$$

Where R is loudspeaker distance and ρ is source distance.

Equation 5

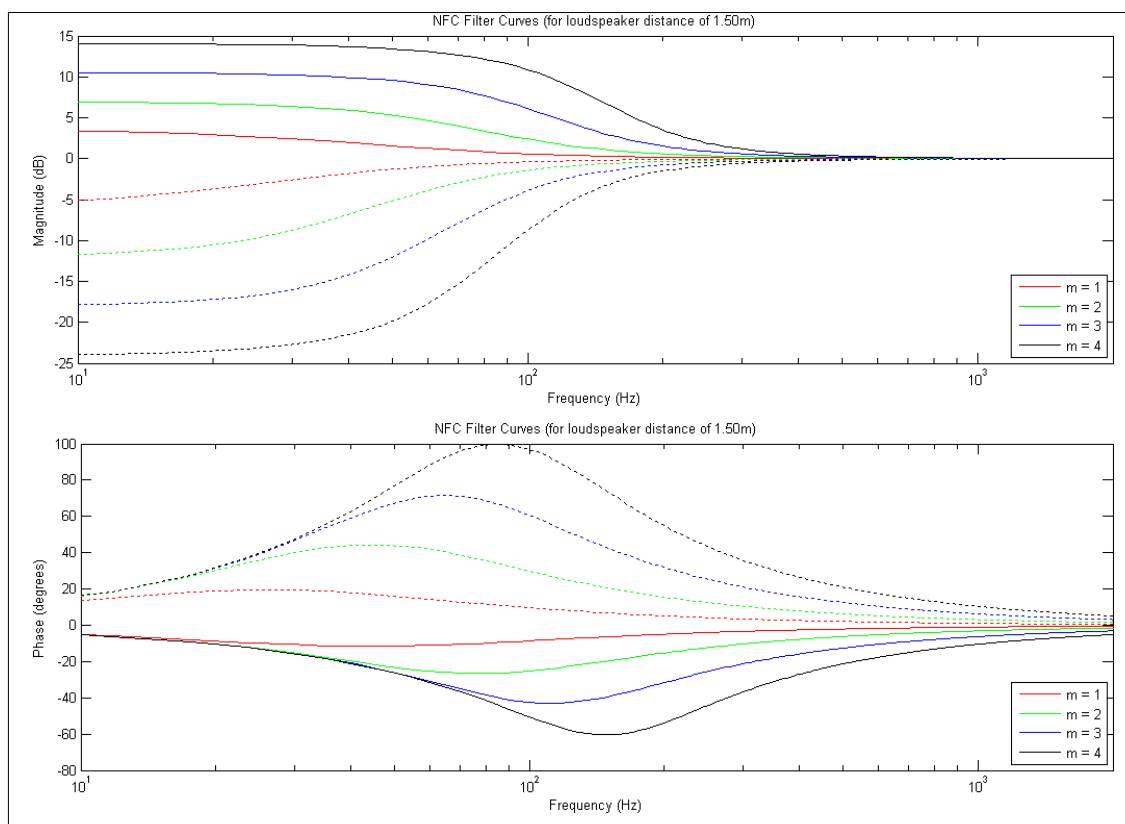


Figure 8 – 1<sup>st</sup> to 4<sup>th</sup> order NFC filters for sources at 1m and 3m, loudspeaker distance of 1.5m.

### 3 TEST PROCEDURE

The impulse responses of the Soundfield ST350 and MKV microphones were measured in the middle of a large, semi acoustically-treated room (a large lecture theatre). Measurements were taken at 250mm intervals between 250mm and 3m from an on-axis (positive X) source. 30-second logarithmic sine sweeps were recorded through the Soundfield microphones, and the impulse responses (for all channels) retrieved by convolution with the inverse filter using Aurora software (<http://www.aurora-plugins.com/>). Additionally, a reference recording was made with an Earthworks omnidirectional calibration microphone, positioned just above and ahead of the Soundfield microphones.

The obtained impulse responses were centred (while preserving the relative delay), truncated to 2601 samples and windowed with a Kaiser window (with a shape parameter value of  $5\pi$ ) to remove unwanted room response.

Anomalous distance-related characteristics were identified on the Y channel of both Soundfield microphones. These resulted from non-ideal loudspeaker positioning during the testing procedure. The loudspeaker was placed on the edge of some folded away raked seating where a horizontal channel seemed to disperse the loudspeaker sound. These anomalies were not present when the test was performed in a smaller room, but here the room response was too close in time to the mic response to be windowed out. Similar characteristics were not found on the Z channel.

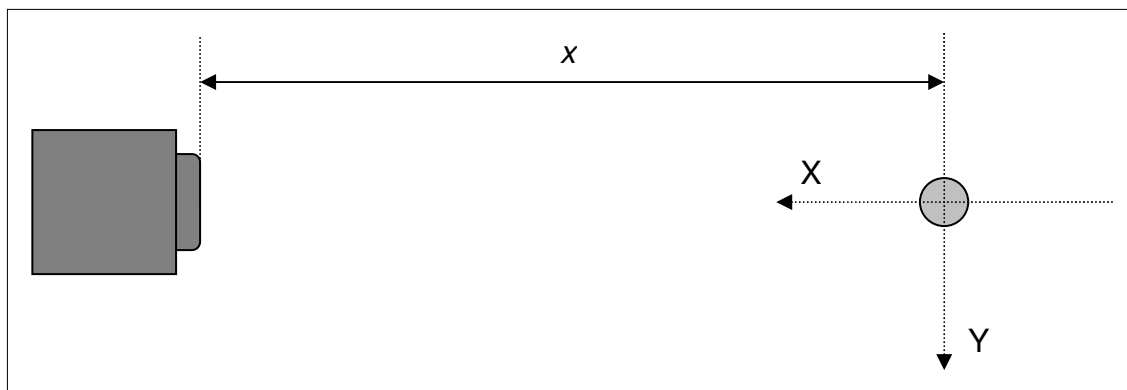


Figure 9 – Test set-up.

## 4 RESULTS

### 4.1 SoundField ST350

Figure 10 shows the W and X channel magnitude responses, with speaker characteristics still present. Figure 11 shows the difference between the X and W channels. Distance-dependent magnitude and phase differences, due to near-field effect, are present below approximately 300Hz and 1kHz, respectively. An absolute gain of 3dB is present in this graph, and all further 'X minus W' graphs, resulting from the expected 3dB level difference between the W channel and all 1<sup>st</sup> order components. This 3dB gain has been removed from all error graphs.



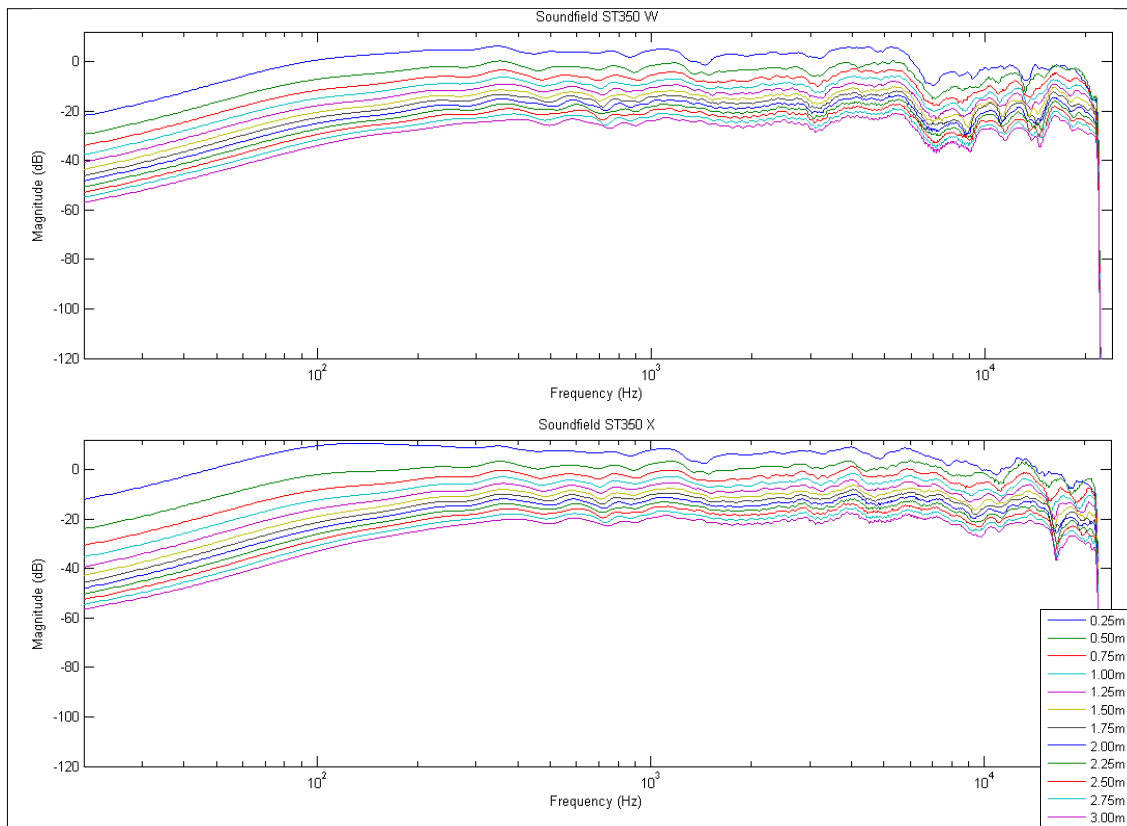


Figure 10 – ST350 W (top) and X (bottom) channel magnitude responses.

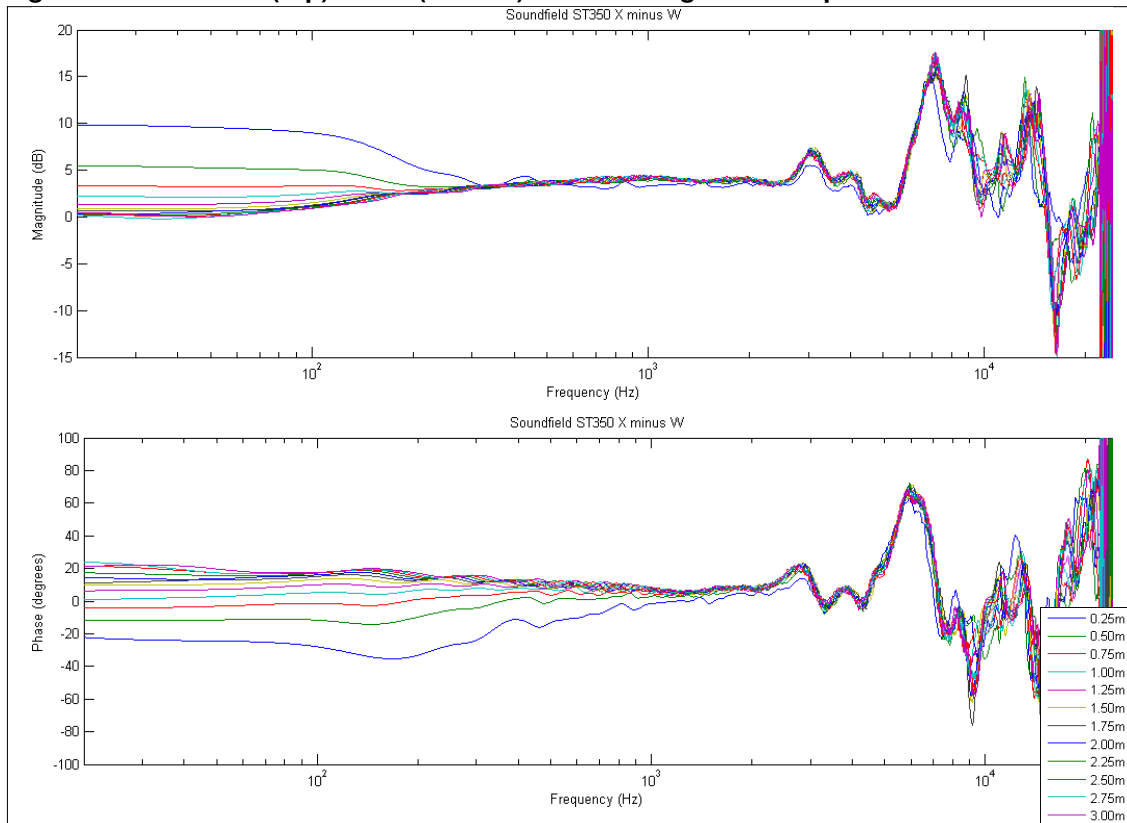
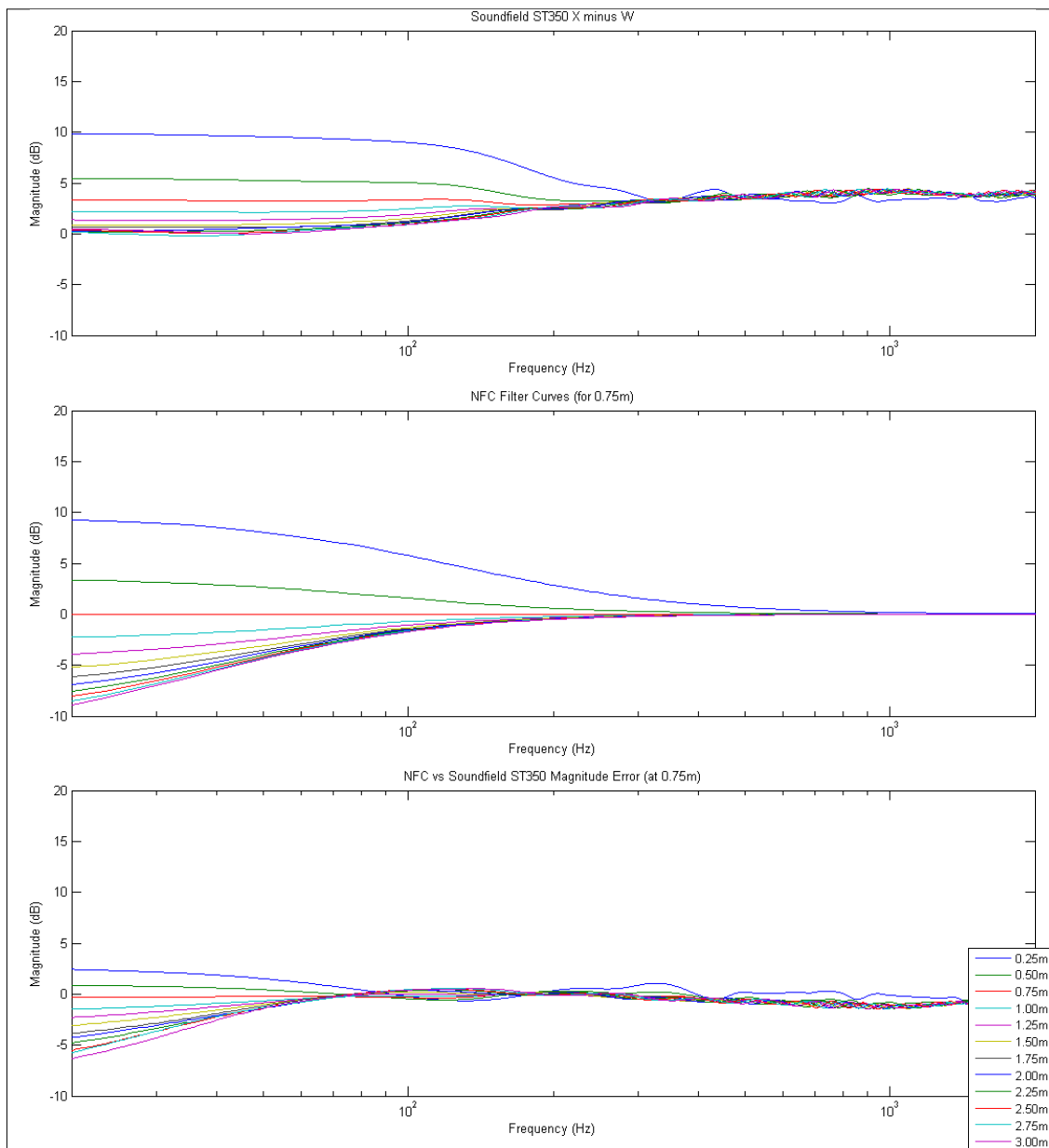


Figure 11 – ST350 magnitude (top) and phase (bottom) differences between X and W channels.



**Figure 12 – ST350 magnitude against 0.75m NFC filter magnitude (empirically matched), with error (3dB gain has been removed).**

Figure 12 plots the Soundfield ST350 low-frequency magnitude response against an (empirically selected) NFC filter with a reference distance of 0.75m. The error between the two is given at the bottom. While the two responses are similar, the ST350 provides both lower amplification for ‘inside’ sources and, in particular, lower attenuation for ‘outside’ sources than the NFC filter. The ST350 phase response is plotted similarly in Figure 13, against a 1.5m NFC filter. Additionally, two simple error minimisation algorithms were applied to attempt to find optimal reference distances, the results of which are given in Table 2. Graphs showing the error between the ST350 responses and these NFC filters can be found in the appendices. The results suggest an optimum reproduction distance of approximately 1m.

Error minimisation method	Magnitude ref. distance	Phase ref. distance
Empirical selection	0.75m	1.50m
Lowest absolute error	0.95m	1.00m
Lowest mean absolute error	1.15m	0.95m

Table 2 – Potential Soundfield ST350 reference distances.

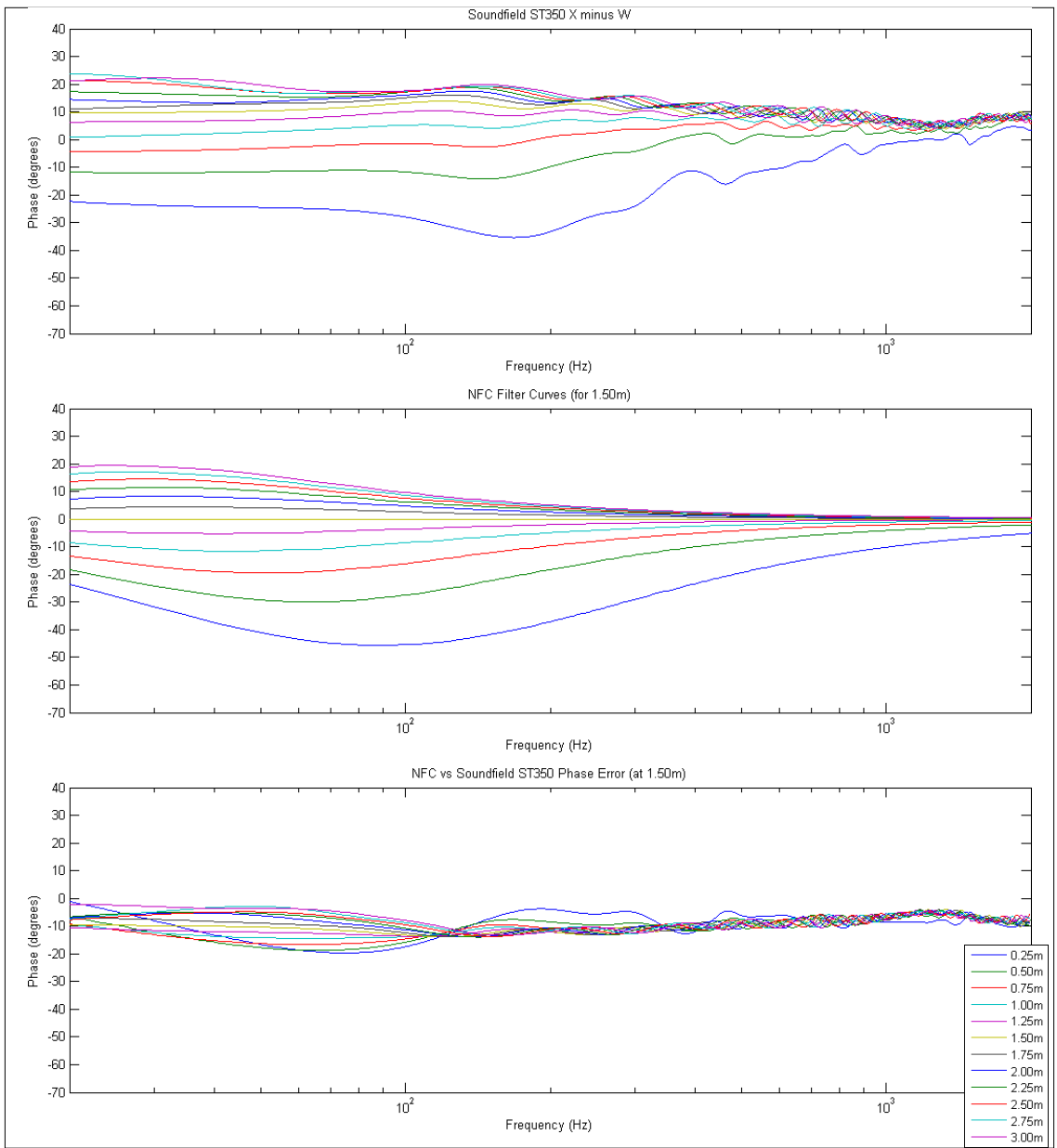


Figure 13 – ST350 phase against 1.5m NFC filter phase (empirically matched), with error.

### 4.2 SoundField Mark V

The Soundfield MKV microphone shows distance-related amplitude and phase differences below 300Hz and 1kHz, respectively. Like the ST350, the MKV magnitude response is similar to that of an NFC filter though with overall lower amplification and attenuation. The possible corresponding NFC filter reference distances are given in Table 3. Unlike the ST350, these results do not immediately suggest a single optimal reference distance, with a significant mismatch between magnitude and phase results.

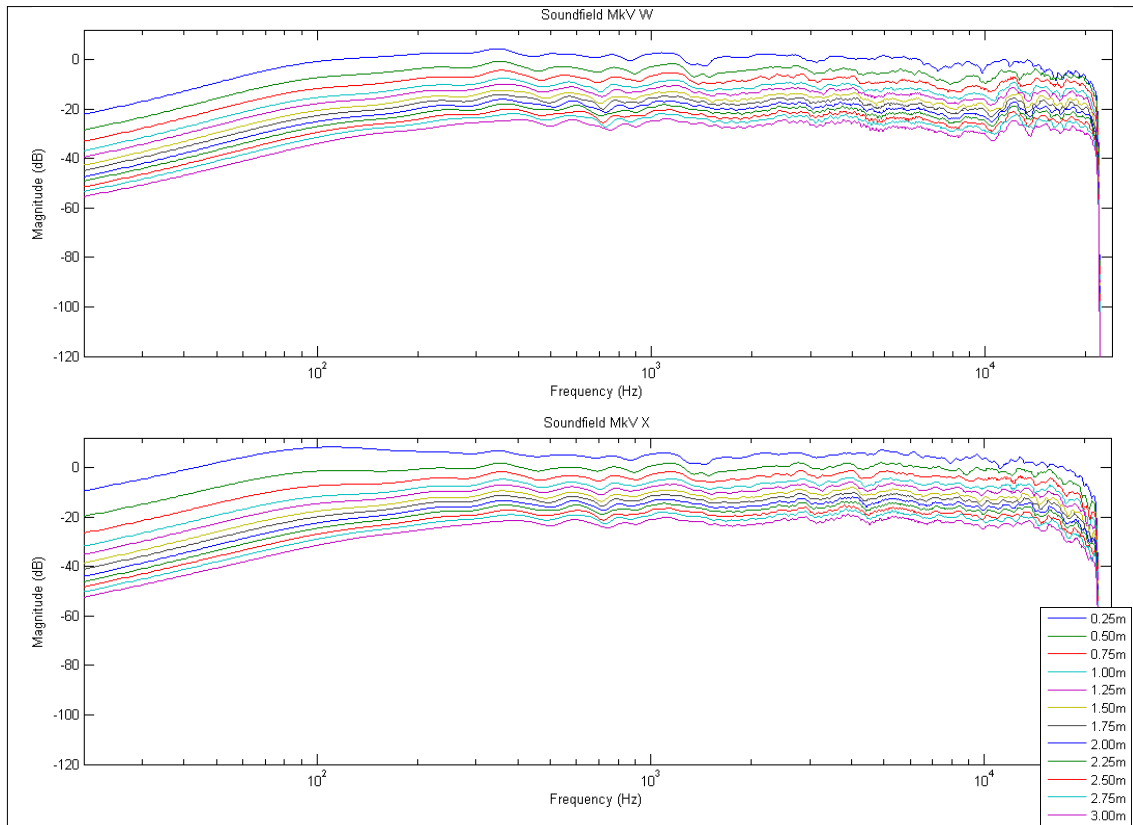


Figure 14 – MKV W (top) and X (bottom) channel magnitude responses.

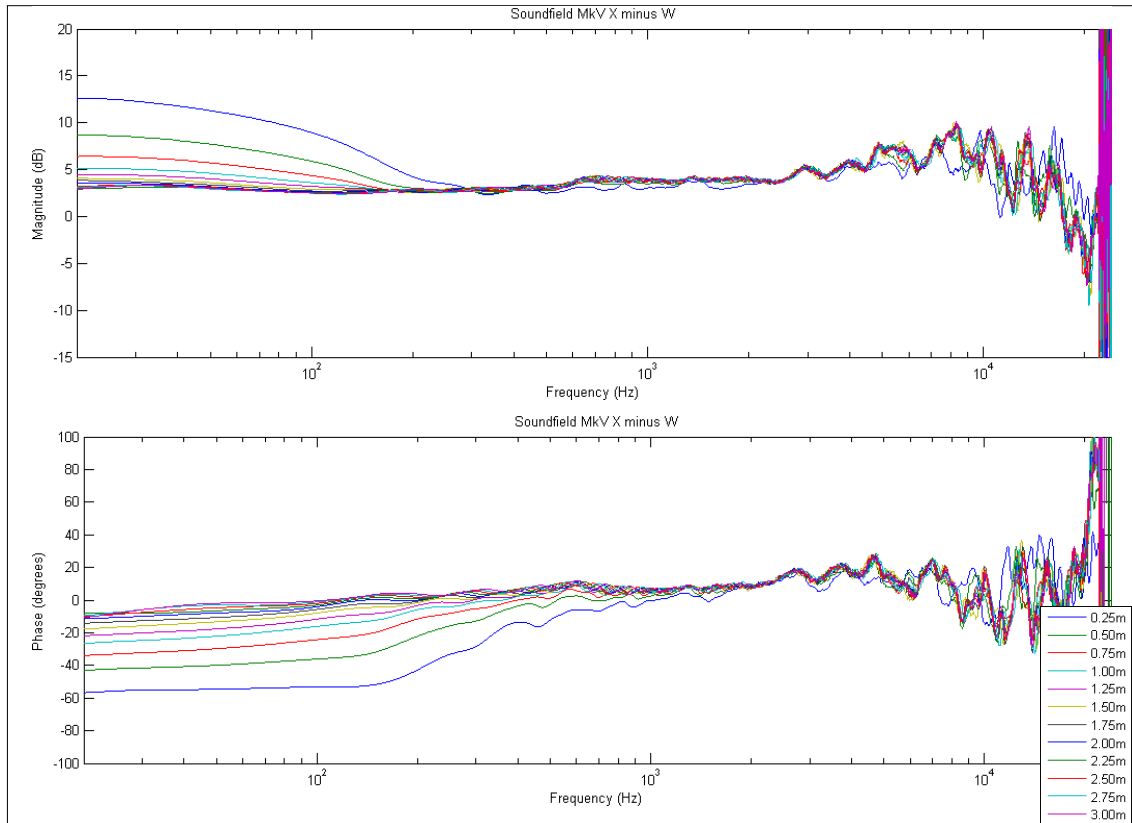


Figure 15 – MKV magnitude (top) and phase (bottom) differences between X and W channels.

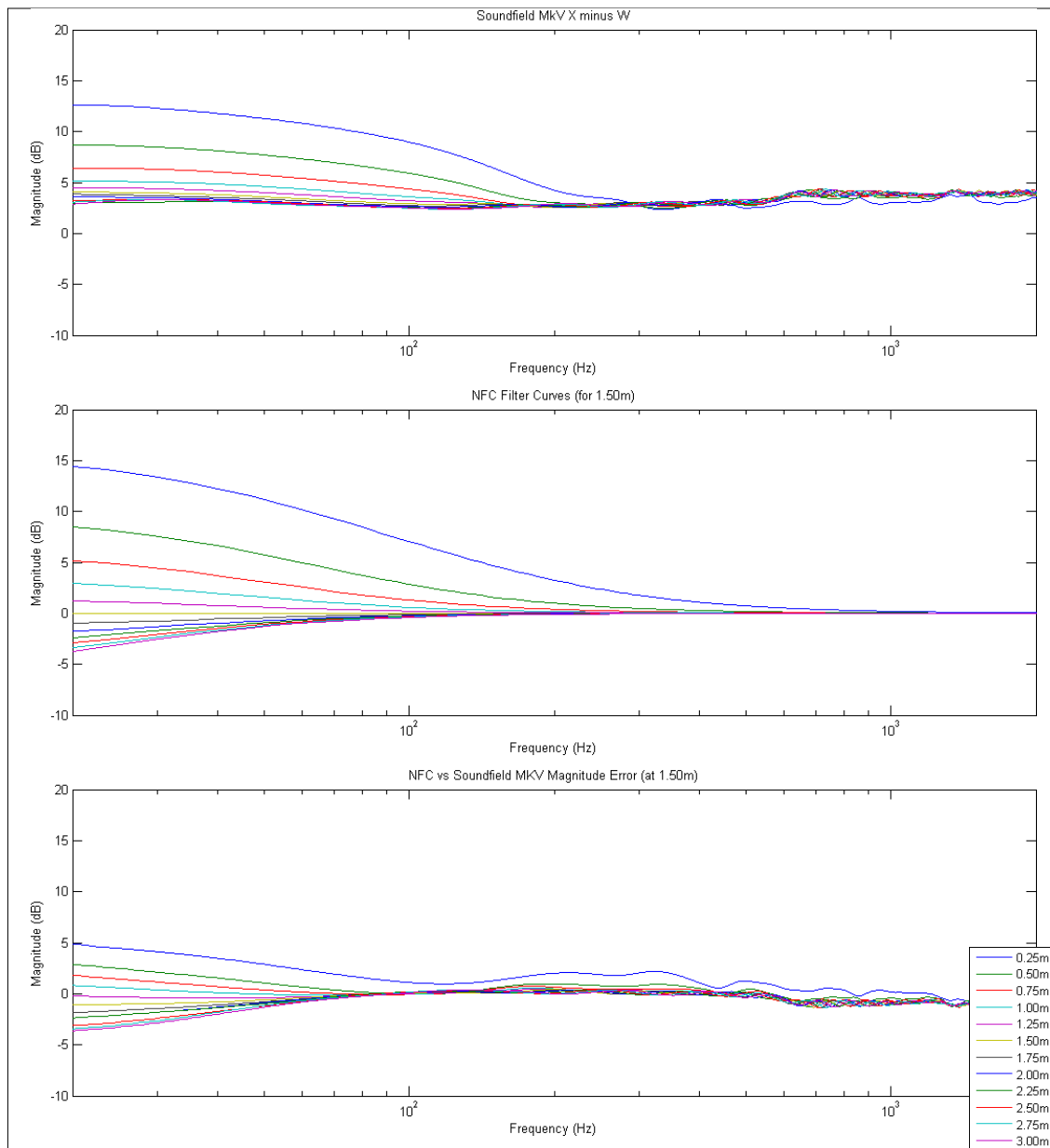


Figure 16 – MKV magnitude against 1.5m NFC filter magnitude (empirically matched), with error (3dB gain has been removed).

Error minimisation method	Magnitude ref. distance	Phase ref. distance
Empirical selection	1.50m	5.00m
Lowest absolute error	1.35m	3.15m
Lowest mean absolute error	1.80m	3.10m

Table 3 – Potential Soundfield MKV reference distances.

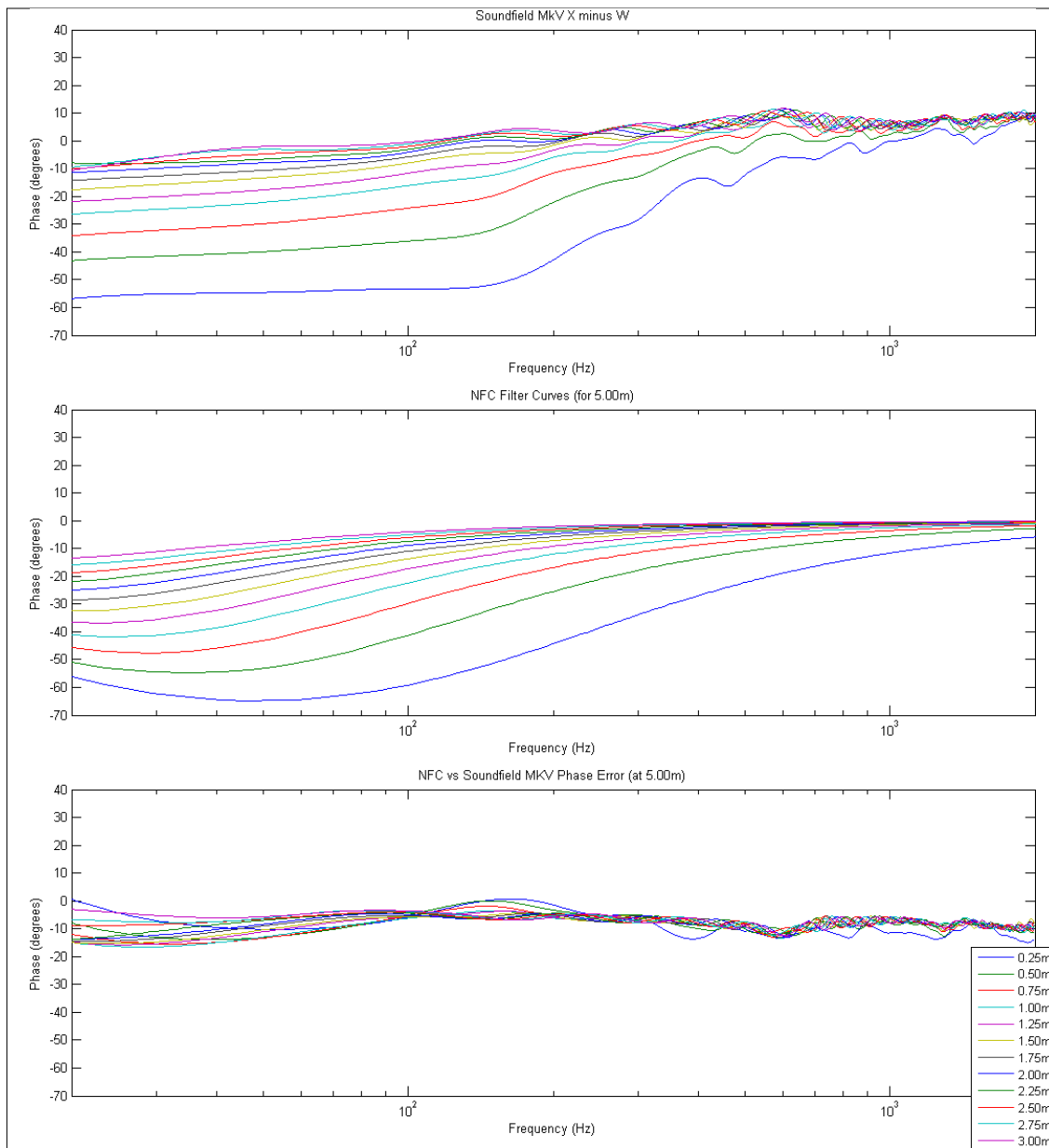


Figure 17 – MKV phase against 5m NFC filter phase (empirically matched), with error.

## 5 CONCLUSIONS

In this report, near-field effect and its relevance in ambisonics was discussed. An equation for the 1<sup>st</sup> order component response to an on-axis sine wave source of given frequency and distance was determined, and from it was derived Equation 4 given by Gerzon<sup>6</sup> for the cut-off frequency of a 1<sup>st</sup> order filter simulating near-field effect. It is important to note that the analysis performed may not be extended to higher-order components.

Impulse response measurements for the Soundfield ST350 and MKV microphones were taken for distances between 250mm and 3m. Using the W channel as a reference, the amplitude and phase differences caused by near-field effect were found. The responses, as a result of calibration for a flat response at a finite distance, were seen to be similar to those of various NFC filters. Simple error-minimising algorithms suggest that the optimum loudspeaker array radius for reproduction of Soundfield ST350 recordings is approximately 1m. A similarly conclusive result is not available for

the Soundfield MKV, which has a magnitude response like that of a 1.5m NFC filter, but a phase response close to that of NFC filters between 3 and 5m. From this a number of conclusions can be made:

1. The ST350 and MKV microphones are using different calibration filter schemes.
2. The ST350 seems to be calibrated flat for both phase and amplitude differences between the pressure and pressure gradient responses for a distance of around 1m. This is equivalent to the ST350 being calibrated for Near Field Compensation of 1m.
3. The MKV is calibrated for a flat frequency response at around 1.5m, but with the aim of not affecting the phase differences between the pressure and pressure-gradient response of the microphone.

In terms of compensating for these responses when the material is to be decoded ambisonically, then the recordings need to be treated differently depending on which microphone was used.

1. If the ST350 is used, then the Mic. Distance Compensation (equivalent to Daniel's NFC) should be set to around 1m. The Speaker Distance Compensation control should then be set to the distance of your speakers from the centre point.
2. If the MKV is used, then Mic. Distance Compensation should be set to 'Off', with the Speaker Distance Compensation control set to the distance of your speakers from the centre point.

The Wigware Ambisonic Decoders have been produced to take into account data regarding the microphone's (or panner's) calibrated distance. An example four speaker square decode setup for a recording made with an ST350 microphone with speakers placed 2 metres from the centre point is shown in Figure 18. The Wigware Ambisonic Plug-ins can be downloaded from <http://www.derby.ac.uk/staff-search/dr-bruce-wiggins>.

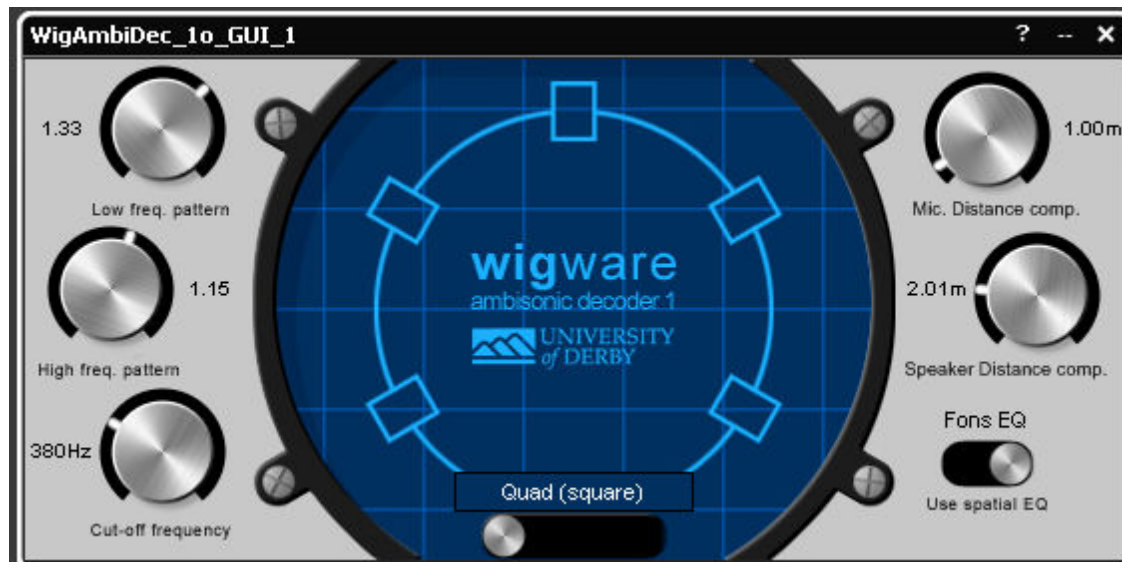


Figure 18 - Mic distance and speaker distance compensation settings for an ST350 mic and speaker array radius of 2m.

## 6 FURTHER WORK

This paper details the start of a project looking into distance panning and reproduction. The microphones' responses were studied for two reasons:

1. To see if the microphones followed the theory with regards to near-field response.
2. To make sure that the microphones weren't exhibiting other distance-related phenomena.



In this regard, this part of the project has been successful with useful and enlightening results found. The natural extension of this work is to ascertain just how important this low frequency cue is with regards to the perception of the reproduced sound field.

The microphones have been proven to follow simulated results by Daniel<sup>1 & 2</sup>; work must now shift towards looking at the other psycho-acoustic cues presented by the recordings that exhibit good distance perception. Again, Gerzon's paper on distance panning is a good starting point for this work<sup>6</sup>.

Much of the current work on 1<sup>st</sup> and higher order ambisonics concentrates on point source representation, but no real source is ever a true point source. As real sources approach the listener, especially if they are mechanical, the individual elements that make up the sound tend to separate, and are perceived as separate, rather than a single sound source. A B-Format recording of a motorbike which can be downloaded at [www.soundfield.com](http://www.soundfield.com) demonstrates this, and exhibits a particularly good sense of coming into the speaker array when ambisonically decoded.

Once the above has been carried out, current panners should be extended to reproduce these cues (the Wigware Ambisonic Panners have already been coded to reproduce near-field effect).

## **7 REFERENCES**

1. Daniel, J. Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. Copenhagen: AES 23<sup>rd</sup> International Conference (2003).
2. Daniel, J. Further Study of Sound Field Coding with Higher Order Ambisonics. Berlin: AES 116<sup>th</sup> Convention (2004).
3. Craven, P.G., Gerzon, M.A. Coincident Microphone Simulation Covering Three Dimensional Space and Yielding Various Directional Outputs, U.S. Patent no 4042779 (1977).
4. Farrah, K. Soundfield Microphone – Design and development of microphone and control unit. *Wireless World*, October, p. 48-50 (1979).
5. Farrar, K. Soundfield Microphone. Parts 1 & 2. *Wireless World*, October & November. p. 48 – 50 & p. 99 – 103 (1979).
6. Gerzon, M. The Design of Distance Panpots. Vienna: AES 92nd Convention (1992).
7. Adriaensen, F. Ambisonics paner [sic] with distance. [Discussion list]. Response to Miguel Nagrao. Sent 3rd May 2008, 13.48. Available at: <https://mail.music.vt.edu/mailman/private/sursound/> [accessed 16th June 2008].

## 8 APPENDICES

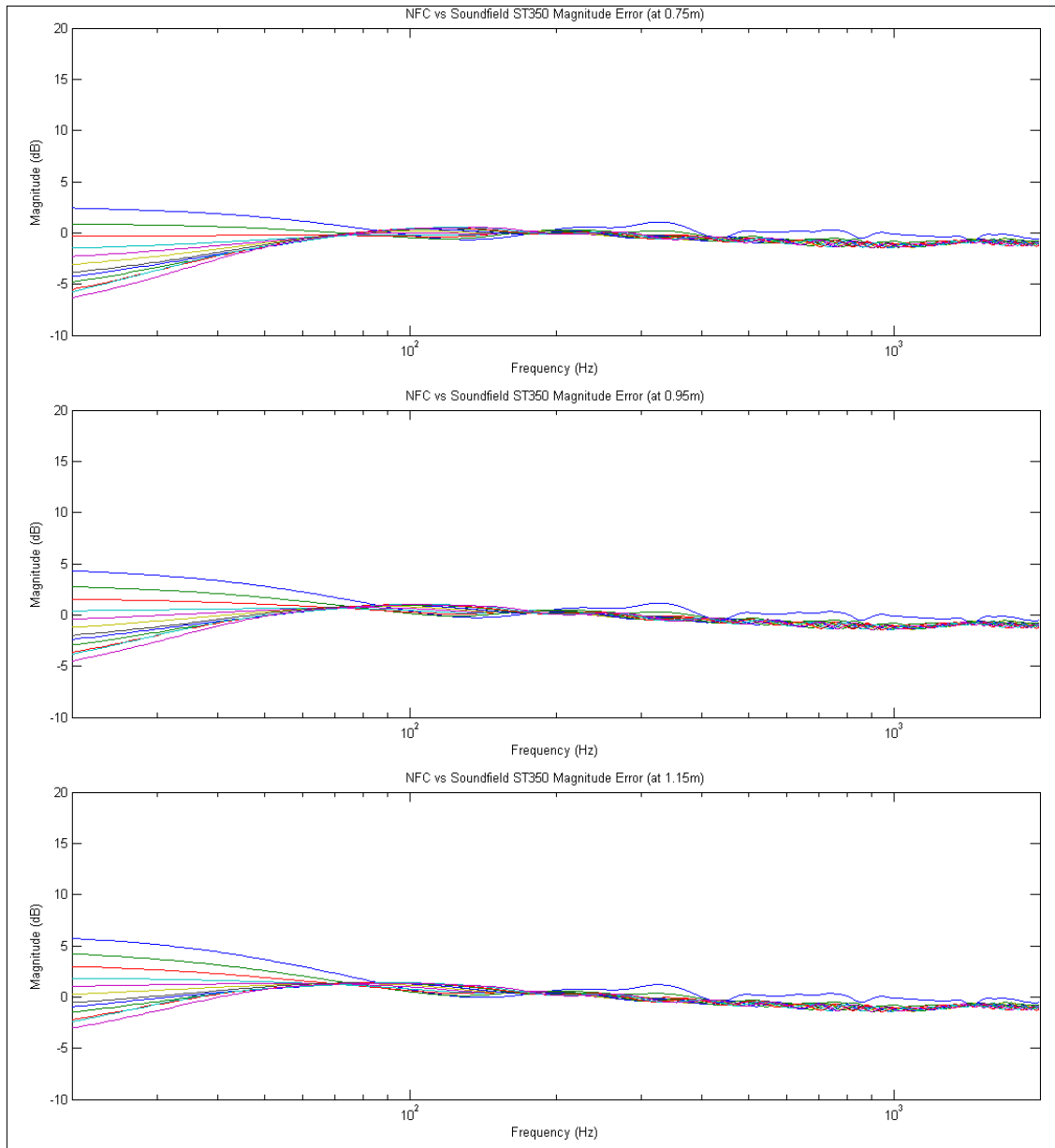


Figure 19 – ST350 magnitude error against 0.75m NFC filter (top), 0.95m NFC filter (middle) and 1.15m NFC filter (bottom)

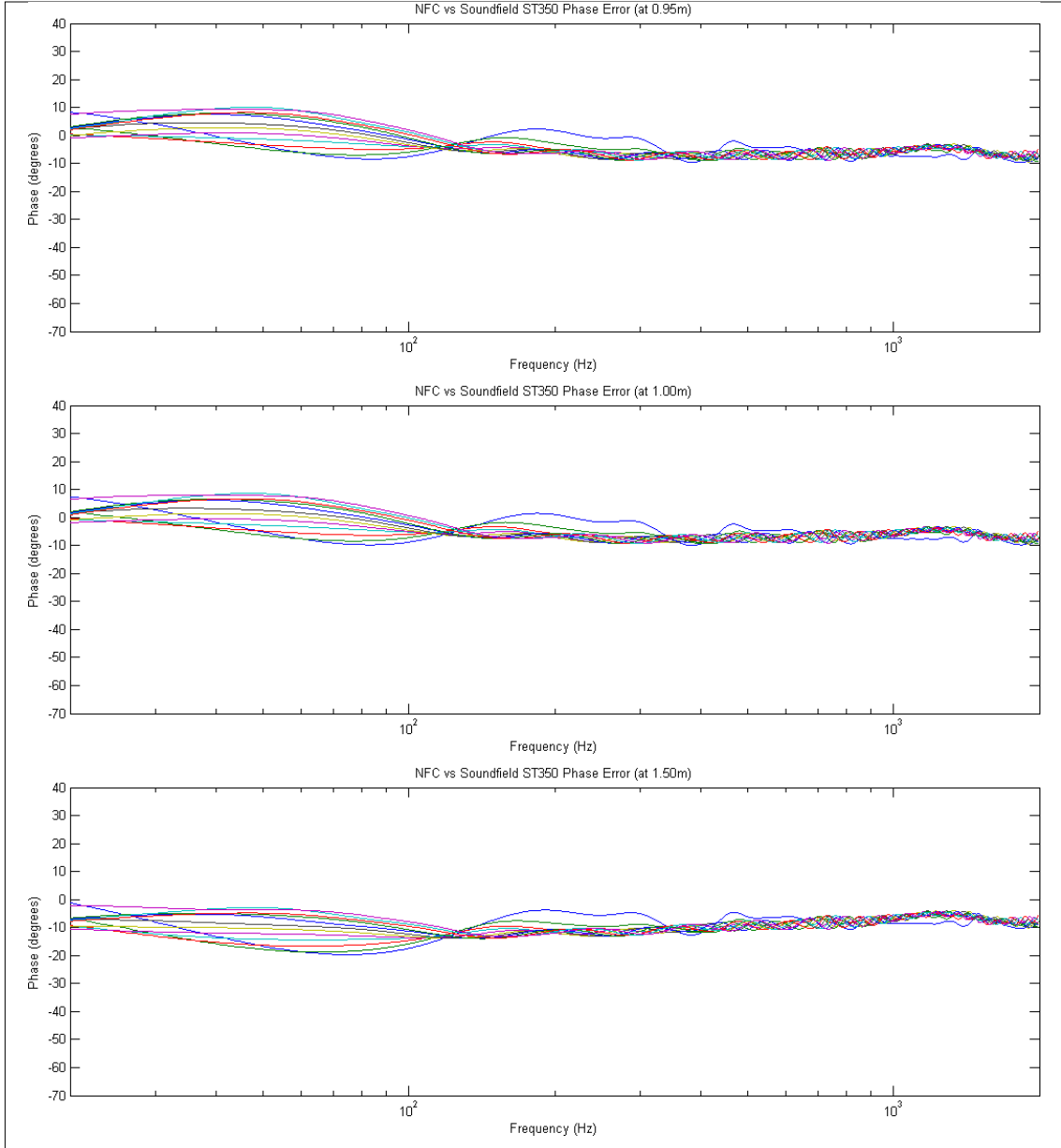


Figure 20 – ST350 phase error against 0.95m NFC filter (top), 1m NFC filter (middle) and 1.5m NFC filter (bottom)

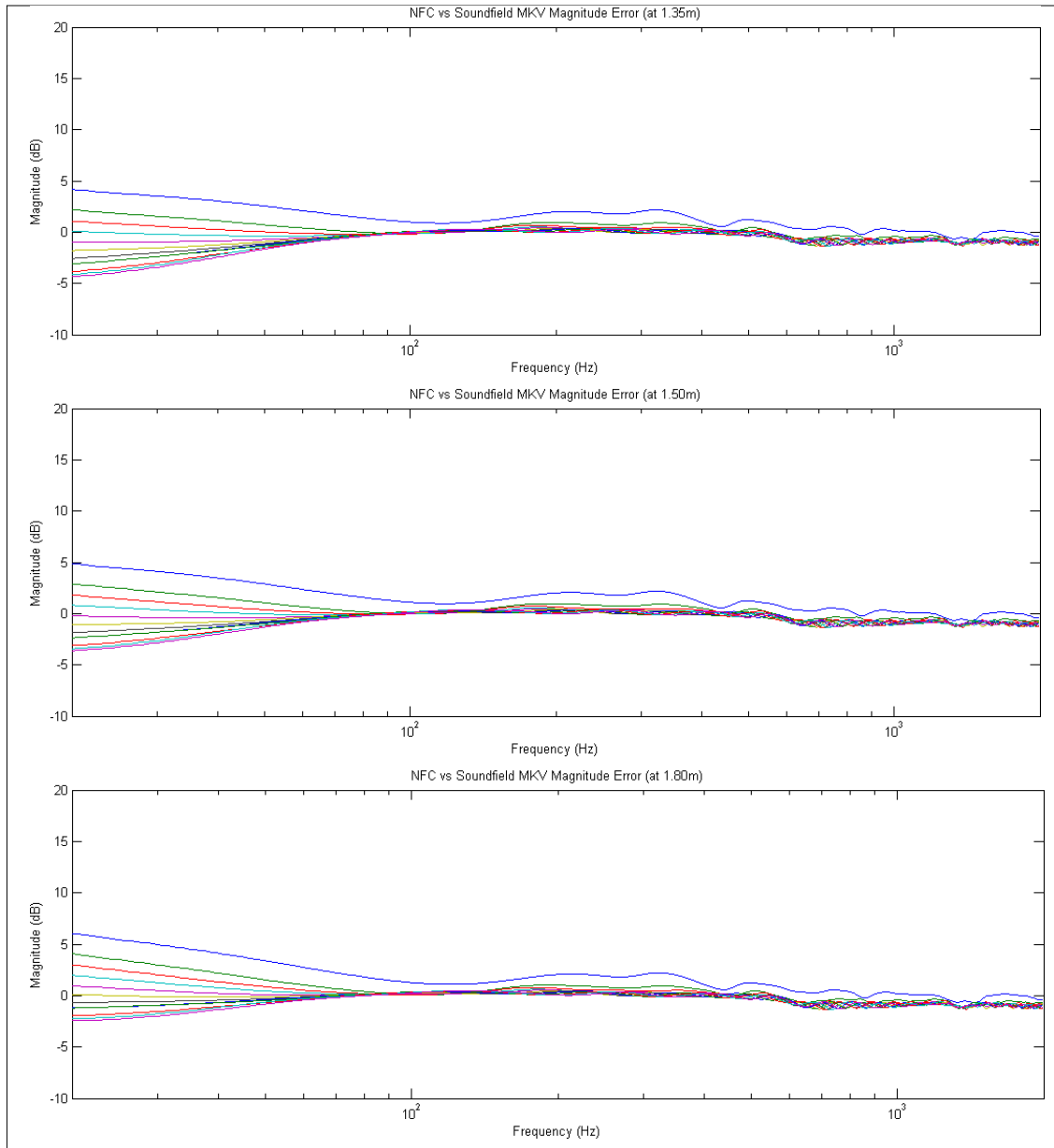


Figure 21 – MKV magnitude error against 1.35m NFC filter (top), 1.5m NFC filter (middle) and 1.8m NFC filter (bottom)

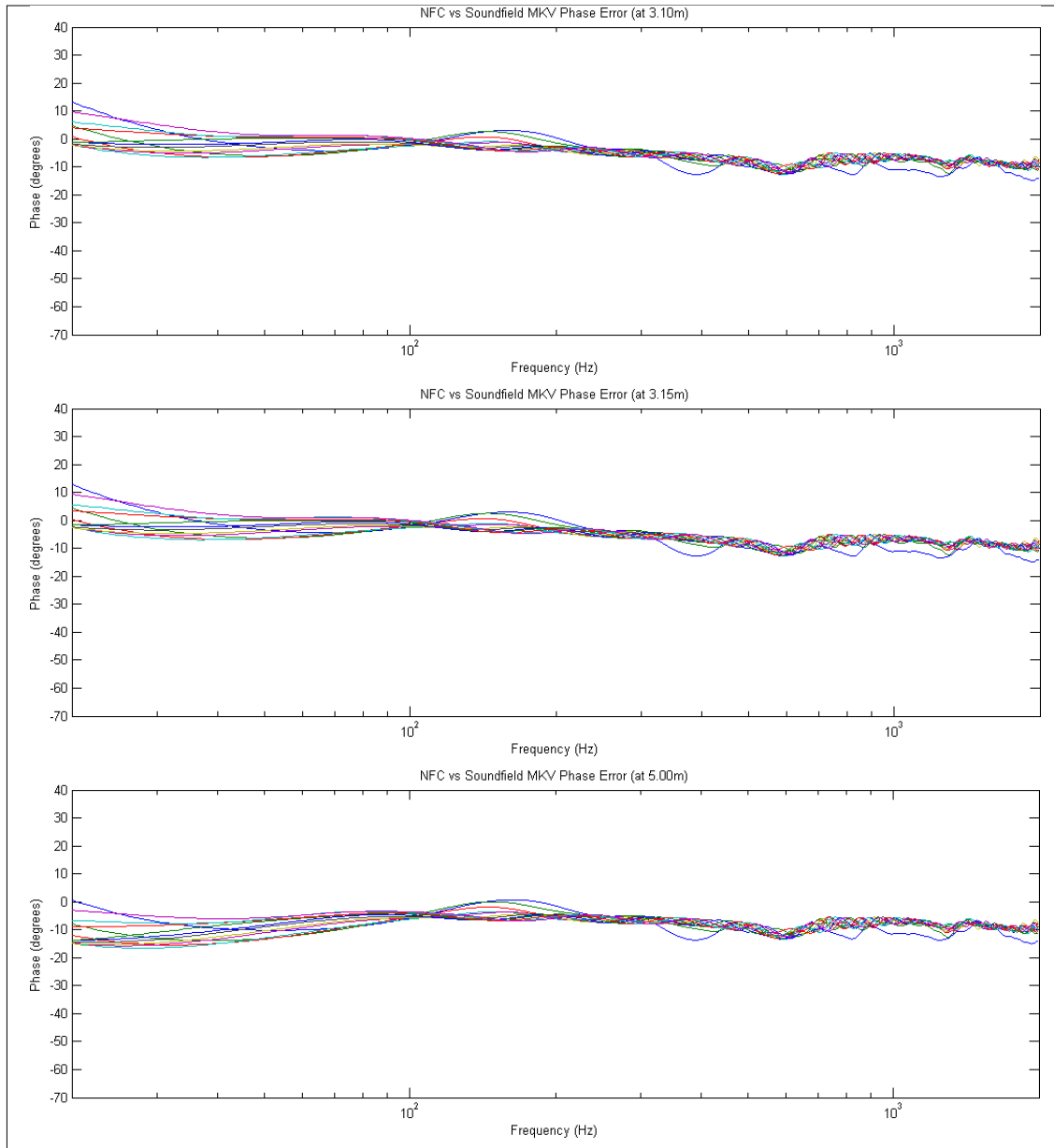


Figure 22 – MKV phase error against 3.1m NFC filter (top), 3.15m NFC filter (middle) and 5m NFC filter (bottom)