

HOW TO SHOOT YOURSELF IN THE FOOT WITH “POINT AND SHOOT” LOUDSPEAKERS.

D Gilfillan	Gilfillan Soundwork, ICE Design, Australia
M Thompson	Hanson Associates ICE Design Australia
J Love	Gilfillan Soundwork, ICE Design, Australia
G. Leembuggen	Acoustic Directions ICE Design, Australia

1 INTRODUCTION

When arraying loudspeakers to increase radiated sound levels or change directionality, the loudspeaker drive elements should be placed sufficiently close to one another to allow additive interactions of their respective outputs in the direction of a listening plane at all frequencies of interest. In contrast, current arraying practice often locates loudspeakers near each other, but not sufficiently close to allow additive interactions in the audience area over the entire frequency range of interest.

This problem usually occurs when loudspeaker boxes that are designed to be stand-alone (often described as “point and shoot” or mistakenly as “point-source” speakers) are placed near each other forming a cluster. In these situations, the physical distances between the various radiating elements result in destructive interference which produces strong comb-filtering across the intended listening area.

Despite many years of accumulated knowledge and the primacy of “line array” systems in the marketplace, this practice of clustering “point and shoot” loudspeakers in arrays is still common, and unfortunate. The rationale behind this practice appears to be that locating individual boxes together in this way will increase both the power output and the area of coverage. In fact, neither of these outcomes can be relied upon. Following a stark reminder earlier this year with a stadium sound system that we were asked to improve, the authors decided that a review of the pitfalls of this practice is warranted. That system was installed in 2023 into a new high-profile stadium.

2 ATTRACTIVENESS OF THE SPEAKER CLUSTER

Speaker clusters were born of the live music industry. Live music engineers and musicians were interested in increasing output and in most cases, single loudspeaker elements could not provide sufficient sound pressure level to give an audience the experience that the band and engineer wanted to provide. Loudspeakers which were primarily designed for stand-alone operation were routinely clustered together, to reduce the number of suspension points and attempt to create a single source of sound output in a venue, with a high enough output level. This practice continues.

The negative consequences of placing multiple sources together – namely the problems caused by acoustic interference of sources, is either not considered problematic or is simply regarded as “collateral damage” in the quest for more output. Ultimately, these clusters only increase the average sound pressure level with a 3 dB per doubling of boxes, as they are unable to provide coherent addition of sound pressure, in the audience area, which would provide 6 dB per doubling. Additionally, this practice makes the direct-field frequency responses very different at different places in the audience area.

Exacerbating the problem, manufacturers made (and continue to make) concert loudspeakers into a trapezoidal shape (rather than rectangular prism) which gave practitioners and musicians the idea that “clustering was meant to be”. Indeed, some big-name manufacturers encouraged users to form clusters of their trapezoidal boxes. Examples of speaker clustering are shown in **Figure 1**. **Figure 2** shows a typical two-element cluster, (analysed in **Figure 3**).

In venues, clusters can appear to “tick lots of boxes”. Clusters are a simple and inexpensive way to put sound into a space. In stadia and larger venues, clusters can be space efficient, enabling sight-lines from the audience to video displays to be maintained. Using the point and shoot mentality, simple modelling in EASE software (without vector addition of sources) can often indicate compliance with STI requirements.

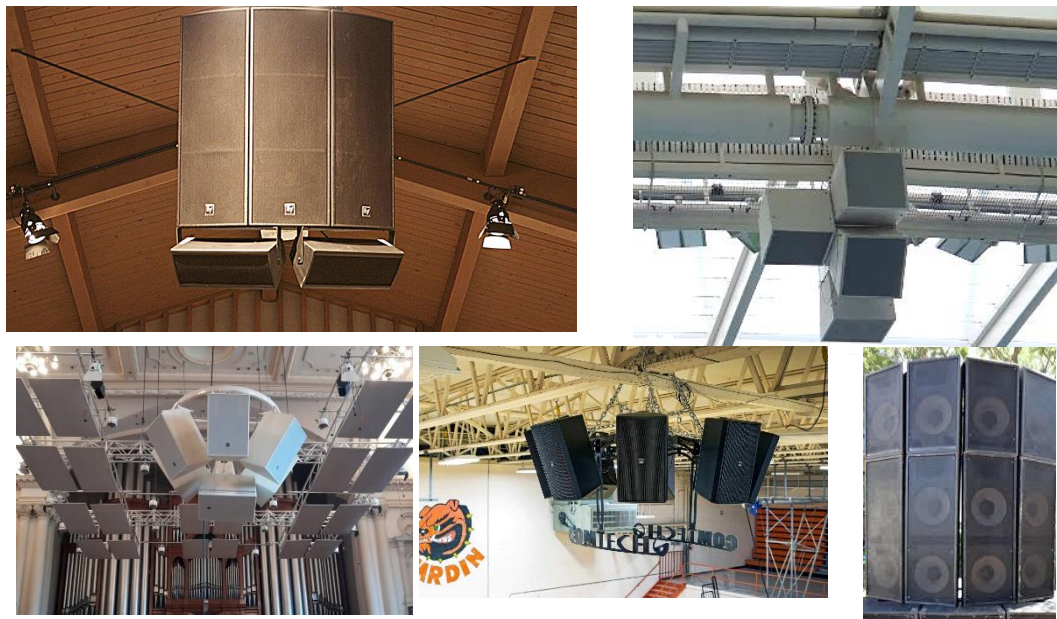


Figure 1. Typical clusters of “point and shoot” speakers.

3 ACOUSTICAL PERFORMANCE

As noted in the introduction, the major acoustical problem with clusters is that the frequency response over the audience area is often very poor. This degradation occurs at frequencies whenever the effective separation between sources whose radiation patterns overlap is greater than half the associated wavelength. **Figure 3** compares the polar responses of the two-element cluster in **Figure 2** (bottom panel in **Figure 3**) with that of a single box (top panel).

One potential benefit of some clusters is the increase in directivity at frequencies at which the effective source separation is less than half a wavelength, often below 300 Hz.



Figure 2. Typical two-element cluster

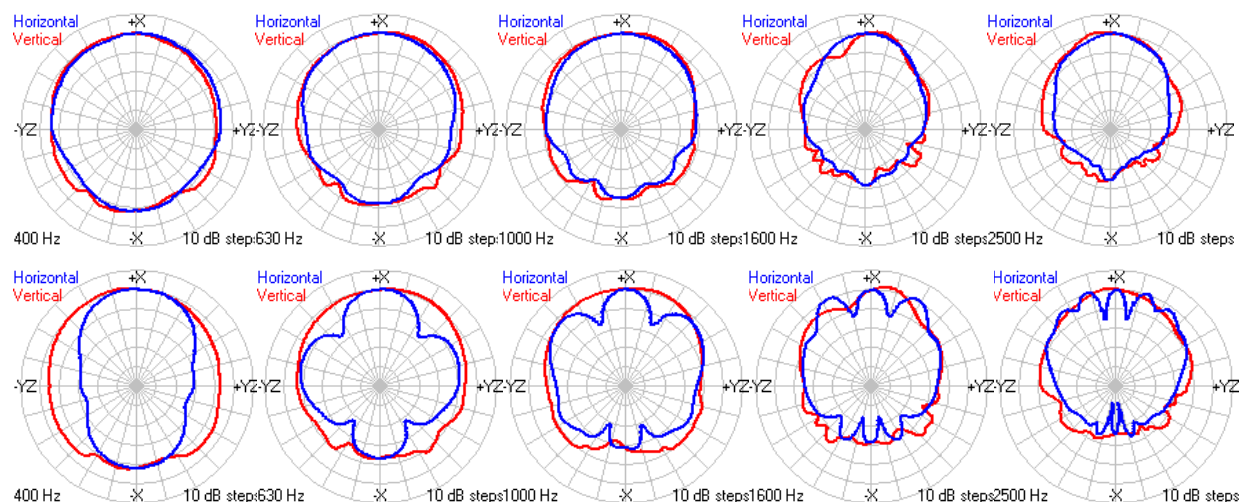


Figure 3. Comparison of the polar response of two-element cluster (bottom panel) with that of a single box (top panel).

With the two-box cluster, the narrowing of the cluster radiation pattern is clearly evident at 400 Hz, and narrower than the coverage area of a single box. Even at 2500 Hz, where the directionality of the high-frequency horn will reduce the overlap of the horns radiation, the radiation pattern in the horizontal plane shows deep nulls.

3.1 Performance of an Example Stadium Cluster

3.1.1 Overview

The majority of the clusters in the example stadium consist of three, two-way loudspeakers with passive crossovers in a vertically arrayed configuration. Each loudspeaker comprises a horn-loaded 380 mm low frequency driver and a co-axially mounted high-frequency compression driver coupled to a 38 mm throat horn. These horns have nominal radiation patterns of $90^\circ \times 50^\circ$ and $60^\circ \times 40^\circ$.

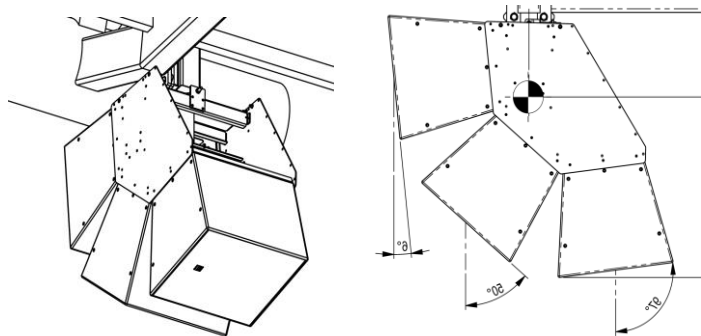
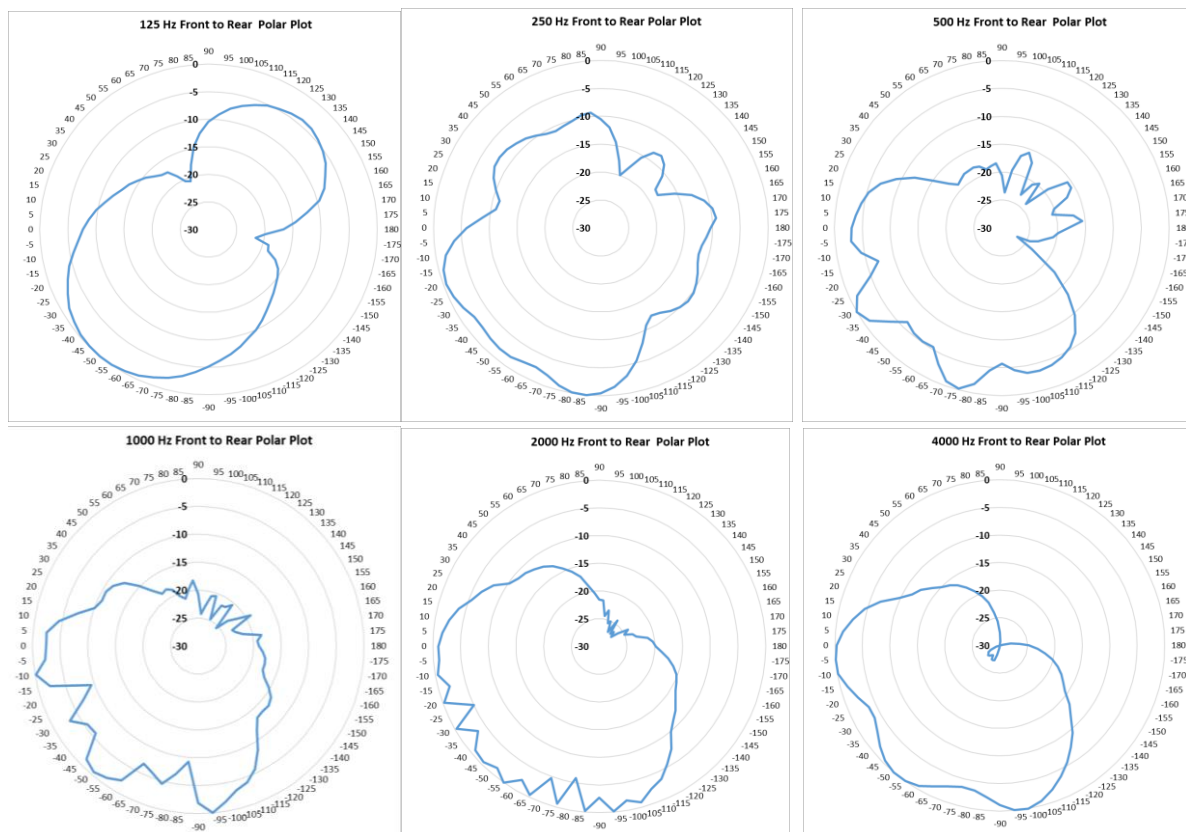


Figure 4. Isometric and side elevations of a typical cluster.

The clusters are mounted around the edge of the stadium roof and are intended to cover audience seated at seats beside the playing surface right up to the top of the top bleachers on all sides of the rectangular stadium. **Figure 4** shows an isometric view and side elevation of a typical cluster. The top two elements (those facing the top bleachers and centre of the stand) are the $90^\circ \times 50^\circ$ boxes. As the audience near the field of play is the most distant from the cluster, the downward-facing box has the $60^\circ \times 40^\circ$ radiation pattern.

3.1.2 Cluster Radiation Patterns

The radiation pattern of this cluster was computed using EASE software at the standard octave-band centre frequencies and the results averaged over a 1/3 octave band. Complex summation was selected. The polar plots of the cluster in the front-to-rear and side-to-side directions were exported and normalised to the peak level in the plot. **Figure 5** below shows polar diagrams of in the front-to-rear direction and examples of the radiation balloons looking from side on.



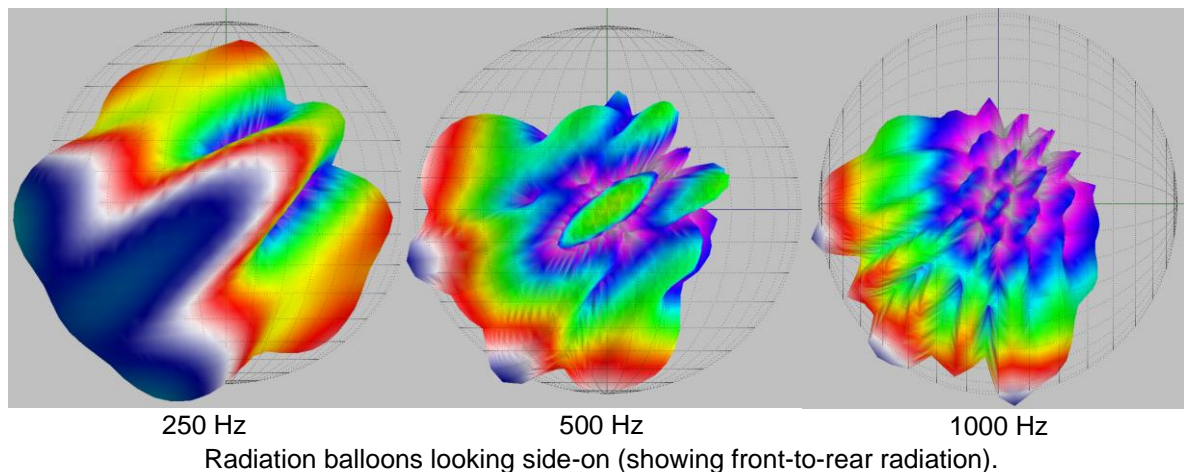


Figure 5. Polar diagrams of the front-to-rear sound radiation of the typical cluster.

The following characteristics of the cluster radiation pattern patterns are noteworthy.

- At 125 Hz, the cluster produces directivity in the front / rear direction compared with a single element. This is highly beneficial since it reduces the sound energy directed towards the grandstands on the other side of the playing field.
- A narrowing of the radiation pattern is particularly evident at 250 Hz.
- At 500 Hz, there is a strong central lobe with bulges either side that are not in the audience area and will contribute to unwanted reverberation.
- At 1 kHz and 2 kHz, there are gaps or notches of between 5 dB and 9 dB in the radiation pattern that is intended to cover the spectators.
- At 4 kHz and 8 kHz, the cluster is essentially operating as desired; point and shoot boxes with radiation patterns that butt up to each other.

Measurements of the frequency responses of the direct field of a cluster that we made in various positions confirmed that these degradations were occurring at frequencies up to 5000 Hz.

3.2 Acoustic environment

One of the main complaints about the sound system was that the speech intelligibility was unsatisfactory in the spectator areas. Measurements of the reverberation times using REW software [1] produced by the stadium surfaces and the distributed sound system revealed that the reverberation times (RTs) were extremely high, as shown in **Figure 6**. The high reverberation time contributed to poor intelligibility.

Our task was to find out whether, in the hostile acoustic environment, the clustered loudspeaker system could be adjusted to (a) meet intelligibility targets and (b) sound good for listeners.

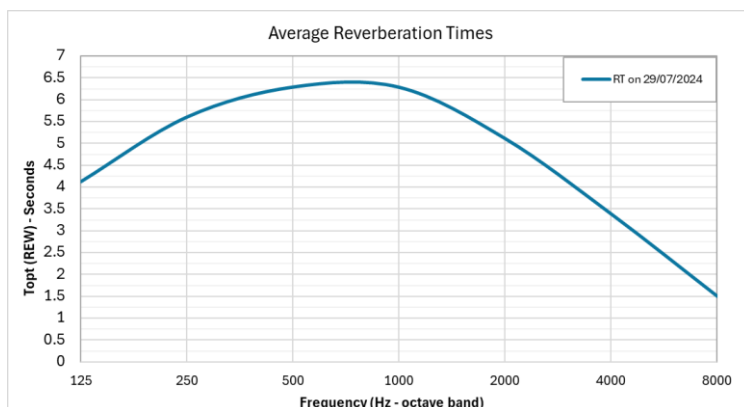


Figure 6. Reverberation times of the stadium with the distributed loudspeaker system.

4 IMPROVING THE CLUSTERS

4.1 Our Approach

The basis for our work to improve the system was that the loudspeaker and their physical mounting arrangement on the edge of the roof could not be altered. Accordingly, the only tool at our disposal was digital signal processing in the digital signal processor. Fortunately, every loudspeaker in the main bowl system was fed by its own amplifier.

With the high reverberation times and the associated late-arriving sound energy creating a strong degrading effect on intelligibility, the only course of action was to maximise the early-to-late ratio of the sound in the spectator area. To maximise this ratio requires that:

- i) the direct field in the spectator area is as high as possible for all frequencies for as many listeners as possible, and
- ii) the sound radiated by the loudspeakers in directions other than the spectator areas is as low as possible.

To achieve these goals required that the problems introduced by the clustering of the speakers were mitigated to the greatest possible extent. Given that the clusters could not be modified, the only course of action was to reduce the severity of the wide interference effects at frequencies that were degrading the frequency response of the direct field in the spectator areas.

Comb-filtering occurs when two identical signals separated by small time intervals in relation to the period of the two signals are added. The resulting phase difference per Hertz is very small and this results in wide, deep notches in the frequency response. Examples of these intervals are 1 ms at 500 Hz, 5 ms at 100 Hz and 250 μ s at 2 kHz. However when the time interval is much larger, for example 10 ms, the phase shift per Hertz is much greater and this increases the number of notches in the frequency response, thereby reducing their bandwidth. The number of 360° rotations in a phase vs frequency plot is a direct indicator of the density of the comb filters in the response.

Given the sensitivity of human hearing to a loudspeaker's direct field response [2, 3], reducing the bandwidth of notches resulting from comb filtering can produce major improvements in perceived sound quality and speech intelligibility in reverberant environments, so long as the two signals are not perceived as separated events.

Since the loudspeakers are individually amplified, increasing the time interval between arrival of the three sources within a cluster can be readily achieved by adding delay or all-pass filters to the two outer speakers in a three element cluster. However, adding delay at low frequencies to the outer speakers to increase the density of the phase cancellations would degrade the beneficial directionality that the cluster intrinsically provides.

Essentially, we needed a signal processing architecture that would:

- i) allow loudspeaker signals at low frequencies to have zero inter-box phase shift (identical processing to maintain improved directionality) and then
- ii) apply a rapid transition above a chosen low-mid frequency to a considerable number of phase rotations between the inner and the two outer loudspeaker signals (to reduce interactions that are harmful to the direct field response at the audience). The low-mid frequency was chosen with consideration of the directivity control of individual elements and the distance between elements.

To avoid unnecessarily extending the duration of early-arriving sound at upper-mid and high frequencies (to prevent further degradation of intelligibility), we elected to limit the amount of additional delay above 1 kHz to 3 ms and use allpass filters to provide more phase-rotations below 1 kHz. We explored the use of temporally diffuse impulses [4] [5] [6] for this decorrelation but concluded we need more delay at low frequencies than they could readily provide, and site time was very limited.

Figure 7 compares the phase response of the seven-stage allpass filter (APF) bank with a three-second signal delay. The APFs are second order with varying resonant frequencies and Q factors. **Figure 8** shows the combined phase response of the APFs and 3 ms delay, with **Figure 9** showing the effective signal delay of the APF bank and 3 ms delay.

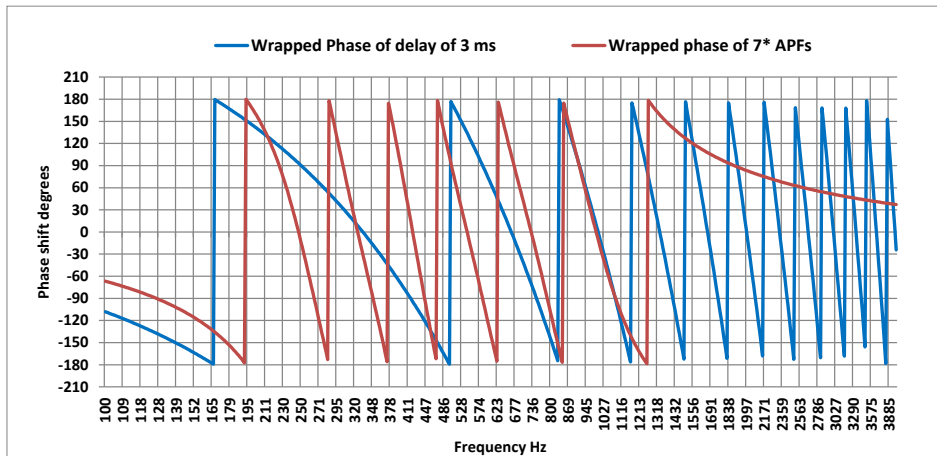


Figure 7. Comparison of the phase responses of the ten-section allpass filter (APF) bank and a three-second signal delay.

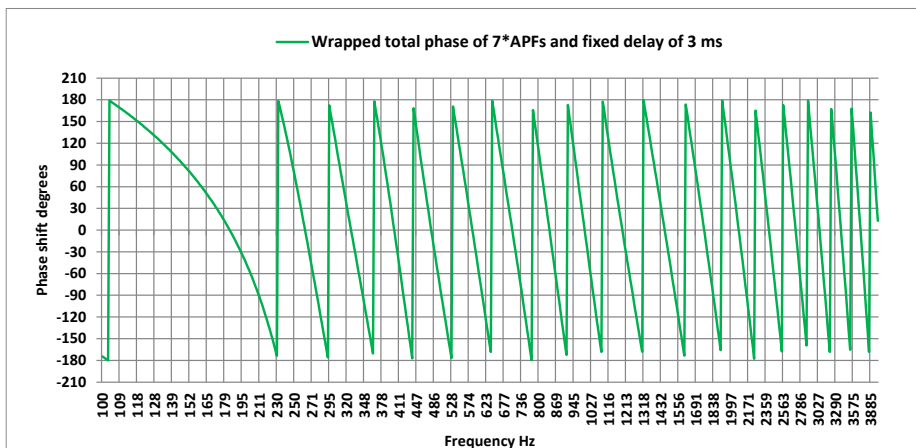


Figure 8. Phase response of the combination of the allpass filters and delay in **Figure 7**.

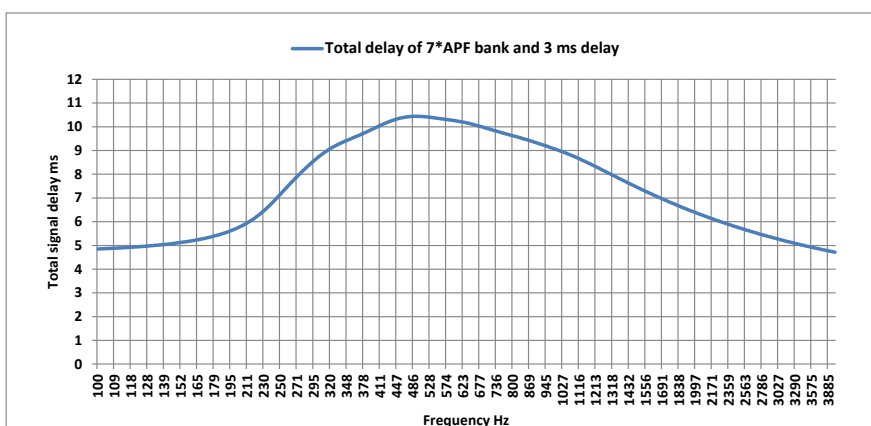


Figure 9. Total delay of the APF bank and 3 ms delay.

4.2 Signal Processing

The signal processing architecture that we employed is shown in **Figure 10**. Features of the signal processing system are:

- At frequencies between 90 Hz (crossover to subwoofer) and 300 Hz, the three speakers are fed with the same signal, except that the middle speaker has a gain of 10 dB to reduce the narrowing at low frequencies of the polar pattern shown in **Figure 5**.
- At frequencies above 300 Hz, all three speakers have equal gain, but the outer two speakers are effectively decorrelated.
- A high shelf with a 300 Hz turnover frequency restores the level of the inner speaker to 0 dB at mid and high frequencies.
- So that the phase response of the shelf is linear, it is implemented as a symmetric FIR filter. To The delay of that filter introduced into the inner box signal is exactly matched by the 5.2 ms delay feeding the outer boxes.
- A disadvantage of this scheme is that the inner box is subject to 10 dB more level from the crossover to the subwoofer at 90 Hz to 300 Hz which limits the overall output.
- The processing blocks are described below.

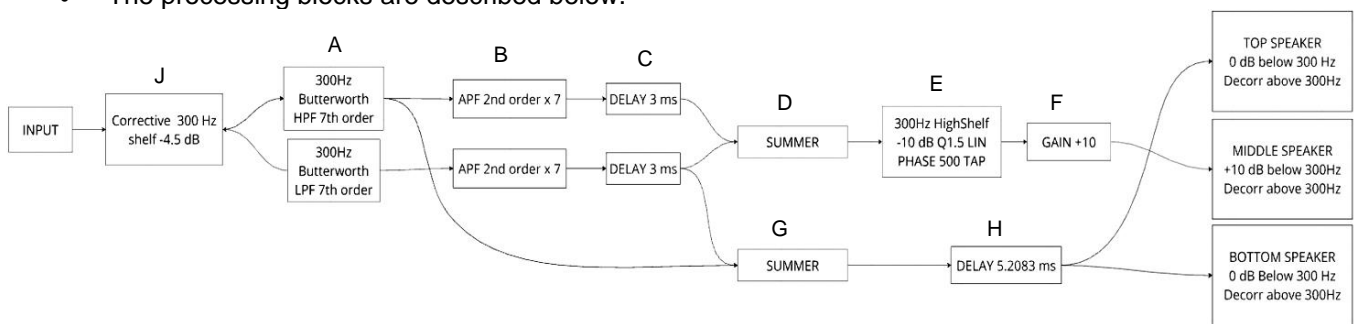


Figure 10. Schematic diagram of the signal processing developed for three-element cluster to reduce the severity of comb-filtering. See **Table 1** for details of labelled blocks.

Table 1. Description of specific blocks in the signal processing,

Block	Function
A	7 th order Butterworth crossover to create low frequency and mid/high frequency chains. 7 th order was selected for its high roll-off rate and the fact that odd-order Butterworth crossover filters sum to unity.
B	Seven stages of second-order all pass filters with turnover frequency and Q selected for reasonable consistency of phase rotations when viewed on a log-frequency scale.
C	The combination of the allpass filter bank and the 3 ms delay provides the decorrelation between the inner and outer boxes at frequencies above 300 Hz
D	The summer recombines the low and mid/high chains with the same phase shift for the inner box, so that below 300 Hz, all boxes have the same phase response.
E	Shelving filter with a gain at high frequencies of -10 dB relative to low frequencies. This is executed as a linear-phase FIR filter so that the filter's delay is constant with frequency and therefore can be matched with a signal delay.
F	Provides 10 dB of gain at low frequencies to the middle speaker and restores 0 dB gain at high frequencies.
G	Sums the low frequency chain with the decorrelated mid/high frequency chain for the outer speakers. (This is a power sum of the low and high signals.)
H	Fixed delay that exactly increases the delay of the outer speakers to exactly match the delay of the high-shelf filter.
J	A low-shelf attenuating the level below 300 Hz by 4.5 dB to make the overall level of the cluster at low frequencies equal to the coherent sum of three boxes at equal drive levels.

Figure 11 shows the calculated differences between the amplitude and phase responses at the outputs of the summing blocks D and G. The phase difference between the outer two boxes and the

inner box is as desired, however, the difference in the amplitude response is not ideal. Calculation of the response differences with a 9th order Butterworth crossover showed only a minor reduction in the amplitude errors.

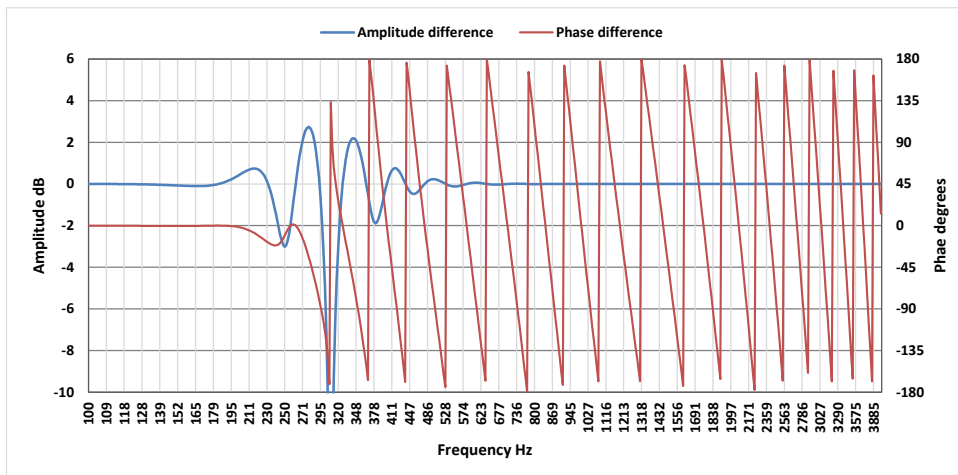


Figure 11. Differences between the amplitude and phase responses at outputs of summing blocks D and G.

5 CONCLUSION

We were unable to achieve a compliant STI result using this method. In fact, modelling indicated that no system could deliver the specified STI with the stadium's geometry, materials and resulting acoustic environment.

However, intelligibility was measurably improved using this decorrelation method compared with the method which applied only cluster equalisation. Also, the listening experience for the audience using this method was substantially improved, both with speech and music.

Loudspeaker clusters such as the ones shown in this paper pose insoluble problems to the electroacoustic engineer and compromised listening experiences for audiences. Systems exist which reduce many of these problems.

The concept of clustering point-and-shoot speakers should be retired. Our advice is to use the advancements that the loudspeaker R&D companies have made in arraying loudspeakers using close driver spacing and appropriate processing, rather than simply clustering speaker boxes.

6 REFERENCES

- [1] "REW Software," [Online]. Available: <https://www.roomeqwizard.com/>.
- [2] D. Griesinger., " The relationship between audience engagement and the ability to perceive pitch, timbre, azimuth and envelopment of multiple sources.," *Proc IOA*, Vols. 33, , no. Pt 2, 2011.
- [3] D. Griesinger, "The Importance of the Direct to Reverberant Ratio in the Perception of Distance, Localization, Clarity, and Envelopment," in *AES Convention: 126*, 2009.
- [4] J. B. Moore, A. J. Hill, "Optimization of Temporally Diffuse Impulses for Decorrelation of Multiple Discrete Loudspeakers," in *142nd Convention of the Audio Engineering Society*, , May 2017.
- [5] J.B. Moore, A.J. Hill. , "Dynamic diffuse signal processing for sound reinforcement and reproduction," *Journal of the Audio Engineering Society*, , vol. 66, no. 11, December, 2018.
- [6] J. Moore, "Dynamic diffuse signal processing in sound reinforcement and reproduction.," Unpublished PhD Thesis, University of Derby, Jan 2019.