Audio Engineering Society

Conference Paper

Presented at the International Acoustics & Sound Reinforcement Conference

2024 January 23-26, Le Mans, France

Signal decorrelation for sound reinforcement system crossovers

Adam Hill

Electro-Acoustics Research Lab, College of Science and Engineering, University of Derby, UK

Correspondence should be addressed to Adam Hill (`a.hill@derby.ac.uk`)

ABSTRACT

While sound reinforcement technology has progressed significantly in recent decades, aspects of system deployment remain largely unchanged, principally the use of stereo (or multiple mono) left/right configurations and crossover networks. As such, the issue of coherent interference between overlapping spatial and spectral coverage remains a challenge to system engineers. This paper focuses on the application of a perceptually transparent method of decorrelation, known as diffuse signal processing (DiSP), to minimize coherent interference within key elements of sound systems. Experiments were conducted with scale model loudspeakers in a hemi-anechoic chamber, mounting the systems onto an automated turntable to inspect the effectiveness of decorrelation over a wide polar range. Results indicate that the application of decorrelation has the potential to significantly reduce spatial variance across an audience area, although further work is necessary to optimize the decorrelation filters to improve performance consistency.

1 Introduction

As emphasized in McCarthy's seminal textbook on sound system design and optimization [1], crossovers are often the most challenging aspect of sound reinforcement. This is because when two or more system elements cover the same spatial or spectral region there is a strong chance of coherent interference, causing an inconsistent response. This results in (1) a position-dependent listening experience and (2) inaccurate reinforcement of the intended timbre of the program material.

While sound reinforcement technology has seen significant progress over the past few decades, most systems still employ a conventional stereo (or multiple mono) left/right configuration with spectral crossovers used to separate frequency bands between different system drive elements. Even with the recent movement towards object-oriented sound systems, which partially overcome such issues with horizontally distributed loudspeakers, such technology is expensive, both in financial and practical terms, making this unlikely to become the norm in the short- to medium-term. As such, it is easier to implement software-based, rather than hardware-based, solutions to the crossover issue.

Although several previously published papers detail decorrelation algorithms for a variety of use-cases [2-11], there has yet to be a focused practical study on the use of decorrelation for live sound system crossovers, which is therefore addressed in this paper.

The paper begins with a brief overview of the decorrelation algorithm used in this research in Section 2, followed by a description of the experimental methods in Section 3. Section 4 provides a critical analysis of the results, and the paper is then concluded in Section 5.

2 Diffuse signal processing

Diffuse signal processing (DiSP), an audio signal decorrelation method first developed to minimize coherent interference within arrays of Distributed Mode Loudspeakers (DML) [11], was previously identified as the most promising technique for use with sound reinforcement due to its inherent flexibility and perceptual transparency, and refined to meet the challenges of such applications [13].

DiSP uses temporally diffuse impulses (TDIs) to achieve decorrelation (Figures 2.1 and 2.2), where each sound producing element in a system has its own unique TDI which contains an initial impulse followed by a rapidly decaying random phase noise tail. TDIs are generated in the frequency domain to allow for precise control over decay to avoid perceptual artifacts. This includes the ability to restrict decorrelation to a specific frequency range, for example between 200 – 2000 Hz as in Figure 2.2.

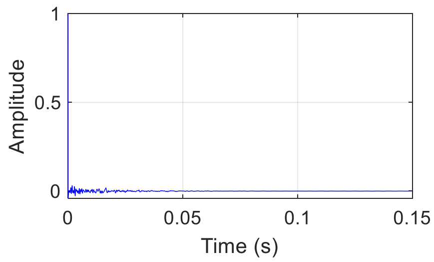

Figure 2.1 Example temporally diffuse impulse (TDI) used within diffuse signal processing (DiSP)

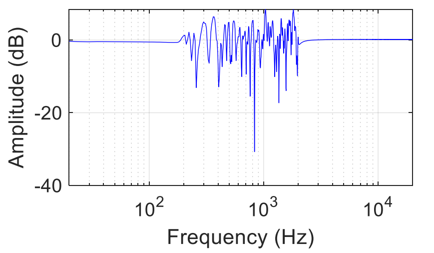

Figure 2.2 Example unequalized magnitude response of a temporally diffuse impulse (TDI) used within diffuse signal processing (DiSP)

DiSP exists in two forms: static DiSP and dynamic DiSP. Static DiSP requires a single TDI per degree of freedom within a sound system which is convolved with the input audio signal to apply the decorrelation. While this is relatively straightforward to implement, the decorrelation only targets the direct sound. Any early reflections within a closed acoustic space will be at least partially correlated with the direct sound, resulting in coherent interference.

Dynamic DiSP [13] addresses this issue by replacing a single TDI with a TDI library. This library is pre-generated and cycled through over time, using interpolation between successive TDIs to avoid perceptual artifacts. If the TDIs are switched at a sufficient rate, the direct sound and early reflections will be decorrelated from each other, therefore minimizing coherent interference.

3 Experimental methods

This research focuses on the reduction of coherent interference within spatial and spectral crossovers and not the reduction of coherent interference due to early reflections. A 5.7 m x 5.3 m x 2.6 m hemi anechoic chamber, certified to provide free-field conditions over a reflecting plane with a cutoff frequency of 100 Hz, in accordance with ISO 26101, was used in this work. As part of this research focuses on subwoofer systems, this test environment would be inappropriate for full-scale subwoofers. Instead, 1:10 scale subwoofers were used, designed to function as scale models of a commercially available medium-format subwoofer (1160 mm x 580 mm x 920 mm with an approximate passband of 30-90 Hz). The passband of the scale model loudspeakers (without filtering) was 300-4000 Hz (corresponding to 30-400 Hz, full scale).

At the core of the experimental setup was an Audiomatica Clio 10 system [14], with a calibrated MIC-01 measurement microphone. The output from Clio was fed into Reaper [15] where two VSTs were used to handle the necessary signal processing. First, a 4th order Linkwitz-Riley crossover was applied using RS-MET's CrossOver plugin [16], with a crossover frequency of 850 Hz (corresponding to 85 Hz, full scale) (Figure 3.1). Secondly, ReaVerb by Cockos (the makers of Reaper) [15] was used to convolve the TDIs with the incoming audio signals.

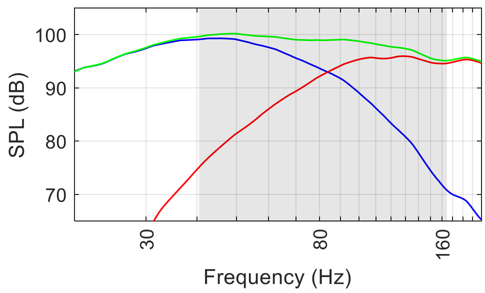

Figure 3.1 Full scale acoustic measurements of loudspeakers (no spacing) driven by the outputs of a 4th order Linkwitz-Riley crossover. The shaded area represents the expected crossover region.

After processing in Reaper, the resulting audio signals were fed into a Quad 306 stereo power amplifier. The amplifier signals were then sent into the hemi-anechoic chamber and connected to the scale model loudspeakers. The following subsections detail the specific methods used for the two experiments.

3.1 Spatial crossover

Two scale model loudspeakers were used as the sound system under inspection to examine the effectiveness of DiSP in resolving spatial crossover coherent interference. The two loudspeakers were placed on a 3.0 m long strip of wood mounted to an Outline ET250 turntable.

Five configurations were measured, where the only variable was the spacing between the loudspeakers. Spacings of 0, 0.2, 0.6, 1.0 and 1.4 m were tested, corresponding to full scale spacings of 0, 2, 6, 10, and 14 m, respectively. The turntable was controlled via Clio, with 5° measurements taken over 180°. The measurement microphone was placed at 2.0 m from the center of the turntable, at a height of 0.17 m (representing the average standing ear height of an adult of 1.7 m) plus 0.13 m, which was the height of the loudspeakers when mounted on the turntable. This meant that the measurement distance between the two loudspeakers would be inconsistent as the turntable rotated, but as this research focused on spectral response, and not absolute level, this was acceptable.

The experimental setup was replicated in a simple model within Matlab. This allowed for optimization of the pair of TDIs. First, when each TDI was generated its magnitude response was checked for significant deviations from flat (0 dB) at the upper and lower frequencies of interest (200 Hz and 2 kHz in this case). If an offset of more than 0.1 dB was observed, then a new TDI would be synthesized.

With the pair of TDIs synthesized and individually checked, the experiment was simulated (without the ground reflection) with the average spatial variance across all measurements calculated. If this variance was calculated to be greater than 1 dB, new TDIs would be generated, and the process repeated until an acceptably performing pair of TDIs was identified. The 1 dB threshold is based on a 1.38 dB audible threshold for low-frequency spatial variance between 20 – 250 Hz as identified in [13].

With the pair of TDIs in place, the physical measurements could be taken using Clio and a swept sine measurement at a sampling rate of 192 kHz. The two scale model loudspeakers' drive signals were independently sent through the low frequency channel of the crossover in Reaper and then convolved with the TDIs. Data was saved directly to .txt files for post-processing.

Data analysis was performed in Matlab, where the only further signal processing applied was 1/9 octave band smoothing to align with the human hearing resolution in the subwoofer range [13].

3.2 Spectral crossover

As with the spatial crossover experiment, the spectral crossover experiment used two scale model loudspeakers. In this case, one loudspeaker would act as a mid- and high-frequency element and the other would act as a subwoofer. The subwoofer would always remain on the floor of the chamber, while the other loudspeaker would be suspended above the subwoofer by 0, 0.4, 0.8, and 1.2 m (corresponding to 0, 4, 8 and 12 m full scale).

On axis measurements were taken across an audience depth spanning 0.2 m to 3.0 m in steps of 0.2 m (corresponding to 2 to 30 m full scale). This would represent the on-axis coverage of the front half of a typical festival audience, where beyond this point the relative difference in propagation distance between the two loudspeakers is less likely to cause significant coherent interference.

As with the spatial crossover experiment, the configurations were each modelled in Matlab to optimize the pair of TDIs (one pair for each of the four loudspeaker spacings under examination). The optimization parameters were the same as detailed in Section 3.1.

Physical measurements were taken and stored using the same settings as detailed in Section 3.1. As before, the data was read into Matlab with a 1/9 octave band filter applied to approximate the resolution of the human hearing system in the subwoofer range.

4 Results and analysis

Considering the hundreds of measurements captured between the two experiments detailed in Section 3, the data was compressed into a smaller collection of spatial distribution plots (Figures 4.1 and 4.2). These plots are presented in the following two subsections along with a critical analysis of the results.

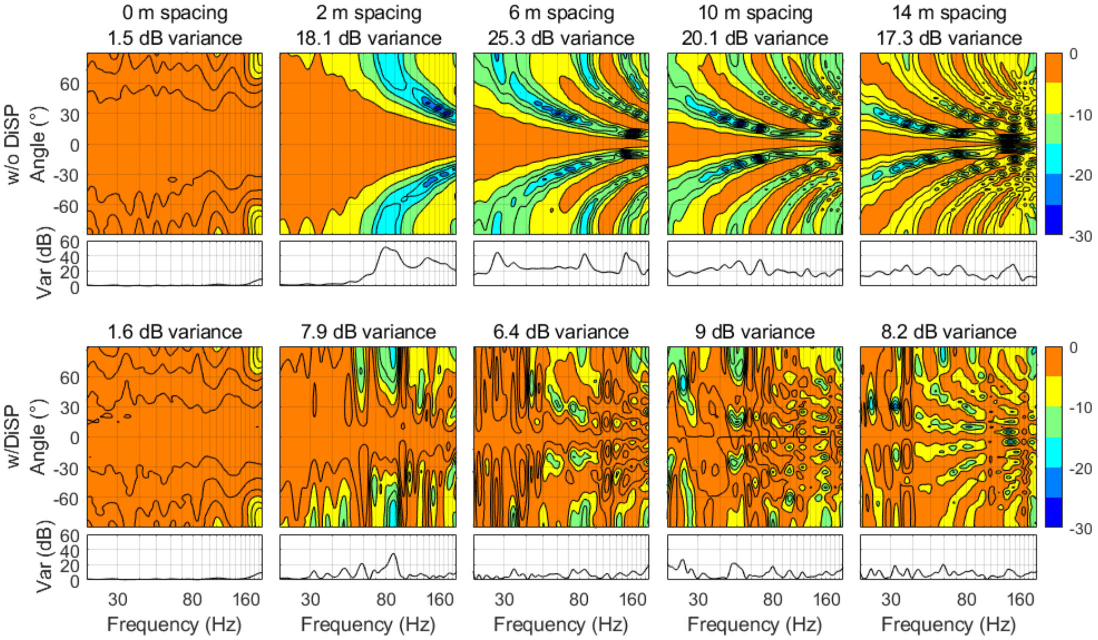

Figure 4.1 Experimental results from the spatial crossover tests. Each column represents the loudspeaker horizontal spacing (full scale), where the top and bottom rows show results without and with DiSP, respectively. The contour plots present normalized sound pressure levels (dB, with the on-axis 0° used as a reference for each configuration) and the accompanying line graphs show spatial variance over frequency (full scale). Variance in the column titles is the average spatial variance from 20-200 Hz (full scale).

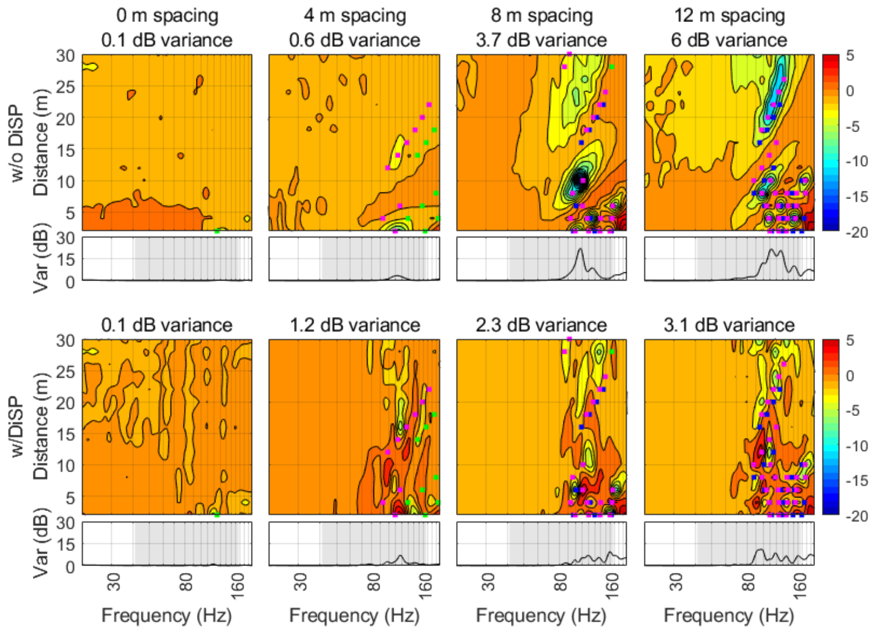

Figure 4.2 Experimental results from the spectral crossover tests. Each column represents the loudspeaker vertical spacing (full scale), where the top and bottom rows show results without and with DiSP, respectively. The contour plots present sound pressure levels (dB, with a single loudspeaker on-axis measurement used as a reference for each configuration) and the accompanying line graphs show spatial variance over frequency (full scale). Variance in the column titles is the average spatial variance from 20-200 Hz (full scale). The shaded region in the variance plots represents the crossover region. Blue, magenta, and green dots represent the expected comb filtered frequencies between the direct sounds, subwoofer and its floor reflection, and high-frequency loudspeaker and its floor reflection, respectively.

4.1 Spatial crossover

The measurements without DiSP (Figure 4.1) show comb filtering due to left/right subwoofer configurations. As subwoofer spacing increases, so does spatial variance, as the loudspeakers' mutual coupling frequency decreases. Practically, this results in highly position-dependent listening experiences for audience members, which goes against the goal of delivering an equally excellent listening experience across a wide audience area [17].

After the application of DiSP, a very different set of coverage patterns emerges. Aside from the 0 m spacing results, where the subwoofers are completely coupled in the operating range (20-200 Hz), variance remains non-zero, but has reduced by over 50% in each configuration. Remembering that DiSP doesn't eliminate the possibility of comb filtering, but instead limits its occurrence to narrow spectral regions and randomizes where it occurs in the audience, the results are as expected. There is still indication of irregularities in the response across the audience, but these follow no set pattern. The deep notches (nearing -30 dB) in the unprocessed results are no longer present, with the worst notch approximately -20 dB with DiSP.

4.2 Spectral crossover

While the spatial crossover experiment focused on the minimization of coherent interference due to uncoupled loudspeakers of the same variety (outputting identical signals), the spectral crossover experiment focused on the minimization of coherent interference due to uncoupled loudspeakers of different varieties – one reproducing low-frequency content and the other reproducing high-frequency content. In this case, DiSP was applied within the crossover network with the hope of providing a more consistent spectral response in the crossover region.

The measurement results without DiSP (Figure 4.2) highlight an increase in spatial variance across the depth of the audience (on-axis) as the spacing between the loudspeaker increases (in real-world terms, as the height of the line array is increased while the subwoofer system remains on the ground). Inspecting the colored dots overlaid on the plots, the worst comb filtering can be observed where theory states that comb filtering should be at its worst, although not all predicted comb filtering frequencies/locations appear to have manifested themselves in these experimental results (which could be due to the spatial and spectral resolution of the analysis, including the spectral smoothing applied).

In all but one of the non-zero spaced configurations, the introduction of DiSP has resulted in significant reductions in spatial variance in the crossover region. As with the spatial crossover results, the peaks/dips in the coverage are still visible, but their location is more random and less pronounced.

A critical observation can be drawn from the results of the 4 m spacing configuration. In this case, spatial variance has increased after the introduction of DiSP. There are two possible reasons for this. First, as can be seen by the predicted comb filtering frequency/location dots, the expected issues only take place in the upper range of the crossover region. As DiSP was applied to the full crossover region, in essence scrambling the response, the lower frequency range has some randomized perturbations which otherwise wouldn't exist. Further work should explore the practicality of basing the DiSP spectral range on the predicted comb filtering frequencies for a given system and audience configuration.

Secondly, the greater spatial variance with DiSP in this instance could also be due to a poor pairing of TDIs. As time didn't permit the test to be re-run with a new pair of TDIs, this should be the focus of further research. It is likely that the TDI optimization routine can be improved to avoid false-positive pairings.

5 Conclusions

Coherent interference is most problematic for live event sound reinforcement when using (1) uncoupled loudspeakers to reproduce the same signal, (2) uncoupled loudspeakers that are partially reproducing the same spectral range, and (3) performance spaces that produce strong early reflections. This paper details research aimed at addressing problems (1) and (2), where a decorrelation algorithm, static diffuse signal processing (DiSP), has been applied to minimize the effects of coherent interference to ensure a consistent listening experience across a wide audience area.

The scale-model experimental results show that DiSP is indeed effective in reducing the prevalence of coherent interference in commonly used sound system configurations at live events, although not all the test cases support this. This inconsistency in results points to further research to (1) limit the spectral range of DiSP to only where it's predicted to be needed and (2) provide a more robust TDI optimization routine. Previous research in this area [13] has suggested that research is required to ensure that DiSP is perceptually transparent. Such early tests indicated that this is feasible, but at the expense of the effectiveness of the decorrelation.

There have also been recent unpublished efforts by the author to minimize the latency of DiSP using partitioned convolution. While early results have been promising (up to a ten-fold decrease in latency), further work is required to ensure latency is suitable for real-time use at live events.

While there is more work to do before DiSP is ready to be implemented at a real-world live event, the results presented in this paper provide a good indication that this procedure should be strongly considered as an effective and easy to implement software-based solution to position-dependent listening experiences at live events.

References

[1] McCarthy, B. Sound systems: Design and optimization: Modern techniques and tools for sound system design and alignment. Focal Press, 2016.

[2] M. Schroeder, "An Artificial Stereophonic Effect Obtained from a Single Audio Signal," *Journal of the Audio Engineering Society*, vol 6, pp. 74-79 (April, 1958).

[3] R. Orban, " A Rational Technique for Synthesizing Pseudo-Stereo from Monophonic Sources," *Journal of the Audio Engineering Society*, vol 18, pp. 157-164 (April, 1970).

[4] M. Wilde, W. Martens and G. Kendall, "*Method and apparatus for creating decorrelated audio output signals and audio recordings made thereby*," US5235656, (August 10th, 1993).

[5] M. Ali, "Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 6, (May, 1998)

[6] G. Kendall, "The Effects of Multi-Channel Signal Decorrelation in Audio Reproduction", *ICMC Proceedings*, pp. 319-326, (1994)

[7] M. Bouéri, C. Kyirakakis, " Audio Signal Decorrelation Based on a Critical Band Approach," presented at the *117th convention* of the Audio Engineering Society (2004 Oct), convention paper 6291.

[8] E. Kermit-Cranfield, J. Abel, "Signal decorrelation using perceptually informed all-pass filters", *Proceedings of the 19th International Conference on Digital Audio Effects*, pp. 225-231, (September, 2016)

[9] R. Penniman, "A General Purpose Decorrelation Algorithm with Transient Fidelity," presented at the *137th Convention* of the Audio Engineering Society, (October 2014), convention paper 9170

[10] B. D. Molle, J. Pinkl and M. Blewett, "Decorrelated Audio Imaging in Radial Virtual Reality Environments," presented at the *139th Convention* of the Audio Engineering Society, (October 2015), e-brief 208.

[11] M.O.J. Hawksford, N. Harris, 'Diffuse signal processing and acoustic source characterization for applications in synthetic loudspeaker arrays,' *112th Convention of the Audio Engineering Society*, convention paper 5612. (2002).

[12] Moore, J.B.; A.J. Hill. "Dynamic diffuse signal processing for low-frequency spatial variance minimization across wide audience areas." 143rd Convention of the Audio Engineering Society, New York, USA. October 2017.

[13] Moore, J.B. Dynamic diffuse signal processing in sound reinforcement and reproduction. PhD Thesis, University of Derby. April 2019.

[14] Audiomatica SRL. "CLIO 10 FW." http://www.audiomatica.com/wp/?page_id=51

[15] Cockos, Inc. Reaper Digital Audio Workstation. https://www.reaper.fm/

[16] Schmidt, R. "RS-MET: CrossOver." http://www.rs-met.com/

[17] Hill, A.J. (chair and editor), "Understanding and managing sound exposure and noise pollution at outdoor events.". AES Technical Document AESTD1007.1.20-05, 2020.