ENHANCED WIDE-AREA LOW-FREQUENCY SOUND REPRODUCTION IN CINEMAS: EFFECTIVE AND PRACTICAL ALTERNATIVES TO CURRENT SUB-OPTIMAL CALIBRATION STRATEGIES

ADAM J. HILL[1], MALCOLM O. J. HAWKSFORD[2], AND PHILIP NEWELL[3]

[1] *Department of Engineering, University of Derby, UK*

a.hill@derby.ac.uk

[2] *School of Computer Science & Electronic Engineering, University of Essex, UK*

mjh@essex.ac.uk

[3] *Acoustics consultant, Moaña, Spain*

philiprnewell@gmail.com

This paper explores strategies for achieving accurate wide-area low-frequency sound reproduction in cinemas. Current standards for B-Chain calibration call for single channel low-frequency equalization aided by either single-point or spatially-averaged response measurements, an approach only applicable to a reasonably spatially invariant low-frequency response. A holistic approach to low-frequency coverage optimization is presented exploiting subwoofer arrays, their positioning and multi-point signal processing. Acoustic-field examples are presented using finite-difference time-domain (FDTD) modeling software that expose a potential for superior wide-area signal reconstruction over that achieved using the current standards and recommendations.

INTRODUCTION

Recent research into the causes of variability in sound reproduction across cinemas reveals that there is an inherent lack of understanding regarding low-frequency sound reproduction [1]. Current standards and recommendations [2,3] suggest using third-octave graphic or parametric equalization to smooth the low-frequency response in cinemas. Often these techniques are based on spatially-averaged response measurements across a seating area [2]. This has been proven to give no benefit regarding uniformity or overall "flatness" in the low-frequency band [1,4,5]. The effect redistributes the problem rather than solves it.

This research aims to resolve misunderstandings regarding low-frequency acoustics and sound reproduction so that a well-informed B-Chain calibration procedure (a B-Chain consists of everything after the volume fader in a system including the power amplifiers, loudspeakers, screen and any acoustical treatment) can be developed and standardized to allow for consistent low-frequency responses in cinemas and dubbing theaters (across all venues and seats within).

A detailed problem definition is laid out in Section 1 covering issues pertaining to acoustics (room-modes, comb-filtering, spatial variance), sound reproduction (individual channel frequency responses, interference, available degrees of freedom, existing standards) and psychoacoustical considerations.

With the problem defined, a detailed examination of current calibration strategies is presented in Section 2. This includes inspecting system responses using various common loudspeaker configurations as well as single point and spatially-averaged response equalization techniques. These approaches are shown to do little (if anything) in reducing seat-to-seat frequency response variations (spatial variance) in cinemas.

A number of improved calibration strategies are demonstrated in Section 3 including optimization algorithms and diffuse signal processing. These techniques provide significant spatial variance reduction across entire seating areas with little chance of human error corrupting the calibration process.

A set of effective and practical recommendations is put forward in Section 4, providing a clear approach regarding the development of a robust strategy to low-frequency calibration in cinemas.

The hope is for these recommendations to be taken into consideration when revising current standards [2] for B-Chain calibration in cinemas and dubbing theaters.

1 PROBLEM DEFINITION

In order to develop an effective and practical methodology for addressing the inherent issues to low-frequency sound reproduction, a detailed analysis is required. Two scenarios are considered in this research: a commercial cinema and a dubbing theater.

1.1 Physical properties

The physical properties of the two spaces were chosen by taking the average dimensions and reverberation times of commercial cinemas and dubbing theaters studied in a recently published SMPTE report [5]. Reverberation times were used to calculate average absorption coefficients using Eq. 1.1 [6].

$$RT_{60}(f) = \frac{0.161V}{S\alpha(f)} \tag{1.1}$$

where $RT_{60}(f)$ is the reverberation time (s) at frequency, f (Hz), which is calculated using the room volume (V, in m^3), surface area (S, in m^2) and average absorption coefficient (α). The chosen properties for the commercial cinema and dubbing theater under inspection are given in Table 1.1.

Property	Commercial cinema	Dubbing theater
Length, L_x	27.0 m	17.0 m
Width, L_y	20.0 m	12.0 m
Height, L_z	10.0 m	6.0 m
Absorption, α	0.287	0.592

Table 1.1 Physical properties for the two closed acoustic spaces under inspection

1.2 Loudspeaker properties

The frequency responses and placement of the loudspeakers in the B-Chain under inspection were designed to align with recommendations [2,3,7,8] and in-situ measurements at various venues [5]. This research tests a B-Chain exhibiting tight properties (in line with recommendations) and relaxed properties (in line with measurement data) to demonstrate differences between optimal and sub-optimal systems. The selected crossover settings are given in Table 1.2.

Channel	Tight	Semi-relaxed	Relaxed
L, C, R	50 Hz (4)	40 Hz (4)	40 Hz (2)
SL, SR	125 Hz (4)	60 Hz (4)	50 Hz (2)
LFE	125 Hz (6)	125 Hz (2)	125 Hz (2)

Table 1.2 Crossover frequencies for tight, semi-relaxed and relaxed B-Chain properties under inspection (crossover point between the low and high-frequency bands with the filter order indicated in brackets)

1.2.1 Screen channels (L, C, R)

Screen channels are typically considered the most important loudspeakers in the B-Chain. The current SMPTE standard states that these channels should have a low-frequency roll-off at 50 Hz [2]. Measurements over multiple venues, however, show screen channels deviate from the standard, where it is common to measure a roll-off around 30 – 40 Hz and in some cases an extension down to 20 Hz [5].

In this research, the crossover points of 50 Hz and 40 Hz are implemented for the tight and semi-relaxed/relaxed B-Chains, respectively. These properties demonstrate the difference between systems perfectly in line with standards and those that are not [5].

1.2.2 Surround channels (SL, SR)

The current standards for surround channel performance indicate that they should follow the characteristics of the screen channels [2]. Experimental data shows that in reality, surround channels typically exhibit a roll-off between 30 Hz and 60 Hz, but there is extremely poor consistency across venues [5]. It is essential to apply delay to the surround channels so that no matter where an individual is located within a cinema or dubbing theater, the screen channel signals will arrive prior to the surround channels. This avoids distraction away from the screen.

Surround channel crossover points were chosen as 125 Hz, 60 Hz and 50 Hz for tight, semi-relaxed and relaxed B-Chains, respectively. The 125 Hz crossover for the tight system is based on manufacturer recommendations [7,8] and although this is not in agreement with the current standard [2], it helps to highlight the advantages of allowing surround channels to reproduce low-frequencies. No time delay was applied in this work, because the highlighted calibration strategies operate regardless of additional system processing.

1.2.3 Subwoofers (LFE)

It is essential that the function of the subwoofers in B-Chains is clearly defined in this work. The acronym LFE has been used regarding the B-Chain for many years, but it has two possible meanings. For modern digital B-Chains, LFE stands for the low-frequency effects channel. This is a separately mixed channel (the ".1" in surround sound configurations). For older analog B-Chains LFE stands for low-frequency extension. In these systems, the LFE is used to extend the low-frequency sound reproduction capabilities of the screen channels [1,3]. This work assumes that modern digital B-Chains are employed, therefore LFE stands for low-frequency effects channel.

In practice it is not uncommon for the sound mixers on films to place content from the LFE into the screen and surround channels to achieve more impact. This work treats the B-Chain under this premise; all loudspeakers within the B-Chain are considered available for any required low-frequency sound reproduction. If the ideas stemming from this research are implemented in practice, it must be understood that some form of post-processing may be required to properly route low-frequency content to the necessary loudspeakers without corrupting the sound designer's artistic intent.

Current standards indicate that the subwoofer should be capable of reproducing sound between 5 Hz and 125 Hz, with a sharp roll off above that [2,3,7,8]. The crossover point was therefore set to 125 Hz, with the semi-relaxed and relaxed B-Chains exhibiting a much more gradual roll-off than that of the tight B-Chain.

On the subject of inter-channel delay, there is currently no standardized fixed-time relationship between the screen/surround channels and the LFE channel. This research does not address the issue of delay, but it must be emphasized that the strategies detailed here will operate regardless of inter-channel delay.

1.2.4 Loudspeaker configuration

Twelve system configurations were chosen. The first six consist of subwoofers while the last six repeat the subwoofer configurations, but with the addition of the screen (L, C, R) and surround (SL, SR) channels. The configurations under inspection are shown in Fig. 1.1, whereby screen, surround and subwoofer units are indicated by squares, diamonds and circles, respectively. An 81-point listening grid is included, with measurement points indicated with crosses.

Screen and surround channels have a height of 6.6 m and 4.5 m, respectively. Subwoofers were placed on the floor. These heights were chosen based on published recommendations [7,8]. All measurements were taken at a height of 1.6 m, which is not perfectly in line with the SMPTE standard [2], but is in line with the practice of many experienced calibration engineers and still effectively demonstrates the characteristics of B-Chain low-frequency sound reproduction, regardless of measurement height.

As critical listening in dubbing theaters is commonly performed in a restricted listening area, the simulations for the dubbing theater were repeated with a 9-point listening grid centered about a central listening location, two-thirds of the room length from the screen. This area covers 10 m^2, as compared to the 112 m^2 covered by the 81-point listening grid. This additional configuration will highlight the effect of listening area size on the effectiveness of calibration strategies.

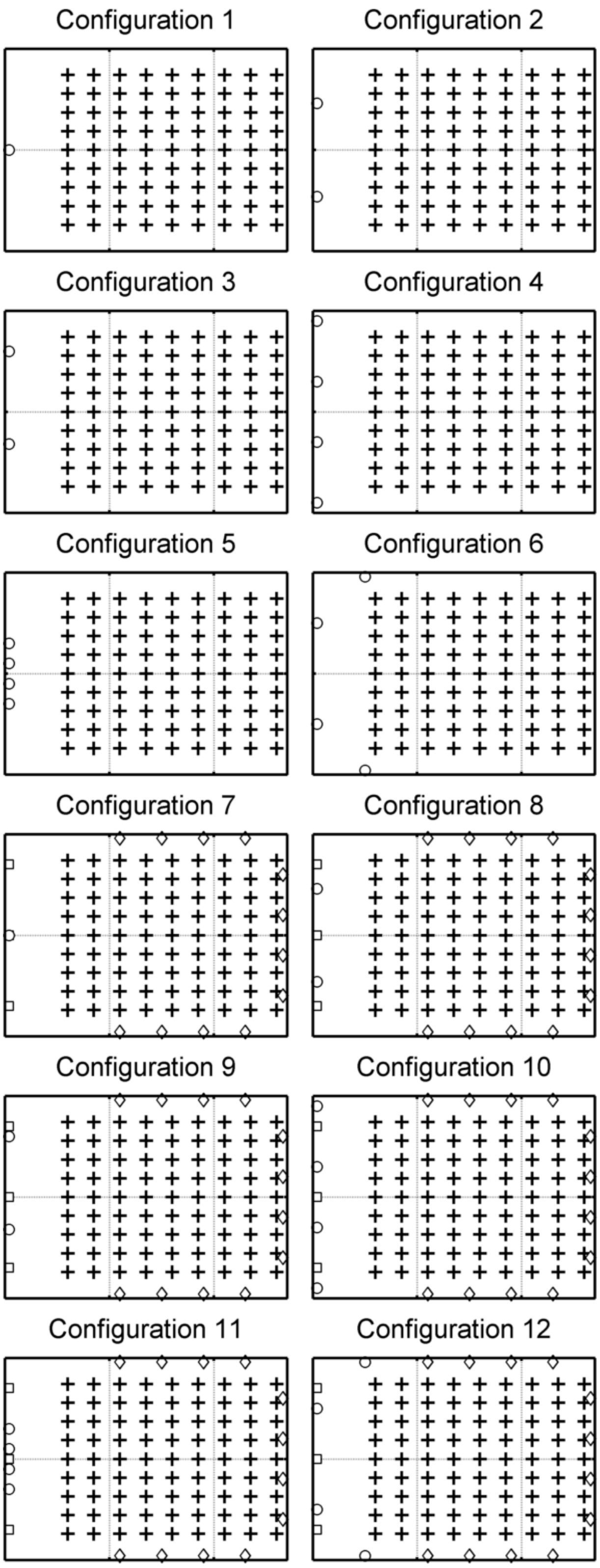

Figure 1.1 Configurations under examination (squares = L, C, R, diamonds = SL, SR and circles = LFE)

1.3 Room-modes

A considerable portion of published discussions regarding calibration of B-Chains for low-frequency optimization focuses on the issue of room-modes. Room-modes are complex standing wave patterns due to multiple reflections between parallel surfaces. Room-mode frequencies are defined (in rectangular topologies) using Equation 1.2 [9].

$$f_m = \frac{c}{2}\sqrt{\left(\frac{\eta_x}{L_x}\right)^2 + \left(\frac{\eta_y}{L_y}\right)^2 + \left(\frac{\eta_z}{L_z}\right)^2} \tag{1.2}$$

where, f_m is the m^{th} room-mode (Hz) which is based on the speed of sound in air (c, in m/s), the modal indices (η_x, η_y, η_z) and the room dimensions, (L_x, L_y, L_z measured in meters).

The published SMPTE standard [2] states that "microphone positions employed in a spatial average will be distributed among a range of positions in lateral and transverse directions to minimize the influence of any particular room mode." A similar statement is given in the corresponding SMTPE recommendation [3].

This postulation is entirely untrue. Room-modes cannot be addressed using spatial averaging because the room-mode pattern is a result of the geometry of a closed acoustic space, resulting in highly position-dependent frequency responses. Response averaging, therefore, does nothing to reduce spatial variance and only causes the average frequency response to match the target equalization curve. Although severe room-modes issues are avoided in the averaged response, they still exist at individual locations and will remain uncorrected. This is demonstrated in [10] and is also highlighted in [1]. There is no published solution to address these issues for B-Chains. The question is whether room-modes are actually an issue. A detailed analysis is required.

The low-frequency range of a closed acoustic space is commonly referred to as the *modal region*. This is the frequency range over which individual spectral resonances can be distinguished. The upper limit of the modal band is most commonly defined as the Schroeder frequency (Eq. 1.3) [11].

$$f_s = 2000\sqrt{\frac{RT_{60}}{V}} \tag{1.3}$$

where, f_s is the Schroeder frequency (Hz), RT_{60} is the average reverberation time (s) and V is the room volume (m^3). The Schroeder frequency is based on the spectral and spatial density of room-modes. Although room-modes exist across the frequency spectrum, above the Schroeder frequency they are sufficiently dense so that the human ear cannot distinguish individual modes due to spatial and spectral masking [11].

Taking the two topologies (commercial cinema and dubbing theater) into consideration (regarding their respective modal regions) will define the frequency band over which room-modes must be addressed. Using Eq. 1.3 along with the average low-frequency RT_{60} values (63 Hz and 125 Hz bands, over all relevant venues) of 1.5s for the commercial cinema and 0.44s for the dubbing theater [5], the Schroeder frequencies can be calculated.

The commercial cinema and dubbing theater have Schroeder frequencies of 33.3 Hz and 37.9 Hz, respectively. As human hearing is insensitive to narrow anomalies in this very low-frequency range, it can be deduced that room-modes are not a central issue.

The chief issue in the low-frequency band, therefore, is comb-filtering between direct sounds from loudspeakers and low-order reflections. Previous publications addressing low-frequency issues clearly state to not attempt comb-filtering correction [12], which is correct when using third-octave graphic or parametric equalizers (although this is a flawed approach in its own right – as will be discussed in Section 2). Comb-filtering also occurs due to the shifting nature of the acoustic center at low-frequencies [17]. While this is an important issue to consider, it is not directly addressed in this research because this effect is more severe outside of the defined subwoofer range (below 125Hz). The configurations modeled here place subwoofers a maximum of 0.4 m away from the nearest boundary. Assuming an acoustic center shift of 0.3 m in front of the drive unit, then the distance between the direct and virtual sources becomes 1.4 m. Cancellation occurs around the frequency with a half wavelength corresponding to this distance, which in this case equals 122.5 Hz.

A range of well-informed low-frequency system calibration approaches exist where comb-filtering is addressed (including the effects of the shifting acoustic center). These approaches are explored in Section 3.

1.4 Spatial variance

In order to adequately address the low-frequency issue, the problem must be quantified (which appears to be largely absent from published standards and recommendations). A common metric used in low-frequency research is known as spatial variance.

Spatial variance takes into consideration a range of frequency responses measured at numerous points across a wide listening area, determines the mean frequency response and then enumerates on average how much each individual frequency response differs from the ensemble mean. This is performed at each frequency bin and an average value is obtained (Eq. 1.4).

$$SV = \frac{1}{N_f} \sum_{i=f_{lo}}^{f_{hi}} \sqrt{\frac{1}{N_p - 1} \sum_{p=1}^{N_p} \left(L_p(p,i) - \overline{L_p(i)} \right)^2} \quad (1.4)$$

where, spatial variance (*SV*, in dB) is calculated based on the number of frequency bins analyzed (N_f), the frequency range (f_{lo} to f_{hi}), the number of measurement points (N_p), the sound pressure level at point *p* and frequency *i* ($L_p(p, i)$) and the mean sound pressure level across all measurement points at frequency *i* ($Lp(i)$). This research uses spatial variance to quantify the variability of the low-frequency response across the seating area in the topologies under examination.

1.5 Psychoacoustical considerations

An area often overlooked when dealing with low-frequency reproduction is whether multichannel reproduction is required in this frequency band. Recent research highlights how previous experiments into the perception of low-frequency directionality give conflicting results [13]. An emerging theory of low-frequency localization in closed spaces surmises that low-frequency localizability depends on room dimensions, source and listener location and source signal characteristics [13]. Ultimately, this implies that a catch-all statement regarding low-frequency localization cannot be made and the issue must be inspected on a case-by-case basis.

What this recent research has not addressed, however, is whether low-frequency directionality is important in the context of a full-range signal. Other published work indicates that incorrect low-frequency localization cues may conflict with the (potentially) correct high-frequency cues, resulting in a degraded sound image [14]. While this may indicate that all loudspeakers in a surround system should be full-range to ensure accurate sound imaging, work is still ongoing towards a definitive conclusion; consequently in this current work it will be assumed that multi-channel low-frequency sound reproduction is not essential in regards to localization.

Further research has been reported on the perceptibility of stereo low-frequency reproduction in a live-sound reinforcement system [15]. While there was no statistically significant effect switching between mono and stereo low-frequency reproduction, what did occur was a noticeable reduction in spatial variance of the frequency response across the audience area (using both measurements and listening tests).

This can be attributed to the decorrelation of left and right channels in a stereo system, whereas a mono signal driving all subwoofers results in correlated signals. The correlated acoustic signals cause position-dependent interference patterns, thus increasing spatial variance. This effect has been recognized by researchers focusing on B-Chain calibration in terms of surround channel interference in the form of comb-filtering [16].

While multi-channel low-frequency sound reproduction is not essential for sound imaging purposes, it is advantageous for sound reproduction uniformity within the low-frequency band. This should be kept in mind when designing a B-Chain calibration strategy.

2 CURRENT CALIBRATION STRATEGIES

The bulk of published literature on B-Chain calibration recommends using third-octave band real-time analysis (RTA) with a graphic or parametric equalizer used to implement the corrections. The current calibration procedure recommends a centrally-located measurement point, located two-thirds of the room length away from the screen. In some situations, multiple microphones are used to estimate a spatially-averaged response, but there is no well-defined standard for this process [4].

2.1 Single-point equalization

The use of third-octave analysis and equalization is inappropriate over any frequency range since it is not (as is often believed) in line with human perception of complex sounds [4]. Indeed, smoothing the frequency response to this extent is likely to cause anomalies in the response to be overlooked.

The focus must be to minimize spatial variance across a seating area. Configuration 1 (tight B-Chain) in the commercial cinema detailed in Sections 1.1 and 1.2 were modeled using a finite-difference time-domain (FDTD) acoustic simulation toolbox [18] with a grid point spacing of 0.4 m and an 11^{th} order MLS signal with a sample rate of 1.486 kHz (calculated to avoid spectral or spatial aliasing [10]).

The frequency responses at all 81 points are plotted in Fig. 2.1 along with the calculated spatial variance. The smoothed responses (with 1/12 octave smoothing) are also presented and are used exclusively for the duration of this paper as the smoothed responses are a better approximation of human hearing than are the unsmoothed responses [4].

Clearly there is severe spatial variance with this configuration, resulting in highly position-dependent listening experiences. The existing calibration strategies using a single-channel graphic or parametric equalizer can now be tested. An idealized case is examined here, whereby an inverse filter is generated based on the complex frequency response at a single measurement point [19]. The single-point equalization method results are shown in Fig. 2.2 with the target measurement point indicated by the thick black line.

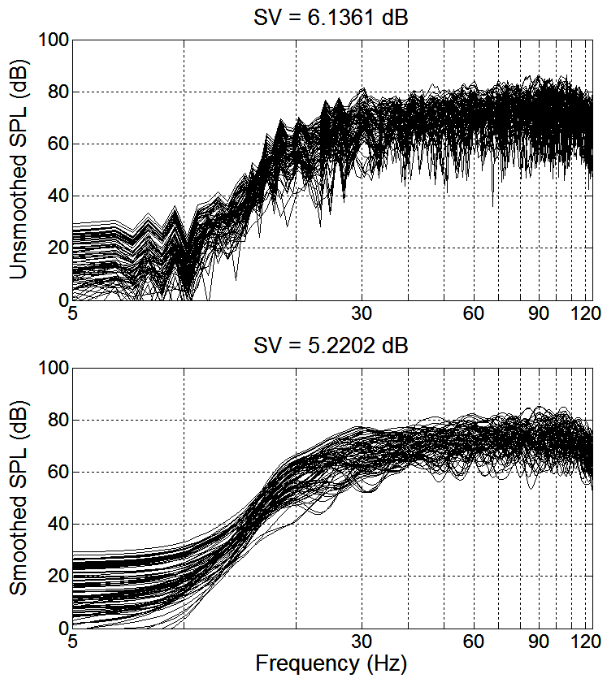

Fig 2.1 81-point measurement grid frequency responses for configuration 1 (tight B-Chain) in a cinema (top = unsmoothed, bottom = 1/12 octave smoothed)

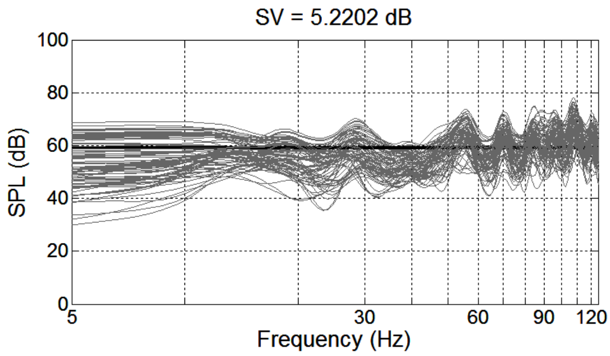

Fig 2.2 81-point measurement grid frequency responses for configuration 1 (tight B-Chain) in a cinema with single-point EQ applied (target point = solid black line)

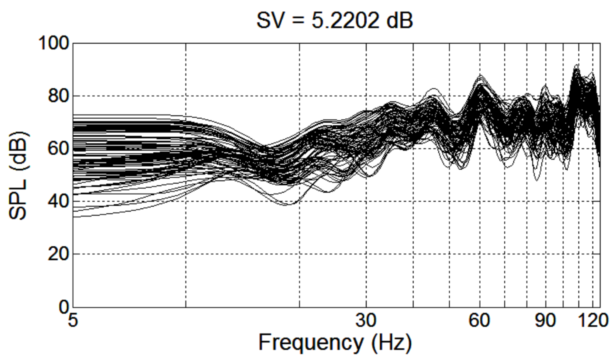

Fig 2.3 81-point measurement grid frequency responses for configuration 1 (tight B-Chain) in a cinema with spatially-averaged response EQ applied

While this is a much more precise form of equalization than is available using a one-third-octave band equalizer, it demonstrates the central issue with single-point correction. The target point indeed shows a perfectly flat frequency response, however the other 80 measurement locations are equally poor as before, with the exception of flattening of the response below 20 Hz. In reality, this would not occur since the acoustic model assumes an ideal loudspeaker, where this form of equalization can be introduced without risking damage to the drive unit. Critically, spatial variance is unchanged, so this approach to B-Chain calibration provides no benefit to the low-frequency response.

2.2 Spatially-average response equalization

Similarly, a spatially-averaged response measurement strategy can be modeled. In this case, the frequency response at each of the 81 measurement locations is taken and then the responses are averaged to generate an inverse filter (Fig. 2.3). As with the single-point method, this calibration strategy provides no reduction in spatial variance.

2.3 Physical configuration

Lastly, subwoofer placements can be inspected to determine if they provide any significant reduction in spatial variance across the listening area. All twelve configurations detailed in Fig. 1.1 were modeled to determine their respective spatial variances (Fig. 2.4). All B-Chains were tested in the commercial cinema and dubbing theater models.

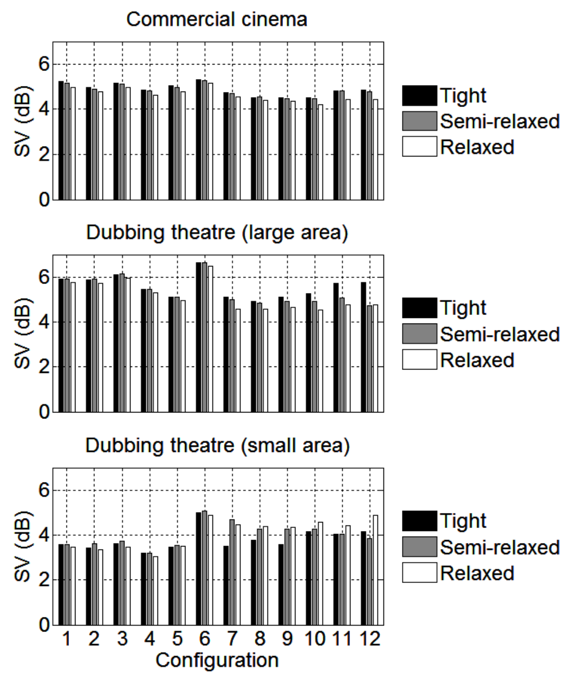

Fig 2.4 Spatial variance (SV) calculations for each B-Chain configuration, as detailed in Fig. 1.1

Regardless of the configuration or B-Chain properties, there is little change in spatial variance. The data indicates, however, that spatial variance over the large areas slightly decreases when allowing for low-frequency content in the screen and surround loudspeakers (configurations 7 – 12). This must be kept in mind to develop an effective calibration strategy.

Systems calibrated with any of the above-mentioned techniques will suffer from roughly (or exactly, in some cases) the same spatial variance as with an uncorrected system. If the goal for B-Chain sound reproduction in the low-frequency band is to achieve an even response across an entire seating area, then a more informed approach must be adopted.

3 IMPROVED CALIBRATION STRATEGIES

Considering the analysis presented in the preceding two sections, it is clear that existing calibration strategies incorrectly address the issue of spatial variance in the low-frequency band. An effective and robust calibration strategy is required that adequately minimizes spatial variance while being simple enough to implement and maintain by a moderately-competent local technician.

3.1 Optimization algorithms

Low-frequency optimization in rooms is a challenge that has been the focus of a large amount of research for many years. There exist numerous approaches to spatial variance minimization (largely targeted for home-cinema applications, but are often applicable to large-scale venues) which typically achieve their results through the application of least mean squares (LMS) based optimization algorithms, including a series of frequency response measurements taken from across the listening area [20-24]. Other methods use loudspeaker polar response control in order to avoid room-mode buildup along certain dimensions and to focus the sound energy towards the listeners [25-27].

It would be excessive and unnecessary to investigate each of these methods within this work. The polar response control methods will be left aside, as they are typically targeted at small-room systems (although the frequency-dependent polar response of certain advanced techniques may be worth future consideration [10, 26]), but it is worthwhile to investigate the usefulness of an optimization routine for B-Chain calibration.

The technique selected to highlight the effectiveness of system optimization is a chameleon subwoofer array (CSA), as described in [10]. This approach takes multiple complex frequency response measurements across a listening area and constructs a correction filter based on the spacing of the measurement points, capabilities of the subwoofers and the acoustical characteristics of the room.

Rather than targeting a flat response, the CSA algorithm targets the spatially-averaged response across the listening area, as it has been argued that people are accustomed to listening to room characteristics and a maximally-flat response may sound unnatural [28]. Whether this is the case or not in cinemas is beside the point, as the CSA system can be reconfigured to target a flat response, if necessary, although this may impact upon system efficiency [10].

As an example of this approach, CSA calibration was applied to configuration 12 from Fig. 1.1. This configuration was chosen as CSAs are most effective with maximally-spaced sources. Tight B-Chain characteristics were maintained, meaning that the screen and surround channels could partially contribute to sound reproduction in the low-frequency band (5 Hz – 125 Hz targeted in this case). The inclusion of all available loudspeakers provides additional degrees of freedom facilitating greater spatial variance reduction across a wide seating area (Fig. 3.1).

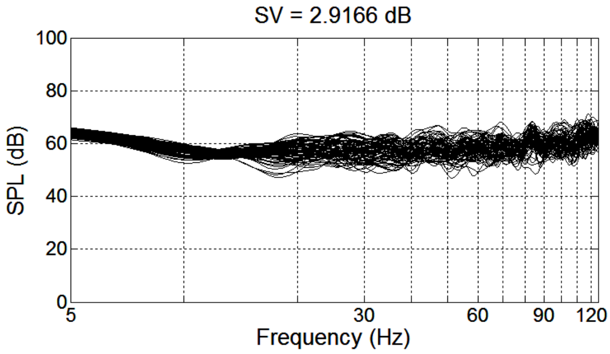

Fig. 3.1 Frequency responses at all 81 measurement points after CSA calibration applied to configuration 12 (Fig. 1.1) in a commercial cinema with tight B-Chain

It is important to note the effective upper frequency limit of CSA processing. This is defined by the largest dimension of the space and the mean listener point spacing (Eq. 3.1) [10].

$$f_H = \frac{cS_p N_m}{2L_x} \tag{3.1}$$

where, the effective upper frequency limit, f_H (Hz), is defined by the speed of sound, c (m/s), the listening point spacing ratio, S_p (1.2 in this case), the width of the listening grid, N_m (in measurement points, 9 in this case), and the largest dimension of the space, L_x (m).

Appling the largest spatial dimension of the commercial cinema (27 m) and the dubbing theater (17 m) results in effective upper frequency limits for CSA correction of 68.6 Hz and 109 Hz, respectively. Above these frequencies, the system runs the risk of spatial aliasing, thus reducing the accuracy of the calibration strategy.

This is essential to keep in mind when examining the full results later (Section 3.4).

Upon inspection of the responses following CSA calibration, it is clear there is significant reduction in spatial variance. Again, because this model assumes an ideal loudspeaker, the low frequency range below 20 Hz shows a significant boost. In reality this would not be the case, but neither would it be necessary, in practice.

While the CSA approach is highlighted here, there exist numerous optimization algorithms that are candidates for use as a B-Chain calibration strategy. However, it must be emphasized that a more uniform frequency response across a wide area is achievable providing a sufficient number of measurements are taken and the system is configured to send low-frequency content (including the LFE channel) to all loudspeakers, regardless of their low-frequency reproduction capabilities. Critically, each channel requires bespoke low frequency signal processing rather than just a single equalizer common to all loudspeakers.

3.2 **Diffuse signal processing**

Optimization algorithms can offer significant levels of spatial variance reduction while simultaneously providing control of the overall frequency response of a system. The drawback to these systems, however, is that they require calibration. As B-Chains in cinemas are likely to be calibrated and maintained by local technicians, there is a danger of incorrect implementation of the optimization algorithm, resulting in non-ideal performance. This problem can be avoided if a system is put in place capable of addressing low-frequency spatial variance without the need for calibration.

Diffuse signal processing (DiSP) was first described in [29] as a means of avoiding interference between correlated acoustic signals emanating from arrays of distributed mode loudspeakers (DMLs). The work alludes to the idea of using DiSP for non-DML applications, such as for the control of low-frequency sound reproduction in order to reduce spatial variance.

DiSP operates by using *temporally diffuse impulses* (TDIs). TDIs consist of an initial impulse followed by a rapid envelope decay whereby the decay segment is noise-like in nature [29]. A unique TDI is generated for each individual loudspeaker in a system which provides significant signal decorrelation to avoid interference. DiSP should result in lower spatial variance. This idea is a logical extension of the work discussed in [15], where it was found that stereo low-frequency sound reproduction provides moderate signal decorrelation, reducing sound energy nulls within a seating area.

The central issue in TDI generation is to avoid perceptible signal coloration. This work utilizes phase noise generated with a uniform probability density function along with linear coefficient interpolation, as described in [29]. As an example, the generated TDIs are 512 samples in length (1.486 kHz sample rate), the random phase values were restricted to ±0.94π and the frequency-dependent decay times ranged from 50 to 100 ms (highest to lowest frequency). The final TDIs were generated by taking the average of eight intermediate TDIs so as to smooth any sharp anomalies within the impulses. A full mathematical description of this TDI generation process can be found in [29].

The 19 TDIs (for 3 screen channels, 12 surround channels and 4 subwoofers) were applied to the MLS signal in the FDTD model and simulated. Configuration 12 was chosen as an example, as the subwoofers are widely spaced, allowing for maximal natural decorrelation of the radiated signals (Fig. 3.2).

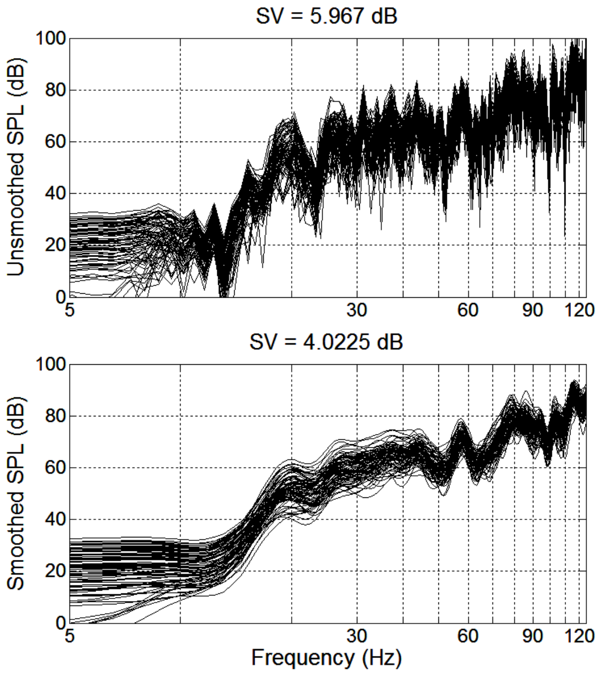

Fig. 3.2 Frequency responses at all 81 measurement points after DiSP applied to configuration 12 (Fig. 1.1) in a commercial cinema with tight B-Chain (top = unsmoothed, bottom = 1/12 octave smoothed)

Inspection of the unsmoothed frequency responses highlights the nature of how TDIs operate. They create a noise-like frequency response, due to the phase noise, resulting in sharp notches. After smoothing, these narrow notches are removed (in line with perception) resulting in a smoother set of responses. The smoothed spatial variance in this case dropped from 4.8325 dB to 4.0637 dB. While not as significant a decrease as with CSA, this example highlights the potential for DiSP use within B-Chains to allow for spatial variance reduction without the need for calibration by local technicians.

The example presented here is meant as proof of concept. Further work should be carried out to fine-tune the set of TDIs for maximum effect and to achieve minimum perceptible signal coloration.

3.3 **Hybrid approach**

The lack of required calibration for the DiSP strategy allows for a straightforward integration into existing systems. Building upon the independent investigations of the CSA and DiSP strategies, the two were combined to form a hybrid correction strategy.

DiSP processing was applied during the CSA calibration routine, which in theory should allow for further source-to-source decorrelation, moving the system closer to exhibiting independent degrees of freedom. The resulting performance was analyzed once again using configuration 12 from Fig. 1.1, with the resulting frequency responses shown in Fig. 3.3.

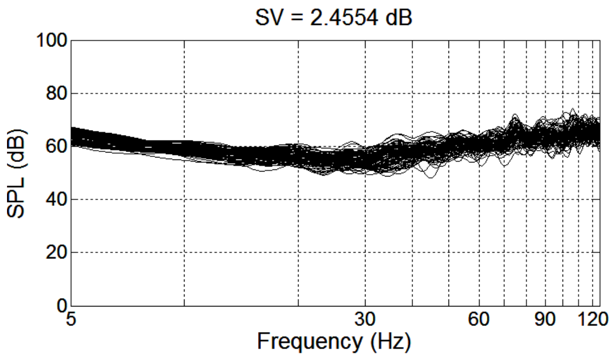

Fig. 3.3 Frequency responses at all 81 measurement points after CSA and DiSP applied to configuration 12 (Fig. 1.1) in a commercial cinema with tight B-Chain

The addition of DiSP to the CSA correction strategy further decreases the spatial variance over the 81-point listening grid by around 0.5 dB. While not a substantial improvement for this configuration, the full results shown in Fig. 3.4 indicate that when utilizing less-strict system properties, the improvement due to the hybrid approach is much more pronounced.

3.4 **Discussion**

The effectiveness of the calibration strategies over all twelve configurations is shown for the modeled commercial cinema and dubbing theater in Fig. 3.4.

3.4.1 Chameleon subwoofer array performance

The CSA calibration strategy is directly related to the available degrees of freedom. In the subwoofer-only systems, spatial variance reduction never exceeds 30%. This is due to the limited available subwoofers being located along the front of the cinema, thus impeding correction over a wide seating area and therefore exhibiting a wildly-varying frequency response.

The best performing calibration in this case is configuration 6 where two of the four subwoofers are placed along the side walls. This supports the argument that CSAs are most effective with wide source spacing [10]. In some cases (such as in the dubbing theater) it can be seen that CSA calibration in fact increases spatial variance. This is likely due to the coarse grid spacing used in the model. In the smaller space of the dubbing theater, it is possible that source and measurement grid points were extremely close to one another, resulting in slight system instabilities.

When the entire B-Chain is taken into consideration for low-frequency sound reproduction, CSA calibration shows its true strength. With 16 – 19 sources available (depending on the configuration), spatial variance reductions approaching 50% are achieved. Unsurprisingly, the non-tight B-Chains performed best in the commercial cinema (due to improved low-frequency reproduction capabilities of the non-subwoofer elements), but the opposite is seen in the dubbing theater results. This contradiction to expectation is likely due to reduced spacing of the loudspeakers in the dubbing theater, resulting in lower source-to-source decorrelation causing the CSA to be less effective when there is more spectral overlap between channels in the B-Chain.

3.4.2 Diffuse signal processing performance

The diffuse signal processing-based calibration strategy exhibits a slightly different behavior to the CSA method. Spatial variance reduction peaks at around 30% over all tested configurations, whereby effectiveness increases with the number of available loudspeakers for low-frequency sound reproduction.

While DiSP calibration does not approach CSAs in terms of effectiveness, it must be stressed that the key advantage of DiSP is that no on-site calibration is required. Once the TDIs have been generated, they are applied to the processing chain for each loudspeaker. No knowledge of venue/system topology is required to implement this strategy, thus lending itself to a universally-robust solution for B-Chain calibration.

Assuming the TDIs are carefully generated to avoid coloration, this form of calibration should not affect the timbre of the system, which therefore circumvents the issues raised in [30] where steady-state based equalization is shown to negatively impact the direct sound from the sources in the form of clearly noticeable coloration.

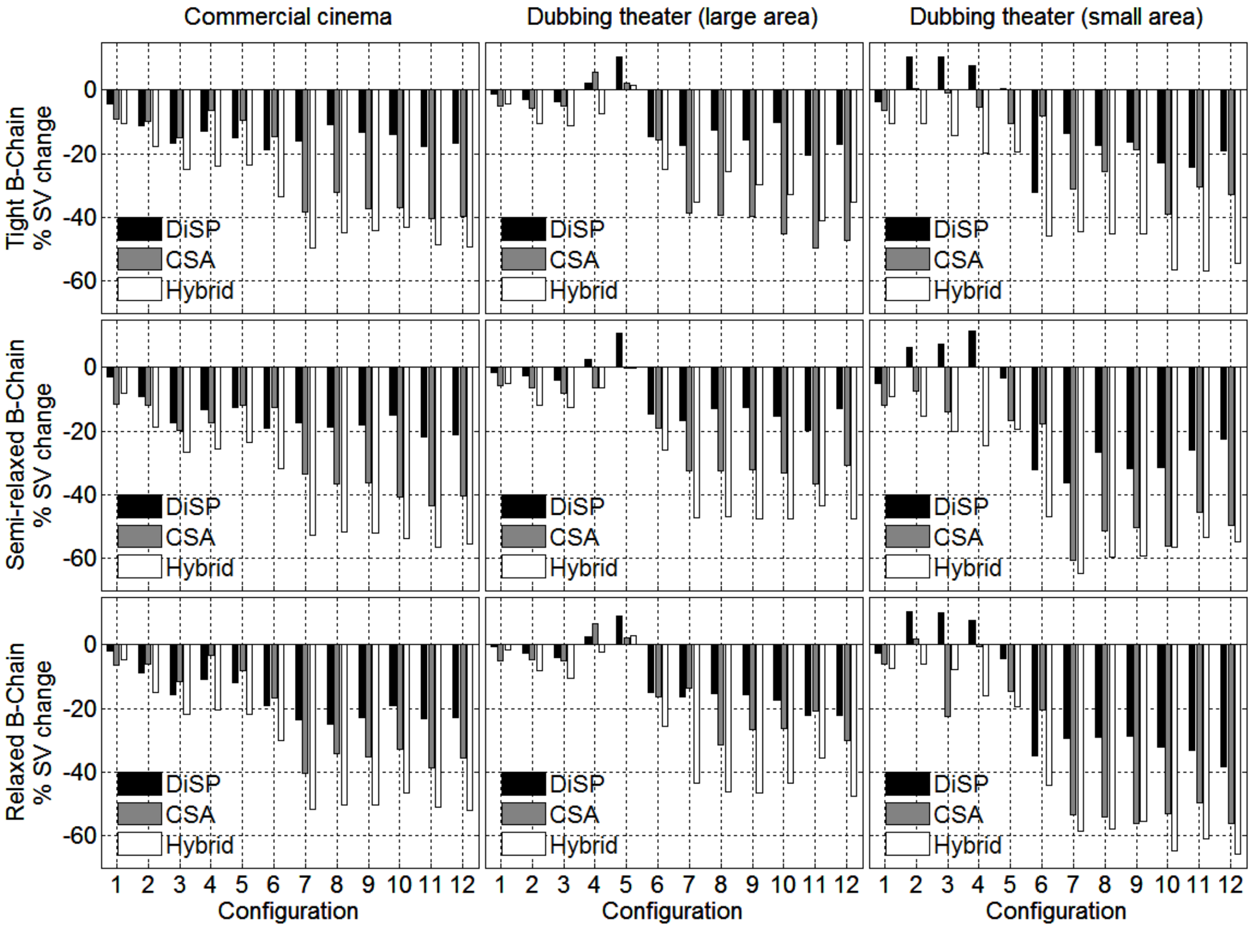

Fig. 3.4 Spatial variance reduction (in reference to spatial variance of the uncorrected systems) due to CSA, DiSP and hybrid calibration strategies for all 12 configurations

4 RECOMMENDATIONS

The research presented in the preceding sections highlights the flaws inherent with current B-Chain calibration strategies in regards to low-frequency performance, as well as suggesting alternative effective and practical calibration strategies which are directly in reference to a clearly defined problem (Section 1).

A set of recommendations can be assembled with the aim of informing new SMPTE standards and recommendations for B-Chain calibration:

1) The low-frequency response of B-Chains must be calibrated in reference to spatial variance reduction. This single metric provides a clear indicator of the effectiveness of a calibration strategy.

2) Low-frequency sound reproduction should not be restricted to the subwoofers. Systems should allow low-frequency content (from the LFE and screen/surround channels) through all available loudspeakers (where their output capability permits). This provides enhanced degrees of freedom for effective calibration.

3) Optimization algorithms achieve extremely low spatial variance. They require precise calibration and careful maintenance. Performance benefits must be weighed against practicality before implementing such an approach.

4) Diffuse signal processing (DiSP) achieves moderately low spatial variance. It requires no calibration or maintenance, allowing for universal implementation at minimal cost. Care must be taken during DiSP development to avoid perceptible signal coloration, ensuring that transient and steady-state sounds maintain their intended timbre.

5) Regardless of the adopted calibration strategy, the approach must be designed with practicality in mind. Local technicians must be able to easily implement and maintain the system without significant risk of human error. Systems must be designed to be stable and to not easily drift out of calibration.

5 CONCLUSIONS

The current strategies for low-frequency calibration of B-Chains are based on a flawed premise of low-frequency acoustics and psychoacoustics. These strategies are shown in this research to provide absolutely no benefit in terms of spatial variance reduction, meaning that the pre- and post-calibrated systems will exhibit equally position-dependent listening experiences. Furthermore, these techniques have been shown by other researchers to be highly prone to human error, resulting in inconsistent performance and often strongly colored transient signals due to excessive equalization [31].

The typical focus on room-modes when designing a low-frequency calibration system is not necessary in the case of modern cinemas, as the dimensions of the space coupled with low reverberation times results in Schroeder frequencies around 35 Hz. Above the Schroeder frequency, the effects of room-modes are not perceptible and therefore do not need to be directly addressed when calibrating a B-Chain. Comb-filtering between sources and low-order reflections is the primary cause of high spatial variance.

Suitable calibration strategies have been presented in Section 3. One possibility is to use an optimization algorithm, based on multiple measurements over a seating area. The included example evidences spatial variance reduction of nearly 50%, assuming a sufficient number of degrees of freedom (i.e. available loudspeakers for low-frequency sound reproduction). This option requires on-site calibration and maintenance.

The second option is diffuse signal processing. While not as effective as optimization algorithms, this method can reduce spatial variance by upwards of 30% (likely more if the TDIs are accurately designed), assuming sufficient degrees of freedom. This option requires no on-site calibration or maintenance.

In the case of an optimization algorithm-based calibration strategy, DiSP can be included within the system without any additional calibration. This hybrid approach decreases correlation between system degrees of freedom, thus maximizing spatial variance reduction (approaching 60% in certain scenarios).

Regardless of the new calibration strategy (hopefully) adopted by the industry, it is clear that a change is long overdue. Current calibration strategies are simply ineffective and are likely to be doing more harm than good. It is the hope of the authors that this research will prompt an informed discussion regarding the revision of current SMTPE standards and will lead to improved and consistent performance of B-Chains across the world.

REFERENCES

[1] Newell, J.; P. Newell; K. Holland. "Room low-frequency response estimation using microphone averaging." Proc. IOA – Reproduced Sound 2014, vol. 36, pt. 3, 2014.

[2] "SMPTE standard for motion pictures – Dubbing stages (mixing rooms), screening rooms and indoor theaters – B-Chain electroacoustic response." SMPTE ST 202:2010. October 20, 2010.

[3] "SMPTE recommended practice – Relative and absolute sound pressure levels for motion-picture multichannel sound systems – Applicable for analog photographic film audio, digital photographic film audio and D-Cinema." SMPTE RP 200:2012. April 18, 2012.

[4] Newell, P.; G. Leembruggen; K. Holland; J. Newell; S. Torres Guijarro; D. Gilfillan; D.S. Dominguez, S. Castro. "Does 1/3rd octave equalization improve the sound in a typical cinema?" Proc. IOA – Reproduced Sound 2011, vol. 33, pt. 6, 2011.

[5] "TC-25CSS B-Chain frequency and temporal response analysis of theatres and dubbing stages." SMPTE theatre testing data report B-Chain study group. October 1, 2014.

[6] Walker, R. "Low-frequency room responses: Part 1 – Background and qualitative considerations." BBC Research Department Report. RD 1992/8. 1992.

[7] Allen, I. "Technical guidelines for Dolby stereo theatres." Dolby Laboratories, rev. 1.33. November, 1994.

[8] "Cinema sound system manual." JBL Professional. 2003.

[9] Toole, F.E. *Sound Reproduction: Loudspeakers and Rooms*, pp.385-389. Focal Press, New York. 2008.

[10] Hill, A.J. "Analysis, modeling and wide-area spatiotemporal control of low-frequency sound reproduction." Ph.D. Thesis, University of Essex, UK. January, 2012.

[11] Schroeder, M.R. "The 'Schroeder frequency' revisited." Journal of the Acoustical Society of America, vol. 99, no. 5, pp. 3240-3241. May, 1996.

[12] Newell, P.; K. Holland; J. Newell; B. Neskov. "New proposals for the calibration of sound in cinema rooms." 130th Convention of the Audio Engineering Society, paper 8383. May, 2011.

[13] Hill, A.J.; M.O.J. Hawksford. "Subjective evaluation of an emerging theory of low-frequency sound source localization in closed acoustic spaces." Proc. IOA – Reproduced Sound 2014, vol. 36, pt. 3, 2014.

[14] Schnupp, J.; I. Nelken; A. King. *Auditory neuroscience: Making sense of sound* (Chapter 5: Neural basis of sound localization). MIT Press, Cambridge, MA, USA. 2011.

[15] Hill, A.J.; M.O.J. Hawksford. "On the perceptual advantage of stereo subwoofer systems in live sound reinforcement." 135th Convention of the Audio Engineering Society, paper 8970. October, 2013.

[16] Newell, P.; K. Holland; S. Torres Guijarro, J. Newell; D.S. Dominguez. "Human factors affecting the acoustic measurement of rooms." Proc. IOA – Reproduced Sound 2012, vol. 34, pt. 4. 2012.

[17] Vanderkooy, J. "The low-frequency acoustic centre: Measurement, theory and application." 128th Convention of the Audio Engineering Society, paper 7992. May, 2010.

[18] Hill, A.J.; M.O.J. Hawksford. "Visualization and analysis tools for low-frequency propagation in a generalized 3D acoustic space." Journal of the Audio Engineering Society, vol. 59, no. 5, pp. 321-337. May, 2011.

[19] Genereux, R.P. "Adaptive loudspeaker systems: correcting for the acoustics environment." 8th International Conference of the Audio Engineering Society. May, 1990.

[20] Kuriyama, J.; Y. Furukawa. "Adaptive loudspeaker system." Journal of the Audio Engineering Society, vol. 37, no. 11, pp. 919-926. November, 1989.

[21] Elliott, S.J.; P.A. Nelson. "Multiple-point equalization in a room using adaptive digital filters." Journal of the Audio Engineering Society, vol. 37, no. 11, pp. 899-907. November, 1989.

[22] Bharitkar, S.; C. Kyriakakis. "Comparison between time delay based and non-uniform phase based equalization for multi-channel loudspeaker-room responses." 119th Convention of the Audio Engineering Society, paper 6607. October, 2005.

[23] Santillan, A.O. "Spatially extended sound equalization in rectangular rooms." Journal of the Acoustical Society of America, vol. 110, no. 4, pp. 1989-1997. October, 2001.

[24] Elliot, S.J.; L.P. Bhatia; D.W. Deghan; A.H. Fu; M.S. Stewart; D.W. Wilson. "Practical implementation of low-frequency equalization using adaptive digital filters." Journal of the Audio Engineering Society, vol. 42, no. 12, pp. 988-998. December, 1994.

[25] Olson, H.F. "Gradient Loudspeakers." Journal of the Audio Engineering Society, vol. 21, no. 2, pp. 86-93. March, 1973.

[26] Backman, J. "Low-frequency polar pattern control for improved in-room response." 115th Convention of the Audio Engineering Society, paper 5867. October, 2003.

[27] Linkwitz, S. "Investigation of sound quality differences between monopolar and dipolar woofers in small rooms." 105th Convention of the Audio Engineering Society, paper 4786. September, 1998.

[28] Vanderkooy, J. "Multi-source room equalization: Reducing room resonances." 123rd Convention of the Audio Engineering Society, paper 7262. October, 2007.

[29] Hawksford, M.O.J.; N. Harris. "Diffuse signal processing and acoustic source characterization for applications in synthetic loudspeaker arrays." 112th Convention of the Audio Engineering Society, paper 5612. May 2002.

[30] Leembruggen, G.; P. Newell; J. Newell; D. Gilfillan; K. Holland; B. McCarty. "Is the X curve damaging our enjoyment of cinema?" SMPTE 2011. Australia.

[31] Newell, P.; K. Holland; M. Desborough; B. Fazenda; B. Neskov; S. Castro; E. Valdigem; J. Newell. "Room-to-room compatibility of the low-frequency content of mixes using resonant and transient signals." Proc. IOA – Reproduced Sound 2009, vol. 31, pt. 4, 2009.