Enhanced Wide-Area Low-Frequency Sound Reproduction in Cinemas: Effective and Practical Alternatives to Current Calibration Strategies

ADAM J. HILL,[1] *AES Member*, **MALCOLM O. J. HAWSKFORD,**[2] *AES Fellow*, **AND**
(a.hill@derby.ac.uk)   (mjh@essex.ac.uk)

PHILIP NEWELL,[3] *AES Member*
(philiprnewell@gmail.com)

[1]*Department of Engineering, University of Derby, United Kingdom*
[2]*School of Computer Science and Electronic Engineering, University of Essex, United Kingdom*
[3]*Acoustic consultant, Moaña, Spain*

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, a methodology 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 and practical experiments that expose a potential for superior wide-area signal reconstruction over that achieved using the current standards and recommendations.

0 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, 2]. Current standards and recommendations [3, 4] suggest using one-third-octave band graphic or parametric equalization to smooth the low-frequency response in cinemas. These techniques are often based on spatially-averaged response measurements across a seating area [3], which has been proven to give minimal benefit regarding uniformity or overall "flatness" in the low-frequency band [1, 2, 5–7]. 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 can be developed and standardized to allow for consistent low-frequency responses in cinemas and dubbing theaters (across all venues and seats within). [A B-chain consists of everything after the volume fader in a system including the power amplifiers, loudspeakers, screen, and any acoustical treatment.]

A detailed problem definition is laid out in Sec. 1 covering issues pertaining to acoustics (room-modes, comb-filtering, spatiotemporal variance), sound reproduction (individual channel frequency responses, interference, available degrees of freedom, existing standards), and psychoacoustic considerations, including the audibility of low-frequency resonances and anti-resonances. Sec. 2 details the characteristics of the venues that have been modeled in this study.

To facilitate comparison, Sec. 3 presents a critical examination of current calibration strategies. This includes inspecting system responses using various common loudspeaker configurations and applying both single-point and spatially-averaged response equalization. However, in contrast to this paper's philosophy, these current techniques contribute nothing to the reduction of seat-to-seat frequency response variations in cinemas. This is because they can only ever seek to globally re-balance the frequency response to that of an averaged listening position, implying that inter-seat response differences remain unchanged.

To address these limitations, Sec. 4 proffers a number of enhanced methods that include distributed optimization algorithms and diffuse signal processing. As such, significant improvements in frequency response uniformity can be achieved across entire seating areas, together with greater overall system robustness, as there is less chance of human error corrupting the calibration process.

Sec. 5 concludes by summarizing a set of recommendations that offer a basis for the development of a robust and spatially more accurate system for low-frequency calibration in cinemas. The intention is for these concepts to be taken into consideration when revising current standards [3] for B-chain calibration in cinemas and dubbing theaters.

1 PROBLEM DEFINITION

Wide-area low-frequency sound reproduction is inherently prone to high spatio-temporal variance across a listening area. This variance is typically caused by a combination of comb-filtering, room-modes, loudspeaker-to-room coupling, and listener location. In the frequency domain, the problem is observed as a position-dependent frequency response, where each listening location experiences different tonality of the reproduced sound [8–10]. In the time domain, the issue manifests itself as a smearing effect, resulting in poor waveform fidelity [9, 11, 12]. Considering the goal to deliver equal listening experiences to all members of an audience, it is crucial to address this problem.

With regard to cinema B-chains, recent efforts by the SMPTE Theatre B-Chain Study Group [6] have highlighted the severity of low-frequency variance across numerous commercial cinemas and dubbing theaters when analyzed in both the time and frequency domains. The group calls for a need to revise current standards and recommendations pertaining to the calibration of cinema B-chains, with a clear requirement to agree on a calibration technique to deliver uniform and temporally-accurate low-frequency sound to all members of an audience.

In order to develop an effective and practical methodology for addressing the identified issues in 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 Room-Modes

A considerable portion of published discussions regarding calibration of B-chains for low-frequency optimization focuses on the issue of room-modes, which are complex standing wave patterns due to multiple reflections between surfaces. Room-mode frequencies are defined (in rectangular topologies) using Eq. (1.1) [8].

$$f_m = \frac{c}{2}\sqrt{\left(\frac{\eta_x}{L_x}\right)^2 + \left(\frac{\eta_y}{L_y}\right)^2 + \left(\frac{\eta_z}{L_z}\right)^2} \qquad (1.1)$$

where, f_m is the mth room-mode (Hz), which is based on the speed of sound in air (c, in m/s), the modal indices ($?_x$, $?_y$, $?_z$), and the room dimensions, (L_x, L_y, L_z, measured in meters).

The low-frequency band 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 to this modal band is most

commonly defined as the Schroeder frequency (Eq. (1.2)) [13].

$$f_s = 2000\sqrt{\frac{RT_{60}}{V}} \qquad (1.2)$$

where, f_s is the Schroeder frequency (Hz), RT_{60} is the average reverberation time (s), and V is the room volume (m³). The Schroeder frequency is based on the spectral and spatial density of room-modes. Although room-modes exist across the entire frequency spectrum, above the Schroeder frequency they are sufficiently dense that the human ear cannot distinguish individual modes due to spatial and spectral masking [13].

The published SMPTE standard [3] 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 [4].

This postulation may be valid in rooms with naturally low spatio-temporal variance (usually due to high levels of acoustical treatment or great size), but in rooms with high variance, room-modes cannot be addressed or subverted using spatial averaging because the room-mode pattern is a result of the geometry of a closed acoustic space. This results in highly position-dependent frequency responses and consequently, measurement position dependent equalization, as highlighted in detail in [7].

Although severe room-mode issues are avoided in the averaged response, they still exist at individual locations and will remain uncorrected. Response averaging, therefore, does little to reduce spatial variance and only leads to the average frequency response matching the target equalization curve. This is demonstrated in [9] and is also highlighted in [1, 2]. There is no published solution to address these issues for B-chains.

The question is, however, whether room-modes are actually an issue. A detailed analysis is required. Taking the two topologies of a commercial cinema and dubbing theater into consideration will define the frequency band over which their respective modal regions must be addressed. Using Eq. (1.2) along with the average low-frequency RT_{60} values (63 Hz and 125 Hz bands, over all relevant venues) of 1.5 s for the commercial cinema and 0.44 s for the dubbing theater [6], the Schroeder frequencies can be calculated (although it is noted that some dubbing theaters now have RT_{60} values closer to the 0.2 s region).

The commercial cinema and dubbing theater have (theoretical) 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 [19], it can be deduced that room-modes are not a central issue.

The chief cause of spatio-temporal variance 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 that comb-filtering correction should not be attempted [14], which is correct when using either one-third-octave band graphic or parametric equalizers.

Comb-filtering also occurs due to the shifting nature of the acoustic center at low-frequencies [15]. 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. The inspected configurations place subwoofer front baffles a maximum of 0.4 m away from the nearest boundary. Assuming an average 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 Sec. 4.

1.2 Variance Quantification

In order to adequately address the low-frequency issue, variance must be quantified in some way, but this appears to be largely absent from published standards and recommendations [3, 4].

One metric commonly used in low-frequency research is known as spatial variance [16]. 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. (1.3)) [16].

$$SV = \frac{1}{N_f} \sum_{i=f_{lo}}^{f_{hi}} \sqrt{\frac{1}{N_p - 1} \sum_{p=1}^{N_p} \left(L_p(p, i) - \overline{L_p(i)} \right)^2} \quad (1.3)$$

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}) using linearly-spaced frequency bins in this work, 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 ($\overline{L_p(i)}$).

Spatial variance is not the only available metric to quantify the low-frequency performance of a sound reproduction system. Mean output level (MOL) is of interest as it provides a good indicator of overall efficiency. MOL probes the sound pressure level at each listening location taken over all frequency bins to give an average level over the audience area (Eq. (1.4)) [17].

$$MOL = \frac{1}{N_f N_p} \sum_{i=f_{lo}}^{f_{hi}} \sum_{p=1}^{N_p} L_p(p, i) \quad (1.4)$$

Consequently, for any calibration method to be deemed acceptable, it is important that the MOL is not significantly compromised.

Additionally, it has been shown that the transient response related to low-frequency reproduction must be con-

trolled to avoid transient coloration [47]. A metric previously proposed for this function is known as phase variance (PV) [11]. As this metric was developed specifically for small room sound reproduction, further work is required to devise an equivalent large room metric.

Variance of spatial average (VSA) is also of interest as it provides a measure of magnitude response flatness, especially for systems that have no access to global equalization [17]. However, because this paper assumes B-chain equalization, VSA is of less interest.

Although it could be considered expedient to employ a combination of temporal and spectral metrics to examine spatio-temporal accuracy, in order to lower the number of experimental parameters, this research focuses primarily on spatial variance to quantify the variability of the low-frequency response across the seating areas in the example listening spaces. However, because of the importance of maintaining system efficiency, MOL is examined in Sec. 4.5.

1.3 Loudspeaker Layout Calibration Effectiveness

In studying multi-loudspeaker systems, it can be instructive to inspect the so-called loudspeaker layout coupling factor (LCF) of various system configurations. This will give a reasonable estimate of the orthogonality of each loudspeaker's response, whereby a lower LCF indicates that each loudspeaker operates more independently within the calibration system.

The method developed for this purpose centers around each loudspeaker position's modal distribution function (MDF), which is calculated using Eq. (1.5) [18].

$$\Psi_{x,y,z}\left(s, \eta_x, \eta_y, \eta_z\right)$$
$$= \cos\left(\frac{\eta_x \pi x}{L_x}\right) \cos\left(\frac{\eta_y \pi y}{L_y}\right) \cos\left(\frac{\eta_z \pi z}{L_z}\right) \quad (1.5)$$

where $\Psi_{x,y,z}$ is the MDF for a source, s, located at coordinates (x, y, z). $?_x$, $?_y$ and $?_z$ are the modal indices for the room resonance under inspection, and L_x, L_y and L_z are the room dimensions (in meters).

The magnitude of an MDF can range from 0 to 1, where 0 indicates no coupling to a room resonance while 1 indicates maximal coupling.

It may appear counterintuitive to examine modal distributions considering the discussion in Sec. 1.1, where it was shown that room-modes will be sufficiently spatially and spectrally dense to be undetectable. Nevertheless, it is important to keep in mind that although the perception of discrete room-modes may not be possible, the room-modes still exist and contribute to the overall low-frequency room response.

Room resonances were calculated up to 150 Hz for this work; as although resonances around this upper limit are beyond the identified subwoofer band, they may nevertheless exhibit a low-Q resonance, and thus extend down into the subwoofer band.

The MDF was calculated for each resonance and then the LCF was calculated for a given configuration using

Eqs. (1.6) and (1.7). These equations calculate the mean MDF value over all sound sources and resonances within the frequency band of interest. The mean is converted onto a logarithmic scale for ease of interpretation. The method of calculation follows that used for mean output level in [16], whereby a value is obtained by taking a mean over multiple parameters.

$$LCF = 10\log_{10}\left(\frac{1}{N_f N_s}\sum \Psi\right) \qquad (1.6)$$

$$\sum \Psi = \sum_{s=1}^{N_s}\sum_{\eta_x=N_1}^{N_2}\sum_{\eta_y=N_1}^{N_2}\sum_{\eta_z=N_1}^{N_2}\Psi_{x,y,z}\left(s, \eta_x, \eta_y, \eta_z\right) \qquad (1.7)$$

where N_f is the number of room resonances analyzed, N_s is the number of loudspeakers in the system, N_1 and N_2 set the range of modal indices to inspect and s represents the source currently under inspection. In this instance, N_1 and N_2 were set to 0 and 25, respectively. Within the calculation, if the resonant frequency is outside the frequency band of interest, that specific MDF is not included in the summation.

The LCF calculation assumes all loudspeakers can reproduce low-frequencies with equal efficiency, so deviations are expected in comparing predictions to experimental data. For the purpose of cinema B-chains, this is not a significant issue, as the 2010 SMPTE standard [3] states that the screen and surround channels should be capable of producing at least –9 dB at 31.5 Hz, meaning that all loudspeakers in the B-chain should be capable of reproducing low-frequency content at sufficient levels to influence the overall system response.

Despite this approximation, LCF remains a good indicator of calibration effectiveness for a given loudspeaker layout. System designers should aim for a low LCF to ensure the system gains maximum benefit from any calibration procedures in use.

1.4 Psychoacoustical Considerations

Considering the various available metrics used to objectively quantify the uniformity and accuracy of low-frequency sound reproduction, it is important to consider how closely these metrics relate to perception at low frequencies. While measurements may indicate a resonance or an anti-resonance in the low-frequency response, if this is undetectable by the human ear it should not be a problem.

Previous investigations into the detection of low-frequency resonances and anti-resonances reveal some important findings [19]. First, if the Q of a resonance doubles, the detection threshold increases by around 3 dB, meaning that very narrow low-frequency resonances are more difficult to detect. If the Q of a resonance is low (around Q = 1), then the detection thresholds are largely independent of center frequency. For mid- to high-Q resonances, detection decreases by 0.5 dB to 2 dB per octave decrease. Only for mid- to high-Q values does the detection of resonances and anti-resonances differ. In these cases the anti-resonances (notches) become much harder to detect.

Last, and potentially most important to this work, is that signal type plays a central role in the detection of reso-

nances. Detection of high-Q resonances was found to be easier when the source signal was pulsed rather than continuous pink noise [19]. This means that for real program material (which will likely have considerable transient content), high-Q resonances will be more noticeable than for steady-state noise-like signals.

These findings support the assertion that the use of one-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 [5]. Indeed, smoothing the frequency response to this extent is likely to cause anomalies in the response to be overlooked. These findings are critical to keep in mind when examining the objective results presented later in this paper.

An area often overlooked when dealing with low-frequency reproduction is whether multichannel reproduction is desirable in the low-frequency band. Recent research highlights how previous experiments into the perception of low-frequency directionality give conflicting results [20]. 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 [20]. Ultimately, this implies that a catch-all statement regarding low-frequency localization cannot be made, and so 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 [21]. 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 multichannel low-frequency sound reproduction is not essential in regards to localization. Indeed, the LFE (low-frequency effects) channel in conventional digital cinemas is mono.

Further research has been reported on the perceptibility of stereo low-frequency reproduction in a live-sound reinforcement system [22]. While there was no statistically significant effect switching between mono and stereo low-frequency reproduction, what did occur was a noticeable reduction in variance of the magnitude response across the audience area (using 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 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 [23], but it has not been noted that its subjective effect is of any consequence (at higher frequencies).

The effect of signal decorrelation has also been recognized by researchers looking into auditory spatial imagery [24–26]. In each investigation the decorrelated subwoofer

Table 2.1. Physical properties for the two closed acoustic spaces under inspection

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
RT_{60}	1.50 s	0.44 s

Table 2.2. 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)

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)

signal increased the potential control of spatial imagery in small-sized rooms. When considering the surround effects required in cinema sound, the additional control and accuracy provided by low-frequency signal decorrelation when utilizing multiple independent low-frequency loudspeakers is quite important. Further research, however, must be carried out to inspect the validity of this effect in large spaces such as cinemas.

However, even if multichannel low-frequency sound reproduction is unessential for sound imaging purposes, it is advantageous for sound reproduction uniformity within the low-frequency band. This is a fact that should be kept in mind when designing any new B-chain calibration strategy.

2 VENUE CHARACTERISTICS

The physical properties of the two spaces to be assessed were chosen by taking the average dimensions and reverberation times of the commercial cinemas and dubbing theaters studied in a recently published SMPTE report [6]. Reverberation times of 1.50 s and 0.44 s (commercial cinema and dubbing theater, respectively, which were taken from the 63 Hz and 125 Hz octave band average RT_{60} measurements in [6]) were used to calculate average absorption coefficients using Eq. (2.1) [10].

$$RT_{60}(f) = \frac{0.161V}{S\alpha(f)} \qquad (2.1)$$

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 2.1.

2.1 Loudspeaker Properties

The frequency responses and placement of the loudspeakers in the B-chain under inspection were designed to align with recommendations [3, 4, 27, 28] and in-situ measurements at various venues [6]. 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 ideal and realistic systems. The selected crossover settings are given in Table 2.2.

For the simulations presented in this paper, all loudspeakers use an ideal frequency response shaped to match the properties highlighted in Table 2.2.

2.1.1 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 gently beginning at 50 Hz [3]. Measurements over multiple venues, however, show screen channels deviate significantly from the standard, where it is common to measure a roll-off around 30 to 40 Hz and, in some cases, an extension down to 20 Hz [6].

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 [6].

2.1.2 Surround Channels (SL, SR)

The current standards for surround channel performance indicate that they should follow the characteristics of the screen channels [3]. Experimental data shows that in reality, surround channels typically exhibit a roll-off anywhere between 30 Hz and 60 Hz, but there is extremely poor consistency across venues [6]. 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 50/60 Hz crossover for the relaxed/semi-relaxed systems are based on manufacturer recommendations [27, 28], and although this is not in agreement with the current standard [3], it helps to highlight the advantages of using surround loudspeakers that are capable of reproducing lower frequencies. No signal delay was applied in this work because the highlighted calibration strategies operate regardless of additional system processing.

2.1.3 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 capabil-

ities of the screen channels [1, 2, 4]. 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 20 Hz and 125 Hz, ±3 dB with a sharp roll off above that [3, 4, 27, 28] (it is also noted that signal content may extend down to 5 Hz [3]). 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.

There also remains the question of the LFE level calibration. Should B-chains be calibrated for uniform LFE level throughout a venue or should they target natural level attention with distance (as naturally occurs with the screen channels)? This topic is not addressed in this work but must be considered in any new B-chain calibration standards and recommendations.

2.1.4 Loudspeaker Configuration

Twelve system configurations were chosen. The first six consist only of LFE 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. 2.1, 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.

The 9×9 grid is centered at (16 m, 10 m) and (10 m, 6 m) in the commercial cinema and dubbing theater respectively. The x- and y-dimension grid point spacing was set with Eqs. (2.2) and (2.3). Tables 2.3, 2.4, and 2.5 indicate the positions of the non-LFE and LFE sources in both venues.

$$dx = \frac{L_x}{S_p G_x} \qquad (2.2)$$

$$dy = \frac{L_y}{S_p G_y} \qquad (2.3)$$

where the x-dimension spacing (dx) is determined by the room length, L_x (m), and the grid length, G_x (in measurement points), while the y-dimension spacing, dy (m), is
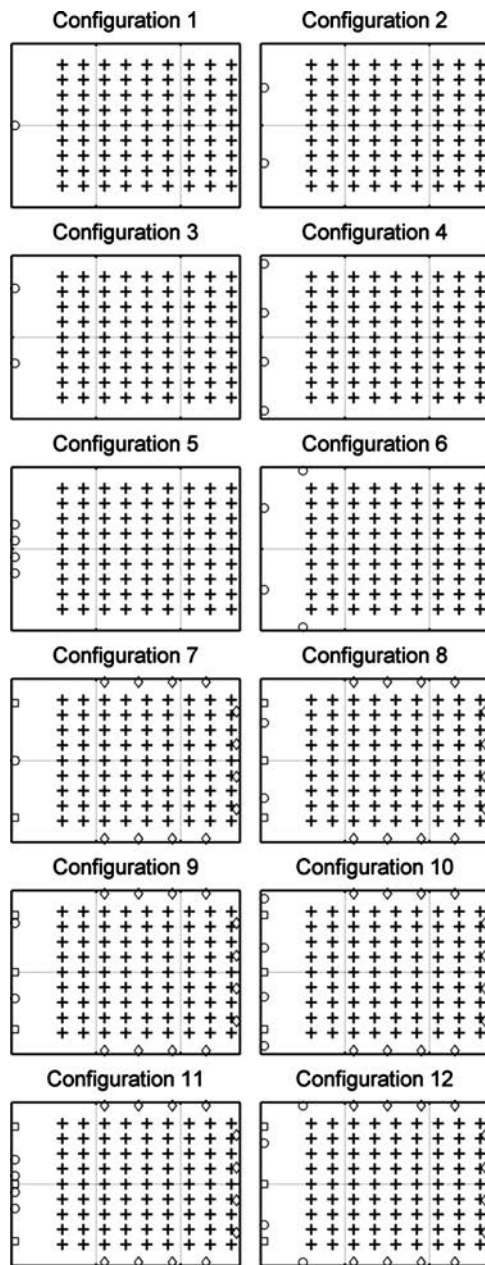

Fig. 2.1. Configurations under examination (squares = L, C, R, diamonds = SL/SR, circles = LFE and crosses = listening grid points).

determined by the room width, L_y (m), and grid width, G_y (in measurement points). The grid length and width are both nine points long in this instance. The spacing constant, S_p, in both denominators must be greater than one to avoid grid points being placed exactly on side walls or too close to the front or rear of the room. The higher the constant, the tighter is the grid spacing. S_p was set to 1.2 for this work. In the commercial cinema, grid points were spaced by 2.5 m and 1.9 m in the x- and y-dimensions, respectively. In the dubbing theater, grid points were spaced at 1.6 m and 1.1 m in the x- and y-dimensions, respectively.

The screen and surround channels have a height of 6.6 m and 4.5 m, respectively. Subwoofers were placed on the

Table 2.3. Non-LFE loudspeaker positions used for configurations 7 – 12, as shown in Fig. 2.1 (all positions given in meters)

Channel	Commercial cinema	Dubbing theater
L	(0.4, 3.0, 6.6)	(0.3, 1.8, 4.0)
C	(0.4, 10.0, 6.6)	(0.3, 6.0, 4.0)
R	(0.4, 17.0, 6.6)	(0.3, 10.2, 4.0)
SL 1	(11.0, 0.4, 4.5)	(6.9, 0.3, 2.7)
SL 2	(15.0, 0.4, 4.5)	(9.4, 0.3, 2.7)
SL 3	(19.0, 0.4, 4.5)	(12.0, 0.3, 2.7)
SL 4	(23.0, 0.4, 4.5)	(14.5, 0.3, 2.7)
SL 5	(26.6, 4.0, 4.5)	(16.8, 2.4, 2.7)
SL 6	(26.6, 8.0, 4.5)	(16.8, 4.8, 2.7)
SR 1	(11.0, 19.6, 4.5)	(6.9, 11.8, 2.7)
SR 2	(15.0, 19.6, 4.5)	(9.4, 11.8, 2.7)
SR 3	(19.0, 19.6, 4.5)	(12.0, 11.8, 2.7)
SR 4	(23.0, 19.6, 4.5)	(14.5, 11.8, 2.7)
SR 5	(26.6, 16.0, 4.5)	(16.8, 9.6, 2.7)
SR 6	(26.6, 12.0, 4.5)	(16.8, 7.2, 2.7)

Table 2.4. LFE loudspeaker positions used in the commercial cinema for the configurations shown in Fig. 2.1 (all positions given in meters, LFE height = 0.4 m)

#	LFE 1	LFE 2	LFE 3	LFE 4
1, 7	(0.4, 1.0)			
2, 8	(0.4, 5.4)	(0.4, 14.6)		
3, 9	(0.4, 6.8)	(0.4, 16.0)		
4, 10	(0.4, 1.0)	(0.4, 7.0)	(0.4, 13.0)	(0.4, 19.0)
5, 11	(0.4, 7.0)	(0.4, 9.0)	(0.4, 11.0)	(0.4, 13.0)
6, 12	(0.4, 5.0)	(0.4, 15.0)	(5.0, 0.4)	(5.0, 19.6)

Table 2.5. LFE loudspeaker positions used in the dubbing theater for the configurations shown in Fig. 2.1 (all positions given in meters, LFE height = 0.4 m)

#	LFE 1	LFE 2	LFE 3	LFE 4
1, 7	(0.3, 6.0)			
2, 8	(0.3, 3.2)	(0.3, 8.9)		
3, 9	(0.3, 4.1)	(0.3, 9.6)		
4, 10	(0.3, 0.6)	(0.3, 4.2)	(0.3, 7.8)	(0.3, 11.4)
5, 11	(0.3, 4.2)	(0.3, 5.4)	(0.3, 6.6)	(0.3, 7.8)
6, 12	(0.3, 3.0)	(0.3, 9.0)	(3.2, 0.3)	(3.2, 11.8)

floor. These heights were chosen based on published recommendations [27, 28] and were proportionally scaled for the shorter height of the dubbing theater. All measurements were taken at a height of 1.6 m, which is not perfectly in line with the SMPTE standard [3], 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 [10].

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 at (10 m, 6 m), roughly two-thirds of the room length from the screen and uses the same spacing distances as all other configurations. This area covers 10 m^2, as compared to the 112 m^2 covered by the 81-point listening grid.
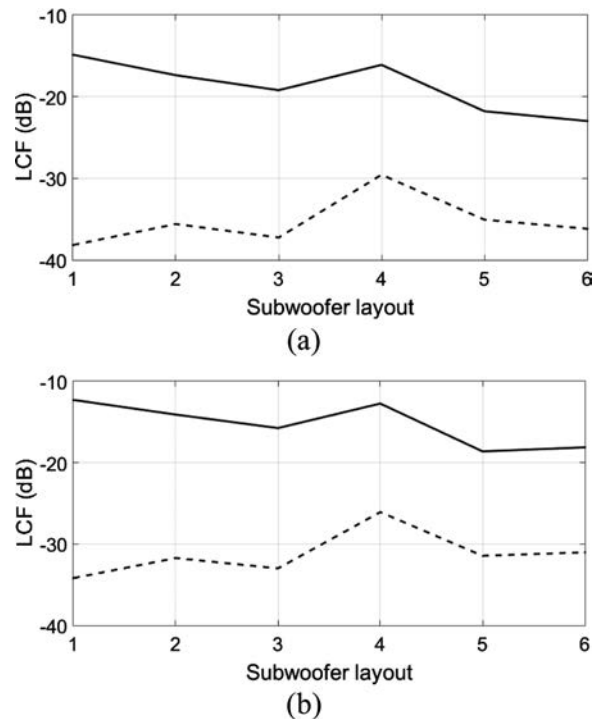

Fig. 2.2. Loudspeaker layout coupling factor for the 12 configurations identified in Fig. 2.1 (solid line = subwoofers only, dotted line = all loudspeakers) for the (a) commercial cinema and (b) dubbing theater.

This additional configuration highlights the effect of listening area size on the effectiveness of calibration strategies.

The LCFs for the commercial cinema and dubbing theater under examination were calculated for each loudspeaker configuration using Eqs.(1.6) and (1.7) and are shown in Fig. 2.2.

As expected, a single central subwoofer (Configuration 1) has the highest LCF. If the subwoofer was moved to a room corner (on the floor), the LCF would equal 0 dB, indicating maximum coupling to all room resonances. As more sources are added and distributed further apart (moving from Configurations 1 to 6) LCF decreases, indicating more effective degrees of freedom available for system calibration.

Configurations 7 to 12 are shown in Fig. 2.2 as dotted lines. The additional 15 loudspeakers dramatically reduce the LCF, providing a significant increase in calibration capability.

3 CURRENT CALIBRATION STRATEGIES

The bulk of published literature on B-chain calibration recommends using one-third-octave real-time analysis (RTA) with a graphic or parametric equalizer used to implement corrections. The current calibration procedure recommends a centrally-located measurement point, positioned two-thirds of the room length away from the screen. In some situations, multiple measurement points are used to estimate a spatially-averaged response, but there is no well-defined standard for this process [5].
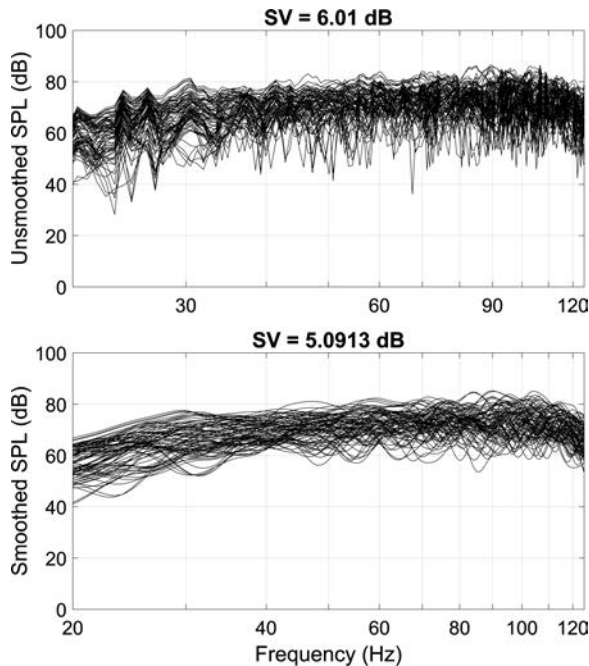
Fig 3.1. 81-point measurement grid frequency responses for configuration 1 (tight B-chain) in a cinema (top = unsmoothed, bottom = 1/12 octave smoothed).
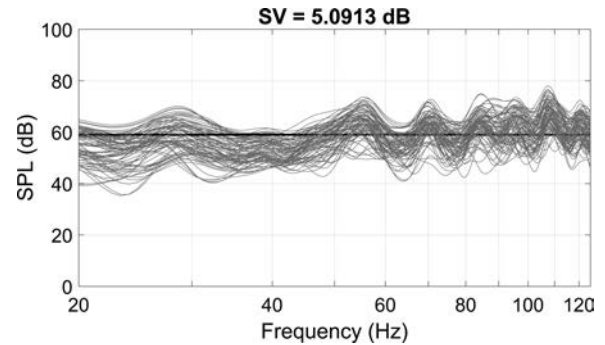

Fig 3.2. 81-point measurement grid frequency responses for configuration 1 (tight B-chain) in a cinema with single-point EQ applied (target point = solid black line).
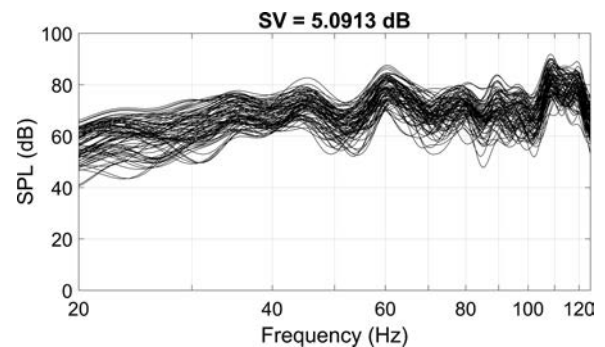

Fig 3.3. 81-point measurement grid frequency responses for configuration 1 (tight B-chain) in a cinema with spatially-averaged response EQ applied.

3.1 Single-Point Equalization

The focus of any calibration strategy must be to minimize variance across a seating area. Configuration 1 (tight B-chain) in the commercial cinema detailed in Sec. 2.1.4 was modeled using a finite-difference time-domain (FDTD) acoustic simulation toolbox [29] with a grid point spacing of 0.4 m and an 11th order MLS signal with a sample rate of 1.486 kHz (calculated to avoid spectral and spatial aliasing [9]).

The frequency responses at all 81 points are plotted in Fig. 3.1 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 [5].

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, one-third-octave band 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 [30]. The single-point equalization method results are shown in Fig. 3.2 with the target measurement point indicated by the thick black line.

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 oc-

cur 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.

3.2 Spatially-Averaged Response Equalization

Similarly, a spatially-averaged response measurement strategy can be modeled. Recently published research [7] provides an in-depth look into the effect of a number of measurement locations and their placement on the effectiveness of spatially-averaged response equalization. The results show that only eight carefully chosen measurement locations are required to generate an average response within 4 dB to 6 dB of the system response across the audience area. Particular emphasis is placed on the issue of measurement location choice, as accidentally choosing particularly poor measurement points will indicate that much more equalization is required than may actually be necessary. The work also notes that the additional effort required when using more than eight measurement points is likely to outweigh any benefits.

In the case being described in this paper, to fully sample the listening area, the frequency response at each of the 81 measurement locations is taken. The responses are then averaged to generate an inverse filter (Fig. 3.3).
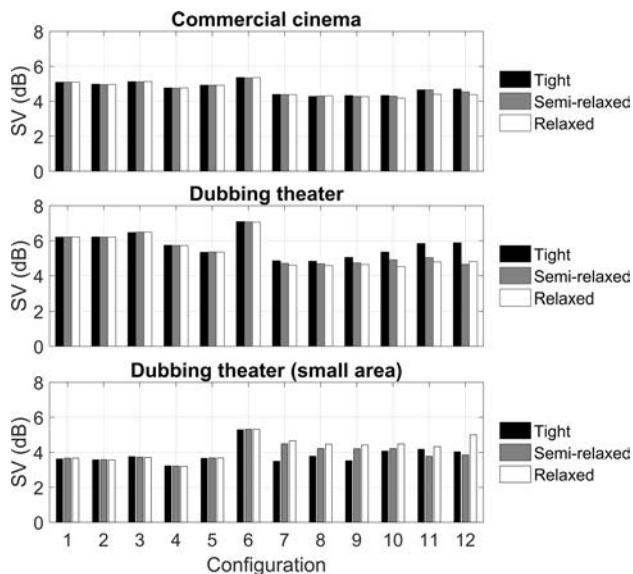
Fig 3.4. Spatial variance (SV) calculations for each B-chain configuration, as detailed in Fig. 2.1.

As with the single-point method, however, this calibration strategy provides no reduction in spatial variance, although it can provide greater overall response "flatness" using global B-chain equalization. Additionally, as mentioned in [7], measuring 81 points within the listening area can be a very long process in practice and is likely to be impractical for a reasonable calibration procedure.

3.3 Physical Configuration

As a further option, subwoofer placements can be inspected to determine if they provide any significant reduction in spatial variance across the listening area. All 12 configurations detailed in Fig. 2.1 were modeled to determine their respective spatial variances (Fig. 3.4). All B-chains were tested in the commercial cinema and dubbing theater simulation models.

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 to 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.

4 IMPROVED CALIBRATION STRATEGIES

Considering the analysis presented in the preceding three sections, it is clear that existing calibration strategies incorrectly address the issue of spatial variance in the low-

frequency band. An effective and robust strategy is required that adequately minimizes spatial variance while being simple enough to implement and maintain by a moderately-competent local technician.

4.1 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) that 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 [31–35]. 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 [12, 36, 37].

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 [9, 37]), 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 [9]. This approach takes multiple complex frequency response measurements across a listening area and constructs a set of correction filters based on the spacing of measurement points, subwoofer capabilities, and the acoustical characteristics of the room. Although the original CSA algorithm is designed to operate using so-called hybrid subwoofers (multi-drive unit devices), it has been shown that the algorithm can be applied to conventional subwoofers [38], which is the focus in this work.

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 [39]. 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 [9].

As an example of this approach, CSA calibration was applied to configuration 12 from Fig. 2.1. This configuration was chosen as CSAs are highly effective due to maximally-spaced sources (–36.19 dB LCF for this configuration). Tight B-chain characteristics were maintained, meaning that the screen and surround channels could partially contribute to sound reproduction in the low-frequency band (20 Hz to 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. 4.1).

It is important to note the effective upper frequency limit of CSA processing. This is defined by the largest dimension
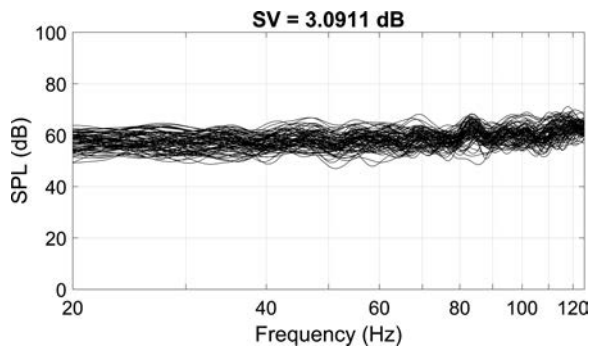
Fig. 4.1. Frequency responses at all 81 measurement points after CSA calibration applied to configuration 12 (Fig. 2.1) in a commercial cinema with tight B-chain.

of the space and the mean listener point inter-spacing (Eq. (4.1)) [9].

$$f_H = \frac{c S_p \max\left(G_x, G_y\right)}{2 L_x} \qquad (4.1)$$

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 maximum dimension of the listening grid, $max(G_x, G_y)$ (in measurement points, 9 in this case since the measurement grid is a 9×9 rectangle), and the largest dimension of the space, L_x (m). The spacing ratio, S_p, is used to determine measurement point spacing, dx and dy, as previously highlighted in Eqs. (2.2) and (2.3) in Sec. 2.1.4 [9].

Applying 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 (Sec. 4.5).

Additionally, it is important to note the calibration performance of CSAs at non-measurement points. This was studied extensively in [9], where it was found that provided measurement point spacing is not wide enough to cause spatial aliasing (points must be placed no further apart than half a wavelength of the highest frequency of interest), non-measurement points within the defined listening area equally benefit from CSA calibration.

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 spatio-temporal 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.

4.2 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, or at least maintained, by local technicians, there is a danger of incorrect implementation of the optimization algorithm or the system drifting out of calibration, resulting in sub-optimal performance. This problem can be avoided by incorporating a process capable of addressing low-frequency variance but without the need for bespoke calibration.

Diffuse signal processing (DiSP) was first described in [40] 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 multiple *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 [40]. A unique TDI is generated for each individual loudspeaker in a system, to facilitate significant signal decorrelation and thus reduce coherent interference. DiSP should therefore result in lower variance. This idea is a logical extension of the work discussed in [22], where it was found that (depending on the signals) 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 triangular probability density function along with linear coefficient interpolation, as described in [40]. As an example, the generated TDIs are 512 samples in length (1.486 kHz sample rate), the random phase values were restricted to $\pm 1.40\pi$ 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 and aid statistical similarity. A full mathematical description of this specific TDI generation process can be found in [40].

Upon inspection of one of the generated TDIs, it is clear that the initial impulse is followed by a decaying low-level noise signal (Fig. 4.2a). This is due to the phase randomization of the TDI with its characteristic frequency-dependent decay times. Summation of all 19 loudspeaker TDIs highlights the decorrelative nature of the TDI generation process, as the initial impulses (direct signal) sum constructively, while the diffuse portions of the impulse responses do not, thus fading into the noise floor (Fig. 4.2b).

Fig. 4.3. Frequency responses at all 81 measurement points after DiSP applied to configuration 12 (Fig. 2.1) in a commercial cinema with tight B-chain (top = unsmoothed, bottom = 1/12 octave smoothed).
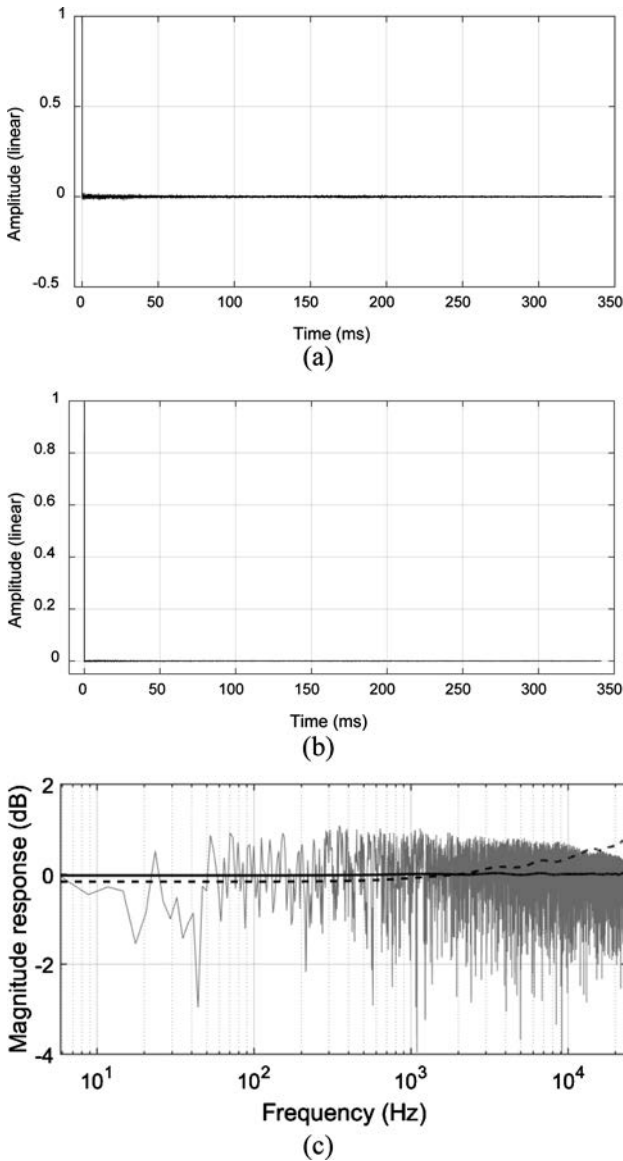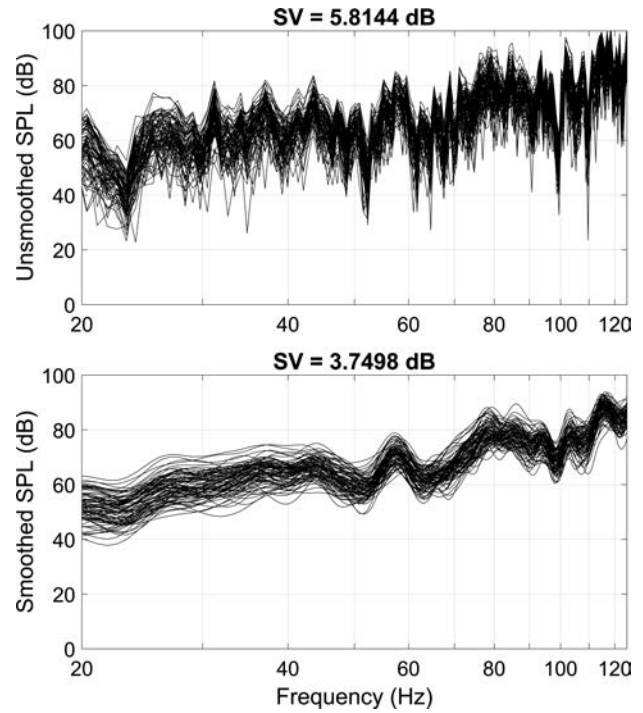

Fig. 4.2. (a) Example TDI (b) Summation of the 19 loudspeaker TDIs (c) Magnitude response of the summed TDIs (gray = unsmoothed, dotted line = smoothed, solid line = smoothed with re-equalization).

The smoothed magnitude response of the summed impulses shows a flat response after the minimum-phase components have been equalized (Fig. 4.2c).

Similar signal decorrelation methods have been previously researched [41, 42]. These methods are similar to DiSP in that they apply a form of all-pass filter to apply random frequency-dependent phase shifts to an input signal. Where these methods differ from DiSP is that they operate using a frequency-independent decay of the decorrelation filter impulse response. This delay is typically around 20 ms, whereby it is noted that a longer decay would result in signal coloration [42]. However, when dealing with low-frequency signals, a decay limit of 20 ms does not allow for adequate phase shifts. This is a strength of the DiSP approach: a frequency-dependent decay is applied to allow for larger phase shifts (if needed) at very low frequencies [40].

It is for this reason that DiSP was chosen as the method for signal decorrelation in this work.

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 high natural decorrelation of the radiated signals (Fig. 4.3), as highlighted by its LCF of –36.19 dB, one of the lowest of any tested configuration.

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 responses in this case show a 20.26% reduction in spatial variance. This example highlights the potential for DiSP use within B-chains, to allow for variance reduction without the need for calibration by local technicians.

The example presented here is meant as proof of concept. Further work must be carried out to optimize the set of TDIs for maximum effectiveness and to achieve minimum perceptible signal coloration.

4.3 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 approach.
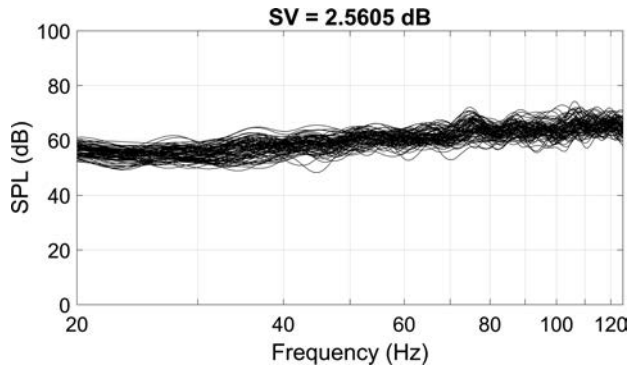
Fig. 4.4. Frequency responses at all 81 measurement points after CSA and DiSP applied to configuration 12 (Fig. 2.1) in a commercial cinema with tight B-chain.

Fig. 4.5. Configuration under examination (squares = L, C, R, diamonds = SL, SR and circles = LFE).

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. 2.1, with the resulting frequency responses shown in Fig. 4.4.

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 later, in Fig. 4.8, indicate that when utilizing less-strict system properties, the improvement due to the hybrid approach is much more pronounced.

4.4 Experimental Validation

The diffuse signal processing (DiSP) calibration approach has not been previously experimentally verified. This section presents experimental validation of DiSP for use in B-chain calibration. The chameleon subwoofer array (CSA) calibration approach has been demonstrated experimentally in previously published research [9, 38] and will not be re-examined here.

A system consisting of 3 screen (L, C, R), 2 LFE, and 6 surround loudspeakers was configured as shown in Fig. 4.5 in an 11.6 m x 10.6 m x 9.1 m space. Vertical distance to the centers of the left, center, right, surround, and LFE channels were 1.4 m, 1.0 m, 1.4 m, 1.3 m, and 0.6 m, respectively.

A 28-point measurement grid was arranged in four rows of seven, with measurement point spacing of 1.15 m and 1.10 m in the x- and y-dimensions, respectively. This spacing was chosen to avoid spatial aliasing in the measurements of the listening area. This corresponds to an upper analysis limit of 149 Hz (x-dimension) and 156 Hz (y-dimension), which adequately covers the defined LFE range in this work (20 to 125 Hz).

All measurements were taken at a height of 0.8 m. As standard B-chains call for loudspeakers to be sufficiently above listeners' heads, this lower measurement height was chosen to maintain the required relationship.
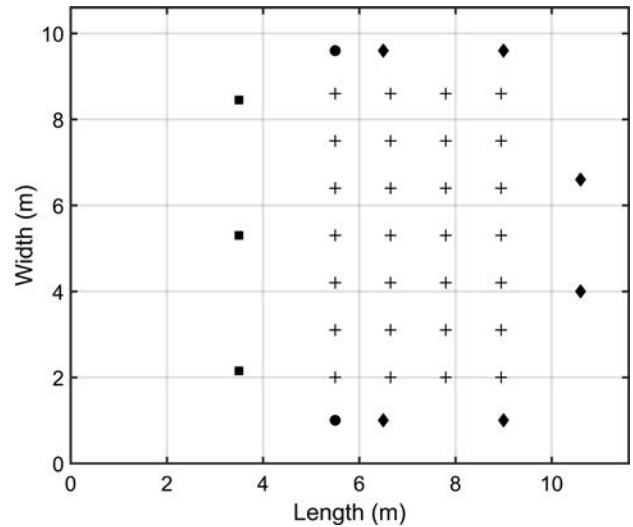
The screen loudspeakers were positioned away from the front wall due to large equipment stored in the front of the test area. This equipment was sectioned off using heavy drapery suspended from the ceiling.

TDIs were generated (as detailed in Sec. 4.2) using a triangular probability density function (PDF), 1.4 phase weighting, and time constants (controlling the frequency-dependent decay of the TDI) of 200 ms (for 20 Hz) and 15 ms (for 125 Hz). A TDI was generated for each loudspeaker (11 in total), with 8 intermediate TDIs averaged to form the final TDI.

Measurements were taken using a Clio FW-01 interface, running Clio 10 software and an Audiomatica MIC-01 measurement microphone [43]. Ten seconds of pink noise (48 kHz, 24 bit) was convolved with each loudspeaker's TDI and the resulting files were played back using Reaper [44], running into a Behringer X32 mixing console [45]. Each input was directly routed to the corresponding output of the console.

The system was first examined with only the LFE units active (two laterally positioned subwoofers, in this case). A transfer function was calculated from the measurements and the raw pink noise signal using Clio, where each measurement was plotted as a separate trace on the same figure (Fig. 4.6). Spatial variance was calculated from this data, following the procedure detailed in Sec. 1.2, showing a spatial variance reduction from 4.20 dB to 3.75 dB (a reduction of 11%).

Restricting the low-frequency calibration to only two independent channels will not result in significant reduction in spatial variance. Furthermore, the LCF for this configuration is –21.7 dB, which is only slightly lower than configuration 2 (2 subwoofers at the front of the room) from the modeled venues. The modeled configuration exhibited a spatial variance reduction of 10.800%. This is in good agreement with the measured spatial variance reduction of 10.607%.
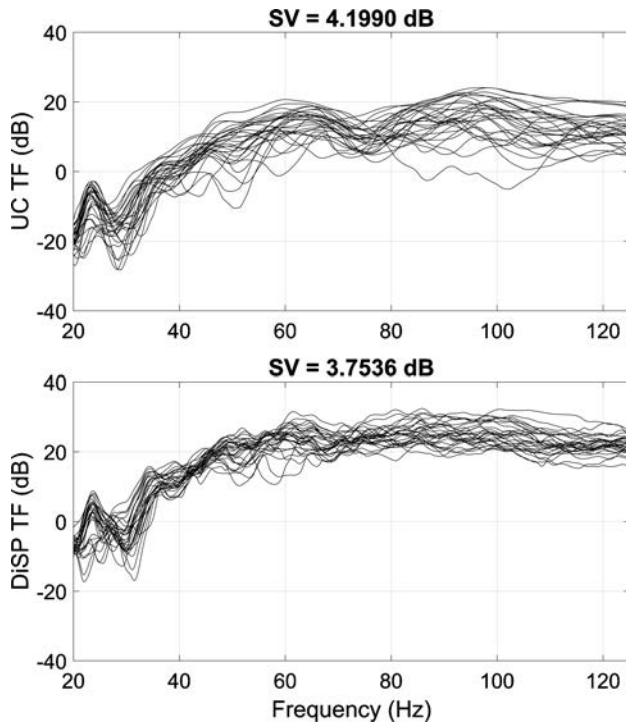
Fig. 4.6. 28-point measurement grid transfer functions from the subwoofer-only experiment for the original system (top) and the DiSP system (bottom).
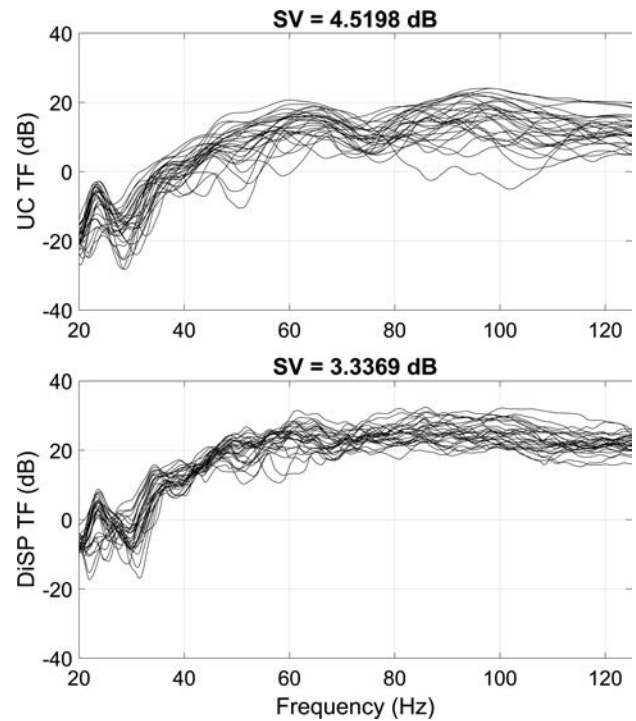

Fig. 4.7. 28-point measurement grid transfer functions from the full 11-loudspeaker system experiment for the original system (top) and the DiSP system (bottom).

The experimental system was then expanded to include all screen, surround, and LFE loudspeakers. Examining the loudspeaker layout gives an LCF of –37.4 dB, which is in agreement with the modeled configuration 12 (consisting of a similar loudspeaker layout). By this measure, the experimental system should provide spatial variance reduction that is in line with the performance of the relaxed version of configuration 12 in the model (25.260% and 23.540% reduction in the modeled cinema and dubbing theater, respectively).

The experimental system (Fig. 4.7) exhibited a reduction in spatial variance of 26.170%, which is in good agreement with the model. This verifies that DiSP is a practical option for B-chain applications, as previously indicated in the modeled venues. Further refinement of the TDI generation algorithm is likely to result in further improvements in calibration performance.

4.5 Discussion

The effectiveness of the calibration strategies over all 12 configurations is shown for the modeled commercial cinema and dubbing theater in Figs. 4.8 and 4.9. Although the modeled spaces give good indication that the investigated techniques are effective, it must be stressed that these methods must be trialed in real cinemas and dubbing theaters in order to fully verify the results. Additionally, the proposed calibration strategies must be thoroughly evaluated with in-situ listening test to ensure there is no unwanted coloration of the source signals post-calibration.

The modeled venues, however, provide a good starting point for identifying an effective and suitable approach to low-frequency calibration for B-chains.

4.5.1 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 (high LCFs, Fig. 2.2) and thus 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 and low LCF [9]. 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 as well as high LCFs. 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 to 19 sources available (depending on the configuration), spatial variance reductions approaching 55% are achieved. Unsurprisingly, the semi-relaxed and relaxed B-chains performed best in the commercial cinema (due to improved low-frequency reproduction capabilities
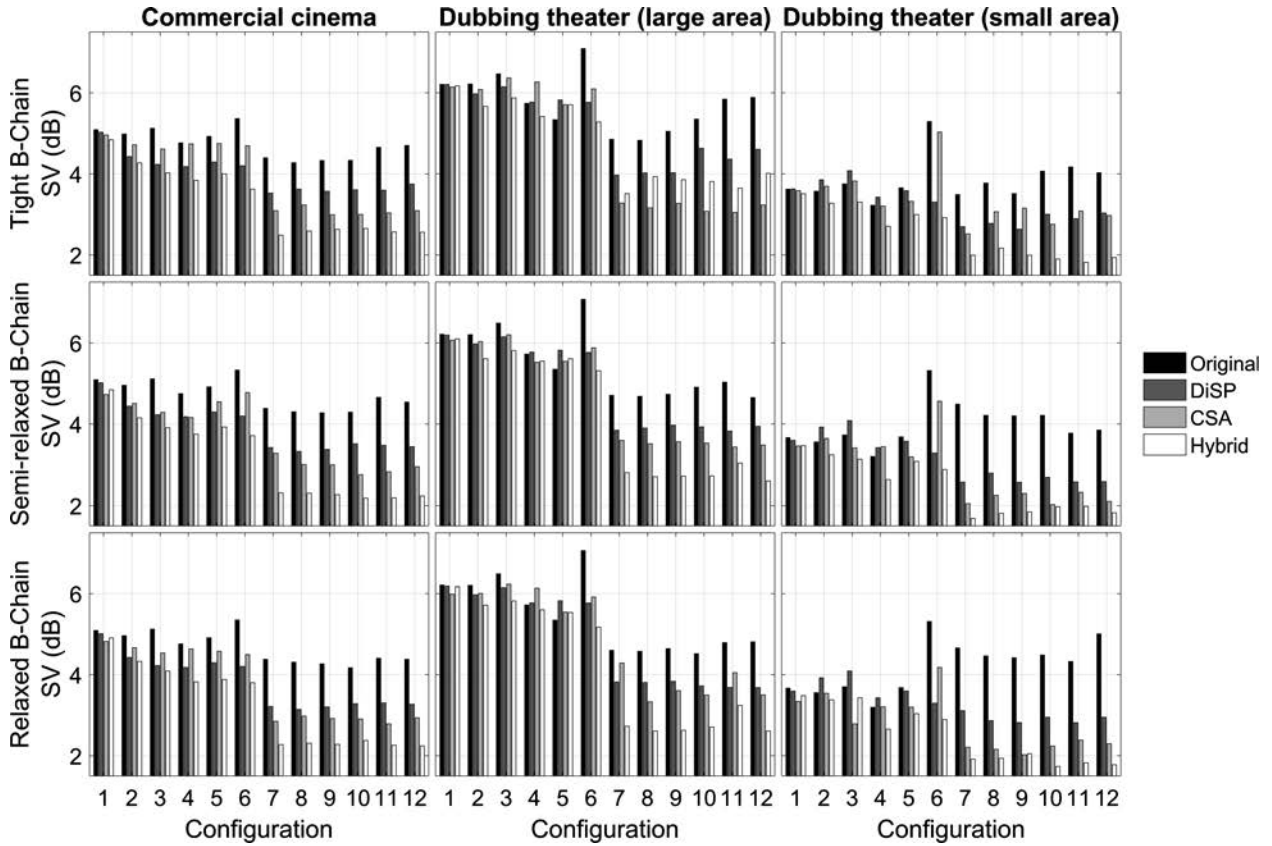
Fig. 4.8   Spatial variance (SV) due to CSA, DiSP, and hybrid calibration strategies for all 12 configurations.
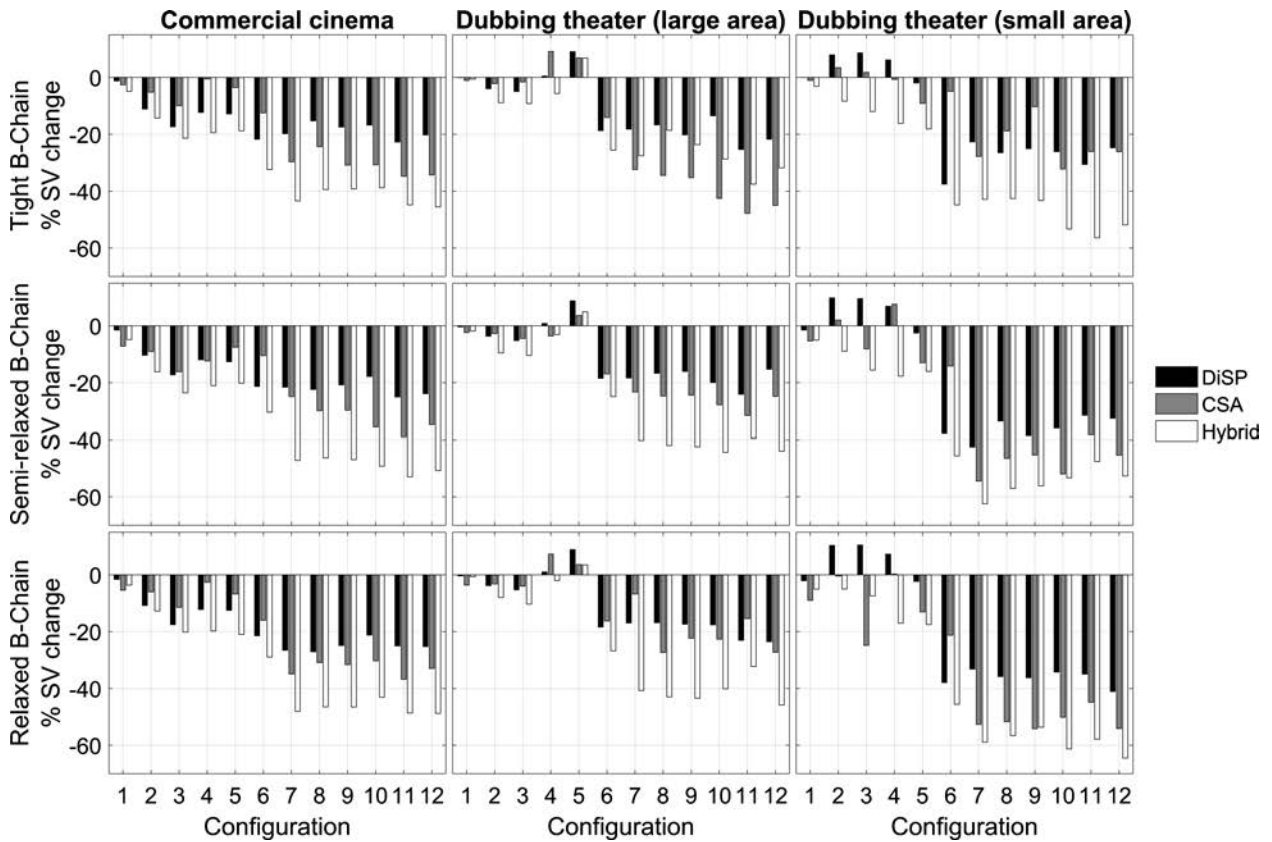

Fig. 4.9. Spatial variance reduction (in reference to spatial variance of the uncorrected systems) due to CSA, DiSP and hybrid calibration strategies for all 12 configurations.

of the non-subwoofer elements), but the opposite is seen in the dubbing theater results. This contradiction to expectation is likely due to the reduced spacing of the loudspeakers in the dubbing theater, resulting in lower source-to-source decorrelation (seen by the higher LCFs in Fig. 2.2) causing the CSA to be less effective when there is more spectral overlap between channels.

4.5.2 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 or system topology is required to implement this strategy, thus lending itself to a universally-robust solution for B-chain set up.

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 [46] where steady-state based equalization is shown to negatively impact the direct sound from the sources in the form of clearly noticeable coloration.

4.5.3 Hybrid Approach Performance

The combination of the CSA and DiSP calibration approaches into a hybrid strategy should, in theory, allow the CSA calibration to exhibit greater effectiveness, since the signal decorrelation provided by the DiSP allows for each channel to have greater independence from the others. This strengthens each degree of freedom available to the calibration strategy.

The results in Figs. 4.8 and 4.9 indicate that this is indeed the case, whereby the hybrid approach reduces spatial variance by nearly 50% in the commercial cinema, 40% in the dubbing theater with a wide listening area, and over 60% in the dubbing theater with the small listening area.

Again, all the caveats of the CSA and DiSP calibration strategies remain, but should a system prove exceptionally problematic, then a hybrid approach of this sort may be a reasonable solution, due to its ability to provide stronger degrees of freedom for low-frequency variance reduction.

4.5.4 Calibration Strategy Efficiency

As highlighted in Sec. 1.2, a calibration strategy that results in a uniform response across a wide audience area, but with a significantly reduced mean output level (MOL), is an unacceptable solution due to inefficient sound reproduction.

The proposed calibration strategies (CSA, DiSP, and hybrid) as well as the current spatially-averaged equalization strategy were examined for MOL over all tested venues.

Results for the relaxed B-chain configuration are shown as this is most closely in line with the system characteristics of installed B-chains (Fig. 4.10).

It is evident that the CSA calibration strategy on its own suffers from poor efficiency in many scenarios. This has been investigated in detail in [9] and is largely attributed to poor choice of measurement and loudspeaker locations. With a more careful physical layout, the issue can be mitigated [9]. A similar issue pertaining to cinema calibration is highlighted in [7]. CSA calibration is particularly inefficient in the dubbing theater with a small, 9-point measurement grid. In this case a poor measurement location will have a much more drastic effect on the calibration efficiency than with one poor location out of an 81-point grid.

Critically, however, the addition of DiSP to CSA to form the hybrid approach removes such sensitivities and MOL is brought back to pre-calibration levels in nearly all cases. This lends good support to the hybrid approach, as the measurement point location sensitivities highlighted in [7, 9] are no longer an issue. Furthermore, upon inspection of the MOL for the spatially-averaged equalization, it can be seen in Fig. 4.10 that this strategy offers no advantage over the hybrid approach and, in fact, provides poorer efficiency in most modeled cases.

5 RECOMMENDATIONS

The research presented in the preceding sections highlights the flaws inherent with current low-frequency optimization strategies in cinema theaters, as well as suggesting alternative, effective, and practical calibration strategies that are directly in reference to a clearly defined problem (Sec. 1).

A set of recommendations can be assembled with the aim of informing future standards and recommendations for B-chain specification and calibration:

1) The low-frequency response of B-chains must be calibrated in reference to a well-defined metric (or collection of metrics), such as spatial variance reduction and mean output level. A well-defined 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) Regardless of the adopted calibration strategy, the approach must be designed with practicality in mind. Local technicians should 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.

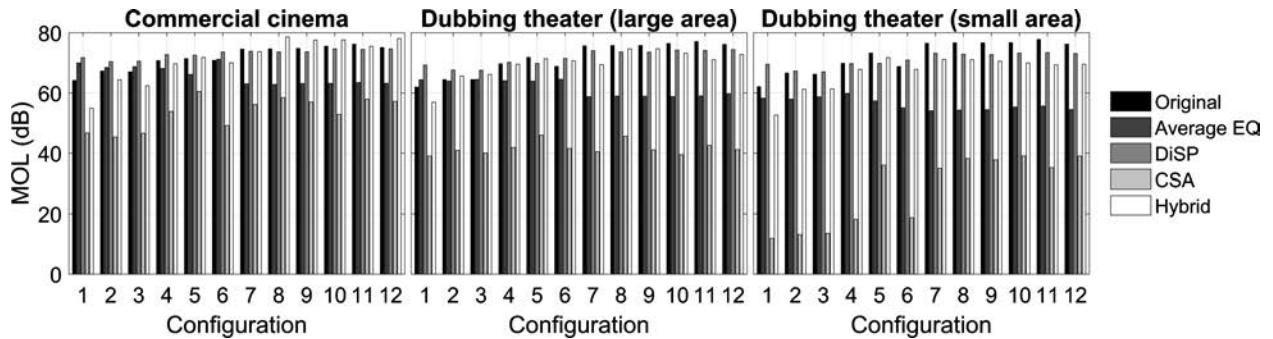There are three identified options for calibration:

Fig. 4.10. Mean output level (MOL) due to spatially-averaged EQ, CSA, DiSP, and hybrid calibration strategies.

1) Optimization routine

This method requires a series of precise measurements in order to generate a set of loudspeaker correction filters. When applied correctly, this approach can achieve extremely low spatio-temporal variance, but this must be weighed against practicality (precise calibration and maintenance required).

2) Diffuse signal processing

This method does not require on-site calibration of any sort. Loudspeaker correction filters are generated off-site and applied to the B-chain with appropriate DSP. This approach achieves moderate spatio-temporal variance reduction (at the moment), although care must be taken to avoid perceptible signal coloration, ensuring that transient and steady-state sounds maintain their intended timbre.

3) Hybrid approach

This is a combination of (1) and (2) that has the potential of achieving the most significant spatio-temporal variance reduction. Again, this approach requires detailed on-site calibration and care must be taken to avoid signal coloration.

The authors have identified diffuse signal processing as the preferred calibration method, as DiSP can be implemented universally, with no required knowledge of loudspeaker layout or venue acoustics as well as no on-site calibration or maintenance required. This approach removes the risk of local technicians allowing a system to drift out of calibration and reduces costs when installing new systems.

6 CONCLUSIONS

The current strategies for the low-frequency calibration of cinema sound systems are based on a flawed premise of low-frequency acoustics and psychoacoustics. These strategies are shown in this research to provide virtually no benefit in terms of spatio-temporal variance reduction, meaning that the pre- and post-calibrated systems will exhibit equally position-dependent listening experience differences. Furthermore, these techniques have been shown by other researchers to be highly prone to human error, resulting in inconsistent performance and often strongly colored sound due to excessive equalization [47].

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 of below 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 Sec. 4. One possibility is to use an optimization algorithm, based on multiple measurements over a seating area. The included example exhibits spatial variance reduction of nearly 55%, assuming a sufficient number of degrees of freedom (i.e., available loudspeakers for low-frequency sound reproduction). However, 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 40% (and 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 65% in certain scenarios).

The current strategies for the generation and calibration of low-frequencies in cinema theaters are unable to reduce the degree of seat-to-seat variations of response over large seating areas. Furthermore, the spatial-averaging techniques that are used for measurement and equalization are by no means always subjectively beneficial. There are far too many limitations inherent in the current practices to allow for much improvement without a significant new approach to the actual generation of the sound being considered. It is the intention of the authors that the research presented in this paper will prompt an informed discussion regarding the revision of the current concepts of

low-frequency generation in cinema theaters in order to achieve a subjectively-improved and more consistent performance.

REFERENCES

[1] J. Newell P. Newell and K. Holland "Room Low-Frequency Response Estimation Using Microphone Averaging," *Proc. IOA – Reproduced Sound 2014*, vol. 36, pt. 3 (2014).

[2] J. Newell and P. Newell K. Holland "Considerations for the Generation and Measurement of Low Frequency Effects in Cinema Rooms" presented at the *AES 57th International Conference: The Future of Audio Entertainment Technology—Cinema, Television, and the Internet* (2015 Mar.), conference paper 2-2.

[3] SMPTE Standard for Motion Pictures – Dubbing Stages (Mixing Rooms), Screening Rooms and Indoor Theaters – B-chain Electroacoustic Response," SMPTE ST 202:2010 (Oct. 20, 2010). http://dx.doi.org/10.5594/SMPTE.ST202.2010

[4] "SMPTE Recommended Practice – Relative and Absolute Sound Pressure Levels for Motion-Picture Multichannel Sound Systems – Applicable for Analog Photographic Film Aaudio, Digital Photographic Film Audio and D-Cinema," SMPTE RP 200:2012 (Apr. 18, 2012). http://dx.doi.org/10.5594/SMPTE.RP200.2012

[5] P. Newell G. Leembruggen K. Holland J. Newell S. Torres Guijarro D. Gilfillan D.S. Dominguez, and S. Castro "Does 1/3rd Octave Equalization Improve the Sound in a Typical Cinema?" *Proc. IOA – Reproduced Sound 2011*, vol. 33, pt. 6 (2011).

[6] "TC-25CSS B-chain Frequency and Temporal Response Analysis of Theatres and Dubbing Stages," SMPTE theatre testing data report B-chain study group (Oct. 1, 2014).

[7] G. Cengarle and   T. Mateos "Effect of Microphone Number and Positioning on the Average of Frequency Responses in Cinema Calibration," *J. Audio Eng. Soc.*, vol. 63, pp. 777–785 (2015 Oct.). http://dx.doi.org/10.17743/jaes.2015.0073

[8] F. E. Toole *Sound Reproduction: Loudspeakers and Rooms*, pp.385–389 (Focal Press, New York, 2008).

[9] A. J. Hill "*Analysis, Modeling and Wide-Area Spatiotemporal Control of Low-Frequency Sound Reproduction,*" *Ph.D. Thesis*, University of Essex, UK (2012 Jan.). http://derby.openrepository.com/derby/handle/10545/230034

[10] P. Newell D. Santos-Dominguez K. Holland S. Torres-Guijarro and S. Castro "The Relationship between Subjective and Objective Response Differences at Different Heights above Cinema Seating," *Proc. IOA – Reproduced Sound 2015*, vol. 37, pt. 4 (2015).

[11] A. J. Hill and M. O. J Hawksford "Low-Frequency Temporal Accuracy of Small-Room Sound Reproduction," presented at the *133rd Convention of the Audio Engineering Society* (2012 Oct.), convention paper 8715.

[12] S. Linkwitz "Investigation of Sound Quality Differences between Monopolar and Dipolar Woofers in Small

Rooms," presented at the *105th Convention of the Audio Engineering Society* (1998 Sep.), convention paper 4786.

[13] M. R. Schroeder "The 'Schroeder Frequency' Revisited," *J. Acous. Soc. Amer.*, vol. 99, no. 5, pp. 3240–3241 (1996 May). http://dx.doi.org/10.1121/1.414868

[14] P. Newell K. Holland J. Newell and B. Leskov "New Proposals for the Calibration of Sound in Cinema Rooms," presented at the *130th Convention of the Audio Engineering Society* (2011 May), convention paper 8383.

[15] J. Vanderkooy "The Low-Frequency Acoustic Centre: Measurement, Theory and Application," presented at the *128th Convention of the Audio Engineering Society* (2010 May), convention paper 7992.

[16] T. Welti "Optimal Configurations for Subwoofers in Rooms Considering Seat to Seat Variation and Low Frequency Efficiency," presented at the *133rd Convention of the Audio Engineering Society* (2012 Oct.), convention paper 8748.

[17] T. Welti "Low-Frequency Optimization Using Multiple Subwoofers," *J. Audio Eng. Soc.*, vol. 54, pp. 347–364 (2006 May).

[18] H. Kuttruff *Room Acoustics*, *5th Edition* (Spoon Press, London, 2009).

[19] S. E. Olive P. L. Schuck J. G. Ryan S. L. Sally and M. E. Bonneville "The Detection Thresholds of Resonances at Low Frequencies," *J. Audio Eng. Soc.*, vol. 45, no. 3, pp. 116–128 (1997 Mar.).

[20] A. J. Hill and 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).

[21] J. Schnupp I. Nelken, and A. King *Auditory Neuroscience: Making Sense of Sound*, Chapter 5: Neural Basis of Sound Localization (MIT Press, Cambridge, MA, USA, 2011).

[22] A. J. Hill and M. O. J. Hawksford "On the Perceptual Advantage of Stereo Subwoofer Systems in Live Sound Reinforcement," presented at the *135th Convention of the Audio Engineering Society* (2013 Oct.), convention paper 8970.

[23] P. Newell K. Holland S. Torres Guijarro, J. Newell, and D. S. Dominguez "Human Factors Affecting the Acoustic Measurement of Rooms," *Proc. IOA – Reproduced Sound 2012*, vol. 34, pt. 4 (2012).

[24] W. L. Martens "The Impact of Decorrelated Low-Frequency Reproduction on Auditory Spatial Imagery: Are Two Subwoofers Better than One?" presented at the *AES 16th International Conference: Spatial Sound Reproduction* (1999 Mar.), conference paper 16-006.

[25] W. L. Martens "Subjective Evaluation of Auditory Spatial Imagery Associated with Decorrelated Subwoofer Signals," *Proc. International Conference on Auditory Displays* (2002 July).

[26] G. S. Kendall "The Effects of Multi-Channel Signal Decorrelation in Audio Reproduction," *Proc. International Computer Music Conference* (1994 Sep.).

[27] I. Allen "Technical Guidelines for Dolby Stereo Theatres," *Dolby Laboratories*, rev. 1.33 (1994 Nov.).

[28] *Cinema Sound System Manual,* JBL Professional (2003).

[29] A. J. Hill and M. O. J. Hawksford "Visualization and Analysis Tools for Low-Frequency Propagation in a Generalized 3D Acoustic Space," *J. Audio Eng. Soc.*, vol. 59, pp. 321–337 (2011 May).

[30] R. P. Genereux, "Adaptive Loudspeaker Systems: Correcting for the Acoustic Environment," *AES 8th International Conference: The Sound of Audio* (1990 May), conference paper 8-032.

[31] J. Kuriyama and Y. Furukawa "Adaptive Loudspeaker System," *J. Audio Eng. Soc.*, vol. 37, pp. 919–926 (1989 Nov.).

[32] S. J. Elliott and P. A. Nelson "Multiple-Point Equalization in a Room Using Adaptive Digital Filters," *J. Audio Eng. Soc.*, vol. 37, pp. 899–907 (1989 Nov.).

[33] S. Bharitkar and C. Kyriakakis "Comparison between Time Delay Based and Non-Uniform Phase Based Equalization for Multichannel Loudspeaker-Room Responses," presented at the *119th Convention of the Audio Engineering Society* (2005 Oct.), convention paper 6607.

[34] A. O. Santillana "Spatially Extended Sound Equalization in Rectangular Rooms," *J. Acous. Soc. Amer.*, vol. 110, no. 4, pp. 1989–1997 (2001 Oct.). http://dx.doi.org/10.1121/1.1401740

[35] S. J. Elliot L. P. Bhatia D. W. Meghan A. H. Fu M. S. Stewart, and D. W. Wilson "Practical Implementation of Low-Frequency Equalization Using Adaptive Digital Filters," *J. Audio Eng. Soc.*, vol. 42, pp. 988–998 (1994 Dec.).

[36] H. F. Olson "Gradient Loudspeakers," *J. Audio Eng. Soc.*, vol. 21, pp. 86–93 (1973 Mar.).

[37] J. Backman "Low-Frequency Polar Pattern Control for Improved In-Room Response," presented at the *115th Convention of the Audio Engineering Society* (2003 Oct.), convention paper 5867.

[38] A. J. Hill and M. O. J. Hawksford "Practical Applications of Chameleon Subwoofer Arrays," presented at the *132nd Convention of the Audio Engineering Society* (2012 Apr.), convention paper 8665.

[39] J. Vanderkooy "Multi-Source Room Equalization: Reducing Room Resonances," presented at the *123rd Convention of the Audio Engineering Society* (2007 Oct.), convention paper 7262.

[40] M. O. J. Hawksford and N. Harris "Diffuse Signal Processing and Acoustic Source Characterization for Applications in Synthetic Loudspeaker Arrays," presented at the *112th Convention of the Audio Engineering Society* (2002 May), convention paper 5612.

[41] G. S. Kendall "The Effects of Multi-Channel Signal Decorrelation in Audio Reproduction," *Proc. International Computer Music Conference, 1994*, pp. 319–326 (1994).

[42] M. D. Wilde W. L. Martens and G. S. Kendall "Method and Apparatus for Creating De-Correlated Audio Output Signals and Audio Recordings Made Thereby," United States Patent #5,235,646, August 10, 1993.

[43] Srl Audiomatica Clio 10 FW and MIC-01, www.audiomatica.com (2015).

[44] Cockos Incorporated, *Reaper Digital Audio Workstation*. www.reaper.fm

[45] Music Group. Behringer X32, www.music-group.com/Categories/Behringer/Mixers/Digital-Mixers/X32/p/P0ASF

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

[47] P. Newell K. Holland M. Desborough B. Fazenda B. Leskov S. Castro E. Valdigem, and 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).

THE AUTHORS

Adam J. Hill   Malcolm O.J. Hawksford   Philip Newell

Adam Hill is currently a lecturer in electrical and electronic engineering at the University of Derby where he runs the MSc Audio Engineering program. He received a Ph.D. from the University of Essex, an M.Sc. with Distinction in acoustics and music technology from the University of Edinburgh, and a B.S.E. in electrical engineering from Miami University. His doctoral research focused on analysis, modeling, and wide-area spatiotemporal control of low-frequency sound reproduction. Other research has included voice coil temperature effects in loudspeakers, acoustically-based classification of coral reef health, and real-time equestrian horse tracking via GPS/IMU integration. Adam also works professionally as a live sound engineer for Gand Concert Sound and is a member of the AES, IOA, IEEE, and IET.

•

Malcolm Hawksford was educated at Aston University in Birmingham, UK, where he received his B.Sc. with first-class honors in 1968, Ph.D. in 1972, and D.Sc. in 2008. He is currently Emeritus Professor within the School of Computing Science and Electronic Engineering at Essex University, Colchester, UK. Early research embraced delta- and sigma–delta modulation (SDM) applied to color TV coding that under the award of a BBC Research Scholarship lead to a method of luminance and chrominance multiplexing exploiting digital time-compression, a forerunner of MAC/DMAC. Current interests include audio engineering,

electronic circuit design, and signal processing focused on loudspeakers, SDM, PWM linearization, spatial audio, and telepresence. Malcolm was recently awarded the Peter Barnett Memorial Award 2014 from the Institute of Acoustics, UK; he is also a recipient of the AES Publications Award for the best contribution by an author of any age for *JAES*, volumes 45 and 46, and holds the AES Silver Medal for major contributions to engineering research in the advancement of audio reproduction. He is a chartered engineer and Fellow of the AES, IET, and IOA. He has been chair of the AES Technical Committee on High-Resolution Audio and was a founding member of Acoustic Renaissance for Audio (ARA).

•

Philip Newell is a consultant in acoustics and electroacoustics who has been designing sound recording studios and performance spaces for over 40 years. He was the technical director of the Virgin Records studio division for 11 years until 1982. Philip has published over 50 papers on electroacoustics-related subjects in journals and conference proceedings and has also written seven books published by Focal Press. He is a Fellow of the Institute of Acoustics, a Member of the Audio Engineering Society, and a Member of the Society of Motion Picture and Television Engineers. As an SMPTE Member he is currently involved in the Technology Committee looking into the whole question of the modernization of cinema sound specification and calibration.