NATURALNESS OF DOUBLE-SLOPE DECAY IN GENERALISED ACTIVE ACOUSTIC ENHANCEMENT SYSTEMS

Will J. Cassidy[1], *Philip Coleman*[2], *Russell Mason*[1], *Enzo De Sena*[1]

[1]Institute of Sound Recording
University of Surrey
Guildford, UK
w.cassidy@surrey.ac.uk | r.mason@surrey.ac.uk | e.desena@surrey.ac.uk

[2]L-Acoustics
67 Southwood Lane, Highgate, London N6 5EG
phil.coleman@l-acoustics.com

ABSTRACT

Active acoustic enhancement systems (AAESs) alter the perceived acoustics of a space by using microphones and loudspeakers to introduce sound energy into the room. Double-sloped energy decay may be observed in these systems. However, it is unclear as to which conditions lead to this effect, and to what extent double sloping reduces the perceived naturalness of the reverberation compared to Sabine decay. This paper uses simulated combinations of AAES parameters to identify which cases affect the objective curvature of the energy decay. A subjective test with trained listeners assessed the naturalness of these conditions. Using an AAES model, room impulse responses were generated for varying room dimensions, absorption coefficients, channel counts, system loop gains and reverberation times (RTs) of the artificial reverberator. The objective double sloping was strongly correlated to the ratio between the reverberator and passive room RTs, but parameters such as absorption and room size did not have a profound effect on curvature. It was found that double sloping significantly reduced the perceived naturalness of the reverberation, especially when the reverberator RT was greater than two times that of the passive room. Double sloping had more effect on the naturalness ratings when subjects listened to a more absorptive passive room, and also when using speech rather than transient stimuli. Lowering the loop gain by 9 dB increased the naturalness of the double-sloped stimuli, where some were rated as significantly more natural than the Sabine decay stimuli from the passive room.

1. INTRODUCTION

When a sound source is recorded in one room and reproduced over loudspeakers in another, both spaces will contribute to the resultant acoustics perception [1]. The system impulse response (IR) is often described as a convolution of the recording and reproduction room impulse responses (RIRs) [1] [2]. Such *room-in-room* effects can occur in physical acoustics, sound reinforcement, and sound reproduction.

Sound energy in a room can be expected to decay exponentially if it has uniformly-distributed absorption [3], as well as sufficient reflective randomisation due to the room geometry and/or diffusion [4] [5]. However, the acoustic energy decay of a space can consist of multiple exponential curves known as *multi-exponential decay*, or the *double-slope effect* (DSE) [6]. This can occur in passive acoustics, such as coupled volumes [7] and rooms where the absorptive materials are non-uniformly distributed [8]. It can also occur in active electroacoustic systems when artificial reverberation is utilised.

Active acoustic enhancement systems (AAESs) introduce sound energy into a room using microphones and loudspeakers to alter the perceived acoustics of the space [9]. The microphone signals are sent to the loudspeakers via a processing unit, which may be as simple as a channel mixing matrix, or artificial reflections may be introduced for more control over perceptual characteristics of the space. Compared to passive variable acoustics, which requires mechanically variable structures, active acoustics can provide a more flexible and cost-effective room enhancement solution [10].

In-line active acoustic systems such as the LARES [11] and SIAP [12] utilise microphones which are in the direct field of the sources, for example, over the stage of a concert hall [10]. The microphone signals are sent through artificial reverberation processing (which may consist only of an early reflections generator [13]) before being reproduced by loudspeakers to distribute the energy evenly across the audience [13]. By physically separating the microphones and loudspeakers and by exploiting their directionalities, the loudspeaker-to-microphone feedback component is minimised, and hence the risk of instability is kept low [14]. Thus, the system acts in a primarily feed-forward manner.

By placing microphones in an AAES near the loudspeakers they are routed to, the chance of feedback increases. This principle can be used to reintroduce energy into the system, known as *regeneration*. To avoid severe resonance, the microphones are typically placed in the reverberant field of each loudspeaker, which ensures the input to the system is diffuse. Instead of relying on artificial reverberation, regenerative AAESs such as the MCR system [15] use the passive space for diffusion. An issue with these systems is that a high channel count is required to achieve a satisfactory increase in reverberation time (RT) without colouration, since using more channels reduces the mean loop gain required to achieve the same power output [11]. At lower channel counts, artificial reverberation can be introduced for regeneration support as used in the Constellation (based on VRAS [16]) and Ambiance [17] systems. This allows the RT to be extended without increasing the loop gain, keeping the system sufficiently below the limit of stability to avoid colouration [18] [19]. However, these systems may exhibit the DSE in certain conditions as a result of the artificial reverberation, so the perceptual impact of this effect should be investigated.

Copyright: © 2024 Will J. Cassidy[1] et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author and source are credited.

The perception of the DSE has been investigated in the context of coupled volumes, e.g. in terms of just-noticeable differences [20], preference [21], suitability for classical stimuli [22], and clarity and reverberance [6]. Bradley and Wang [6] found that aspects related to clarity were not significantly affected. Ermann [21] had insufficient evidence to suggest that subjects preferred Sabine decays to double sloping, but Luizard et al. [22] found that double-slope decay was preferred for solo instruments and choirs, and Sabine decay better suited a symphony orchestra. From these studies, it is still not clear as to whether the DSE is perceived as unnatural or not, and it has since been questioned why a single exponential decay seems to be expected in auditoria [7].

Listeners in concert halls can be put off by the presence of an active system (even if the system is turned off [18]). Hence, the naturalness of a room using an active system is an important metric to consider, but double sloping has not been extensively explored before in the context of AAESs. Existing perceptual evaluation of AAESs has involved both the musicians and the audience, and naturalness has been a topic of investigation [23] [24]. However, prior work regarding naturalness has yet to quantify the amount of double sloping and its influence on the listeners.

In order to investigate the perception of the DSE in AAESs, the system conditions in which this effect occurs were first identified. Then, a listening test was run to evaluate the naturalness of these conditions for varying degrees of double sloping. The aim of this listening test was to determine if the DSE caused by a regenerative AAES using artificial reverberation reduces the naturalness of the perceived acoustics.

This paper is organised as follows. Section 2 presents and validates a regenerative model used for simulating a closed-loop AAES IR. A parameterised simulation based on a room acoustic model is detailed in Section 3. The simulated RIRs are then analysed objectively in Section 4, motivating a listening test based on naturalness outlined in Section 5. The results of this listening test are presented in Section 5.3 and are discussed with the objective results in Section 6. The conclusions of the paper are then summarised in Section 7 with intentions for further work.

2. REGENERATIVE MODEL

In order to explore the effects of regeneration in AAESs, Poletti's general enhancement model [25] was implemented in MATLAB (available on GitHub[1]). Given measured transfer functions between each transducer in the system, as well as the reverberation matrix, the model can calculate a prediction of the closed-loop signal received at one or more points in the room for one or more arbitrary sound sources. This model assumes that the AAES is a linear time-invariant (LTI) system.

2.1. Signal Paths

Figure 1 represents a room in which a general AAES is installed. This system consists of L microphones, K loudspeakers, N room sources and M receivers. The AAES microphone signals are fed to an artificial reverberator defined by the transfer function matrix, $\boldsymbol{X}(z) \in \mathbb{C}^{K\times L}$, scaled by a (frequency independent) loop gain coefficient, $\mu \in \mathbb{R}$. The output transfer function vector, $\boldsymbol{v}(z) \in \mathbb{C}^{M}$, can be defined in terms of the input vector, $\boldsymbol{u}(z) \in \mathbb{C}^{N}$, the reverberator matrix, $\mu\boldsymbol{X}(z)$, and the transfer functions from each

[1] `https://github.com/IoSR-Surrey/AAESToolbox`

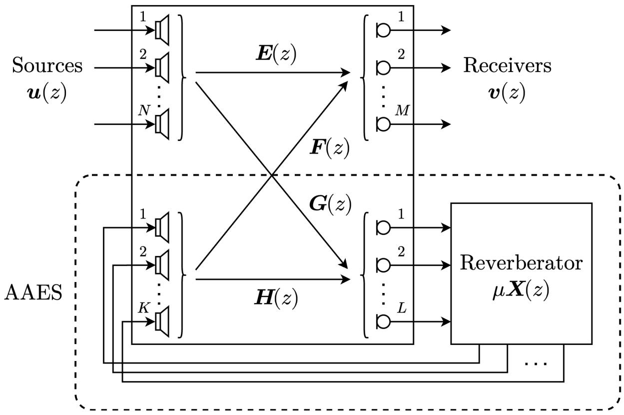

Figure 1: *The transfer function paths for a general AAES using artificial reverberation in a room, based on Poletti's general enhancement model [25]. The active system is labelled "AAES".*

loudspeaker to each microphone of the system, shown on Figure 1 as matrices $\boldsymbol{E}(z) \in \mathbb{C}^{M\times N}$, $\boldsymbol{F}(z) \in \mathbb{C}^{M\times K}$, $\boldsymbol{G}(z) \in \mathbb{C}^{L\times N}$ and $\boldsymbol{H}(z) \in \mathbb{C}^{L\times K}$. The output vector, $\boldsymbol{v}(z)$, can be expressed as a summation of the direct path of the room sources to the receivers and the propagation of the sources through the AAES:

$$\begin{aligned}\boldsymbol{v}(z) &= \boldsymbol{E}(z)\boldsymbol{u}(z)\\ &\quad + \mu\boldsymbol{F}(z)\big[\boldsymbol{I} - \mu\boldsymbol{X}(z)\boldsymbol{H}(z)\big]^{-1}\boldsymbol{X}(z)\boldsymbol{G}(z)\boldsymbol{u}(z), \quad (1)\end{aligned}$$

where $\boldsymbol{I}$ is the identity matrix of dimensions K×K.

2.2. Gain Before Instability

The stability of a multichannel LTI system with feedback can be defined by considering the open-loop transfer function matrix of the feedback loop. In Figure 1, the feedback loop of the model can be identified as the section labelled "AAES". Thus, the open-loop matrix of the feedback component can be written as $\mu\boldsymbol{X}(z)\boldsymbol{H}(z) \in \mathbb{C}^{K\times K}$. The K characteristic functions of this matrix, denoted by $\lambda_k(z) \in \mathbb{C}$ for $k = 1, ..., K$, can be obtained by performing eigenvalue decomposition across frequency [19]:

$$\mu\boldsymbol{X}(z)\boldsymbol{H}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{-1}(z), \quad (2)$$

where $\boldsymbol{\Lambda}(z) = \mathrm{diag}[\lambda_1(z), \lambda_2(z), ...\lambda_K(z)] \in \mathbb{C}^{K\times K}$ and $\boldsymbol{Q}(z) \in \mathbb{C}^{K\times K}$ contains the K eigenvectors in its columns. The LTI system will be stable if $|\lambda_k(e^{j\omega})| < 1$ for all k. It should be noted that the time variation of a real room introduces a risk of instability [26].

The mean *gain before instability* (GBI) is often calculated using the maximum value of either the real part or the magnitude of the eigenvalues, $\lambda_k(e^{j\omega})$. Poletti [26] suggested that using the real part may yield a closer value to the true limit of stability, but this paper uses the magnitude for a safer approach [19].

2.3. Validation

To validate the regenerative IR model, microphones and loudspeakers installed at the L-Acoustics Immersive Lab in London were used to generate a realtime 16-channel AAES with artificial reverberation. RIR measurements were taken between an omni-directional dodecahedron source and an omni-directional microphone with the AAES active, placed one-third and two-thirds along

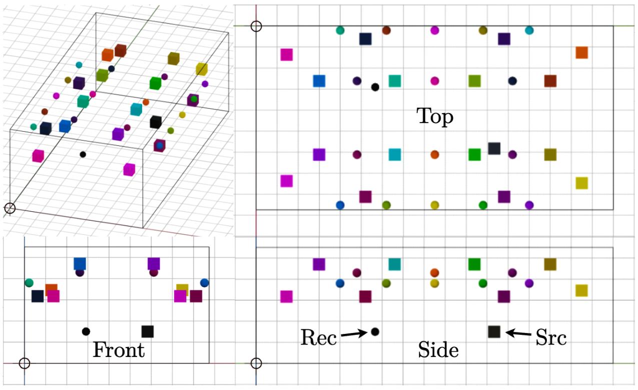

Figure 2: *The 16-channel microphone and loudspeaker layout used for this experiment where one square = 1 m^2. Small spheres represent microphones and large cubes represent loudspeakers. The colours are matched for each microphone-to-loudspeaker routing.*

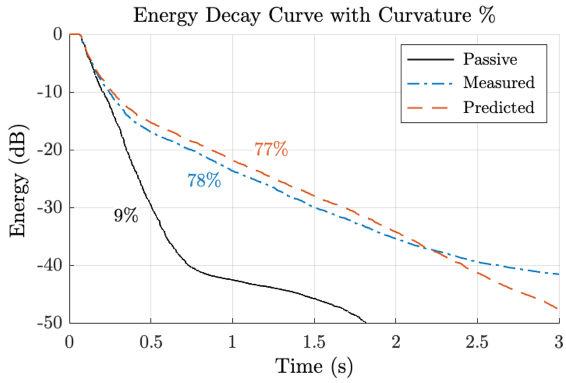

Figure 3: *The EDCs for the 1 kHz octave band of the passive room, the measured RIR with the AAES active, and the predicted RIR. The RT of the reverberator was four times that of the room, and the loop gain of the system was set to 0 dB relative to the GBI.*

the floor diagonal at approximately 1.5 metres high. This setup is depicted in Figure 2. The RIRs between all system loudspeakers and microphones were also measured, which were used to feed the regenerative model to predict the RIR of the live system.

For the purposes of validation, the reverberator ($\boldsymbol{X}(z)$ in Figure 1) was defined as a diagonal matrix with each non-zero element given by the time-domain representation

$$e^{(nT_s \ln 10^{-3})/T_{60}} W_l(nT_s), \tag{3}$$

where n is the time index, T_s is the sampling period, and $W_l(nT_s)$ represents a Gaussian white noise sequence of length $2T_{60}$, which was re-seeded for each input channel, l. This matrix represents a one-to-one mapping between the microphones and loudspeakers, depicted in Figure 2. Five values of T_{60} were chosen as multiples of the passive room RT at 1 kHz, namely 0, 0.5, 1, 2 and 4. For the RT ratio of 0, the reverberator matrix was replaced with an identity matrix, resembling the MCR system [15].

For the realtime AAES setup, the system microphones were routed through Pro Tools where each channel was convolved with its respective noise IR. An attenuation was also applied to each channel according to the worst-case GBI calculation from the model, recalculated for each noise RT ratio. Due to the conservative nature of the GBI calculation, the master gain could be set to 0 dB for each of the RT ratios without causing the system to become critically stable. At this loop gain, exponential sine sweeps were run through the dodecahedron over six seconds to capture an RIR for each of the RT ratios.

Figure 3 shows the 1 kHz octave band energy decay curves (EDCs) for the measured and predicted RIRs. This octave band is used throughout this paper (as performed by Jagla and Chervin [27]) to disregard curvature due to resonance of low frequency energy. This curvature makes it difficult to distinguish between the DSE and narrowband resonance, the latter of which is not considered to be true double sloping.

The objective curvature, a metric used to determine the degree of double sloping, was compared between the EDCs in Figure 3. Curvature can be used to evaluate the change in gradient of an EDC as a percentage, calculated as

$$c = \left| \frac{m_l}{m_e} - 1 \right| \times 100\,\%, \tag{4}$$

where m_e and m_l represent the linear gradient of the early and late regions of the EDC, respectively. This returned a value of 9 % for the passive room, 78 % for the measured RIR and 77 % for the prediction. Despite matching the curvature closely, the late energy is around 2 dB higher for the predicted decay compared to the measured system. This difference is likely due to the GBI being calculated for the predicted feedback loop (see Section 2.2), rather than using a measurement. This GBI was used as the loop gain, μ, for both the live system and the model, but the true GBI of the live system could have been influenced by the *signal-to-noise ratio* (SNR) and time-varying characteristics of $\boldsymbol{H}(z)$. The difference in curvature below -35 dB can be explained by the live system measurement having a higher SNR than the combined individual RIRs in the prediction.

3. AAES RIR SIMULATION

By using the regenerative model from Section 2 together with simulated RIRs, the parameter space of factors affecting the double slope effect can be further explored. In this section, details of a set of simulations with different AAES conditions are presented to provide an insight into which cause double sloping. These modelled conditions are then analysed objectively in Section 4 and subjectively in Section 5.3.

3.1. System Parameters

This section explains the parameters and arguments that were chosen to form a representative set of simulated AAES conditions.

Three sets of shoebox room dimensions were selected to represent small, medium and large acoustic spaces. The three sets were based on real rooms for reference, where the room volumes were approximately matched. The first room was based the Institute of Sound Recording's ITU-R BS 1116 listening room at the University of Surrey, with the dimensions 5.70×7.35×2.50 metres. The second room was based on the aforementioned L-Acoustics Immersive Lab in London, approximately 8.74×17.00×5.50 metres.

Lastly, the largest room was based on an opera house with dimensions 19.52×30.83×15.00 metres.

Three sets of absorption coefficients were selected from Vorländer [28], available in the GitHub repository mentioned in Section 2, to represent three room types: dry, moderately reflective and very reflective. The first two sets of absorption coefficients were intended to match those of TB7 (high absorption) and the L-Acoustics Immersive Lab (medium absorption). The third set was chosen to be arbitrarily more reflective to represent low absorption. Coefficients for seven octave bands were used with small randomisations to reduce the evenness of absorption around the room for a more realistic energy decay. An RIR was then generated for the corresponding room dimensions (using the room acoustic model detailed later in Section 3.2). The RTs at 1 kHz and frequency responses of the resulting RIRs were then compared to the original real rooms, and the absorption coefficients were adjusted by hand.

The chosen channel counts were 8, 12 and 16 since the benefits of using artificial reverberation are greater for AAESs with lower channel counts. The transducer positions were based on the electroacoustic installations in each of the three reference spaces, where the positions were selected to be distributed as evenly as possible. The one-to-one loudspeaker-to-microphone distances were varied to increase the randomisation of delays between transducers, demonstrated in Figure 2.

The reverberator RTs were set to multiples of the passive RT for each room condition, namely 0, 0.5, 1, 2, and 4. It has been suggested by [13] that double sloping is reduced when the reverberator RT decreases below that of the passive room, hence the inclusion of RT ratios 0, 0.5 and 1. The RT ratios of 2 and 4 were chosen to consider RTs slightly and significantly above the suggested double sloping threshold, respectively.

Four loop gains relative to the estimated GBI were used, namely 0, -2, -4 and -6 dB. This allowed the simulations to be grouped in terms of similar amounts of regeneration. The 0 dB relative loop gain was around 2 dB below the true limit of stability due to the conservative nature of the GBI calculation. In total, 540 AAES conditions were created.

3.2. Implementation

The room simulation toolbox from AKtools [29] was used to generate RIRs in MATLAB to populate $\boldsymbol{E}(z)$, $\boldsymbol{F}(z)$, $\boldsymbol{G}(z)$ and $\boldsymbol{H}(z)$ from the model in Section 2. AKtools provides a hybrid model based on the image-source method and stochastic reverberation for shoebox rooms. This open-source toolbox was chosen for the capability of specifying frequency-dependent absorption, and as a more accessible alternative to models such as CATT-acoustic. For this paper, omni-directional sources and receivers with a flat frequency response were used. Finally, closed-loop RIR simulations were generated using the model from Section 2.

To allow the generated RIRs to also be evaluated subjectively, the white noise in the reverberator was replaced with pink noise to reduce the bias caused by the timbral change of the secondary slope, while keeping the reverberator as general as possible. During the model validation session (Section 2.3), double sloping was clearly audible due to the high-frequency sustain observed when the decaying white noise dominated the reverberation. Some transducer paths were particularly short (see Figure 2) where microphones were likely in the direct field of a loudspeaker. Therefore, the frequency independence of the white noise decay resulted in little high-frequency attenuation in these paths, causing

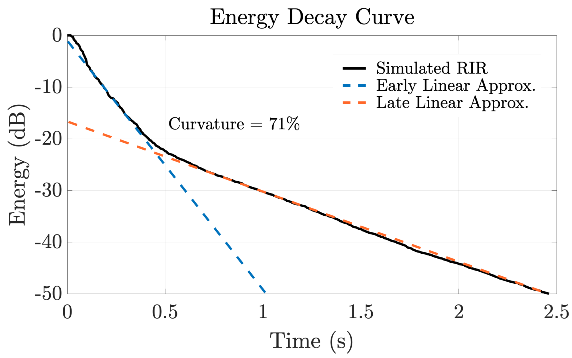

Figure 4: *1 kHz octave band curvature measurement using two linear gradient approximations for the RIR simulated with the following arguments. Channel count: 16, room size: 2, absorption: med, relative loop gain: 0 dB, RT ratio: 4.0.*

unnatural sustain. While both the white and pink noise reverberators result in frequency-dependent decay due to the room absorption, the pink noise provides extra high frequency attenuation in this feedback loop.

4. OBJECTIVE CURVATURE

The objective curvature was calculated using Eq. 4. The early decay gradient was calculated from -10 to -15 dB and the late from -30 to -50 dB. The decay level is used rather than time regions since the two gradients may vary significantly in time depending on the overall RT. These ranges were chosen to best fit the differences seen in the simulated energy decays, an example of which is shown in Figure 4. Figure 5 (a) shows the curvature for the different simulated RIRs from Section 3 for the 1 kHz octave band and for the 16-channel model at a relative loop gain of -6 dB. Figure 5 (b) shows the average of these curvatures as a function of channel counts and loop gains.

A number of observations can be made on the basis of these figures. The objective results show that the curvature metric is strongly affected by the ratio between the RT of the passive room and the RT of the artificial reverberator. It can be seen in Figure 5 (a) that for RT ratios ≤ 1, most room conditions exhibit little double sloping. This supports the hypothesis that keeping the ratio of RTs below 1 results in a single exponential decay [13]. There seems to be more curvature for low RT ratios in the smallest room ("Room 1" in Figure 5 (a)), but this is an artefact of the linear gradient measurements being sensitive to energy fluctuations over short RTs.

In the 8-channel cases, the average curvature was brought down by the smallest, most absorptive room condition showing little RT extension in the 1 kHz octave band. In this situation, the axial room modes (around 24-69 Hz) dominated the feedback loop such that the enhancement did not affect the 1 kHz octave band. In some cases, significant curvature is evident at an RT ratio of zero, e.g., at 0 dB loop gain in Figure 5 (b). In this case, room mode resonances dominate the feedback loop which become more prominent at higher loop gains (see [30, Figure 3]), causing the EDC to change gradient as these frequencies sustain beyond the average energy decay.

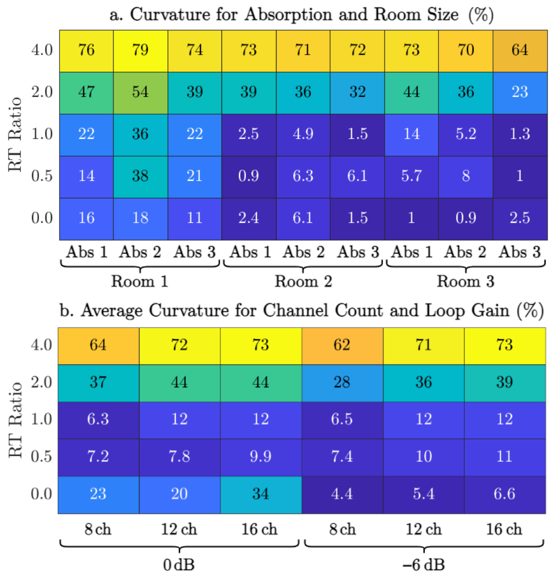

Figure 5: *Heatmap of 1 kHz octave band curvatures for (a) the 16-channel RIRs at −6 dB relative loop gain, and (b) averages across channel count and loop gain. Values closer to zero indicate less double sloping. In both figures, the vertical axis shows the RT ratio of the reverberator. The horizontal axis of (a) shows absorption and room sizes, while the one of (b) shows channel count and loop gain.*

5. LISTENING TEST

From the objective analysis, it can be seen that the RT ratio (that is, the ratio between the RT of the reverberator and the RT of the passive room) has a strong effect on the amount of double sloping that occurs across the room conditions. A listening test was designed to determine whether this objective double sloping affects the perceived naturalness of reverberation.

5.1. Stimuli

The listening test stimuli were based on simulated AAES RIRs generated using the model from Section 2 with the simulated rooms described in Section 3. To keep the duration of the listening test manageable, only the RIRs associated to the 16-channel simulation with the Immersive Lab room dimensions were used, as this configuration was most representative of a general AAES.

The main factor affecting the objective curvature was the RT ratio, so the perceptual difference was investigated for RT ratios of 0 to 4 in intervals of 0.5. After conducting an informal listening test with the IRs convolved with speech, it was found that the difference between RT ratios of 0.5 and 1 was negligible, and that the double sloping was equally as audible for RT ratios above 3. Therefore, the RT ratios were chosen as 1-3 with 0.5 increments. The direct source to receiver RIR was also included to represent the case where the active system was turned off.

Despite not having a clear effect on the curvature, the absorption seemed to influence how natural the reverberation sounded when listening informally to the generated RIRs. To provide some variation in absorption, the *high* and *medium* absorption coefficients were used to provide a difference in passive RT and frequency dependence. Similarly, the loop gain, which did have an effect on curvature in some cases, was considered to have a potential influence on naturalness and was therefore included. Loop gains of 0 dB and −9 dB relative to the calculated GBI were selected to capture the largest difference possible while ensuring double sloping was still audible.

Lastly, two programme items were selected to be convolved with the simulated RIRs to create the stimuli. Anechoic speech was taken from the University of York Openair dataset [31], and an anechoic recording of a bongo was used from the Bang & Olufsen Archimedes dataset [32]. The excerpts were then edited to around 7 seconds long, and the convolutions were performed with sufficient tail time to preserve the decay. A total of 44 monaural stimuli were created consisting of 22 combinations of absorption, loop gain and RT ratio, for each of the programme items.

5.2. Methodology

A listening test was made to compare the above 44 stimuli over 16 pages in terms of the naturalness of reverberation. The test used a multiple stimulus presentation with high and low hidden anchors. The anchors were selected using a pilot test to elicit the most and least natural excerpts for each of the two programme items. Five subjects listened to the 44 stimuli and were asked to select the most and least natural for both programme items. Each of the four judgements were agreed on by two subjects, and the others were generally similar. These anchors were included to increase the consistency between pages.

Each page of the main test featured seven stimuli: the two (high and low) anchors, plus another five stimuli which were selected randomly from the remaining 20 stimuli of that programme item. The random seed used was different for each subject. The programme items were alternated between each page to reduce fatigue, and the stimuli were repeated once such that each randomised stimulus appeared twice, and the anchors appeared eight times. The subjects were asked to rate the “naturalness of reverberation” based on their experience of listening to naturally reverberant recordings. In this context, naturalness refers to whether the reverberation sounds artificial, rather than whether the space sounds man-made. Subjects rated at least one excerpt at 100 and at least one at 0 to normalise the range of results. Since the anchors were hidden, the subjects weren’t forced to rate these as the most and least natural. It was verbally emphasised that the subjects should *not* rate according to their preference.

The listening test was conducted by 20 trained listeners (a range of Tonmeister students, PhD students and post-doctoral staff) over Sennheiser HD600 headphones in an Institute of Sound Recording PhD office with a background noise level of 38 dB(A) SPL. The subjects were first asked to listen to all of the stimuli used in the test during a familiarisation stage to present the full range of naturalness. They then conducted a practice page where the researcher ensured they understood the task. The test took around 45 minutes to complete on average.

5.3. Subjective Results

This section details the statistical analysis of the naturalness results, and interprets the meaning of significant comparisons. The results have been split by programme item since the anchors were selected separately for each, and hence only half of the listening test pages had common stimuli. Therefore, no statistical comparisons can be made between the programme items.

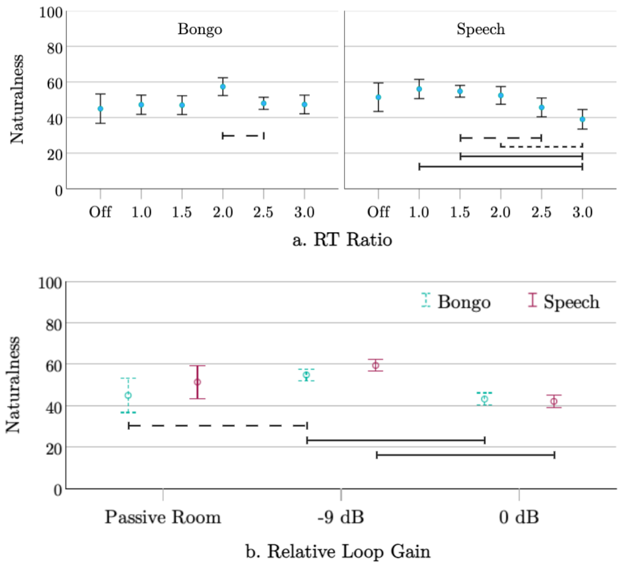

Figure 6: *The means and 95 % confidence intervals of the naturalness ratings for (a) RT ratio, and (b) loop gain, split by programme item. Statistical significances are displayed as follows. Sparsely dashed: p < 0.05, densely dashed: p < 0.01, solid: p < 0.001.*

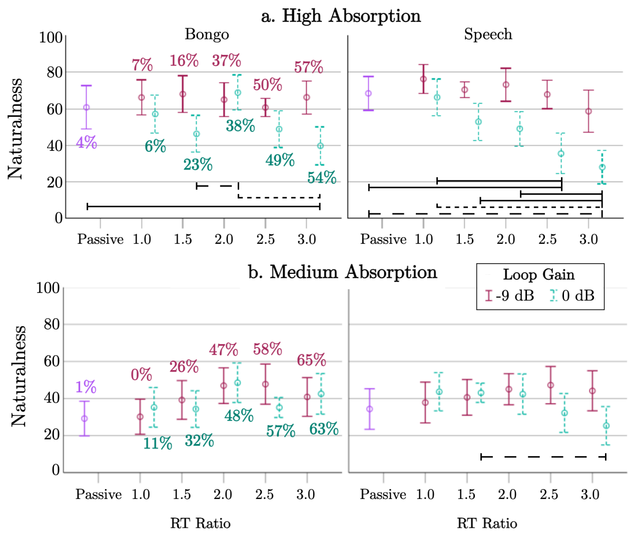

Figure 7: *The means and 95 % confidence intervals of the naturalness ratings for RT ratio (x-axis) and loop gain (clusters) split for high and medium absorption (a and b). All of the significant differences relate to the 0 dB relative loop gain. Curvature values are annotated (left) which are the same for the speech stimuli (right).*

5.3.1. Overview: RT Ratio and Loop Gain

The ratio of RTs between the passive room and the artificial reverberator has been shown to affect the amount of double sloping in Section 4. To begin to answer the research question of to what extent the DSE affects the naturalness of reverberation, the means of the naturalness ratings have been compared when varying RT ratio as well as the relative loop gain.

For each programme item, an analysis of variance (ANOVA) was performed to test the following null hypothesis: *the RT ratio has no effect on the naturalness*. At the 0.05 significance level, there was sufficient evidence to reject the null hypothesis for both programme items ($p = 0.032$ for bongo, and $p < 0.001$ for speech). Therefore, post-hoc tests were run using Tukey's honestly significant difference (HSD) multiple comparison adjustment. Figure 6 (a) shows error bars for the naturalness ratings across variations in RT ratio, where all significant differences are displayed.

Overall, the RT ratio had a significant effect on the naturalness of reverberation for the speech programme item. The RT ratios above 2.0 were significantly less natural than those closer to 1.0 (where mean naturalness decreased by 17 % from an RT ratio of 1.0 to 3.0), suggesting the increased double sloping was not only noticeable, but detrimental to the acoustic naturalness. Surprisingly, this effect was not seen for the bongo programme item despite being more wideband and transient in nature than the speech, which would typically be more revealing.

The effect of the different loop gains was tested using ANOVA tests. There was sufficient evidence to reject the null hypothesis — *the loop gain has no effect on naturalness* — for the bongo and speech ($p < 0.001$ for both). Post-hoc tests revealed that the 0 dB relative loop gain was significantly less natural than -9 dB for both programme items, shown in Figure 6 (b), where the mean naturalness decreased by an average of 14 %. Also, the -9 dB loop gain was 10 % *more* natural than the passive room.

5.3.2. Interactions: RT Ratio, Absorption and Loop Gain

To further investigate the differences seen in the naturalness ratings for increasing RT ratios, this relationship was tested when splitting the results for absorption, loop gain and programme item.

For high absorption (Figure 7 (a)), ANOVAs were run to test the interactions between RT ratio and naturalness for each loop gain, split for each programme item. For the -9 dB loop gain, there was insufficient evidence to reject the null hypothesis: *the RT ratio does not influence naturalness for the -9 dB relative loop gain* ($p = 0.682$ and $p = 0.106$ for bongo and speech, respectively). For the 0 dB relative loop gain, the ANOVAs did return sufficient evidence to reject H_0 (bongo: $p = 0.001$; speech: $p < 0.001$).

Similarly, ANOVAs were run for the medium absorption to test the interaction between RT ratio and naturalness for each loop gain and programme item. The error bars of these ratings are presented in Figure 7 (b). For the -9 dB relative loop gain, only the bongo programme item returned enough evidence to reject H_0 ($p = 0.027$), however post-hoc tests did not reveal any significant interactions using Tukey's HSD. For the 0 dB relative loop gain, only the speech programme item returned sufficient evidence to reject H_0 ($p = 0.027$), and the post-hoc tests revealed that the RT ratio of 3.0 was significantly less natural than the RT ratio of 1.5.

Figure 7 (a) shows that for the speech programme item (right) with 0 dB relative loop gain in the highly absorptive room, the naturalness drops approximately linearly with the RT ratio. The mean naturalness decreases by 41 % from the passive room to an RT ratio of 3.0, which supports the alternate hypothesis: *a change in RT ratio affects the reverberation naturalness.* A similar effect is seen for the bongo programme item (Figure 7 (a), left), where higher RT ratios were rated as significantly less natural. However, this level of significance is not seen for the more reverberant room (Figure 7 (b)), which suggests that double sloping has more of an effect on naturalness in drier passive rooms.

The fact that there are no significant differences between the mean naturalness ratings for RT ratios of the -9 dB relative loop gains (Figure 7) suggests that reducing the loop gain by 9 dB causes the DSE to either drop to a more natural level, or to become unnoticeable. The former is supported by the fact that the -9 dB bongo programme item was rated as significantly more natural than the passive room overall, indicating there was in fact an audible difference. The curvatures of the high and medium absorption EDCs are similar across the RT ratios, but some naturalness means vary significantly. For example, comparing RT ratios of 1.0 to 3.0, the curvature increases by 48 % for high absorption resulting in -38 % in mean naturalness. However, for the medium absorption, a 52 % increase in curvature results in only -18 % in mean naturalness. Hence, the amount of double sloping does not explain all of the differences in naturalness.

6. DISCUSSION

Above an RT ratio of 1, the curvature increased significantly, which aligns with the findings of Poletti [13], stating that double sloping becomes evident when the RT of the reverberator exceeds that of the passive room. In the listening test, the naturalness started to decrease with statistical significance above an RT ratio of 2 for the speech programme item. Due to the close relationship between RT ratio and degree of double sloping, the subjective results suggest that the DSE reduces the naturalness of reverberation once the curvature is $\gtrsim 40\%$. Further investigation should be conducted to determine if the results of this paper translate across different reverberator parameters.

The relationship between the amount of double sloping and subjective naturalness was sensitive to the programme item, where speech revealed more significant differences than a bongo. Programme items affecting the judgement of double sloping was also found by Luizard et al. [22], where Sabine decays were preferred for larger musical ensembles. For the speech stimuli in this paper, the amount of double sloping was negatively correlated to reverberation naturalness in certain situations. This effect was not found by Ermann [21] when preference was rated, but it should be noted that multiple subjects in this study emphasised that they were *not* rating preference, suggesting these ratings would have been different. Therefore, the results in this paper do not necessarily mean that a less natural stimulus would be rated as less preferable.

The curvature metric displayed similar values for different loop gains which was expected since curvature neglects the slope height, which is the main EDC factor influenced by in-line loop gain [27]. When the loop gain (and thus, secondary slope energy level) was higher, the relationship between RT ratio and naturalness was more critical. Luizard et al. [22] explored the suitability of reverberation when varying the simulated aperture width of coupled volumes, which has a similar effect on the EDC as loop gain in AAESs. Their results show that a larger aperture (similar to an increased loop gain) does not necessarily result in lower suitability. In this paper, lowering the loop gain by 9 dB resulted in double sloping conditions being rated as significantly more natural than at 0 dB. If suitability is considered similar to naturalness, then the difference between these results may not be due to the height of the secondary decay, but rather an artefact of the AAES versus a coupled volume model. This could be related to colouration which has been shown to be worse at higher loop gains [30].

Subjects also became more critical of increasing RT ratio when increasing the absorption in the passive room, despite the objective curvature not being influenced. This shows that another factor of the reverberation contributed to the naturalness as a result of the absorption, which worsened for higher degrees of double sloping. This was only seen for the higher loop gain, suggesting it might also be related to colouration. Jagla and Chervin [27] have shown that regeneration in an almost-anechoic room is more limited than in a reflective space, causing the AAES to act as in-line. Also, Coleman et al. [30] showed that more absorptive rooms are prone to worse colouration. Similar effects may have occurred in this paper, resulting in a more decoupled and coloured room response for the highly absorptive room.

7. CONCLUSIONS AND FURTHER WORK

This paper investigates factors affecting the double-slope effect in active acoustic enhancement systems, and explores to what extent double sloping resulting from these conditions affects the naturalness of reverberation. Monaural RIRs have been generated for varying room sizes, absorptions, channel counts, reverberator RTs and loop gains. The simulations are analysed objectively for double sloping, and subjectively for naturalness.

The objective curvature measurements show a strong influence of the ratio between the RTs of the reverberator and passive room on the amount of double sloping. This effect is most prominent above a ratio of 1, whereby the artificial reverberator RT is longer than the room RT, which supports previous work [13]. The channel count, absorption, room size and loop gain did not have a clear effect on the objective curvature of the decay. However, the relationship between the amount of double sloping and subjective naturalness was seen to change under different AAES conditions.

RIRs for the 16-channel medium room configuration were convolved with anechoic speech and bongo recordings to form 44 listening test stimuli. It was found that the naturalness significantly decreased when the reverberator RT was greater than two times the passive room RT for the speech programme item. The strong connection between RT ratio and amount of double sloping suggests that more double sloping can lead to less natural reverberation. This effect was strongest in the most absorptive room and when the loop gain was closest to the GBI.

In some cases, the naturalness did not significantly decrease when the amount of double sloping increased. For the -9 dB loop gain, the passive room was often rated lower than the active system which may mean the room acoustic model was not sufficiently natural-sounding. Also, the more reverberant room exhibited less decrease in naturalness for more pronounced double sloping. This suggests that installing an AAES in a more reverberant room may result in more natural-sounding double sloping, and that double sloping may become more acceptable in terms of naturalness if the loop gain is reduced.

Further work will include investigating the perceptual effects of the artificial reverberator when varying parameters such as echo density, timbre and channel routing. In addition, the spatial impression of active acoustic systems should be considered subjectively, for example using binaural rendering.

8. ACKNOWLEDGMENTS

Thanks to the listening test participants and the project funders: L-Acoustics and the University of Surrey Doctoral College.

9. REFERENCES

[1] J. Grosse, "Perceptual optimization of room-in-room reproduction with spatially distributed loudspeakers," *Proc. Euro. Acoust. Assoc. Joint Symp. on Auralization and Ambisonics, Berlin, Germany*, pp. 7–13, Apr. 2014.

[2] A. Haeussler and S. van de Par, "Crispness, speech intelligibility, and coloration of reverberant recordings played back in another reverberant room (room-in-room)," *J. Acoust. Soc. Am.*, vol. 145, no. 2, pp. 931–944, Feb. 2019.

[3] F. Mortessagne, O. Legrand, and D. Sornette, "Role of the absorption distribution and generalization of exponential reverberation law in chaotic rooms," *J. Acoust. Soc. Am.*, vol. 94, no. 1, pp. 154–161, Jul. 1993.

[4] N. Kanev, "Sound decay in a rectangular room with specular and diffuse reflecting surfaces," in *Forum Acusticum*, Aalborg, Denmark, Jul. 2011, pp. 1935–1940.

[5] W. B. Joyce, "Exact effect of surface roughness on the reverberation time of a uniformly absorbing spherical enclosure," *J. Acoust. Soc. Am.*, vol. 64, no. 5, pp. 1429–1436, Nov. 1978.

[6] D. T. Bradley and L. M. Wang, "The effects of simple coupled volume geometry on the objective and subjective results from nonexponential decay," *J. Acoust. Soc. Am.*, vol. 118, no. 3, pp. 1480–1490, Sep. 2005.

[7] J. Balint and F. Kaiser, "Multi-exponential decay curves in auditoriums," *Proc. Inst. Acoust.*, vol. 40, pp. 418–426, 2018.

[8] E. Nilsson, "Decay processes in rooms with non-diffuse sound fields part I: Ceiling treatment with absorbing material," *Building Acoust.*, vol. 11, no. 1, pp. 39–60, Mar. 2004.

[9] G. Dodd, "Active or passive acoustics for auditoria? – a preliminary survey," in *2nd Euro. Conf. on Architecture*, Paris, France, 1990, pp. 227–229.

[10] F. Rumsey and B. Kok, "Acoustic enhancement systems," *J. Audio Eng. Soc.*, vol. 62, no. 6, pp. 449–453, 2014.

[11] D. Griesinger, "Improving room acoustics through time-variant synthetic reverberation," in *90th Conv. Audio Eng. Soc.*, Paris, France, Feb. 1991.

[12] W. Prinssen and M. Holden, "System for improved acoustic performance," *Proc. Inst. Acoust.*, vol. 14, no. 2, pp. 93–101, 1992.

[13] M. Poletti, "The control of early and late energy using the variable room acoustics system," in *Proc. of Acoust.*, Christchurch, New Zealand, Nov. 2006, pp. 215–218.

[14] M. Poletti, "An assisted reverberation system for controlling apparent room absorption and volume," in *101st Conv. Audio Eng. Soc.*, Los Angeles, California, Nov. 1996.

[15] S. H. de Koning, "The MCR system – multiple-channel amplification of reverberation," *Philips Tech Review*, vol. 41, pp. 12–23, 1984.

[16] M. Poletti, "The performance of a new assisted reverberation system," *Acta Acustica*, vol. 2, pp. 511–524, 1994.

[17] F. Roskam, J. Laval, E. Corteel, and S. Willsallen, "ABBA voyage: Object-based mixing meets acoustic enhancement," in *Presented at Inst. Acoust. Reproduced Sound*, Bristol, Nov. 2022.

[18] M. Poletti, "Active acoustic systems for the control of room acoustics," in *Proc. Intl. Symp. on Room Acoust.*, Melbourne, Australia, Aug. 2010.

[19] S. J. Schlecht and E. A. P. Habets, "Reverberation enhancement systems with time-varying mixing matrices," in *59th Intl. Conf. Audio Eng. Soc.*, Montreal, Canada, Jul. 2015.

[20] P. Luizard, B. F. G. Katz, and C. Guastavino, "Perceptual thresholds for realistic double-slope decay reverberation in large coupled spaces," *J. Acoust. Soc. Am.*, vol. 137, no. 1, pp. 75–84, Jan. 2015.

[21] M. Ermann, "Double sloped decay: Subjective listening test to determine perceptibility and preference," *Building Acoust.*, vol. 14, no. 2, pp. 91–107, Jun. 2007.

[22] P. Luizard, B. F. G. Katz, and C. Guastavino, "Perceived suitability of reverberation in large coupled volume concert halls," *Psychomusicology: Music, Mind, and Brain*, vol. 25, no. 3, pp. 317–325, Sep. 2015.

[23] D. Ko, W. Woszczyk, and S. H. Chon, "Evaluation of a new active acoustics system in performances of five string quartets," in *132nd Conv. Audio Eng. Soc.*, Budapest, Hungary, Apr. 2012.

[24] S. H. Chon, D. Ko, and S. Kim, "Listeners' response to string quartet performances recorded in virtual acoustics," in *59th Intl. Conf. Audio Eng. Soc.*, Montreal, Canada, Jul. 2015.

[25] M. Poletti, "The analysis of a general assisted reverberation system," *Acta Acustica united with Acustica*, vol. 84, pp. 766–775, 1998.

[26] M. Poletti, "The stability of single and multichannel sound systems," *Acta Acustica united with Acustica*, vol. 86, no. 1, pp. 163–178, Jan. 2000.

[27] J. Jagla and P. Chervin, "The effect of electronic reverberation in a hybrid reverberation enhancement system: CarmenCita," in *Forum Acusticum*, Lyon, France, Dec. 2020, pp. 1223–1227.

[28] M. Vorländer, "Back matter," in *Auralization*, 1st ed. Springer Berlin, Heidelberg, 2008, pp. 303–335.

[29] F. Brinkmann and S. Weinzierl, "AKtools – an open software toolbox for signal acquisition, processing, and inspection in acoustics," in *142nd Conv. Audio Eng. Soc.*, Berlin, Germany, e-Brief 309, 2017.

[30] P. Coleman, N. Epain, S. Venkatesh, and F. Roskam, "Exploring perceptual annoyance and colouration assessment in active acoustic environments," in *Audio Eng. Soc. Intl. Conf. on Acoust. & Sound Reinforcement*, Le Mans, France, Jan. 2024.

[31] University of York, "OPENAIR anechoic speech," 2019. [Online]. Available: https://webfiles.york.ac.uk/OPENAIR/Anechoic/speech-macbeth/

[32] Bang & Olufsen, "Music for Archimedes," CD 101, 1992. [Online]. Available: http://www.angelofarina.it/Public/Anecoic/Archimedes/