Comparison of Room Acoustic Measurements of Diffuseness with a Theoretical Model

David Spargo, William L. Martens and Densil Cabrera

Faculty of Architecture, Design and Planning, The University of Sydney, NSW 2006, Australia

ABSTRACT

Predicting the diffuseness of a sound field in a small room through theoretical modelling is a task currently receiving much attention. The emphasis of such work has been upon establishing a relationship between the measurable properties of diffusing surfaces and their effects on the sound field they contain. Although several theoretical perspectives have been taken on this problem, one currently popular theoretical model lends itself to rigorous testing using sets of room acoustic measurements made in rooms allowing variable diffusion treatments. This is the focus of the work presented in this paper. It is proposed here that the success of theoretical modelling for room acoustics is best evaluated using such diffuseness measurements, as they provide the opportunity to test predictions in real world case studies.

INTRODUCTION

It is widely agreed that in rooms designed for supporting performances, be they live or reproduced, a degree of diffuseness in the sound field is desirable as it improves the listening experience (Toole 2008; Cox & D'Antonio 2009; Long 2006). It is of no surprise then, that methods for evaluating and predicting the cause and effect of diffusers and their effect on the diffuseness of the sound field in a room are of great interest to room designers and acoustics practitioners, as these would provide valuable tools for improved room design and verification.

The acoustical phenomenon of reverberation can be described in terms of the well-defined relationship between absorption coefficient (α), room dimensions and reverberation decay time (RT) expressed in the ubiquitous reverberation time formulae. With knowledge of the room's dimensions and α, RT can be predicted. Conversely, measurement of RT using interrupted noise or room impulse response (RIR) is used to derive $\overline{\alpha}$. While several methods for characterising diffusing surfaces have been suggested, description of a similar relationship between diffusing coefficients and the diffuseness of the sound field is still incomplete. The recent adoption of a standard for characterising the scattering properties of a surface (ISO 2004), goes some way to bridging this gap but only insofar as it provides a quantitative description suitable for use in room modelling. It does not describe the diffuser's quality nor is there a link between the properties of diffusing elements and the diffuseness of the sound field they influence like that expressed in the reverberation time formulae (in relation to $\overline{\alpha}$).

There have been put forward a number of theories describing the diffuse field decay of energy in rooms, (H. Kuttruff 2000; Hodgson 1996) and more recently (Jing & Xiang 2008) for example, but empirical application is yet to be realised. Hanyu recently suggested a theoretical approach that offers a more direct link between the scattering coefficient and the decay of the sound field in a way that encourages testing in real rooms (Hanyu 2010).

Practical and robust methods for measuring diffusion in the sound field must also be employed if theoretical models are to be tested against real rooms. For example, (Barron 2011) has presented a means for quantifying diffuseness through the spread of reverberation times as quantified through standard deviation calculated for spatially distributed measurement points within a given space. This is not the approach taken here, since a more direct measurement of diffuseness is the goal of the current investigation. More closely related to this goal would be the extraction of diffusion data from individual RIRs, for which several statistical methods have been suggested (Abel & Huang 2006; Stewart & Sandler 2007; Defrance & Polack 2008).

The aims of this study are to:

- Test if changes in the sound field suggested by Hanyu's model are detectable.
- Evaluate several purported methods for detecting diffusion using single source/receiver in real room impulse responses.
- Test the robustness of these methods on a range of rooms both large and small.

In the following sections of this paper, we will first outline the concept of the diffuse sound field and Hanyu's theoretical model as it relates to this study. Then, the three statistical analysis methods are described followed by examples of the analysis applied to a virtual room and two real rooms and the results discussed.

THE DIFFUSE SOUND FIELD

Reverberation is the most commonly recognised and analysed acoustical phenomenon occurring in rooms. This is partly due to the fact that it is readily detectable by ear and instruments and partly because it is well described in physical terms. It is the result of sound energy, radiated from a source, reaching the listener (or receiver) after being reflected by the room's boundary surfaces and by surfaces of objects that may be within the room. Each time the sound energy interacts with a surface there is the possibility that it may be reflected, absorbed, diffused or, more likely, a combination of all three. Sound energy that is not absorbed during this interaction with the surface is then redirected back into the room where it will likely interact with yet more surfaces. This process may, depending on the surface properties, occur many times until all the sound energy is absorbed and in doing so, an increasingly complex sound field evolves that decays exponentially and becomes increasingly diffuse over time, that is it becomes more random or chaotic in terms of the energy distribution (both isotropic and homogeneous). Of course, in some particularly dead rooms, or those with focussing geometry, this may not be the case. The mechanisms for reflection and absorption are well known and are properly defined by the absorption coefficient. However, a complete description for diffusion is yet to be defined although the scattering coefficient, recently adopted, provides a partial quantitative description suitable for modelling purposes (ISO 2004).

In developing a theoretical perspective of the evolving sound field, an understanding of two main aspects of the how the energy in the field evolves is fundamental; these are frequency and time. In the frequency domain, modal density increases exponentially with frequency (H. Kuttruff 2000). According to Schroeder (Schroeder 1996), once the modal density exceeds the critical frequency (f_c), the reverberation process can be considered stochastic because the modes overlap to such a degree that they become individually indistinguishable from one another. This is opposed to the deterministic character of individual modes below f_c where individual modes are easily distinguished both by measurement and by ear. In the time domain, echo density also increases exponentially with time. Here, discrete reflection or echo events in the early part of the time signal are separated from the growing incidence and randomness of echoes in the latter part of the time signal by what some referred to as the 'Characteristic Time' or 'Critical Time' (t_c) in reverberation synthesis circles (K. Kuttruff 1993; Farina n.d.), or 'Mixing Time' (T_m) by others interested in analysing RIR (Defrance & Polack 2008; Stewart & Sandler 2007). At present, there is no agreed definition or method for measuring mixing time but the concept has significance as it is deemed to represent the point in time where the evolution of the sound field becomes stochastic and diffuse.

Modelling diffuseness in the sound field

Recently, Hanyu presented a simple theoretical model that seeks to relate some of the diffusion characteristics of room surfaces, in the form of the absorption coefficients and scattering coefficients as described in (ISO 2004), and the diffuseness of the sound field contained by the room's surfaces. The theory presents a convenient method for predicting the rate at which sound energy is converted from specular energy into diffuse energy during successive reflections from the room surfaces.

The theory is based on a probabilistic model of surface reflectivity derived from two components, the absorption coefficient α and scattering coefficient s, expressed in the following equation:

$$\alpha + (1-\alpha)s + (1-\alpha)(1-s) = 1 \tag{1}$$

The probability that sound energy will be reflected and scattered ranges from 0 to 1, with 0 indicating 0% probability and 1 indicating 100% probability.

This leads to the formulation of expressions for the rate of decay in the energy in a reverberant space due to probable absorption

$$P_a(t) = \exp\left[\frac{\ln(1-\overline{\alpha})}{d}ct\right] \tag{2}$$

where $\overline{\alpha}$ is the average absorption coefficient and d is the mean free path, $4V/S$, t is time in seconds and c is the speed of sound in m/s (ct is therefore the distance travelled by the sound). The rate of decay for probable specularly reflected energy is

$$P_{spec}(t) = \exp\left[\frac{\ln\overline{\mu}}{d}ct\right] \tag{3}$$

where $\overline{\mu}$ is the average of the specular reflection coefficient, that is the probability of incident sound being specularly reflected given by $(1-\alpha)(1-s)$. The rate of decay for probable scattered energy is then

$$P_d(t) = P_a(t) - P_{spec}(t)\,. \tag{4}$$

Figure 1 shows a plot of these three decay functions over time for a room 20 m x 20 m x 20 m with $\alpha = 0.2$ and $s = 0.5$.

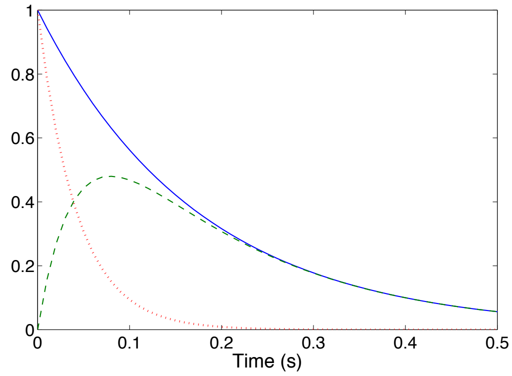


Figure 1 Typical energy decay curves after Hanyu. The dashed line shows the evolution over time (in seconds) of energy in the room attributed to diffuse reflections, $P_d(t)$. The dotted line shows energy attributed to specular reflections, $P_{spec}(t)$. The solid line shows the combined energy in the room from all reflections, $P_a(t)$.

The convenience of this model is in the way that specular and diffuse energy decays can be described individually and separately from the combined energy decay, typically described in terms of the reverberation time (RT) metric. Although the model considers the sum of energy in the entire room at any time during the evolution of the room response, it is reasonable to expect that indications of this behaviour should be detected in individual room impulse responses. Therefore, a method for detecting the diffuseness data within RIR is needed.

DETECTING DIFFUSION IN RIR

The effects of diffusion may be captured indirectly in many of the standard room acoustic metrics such as Reverberation Time (RT), Early Decay Time (EDT) and Clarity Index (C_{50}, C_{80}) all based on single source/receiver relationships. Other techniques based on multichannel directional receivers such as Interaural Cross Correlation (IACC) and Lateral Fraction (LF) are thought to be more sensitive to diffusion effects but there are as yet no direct measures for diffusion.

Recently, some statistical approaches for the analysis of single channel RIRs have been proposed (Stewart & Sandler 2007; Abel & Huang 2006; Defrance & Polack 2008). Each of these methods has been suggested as an indicator of time domain echo density in single channel RIRs. They are all variations of Gaussian estimators based on the premise that as time progresses the sound energy reaching a receiver resulting from sound emitted from source in the same room will contain ever-increasing numbers of reflections from the room's surfaces giving rise to an increasingly random distribution of energy that approaches Gaussian distribution. The statistical function is created by recording the analysis result of a Hanning window function stepped in time along a RIR. The window length and the step time can be varied. From this work, three statistical methods are used. The first is to calculate the standard deviation σ of the time domain signal. The second method calculates the kurtosis of the same signal, and the third method calculates the normalised echo density (NED) as described in (Abel & Huang 2006). Essentially, it is suggested in each of the studies mentioned above that these methods are able to identify in the RIR the point or region in time around which the sound field changes from non-diffuse to diffuse. This is indicated by sharp transitions in the statistical functions. In the case of the standard deviation and kurtosis, this is a rapid decrease in the function and for NED it is an elbow point where the function stops rising and flattens out at maximum value. These points are then deemed to indicate the transition between early reflection and late reverberation, called "mixing time" by (Defrance & Polack 2008) and "transition time" by (Stewart & Sandler 2007).

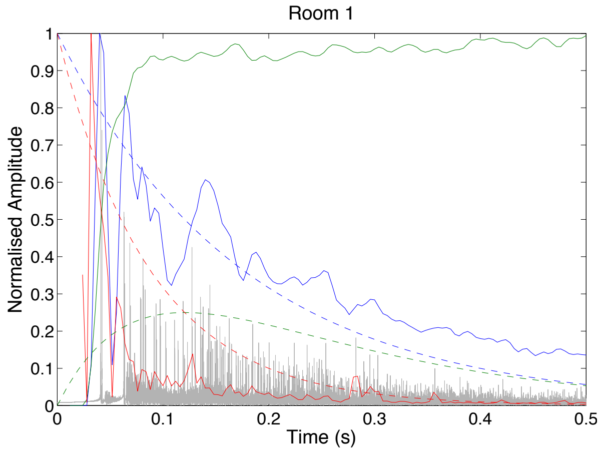
In this study the analysis methods described above are applied to a variety of RIRs to test robustness with regard to room size and sensitivity to changes in surface scattering. Three room types of various sizes are used, one virtual and two real, where the amount of scattering can be changed with other factors held constant. It is important to note that in all three of the aforementioned studies, only larger rooms with relatively long RTs were tested. The smaller rooms studied below present several problems. In comparison, the time scale of the RIR is compressed making the analysis period much shorter. Also, the portion of the spectrum where stochastic behaviour is likely is limited to the higher frequencies. To compensate for this, frequency scaling is required. This achieved using fourth order zero-phase high-pass Butterworth filtering to ensure that only the portion well above the Schroder frequency is analysed.

Simulated room

The energy decay in the room can be calculated using the room's dimensions and the surface absorption coefficients using the well-known reverberation time equations from Sabine and others. Geometrical models using these values are commonly used to generate room simulations and simulated room impulse responses for analysis of source/receiver relations within rooms. Incorporating scattering coefficients into these modelling processes provides an approximation of the diffusing qualities of surfaces and their effect on the sound field.

A room simulation provides a convenient first step in testing the efficacy of the statistical methods described above. This provides controlled conditions for the detection of possible variations in energy attributed purely to specular and diffuse reflections described by Hanyu's theory. To this end, computer models of seven rooms described in Hanyu's study were constructed in MCRoomSim (Wabnitz et al. 2010). An arbitrarily located set of single channel omnidirectional source and receiver were configured and the same arrangement was used for each room for consistency. Simulated room impulse responses were then generated for each room/source/receiver combination. This model provides a convenient platform for testing simple room configurations because the specular and diffuse components of the sound field can be output separately for analysis in a way similar to the verification modelling in Hanyu's study.

Figure 2 shows an example of analysis of room 1 and room 2. Both rooms have the same dimensions, 20 m x 20 m x 20 m, the same average absorption coefficient of 0.2 and therefore roughly the same RT of about 2.4 seconds. However, room 1 has an average scattering coefficient of 0.2 while room 2 is 0.5. In both cases a window length of 24 ms with a step size of 4 ms is used for the statistical analysis.



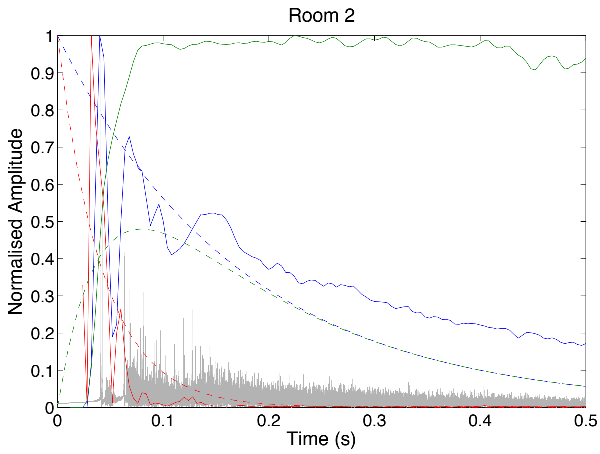


Figure 2 Analysis of simulated RIR from MCRoomSim of rooms 1 & 2 from Hanyu. ETC of RIR in grey, dashed green line shows $P_d(t)$, dashed red line shows $P_{spec}(t)$, dashed blue line shows $P_a(t)$, blue line shows σ, red line shows kurtosis, green line shows NED. Window size 24 ms, step size 4 ms.

In this example the differences caused by changing only the scattering coefficient are readily detectable by eye. This is evident in the energy time curve (in grey), in the modelling of the energy decay using Hanyu's method (dotted lines) and in the analysis using the three statistical methods. It is clear that in this case each of the statistical analysis methods shows a fair degree of sensitivity to the change in scattering coefficient. The NED analysis seems to provide the clearest indication of the transition from non-diffuse to diffuse with a knee point at around 90 ms for room 1 and around 75 ms for room 2. It also seems to agree with the predicted $P_d(t)$ value from the theoretical model. While the standard deviation and kurtosis function also seem to be indicating diffusion they are not nearly so well defined. For the purpose of verification of the methods, these results are consistent with those of previous studies.

Reverberation room

In laboratory reverberation rooms the diffuseness of the sound field is a topic of frequent investigation since the accuracy of a variety of measurements made in the rooms depends on a completely diffuse sound field.

In this second analysis set, several RIRs are taken from a large group of measurements conducted as part of a previous study on measurement of the diffuse field in the reverberation room at the Faculty of Architecture, University of Sydney (Bassett et al. 2010). In that study volumetric diffuser elements were progressively added to the room between measurements to investigate how this changed the sound field diffusivity. This data set is very useful in the context of the current study, since absorption coefficients and source/receiver locations remained approximately constant while diffusion was varied. A broadband porous absorber occupied about half of the floor of the room.

Since the reverberation room is considerably smaller than the previous room tested, only 130 m^3 compared to 8000 m^3, it is frequency-scaled with a cut-off frequency of 1 kHz.

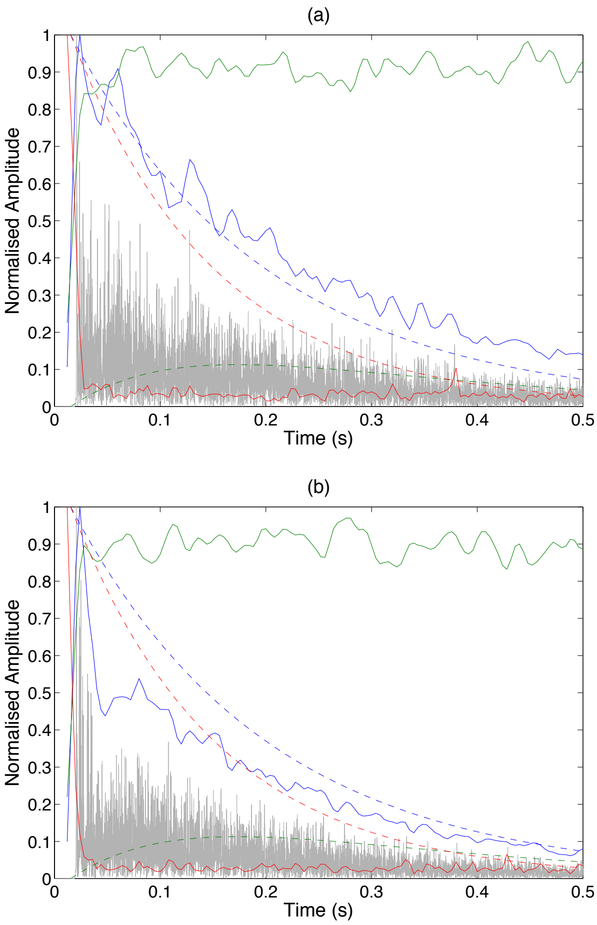


Figure 3 Analysis of reverberation room with no diffusors (a) and 15 diffusors (b). ETC of RIR in grey, dashed green line shows $P_d(t)$, dashed red line shows $P_{spec}(t)$, dashed blue line shows $P_a(t)$, blue line shows σ, red line shows kurtosis, green line shows NED. Window size 24 ms, step size 4 ms.

A comparison of the analysis of two reverberation room states one with no diffusers is shown in Figure 3(a) and one with 15 diffusers is shown in Figure 3(b). The analysis window size and step size are the same as for the previous case, 24 ms and 4 ms respectively. Here, the values from the theoretical model provide only a reference point since the scattering coefficient of the diffuser panels is unknown although $P_a(t)$ is correctly modelled for this room. In this example, the standard deviation function shows the strongest indication of an increase in sound field diffusion. There is also a possible indication from the NED but the roughness of the function makes it hard to interpret. Neither provides a clear indication of a transition point however.

Critical listening room

In a critical listening room a high degree of precision is expected from all sound reproduction components including the room. The correct placement of scattering surfaces is especially important as it has a strong influence over the quality of sound the listener perceives. Compared to a concert hall, critical listening rooms are usually quite small and have very short RTs.

The third selection of RIRs was taken for a larger set measured in a variable acoustic listening room A816 at CIRMMT, McGill University in Montreal. This room utilises interchangeable acoustic modules over all wall and ceiling surfaces to allow the setting of a wide range of room acoustic conditions. Two types of modules are used, RPG Skylines and RPG MelaFlex Wedges. Using the models, room states can be varied from reverberant (no surface treatment) to diffuse to near anechoic or combinations of treatments. This makes comparisons interesting since the absorption, scattering and diffusion coefficients of the acoustic modules are well documented and make detailed computer modelling possible.

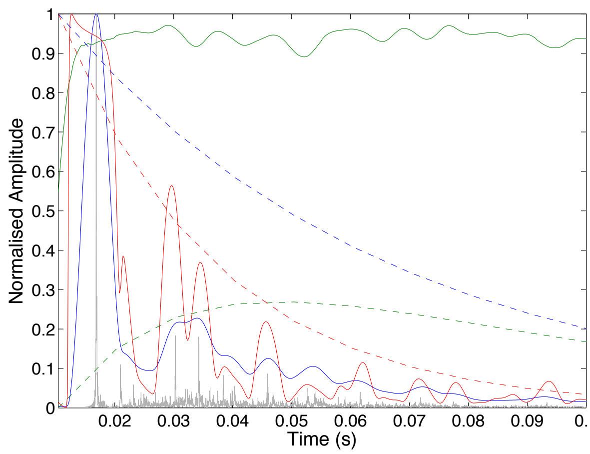


Figure 4 An example of the analysis of room A816 in reverberant state. ETC of RIR in grey, dashed green line shows $P_d(t)$, dashed red line shows $P_{spec}(t)$, dashed blue line shows $P_a(t)$, blue line shows σ, red line shows kurtosis, green line shows NED. Window size 10 ms, step size 0.2 ms.

It appears that in rooms this small, extracting meaning from the statistical analysis method becomes very difficult. Figure 4 illustrates this with an example of a RIR from room A816 in its reverberant state, that is, with reflective walls and absorbent ceiling. Even with a high degree of frequency scaling, 4 kHz cut off, and reduced analysis window size, the analysis methods do not seem to work very well. For example, the NED analysis in Figure 4 indicates the sound field is diffuse by about 10 ms but this cannot be the case since the direct sound has not yet had time reach the receiver or reflect off other surfaces. After trying a range of window sizes, 10 ms with 0.2 ms step size seemed to provide the most useful data, particularly from the standard deviation analysis.

For this room, direct comparisons from one room state to another are made difficult because the analysis is so sensitive to small features in the RIR. However, the standard deviation appears more robust than the NED and kurtosis. Figure 5 shows the standard deviation for the three room states. This does seem to resemble expectations.

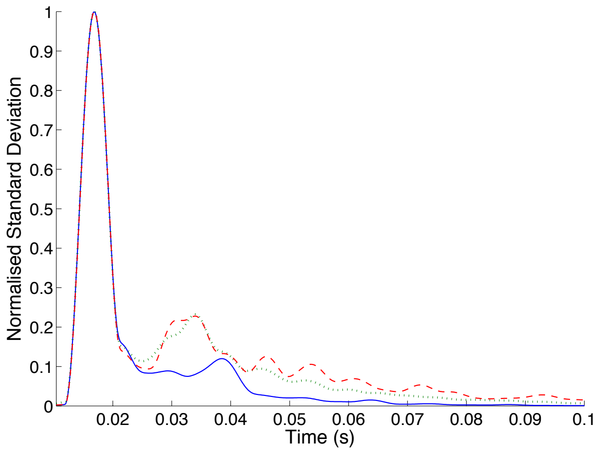


Figure 5 A comparison of the standard deviation analysis for room A816 in three states, reverberant (red dashed line), diffuse (green dotted line) and anechoic (solid blue line). Window size 10 ms, step size 0.2 ms.

DISCUSSION

A practicable link between scattering coefficients and sound field diffusion will provide a valuable tool for acoustics practitioners particularly in rooms where accuracy in sound reproduction is critical. In this study a theoretical model that suggests such a link has been evaluated by comparing it to data analysis of RIRs from simulated and real rooms. Three statistical analysis methods have also been evaluated which potentially measure diffuseness in a sound field through analysis of RIRs. Both the theoretical model and the analysis methods have been evaluated on small and large rooms to test their robustness to variation in room size.

In large rooms both the theoretical model and the analysis appear to work well. Good agreement seems to be found between the models of diffuse energy in the sound field and the results of the statistical methods of detecting it. However, it is clear that as room size is reduced, the theoretical model and the analysis methods appear to break down rapidly. Of the three analysis methods, the standard deviation seems to be most robust but it is very sensitive to window and step size and there is as yet no clear method for setting this other than guessing. An apparent lack of robustness in the analysis method for small rooms also brings into question the applicability of the methods as psychoacoustic measures, which is one of the main purposes purported of the NED method especially.

Using these statistical analysis methods for detecting mixing time needs further investigation. Defrance & Polack (Defrance & Polack 2008) state that mixing time is linked to the room volume but it must also be linked to surface diffusion. This is evident in the theoretical models, certainly in the larger rooms, from the way the peak in the $P_d(t)$ shifts in time relative to the scattering coefficient and in the corresponding time shifts in the supposed transition points estimated from the statistical analysis.

This study demonstrates the value of testing theoretical modelling predictions and analysis methods against a wide variety of measurements from real world case studies.

REFERENCES

Abel, J.S. & Huang, P., 2006. A Simple, Robust Measure of Reverberation Echo Density. *AES 121st Convention*, pp.1–10.

Barron, M., 2011. Non-linear decay in simple spaces and their possible exploitation. *Proceedings of the Institute of Acoustics*, 34, pp.1–12.

Bassett, J., Spargo, D. & Cabrera, D., 2010. Using an ambisonic microphone for measurement of the diffuse state in a reverberant room. *Proceedings of 20th International Congress on Acoustics, ICA 2010*.

Cox, T. & D'Antonio, P., 2009. *Acoustic absorbers and diffusers: theory, design and application*, Spon Press.

Defrance, G. & Polack, J.D., 2008. Measuring the mixing time in auditoria. *Journal of the Acoustical Society of America*.

Farina, A., Pyramid Tracing vs. Ray Tracing for the simulation of sound propagation in large rooms. *Computational Acoustics and its Environmental Applications*.

Hanyu, T., 2010. A framework for characterizing sound field diffusion based on scattering coefficient and absorption coefficient of walls. In *Proceedings of the International Symposium on Room Acoustics, ISRA 2010*.

Hodgson, M., 1996. When is diffuse-field theory applicable? *Applied Acoustics*, 49(3), pp.197–207.

ISO, 2004. *ISO 17497-1:2004: Acoustics - Sound-scattering properties of surfaces - Part 1: Measurement of the random-incidence scattering coefficient in a reverberation room*,

Jing, Y. & Xiang, N., 2008. On boundary conditions for the diffusion equation in room-acoustic prediction: Theory, simulations, and experiments. *The Journal of the Acoustical Society of America*, 123(1), p.145.

Kuttruff, H., 2000. *Room acoustics* Fourth. Spon Press.

Kuttruff, K., 1993. Auralization of impulse responses modeled on the basis of ray-tracing results. *Journal of The Audio Engineering Society*.

Long, M., 2006. *Architectural acoustics*, Academic Press.

Schroeder, M., 1996. The "Schroeder frequency" revisited. *The Journal of the Acoustical Society of America*.

Stewart, R. & Sandler, M., 2007. Statistical measures of early reflections of room impulse responses. In *Proceedings of the 10th International Conference on Digital Audio Effects*.

Toole, F., 2008. *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*, Focal Press.

Wabnitz, A., Epain, N. & Jin, C., 2010. Room acoustics simulation for multichannel microphone arrays. *Proceedings of the International Symposium on Room Acoustics, ISRA 2010*.