Western Australian
Telecommunications Research Institute

A joint venture between:
The University of Western Australia &
Curtin University of Technology

E-mail: ericl@watri.org.au
Tel.: +61 (0)8 6488 4642

Classification:
Public

Document Type:
Scientific Report

Document Reference:
PRJ-NICTA-PM-025

Status:
Draft

Revision:
2.2

Date:
June 15, 2008

Title:

Prediction of Energy Decay in Room Impulse Responses Simulated with the Image-Source Model

Author(s):

Eric A. Lehmann and Anders M. Johansson

Document History:

Revision	Date	Comments
1.0	06-12-06	Initial full report draft
1.1	08-01-07	Corrections from internal review
1.2	09-01-07	Added plot of $\hat{h}_P(t)$ to show decrease to 0 as $t \to \infty$. This version was used for the creation of 2 separate journal/conference papers.
1.3	16-01-07	Plot of $\hat{h}_P(t)$ now plotted in log-scale
2.0	25-02-08	Included corrections from review process on JASA submission
2.1	24-05-08	Version made public (papers published/to appear in WASPAA'07 and JASA)
2.2	15-06-08	Appendix updated to explain coordinate rotation

Contents

Abstract

Image source methods have become a widely-used analysis tool in many fields of acoustics and engineering. In this paper, a technique is proposed for approximating the energy decay (energy–time curve) in the room impulse responses simulated using the image source model. To this purpose, a modified version of the image method is considered: the computations are carried out in the frequency domain and negative reflection coefficients are used, which leads to more natural-looking impulse responses. A geometrical analysis of the image source principle leads to an analytical formula describing the energy decay curve, which is valid for either a uniform or non-uniform definition of the enclosure's six absorption coefficients. The proposed approximation is investigated and compared to image source results on the basis of simulations involving various room sizes and reverberation levels, and with uniform as well as non-uniform sound absorption coefficients. It is shown that the proposed formula provides a good approximation of the energy–time curve computed from a simulated room impulse response: the predicted curve accurately "mimics" the overall slope as well as the specific curvature of the energy decay. The result presented in this work thus enables designers to undertake a preliminary analysis of a simulated reverberant environment without the need for time-consuming image method simulations.

A potential implementation example for the proposed method is also considered in this work. Currently available formulae for the prediction of an enclosure's reverberation time, such as the well-known formulae by Sabine and Eyring, do not provide accurate results when used in conjunction with the image method. In this paper, it is shown how the proposed energy decay approximation can be used to effectively determine the enclosure's absorption coefficients in order to achieve a desired reverberation time. This approach hence ensures that the image source model effectively generates impulse responses with a correct level of reverberation, which is of particular importance, for instance, for the purpose of assessing the performance of acoustic signal processing algorithms operating in reverberant conditions.

1 Introduction

The image source model (ISM) has received enormous attention in many fields of engineering and acoustics over the last few decades. Together with other modeling techniques of geometrical room acoustics, such as ray and beam tracing, the ISM undoubtedly represents an established tool of considerable importance in the research community. This huge gain in popularity can be linked mainly to the dramatic increase in processing power of personal computers, which partly alleviates the computational burden required by the ISM technique. Its success also relies on its conceptual simplicity, thus making ISM-based algorithms relatively straightforward to implement.

As a result, the ISM approach has been used, mostly in the domain of room acoustics and architectural design, as a basis principle for a variety of purposes, such as generic simulation of acoustic transfer functions in reverberant enclosures [1–5], prediction of sound propagation in closed or open environments [6–8], and modeling and design of architectural spaces [9–11], including specific spaces such as long enclosures (road tunnels, narrow "street canyons", etc.) [12–14] and factories (noise control) [15]. The image source model also provides insight into the fundamental acoustical properties of various enclosures, and can be used for the analysis of perceptual properties such as reverberation time, speech intelligibility and speech transmission index [16–18], or for the purpose of auralization and binaural reproduction [19–22]. Recently, the ISM technique has also been implemented in relation to spatialized sound rendering in virtual environments [23–26], immersive or interactive systems [27, 28], and augmented reality video games [29].

Another important domain of application of the image source technique is as a tool used in the process of assessing the performance of many acoustic signal processing algorithms operating in reverberant environments. To name but a few examples, the ISM approach has been used in order to validate algorithms for blind source separation [30–32], channel identification and equalization [33–36], acoustic source localization and tracking [37–39], speech recognition [40, 41], speech enhancement [42, 43], and various other array signal processing techniques [44–48]. Typically, the image source model is used to test the algorithms under consideration, in order to determine their robustness against increasing levels of environmental reverberation. A significant difficulty during this assessment process is related to predicting the reverberation time (RT) in the room impulse responses generated with the ISM. Well-established formulae, such as Sabine's or Eyring's reverberation time, do not provide good results when attempting to determine the environment's sound absorption coefficients in order to achieve a given reverberation level in the ISM results, especially when used in conjunction with a non-uniform definition of the sound reflection coefficients in the considered room. The current solution to this problem is a computationally demanding trial-and-error process, where the reverberation time is determined from a few sample transfer functions. In settings where many impulse responses are to be computed, such as, for instance, when simulating a source moving in the environment [37, 38], this can lead to a prohibitively time-consuming process. A solution to this specific problem was the motivating idea behind the work presented here.

This paper describes a numerical strategy for predicting the energy decay in a room impulse response (RIR) simulated with the ISM. The proposed approximation method is based on a geometrical consideration of the ISM principle: the power in the transfer function at a specific time lag can be seen as the addition of the contributions from the image sources located on a sphere centered around the receiver. This approach leads to a closed-form equation which then allows a numerical prediction of the energy decay curve (EDC), thus alleviating the need to practically simulate the RIR of interest. Furthermore, this method efficiently deals with situations where current RT prediction techniques experience significant inaccuracies, such as in the case of highly non-uniform absorption coefficients. The ability to accurately predict the energy decay thus allows a simplified design process when dealing with the image source technique, and typically provides an efficient tool for the above-mentioned problem of determining the value of absorption coefficients achieving a desired reverberation time. It must be emphasized here that the purpose of this work is not towards improving the performance of the ISM or the accuracy of current RT prediction formulae in comparison to practically-measured responses in real rooms or in scale models. Instead, the proposed technique can be seen as a complementary tool developed to make the ISM more efficient to deal with, thus facilitating the ISM-based analysis of acoustical spaces, in particular by reducing the computational burden, and hence the computation time, required for such an analysis.

In the next section, the basic principles underlying the image source model are first reviewed, followed by a description of the modified ISM technique that will be used throughout the rest of this work. Section 3 then presents the details of the proposed EDC approximation technique, where parts of the

derivations can be found in the Appendix for clarity. The accuracy of this approximation method is also demonstrated using a series of numerical evaluation scenarios. In Section 5, a typical application example is considered: it is shown how the proposed technique can be efficiently applied to the problem of setting the environment's absorption coefficients in order to achieve a specific reverberation time in the simulated RIRs. The superiority of this approach over several formulae used currently in the literature is demonstrated by means of ISM simulations in various room settings. Finally, Section 6 concludes with a discussion of the main concepts and results presented in this work.

2 Image Source Method

This section briefly reviews the basic principles underlying the image source simulation technique, and establishes the corresponding notational conventions used throughout the rest of the paper. It first summarizes the developments leading to the conventional image source method, as presented originally in a landmark paper by Allen and Berkley [1]. It then presents a modified version of this technique that will be used in relation to the proposed energy decay approximation in Section 3.

2.1 Original Approach

Allen and Berkley's implementation of the image source method [1] is a well-established algorithm for generating simulated RIRs in a given room. Assume that a Cartesian coordinate system with coordinates (x, y, z) is defined in the considered enclosure, with its origin corresponding to one of the room corners. Let $\boldsymbol{p}_s$ and $\boldsymbol{p}_r$ denote the position vectors of a source and a receiver in this setting, respectively, as follows:

$$\boldsymbol{p}_s = [x_s, y_s, z_s]^{\mathrm{T}}, \tag{1}$$

$$\boldsymbol{p}_r = [x_r, y_r, z_r]^{\mathrm{T}}, \tag{2}$$

where $[\,\cdot\,]^{\mathrm{T}}$ denotes the matrix transpose operator. Similarly, let

$$\boldsymbol{r} = [L_x, L_y, L_z]^{\mathrm{T}} \tag{3}$$

represent the vector of room dimensions, with length L_x, width L_y and height L_z. It is also assumed that the acoustical characteristics of each surface in the enclosure are characterized by means of a sound reflection coefficient β, related to an absorption coefficient α according to the well-known formula

$$\alpha = 1 - \beta^2. \tag{4}$$

The reflection coefficients for each surface are denoted $\beta_{x,i}$, $\beta_{y,i}$ and $\beta_{z,i}$, respectively, with $i = 1, 2$, where the subindex 1 refers to the wall closest to the origin. For simplicity, this work makes use of the usual geometrical room acoustics assumption that the reflection coefficients are frequency-independent as well as angle-independent (specular reflection on a locally reacting surface).[1]

The room impulse response from the source to the receiver can be determined by considering image sources on an infinite grid of mirror rooms expanding in all three dimensions. The contribution of each image source to the receiver signal is a replica of the source signal delayed by a lag τ, corresponding to the sound propagation time from the image source to the receiver, and attenuated by a factor A corresponding to the number of reflections on each wall as well as the sound intensity decay along the path from the source to the receiver. The RIR $h(\cdot)$ hence follows as

$$h(t, \boldsymbol{p}_s, \boldsymbol{p}_r) = \sum_{\boldsymbol{u}=0}^{1} \sum_{\boldsymbol{l}=-\infty}^{\infty} A(\boldsymbol{u}, \boldsymbol{l}) \cdot \delta\big(t - \tau(\boldsymbol{u}, \boldsymbol{l})\big)\,, \tag{5}$$

where t denotes time, $\delta(\cdot)$ is the Dirac impulse function, and the triplets $\boldsymbol{u} = (u, v, w)$ and $\boldsymbol{l} = (l, m, n)$ are parameters controlling the image source indexing in all dimensions. For conciseness, the sum over $\boldsymbol{u}$

[1] For the purpose of clarity, this work only considers frequency-independent and angle-independent coefficients. Both the angle and frequency dependence could be included in the acoustic model, but only at the expense of significantly complicating the derivations presented in the following.

(respectively $\boldsymbol{l}$) in (5) is used to represent a triple sum over each of the triplet's internal indices. The attenuation factor $A(\cdot)$ and time delay $\tau(\cdot)$ in (5) are defined as follows:

$$A(\boldsymbol{u},\boldsymbol{l}) = \frac{\beta_{x,1}^{|l-u|}\beta_{x,2}^{|l|}\,\beta_{y,1}^{|m-v|}\beta_{y,2}^{|m|}\,\beta_{z,1}^{|n-w|}\beta_{z,2}^{|n|}}{4\pi\cdot d(\boldsymbol{u},\boldsymbol{l})}\,, \tag{6}$$

$$\tau(\boldsymbol{u},\boldsymbol{l}) = d(\boldsymbol{u},\boldsymbol{l})/c\,, \tag{7}$$

where $c = 343\text{m/s}$ is the sound propagation velocity, and $d(\cdot)$ represents the distance from the image source to the receiver, computed according to

$$d(\boldsymbol{u},\boldsymbol{l}) = \big\|\,\mathrm{diag}(2u-1, 2v-1, 2w-1)\cdot\boldsymbol{p}_s + \boldsymbol{p}_r - \mathrm{diag}(2l, 2m, 2n)\cdot\boldsymbol{r}\,\big\|\,, \tag{8}$$

where $\|\cdot\|$ is the Euclidean norm and with $\mathrm{diag}(\cdot)$ denoting a diagonal matrix with the arguments as diagonal elements. According to (5), the resulting transfer function $h(\cdot)$ can be seen as a positive (constructive) addition of the sound amplitude associated with each of the image sources.

Finally, note that the number of image sources to include in the summation of (5) grows exponentially with the considered order of reflections. The simulation of a full-length RIR using an image-source approach can thus lead to a considerable computational load in practice.

2.2 Modified ISM Technique

The basic image source simulation method can be improved in a number of different ways. This section presents two such modifications, which lead to more efficient simulations and more accurate results. The resulting algorithm will be used as a basis for the simulations presented at the end of this document.

2.2.1 Frequency-Domain Computations

The ISM implementation presented in Section 2.1 is typically inappropriate in practice when dealing with discrete-time signals, since the time delay $\tau(\cdot)$ in (7) does not usually correspond to an integer multiple of the sampling period. In Allen and Berkley's approach [1], this problem is dealt with by using nearest-integer rounding of the image source's propagation time, resulting in a slight shift of the corresponding impulse in the RIR. Computing the desired RIR according to (5) hence results in a coarse, histogram-like representation [3, 5], which subsequently requires high-pass filtering in order to remove a non-physical defect of the model resulting at zero frequency.

A more accurate solution to this problem is to carry out the ISM computations in the frequency domain, which allows the representation of fractional delays that are not necessarily integer multiples of the sampling period. In the frequency domain, a time shift τ can be represented as $\exp(-j\omega\tau)$, with $j = \sqrt{-1}$ and ω denoting the frequency variable. The frequency-domain RIR $H(\cdot)$ hence results from (5) as

$$H(\omega,\boldsymbol{p}_s,\boldsymbol{p}_r) = \sum_{\boldsymbol{u}=0}^{1}\sum_{\boldsymbol{l}=-\infty}^{\infty} A(\boldsymbol{u},\boldsymbol{l})\cdot \mathrm{e}^{-j\omega\,\tau(\boldsymbol{u},\boldsymbol{l})}, \tag{9}$$

where $A(\cdot)$ and $\tau(\cdot)$ are computed according to (6) and (7), respectively. The time-domain RIR then results as the inverse Fourier transform of $H(\cdot)$:

$$h(t,\boldsymbol{p}_s,\boldsymbol{p}_r) = \mathcal{F}^{-1}\big\{H(\omega,\boldsymbol{p}_s,\boldsymbol{p}_r)\big\}. \tag{10}$$

Note that for time-sampled, and hence band-limited signals, the contribution of each image source in the time domain then results as a (truncated) sinc-like fractional-delay filter that accounts for non-integer time delays; the computation of $h(\cdot)$ in (5) hence effectively becomes

$$h(t,\boldsymbol{p}_s,\boldsymbol{p}_r) = \sum_{\boldsymbol{u}=0}^{1}\sum_{\boldsymbol{l}=-\infty}^{\infty} A(\boldsymbol{u},\boldsymbol{l})\cdot \mathrm{sinc}\big(t-\tau(\boldsymbol{u},\boldsymbol{l})\big)\,, \tag{11}$$

with $\mathrm{sinc}(\xi) = \sin(2\pi F_s\xi)/(2\pi F_s\xi)$ and F_s denoting the sampling frequency. This approach, which was used previously by various authors [18, 33], essentially represents the frequency-domain equivalent to

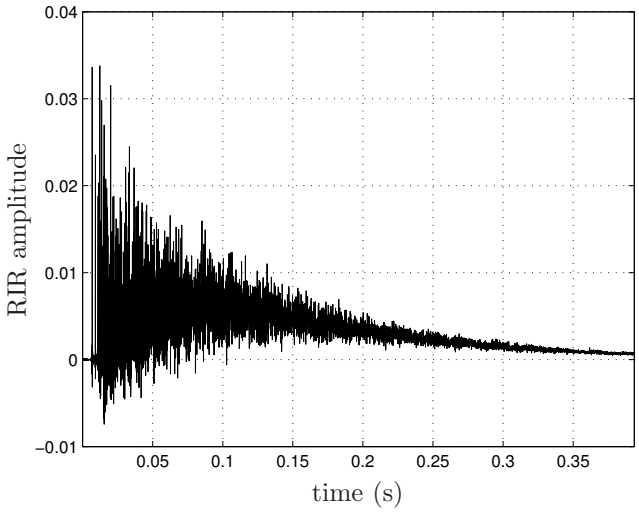

Figure 1: Typical RIR example obtained using frequency-domain ISM computations with *positive* reflection coefficients. This example was computed for a $3.2\text{m} \times 4\text{m} \times 2.7\text{m}$ room, with $\boldsymbol{p}_r = [1.1\text{m}, 1\text{m}, 1.2\text{m}]^{\mathrm{T}}$, $\boldsymbol{p}_s = [2\text{m}, 3\text{m}, 2\text{m}]^{\mathrm{T}}$, $F_s = 16\text{kHz}$ and uniform reflection coefficients $\beta = 0.92$ for each enclosure surface.

low-pass impulse method (with infinite window duration) proposed by Peterson [49].

In the case of a practical implementation, the considered signal data is typically discretized following an anti-aliasing step, which ensures that the signals are band-limited. As a result, a significant advantage of the frequency-domain approach described by (9) and (10) is the ability to speed up the ISM computations according to the following principle. With time-sampled signals, the Fourier transform in (10) is replaced by a fast Fourier transform (FFT), and the function $H(\cdot)$ in (9) is computed at a number of discrete frequency values ω_k:

$$\omega_k = k \cdot \underbrace{2\pi F_s/K}_{\Delta\omega}, \quad k = 1, \ldots, K, \tag{12}$$

where K the FFT length. It follows that the complex exponential function in (9) can be rewritten as

$$\mathrm{e}^{-j\omega_k \tau} = \mathrm{e}^{-j(\omega_{k-1}+\Delta\omega)\tau} \tag{13}$$

$$= \mathrm{e}^{-j\omega_{k-1}\tau} \cdot \mathrm{e}^{-j\Delta\omega\,\tau}. \tag{14}$$

Thus, the exponential term for the k-th frequency value can be replaced with a recursive computation involving the previous term with index $k-1$ multiplied by the factor $\exp(-j\Delta\omega\,\tau)$, which is a constant calculated once only for each image source. This results in a significant reduction of the computational complexity in the ISM implementation, since otherwise the complex exponential function is executed K times for each one of a usually very large number of image sources.

2.2.2 Negative Reflection Coefficients

Given a specific absorption coefficient α characterizing any room surface, the corresponding reflection coefficient β follows from (4) as

$$\beta = \pm\sqrt{1-\alpha}\,. \tag{15}$$

The ISM implementation in Section 2.1 makes use of the positive definition of the β parameter, which results in a positive addition of the amplitude contribution from all the considered image sources. However, when used in conjunction with a frequency-domain implementation (i.e., Peterson's method), this method generates RIRs showing a distinctively anomalous (i.e., non-physical) tail decay, as depicted in Figure 1: the RIR is shifted upwards by a time-varying offset value.

An alternative approach is to use the negative definition of the parameter β in (15); this can be explained by considering the well-known angle-dependent reflection coefficient formula for a boundary

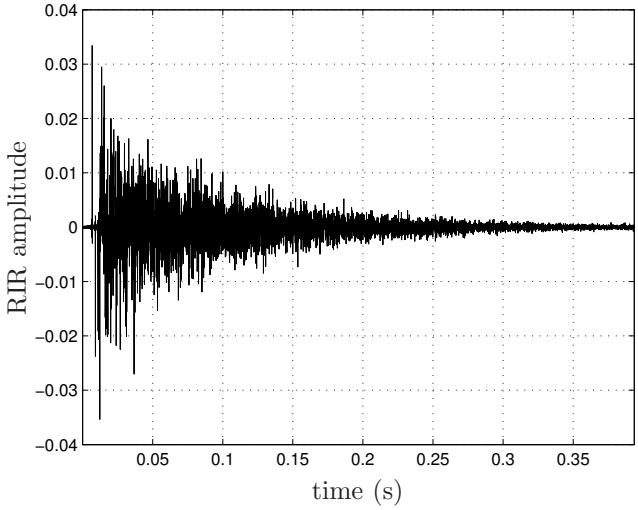

Figure 2: Typical RIR example obtained using frequency-domain ISM computations with *negative* reflection coefficients. This example was computed using the same environmental setup as for Figure 1.

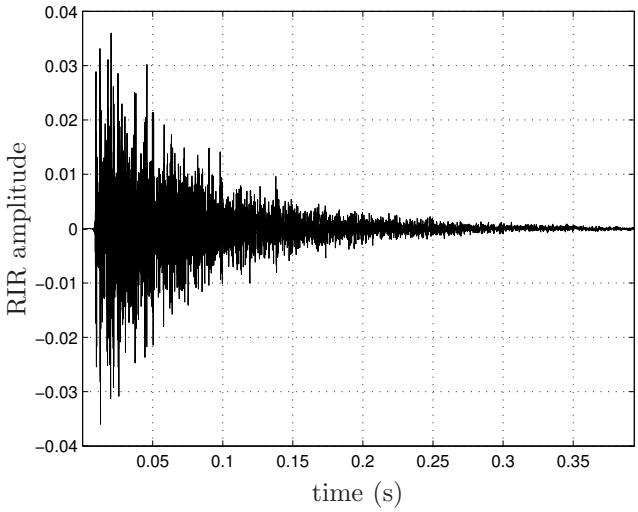

Figure 3: Typical example of real RIR, practically recorded in a room with reverberation time $T_{60} \approx 0.6$s.

with impedance ζ [50]:

$$\beta = \frac{\zeta \cos(\psi) - 1}{\zeta \cos(\psi) + 1}, \tag{16}$$

which can become negative for a certain range of incidence angle ψ.[2] Using a negative β coefficient implements a phase inversion (180° phase shift) upon every sound reflection on the enclosure surfaces, which effectively replicates the effect of a soft surface. This phase inversion typically represents an approximation of what happens in practice, where a sound-reflecting material is usually characterized in terms of sound absorption and phase shift by means of a complex reflection coefficient (acoustic impedance) [11, 50–53]. The main reason for using negative β coefficients however is that this results in better RIRs, as shown in Figure 2 which depicts an example of RIR resulting from this approach. By comparison with Figure 1, it can be seen that the frequency-domain ISM algorithm using negative reflection coefficients effectively achieves better practical results: the overall shape of the RIR is not biased as in Figure 1, and the computed RIR looks more "natural", i.e., compares well with a practically-measured transfer function recorded in real acoustic setting [54, 55], an example of which is displayed in

[2] Please note that (16) is only provided here in order to provide some insight into the physical meaning of using negative β coefficients. As mentioned in Section 2.1, the derivations in this work assume angle-independent reflection coefficients.

Figure 3. Thus, because this model can be seen as being more accurate in replicating the effects of a real acoustic environment, the ISM algorithm used in the remainder of this work will be based on (9) with, instead of the equation previously given in (6), the following definition of an image source's amplitude factor:

$$A(\boldsymbol{u},\boldsymbol{l}) = \frac{1}{4\pi\, d(\boldsymbol{u},\boldsymbol{l})} \cdot (-\beta_{x,1})^{|l-u|}(-\beta_{x,2})^{|l|}(-\beta_{y,1})^{|m-v|}(-\beta_{y,2})^{|m|}(-\beta_{z,1})^{|n-w|}(-\beta_{z,2})^{|n|}\,, \tag{17}$$

where $d(\cdot)$ is computed according to (8) and with the β parameters corresponding to the usual definition of sound reflection coefficients. Finally, it must be noted that this "negative coefficient approach" was previously studied and used by António *et al.* [18].

2.3 Energy Decay

Given a RIR $h(t)$ computed for a specific environment according to (10), the energy decay envelope $E(t)$, known in the literature as energy–time curve (ETC) or energy decay curve (EDC), can be computed using a normalized version of the Schroeder integration method [16, 56, 57]:

$$E(t) = 10 \cdot \log_{10}\left(\frac{\int_t^\infty h^2(\xi)\,\mathrm{d}\xi}{\int_0^\infty h^2(\xi)\,\mathrm{d}\xi}\right), \tag{18}$$

where $E(\cdot)$ is expressed in dB. The result from (18) can then be used as a basis for deriving an estimate of the reverberation time, such as T_{20} or T_{60} for instance.

A common problem with the integration method described here is that of slope biasing in the tail of the energy–time curve [57, 58]. In practical RIR measurements, the presence of background noise will bias the curve upwards in the late part of the energy decay, in case the upper integration limit in (18) is not chosen properly. It is however worth noting that this problem is here avoided altogether since the Schroeder integration method is applied directly to the computed RIR itself; the background noise is therefore nonexistent in the considered application.

3 Proposed Energy Decay Approximation

This section presents the developments leading to the proposed method for energy decay approximation. The derivations are first carried out in a two-dimensional (2D) (x, y)-plane, and the results are then extended to the 3D case. For clarity, some derivations are presented in the Appendix.

3.1 Theoretical Developments

Similarly to a ray-tracing model [59], each image source in the ISM technique described earlier can be seen as releasing a single sound "particle" (impulse) traveling in the direction of the receiver at the speed of sound. Each particle carries a unit amount of acoustic power, which decreases upon each reflection on a boundary surface according to its absorption coefficient, and as a function of the distance traveled to the receiver. These sound particles are then summed at the receiver at the corresponding time delays to create the impulse response.

As a result of the above consideration, the value of the impulse response $h(t)$ for a given time t corresponds to the addition of the sound amplitude contributions from all the image sources located on (or very close to) a circle of radius ϱ around the receiver, with $\varrho = c \cdot t$. This principle is demonstrated in Figure 4 for one quadrant of the (x, y)-plane. Dashed rectangles represent images of the considered room, which is shown as a shaded rectangle, and circles denote the image sources of interest for the current developments. The original source has been placed in the center of the room for illustration purposes. The analysis presented in this section is based on the assumption of a large radius ϱ, that is

$$\varrho \gg \max\{L_x, L_y, L_z\}\,. \tag{19}$$

For simplicity of the following derivations, it is also assumed that the receiver is located at the center of the coordinate system, as depicted in Figure 4. It is obvious that this assumption, together with the fact

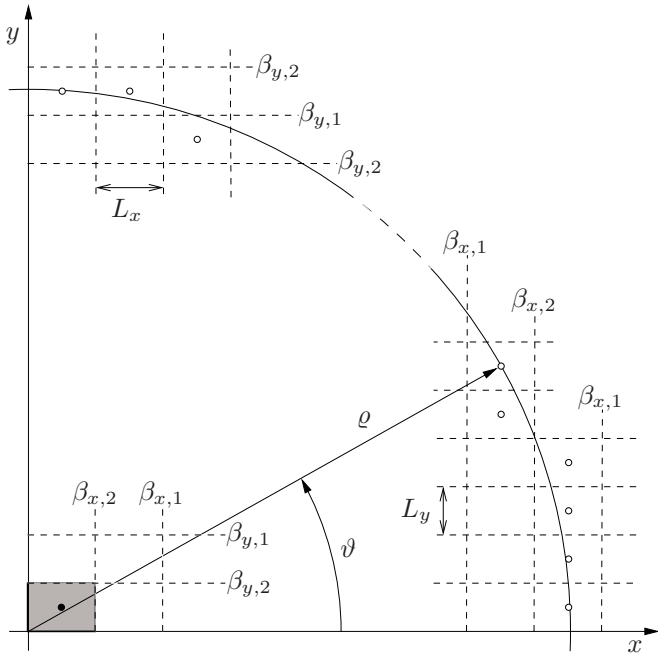

Figure 4: Two-dimensional representation of the ISM principle. Dashed lines represent the grid of images of the original room, which is displayed as a shaded rectangle. The β parameters indicate the reflection coefficients of the corresponding boundaries, and circles ($\circ$) represent the considered image sources.

that some image sources do not lie perfectly on the considered circle, typically lead to approximation errors that become negligible as the radius ϱ of the considered circle increases.

Let us now consider the i-th source located at an angle ϑ_i along the considered circle. Prior to reaching the receiver, its sound impulse traverses a number $W_x \triangleq W_x(\vartheta_i)$ of walls in the x-direction, and $W_y \triangleq W_y(\vartheta_i)$ walls in the y-direction, which can be determined in a straightforward manner on the basis of the known position of the source. Consequently, the power contribution $P_i(\cdot)$ made by the i-th source to the transfer function can be expressed as

$$P_i(\varrho, \vartheta_i) = \frac{(\beta_{x,1}^2)^{\frac{W_x}{2}} (\beta_{x,2}^2)^{\frac{W_x}{2}} (\beta_{y,1}^2)^{\frac{W_y}{2}} (\beta_{y,2}^2)^{\frac{W_y}{2}}}{(4\pi\varrho)^2} \,. \tag{20}$$

This expression makes use of the assumption that along the path to the receiver, the number of walls with coefficient $\beta_{x,1}$ (respectively $\beta_{y,1}$) is approximately equal to the number of walls with coefficient $\beta_{x,2}$ (respectively $\beta_{y,2}$), that is, approximately equal to half the number of walls $W_x/2$ (respectively $W_y/2$). This condition essentially becomes valid as the radius ϱ becomes large. Also, note that (20) uses squared-amplitude coefficients compared to (6) and (17), since the current developments are based on acoustic power rather than amplitude [11].

From this analysis, it follows that the value of the power impulse response $h_P(t)$ at time t, where the subscript in $h_P(\cdot)$ emphasizes the fact that the RIR is power-based, can be determined as

$$h_P(t) = \sum_{i \in \mathcal{I}_c} P_i(\varrho, \vartheta_i) \,, \tag{21}$$

with $\varrho \triangleq \varrho(t) = c \cdot t$, and $\mathcal{I}_c$ representing the index set of the sources located on, or very close to the considered circle. The basis of the proposed approximation is then to consider (21) as a Riemann sum that can be represented as the integral of a continuous function over the angle ϑ:

$$h_P(t) \cdot \Delta\vartheta = \sum_{i \in \mathcal{I}_c} P_i(\varrho, \vartheta_i) \cdot \Delta\vartheta \tag{22}$$

$$\approx \int_0^{2\pi} P(\varrho, \vartheta) \, \mathrm{d}\vartheta \,. \tag{23}$$

With (20), the solution to this integral leads to an analytical expression, which can then be extended to the 3D case. These derivations, which are provided in the Appendix, lead to the following estimate $\hat{h}_P(\cdot) \approx h_P(\cdot)$ of the power transfer function:

$$\hat{h}_P(t) = \frac{1}{8\varrho\,\bar{r}} \cdot \begin{cases} \frac{B_z}{\log\left(\frac{B_y}{B_x}\right)} \cdot \left(\mathrm{Ei}_1\big(\log\big(\frac{B_z}{B_x}\big)\big) + \log\big(\log\big(\frac{B_z}{B_x}\big)\big) - \mathrm{Ei}_1\big(\log\big(\frac{B_z}{B_y}\big)\big) - \log\big(\log\big(\frac{B_z}{B_y}\big)\big) \right) & \text{if } B_x \neq B_y \neq B_z\,, \\ \frac{B_z}{\log\left(\frac{B_z}{B}\right)} \cdot \left(\mathrm{Ei}_1\big(\log\big(\frac{B_z}{B}\big)\big) + \log\big(\log\big(\frac{B_z}{B}\big)\big) + \gamma \right) & \text{if } B_z = B_y \neq B_x \triangleq B \text{ or } B_z = B_x \neq B_y \triangleq B\,, \\ \frac{B - B_z}{\log\left(\frac{B}{B_z}\right)} & \text{if } B_z \neq B_x = B_y \triangleq B\,, \\ B & \text{if } B_x = B_y = B_z \triangleq B\,, \end{cases} \tag{24}$$

where $\varrho \triangleq \varrho(t) = c \cdot t$ as defined earlier, and with the following definitions:

$$B_x = (\beta_{x,1}\beta_{x,2})^{\varrho/L_x}, \tag{25}$$

$$B_y = (\beta_{y,1}\beta_{y,2})^{\varrho/L_y}, \tag{26}$$

$$B_z = (\beta_{z,1}\beta_{z,2})^{\varrho/L_z}, \tag{27}$$

$$\bar{r} = \frac{L_x + L_y + L_z}{3}\,, \tag{28}$$

with $\gamma = 0.5772157\ldots$ the Euler–Mascheroni constant, and with $\mathrm{Ei}_n(\xi)$ denoting the exponential integral, defined in terms of the incomplete Gamma function for $\xi \neq 0$ as [60]

$$\mathrm{Ei}_n(\xi) = \xi^{n-1} \cdot \Gamma(1-n, \xi)\,. \tag{29}$$

The exponential integral is a common mathematical function, implemented for instance as `expint` in MATLAB, `Ei` in MAPLE, and `ExpIntegralEi` in MATHEMATICA; alternatively, it can also be evaluated numerically using a power series representation [60, 61].

An estimate of the energy decay curve can then be computed on the basis of (18) as

$$\widehat{E}(t) = 10 \cdot \log_{10}\left(\frac{\int_t^\infty \hat{h}_P(\xi)\,\mathrm{d}\xi}{\int_0^\infty \hat{h}_P(\xi)\,\mathrm{d}\xi} \right). \tag{30}$$

In practice, the integrals on the right-hand side of (30) can be replaced with a Riemann sum as follows:

$$\int_t^\infty \hat{h}_P(\xi)\,\mathrm{d}\xi \approx T \cdot \sum_{i=0}^{\infty} \hat{h}_P(t + iT)\,. \tag{31}$$

with an appropriate discretisation step T. As specified by fundamental principles of numerical integration, the validity of this approximation depends upon the function $\hat{h}_P(\cdot)$ being smooth and bounded in the considered interval, which is supported by the plots in Figure 5; it is also shown in the next section that (31) indeed holds true for the type of function defined in (24). Thus, the estimated energy–time curve can be finally computed according to (30) and (31), and for $t > t_0$, as

$$\widehat{E}(t) \approx 10 \cdot \log_{10}\left(\frac{\sum_{i=0}^{\infty} \hat{h}_P(t + iT)}{\sum_{i=0}^{\infty} \hat{h}_P(t_0 + iT)} \right), \quad t > t_0. \tag{32}$$

The introduction of the parameter t_0 in (32) can be explained as follows. According to the assumptions made in this work, the EDC approximation is expected to be inaccurate for small ϱ values, that is, for $t \to 0$. Therefore, the approximation formula in (32) can be considered as relevant only for values of t greater than a specific threshold, denoted here as t_0. Section 4.1 will provide more detail regarding an appropriate setting of the t_0 parameter for numerical simulation purposes.

3.2 Discussion

Two distinct sources of error can be identified in relation to the expression proposed in (32). As mentioned above, the assumption of a large radius ϱ will typically lead to a poor approximation of the true EDC as

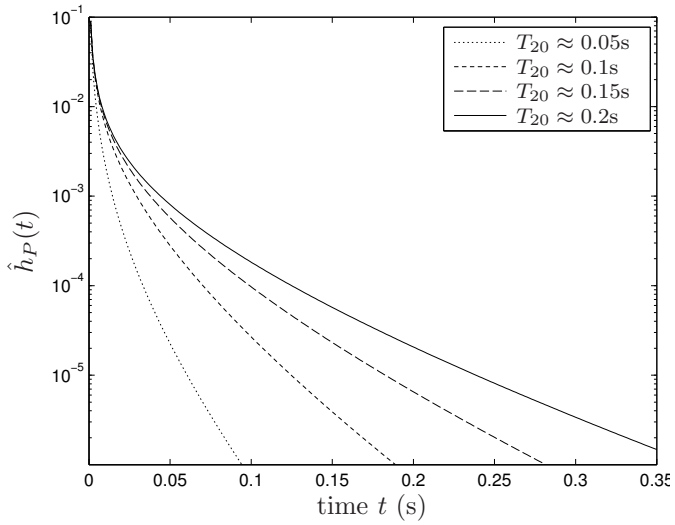

Figure 5: Numerical evaluation of the approximate power transfer function $\hat{h}_P(\cdot)$, for various levels of reverberation, i.e., various β coefficients, in a room with dimensions $\boldsymbol{r} = [4.0\text{m}, 5.0\text{m}, 2.9\text{m}]^{\mathrm{T}}$.

	Figure 6	Figures 7 and 8
$\boldsymbol{r}$ (m)	$[4.0, 5.0, 2.9]^{\mathrm{T}}$	$[3.2, 4.0, 2.7]^{\mathrm{T}}$
$\boldsymbol{p}_s$ (m)	$[1.5, 1.0, 1.0]^{\mathrm{T}}$	$[1.1, 1.0, 1.2]^{\mathrm{T}}$
$\boldsymbol{p}_r$ (m)	$[3.5, 3.8, 1.9]^{\mathrm{T}}$	$[2.0, 3.0, 2.0]^{\mathrm{T}}$
F_s (Hz)	16000	16000

Table 1: Environmental parameter setup for the results presented in Figures 6, 7 and 8.

$t \to 0$. In addition, the parameter t_0 effectively introduces an additive error term Δ in the denominator of (32). This error term is related to the missing initial part of the denominator integration:

$$\Delta \triangleq \Delta(t_0) = \int_0^{t_0} h^2(\xi)\,\mathrm{d}\xi\,, \tag{33}$$

which is impossible to calculate on the basis of $\hat{h}_P(\cdot)$ due to the intrinsically poor approximation of $h^2(\cdot)$ provided by $\hat{h}_P(\cdot)$ for values of t below the threshold t_0. The error term Δ is however independent of the time variable t, and thus potentially creates a constant offset in the EDC approximation curve $\widehat{E}(\cdot)$. These two distinct effects (i.e., poor approximation at low values of t and offset due to Δ) will be illustrated more specifically in the following section.

Finally, the infinite sums in (32) have to be truncated to a finite set of indices in practice. As shown in Figure 5, the function $\hat{h}_P(\xi)$ tends towards 0 very quickly as ξ increases, and as a result, the summation can be terminated relatively early while still providing a good approximation for practical purposes.

4 Experimental Results

4.1 Numerical Evaluations

This section provides some examples of the results obtained with the proposed approximation method. Figure 6 considers a typical enclosure setup, the details of which are provided in Table 1, for three different levels of reverberation and assuming uniform reflection coefficients β for all enclosure surfaces. The solid lines represent the energy decay lines computed via (18) on the basis of the impulse responses simulated with the ISM technique of Section 2.2. Circle markers ($\circ$) indicate the values obtained via (24) and (32) computed at several discrete values of time. Figure 7 shows similar results, obtained using a different room setup (see Table 1) in the case of non-uniform wall reflection coefficients, the values of which are given in Table 2. Note that the curves in Figure 7 correspond to a scenario where a pair

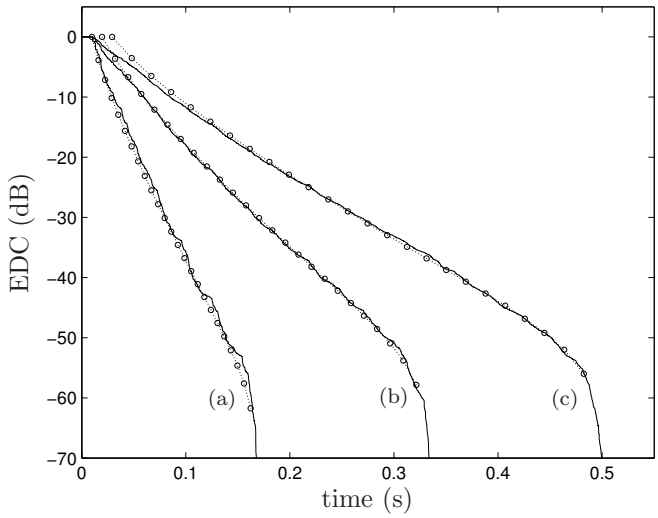

Figure 6: Examples of energy decay curves with uniform reflection coefficients. (a) $\beta = 0.669$ ($T_{20} \approx 0.05$s), (b) $\beta = 0.831$ ($T_{20} \approx 0.1$s), and (c) $\beta = 0.889$ ($T_{20} \approx 0.15$s). Solid lines represent $E(t)$ obtained from ISM computations; circles ($\circ$) indicate values obtained with the proposed approximation $\widehat{E}(t)$.

Curve	$\beta_{x,1}$	$\beta_{x,2}$	$\beta_{y,1}$	$\beta_{y,2}$	$\beta_{z,1}$	$\beta_{z,2}$
(a) $T_{20} \approx 0.05$s	0.032	0.032	0.548	0.548	0.837	0.837
(b) $T_{20} \approx 0.1$s	0.675	0.675	0.787	0.787	0.915	0.915
(c) $T_{20} \approx 0.15$s	0.802	0.802	0.866	0.866	0.945	0.945

Table 2: Values of reflection coefficients for each boundary surface, for the curves displayed in Figure 7.

of opposing walls are significantly different in reflectivity compared to other surfaces; this specific case was found to lead to discrepancies between estimated and measured reverberation time in Allen and Berkley's original publication [1].

Despite several simplifying assumptions made in this work, Figures 6 and 7 demonstrate that the proposed EDC approximation technique is quite accurate when estimating the energy decay in RIRs produced with the image source method. The overall decay rate, as well as the shape (curvature) of the decay line for non-uniform β coefficients, matches the practical results relatively well. With respect to the effects of the large-radius assumption mentioned previously, these numerical results clearly illustrate the discrepancy between the approximated and the practical results for low values of t, which appears as a slight upward bias at the beginning of the approximation curves. This effect becomes more pronounced for larger reverberation times, but remains nonetheless relatively negligible for practical purposes.

It must be noted here that the results in Figures 6 and 7 have been obtained with an optimal setting of the variable t_0 (i.e., the time lag of the first value on the approximation curves). This effectively compensates for the constant error term mentioned in Section 3.2, and thus enables a better visual comparison of the displayed results. In practice, a non-optimal setting of t_0 would hence result in a slight offset in the corresponding EDC approximation curve. It was found empirically that choosing $t_0 = 1.5 \cdot \|\boldsymbol{p}_s - \boldsymbol{p}_r\|/c$ or $t_0 = 1.5 \cdot \bar{r}/c$ achieves a relatively good match for a large number of enclosure sizes and reflection coefficients. Note also that this offset only has a marginal effect when assessing the overall energy decay of the considered RIR or when measuring quantities such as the reverberation time. If necessary, Ref. [62] further demonstrates the practical accuracy of the proposed method by means of more extensive numerical simulations. The development of a more accurate definition of the parameter t_0 is left as matter for further research, as a result of its minor impact on the considered analysis.

Finally, the different plots in Figure 8 provide an insight into the influence of the integration interval T in (32). This figure displays the approximation results for three different integration lengths, computed with $T_{20} \approx 0.1$s and for an environmental setup as described in Table 1. These results clearly demonstrate the fact that the accuracy of the approximation remains very good regardless of the number of points considered along the curve, which corroborates the validity of the approximation in (31). If necessary, the

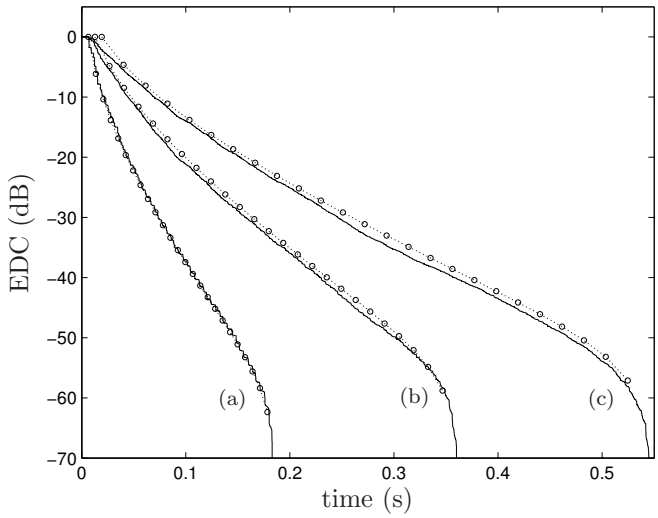

Figure 7: Examples of energy decay curves with non-uniform reflection coefficients, for (a) $T_{20} \approx 0.05$s, (b) $T_{20} \approx 0.1$s, and (c) $T_{20} \approx 0.15$s. See Table 2 for corresponding β values. Solid lines represent $E(t)$ obtained from ISM computations; circles ($\circ$) are values from the proposed approximation $\widehat{E}(t)$.

computations can hence be made more efficient by reducing the number of points on the approximation curve, with only a marginal reduction of the representation accuracy.

4.2 Discussion

The accuracy of the proposed EDC approximation method was tested and confirmed for a number of scenarios considering different enclosure volumes, source and sensor positions, uniform and non-uniform reflection coefficients, and so forth. As shown in Section 4.1 and despite several simplifying assumptions made in Section 3.1, it is interesting to see that the discrepancy between the approximation and the practical results remains relatively small even for very low radius values.

The proposed approximation technique hence enables designers to efficiently investigate the acoustical characteristics of a simulated room without the need to generate the RIRs of interest. Due to the considerable computational demands usually associated with the ISM technique, this consequently represents a substantial reduction of the resulting computation time during the analysis.

5 Application Example

This section describes a possible application of the developments presented in Section 3. It considers the example of how to accurately determine the value of reflection coefficients in order to achieve a given reverberation time in the considered virtual enclosure. As mentioned earlier, the image source method constitutes a major tool in the process of assessing the performance of many acoustic signal processing algorithms operating in reverberant environments. Typically, the image source model is used to test the algorithms under consideration, in order to determine their robustness against increasing levels of environmental reverberation. For the sake of a fair and consistent comparison, it is hence important to ensure that the algorithms are assessed using the same measure of reverberation across all test simulations.

In recent literature [37, 42, 43, 46], a common approach to this assessment process involves the use of the well-known Sabine or Eyring formulae to determine reflection coefficients given a desired reverberation time. The RIRs are then generated with the ISM technique (possibly using Peterson's improvement [49] to Allen and Berkley's original method [1]), and the performance results are finally displayed against the reverberation time that was used as a basis for the derivation of the β coefficients. As shown in the next subsection, this approach is however subject to significant inaccuracies, especially for a non-uniform definition of the reflection coefficients in the considered room. This discrepancy between predicted and

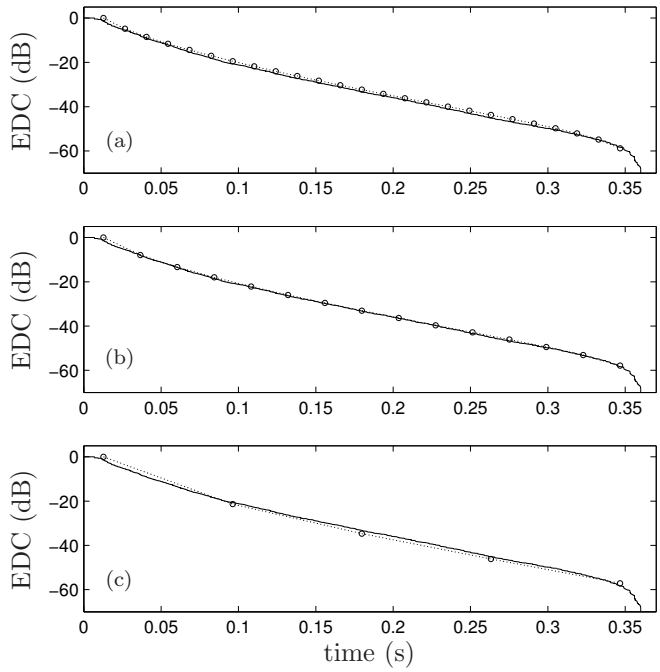

Figure 8: EDC approximation results with varying integration intervals, for $T_{20} \approx 0.1$s and using the same setup as in Figure 7. The integration interval is defined as (a) $T = 0.014$s, (b) $T = 0.024$s, and (c) $T = 0.083$s.

measured reverberation time was also observed in several publications of the acoustics literature [18, 63]. As a result, it is very likely that the performance outcomes are ultimately presented for a reverberation level that does not correspond to the actual testing conditions.

An alternative approach chosen by various authors is to present the performance results versus the reflection coefficient itself [30–33, 36]. However, contrary to more intuitive measures such as the reverberation time T_{60} for instance, this way of presenting results does not provide any real insight into the practical reverberation characteristics of the considered environment.

The EDC approximation method proposed in Section 3 can be used efficiently to alleviate these problems altogether, by providing an accurate and unequivocal correspondence between the reflection coefficients β and the resulting reverberation time.

5.1 Reverberation Time Prediction Formulae

Let us consider an enclosure with dimensions $L_x \times L_y \times L_z$ where each boundary surface is assigned an absorption coefficient as follows: $\alpha_{x,1} \triangleq \alpha_1$, $\alpha_{x,2} \triangleq \alpha_2$, $\alpha_{y,1} \triangleq \alpha_3$, and so forth. For the sake of simplicity in the following presentation, it will be assumed that the coefficients are identical for each pair of opposing walls, i.e.,

$$\alpha_1 = \alpha_2 \triangleq \alpha \cdot w_x \,, \tag{34}$$

$$\alpha_3 = \alpha_4 \triangleq \alpha \cdot w_y \,, \tag{35}$$

$$\alpha_5 = \alpha_6 \triangleq \alpha \cdot w_z \,, \tag{36}$$

where w_x, w_y and w_z are absorption weighting factors for the walls in the x, y and z dimensions, respectively. This representation allows for clearer derivations related to non-uniform absorption coefficients, which can be simply characterized in terms of the single parameter α used in conjunction with the weights vector $\boldsymbol{w} = [w_x, w_y, w_z]$ or, alternatively, the weighting ratio $w_x : w_y : w_z \equiv \alpha_1 : \alpha_3 : \alpha_5$. It must be stressed however that this restriction is without loss of generality, since the following derivations can be easily extended to the case where all coefficients have different values.

As mentioned above, previous literature works have made an extensive use of well-known formulae by Sabine or Eyring in order to determine the value of absorption coefficients achieving a desired reverberation time T_{60}. Many other formulae can also be found in the acoustics literature [64]. The present work will investigate several of these definitions, namely Sabine's formula [65]:

$$T_{60,\mathrm{S}}(\alpha, \boldsymbol{w}) = \frac{0.161 \cdot V}{\sum_{i=1}^{6} S_i \, \alpha_i} \,, \tag{37}$$

Eyring's formula [66]:

$$T_{60,\mathrm{E}}(\alpha, \boldsymbol{w}) = \frac{0.161 \cdot V}{-S \cdot \log\bigl(1 - \sum_{i=1}^{6} S_i \, \alpha_i / S\bigr)} \,, \tag{38}$$

Millington–Sette's formula [67, 68]:

$$T_{60,\mathrm{MS}}(\alpha, \boldsymbol{w}) = \frac{0.161 \cdot V}{-\sum_{i=1}^{6} S_i \cdot \log(1 - \alpha_i)} \,, \tag{39}$$

and Fitzroy's formula [69]:

$$T_{60,\mathrm{F}}(\alpha, \boldsymbol{w}) = \frac{0.161 \cdot V}{S^2} \cdot \left(\frac{-2 L_y L_z}{\log\bigl(1 - (\alpha_1 + \alpha_2)/2\bigr)} - \frac{2 L_x L_z}{\log\bigl(1 - (\alpha_3 + \alpha_4)/2\bigr)} - \frac{2 L_x L_y}{\log\bigl(1 - (\alpha_5 + \alpha_6)/2\bigr)} \right), \tag{40}$$

where S is the total surface area of the enclosure, and S_i, $i = 1, \ldots, 6$, are the surface areas of each individual wall.

Given a specific weighting vector $\boldsymbol{w}$, the problem of determining the value of α that achieves a desired reverberation time, denoted here as $T_{60,\mathrm{des}}$, can be seen as a nonlinear optimization problem:

$$\alpha_{(\cdot)} = \arg \min_{\alpha \in [0,1]} \bigl| T_{60,\mathrm{des}} - T_{60,(\cdot)}(\alpha, \boldsymbol{w}) \bigr| \,, \tag{41}$$

where the subindex in $\alpha_{(\cdot)}$ indicates one of the formulae in (37)–(40). This minimization problem can be solved numerically using, for instance, a golden section search algorithm with parabolic interpolation [70], as implemented by the function `fminbnd` in MATLAB.

In the following, the reverberation time will be characterized using T_{20} rather than T_{60}, where T_{20} is here defined according to the classical formula as the time required by the RIR energy $E(\cdot)$, defined by (18), to decay from -5dB to -25dB:

$$T_{20} = E^{-1}(-25) - E^{-1}(-5) \,, \tag{42}$$

where $E^{-1}(\cdot)$ represents the inverse of the EDC function $E(\cdot)$, i.e., $E^{-1}(\xi)$ corresponds to the time lag t_ξ for which $E(t_\xi) = \xi$. The reason for using T_{20} instead of T_{60} is solely in order to reduce the computational load when measuring the reverberation time in simulated RIRs in Section 5.3. As a result of using non-uniform absorption coefficients, the EDC will typically display a non-negligible curvature, which in turn means that it would be inaccurate to measure T_{60} by interpolation of the initial decay slope in the energy–time curve. By using the parameter T_{20} instead of T_{60}, it is hence only necessary to compute the EDC down to -25dB (or slightly below), which involves significantly less image sources during the RIR computations. It is however emphasized once more that this substitution is only to facilitate the display of the comparison results; the present developments are valid for any quantity of interest defined on the basis of the EDC, such as T_{60}, T_{30}, early decay time, etc.

Because the various classical formulae given in (37)–(40) have been derived on the basis of the average sound absorption within the room, or assuming a homogenous spatial distribution of the sound power (diffuse field), they implicitly assume a linear energy decay in the resulting EDC. It follows that for these cases, the T_{20} value is simply defined as

$$T_{20,(\cdot)}(\alpha, \boldsymbol{w}) = T_{60,(\cdot)}(\alpha, \boldsymbol{w}) / 3 \,. \tag{43}$$

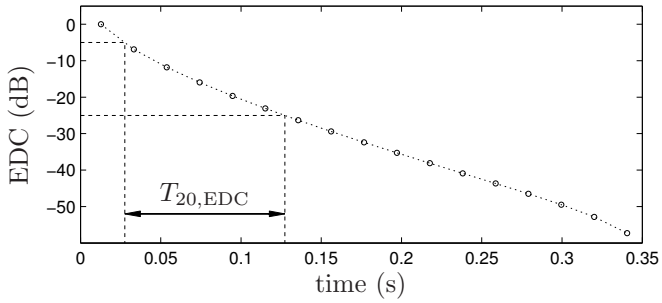

Figure 9: Determining the reverberation time T_{20} from the proposed EDC approximation results.

5.2 Reverberation Time Prediction using EDC Approximation

The EDC approximation method described in Section 3 can be used in a straightforward manner in order to determine the value of the absorption coefficients α_i achieving a desired reverberation time. As depicted in Figure 9, the numerical value of $T_{20,\mathrm{EDC}}(\alpha, \boldsymbol{w})$ for a given absorption parameter α and weighting vector $\boldsymbol{w}$ is obtained by simply computing the EDC approximation curve according to (24) and (32), and determining the corresponding reverberation time directly from it. A first-order interpolation between points is also used in this process to refine the estimate. Using this approach, the α parameter value that yields the desired reverberation time $T_{20,\mathrm{des}}$ is finally obtained as the solution to the same optimization problem as given previously in (41).

5.3 Numerical Results

In this section, the accuracy of the proposed EDC approximation method is assessed and compared to the currently established formulae using the following test. Given a specific target reverberation time $T_{20,\mathrm{des}}$ and weight vector $\boldsymbol{w}$, the absorption parameter α is determined via (41) for each of the considered reverberation time approximation formulae in (37)–(40), as well as the proposed technique as described in Section 5.2. The image source method is then used with the resulting α values to generate a number of RIRs in the considered environment, and the "true" reverberation time value $T_{20,\mathrm{meas}}$ is then measured directly from the RIR using the definition in (42). For the proposed EDC approximation method, the frequency-domain ISM algorithm with negative reflection coefficients is used to simulate the RIRs, as described in Section 2.2. The original implementation of Allen and Berkley (positive reflection coefficients) is used in conjunction with the classical reverberation prediction methods, since this represents the approach that is currently widely used in the literature. The resulting error ε is then simply defined, for each method, as

$$\varepsilon = \left| T_{20,\mathrm{des}} - T_{20,\mathrm{meas}} \right| . \tag{44}$$

For a given $T_{20,\mathrm{des}}$, this process is repeated for a total of 30 randomly selected source and receiver positions, in each of eight different rooms, ranging from a size of 2.7m × 3m × 2.5m (volume of 20.25m^3) to 6.53m × 7.26m × 4.27m (volume of 202.5m^3). Finally, Figure 10 presents the distribution of the resulting 240 error values obtained with each method under consideration and for each desired value of reverberation time, in the form of a median and interquartile-range plot. Each plot in this figure presents the results obtained with a different ratio of absorption coefficient $w_x : w_y : w_z$, which represent various levels of non-uniformity amongst the specific absorption coefficients, including the special case of equal coefficient values.

Note that these graphs only display results for simulation conditions that are considered to be "physically" attainable: assuming non-uniform wall coefficients results in some surfaces having non-total absorption (i.e., $\alpha \neq 1$), which in turn will generate a minimum, non-vanishing amount of reverberation in the enclosure. In other words, some values of T_{20} below a certain limit will be impossible to generate given a specific room $\boldsymbol{r}$ and non-uniform weighting vector $\boldsymbol{w}$. Error results below this limit are not displayed in Figure 10, since it is known that such simulation conditions will yield large errors by definition.

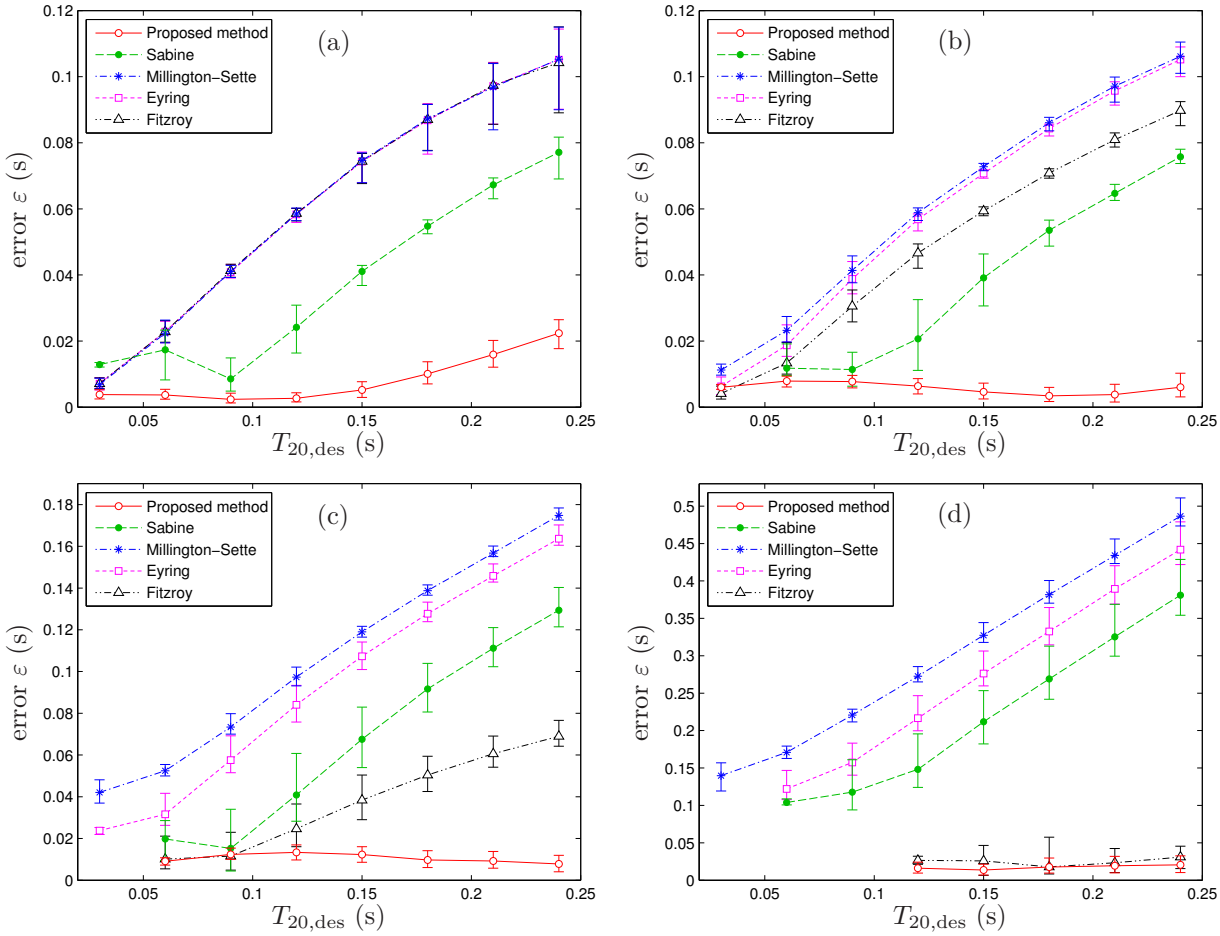

Figure 10: Distribution of the resulting T_{20} prediction error, for each considered method: error bars represent the interquartile range (25-th and 75-th percentile, respectively), whereas center markers correspond to the median of the distribution. The absorption coefficient ratio $w_x : w_y : w_z$ is defined as (a) 1.0 : 1.0 : 1.0, (b) 1.0 : 0.8 : 0.6, (c) 1.0 : 0.6 : 0.3, and (d) 1.0 : 0.5 : 0.1.

5.4 Discussion

Figure 10 shows that the proposed EDC approximation method is able to maintain a low level of error for the considered test scenarios, and provides the best results among the five considered T_{20} prediction methods. While most of the other methods display a large resulting error, some of them also have a widely spread distribution of their error values (large interquartile range). This results from the fact that the accuracy of such methods varies significantly as a function of the considered room size: typically, such methods will generate larger errors as the size of the considered enclosure becomes smaller. On the other hand, the technique proposed in this paper is able to maintain a relatively concentrated distribution throughout the simulations, which highlights the fact that it provides good results for a wide range of room volumes.

When considering the various formulae described in Section 5.1, it can be seen on the basis of Figure 10 that none of them provides a consistently low level of estimation error. On the contrary, for most of them the error becomes larger as the desired value of reverberation time increases. This consequently raises some doubts regarding their usefulness in predicting the reverberation level in simulated RIRs, which can be regarded as being of some concern since several publications have been published which provide assessment results based on these formulae [42, 43, 45–47]. On the other hand, the results obtained with the proposed EDC approximation are consistently better than any other method. For the proposed technique, the largest error in fact occurs in the case of uniform absorption coefficients and large T_{20} values, as shown in the top left-hand graph of Figure 10. This is a direct result of the upwards biasing problem

of the EDC approximation curve, as observed in Section 4.1, and which becomes more pronounced when the "true" EDC is close to a straight line, i.e., in the case of uniform α coefficients. Finally, it is also interesting to notice how Fitzroy's formula becomes more accurate as the level of non-uniformity in the absorption coefficients increases. This formula was empirically derived to account for a non-uniform definition of the absorption coefficients in the room [69], which explains this effect in the results of Figure 10.

In summary, the results displayed in Figure 10 demonstrate that the proposed EDC approximation method can be used as an efficient tool in the process of analyzing acoustic signal processing algorithms, evaluated on the basis of image source simulations. The results in this section show that the accuracy of this approximation technique allows to generate simulated RIRs with a reverberation time within a small percentage of the desired T_{20} value. Consequently, this approach can be efficiently implemented in replacement of the only current solution to this specific issue, i.e., a lengthy trial-and-error process using time-consuming ISM simulations.

6 Conclusion

This paper proposes a new method for approximating the energy decay in simulated room impulse responses. It was demonstrated that the proposed technique provides an accurate approximation of the power transfer function generated on the basis of a modified version of the widely-used image source model: this modification involves computations in the frequency domain to account for non-integer time delays, and implements a phase inversion upon each reflection of the sound rays off the enclosure boundaries. As a result, the simulated transfer functions look more natural and constitute a better representation of practical impulse response measurements, compared to Allen and Berkley's original method.

This paper also considered an application example where the proposed approximation method is proving useful, namely when implemented during the testing phase of single or multi-channel acoustic signal processing algorithms. Using the approach described in this work, it is possible to accurately determine the value of absorption coefficients required to achieve a desired reverberation level in the environment, which results in two valuable benefits. Since this process can now be carried out without the need for time-consuming image method simulations, the time required for the design or analysis of algorithms using the image source method is drastically reduced. Also, it ensures that the algorithms are tested, and performance results are presented, versus a reverberation level that corresponds more or less exactly to what was used during the algorithm simulations. This ensures uniformity of the results when comparing the algorithm performance presented by various authors in different publications.

Finally, the technique described in this work can also be of potential interest in other fields of acoustic design and engineering, such as, for instance, virtual auditory environments, perceptual acoustics, interactive architectural acoustics, reverberant sound field modeling, etc.

Acknowledgments

This work was supported by National ICT Australia (NICTA). NICTA is funded through the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

Appendix

This Appendix contains the derivations leading to the result of (24), starting from a two-dimensional consideration of the problem (see Figure 4). From (20), the contribution to the power RIR from an image source, located at angle ϑ on a circle with radius ϱ, is

$$P(\varrho,\vartheta) = \frac{(\beta_{x,1}\beta_{x,2})^{W_x}(\beta_{y,1}\beta_{y,2})^{W_y}}{(4\pi\varrho)^2}\,, \tag{45}$$

with $\varrho \triangleq \varrho(t) = c\cdot t$. On the basis of Figure 4, and using a first-order approximation, the number of boundary surfaces between the source and the receiver, in the x and y dimensions, are defined as follows (for $\vartheta \in [0,\pi/2]$):

$$W_x = \frac{\varrho}{L_x}\cdot\left(1-\frac{2\vartheta}{\pi}\right), \tag{46}$$

$$W_y = \frac{\varrho}{L_y}\cdot\frac{2\vartheta}{\pi}\,. \tag{47}$$

From (23), the estimated power transfer function $\hat{h}_P(\cdot)$ results as follows:

$$h_P(t) \approx \hat{h}_P(t) = \frac{4}{\Delta\vartheta}\int_0^{\frac{\pi}{2}} P(\varrho,\vartheta)\,\mathrm{d}\vartheta\,, \tag{48}$$

where the analysis has been restricted to a quarter of the circle, for $\vartheta \in [0,\pi/2]$, as a result of the symmetry in the problem definition. The angular variable $\Delta\vartheta$ can be determined as

$$\Delta\vartheta = \frac{2\pi}{N_s}\,, \tag{49}$$

where N_s corresponds to the total number of image sources considered on the circle (or located very close to it). From Figure 4, it can be seen that for $\vartheta \to 0$, the average distance between the image sources on the circle approaches L_y; similarly, for $\vartheta \to \pi/2$, it approaches L_x. Consequently, the parameter N_s is here defined as the circle's circumference divided by the average room dimension, i.e.,

$$N_s = \frac{2\pi\varrho}{\bar{r}}\,, \tag{50}$$

with $\bar{r} = (L_x + L_y)/2$. The value of $\Delta\vartheta$ then follows as $\Delta\vartheta = \bar{r}/\varrho$, which leads to the final power RIR approximation for the 2D setting:

$$\hat{h}_P(t) = \frac{4\varrho}{\bar{r}}\int_0^{\frac{\pi}{2}} P(\varrho,\vartheta)\,\mathrm{d}\vartheta\,. \tag{51}$$

In effect, the various approximations used in the developments so far take into account the *average* distribution of the image sources located in the vicinity of the considered circle, rather than only those sources located *exactly* at a radius ϱ from the receiver.

The extension of the above derivations to the 3D case is based on the same approach, as illustrated in Figure 11. The total power in the RIR for a given time lag t corresponds to the sum of the contributions from image sources located on, or very close to a sphere with corresponding radius $\varrho = c\cdot t$, centered on the origin. With the introduction of the polar angle $\varphi \in [0,\pi]$ (see Figure 11), the corresponding Riemann sum then follows as

$$\sum_{i\in\mathcal{I}_s} P_i(\varrho,\vartheta_i,\varphi_i)\cdot\Delta\vartheta\Delta\varphi = \int_0^{\pi}\int_0^{2\pi} P(\varrho,\vartheta,\varphi)\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,, \tag{52}$$

where $\mathcal{I}_s$ represents the index set of the sources located on the sphere. The main issue with (52) is regarding the computation of the term $\Delta\vartheta\Delta\varphi$. As demonstrated in Figure 11, simply defining uniformly spaced points in ϑ and φ (as done in the 2D derivations), leads to a grid on the sphere that obviously does not correspond to the distribution of image sources in the ISM technique. The "true" distribution of the

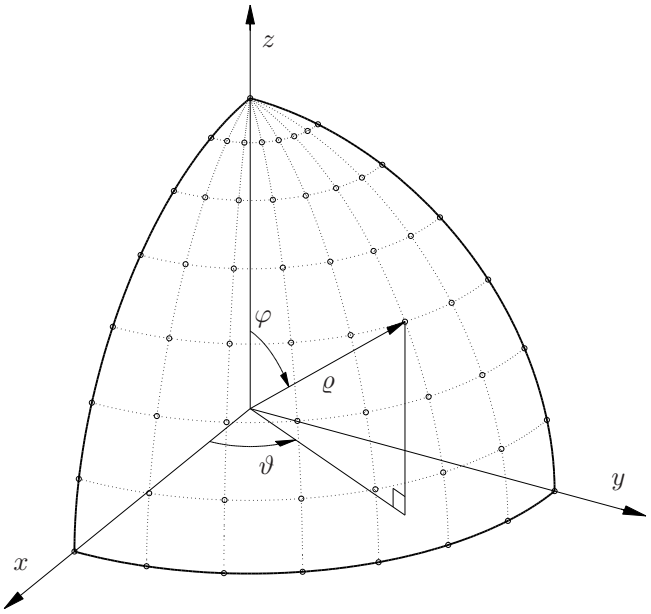

Figure 11: Definition of variables in a spherical coordinate system. The grid of points on the sphere is *not* representative of the 3D distribution of image sources in the ISM technique.

sources corresponds to a nonlinear spherical grid, which is quite complex to derive mathematically from the grid of mirror rooms defined by the image source method. One possible approach for a 3D scenario is to update the expression for the 2D case obtained in (51); this approximation has been empirically found to be sufficiently accurate, due mainly to the normalization process carried out in (30) when computing the energy decay values. Again limiting the analysis to one eighth of the sphere, (51) becomes

$$\hat{h}_P(t) = \frac{8\varrho}{\bar{r}} \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} P(\varrho, \vartheta, \varphi)\, \mathrm{d}\vartheta\, \mathrm{d}\varphi\,, \tag{53}$$

where $\bar{r}$ now includes the room dimension for the z coordinate:

$$\bar{r} = \frac{L_x + L_y + L_z}{3}\,, \tag{54}$$

and with the 3D extension of the power amplitude coefficient and number of walls:

$$P(\varrho, \vartheta, \varphi) = \frac{(\beta_{x,1}\beta_{x,2})^{W_x} (\beta_{y,1}\beta_{y,2})^{W_y} (\beta_{z,1}\beta_{z,2})^{W_z}}{(4\pi\varrho)^2}\,, \tag{55}$$

$$W_x = \frac{\varrho}{L_x} \cdot \left(1 - \frac{2\vartheta}{\pi}\right) \cdot \frac{2\varphi}{\pi}\,, \tag{56}$$

$$W_y = \frac{\varrho}{L_y} \cdot \frac{2\vartheta}{\pi} \cdot \frac{2\varphi}{\pi}\,, \tag{57}$$

$$W_z = \frac{\varrho}{L_z} \cdot \left(1 - \frac{2\varphi}{\pi}\right). \tag{58}$$

Inserting (55)–(58) into (53) and analytically solving the double integral finally leads to the result in (24). As shown in this equation, the resulting expression can be simplified further in case equality holds between the value of certain parameters.

In Equation (24), rotation of the coordinates x, y and z can be applied in order to avoid cases that would otherwise lead to negative arguments for the exponential integral $\mathrm{Ei}_1(\cdot)$ or for the logarithm function. For instance, if a given environmental setup leads to $B_z = B_x \neq B_y$ with $B_y > B_z$, Equation (24) involves the logarithm of a negative argument, which in turn generates complex values for the power transfer function. Such a case can be dealt with by swapping the coordinates y and z, i.e., by carrying

out the computations with the newly defined variables $B'_y \triangleq B_z$ and $B'_z \triangleq B_y$. As a result, this leads to a case where $B'_z \neq B'_y = B_x$, which can be successfully handled by Equation (24). This type of coordinate rotation effectively corresponds to a simple re-definition of the axis coordinates (here, e.g., $y' \triangleq z$ and $z' \triangleq y$) with respect to the spherical coordinate system shown in Figure 11, and does not change the generality or validity of the main result provided in (24).

References

[1] J. Allen and D. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, April 1979.

[2] J. Borish. Extension of the image model to arbitrary polyhedra. *Journal of the Acoustical Society of America*, 75(6):1827–1836, June 1984.

[3] H. Lee and B.-H. Lee. An efficient algorithm for the image model technique. *Applied Acoustics*, 24(2):87–115, 1988.

[4] U. Kristiansen, A. Krokstad, and T. Follestad. Extending the image method to higher-order reflections. *Applied Acoustics*, 38(2–4):195–206, 1993.

[5] M. Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *Journal of the Acoustical Society of America*, 86(1):172–178, July 1989.

[6] S. Dance, J. Roberts, and B. Shield. Computer prediction of sound distribution in enclosed spaces using an interference pressure model. *Applied Acoustics*, 44(1):53–65, 1995.

[7] M. Hodgson. On the accuracy of models for predicting sound propagation in fitted rooms. *Journal of the Acoustical Society of America*, 88(2):871–878, August 1990.

[8] B. Gibbs and D. Jones. A simple image method for calculating the distribution of sound pressure levels within an enclosure. *Acustica*, 26:24–32, 1972.

[9] Y. Lam. Issues for computer modelling of room acoustics in non-concert hall settings. *Acoustical Science and Technology*, 26(2):145–155, 2005.

[10] M. Kleiner, R. Orlowski, and J. Kirszenstein. A comparison between results from a physical scale model and a computer image source model for architectural acoustics. *Applied Acoustics*, 38(2–4):245–265, 1993.

[11] J. Suh and P. Nelson. Measurement of transient response of rooms and comparison with geometrical acoustic models. *Journal of the Acoustical Society of America*, 105(4):2304–2317, April 1999.

[12] K. Li and K. Iu. Propagation of sound in long enclosures. *Journal of the Acoustical Society of America*, 116(5):2759–2770, November 2004.

[13] K. Li and K. Iu. Full-scale measurements for noise transmission in tunnels. *Journal of the Acoustical Society of America*, 117(3):1138–1145, March 2005.

[14] K. Iu and K. Li. The propagation of sound in narrow street canyons. *Journal of the Acoustical Society of America*, 112(2):537–550, August 2002.

[15] S. Dance and B. Shield. The complete image-source method for the prediction of sound distribution in non-diffuse enclosed spaces. *Journal of Sound and Vibration*, 201(4):473–489, 1997.

[16] K. Li and P. Lam. Prediction of reverberation time and speech transmission index in long enclosures. *Journal of the Acoustical Society of America*, 117(6):3716–3726, June 2005.

[17] J. Kang. A method for predicting acoustic indices in long enclosures. *Applied Acoustics*, 51(2):169–180, June 1997.

[18] J. António, L. Godinho, and A. Tadeu. Reverberation times obtained using a numerical model versus those given by simplified formulas and measurements. *Acta Acustica united with Acustica*, 88(2):252–261, March/April 2002.

[19] J. Rindel and C. Lynge Christensen. Room acoustic simulation and auralization—How close can we get to the real room? In *Proceedings of the Eighth Western Pacific Acoustics Conference*, Melbourne, Australia, April 2003. Keynote lecture, 8 pages.

[20] H. Lehnert and J. Blauert. Virtual auditory environment. In *Proceedings of the IEEE International Conference on Advanced Robotics*, volume 1, pages 211–216, Pisa, Italy, June 1991.

[21] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36(3–4):171–218, 1992.

[22] L. Savioja, T. Lokki, and J. Huopaniemi. Auralization applying the parametric room acoustic modeling technique—the DIVA auralization system. In *Proceedings of the International Conference on Auditory Display*, pages 219–224, Kyoto, Japan, July 2002.

[23] T. Lokki, L. Savioja, R. Väänänen, J. Huopaniemi, and T. Takala. Creating interactive virtual auditory environments. *IEEE Computer Graphics and Applications*, 22(4):49–57, July 2002.

[24] N. Tsingos and J.-D. Gascuel. Soundtracks for computer animation: sound rendering in dynamic environments with occlusions. In *Proceedings of the Graphics Interface Conference*, pages 9–16, Kelowna, BC, Canada, May 1997.

[25] J. Miller and E. Wenzel. Recent developments in SLAB: a software-based system for interactive spatial sound synthesis. In *Proceedings of the Eighth International Conference on Auditory Display*, pages 403–408, Kyoto, Japan, July 2002.

[26] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal on Advances in Signal Processing*, 2007, 2007. Article ID 70540, 19 pages.

[27] E. Wenzel, J. Miller, and J. Abel. A software-based system for interactive spatial sound synthesis. In *Proceedings of the International Conference on Auditory Display*, pages 151–156, Atlanta, GA, USA, April 2000.

[28] J.-M. Jot. Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems*, 7(1):55–69, 1999.

[29] D. Zotkin, R. Duraiswami, and L. Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on Multimedia*, 6(4):553–564, August 2004.

[30] M. Ikram and D. Morgan. A multiresolution approach to blind separation of speech signals in a reverberant environment. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 2757–2760, Salt Lake City, UT, USA, May 2001.

[31] M. Ikram and D. Morgan. A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 881–884, Orlando, FL, USA, May 2002.

[32] J. Hopgood, P. Rayner, and P. Yuen. The effect of sensor placement in blind source separation. In *Proceedings of the IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pages 95–98, New Paltz, NY, USA, October 2001.

[33] B. Radlović, R. Williamson, and R. Kennedy. Equalization in an acoustic reverberant environment: robustness results. *IEEE Transactions on Speech and Audio Processing*, 8(3):311–319, May 2000.

[34] F. Talantzis and D. Ward. Robustness of multichannel equalization in an acoustic reverberant environment. *Journal of the Acoustical Society of America*, 114(2):833–841, August 2003.

[35] D. Ward. On the performance of acoustic crosstalk cancellation in a reverberant environment. *Journal of the Acoustical Society of America*, 110(2):1195–1198, August 2001.

[36] Y. Huang and J. Benesty. A class of frequency-domain adaptive approaches to blind multichannel identification. *IEEE Transactions on Signal Processing*, 51(1):11–24, January 2003.

[37] E. Lehmann and A. Johansson. Particle filter with integrated voice activity detection for acoustic source tracking. *EURASIP Journal on Advances in Signal Processing*, 2007, 2007. Article ID 50870, 11 pages.

[38] D. Ward, E. Lehmann, and R. Williamson. Particle filtering algorithms for tracking an acoustic source in a reverberant environment. *IEEE Transactions on Speech and Audio Processing*, 11(6):826–836, November 2003.

[39] J. Vermaak and A. Blake. Nonlinear filtering for speaker tracking in noisy and reverberant environments. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 3021–3024, Salt Lake City, UT, USA, may 2001.

[40] K. Palomäki, G. Brown, and J. Barker. Missing data speech recognition in reverberant conditions. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 65–68, Orlando, FL, USA, May 2002.

[41] K. Palomäki, G. Brown, and D. Wang. A binaural processor for missing data speech recognition in the presence of noise and small-room reverberation. *Speech Communication*, 43(4):361–378, 2004.

[42] S. Doclo and M. Moonen. GSVD-based optimal filtering for single and multimicrophone speech enhancement. *IEEE Transactions on Signal Processing*, 50(9):2230–2244, September 2002.

[43] P. Aarabi and G. Shi. Phase-based dual-microphone robust speech enhancement. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 34(4):1763–1773, August 2004.

[44] J. Chen, J. Benesty, and Y. Huang. Robust time delay estimation exploiting redundancy among multiple microphones. *IEEE Transactions on Speech and Audio Processing*, 11(6):549–557, November 2003.

[45] P. Aarabi. Self-localizing dynamic microphone arrays. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 32(4):474–484, November 2002.

[46] M. Brandstein and H. Silverman. A robust method for speech signal time-delay estimation in reverberant rooms. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 375–378, Munich, Germany, April 1997.

[47] S. Doclo and M. Moonen. Robust adaptive time delay estimation for speaker localization in noisy and reverberant acoustic environments. *EURASIP Journal on Applied Signal Processing*, 2003(11):1110–1124, 2003.

[48] Y. Zheng, R. Goubran, and M. El-Tanany. Robust near-field adaptive beamforming with distance discrimination. *IEEE Transactions on Speech and Audio Processing*, 12(5):478–488, September 2004.

[49] P. Peterson. Simulating the response of multiple microphones to a single acoustic source in a reverberant room. *Journal of the Acoustical Society of America*, 80(5):1527–1529, November 1986.

[50] H. Kuttruff. *Room acoustics*. Spon Press, London – New York, Fourth edition, August 2000.

[51] C. Nocke. In-situ acoustic impedance measurement using a free-field transfer function method. *Applied Acoustics*, 59(3):253–264, March 2000.

[52] M. Delany and E. Bazley. Acoustical properties of fibrous absorbent materials. *Applied Acoustics*, 3(2):105–116, April 1970.

[53] B. Brouard, D. Lafarge, J.-F. Allard, and M. Tamura. Measurement and prediction of the reflection coefficient of porous layers at oblique incidence and for inhomogeneous waves. *Journal of the Acoustical Society of America*, 99(1):100–107, January 1996.

[54] M. Schroeder. Integrated-impulse method measuring sound decay without using impulses. *Journal of the Acoustical Society of America*, 66(2):497–500, August 1979.

[55] B. Katz. International round robin on room acoustical impulse response analysis software 2004. *Acoustics Research Letters Online*, 5(4):158–164, October 2004.

[56] M. Schroeder. New method of measuring reverberation time. *Journal of the Acoustical Society of America*, 37(6):409–412, June 1965.

[57] M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, and V. Välimäki. Estimation of modal decay parameters from noisy response measurements. *Journal of the Audio Engineering Society*, 50(11):867–878, November 2002.

[58] D. Morgan. A parametric error analysis of the backward integration method for reverberation time estimation. *Journal of the Acoustical Society of America*, 101(5):2686–2693, May 1997.

[59] A. Krokstad, S. Strøm, and S. Sørsdal. Calculating the acoustical room response by use of a ray tracing technique. *Journal of Sound and Vibration*, 8(1):118–125, 1968.

[60] M. Abramowitz and I. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications, New York, 1965.

[61] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical recipes in C: the art of scientific computing*. Cambridge University Press, Cambridge, Second edition, 1992.

[62] E. Lehmann, A. Johansson, and S. Nordholm. Reverberation-time prediction method for room impulse responses simulated with the image-source model. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'07)*, pages 159–162, New Paltz, NY, USA, October 2007.

[63] S. Dance and B. Shield. Modelling of sound fields in enclosed spaces with absorbent room surfaces—Part II: absorptive panels. *Applied Acoustics*, 61(4):373–384, December 2000.

[64] R. Neubauer and B. Kostek. Prediction of the reverberation time in rectangular rooms with non-uniformly distributed sound absorption. *Archives of Acoustics*, 26(3):183–202, 2001.

[65] W. Sabine. *Collected papers on acoustics*. Dover Publications, New York, 1964.

[66] C. Eyring. Reverberation time in "dead" rooms. *Journal of the Acoustical Society of America*, 1(2A):217–241, January 1930.

[67] W. Sette. A new reverberation time formula. *Journal of the Acoustical Society of America*, 4(3):193–210, January 1933.

[68] G. Millington. A modified formula for reverberation. *Journal of the Acoustical Society of America*, 4(1A):69–82, July 1932.

[69] D. Fitzroy. Reverberation formula which seems to be more accurate with nonuniform distribution of absorption. *Journal of the Acoustical Society of America*, 31(7):893–897, July 1959.

[70] R. Brent. *Algorithms for minimization without derivatives*. Prentice-Hall, Englewood Cliffs, NJ, 1973.