Prediction of energy decay in room impulse responses simulated with an image-source model

Eric A. Lehmann[a] and Anders M. Johansson
Western Australian Telecommunications Research Institute, 35 Stirling Highway, Crawley WA 6009 Australia

(Received 21 January 2007; revised 25 February 2008; accepted 4 May 2008)

A method is proposed that provides an approximation of the acoustic energy decay (energy–time curve) in room impulse responses generated using the image-source technique. A geometrical analysis of the image-source principle leads to a closed-form expression describing the energy decay curve, with the resulting formula being valid for a uniform as well as nonuniform definition of the enclosure's six absorption coefficients. The accuracy of the proposed approximation is demonstrated on the basis of impulse-response simulations involving various room sizes and reverberation levels, with uniform and nonuniform sound absorption coefficients. An application example for the proposed method is illustrated by considering the task of predicting an enclosure's reflection coefficients in order to achieve a specific reverberation level. The technique presented in this work enables designers to undertake a preliminary analysis of a simulated reverberant environment without the need for time-consuming image-method simulations.
 [DOI: 10.1121/1.2936367]

PACS number(s): 43.55.Ka, 43.55.Br, 43.60.−c [WMC] Pages: 269–277

I. INTRODUCTION

The image-source model (ISM) has become an ubiquitous tool in many fields of acoustical and engineering research over the past few decades. Its success relies mainly on its conceptual simplicity, which makes ISM-based algorithms relatively straightforward to implement. As a result, the ISM approach has been used as a basis principle for a wide range of purposes including, for instance, prediction of sound propagation in enclosed environments[1,2] and long tunnels,[3] architectural modeling and design,[4,5] noise control in large spaces,[6] and analysis of perceptual properties such as speech intelligibility and speech transmission index in various enclosures.[7] In recent times, the image-source model has also been implemented for the purpose of binaural auralization,[8] spatialized sound rendering in virtual environments[9] and interactive systems,[10] and augmented-reality applications such as video games.[11]

Another important domain of application of the image-source technique is in order to assess the performance of various signal processing algorithms operating in reverberant environments. Some application examples include algorithms for blind source separation,[12] channel identification and equalization,[13,14] acoustic source localization and tracking,[15] speech recognition,[16] and speech enhancement.[17] In this context, the image model is generally used to determine the algorithm's robustness against increasing levels of environmental reverberation. Although not usually addressed in literature, a significant issue during this process is related to predicting the reverberation time in the resulting room impulse responses generated with the ISM. Well established formulas, such as Sabine or Eyring's reverberation time, do not provide accurate results when used to determine the enclosure's sound absorption in order to achieve a desired reverberation level.[18,19] This discrepancy between predicted and actual reverberation time, which is highlighted in this paper, is especially pronounced with a nonuniform definition of the sound reflection coefficients in the considered ISM room setup. As a result, a risk exists that the algorithm under test may be ultimately simulated with a reverberation level different from what is assumed during the experiment, leading to a potentially significant bias in the performance results. Finding an efficient solution to this specific issue was the original motivation behind the research presented in this work.

This paper describes a method 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 acoustic 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 expression which then allows a numerical prediction of the energy decay curve (EDC); this alleviates the need to practically simulate the RIR of interest, which represents a computational advantage, for instance, during the process of adjusting the RIR's reverberation time. Furthermore, this method efficiently deals with situations where current reverberation-time prediction techniques experience significant inaccuracies, such as in the case of nonuniform absorption coefficients. The ability to accurately predict the acoustic energy decay in a given enclosure hence provides a computationally efficient solution to the above mentioned problem of predicting the RIR's reverberation level, and also represents a general development tool providing some preliminary insight into the acoustical characteristics of the simulated environment.

[a]Tel.: +61 (0)8 6488 4642. FAX: +61 (0)8 6488 7254. Electronic mail: eric.lehmann@watri.org.au

In the following, the basic principles underlying the image-source method are first briefly reviewed in Sec. II, which also presents a modification of the original ISM algorithm leading to more practically accurate impulse responses. This modified ISM technique is used as a basis for the proposed EDC approximation method, the details of which are then presented in Sec. III. The accuracy and practical relevance of this approximation method are then demonstrated using a series of numerical evaluations in Sec. IV, which considers a typical application example and provides a computational efficiency assessment. Finally, Sec. V concludes with a discussion of the main concepts presented in this work.

II. IMAGE-SOURCE METHOD

This section briefly reviews the basic principles of the image-source technique and establishes the notation used throughout the rest of the paper.

A. Original approach

The conventional image-source method, as presented originally in a landmark paper by Allen and Berkley,[20] 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, respectively, in this setting:

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

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

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 assumed that the acoustical property of each surface in the enclosure is characterized by means of a sound reflection coefficient β, related to the absorption coefficient α according to

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

The reflection coefficients for each surface are denoted as $\beta_{x,i}$, $\beta_{y,i}$, and $\beta_{z,i}$, with $i=1,2$, where the subindex 1 refers to the wall closest to the origin. As commonly assumed, this work is based on geometrical room-acoustics principles and assumes that the reflection coefficients are frequency independent as well as angle independent.[21]

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 and attenuated by an amplitude factor A. The RIR $h(\cdot)$ hence follows as[20]

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

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 indexing of the image sources in all dimensions. For conciseness, the sum over $\boldsymbol{u}$ or $\boldsymbol{l}$ in Eq. (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 Eq. (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 is the sound propagation velocity and $d(\cdot)$ represents the distance from the image source to the receiver,

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

with $\|\cdot\|$ as the Euclidean norm and $\mathrm{diag}(\cdot)$ denoting a diagonal matrix with the arguments as diagonal elements.

Finally, note that the number of image sources to include in the summation of Eq. (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.

B. 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 better practical results. The resulting algorithm will be used as a basis for the simulations presented at the end of this paper.

1. Frequency-domain computations

The ISM implementation presented in Sec. II A typically needs to be updated in practice when dealing with discrete-time signals since the time delay $\tau(\cdot)$ does not usually correspond to an integer multiple of the sampling period. In the approach of Allen and Berkley,[20] this problem is dealt with by using nearest-integer rounding of each image source's propagation time, resulting in a shift of the corresponding impulse in the RIR. This approach thus leads to a coarse histogramlike representation of the desired RIR, which subsequently requires high-pass filtering in order to remove the nonphysical defect of this 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 delays that are not necessarily integer multiples of the sampling period. In the frequency domain, a time shift τ is 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 Eq. (5) as

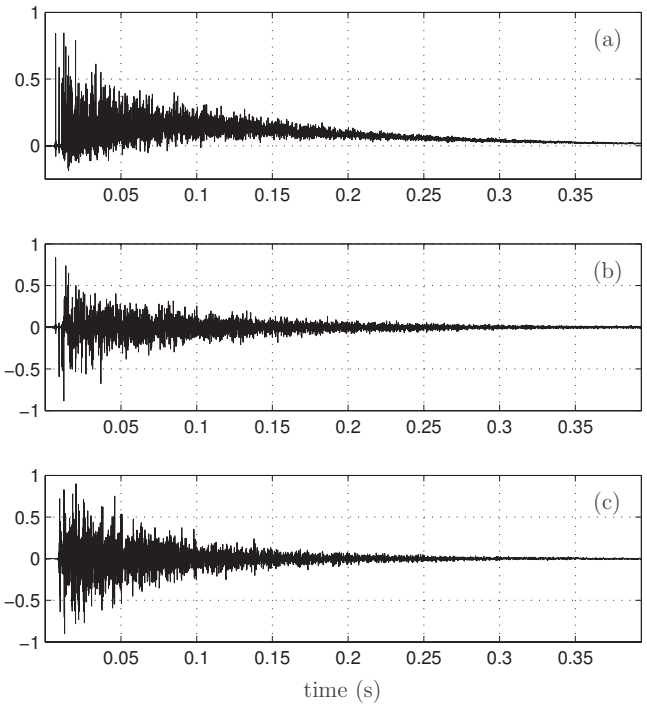

FIG. 1. (a) Typical RIR obtained using frequency-domain ISM (equivalent to Peterson's method) with positive reflection coefficients, $\boldsymbol{r}=[3.2\ \mathrm{m}, 4\ \mathrm{m}, 2.7\ \mathrm{m}]^{\mathrm{T}}$, $\boldsymbol{p}_r=[1.1\ \mathrm{m}, 1\ \mathrm{m}, 1.2\ \mathrm{m}]^{\mathrm{T}}$, $\boldsymbol{p}_s=[2\ \mathrm{m}, 3\ \mathrm{m}, 2\ \mathrm{m}]^{\mathrm{T}}$, sampling frequency $F_s=16$ kHz, and uniform reflection coefficients $\beta=0.92$. (b) Typical RIR obtained with the same environmental setup but using negative reflection coefficients. (c) Typical measurement of a real RIR recorded in a room with reverberation time $T_{60}\approx 0.6$ s.

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

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

$$h(t)=\mathcal{F}^{-1}\{H(\omega)\}. \tag{10}$$

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 noninteger propagation times. This approach, which was previously used by various authors,[13,18] essentially represents the frequency-domain equivalent to the low-pass impulse method (with infinite window duration) proposed by Peterson.[22]

2. Negative reflection coefficients

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

$$\beta=\pm\sqrt{1-\alpha}. \tag{11}$$

The original ISM implementation makes use of the positive definition of the β parameter. However, when used in conjunction with a frequency-domain implementation (or equivalently, Peterson's method), this approach generates anomalous RIRs showing a distinctively nonphysical tail decay, as depicted in Fig. 1(a).

An alternative approach is to use the negative definition of the parameter β in Eq. (11). This can be explained by considering the angle-dependent formula[23] for the reflection coefficient of a boundary with impedance ζ,

$$\beta=\frac{\zeta\cos(\psi)-1}{\zeta\cos(\psi)+1}, \tag{12}$$

which can become negative for a certain range of incidence angle ψ.[24] As shown in Fig. 1(b), and contrary to Fig. 1(a), this alternative approach results in RIRs whose shape compares well to that of practical RIR measurements, an example of which is displayed in Fig. 1(c). 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 Eqs. (9) and (10) with the following definition of each image source's amplitude factor:

$$A(\boldsymbol{u},\boldsymbol{l})=\frac{1}{4\pi d(\boldsymbol{u},\boldsymbol{l})}(-\beta_{x,1})^{|l-u|}(-\beta_{x,2})^{|l|}(-\beta_{y,1})^{|m-v|}\times(-\beta_{y,2})^{|m|}(-\beta_{z,1})^{|n-w|}(-\beta_{z,2})^{|n|}, \tag{13}$$

where $d(\cdot)$ is computed according to Eq. (8), and with the β parameters corresponding to the usual definition of sound reflection coefficients.[25] Finally, it must be noted that this "negative-coefficient approach" was previously studied and used by António *et al.*[18]

C. Energy decay

Given a RIR $h(t)$ computed for a specific environment according to Eq. (10), the energy decay envelope $E(t)$, known in literature as energy–time curve or energy decay curve (EDC), can be computed using a normalized version of the Schroeder integration method,[7,26]

$$E(t)=10\log_{10}\left[\frac{\int_t^\infty h^2(\xi)\mathrm{d}\xi}{\int_0^\infty h^2(\xi)\mathrm{d}\xi}\right], \tag{14}$$

where $E(\cdot)$ is expressed in dB. The result from Eq. (14) can then be used as a basis for deriving an estimate of the reverberation time, such as T_{20} or T_{60}.

III. PROPOSED ENERGY DECAY APPROXIMATION

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

A. Two-dimensional case

With the ISM technique, each image source 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

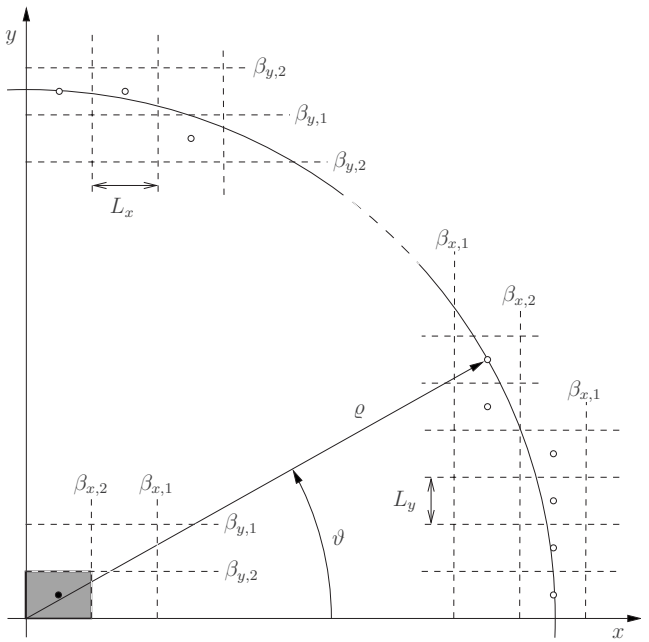

FIG. 2. Two-dimensional representation of the ISM principle. Dashed lines represent the grid of images of the original room, which is displayed as a shaded rectangle. The β parameters indicate the reflection coefficients of the corresponding boundaries, and circles (∘) represent the considered image sources.

traveled to the receiver. At the receiver, these sound particles are then added together at the corresponding time lags to create the impulse response.

Based on this principle, 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 $\varrho \triangleq \varrho(t) = ct$. This principle is demonstrated in Fig. 2 for one quadrant of the (x,y)-plane. Dashed rectangles represent images of the considered room, and circles denote the image sources of interest for the given radius ϱ. The original source has been placed at the center of the room for illustration purposes. The analysis presented in this section is based on the assumption of a large radius, that is,

$$\varrho \gg \max\{L_x, L_y, L_z\}. \tag{15}$$

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 Fig. 2. This assumption, together with the fact that some image sources do not lie perfectly on the considered circle, typically leads to approximation errors that become negligible as the radius ϱ of the considered circle increases.

Let us now consider an image source located at an angle ϑ along the considered circle, see Fig. 2. Prior to reaching the receiver, its sound impulse traverses a number $W_x \triangleq W_x(\varrho, \vartheta)$ of walls in the x direction, and $W_y \triangleq W_y(\varrho, \vartheta)$ walls in the y direction, which can be determined in a straightforward manner on the basis of the known position of the image source. Consequently, the power contribution $P(\cdot)$ made by this source to the transfer function can be expressed as

$$P(\varrho, \vartheta) = \frac{(\beta_{x,1}^2)^{W_x/2}(\beta_{x,2}^2)^{W_x/2}(\beta_{y,1}^2)^{W_y/2}(\beta_{y,2}^2)^{W_y/2}}{(4\pi\varrho)^2}. \tag{16}$$

Note that in contrast to Eq. (13), this expression effectively corresponds to a squared amplitude coefficient since the current developments are based on acoustic power rather than amplitude.[5] Equation (16) makes use of the assumption that along the path to the receiver, the number of walls with coefficient $\beta_{x,1}$ ($\beta_{y,1}$) is approximately equal to the number of walls with coefficient $\beta_{x,2}$ ($\beta_{y,2}$), that is, approximately equal to half the number of walls $W_x/2$ ($W_y/2$). Once again, this condition essentially becomes valid as the radius ϱ becomes large. In order to ultimately achieve a closed-form expression of the desired EDC approximation, the number of boundaries between the source and the receiver, in both dimensions, is determined according to a first-order approximation based on Fig. 2 (for $\vartheta \in [0, \pi/2]$),

$$W_x = \frac{\varrho}{L_x}\left(1 - \frac{2\vartheta}{\pi}\right), \tag{17}$$

$$W_y = \frac{\varrho}{L_y}\frac{2\vartheta}{\pi}. \tag{18}$$

Using the approach described so far, 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 here power-based, can be determined as

$$h_P(t) = \sum_{i \in \mathcal{I}_c} P(\varrho, \vartheta_i), \tag{19}$$

with $\varrho = ct$, the variable ϑ_i denoting the angle of the ith image source, and $\mathcal{I}_c$ representing the index set of the sources located on the considered circle. The basis of the proposed approximation is then to consider Eq. (19) as a Riemann sum, which can be represented as the integral of a continuous function over the angle ϑ,

$$h_P(t) \cdot \Delta\vartheta = \sum_{i \in \mathcal{I}_c} P(\varrho, \vartheta_i) \cdot \Delta\vartheta \tag{20}$$

$$\approx \int_0^{2\pi} P(\varrho, \vartheta)\mathrm{d}\vartheta. \tag{21}$$

As a result of the symmetry in the problem definition, the analysis can be restricted to a quarter of the circle, i.e., for $\vartheta \in [0, \pi/2]$. An estimate $\hat{h}_P(\cdot)$ of the power transfer function then follows from Eq. (21) as

$$h_P(t) \approx \hat{h}_P(t) = \frac{4}{\Delta\vartheta}\int_0^{\pi/2} P(\varrho, \vartheta)\mathrm{d}\vartheta. \tag{22}$$

The angular variable $\Delta\vartheta$ can be determined as

$$\Delta\vartheta = \frac{2\pi}{N_s}, \tag{23}$$

where N_s corresponds to the total number of image sources considered on the circle, or located very close to it. From Fig. 2, it can be seen that for $\vartheta \to 0$, the average distance between the image sources on the circle approaches L_y; simi-

larly, it approaches L_x for $\vartheta \to \pi/2$. Consequently, the parameter N_s is defined here as the circle's circumference divided by the average room dimension, i.e.,

$$N_s = \frac{2\pi\varrho}{\bar{r}}, \tag{24}$$

with $\bar{r}=(L_x+L_y)/2$. The value of $\Delta\vartheta$ then follows as $\Delta\vartheta = \bar{r}/\varrho$, which, together with Eq. (22), leads to the following approximation of the power RIR for the 2D setting:

$$\hat{h}_P(t) = \frac{4\varrho}{\bar{r}} \int_0^{\pi/2} P(\varrho, \vartheta) \mathrm{d}\vartheta. \tag{25}$$

B. Extension to three dimensions

An extension to the 3D case is obtained via a generalization of the above developments and the introduction of the polar angle $\varphi \in [0, \pi]$, as depicted in Fig. 3. Limiting here again the analysis to one eighth of the sphere, this results in Eq. (25) then becoming

$$\hat{h}_P(t) = \frac{8\varrho}{\bar{r}} \int_0^{\pi/2} \int_0^{\pi/2} P(\varrho, \vartheta, \varphi) \mathrm{d}\vartheta \mathrm{d}\varphi, \tag{26}$$

where $\bar{r}$ now includes the third room dimension,

$$\bar{r} = \frac{L_x + L_y + L_z}{3}, \tag{27}$$

and with the 3D extension of the power coefficient and number of walls,

$$P(\varrho, \vartheta, \varphi) = \frac{(\beta_{x,1}\beta_{x,2})^{W_x}(\beta_{y,1}\beta_{y,2})^{W_y}(\beta_{z,1}\beta_{z,2})^{W_z}}{(4\pi\varrho)^2}, \tag{28}$$

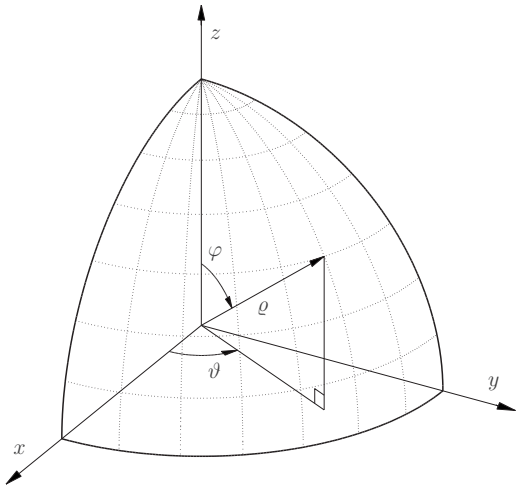

FIG. 3. Definition of the spherical coordinate system.

$$W_x = \frac{\varrho}{L_x}\left(1 - \frac{2\vartheta}{\pi}\right)\frac{2\varphi}{\pi}, \tag{29}$$

$$W_y = \frac{\varrho}{L_y}\frac{2\vartheta}{\pi}\frac{2\varphi}{\pi}, \tag{30}$$

$$W_z = \frac{\varrho}{L_z}\left(1 - \frac{2\varphi}{\pi}\right). \tag{31}$$

Inserting Eqs. (28)–(31) into Eq. (26) and analytically solving the double integral finally leads to the following estimate $\hat{h}_P(\cdot) \approx h_P(\cdot)$ of the power transfer function:

$$\hat{h}_P(t) = \frac{1}{8\varrho\bar{r}} \begin{cases} \dfrac{B_z}{\log\left(\frac{B_y}{B_x}\right)} \left\{ \mathrm{Ei}\left[\log\left(\frac{B_z}{B_x}\right)\right] + \log\left[\log\left(\frac{B_z}{B_x}\right)\right] - \mathrm{Ei}\left[\log\left(\frac{B_z}{B_y}\right)\right] - \log\left[\log\left(\frac{B_z}{B_y}\right)\right] \right\} & \text{if } B_x \neq B_y \neq B_z \\ \dfrac{B_z}{\log\left(\frac{B_z}{B}\right)} \left\{ \mathrm{Ei}\left[\log\left(\frac{B_z}{B}\right)\right] + \log\left[\log\left(\frac{B_z}{B}\right)\right] + \gamma \right\} & \text{if } B_z = B_y \neq B_x \triangleq B \text{ or } B_z = B_x \neq B_y \triangleq B \\ \dfrac{B - B_z}{\log\left(\frac{B}{B_z}\right)} & \text{if } B_z \neq B_x = B_y \triangleq B \\ B, & \text{if } B_x = B_y = B_z \triangleq B. \end{cases} \tag{32}$$

with $\gamma = 0.577\,215\,7\cdots$ the Euler–Mascheroni constant, $\mathrm{Ei}(\cdot)$ denoting the first-order exponential integral,[27] $\varrho \triangleq \varrho(t) = ct$ as defined earlier, and with the following definitions:

$$B_x = (\beta_{x,1}\beta_{x,2})^{\varrho/L_x}, \tag{33}$$

$$B_y = (\beta_{y,1}\beta_{y,2})^{\varrho/L_y}, \tag{34}$$

$$B_z = (\beta_{z,1}\beta_{z,2})^{\varrho/L_z}. \tag{35}$$

In Eq. (32), 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 logarithm function or the exponential integral.

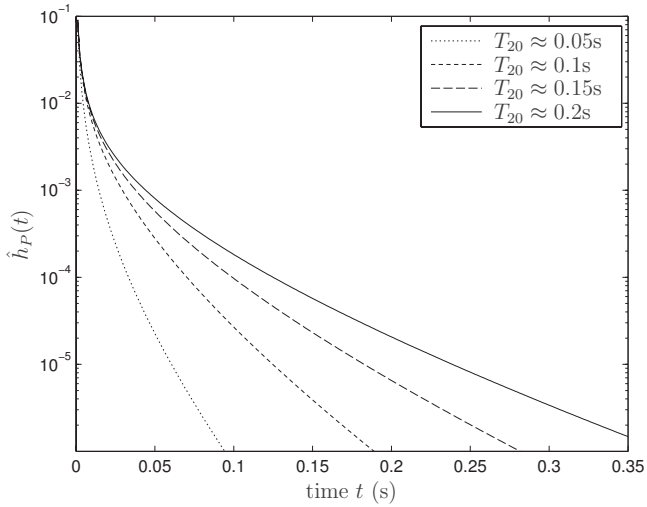

FIG. 4. Numerical evaluation of the approximate power transfer function $\hat{h}_P(\cdot)$, for various levels of reverberation, i.e., various values of β coefficients, and simulated with $\boldsymbol{r}=[4\text{ m}, 5\text{ m}, 2.9\text{ m}]^{\mathrm{T}}$.

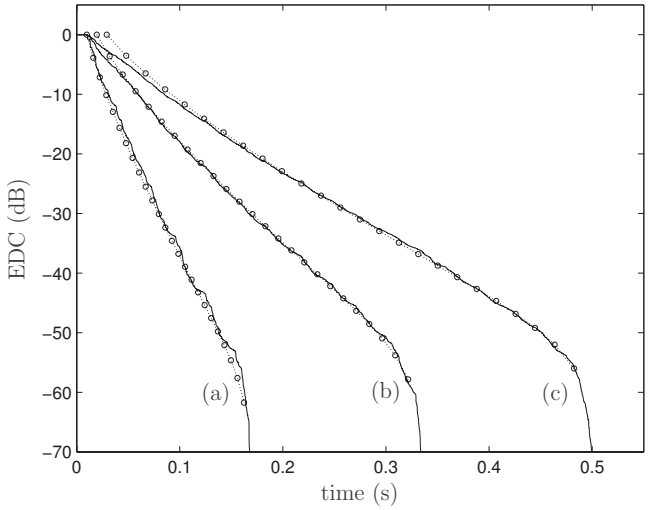

FIG. 5. Examples of energy decay curves with uniform reflection coefficients. (a) $\beta=0.669$ ($T_{20}\approx 0.05$ s), (b) $\beta=0.831$ ($T_{20}\approx 0.1$ s), and (c) $\beta=0.889$ ($T_{20}\approx 0.15$ s). Solid lines represent $E(t)$ obtained from ISM computations, circles (○) indicate values obtained with the proposed approximation $\hat{E}(t)$.

C. Approximation of energy decay

On the basis of Eq. (14), an estimate of the energy decay curve can be computed as

$$\hat{E}(t) = 10\log_{10}\left[\frac{\int_t^\infty \hat{h}_P(\xi)\mathrm{d}\xi}{\int_0^\infty \hat{h}_P(\xi)\mathrm{d}\xi}\right]. \tag{36}$$

In practice, the integrals in Eq. (36) can be replaced with Riemann sums as follows:

$$\int_t^\infty \hat{h}_P(\xi)\mathrm{d}\xi \approx T\cdot\sum_{i=0}^{\infty}\hat{h}_P(t+iT), \tag{37}$$

with an appropriate discretization step T. The validity of this approximation depends on the function $\hat{h}_P(\cdot)$ being smooth and bounded in the considered interval, which is supported by the plots in Fig. 4; it is also shown in the next section that Eq. (37) indeed holds true for the type of function defined in Eq. (32). Thus, the estimated energy–time curve can be finally computed according to Eqs. (36) and (37), and for $t > t_0$, as

$$\hat{E}(t) \approx 10\log_{10}\left[\frac{\sum_{i=0}^{\infty}\hat{h}_P(t+iT)}{\sum_{i=0}^{\infty}\hat{h}_P(t_0+iT)}\right], \quad t > t_0. \tag{38}$$

The introduction of the parameter t_0 in Eq. (38) 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 Eq. (38) can be considered as relevant only for values of t greater than a specific threshold, denoted here as t_0. Section IV will provide more detail regarding an appropriate setting of the t_0 parameter for numerical simulation purposes.

D. Discussion

Two distinct sources of error can be identified in relation to the expression proposed in Eq. (38). As mentioned above, the assumption of a large radius ϱ will typically lead to a poor approximation of the true EDC as $t\to 0$. In addition, the parameter t_0 effectively introduces an additive error term related to $\int_0^{t_0} h^2(\xi)\mathrm{d}\xi$ in the denominator of Eq. (38). This error term is, however, independent of the time variable t, and thus potentially creates a constant offset in the EDC approximation curve $\hat{E}(\cdot)$. These two distinct effects will be illustrated more specifically in the following section.

Finally, the infinite sums in Eq. (38) have to be truncated to a finite set of indices in practice. As shown in Fig. 4, the function $\hat{h}_P(t)$ tends toward 0 very quickly as t increases, and as a result, the summation can be terminated relatively early while still providing a good approximation for practical purposes.

IV. EXPERIMENTAL RESULTS

A. Numerical evaluations

This section provides some examples of the results obtained with the proposed EDC approximation method. Figure 5 considers a typical enclosure setup, the details of which are provided in Table I, 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 Eq. (14) on the basis of the impulse responses simulated with the ISM technique of Sec. II B. Circle markers (○) indicate the values obtained via Eqs. (32) and (38) computed at several discrete values of time. Figure 6 shows similar results, obtained using a different room setup

TABLE I. Parameter setup for the results presented in Figs. 5–7. F_s denotes the sampling frequency.

	Figure 5	Figures 6 and 7
$\boldsymbol{r}$ (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

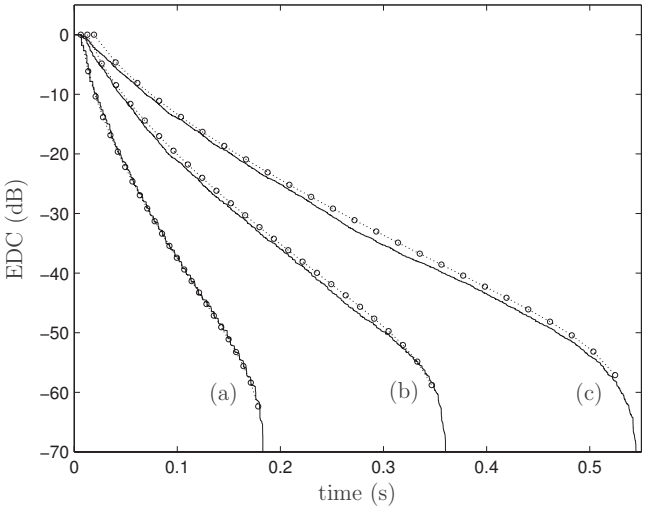

FIG. 6. Examples of energy decay curves with nonuniform reflection coefficients, for (a) $T_{20} \approx 0.05$ s, (b) $T_{20} \approx 0.1$ s, and (c) $T_{20} \approx 0.15$ s. See Table II for the corresponding β values. Solid lines represent $E(t)$ obtained from ISM computations, circles (∘) are values from the proposed approximation $\hat{E}(t)$.

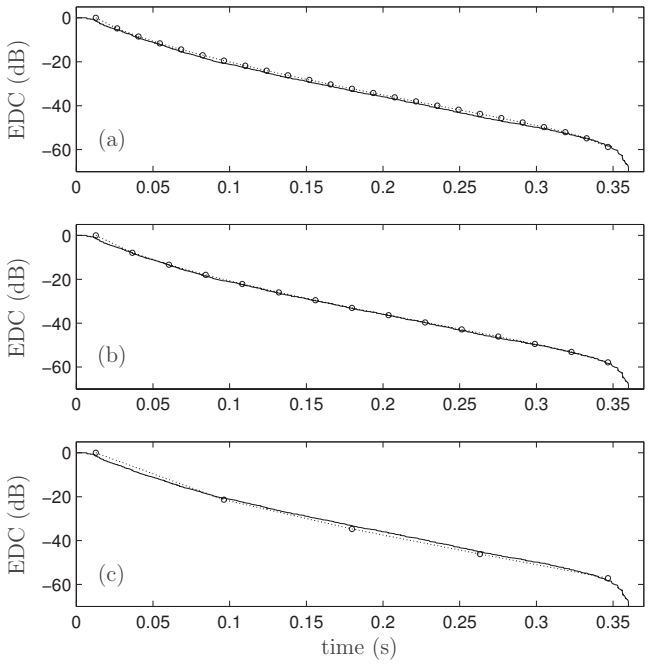

FIG. 7. EDC approximation results with varying interval lengths, for $T_{20} \approx 0.1$ s (see Table I for setup parameters). The discretization interval is defined as (a) $T=0.014$ s, (b) $T=0.024$ s, and (c) $T=0.083$ s.

(see Table I) in the case of nonuniform wall reflection coefficients, the values of which are given in Table II. Note that the curves in Fig. 6 correspond to a scenario where a pair of opposing walls is significantly different in reflectivity compared to other surfaces; this specific case was found to lead to discrepancies between the estimated and measured reverberation times in the publication of Allen and Berkley.[20]

Despite several simplifying assumptions made in this work, Figs. 5 and 6 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 lines for nonuniform β coefficients, match the practical results relatively well. With respect to the effects of the large-radius assumption mentioned previously, these numerical results also illustrate the discrepancy between the approximated and the practical results at low t values, 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 most practical purposes.

It must be noted here that the results in Figs. 5 and 6 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 discussed in Sec. III D, and thus enables a better visual comparison of the displayed results. In practice, a nonoptimal 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 array 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. Reference 19 further demonstrates the practical accuracy of the proposed method by providing more results from extensive numerical simulations.

Finally, the different plots in Fig. 7 provide an insight into the influence of the discretization interval T in Eq. (38). This figure displays the approximation results for three different time lengths, computed with $T_{20} \approx 0.1$ s and for an environmental setup as described in Table I. 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 Eq. (37). If necessary, the calculations can hence be made more computationally efficient by reducing the number of points on the approximation curve, with only a marginal reduction of the representation accuracy.

In addition to the practical results presented in this section, the accuracy of the proposed technique has been further tested and confirmed for a wide range of scenarios involving different enclosure volumes, various source and sensor positions, as well as uniform and nonuniform reflection coefficients.

TABLE II. Values of reflection coefficients for each boundary surface, for the curves displayed in Fig. 6.

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 III. Desired vs measured reverberation times with nonuniform sound absorption coefficients. The relative error percentage is displayed in brackets.

	Measured T_{20} (s)		
Desired T_{20} (s)	Proposed	Sabine	Eyring
0.05	0.045 (10%)	0.048 (4%)	0.074 (48%)
0.10	0.092 (8%)	0.131 (31%)	0.161 (61%)
0.15	0.144 (4%)	0.212 (41%)	0.238 (59%)
0.20	0.201 (1%)	0.284 (42%)	0.311 (56%)
0.25	0.255 (2%)	0.352 (41%)	0.371 (48%)
0.30	0.318 (6%)	0.416 (39%)	0.438 (46%)

TABLE IV. Average CPU times required for EDC computations at various reverberation times T_{60}, for ISM-based simulations and the proposed EDC approximation method.

	CPU time (s)	
T_{60} (s)	Proposed	ISM-based
0.1	0.0082	2.72
0.2	0.015	7.75
0.3	0.022	32.96
0.4	0.029	101.18
0.5	0.036	236.17
0.6	0.042	466.92

B. Application example

As discussed in Sec. I, the proposed EDC approximation method can be used to efficiently determine which values of reflection coefficients must be chosen to achieve a desired reverberation level T_{20} in a given environment. This is illustrated here with the typical example of a $4\times5\times2.9$ m^3 room with nonuniform wall absorption coefficients defined by the following ratios:

$$(\alpha_{x,1}:\alpha_{x,2}:\alpha_{y,1}:\alpha_{y,2}:\alpha_{z,1}:\alpha_{z,2}) = (1.0:0.9:0.7:0.6:0.4:0.3), \tag{39}$$

where the notation $(\mu_1:\mu_2:\mu_3:\cdots)$ is used to describe a series of ratios between multiple variables, i.e., $\mu_1:\mu_2,\mu_2:\mu_3$, etc. Using the method proposed in Sec. III, the approximate EDC can be computed with varying values of absorption coefficients [while ensuring that the ratios in Eq. (39) remain satisfied], until α parameters are found that yield the correct reverberation time, i.e., the correct slope in the predicted energy–time curve. Note that this optimization process is very fast since each iteration consists only of a few numerical evaluations of the expressions in Eqs. (32) and (38). The frequency-domain ISM computation is then carried out according to Sec. II B with the resulting α values, and the "true" reverberation time is measured directly from the computed RIR. For comparison, the same process is carried out with two commonly used formulas for predicting the reverberation time, namely, Sabine's equation,[28]

$$T_{20,\mathrm{Sab}} = \frac{1}{3}\frac{0.161V}{\sum_{i=1}^{6} S_i\alpha_i}, \tag{40}$$

and Eyring's equation,[29]

$$T_{20,\mathrm{Eyr}} = \frac{1}{3}\frac{0.161V}{-S\log(1-\sum_{i=1}^{6} S_i\alpha_i/S)}, \tag{41}$$

where V is the room volume, S is the total surface area of the enclosure, and S_i and α_i, $i=1,\ldots,6$, represent the surface areas and absorption coefficients of each individual wall, respectively. In the simulations, Eqs. (40) and (41) are used in conjunction with Peterson's ISM implementation[22] with positive reflection coefficients, as this represents the approach chosen by many authors in order to assess the performance of various signal processing algorithms.

Table III presents the assessment results, which correspond to an average over 20 randomly selected source–receiver configurations in the considered environment. These results are shown for various desired values of the reverberation time T_{20}, which is defined here as the time required by the acoustic energy $E(t)$ in the RIR to decrease from -5 to -25 dB. Note that any other quantity could have been chosen as a measure of reverberation: the developments presented in this paper are valid for any parameter of interest defined on the basis of the EDC, such as T_{60}, T_{30}, early decay time, etc.

As mentioned in Sec. I, Table III highlights the significant discrepancies existing between the desired reverberation level and that predicted with Sabine or Eyring's formulas; these expressions severely underestimate the value of the α parameters, leading to reverberation times that are significantly higher than targeted. On the other hand, because the proposed method is based on a relatively accurate prediction of the resulting acoustic energy decay, its reverberation time estimates remain within a small percentage of the desired T_{20} values, even when the sound reflection coefficients are not defined uniformly in the environment. In the specific context of testing signal processing algorithms using ISM simulations, this consequently minimizes the discrepancy between the assumed and actual reverberation times and ensures a minimal bias in the resulting performance assessment.

C. Computational requirements

As mentioned earlier, the computation of energy–time curves by means of ISM-based simulations can be computationally very demanding. The EDC approximation method proposed in this work alleviates this heavy computational burden by providing a closed-form expression allowing the quasi-instantaneous computation of a room's EDC curve.

To highlight the computational advantages of the proposed approach, Table IV presents the average CPU (central processing unit) time required by each method for the computation of an energy decay curve, as a function of the reverberation time T_{60}. These results represent the average over 80 EDC simulations carried out for various enclosures with randomly generated dimensions and source–sensor configurations. For the proposed approach, all EDCs were approximated with a constant discretization step $T=0.005$ s. Both methods were implemented in the MATLAB programing language with a sampling frequency $F_s=8$ kHz, and the computations were carried out on a modern 2.0 GHz computer. For both methods, the computation of each separate

EDC was terminated once the energy level in the RIR had decreased by approximately 60 dB.

As demonstrated by the results in Table IV, the proposed technique considerably reduces the computational requirements compared to ISM simulations. This computational advantage is even more significant with the multiple computation of one RIR, for instance, in the process of adjusting sound reflection coefficients to achieve a desired reverberation time (as in the example of Sec. IV B). Note also that the computational requirements of the ISM-based simulations further increase dramatically as the sampling frequency F_s is increased.

V. CONCLUSIONS

This paper proposes a method for approximation of the acoustic energy decay in simulated room impulse responses, and demonstrates that this technique provides an accurate prediction of the energy–time function generated on the basis of a modified version of the widely used image-source model. This method thus enables designers to efficiently investigate some of the acoustical characteristics of a simulated room without the need to generate the impulse responses of interest. Due to the considerable computational demands usually associated with the image-source model, this consequently represents a substantial reduction in the resulting computational burden. Furthermore, the developments presented here explicitly establish an unequivocal connection between environmental factors such as the walls' absorption coefficients and the level of reverberation resulting in the considered enclosure. As shown in this work, this relation is not currently well modeled by classical reverberation-time formulas, especially with a nonuniform definition of the sound absorption coefficients. In order to test audio processing algorithms, the proposed method hence provides engineers with a valuable tool to generate realistic impulse responses, whose main parameter of interest, namely the reverberation level, can be effortlessly and accurately tuned. The technique described in this paper can also be of potential interest in other application fields such as virtual auditory environments, perceptual acoustics, architectural design, 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.

[1]S. Dance, J. Roberts, and B. Shield, "Computer prediction of sound distribution in enclosed spaces using an interference pressure model," Appl. Acoust. **44**, 53–65 (1995).
[2]M. Hodgson, "On the accuracy of models for predicting sound propagation in fitted rooms," J. Acoust. Soc. Am. **88**, 871–878 (1990).
[3]K. Li and K. Iu, "Full-scale measurements for noise transmission in tunnels," J. Acoust. Soc. Am. **117**, 1138–1145 (2005).
[4]Y. Lam, "Issues for computer modelling of room acoustics in non-concert hall settings," Acoust. Sci. & Tech. **26**, 145–155 (2005).
[5]J. Suh and P. Nelson, "Measurement of transient response of rooms and comparison with geometrical acoustic models," J. Acoust. Soc. Am. **105**, 2304–2317 (1999).
[6]S. Dance and B. Shield, "The complete image-source method for the prediction of sound distribution in non-diffuse enclosed spaces," J. Sound Vib. **201**, 473–489 (1997).
[7]K. Li and P. Lam, "Prediction of reverberation time and speech transmission index in long enclosures," J. Acoust. Soc. Am. **117**, 3716–3726 (2005).
[8]L. Savioja, T. Lokki, and J. Huopaniemi, "Auralization applying the parametric room acoustic modeling technique—the DIVA auralization system," Proceedings of the International Conference on Auditory Display, Kyoto, Japan, 2002, pp. 219–224.
[9]J. Miller and E. Wenzel, "Recent developments in SLAB: a software-based system for interactive spatial sound synthesis," Proceedings of the International Conference on Auditory Display, Kyoto, Japan, 2002, pp. 403–408.
[10]T. Lokki, L. Savioja, R. Väänänen, J. Huopaniemi, and T. Takala, "Creating interactive virtual auditory environments," IEEE Comput. Graphics Appl. **22**, 49–57 (2002).
[11]D. Zotkin, R. Duraiswami, and L. Davis, "Rendering localized spatial audio in a virtual auditory space," IEEE Trans. Multimedia **6**, 553–564 (2004).
[12]M. Ikram and D. Morgan, "A multiresolution approach to blind separation of speech signals in a reverberant environment," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, 2001, pp. 2757–2760.
[13]B. Radlović, R. Williamson, and R. Kennedy, "Equalization in an acoustic reverberant environment: robustness results," IEEE Trans. Speech Audio Process. **8**, 311–319 (2000).
[14]F. Talantzis and D. Ward, "Robustness of multichannel equalization in an acoustic reverberant environment," J. Acoust. Soc. Am. **114**, 833–841 (2003).
[15]E. Lehmann and A. Johansson, "Particle filter with integrated voice activity detection for acoustic source tracking," EURASIP Journal on Advances in Signal Processing **2007**, Article ID 50870, 11 pages (2007).
[16]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 Commun. **43**, 361–378 (2004).
[17]S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," IEEE Trans. Signal Process. **50**, 2230–2244 (2002).
[18]J. António, L. Godinho, and A. Tadeu, "Reverberation times obtained using a numerical model versus those given by simplified formulas and measurements," Acust. Acta Acust. **88**, 252–261 (2002).
[19]E. Lehmann, A. Johansson, and S. Nordholm, "Reverberation-time prediction method for room impulse responses simulated with the image-source model," Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 2007, pp. 159–162.
[20]J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**, 943–950 (1979).
[21]This assumption is mainly for conciseness reasons. 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 this work.
[22]P. Peterson, "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," J. Acoust. Soc. Am. **80**, 1527–1529 (1986).
[23]H. Kuttruff, *Room Acoustics*, 4th ed. (Spon, London, 2000).
[24]Please note that Eq. (12) is only provided here in order to provide some insight into the physical meaning of using negative β coefficients. As mentioned in Sec. II A, the derivations in this work assume angle-independent reflection coefficients.
[25]Note further that the use of negative β coefficients in Eq. (13) when computing simulated RIRs does not impact on the subsequent developments since this work focuses on acoustic energy which, in turn, relies on squared amplitude coefficients (i.e., acoustic power).
[26]M. Schroeder, "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**, 409–412 (1965).
[27]The exponential integral Ei(·) is a standard function built into several mathematical packages such as MATLAB, MAPLE, and MATHEMATICA.
[28]W. Sabine, *Collected Papers on Acoustics* (Dover, New York, 1964).
[29]C. Eyring, "Reverberation time in "dead" rooms," J. Acoust. Soc. Am. **1**, 217–241 (1930).