Hybrid prediction tool implemented for acoustic design studies of open plan office spaces

Citation for the original published paper (version of record):
Forssén, J., Müller, L., Hedlund, E. et al (2026). Hybrid prediction tool implemented for acoustic design studies of open plan office spaces. Applied Acoustics, 248.
http://dx.doi.org/10.1016/j.apacoust.2026.111284

Applied Acoustics 248 (2026) 111284

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Hybrid prediction tool implemented for acoustic design studies of open plan office spaces

Jens Forssén*, Leon Müller, Elin Hedlund, Wolfgang Kropp

Division of Applied Acoustics, Department of Architecture and Civil Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

HIGHLIGHTS

- A hybrid room acoustics prediction model was developed and implemented for open-plan office spaces.
- The prediction model combines phased image sources and acoustic radiosity.
- Shielding effects of acoustic screens are modelled, enabling frequency-dependent diffraction calculations.
- Impulse responses and frequency-domain insertion losses show reasonable agreement between the model and measurements.
- Listening test results indicate the perceptual plausibility of the model.

ARTICLE INFO

Keywords:
Room acoustics
Open office
Phased image sources
Acoustic radiosity
Hybrid model

ABSTRACT

A hybrid room acoustics prediction tool using image sources and acoustic radiosity has been implemented. The theory and implementation are described including the specular reflection modelling in frequency domain using phased image sources, the diffuse reflection modelling using a downsampled time-domain acoustic radiosity approach, and how the models are combined with a successive transfer of energy from the image source model to the radiosity model. In particular, modelling of the shielding effect of screens is described in detail. The predicted results are compared with measured data, with and without shielding screens, showing reasonable agreement in the time-domain response as well as in the frequency-domain insertion loss of one screen. The results of a listening experiment show reasonably good agreement in perceived similarity between measurements and model predictions, both with and without a shielding screen, indicating the usefulness of the model.

1. Introduction

A poor sound environment in open-plan offices has been shown to be detrimental to the work environment in terms of annoyance, work performance, and well-being. Studies show that about a third of people working in open-plan offices are annoyed by noise [1]. The main type of disturbing sound in workplaces is speech. Irrelevant speech reduces work performance [2] and causes stress [3]. Research has shown that not only irrelevant speech, but also irrelevant sound effects might have a cognitive impact due to certain properties related to changes in pitch and amplitude [4]. The question arises as to how far the acoustic design of open-plan office spaces can be utilised to minimise the negative impact of irrelevant speech or other sounds. Tools such as layout, absorption, screens, and background noise could be seen as possibilities to improve the work environment. For example, there are studies discussing the influence of reverberation time on cognitive performance in open-plan offices (e.g., Yadav et al. [5]) where the influence of different degrees of absorption on the resulting work performance tasks is investigated. Of importance in this discussion is also the combination of room reverberation and the number of irrelevant speech voices [6–8]. A part of the discussion also concerns the validity of such studies as it requires that the experiments are carried out under coherent conditions. This is not always the case, e.g., when using recordings of speech in rooms that are different from the room of interest for the study. A possible solution would be to carry out such experiments in virtual environments where room acoustic design can easily be changed under controlled conditions. To create such virtual environments, room acoustics models are needed

* Corresponding author.
Email address: jens.forssen@chalmers.se (J. Forssén).
URL: https://www.chalmers.se/en/persons/jf/ (J. Forssén).

https://doi.org/10.1016/j.apacoust.2026.111284
Received 31 March 2025; Received in revised form 17 February 2026; Accepted 20 February 2026
Available online 24 February 2026

which are perceptually plausible and which allow variation of acoustic properties such as reverberation time and the inclusion of screens.

The current paper describes the implementation and validation of a hybrid room acoustics prediction tool using phased image sources and acoustic radiosity. The purpose of the tool is to enable investigation of the effects of various acoustic design solutions, both in terms of acoustic metrics and for auralization for perceptual judgement. The developed tool aims to balance the numerical cost with the flexibility to consider various design solutions of interest, including the spatial distribution of absorbing materials and their spectral behaviour, surface roughness, and screens. At present, simplifying assumptions include the limitation to shoebox shaped rooms.

Ray tracing methods have been used extensively to model impulse responses of rooms [9–11]. In classical ray-tracing, energy sound rays are tracked from a source location to a receiver sphere via a number of orders of reflections on the surfaces of the room. At each reflection, a portion of the ray's energy may be absorbed or converted from specular to diffuse reflection, where the energy from the diffuse reflections is traced further but spread out over emission angles (deterministically or randomised). To model frequency-dependent absorption and diffusion properties, the calculations are rerun in several frequency bands, e.g., in 1/1 or 1/3-octave bands.

For rooms with many surface elements, calculations become costly for long impulse responses, whereby the late response needs to be treated by other models. For this, various approaches and combinations of approaches have been developed and tested (e.g., [12,13]), and in practical cases, the image source method is often limited to a reflection order of two or three (e.g., Lynge Christensen et al. [14]), whereafter the late response modelling takes over.

One of the drawbacks of classical ray-tracing is that the delay that occurs at a reflection is disregarded. It has been shown that this may have an audible influence [15] and, to capture this effect, the boundary materials of the room can be modelled by their surface impedance, which determines the phase shift of a reflected wave as a function of frequency. When working with image sources in frequency domain (*phased* image sources), the effect of surface impedance can be considered (e.g., Marbjerg et al. [15]) and the frequency dependence can be captured in a single run, whereafter the impulse response in time domain is given by using an inverse Fourier transform. The implemented method supports the use of complex-valued surface impedance. Since the method presented here is limited to shoebox-shaped rooms, it allows for fast image source calculations up to relatively high orders (e.g., order 16 or 32). Thereby, the modal response of the room can also be resolved with a higher accuracy, predicting the frequencies of the resonance peaks and the pressure amplitude distribution in the room at lower frequencies, i.e., below the Schroeder frequency.

The late part of the impulse response can be modelled, in an approximate approach, as an exponentially decaying energy tail using the early part result as input [16]. However, the acoustic radiosity model (e.g., [9,17–20]) has received renewed interest in recent years and is deemed useful for modelling the late part (e.g., [14,15,21]). In the hybrid approach, the fraction of energy that is converted to diffuse field at each reflection is removed from the image source calculation and inserted into the radiosity calculation (e.g., Koutsouris et al. [21]). In the acoustic radiosity model, the room boundaries are discretised into surface elements that successively emit and receive energy from and to the other surface elements. Up to this point, both the phased image source model and the radiosity model are independent of receiver location, which can be utilised for efficient application to situations with multiple receiver locations, e.g., for a moving receiver. A drawback of the acoustic radiosity model is that any frequency-dependent absorption needs to be modelled using reruns in frequency bands. However, this can be sped up by using parallel computing. A benefit of the acoustic radiosity model is that the energy response function may be sampled at a much lower rate (e.g., by a factor of 20) than what the Nyquist criterion stipulates for the corresponding sound pressure signal, as further described below.

Section 2 describes the methods used here. First, the phased image source model, the acoustic radiosity model and the hybrid approach are presented. Thereafter selected implementation details are clarified, followed by a description of the screen modelling, further detailed in Appendix A. In Section 3, numerical results are presented in comparison with measured data, and in Section 4 the listening experiment and the results therefrom are described. Section 5 and Section 6 contain discussion and conclusion.

2. Method

2.1. Theoretical description of the hybrid image source and acoustic radiosity prediction model

The hybrid model combines an image source model and an acoustic radiosity model to calculate impulse responses for a sound source in a room. The approach mainly follows that described by Koutsouris et al. [21]. Further details of the implementation used here are highlighted in Section 2.1.4.

2.1.1. Image source method

In the image source model, the pressure at a receiver position is the sum of contributions from the source and a finite set of image sources, where the latter represents the effect of reflections at the room boundaries.

For a shoebox-shaped room with sound-hard (or sound-soft) boundaries, an infinite set of image sources would lead to an exact solution, where each image source is located at the reflection of the point of the original source across each boundary plane and its reflections. However, for (most) other room shapes or boundary conditions, the image source model would not approach the exact solution but may still provide relevant and useful solutions.

For a time-domain approach, the sound pressure, p, in free space, due to a point source formulated as $-4\pi\delta(x-x_S, y-y_S, z-z_S, t-t_0)$, can be written as

$$p(x_R, y_R, z_R, t) = \frac{\delta(t - t_0 - R/c)}{R} \tag{1}$$

where δ is the delta function, x_S, y_S, z_S and x_R, y_R, z_R are the coordinates of the source and the receiver, R is the distance between them, t is time, with t_0 the timing of the source impulse, and c is the sound speed.

The above formulation of the sound pressure can be identified as a Green's function for which the frequency domain equivalent can be written

$$P(x_r, y_r, z_r) = \frac{\mathrm{e}^{-\mathrm{j}kR}}{R} \tag{2}$$

see e.g., [22, Ch.7]. (Note that this reference uses a time factor $\exp(-\mathrm{i}\omega t)$ instead of $\exp(\mathrm{j}\omega t)$ as is used here.)

In the phased image source model implemented here, the summation of pressure contributions from the original source and all image sources is performed in frequency domain, which may include any complex-valued reflection factor at each reflection, whereafter an inverse Fourier transform (IFFT) is applied to obtain the time-domain pressure signal. This approach enables efficient modelling of boundary surfaces characterised by their complex-valued impedance and the plane-wave reflection factor. The number of sources used (original source plus image sources), N_s, is here determined by the highest modelled reflection order, N_{ord}, such that $N_s = (2N_{\mathrm{ord}}+1)^3$, where the zeroth reflection order, $N_{\mathrm{ord}} = 0$, corresponds to the total absence of boundary reflections.

2.1.2. Radiosity model

The acoustic radiosity model is energy based and as such cannot model any interference effects. That is, phase relations between propagating waves are assumed incoherent [9,20]. The acoustic radiosity model as used here is therefore assumed to be constrained to modelling the part of the sound field that has been converted from the original perfectly coherent wave from the source into incoherent contributions, i.e.,

the increasingly diffuse part of the boundary reflections at increasing reflection order.

The theoretical framework can be initiated by recapitulating the relations between the time-dependent acoustical quantities of pressure, $p(t)$, particle velocity, $v(t)$, intensity, $I(t)$, and energy density, $E(t)$. It should be noted that the particle velocity is generally a vector (i.e., has a direction in addition to its amplitude), $\mathbf{v}(t)$, as well as the intensity, $\mathbf{I}(t)$, such that $\mathbf{I}(t) = p(t)\mathbf{v}(t)$. However, we will here use the scalar quantities (absolute values) of the particle velocity and intensity ($v(t) = |\mathbf{v}(t)|$ and $I(t) = |\mathbf{I}(t)|$) and infer the effects of the vectorial properties explicitly. The relations of interest can be written as $I(t) = p(t)v(t)$ and $E(t) = I(t)/c$, where c is the sound speed of the medium. Assuming a far-field condition (i.e., pressure and particle velocity are in phase), we have $p(t)/v(t) = \rho_0 c$, where ρ_0 is the density of the medium. By using the above far-field expression, one can find $I(t) = \frac{p^2(t)}{\rho_0 c}$ and $E(t) = \frac{p^2(t)}{\rho_0 c^2}$.

We will also need to introduce the power, W, by which we excite the system. From energy conservation we can write a (steady state) balance between emitted power, W, from an omnidirectional source and the sum of intensity, I, passing through a fictitious spherical surface in the normal direction (assuming no losses due to propagation through the medium, i.e., zero air attenuation), $W = I \cdot 4\pi R^2$, where R is the radius of the sphere with the source at its centre. Using the above equation and the fact that $E = I/c$, we can find an expression for the direct energy density, E_0, due to an omnidirectional source with power W_0 at distance R: $E_0 = W_0/(4\pi c R^2)$.

The radiosity model is usually formulated using a quantity of *radiation density* (or sometimes *irradiation density*), here denoted B. The direct contribution from the source to the radiation density at all surface elements, sometimes called *first shot*, can be written as

$$B_{i,0} = \frac{\rho_i}{4\pi} \frac{\cos\theta_{i,0}}{R_{i,0}^2} W_0 \tag{3}$$

where i denotes the receiving surface element and 0 denotes the source, ρ_i is the reflection coefficient of element i, and $\theta_{i,0}$ is the angle between the normal of the surface element and the line, with length $R_{i,0}$, from the source to the centre of the element.

The radiation density, B_i, of a surface element i can therefore be seen as the projection of the incoming intensity onto the normal direction of the surface element (in the above equation, the effect of the projection can be identified as the factor $\cos\theta_{i,0}$), multiplied by the reflection coefficient, ρ_i.

The above expression for the radiation density is approximate, assuming small variation over a surface element. A more exact formulation can be written, assuming variation of $\theta_{i,0}$ and $R_{i,0}$ over the element, as

$$B_{i,0} = \frac{\rho_i}{4\pi S_i} W_0 H_{i,0} \tag{4}$$

where

$$H_{i,0} = \int_{S_i} \frac{\cos\theta_{i,0}}{R_{i,0}^2} \mathrm{d}S. \tag{5}$$

In the remaining shots, each surface element, i, may have contributions from all other surface elements, $j = 1 \ldots N$, where N is the number of elements. The corresponding equation may be written as (e.g., Koutsouris et al. [21])

$$B_i = \rho_i \sum_{j=1}^{N} B_j F_{i,j} + B_{i,0} \tag{6}$$

where

$$F_{i,j} = \frac{V_{i,j}}{S_i} \int_{S_i} \int_{S_j} \frac{\cos\theta_{S_i} \cos\theta_{S_j}}{\pi R_{i,j}^2} \mathrm{d}S_j \mathrm{d}S_i. \tag{7}$$

Here, a visibility factor, $V_{i,j}$, is also included, usually with a value of 0 or 1, which may be used to model the effect of shielded paths between elements, e.g., for non-convex domains. It should be noted that inherent in the above formulation is the assumption that the radiation from each element follows Lambert's law, that is, that the radiated intensity varies with the angle θ from its normal as $\cos\theta$, independent of the angle of the incident intensity. (This can be identified by the factor $\cos\theta_{S_j}$ in the above equation.)

If the equation system of Eq. (6) is solved to determine the values of B_i for all surface elements, the energy density at a receiver (denoted with the subscript r) can be calculated as

$$E_r = \frac{V_{r,j}}{\pi c} \sum_{j=1}^{N} B_j H_{r,j} + E_0 \tag{8}$$

where

$$H_{r,j} = \int_{S_j} \frac{\cos\theta_r}{R_{r,j}^2} \mathrm{d}S. \tag{9}$$

Including time dependence and air attenuation (with coefficients in the form of $\exp(-\alpha_m R)$, where α_m is the air absorption exponent), the acoustic radiosity equations can be written as follows (with integrals F and H and other variables as above):

$$B_i(t) = \rho_i \sum_{j=1}^{N} B_j \left(t - \frac{R_{i,j}}{c}\right) F_{i,j} \mathrm{e}^{-\alpha_m R_{i,j}} + B_{i,0}(t) \tag{10}$$

$$B_{i,0}(t) = \frac{\rho_i}{4\pi S_i} W_0 \left(t - \frac{R_{i,0}}{c}\right) H_{i,0} \mathrm{e}^{-\alpha_m R_{i,0}} \tag{11}$$

$$E_r(t) = \frac{V_{r,j}}{\pi c} \sum_{j=1}^{N} B_j \left(t - \frac{R_{r,j}}{c}\right) H_{r,j} \mathrm{e}^{-\alpha_m R_{r,j}} + E_0(t) \tag{12}$$

$$E_0(t) = \frac{W_0\left(t - \frac{R_0}{c}\right)}{4\pi c R_0^2} \mathrm{e}^{-\alpha_m R_0}. \tag{13}$$

2.1.3. Hybrid model

In the hybrid model, a scattering coefficient, σ, determines the amount of energy transferred from the image source model to the acoustic radiosity model in each reflection order. For the room acoustical problem of interest, an initial pulse is emitted by the source and modelled by the image sources. At each reflection on the boundary, a fraction of the energy is removed due to both absorption, α_i, and scattering, σ_i, which may have different values for different surface elements, i. The remaining reflected fraction in the image source model is therefore $\rho_i = (1-\alpha_i)(1-\sigma_i)$ (which is approximately $1-\alpha_i-\sigma_i$ for $\alpha_i, \sigma_i \ll 1$). The scattered energy, removed from the image source model at each reflection, is assumed to be fully diffuse and is fed into the radiosity model. Therefore, there will be a build-up of energy in the radiosity model, eventually decaying due to the absorption in the acoustic radiosity model.

The relation between a pressure pulse of the image source model and an energy pulse of the acoustic radiosity model is needed. Starting with a pressure signal, as in Eq. (1), multiplied by an amplitude, A, the pressure pulse can be made to match a given starting power pulse of $W_0 = 1$ (W) for one sample: $p(x_r, y_r, z_r, t) = A\,\delta(R/c - (t - t_0))/R$. Using the fact that the intensity of the first pulse (i.e., the direct sound) is $I_0 = p^2/\rho_0 c$ gives $I_0 = p^2/(\rho_0 c) = A^2/(R^2 \rho_0 c)$. Using $W_0 = I_0 \cdot 4\pi R^2$ gives $W_0 = 4\pi A^2/(\rho_0 c)$ whereby the sought relation, with $W_0 = 1$, can be written:

$$A = \sqrt{W_0 \frac{\rho_0 c}{4\pi}} = \sqrt{\frac{\rho_0 c}{4\pi}}. \tag{14}$$

2.1.4. Implementation details of the hybrid model

In the following, relevant details of the implementation are described. In the numerical implementation of the acoustic radiosity model, the surface of the room is discretised, here using triangular elements. To keep the numerical cost low, a low number of surface elements is preferable. From initially conducted numerical convergence tests, it was concluded that triangle sizes of about $\ell = 1$ m edge length are sufficient. In cases where the distance from any room boundary to the source or receiver is shorter, the triangle size needs to be accordingly smaller.

To describe the energy transfer in the hybrid model, the mean reflection coefficient, $\bar{\rho} = 1 - \bar{\alpha}$, is used, where $\bar{\alpha}$ is the mean surface absorption. For reflection order n (where $n = 0$ represents the direct sound of the original source), the total acoustic energy in the system is $E_{\mathrm{HBRD}}(n) = E_0\bar{\rho}^n$, where E_0 is the starting energy of the initial pulse. Denoting the mean scattering coefficient $\bar{\sigma}$, the energy remaining in the image source model is $E_{\mathrm{IS}}(n) = E_0\bar{\rho}^n(1-\bar{\sigma})^n$. Since the sum of energy in the image source model and in the radiosity model should equal that of the whole system, the energy in the radiosity model can be formulated as $E_{\mathrm{AR}}(n) = E_{\mathrm{HBRD}}(n) - E_{\mathrm{IS}}(n)$. At $n = 0$, all energy is in the image source model, i.e., $E_{\mathrm{IS}}(0) = E_0$ and $E_{\mathrm{AR}}(0) = 0$. At $n = 1$, when the first energy enters the radiosity model, we get $E_{\mathrm{AR}}(1) = E_{\mathrm{HBRD}}(1) - E_{\mathrm{IS}}(1) = E_0\bar{\rho} - E_0\bar{\rho}(1-\bar{\sigma}) = E_0\bar{\rho}\bar{\sigma}$. Looking at how $E_{\mathrm{AR}}(n)$ is updated with n, for $n \geq 2$, it can be shown that

$$E_{\mathrm{AR}}(n) = \bar{\rho}\,\frac{1-(1-\bar{\sigma})^n}{1-(1-\bar{\sigma})^{n-1}}\,E_{\mathrm{AR}}(n-1). \tag{15}$$

The result of Eq. (15) is used in the implementation of the acoustic radiosity model. This provides a simplified and computationally less expensive method to update the energy in the radiosity model independently of the reflection update in the image source model.

In the implementation of the acoustic radiosity model, the discretised version of the surface integrals (Eqs. (5), (7) and (9)) is solved using the *spherical-triangle method* [20] based on Gerard's theorem, which provides high accuracy at a low numerical cost.

For the sampling of the energy response in the acoustic radiosity model, a sampling frequency of $q = 24$ times lower is used compared to that of the image source model and of the final pressure impulse response, which here uses a sampling frequency of $F_s = 48$ kHz. In a post-processing algorithm, the results of the radiosity model at a low sample rate are converted to a pressure impulse response at a high sample rate. First, the energy-time signal is extrapolated to the desired time-extent (e.g., to one or a few seconds duration, depending on the room reverberation), using the late part of the radiosity result. Second, the energy-time signal is upsampled in such a way that each sample is, with uniform probability, located at a randomly selected slot of the possible q slots of the upsampled version. Finally, a Gaussian white noise signal at a high sample rate is multiplied by the square root of the upsampled energy response. That is, the pressure impulse response created in the post-processing of the radiosity results is a noise signal with an envelope determined by the radiosity model. In the hybrid model, the pressure impulse response resulting from the radiosity model is added to that of the image source model. The radiosity model is rerun for the octave bands 31.5 Hz to 8 kHz, i.e., for the corresponding frequency changes in the reflection and scattering coefficients of the surface elements (implemented for computation in parallel).

Throughout the algorithm of the hybrid model, the direction of incidence in the horizontal plane to the receiver is stored, here in increments of 5° angles, producing a total of 72 pressure impulse responses, allowing binaural listening by convolving them with the corresponding head-related transfer functions.

In the image source model, pressure contributions from the original source and the image sources, in the form of $P = A\exp(-\mathrm{j}kR)/R$, are added, where the amplitude, A, is affected by the reflections at the boundaries. For each order of reflection, the amplitude is multiplied by the reflection factor, r, of the reflecting surface. For an impedance description of the boundary (e.g., Marbjerg et al. [15]), r is generally a complex number and is linked to the reflection coefficient in the acoustic radiosity model as $|r|^2 = \rho$. Also, since the angle of incidence is known, angular-dependent reflection factors can be modelled.

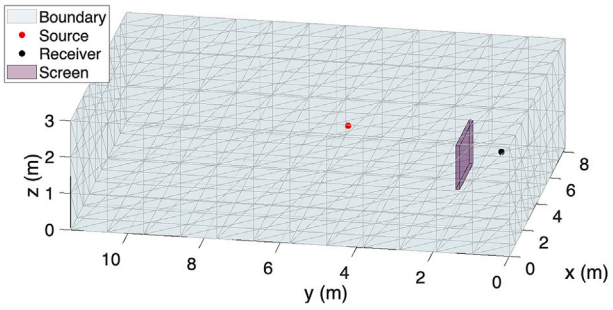

Fig. 1. Geometry of a room with source, receiver and screen.

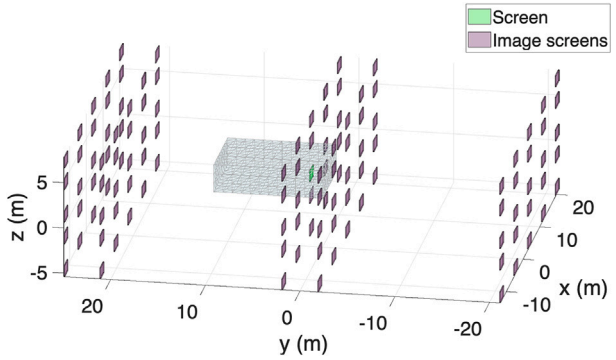

Fig. 2. Geometry for the same situation as in Fig. 1, showing also image screens exemplified for reflection orders 1 and 2.

2.1.5. Modelling the acoustic shielding due to an installed screen in the room

The implementation of a screen in the hybrid model is described here. A single flat screen is assumed to be placed in the interior of the room, similar to an office environment where a screen is used for the acoustic shielding of a workplace. The screen is defined by its centre coordinates x, y and z, its horizontal rotation, its width, height, and thickness, and a possible floating gap between the floor and the screen.

In the image source model, mirror images of the screen are determined up to the same order as for the image sources. In Fig. 1(a) the screen is shown with a thickness of 0.1 m, a width of 2 m, a height of 1.2 m and a float gap between the floor and the bottom edge of 0.6 m, such that the top edge of the screen is at $z = 1.8$ m. The width dimension of the screen is assumed here to be parallel with the x axis but could be rotated in the horizontal plane. Fig. 2 also includes the image screens up to reflection order 2.

The propagation from the source or an image source to the receiver is assumed to be affected by shielding if the straight-line propagation path intersects with the volume of the screen or any image screen. In the algorithm, for each source (original source or image source), all screens (original screen or image screen) that intersect with the source–receiver line are collected. For each source–receiver pair, the collection of shielding screens provides the input to calculate the sound attenuation due to the shielding. First, from the collection of shielding screens, the convex hull is determined for each of the four paths from the source to the receiver: above, below, and around the two sides of the collection of shielding screens, as shown for only the original source and screen in Fig. 3. In Fig. 4 the paths around the sides of the screen and image-screens are shown up to second-order reflection. In the upper part of the

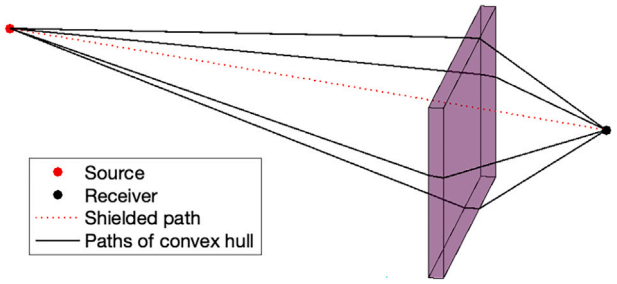

Fig. 3. The four paths of the convex hull (solid line) for reflection order zero. The now shielded path is shown as a dotted line.

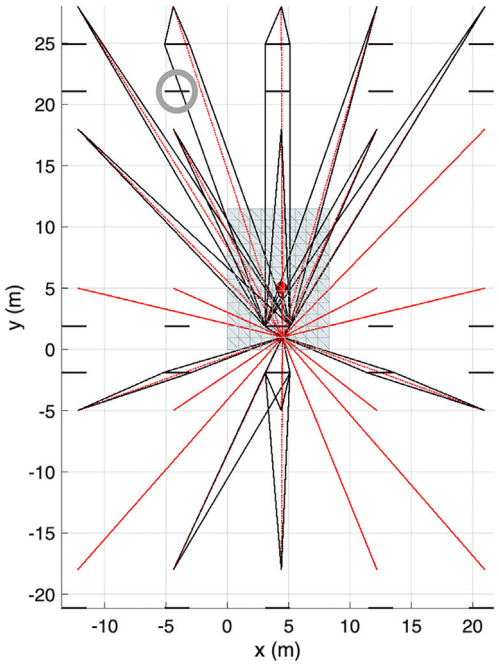

Fig. 4. Top view of the convex hulls (solid black lines) for reflection orders 0–2. The image screen that is marked by a circle is not included in the set of shielding screens, as described in the text. (In coloured figure: direct paths are shown as red lines; solid lines when not shielded and dashed lines when shielded; the source is marked in red.)

figure, it can be seen how the convex hulls are formed around several screening objects (screen and image screens).

A drawback of the algorithm used here to identify shielding objects is that the convex hull is determined only from the set of objects that screen the direct path; the updated paths of the convex hull may intersect with new objects that are not included in the previously determined set of shielding objects. One such example is shown in Fig. 4, in the upper left, where a convex hull path intersects a screen that does not interrupt the direct path from image source to the receiver. Since this kind of omission only takes place for paths that are already shielded at least once by the screen and its images, the error is assumed to be acceptably small.

In the image source model, the calculation of the diffraction, i.e., the sound reducing effect of the screen, is based on the *monopole Rayleigh integral* [23]. The increase in path length, ΔR, when travelling along a convex hull compared to the direct distance from the source (or image source) to the receiver, as described above, is fulfilled by an effective screen with a surface-normal parallel to the source–receiver line. The Rayleigh integral solution is formulated as a single-variable integral. By using Babinet's principle, i.e., that the integral is calculated on the screened domain instead of on the unscreened domain, a numerically fast diffraction calculation is enabled. More details are available in Appendix A.

Calculating the effect of screening in the acoustic radiosity model is performed using a more engineering kind of model based on the one suggested by the Harmonoise project [24], but with zero attenuation assumed when the direct path is unshielded. The frequency-dependent diffraction is based on the Fresnel number, $N_{\mathrm{F}} = 2\Delta R/\lambda$, where $\lambda = c/f$ is the wavelength of the sound, with f being the frequency. The reduction in sound pressure level for one path due to the shielding can then be written as

$$\begin{aligned} \Delta L_D &= -6 - 12\sqrt{N_{\mathrm{F}}}, && N_{\mathrm{F}} < 0.25 \\ &= -8 - 8\sqrt{N_{\mathrm{F}}}, && 0.25 < N_{\mathrm{F}} < 1 \\ &= -16 - 10\log_{10}(N_{\mathrm{F}}), && N_{\mathrm{F}} > 1. \end{aligned} \tag{16}$$

In the acoustic radiosity model, the reflections in all surface elements are assumed to be fully diffuse, with Lambertian reflectance, whereby a screen's shielding effect is blurred and quickly diminishes with increasing reflection order. To model diffraction in the radiosity model, the visibility factor, $V_{i,j}$, as in Eq. (7), is used; however, the value is not restricted to 0 or 1 but is a real number between 0 and 1. For this, Eq. (16) is used assuming that the screen does not introduce an additional time delay and with a simplified input to the Fresnel number, N_F; only one value of N_F is calculated as described below.

For the line between a pair of sending and receiving surface element centres, the distance from the line to the centre coordinate of the screen is calculated. If this distance is shorter than a chosen value, D, a shielding is calculated assuming an object with height and width equal to $2D$ placed with its centre on the line. Thus, the diffraction calculation is simplified to at most using Eq. (16) once for each pair of sending and receiving surface elements, with a value of N_F assuming an effective size, $2D$, of the shielding object; here, $2D$ is set to the smaller of the height and width dimensions of the screen. (The same approach is also used for the energy transfer from the surface elements to the receiver.) Hence, the four paths around the object become identical and are summed up energetically, leading to

$$V_{i,j} = 4 \cdot 10^{\Delta L_D/10} \tag{17}$$

where L_D is calculated for one path. If the path is not shielded, $V_{i,j} = 1$. With this approach, the acoustic radiosity model also considers a diffraction effect that depends on both geometry and frequency. (For calculation in octave bands, the centre frequency of each band is used as input to the diffraction calculation.) It can be noted that in the limit of small detour ($\Delta R \to 0$) or low frequency, we have that $N_F \to 0$ and $\Delta L_D \to -6$ (dB), resulting in $V_{i,j} = 4 \cdot 10^{-0.6} \approx 1$, as desired.

The diffraction modelling used here is insensitive to the detailed effect of each edge on amplitude and phase; however, it captures a physically reasonable overall dependence on geometry and frequency and is attractive with respect to its adaptability to different situations of sound shielding and to the low numerical cost. No reflection on the screen surface is modelled, which is why the current implementation can be seen as modelling screens with fully absorbing surfaces.

The modelling of a single screen, as described above, can be extended to situations with multiple screens. In the following, results are presented for situations including shielding by one or two screens.

3. Numerical results and comparison with measurements

As a test case, a room with a volume of about 300 m^3 was used; width $L_x = 8.4$ m, length $L_y = 11.8$ m, and height $L_z = 3.1$ m, in an (x, y, z)-coordinate system. The source was located at $(x, y, z) = (3.8, 2.1, 1.62)$ m and two receiver positions were used: Position 1 at $(3.8, 5.1, 1.20)$ m and Position 2 at $(3.8, 8.8, 1.20)$ m. Situations were investigated with and without two identical rectangular-shaped screens shielding the direct

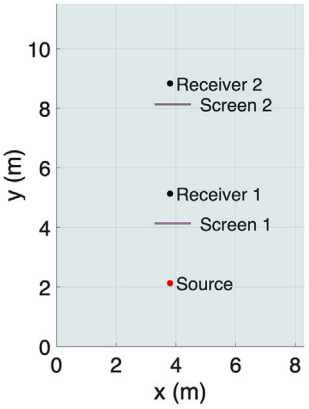

Fig. 5. Locations of source, receivers and screens.

source–receiver path. The screen height is 1.7 m, with a float gap of 0.1 m above the floor such that the top edge is at $z = 1.8$ m. The screen width is 1.2 m and the screen thickness is 5 cm. Screen 1 was located with its centre point at $(x, y) = (3.9, 4.1)$ m, i.e., distanced 2 m from the source and 1 m from receiver 1 along the y axis, and offset by 0.1 m edgeways along the x axis. Screen 2 was located at the same y-coordinate but placed 0.7 m in front of the receiver at Position 2, i.e., at $(x, y) = (3.9, 8.1)$ m. The locations of the source, receivers and screens are shown in Fig. 5.

Using reflection order 8 in the image source model and 30 shots in the radiosity model, the resulting impulse response is about 0.3 s, at which point the decay is exponential and extended in the radiosity model to about 1.5 s. (The calculation time for each case is about 10 minutes using Matlab version 2023b on a 2021 MacBook Pro with M1 Max chip and 64 GB RAM.)

Calculations and measurements were made for the room, with and without the shielding screens. Photos of the real room are shown in Fig. 6, with and without the screens shielding the direct paths between the source and receivers; the locations of the receivers used for the comparison are marked in the photo for the non-shielded case. (The third microphone visible in the photo is used to observe the amplitude of the direct sound.) The ceiling has a 5 cm cover of porous absorbers and the room is furnished with desks and chairs. (The room is used for teaching, but it could alternatively be used as an open-space office.) Typical absorption data in octave bands were used as the first input to the model, distinguishing between the ceiling and the remaining surfaces. (The linoleum floor is assumed to have the same absorption as the walls.) Measured reverberation times were then used to update the absorption data. The resulting reverberation times of the model are plotted in Fig. 7 together with the measured reverberation times for receiver positions 1 & 2. Comparing the measured and modelled reverberation times in Fig. 7, the agreement is good at mid- and high frequencies, whereas at the lower frequency bands of 63 Hz and 125 Hz there are significant deviations. At lower frequencies the uncertainty is expected to increase and at 31.5 Hz large deviations are shown. Further iterations of updating the model's absorption data were not pursued in the current work, even though better agreement with measured reverberation times may have been attained. The values used for the absorption coefficients in the model are shown in Table 1 together with the scattering coefficients used. The ceiling absorption decreases above 500 Hz, which may be related to the surface paint of the porous absorbers. Air attenuation was not included in the current version of the model.

3.1. Influence of screen 1 on receiver 1

In Fig. 8 the measured impulse responses are shown for Receiver 1 with and without Screen 1. Without the screen, the direct sound is

Fig. 6. Photo of room without screening (a) and with screening (b). The receiving microphones at positions 1 & 2 are marked by gray circles in (a), where position 2 is furthest away from the source.

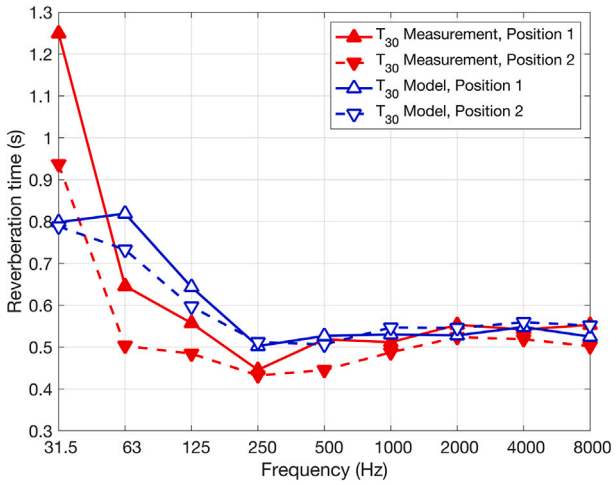

Fig. 7. Reverberation time, $T_{30}(s)$, estimated from the modelled impulse response and measured in the room without screen at positions 1 and 2.

Table 1
Table of used absorption coefficients, α, and scattering coefficients, σ.

Frequency (Hz)	31.5	63	125	250	500	1k	2k	4k	8k
α, ceiling (-)	.40	.20	.22	.49	.58	.55	.36	.36	.36
α, other surfaces (-)	.20	.20	.18	.33	.39	.37	.24	.24	.36
σ, all surfaces (-)	.20	.20	.20	.20	.20	.27	.47	.67	.80

observed to arrive at about 8.8 ms. With the screen, the first arrival occurs at the same time, identified as being due to transmission through the screen. Thereafter, three pulses can be distinguished before 10 ms, which are identified as being due to paths via the top edge and via the left and right edges of the screen. At around 12 ms, contributions via the screen's bottom edge and via a floor reflection are concluded to arrive, whereas at around 13 ms unshielded paths via a reflection in the ceiling are identified both with and without the screen. The existence of two large ventilation ducts below the ceiling (visible in Fig. 6) is concluded to cause additional contributions at around both 12 ms and 14 ms.

In Fig. 9 the measured and modelled results are compared without a screen (top plot) and with a screen (bottom plot). The arrival times of the modelled contributions are in concordance with those described above for the measured results except for the effects of the ventilation ducts and the direct transmission through the screen, which are not included in the model; however, their time structures are generally more precise (less spread out). The direct sound of the modelled impulse response can be seen to show small oscillating tails (see the upper plot in Fig. 9). These oscillations are due to the image source modelling and its

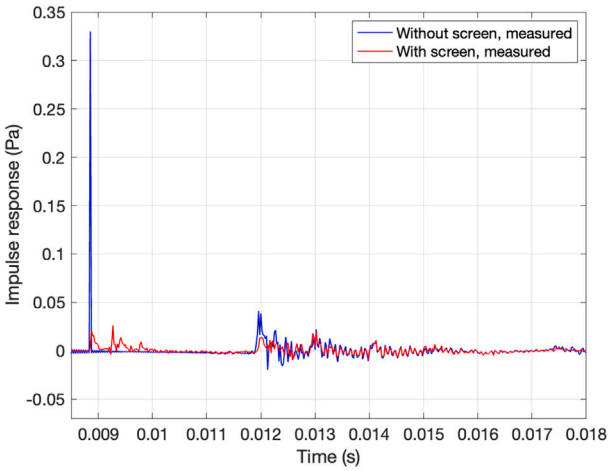

Fig. 8. Measured impulse responses at receiver 1 with and without screen 1.

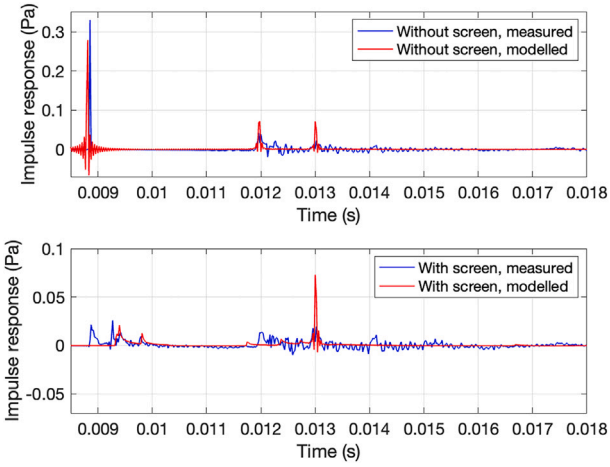

Fig. 9. Comparison between measured and modelled impulse responses at receiver 1 without screen 1 (upper plot) and with screen 1 (lower plot).

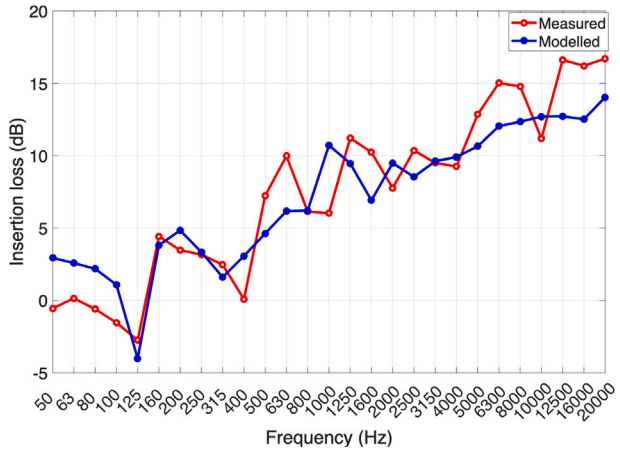

Fig. 10. Measured and modelled insertion loss of screen 1, plotted as function of frequency in 1/3-octave bands based on early part impulse responses at receiver 1, before first reflection in room boundary.

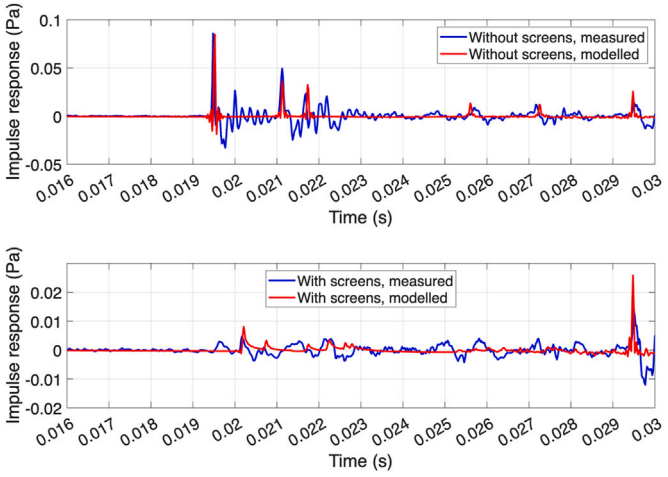

Fig. 11. Comparison between measured and modelled impulse responses at receiver 2 without screens 1 and 2 (upper plot) and with screens 1 and 2 (lower plot).

conversion from frequency domain to time domain and are concluded to not make any audible effect. (The oscillating tails can be made to vanish by adjusting the source–receiver travel time to be a whole number of samples.)

The very early part of the results, up to 11 ms, is used for studying the measured and predicted insertion loss in 1/3-octave bands, which includes the paths via the screen's top and left and right edges. In Fig. 10 the resulting insertion losses for both measurements and modelling are displayed; the agreement is fairly good and shows an expected trend of increased insertion loss with frequency.

3.2. Influence of two screens

To investigate a situation with multiple screens, the response at Receiver 2 is studied with and without the insertion of two screens. The geometry is shown in Fig. 5 and the impulse responses are shown in Fig. 11. For the results without screens (upper plot in Fig. 11), it can be seen that the direct sound (at 19.5 ms) arrives later than for Receiver 1 (compare with Fig. 9), as expected, and is followed by reflections from the floor (at 21.1 ms), the ceiling (at 21.7 ms), the ceiling and floor (at 25.6 ms), and the floor and ceiling (at 27.3 ms). These contributions are reduced by the screens, as can be seen in the lower plot in Fig. 11, as expected. (Notice the different plot scales of the upper and lower plots.) The peak at 29.5 ms is not reduced by the screens, which is as expected since the corresponding path via the side wall is not obstructed by the screens.

3.3. Study of phased reflection

Instead of using real-valued reflection factors, derived from given absorption coefficients, a study is made with the ceiling modelled as a finite-impedance plane, which allows a physically more correct modelling that provides possible phase delays of the reflections. A spherical reflection factor is used, which depends on frequency, impedance, angle of incidence and source–receiver distance. The impedance is given by the Miki model for a 50 mm thick porous absorber with flow-resistivity 10 kNs/m^4 on hard backing. Further details on the impedance modelling and the spherical reflection factor can be found in [25].

In Fig. 12 calculated impulse responses with and without the phased model are shown for Receiver 2 without any screen. The phased model of the ceiling uses the complex-valued spherical reflection factor, as described above, whereas the un-phased model uses the absolute value of the spherical reflection factor. It can be seen that the third peak (at 21.7 ms), identified as the first ceiling reflection, gets an altered response when changing from the un-phased model to the phased model; the initial amplitude is decreased and the response is spread out in time,

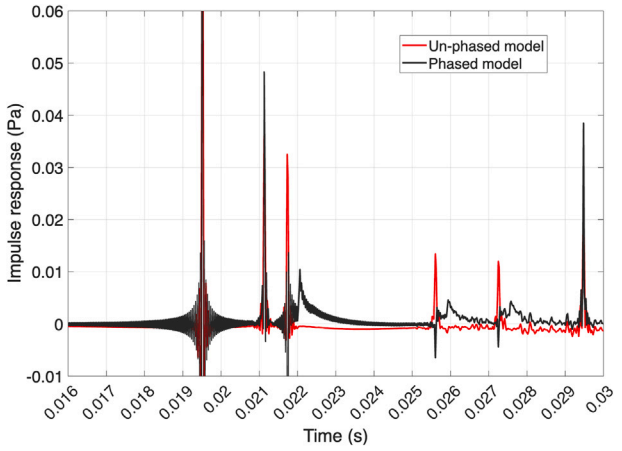

Fig. 12. Calculated impulse responses with and without phased model (receiver 2, no screen).

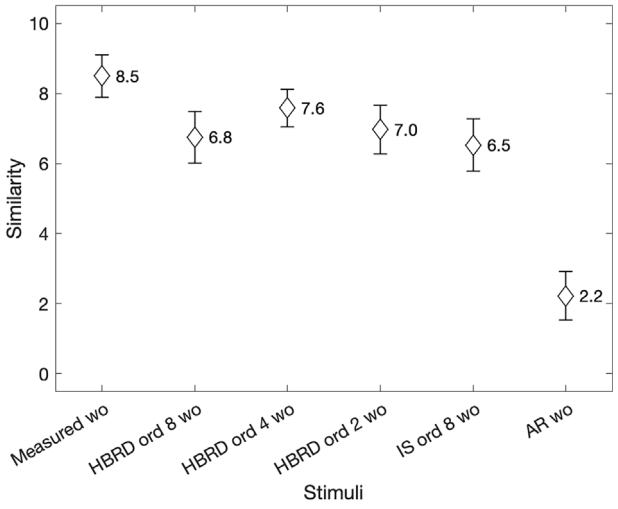

Fig. 13. Rated similarity for speech sound without screen, with measurement without screen (*Measured wo*) as reference. IS stands for image source model, AR stands for acoustic radiosity model, and HBRD stands for hybrid model. The plot shows mean value (◊) and estimated 95% confidence interval.

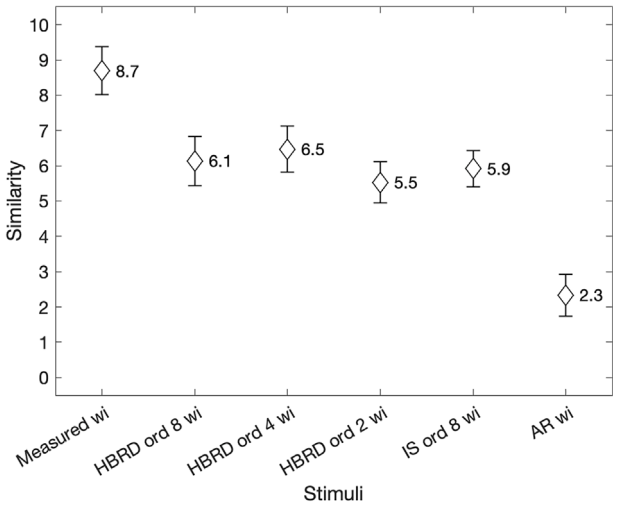

Fig. 14. Rated similarity for speech sound with screen, with measurement with screen (*Measured wi*) as reference.

also showing a second peak (delayed by about 0.3 ms), attributed to the reflection in the hard backing. The two following peaks (at 25.6 and 27.3 ms) are also altered due to a ceiling reflection; the initial amplitude decreases, and the response spreads out over time.

4. Listening test

4.1. Procedure of the listening test

A listening test was carried out at the facilities of Applied Acoustics, Chalmers. Thirty-six students participated in the test, 23 male and 13 female, with an average age of 24 years. All participants reported normal hearing. When asked about their previous experience with listening tests, 23 participants reported never having participated in one before, seven had participated in 1–2 tests, three in 3–5 tests, and the remaining three participants had taken part in more than five tests. Hence, the group of participants can be considered as non-expert listeners.

Each session of the test lasted around 15 minutes, with three or four participants at a time. They performed the test independently of each other, with visual and acoustic shielding that prevented disturbance from the others. The test was implemented using a HEAD acoustics SQala jury testing system, and the stimuli were played back via calibrated headphones (Sennheiser HD 650 with HEAD acoustics BPU interface). For each step of the test, the test person was asked to rate the similarity between one reference sound and six other sounds. The rating was made on a scale in 11 steps, from *Not at all* similar to *Extremely* similar, later converted to a numerical scale of 0–10.

To produce the sound stimuli, measured and calculated impulse responses in mono at Receiver 1 were convolved with two dry sounds of different character, *Male speech* and *Funky drums*, where the temporal envelope largely differs between the speech sound and the impulse-like drum sound. The A-weighted equivalent level of the sounds without the screen was 68 dB for the speech sounds and 73 dB for the drum sounds, and 1 dB lower for the sounds with Screen 1.

The first three steps of the listening test used the speech sound and the remaining three steps used the drum sound. In the first step the measured impulse response without a screen was used as a reference and the six sounds to be rated contained (in random order) a copy of the reference sound and five sounds using different calculation settings: the hybrid model with reflection order 2, 4 or 8 of the image sources part; pure image source model with reflection order 8; and pure radiosity model. The second step was set up analogously to the first step but for impulse responses with the screen.

In the third step, situations with and without a screen were compared. The measured impulse response with the screen was used as a reference and the six sounds to be rated contained: a copy of the reference; the sound from the measured impulse response without the screen; two sounds from the hybrid model (image source model with reflection order 8), with and without the screen; and two sounds from the pure image source model (reflection order 8), with and without the screen. The remaining three steps were set up in the same way as the first three steps but with the speech sound exchanged for the drum sound.

4.2. Results of the listening test

An initial overall statistical analysis was performed using the Friedman test showing that there is a main effect of the stimulus type on the similarity rating ($p < 10^{-17}$, $\chi^2(5) > 90$). Figs. 13–18 show mean values and their corresponding 95% confidence intervals of rated similarities. (The calculated 95% confidence interval assumes that the mean value follows a normal distribution.)

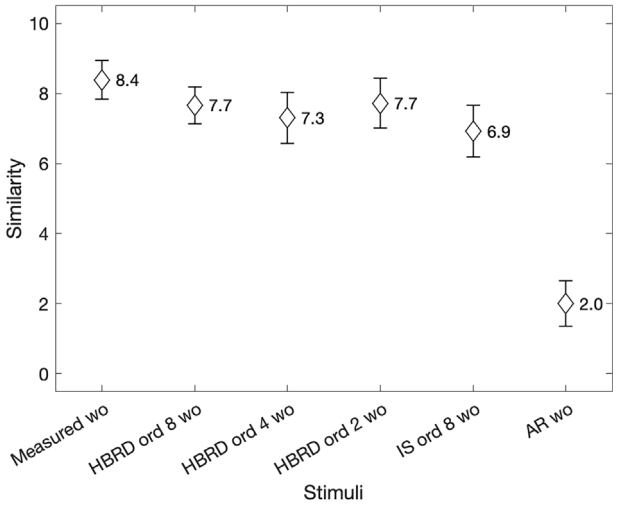

Fig. 15. Rated similarity for drum sound without screen, with measurement without screen (*Measured wo*) as reference.

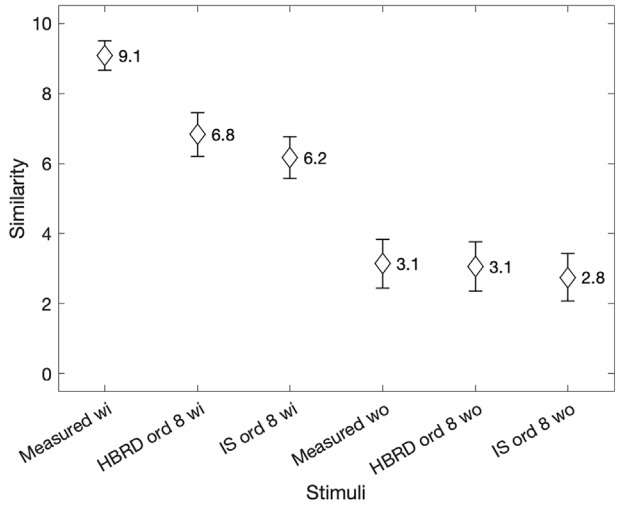

Fig. 17. Rated similarity for speech sound with and without screen, with measurement with screen (*Measured wi*) as reference.

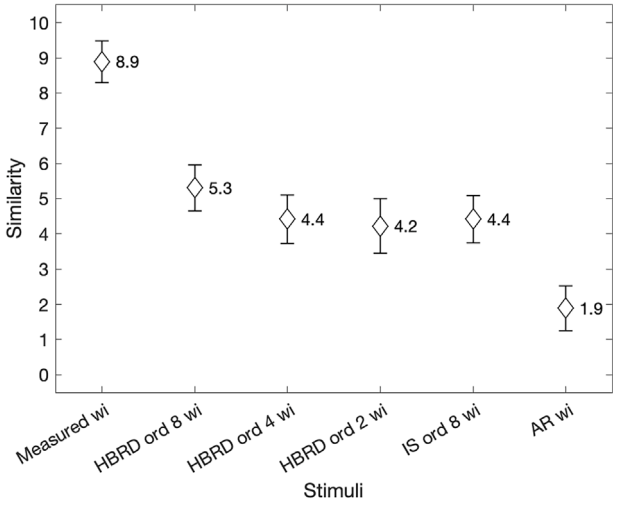

Fig. 16. Rated similarity for drum sound with screen, with measurement with screen (*Measured wi*) as reference.

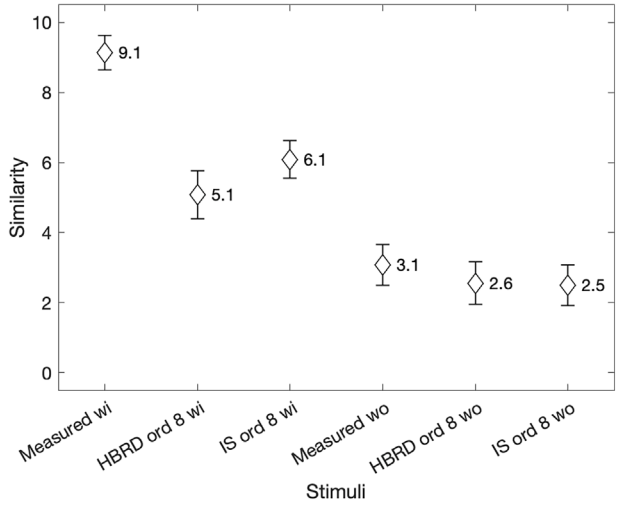

Fig. 18. Rated similarity for drum sound with and without screen, with measurement with screen (*Measured wi*) as reference.

4.2.1. Results of separate similarity ratings for screen and no-screen conditions

In Figs. 13–16, the results of Steps 1, 2, 4, and 5 are shown. For *male speech* (Figs. 13 and 14) the mean similarity for the hybrid model with image source reflection order 8 was 6.8 without a screen and 6.1 with a screen compared to its respective reference, while the mean similarity for the reference sound to itself was 8.5 without a screen and 8.7 with a screen. For *funky drums* (Figs. 15 and 16) the corresponding mean similarity was 7.7 without a screen and 5.3 with a screen, while the mean similarity for the reference was 8.4 without a screen and 8.9 with a screen.

Hence, the mean rated similarity between the measurement and the hybrid model (with image source reflection order 8) reached 6–7 with and without the screen for the speech sounds. For the drum sounds the rating was higher without the screen (7.7) and lower with the screen (5.3). When comparing the reference sound with itself, the mean rating was 8–9 with and without the screen. The poorest ratings resulted from using only the radiosity model, ranging from 1.9 to 2.3. Thus, rated values of about 5 or higher can be seen to correspond to fairly high similarity. The other calculation settings, i.e., the hybrid model with image source reflection orders 2 and 4, and the pure image source model with reflection order 8, gave similar results as the hybrid model with image-source reflection-order 8.

4.2.2. Results of similarity ratings between screen and no-screen conditions

It is also of interest to study the model's ability to distinguish between the situations with and without a screen. Figs. 17 and 18 show results from steps 3 and 6. For *male speech* (Fig. 17), with the reference sound being with the screen, a paired comparison showed a clear difference between the hybrid model with the screen, rated 6.8, and the hybrid model without the screen, rated 3.1. To confirm statistical significance, a Wilcoxon signed rank test was used, resulting in a p-value of $4.5 \cdot 10^{-8}$ after Bonferroni correction. (Also, the measured impulse response without the screen gave a mean value of 3.1.)

For *funky drums* (Fig. 18) the corresponding difference in similarity for the hybrid model was not as large, 5.1 with screen and 2.6 without screen, however, still highly significant with a resulting p-value of $6.3 \cdot 10^{-5}$ from the Wilcoxon signed rank test and after Bonferroni correction. The reason for the relatively high rating of the reference compared to itself shown here (mean value 9.1 for both speech and drum sounds) is

likely due to the fact that the poorest sounds, produced by using only radiosity, are not included here among the stimuli.

When comparing the ratings for the hybrid model with those for the pure image source model in Figs. 17 and 18, it can be seen that they are about the same. No statistically significant differences were found. This suggests that, for the relatively high reflection order used here, i.e., up to eight-order reflections in both the hybrid model and the pure image source model, the test participants did not perceive a significant difference between the two model settings in the studied situations. For implementations using lower-order reflections, however, it is reasonable to assume that a pure image source model would deviate more noticeably from the hybrid model.

5. Discussion

The modelled impulse responses agree reasonably well with the measured ones, both with and without shielding, using one or two screens. In terms of the arrival times of the direct sound and the identified early reflections, the agreement is good. However, the shapes of the modelled responses are more distinct than those of the measured responses. This is attributed to room details that are not included in the model, such as furniture and surface unevenness (e.g., window niches). In addition, the ventilation ducts below the ceiling were concluded to contribute to pulses in the measured responses that are not included in the model. Concerning the insertion loss of a single screen, the model shows fair agreement with the measured results when evaluated using the early contributions, before reflections in room boundaries.

The listening experiment shows an overall reasonably high similarity rating between the results of the developed hybrid model and the measured impulse responses. However, when comparing the ratings for the hybrid model with those for pure image sources, no significant differences were found for the studied situations and model settings, which included reflections up to eighth order. Thus, the listening test results are inconclusive in determining whether the developed hybrid model performs better than the pure image source model. Nevertheless, for broader applications of the hybrid approach, e.g., to larger and more irregularly shaped rooms, implementations with lower-order reflections would be of interest, where it is reasonable to assume that a pure image source model would deviate more noticeably from the hybrid model.

That the mean ratings for comparisons of reference sounds with themselves were in the range of 8–9, rather than 10, may partly be explained by the participants being non-experts. Additionally, the absence of very poor-quality sounds in the test may have contributed to this result.

For the modelling of a screen between the source and receiver, a Rayleigh integral approach was used as an extension of the image source modelling, as detailed in Appendix A. The implementation as a single-variable integral gives short calculation times, but the accuracy may be insufficient; for future work, a higher accuracy may be achieved by using a more exact model, e.g., as provided by the Edge Diffraction Matlab toolbox (EDtoolbox) [26].

The developed model represents a compromise between physical detail and computational efficiency, aimed to be suitable for design-oriented studies and perceptual investigations of open-plan office acoustics. Its hybrid formulation allows early reflections and shielding effects to be modelled with a level of temporal and spectral detail that would be difficult to achieve with a purely energetic approach, while still providing an efficient description of the later sound field. At the same time, the present implementation involves simplifying assumptions that limit its accuracy in specific respects, including the restriction to shoebox geometries, the omission of interior objects, and the use of approximate treatments of scattering, diffraction, and diffuse reflection. These simplifications partly explain the remaining discrepancies with measurements, particularly in the fine temporal structure of the impulse responses. Nevertheless, the overall agreement with measurements and listening-test results indicates that the model provides a useful basis for further development, for example towards more general room geometries, refined diffraction and scattering models, and inclusion of additional physical details where required.

6. Conclusion

The implementation of a hybrid room acoustics prediction tool using phased image sources and acoustic radiosity is described and evaluated using measurements and a listening test for situations with and without screens installed in a room, shielding the direct path between source and receiver. The implementation includes a lower sample rate for the acoustic radiosity model, a factor of 24 lower than for the image source model (48 kHz used here), as well as a derived update factor for the radiosity solution depending on the mean scattering and reflection coefficients. In addition, new approaches to modelling screen diffraction in the image source model and in the acoustic radiosity model are described. The evaluation using measured impulse responses shows reasonable agreement in the time-domain responses both with and without screens and in the frequency-domain insertion loss of a single screen for the first few diffraction contributions. A listening experiment has demonstrated a reasonably high similarity rating between the outcomes of the developed hybrid model and the measured impulse responses. Subsequently, the model could be applied to investigate the effects of office screens and contribute to the assessment of open office work environments. Future work may involve adapting the model for applicability to larger rooms with general geometries and incorporating air absorption.

CRediT authorship contribution statement

Jens Forssén: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Leon Müller:** Writing – review & editing, Investigation. **Elin Hedlund:** Writing – review & editing, Validation, Methodology, Investigation, Data curation. **Wolfgang Kropp:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

AI declaration

During the preparation of this work the authors used the AI assistant in Overleaf for language improvements. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Wolfgang Kropp reports that financial support was provided by Afa Försäkring. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The project is funded by Afa Försäkring, project nr 190278. The listening test was enabled by resources funded by the HEAD Genuit Foundation under the agreement P-22/01-W.

Appendix A. Implementation of the Rayleigh integral approach as a single-variable integral for diffraction calculation

The Rayleigh integral is an exact solution to the wave equation where the acceleration is known on a plane boundary of a semi-infinite domain, e.g., Pierce [27]. The received pressure may be formulated as

$$p_{\mathrm{Rayl}}(x_R, y_R, z_R, t) = \frac{\rho_0}{2\pi} \iint_S \frac{a_n(t - R_R/c)}{R_R} \mathrm{d}S \tag{18}$$

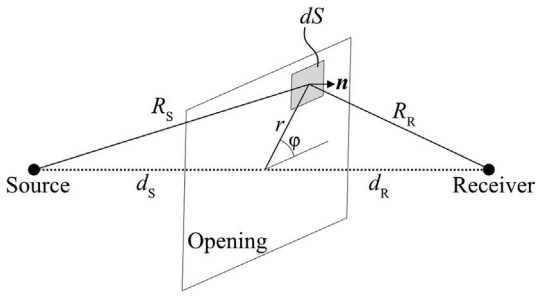

Fig. 19. A sketch of the geometry used for the Rayleigh integral calculation.

where ρ_0 is the density of the medium, $a_n(t)$ is the time function of the acceleration in the normal direction to the boundary, and R_R is the distance between a boundary point and the receiver.

Using Babinet's principle (also called the Bouwkamp–Babinet principle) to calculate the effect of a screen, the Rayleigh integral can be calculated over the surface of the screen, with $a_n(t)$ as due to the incident field, and the resulting pressure p_{Rayl} is subtracted from the incident pressure p (given by Eq. (1)) (see Ref. [23]). This models a screen with a zero-pressure boundary condition [23]. To calculate the acceleration due to the incident field, conservation of momentum gives

$$a_n(t) = -\frac{1}{\rho_0}\frac{\partial p(t)}{\partial n} \tag{19}$$

where $\partial p/\partial n$ is the gradient of the incident pressure in normal direction to the surface. Using the notation in Fig. 19, the result of inserting Eqs. (19) and (1) in Eq. (18) can be written as

$$p_{\text{Rayl}}(\mathbf{x}_R, t) = \frac{-1}{2\pi}\iint_S \frac{\partial}{\partial n}\left(\frac{\delta(t - t_0 - R_S/c - R_R/c)}{R_S}\right)\frac{1}{R_R}\,\mathrm{d}S \tag{20}$$

where $\mathbf{x}_R = (x_R, y_R, z_R)$. Using that $\frac{\partial}{\partial n}$ is the same as $\frac{d_S}{R_S}\frac{\partial}{\partial R_S}$, Eq. (20) can be further developed as

$$\begin{aligned} p_{\text{Rayl}}(\mathbf{x}_R, t) &= \frac{-d_S}{2\pi}\iint_S \frac{\partial}{\partial R_S}\left(\frac{\delta(t - t_0 - R_S/c - R_R/c)}{R_S}\right)\frac{1}{R_S R_R}\,\mathrm{d}S \\ &= \frac{-d_S}{2\pi}\iint_S \left(\frac{\delta'(t - t_0 - R_S/c - R_R/c)(-\frac{1}{c})}{R_S}\right. \\ &\quad \left. - \frac{\delta(t - t_0 - R_S/c - R_R/c)}{R_S^2}\right)\frac{1}{R_S R_R}\,\mathrm{d}S. \end{aligned} \tag{21}$$

The second term in the parenthesis of Eq. (21) is identified as a near-field component; keeping only the far-field component, Eq. (21) is rewritten as follows:

$$p_{\text{Rayl}}(\mathbf{x}_R, t) = \frac{d_S}{2\pi c}\iint_S \frac{\delta'(t - t_0 - R_S/c - R_R/c)}{R_S^2 R_R}\,\mathrm{d}S. \tag{22}$$

To exploit that the integrand is symmetric around the source–receiver axis, a change of variables is made according to $\mathrm{d}S = \mathrm{d}\varphi\, r\, \mathrm{d}r$, which gives the following:

$$\begin{aligned} p_{\text{Rayl}}(\mathbf{x}_R, t) &= \frac{d_S}{2\pi c}\int_0^{r_{\max}}\int_0^{\varphi_{\max}(r)} \frac{\delta'(t - t_0 - R_S/c - R_R/c)}{R_S^2 R_R}\,\mathrm{d}\varphi\, r\, \mathrm{d}r \\ &= \int_0^{r_{\max}} \frac{d_S}{2\pi c}\,\varphi_{\max}(r)\,\frac{\delta'(t - t_0 - R_S/c - R_R/c)}{R_S^2 R_R}\, r\,\mathrm{d}r. \end{aligned} \tag{23}$$

In Eq. (23) $\varphi_{\max}(r)$ is the angular extent of the screening surface at radius r; for a circular disc placed symmetrically, $\varphi_{\max}(r) = 2\pi$ for all values of r up to the edge of the disc, while for another screening shape, the angular extent would vary with r.

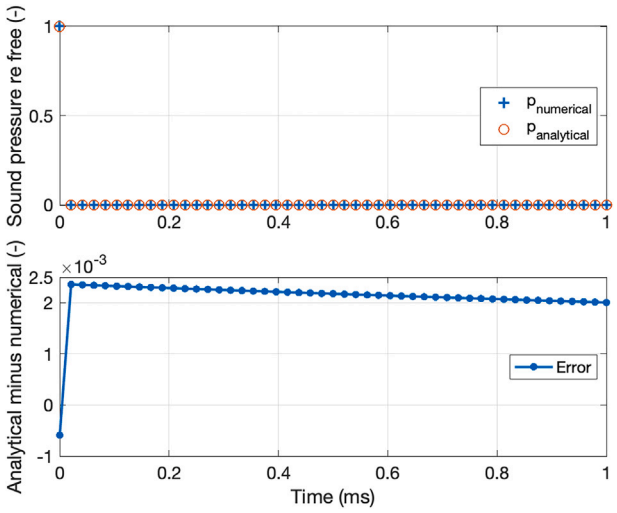

Fig. A.20. Analytical and numerical solution for free-field impulse response plotted relative to free field (upper) and their difference (lower).

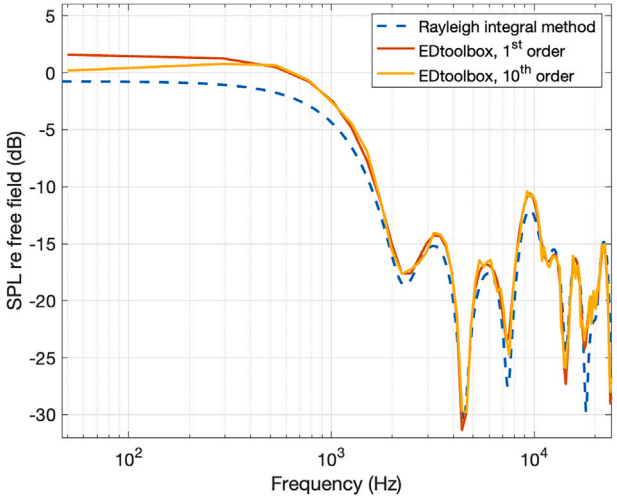

Fig. A.21. Sound pressure level as function of frequency from using the Rayleigh integral approach to model screen diffraction compared with those from the EDtoolbox with diffraction orders 1 and 10.

The final analytical result (Eq. 23) is discretised and calculated numerically to provide the diffraction impulse response. The discretisation is made equidistant in time, $\Delta t = 1/F_s$, such that the discrete time is $t_i = i\,\Delta t, i = 0 \ldots N-1$ (where N is the number of samples), which makes the steps in r non-equidistant, i.e., Δr varies with r. The discrete delta function is modelled as non-zero over an interval of Δt, with an amplitude of $1/\Delta t$, such that the time integration gives a unit result. Consequently, the derivative of the delta function is modelled as a positive and a negative pulse (each with duration Δt) with amplitudes $\pm 1/\Delta t^2$. Except for the derivative of the delta function, the remainder of the integrand in Eq. (23) can be written in discretised form as

$$I_i = \frac{d_S\ \varphi_{\max}(r(t_i))\ r(t_i)\ \Delta r(t_i)}{2\pi\ c\ R_S^2(t_i)\ R_R(t_i)}. \tag{24}$$

For each screening situation (i.e., for each source–receiver pair where the direct line of sight is shielded), $I_i, i = 1 \ldots N$, is calculated and the discrete function $p_{\text{Rayl}}(\mathbf{x}_R, t_i)$ is calculated as the numerical difference

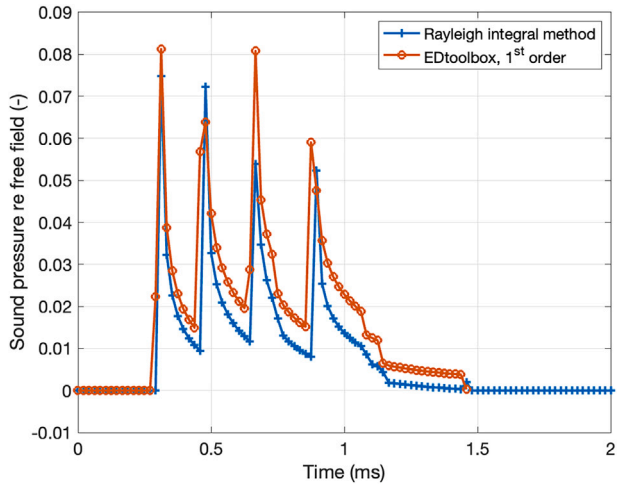

Fig. A.22. Sound pressure in time-domain from using the Rayleigh integral approach to model screen diffraction compared with that of the EDtoolbox with diffraction order 1.

of $I_i/\Delta t^2$, i.e., $p_{\text{Rayl}}(\mathbf{x}_R, t_i) = \texttt{diff}(I_i/\Delta t^2)$. The operation `diff` takes the difference of consecutive samples and, to fulfil the desired integral, a leading zero is entered in I_i. The variable values needed in each time step t_i are found by first calculating the total distance, $R = R_S + R_R$, from source to receiver via the integration surface, as $R(t_i) = d_S + d_R + c\,t_i$, for which $t = 0$ corresponds to the arrival of the first pulse to the receiver. From $R(t_i)$, $r(t_i)$ can be calculated as follows:

$$r(t_i) = \frac{1}{2}\sqrt{R(t_i)^2 - 2(d_S^2 + d_R^2) + \frac{d_S^2 - d_R^2}{R(t_i)^2}}. \tag{25}$$

From $r(t_i)$ the remaining variables R_S and R_R are given by $R_S(t_i) = \sqrt{r^2(t_i) + d_S^2}$ and $R_R(t_i) = \sqrt{r^2(t_i) + d_R^2}$. To improve the numerical accuracy, the values of $r_i = r(t_i)$ are exchanged for the midpoint: $r_i = (r(t_i) + r(t_{i+1}))/2$.

To validate the analysis and numerical implementation, the accuracy in reproducing the free-field impulse response is first evaluated. For this, a test case with a source–receiver distance of 6 m is chosen with the plane of integration midway (i.e., $d_S = d_R = 3$ m). In Fig. A.20 the upper plot shows the analytical and numerical solution in the time steps t_i from zero up to 1 ms, where zero time is set when the pulse reaches the receiver. The solution is plotted relative to the free field and it displays a unit pulse at time zero followed by zeros, as expected. The lower plot shows the difference between analytical and numerical results in which the size of the error is visible; the first sample of the numerical solution is overestimated by about $6 \cdot 10^{-4}$ and the remaining samples are underestimated by about $2.5 \cdot 10^{-3}$ or less. Hence, in relation to the free field, the relative error is less than 2.5‰.

For a screen, the Rayleigh integral is calculated over the extent of the screened area, and the result is subtracted from that of the free-field solution. To validate the implementation for a screened case, the results are compared with those obtained by using the freely available Edge Diffraction Matlab toolbox (EDtoolbox) [26]. The calculated case is for a screen placed midway between the source and receiver distanced by 6 m where the screen is 1 m wide and 1.2 m tall with its centre offset both upward and edgeways by 0.1 m. In addition to first-order diffraction, results for 10th-order diffraction were calculated using the EDtoolbox. The comparisons with the Rayleigh integral calculation are plotted in the frequency domain in Fig. A.21. Compared with the EDtoolbox results, the Rayleigh integral results are shown to generally provide underestimates with differences of about 1–2 dB at the interference peaks at higher frequencies, above 1 kHz, and larger differences at lower frequencies. In the low-frequency limit it can be seen that the higher-order results from the EDtoolbox tend to lower values thereby reducing the difference to the Rayleigh integral results. The overall agreement of using the Rayleigh integral is not very good, but the accuracy is deemed sufficient for the purpose of the work. It should also be noted that in the EDtoolbox a rigid screen surface is modelled whereas in the Rayleigh integral approach used here the screen surface is seen as having a zero-pressure boundary condition, whereby the two modelled situations are slightly different and expected to provide different results.

The corresponding time-domain signals obtained using the Rayleigh integral and the EDtoolbox at first-order diffraction are plotted in Fig. A.22. The agreement in the time structure is reasonable, and one can observe how each of the four edges of the screen provides a tail-shaped pulse.

Data availability

The sound stimuli used in the listening experiment of this study are openly available on Zenodo at http://doi.org/10.5281/zenodo.17376995.

References

[1] Radun J, Hongisto V. Perceived fit of different office activities – the contribution of office type and indoor environment. J Environ Psychol 2023;89:102063. https://doi.org/10.1016/j.jenvp.2023.102063
[2] Jahncke H. Cognitive performance and restoration in open-plan office noise [Ph.D. thesis], Sweden: Luleå University of Technology; 2012.
[3] Radun J, Maula H, Rajala V, Scheinin M, Hongisto V. Speech is special: the stress effects of speech, noise, and silence during tasks requiring concentration. Indoor Air 2021;31:264–74. https://doi.org/10.1111/INA.12733
[4] Ellermeier W, Zimmer K. The psychoacoustics of the irrelevant sound effect. Acoust Sci Technol 2014;35:10–16. https://doi.org/10.1250/ast.35.10
[5] Yadav M, Georgi M, Leist L, Klatte M, Schlittmeier SJ, Fels J. Cognitive performance in open-plan office acoustic simulations: effects of room acoustics and semantics but not spatial separation of sound sources. Appl Acoust 2023;211:109559. https://doi.org/10.1016/j.apacoust.2023.109559
[6] Vachon F, Winder E, Lavandier M, Hughes RW. The bigger the better and the more the merrier? realistic office reverberation levels abolish cognitive distraction by multiple-voice speech. In: 12th ICBEN Congress on noise as a public health problem. Zurich, Switzerland; 2017, June. p. 1–9.
[7] Yadav M, Kim J, Cabrera D, de Dear R. Auditory distraction in open-plan office environments: the effect of multi-talker acoustics. Appl Acoust 2017;126:68–80. https://doi.org/10.1016/j.apacoust.2017.05.011
[8] Reinhart P, Zahorik P, Souza PE. Effects of reverberation, background talker number, and compression release time on signal-to-noise ratioa. J Acoust Soc Am 2017;142:EL130–EL135. https://doi.org/10.1016/10.1121/1.4994683
[9] Kuttruff H ed Room Acoustics. 6th ed. CRC Press; 2016. https://doi.org/10.1201/9781482286632
[10] Hodgson M. Evidence of diffuse surface reflections in rooms. J Acoust Soc Am 1991;89:765–71. https://doi.org/10.1121/1.1894636
[11] Savioja L, Svensson UP. Overview of geometrical room acoustic modeling techniques. J Acoust Soc Am 2015;138:708–30. https://doi.org/10.1121/1.4926438
[12] Naylor GM. Odeon—another hybrid room acoustical model. Appl Acoust 1993;38:131–43. https://doi.org/10.1016/0003-682X(93)90047-A
[13] Wittebol W, Wang H, Hornikx M, Calamia P. A hybrid room acoustic modeling approach combining image source, acoustic diffusion equation, and time-domain discontinuous Galerkin methods. Appl Acoust 2024;223:110068. https://doi.org/10.1016/j.apacoust.2024.110068
[14] Lynge Christensen C, Koutsouris G, Gi J. ODEON room acoustics software user manual, Technical Report, Odeon A/S; 2023. www.odeon.dk.
[15] Marbjerg G, Brunskog J, Jeong C-H, Nilsson E. Development and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms. J Acoust Soc Am 2015;138:1457–68. https://doi.org/10.1121/1.4928297
[16] Moorer JA. About this reverberation business. Comput Music J 1979;3:13. https://doi.org/10.2307/3680280
[17] Carroll MM, Chien CF. Decay of reverberant sound in a spherical enclosure. J Acoust Soc Am 1977;62:1442–6. https://doi.org/10.1121/1.381662
[18] Miles RN. Sound field in a rectangular enclosure with diffusely reflecting boundaries. J Sound Vib 1984;92:203–26. https://doi.org/10.1016/0022-460X(84)90556-X
[19] Lewers T. A combined beam tracing and radiant exchange computer model of room acoustics. Appl Acoust 1993;38:161–78. https://doi.org/10.1016/0003-682X(93)90049-C
[20] Nosal E-M, Hodgson M, Ashdown I. Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms. J Acoust Soc Am 2004;116:970–80. https://doi.org/10.1121/1.1772400
[21] Koutsouris GI, Brunskog J, Jeong C-H, Jacobsen F. Combination of acoustical radiosity and the image source method. J Acoust Soc Am 2013;133:3963–74. https://doi.org/10.1121/1.4802897

[22] Morse PM, Feshbach H. Methods of theoretical Physics. McGraw-Hill; 1953.
[23] Beranek LL, Mellow TJ. Radiation and scattering of sound by the boundary integral method, in: Acoustics: Sound Fields and Transducers, 2012:535–631. https://doi.org/10.1016/b978-0-12-391421-7.00013-0
[24] Van Maercke D, Defrance J. Development of an analytical model for outdoor sound propagation within the harmonoise project. Acta Acust united Acust 2007;93:201–12. https://api.semanticscholar.org/CorpusID:113437015.
[25] Waaranperä M, Forssén J. Time-domain model for spherical wave reflection in a flat surface with absorber character – application to the SOPRA measurement method. Appl Acoust 2025;227:110251. https://doi.org/10.1016/J.APACOUST.2024.110251
[26] EDtoolbox. Edge diffraction matlab toolbox software. https://github.com/upsvensson/Edge-diffraction-Matlab-toolbox [8 January 2025].
[27] Pierce AD. Acoustics: an introduction to its physical principles and applications. 3rd ed. 2019. https://doi.org/10.1007/978-3-030-11214-1.