
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Götz, Georg; Schlecht, Sebastian J.; Pulkki, Ville

Common-slope modeling of late reverberation in coupled rooms

Published in:

Proceedings of the 24th International Congress on Acoustics: A12: Room Acoustics

Published: 01/01/2022

Document Version

Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:

Götz, G., Schlecht, S. J., & Pulkki, V. (2022). Common-slope modeling of late reverberation in coupled rooms. In *Proceedings of the 24th International Congress on Acoustics: A12: Room Acoustics* (pp. 1-12). (Proceedings of the ICA congress). Acoustical Society of Korea (ASK). https://ica2022korea.org/data/Proceedings_A12.pdf

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

ABS-0161

Common-slope modeling of late reverberation in coupled rooms

Georg GÖTZ^(1*), Sebastian J. SCHLECHT^(1,2), Ville PULKKI⁽¹⁾

⁽¹⁾Aalto Acoustics Lab, Department of Signal Processing and Acoustics, Aalto University, Finland

⁽²⁾Media Lab, Department of Art and Media, Aalto University, Finland

^(*)Correspondence: georg.gotz@aalto.fi

ABSTRACT

Coupled rooms have a distinct sound energy decay behavior, which exhibits more than one decay time under certain conditions. The sound energy decay analysis in such scenarios requires decay models consisting of multiple exponentials with distinct decay rates and amplitudes. While multi-exponential decay analysis is commonly used in room acoustics, the spatial and directional sound energy decay variations in coupled rooms have received little attention. In this work, we introduce the common-slope model of late reverberation for coupled rooms. Common slopes are spatially and directionally invariant decay functions over time, whose amplitudes model all decay variations with respect to the source-receiver configuration. For example, in a scene consisting of two coupled rooms, it is possible to determine two common decay times that approximate the decay for all source-receiver configurations in the scene. Consequently, all spatial and directional decay variations are expressed via decay amplitudes only. We apply the common-slope analysis to measurements of room transitions between coupled rooms. Our analysis shows that the common-slope model approximates the measured sound energy decay with little error. The proposed common-slope model can be used for room acoustic analysis and the efficient synthesis of artificial late reverberation tails.

Keywords: late reverberation model, sound energy decay, coupled rooms, multi-exponential decay

1 INTRODUCTION

The sound energy decay of coupled rooms has received much attention in room acoustic literature [1–5]. It is well documented in those studies that, under certain conditions, the sound energy in coupled rooms decays with more than one decay rate. Multi-exponential models are commonly used to analyze such multi-rate energy decays in terms of decay times and amplitudes [2]. For this purpose, different approaches to determine the model parameters have been proposed [6, 7].

The diffuse field assumption is the foundation for many room acoustic studies and theories. In a diffuse sound field, the energy is assumed to be spatially and directionally uniformly distributed. In more complicated room geometries, such as coupled rooms or rooms with non-uniform absorption distribution, this assumption may be violated. More recently, there has been a growing interest in analyzing spatial and directional sound field variations with respect to sound energy decays [8–14, 3, 15, 5, 16].

This paper builds upon the previously cited studies on spatial and directional sound energy decay analysis, and combines them with some ideas from the common-acoustical-pole and residue (CAPR) model by Haneda et al. [17]. We introduce the common-slope model of late reverberation, which assumes a fixed set of common decay times for multiple source-receiver configurations. Consequently, all spatial and directional sound energy decay variations are described only with decay and noise amplitudes. This model yields highly interpretable room acoustic analysis results, because the degrees of freedom (i.e., the number of model parameters) are approximately halved. Furthermore, the compactness of the common-slope model may be beneficial for all acoustic applications relying on sound energy decay models, such as echo cancellation, source separation, dereverberation, sound field equalization, and parametric spatial audio rendering.

The remainder of this paper is structured as follows. Section 2 provides an overview of the acoustic fundamentals of this work. Section 3 introduces the common-slope model of reverberation and Section 4 demonstrates its usage on a large number of acoustic measurements conducted in coupled rooms. Section 5 concludes this work.

More details, background, and a thorough derivation of the common-slope model can be found in a recently submitted journal article by the authors [18]. This paper extends the journal paper by demonstrating the common-slope analysis on further transitions between coupled rooms, including various sound source positions.

2 BACKGROUND

The time-domain transfer function for sound traveling between a sound source at position $\mathbf{x}_s = (x_s, y_s, z_s)$ and a receiver at position $\mathbf{x}_r = (x_r, y_r, z_r)$ can be given in terms of a room impulse response (RIR). When dealing with directional sound sources or receivers, the propagation path's direction of departure (DOD) from the source and direction of arrival (DOA) at the receiver must be considered. They are denoted by $\mathbf{\Omega}_s = (\phi_s, \theta_s)$ and $\mathbf{\Omega}_r = (\phi_r, \theta_r)$, respectively, with the azimuth angle ϕ and the elevation angle θ . In this paper, we define the *source-receiver configuration* $\mathbf{x} = (\mathbf{x}_s, \mathbf{x}_r, \mathbf{\Omega}_s, \mathbf{\Omega}_r)$ as the combination of source and receiver position and the propagation directions.

The sound energy decay of rooms is commonly described in terms of energy decay functions (EDFs). To this end, the Schroeder backwards integration procedure [19] can be used to calculate an unnormalized EDF $d(\mathbf{x}, t)$ from an RIR $h(\mathbf{x}, t)$ as

$$d(\mathbf{x}, t) = \sum_{l=t}^L h^2(\mathbf{x}, l), \quad (1)$$

where L is the number of samples in the EDF. We assume discrete time throughout this paper, i.e., t is the discrete-time sample index.

Coupled rooms have a distinct sound energy decay behavior, which exhibits more than one decay rate under certain conditions [1, 20]. In such scenarios, the sound energy decay is typically modeled as a superposition of multiple exponentials with individual decay rates and amplitudes [2]. The multi-exponential model with noise is given by [2]

$$d_{\kappa}^{(\text{tr.})}(\mathbf{x}, t) = N_{0,\mathbf{x}} \Psi_{0,\mathbf{x}}^{(\text{tr.})}(t) + \sum_{k=1}^{\kappa} A_{k,\mathbf{x}} [\Psi_{k,\mathbf{x}}^{(\text{tr.})}(t) - \Psi_{k,\mathbf{x}}^{(\text{tr.})}(L)], \quad (2)$$

with the decay kernel

$$\Psi_{k,\mathbf{x}}^{(\text{tr.})}(t) = \begin{cases} L-t, & \text{if } k=0 \\ \exp\left(\frac{-13.8}{f_s} \frac{t}{T_{k,\mathbf{x}}}\right), & \text{if } k>0 \end{cases}. \quad (3)$$

In the above model, $T_{k,\mathbf{x}}$ and $A_{k,\mathbf{x}}$ are the decay times and amplitudes of the k th exponential, respectively, $N_{0,\mathbf{x}}$ is the amplitude of the noise term, f_s is the sampling frequency, and $-13.8 = \ln(10^{-6})$ is a constant ensuring that the sound energy has decayed by 60 dB after $T_{k,\mathbf{x}}$ seconds, where $\ln(\cdot)$ denotes the natural logarithm. The square brackets in Eq. (2) include the constant term $\Psi_{k,\mathbf{x}}^{(\text{tr.})}(L)$, which accounts for the finite upper limit of integration in the Schroeder backwards integration and can be dropped for large L [21].

We will refer to this model throughout the paper as the *traditional multi-exponential model*. It is evident from Eqs. (2) and (3) that the κ th order traditional multi-exponential model features $(2\kappa + 1)$ free parameters that have to be determined for every source-receiver configuration \mathbf{x} , namely, κ decay times, κ decay amplitudes, and 1 noise amplitude. Standard decay analysis methods like the DecayFitNet [6] or Bayesian analysis [7] can be used for this purpose.

Position- and direction-dependent sound energy decay analysis was extensively researched during the previous years [8–14, 3, 15, 5, 16]. Among these studies, some have investigated variations of ISO 3382 parameters [22, 23] like reverberation time, early decay time, or clarity [8–14, 3], while others are based on multi-exponential models [15, 5] or entire EDFs [16]. The studies can be divided into investigations of spatial [8–10, 13–15, 3] or directional [11, 12] sound energy decay variations, and some of them deal specifically with coupled rooms [3, 15, 5]. In summary, the aforementioned studies found directional and spatial sound energy

decay variations of various magnitudes, and they make clear that such variations should be considered in room acoustic modeling.

3 COMMON-SLOPE MODEL OF REVERBERATION

This section introduces the common-slope model of reverberation and outlines how it can be used to describe sound energy decay variations in rooms.

As elaborated before, EDFs can be modeled as a linear combination of one or more exponentials and a noise term. The analysis is usually carried out in frequency bands under the assumption that a specific frequency band features only a limited number of decay rates. Earlier work by Kuttruff investigates the distribution of room modes over the frequency range and establishes their relationship to EDFs [20, 24]. More precisely, when the room modes within a frequency band have similar decay times, the resulting EDF follows an exponential, whose decay time is the average of the individual mode decay times [20]. In coupled rooms, multiple mode groups with decay times scattered around distinct means may occur. This phenomenon motivates multi-exponential decay models, in which the number of slopes corresponds to the number of mode groups.

The relationship between room modes and energy decay slopes is a central part of this work. It is particularly useful to recognize that the mode decay times only depend on the room geometry and wall properties [20]. In other words, for varying source-receiver configurations, the mode decay times stay constant, whereas all acoustic changes can be modeled with mode decay amplitudes [20, 17]. In the following, we will refer to this property as the *common-decay property (CDP)*. A similar line-of-thought can be found in previous work by Haneda et al. [17], which use the CDP in their common-acoustical-pole and residue (CAPR) model to interpolate between RIRs. Another study relying on the CDP was presented by Das et al. [25], who extend the CAPR toward frequency-band-wise processing. In this paper, we will utilize the CDP and combine it with the previously elaborated insight that room modes and EDFs are closely connected.

We propose the common-slope model of reverberation, which is given by

$$d_{\kappa}(\mathbf{x}, t) = N_{0,\mathbf{x}} \Psi_0(t) + \sum_{k=1}^{\kappa} A_{k,\mathbf{x}} [\Psi_k(t) - \Psi_k(L)], \quad (4)$$

with the decay kernel

$$\Psi_k(t) = \begin{cases} L-t, & \text{if } k=0 \\ \exp\left(\frac{-13.8}{f_s} \frac{t}{T_k}\right), & \text{if } k>0 \end{cases}. \quad (5)$$

At first sight, this model considerably resembles the traditional multi-exponential decay model in Eqs. (2) and (3). However, it is important to note that the decay kernel Ψ_k is now independent of the source-receiver-configuration \mathbf{x} , as suggested by the CDP. More precisely, the common-slope model assumes that EDFs of multiple source-receiver-configurations can be modeled with a common set of exponential decay times T_k . We therefore refer to the common T_k values as *common decay times* and to the corresponding EDF slopes as *common slopes*. Consequently, all variations with respect to the source-receiver-configuration are described in terms of the amplitude values $A_{k,\mathbf{x}}$ and noise values $N_{0,\mathbf{x}}$ only. Please note that the common-slope model can only be applied if the different source-receiver-configurations are part of the same scene. For example, a coupled room geometry consisting of multiple connected rooms counts as one scene, but multiple non-connected rooms in different buildings do not.

The common decay times can be obtained in three steps. Firstly, a standard decay analysis approach like the DecayFitNet [6] or Bayesian analysis [7] is used to determine the decay times $T_{k,\mathbf{x}}$ of the traditional multi-exponential model [c.f. Eqs. (2) and (3)]. Secondly, the k-means clustering algorithm [26, 27] is applied to the decay times $T_{k,\mathbf{x}}$ to obtain κ decay time clusters corresponding to the κ assumed mode groups. Each $T_{k,\mathbf{x}}$ value is assigned to a cluster, such that it has the smallest absolute distance to the cluster mean. The number of clusters can be visually determined from a histogram of the $T_{k,\mathbf{x}}$ values. Lastly, one common decay time is determined for each cluster. The common decay time is defined as the center of the histogram bin containing the largest number of $T_{k,\mathbf{x}}$ values. Figure 1 demonstrates this procedure for all three transitions analyzed in this paper (c.f. Section 4.1).

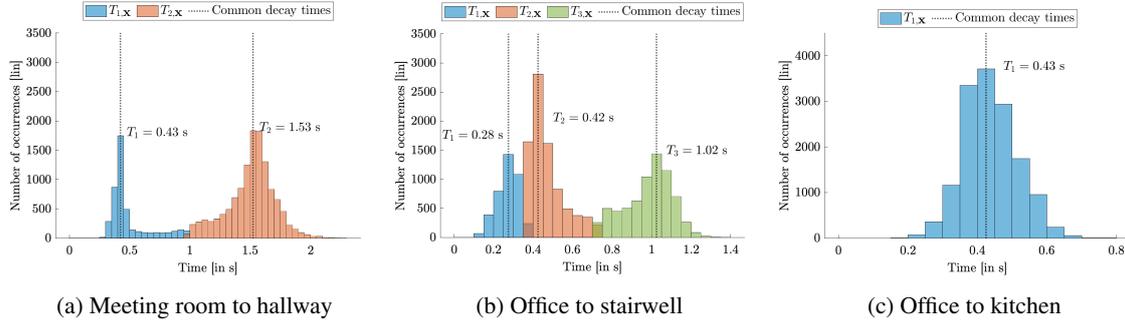


Figure 1. K-means clustering of decay times $T_{k,\mathbf{x}}$ to obtain the common decay times T_k . The subplots demonstrate the clustering on all three transitions that are analyzed in this paper (c.f. Section 4.1). Each subplot is based on 9696 source-receiver configurations (101 receiver positions \times 4 source positions \times 24 beamformer directions, c.f. Section 4.2 for more details).

After the common decay times T_k have been determined, the remaining model parameters $A_{k,\mathbf{x}}$ and $N_{0,\mathbf{x}}$ need to be estimated for all source-receiver configurations \mathbf{x} . Due to the common decay times, all non-linearities of the traditional multi-exponential model have been eliminated. Consequently, the remaining estimation problem becomes solvable as a constrained linear least-squares problem with $A_{k,\mathbf{x}} \geq 0$ and $N_{0,\mathbf{x}} \geq 0$.

4 RESULTS

In the following section, we will demonstrate the common-slope model and show how it can be used to obtain interpretable room acoustic analysis results. We will focus specifically on coupled room geometries.

4.1 Dataset under investigation

The analyzes in this section will be based on the Room Transition dataset by McKenzie et al. [28, 29]. The Room Transition dataset contains higher-order Ambisonics RIRs, which were measured with an Eigenmike microphone array. During the measurements, the microphone array was placed in 5 cm intervals on 5 m long, straight transition lines centered around the aperture between the coupled rooms. Each transition was measured with four different source positions: two in each room, one of which has clear line-of-sight to all receiver positions (CLOS), while the other one has no clear line-of-sight (NLOS). Consequently, there are $101 \times 4 = 404$ RIRs per transition, corresponding to 404 different source-receiver configurations \mathbf{x} .

In this paper, we will demonstrate the common-slope model on the transitions “Meeting room to hallway”, “Office to stairwell”, and “Office to kitchen”. Table 1 briefly summarizes the dimensions and properties of the individual rooms. For more information, please refer to the dataset and its accompanying publication, which also include the floor plans of all scenes [28, 29].

4.2 Room transition along a straight line: spatial and directional analysis

In this section, we extend the preceding analysis with directional information. To this end, we use the directional information captured by the higher-order Ambisonic RIRs and perform beamforming into different directions. In the present analysis, we beamform with a 15° azimuth resolution into directions on the horizontal plane. After beamforming into a certain direction, EDFs can be calculated from the directional RIRs via the Schroeder backwards integration procedure [19] to yield directional EDFs (DEDFs). This procedure is analogous to the methodology in prior work [16, 30].

We obtain beamformer output signals $\mathbf{S} \in \mathbb{R}^{J \times L}$ for J analysis directions as

$$\mathbf{S} = \mathbf{A}\mathbf{h}, \quad (6a)$$

$$\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_J]^T, \quad (6b)$$

where \mathbf{s}_ξ is the ξ th beamformer output, $\mathbf{A} \in \mathbb{R}^{J \times (N+1)^2}$ and $\mathbf{h} \in \mathbb{R}^{(N+1)^2 \times L}$ denote the analysis matrix and the N th-order Ambisonic RIR¹, respectively, and $(\cdot)^T$ denotes the matrix transpose. We assume axisymmetric beamformers, i.e., the analysis matrix \mathbf{A} can be obtained from the beamformer weights $\mathbf{c}_N \in \mathbb{R}^{N \times 1}$ as

$$\mathbf{A} = \mathbf{Y} \text{diag}_N(\mathbf{c}_N), \quad (7a)$$

$$\mathbf{Y} = [\mathbf{y}(\boldsymbol{\Omega}_1), \mathbf{y}(\boldsymbol{\Omega}_2), \dots, \mathbf{y}(\boldsymbol{\Omega}_J)]^T, \quad (7b)$$

$$\mathbf{c}_N = [c_0, c_1, \dots, c_N]^T, \quad (7c)$$

where $\mathbf{y}(\boldsymbol{\Omega}_\xi) \in \mathbb{R}^{(N+1)^2 \times 1}$ denote spherical harmonics (SHs) evaluated at the beamformer steering directions $\boldsymbol{\Omega}_\xi$. Due to the axisymmetry of the beamformers, we repeat the v th beamformer weight μ times, with μ and v being the SH degree and order, respectively. This operation is formalized by $\text{diag}_N(\cdot)$.

Previous studies on room acoustic analysis and Ambisonic RIR processing showed that a great front-to-back-separation is important for resolving energy differences along room axes [30, 31]. Therefore, we employ a spatial Butterworth beamformer [32, Table 3.1] in this work, whose axisymmetric weights are given by

$$c_v^{\text{Butterworth}} = \frac{1}{\sqrt{1 + (v/v_c)^{2\gamma}}}, \quad (8)$$

where we set the Butterworth beamformer order $\gamma = 5$, and the cuton SH order $v_c = 3$.

In the actual common-slope analysis, we apply the DecayFitNet [6] on all 9696 DEDFs of a specific transition (101 receiver positions \times 4 source positions \times 24 beamformer directions) to obtain the $T_{k,\mathbf{x}}$ values of the traditional multi-exponential model [c.f. Eqs. (2) and (3)]. Afterwards, we determine the common decay times T_k based on all 9696 $T_{k,\mathbf{x}}$ values as outlined in Section 3. Finally, the decay amplitudes $A_{k,\mathbf{x}}$ and noise amplitudes $N_{0,\mathbf{x}}$ are calculated via a linear least-squares fit of the common-slope model [c.f. Eqs. (4) and (5)] to the measured DEDFs.

Figure 2 shows the common-slope analysis results for the transition ‘‘Meeting room to hallway, source in meeting room, clear line-of-sight’’. Two common decay times were determined (c.f. Figure 1a) and they amount to $T_1 = 0.43$ s and $T_2 = 1.53$ s. Figure 2a illustrates how the $A_{1,\mathbf{x}}$ values change for various positions on the transitions and beamforming directions. The values are generally higher in the meeting room, and gradually fade out when transitioning toward the hallway. Slightly increased amplitudes can be observed near the hallway wall (i.e., $x_r = 500$ cm), which can be attributed to reflections from the wall. Distinct peaks can be observed for $\phi_r = \pm 90^\circ$. The amplitude variations for different look directions indicate that the reverberation is considerably

¹The channel ordering is not explicitly defined here because it is not relevant for this work as long as it is used consistently throughout the analysis pipeline. Please note that the maximum spherical harmonic order N should not be confused with the noise term N_0 of the decay model.

Transition	Room description	Dimensions (w \times l \times h)	Volume
Meeting room to hallway	Acoustically treated meeting room Reverberant hallway	6.6 m \times 4.6 m \times 2.8 m 18 m \times 4.5 m \times 2.8 m	85 m ³ 226.8 m ³
Office to stairwell	Storage-room-like office Stairwell connecting three floors in total	5 m \times 11.7 m \times 3.6 m 6.3 m \times 3.3 m \times 14.4 m	210.6 m ³ 299.4 m ³
Office to kitchen	Acoustically treated office Acoustically treated office kitchen	9.1 m \times 3.5 m \times 2.9 m 13.7 m \times 4.4 m \times 2.9 m	92.4 m ³ 174.8 m ³

Table 1. Summary of all analyzed room transitions. The measurements are part of the Room Transition dataset by McKenzie et al. [28, 29].

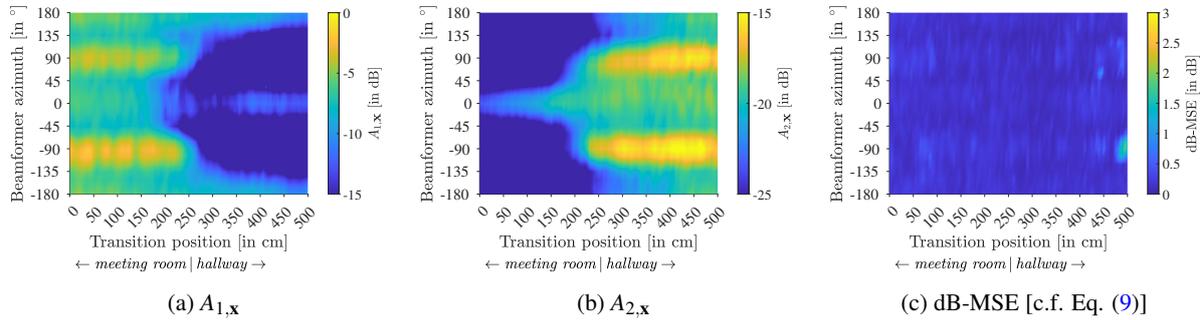


Figure 2. Common-slope analysis results (1 kHz octave band) of the “Meeting room to hallway, source in meeting room, clear line-of-sight” transition. The plots show the a) $A_{1,x}$, b) $A_{2,x}$, and c) dB-MSE [c.f. Eq. (9)] values for a common-slope analysis with the common decay times $T_1 = 0.43$ s and $T_2 = 1.53$ s.

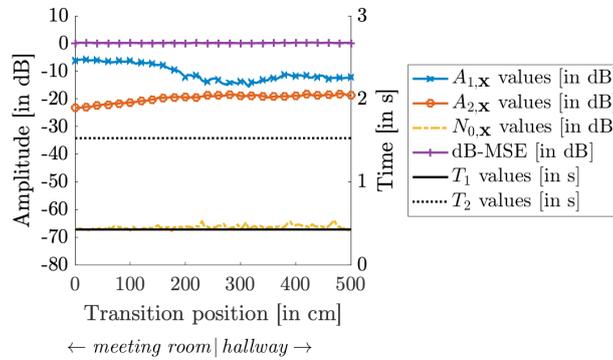


Figure 3. Common-slope analysis results (1 kHz octave band, $\phi_r = 0^\circ$, $\theta_r = 0^\circ$) of the “Meeting room to hallway, source in meeting room, clear line-of-sight” transition. The plots show the $A_{1,x}$, $A_{2,x}$, $N_{0,x}$, and dB-MSE [c.f. Eq. (9)] values for all positions along the transition line. The common-slope analysis is based on the common decay times $T_1 = 0.43$ s and $T_2 = 1.53$ s.

anisotropic, which could be explained by an uneven distribution of absorption material in the room and the energy transfer through the door. The $A_{2,x}$ values corresponding to the common decay time T_2 are depicted in Figure 2b. They gradually fade in while transitioning into the hallway. For positions closer to the door, one can observe how the energy leaks into the meeting room, and that this effect is considerably directional. Just like the $A_{1,x}$ values, the $A_{2,x}$ exhibit anisotropy, where clear peaks can be observed for $\phi_r = \pm 90^\circ$. Figure 2c depicts the dB-MSE between the common-slope fit $d_\kappa^{(\text{dB})}(\mathbf{x}, t)$ and the true DEDFs $d^{(\text{dB})}(\mathbf{x}, t)$, which is defined as

$$\text{dB-MSE} = \frac{1}{L} \sum_{t=1}^L [d_\kappa^{(\text{dB})}(\mathbf{x}, t) - d^{(\text{dB})}(\mathbf{x}, t)]^2, \quad (9)$$

where both DEDFs are represented on a logarithmic scale in dB. We exclude some portions of the EDF because they are not representative for the actual late reverberation decay:

1. The first $\frac{2m}{343 \text{ m/s}} \approx 5.8$ ms, because this part of the EDF only includes the direct sound and possibly one or two reflections, which show up as a very steep energy drop. This part will inevitably introduce an error, because the multi-exponential model cannot properly model this part.
2. The last 5% of the EDF (i.e. $t > 0.95L$), because this part mostly features noise, which exhibits statistical uncertainty. Although the Schroeder backwards integration procedure converges to a steady noise floor after a while, it still requires to integrate a larger number of samples to account for this uncertainty.

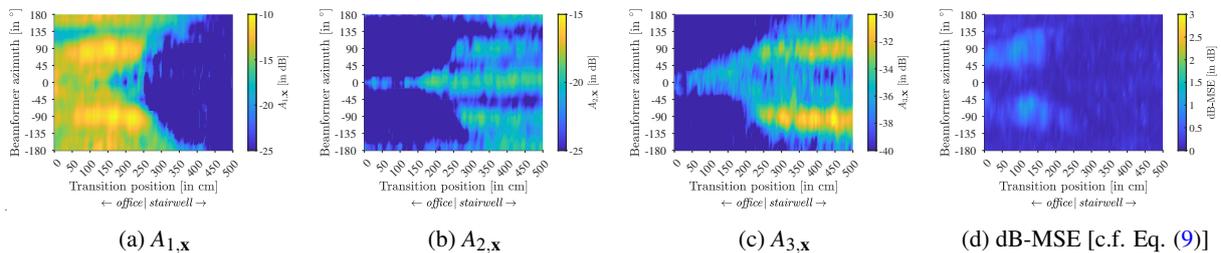


Figure 4. Common-slope analysis results (1 kHz octave band) of the “Office to stairwell, source in office, no line-of-sight” transition. The plots show the a) $A_{1,x}$, b) $A_{2,x}$, c) $A_{3,x}$, and d) dB-MSE [c.f. Eq. (9)] values for a common-slope analysis with the common decay times $T_1 = 0.28$ s, $T_2 = 0.42$ s, and $T_3 = 1.02$ s.

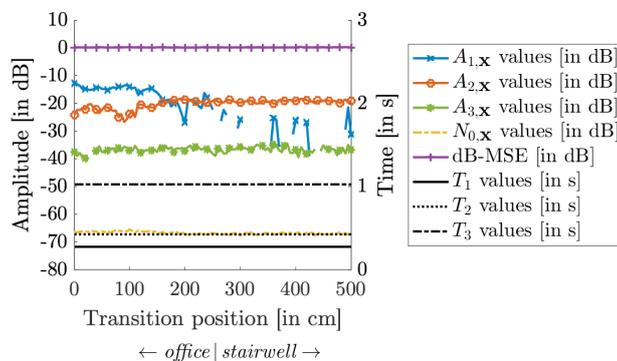


Figure 5. Common-slope analysis results (1 kHz octave band, $\phi_r = 0^\circ$, $\theta_r = 0^\circ$) of the “Office to stairwell, source in office, no line-of-sight” transition. The plots show the $A_{1,x}$, $A_{2,x}$, $A_{3,x}$, $N_{0,x}$, and dB-MSE [c.f. Eq. (9)] values for all positions along the transition line. The common-slope analysis is based on the common decay times $T_1 = 0.28$ s, $T_2 = 0.42$ s, and $T_3 = 1.02$ s.

The dB-MSE is well below 3 dB for all source-receiver configurations, with median and 99 % values of 0.18 dB and 0.51 dB, respectively. The low dB-MSE values indicate that the common-slope model is suitable for describing the energy decay behavior along the entire transition, despite having fewer degrees-of-freedom.

In Figure 3, we focus only on the beamformer direction $\phi_r = 0^\circ$, $\theta_r = 0^\circ$ to highlight the cross-fade between the two different decay processes. It becomes clear from the figure, how the $A_{1,x}$ values gradually fade out, while the $A_{2,x}$ values are getting stronger during the transition into the hallway. Interestingly, the $A_{1,x}$ values are slightly increasing again toward the 500 cm transition position. This observation can be attributed to stronger reflections from the hallway wall. The $N_{0,x}$ values remain approximately constant throughout the transition. Finally, the low dB-MSE values [c.f. Eq. (9)] indicate that the common-slope model is a suitable description for the analyzed transition.

Figure 4 shows the common-slope analysis results for the transition “Office to stairwell, source in office, no line-of-sight”. This transition required three common decay times to describe all source-receiver configurations (c.f. Figure 1b), and they amount to $T_1 = 0.28$ s, $T_2 = 0.42$ s, and $T_3 = 1.02$ s, respectively. Just as in the previously described transition, the $A_{1,x}$ values gradually fade out while transitioning into the second room, whereas the $A_{2,x}$ and $A_{3,x}$ values are getting stronger. Distinct peaks at lateral directions highlight the anisotropy of the sound energy decay. Furthermore, the low dB-MSE median and 99 % quantile values of 0.13 dB and 1.08 dB, respectively, demonstrate that only little error between true DEDFs and common-slope model can be observed. This result indicates that the common-slope model is suitable for describing the analyzed room transition.

Figure 5 focuses on the analysis results for the beamformer direction $\phi_r = 0^\circ$, $\theta_r = 0^\circ$. It becomes more

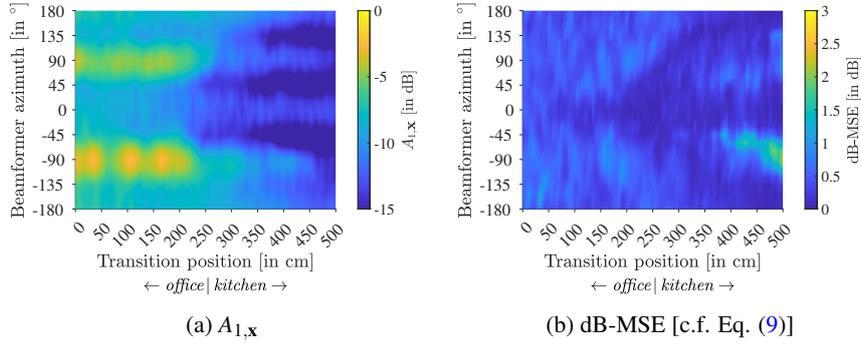


Figure 6. Common-slope analysis results (1 kHz octave band) of the “Office to kitchen, source in office, no line-of-sight” transition. The plots show the a) $A_{1,x}$ and b) dB-MSE [c.f. Eq. (9)] values for a common-slope analysis with the common decay time $T_1 = 0.43$ s.

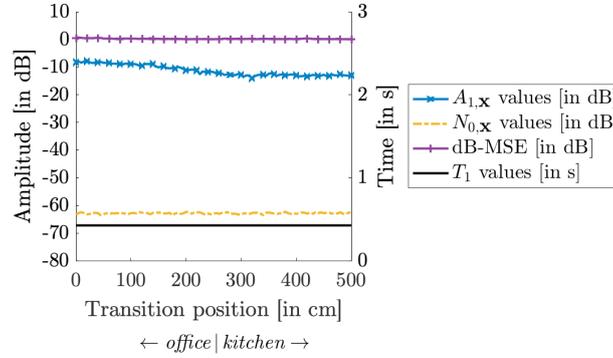


Figure 7. Common-slope analysis results (1 kHz octave band, $\phi_r = 0^\circ$, $\theta_r = 0^\circ$) of the “Office to kitchen, source in office, no line-of-sight” transition. The plots show the $A_{1,x}$, $N_{0,x}$, and dB-MSE [c.f. Eq. (9)] values for all positions along the transition line. The common-slope analysis is based on the common decay time $T_1 = 0.43$ s.

evident from this figure, how the $A_{1,x}$ values gradually fade out during the transition into the stairwell, while the $A_{2,x}$ and $A_{3,x}$ values are getting stronger. For positions in the stairwell, the slope of the shorter decay time T_1 is sometimes masked by the slower decaying slopes with decay times T_2 and T_3 . In these cases, the T_1 slope may be hard to detect, thus resulting in near-zero amplitude values $A_{1,x}$. In Figure 5, we therefore excluded all $A_{1,x}$ values below 40 dB to make the plot more understandable. The noise level $N_{0,x}$ is approximately constant along the transition. Lastly, the dB-MSE values [c.f. Eq. (9)] are low for all positions, thus indicating that the common-slope model is a suitable model for describing this transition.

Figure 6 depicts the common-slope analysis results for the transition “Office to kitchen, source in office, no line-of-sight”. For this transition, only one common decay time could be determined with the k-means method (c.f. Figure 1c) and it amounts to $T_1 = 0.43$ s. The acoustic properties of both rooms are very similar, and consequently only one decay rate could be determined for all source-receiver configurations. The $A_{1,x}$ values once again show considerable anisotropy, and they gradually fade out while transitioning into the room without the sound source. Slightly increased errors can be observed for this transition, with median and 99 % quantile values of 0.39 dB and 1.45 dB, respectively. The histogram in Figure 1c shows that the decay times $T_{k,x}$ vary between 0.3 s and 0.6 s. Such a small decay time variability could be easily accommodated for the previous two transitions, because the multi-exponential model is very versatile and can compensate for slightly wrong decay times by adapting the amplitudes accordingly. This has already been found by Lanczos, who states that the decomposition of a decay function into a linear combination of exponentials is a highly ill-conditioned

problem [33]. However, with only one active decay time or slope, the margin for compensation is considerably reduced. Consequently, slightly higher errors are observed for this transition.

	dB-MSE [c.f. Eq. (9)]					
	Common-slope model			Traditional multi-exponential model		
	mean	median	99 % q.	mean	median	99 % q.
Meeting room to hallway						
Source in meeting room, NLOS	0.15 dB	0.13 dB	0.38 dB	0.16 dB	0.14 dB	0.40 dB
Source in meeting room, CLOS	0.18 dB	0.16 dB	0.51 dB	0.19 dB	0.17 dB	0.45 dB
Source in hallway, NLOS	0.38 dB	0.31 dB	1.44 dB	0.23 dB	0.19 dB	0.88 dB
Source in hallway, CLOS	0.26 dB	0.22 dB	1.12 dB	0.27 dB	0.22 dB	1.04 dB
Office to stairwell						
Source in office, NLOS	0.21 dB	0.13 dB	0.94 dB	0.18 dB	0.16 dB	0.53 dB
Source in office, CLOS	0.18 dB	0.16 dB	0.42 dB	0.21 dB	0.20 dB	0.51 dB
Source in stairwell, NLOS	0.15 dB	0.13 dB	0.51 dB	0.16 dB	0.14 dB	0.51 dB
Source in stairwell, CLOS	0.21 dB	0.17 dB	0.85 dB	0.17 dB	0.15 dB	0.51 dB
Office to kitchen						
Source in office, NLOS	0.39 dB	0.36 dB	1.45 dB	0.17 dB	0.16 dB	0.41 dB
Source in office, CLOS	0.46 dB	0.40 dB	1.25 dB	0.21 dB	0.18 dB	0.61 dB
Source in kitchen, NLOS	0.18 dB	0.17 dB	0.40 dB	0.16 dB	0.14 dB	0.38 dB
Source in kitchen, CLOS	0.19 dB	0.17 dB	0.40 dB	0.17 dB	0.15 dB	0.52 dB

Table 2. Decibel-based mean squared error [dB-MSE, c.f. Eq. (9)] between directional energy decay functions of the room transition dataset [28, 29], the common-slope model [c.f. Eqs. (4) and (5)] and the traditional multi-exponential model [c.f. Eqs. (2) and (3)]. The table is based on the spatial and directional analysis, where directional information is obtained via beamforming. A perfect fit would yield a dB-MSE value of 0 dB. We refer to the different analyzed source positions as NLOS = no line-of-sight, and CLOS = clear line-of-sight.

Figure 7 shows only the analysis results for the beamformer direction $\phi_r = 0^\circ$, $\theta_r = 0^\circ$. It demonstrates how the $A_{1,x}$ values are gradually getting smaller while transitioning into the kitchen. In contrast, the noise amplitudes $N_{0,x}$ remain approximately constant throughout the transition. The dB-MSE [c.f. Eq. (9)] remains low for the entire transition, thus indicating that the common-slope model can accurately fit the sound energy decay of this scene.

Lastly, Table 2 compares the fitting performance of the common-slope model [c.f. Eqs. (4) and (5)] and the traditional multi-exponential model [c.f. Eqs. (2) and (3)]. It features the mean, median, and 99 % quantile dB-MSE values [c.f. Eq. (9)] of all transitions and source positions. The mean and median values of the common-slope model lie between 0.13 dB and 0.46 dB, and the 99 % quantile values are below 1.5 dB for all evaluated cases, thus indicating that the common-slope model is suitable for describing all evaluated transitions. Furthermore, the fitting performance is comparable to, but sometimes less robust than the traditional multi-exponential model, despite requiring only approximately half of its parameters.

5 CONCLUSIONS

This paper introduced the common-slope model of reverberation, which uses a common set of decay times to model sound energy decay functions (EDFs) of multiple source-receiver configurations within the same environment. Consequently, all position- and direction-dependent EDF variations are described in terms of decay and noise amplitudes only. In the present study, we used the common-slope model to analyze sound energy decays

in coupled rooms. We showed that the common-slope model is suitable for describing multi-exponential EDFs with varying source-receiver configurations, while requiring considerably fewer parameters than the traditional multi-exponential model. By using the same set of decay times to model all source-receiver configurations, the common-slope model yields easily interpretable room acoustic analysis results. For example, we demonstrated that the common-slope model can be used to analyze how the decay behavior gradually changes when transitioning from one room to another through a connecting door. Furthermore, we showed that the common-slope model introduces only little error between the true and modeled EDF.

The common-slope model will benefit future research efforts in room acoustic analysis and modeling. Additionally, it can be used in all acoustic applications relying on sound energy decay or late reverberation models, such as echo cancellation, source separation, dereverberation, sound field equalization, and parametric spatial audio rendering.

ACKNOWLEDGMENTS

This research has received funding from the Aalto University Doctoral School of Electrical Engineering and the Academy of Finland, project no. 317341.

REFERENCES

- [1] C. F. Eyring, "Reverberation time measurements in coupled rooms," *J. Acoust. Soc. Am.*, vol. 3, no. 2, pp. 181–206, 1931.
- [2] N. Xiang and P. M. Goggans, "Evaluation of decay times in coupled spaces: Bayesian parameter estimation," *J. Acoust. Soc. Am.*, vol. 110, no. 3, pp. 1415–1424, 2001.
- [3] A. Billon, V. Valeau, A. Sakout, and J. Picaut, "On the use of a diffusion model for acoustically coupled rooms," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 2043–2054, 2006.
- [4] H. Pu, X. Qiu, and J. Wang, "Different sound decay patterns and energy feedback in coupled volumes," *J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 1972–1980, 2011.
- [5] N. Xiang, J. Escolano, J. M. Navarro, and Y. Jing, "Investigation on the effect of aperture sizes and receiver positions in coupled rooms," *J. Acoust. Soc. Am.*, vol. 133, no. 6, pp. 3975–3985, 2013.
- [6] G. Götz, R. Falcón Pérez, S. J. Schlecht, and V. Pulkki, "Neural network for multi-exponential sound energy decay analysis," *J. Acoust. Soc. Am.*, vol. 152, no. 2, pp. 942–953, 2022.
- [7] T. Jasa and N. Xiang, "Efficient estimation of decay parameters in acoustically coupled-spaces using slice sampling," *J. Acoust. Soc. Am.*, vol. 126, no. 3, pp. 1269–1279, 2009.
- [8] T. Akama, H. Suzuki, and A. Omoto, "Distribution of selected monaural acoustical parameters in concert halls," *Appl. Acoust.*, vol. 71, no. 6, pp. 564–577, 2010.
- [9] X. Pelorson, J.-P. Vian, and J.-D. Polack, "On the variability of room acoustical parameters: Reproducibility and statistical validity," *Applied Acoustics*, vol. 37, no. 3, pp. 175–198, 1992.
- [10] K. Fujii, T. Hotehama, K. Kato, R. Shimokura, Y. Okamoto, Y. Suzumura, and Y. Ando, "Spatial Distribution of Acoustical Parameters in Concert Halls: Comparison of Different Scattered Reflections," *Journal of Temporal Design in Architecture and the Environment*, vol. 4, no. 1, pp. 59–68, 2004.
- [11] M. Berzborn, J. Balint, and M. Vorländer, "On the Estimation of Directional Decay Times in Reverberation Rooms," in *Euronoise 2021*, (Online conference), pp. 1–9, 2021.
- [12] B. Alary and V. Välimäki, "A Method for Capturing and Reproducing Directional Reverberation in Six Degrees of Freedom," in *International Conference on Immersive and 3D Audio (I3DA)*, (Online Conference), pp. 1–8, 2021.

- [13] M. Garai, S. De Cesaris, and D. D’Orazio, “Spatial distribution of monaural acoustical descriptors in historical Italian theatres,” in *International Symposium on Room Acoustics (ISRA)*, (Toronto, Canada), pp. 1–13, 2013.
- [14] M. Horvat, K. Jambrosic, and H. Domitrovic, “The examination of the influence of standing waves on reverberation time measurements in small reverberant rooms,” in *Proc. of Acoustics’08*, (Paris, France), pp. 8465–8470, 2008.
- [15] A. Billon, “On the spatial variations of double-sloped decay in coupled-volume concert halls,” in *International Symposium on Room Acoustics (ISRA)*, (Seville, Spain), pp. 1–5, 2007.
- [16] M. Berzborn and M. Vorländer, “Directional sound field decay analysis in performance spaces,” *Building Acoustics*, pp. 1–15, 2021.
- [17] Y. Haneda, Y. Kaneda, and N. Kitawaki, “Common-acoustical-pole and residue model and its application to spatial interpolation and extrapolation of a room transfer function,” *IEEE Trans. Speech Audio Process.*, vol. 7, no. 6, pp. 709–717, 1999.
- [18] G. Götz, S. J. Schlecht, and V. Pulkki, “Common-slope modeling of late reverberation.” submitted to *IEEE/ACM Trans. Audio, Speech, Language Process.*, Available Online: <http://doi.org/10.36227/techrxiv.20482767>, 2022.
- [19] M. R. Schroeder, “New Method of Measuring Reverberation Time,” *J. Acoust. Soc. Am.*, vol. 37, no. 3, pp. 409–412, 1965.
- [20] H. Kuttruff, *Room Acoustics*. London, UK: Spon Press, 4th ed., 2000.
- [21] N. Xiang, P. M. Goggans, T. Jasa, and M. Kleiner, “Evaluation of decay times in coupled spaces: Reliability analysis of Bayesian decay time estimation,” *J. Acoust. Soc. Am.*, vol. 117, no. 6, pp. 3707–3715, 2005.
- [22] ISO 3382-1, “Acoustics - Measurement of room acoustic parameters - Part 1: Performance Spaces,” standard, International Organization for Standardization (ISO), Geneva, Switzerland, 2009.
- [23] ISO 3382-2, “Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms,” standard, International Organization for Standardization (ISO), Geneva, Switzerland, 2008.
- [24] H. Kuttruff, “Eigenschaften und Auswertung von Nachhallkurven [Characteristics and Analysis of Sound Energy Decay Curves],” *Acta Acustica united with Acustica*, vol. 8, no. Supplement 1, pp. 273–280, 1958.
- [25] O. Das, P. Calamia, and S. V. Amengual Garí, “Room Impulse Response Interpolation from a Sparse Set of Measurements Using a Modal Architecture,” in *Int. Conf. Acoust., Speech, Sig. Proc. (ICASSP)*, (Online Conference), pp. 960–964, 2021.
- [26] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proc. of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, (Berkeley, CA, USA), pp. 281–297, 1967.
- [27] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press, 4th ed., 2005.
- [28] T. McKenzie, S. J. Schlecht, and V. Pulkki, “A dataset of measured spatial room impulse responses for the transition between coupled rooms (Version 1.2),” Feb. 2021. Zenodo, Available online: <https://doi.org/10.5281/zenodo.4636068>.
- [29] T. McKenzie, S. J. Schlecht, and V. Pulkki, “Acoustic Analysis and Dataset of Transitions Between Coupled Rooms,” in *Int. Conf. Acoust., Speech, Sig. Proc. (ICASSP)*, (Online conference), pp. 481–485, 2021.

- [30] G. Götz, C. Hold, T. McKenzie, S. J. Schlecht, and V. Pulkki, “Analysis of multi-exponential and anisotropic sound energy decay,” in *Proc. of the 48th Jahrestagung für Akustik (DAGA)*, (Stuttgart, Germany), pp. 1055–1058, 2022.
- [31] C. Hold, T. McKenzie, G. Götz, S. J. Schlecht, and V. Pulkki, “Resynthesis of Spatial Room Impulse Response Tails With Anisotropic Multi-Slope Decays,” *J. Audio Eng. Soc.*, vol. 70, no. 6, pp. 526–538, 2022.
- [32] B. Devaraju, *Understanding filtering on the sphere*. PhD thesis, University of Stuttgart, 2015.
- [33] C. Lanczos, *Applied Analysis*. New York, US: Dover Publications, reprint ed., 1988.