DOI: 10.61782/fa.2025.0830

FORUM ACUSTICUM EURONOISE 2025

STOCHASTIC VARIATIONAL INFERENCE OF DIRECTIONAL DECAY TIMES IN A REVERBERATION ROOM

Marco Berzborn and Michael Vorländer
Institute for Hearing Technology and Acoustics, RWTH Aachen University, Germany

ABSTRACT

In rooms with non-uniform distribution of absorption, the sound field is generally subject to multi-exponential energy decay. This is especially prominent if absorption is concentrated on a single surface, such as during the measurement of Sabine's absorption coefficient in reverberation rooms. Accordingly, the sound field is subject to directionally dependent damping. In this work, an experimental inference method to identify the directional damping distributions from microphone array measurements is presented. The method is based on a stochastic directional energy decay model. The model hinges on the assumption that the decay process is sufficiently described by a small number of damping constants, assumed to be representative of groups of modes with shared damping properties. A Bayesian variational inference approach is used to jointly infer the model parameters - namely the damping constants and the directional energy distributions of each group of modes - from experimental data.

Keywords: reverberation, directional energy decay, microphone array

1. INTRODUCTION

The measurement of Sabine's absorption coefficient according to ISO-354 [1] is performed in a reveberation room. Fundamental requirements for Sabine's theory of reverberation are the presence of a diffuse sound field and uniform distribution of absorption. The latter requirement is typically not easy to fulfill as the absorbing specimen is usually concentrated on the floor area of the room. Such non-uniform distribution of absorption due to the presence of the absorbing specimen results in a multi-exponential decay with distinct directional properties. This was experimentally quantified by Balint *et al.* who used a Bayesian framework to estimate the decay constants from omnidirectional Energy Decay Curves (EDCs) measured in a reverberation room [2]. Similarly, Xiang *et al.* estimated the decay parameters for a multi-exponential decay model based on the omnidirectional EDC measured in a coupled room scenario [3]. A neural network based approach was proposed by Götz *et al.* [4] and extended to the Common Slope Model (CSM) in Ref. [5] to investigate the multi-exponential decay behavior in coupled rooms.

In this paper, an inference method to identify the directional damping distributions from the experimental Directional Energy Decay Curve (DEDC) captured using a spherical microphone array is proposed. Inspired by the stochastic energy decay model by Kuttruff [6], a directionally dependent model using the Laplace transform of the damping distribution is presented. A Bayesian variational inference approach is used to jointly infer the model parameters, namely the damping constants and the directional energy distributions for groups of modes. The model and parameter estimation methods are presented in Sections 3 and 4. The experimental and inference setups are introduced in Sections 5 and 6, respectively. Sections 7 and 8 present the results and conclusions of the present study.

2. THE DIRECTIONAL ENERGY DECAY CURVE

Assuming that the modes constituting the sound field within a room are well represented as propagating plane waves, we decompose the sound field into a basis of plane waves using a spherical microphone array [7]. The resulting spatio-temporal density function describing the sound

**Corresponding author*: *mbe@akustik.rwth-aachen.de.*

field is given as the vector

$$\mathbf{g}(t) = [g(t, \mathbf{\Omega}_1), \ldots, g(t, \mathbf{\Omega}_q)]^\mathsf{T}, \tag{1}$$

where $\mathbf{\Omega}_q$ contains the co-latitude and azimuth angles (θ_q, ϕ_q) for the q-th steering direction of the plane wave decomposition beamformer. Using the Schroeder integral [8], the DEDC is calculated as [9] which describes the energy decay of the sound field with respect to time and steering direction

$$\mathbf{d}(t) = \int_t^{t_\chi} |\mathbf{g}(\tau)|^2 \, \mathrm{d}\tau, \tag{2}$$

where t_χ is the time of truncation, which is typically chosen as the time where the Directional Room Impulse Response (DRIR) and the noise floor intersect. For a more compact notation in discrete time, the DEDC is represented as a matrix $\mathbf{D} \in \mathbb{R}^{Q\times T}$, where T is the number of discrete time samples and Q is the number of steering directions.

3. THE DIRECTIONAL COMMON DAMPING MODEL

Close to and above the Schroeder where a sufficient number of modes are excited, the energy decay of the sound field can be represented using stochastic models [10]. Kuttruff showed that the energy decay function can be calculated using the Laplace transform of the energy weighted discrete density distribution of modal damping constants [6]. Extending the model by Kuttruff to account for the directional properties of the damping distributions, a stochastic model representing the energy decay captured in Eq. (2) can be written as

$$\mathbf{M} = \mathbf{SHE}, \tag{3}$$

where $\mathbf{H} \in \mathbb{R}^{V\times K}$ is a matrix representing the histogram for V directions $\mathbf{\Theta}_v$ and K damping constants $\boldsymbol{\delta}_k$, $\mathbf{E} \in \mathbb{R}^{K\times T}$ is a matrix containing the Laplace transform kernel for the respective damping constants, and $\mathbf{S} \in \mathbb{R}^{Q\times V}$ is a filter matrix representing the energy response of the beamformer for all steering directions $\mathbf{\Omega}_q$ and directions of propagation represented in the histogram $\mathbf{\Theta}_v$.

For sufficiently narrow frequency bands, the histograms in Eq. (3) are usually compact and in the case of non-uniform distribution of absorption multi-modal [6]. It can be assumed that these compact histograms are well represented by their respective mean damping constants. As a result the number of candidate damping constants for inverse estimation can be chosen much lower than the number of modes constituting the sound field. This is equivalent to grouping a large number of modes with respect to their shared damping properties, while not imposing further assumptions on their directional energetic properties. This approach is inspired by the multi-exponential decay model by Xiang *et al.* [3] and is similar to the recently developed CSM by Götz *et al.* [11].

While these simplifications reduce the number of parameters to be estimated to fewer than the number of observations, the problem is still severely ill-conditioned for two reasons: Firstly, the entries of both model matrices are unknown, resulting in a non-linear system of equations. Secondly, including the filter matrix $\mathbf{S}$ introduces a directional smoothing of the DEDC, which results in an ill-posed inverse problem. For this reason, the filter matrix is replaced by a modified histogram matrix $\mathbf{A} = \mathbf{SH}$ resulting in

$$\tilde{\mathbf{M}} = \mathbf{AE}, \tag{4}$$

which will be inferred instead of the histogram matrix $\mathbf{H}$. It is important to note here that $\tilde{A} \in \mathbb{R}^{Q\times K}$ now represents a smoothed angular energy distribution relating the damping constants to the steering directions.

Typically, the experimentally observed DEDC has finite length and is subject to additive noise. Here, the truncation is performed for each steering direction individually – at the time $t_{\chi,v}$ (see Eq. (2)) – where the DRIR and the noise floor intersect. The resulting error terms due to noise and temporal truncation can be calculated analogous to the multi-exponential decay model by Xiang *et al.* [3]. Both terms can be represented as additive matrix contributions $\mathbf{R}_{\boldsymbol{\chi}} \in \mathbb{R}^{Q\times T}$ and $\mathbf{R}_\mathbf{n} \in \mathbb{R}^{Q\times T}$ representing the directionally dependent truncation error and the noise terms, respectively. The full model finally reads as

$$\tilde{\mathbf{M}} = \mathbf{AE} - \mathbf{R}_{\boldsymbol{\chi}} + \mathbf{R}_\mathbf{n}. \tag{5}$$

In analogy to the CSM which relies on pre-evaluated and exponential functions with fixed slopes, the model is referred to as the Directional Common Damping Model (DCDM) in the following, as it is based on groups of damping constants which are common to multiple modes. It should be highlighted here that, aside from the estimation of the damping constants jointly with the directional energy distributions, the model formulation is analogous to the CSM proposed by Götz *et al.* [5] when using a microphone array.

4. VARIATIONAL PARAMETER ESTIMATION

The vector of all DCDM parameters

$$\Xi = (a_{1,1}, \ldots, a_{V,1}, \ldots, a_{V,K}, \delta_1, \ldots, \delta_K) \quad (6)$$

is jointly inferred within a Bayesian framework. The posterior distribution $p(\Xi, \sigma|\mathbf{D})$ describing the probability density of the model parameters Ξ and hyper-paramter σ, given the observed DEDC $\mathbf{D}$, is approximated using Automatic Differentiation Variational Inference (ADVI) [12]. The likelihood – which is proportional to the error between the model and the data – is assumed to be normally distributed with mean $\tilde{\mathbf{M}}(\Xi)$ – i.e. the result of Eq. (5) – and standard deviation σ. The standard deviation of the likelihood is jointly estimated with the model parameters, assuming a half-normal prior distribution with zero mean,

$$\sigma \sim \hat{\mathcal{N}}(0, 0.1). \quad (7)$$

The prior distributions for the model parameters are equally chosen half-normal distributions

$$\begin{aligned} \tilde{a}_{v,k} &\sim \hat{\mathcal{N}}(0, 0.1), \\ \tilde{\delta}_k &\sim \hat{\mathcal{N}}(0, 1). \end{aligned} \quad (8)$$

The priors are only weakly informative and purely chosen to regularize the estimation problem analogous to Tikhonov regularization. The introduced bias towards small damping constants and small energy values is designed to avoid over-fitting. In particular on specific reflections during the very early part of the decay process. The limitation to positive values follows the physical constraints of the problem, i.e. positive energy and stable system.

For additional information on stochastic inference, the reader is referred to Gelman *et al.* [13].

5. EXPERIMENTAL SETUP

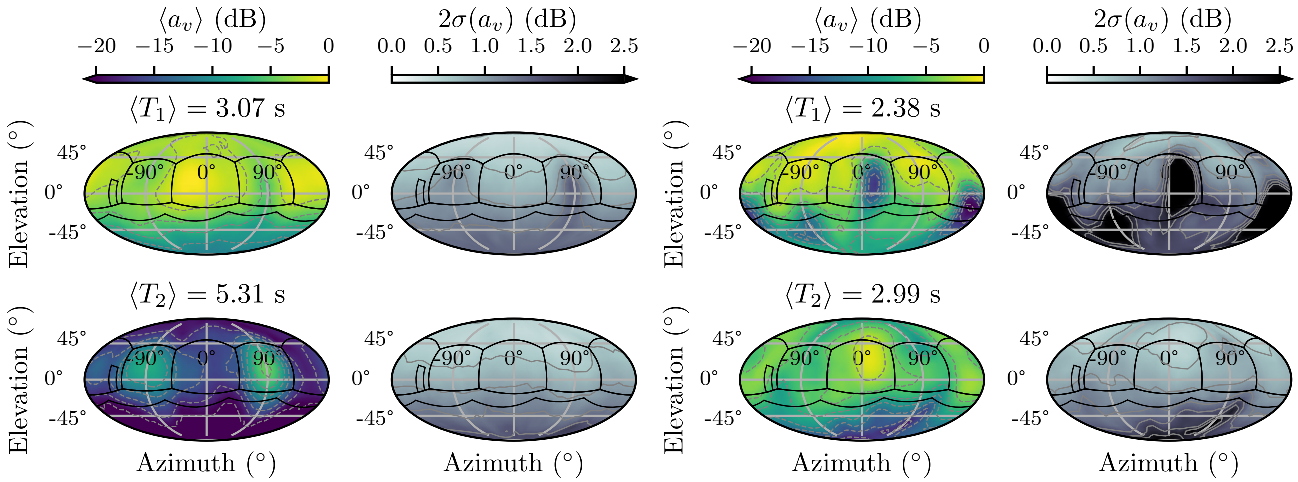
Experimental data was collected in a rectangular reverberation room at the Technical University of Denmark (2800 Kgs. Lyngby, Denmark) in two configurations: The room equipped with and without panel diffusers and including an absorbing sample of glass wool (Ecophon Industry Modus S with a thickness of $100\,\mathrm{mm}$ and a surface area of $10.8\,\mathrm{m}^2$). The room has a volume of $245\,\mathrm{m}^3$ and a Schroeder frequency slightly above $300\,\mathrm{Hz}$.

The two layers of the array with $0.25\,\mathrm{m}$ and $0.45\,\mathrm{m}$ radii consist of 144 sampling positions chosen according to an equal-area grid on the sphere [14]. Additional sampling positions inside the respective spheres were used to stabilize the eigenfrequencies of the spheres [15]. Results are presented for a microphone array position approximately centered above the absorbing specimen.

The DEDCs were calculated for third octave bands from $250\,\mathrm{Hz}$ to $2\,\mathrm{kHz}$, from which the results for the $315\,\mathrm{Hz}$ frequency band are presented here. The plane wave decomposition is performed using a spherical harmonic order of $N = 5$, additionally including Dolph-Chebyshev beamforming weights to achieve a uniform angular resolution. This results in a to main-lobe width of $60°$. A total of 500 steering directions were used, which were uniformly distributed on the sphere using the method presented in Ref. [14]. The truncation times for the Schroeder integration in Eq. (2) were estimated using the algorithm proposed by Lundeby *et al.* [16]. All DEDCs were additionally truncated with a $20\,\mathrm{dB}$ headroom above the noise floor to avoid influences on the DEDCs.

6. INFERENCE SETUP

In order to reduce the inference time, the DEDCs were down-sampled by a factor of $\frac{4 f_s T_{20}}{100}$.The model was implemented using the probabilistic programming framework *PyMC* [17], which also provides an implementation of the used ADVI algorithm [12]. In a first step, a mean-field approximation neglecting the correlation between the parameters was used to estimate an initial fit for the posterior distributions of the model parameters. In a second step, an extended posterior approximation using multi-variate Gaussian distributions for the entries of the smoothed histogram matrix $\tilde{\mathbf{A}}$ was used to take into account the correlation between them. This is crucial as the entries are correlated through beamformer response represented by the squared filter matrix $\mathbf{S}$. The ADVI algorithm was configured using the smooth gradient and momentum based optimizer *ADAM* [18]. In order to achieve faster and more robust convergence avoiding local minima, the optimization target – that is the experimental DEDC – was provided to the algorithm as a mini-batch with a size of 300 observations, randomly selected in each iteration step. The best matching model order was selected based on the temporal evolution of the median of the squared error between the experimental DEDC and the DCDM for each configuration and frequency band.



(a) Without panel diffusers.

(b) With panel diffusers.

Figure 1: Mean and standard deviation of the energy distribution for the two damping groups at 315 Hz. Grey contour lines are proportional to color bar ticks. Black lines mark the outline of the room and absorbing specimen.

7. RESULTS

Since in ADVI the posterior distributions are approximated by multivariate Gaussian distributions, it is sufficient to present the mean $\langle\cdot\rangle$ and standard deviation $\sigma(\cdot)$ of the damping constants and directional energy distributions inferred using the DCDM. Note that the standard deviation is scaled by a factor of two to represent the extended standard uncertainty. For both configurations, two distinctly different damping groups are found. In the following, the mean and standard deviation are calculated on the respective logarithmically scaled energy distributions in decibel. The resulting mean energies are scaled by the maximum energy across all steering directions and decay constant groups. Values below -20 dB are clamped to the same color encoding as the minimum of the color bar for clarity. The inferred directional energy distributions and uncertainties are presented using the *Mollweide* projection of the sphere in Figs. 1a and 1b. The mean decay time of each group – i.e. $\langle T_k \rangle = 6 * \log 10/\langle \delta_k \rangle$ – is indicated in the respective figure above the energy distributions. Black lines in Fig. 1 represent a projection of the room and the absorbing specimen. Grey contour lines mark the ticks in the color bar.

For the configuration not including panel diffusers (see Fig. 1a) it is most evident that the directional energy distributions corresponding to the two different reverberation time groups are well separated. The directional energy distributions for the short reverberation time is more uniform over all directions of incidence, albeit with clear minimum in the direction of the absorbing specimen. The directional energy distribution for the long reverberation time group shows distinct maxima at $(\theta, \phi) = (0°, \pm 90°)$, corresponding well to the directions of axial modes in the *y*-direction. Additionally, the directional energy distribution for the long reverberation time group shows clear minima towards north and south poles. Accordingly, these directions are not associated with short reverberation times, which is in agreement with the expected distribution of axial modes in the *y*-direction.

The uncertainty of the directional energy distributions is below 1 dB for most directions in both groups. Yet, clear angular variations are visible. While the uncertainty is less than 1 dB for directions close to $(\theta, \phi) = (0°, \pm 90°)$ for the long reverberation time group, it is increased to up to 2.5 dB for the respective other decay time group. This increase in the parameter uncertainty is most probably caused by the imperfect separation of the directional energy distributions in the experimental DEDC caused by the insufficient side-lobe attenuation of the beamformer which is not considered in the DCDM.

Additionally, a general increase in uncertainty is found for directions close to the absorbing specimen due to small quantities of incident energy in the experimental DEDC.

For the configuration including panel diffusers (see Fig. 1b), the directional energy distributions are less well angularly separated. This is to be expected, as the panel diffusers are designed to increase the diffusion of the sound field. Yet, distinct angular patterns are still visible for the group corresponding to the longer reverberation time. Most prominently, the maxima are observed at $\phi = 0°$ and $\phi = 180°$, which is in agreement with axial modes in the x-axis. Note that the elevation of the maxima is slightly tilted towards the north pole for $\phi = 0°$ and the south pole for $\phi = 180°$ which is most probably due to local diffraction effects. For the shorter reverberation time group, the directional energy distribution varies less than 10 dB for all directions except for directions corresponding to maxima in the distribution of the longer decay time group. Interestingly, for these directions uncertainties above 2.5 dB are observed, indicating that the model neglecting the beamformer response does not represent the experimental data as well as for the configuration not including diffusing elements. Note that the uncertainty is lower than 1.5 dB for the angular distribution corresponding to the longer reverberation time, which is comparable to the results for the configuration without panel diffusers.

8. CONCLUSION

The present work presents a stochastic model for directional directionally dependent sound field decay based on the Laplace transform of the modal damping distribution. A simplification capable of inferring the damping constants as well as the respective directional energy distributions is proposed. Present results show that the proposed DCDM can be used to infer common groups of damping constants from experimentally captured DEDCs. Physically plausible directional energy distributions reflecting dominant axial modes were obtained for the long reverberation time groups for both configurations. Yet, the directional energy distribution for the short reverberation time groups were not separated sufficient well. As a result, higher uncertainty in the inferred directional energy distributions was typically observed for directions with insufficient separation. Consequently, future work should focus on considering the effect of the beamformer response on the DCDM to improve the separation of the directional energy distributions for the two reverberation time groups.

9. ACKNOWLEDGMENTS

The authors would like to thank Melanie Nolan and Efren Fernandez Grande for the collaboration on the measurement campaign and Jamilla Balint for the fruitful discussions on multi-exponential energy decay.
We would like to thank the Deutsche Forschungsgemeinschaft (DFG) for funding parts of this work under grant number 298797807.

10. REFERENCES

[1] ISO354:2003, *ISO 354:2003 - Measurement of sound absorption in a reverberation room*, 2003.

[2] J. Balint, F. Muralter, M. Nolan, and C.-H. Jeong, "Bayesian decay time estimation in a reverberation chamber for absorption measurements," *J. Acoust. Soc. Am.*, vol. 146, no. 3, pp. 1641–1649, Sep. 2019. DOI: 10/gf82xq.

[3] N. Xiang, P. Goggans, T. Jasa, and P. Robinson, "Bayesian characterization of multiple-slope sound energy decays in coupled-volume systems," *J. Acoust. Soc. Am.*, vol. 129, no. 2, pp. 741–752, Feb. 2011. DOI: 10.1121/1.3518773.

[4] 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, Aug. 1, 2022. DOI: 10.1121/10.0013416.

[5] G. Götz, S. J. Schlecht, and V. Pulkki, "Common-Slope Modeling of Late Reverberation," *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, vol. 31, pp. 3945–3957, 2023. DOI: 10.1109/TASLP.2023.3317572.

[6] H. Kuttruff, "Eigenschaften und Auswertung von Nachhallkurven," *Acustica*, vol. 8, no. 4, pp. 273–280, 1958.

[7] B. Rafaely, *Fundamentals of Spherical Array Processing* (Springer Topics in Signal Processing), 1st ed. Springer-Verlag GmbH Berlin Heidelberg, 2015, vol. 8.

[8] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, no. 6, pp. 1187–1187, 1965. DOI: 10.1121/1.1939454.

[9] M. Berzborn, J. Balint, and M. Vorlaender, "Directional sound field decay analysis in coupled spaces," *J. Acoust. Soc. Am.*, vol. 149, no. 4, A116–A116, Apr. 2021. DOI: 10 / gqj4cb.

[10] M. R. Schroeder, "Eigenfrequenzstatistik und Anregungsstatistik in Räumen," *Acta Acustica united with Acustica*, vol. 4, no. 4, pp. 456–468, 1954.

[11] G. Gotz, S. J. Schlecht, and V. Pulkki, "Common-slope modeling of late reverberation," 2022.

[12] A. Kucukelbir, D. Tran, R. Ranganath, A. Gelman, and D. M. Blei, "Automatic Differentiation Variational Inference," *J Mach Learn Res*, vol. 18, no. 1, pp. 430–474, 2017. arXiv: 1603.00788.

[13] A. Gelman, J. B. Carlin, H. S. Stern, *et al.*, *Bayesian Data Analysis*, 3rd. Chapman and Hall/CRC, 2021.

[14] P. Leopardi, "A partition of the unit sphere into regions of equal area and small diameter," *Electron. T. Numer. Ana.*, vol. 25, no. 12, pp. 309–327, 2006.

[15] G. Chardon, W. Kreuzer, and M. Noisternig, "Design of spatial microphone arrays for sound field interpolation," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 5, pp. 780–790, 2015. DOI: 10.1109/JSTSP.2015.2412097.

[16] A. Lundeby, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of Measurements in Room Acoustics," *Acustica*, vol. 81, no. 4, pp. 344–355, Jul. 1, 1995.

[17] O. Abril-Pla, V. Andreani, C. Carroll, *et al.*, "PyMC: A modern, and comprehensive probabilistic programming framework in Python," *PeerJ Comput. Sci.*, vol. 9, e1516, Sep. 1, 2023. DOI: 10.7717/peerj-cs.1516.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the International Conference on Learning Representations*, San Diego, CA, 2015.