Version: Accepted Version

Sensitivity of Room Impulse Responses in Changing Acoustic Environment

Karolina Prawda

AudioLab, School of Physics, Engineering and Technology,
University of York, York, UK
karolina.prawda@york.ac.uk

Abstract—Changes in room acoustics, such as modifications to surface absorption or the insertion of a scattering object, significantly impact measured room impulse responses (RIRs). These changes can affect the performance of systems used in echo cancellation and active acoustics and support tasks such as navigation and object tracking. Recognizing and quantifying such changes is, therefore, critical for advancing technologies based on room acoustics. This study introduces a method for analyzing acoustic environment changes by evaluating the similarity of consecutively recorded RIRs. Short-time coherence is employed to characterize modifications, including changes in wall absorption or the presence of a moving person in the room. A sensitivity rating is further used to quantify the magnitude of these changes. The results clearly differentiate between types of modifications—atmospheric variation, changes in absorption, and human presence. The methods described provide a novel approach to analyzing and interpreting room acoustics, emphasizing RIR similarity and extracting information from temporal and spectral signal properties.

Index Terms—room impulse responses, similarity, short-time coherence, acoustic measurements, sensitivity

I. INTRODUCTION

The acoustics of a room are typically characterized by parameters such as reverberation time (RT), clarity, and definition, which are derived from room impulse responses (RIRs) [1], [2]. However, these objective measures do not fully capture the variability of RIRs over time and space [3]. Consequently, multiple measurements are often required to understand spatio-temporal changes in acoustic conditions. This study proposes an alternative approach to analyzing RIRs in time-varying environments.

Changes in room geometry and acoustic properties introduce differences in RIRs [4]–[7], which depend on the specific nature of the modification. Understanding such changes in room transfer functions is essential for designing active acoustics and echo cancellation systems, as these systems must be precisely calibrated to avoid feedback loop instability [6]–[9]. Analyzing RIRs in response to room geometry changes can aid in identifying and tracking modifications, such as the presence and movement of people [10], [11]. Additionally, acoustic signals can enhance the performance of tracking and navigation systems for mobile robotics, particularly in environments with transparent surfaces [12] or occluding objects [13]. In all these applications, it is crucial to distinguish stochastic variations in the measurement environment [14] from changes

in its properties [15] to ensure robust system calibration and reliable result analysis.

To address this, Prislan *et al.* introduced a cross-correlation-based sensitivity measure to quantify differences in RIRs recorded in the same room with altered sound source positions [15]. Their experiments revealed that high frequencies are more sensitive to changes in measurement conditions. Similarly, Satoh used cross-correlation to attribute changes in repeated RIRs to factors such as air temperature drift or air movement [16]. More recently, Prawda *et al.* extended this concept using frequency-dependent short-time coherence [14]. Her work demonstrated that RIR variability, caused by the inherent instability of the propagation medium, increases with frequency and RIR duration [14].

This study introduces a method to analyze differences between repeated RIRs when substantial changes occur in the measurement environment. Short-time coherence is employed to characterize temporal and frequency-domain changes, while a sensitivity rating provides a single-parameter descriptor of these modifications.

The paper is organized as follows: Section II describes the RIR analysis methodology. Section III presents the results of applying the proposed approach to measurements with varying surface absorption, while the effect of a moving person is discussed in Section IV. Section V concludes the study.

II. METHODOLOGY

Differences in consecutively measured RIRs in the same space can stem from several factors: time variance, manifested, for example, through changes in atmospheric conditions and medium movements [14], [17], [18], and modifications in the acoustics of the space, such as when objects are added or moved within the room [5], [19]–[21]. Here, the focus is on the latter. The discussion revolves around the coherence between repeated RIR measurements and the related sensitivity rating.

A. Short-time coherence

Two repeated acoustic measurements are modeled as follows:

$$
\begin{aligned}
x(t, f) &= h(t, f) + u(t, f), \\
x'(t, f) &= h'(t, f) + u'(t, f),
\end{aligned}
\tag{1}
$$

where h and h' are RIRs and u and u' are stationry background noise terms at time t and frequency f. It is assumed that h

and h' have the same distribution and analogously u and u'. The coherence between x and x' is

$$\rho_{x,x'}^2(t,f) = \frac{\left|\mathbb{E}\big[(h+u)(h'+u')^*\big]\right|^2}{\mathbb{E}[|h+u|^2]\,\mathbb{E}[|h'+u'|^2]}, \qquad (2)$$

where $\mathbb{E}$ is the expected value across time and $()^*$ denotes the complex conjugate. For discrete signals, $\mathbb{E}$ is approximated by a short-time average. Here, the time-frequency dependency is omitted where appropriate for the sake of conciseness.

$\rho_{x,x'}^2(t,f)$ can be decomposed into two terms: the loss of coherence due to changes in the measurement environment, $\rho_{h,h'}^2(t,f)$, and the expected loss of coherence resulting from sound energy decay and increasing contribution of uncorrelated background noise [14], [15], [22], $\rho_{\text{SNR}}^2(t,f)$:

$$\rho_{x,x'}^2(t,f) = \rho_{h,h'}^2(t,f)\rho_{\text{SNR}}^2(t,f), \qquad (3)$$

where

$$\rho_{h,h'}^2(t,f) = \left(\frac{\left|\mathbb{E}\big[hh'^*\big]\right|}{\mathbb{E}[|h|^2]}\right)^2, \qquad (4a)$$

$$\rho_{\text{SNR}}^2(t,f) = \left(\frac{\mathbb{E}\big[|h|^2\big]}{\mathbb{E}[|h|^2] + \mathbb{E}[|u|^2]}\right)^2. \qquad (4b)$$

Previous research [14] shows that in the case of stochastic changes that affect all the RIR reflections (temperature fluctuations, turbulence in the propagation medium), the effect of atmospheric variation can be assessed independently. However, when modified acoustics result in different reflection amplitudes (due to absorption changes), changes in the nature of the reflections (scattering changes), or the absence of some reflections (occlusion) in a RIR, disentangling the effects in (4a) becomes significantly more complex. Tracking modified reflections is particularly challenging when the RIRs become increasingly diffuse. To address this, the present study focuses on a holistic assessment of the impact of room acoustics on coherence and aims to provide empirical insights into this issue.

B. Sensitivity rating

To quantify the amount of change that has occurred within the acoustic environment under test, we use the sensitivity rating Γ as a single-parameter descriptor [15]. Unlike correlation, Γ quantifies how different or uncorrelated two RIRs are. It assumes values between 0 and 1, where 0 indicates that the RIRs are not sensitive to variations in the measurement environment and thus identical, and 1 indicates that the changes greatly affect the measured RIRs, prompting the signals to become completely uncorrelated between the measurements.

In this study, the original definition of Γ is used, however, the sensitivity function from [15] is replaced with the short-time coherence

$$\Gamma(f) = \frac{\sum_t(1 - \rho_{x,x'}^2(t,f))\sqrt{\mathbb{E}[|h+u|^2]\,\mathbb{E}[|h'+u'|^2]}}{\sum_t\sqrt{\mathbb{E}[|h+u|^2]\,\mathbb{E}[|h'+u'|^2]}}. \qquad (5)$$
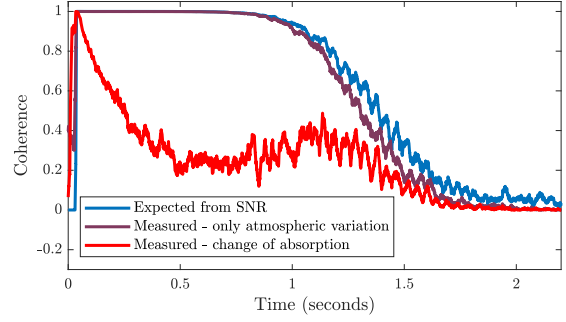

Fig. 1. Short-time coherence between repeated RIRs from *Arni* database. The coherence loss resulting from time-variant measurement environment is more similar to the expected coherence curve obtained through (4b) than when the absorption distribution in the room has changed.

Here, $1 - \rho_{x,x'}^2(t,f)$ is non-negative, while the sensitivity function can be below zero, e.g., when the phases of reflections between a pair of RIRs are reversed. As this study focuses on the presence, amplitude and time-of-arrival of reflections, using coherence allows for easier analysis of frequency-dependent Γ and offers robustness towards negative values.

When computing (5), attention must be paid to the summation limit and to analyzing only the useful RIR portions, i.e., those with sufficiently high SNR [15]. Thus, similarly to [14], the sensitivity rating in this study is calculated for the parts of the RIRs that have an SNR of 30 dB or higher.

III. CHANGE IN SURFACE ABSORPTION

First, the effect of changing surface absorption was studied using the proposed methodology on a set of RIRs measured in the variable acoustic laboratory *Arni* [23]. *Arni* features 55 variable acoustic panels, which can be configured in either a closed-reflective or open-absorptive state. Details about the room and the measurement procedures can be found in [21]. In this study, the focus was placed on broadband coherence and sensitivity rating.

This experiment considered the combinations of RIRs with the same room equivalent absorption area $A = \sum_i \alpha_i\,S_i$, where α_i is the absorption coefficient and S_i is the surface area of the ith wall. While the total absorption was constant across RIRs, the distribution varied, as the number of absorptive and reflective panels remained the same, but their placement in the room differed for each measurement.

The broadband short-time coherence of two reverberant combinations (#5143 and #5144 [23]) in *Arni*, each with 52 panels closed (RT = 1.2 s). Figure 1 compares their coherence with that measured for RIRs subject to only atmospheric variation (the same panel configuration, repeated measurements) and the expected coherence loss calculated from (4b). Both the SNR-motivated and atmospheric variation-based coherence losses were gradual, accelerating only late in the RIR. In contrast, the coherence for the absorption change dropped rapidly to around 0.3, stabilizing until the sound energy decayed. This sharp distinction highlights how changes in
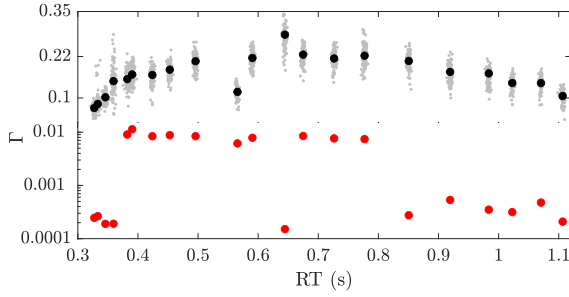
Fig. 2. Sensitivity of RIRs from *Arni* database. Grey dots show single values of Γ for each RIR pair, whilst black dots represent median values for each analyzed condition of A. The red dots are the sensitivity values for when the change in RIRs are caused only by atmospheric variation. Note that the top half of the Y-axis is linear, while the bottom half is logarithmic for enhanced readability of the results.

room acoustics are clearly identifiable and easily distinguished from stochastic variations in RIRs.

The sensitivity rating Γ for the example from Fig. 1 is 0.096 for the absorption distribution change and 0.0002 for atmospheric variation only. This indicates that the RIRs are nearly 500 times more sensitive to changes in surface absorption than to the inherent instability of the measurement environment.

To further investigate the short-time coherence and sensitivity ratings for changes in the measurement environment within *Arni*, several A values were analyzed. For each open-closed panel configuration, all 100 measured RIRs were utilized, with the first RIR from each condition serving as the reference. The remaining 99 signals were correlated with the reference, and Γ values were calculated for each coherence curve.

Figure 2 shows the sensitivity rating for the considered conditions, ordered according to their RT values. Additionally, a median sensitivity rating was computed for each condition.

The results reveal a rise in Γ as RT increases, peaking around the midpoint of the total RT range, followed by a subsequent decline. This trend aligns with the extent of absorption distribution change in each condition. At low RTs, only a few panels transition from reflective to absorptive states between consecutive RIRs (e.g., three out of 55 panels on the left side of the X-axis in Fig. 2), resulting in minimal absorption distribution variation. Similarly, at high RTs, where the majority of panels are reflective (e.g., 53 out of 55 panels on the right side of the X-axis), the room's absorption distribution shows limited variability. However, at intermediate RTs, where 20 to 30 out of 55 panels switch states between RIR measurements (middle of the X-axis), the absorption distribution changes significantly, yielding higher sensitivity ratings. Thus, Γ not only identifies the source of change in room acoustics but also quantifies the extent of that change.

Figure 2 also compares Γ values for absorption-induced changes and atmospheric-variation-induced changes between RIRs, where the same panel configuration was measured multiple times. The results corroborate those from Fig. 1, showing that Γ for atmospheric variation is orders of magnitude lower than for absorption changes, making these variations easily
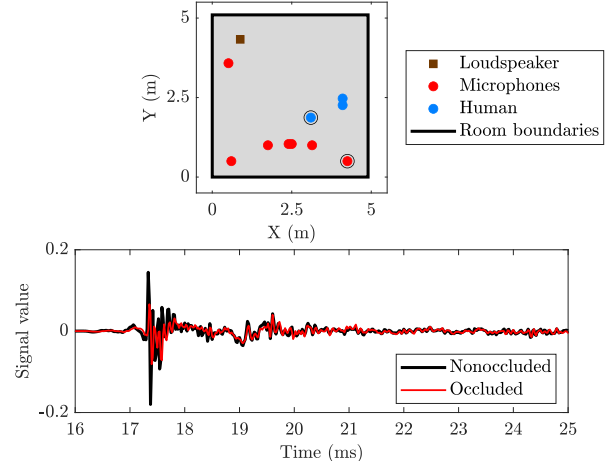

Fig. 3. (Top) Layout of the analyzed room from the SoundCam database. The circled microphone position is occluded in one of the considered configurations, when the human is in the circled position. (Bottom) A close-up of first 10 ms of two RIRs measured with the circled microphone: (black) direct path nonoccluded, (red) direct path occluded.

distinguishable. Moreover, Γ resulting from atmospheric variation is independent of RT or panel configuration changes but is instead linked to the environmental instability during the measurement. For instance, measurements with RTs between 0.38 s and 0.8 s were captured in conditions with greater variability, such as stronger air movement, compared to other measurements.

IV. HUMAN PRESENCE IN THE ROOM

In addition to analyzing the impact of absorption distribution, this study examined the effect of a human presence on RIR sensitivity. A person introduces scattering and absorption into the measurement space [24] and may occlude certain microphones when positioned in the line of sight between the sound source and receiver. To investigate this, we utilized the SoundCam database [10], which contains RIR measurements from various rooms, both empty and with a human present who changes location between recordings.

Here, we used the RIRs captured in the *Treated Room* scenario. The top pane of Fig. 3 depicts the measurement space with the sound source, receiver, and human's positions in three consecutive measurements. In one of those, a microphone is occluded, as marked by the circle around the microphone and the occluder positions. The bottom pane of Fig. 3 illustrates an example of two RIRs from the *Treated Room*. The direct path between the sound source and the microphone is occluded in one of the RIRs, causing considerable differences compared to the nonoccluded RIR, especially in the direct sound and the first few reflections.

To account for the human body's frequency-dependent effects on sound waves, $\rho_{x,x'}^2$ and Γ were analyzed across different frequency bands. A constant bandwidth of 1 kHz was chosen, centered ±500 Hz around frequencies ranging from 1 to 19 kHz [14], to ensure adequate resolution at high frequencies.
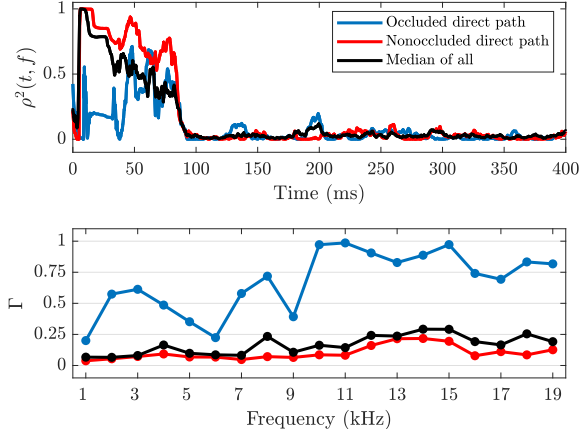
Fig. 4. (Top) Short-time coherence curves for occluded (blue) and nonoccluded (red) microphones compared to the median of all receivers for the 19-kHz band. (Bottom) Corresponding Γ for frequency bands between 1–19 kHz.

Fig. 5. (Top) Time-frequency short-time coherence for two RIRs from Sound-Cam dataset [10]. (Bottom) Corresponding frequency-dependent sensitivity Γ.

The top pane of Fig. 4 illustrates the short-time coherence values for the 19-kHz band. When the direct path is occluded in one RIR of a pair, the coherence curve drops to near-zero values at the beginning but increases over time. This pattern allows for straightforward identification of the person's position, as the timing of low-coherence values corresponds to the reflections' time-of-arrival and the room geometry.

For the assessment of the total change in room acoustics, however, a near-zero coherence may introduce out-of-distribution values that destabilize the analysis and the resulting sensitivity rating. Comparing the coherence curves for occluded and nonoccluded direct paths in the top pane of Fig. 4, the estimation of the effect of human position during this particular measurement session is overestimated by looking just at the results from the occluded microphone. Therefore, we propose to use the median of the short-time coherence of all the microphones that simultaneously recorded the same scene from different positions. The median is a stable estimate that is hardly affected by outliers [22], [25]–[28], therefore it allows to more accurately assess the room-wide effect in place of focusing on a specific sound path. In the top pane of Fig. 4, the median $\rho^2_{x,x'}$ positions itself between the coherence curves for the occluded and nonoccluded RIRs.

A similar pattern is visible in the sensitivity ratings, depicted in the bottom pane of Fig. 4. The Γ values for the RIRs captured with the occluded microphone are much higher than those for the nonoccluded microphone, reaching close to unity in the 10, 11, and 15-kHz bands. Similar to the short-time coherence, the sensitivity rating helps to estimate the location of the occluder in the room, albeit less accurately due to the lack of a direct connection between Γ and room geometry. Again, using the median sensitivity rating allows for an estimation of the room-wide effect.

To obtain more detailed information about the time-frequency characteristics of the short-time coherence and sensitivity rating, we calculate (2) and (5) on short-time Fourier
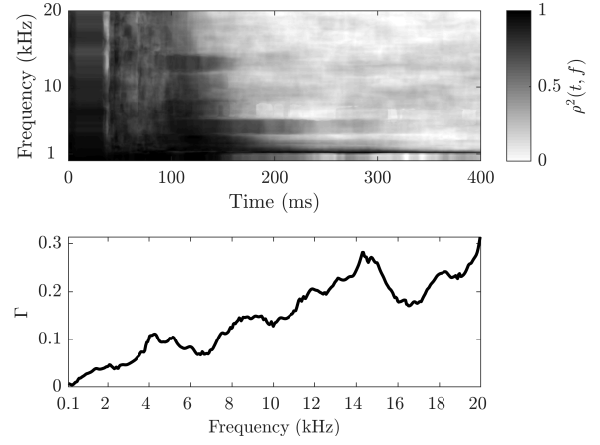
transform of a pair of RIRs with nonoccluded direct path. The results are displayed in Fig. 5.

The top pane of Fig. 5 shows the temporal evolution of the short-time coherence for frequencies linearly spaced between 100 Hz and 20 kHz. The initial high values of $\rho^2_{x,x'}$ are followed by a sudden drop around 38 ms, which is consistent with the coherence plots in the top pane of Fig. 4, and might result from the difference in the reflection pattern due to the changed position of a human between the two measurements. The drop widens with increasing frequency, suggesting a stronger effect of the occluder on shorter wavelengths.

The sensitivity rating in the bottom pane of Fig. 5 shows a nearly linear increase in Γ over frequency and is similar to the values for nonoccluded microphone in Fig. 4, although differences associated with the bandwidth appear [15]. The level of detail in the high frequencies over 10 kHz justifies the use of linear frequency spacing and constant bandwidth over the more common logarithmic approach.

V. CONCLUSION

This paper presents a method to analyze the differences between repeated RIR measurements, focusing on the changes in absorption distribution in the room and the presence of scattering and absorbing object, in this case, a human, which changes location between consecutive recordings. The results demonstrate that using short-time coherence and sensitivity rating effectively indicate changes in the measured conditions and quantify the extent of these changes. The proposed methodology is a valuable analysis tool in application scenarios such as acoustic echo cancellation, active acoustic systems, and other areas that require detailed knowledge of the acoustics of the measurement environment and its stability.

REFERENCES

[1] M. Vorländer, "Objective characterization of sound fields in small rooms," in *Proc. of the Audio Eng. Soc. 15th Int. Conf.: Audio, Acoustics & Small Spaces*, Copenhagen, Denmark, Oct. 1998.

[2] J. Pätynen, S. Tervo, and T. Lokki, "Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses," *J. Acoust. Soc. Am.*, vol. 133, no. 2, pp. 842–857, 2013.

[3] X. Pelorson, J.-P. Vian, and J.-D. Polack, "On the variability of room acoustical parameters: Reproducibility and statistical validity," *Appl. Acoust.*, vol. 37, no. 3, pp. 175–198, 1992.

[4] M. Guski and M. Vorländer, "The influence of single scattering objects for room acoustic measurements," in *Annual German Conf. Acoust. (DAGA)*, 2015, pp. 334–335.

[5] T. Becker, E. Hänsler, and U. Schultheiss, "Probleme bei der Kompensation akustischer Echos," *Frequenz*, vol. 36, no. 6, pp. 142–148, 1984.

[6] W. K. Connor, "Experimental investigation of sound-system-room feedback," *J. Audio Eng. Soc*, vol. 21, no. 1, pp. 27–32, 1973.

[7] M. Green, J. Szymanski, and M. Speed, "Assessing the suitability of the magnitude slope deviation detection criterion for use in automatic acoustic feedback control," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Brno, Czech Republic, Sept. 2016.

[8] C. Antweiler and H.-G. Symanzik, "Simulation of time variant room impulse responses," in *Proc. Int. Conf. Acoust., Speech, Signal Proc. (ICASSP)*, 1995, vol. 5, pp. 3031–3034.

[9] G. W. Elko, E. Diethorn, and T. Gänsler, "Room impulse response variation due to thermal fluctuation and its impact on acoustic echo cancellation," in *Proc. Int. Workshop on Acoustic Echo and Noise Control*, 2003.

[10] M. Wang, S. Clarke, J.-H. Wang, R. Gao, and J. Wu, "Soundcam: A dataset for finding humans using room acoustics," in *Advances in Neural Inf. Process. Syst.*, 2023.

[11] P. Marmaroli, M. Allado, and R. Boulandet, "Towards the detection and classification of indoor events using a loudspeaker," *Appl. Acoust.*, vol. 202, pp. 109161, 2023.

[12] R. Singh and K. S. Nagla, "Multi-data sensor fusion framework to detect transparent object for the efficient mobile robot mapping," *Int. J. Intelligent Unmanned Systems*, vol. 7, no. 1, pp. 2–18, 2019.

[13] D. B. Lindell, G. Wetzstein, and V. Koltun, "Acoustic non-line-of-sight imaging," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[14] K. Prawda, S. J. Schlecht, and V. Välimäki, "Short-time coherence between repeated room impulse response measurements," *J. Acoust. Soc. Am.*, vol. 156, no. 2, pp. 1017—-1028, 2024.

[15] R. Prislan, J. Brunskog, F. Jacobsen, and C.-H. Jeong, "An objective measure for the sensitivity of room impulse response and its link to a diffuse sound field," *J. Acoust. Soc. Am.*, vol. 136, no. 4, pp. 1654–1665, 10 2014.

[16] F. Satoh, "Acceptable temperature changes during synchronous averaging for reverberation time measuring by swept-sine method," in *Proc. 19th Int. Congr. Acoust. (ICA)*, Madrid, Spain, Sept. 2007, pp. 1–6.

[17] L. Tronchin, "Variability of room acoustic parameters with thermo-hygrometric conditions," *Appl. Acoust.*, vol. 177, pp. 107933, 2021.

[18] C. Othmani, N. S. Dokhanchi, S. Merchel, A. Vogel, M. Ercan Altinsoy, C. Voelker, and F. Takali, "Acoustic tomographic reconstruction of temperature and flow fields with focus on atmosphere and enclosed spaces: A review," *Appl. Therm. Eng.*, vol. 223, pp. 119953, 2023.

[19] J. Cucharero, T. Hänninen, and T. Lokki, "Influence of sound-absorbing material placement on room acoustical parameters," *Acoustics*, vol. 1, no. 3, pp. 644–660, 2019.

[20] G. Götz, S. J. Schlecht, and V. Pulkki, "A dataset of higher-order ambisonic room impulse responses and 3d models measured in a room with varying furniture," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2021, pp. 1–8.

[21] K. Prawda, S. J. Schlecht, and V. Välimäki, "Calibrating the Sabine and Eyring formulas," *J. Acoust. Soc. Am.*, vol. 152, no. 2, pp. 1158–1169, 2022.

[22] K. Prawda, S. J. Schlecht, and V. Välimäki, "Robust selection of clean swept-sine measurements in non-stationary noise," *J. Acoust. Soc. Am.*, vol. 151, no. 3, pp. 2117–2126, 2022.

[23] K. Prawda, S. J. Schlecht, and V. Välimäki, "Dataset of impulse responses from variable acoustics room Arni at Aalto Acoustic Labs," Aug. 2022.

[24] S. G. Conti, P. Roux, D. A. Demer, and J. de Rosny, "Measurement of the scattering and absorption cross sections of the human body," *Appl. Phys. Lett.*, vol. 84, no. 5, pp. 819–821, 02 2004.

[25] D. L. Donoho and P. J. Huber, "The notion of breakdown point," *A festschrift for Erich L. Lehmann*, vol. 157184, 1983.

[26] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, 2013.

[27] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Am. Stat. Assoc.*, vol. 88, no. 424, pp. 1273–1283, 1993.

[28] K. Prawda, S. J. Schlecht, and V. Välimäki, "Non-stationary noise removal from repeated sweep measurements," *J. Acoust. Soc. Am. Express Letters*, vol. 4, no. 8, pp. 081601–1–7, 2024.