TP1: DEGREE OF ANISOTROPY

Aaron Joel Petrini

Universidad Nacional de Tres de Febrero, Ingeniería de Sonido, Caseros, Buenos Aires, Argentina.
email: aaron.petrini96@gmail.com

Iván Kaspierowicz

Universidad Nacional de Tres de Febrero, Ingeniería de Sonido, Caseros, Buenos Aires, Argentina
email: kaspierowicz@gmail.com

1. Introduction

In the present work the development of a quantification method for the degree of anisotropy of a room is presented.
5 different scenarios for the same room were prepared, and 10 room impulse responses were measured, each for a different microphone position, for each one of the scenarios. From this data the results were obtained processed by a script developed in the software Phyton. Four different methods are compared to quantify the anisotropy of the room: determinates DTW and Bhattacharyya distance between the actual and the ideal energy decay function, evaluate the k threshold for the fluctuation decay curves and compare the transition times in every point.
Results showed that chairs scenario has the most diffuse sound field and that the wood board case is the worst in the 4 methods. Furthermore, an analysis of the best method for quantifying the anisotropy is presented.

Keywords: Anisotropy, Texture, Phyton, EDF, FDC.

2. Theory & state of art

2.1 Sound diffuseness.

Traditional theory divides the temporal limit for direct and reflected sound with an interval set at 50-80 ms. Transition time can be interpreted as the time after which all the reflected pulses are heavily overlapping, and therefore the envelope becomes more important. [1]

As described by J.D. Polack, impulses responses are gaussian processes. If the early part with strong reflections is discard the reverberation tail exhibits a gaussian distribution of amplitudes decaying exponentially. Later Abel suggests room impulse texture as a descriptor for reverberation quality and propose the echo density profile (EDP). Those strong reflections which describe Polack is considered now as an outlier of the Gaussian distribution.

In this paper two alternatives to quantified the isotropy of a room are given. The first based on the study of A. Bidondo and L. Pepino. The second based on the Hanyu paper.

Transition time (1) is defined based on the cumulative energy over the reflections of the normalized decay cancelled impulse response. The decay cancelled impulse response (The outliers of the IR) is obtained by subtracting the median filter of the IR. Then, the transition time is when the cumulative energy reaches the 99%. [2]

$$T_t(t) = t_t \ : \ edf(t) = \frac{cumsum\big(RIR_{out}(t)\big)}{\max\Big(cumsum\big(RIR_{out}(t)\big)\Big)} = 0.99 \tag{1}$$

$$RIR_{out} = EDC(t) - RIR_{median}(t) \tag{2}$$

Where t1 is the time for the first minimum after the maximum of the impulse response and tt the time at which the energy is equal to 0.99.

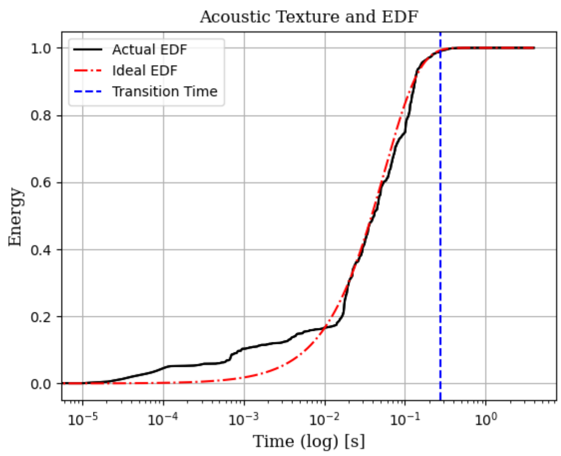

Figure 1. Transition time, Actual EDF and the ideal EDF.

The method implemented by Hanyu uses the Schroeder integration decay curve as detection of the outliers' reflections. Then after obtain the normalized energy decay cancelled impulse response $h^2(t)$ the total amount of energy is defined as:

$$R_{total} = \int_{t_1}^{t_2} h^2(t)dt \tag{3}$$

On the other hand, the energy that reaches a determinate threshold k is defined as R(k). The ratio z is defined as the relation between R(k) and the total energy which is also called the fluctuation decay curve.

$$z(k) = \frac{R(k)}{R_{total}} \tag{4}$$

Then an evaluation over a range of N threshold k is done to obtain the fluctuation decay curve. A steeper fluctuation decay curve means that the time fluctuation of the sound energy in the impulse response is small. Therefore, diffuseness of sound field might be evaluated by the fluctuation decay curve.

The threshold k_t at which z(k) becomes 0.01 is defined as the "degree of time series fluctuation" of reflected sound energy (in other words the "transition threshold"). [3]

Several methods are proposed to quantified the isotropy of the room. The first is calculate the mean of the transition time in the positions. Second is calculate the mean of the DTW distance or Bhattacharyya distance. [4,5] The last is evaluate the mean of the thresholds k.

Dynamic Time Warping (DTW) is a way to compare two -usually temporal- sequences that do not sync up perfectly. It is a method to calculate the optimal matching between two time series sequences. [4]

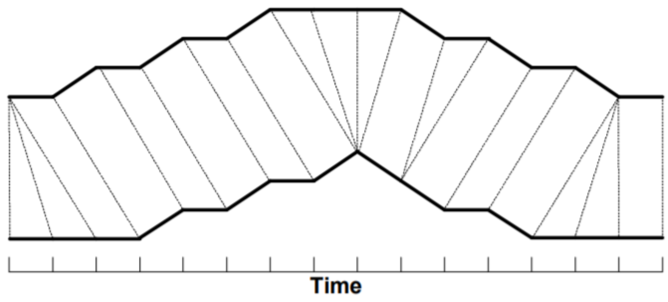

Figure 2. DTW example.

If both of the time series in Figure 2 were the same, all of the lines would be straight vertical lines because no warping would be necessary to 'line up' the two-time series. The warp path distance is a measure of the difference between the two-time series after they have been warped together. If two sequences are the same DTW is zero.

Dynamic time warping is often used in speech recognition to determine if two waveforms represent the same spoken phrase.

In other direction, The Bhattacharyya distance is a measure of divergence. This is very used to determinate a possible gaussian process. The Bhattacharyya distance is defined as:

$$B(1,2) = -ln\rho(P_1, P_2), \quad 0 \leq B(1,2) \leq \infty \tag{5}$$

Where P1 and P2 are the Bhattacharyya coefficients. [6]

Other ways to characterize the isotropy of a room is evaluating directional properties of the sound fields and its angular symmetry from a wavenumber decomposition of a spatial-temporal measurements.

"The results demonstrate how isotropy tends to increase or decrease as a function of time, depending on the disposition of the diffusing and absorbing elements. Diffusers are found to effectively redirect the energy in the room, although they do not succeed in generating a uniform incidence on the sample". [7]

2.2 Room impulse response

The impulse response is defined in the ISO 3382 [8] standards as the temporal evolution of the sound pressure observed at a point in a room as a result of the emission of a Dirac impulse at another point in the room, while it is impossible in practice to create and radiate true Dirac delta functions, short transient sounds can offer close enough approximations for practical measurement. The characteristics of a linear, time-invariant component of any system are fully described by its impulse response *h(t)*. It is desirable to obtained said impulse response as accurately as possible. In traditional impulse response measurement, periodic pulse and Maximal-Length Sequence (MLS) are often used as excitation signals [9]. An alternative impulse response measurement method has been developed by Angelo Farina [10]. In this method, a logarithmic sine sweep stimulus was employed, where the frequency varies exponentially as a function of time over the range of interest. Unlike linear-swept sine signal, the logarithmic sine sweep signal sweeps much faster in higher frequency regions than lower frequency regions. The output obtained from the system by the used of this kind of stimuli happens to be of both linear response to the excitation and harmonic distortion. The deconvolved output presents a clean separation of linear response and harmonic distortion, which allows the linear response to be delineated. In addition, this technique provides a considerable advantage with regards to SNR compared with linear sine sweep, periodic pulse or MLS techniques [10].

Finally, to obtain the room impulse response it is necessary to convolve the measured signal with the corresponding inverse sweep. The generation of the inverse filter is simply matter of time-reversing the excitation signal, and then applying to it (the reversed impulse) an amplitude envelope to reduce the level by 6 dB/octave [9].

3. Procedure

3.1 Measurement

The first step to start this work is the measurement of the room impulse responses (RIRs). It is decided to measure 10 different impulse responses for each one of the 5 different scenarios proposed for the same room. The 10 different measurements correspond to the different microphone positions. 5 different positions with a distance of 20 cm between each were marked in the floor to avoid placing the microphone in an incorrect position, and all the recordings were done with two different microphone heights, 1.2 m and 1.7m. The sound source was placed in a corner of the room, facing the corner as well, to obtained a more equal distribution of the energy in the room, simulating an omnidirectional source. All the microphone positions were more than 1 m apart from the reflecting rooms, following the recommendations from the ISO 3382 standard.

A scheme of the room, and the different microphone positions labeled with circled numbers from 1 to 5, can be seen in figure 3. All distances are in meters, and all the doors remained closed during the measurement.

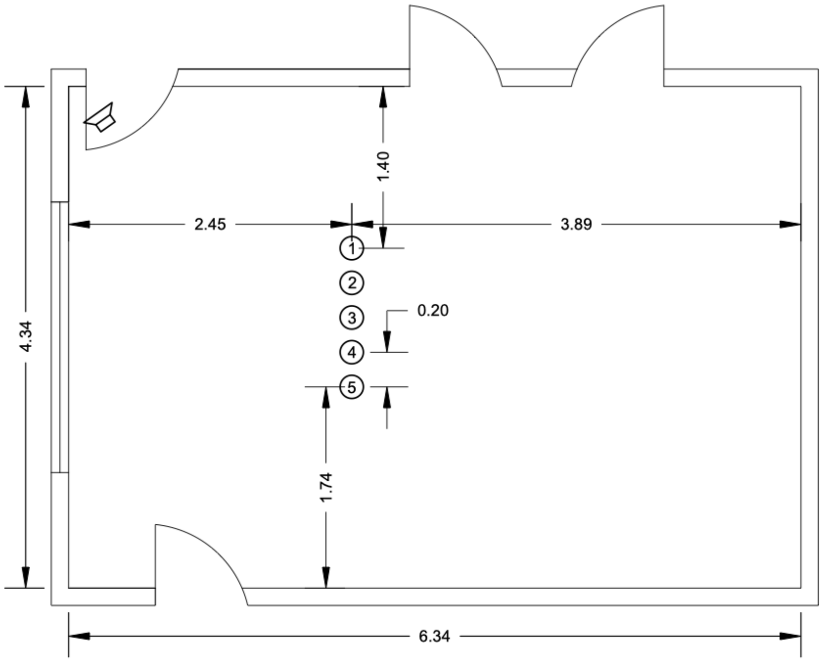

Figure 3: Scheme of the placement of the different objects needed for the measurements in the room.

All the measurements were done with the use of the app “Impulse Response Measurer” [11], available in the software MATLAB. For the choosing of the frequency range for the excitation signal, certain limitations were taken into account. It must be considered that the frequency range will be limited by the different instruments used in the measurements. In first instance, the JBL 305p MKII studio monitor was used as the sound source. The frequency range of this speaker goes from 49 HZ to 20 kHz according to the specifications provided by the manufacturer [12]. On the other hand, a Behringer ECM8000 was used as the recording microphone, its frequency range goes from 20 Hz to 20kHz according to the specifications provided by the manufacturer [13]. In order to not exceed this frequency limitations, a logarithmic sine sweep was chosen to be the excitation signal, with a length of 12 seconds and a frequency range that goes form 50 Hz to 20 kHz.

Five different scenarios were prepared for the same room. In table 1, the different scenarios are shown.

Table 1. Different proposed scenarios for the same room.

Case A	Reflecting table
Case B	Stools + Chairs
Case C	Chairs
Case D	Tied Curtains
Case E	Loose Curtains

In the following figures, pictures taken during the measurement process displaying the different scenarios are shown.

Figure 4: Case E, loose curtains.

Figure 5: Case B, chairs + stools.

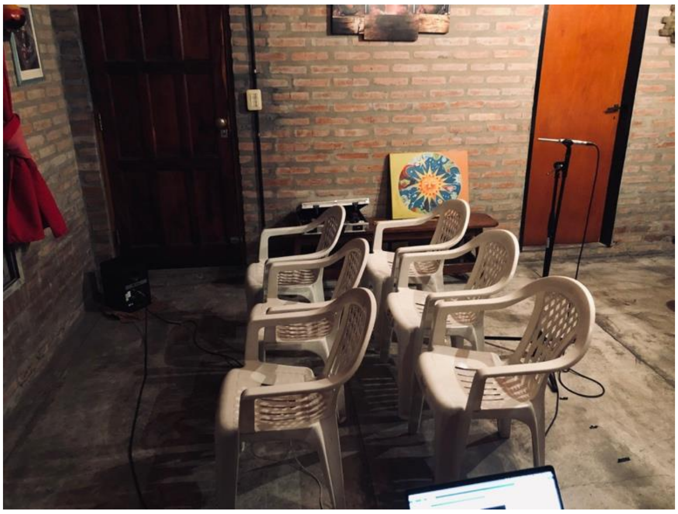

Figure 6: Case C, chairs.

Figure 7: Case D, tied curtains.

Figure 8: Case A, reflecting table.

In table 2, the different ambient conditions during the measurement process are shown.

Table 2: Ambient conditions during measurement process.

	Start	End
Time	8:27 pm	9:30 pm
Temperature	14º C	14º C
Humidity	82%	81%
Pressure	1015 hPa	1014 hPa

3.2 Data processing

After obtaining the RIRs, the necessary calculations to obtain the desired results are processed by a script developed in Phyton programming language. The figure below descripts the general flow processing of the script.

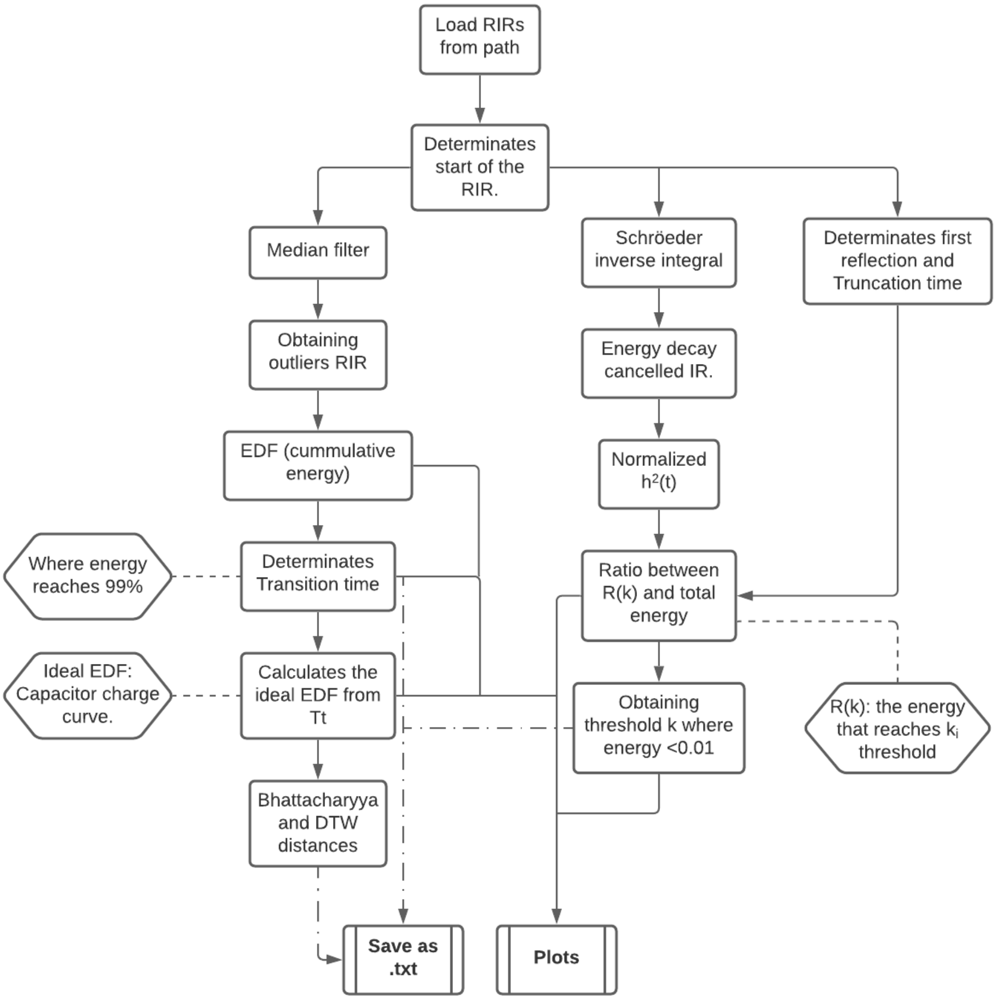

Figure 9. Diagram block for the RIR processing.

The outputs of the script are, in first place, a series of .txt files that contains the transition time, DTW distance, Bhattacharyya distance, k thresholds and transition time for all the positions. In second place, two plots: one from the actuals EDF, ideals EDF and transition times for every position and other from FDC and k thresholds for all positions.

Then, in order to obtain the results, the data files are processed in Excel to creates the tables and calculate the mean deviation for the proposed methods.

4. Results

The following table shows the results for case A. It can be observed that the position 0 has the greater value for DTW distance, k threshold and transition time. Position 3 has the highest value of Bhattacharyya distance. The mean for DTW distance is 8.312, Bhattacharyya distance: -10.662, k threshold: 40.68 and Transition time: 80.65 ms.

Table 3. Parameters obtained for case A. Max (orange), Min (blue).

A(Table)	DTW distance	Bhat. distance	k threshold	tt [ms]
0	11.220	-10.6591	88.8	88.542
1	8.806	-10.6615	33.1	75.333
2	9.867	-10.6609	65.9	83.458
3	8.205	-10.6677	17.8	76.083
4	8.888	-10.6594	28.3	82.354
5	7.542	-10.6673	27.5	75.458
6	8.517	-10.6600	26.3	79.313
7	5.575	-10.6604	20.8	85.125
8	8.471	-10.6603	55.2	78.271
9	6.032	-10.6589	43.1	82.563

The next case is B (Stools). The position 0 has the greater value for DTW distance. Position 2 has the higher k threshold and B. distance and the lower transition time. Transition time is maximum at position 7 when DTW distance is minimum. The mean for DTW distance is 6.497, Bhattacharyya distance: -10.6657, k threshold: 31.59 and Transition time: 89.148 ms.

Table 4. Parameters obtained for case B. Max (orange), Min (blue).

B(Stools)	DTW distance	Bhat. distance	k threshold	tt [ms]
0	10.449	-10.6596	43	91.375
1	7.122	-10.6568	31.9	93.417
2	10.274	-10.6612	82.3	82.917
3	7.709	-10.6585	26.1	86.771
4	6.904	-10.6594	18.5	84.667
5	5.863	-10.6583	24.9	86.667
6	6.790	-10.6583	25.9	83.250
7	2.857	-10.6501	17.5	97.354
8	4.351	-10.6574	30.4	87.896
9	2.654	-10.6489	15.4	97.167

The next case is C (Chairs). Position 0 has the higher threshold meanwhile position 2 the greater DTW distance and position 4 the maximum B. distance. Position 7 has the minimum in all parameters except for transition time which is maximum. The mean for DTW distance is 6.640, Bhattacharyya distance: -10.658, k threshold: 26.280 and Transition time: 87.965 s.

Table 5. Parameters obtained for case C. Max (orange), Min (blue).

C(Chairs)	DTW distance	Bhat. distance	k threshold	tt [ms]
0	9.769	-10.6593	39.4	90.438
1	6.886	-10.6575	31.3	89.146
2	10.172	-10.6604	32.4	84.125
3	7.804	-10.6589	35.9	85.854
4	5.229	-10.6602	19.7	82.271
5	4.971	-10.6576	20.4	85.479
6	6.994	-10.6584	21.9	83.896
7	2.666	-10.6489	15.8	105.104

8	4.942	-10.6575	25.8	86.438
9	4.967	-10.6574	20.2	86.896

The next case is D (Tied curtains). Position 0 has the higher threshold and DTW distance. position 2 has the maximum B. distance. Position 7 has the minimum in all parameters except from transition time which is maximum again. The mean for DTW distance is 7.022, Bhattacharyya distance: -10.657, k threshold: 34.220 and Transition time: 90 ms.

Table 6. Parameters obtained for case D. Max (orange), Min (blue).

D(Tied)	DTW distance	Bhat. distance	k threshold	tt [ms]
0	11.187	-10.6592	90.5	91.479
1	7.474	-10.6560	14.7	98.813
2	11.148	-10.6600	68.9	84.354
3	7.077	-10.6595	28	85.708
4	5.694	-10.6599	22.5	83.396
5	5.492	-10.6565	25.3	94.208
6	6.901	-10.6584	27.6	82.313
7	3.516	-10.6494	13.2	106.333
8	7.290	-10.6586	33.5	82.979
9	4.436	-10.6566	18	90.438

The next case is E (curtains). Position 0 has the higher DTW distance and position 2 the maximum B. distance and k threshold. Position 7 once again, has the minimum in all parameters except from transition time which is maximum. The mean for DTW distance is 6.777, Bhattacharyya distance: -10.658, k threshold: 36.190 and Transition time: 85.290 s.

Table 7. Parameters obtained for case E. Max (orange), Min (blue).

E(Curtains)	DTW distance	Bhat. distance	k threshold	tt [ms]
0	11.079	-10.6593	50.3	88.583
1	6.566	-10.6573	19.8	90.833
2	9.930	-10.6606	114	77.542
3	7.397	-10.6588	27.2	85.688
4	8.788	-10.6603	39.5	77.333
5	5.008	-10.6579	24.8	85.479
6	4.990	-10.6591	25.7	84.083
7	2.527	-10.6504	13.3	97.479
8	7.482	-10.6596	27.8	78.646
9	4.008	-10.6570	19.5	87.229

In order to compare the results and can make any conclusions the next tables show the mean deviation for every case from the global mean value.

Table 8. Ranking of the mean values in every case.

Mean	DTW distance	Bhat. distance	k threshold	tt [ms]
1	Case C	Case B/D	Case C	Case D
2	Case B	Case B/D	Case B	Case B

3	Case E	Case C	Case D	Case C
4	Case D	Case E	Case E	Case E
5	Case A	Case A	Case A	Case A

Table 9. Mean deviation values from the global mean in every case.

Mean	DTW distance	Bhat. distance	k threshold	tt [ms]
0	0.919	0.9999	0.778	1.039
1	0.927	0.9999	0.935	1.029
2	0.967	0.9999	1.013	1.016
3	1.002	1.0000	1.071	0.985
4	1.186	1.0003	1.204	0.931

Note: The tt values from the Table 7 are inversed because the less values of transition times, the poor diffuseness.

Tables 6 and 7 reveals that the wood board is the worst scenario for every parameter. The methods of DTW distances and the threshold k are almost coincident in the rankings except from the third and fourth places which are in inverse order. In the Bhattacharyya distance method, the values of the mean to the global one are close to each other, and are equal in some cases (first and second places). Case B is unanimous second place and Case E is the same for Bhattacharyya, k threshold and Tt.

Observing the figures, it can be determined that the LP1 presents the most deviation from ideal texture and the highest value of k. The point LP1 is the farther measurement position from the reflection superficies.

5. Conclusions.

By the results from section 4, it can be determinate that case A is the worst for all proposed methods. All the results from table 6 and figures in "appendix A" show a correlation and coherence if a visual inspection is realized.

The threshold from fluctuation decay curve and DTW distance from acoustic texture are coincident in their results over the tests. Based on previous studies [2] and the given results, it can be demonstrated that the difference between the ideal (gaussian) Energy Decay Function and the actual EDF characterizes the diffuseness of the room in that measurement point.

The results prove that in most cases the higher the transition time the lowest the DTW, B. distance and k threshold. This means that a higher transition time reveals a greater diffuseness in that particular point. In the opposite way, no correlation has been found, a lower transition times don't show a correlation with lower diffuseness.

DTW distance particularly results in the best method for characterizing the room in this paper. In order to prove consistency of all methods and make more conclusions, a measurement in several rooms and reverberation chambers must be done. Furthermore, the measurement points done for in this paper are not enough to make a full characterization of the isotropy in a room.

It should be noted that the results are for the full IR. If quantification of isotropy in the octave band wants to be done, the same process is repeated for the IR filtered in every frequency band.

With the purpose to enrich the paper, the method proposed in [2] with the diffusion coefficient could be an interesting approach to quantities the isotropy in a room.

6. Bibliography

[1] A. Bidondo, J. Vazquez, S. Vazquez, M. Arouxet, G. Heinze. A new and simple method to define the time limit between the early and late sound fields. Audio Engineering Society. Paper 9631. (2016)

[2] A. Bidondo, L. Pepino. Room acoustic texture: a methodology for its quantification. Proceedings of the 23rd international congress on acoustics. Germany. (2019)

[3] T. Hanyu. Analysis Method for Estimating Diffuseness of Sound Fields by Using Decay-Cancelled Impulse Response. Building acoustics. Vol 21, No 2, pp 125-134. (2014)

[4] S. Salvador, P. Chan. FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space. Florida institute of technology. Melbourne.

[5] A. Bhattacharyya, On a measure of divergence between two multinomial populations. Indian journal of statistics. Vol.7, No 4 (Jul 1946), pp. 401-406. (1946).

[6] F. Schweppe. On the Bhattacharyya distance and the divergence between gaussian processes. Department of electrical engineering. M.I.T. Cambridge. (1967).

[7] M. Nolan, M. Berzborn, E. Fernandez-Grande. Isotropy in decaying sound fields. Acoustics Society of America. Pp 1077-1088, August. (2020).

[8] ISO 3382-2: 2008. Acoustics – Measurement of room acoustic parameters- Part 2: Reverberation time in ordinary rooms.

[9] Angelo Farina, “Simultaneous measurement of impulse response and distortion with a swept sine technique,” Presented at the 108th AES Convention, Paris, France, 2000.

[10] Angelo Farina, “Advancements in impulse response measurements by sine sweeps,” Presented at the 122ed AES Convention, Vienna, Austria, May 2007.

[11] MATLAB, Official Website. Measure Impulse Response of an Audio System. https://www.mathworks.com/help/audio/ug/measure-impulse-response-of-an-audio-system.html

[12] JBL Official Website. https://jblpro.com/products/305p-mkii

[13] Behringer Official Website. https://www.behringer.com/product.html?modelCode=P0118

7. Appendix A. Figures

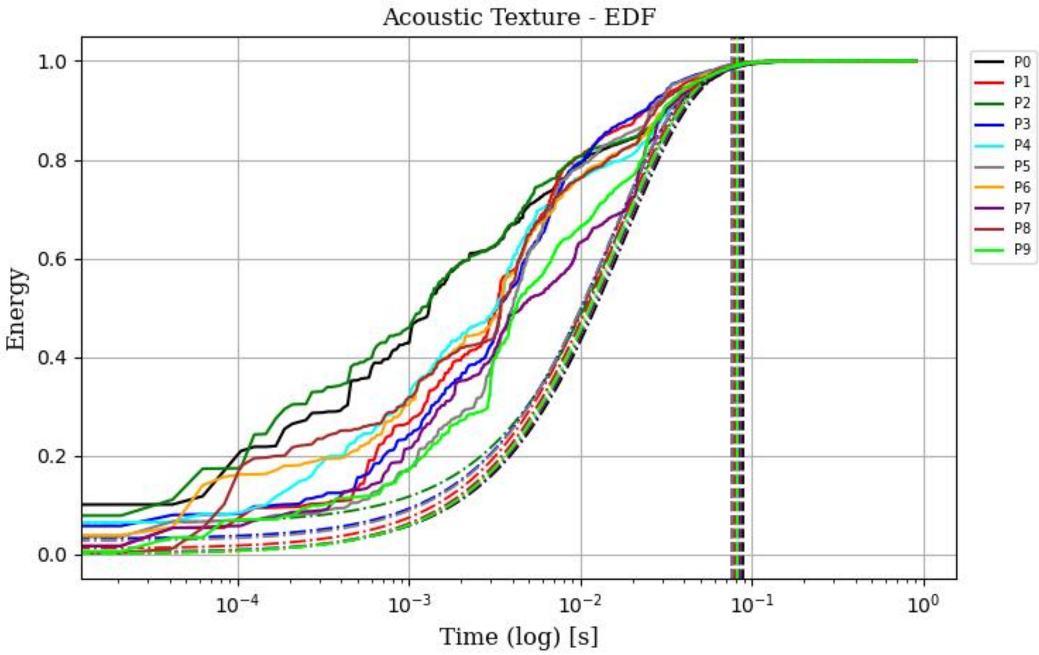

Figure 10. Ideal EDF (dotted line) and actual EDF for the case A (wood board).

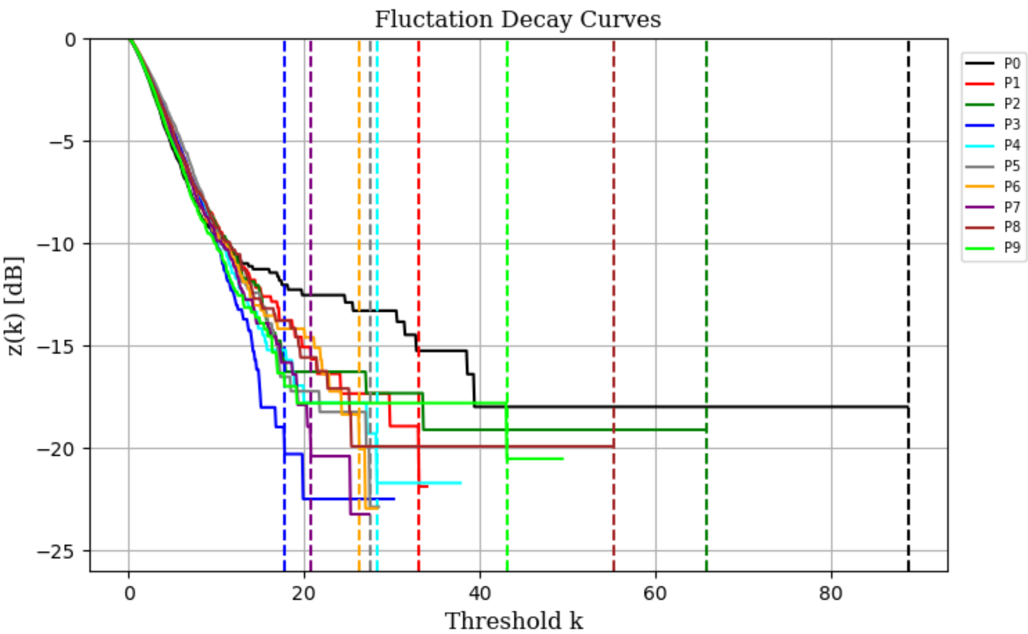

Figure 11. FDC for the case A (wood board).

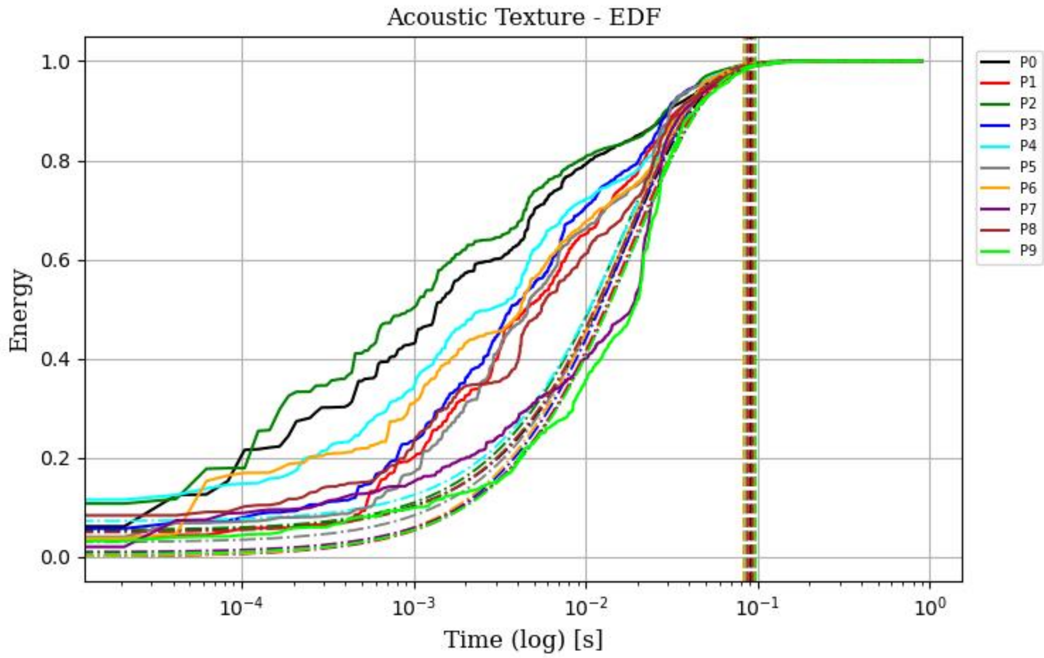

Figure 12. Ideal EDF (dotted line) and actual EDF for the case B (stools).

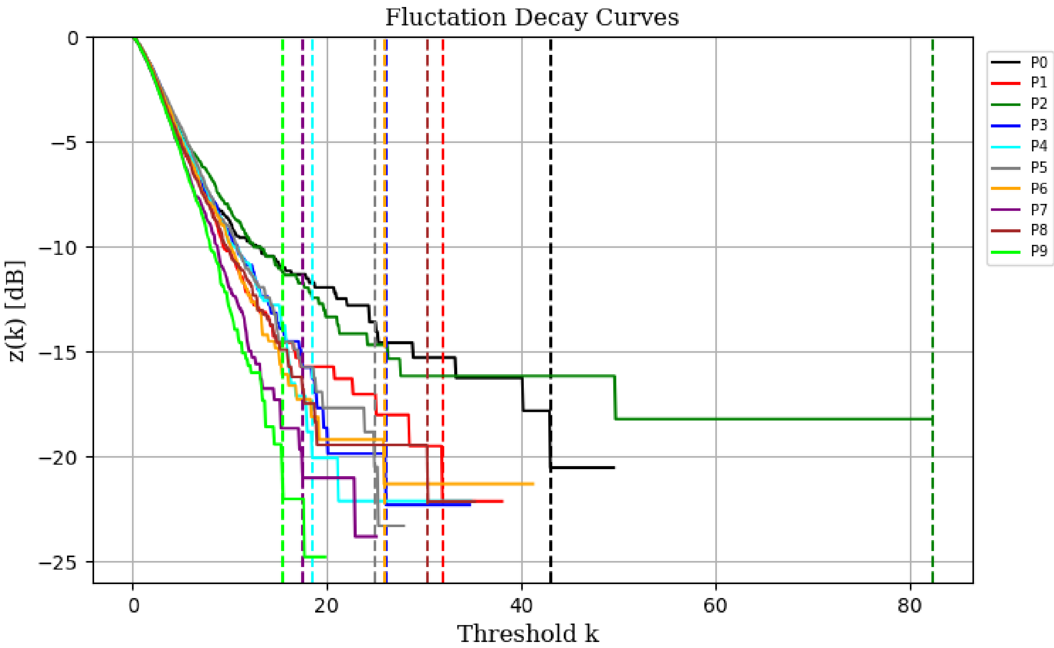

Figure 13. FDC for the case B (stools).

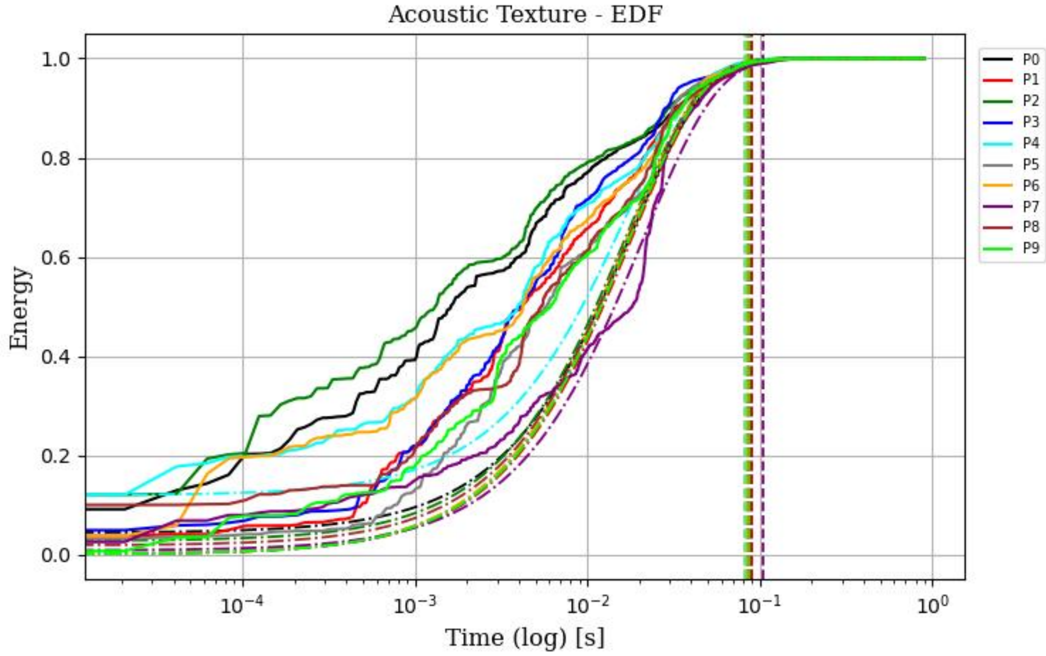

Figure 14. Ideal EDF (dotted line) and actual EDF for the case C (chairs).

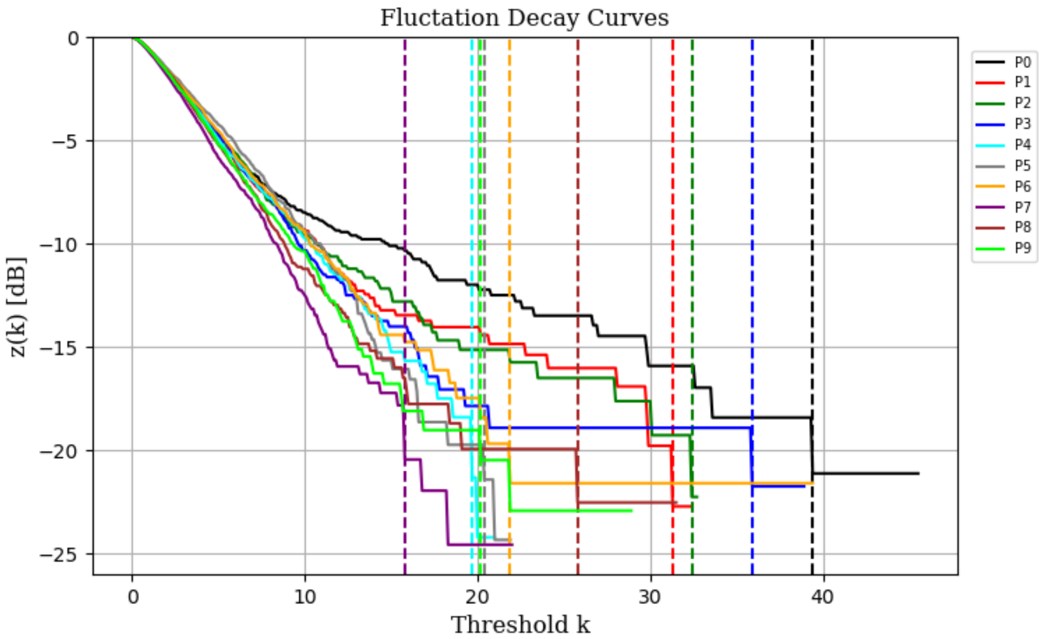

Figure 15. FDC for the case C (chairs).

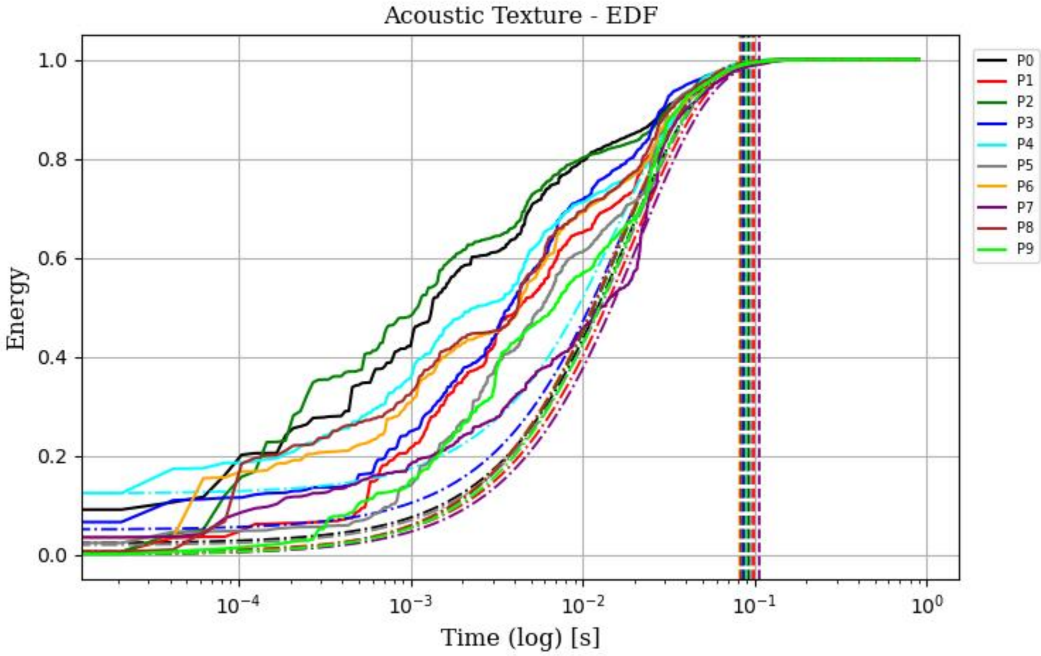

Figure 16. Ideal EDF (dotted line) and actual EDF for the case D (tied curtains).

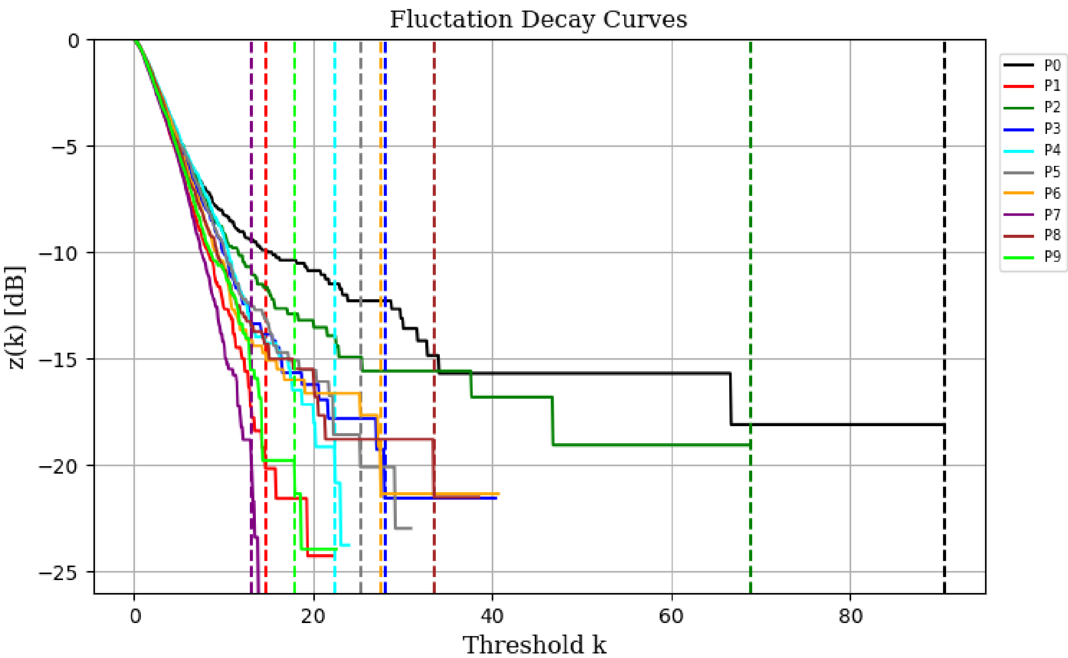

Figure 17. FDC for the case d (tied curtains).

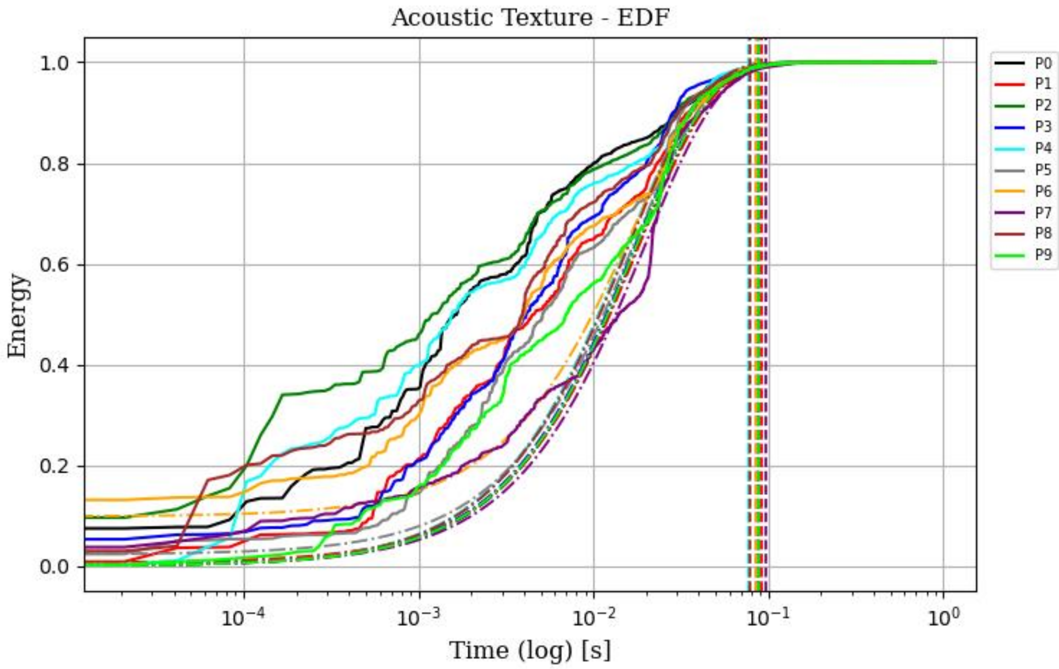

Figure 18. Ideal EDF (dotted line) and actual EDF for the case e (curtains).

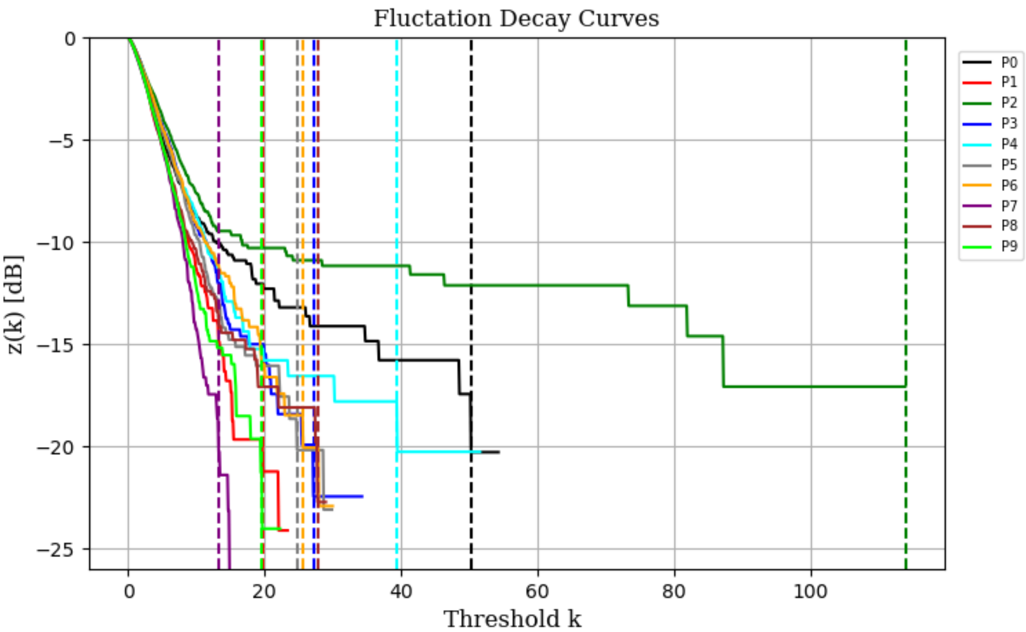

Figure 19.FDC for the case e (curtains).