**Title**
An Acoustic Experiment for Generating Pure Material Impulse Response



**Author**
Wang, Zhongran

**Publication Date**
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

# An Acoustic Experiment for Generating Pure Material Impulse Response

A Thesis submitted in partial satisfaction of the
requirements for the degree Master of Fine Arts

in

Theatre and Dance (Design)

by

Zhongran Wang

Committee in charge:

Professor Shahrokh Yadegari, Chair
Professor Robert Brill
Professor Victoria Petrovich

2019

The Thesis of Zhongran Wang is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Thank you to Professor Shahrokh Yadegari for his mentorship and encouragement of embracing the beauty of art and science.

Thank you to UCSD for offering an academia environment for multi-culture and interdisciplinary study.

ABSTRACT OF THE THESIS

An Acoustic Experiment for Generating Material Impulse Response

by

Zhongran Wang

Master of Fine Arts in Theatre and Dance (Design)

University of California San Diego, 2019

Professor Shahrokh Yadegari, Chair

This paper will present an approach for measuring the reflection impulse response (IR) of various physical materials. Digital signal processing (DSP) techniques are applied to minimize the acoustic influences by the testing environment and test devices, such as speakers, microphones and digital converters. The result IR represents a physical material's acoustic characteristics and can be used for platforms such as virtual-reality sound generation, acoustic space simulation etc.

# Chapter 1

# Introduction

With the development of integrated multimedia technology, virtual-reality platforms have become popular for games, animations, and brand-new ways of storytelling. The need for virtual reality audio reproduction has become high. The audio should be able to provide users a natural sounding environment that provides spatial characteristics that matches the visual cue. Besides providing the aural cues of simulated sound source according to the head position of the listener, spatial audio also presents acoustic space/room information: For example, a wooden cabin and a stone cave, should present different aural characteristics.

There are techniques for virtual acoustics simulation for room modeling that consider physical material reflection. For example, in DIVA system, an interactive virtual acoustic environment, low order filter IIR filter based on material's absorption coefficients are implemented to simulate wave reflectance from specific material [1]. There is a database of absorption coefficients widely used for synthesizing an impulse response (IR) by calculating the spectral roll-off and then taking the inverse Fourier transform. The library facilitates the acquisition of synthesized material IR based on ISO 354:2003 [2][3]. However, synthesizing the material's acoustic response, which models absorptions, scattering, diffraction, etc. is usually complicated. Physically, a material IR should involve two parts: the specific material's impedance and the surface's geometric shape [4] [5]. The impedance could be easily modeled as the absorption and reflection coefficients of each octave frequency band, while scattering effects depending on material's physical surface shape and corrugations introduce complex micro-delays that affect the frequency response. It is difficult to encode the scattering-effect information into a synthesized IR.

For acquiring more natural and precise aural experience with material reflection, higher resolution of absorption and scattering characteristics, we propose a method combining acoustic measurement and post DSP processing to characterize physical material's acoustic response, which minimizes the acoustic influences from the testing room and speaker coloring. The block diagram below (Figure 1) shows the general steps, and it is being discussed in later chapters.

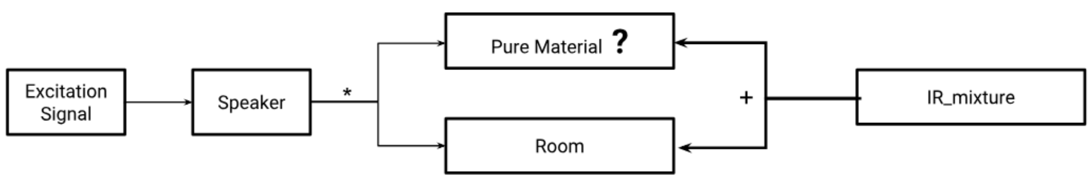


**Figure 1:** General model for acquiring pure physical material IR

# Chapter 2

# Approaches

## 2.1 The Configuration of the Environment

The experiment consists of two parts: acoustic tests in Spat Lab and analysis and post processing using MATLAB. We use a single Genelec 8030B speaker, an omni-directional microphone and an RME BabyFace audio interface. In the acoustic measurement, the speaker is placed 0.6 meters from microphone with an absorptive partition mat in between. For blocking undesirable reflections from reflective components such as the back wall and the floor, we added absorptive materials behind the microphone as well as between the tested material and the microphone / speaker. Detailed configurations in both top-view and section plan are shown in Figure 2.

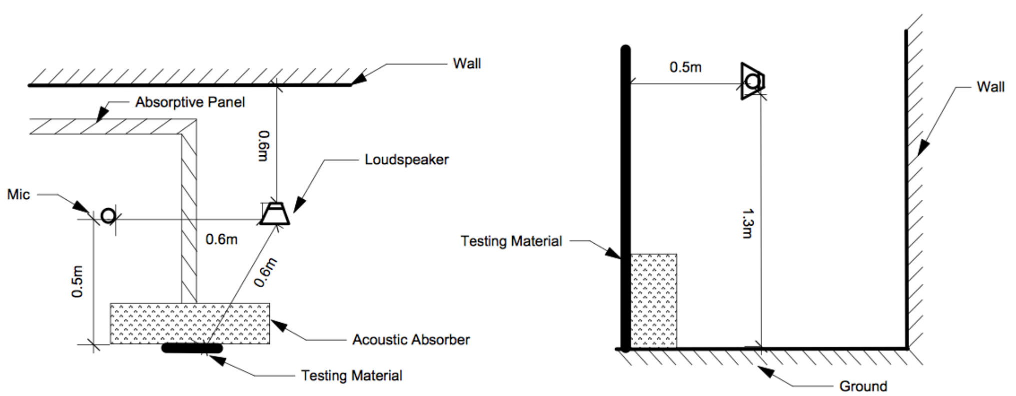


**Figure 2:** Acoustic measuring configuration, left is top-view, right is section plan

## 2.2 Software Setup

We use MATLAB Impulse Response Measurer package for testing IRs. It returns a normalized IR in the range of -1 to 1 as well as the raw recorded audio data. A normalized IR is suitable for showing the energy ratio between the frequency bins. However, when comparing the reflections of different materials, we need unnormalized IRs so that the overall reflected and absorption energies may be compared. Therefore, we plan to have a step for denormalization in the later phase. In this software, we use the MLS (Maximum Length Sequence) as the excitation signal. MLS is a pseudo-random binary sequence, with which the auto-correlation of MLS as an impulse produces a flat-spectrum impulse, and its circular cross-correlation with the output of the tested system returns the IR. The excitation signal's level is set to -13 dBFS, and the noise floor of the general testing environment was observed to be around -40 dBFS.

## 2.3 Acoustic Experiment and Measurement

We placed the material, a plexiglass covered by a sheet of paper, 0.5 meter from the microphone and speaker. While the speaker plays the MLS signal, the microphone receives the reflected sound from the test material together with the room's influence. We call this result an "IR mixture." Then, we tested the IR in the same configuration without the material in front, which provided an open-space room IR. Lastly, the speaker's IR was also tested in a different configuration. In this case, the speaker and microphone were placed at the height of 0.5 meter, while their distance was 1.0 meter from each other (Figure 3).

# Chapter 3

# Analysis and Post Processing

## 3.1 Denormalization and Subtraction

Now we obtained three IRs, with which we needed to find the relationships among them. The system was modeled as shown in Figure 1. In order to create a pure material IR, we subtracted the room's influence ("IR room") from the IR mixture in the time domain, and deconvolved the result with the speaker IR in the frequency domain to remove the coloring by the speaker.

After subtracting the IR room from the IR mixture, we have encountered two challenges: analyzing and discerning the peaks of the reflection from the real IR data, and finding the unnormalized magnitude values of each IR.

First, we look into the real IR mixture data by comparing the situations when the material is spaced 0.5, 1 and 2 meters away. Figure 3 shows the configuration in a top-view floor plan.

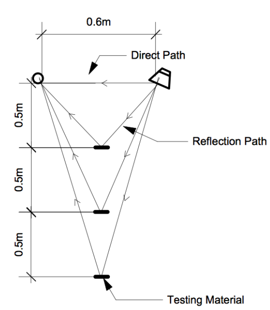


Figure 3

**Table 1**

| Distance (meter) | Path Distance Difference (meter) | Delay Time (second) | Delay in Samples |
|---|---|---|---|
| 0.5 | ~0.6 | ~0.0017 | ~75 |
| 1 | ~1.5 | ~0.0044 | ~195 |
| 2 | ~3.4 | ~0.01 | ~440 |

Figure 3 indicates the direct path from the speaker to the microphone passing through the partition mat, and the reflection path from the speaker to the material and then reflected to the microphone. The material position is shown with the bold short line. For three different material spacing, the differences of the distance and the delay time between the direct path and the reflection path are calculated in the prediction table (Table 1). Note that, in the table, calculations are based on the sampling rate of 44100 samples/sec and the speed of sound in the air at a normal room temperature (343 meters/sec).

Figure 4 is the plot for all these 3 IRs' real data under different spacing configurations. The comparison of them under the same time axis (X axis) reveals that all IRs' first smaller peak overlaps, and its distance in sample to the second peak varies in a relationship which matches the value of delay in samples in prediction table (Table 1). We hypothesize that the first peak in the real data shows the direct path from the microphone to the speaker, and the second peak indicates the reflection path from the test material. Since we are only interested in the information of the material's reflection, it is expected that the first peak will be removed by the subtraction after having unnormalized IRs.

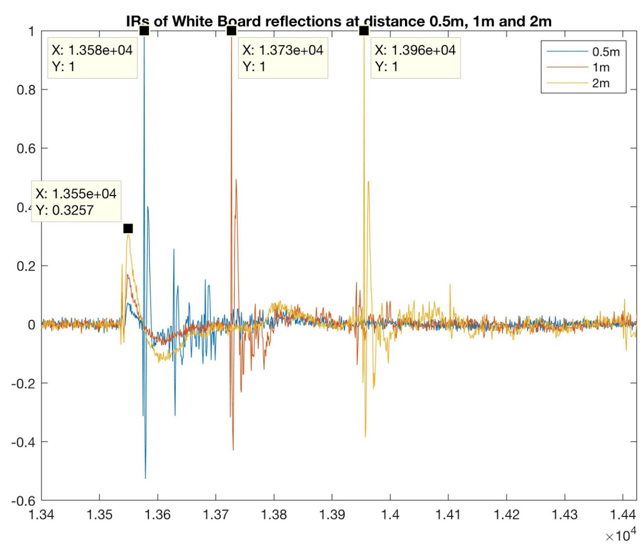


**Figure 4**

As mentioned in the section Approaches and Software Setup, we will need to de-normalize the IRs produced by MATLAB Impulse Response Measurer to find the real magnitude values of each IR. Figure 5 illustrates the planned method. Since we have the normalized IR and raw recorded audio, we may be able to find the scaling factor between the magnitude of the convolved result (i.e., the normalized IR with the input MLS and the magnitude of the raw recorded audio by a simple division. This scaling factor will then be used to acquire the unnormalized IR. Figure 6 plots denormalized IR mixture (with the material reflection) and denormalized room IR without the material. Note that the time information is not processed. It is shown that the room IR, both in the time and magnitude perspective, fits in the first peak of an IR mixture. By subtracting them, the material-reflection IR is acquired (Figure 7).

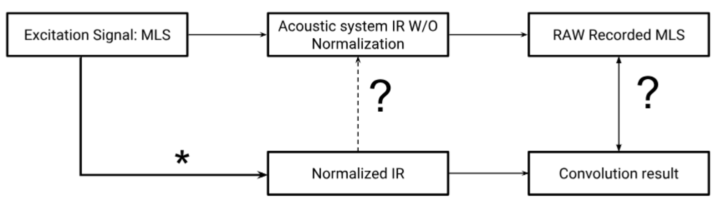


**Figure 5:** denormalization

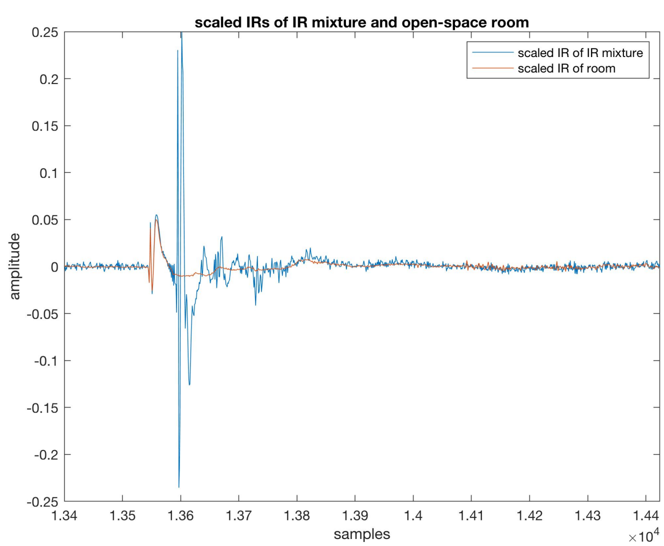


**Figure 6**

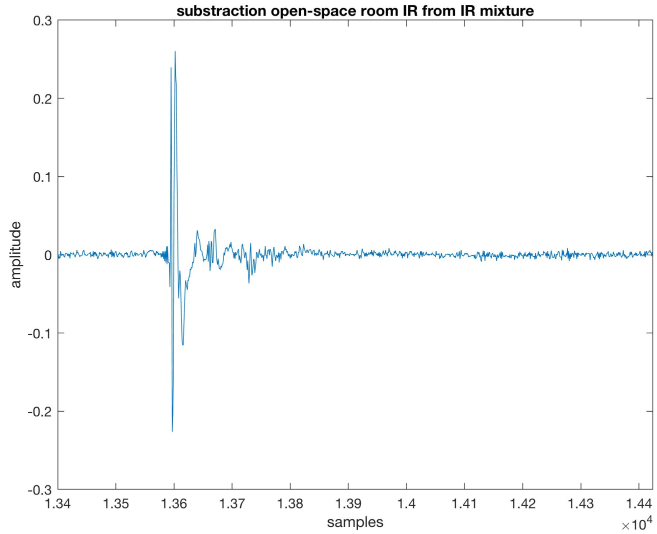


**Figure 7**

## 3.2 Deconvolution

### 3.2.1 Naïve Inverse Filtering

The subtraction result still contains the influence of the speaker. We chose to deconvolve the speaker's IR from the subtraction result in the frequency domain. The most straightforward way for obtaining a deconvolved IR is the inverse filtering (Equation 1), which takes the fast Fourier transform (FFT) of both signals and divides their spectrum followed by the inverse FFT. H(f) is an inverse filter for undoing the influence from testing device, i.e., the speaker. However, this naïve inverse filter may boost the noise in frequency response of the final result. As in the speaker IR measurement process, noise mixed into signal corrupts the higher frequencie more when it is being divided for acquiring the true IR, the low-level noise got boosted in the final result. Also, the speaker has its own produced frequency range, so as the real measurement shows the energy distributed on the low and high frequency end, which are outside of the range, are low values, while the inverse filtering will increase the gain there. For above reasons, specialized approaches such that inverse filter's spectral gain could be controlled over different frequency band are needed.

$$H(f)_{naive} = \frac{1}{\mathrm{FFT}\{\mathrm{IR}_{spk}\}}$$

$$pure\ material\ IR = \mathrm{iFFT}\{\mathrm{FFT}\{\mathrm{IR}_{subtraction}\}\mathrm{H(f)}_{naive}\}$$

**Equation 1**

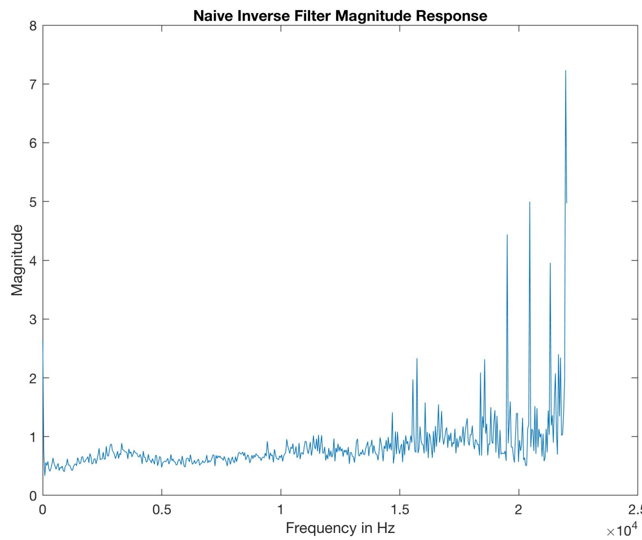


**Figure 8**

## 3.2.2 Kirkeby Algorithm and Weiner Deconvolution

Kirkeby Algorithm is a classical method usually used for removing unwanted influence, in our case, the testing speaker's coloring, from the measurement system, without over-correcting and boosting the gain of the frequency contents outside of the range. The regularization mechanism introduces an equalizer, i.e., a controllable gain over different frequencies, to adjust inverse filtering result [6].

$$H(f)_{Kirkeby} = \frac{\mathrm{FFT}\{\mathrm{IR}_{spk}\}^{*}}{|\mathrm{FFT}\{\mathrm{IR}_{spk}\}|^{2} + \beta\,|\mathrm{Reg(f)}|}$$

**Equation 2**

Equation 2 above shows how Kirkeby inverse filter works for deconvolving the loudspeaker's influence. $\beta\,|Reg(f)|$ is the regularization equalizer, $\beta$ is the adjustable strength parameter [7]. When the gain of regularization equalizer is small enough, the inverse filter retrieves all frequency contents. When it's getting bigger in gain, it starts to take effect in the denominator. Depending on the frequency of interest, the inverse filter reduced the gain of those unwanted frequencies. For example, if Reg(f)=1 over a frequency range of 60 Hz to 2200Hz is set as shown in Figure 9, the inverse filter's gain boosting outside of the range will be reduced, the reduced level is decided by $\beta$.

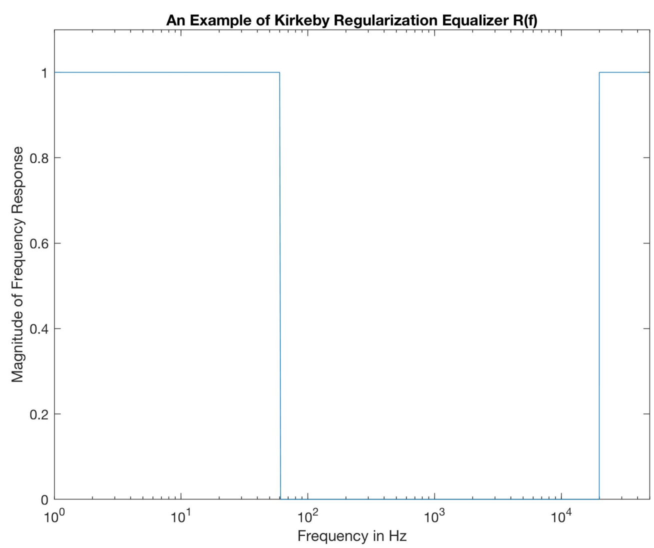


**Figure 9**

The classical Kirkeby is effective for the problem of over-correcting the system's true impulse response by solving the boosting gain outside of a frequency range. However, the function of a simple regularization equalizer like above might be limited in solving the gain boosting issue within the frequency range we purse. For example, in Figure 8, the peaky noise starts to corrupt frequency contents around 15k Hz, and we might be still interested in the contents higher than that. The classical Kiekeby method can define the regularization frequency range up till 15kHz, but a more precise examination of the frequency response outside of range is lost.  Furthermore, to define a regularization equalizer with higher frequency resolution, which can act to apply a weighted gain for each frequency based on the noise corruption level, is expected.

Weiner deconvolution takes the consideration of noise corruption problem over different frequencies by introducing a signal-to-noise ratio (SNR) weighted spectrum gain controller [8]. The Weiner inverse filter thus might be considered as a weighted regularization equalizer (Equation 3). As we measured the speaker's IR in the room with the noise floor of around -40 dBFS, the noise may have degraded the spectrum of the speaker especially in higher frequencies. As a result, in Figure 10, we observe that with a naïve dividing approach, the undesirable noise in the higher-frequency is boosted.

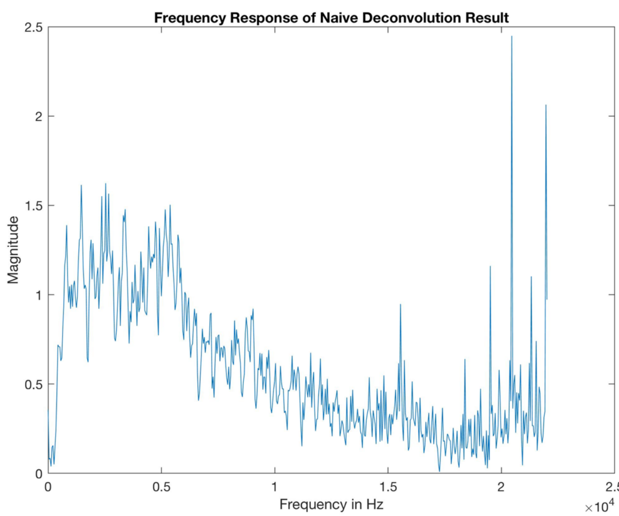


**Figure 10**

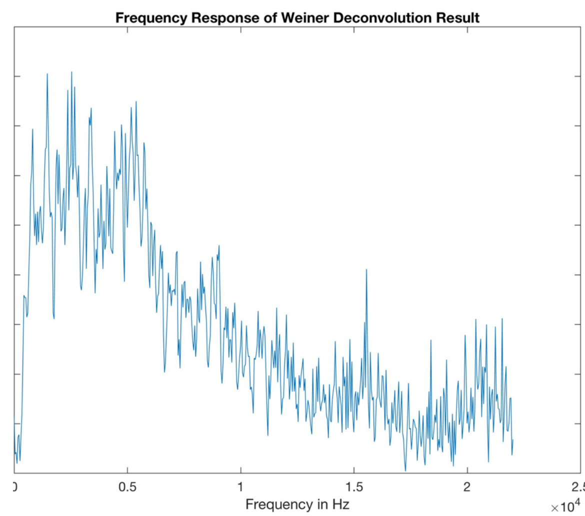


**Figure 11**

As shown in Equation 3, by taking the power-spectrum density of the speaker IR and the frequency-dependent SNR, the left part of the equation acts as a spectrum gate. In higher frequencies, where SNR is low, the spectrum of the subtraction result is weighted more while the speaker spectrum is weighted less, and vise versa. Figure 11, in comparison to Figure 10, shows the result after a Weiner deconvolution. For higher-frequency parts, the peaks caused by the noise amplification were effectively reduced. By taking the inverse FFT, we acquire pure material IR presented in Figure 12.

$$H(f)\_weiner = \mathrm{iFFT}\left\{\frac{\left|\mathrm{FFT}\{\mathrm{IR}_{spk}\}\right|^2}{\left|\mathrm{FFT}\{\mathrm{IR}_{spk}\}\right|^2 + 1/SNR(f)} \frac{1}{\mathrm{FFT}\{\mathrm{IR}_{spk}\}}\right\}$$

$$pure\ material\ IR = \mathrm{iFFT}\left\{H(f)_{weiner}\mathrm{FFT}\{\mathrm{IR}_{subtraction}\}\right\}$$

**Equation 3**

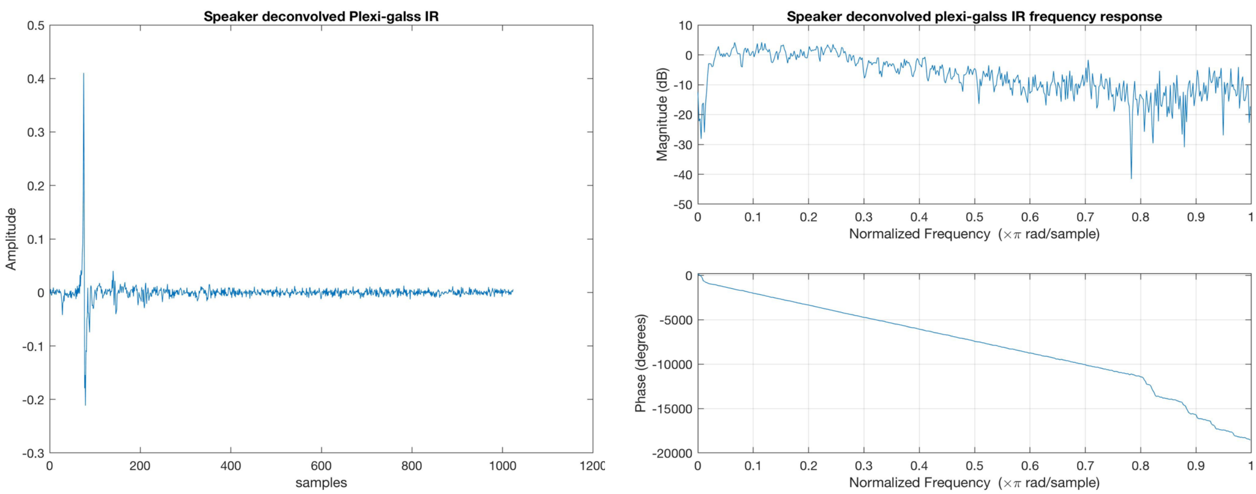


**Figure 12:** Left is material IR in time domain, right is this IR's frequency response

# Chapter 4

# Results and Discussion

## 4.1 Problems in Acoustic Measurement and Post-processing

As examined, the acquired impulse response is a stable FIR system with all zeros inside the unit circle. As shown in Figure 12, the phase approximately follows a linear curve, with noise distortion in phase occurring in higher frequencies. Linear phase filter is generally used as audio filter for re-constructing the magnitude spectrum. As linear phase produces same delay amount for each frequency, it can maximumly preserve the time domain waveform [9].

However, the noise corruption observed in higher frequencies may come from the nature that, in measurement, higher frequency experiences higher rate of transients in time and thus contains more information than lower frequency. Noise energy spread evenly on different frequencies, so the same amount of noise is easier to affect the high frequencies. Also, in practice, MLS as the excitation signal with sharp transitions inherent in the signal actually brings higher crest factor, which degrades the SNR of the result IR [10]. It is suggested to have more runs of MLS sequence in the excitation process, so that the averaging gives higher SNR performance.

The post processing such as the inverse filtering process, and linear phase may introduce pre-ringing artifacts, under the effect of which, pre-echo might show near the sharp transient of sound. It is concerned that precedent temporal masking is noticeable for auditory perception. Below, Figure 13 shows one example of pre-ringing IR result after deconvolution.

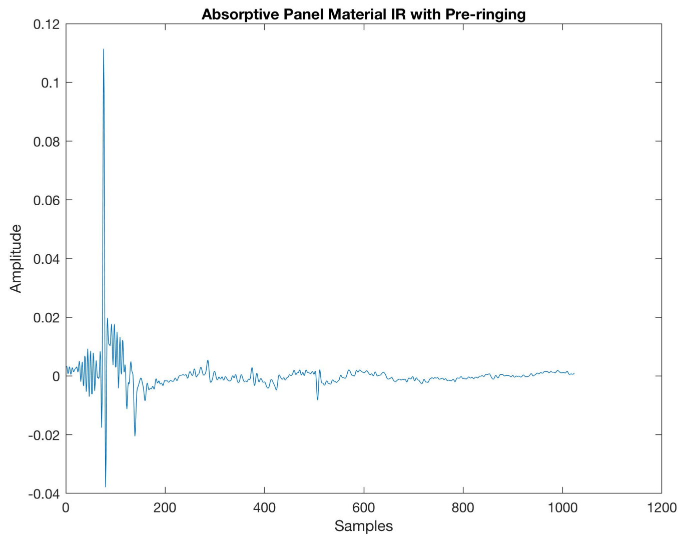


**Figure 13**

## 4.2 Minimum Phase Filter

For removing these pre-ringing artifacts and enabling the material filter to be a causal and stable system, a minimum phase filter, without changing the frequency response's magnitude spectrum, is desired. The attribute that a filter has minimum phase if and only if its cepstrum (i.e., inverse Fourier Transform of the log spectrum) is causal [11]. Hilbert transform is then implemented to the complex cepstrum to construct an analytic signal, which guarantees the causality.  Through the Hilbert transformed cepstrum approach, in time domain, the pre-ringing effects of material IR is diminished (Figure 14). Moreover, all acquired IR can also remain standardized form of phase, which is beneficial for building the IR library.

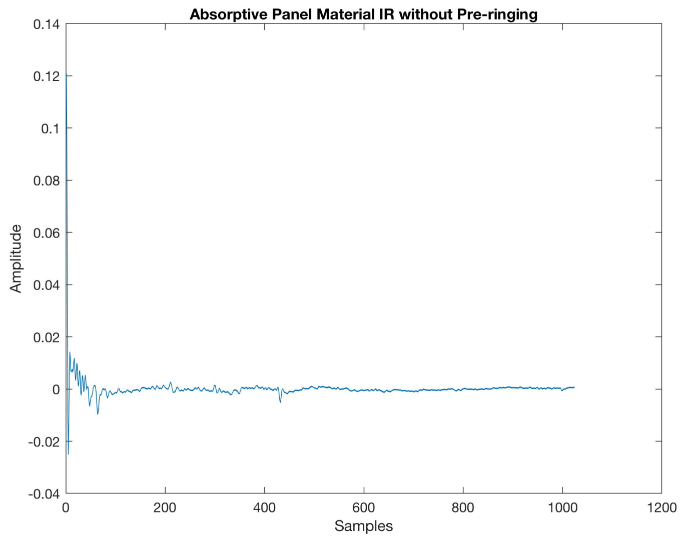
Absorptive Panel Material IR without Pre-ringing
Amplitude
Samples
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
0
200
400
600
800
1000
1200

**Figure 14**

# Chapter 5

# Future Work

In terms of this project's scope, the goal is to build a pure material IR library that could be used for convolving with, for example, the sound rays generated by a ray-tracing algorithm, or HRIR (Head-Related Impulse Response) in a virtual-reality environment. For the evaluation of methodology and development, it may be useful to compare if, for identical materials, the data from an architecture company using an ISO standard test matches our acoustic experiment results. Furthermore, synthesizing IRs based on the absorption coefficients and scattering coefficients of various materials offered by architecture companies may provide useful references for evaluating our acoustic experiment results. Subjective listening tests involving human listeners may be conducted for further result evaluation.

# References


[1] Ashok, M., King, R., Kamekawa, T., & Kim, S. (2018, May). Acoustic and Subjective Evaluation of 22.2-and 2-Channel Reproduced Sound Fields in Three Studios. In *Audio Engineering Society Convention 144*. Audio Engineering Society.

[2] Pohl, A., Drechsler, S., Stephenson, U. M., Schroder, D., Svensson, U. P., & Vorlander, M. (2012, August). openMat-an open XML-database for acoustical properties of materials and objects. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 2012, No. 5, pp. 6693-6703). Institute of Noise Control Engineering.

[3] International Organization for Standardization. (2003). *Acoustics: Measurement of Sound Absorption in a Reverberation Room*. ISO.

[4] Vorländer, M. (2007). *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer Science & Business Media.

[5] Cox, T., & d'Antonio, P. (2016). *Acoustic absorbers and diffusers: theory, design and application*. Crc Press.

[6] Farina, A. (2007, May). Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*. Audio Engineering Society.

[7] Bouchard, M., Norcross, S. G., & Soulodre, G. A. (2006, October). Inverse filtering design using a minimal-phase target function from regularization. In *Audio Engineering Society Convention 121*. Audio Engineering Society.

[8] Oppenheim, A. V., & Verghese, G. C. (2010). *Signals, Systems and Inference: Class Notes for 6.011, Introduction to Communication, Control and Signal Processing*. Chapter 11

[9] Smith, J. O. (2007). *Introduction to digital filters: with audio applications* (Vol. 2). Julius Smith.

[10] Chan, I. H. (2010). Swept Sine Chirps for Measuring Impulse Response [J]. *Power (dBVrms)*, *50*(40), 30.

[11] Maurya, A. (2017). Interference Unmixing and Estimation Technique for Improvement of Speech Separation Performance.