

SCATTERING DELAY NETWORKS WITH APERTURE SIZE CONTROL FOR
SIMULATING COUPLED VOLUME ACOUSTICS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

TİMUÇİN BERK ATALAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MODELLING AND SIMULATION

JULY 2019

Approval of the thesis:

**SCATTERING DELAY NETWORKS WITH APERTURE SIZE CONTROL
FOR SIMULATING COUPLED VOLUME ACOUSTICS**

Submitted by **TİMUÇİN BERK ATALAY** in partial fulfillment of the requirements for the degree of **Master of Science in Modelling and Simulation Department, Middle East Technical University** by,

Prof. Dr. Deniz Zeyrek Bozşahin
Dean, **Graduate School of Informatics**

Assist. Prof. Dr. Elif Sürer
Head of Department, **Modelling and Simulation, METU**

Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu
Supervisor, **Modelling and Simulation, METU**

Examining Committee Members:

Prof. Dr. Alptekin Temizel
Modelling and Simulation Department, METU

Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu
Modelling and Simulation Department, METU

Assoc. Prof. Dr. Berke Gür
Mechatronics Engineering, BAU

Assoc. Prof. Dr. Banu Günel Kılıç
Information Systems, METU

Assist. Prof. Dr. Mehmet Koray Pekerçiçi
Architecture, METU

Date: 26.07.2019



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: TİMUÇİN BERK ATALAY

Signature :

ABSTRACT

SCATTERING DELAY NETWORKS WITH APERTURE SIZE CONTROL FOR SIMULATING COUPLED VOLUME ACOUSTICS

ATALAY, TİMUÇİN BERK

M.S., Department of Modelling and Simulation

Supervisor: Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu

July 2019, 64 pages

Artificial reverberators provide a computationally viable alternative to full-scale room acoustics simulation methods when used in interactive, immersive systems. Scattering delay network (SDN) is a relatively recent artificial reverberator that allows direct parametric control over the geometry of a simulated rectangular enclosure as well the directional characteristics of the simulated sound sources and microphones. Coupled volume SDN aims to extend the SDN structure for enclosures coupled via an aperture. The extension allows independently controlling the acoustical properties of the coupled enclosures and the size of the connecting aperture. The utility of the proposed method is demonstrated by comparing the acoustical features, initial level difference and decay ratio derived from the energy decay curves calculated using the impulse responses obtained from coupled volume SDNs.

Keywords: Artificial reverberators, room acoustics simulation, delay networks, coupled volumes, double-sloped decay

ÖZ

AÇIKLIK KONTROLLÜ SAÇILIMLI GECİKME AĞLARI İLE ÇİFT HACİM AKUSTİĞİ SİMULASYONU

ATALAY, TİMUÇİN BERK

Yüksek Lisans, Çokluortam Bilişimi Bölümü
Tez Yöneticisi: Doç. Dr. Hüseyin Hacıhabiboğlu

Temmuz 2019 , 64 sayfa

Yapay çınlanımcılar, etkileşimli, kapsayıcı sistemlerde kullanıldığında, tam ölçekli oda akustiği simülasyon yöntemlerine hesaplamalı olarak uygulanabilir bir alternatif sunar. Saçılımlı gecikme ağları (SGA), bir dikdörtgen odanın geometrisi üzerinde doğrudan parametrik kontrolün yanı sıra, simüle edilmiş ses kaynaklarının ve mikrofonların yönsel karakteristiklerini de kontrol etmeye yardımcı olan nispeten yeni bir yapay çınlanımlayıcıdır. Çift hacimli SGA, bir açıklık ile birleştirilen hacimler için SGA yapısını geliştirmeyi amaçlar. Bu geliştirme, bağlı hacimlerin akustik özelliklerini ve bağlantıyı sağlayan açıklığın boyutunu bağımsız olarak kontrol etmeyi sağlar. Önerilen yöntemin faydası, çift hacimli SGA'lerden elde edilen dürtü tepkileri kullanılarak hesaplanan enerji kayıp eğrilerinden çıkarılan akustik özellikler, başlangıçtaki seviye farkı ve eksilme oranı karşılaştırılarak gösterilmiştir.

Anahtar Kelimeler: yapay yankıyımcılar, oda akustiği simülatörleri, gecikim ağları, çift hacimler, çift-eğimli eksilme



For my family...

ACKNOWLEDGMENTS

I would like to thank my advisor Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu for his support and guidance. It was due to his help and knowledge that I was able to build up my research which eventually evolved into this thesis. It was a thrilling journey to study and learn more about room acoustics.

I would also like to thank my thesis committee: Assoc. Prof. Alptekin Temizel, Assoc. Prof. Banu Günel Kılıç, Assoc. Prof. Berke Gür, Assist. Prof Mehmet Koray Pekerçli for accepting my invitation. Their helpful comments and challenging questions will undoubtedly improve my research.

I would like to thank Müller-BBM and Amplified Parts for letting me use their images in my thesis.

I would also like to thank Mr. Vatan Emen for useful discussions and preliminary work on the extension of scattering delay network models to coupled volume acoustics.

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ	v
ACKNOWLEDGMENTS	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xvi
CHAPTERS	
1 INTRODUCTION	1
1.1 Contributions	3
1.2 The Outline of the Thesis	3
2 BACKGROUND	5
2.1 Room Acoustics	5
2.1.1 Acoustics of Single Volumes	7
2.1.2 Acoustics of Coupled Volumes	8
2.2 Room Acoustics Modelling	10
2.2.1 Scale Physical Models	10
2.2.2 Image Source Method	11

2.2.3	Ray Tracing	12
2.2.4	Beam Tracing	12
2.2.5	Digital Waveguide Mesh	13
2.3	Artificial Reverberators	15
2.3.1	Analog Artificial Reverberators	15
2.3.1.1	Reverberation Chambers	15
2.3.1.2	Spring Reverberators	15
2.3.1.3	Plate Reverberators	16
2.3.2	Digital Artificial Reverberators	16
2.3.2.1	Schroeder's Reverberator	16
2.3.2.2	Moorer's Reverberator	18
2.3.2.3	FDN	19
2.3.2.4	Jot's Reverberator	19
2.3.2.5	DWN	19
2.3.2.6	SDN	20
3	SCATTERING DELAY NETWORKS	23
3.1	SDN design overview	23
3.1.1	Source to Microphone Connection	25
3.1.2	Source to Node Connections	26
3.1.3	Node to Node Connections	27
3.1.4	Node to Microphone Connections	28
3.2	Transfer Function of SDN	29
3.3	Evaluation of SDN	30

3.4	Extensions of SDN	31
4	COUPLED VOLUME SCATTERING DELAY NETWORKS	33
4.1	CV-SDN Settings	33
4.1.1	Same Volume CV-SDN Setting	34
4.1.2	Different Volume CV-SDN Setting	35
4.2	CV-SDN Design Overview	35
4.2.1	Source to Aperture Connection	36
4.2.2	Node to Aperture Connections	37
4.2.3	Aperture to Node Connections	38
4.2.4	Aperture to Microphone Connection	38
4.3	Transfer Function of CV-SDN	39
4.3.1	Transfer Function of the Same Volume CV-SDN	39
4.3.2	Transfer Function of Different Volume CV-SDN	42
5	EVALUATION AND DISCUSSION	45
5.1	Same Volume CV-SDN	45
5.1.1	Test Setup	46
5.1.2	Results and Discussion	47
5.2	Different Volume CV-SDN	52
5.2.1	Test Setup	53
5.2.2	Results and Discussion	53
6	CONCLUSION AND FUTURE WORK	57
6.1	Conclusion	57
6.2	Possible use cases of CV-SDN	58

6.3 Possible Future Work	59
REFERENCES	61



LIST OF TABLES

TABLES

Table 5.1	Aperture Opening Size Results compared to Theoretical Values . . .	47
Table 5.2	Aperture Opening Size Results compared to ODEON Values	47
Table 5.3	Absorption Coefficient Results Compared to Theoretical Values . . .	51
Table 5.4	Absorption Coefficient Results Compared to ODEON Values	51
Table 5.5	Coupled Volume Ratio Results Compared to Theoretical Values . . .	51
Table 5.6	Coupled Volume Ratio Results Compared to ODEON Values	53
Table 5.7	Different volume CV-SDN EDT Parameter comparison	56

LIST OF FIGURES

FIGURES

Figure 2.1	A room impulse response (RIR) graphic	6
Figure 2.2	EDC of a simple rectangular volume.	8
Figure 2.3	Convex EDC of a coupled volume structure where source and receiver are placed in separate volumes.	9
Figure 2.4	Concave EDC of a coupled volume structure where source and receiver are placed in the same volume.	10
Figure 2.5	Decay Ratio and ΔdB properties.	11
Figure 2.6	Scale model of a chamber music hall.	12
Figure 2.7	Image Source Method.	13
Figure 2.8	Raytracing method.	14
Figure 2.9	DWM junctions with incoming and outgoing wave variables	14
Figure 2.10	A Reverberation Chamber.	16
Figure 2.11	Spring reverberator.	17
Figure 2.12	Plate reverberator with two pickups producing stereo output.	18
Figure 2.13	Comb Filter.	18
Figure 2.14	All-pass filter (All-pass Reverberator).	19
Figure 2.15	Moorer's Reverberator.	20

Figure 2.16	FDN reverberator.	21
Figure 2.17	Jot's reverberator.	21
Figure 3.1	2D representation of SDN.	24
Figure 3.2	Source to Microphone Connection.	25
Figure 3.3	Source to Node Connections.	27
Figure 3.4	Node to Node Connections.	27
Figure 3.5	Node to Microphone Connections.	28
Figure 3.6	Transfer function of SDN.	29
Figure 3.7	Room impulse response of SDN.	30
Figure 3.8	EDC of SDN.	31
Figure 4.1	2D representation of same volume CV-SDN.	35
Figure 4.2	2D representation of different volume CV-SDN.	36
Figure 4.3	Source to Aperture Connection.	36
Figure 4.4	Node to Aperture Connections.	37
Figure 4.5	Aperture to Node Connections.	38
Figure 4.6	Aperture to Microphone Connection.	38
Figure 4.7	Same Volume CV-SDN System Diagram.	40
Figure 4.8	Different Volume CV-SDN System Diagram.	42
Figure 5.1	Test Setup of the same volume CV-SDN with (a) showing 3D and (b) shown 2D representation.	48
Figure 5.2	Same Volume CV-SDN EDC results with test setup.	49
Figure 5.3	Same Volume CV-SDN Absorption EDC results with test setup.	50

Figure 5.4	Same Volume CV-SDN Coupled Volume Ratio EDC results with test setup.	52
Figure 5.5	Test Setup of the different volume CV-SDN with (a) showing 3D and (b) shown 2D representation.	54
Figure 5.6	Different volume CV-SDN EDC result with test setup.	55
Figure 5.7	Different volume CV-SDN EDC result with different aperture openings.	56



LIST OF ABBREVIATIONS

GPU	Graphics Processing Unit
LTI	Linear Time Invariant
DSP	Digital Signal Processing
RIR	Room Impulse Response
MLS	Maximum Length Sequences
IRS	Inverse Repeated Sequence
RT60	Reverberation Time
EDT	Early Decay Time
LDT	Late Decay Time
DSD	Double-Sloped Decay
EDC	Energy Decay Curve
ISM	Image Source Method
DWN	Digital Waveguide Network
DWM	Digital Waveguide Mesh
FDN	Feedback Delay Network
SDN	Scattering Delay Networks
CV-SDN	Coupled Volume Scattering Delay Networks
CW	Common Wall
DR	Decay Ratio
XR	Extended Reality

CHAPTER 1

INTRODUCTION

Simulating acoustics of enclosures is a typically costly operation which is not always appropriate in games and extended reality applications. More typically, direct path and early specular reflections are calculated and the late reverberation is added via algorithms, called artificial reverberators, which provide perceptual (as opposed to physical) acuity, at a minor fraction of the computational cost of full-scale room acoustic simulations [1, 2, 3].

Direct sound, early and late reflections constitute the acoustic response of a room. The response is called room impulse response (RIR) and can be observed in interactive environments such as games and VR by generation a short burst of noise like a gunshot or a clap. Generation and manipulation of RIR are done with two types of artificial reverberators. Analog artificial reverberators make use of mechanical or electronic apparatus in the real world. Some of the examples in the field are reverberation chambers, spring reverberators, and plate reverberators. Analog reverberators can be useful in certain respects, however, they have several drawbacks. First of all, acquiring equipment with the addition of creating raw audio is costly. Also, it is a challenging task to arrange analog reverberators to have the desired RIR. Finally, recording equipment of good quality should be present to capture the reverberated output, which is not often the case.

Digital reverberators, on the other hand, make use of digital signal processing algorithms to generate RIRs. The use of digital artificial reverberators brings many possibilities in video games, extended reality applications, and acoustical design and prototyping of architectural structures such as concert halls. These reverberators typically aim to capture the most prominent perceptual features of room acoustics. Earliest examples used two different mechanisms to simulate the reflection and modal properties of the acoustics of rooms by combining comb filters to simulate the frequency modes of the room response and allpass filters to simulate its diffuse characteristics [4, 5]. This idea was later generalized to multiple channels [6] and subsequently reformulated to allow parametric control over different aspects of reverberation [7]. Despite their computational efficiency, these artificial reverberators do not embody any direct physical correspondence with the acoustics of a given room geometry. Borrowing the idea of a scattering junction from digital waveguide mesh (DWM) models [8, 9, 10], Smith proposed a closed network of scattering junctions that provided the blueprint for circulant artificial reverberators [11]. This model was later formulated by Karjalainen *et al.* into an artificial reverberator model that aimed to accurately simulate axial modes of a room [12].

Overall, there is a trade-off between accuracy and computational complexity in digital reverberators. When generating RIRs, late reflections are generally approximated since it is not possible to hear the difference with the human ear. However, with the use of GPUs and increased processing power of computers, digital reverberators have room for improvement.

Scattering delay network (SDN) is a type of digital artificial reverberator which combines elements of geometric room acoustics with computationally efficient delay networks to provide an impulse response which has a one-to-one correspondence with the room geometry [13]. SDN was later reformulated for frequency-domain operation [14] and analyses of its performance and stability were presented in [15]. The major advantage of SDN is that it can accurately provide perceptually important features of the room response such as the direct path and first-order reflections and has a gracefully degrading accuracy for higher order reflections and reverberation. It can provide a very natural sounding room acoustics simulation and can run at interactive rates, which makes it suitable for extended reality and other interactive applications. Very recently, SDN was used for augmented reality applications [16, 17]. Moreover, SDN is shown to provide levels of naturalness and pleasantness that are comparable to full-scale room acoustic simulators [18]. SDN has also expanded to an artificial reverberator called Waveguide Web to include more accurate second-order reflections [19].

Coupled volumes occur in real life as rooms connected through an aperture such as a door or a window and their energy decay curves (EDCs) have distinct properties that depend on the individual decay characteristics of the volumes themselves as well as the size of the aperture in between them. Single volumes (i.e. rooms) have a constant sloped EDCs, whereas coupled volumes exhibit much different behavior where EDC consists of two different slopes. We will, henceforth, call this behavior, double sloped decay (DSD). Coupled volumes also frequently occur in virtual environments such as immersive 3D games, mixed and virtual reality applications, and concert halls. One type of DSD pattern (i.e. Concave DSD) is quite distinguishable in concert halls [20]. Hence, an interactive and customizable artificial reverberator which can simulate a coupled volume case could be used in computer applications due to its ease of use. Moreover, the results observed with such reverberators could be used to examine or even design and prototype architectural structures like concert halls.

As SDN is limited as it only performs on shoebox type of rooms that have six sides, the extension of the model to coupled volumes is required since there are many mixed reality and video game applications where more than one room is utilized. Therefore, it can be beneficial to analyze the basic properties of SDN algorithm to generate a model which is capable of simulating the acoustics of coupled volumes.

This thesis proposes the coupled-volume SDN (CV-SDN) as an extension to the original SDN method. CV-SDN allows 1) the simulation of the acoustics of coupled volumes, and 2) the parametric control of the aperture size between them. CV-SDN can replicate the behavior of acoustical characteristics of coupled volumes by using several SDN models and an aperture model whose size can be adjusted. Two settings of CV-SDN is possible depending on the positions of the sound source and the receiver. Both settings can emulate the DSD phenomena observed with real coupled rooms. Mathematical representation of CV-SDN can be easily generated as SDN's

transfer function is clearly defined. Moreover, SDN being a stable system allows CV-SDN to be stable as well.

1.1 Contributions

The contributions of CV-SDN are as follows:

- An extension of SDN to two different settings of coupled volumes with the addition of an aperture connecting the two volumes.
- Mathematical representation of the proposed structure as well as the energy transmission modeling of the aperture.
- Same Volume CV-SDN setting which simulates the case where source and microphone are placed in the same volume. The model is preferred for its energy decay characteristics, which same volume CV-SDN can generate. The setting is suitable for the initial design and prototype of architectural structures like concert halls.
- Different Volume CV-SDN setting which simulates the case when source and microphone are placed in the separate volumes. Although the setting is not widely studied due to energy decay characteristics, it can generate outputs on par with the studies in the field.
- The study is under review for publication in Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA-19), Oct. 2019, New Paltz, New York, USA.

1.2 The Outline of the Thesis

The thesis is organized as follows. Background is given in Chapter 2. Room acoustics, studies in coupled volumes, and both analog and digital artificial reverberators are thoroughly discussed. Chapter 3 explains the Scattering Delay Networks in detail as it forms the backbone of CV-SDN. In Chapter 4, the proposed CV-SDN reverberator is defined. Two different settings of CV-SDN are proposed and a transfer function is derived for each one. Additionally, the stability of both settings is analyzed. Chapter 5 first discusses the test setups used to evaluate both settings of CV-SDN. Obtained results of same volume CV-SDN are compared with the theoretical values and another computational acoustics simulator in the field. The results of different volume CV-SDN is compared with the scale physical model in another study. Finally, Chapter 6 concludes the thesis and discusses the possible future work of CV-SDN.



CHAPTER 2

BACKGROUND

In this chapter, room acoustics and the differences in the single and coupled volume acoustics are discussed. Additionally, background on the history of room acoustics modeling methods and artificial reverberators are given.

2.1 Room Acoustics

The acoustics of a room is determined by several parameters such as, the geometry of the room, the absorption of the walls which scales the energy of sound during reflection, the response of the room to the reverberation of a single unit impulse signal.

When a sound wave impinges on absorptive surfaces, part of its energy is absorbed and another part is reflected. The reflecting wave is specularly and/or diffusely reflected from the wall. The reflected amount of energy is represented with the wall reflection coefficient defined as:

$$\beta = \sqrt{1 - \alpha}, \quad (2.1.1)$$

where α is the absorption coefficient representing the fraction of energy absorbed by the surface and depends on the material of the surface.

$1/R$ distance attenuation law is also an important factor in room acoustics. Depending on the distance between two points in space, attenuation causes the reduction in the amplitude of the sound signal. The attenuation coefficient can be set as:

$$g_{S,M} = \frac{1}{\|\mathbf{x}_S - \mathbf{x}_M\|}, \quad (2.1.2)$$

where \mathbf{x}_S and \mathbf{x}_M expresses the positions of the source and the receiver (i.e. microphone).

A room is generally defined as a linear and time-invariant (LTI) system from a signal processing point of view. In other words, it can be thought of as a system which receives input from a sound source and generates the reverberated sound at a measurement point. Hence, the acoustics of a room is defined by its impulse response. More clearly, it is referred to as the room impulse response (RIR).

There are three parts of an RIR as shown in Fig. 2.1:

1. Direct Sound corresponds to the first signal received at the measurement point. It is only subject to distance attenuation. Therefore, no reflections are considered. In Fig. 2.1 it is the first impulse in the plot. The figure is normalized to have a maximum value of 1.
2. The second part of RIR is the early reflections. They represent the first-order reflection by the walls. Early reflections are comparable to the direct sound impulse as they are only affected by a single reflection and the distance attenuation.
3. The final part is called the Late Reverberation and it is the tail of the RIR. It represents the exponential decay of the impulse signal.

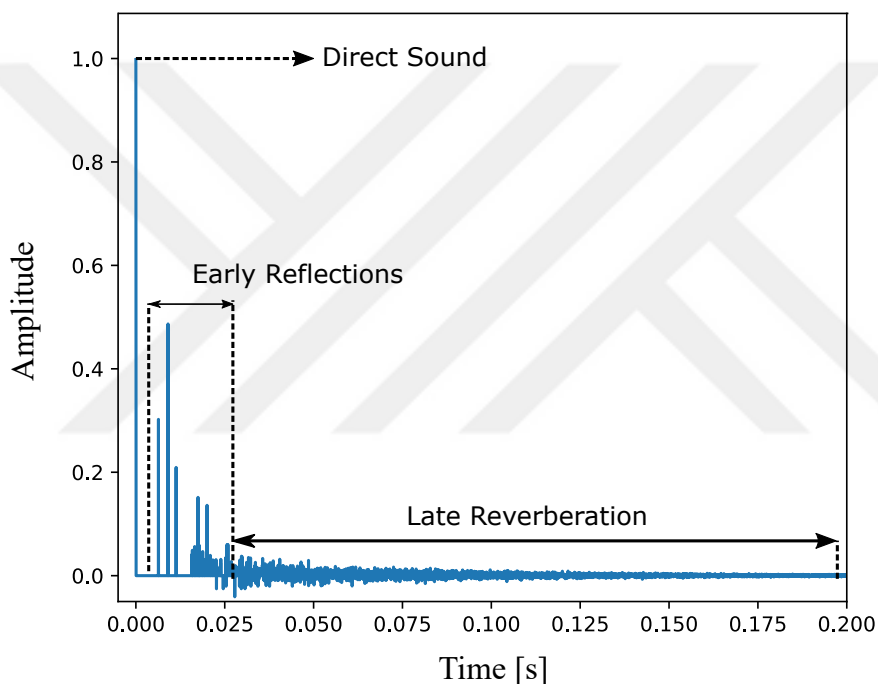


Figure 2.1: A room impulse response (RIR) graphic

There are two ways of obtaining RIRs.

Recording where an impulsive source is used as excitation is the first way. Claps, balloons, and pistols are examples of impulsive sounds. This method is easily applied through the recording environment and equipment may affect the results unintentionally. For example, the directivity of the source can change the output unexpectedly.

Secondly, deterministic measurement signals could be used. There are several techniques in this field. The common ones in the field are Maximum Length Sequences (MLS), Inverse Repeated Sequence (IRS), and Time-Stretched pulses and (linear or logarithmic) sine sweep techniques (linear or logarithmic) sine sweep techniques [21].

Energy Decay Curve (EDC) is used for assessing the energy decay behavior of the RIR. EDC is the tail integral of squared impulse responses [22] and can be used for finding the reverberation time (RT) of a room. It can be defined as:

$$EDC(t) = 10 \log_{10} \left(\frac{\int_t^{\infty} |h(t)|^2 dt}{\int_0^{\infty} |h(t)|^2 dt} \right) \quad (2.1.3)$$

where $h(t)$ represents the room impulse response (RIR). EDC can be used to analyze the time dependence of energy decay in rooms [23].

Reverberation time is the time delay at which EDC falls a certain amount below its initial value. RIR values after a certain amount are assumed as noises. The most common measure for reverberation time is RT_{60} , which tells us the time at which EDC falls 60 dB below its initial value.

2.1.1 Acoustics of Single Volumes

Single volumes (i.e. rooms) have been the first testbed of major concepts in room acoustics. Most of the theories are based on the acoustics of a single room since they are easy to test due to their structure and provides a baseline from which more complex models can be created. Although acoustics was an interesting subject for philosophers such as Pythagoras and Aristotle [24] and continued to develop over the history of time, the major findings of the acoustics of single rooms in last century has started with Sabine's well known RT_{60} formula in 1922 [25]. The formula is still used for reverberant rooms given as:

$$RT_{60} = \frac{0.161V}{\sum_i A_i \alpha_i}, \quad (2.1.4)$$

where V is the total volume of the room, A_i is the surface of the wall i , and α_i is the absorption coefficient of the wall i . Although the formula presents an easy way of calculating the RT_{60} , it is not reliable for rooms with a short reverberation time. In order to cover such rooms, Eyring introduced a new formula in 1930 [26]. Eyring's formula for reverberation time is:

$$RT_{60} = \frac{0.161V}{\left(\sum_i A_i \right) \ln \left(1 - \left(\sum_i A_i \alpha_i / \sum_i A_i \right) \right)}. \quad (2.1.5)$$

The preliminary studies in the field show that EDC of a single volume is linear excluding the discrete samples in the first few milliseconds. Fig. 2.2 shows the EDC of a single volume. The rate of decay observed on the graph is directly related to the absorption coefficient of the walls of the room, and directivity of the source and microphone.

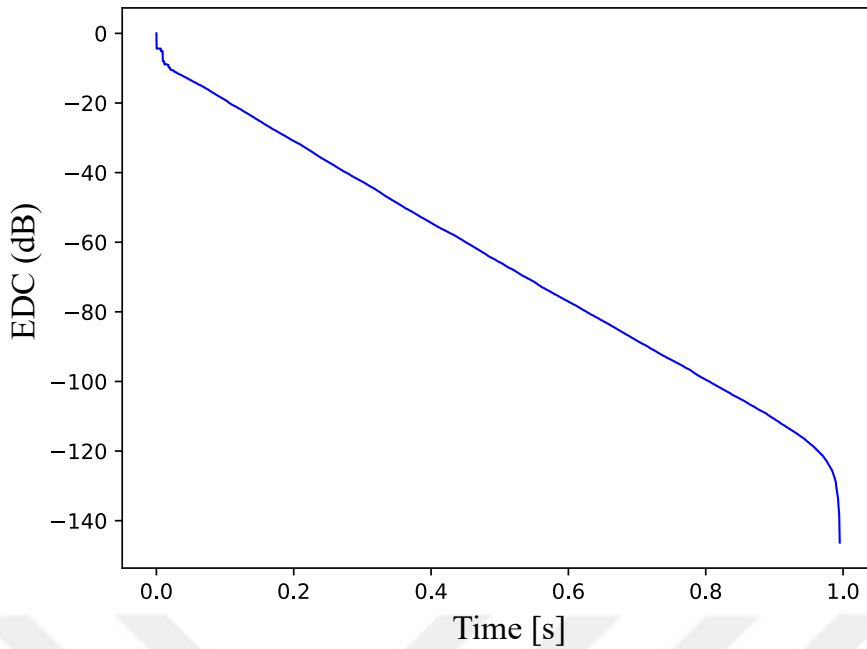


Figure 2.2: EDC of a simple rectangular volume.

2.1.2 Acoustics of Coupled Volumes

Coupled Volumes consist of multiple volumes (i.e. rooms) which are connected through apertures. Coupled volumes behave differently than single volumes. The main difference can be observed by examining the EDC of the coupled volumes.

As mentioned in the previous section, a simple rectangular enclosure has a linear decay (i.e. EDC has a single slope). When two volumes with different reverberation characteristics are coupled via an aperture, the energy decay characteristics of the combined acoustical system becomes different from either one of the volumes and two distinct decay rates can be observed [27]. This phenomenon is called *double-sloped decay* (DSD). There are two different types of DSD profiles. Convex DSD has a slow early decay rate and a fast late decay rate. Concave DSD, on the other hand, starts with a rapid decay rate in the initial stages where the slope is negative and slows down in the tail [23].

Type of DSD depends on the placements of the source and receiver. Although some observations show that a demarcation line which determines where Convex and Concave DSD occur exists near the aperture [28], as long as source and receiver are not close to aperture, positioning determines the DSD phenomena. Convex DSD is observed when source and receiver (e.g. microphone) are placed in the different volumes. Fig. 2.3 shows the convex DSD phenomena. In this case, the source is placed in a volume with lower reverberation. Whereas the receiver is in a volume with a higher reverberation. Concave DSD occurs when the source and the receiver are placed in the volume that has the lower reverberation [28]. In Fig. 2.4, concave DSD phenomena is given. In both cases, late decay corresponds to the decay characteristics of the coupled volume, whereas early decay to the main volume.

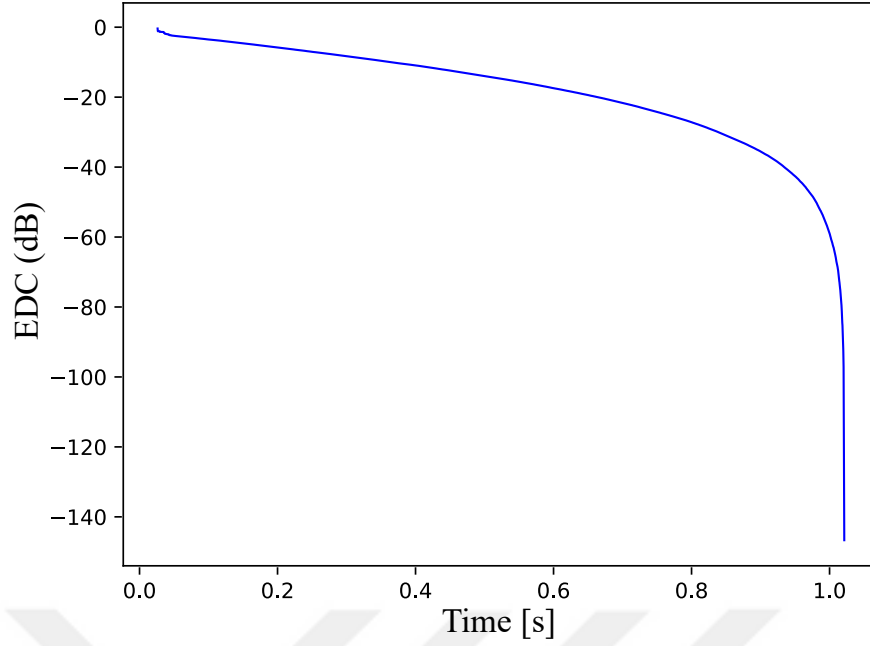


Figure 2.3: Convex EDC of a coupled volume structure where source and receiver are placed in separate volumes.

Aperture size also affects the properties of the specific decay pattern [29]. The transition time from one slope to the other shifts with the relative size of the aperture. More clearly, with a larger aperture size the energy transmitted from the coupled volume increases. This results in the coupled volume (i.e. the volume that determines the tail slope) dominating EDC at an earlier rate.

The relation of the aperture size to the coupled volume acoustics becomes evident when acoustics of concert halls, theatres, and opera houses are considered. For such venues, listeners typically prefer low and middle levels of DSD [20]. To study concave DSD, parameters such as early decay time (EDT), late decay time (LDT), and coupling coefficient ratio are used [29, 30]. EDT is the time of the first 10 dB drop. LDT is the decay time between 25 dB and 35 dB below the initial level. Coupling coefficient is defined as the ratio of T_{30}/T_{15} . T_{30} represents the decay time between -5 dB and -35 dB, whereas T_{15} represents the decay time from -5 dB to -20 dB. Bradley and Wang [31] introduced two additional parameters, the decay ratio, and ΔdB which are used in this thesis to compare the proposed system with theoretical values. Decay ratio refers to the ratio of the slope of the early part of the EDC to the slope of the late part of the EDC. ΔdB refers to the sound pressure level difference between the initial levels of the two single-slope EDC profiles that make up the DSD profile. Fig. 2.5 shows a visual description of Decay Ratio and ΔdB .

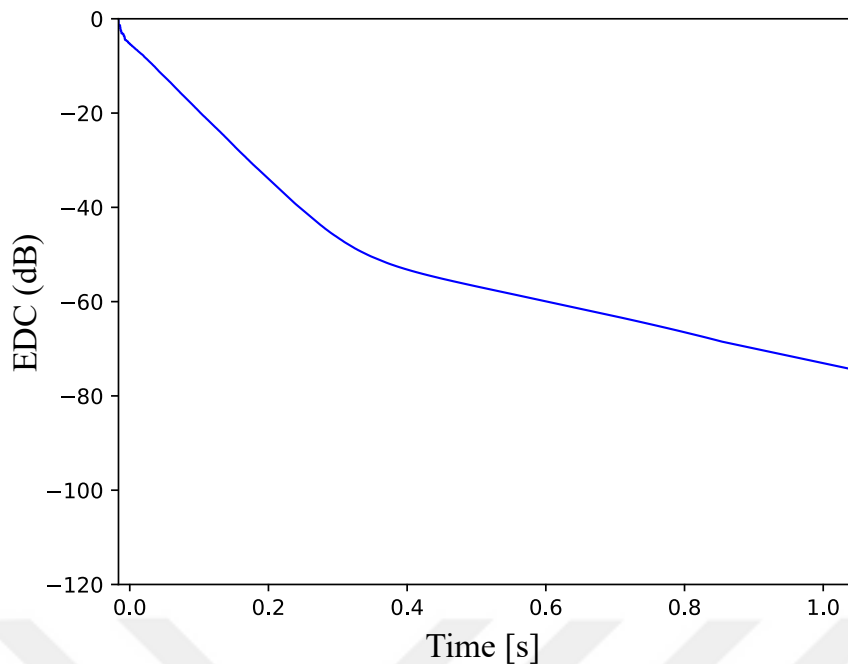


Figure 2.4: Concave EDC of a coupled volume structure where source and receiver are placed in the same volume.

2.2 Room Acoustics Modelling

The acoustics of a room can be modeled with its geometrical properties by numerical methods. There are several methods of room acoustics modeling explained in this section.

2.2.1 Scale Physical Models

Compared to the rest of the methods, scale physical model of a room is a physical one. It is becoming outdated with improvements in digital methods. The aim is to construct a smaller room which has the same geometry as the original enclosure. In Fig. 2.6¹ scale physical model of a chamber hall is given. Reverberation is predicted by employing ultrasound wave method as these waves have short wavelengths with high frequencies. In other words, all components of the room acoustics, including the wavelength are scaled down.

Instead of ultrasound wave method, lighting can be used [23]. Depending on whether objects receive light or the intensity of light reflected upon the objects, the acoustics of the hall can be predicted. Correlation between the absorption of the light capability of material with the reflection coefficient of sound should be made carefully to generate the acoustics of the hall. Nevertheless, this is an unusual method.

¹ COPYRIGHT of Müller-BBM GmbH

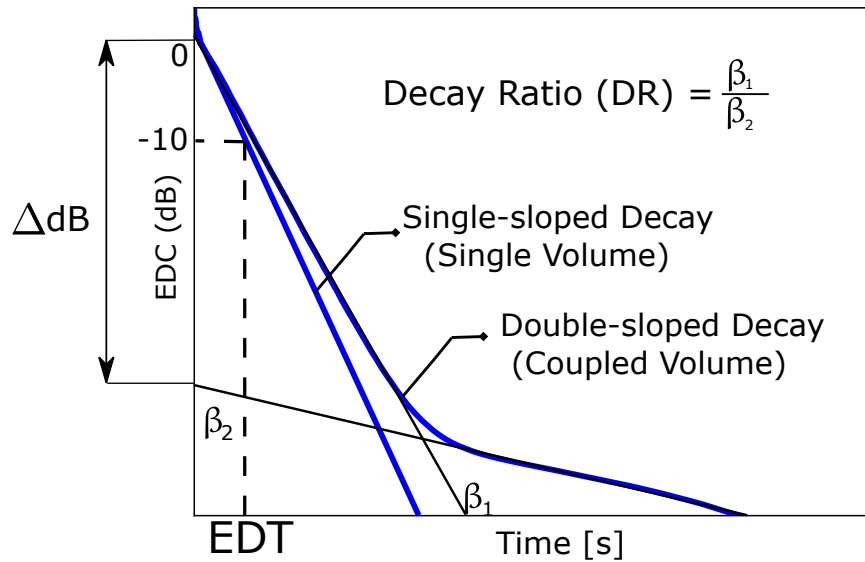


Figure 2.5: Decay Ratio and ΔdB properties.

2.2.2 Image Source Method

Image Source Method (ISM) is a simple geometric acoustics method [32]. Relative to the position of the sound source, reflections are placed as mirror image sources. In the method, it is assumed that walls reflect incident sound just like a mirror reflects incident light rays. More clearly, reflective surfaces are considered as a mirror. By superposing contributions from image sources (i.e. virtual sources) to a measurement point, calculation of the RIR can be performed.

The source images obtained from the first reflections are easy to calculate; however, higher order reflections are calculated with first-order image sources. Therefore, it becomes much more complex as each image source generates additional image sources. This makes simulating late reverberation challenging as the computational cost of the method increases exponentially [33]. If the model is designed to be run in real-time, adjusting source position adds additional cost as all the images will be recalculated as well. Hence, calculating higher order reflections using ISM is not suitable for games and extended reality applications. Computational cost problem can be overcome either by pre-computing higher order image sources or omitting them. Nevertheless, pre-computation may produce unnecessarily large data.

Borish extended the ISM for arbitrary room shapes in 1984 [34]. The computational cost of ISM increases dramatically as visibility and validity check of image sources are considered as well as their positions. Visibility checks are employed to assure that an image source is visible by the measurement point. Validity check ensures that an image source is created from the reflective surface of a wall.

A suitable model for ISM can be a rectangular room where only first reflections are considered. Fig. 2.7 shows a simple model which has a low computational cost. A source and a measurement point are considered with walls as reflective surfaces creating image sources. The direct path between source and measurement point cor-



Figure 2.6: Scale model of a chamber music hall.

responds to direct sound. Paths between image sources and measure point contribute to early reverberation. Although higher order image sources are not shown, they can be computed with the current image sources, the contribution of which corresponds to late reverberation. In the figure, only four images are drawn as the figure represents the 2D case. However, there is a total of six first-order-images.

2.2.3 Ray Tracing

Ray tracing is a geometric room acoustic method (i.e. utilizes the geometry of the room to calculate the reflections of the sound) [35] used in optics, acoustics, and computer graphics. The method can compute both specular and diffuse reflections. It was shown by M. R. Schroeder that reverberation time of ray tracing is comparable to Sabine's and Eyring's reverberation time formulas [36].

Fig. 2.8 shows the paths from the source to a measurement point. Only one of the rays connected to the measurement point is displayed. The number of rays cast from the source can be pre-defined. By weighing the amplitude of rays, one can account for the directivity if the source is not omnidirectional. To have a more accurate model, it is better to use an increased number of rays so. However, as each ray is computed during the operation, the method's computational cost might become prohibitively high depending on the hardware used.

2.2.4 Beam Tracing

Beam tracing is an extension of ray tracing to solve the computational cost problem. Beams of rays are used in this method [37]. The model has advantages over image

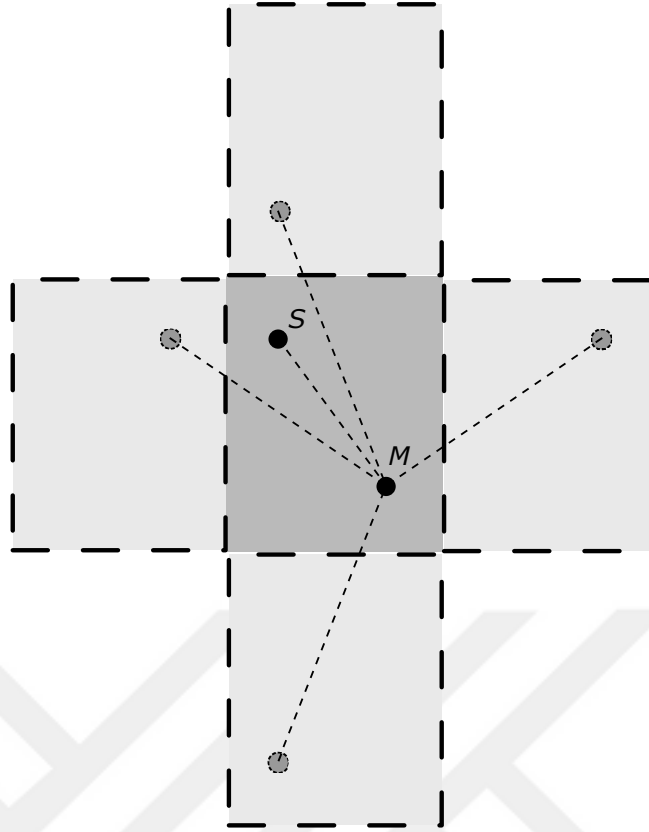


Figure 2.7: Image Source Method.

source and ray tracing methods. For example, in the ray tracing method, not all calculated rays intersect with the measurement point, hence some of the calculations are often obsolete. Thus, instead of ray tracing, beam rays are used in volumes overcome the computational cost problem. However, the beam tracing method is disadvantageous for rooms with curved surfaces and indents as it will be too difficult to trace the beam from the source to the measurement point.

2.2.5 Digital Waveguide Mesh

Digital Waveguide Mesh (DWM) is a room acoustics simulation model which uses bidirectional delay lines (i.e. branches) between junctions (i.e. nodes) [11]. Hence, DWM consists of junctions and lossless branches. In each junction, pressure from incoming branches is scattered to outgoing branches. The scattering operation is performed by the scattering matrix. A bidirectional branch connects the nodes and is called waveguide. DWM can solve the acoustic wave equation in time domain.

Two DWM junctions can be observed in Fig. 2.9. DWM is performed in two steps, the scattering and propagation steps. The scattering operation is performed between two junctions and bidirectional delay lines are used. The incoming pressure coming from DWM junction i is indicated as: p_{ij}^+ , where $i = 1, 2, \dots, N$ and N represents the number of junctions connected to junction i . After a certain delay, which is indicated

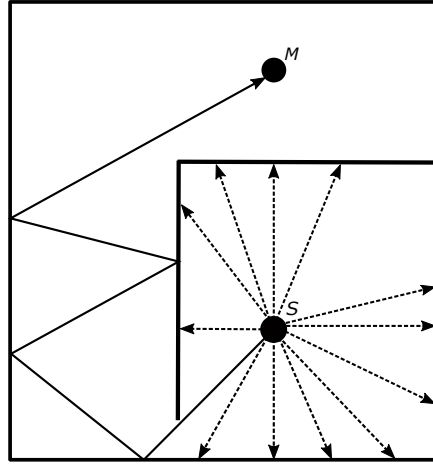


Figure 2.8: Raytracing method.

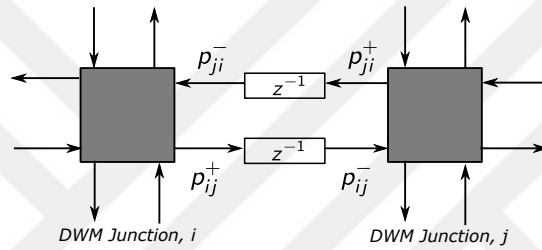


Figure 2.9: DWM junctions with incoming and outgoing wave variables

by z^{-1} between the nodes, outgoing pressure, p_{ij}^- to DWM junction j is received. In the propagation operation, waves are fed into bidirectional delay lines for scattering operation. During this process, pressure in any node is scaled by $2/N$, where N is the number of nodes in the network. Overall, as DWN junctions are lossless, a scattering matrix A is defined such that:

$$\mathbf{p}_{ij}^- = \mathbf{A} \mathbf{p}_{ij}^+. \quad (2.2.1)$$

Eq. 2.2.1 is the basis of Scattering Delay Networks, which will be explained in detail within Chapter 3.

The computational cost of the model increases proportionally with the spatial and temporal sampling rates as well as the room volume. Unless the computation is accelerated, for example with parallel computing, DWM is not a feasible option for complex real-time models. However, if the model is relatively simple, and there are some approximations made during calculations, DWM can be a viable option. Some of the artificial reverberators, including scattering delay networks (SDN), is constructed upon the idea of DWM.

2.3 Artificial Reverberators

Artificial reverberators have been used to simulate room effect on the raw sound. There are two types of artificial reverberators: (i) analog and (ii) digital artificial reverberators. Analog reverberation is achieved by mechanic contraptions and vibrations. Analog reverberators were used at the early 60s with reverberation chambers being the first amongst them.

Digital artificial reverberators allow the manipulation of parameters that can define the RIR. Digital artificial reverberators do not aim to replicate reverberation physically. Instead, generated synthetic effect is created which perceptually resembles the sound in a real room. Some approximations are necessary in the model to reduce the computational cost. This is done as long as the perceptual quality of the sound is kept at an acceptable level.

2.3.1 Analog Artificial Reverberators

Analog artificial reverberators mostly use physical methods such as electronic devices or mechanic apparatus to achieve reverberation operation. There are three types of analog artificial reverberators.

2.3.1.1 Reverberation Chambers

Reverberation chambers are rooms with reflective plates (i.e. reflectors) placed in them as shown in Fig. 2.10. These chambers are used since the beginning of 1920s [38, 39]. A source is placed in an acoustic space and the response is recorded with a microphone. Convex reflectors can be used the room more reverberant by allowing sound to scatter. The result gained by the reverberation chamber produces little or no artifacts. Since a direct replication of a reverberant room is not easy, reverberation chambers are used only to add reverberation effect on raw audio. Moreover, constructing a reverberation chamber is pricey and it is not preferred when portability is concerned.

2.3.1.2 Spring Reverberators

Spring reverberators achieve the reverberation effect with the vibration of springs. They were invented by Hammond [40]. Inside a spring reverberator, input and output transducers with several transmission springs exist. Fig. 2.11² shows the drawing of a spring reverberator with three transmission strings. Due to the excitation of the input transducer, magnetic field initiates the vibrations that, in turn, move the springs. Then, springs excite the output transducer by vibrating. The output transducer then picks up the transmission and converts them to current to be processed as output [41].

² Image courtesy of Amplified Parts (www.amplifiedparts.com)



Figure 2.10: A Reverberation Chamber.

2.3.1.3 Plate Reverberators

Plate reverberators were introduced in late 1950s [42]. Plate reverberators function by the excitation of a metal plate by a transducer. Reverberation time is controlled with the use of a damping pad. Reverberator is held onto a frame and audio signal is used to feed the driver. The created vibrations are two-dimensional as opposed to spring reverberators. The output of the system (mono) can be gathered by a pickup and is comparable to the reverberation of a rectangular room. To have stereo output, two pick-ups can be utilized such as in Fig. 2.12. Although plate reverberators are useful in producing two-dimensional vibration, they are less practical compared to spring reverberators. Different materials for the plate reverberators were used later on. For example, instead of steel, gold is used to reduce the speed of sound. This change resulted in a reverberated sound that resembles a bigger room.

2.3.2 Digital Artificial Reverberators

Digital artificial reverberators aim to process anechoic sound and generate reverberated output. In doing so, they also need to allow the adjustments of the decay characteristics of the model. Hence, RIR is represented as accurate as possible.

2.3.2.1 Schroeder's Reverberator

The first digital reverberation algorithm has been developed by Schroeder in 1961 [4, 5]. Schroeder used comb filters and all-pass filters. Comb filters represent the frequency modes of the room. Allpass filters represent the time-domain decay characteristics. Although the output generated by the proposed algorithm is not natural

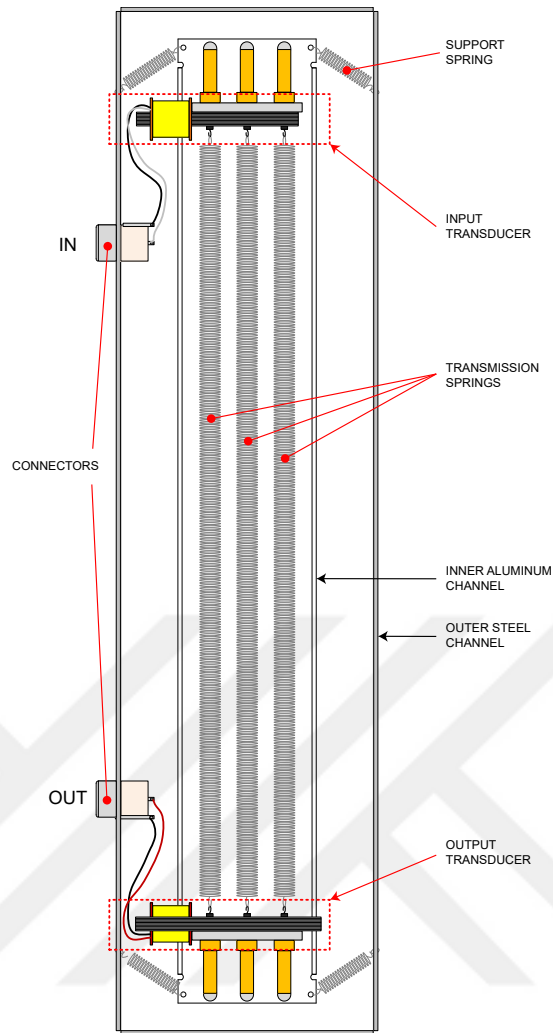


Figure 2.11: Spring reverberator.

sounding, its low computational cost has been crucial in terms of digital artificial reverberator development.

Comb filters are constructed with adding a delay of a signal to itself. The resulting output includes both the original signal component and the delayed version with its amplitude scaled. As seen in Fig. 2.13, the delay is used as a feedback loop with a gain g , that must be lower than 1 to ensure stability. The resulting frequency response is not flat however, which is not the expected response for a reverberator. Schroeder suggested that by employing several feedback loops, it is possible to obtain a flat frequency response. Thus, the all-pass filter can be used to produce artificial reverberation effect.

The resulting all-pass reverberator is not sufficient still. The reason is that it does not generate a suitable echo density as it is not aperiodic and the time interval between two reflections is not low enough. To overcome this problem, several all-pass filters are connected in series. This operation allows the combined frequency response to become flat while increasing echo density and making it aperiodic.

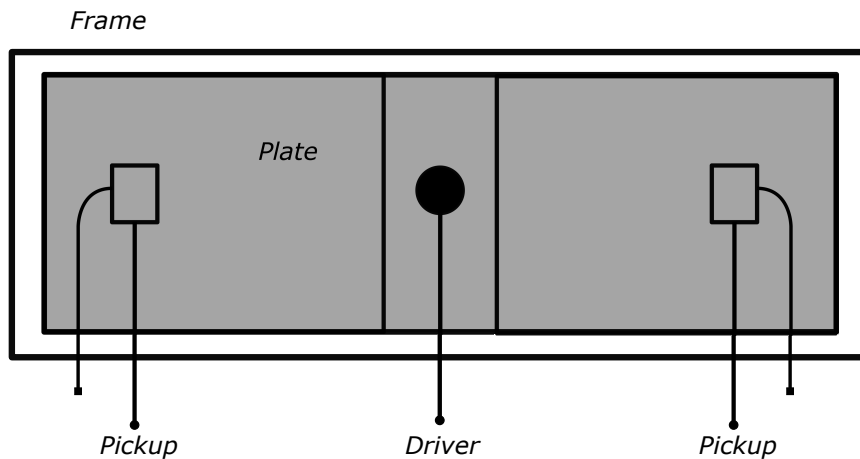


Figure 2.12: Plate reverberator with two pickups producing stereo output.

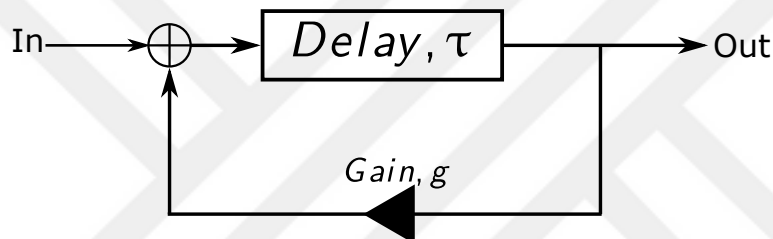


Figure 2.13: Comb Filter.

2.3.2.2 Moorer's Reverberator

Moorer's reverberator advances the work of Schroeder [43]. Similar to Schroeder, Moorer employed comb and all-pass filters. Additional delay lines whose lengths are determined by the positions of the walls are included with these filters. Moreover, Moorer added low-pass filters to delay-loops. Due to improvements, Moorer's model can simulate the early reflections compared to Schroeder's. Overall, six comb-filters are used in parallel with first-order low pass filters on delay loops.

Fig. 2.15 displays the Moorer's reverberator. In the figure, z^{-C_i} denotes the i^{th} comb filter. Low pass filters are employed on the delay loops.

Although the acoustics of a concert hall is not achieved perceptually, the output of Moorer's reverberator is substantially free of artifacts and thus an improvement on Schroeder's reverberator. Moorer suggested that enhancements on the model could be made by modeling the geometry of concert halls, employing additional filters, and lowering computational cost.

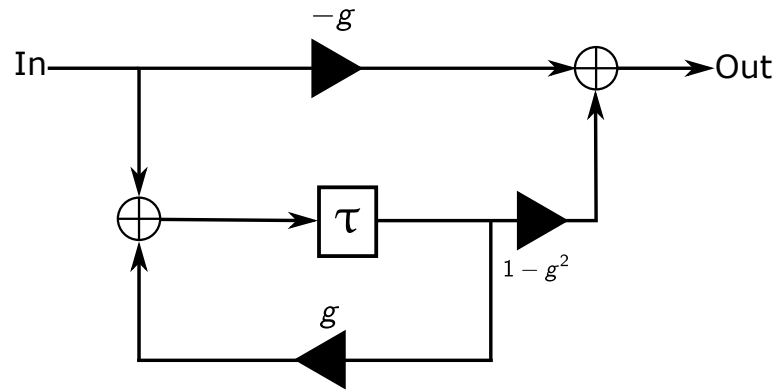


Figure 2.14: All-pass filter (All-pass Reverberator).

2.3.2.3 FDN

FDN is an artificial multichannel reverberator proposed by Stautner and Puckette with four channels [6]. A multi-channel input signal is fed into parallel delay lines that connect to the corresponding output channel. Additionally, delay lines are recursively connected over a feedback matrix. The feedback matrix is unitary (i.e. lossless) so that the system preserves the energy and is stable. Essentially, FDN can be regarded as a reverberation algorithm that is the linear combination of several comb-filters. Connection of low-pass filters to the delay lines allows frequency dependent reverberation.

Fig. 2.16 shows a four channel FDN reverberator. Each input, $x_i(n)$ is fed into a delay line with delay, z^{-M_i} and a $N \times N$ feedback matrix A completes the feedback loop.

2.3.2.4 Jot's Reverberator

Jot's reverberator is a derivative of FDN which uses only a single channel for both input and output [7]. The reverberator is shown in Fig. 2.17 and employs absorptive filters to model air absorption. Hence, the model allows the adjustment of reverberation time with respect to frequency unlike the previous ones proposed by Schroeder and Moorer. Jot's reverberator is widely used in the industry and used as the comparison for the later reverberators.

2.3.2.5 DWN

Digital Waveguide Network consists of numerous DWM nodes which are interconnected with bidirectional delay lines [11]. As shown in Fig. 2.9, each node can act as an input and output. A DWN can consist of many DWM to simulate different types of room. As the network gets more complex with an increasing number of nodes, simple geometries are preferred to construct DWN. Therefore, a shoebox type of room is

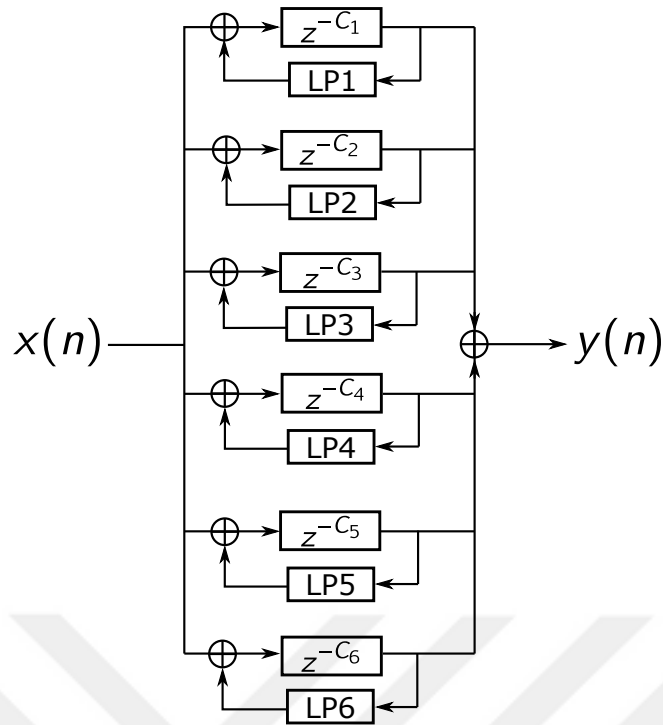


Figure 2.15: Moorer's Reverberator.

suitable for a DWN.

2.3.2.6 SDN

Scattering Delay Networks is an artificial reverberator which simulates the acoustics of single, rectangular volumes [13]. The algorithm can provide RIR of a rectangular room. SDN is also designed for frequency-domain operations [14]. Moreover, SDN is proven to be stable (i.e. room impulse response has decay behavior in rectangular rooms) in [15].

SDN is advantageous as it can provide room response with the direct path and first-order reflections while gracefully degrading accuracy for higher order reflections. SDN is proven to be perceptually comparable to other room-acoustics methods [18].

SDN is explained in detail in the next chapter as it is the basis of CV-SDN algorithm.

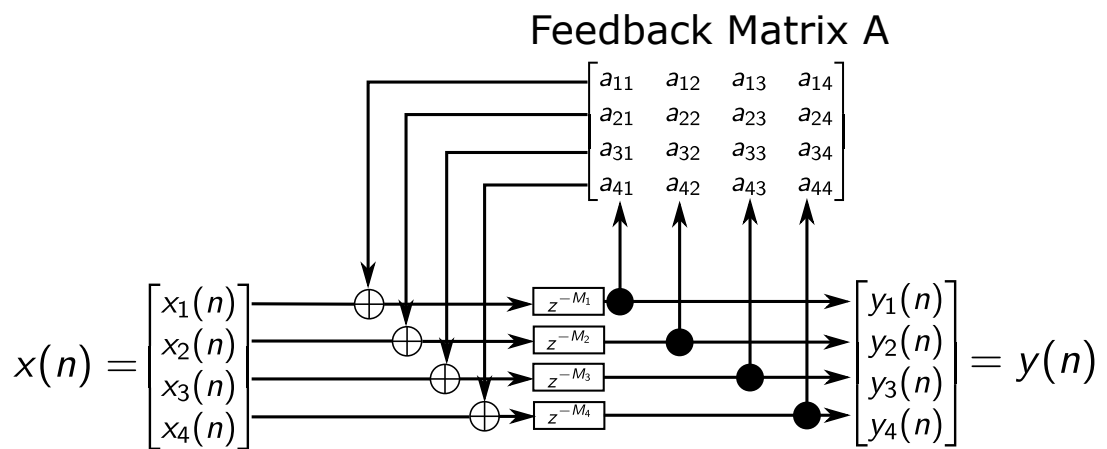


Figure 2.16: FDN reverberator.

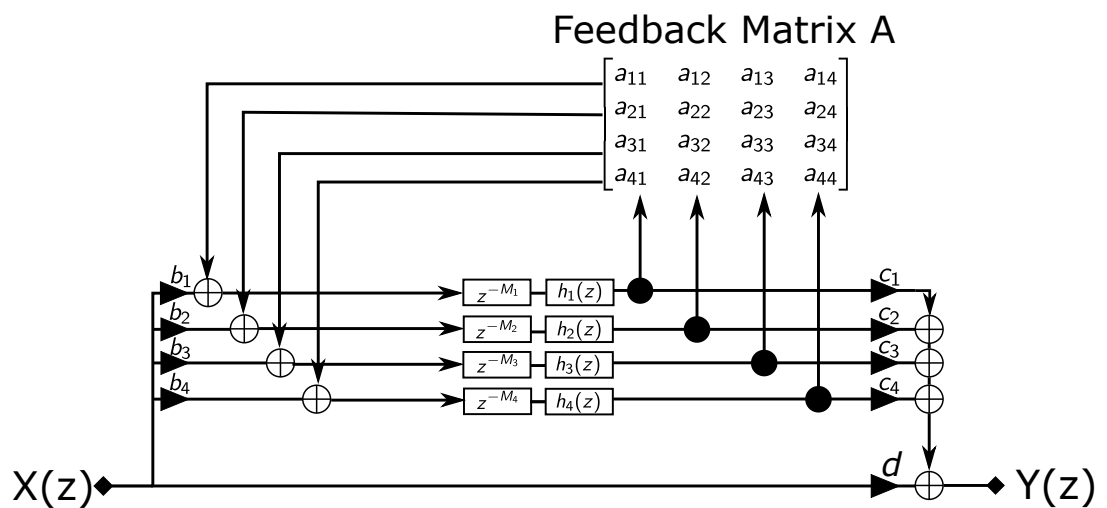


Figure 2.17: Jot's reverberator.



CHAPTER 3

SCATTERING DELAY NETWORKS

Scattering delay network (SDN) is a fully connected sparse digital waveguide network (DWN) with a minimal number of scattering junctions located at points of incidence on simulated walls. This allows the exact simulation of first-order early reflections in terms of their time delays, amplitudes, and a gracefully degrading simulation of second and higher-order reflections. The approximations made in second and higher-order reflections help SDN have relatively low computational cost.

Properties such as source and microphone positions, directivity being customizable make it scalable to different audio configurations. Additionally, wall absorption and room geometry can be adjusted to obtain the desired RIR. The customizability of SDN with the addition of its low computational operation provides a valuable addition to room acoustics field [3].

3.1 SDN design overview

Simulating room acoustics can be performed when images of the source are placed similarly to the image method [32]. First-order reflections of the sound source are imagined and the scattering nodes are placed at the point where the image to microphone vector intersects with the walls. Overall, six scattering nodes are necessary for a rectangular room. Although it is preferable to have more nodes to simulate higher order reflections accurately, approximation should be made here so that the algorithm has a low computational cost. Hence, the model uses the same node set for higher order reflections.

The basic SDN model consists of a source and a microphone node as well as six scattering nodes for a rectangular enclosure. Fig. 3.1 shows a simplified 2D diagram of the SDN structure. Source (black circle) and microphone (white circle) nodes determine the positions of the scattering nodes (black squares) on the simulated reflecting surfaces. Dashed lines represent the unidirectional connection originating from the sound source. Solid black lines between nodes represent the bidirectional connections where a lossless scattering operation carried out by scattering nodes is followed by a filter (or a scalar gain) emulating wall absorption. Finally, solid black lines extending from the scattering nodes to the microphone node are unidirectional connections. The consistency of scattering node positions with actual specular reflections allows simulating the delays and directions of first-order reflections very accurately.

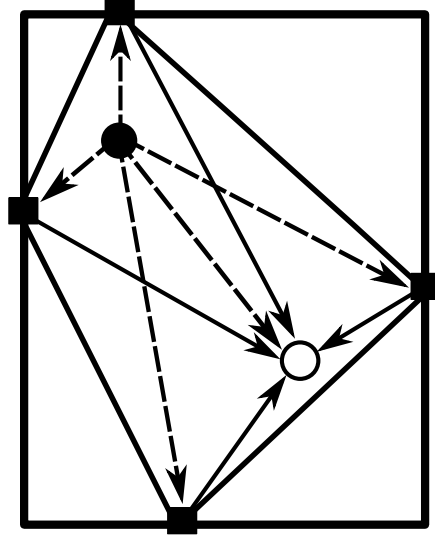


Figure 3.1: 2D representation of SDN.

Unlike perceptually motivated reverberators such as Jot's reverberator [44], one-to-one correspondence with the room geometry allows a trivial selection of the delay line lengths in the SDN model to model propagation delay between two points. As SDN can be regarded as a fully connected DWN, each node with the addition of sound source and microphone is connected with delay-lines. For the source to node and node to microphone connections, unidirectional delay lines are used. On the other hand, scattering nodes on the walls are connected via bidirectional delay lines as the sound should reflect back and forth to create a reverberation effect. Delay line length between two points, a and b is defined as:

$$D_{a,b} = \lfloor F_s \|\mathbf{x}_a - \mathbf{x}_b\| / c \rfloor, \quad (3.1.1)$$

where F_s is the sampling rate of the audio sample given to the system, $\|\mathbf{x}_a - \mathbf{x}_b\|$ is the distance between node a and node b , and c is the speed of sound in dry air.

Attenuation is employed to model the reduction in the amplitude of the sound signal due to distance. As discussed in Chapter 2, Eq. (2.1.2) expresses the distance between two points in space.

SDN performs the scattering operation by using scattering matrix. For $K + 1$ nodes in a volume, each node scatters to the other K nodes in a room. Therefore, a $K \times K$ scattering matrix \mathbf{S} can be used for scattering operation such that:

$$\mathbf{S} = \frac{2}{K} \times \mathbf{1}_{K \times K} - \mathbf{I}_{K \times K}. \quad (3.1.2)$$

As there are six scattering nodes placed for a rectangular enclosure, scattering matrix

can be adjusted as:

$$\mathbf{S} = \begin{bmatrix} -0.6 & 0.4 & 0.4 & 0.4 & 0.4 \\ 0.4 & -0.6 & 0.4 & 0.4 & 0.4 \\ 0.4 & 0.4 & -0.6 & 0.4 & 0.4 \\ 0.4 & 0.4 & 0.4 & -0.6 & 0.4 \\ 0.4 & 0.4 & 0.4 & 0.4 & -0.6 \end{bmatrix} \quad (3.1.3)$$

Scattering operation completed by multiplying scattering matrix with the incoming signals, and obtaining the outgoing ones. Thus, multiplication of inputs with scattering matrix \mathbf{S} provides the outputs:

$$\mathbf{p}^- [n] = \mathbf{S} \mathbf{p}^+ [n], \quad (3.1.4)$$

where $\mathbf{p}^- [n]$ and $\mathbf{p}^+ [n]$ denote outgoing and incoming waves respectively.

Finally, to define the energy circulation of energy within a single room, absorption must be discussed. As discussed in Chapter 2, reflection coefficient β is multiplied when a scattering operation is performed to obtain the scaled wave variables.

Finally, SDN can simulate the directional properties of the source(s) and microphone(s) as well as their rotations if required.

3.1.1 Source to Microphone Connection

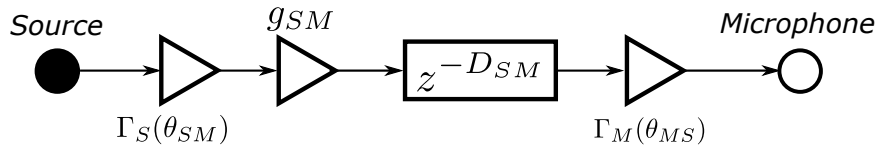


Figure 3.2: Source to Microphone Connection.

The connection between source and microphone accounts for directivity, attenuation and propagation delay. The source is specified by position \mathbf{x}_S , directivity $\Gamma_S(\theta)$, and acoustical axis with the unit vector of \mathbf{n}_S . Similarly, the microphone is specified by position \mathbf{x}_M , directivity $\Gamma_M(\theta)$, and acoustical axis with the unit vector of \mathbf{n}_M .

In the connection, first, the directivity of the source is considered. Directivity of source is defined as:

$$\Gamma_S(\theta_{SM}) = \sum_{p=0}^P a_p \cos^p \theta_{SM}, \quad (3.1.5)$$

where θ_{SM} is the incidence angle defined as:

$$\theta_{SM} = \cos^{-1} \left(\frac{\mathbf{n}_{SM} \cdot \mathbf{n}_S}{|\mathbf{n}_{SM}|} \right), \quad (3.1.6)$$

with $\mathbf{n}_{SM} = \mathbf{x}_M - \mathbf{x}_S$ being the normal vector from source to microphone.

Then, the attenuation is calculated as:

$$g_{S,M} = \frac{1}{\|\mathbf{x}_S - \mathbf{x}_M\|}, \quad (3.1.7)$$

to represent the reduction in the amplitude of signal between source and microphone.

Additionally, due to distance, the delay between source and microphone is represented by $z^{-D_{SM}}$. It is calculated as $D_{SM} = F_S \|\mathbf{x}_S - \mathbf{x}_M\| / c$, where F_S is the sampling rate and c is the speed of sound in dry air.

Finally, the microphone directivity is employed. It is defined as:

$$\Gamma_M(\theta_{MS}) = \sum_{q=0}^Q a_q \cos^q \theta_{MS}, \quad (3.1.8)$$

where θ_{MS} is the incidence angle defined as:

$$\theta_{MS} = \cos^{-1} \left(\frac{\mathbf{n}_{MS} \cdot \mathbf{n}_M}{|\mathbf{n}_{MS}|} \right), \quad (3.1.9)$$

with $\mathbf{n}_{MS} = \mathbf{x}_S - \mathbf{x}_M$ being the normal vector from microphone to source.

The resulting output from source and microphone connection is represented as the direct path in RIR.

3.1.2 Source to Node Connections

There is a total of six connections between source and nodes. In each connection, directivity, attenuation and propagation delay are considered. Additionally, a correction gain factor of 1/2 is utilized so that the early and late reflections are represented accurately. Directivity is only calculated from the source side as there are no directional or rotational properties of nodes. Directivity of the source is defined as:

$$\Gamma_S(\theta_{Sk}) = \sum_{p=0}^P a_p \cos^p \theta_{Sk}, \quad (3.1.10)$$

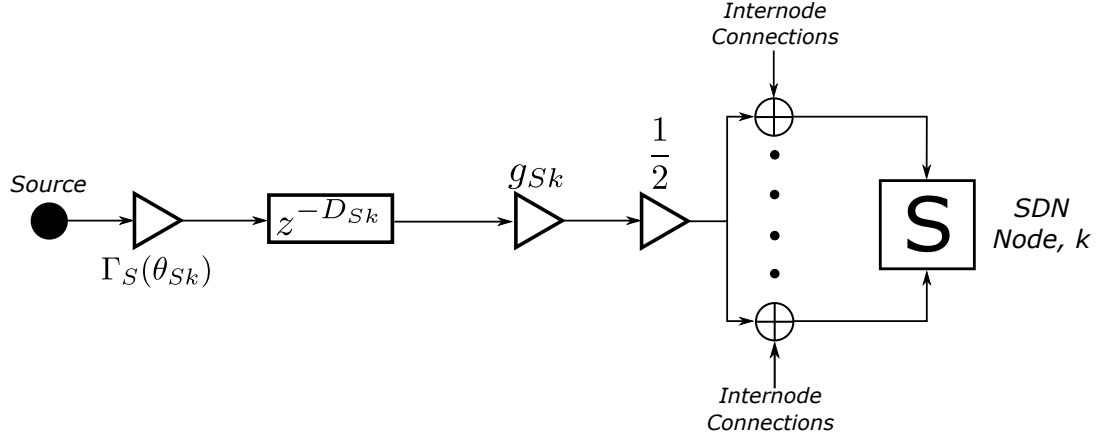


Figure 3.3: Source to Node Connections.

where θ_{Sk} is the incidence angle between source and node k as:

$$\theta_{Sk} = \cos^{-1} \left(\frac{\mathbf{n}_{Sk} \cdot \mathbf{n}_S}{|\mathbf{n}_{Sk}|} \right), \quad (3.1.11)$$

with $\mathbf{n}_{Sk} = \mathbf{x}_S - \mathbf{x}_k$ being the normal vector from source to a node.

Propagation delay $z^{-D_{Sk}}$ and attenuation are both calculated concerning positions of the source and scattering nodes and are defined as $D_{Sk} = F_S \|\mathbf{x}_S - \mathbf{x}_k\| / c$, and $g_{S,k} = \frac{1}{\|\mathbf{x}_S - \mathbf{x}_k\|}$ respectively.

For each scattering node, there are six incoming wave variables. One of those variables is coming from the source side, and the rest are from the other scattering nodes. The output of the source side in Fig. 3.3, is added to all of the incoming lines p^+ of the respective SDN node. This operation is called pressure insertion.

3.1.3 Node to Node Connections

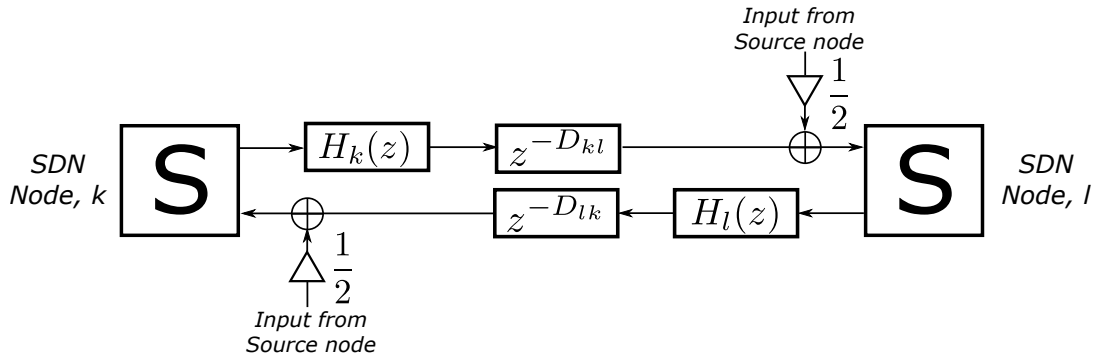


Figure 3.4: Node to Node Connections.

Node interconnections utilize propagation delay, $z^{-D_{kl}}$, absorption filters, $H_k(z)$ and $H_l(z)$ each representing their respective wall absorption characteristic. Absorption filter corresponds to losses when waves are incident on walls of the enclosure. The calculation is performed as explained in Eq. (2.1.1). Directivity and attenuation between nodes are not considered. Propagation delays are calculated with node to node distances such that $D_{kl} = F_S ||\mathbf{x}_k - \mathbf{x}_l|| / c$. Fig. 3.4 also shows input from source node due to pressure insertion.

3.1.4 Node to Microphone Connections

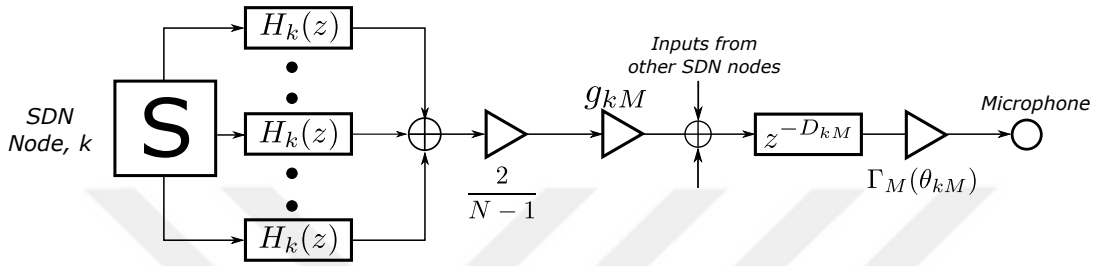


Figure 3.5: Node to Microphone Connections.

Node to microphone connections consists of absorption filter, a normalization gain, attenuation, propagation delay, and directivity. First, from output vector $p^-[n]$, every index is taken and summed up. This is called pressure extraction.

Afterward, a normalization gain of $\frac{2}{N-1}$ is used to scale the signal. For a rectangular enclosure, N is set to 6. Attenuation based on the distance between the node and the microphone is employed. It differs from the previous definitions and defined as:

$$g_{k,M} = \frac{1}{1 + \frac{||\mathbf{x}_k - \mathbf{x}_M||}{||\mathbf{x}_S - \mathbf{x}_k||}}. \quad (3.1.12)$$

Then, inputs from other SDN nodes are collected and transferred with a delay of $z^{-D_{kM}}$ calculated with $D_{kM} = F_S ||\mathbf{x}_k - \mathbf{x}_M|| / c$.

Finally, microphone directivity is exercised. It is defined as:

$$\Gamma_M(\theta_{kM}) = \sum_{q=0}^Q a_q \cos^q \theta_{kM}, \quad (3.1.13)$$

where θ_{kM} is the incidence angle between microphone and k^{th} node defined as:

$$\theta_{kM} = \cos^{-1} \left(\frac{\mathbf{n}_{kM} \cdot \mathbf{n}_M}{|\mathbf{n}_{kM}|} \right), \quad (3.1.14)$$

with $\mathbf{n}_{kM} = \mathbf{x}_k - \mathbf{x}_M$ being the normal vector from k^{th} node to microphone.

3.2 Transfer Function of SDN

The transfer function of SDN can be analyzed in z-domain. A source node and a microphone node can be defined with positions at \mathbf{x}_s and \mathbf{x}_m with the directivity patterns $\Gamma_s(\theta)$ and $\Gamma_m(\theta)$. As discussed previously, six scattering nodes can be created with the image method. Room geometry is imagined to be a rectangular one for SDN to be defined.

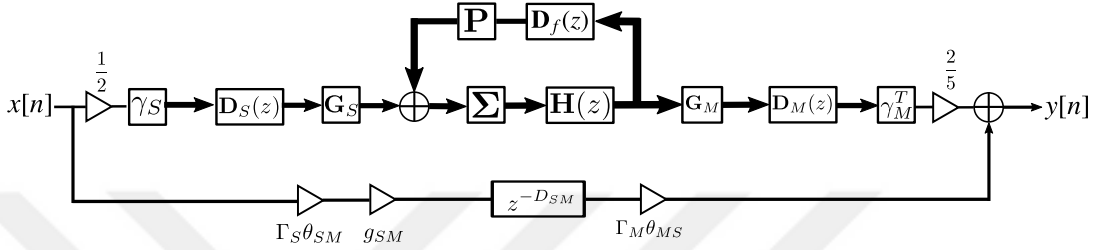


Figure 3.6: Transfer function of SDN.

Transfer function of SDN is shown at Fig. 3.6. Transfer function is adapted from the ones given in [14, 15]. Thus, SDN can be represented as:

$$H_{SDN}(z) = \bar{g}z^{-D_{sm}} + \frac{1}{5}\mathbf{k}_m(z)\mathbf{F}(z)\mathbf{k}_s(z) \quad (3.2.1)$$

where $\mathbf{F}(z) = [\mathbf{\Sigma}^{-1}\mathbf{H}(z^{-1}) - \mathbf{P}\mathbf{D}_f(z)]^{-1}$.

It should be noted that in the definitions of the transfer function below, $\mathbf{A} \otimes \mathbf{B}$ represents the Kronecker product of two matrices, θ_{si} (or θ_{im}) represent the angles between the acoustic axes of the source node (or the microphone node) and the i -th scattering node.

The terms of the transfer function in Eq. (3.2.1) are:

1. $\mathbf{\Sigma} = \mathbf{I}_{5 \times 5} \otimes \mathbf{S}$ is the block diagonal scattering matrix with $\mathbf{S} = \frac{2}{5}\mathbf{1}_{5 \times 5} - \mathbf{I}_{5 \times 5}$. Scattering matrix is used as given in Eq. (3.1.3).
2. \mathbf{P} is a permutation matrix representing the connectivity of the scattering nodes.
3. $\mathbf{D}_f(z) = \text{diag}(z^{-D_{1,2}}, \dots, z^{-D_{6,5}})$ is the internode delay matrix where $D_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\| \cdot F_s/c$ is the delay between two scattering nodes, i and j , positioned at \mathbf{x}_i and \mathbf{x}_j , respectively, with c as the speed of sound in dry air and F_s the sampling rate.
4. $\mathbf{H}(z) = \mathbf{H}_w(z) \otimes \mathbf{I}_{5 \times 5}$ is the reflectance matrix with $\mathbf{H}_w(z) = \text{diag}(H_1(z), \dots, H_6(z))$ where $H_i(z)$ represents the reflectance transfer function the i -th wall. The reflectance terms are related with wall reflection coefficient given in Eq. (2.1.1).

5. $\mathbf{k}_s(z) = \mathbf{G}_s \mathbf{D}_s \gamma_s$, and $\mathbf{k}_m(z) = \gamma_m^T \mathbf{D}_m \mathbf{G}_m$ are the source-model and model-microphone adaptor vectors, where:
- $\gamma_s = \mathcal{G}_s \otimes \mathbf{1}_{5 \times 1}$ with $\mathcal{G}_s = [\Gamma_s(\theta_{s1}), \dots, \Gamma_s(\theta_{s6})]^T$ comprise source to node directivity gain. Similarly, $\gamma_m = \mathcal{G}_m \otimes \mathbf{1}_{5 \times 1}$ with $\mathcal{G}_m = [\Gamma_m(\theta_{1m}), \dots, \Gamma_m(\theta_{6m})]^T$ is the node to microphone directivity gain,
 - $\mathbf{D}_s = \mathbf{d}_s \otimes \mathbf{I}_{5 \times 5}$ where $\mathbf{d}_s = \text{diag}(z^{-D_{s,1}}, \dots, z^{-D_{s,6}})$ with $D_{s,i} = \|\mathbf{x}_s - \mathbf{x}_i\| \cdot F_s / c$ is the source-node delay matrix. Equivalently, $\mathbf{D}_m = \mathbf{d}_m \otimes \mathbf{I}_{5 \times 5}$ where $\mathbf{d}_m = \text{diag}(z^{-D_{1,m}}, \dots, z^{-D_{6,m}})$ with $D_{i,m} = \|\mathbf{x}_i - \mathbf{x}_m\| \cdot F_s / c$ is the node-microphone delay matrix,
 - $\mathbf{G}_s = \mathbf{g}_s \otimes \mathbf{I}_{5 \times 5}$ and $\mathbf{G}_m = \mathbf{g}_m \otimes \mathbf{I}_{5 \times 5}$ with $\mathbf{g}_s = \text{diag}(g_{s1}, \dots, g_{s6})$ is the source gain matrix. Correspondingly, $\mathbf{g}_m = \text{diag}(g_{1m}, \dots, g_{6m})$ is the microphone gain matrix. Each matrix depends on the relative physical positions of the nodes.
6. Finally, $\bar{g} = g_{sm} \Gamma_s(\theta_{sm}) \Gamma_m(\theta_{ms})$ is the gain emulating distance attenuation, and $D_{s,m} = \|\mathbf{x}_s - \mathbf{x}_m\| \cdot F_s / c$ is the delay between the source and microphone nodes. As before, c is the speed of sound in dry air and F_s is the sampling rate.

3.3 Evaluation of SDN

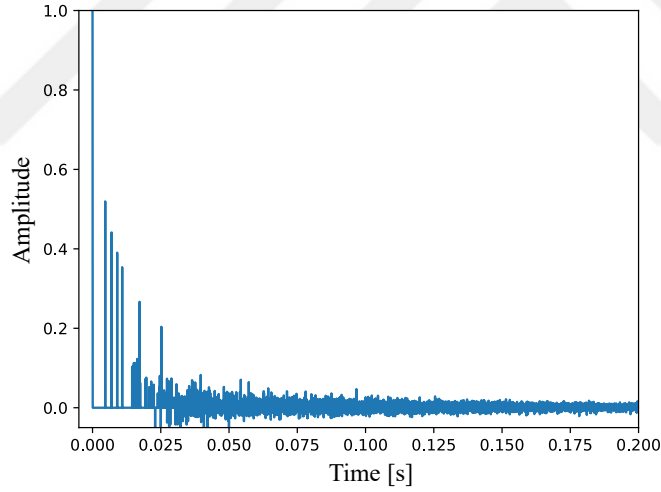


Figure 3.7: Room impulse response of SDN.

A simulation setup with a rectangular volume with dimensions $9 \times 7 \times 4$ is created for SDN. Taking left bottom of the enclosure at the origin, an omnidirectional source and microphone are placed at $\mathbf{x}_s = (4.5, 3.5, 2)$ and $\mathbf{x}_m = (2, 2, 1.5)$. The absorption coefficient of the walls are set as 0.2, which results in wall reflection coefficient of 0.89 with Eq. (2.1.1).

With the above configuration, RIR and EDC of SDN can be obtained as shown in Fig. 3.7 and Fig. 3.8 respectively. From EDC, RT_{60} of the system is found to be 0.94 seconds. The results are compared with two well-known formulas developed by

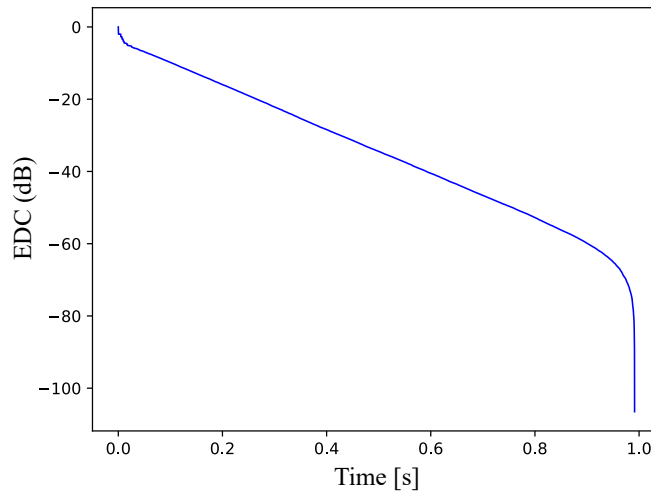


Figure 3.8: EDC of SDN.

Sabine [45] and Eyring [26] and were shown to agree with these theoretically derived formulae.

3.4 Extensions of SDN

SDN is extended to include second-order reflections with the addition of directionally dependent filtering for each scattering node [19]. The extended method is called Waveguide Web and it is an important artificial reverberator to be used in open acoustic scenes which SDN is not capable of.

In an augmented reality audio application, SDN is used as a reverberator [16]. It was the first augmented reality action where SDN is used. Perceptual evaluation of SDN is also undertaken with the augmented reality application [17]. Waveguide Web method is used for comparison. It is seen that Waveguide Web method is preferred as it offers more accurate second-order reflections relative to SDN.

Perceived naturalness and pleasantness of SDN were compared to other reverberation methods such as FDN and CATT-Acoustic models [18]. It is found that SDN achieved higher mean ratings both in naturalness and pleasantness. Moreover, similar results were obtained when SDN is compared with binaural room impulse responses (BRIRs).



CHAPTER 4

COUPLED VOLUME SCATTERING DELAY NETWORKS

Coupled Volume SDN (CV-SDN) is an extension of the classic SDN which uses multiple SDNs connected via an aperture node to emulate the effect of sound energy exchange between connected volumes. On top of SDN, CV-SDN introduces a second volume and an aperture. The second volume will be referred to as the coupled volume, whereas the original room (i.e. volume) will be referred to as the main volume. Coupled volume can have a different volume and absorption characteristics from the main volume. The real life use cases for CV-SDN are given in Chp. 6.

The aperture is imagined as a rectangular opening like a door or a window with a position and an area that can be specified. We assume that the aperture acts as an anechoic rectangular patch on the shared surface of the two volumes, which we call the common wall (CW). While this assumption is not entirely physically accurate in terms of the acoustic impedance at the aperture boundary, it facilitates the formulation of the energy exchange between the SDN models. Aperture size control and energy modeling are further described in transfer function derivations that will be explained below.

4.1 CV-SDN Settings

It is important to define different settings that CV-SDN can have as we are not dealing with a single volume. These settings are determined by the positions of the source and the microphone. There is a total of two combinations of source and microphone positioning where:

1. Source and microphone in the main volume.
2. Source and microphone are in different volumes with the source being in the main volume.

It should be noted that as suggested by Pu et. al [28] and Eyring [27], there are two different DSD patterns for coupled volumes. Each pattern can only be obtained when the source is placed into a less reverberant volume. Depending on the positioning of the microphone, concave or convex DSD patterns can be observed. The concave decay profile occurs when both the source and microphone are placed in the same volume (i.e. main volume) which has a lower reverberation than the coupled volume.

The convex decay case occurs when source and microphone are placed in different volumes with the main volume being the less reverberant one. Hence, two settings are considered given that the main volume is the more reverberant one.

Additionally, there can be only two different CV-SDN transfer functions based on microphone and source placements regardless of the reverberation characteristics of both volumes, which will be explained in the transfer function of CV-SDN settings. The transfer function and system diagram of both settings will be clarified in Section 4.3.

4.1.1 Same Volume CV-SDN Setting

Same volume CV-SDN emulates the condition for which the source and the microphone are both in the same volume (i.e. the main volume) as shown in Fig. 4.1. Source (black circle) and microphone (white circle) nodes determine the positions of the scattering nodes (black squares) on the simulated reflecting surfaces. Similarly, other sets of nodes are present due to connections between source to aperture, and aperture to the microphone. For simplicity, only one set of nodes (i.e. the node set created with source and microphone) are shown in the figure. Like SDN, nodes are generated using the image method. Plus sign on the CW represents the aperture node and connects two volumes. In the coupled volume, there is not a source and receiver pair whose positions are distinct. Hence, the node set in the coupled volume is not created with ISM. Therefore, an approximation is made and nodes are placed at the geometric center of bounding surfaces of the coupled volume.

Five SDN components are needed for the same volume CV-SDN to function. The first component emulates the main volume which contains both the source and the microphone. In the second connected SDN component that also emulates the main volume, the aperture node is designated as a microphone node. The third connected SDN component emulates the coupled volume where the aperture node acts as both the source and the microphone nodes and is connected to the associated scattering nodes emulating the coupled volume. The fourth SDN component is constructed with the aperture being source and microphone in the main volume. This SDN establishes the feedback channel between the two volumes. In the fifth SDN component, the aperture node acts as a source in the main volume and transmits the signal it receives from the coupled volume to the actual microphone node.

Five SDN structures result in five different node sets (i.e. 30 scattering nodes). However, these node sets are only considered within SDN structure they are connected to. For example, in Fig. 4.1, node set created between source and microphone are shown in the main volume. It should be emphasized that the other node sets are present as well and are not shown in the figure to keep it simple and understandable.

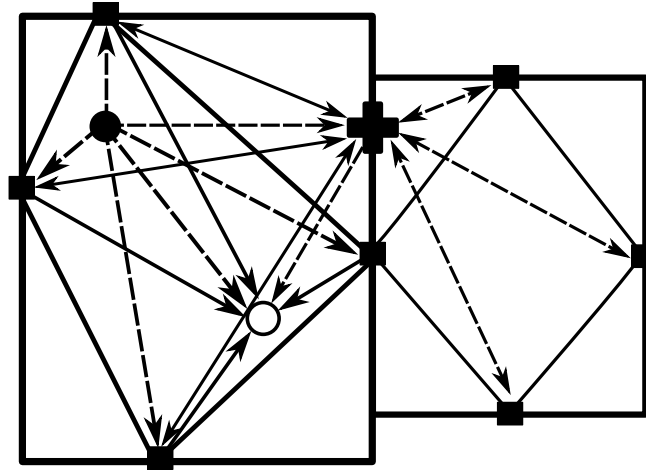


Figure 4.1: 2D representation of same volume CV-SDN.

4.1.2 Different Volume CV-SDN Setting

Different volume CV-SDN emulates the condition for which the source and the microphone are placed in different volumes as shown in Fig. 4.2. The source (black circle) is alone in the main volume. Therefore, a node set within the main volume is generated due to source and aperture acting as the microphone. Aperture (i.e. plus sign) connects the two volumes. Another node-set (i.e. the node set between aperture and microphone) is shown in the coupled volume although two additional set of nodes exists due to the feedback path caused by the aperture. In both of the volumes, the aperture to aperture connection is present. This creates the feedback loop of different volume CV-SDN. The scattering nodes of the invisible node sets are placed at the geometric center of the bounding surfaces of the main and coupled volumes.

Overall, Four SDN components are needed for different volume CV-SDN to function. The first component emulates the main volume which contains the source and aperture acting as a microphone. The second component emulates the coupled volume, where aperture node is designated as the source node with the microphone as the receiver. The third SDN component establishes a feedback channel in the coupled volume and accepts the aperture as both the source and the microphone. The fourth SDN component completes the feedback path in the main volume with aperture acting as both source and microphone.

4.2 CV-SDN Design Overview

Before defining the transfer function of the two settings, the new possible connections of CV-SDN should be discussed. The connections defined in Section 3.1 of Chapter 3 are still utilized in CV-SDN models. The connections between the source to the microphone, source to nodes, nodes to nodes, and nodes to the microphone are the same as classic SDN. However, new connections are included with the introduction of aperture node. It should also be mentioned that the positions of the scattering nodes

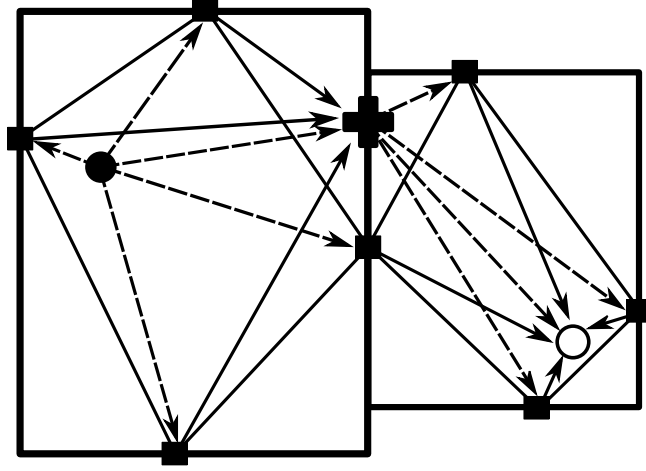


Figure 4.2: 2D representation of different volume CV-SDN.

are still determined with the image method, except in the volume where neither source or microphone exists.

4.2.1 Source to Aperture Connection

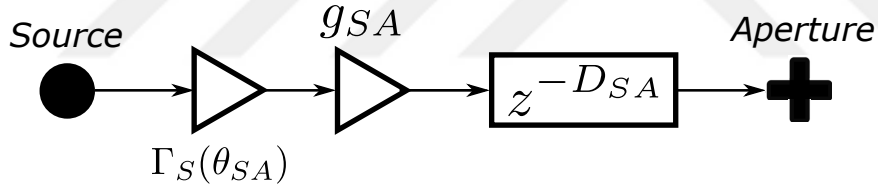


Figure 4.3: Source to Aperture Connection.

Fig. 4.3 shows the connection between the source and the aperture. In this connection, aperture acts as a sink (i.e. microphone). Therefore, the connection employs a source to aperture directivity, attenuation, and propagation delay. Similar to SDN connection of source to the microphone, the source is specified by the position \mathbf{x}_S , directivity $\Gamma_S(\theta)$, and acoustical axis with the unit vector of \mathbf{n}_S . However, the aperture is only defined by its position, \mathbf{x}_A and unit vector, \mathbf{n}_A which is perpendicular to the CW. Directivity of the aperture is not defined. It should be noted that the position of the aperture must be on the same plane with CW.

First, directivity of source is considered. Directivity of source is defined as:

$$\Gamma_S(\theta_{SA}) = \sum_{p=0}^P a_p \cos^p \theta_{SA}, \quad (4.2.1)$$

where θ_{SA} is the incidence angle between the source and the aperture defined as:

$$\theta_{SA} = \cos^{-1} \left(\frac{\mathbf{n}_{SA} \cdot \mathbf{n}_S}{|\mathbf{n}_{SA}|} \right), \quad (4.2.2)$$

where $\mathbf{n}_{SA} = \mathbf{x}_S - \mathbf{x}_A$ is the normal vector from source to aperture.

Then, the attenuation is calculated as:

$$g_{S,A} = \frac{1}{\|\mathbf{x}_S - \mathbf{x}_A\|}, \quad (4.2.3)$$

indicating the reduction in the amplitude of the signal between source and aperture.

The delay between source and aperture is represented with $z^{-D_{SA}}$ and it can be calculated as $D_{SA} = F_S \|\mathbf{x}_S - \mathbf{x}_A\| / c$, where F_S is the sampling rate and c is the speed of sound in dry air.

The directivity of the aperture is disregarded since it is not defined.

4.2.2 Node to Aperture Connections

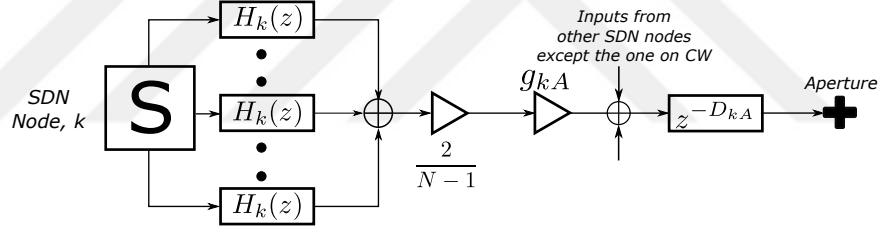


Figure 4.4: Node to Aperture Connections.

Node to Aperture connections is shown in Fig. 4.4, and consist of absorption filters, a normalisation gain factor, attenuation, and propagation delay. From each scattering node, each index of output vector $p^- [n]$ is taken and summed up. As explained previously, the operation is called pressure extraction.

Afterward, a normalisation gain of $\frac{2}{N-1}$ is used to scale the signal. Since both volumes are rectangular enclosures, N is set to 6 for both volumes. Attenuation based on the distance between the node and the aperture is also emulated. The calculation is different from the other cases and given as:

$$g_{k,A} = \frac{1}{1 + \frac{\|\mathbf{x}_k - \mathbf{x}_A\|}{\|\mathbf{x}_S - \mathbf{x}_k\|}}. \quad (4.2.4)$$

Then, the inputs from other SDN nodes except the one on the CW is summed up. The reason why one of the nodes is not considered is that it lies on the same surface (i.e.

CW) with the aperture node and transmission from the node on CW to the aperture is physically not possible.

Finally, the delay, $z^{-D_{kA}}$ is employed. It is calculated with $D_{kA} = F_S \|\mathbf{x}_k - \mathbf{x}_A\| / c$.

4.2.3 Aperture to Node Connections

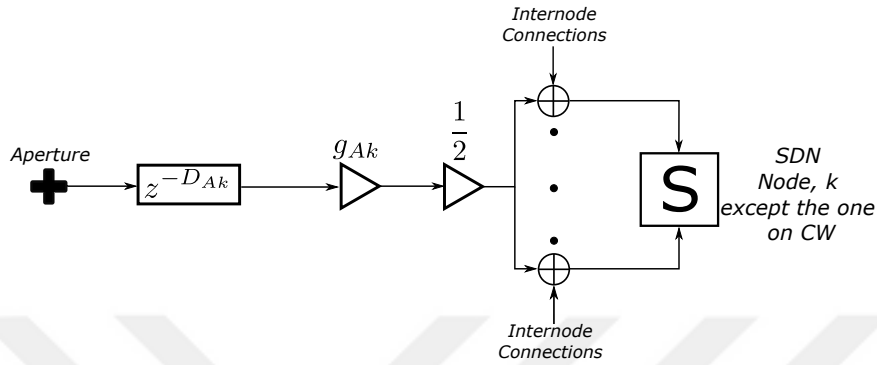


Figure 4.5: Aperture to Node Connections.

Aperture to Node connections, shown in Fig. 4.5, employ propagation delay, attenuation, and a scale factor of $1/2$. In the connections, aperture acts as a source. As before, the directivity of the aperture is not employed. Attenuation is defined as $g_{A,k} = \frac{1}{\|\mathbf{x}_A - \mathbf{x}_k\|}$.

The propagation delay $z^{-D_{Ak}}$ concerning aperture and node positions calculated with $D_{Ak} = F_S \|\mathbf{x}_A - \mathbf{x}_k\| / c$.

Similar to SDN, for each scattering node except the one on CW, there are six incoming wave variables. One of those variables is coming from the aperture side, and the rest are from the other scattering nodes. For the scattering node on CW, there are five incoming wave variables only. Hence, there is a total of five connections between aperture and nodes. The connection between the node on the CW and aperture is disregarded as they lie on the same surface and transmission from aperture to the node on the CW is not physically possible. As previously, this is called pressure insertion.

4.2.4 Aperture to Microphone Connection

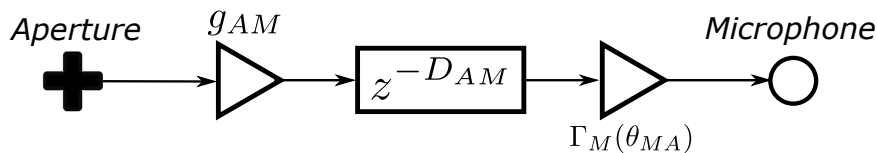


Figure 4.6: Aperture to Microphone Connection.

The Fig. 4.6 represents the connection between aperture and microphone. This connection is similar to the source to aperture connection. Aperture acts as a source, and directivity of it is not employed. It is defined only by its position \mathbf{x}_A and normal vector \mathbf{n}_A . The microphone is defined similarly as the source. It has a position, \mathbf{x}_M , directivity $\Gamma_M(\theta)$, and acoustical axis with the unit vector of \mathbf{n}_M .

First attenuation is emulated between aperture and the microphone as $g_{A,M} = \frac{1}{\|\mathbf{x}_A - \mathbf{x}_M\|}$. Then propagation delay $z^{-D_{AM}}$ is employed as $D_{AM} = F_S \|\mathbf{x}_A - \mathbf{x}_M\| / c$.

Finally, directivity of the microphone is utilized. It is defined as:

$$\Gamma_M(\theta_{MA}) = \sum_{q=0}^Q a_q \cos^q \theta_{MA}, \quad (4.2.5)$$

where θ_{MA} is the incidence angle between microphone and the aperture defined as:

$$\theta_{MA} = \cos^{-1} \left(\frac{\mathbf{n}_{MA} \cdot \mathbf{n}_M}{|\mathbf{n}_{MA}|} \right), \quad (4.2.6)$$

with $\mathbf{n}_{MA} = \mathbf{x}_A - \mathbf{x}_M$ being the normal vector from microphone to aperture.

4.3 Transfer Function of CV-SDN

Similar to the classic SDN, the source and microphone nodes only act as input (i.e. source) and output (i.e. sink) points, respectively. In contrast, aperture nodes act as both an input and an output point. While the depiction of CV-SDN as shown in Fig. 4.1 and Fig. 4.2 can be intuitively understood, it is more instructive and amenable to further analysis to express the transfer functions of both settings of CV-SDN using a linear combination of classic SDN models. Hence, the Transfer functions of CV-SDN consist of the combination of several transfer functions of SDN. The terms introduced in Section 3.2 of Chapter 3 are expanded upon.

It must be pointed out that in the discussions of the transfer function, subscripts denote the source and sink, and the superscript denotes the volume that the SDN component models. For example, $H_{s,a}^{(1)}(z)$ represents the transfer function from the source to the aperture node, with model parameters determined according to the geometry of the main volume.

4.3.1 Transfer Function of the Same Volume CV-SDN

The same volume CV-SDN system diagram is shown in Fig. 4.7. The components comprising the system are:

- $H_{s,m}^{(1)}(z)$ representing the main volume transfer function,

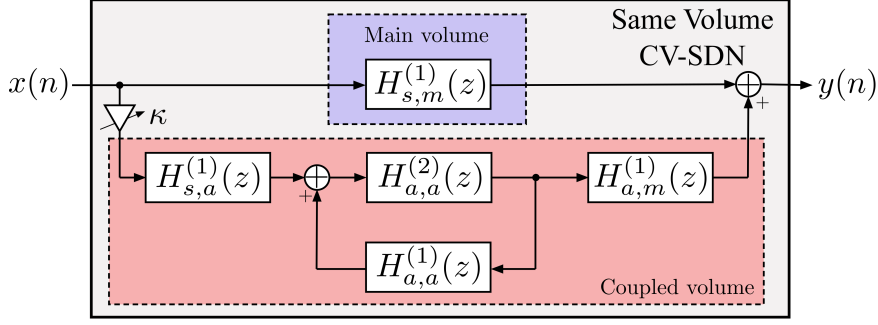


Figure 4.7: Same Volume CV-SDN System Diagram.

- $H_{s,a}^{(1)}(z)$ representing the transfer function between the source and the aperture,
- $H_{a,a}^{(2)}(z)$ representing the reverberation in the coupled volume after energy exchange,
- $H_{a,a}^{(1)}(z)$ representing the reverberation in the main volume due to energy transmitted back from the coupled volume,
- $H_{a,m}^{(1)}(z)$ representing the transfer function between the aperture and the microphone node in the main volume.

Further, the energy transferred to the coupled volume is controlled by the linkage coefficient κ which will be discussed further below. In all transfer functions, the recursive parts of SDN components consist of fully connected delay networks. However, the adaptor vectors used in connecting the source and sink nodes in each case are different depending on whether or not the sink or the source is an aperture node.

Source-to-microphone: The component $H_{s,m}^{(1)}(z)$ is identical to a classic SDN except for an adjustment of one of the terms of the reflectance matrix due to the aperture being modelled as an anechoic patch. In contrast with the classic SDN, the reflectance matrix is defined as:

$$\mathbf{H}_w^{(1)}(z) = \text{diag}(H_1^{(1)}(z), \dots, H_{cw}^{(1)}(z), \dots, H_6^{(1)}(z)), \quad (4.3.1)$$

where $H_{cw}^{(1)}(z)$ is the adjusted reflectance term for the scattering node that shares the same boundary surface with the aperture node (i.e. scattering node on the CW). This adjustment is discussed in the following discussions.

Overall, the transfer function of this component is:

$$H_{s,m}^{(1)}(z) = \bar{g}_{s,m} z^{-D_{s,m}} + \frac{1}{5} \mathbf{k}_m(z) \mathbf{F}_{s,m}^{(1)}(z) \mathbf{k}_s(z), \quad (4.3.2)$$

where $\mathbf{F}_{s,m}^{(1)}(z) = [\mathbf{\Sigma}^{-1} \mathbf{H}_w^{(1)}(z^{-1}) - \mathbf{P} \mathbf{D}_{sm,f}^{(1)}(z)]^{-1}$.

Source-to-aperture: The main differences of this component from a classic SDN appear in (i) the direct path from the source to aperture, (ii) the reflectance matrix, and (iii) the aperture adaptor vector. The aperture is assumed to behave like an omnidirectional microphone. Therefore the direct path gain includes only the source directivity and distance attenuation terms such that $\bar{g}_{sa} = g_{sa}\Gamma_s(\theta_{sa})$ where $g_{sa} = (\|\mathbf{x}_s - \mathbf{x}_a\|)^{-1}$. The reflectance matrix is modified as in the source-to-microphone component, and the aperture adaptor matrix is defined as:

$$\mathbf{k}_{as} = \mathbf{I}_{as}\mathbf{G}_{sa}\mathbf{D}_{sa}\mathbf{1}_{6\times 1} \quad (4.3.3)$$

with $\mathbf{I}_{as} = \text{diag}(1 - \delta_{1a}, \dots, 1 - \delta_{6a})$ where δ_{ia} is an indicator function which is unity iff. aperture and the connected scattering node, i , are on the same boundary of the volume and zero otherwise. The position of the scattering node on the same surface is assumed to lie at the centroid of the bounded surface.

Overall, the transfer function of this component is:

$$H_{s,a}^{(1)}(z) = \bar{g}_{sa}z^{-D_{s,a}} + \frac{1}{5}\mathbf{k}_a^{(1)}(z)\mathbf{F}_{s,a}^{(1)}(z)\mathbf{k}_s(z), \quad (4.3.4)$$

where $\mathbf{F}_{s,a}^{(1)}(z) = [\Sigma^{-1}\mathbf{H}_w^{(1)}(z^{-1}) - \mathbf{PD}_{sa,f}^{(1)}(z)]^{-1}$.

Aperture-to-aperture: The component uses the aperture node both as a source and as a microphone. However, there is no direct path component, but only the recursive part. There are two aperture-to-aperture components in the CV-SDN model, $H_{a,a}^{(1)}(z)$ and $H_{a,a}^{(2)}(z)$ such that:

$$H_{a,a}^{(q)}(z) = \frac{1}{5}\mathbf{k}_a^{(q)}(z)\mathbf{F}_{a,a}^{(q)}(z)\mathbf{k}_a^{(q)}(z), \quad q = 1, 2 \quad (4.3.5)$$

where $\mathbf{F}_{a,a}^{(q)}(z) = [\Sigma^{-1}\mathbf{H}_w^{(q)}(z^{-1}) - \mathbf{PD}_{aa,f}^{(q)}(z)]^{-1}$ with $\mathbf{D}_{aa,f}^{(q)}(z)$ calculated using scattering nodes positioned on geometric centroids of bounding surfaces of the main (i.e. $q = 1$) and coupled volumes (i.e. $q = 2$), respectively. Notice that the main volume component is connected to the coupled volume component via a feedback loop.

Aperture-to-microphone: The component models the transfer function from the aperture back to the microphone. Here, the aperture node acts as a source node. The positions of scattering nodes that are not on the same boundary with the aperture node are calculated using microphone and aperture node positions. Similar to the source-to-aperture component, the aperture acts as an omnidirectional source resulting in a direct path gain of $\bar{g}_{am} = g_{am}\Gamma_m(\theta_{am})$ where $g_{am} = (\|\mathbf{x}_a - \mathbf{x}_m\|)^{-1}$. The aperture adaptor vector is given as $\mathbf{k}_a = \mathbf{I}_a^{(1)}\mathbf{G}_{am}\mathbf{D}_{am}\mathbf{1}_{6\times 1}$.

Thus, the transfer function of this component is:

$$H_{a,m}^{(1)}(z) = \bar{g}_{a,m}z^{-D_{s,m}} + \frac{1}{5}\mathbf{k}_m(z)\mathbf{F}_{a,m}(z)\mathbf{k}_a(z), \quad (4.3.6)$$

where $\mathbf{F}_{a,m}^{(1)}(z) = [\Sigma^{-1}\mathbf{H}_w^{(1)}(z^{-1}) - \mathbf{PD}_{am,f}^{(1)}(z)]^{-1}$.

Aperture size control: Acoustic energy transmission from the main volume to the coupled volume via an aperture is modeled via the linkage coefficient. The linkage coefficient is defined as:

$$\kappa = \sqrt{S_A/S_T} \quad (4.3.7)$$

where S_A is the aperture surface area, and $S_T = S_a + S_{cw}$ is the total area of the respective volume boundary and S_{cw} is the area of the CW. In other words, the energy transferred to the coupled volume depends on the relative area of the aperture.

Considering that the aperture would act as an anechoic patch, the effective absorption coefficient of the wall that contains the aperture also has to be changed accordingly. This adjustment would result in the reflectance transfer function, given as:

$$H_{cw}^{(q)}(z) = \frac{S_A + H_i^{(q)}(z)S_{CW}}{S_{CW} + S_A}, \quad (4.3.8)$$

where $H_i^{(q)}(z)$ emulates the reflectance transfer function for the material of the common surface on the side of volume q . This adjustment has the effect of increasing the energy decay rate and reducing the reverberation time.

Once these components are defined, the transfer function of the same volume CV-SDN can be expressed as:

$$H_{SV-CV-SDN}(z) = H_{s,m}^{(1)}(z) + \kappa \frac{H_{s,a}^{(1)}(z)H_{a,a}^{(2)}(z)H_{a,m}^{(1)}(z)}{1 - H_{a,a}^{(1)}(z)}. \quad (4.3.9)$$

4.3.2 Transfer Function of Different Volume CV-SDN

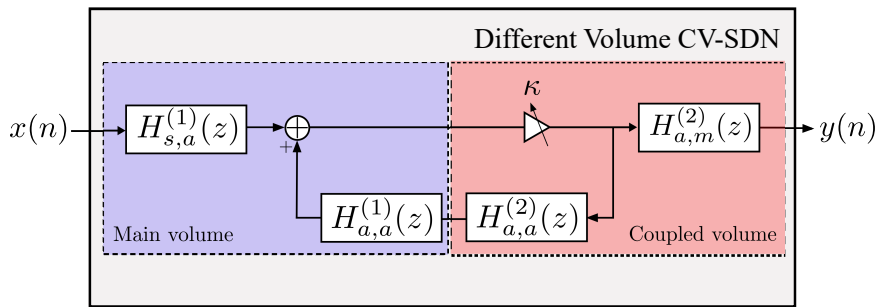


Figure 4.8: Different Volume CV-SDN System Diagram.

The system diagram of the different volume CV-SDN is given in Fig. 4.8. The different volume CV-SDN has the following components:

- $H_{s,a}^{(1)}(z)$ representing the main volume,

- $H_{a,a}^{(2)}(z)$ representing the reverberation in the coupled volume due to aperture,
- $H_{a,a}^{(1)}(z)$ representing the reverberation in the main volume due to aperture,
- $H_{a,m}^{(2)}(z)$ representing the transfer function between the aperture and the microphone node in the coupled volume.

Similar to the same volume CV-SDN, the energy transferred to the coupled volume is controlled by the linkage coefficient κ . As before, the recursive parts of SDN components constructed with fully connected delay networks. Nevertheless, the adaptor vectors used for connecting source and sink nodes are different depending on whether or not the aperture node is the source or the sink.

Source-to-aperture: The component $H_{s,a}^{(1)}(z)$ is different from a classic SDN due to (i) the direct path from the source to aperture, (ii) the reflectance matrix, and (iii) the aperture adaptor vector. As always, aperture is accepted as an omnidirectional microphone. Hence, only the source directivity and distance attenuation terms such that $\bar{g}_{sa} = g_{sa}\Gamma_s(\theta_{sa})$ where $g_{sa} = (\|\mathbf{x}_s - \mathbf{x}_a\|)^{-1}$ are included in the direct path.

The reflectance matrix is modified so that the adjusted reflectance term in the reflectance matrix represents the CW scattering node. The aperture adaptor matrix is defined as:

$$\mathbf{k}_{as} = \mathbf{I}_{as}\mathbf{G}_{sa}\mathbf{D}_{sa}\mathbf{1}_{6\times 1} \quad (4.3.10)$$

with $\mathbf{I}_{as} = \text{diag}(1 - \delta_{1a}, \dots, 1 - \delta_{6a})$ where δ_{ia} is an indicator function. As discussed before, the indicator function is unity only on the condition that aperture and the connected scattering node, i , are on the same boundary of the volume (i.e. common wall).

Overall, the transfer function of this component is:

$$H_{s,a}^{(1)}(z) = \bar{g}_{sa}z^{-D_{s,a}} + \frac{1}{5}\mathbf{k}_a^{(1)}(z)\mathbf{F}_{s,a}^{(1)}(z)\mathbf{k}_s(z) \quad (4.3.11)$$

where $\mathbf{F}_{s,a}^{(1)}(z) = [\Sigma^{-1}\mathbf{H}_w^{(1)}(z^{-1}) - \mathbf{P}\mathbf{D}_{sa,f}^{(1)}(z)]^{-1}$.

Aperture-to-aperture: The component uses the aperture node both as a source and as a microphone. Only the recursive path is exercised. Like same volume CV-SDN aperture to aperture component exists both in the main and the coupled volume. $H_{a,a}^{(1)}(z)$ and $H_{a,a}^{(2)}(z)$ can be defined as:

$$H_{a,a}^{(q)}(z) = \frac{1}{5}\mathbf{k}_a^{(q)}(z)\mathbf{F}_{a,a}^{(q)}(z)\mathbf{k}_a^{(q)}(z), \quad q = 1, 2 \quad (4.3.12)$$

where $\mathbf{F}_{a,a}^{(q)}(z) = [\Sigma^{-1}\mathbf{H}_w^{(q)}(z^{-1}) - \mathbf{P}\mathbf{D}_{aa,f}^{(q)}(z)]^{-1}$ with $\mathbf{D}_{aa,f}^{(q)}(z)$ calculated using scattering nodes positioned on geometric centroids of bounding surfaces of the main (i.e. $q = 1$) and coupled volumes (i.e. $q = 2$), respectively. Reverberation in the

coupled volume is fed-back into the main volume and then back to the coupled volume.

Aperture-to-microphone: The component represents the transfer function from the aperture to the microphone. The aperture node acts as a source. The positions of scattering nodes that are not on the same boundary with the aperture node are calculated using microphone and aperture node positions. Similar to the source-to-aperture component, the aperture acts as an omnidirectional source resulting in a direct path gain of $\bar{g}_{am} = g_{am}\Gamma_m(\theta_{am})$ where $g_{am} = (\|\mathbf{x}_a - \mathbf{x}_m\|)^{-1}$. The aperture adaptor vector is given as $\mathbf{k}_a = \mathbf{I}_a^{(2)} \mathbf{G}_{am} \mathbf{D}_{am} \mathbf{1}_{6 \times 1}$.

The transfer function of the component is:

$$H_{a,m}^{(2)}(z) = \bar{g}_{a,m} z^{-D_{s,m}} + \frac{1}{5} \mathbf{k}_m(z) \mathbf{F}_{a,m}^{(2)}(z) \mathbf{k}_a(z). \quad (4.3.13)$$

where $\mathbf{F}_{a,m}^{(2)}(z) = [\mathbf{\Sigma}^{-1} \mathbf{H}_w^{(2)}(z^{-1}) - \mathbf{P} \mathbf{D}_{am,f}^{(2)}(z)]^{-1}$.

Aperture size control: Acoustic energy transmission is defined as discussed in 4.3. Transmission is modeled with the linkage coefficient, $\kappa = \sqrt{S_A/S_T}$ where S_A is the aperture surface area, and $S_T = S_a + S_{cw}$ is the total area of the respective volume boundary and S_{cw} is area of the CW shared between the volumes.

Reflectance transfer function is modified and given as:

$$H_{cw}^{(q)}(z) = \frac{S_A + H_i^{(q)}(z) S_{CW}}{S_{CW} + S_A} \quad (4.3.14)$$

where $H_i^{(q)}(z)$ emulates the reflectance transfer function for the material of the common surface on the side of volume q . This adjustment has the effect of increasing the energy decay rate and reducing the reverberation time.

With the components defined above, the transfer function of different volume CV-SDN can be expressed as:

$$H_{DV-CV-SDN}(z) = \kappa \frac{H_{s,a}^{(1)}(z) H_{a,m}^{(2)}(z)}{1 - H_{a,a}^{(1)}(z) H_{a,a}^{(2)}(z)}. \quad (4.3.15)$$

CHAPTER 5

EVALUATION AND DISCUSSION

In this chapter, evaluation of CV-SDN is performed and results are discussed in detail. Both settings of CV-SDN are evaluated in different settings, which are explained in detail.

Among the two settings of CV-SDN, same volume configuration is more widely examined. The reason for that is the Concave DSD pattern can only be obtained when both source and microphone are placed in the same volume. Concave DSD pattern is especially important for designing and prototyping architectural structures like concert halls as the performer and the audience are both in the same concert room, whereas several empty balconies are used to create the coupled volume effect. From a perceptual standpoint, the steeper early decay at the start of EDC provides clarity and the smaller late decay (i.e. tail decay) contributes to the reverberance effect. The perceptual evaluation in one of the studies indicates that the audience prefers a sharp decay at the start in EDC and a slower one after that [20]. In addition to the perceptual viewpoint, the adjustable aperture size is important for these studies because in concert halls there are doors whose opening can be changed.

For the evaluation of same volume CV-SDN, decay ratio (DR) and ΔdB parameters are used to assess the validity of CV-SDN algorithm. These parameters were introduced back in Section 2.1.2 of Chapter 2. Additionally, Fig. 2.5 is given to provide a clear understanding of the calculation of these parameters.

Different volume CV-SDN, on the other hand, outputs a Convex DSD pattern. The pattern is not widely investigated in terms of changing aperture size and its behavior. Nonetheless, different volume CV-SDN model results are gathered to observe the convex DSD pattern. For the evaluation, EDT is used.

For both of the settings, different volume configurations are used. The reason for this is that same volume CV-SDN settings are constructed to study concert halls. Therefore, the volumes are much bigger than a conventional coupled volume configuration. In different volume CV-SDN case, much smaller volume sizes are employed.

5.1 Same Volume CV-SDN

For Same Volume CV-SDN, a coupled volume configuration presented by Bradley and Wang [31] is used for evaluation. The parameters are compared with theoretical

values and ODEON results.¹

There are three different experiments made with the test setup. First, the effect of aperture opening size change is observed. Secondly, the absorption coefficient of the coupled volume is changed. Finally, the size of the coupled volume is changed to see its effect on decay ratio (DR) and ΔdB parameters as they are given in [31].

Compared to ODEON, CV-SDN can not be classified as computational acoustics program as in its current form it only functions as a digital artificial reverberator. Moreover, CV-SDN does not employ ray tracing and performs room acoustics modeling with ISM as explained in Chapter 4. It should be inserted that CV-SDN approximates the second and higher order reflections similar to SDN. The parameters calculated with ODEON are taken from [31].

Both theoretical values and ODEON results are provided in the study of Bradley and Wang [31]. In the paper, decay ratio (DR) is calculated as:

$$DR = \frac{Decay\ 2}{Decay\ 1}, \quad (5.1.1)$$

where Decay 1 is the time decay of the main, Decay 2 is the time decay of the coupled volume. Each time decay can be calculated with RT_{60} of both main and coupled volumes. As time decay is inversely proportional to the slope, DR can be more easily calculated with:

$$DR = \frac{m_1}{m_2}, \quad (5.1.2)$$

where m_1 represents the slope of early decay and m_2 represents the slope of late decay.

The calculation of ΔdB in the paper is given as:

$$\Delta dB = 10 \log\left(\frac{A_{1S}A_{2S}}{S^2}\right), \quad (5.1.3)$$

where S is the surface area of the aperture and $A_{iS} = A_i\alpha_i + S$ with A_i and α_i being the surface area and average absorption coefficient of the volume i respectively.

5.1.1 Test Setup

In the configuration, the main volume which contains both the source and the microphone has the dimensions $34 \times 28 \times 26$ m and the coupled volume has the dimensions $21 \times 18 \times 13$ m. Taking the left bottom corner of the main volume as the origin point, the omnidirectional source and the microphone are positioned at $\mathbf{x}_s = (25, 19, 8)$ and

¹ ODEON is a computational acoustics simulation program that utilizes ray-tracing to model the acoustics of a room [46]. The program is tested and confirmed to be suitable as an acoustics tool [47].

$\mathbf{x}_m = (5, 19, 8)$. The main volume is set to be less reverberant with $\alpha = 0.25$, whereas coupled volume is more reverberant with $\alpha = 0.02$. The aperture was positioned at the center of the CW. Different aperture areas of from 2%, 5%, 10%, 20%, and 40% were simulated and EDCs are obtained from the calculated room impulse responses. Two metrics, DR and ΔdB , were calculated and compared with the theoretical and ODEON results.

Fig. 5.1 shows the 3D model of the test setup used. In the drawing, x , y , and z axes are shown via vectors. Source and the microphone are placed in the main volume. With such positioning, same volume CV-SDN algorithm is used. Additionally, having a less reverberant main volume with $\alpha = 0.25$ is fit to output a concave DSD profile. Aperture depicted in the figure can be thought of as a window whose size can be adjusted.

5.1.2 Results and Discussion

Fig. 5.2 shows the EDC curves obtained for the different aperture sizes. The figure clearly shows the double-sloped decay behavior and the shift in the transition time between the two decays. Each EDC is represented with a different color.

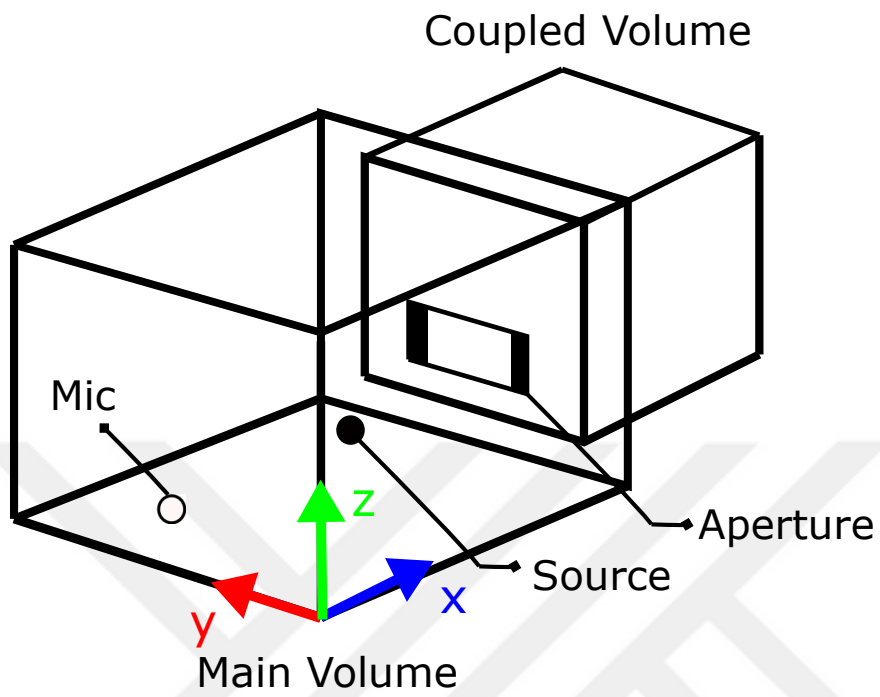
Table 5.1: Aperture Opening Size Results compared to Theoretical Values

Aperture Opening	ΔdB (CV-SDN)	ΔdB (Theoretical)	DR (CV-SDN)	DR (Theoretical)
2%	43.03	34.23	7.12	6.11
5%	36.14	26.56	6.06	5.27
10%	30.51	21.73	5.2	4.33
20%	25.32	17.32	4.38	3.23
40%	21.15	13.23	3.04	2.32

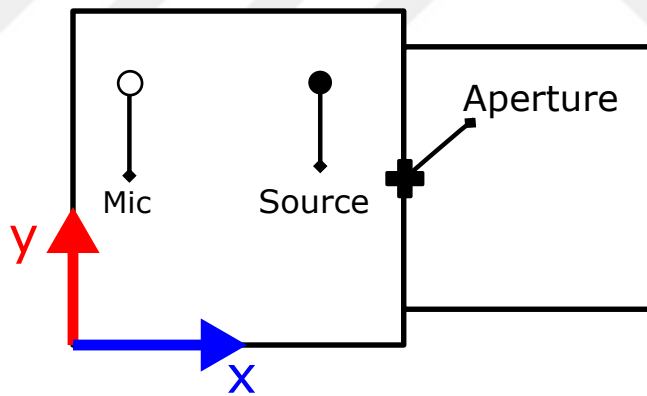
Table 5.2: Aperture Opening Size Results compared to ODEON Values

Aperture Opening	ΔdB (CV-SDN)	ΔdB (ODEON)	DR (CV-SDN)	DR (ODEON)
2%	43.03	12.10	7.12	1.91
5%	36.14	0	6.06	1
10%	30.51	0	5.2	1
20%	25.32	0	4.38	0.98
40%	21.15	15.42	3.04	3.44

There is a total of five decay curves each representing a different aperture opening size. For each curve, starting decays are pretty similar since only the reverberation in the main volume is present. The slight change can be seen due to the modified



(a)



(b)

Figure 5.1: Test Setup of the same volume CV-SDN with (a) showing 3D and (b) shown 2D representation.

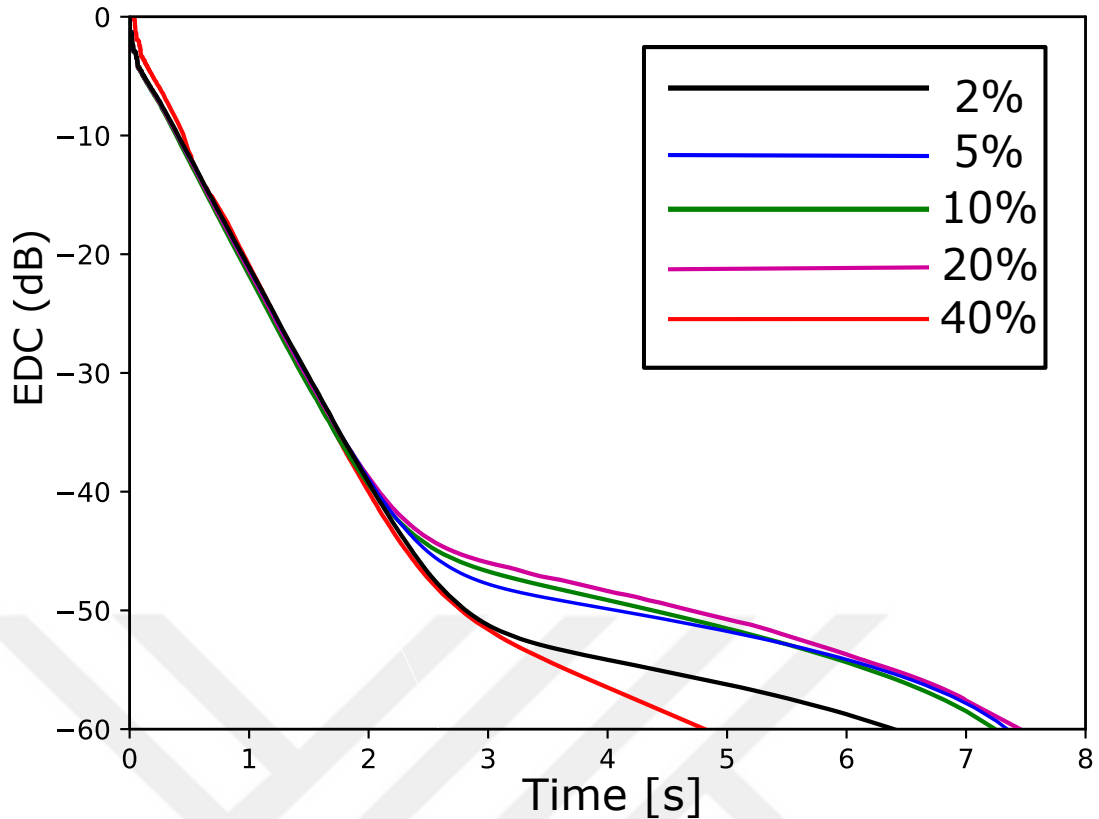


Figure 5.2: Same Volume CV-SDN EDC results with test setup.

absorption coefficient of CW via the adjustable aperture. However, the absorption coefficient of the CW on the main volume side does not make a dramatic change as the main volume is not reverberant and has an absorption coefficient of 0.25. In other words, an increase in aperture opening size results in a higher reverberant surface on CW. This results in a slight increase in the absorption coefficient of the main volume and a slightly steeper slope at the beginning of EDC.

As the aperture area expands, the absorption coefficient of the CW on the coupled volume side changes as well. Although this increases the decay slope in both volumes, it has a more pronounced effect on the coupled volume which has an absorption coefficient of 0.2, which is much smaller compared to the main volume. The rate of decay in the tail gets higher as the aperture size is increased. This results in the decreasing trend in the DR as also predicted by the theoretical results given in [31].

Further, the transmitted energy from main to coupled volume modeled with linkage coefficient κ increases with increasing aperture size causing the coupled volume to dominate EDC at an earlier time. As the surface area of CW stays constant, the aperture opening size is directly proportional to κ . In the case of aperture opening size being zero, same volume CV-SDN model only simulates a single volume enclosure, which is equivalent to the classic SDN.

Table 5.1 shows the exact values of ΔdB and DR values obtained by CV-SDN compared with theoretical values. While the DR obtained with the CV-SDN is slightly

overestimated, it closely follows the general trend in the theoretical results. Similarly, although ΔdB values shown in Table 5.1 are overestimated by approximately $8dB$ for all cases, the results also follow the decreasing trend in the theoretical results.

Table 5.2 displays the comparison between CV-SDN and ODEON values. From the data, it can be seen that ODEON fails to generate a concave DSD pattern for aperture openings of 5%, 10%, and 20% since DR is equal to 1. On other opening sizes, only 40% shows a comparable DR to theoretical values. Whereas CV-SDN can output a concave DSD pattern for all aperture openings. Similarly, ΔdB values indicate the same phenomenon. The reason why ODEON produces 0 is due to both early and late decays coinciding at the same y-intercept, which is only explained with a single sloped decay. The lack of DSD profile output of ODEON can be explained by the room acoustic modeling method, which is ray tracing. As the volume used in the test setup gets larger, the cost of calculating every reflection increases. Approximations made in the model of ODEON may result in the loss of DSD profile.

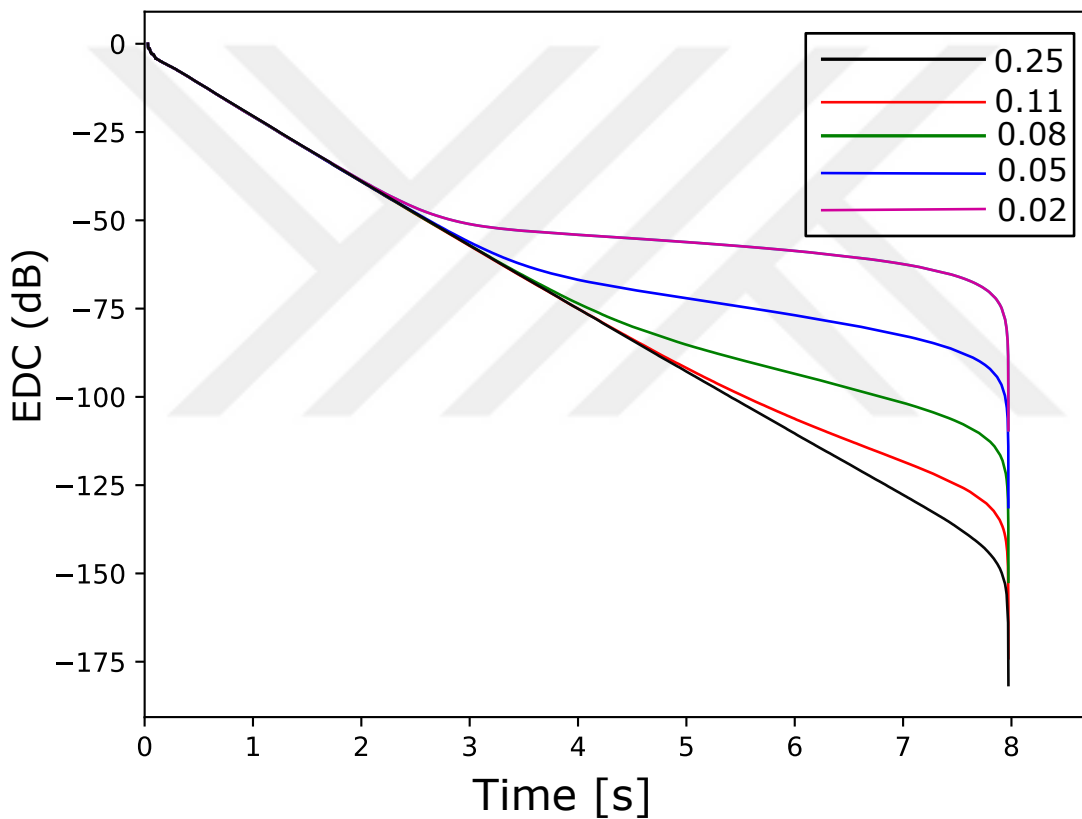


Figure 5.3: Same Volume CV-SDN Absorption EDC results with test setup.

Fig. 5.3 shows the impact of change in absorption coefficient of coupled volume. Four different values as 0.02, 0.05, 0.08, 0.11 are used. The fifth one, shown as the black line in the figure represents the case when the absorption coefficient of both volumes are the same. As we increase the absorption coefficient, first and late decays converge to each other. Moreover, the transition point shifts to the right with increased absorption as the samples coming from the coupled volume are weaker compared to cases where the absorption coefficient of the coupled volume is low.

Table 5.3: Absorption Coefficient Results Compared to Theoretical Values

Absorption Coefficient	ΔdB (CV-SDN)	ΔdB (Theoretical)	DR (CV-SDN)	DR (Theoretical)
0.02	43.03	34.28	7.12	6.12
0.05	44.08	37.98	3.45	2.82
0.08	46.77	39.83	2.16	1.79
0.11	38.37	41.07	1.47	1.28

Table 5.4: Absorption Coefficient Results Compared to ODEON Values

Absorption Coefficient	ΔdB (CV-SDN)	ΔdB (ODEON)	DR (CV-SDN)	DR (ODEON)
0.02	43.03	12.10	7.12	1.91
0.05	44.08	11.02	6.06	1.76
0.08	46.77	11.76	5.2	1.79
0.11	38.37	10.19	4.38	1.708

Overall, increasing absorption coefficient results in decreased DR values. For ΔdB values, a slight increase is observed due to late decay dominating EDC at a later time.

Table 5.3 shows the exact values of ΔdB and DR values obtained by CV-SDN compared with theoretical values. While the DR obtained with the CV-SDN is slightly overestimated, it closely follows the general trend in the theoretical results.

Table 5.4 displays the comparison between CV-SDN and ODEON values. As previously, ODEON fails to generate meaningful results with changing absorption coefficients.

Table 5.5: Coupled Volume Ratio Results Compared to Theoretical Values

Volume Ratio	ΔdB (CV-SDN)	ΔdB (Theoretical)	DR (CV-SDN)	DR (Theoretical)
20%	43.03	34.28	7.12	6.12
35%	45.16	37.75	8.62	7.20
50%	50.85	39.9	9.14	8.57

Fig. 5.4 shows the results due to increased size of the coupled volume. Three different ratios as 20%, 35%, and 50% are used. As the coupled volume ratio (i.e. ratio of coupled volume size to the main volume size) increases, we see tail decay dominates at a later time. In short, the effect is minimal compared to other test setups.

Table 5.5 shows the exact values of ΔdB and DR values obtained by CV-SDN compared with theoretical values. While the DR obtained with the CV-SDN is slightly

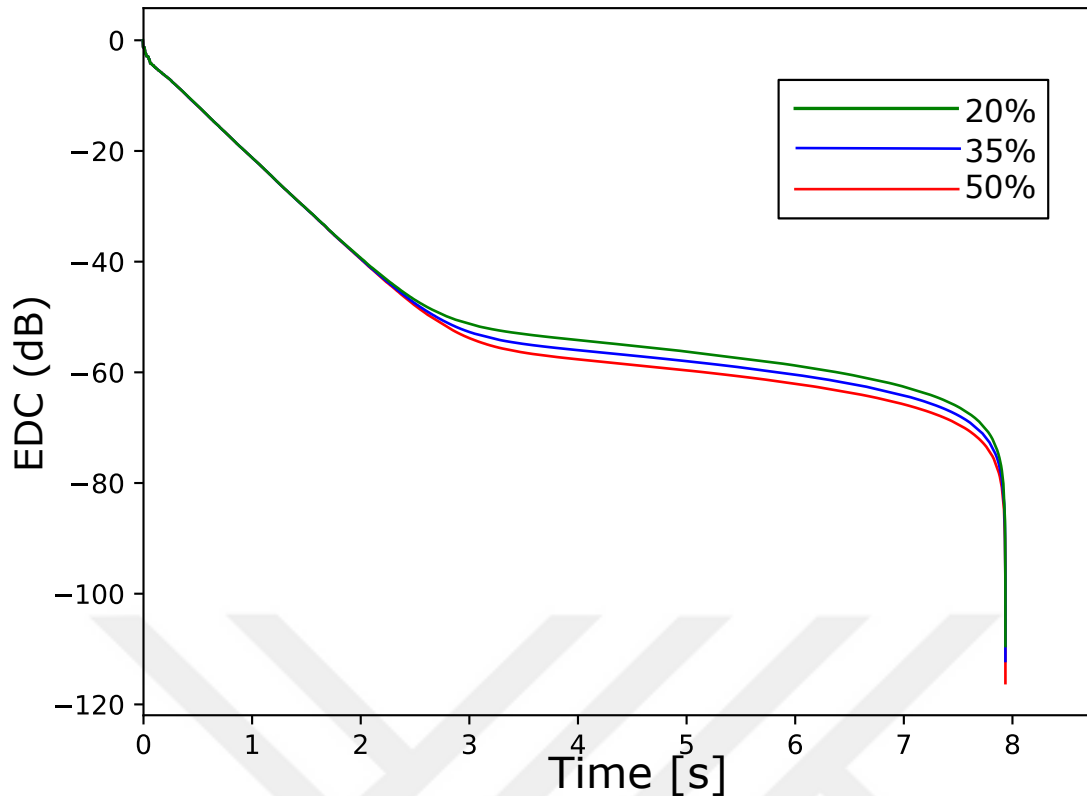


Figure 5.4: Same Volume CV-SDN Coupled Volume Ratio EDC results with test setup.

overestimated, it closely follows the general trend in the theoretical results.

Table 5.6 displays the comparison between CV-SDN and ODEON values. As previously, ODEON fails to generate meaningful results with changing coupled volume ratio.

5.2 Different Volume CV-SDN

For different volume CV-SDN, coupled volume configuration presented by Pu. et al [28] is used. Although the study does not focus on the parametrical evaluation of convex DSD patterns, it presents a model which can output an EDC where the early decay is slow and late decay is steep. In the study, a scale model is used and the results are compared with experimental findings and with geometrical acoustical modeling used in ODEON. Similarly, ODEON results are taken from the paper.

In addition to the observation of convex DSD pattern, the effect of aperture size adjustment of CV-SDN is discussed. It should be noted that the aperture opening size adjustment is not investigated for the configurations where source and microphone are placed in the separate volumes.

Table 5.6: Coupled Volume Ratio Results Compared to ODEON Values

Volume Ratio	Δ dB (CV-SDN)	Δ dB (ODEON)	DR (CV-SDN)	DR (ODEON)
20%	43.03	11.55	7.12	1.92
35%	45.16	15.75	8.62	2.17
50%	50.85	8.4	9.14	2.03

5.2.1 Test Setup

The test configuration utilizes the source placed in the main volume and the microphone in the coupled volume. The main volume has the dimensions $2.75 \times 5 \times 4.75$ and the coupled volume has the dimensions $5.5 \times 5 \times 4.75$ m. The left bottom corner is taken as the origin point. The omnidirectional source and microphone are placed at $\mathbf{x}_s = (1, 1, 2.38)$ and $\mathbf{x}_m = (4, 3.75, 2.38)$. To obtain a convex DSD profile, the main volume is set to be the less reverberant volume with $\alpha = 0.6$. The coupled volume is the more reverberant one with $\alpha = 0.05$. The aperture is a rectangular structure and positioned at the center of CW. The study presented by Pu. et al [28] uses a single aperture opening. In the configuration, the aperture has the dimensions of 1.25×1.25 , which results in the opening size of 7% of the CW. Apart from the study, five additional opening sizes selected in same volume CV-SDN evaluation are included to examine the behavior of different volume CV-SDN with changing aperture size. EDCs of the configuration are obtained with the calculation of room impulse responses.

Fig. 5.5 displays the 3D and 2D representation of the same configuration of the different volume CV-SDN. The source is placed at the bottom left corner of the main volume. The microphone is closer to the aperture and placed at the upper left corner of the coupled volume. Aperture is at the center of the CW, and quite small compared to the one used in same volume CV-SDN.

5.2.2 Results and Discussion

The model used by Pu. et al [28] has an EDT of (i.e. first 10 dB drop below initial level) within 0.4 seconds. Although the initial constant decay is very brief, different volume CV-SDN has an EDT of approximately 4.1 seconds. The results obtained are in agreement with the experimental convex DSD curve given in [28].

Fig. 5.6 shows the convex EDC of different volume CV-SDN. It may be observed that the discrete samples in the initial decay are not displayed as exercised by Eyring in [27]. These samples disrupt the initial part of EDC and make it harder to see the constant slope observed in the initial stages. Concavity of the graph can be seen as the slope gets steeper at higher lags. Concavity obtained in different volume CV-SDN can be explained by the feedback path that generates between the coupled and the main volume. After a certain time, the dominant samples are the ones looping through the

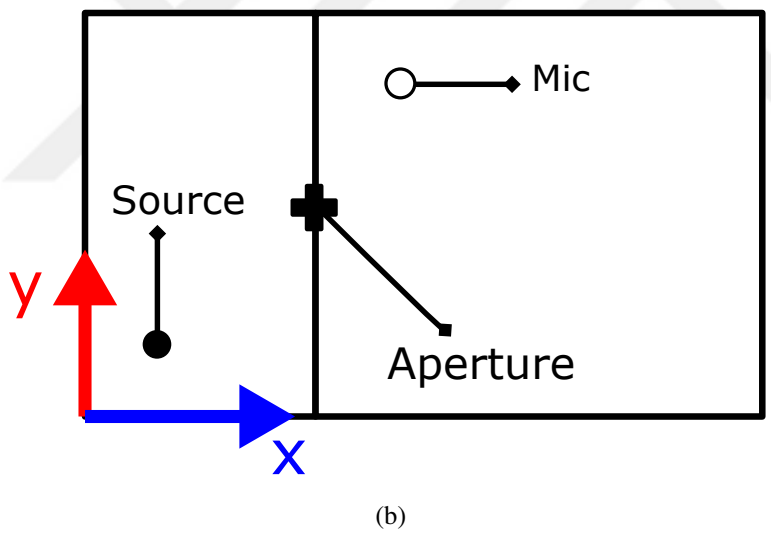
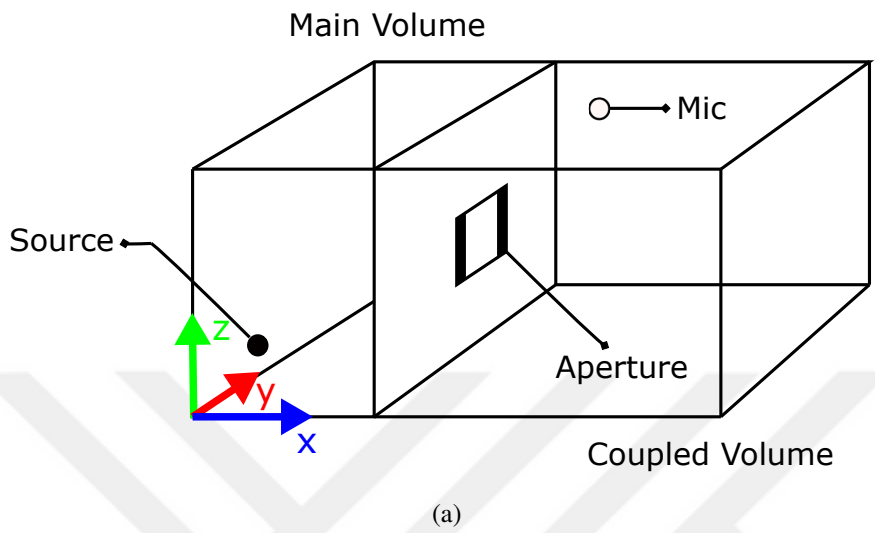


Figure 5.5: Test Setup of the different volume CV-SDN with (a) showing 3D and (b) shown 2D representation.

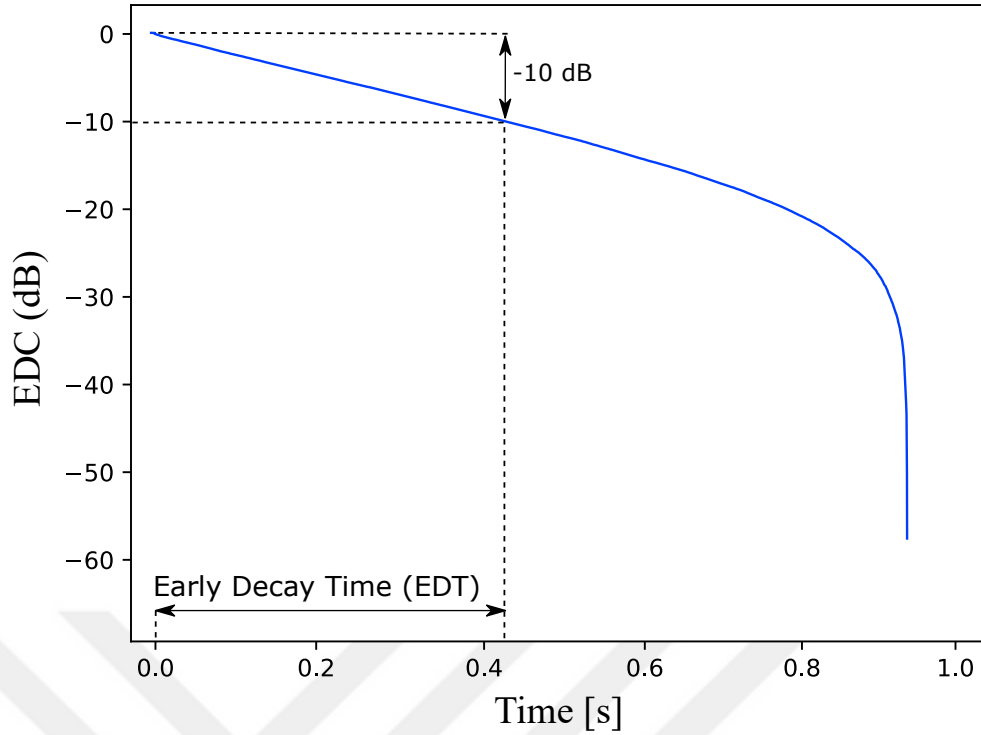


Figure 5.6: Different volume CV-SDN EDC result with test setup.

feedback path. In the path, the rate of decay is higher as samples are reverberated both in the coupled volume and the main volume.

In addition to the opening size of 7% used in [28], two additional openings with 5% and 10% are tested to examine the impact of adjustable aperture. The results are shown in Fig. 5.7. The solid line at the bottom, long dashed line, short dashed line, dotted line, dashed-dotted line, and solid line at the top represent 2%, 5%, 7%, 10%, 20%, and 40% respectively. The discrete samples in the initial decay are included to show the dB drop in the initial stages. As the aperture opening size decreases, the dB drop is increased. This is due to decreased linkage coefficient κ . In other words, the energy exchange from the main volume to the coupled volume is reduced. Moreover, the decay rate also increases with decreasing aperture. As the varying aperture size changes the absorption coefficient of the CW, the coupled volume gets more reverberant with decreasing aperture opening size. More reverberant environment results in a less steep decay, which is indicated by the solid line. Hence, EDT decreases with increasing aperture opening size. If the aperture opening size is higher, the overall absorption coefficient of the coupled volume increases dramatically. This effect disturbs the convex DSD effect and we see almost a flat line for the solid line representing 40% opening size in the figure. EDT parameter values are given in 5.7.

Similar to same volume CV-SDN, change in aperture opening size does not make a drastic change in the overall absorption coefficient of the main volume. Thus its effect cannot be observed from EDC.

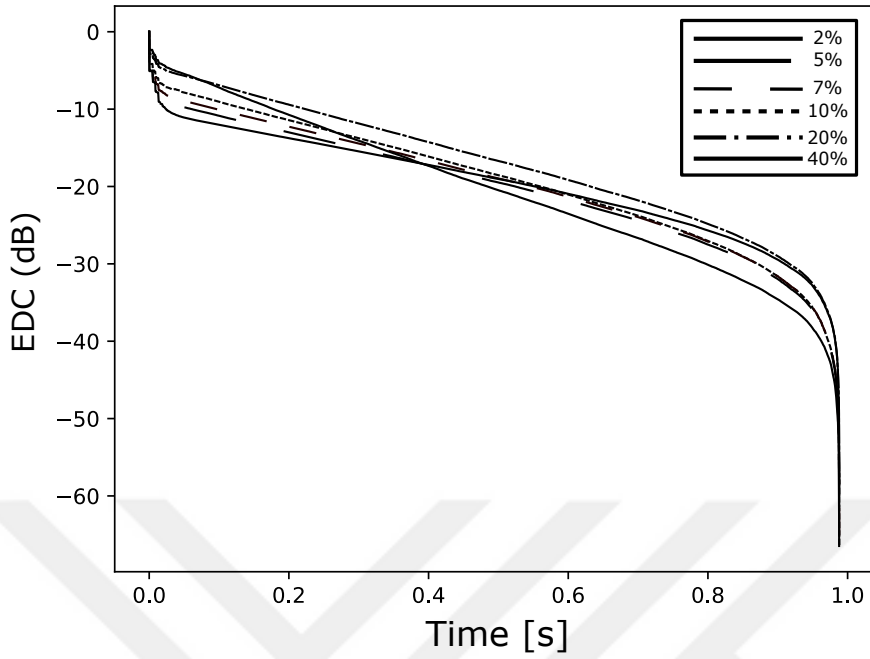


Figure 5.7: Different volume CV-SDN EDC result with different aperture openings.

Table 5.7: Different volume CV-SDN EDT Parameter comparison

Aperture Opening	EDT (CV-SDN)	EDT (ODEON)	EDT (Theoretical)
2%	0.44	—	0.41
5%	0.42	—	0.41
7%	0.41	0.39	0.40
10%	0.40	—	0.40
20%	0.39	—	0.39
40%	0.37	—	0.38

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

The rise in computer processing power has led to more realistic and reliable graphics and audio in the video-games and extended reality applications. Hence, yesterdays costly procedural generation of content is becoming more popular than ever. Due to the need for realistic audio in such applications, acoustic simulation of enclosures and thus reverberation is important as well. Before, the audio was used from pre-recorded samples or recorded in the real environment. This method of reverberation is not only costly but cumbersome as well. Digital reverberation, on the other hand, offers a more convenient platform in which reverberated output of the sound can be adjusted with parameters such as source and microphone positions, directivities, and room properties. By using a digital artificial reverberator, not only customizable audio is possible but also the reduced cost of the generation of audio.

Development of digital reverberators has gained momentum with the model proposed by Schroeder and they are still being improved to-date. SDN is one of those digital reverberators that can simulate first-order reflections very accurately and has a gracefully degrading accuracy for the higher order reflections. It can successfully simulate the acoustics of a rectangular enclosure and the results, such as RT_{60} match the theoretical values. However, the drawback of SDN lies in its strength. It can only function on rectangular and single enclosures. The need for coupled volume settings is apparent as there are studies on concert halls where several volumes are connected with an aperture to each other [20]. Additionally, applications in mixed reality and video games often simulate more than one room and depend on the generation of realistic reverberation for multiple volumes.

A new model called CV-SDN was proposed in this thesis by expanding upon the classic SDN architecture. CV-SDN introduces a coupled volume and an aperture. The aperture is modeled as a rectangular opening like a window or a door, which has height and width values. The model allows the adjustment of aperture size and successfully generates the expected EDC curves where convex and concave DSD profiles can be observed.

There are two possible settings of CV-SDN depending on the positions of the source and the microphone. First setting simulates coupled volumes where source and microphone are placed in the main volume and is called same volume CV-SDN. The first setting is constructed with the connection of five different SDNs. The second setting,

different volume CV-SDN, is realized when the source and the microphone are placed in the separate volumes. Four SDNs are required to model the second setting. In both of the settings, the main volume is set to be more reverberant one so that both DSD patterns can be observed from EDCs.

The transfer function of each setting is derived in detail. Test setups are created for each setting and emphasis is put on the same volume CV-SDN as it is more widely investigated in the room acoustics field. Test setup of the same volume CV-SDN observes an aperture with different opening sizes. Two parameters, DR , and ΔdB are calculated from EDCs and compared with theoretical results. Additionally, parameter values generated with ODEON, a computational acoustics simulation program, are compared with those of CV-SDN. It is seen that CV-SDN matches the pattern observed with the theoretical values. Same Volume CV-SDN can maintain the concave DSD pattern through all of the aperture openings, whereas ODEON fails to generate a DSD profile for some of the openings. This proves that CV-SDN can be used as a basis for the architectural modeling of concert halls.

The different volume CV-SDN is tested to show that it can generate convex DSD profile. A test setup is created and EDT parameters are calculated. The results are compared both with theoretical and ODEON values and proved that different volume CV-SDN can simulate the convex DSD phenomenon. Additionally, the effect of the aperture opening size is discussed. However, the accuracy of aperture area adjustment in different volume CV-SDN is not compared to other studies as the focus in the field mainly lies in the configurations where source and microphone rest together in the same volume.

6.2 Possible use cases of CV-SDN

CV-SDN in its current form can be used in different ways.

The most prominent one coming from simulation viewpoint is the XR (Extended Reality) applications which consist of virtual, augmented and mixed reality. In a computer simulation, it is quite important to create environments where audio and hence reverberation is realistic. Using a single room artificial reverberator such as SDN is possible but simulations rarely end up with a single volume setting. Hence, employing CV-SDN in rooms with openings such as doors or windows can lead to a more accurate representation of reverberated audio. Increased realism in such settings could provide more entertainment and value for the customers.

The second use field is architectural acoustics. As it stands, with CV-SDN it can employ a concert hall effect on raw audio due to DSD phenomena. By placing the source (i.e. performer) and microphone (i.e. audience) in the main volume, we can create an effect which transforms the audio as if it is recorded in a coupled volume setting such as a concert hall. More importantly, with CV-SDN we can create coupled volumes with different absorption coefficients and sizes, and then study the energy decay curve of the given room. The model proposed with this thesis can be also used in Building Information Modelling (BIM) Framework [48]. By changing aperture size and area, the absorption coefficient of coupled volume, and coupled volume size

we can manipulate EDC to suit our needs. CV-SDN can be used by architects in the field to gather an approximation on what the response of a concert hall would be.

The third area of use of CV-SDN is in crisis scenarios and forensic studies. XR applications are already being utilized to train firefighter staff [49]. Additionally, audio recordings of gunshots are being studied in simulation settings [50]. Similarly to first use area given, XR applications can be constructed where individuals can be trained on how hazardous things sound. The simulations could be constructed with different volume CV-SDN settings. For example, a simulation environment for firefighters can be utilized so that they can be trained to differentiate the sound whether there is a fire in the next room or not. Moreover, gunshot recordings could be created with raw gunshot audio to identify in what type of room a gun is fired.

6.3 Possible Future Work

CV-SDN uses an aperture but does not make a distinction between the different shapes of it. As the only impact of the aperture is due to its opening size, improvements on CV-SDN model can be made to differentiate between aperture openings that have the same opening area but distinct shapes. The addition of edge diffraction modeling can create a more accurate system which may give results much closer to the theoretical values. The edge diffraction modeling at the same time could make the acoustics of each volume more accurate individually.

In its current form, CV-SDN is not tested to be run in interactive rates. With the introduction of real-time simulation to CV-SDN, use cases mentioned in the previous section can be more easily employed. The original SDN was able to run at interactive rates with iPhone 5. It is quite possible to manage CV-SDN which employs four to five SDN structures so that it runs in real-time.

Similar to SDN, CV-SDN only functions with rectangular enclosures. Although the model is a rough approximation, extensions could be made so that arbitrarily shaped coupled volumes can be simulated as well. For this, edge diffraction and generation of image sources will be important, and the placement of aperture size will be crucial for the results. Afterward, The improved model can be used to simulate concert halls that have complex interiors. This could lead to more accurate modeling of such architectural structures.

Moreover, as CV-SDN is based on two volumes, implementation of additional enclosures may be discussed. By doing this, the acoustical design of more complex structures can be more accurately created as there are several balconies connected to the main hall. The introduction of additional volumes will require the use of further SDN models. It would be a challenge to keep a computationally efficient model while adding SDN models to the whole architecture. When converting SDN to CV-SDN, the same volume CV-SDN required four, and different volume CV-SDN structure required three additional SDN structures. These inclusions in the model were due to feedback path emerging between the rooms. If we are to introduce even more coupled volumes to the model, the number of SDNs could increase exponentially due to feedback paths forming between the volumes. Furthermore, each included aperture

will also change the absorption coefficient of the surface they are attached to. For example, if we connect four volumes as coupled volumes to the main volume to simulate a concert hall setting. Each surface of the main volume will have walls with different absorption coefficients. Perhaps a frequency-domain implementation can be undertaken to overcome the computational complexity problem. Nevertheless, if the expanded model results in a high and unmanageable computational complexity, it could be possible to make approximations so that feedback paths are minimized and the resulting number of additional SDNs could increase linearly instead of exponentially.

Perceptual evaluation of CV-SDN can be performed and is especially important for same volume CV-SDN setting. The design of architectural structures such as concert halls takes auditory pleasure of the audience into account. As the concave DSD response is preferred according to studies, audio outputs of CV-SDN can be tested on individuals to see which aperture opening size they prefer. Moreover, the comparison of the sound generated in both settings of CV-SDN can also be made. Recordings of performers taken in concert halls could be compared to the raw audio of performers simulated to create reverberance effect in CV-SDN.

REFERENCES

- [1] V. Valimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, “Fifty years of artificial reverberation,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 5, pp. 1421–1448, 2012.
- [2] W. G. Gardner, “Reverberation algorithms,” in *Applications of digital signal processing to audio and acoustics*, pp. 85–131, Springer, 2002.
- [3] V. Välimäki, J. Parker, L. Savioja, J. O. Smith, and J. Abel, “More than 50 years of artificial reverberation,” in *Audio Engineering Society Conference: 60th International Conference: DREAMS (Dereverberation and Reverberation of Audio, Music, and Speech)*, Audio Engineering Society, 2016.
- [4] M. R. Schroeder and B. F. Logan, ““colorless” artificial reverberation,” *IRE Transactions on Audio*, no. 6, pp. 209–214, 1961.
- [5] M. R. Schroeder, “Natural sounding artificial reverberation,” *Journal of the Audio Engineering Society*, vol. 10, no. 3, pp. 219–223, 1962.
- [6] J. Stautner and M. Puckette, “Designing multi-channel reverberators,” *Computer Music Journal*, vol. 6, no. 1, pp. 52–65, 1982.
- [7] J.-M. Jot and A. Chaigne, “Digital delay networks for designing artificial reverberators,” in *Audio Engineering Society Convention 90*, Audio Engineering Society, 1991.
- [8] S. Bilbao and J. Smith, “Finite difference schemes and digital waveguide networks for the wave equation: Stability, passivity, and numerical dispersion,” *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 3, pp. 255–266, 2003.
- [9] D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, “Acoustic modeling using the digital waveguide mesh,” *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 55–66, 2007.
- [10] M. Karjalainen and C. Erkut, “Digital waveguides versus finite difference structures: Equivalence and mixed modeling,” *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 978–989, 2004.
- [11] J. O. Smith, “A New Approach to Digital Reverberation Using Closed Waveguide Networks,” in *Proc. 1985 Int. Computer Music Conf.*, (Vancouver, Canada), pp. 47–53, 1985.
- [12] P. Huang, M. Karjalainen, and J. O. Smith, “Digital waveguide networks for room response modeling and synthesis,” in *Audio Engineering Society Convention 118*, Audio Engineering Society, 2005.

- [13] E. De Sena, H. Hacıhabiboglu, and Z. Cvetkovic, “Scattering delay network: An interactive reverberator for computer games,” in *Audio Engineering Society Conference: 41st International Conference: Audio for Games*, Audio Engineering Society, 2011.
- [14] H. Hacıhabiboglu, E. De Sena, and Z. Cvetkovic, “Frequency-domain scattering delay networks for simulating room acoustics in virtual environments,” in *2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems*, pp. 180–187, IEEE, 2011.
- [15] E. De Sena, H. Hacıhabiboğlu, Z. Cvetković, and J. O. Smith, “Efficient synthesis of room acoustics via scattering delay networks,” *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 23, no. 9, pp. 1478–1492, 2015.
- [16] A. D. Baldwin, S. Serafin, and C. Erkut, “Scattar: a mobile augmented reality application that uses scattering delay networks for room acoustic synthesis,” in *23rd ACM Symposium on Virtual Reality Software & Technology Virtual Reality Software and Technology*, Association for Computing Machinery, 2017.
- [17] A. Baldwin, S. Serafin, and C. Erkut, “Towards the design and evaluation of delay-based modeling of acoustic scenes in mobile augmented reality,” in *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, pp. 1–5, IEEE, 2018.
- [18] S. Djordjević, “Perceptual evaluation of room acoustics simulation for computer games via scattering delay networks,” 2019.
- [19] F. Stevens, D. T. Murphy, L. Savioja, and V. Välimäki, “Modeling sparsely reflecting outdoor acoustic scenes using the waveguide web,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 8, pp. 1566–1578, 2017.
- [20] M. Ermann, “Coupled volumes: Aperture size and the double-sloped decay of concert halls,” *Building Acoustics*, vol. 12, no. 1, pp. 1–13, 2005.
- [21] G.-B. Stan, J.-J. Embrechts, and D. Archambeau, “Comparison of different impulse response measurement techniques,” *Journal of the Audio Engineering Society*, vol. 50, no. 4, pp. 249–262, 2002.
- [22] J. O. Smith, *Physical audio signal processing: For virtual musical instruments and audio effects*. W3K publishing, 2010.
- [23] H. Kuttruff, *Room acoustics*. Crc Press, 2016.
- [24] C. B. Boyer and U. C. Merzbach, *A history of mathematics*. John Wiley & Sons, 2011.
- [25] W. C. Sabine, *Collected papers on acoustics*. Harvard University Press Cambridge, MA, 1927.
- [26] C. F. Eyring, “Reverberation time in “dead” rooms,” *The Journal of the Acoustical Society of America*, vol. 1, no. 2A, pp. 217–241, 1930.

- [27] C. F. Eyring, "Reverberation time measurements in coupled rooms," *The Journal of the Acoustical Society of America*, vol. 3, no. 2A, pp. 181–206, 1931.
- [28] H. Pu, X. Qiu, and J. Wang, "Different sound decay patterns and energy feedback in coupled volumes," *The Journal of the Acoustical Society of America*, vol. 129, no. 4, pp. 1972–1980, 2011.
- [29] N. Xiang, J. Escolano, J. M. Navarro, and Y. Jing, "Investigation on the effect of aperture sizes and receiver positions in coupled rooms," *The Journal of the Acoustical Society of America*, vol. 133, no. 6, pp. 3975–3985, 2013.
- [30] B. W. Harrison, G. Madaras, and R. D. Celmer, "Computer modeling and prediction in the design of coupled volumes for a 1000-seat concert hall at goshen college, indiana," *The Journal of the Acoustical Society of America*, vol. 109, no. 5, pp. 2388–2388, 2001.
- [31] D. T. Bradley and L. M. Wang, "The effects of simple coupled volume geometry on the objective and subjective results from nonexponential decay," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1480–1490, 2005.
- [32] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.
- [33] G. Kendall, W. Martens, D. Freed, D. Ludwig, and R. Karstens, "Image model reverberation from recirculating delays," in *Audio Engineering Society Convention 81*, Audio Engineering Society, 1986.
- [34] J. Borish, "Extension of the image model to arbitrary polyhedra," *The Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1827–1836, 1984.
- [35] A. Krokstad, S. Strom, and S. Sørdsal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, no. 1, pp. 118–125, 1968.
- [36] M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *The Journal of the acoustical society of america*, vol. 47, no. 2A, pp. 424–431, 1970.
- [37] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan, "A beam tracing method for interactive architectural acoustics," *The Journal of the acoustical society of America*, vol. 115, no. 2, pp. 739–756, 2004.
- [38] M. Rettinger, "Reverberation chambers for broadcasting and recording studios," *Journal of the Audio Engineering Society*, vol. 5, no. 1, pp. 18–22, 1957.
- [39] B. A. Blesser, "An interdisciplinary synthesis of reverberation viewpoints," *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 867–903, 2001.
- [40] H. Laurens, "Electrical musical instrument," Feb. 4 1941. US Patent 2,230,836.
- [41] J. Parker and S. Bilbao, "Spring reverberation: A physical perspective," in *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx'09)*, pp. 416–421, 2009.

- [42] W. Kuhl, “The acoustical and technological properties of the reverberation plate,” *EBU Review, Part A-Technical*, vol. 49, pp. 8–14, 1958.
- [43] J. A. Moorer, “About this reverberation business,” *Computer music journal*, pp. 13–28, 1979.
- [44] J.-M. Jot, “An analysis/synthesis approach to real-time artificial reverberation,” in *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 221–224, IEEE, 1992.
- [45] W. Sabine, “Collected papers on acoustics: Paper 1—reverberation,” 1922.
- [46] C. L. Christensen, “Odeon, a design tool for auditorium acoustics, noise control and loudspeaker systems,” in *Proceedings of Reproduced Sound 17: Measuring, Modelling or Muddling*, pp. 137–144, 2001.
- [47] D. T. Bradley and L. M. Wang, “Comparison of real world measurements and computer model results for a dedicated coupled volume system,” *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2552–2552, 2004.
- [48] B. Succar, “Building information modelling framework: A research and delivery foundation for industry stakeholders,” *Automation in construction*, vol. 18, no. 3, pp. 357–375, 2009.
- [49] J. M. Houtkamp, A. Toet, and F. A. Bos, “Task-relevant sound and user experience in computer-mediated firefighter training,” *Simulation & Gaming*, vol. 43, no. 6, pp. 778–802, 2012.
- [50] B. M. Brustad and J. C. Freytag, “A survey of audio forensic gunshot investigations,” in *Audio Engineering Society Conference: 26th International Conference: Audio Forensics in the Digital Age*, Audio Engineering Society, 2005.