

Towards Silent Open Windows

An Open-loop Wave-domain Algorithm for Control of Noise through Apertures

D. Ratering

Master of Science Thesis

Towards Silent Open Windows

An Open-loop Wave-domain Algorithm for Control of Noise through Apertures

MASTER OF SCIENCE THESIS

For the degree of Master of Science in Systems and Control at Delft
University of Technology

D. Ratering

October 26, 2021

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of
Technology

GN

The work in this thesis was supported by GN Group. Their cooperation is hereby gratefully acknowledged.



Copyright © Delft Center for Systems and Control (DCSC)
All rights reserved.



DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
DELFT CENTER FOR SYSTEMS AND CONTROL (DCSC)

The undersigned hereby certify that they have read and recommend to the Faculty of
Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis
entitled

TOWARDS SILENT OPEN WINDOWS

by

D. RATERING

in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE SYSTEMS AND CONTROL

Dated: October 26, 2021

Supervisor(s):

prof.dr. W.B. Kleijn

dr.ir. R. Ferrari

ir. J. Gonzalez Silva

Reader(s):

dr.ir. J.A. Martínez Castañeda

Abstract

Noise pollution is a major health threat to society. Active noise control systems that attenuate noise through open windows (apertures) have the potential to create quieter homes while maintaining ventilation and sight. Such systems are commonly realized with closed-loop least-mean-square (LMS) algorithms. However, these algorithms require a large number of error microphones inside the room and provide only local attenuation. Using many error microphones leads to slow convergence and high computational effort, having additional disadvantages. Therefore, closed-loop active noise control algorithms are undesired for real-world application. In this study, we develop an open-loop wave-domain algorithm that converges instantaneously and operates with low computational effort. As it is open-loop, it does not require error microphones. We position a control region in the far-field that covers all directions from the aperture into the room. In the algorithm, we minimize the sound in that control region. Hence, it inherently ensures cancellation in the whole room. We derive acoustic transfer functions to obtain frequency responses of the aperture and loudspeakers. Those are used for soundfield calculation. The sum of the soundfields, from the aperture and the loudspeaker array, is then expressed in orthonormal basis functions. By minimizing this sum in least mean square sense, we can calculate the filter-weights that minimize the sound energy in the control region. Implementation of these filter-weights with block-wise processing using the Short-Time Fourier Transform generates the signals for the loudspeaker array. However, this processing induces a delay. To compensate for this algorithmic delay, two methods are compared. The first is positioning a reference microphone further in front of the aperture. The second method uses an autoregressive model for signal prediction. Both lead to a loss in attenuation performance compared to the optimal algorithm. We compare the optimal wave-domain algorithm with a LMS-based reference algorithm, as well as both the algorithmic delay compensation methods. The algorithms are tested with a sparse and grid loudspeaker array, and we use rumbler-siren, airplane, and white noise signals as incoming noise. Furthermore, we compare performance for three incident angles. Our simulation results indicate that wave-domain processing has the potential to outperform LMS-based methods in practical active noise control for apertures. More specifically, we obtain an average -10 dB global reduction up to 2 kHz for all signal types with the optimal wave-domain approach. In comparison, the performance of the closed-loop algorithm ranges between -5.2 and -9.2 dB, depending on the

signal type. Furthermore, we indicate the limited impact of the incident angle on performance for the wave-domain algorithm. Positioning a reference microphone in front of the aperture outperforms the predictor approach in all scenarios, and its performance compares to that of the closed-loop LMS algorithms. Eventually, the absence of error microphones and inherent global control ensure that the wave-domain algorithm can be used for a practical active noise control system for apertures. Future work could improve the algorithm by reducing loss of performance with small window-sizes due to time-delay wrapping in the block-wise processing. Furthermore, a natural continuation of this study is to develop and test the wave-domain algorithm for scenarios with a moving primary noise source to further emphasize its advantage over the closed-loop LMS algorithm.

Table of Contents

Acknowledgements	xiii
1 Introduction	1
1-1 Motivation	1
1-2 Introduction to Active Noise Control for Apertures	1
1-3 Definition of Research Gap	3
1-4 Problem Formulation	4
1-5 Thesis Outline	5
2 Theoretical Background	7
2-1 Preliminaries	7
2-2 Signal Processing Concepts	8
2-2-1 Linear Transformation and Manipulation	8
2-2-2 Frame Theory	9
2-2-3 Discrete Fourier Transform	10
2-2-4 Short-time Fourier Transform	11
2-2-5 Window Functions	13
2-3 Huygens Principle	14
2-4 Multiple-error Normalized LMS Algorithm	15
2-4-1 Motivation	15
2-4-2 Algorithm	15
2-4-3 Stability and Convergence	18
2-4-4 Algorithm Parameters	18
2-5 Summary	19

3	Simulation Model and Acoustic Transfer Functions	21
3-1	Simulation Environment	21
3-1-1	3D Representation	21
3-1-2	2D Representation	23
3-2	Acoustic Transfer Functions	24
3-2-1	Aperture ATF	24
3-2-2	Loudspeaker ATF	26
3-3	Block Processing	27
3-3-1	Implementation	27
3-3-2	Time-delay Wrapping	27
3-3-3	Window-size and Frequency Resolution	29
3-4	Schematic Overview	31
3-5	Summary	32
4	Wave-Domain Algorithm Design	33
4-1	Control Region	33
4-2	Algorithm development	34
4-2-1	Soundfield Basis Expansion	34
4-2-2	Orthonormalization of a Set of Basis Functions	35
4-2-3	Soundfield Expressions	37
4-2-4	Filter-Weight Calculation	38
4-2-5	Algorithmic Delay Compensation	39
4-3	Parameters	40
4-3-1	Number of Basis Functions	40
4-3-2	Window-size	41
4-4	Summary	41
5	Simulation Results	43
5-1	Experimental Setup	43
5-1-1	Implementation	43
5-1-2	General Parameters	44
5-1-3	Primary Noise Signals	44
5-2	Evaluation Methods	45
5-2-1	Segmental SNR	45
5-2-2	Animations	46
5-3	Soundfield Modeling Results	46
5-4	3D Algorithm Results	47
5-4-1	Graphical Illustrations	48
5-4-2	Number of Basis Functions	49
5-4-3	Wave-domain algorithm with various Window-sizes	50
5-4-4	Multiple-error Normalized Least Mean Squares (NLMS) algorithm versus Optimal Wave-domain Algorithm	52
5-4-5	Grid and Sparse Array Comparison	54
5-4-6	Algorithmic Delay Compensation	56
5-5	Summary Table	58

6 Discussion	61
6-1 Simulation Environment and Acoustic Transfer Functions	61
6-2 Reference Algorithm and Optimal Wave-domain Algorithm	63
6-3 Algorithmic Delay Compensation	67
7 Conclusion	69
A Conference Paper	71
B Incoming Plane Wave Modeling	77
C Extended Soundfield Modeling Results	79
D Additional Simulation Results	81
D-1 Other 3D Array Types	81
D-2 2D Results	82
Bibliography	89
Glossary	93
List of Acronyms	93

List of Figures

2-1	The cartesian coordinate system and its related spherical coordinate system. . . .	8
2-2	Sequence of a linear transformation and coder using an invertible $n \times n$ matrix U .	9
2-3	Sequence of a linear transformation and coder using an invertible $k \times n$ matrix U .	9
2-4	Schematic form of the full linear transformation matrix UW of the Short-time Fourier Transform (STFT), where each square matrix is $U_m W_M$	12
2-5	No perfect reconstruction is obtained with a regular Hann window as $P(n) \neq 1 \forall n$.	14
2-6	Perfect reconstruction is obtained with a normalized Hann window as $P(n) = 1 \forall n$.	14
2-7	Huygens' principle	15
2-8	Block diagram multiple-error Normalized LMS.	16
3-1	Picture of physical test environment and a graphical schematic of the aperture. .	22
3-2	3D loudspeaker array schematics.	23
3-3	2D simulation environment and array schematics.	23
3-4	The time delay wrapping issue that occurs when long delays are implemented with short STFT blocks. Due to the periodicity assumption of the Fourier transform, the time-delay shift causes the end of the signal block to wrap to the beginning of the block, visible when taking the following steps. The original signal (1) is windowed to obtain a windowed signal (2). Then, the signal is transformed to the frequency-domain, a time-delay is applied and transformed back to the time-domain (3). Finally, the window is applied again, resulting in a wrapped signal (4).	28
3-5	Aperture Acoustic Transfer Function (ATF) frequency responses for the high (blue) and low-resolution (red) scenario. The low-resolution line shows a smoothed version and corresponds with the high-resolution version at the grey lines, that indicate the frequency bins.	29
3-6	The impulse response of the aperture ATF. The solid blue line shows the original impulse response. The dashed red line shows the filtered case, where a rectangular window is applied due to the low frequency resolution. Only the first part of the middle part of the impulse response remains, which directly shows the issue of a short block length.	30

3-7	The schematic overview of the error analysis procedure with frequency weighting $\mathbf{h}(k)$, weighted frequency response $\mathbf{y}(k)$, it's low frequency resolution version $\hat{\mathbf{y}}(k)$ and the error $e(k)$. The $*$ denotes convolution.	31
3-8	Block diagram of the simulation environment.	32
4-1	2D cross-section of the environment with control region \mathbb{D} , a hemisphere in the far-field, between r_{min} and r_{max} from the aperture.	34
4-2	Any soundfield can be described by a finite weighted sum of simple waves.	35
5-1	Rumbler-siren characteristics, one of the signals that is used in the simulations.	44
5-2	Aperture soundfields for various frequencies at 0° incident angle.	46
5-3	Aperture soundfields for various frequencies at 30° and 60° incident angle.	46
5-4	Comparison between the soundfields generated with the old and new Aperture Acoustic Transfer Function. The soundfields by the old version show clearly that the sound originates from the aperture. Because of that, there is a clear line between the side and main lobes visible. The new soundfield shows that the waves originate from the aperture sections and has more smeared out overlap between the main and sidelobes. The latter is a more realistic representation.	47
5-5	2D animation for a pure 1kHz sine signal controlled with the Optimal Wave-Domain Algorithm (WDC-O) and a grid array. The algorithm is able to cancel the soundfield as we see less amplitude in graph Figure 5-5c.	48
5-6	3D results for various scaling of G , using white noise, 0° incident angle, the grid array and a window-size of $N=3.91$ ms. The attenuation performance is similar for all scaling factors. a smaller scaling factor is beneficial for the computational load. Also, $\beta = 1/32$ showed less performance at the mid-frequency range. Hence, we chose $\beta = 1/16$	49
5-7	3D Fundamental error analysis, with the Signal-to-Noise Ratio (SNR) over the evaluation microphones. We used a flat frequency weighting $s(k)$, and evaluated for 20 random incident angle scenarios, leading to 20 different $\mathbf{h}(k)$. We can conclude that increasing the window-size leads to less fundamental error in the aperture ATF. There is major difference between $N = 2.0$ ms and $N = 3.9$ ms, indicating that the ATF impulse response for a window-size smaller than $N = 3.9$ ms is cut off drastically, leading to major errors.	49
5-8	3D Grid array SEG_f results with narrow 99% confidence interval for various window-sizes. Increasing the window-size increases performance. The major difference is visible at the lower frequencies.	50
5-9	3D Sparse array SEG_f results with narrow 99% confidence interval for various window-sizes. A larger window-size resulted in better performance. The major difference is visible at low frequencies.	51
5-10	3D phase-shift values in $W_q(k)$, for all speakers for 0° incident angle. As expected, they are concentrated around 180° , which leads to large phase-shifts in the block processing.	51
5-11	3D Grid results for the NLMS and WDC-O algorithms. Performance is comparable. However, WDC-O outperforms in terms of convergence time, as it converges instantaneously. In contrast, the NLMS needs 1 s to converge.	52
5-12	3D Sparse results for the NLMS and WDC-O algorithms. Performance is comparable. However, WDC-O outperforms in terms of convergence time, as it converges instantaneously. In contrast, the NLMS needs almost 2 s to converge.	53

5-13	3D grid array SEG_f results with narrow 99% confidence interval for WDC-O and NLMS at three incident angles. In general, a larger incident angle led to less attenuation. The performance of the NLMS is more sensitive to the incident angle compared to the WDC-O results.	53
5-14	3D sparse array SEG_f results with narrow 99% confidence interval for WDC-O and NLMS at three incident angles. In general, a larger incident angle led to less attenuation. The performance of the NLMS is more sensitive to the incident angle compared to the WDC-O results. The major difference is visible at lower frequencies.	54
5-15	3D grid and sparse array SEG_f results with narrow 99% confidence interval for WDC-O, white noise, 0° incident angle. Performance of the grid array was better for all frequencies. The grid array was able to cancel noise up to 2 kHz, while the sparse array only performed up to 1.2 kHz.	55
5-16	Animation of aperture, grid and sparse array soundfields for 1.5 kHz sine wave. The sparse array soundfield does not correspond to the soundfield of the aperture array, while the grid array soundfield is very similar. This shows the major difference in performance between the two arrays.	55
5-17	3D SNR value with 99% confidence interval for both algorithms cancelling rumbler-siren noise to find the optimal window-size. The optimal point for the Wave-Domain Algorithm with Predictor (WDC-P) (Sparse) lies at 15.6 ms, while the optimal window-size for the grid array is 7.8 ms.	56
5-18	3D grid array SEG_f results with narrow 99% confidence interval, for all algorithms, 0° angle and rumbler-siren noise. The best performing algorithm is the WDC-O. Only the WDC-M and WDC-O are able to attenuation noise over the full frequency range up to 2 kHz.	56
5-19	3D grid SEG_t results for all algorithms, 0° incident angle and rumbler-siren noise. The WDC-O was clearly the best performing algorithm as the attenuation was kept steady at -10 dB. Attenuation of the other methods fluctuated over time, depending on the signal content. Moreover, the long convergence time of the NLMS algorithm is visible in Figure 5-19b, while the other had instantaneously or very fast convergence.	57
B-1	The perpendicular distance between the measure location and the primary source location plane is denoted with r and can be calculated by $r = \frac{l^2+q^2-p^2}{2l}$	78
C-1	Aperture soundfields for various frequencies at 0° incident angle.	79
C-2	Aperture soundfields for various frequencies at 0° incident angle.	79
C-3	Aperture soundfields for various frequencies at 30° incident angle.	80
C-4	Aperture soundfields for various frequencies at 30° incident angle.	80
C-5	Aperture soundfields for various frequencies at 60° incident angle.	80
C-6	Aperture soundfields for various frequencies at 60° incident angle.	80
D-1	Four arrays tested in 3D simulations.	81
D-2	3D results for 4 types of arrays, the sparse array, sparse array by Downey with 4 additional speakers, sparse array by Miller, and the 7×7 grid array. Simulations with white noise, 0° incident angle and a window-size of $N=7.8$ ms.	82
D-3	2D Grid array SEG_f results for various window-sizes.	82
D-4	2D Sparse array SEG_f results for various window-sizes.	83
D-5	2D Grid array SEG_f results for the NLMS and WDC-O algorithms.	83

D-6	2D Grid time results for the NLMS and WDC-O algorithms.	84
D-7	2D Sparse array SEG_f results for the NLMS and WDC-O algorithms.	84
D-8	2D Sparse time results for the NLMS and WDC-O algorithms.	85
D-9	2D Grid array SEG_f results for WDC-O and NLMS for three different incident angles.	85
D-10	2D Sparse array SEG_f results for WDC-O and NLMS for three different incident angles.	86
D-11	2D Grid and sparse array SEG_f results for WDC-O, white noise, 0° incident angle.	86
D-12	Optimal window-size for 2D setup. The optimal point for the WDC-P (Sparse) lies at 15.6 ms, while the optimal window-size for the grid array is 7.8 ms.	87
D-13	2D Grid array SEG_f results for all controllers, 0° angle and rumbler-siren noise.	87
D-14	2D Grid time results for all controllers, 0° angle and rumbler-siren noise.	88

List of Tables

5-1	3D Results for rumbler-siren noise at 0°angle, in decibels. Given over four frequency ranges, for all algorithms. In general, less attenuation was obtained for higher frequencies. Results are similar between the sparse and grid array, but the grid array outperformed the sparse for all algorithms at higher frequencies. WDC-O performed best. Also, the WDC-M and WDC-O with grid array were able to obtain decent performance above 1.5 kHz while the other algorithms did not. . .	58
5-2	3D results for all algorithms, for both grid (G) and sparse (S) arrays. Attenuation is given in decibel (dB). WDC-P did not attenuate airplane and white noise as it could not predict the signal. The grid array outperformed the sparse array in any scenario. The WDC-O algorithm is the clear winner. NLMS and WDC-M performed fairly similar.	59
D-1	2D Results for rumbler-siren noise at 0°angle, in decibels. Given over four frequency ranges, for all controllers.	83

Acknowledgements

This document is my final work towards the completion of the Master Systems and Control at the Delft University of Technology. It was an honor to collaborate with the GN research group designing a ‘silent window’. I have learned a lot, not only about the topic of active noise control, but I also grew personally in presenting complex theory, conducting structured thesis research, and working together online during the circumstances of the COVID-19 pandemic. I positively look back at this challenging year.

First of all, I would like to thank my daily supervisor, prof.dr. W.B. Kleijn. Bastiaan, you really inspired me and helped me achieve the most out of this project. Our weekly discussions were in-depth, and you were very responsive when I had questions. Your way of guiding me should be used by any thesis supervisor. It gave the feeling that I was in control. Also, I enjoyed giving you regular updates on the news in the Netherlands, learning about New Zealand, and learning about cultures in general. Hopefully, we will be able to meet in person someday. Next, a special thanks to my supervisors at Delft Center for Systems and Control (DCSC): Dr.ir. R. Ferrari and ir. J. Gonzalez Silva. Riccardo, our monthly meetings were encouraging and steered me in the right direction. Moreover, I appreciate the feedback you gave me. Jean, we had many interesting discussions online and in the office. I enjoyed working with you and admire your mindset towards complex subjects. Also, thanks for the feedback and helping me with using the cluster. I would also like to thank the GN research group, mostly located at PennState University, Pennsylvania, for our monthly encounters. It was valuable to discuss my work, and I hope my contributions lead to a working aperture ANC system at your lab. Good luck working on the next challenges. Furthermore, thanks to Almer and Erik, from GN Group Eindhoven office, for making the whole thesis trajectory at GN possible, and our monthly meetings. Without your efforts, it would not have worked out.

I would like to thank my friends and family. First, thanks for reviewing my thesis report, Kim, Huub, Dik and Mees. Then, a big thanks to my parents for their continuous help and proudness during my years as a student. It has been a wonderful time that would not have been possible without your support. And finally, thank you Kim, for your positive energy, listening to my endless technical stories, and supporting me during this busy year.

Delft, University of Technology
October 26, 2021

D. Ratering

Master of Science Thesis

D. Ratering

Chapter 1

Introduction

1-1 Motivation

Noise pollution is a major health threat to society [1]. Active Noise Control (ANC) systems that attenuate noise propagating through open windows (apertures) have the potential to create quieter homes while maintaining ventilation and sight [2]. Existing ANC systems deploy loudspeakers to produce anti-noise soundfields that reduce sound energy in local regions, for instance with noise cancellation headphones [3] or over large regions such as in airplane cabins [4]. ANC systems that actively control sound propagating through open windows have been studied [5, 6]. Their objective is to reduce the sound energy in all directions from the aperture into the room. Current methods deploy closed-loop algorithms, leading to long convergence times [7], heavy computational load [8] and the need for a large number of error microphones being positioned in the room. These drawbacks limit the feasibility of such systems [9]. Recent studies focus on reducing the computational load by improved closed-loop algorithm [10, 11]. However, these studies do not solve the issue of error microphone placement. As such, we propose an open-loop wave-domain approach that addresses the shortcomings of the closed-loop approach because it converges instantaneously, works with offline weight calculation, and does not require error microphones.

1-2 Introduction to Active Noise Control for Apertures

In this section, we briefly cover relevant studies, to position the proposed method in current literature. For a detailed literature review on ANC for apertures, the reader is referred to the literature study executed prior to this thesis work [12]. The study presented in this thesis is a continuation of work that is currently being carried out by a research group within GN Group [13, 14].

The method of ANC originates from the late 1800s [15] but was first patented in 1936, when Lueg [16] proposed a system that reduces sound, based on the superposition principle. This

principle denotes the linearity of the wave equation where two soundfields can be summed. More specifically, ANC systems aim for destructive interference [17], where the sum of these two soundfields reduces the acoustic pressure, resulting in less undesired sound (noise). A soundfield is a region of an elastic medium that contains sound waves [18]. Thus, if we create a so-called anti-noise soundfield, that has the same amplitude and opposite phase compared to the noise soundfield, we can sum the soundfields of the noise and anti-noise resulting in a residual noise of zero acoustic pressure [2] and destructive interference is obtained.

Most ANC systems for apertures deploy Least Mean Squares (LMS) algorithms, such as the well-known Filtered-x Least Mean Squares (FXLMS) algorithm [7], or its multi-channel equivalent, the multiple-error LMS [7,10]. These closed-loop algorithms aim to minimize error signals measured by error microphones positioned in the room by using the sound generated by loudspeakers in the aperture. By minimizing the error at these error microphones, such systems aim to create a destructive soundfield in all directions from the aperture. This approach is effective at the error microphone locations. However, the consistency over a large area (such as a room) is low [19]. Therefore, the number of error microphones needs to be substantial to ensure global control, which is undesirable and limits the feasibility of such systems. Current studies indicate that LMS-based ANC for apertures can obtain a -10 dB global reduction between 0.5 and 2 kHz [20]. Many studies focus on reducing complexity by improving the LMS algorithms [10,11]. Positioning error microphones close to the aperture and loudspeaker reduces complexity and generates more feasible systems. However, these systems do not ensure global control intrinsically, and performance degrades when the incident angle of the incoming noise is increased [20]. Others proposed a feedforward method with reasonable performance between 0.4 and 1 kHz [21], and an open-loop method [9] that is promising in a simulation where pure sinusoidal noise is cancelled. However, those methods do not meet the performance of state-of-the-art LMS-based algorithms.

Supplementary to the development of algorithms, various studies extensively cover the physical setup used in ANC systems for apertures. In general, ANC systems for apertures with loudspeakers distributed over the aperture [6] outperformed those with loudspeakers placed on the boundary of the aperture [5]. A compromise between both setups is a sparse array that consists of a window with a crossbar containing loudspeakers [13,14].

We propose an wave-domain algorithm for ANC for apertures as it ensures global control in all directions intrinsically. It does not require microphones positioned in the room. As such, it is an open-loop controller. To our knowledge, this study is the first to investigate a wave-domain approach to ANC for apertures. The control problem has similarities with wave-domain spatial control of sound produced by multi-speaker sound systems, e.g., [22–24]. Similarly, the active cancellation of wall reflection of sound produced by sources such as people talking [25], and wave-field synthesis [26] show resemblance. Our wave-domain algorithm uses a wave-domain basis function expansion over a control region. Both the soundfields from the primary noise and loudspeaker array are expressed in these basis functions, and their sum is minimized in a least mean square sense [22–24]. Applying this minimization allows calculating optimal filter-weights for the loudspeaker array that minimize far-field sound energy for any wavenumber.

The wave-domain approach is a model-based method. Consequently, an Acoustic Transfer Function (ATF) for the soundfields is necessary for such algorithms. Loudspeakers are often modeled by monopoles [27]. Recent studies focused on expressing aperture soundfields with

single equations [13,14]. However, the developed equation ATF is only valid in far-field, which is too far away for the developed designs.

Finally, there is an important artifact of the wave-domain algorithm as it operates in the wave-domain. Hence, it requires to transform signals with the Short-time Fourier Transform (STFT) to get their wave-domain representation. This operation induces an algorithmic delay equal to the window-size of the STFT. Consequently, we need to compensate for this algorithmic delay. In previously investigated wave-domain algorithms, this compensation is done by deploying a signal predictor [24]. Positioning the reference microphone further away from the aperture is also a possible solution for this algorithmic delay compensation. If a predictor is used, the selection of the STFT window must be based on a trade-off between prediction performance and the wave-domain algorithm performance [24].

1-3 Definition of Research Gap

Following the literature study [12] and previous work on the project [13,14] we identified a research gap in the domain of ANC systems for apertures. As far as we know, there is no study that investigates the wave-domain approach for this control problem. Specifically, the following is missing in current literature:

- The derivation of a wave-domain algorithm that calculates loudspeaker filter-weights for global control of soundwaves propagating through an aperture, operating within the following constraints:
 - no error microphones,
 - no convergence time,
 - limited computational effort.
- The derivation of an aperture ATF that are valid within a reasonable distance from the aperture.
- The comparison of wave-domain algorithm with LMS-based algorithms.
- An investigation into the performance of wave-domain algorithms with sparse and grid arrays.
- An study into the performance of a wave-domain algorithm for ANC for apertures with primary noise coming from various incident angles.
- A study investigating compensation methods for the algorithmic delay of the wave-domain algorithm for ANC for apertures.

In the following section, we propose a problem formulation that will help us in filling this research gap.

1-4 Problem Formulation

As depicted in the prior sections, the current state-of-the-art ANC for apertures does not cover wave-domain algorithms. The general concept of wave-domain algorithms in real-time has already been proven in simulations and real-world tests [23, 24]. Due to its open-loop character, it has large potential to suit as an algorithm for filter-weight calculation in feasible ANC setups as there is no need for a large number of error microphones in the room. To calculate attenuation performance, a simulation environment needs to be established in which the algorithms can be tested. Moreover, the algorithm itself needs to be developed. The algorithmic delay, as an artifact of the wave-domain approach, is a limiting factor. An initial attempt to compensate for this is necessary to show that algorithmic delay compensation is possible. If that works, we can prove that the wave-domain algorithm is implementable in real-time. Finally, we need a reference algorithm. A LMS-based algorithm is used as reference method.

With this in mind, we seek to answer the following research question in this study:

Can an active noise control system with an open-loop wave-domain algorithm achieve superior performance compared to state-of-the-art algorithms when attenuating undesired sound propagating through open windows?

To answer this question, we define the subsequent research objectives that resemble with the identified research gaps:

- Develop a simulation environment (2D and 3D) in which the new wave-domain algorithm, as well as the baseline algorithm, can be tested.
- Implement a closed-loop LMS-based algorithm as the state-of-the-art reference method.
- Develop and implement an open-loop wave-domain algorithm.
- Develop initial algorithmic delay compensation methods, one by microphone placement and the other with a signal predictor.
- Test and compare developed algorithms by using sparse and grid loudspeaker arrays, as well as investigate performance for various primary noise incident angles.

We supplement these objectives with constraints, conditions, and assumptions to demarcate a clear scope of this study. First of all, the main goal is to develop the algorithm for weight calculation in a wave-domain approach. To keep it achievable, we constrain the study to simulations only. A simulation environment is build to test the algorithm. Secondly, we are not interested in seeking the best arrangement in terms of loudspeakers and microphone placement. Thus, we limit our study to two loudspeaker array setups. On the one hand, we have the sparse array, developed by Downey [14] and on the other hand, we deploy a straightforward grid array, similar to most studies on ANC for apertures [6].

In addition to the given constraints, we work with the following conditions. The undesired sound, denoted as primary noise, propagates from an outside noise source through the aperture into the room. The position of the outside noise source is fixed during simulation. We follow prior studies by looking into a noise coming from three fixed angles, 0° , 30° , and 60° . In

addition, this source is positioned in the far-field, allowing us to model the incoming wave as a plane wave. Lastly, we will test the system for three types of primary noise: white noise, airplane noise, and a rumbler-siren sound.

For our study it was assumed that acoustic far-field equations hold. So, we model an infinitely large empty room without walls and no reflection. Most importantly, we focus on algorithm development and not on any calibration of hardware components. Hence, we assume that we have loudspeakers and microphones with a perfectly flat frequency response.

Using these limits, we set a clear scope of this thesis study. This chapter is concluded with the outline of this thesis report, described in the next section.

1-5 Thesis Outline

The report starts with Chapter 2, containing theoretical background information on relevant theoretical concepts and the deployed reference algorithm. Chapter 3 covers the simulation environment. Moreover, we derive the Acoustic Transfer Functions (ATFs) necessary for wave-propagation modeling and used in the wave-domain algorithm. The proposed wave-domain algorithm is described in Chapter 4. The results are presented in Chapter 5. An elaborate discussion on the results, limitations, and suggestions for future work are given in Chapter 6. Finally, the thesis is briefly concluded in Chapter 7. Attached to the report are Appendices, including the conference paper, extended theory, and extended results.

The reader is referred to the conference paper, which we submitted to IEEE ICASSP 2022, in Appendix A for a concise version of this thesis work. It contains the developed algorithm and briefly covers the results.

Chapter 2

Theoretical Background

In this chapter, we begin by establishing notations and definitions in Section 2-1. Furthermore, we examine the relevant extant theory that forms the basis of the study (Section 2-2), such as basis signal processing concepts like the Short-time Fourier Transform (STFT) and window function requirements. In Section 2-3, Huygens' Principle is briefly introduced. We present an extensive derivation of the multiple-error Normalized Least Mean Squares (NLMS) algorithm, which we deploy as the reference algorithm, in Section 2-4. Finally, a brief summary of this chapter is given in Section 2-5.

2-1 Preliminaries

In this section, we cover assumptions and definitions that are used throughout the report. The following mathematical definitions are established. We denote the imaginary number with $j = \sqrt{-1}$. The sinc-function is defined as unnormalized: $\text{sinc}(x) = \sin(x)/x$ [27]. Furthermore, $[\cdot]^H$ is the conjugate transpose, the euclidean norm is denoted using $\|\cdot\|$, $*$ means linear convolution and \odot is the element-wise multiplication. We define the wavenumber as $k = 2\pi f/c$, with f frequency in Hz and c the speed of sound in m/s. Finally, we have the density of air, ρ_0 in kg/m^3 .

To describe points in spatial-domain, we need coordinate systems. We use both the cartesian and spherical coordinate systems as these are convenient when modeling plane waves and spherical waves. The coordinate systems are visualized in Figure 2-1. In this graph, the $x - y$ plane corresponds with the horizontal plane while the z -direction points in the vertical direction. For the spherical coordinate system we use the physics notation. Thus, θ is the elevation: the angle from the vertical z -axis and ϕ denotes the azimuth: the angle from the positive x -axis. Furthermore, r denotes the distance from the origin, o . The relation between the coordinate systems is described by:

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta, \quad (2-1)$$

and are shown in Figure 2-1. With these definitions, any point in 3D-space can be described in both spherical and Cartesian coordinates. In the next section, we discuss basic signal processing concepts that we need in our algorithm development.

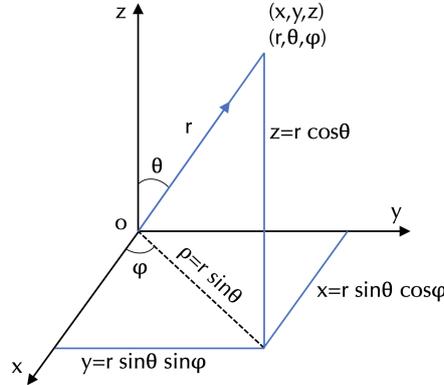


Figure 2-1: The cartesian coordinate system and its related spherical coordinate system.

2-2 Signal Processing Concepts

When modeling wave-propagation and soundfields, it is necessary to process signals in both time-domain and wave-domain (time-frequency domain). Therefore, the use of Fourier transformation is evident. The STFT, closely related to the Gabor transform, allows us to represent a discrete signal in wave-domain where wave propagation equations are applied. Thus, we must ensure that the manipulation of the signal in the wave-domain corresponds to its time-domain relative. For this, we study the STFT in more detail and derive conditions that ensure a tight frame STFT. Preparatory to discussing the theory of STFT, we cover relevant concepts such as linear transformations, frame theory, and the Discrete Fourier Transform (DFT). For this, we closely follow some sections from [28].

2-2-1 Linear Transformation and Manipulation

Let us first discuss linear transformations. A simple linear transformation sequence is shown in Figure 2-2. Here, a column vector $x(n)$ is transformed by using a square and invertible matrix U with size $n \times n$. The result is another column vector $y(n) = Ux(n)$. In this transformation, both $x(n)$ and $y(n)$ are of length n . Similar to this matrix multiplication, we can denote the transformation as a standard inner product. We denote the inner product between the n -dimensional column vectors $x(n)$ and $z(n)$ as $\langle x(n), z(n) \rangle \equiv \sum_{m=1}^n z_m(n)^* x_m(n)$ with $*$ the complex conjugate. This inner product has a simple relation to the matrix multiplication: $y_l(n) = \langle x(n), U_l^H \rangle$ with H the Hermitian transpose and U_l^H the l -th column of U^H . Hence, both methods can be used to calculate $y(n)$. The obtained vector $y(n)$ is a vector in subspace \mathbb{C}^n . The inverse transform is $x(n) = U^{-1}y(n)$. Since U is invertible, the columns of U^{-1} must span \mathbb{C}^n , that is, the columns of U^{-1} must form a basis of \mathbb{C}^n . $x(n) = U^{-1}y(n)$ can be interpreted as an expansion of $x(n)$ into basis vectors. Such a basis expansion of the inverse transform is explicitly stated as:



Figure 2-2: Sequence of a linear transformation and coder using an invertible $n \times n$ matrix U .

$$x(n) = \sum_{l=1}^n y_l(n) U_l^{-1} = \sum_{l=1}^n \langle x(n), U_l^H \rangle U_l^{-1}, \quad (2-2)$$

with U_l^{-1} is the l -th column of U^{-1} [28]. After the linear transformation, vector $y(n)$ is manipulated by \mathcal{Q} . The new vector $\mathcal{Q}(y(n))$ is then transformed back with the inverse matrix U^{-1} . We require that the Euclidean norm of the quantization error is invariant with respect to the inverse transformation U^{-1} which results in $I = U^H U$. This resolves to $U^H = U^{-1}$ as we can denote the inverse mapping of $x(n) = U^H y(n)$ as an orthonormal basis expansion:

$$x(n) = \sum_{l=1}^n x_l(n) U_l^H = \sum_{l=1}^n \langle x(n), U_l^H \rangle U_l^H. \quad (2-3)$$

Eq. (2-3) is equal to Eq. (2-2) for the specific case of a unitary matrix U .

Instead for a square matrix U , we could the case of a linear transformation where U is non-square with dimensions $k \times n$. A sequence is shown in Figure 2-3 where $y(k) = Ux(n)$. The rows of U span the input vector space, so we must have $k \geq n$ and $Ux(n)$ is only zero if $x(n)$ is zero. Because of $k \geq n$, we have dimension redundancy. The columns of U span a n -dimensional subspace of \mathbb{C}^k . Hence, we have one-to-one mapping from \mathbb{C}^n to \mathbb{C}^k : if both $z(k) = Ux(n)$ and $y(k) = Ux(n)$, then $z(k) - y(k) = 0$ and thus $z(k) = y(k)$. Since the forward mapping is one-to-one, it is invertible. However, we have some problems. Firstly, if we manipulate $y(k)$ with $\mathcal{Q}(\cdot)$, we are not sure the resulting $\mathcal{Q}(y(k))$ is in the column space of U . Secondly, the inverse of U , shown here by \tilde{U} is generally not unique [28]. These two issues are solved by implementing the pseudo-inverse (Moore-Penrose inverse), defined as $U^\dagger = (U^H U)^{-1} U^H$, as that is the optimal inverse in a mean squared-error sense for the transformation, manipulation, inverse-transformation operation sequence shown in Figure 2-3. The derivation of the pseudo-inverse is found in various textbooks [29, 30].



Figure 2-3: Sequence of a linear transformation and coder using an invertible $k \times n$ matrix U .

2-2-2 Frame Theory

Let us continue from the prior section with matrix U with size $k \times n$, whose rows span \mathbb{C}^n . If these rows are finite, they form a frame. The row vectors of U , or the column vectors of U^H form a frame if the so-called frame condition holds. This conditions is defined as:

$$Ax^H(n)x(n) \leq x^H(n)U^H Ux(n) \leq Bx^H(n)x(n), \quad \forall x(n) \in \mathbb{C}^n, \quad (2-4)$$

where frame bounds $A > 0$ and $B < \infty$ are both scalars [28]. Because $A > 0$, the frame vectors (columns of U^H) must span \mathbb{C}^k and $B < \infty$ results in that the vector norms of the frame vectors are finite. In case of $A = B$, the frame is defined as tight. Let us define the nomenclature analysis frame for the columns of U^H and synthesis frame for the columns of U^\dagger . The names are related to their function. We use the inner product notation, $y_n^m = \langle x_n, U_n^H \rangle$. The analysis frame is used to analyze the input x_n . Moreover, the synthesis frame is used to resynthesize the original vector. Here, written in frame expansion form: $x_n = U^\dagger y^m = \sum_{l=1}^m y_l^m U_l^\dagger = \sum_{l=1}^m \langle x_n, U_l^H \rangle U_l^\dagger$, where U_l^\dagger are the columns of U^\dagger . From the frame condition it follows immediately that, for a tight analysis frame, with frame bound A , we have $U^H U = AI$. The corresponding synthesis frame has a frame bound of $1/A$. Furthermore, in case of a tight frame, the pseudo-inverse reduces to the following form: $U^\dagger = (U^H U)^{-1} U^H = \frac{1}{A} U^H$. This implies that the analysis and synthesis frames are identical except for a constant.

2-2-3 Discrete Fourier Transform

We now have an understanding of linear transformations and tight frames. Before discussing the STFT, we briefly recap what the DFT is, as the STFT consults the DFT. In literature [27–29], the DFT of a vector $x(n)$, with discrete time index n is often defined as:

$$y(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j \frac{2\pi}{N} kn}, \quad (2-5)$$

where $y(k)$ is a new vector, with wavenumber k . Furthermore, we have N the total number of samples in $x(n)$ and j the imaginary number. For completeness, we define the Inverse Discrete Fourier Transform (I-DFT) as well. This inverse linear transformation is denoted as:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} y(k) \cdot e^{j \frac{2\pi}{N} kn}. \quad (2-6)$$

Next to this conventional notation as a sum, we can denote the DFT as a linear transformation in form of matrix multiplication, as we discussed in the prior section. So we have:

$$y(k) = U^{DFT} \cdot x(n), \quad (2-7)$$

with the transformation matrix U^{DFT} defined as follows:

$$U^{DFT} = \frac{1}{\sqrt{k}} \cdot \begin{bmatrix} \mathcal{W}_0 & \mathcal{W}_0 & \mathcal{W}_0 & \cdots & \mathcal{W}_0 \\ \mathcal{W}_0 & \mathcal{W}_{-1} & \mathcal{W}_{-2} & \cdots & \mathcal{W}_{-(k-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{W}_0 & \mathcal{W}_{-(k-1)} & \mathcal{W}_{-2(k-1)} & \cdots & \mathcal{W}_{-(k-1)(k-1)} \end{bmatrix}, \quad (2-8)$$

in which the so-called twiddle factor is defined as:

$$\mathcal{W} = e^{j\frac{2\pi}{k}}. \quad (2-9)$$

From now, for ease of notation, we denote U^{DFT} simply with U . We can easily see that the rows of U form an orthonormal basis in \mathbb{C}^k . Thus, the inverse of U is simply its conjugate transpose, leading to the fact that $U^H U = I$.

The notation we elaborated is useful later on, however, note that the actual implementation of the DFT by using matrix multiplication is not a computationally friendly method. Therefore, when implemented in the simulations, the more efficient Fast Fourier Transform (FFT) is deployed [31].

2-2-4 Short-time Fourier Transform

With the prior discussed theory at hand, we investigate the STFT. We describe the STFT as follows. Instead of applying a DFT to all sample points of a signal $x(n)$, in a STFT we split up this signal in blocks of N samples, and apply the DFT to each block. By doing so, the result is a sequence of DFT transformed vectors that tell us something about the change of the frequency content over time. Even more important, it allows us to manipulate the frequency content of a signal over time. This is fundamental to our open-loop algorithm and will be more clear later on. Each time-block in the STFT is windowed by a window function to prevent spectral leakage. However, by applying a window, we remove important content of that block of signal as we manipulate the amplitude. We ensure an overlap of subsequent blocks such that we do not lose information or overcome this issue. This overlap is created by using a certain hop size (H), smaller than the block size, and, therefore, window-size (N). Thus, $N > H$. Let us formally define the STFT of a signal $x(n)$ with total length L , in sum notation as

$$\mathbf{Y}(k, m) = \sum_{n=0}^{N-1} x(n)w(n - mH)e^{-j\frac{2\pi}{N}kn}, \quad (2-10)$$

where $\mathbf{Y}(k, m)$ is a stacked column vector consisting of M smaller column vectors $y(k)$, having a total length NM [32]. In Eq. (2-10), m denotes the hop, with total number of hops being $M = L/H$. For convenience, it is assumed that our signal $x(n)$ is periodic.

A simple and commonly used STFT method introduces windowing only to the analysis step where the window function is selected such that the segments sum to one ($\sum w(n + mH) = 1$). However, while this is straightforward and provides perfect reconstruction of a non-manipulated signal, after synthesis without windowing, using the STFT without windowing in the synthesis step has a number of disadvantages:

1. The approach is inherently oversampled since the segments overlap.
2. Manipulations (denoted in Figure 2-2 and Figure 2-3 with $\mathcal{Q}(\cdot)$) result in discontinuities in the time-domain, since the synthesis functions are not tapered.

$$P(n) = \sum_{m \in \mathcal{Z}} w(n + mH)^2 = \frac{A}{N}, \quad n \in \{0, \dots, H - 1\}. \quad (2-13)$$

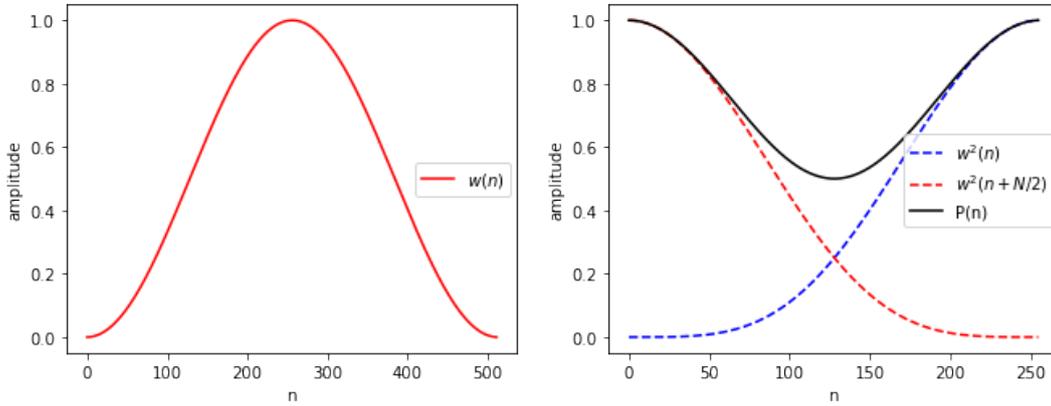
Note that the $1/N$ term is generally part of the inverse DFT transform, and we are left with a scaling factor of A , we found in our frame theory discussion in Section 2-2-2. Consequently, to obtain perfect reconstruction, we choose A such that $\frac{A}{N} = 1$. Let us discuss window functions in more detail next.

2-2-5 Window Functions

A commonly used window in time-frequency analysis is the Hann window [33], shown in Figure 2-5a. However, this window does not fulfill the power complementarity condition, shown in Eq. (2-13). We see in Figure 2-5a that the sum of the squared windows ($P(n)$) is not constant. The cosine window function does fulfill the condition when an overlap of 2 is used. However, this window is undesired as the tails are too steep. Therefore, we are seeking a window that fulfills the requirements but does have smooth tails. To make this process more simple, we design a normalization factor that scales every type of window (combined with an arbitrary hopsize) to a window that fulfills the condition. By scaling the amplitude and choosing a desired window-shape, we obtain a perfect window for the STFT. We make the amplitude normalization function dependent on $P(n)$. And define the normalized window as $\hat{w}(n) = \frac{w(n)}{\sqrt{P(n)}}$ [34]. This normalized window has its own periodic $\hat{P}(n)$. To proof it's unity, we take the following steps:

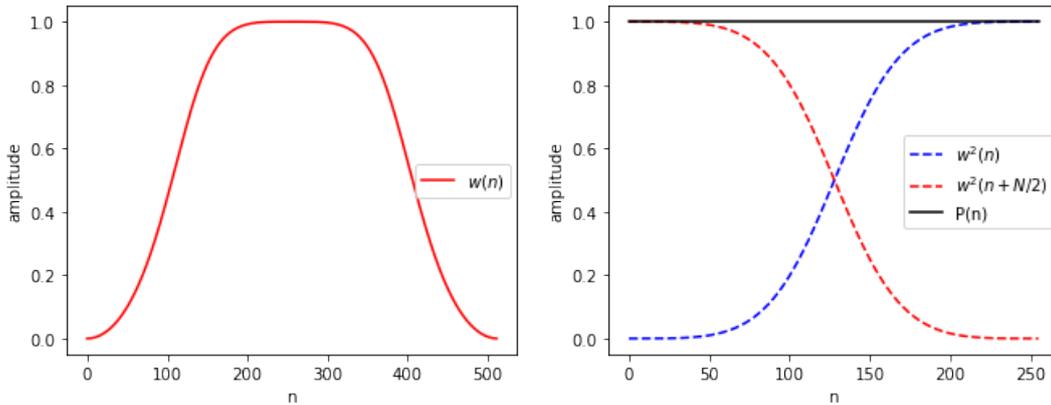
$$\begin{aligned} \hat{P}(n) &= \sum_m \hat{w}(n - mH)^2 \\ &= \sum_k \frac{w(n - mH)^2}{P(n - mH)} \\ &= \sum_k \frac{w(n - mH)^2}{P(n)} && \text{(because of periodicity of P with H)} \\ &= \frac{1}{P(n)} \sum_k w^2(n - mH) \\ &= \frac{1}{P(n)} P(n) = 1. \end{aligned} \quad (2-14)$$

By applying the normalization function, we can obtain a normalized Hann window, shown in Figure 2-6a, that fulfills the power complementarity conditions and has smooth tails. Figure 2-6b shows that the window fulfills the conditions as $P(n) = 1 \forall n$. Now that we have a clear understanding of the signal processing concepts, we move on to a concept that is applied in Active Noise Control (ANC) systems.



(a) Plot of the regular Hann window of window-size $N = 512$ samples. (b) Clearly, $P(n) \neq 1 \forall n$ for a regular Hann window with overlap of 50%.

Figure 2-5: No perfect reconstruction is obtained with a regular Hann window as $P(n) \neq 1 \forall n$.



(a) Normalized Hann window of window-size $N = 512$ samples and normalized for a 50% overlap. (b) Clearly, $P(n) = 1 \forall n$ for a normalized Hann window with overlap of 50%.

Figure 2-6: Perfect reconstruction is obtained with a normalized Hann window as $P(n) = 1 \forall n$.

2-3 Huygens Principle

An important acoustic principle is a theory developed by Dutch physicist Christiaan Huygens in 1678. This concept states that the propagation of a wavefront can be visualized by considering the wavefront as a set of closely spaced point sources, each generating spherical waves. The combination of these generated waves represents the new wavefront surface with the correct intensity and direction [20, 35]. This is visualized in Figure 2-7. If these point sources are positioned closer to each other than the wavelength of the sound, one can exactly describe the wave [20].

With this concept in mind and the linearity of the wave-equation, one can sum soundfields from various sources. Any wavefront that is a sum of wavefronts by multiple sources can be used to describe a new wavefront. Therefore, if destructive interference is generated in a current wavefront, the Huygens principle tells us that we have less energy in the new wavefront

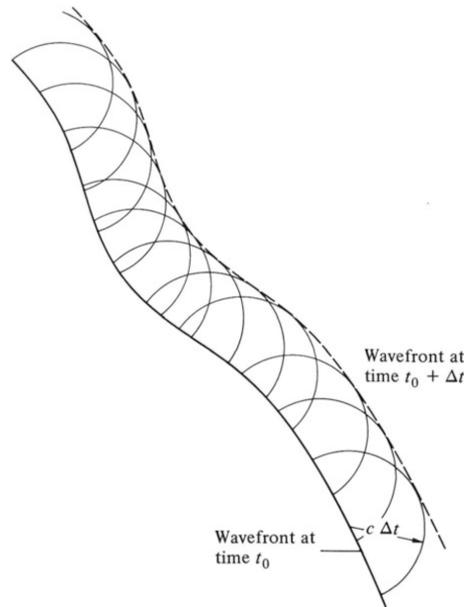


Figure 2-7: Visualization of Huygens' principle where the wavefront at $t_0 + \Delta t$ can be fully described by spherical waves from points on the wavefront at t_0 .

generated by the element waves from the current wavefront.

2-4 Multiple-error Normalized LMS Algorithm

This section describes the theory of the reference algorithm, the Multiple-error NLMS algorithm. As it is the reference algorithm, we deploy an existing algorithm that is well-known and used in most ANC systems for apertures.

2-4-1 Motivation

The multiple-error NLMS algorithm is used as the reference algorithm because of the extensive use in ANC for aperture literature. It is model-independent, which makes the implementation straightforward. Furthermore, it is a robust algorithm [36]. Disadvantages of such algorithms are slow convergence and heavy computations, especially when the number of loudspeakers and error microphones are large. Moreover, for actual implementation in an ANC aperture, the closed-loop approach is less feasible. The error microphones should be positioned distributed around the aperture, leading to less functionality of the aperture itself. Additionally, the error microphones in the room pick up sound from inside as well, reducing the performance of the algorithm. Next, we discuss the implementation of the algorithm.

2-4-2 Algorithm

The multiple-error NLMS algorithm is commonly used in various ANC applications. It is the Multiple-input Multiple-output (MIMO) and normalized variation on the well-known

Filtered-x Least Mean Squares (FXLMS) algorithm. Many improved algorithms exist to reduce computational efforts and increase convergence speed. Reducing computational complexity is not the scope of this study. Hence, we set up the standard multiple-error LMS algorithm by [10] and extend it with the NLMS algorithm [36,37], in order to ensure stability of the algorithm. In this section, we derive the necessary equations and discuss the control scheme.

In the multiple-error NLMS algorithm, shown in Figure 2-8, we denote the number of reference microphones, loudspeakers and error microphones with R , L and E , respectively. Hence, it is called a (R,L,E) ANC system. Furthermore, we denote the filter length by F and impulse response length by J .

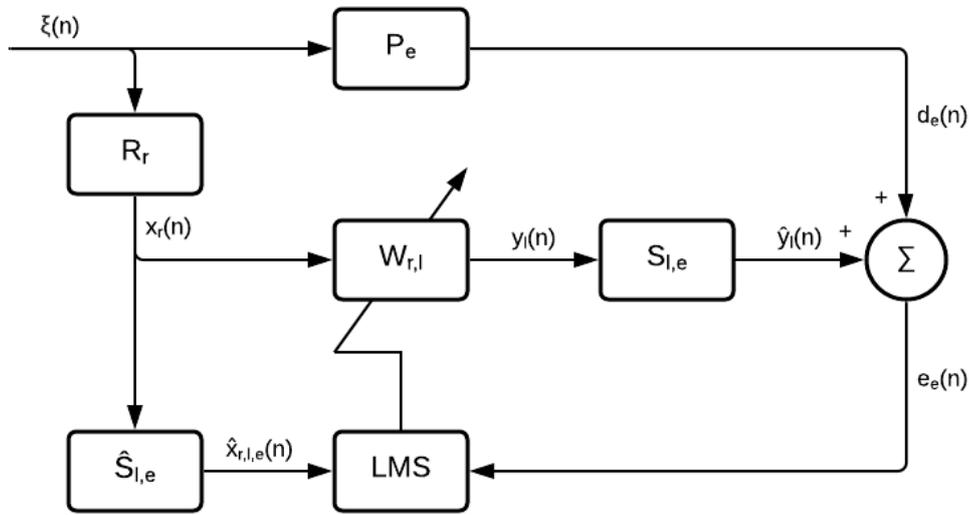


Figure 2-8: Block diagram of the Multiple-error Normalized LMS algorithm.

The primary path of the primary noise signal at the e -th error mic is denoted as $d_e(n) = P_e * \xi(n)$, which is the convolution between the impulse responses P_e and primary noise signal $\xi(n)$ at the origin. Here, P_e is the impulse response of the aperture, which is the Fourier transform of the frequency response of the aperture Acoustic Transfer Function (ATF) derived in Chapter 3. The incoming signal at the r -th reference microphone, is given by:

$$x_r(n) = [x_r(n), x_r(n-1), \dots, x_r(n-F+1)]^T. \quad (2-15)$$

Furthermore, we obtain the estimated secondary path signal \hat{x}_{rle} between loudspeaker l and error microphone e for each reference microphone r with:

$$\hat{x}_{rle}(n) = \hat{S}_{le} * x_r(n), \quad (2-16)$$

where \hat{S}_{le} is the impulse response between loudspeaker l and error microphone e . This impulse response is the Fourier transform of the frequency response of the loudspeaker ATF, also

derived in Chapter 3. Noteworthy is that $\hat{S}_{le} = S_{le}$ in our model, as we do not have an estimated secondary path in our simulation environment. It is denoted as:

$$\hat{S}_{le} = S_{le} = [S_{le}^{(0)}, S_{le}^{(1)}, \dots, S_{le}^{(J-1)}]^T. \quad (2-17)$$

Next, we note the filter of the l -th loudspeaker based on reference microphone r as:

$$w_{rl}(n) = [w_{rl}^{(0)}(n), w_{rl}^{(1)}(n), \dots, w_{rl}^{(F-1)}(n)]^T. \quad (2-18)$$

We obtain the l -th loudspeaker signal at time n with:

$$y_l(n) = \sum_{r=1}^R w_{rl}^T(n) x_r(n). \quad (2-19)$$

Subsequently, by convoluting the signals from all loudspeakers with their secondary path impulse response and sum them, we obtain the secondary noise signal at error microphone e with:

$$\hat{y}_e(n) = \sum_{l=1}^L S_{le}^T * y_l(n). \quad (2-20)$$

Finally, we can calculate the error signal at error microphone e with:

$$e_e(n) = d_e(n) + \hat{y}_e(n). \quad (2-21)$$

Each iteration, the filter weights are updated. For a regular Least Mean Squares (LMS) algorithm, this is done with [7]:

$$w_{rl}(n+1) = w_{rl}(n) - \mu \sum_{e=1}^E e_e(n) \hat{x}_{rle}(n). \quad (2-22)$$

The LMS algorithm deploys the steepest descent method to find the optimal filter weight. Here, μ is the learning rate. However, the correction $\mu \sum_{e=1}^E e_e(n) \hat{x}_{rle}(n)$ that is applied to the filter weights is directly proportional to the input $\hat{x}_{rle}(n)$. In case of a large $\hat{x}_{rle}(n)$, the algorithm faces a gradient noise amplification problem [36]. Therefore, the NLMS algorithm is employed. This algorithm normalizes the correction applied to the filter update. For a Single-input Single-output (SISO) algorithm, the update equation is:

$$w(n+1) = w(n) - \frac{\tilde{\mu}}{\epsilon + \|\hat{x}(n)\|^2(F + \Delta)} e(n) \hat{x}(n), \quad (2-23)$$

with adaptation constant $0 < \tilde{\mu} < 1$, determining the convergence speed and $\epsilon > 0$, a positive constant that prevents numerical difficulties when $\|\hat{x}(n)\|^2$ becomes very small, F the filter length and Δ the overall delay in the secondary path [7, 36].

However, in a MIMO system as we deploy, there are E error microphones, leading to E filtered input signals $\hat{x}_{rle}(n)$ for each loudspeaker. Therefore, we adjust Eq. (2-23) to include the power of the most powerful input signal, as follows:

$$w_{rl}(n+1) = w_{rl}(n) - \frac{\tilde{\mu}}{\epsilon + \max_e(\|\hat{x}_{rle}(n)\|^2)(F + \Delta)} \sum_{e=1}^E e_e(n) \hat{x}_{rle}(n). \quad (2-24)$$

2-4-3 Stability and Convergence

It is important to ensure the stability of the algorithm. The LMS algorithm is discussed extensively in the literature. The reader is referred to Haykin [36] for complete derivations of the LMS and NLMS algorithms and their stability. Most important for implementation are the parameters that are set. The LMS algorithm is stable if $0 < \mu < \mu^{max}$ [17], where

$$\mu^{max} = \frac{\tilde{\mu}}{\epsilon + \|\hat{x}(n)\|^2(F + \Delta)}. \quad (2-25)$$

The stability criteria is extended to the MIMO case, where we take into account the power of all filtered input signals. Therefore, stability is (in the most conservative sense) maintained when using the following learning rate:

$$\mu_l^{max} = \frac{\tilde{\mu}}{\epsilon + \max_e(\|\hat{x}_{rle}(n)\|^2)(F + \Delta)}, \quad (2-26)$$

for each loudspeaker l , where we seek the maximum power input of all filtered reference signals.

As stated, for stability $0 < \tilde{\mu} < 1$. Clearly, if $\tilde{\mu}$ large, the algorithm converges faster [36]. However, this could lead to a worse algorithm's performance in steady-state. Optimizing the performance of this algorithm is not the main scope of the study. However, an informal parameter optimization with various values for $\tilde{\mu}$ is executed.

2-4-4 Algorithm Parameters

One of the drawbacks of MIMO LMS based algorithms is their complexity. Increasing the number reference microphones R , loudspeakers L , or error microphones E drastically increases the computational load of the algorithm. Therefore, many studies investigate the reduction of complexity of MIMO LMS based algorithms [10, 20, 38] to make them feasible for ANC for apertures. However, as the NLMS algorithm is used as reference algorithm, reducing computational load does not fit in the scope of this thesis. Nevertheless, we need to choose parameters for the algorithm. Firstly, the system implemented will be deployed with one reference microphone, and we compare two loudspeaker arrays. The number of error microphones should be enough to cover all directions from the aperture. Secondly, increasing the filter length F indisputably increases computational load. Moreover, the filter length influences the performance of the filter. Therefore, choosing the correct filter length is important. It should have substantial length to include all the significant taps of the impulse

response of the system. Thirdly, the length of the impulse responses of the aperture and loudspeakers should be long enough to include all significant information. The number of error microphones, the filter length, and the optimal impulse response length are chosen after an informal parameter optimization.

2-5 Summary

The purpose of this chapter was to discuss important existing theories we need to investigate for ANC for apertures. It established the fundamental signal processing concepts necessary for the simulation environment and wave-domain algorithm. Furthermore, we covered the reference algorithm. The subsequent points summarize this chapter:

- We set some basic notations, definitions, and coordinate systems to ensure consistency throughout the report.
- Modeling in the wave-domain is done by using the STFT. To ensure perfect reconstruction, we need to make sure we have a tight frame window function. This means, it has to fulfill the following condition: $\sum_m \hat{w}(n - mH)^2 = 1$. We presented a method to normalize any window to fulfill this condition.
- Huygens' Principle tells us that we can describe any new wavefront, by a set of closely spaced point sources on the current wavefront, that emit spherical waves.
- We clarified the necessary equations for implementation of the multiple-error NLMS algorithm. This algorithm is extensively used in ANC in general and for ANC for apertures, specifically. Therefore, it is used as a baseline method.
- As prior studies indicate, we expect good attenuation performance with the multiple-error NLMS algorithm. However, we do expect slow convergence and heavy computational load as presented in the state-of-the-art literature. Moreover, the use of error microphones makes it a rather non-feasible algorithm for practical purposes of ANC for apertures.

Simulation Model and Acoustic Transfer Functions

As we determined in the introduction, we seek to develop a simulation environment. In this environment, we analyze the performance of the wave-domain algorithm, and the reference algorithm, for various loudspeaker arrays and incident angles. Furthermore, we need to derive an aperture Acoustic Transfer Function (ATF) that is valid within a reasonable distance from the aperture. Both aspects are covered in this chapter. Firstly, we cover the general characteristics of the simulation environment in Section 3-1. Thereafter, the inferring of Acoustic Transfer Functions (ATFs) for the aperture and loudspeakers is given in Section 3-2. In Section 3-3, we discuss the implementation of the block processing we deploy in the simulation environment and cover relevant limitations. We give a schematic overview of the simulation implementation in Section 3-4 and summarize the chapter in Section 3-5.

3-1 Simulation Environment

To test the algorithm developed in this study, we need a simulation environment in which we can measure its performance and compare it to the existing LMS-based algorithm. We realize this simulation environment with a signal-processing approach that calculates soundfields based on the frequency responses from ATFs. This is applied three times to calculate soundfields from (1) a plane wave propagating from a fixed primary source towards the aperture, (2) the aperture, and (3) the loudspeaker array. Before we discuss the used ATFs, we discuss general properties of the environment in the following subsections.

3-1-1 3D Representation

As this thesis study is a continuation of the work by Miller and Downey [13,14], the simulation environment should resemble their developed physical setup. Hence, for 3D modeling, we adopted the same physical properties. Figure 3-1a is a picture of the physical setup that is

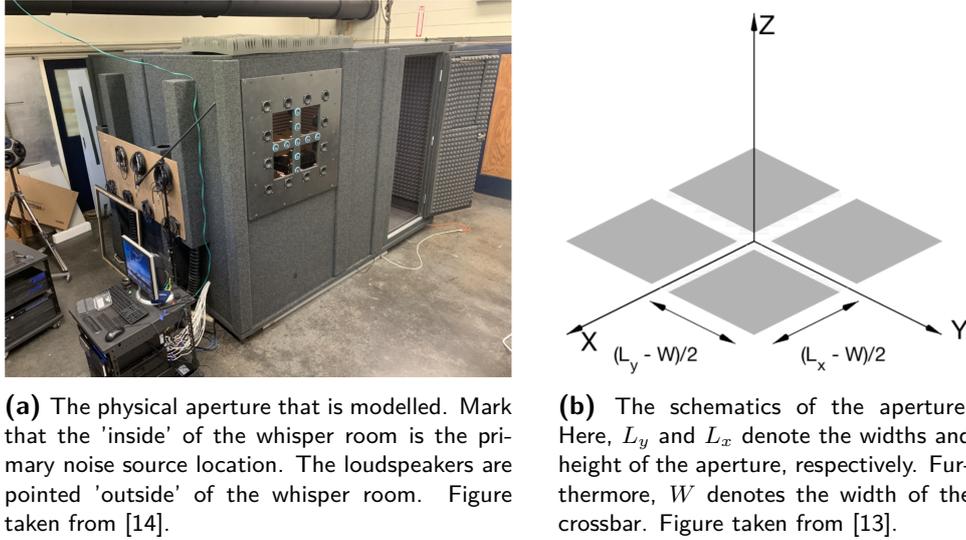


Figure 3-1: Picture of physical test environment and a graphical schematic of the aperture.

modeled here. The aperture contains a crossbar, and it has loudspeakers positioned on the crossbar and the boundary of the aperture. Figure 3-1b shows a graphical representation of the aperture and has the following dimensions: the height and width are L_x and L_y , respectively, and the crossbar has a width of W_+ .

Both the open-loop wave-domain algorithm and the multiple-error LMS system use a reference microphone. In a physical experiment, the reference microphone must fulfill certain properties, as well as be calibrated. However, this is out of the scope of this study. Hence, we assume that the reference microphone has an ideal frequency response, and only one microphone is enough for modeling the incoming noise. We position the microphone at the origin $((x, y, z) = (0, 0, 0))$, in the middle of the aperture. Furthermore, we assume the incident angle of the plane wave, denoted with θ_0 and ϕ_0 , of the incoming primary noise plane wave, is known a priori. Methods for calculating this angle are already available, e.g. [39] or [40] and will not be covered here.

In addition to the reference microphone, the loudspeaker array is modeled. Two types of arrays are compared. The first, denoted as the sparse array, is the array used in the study by Downey [14], containing 21 loudspeakers, sparsely positioned on the crossbar and aperture boundaries. Secondly, we model a grid array that has 49 loudspeakers distributed over the entire aperture. This array is implemented because prior studies showed that grid-based arrays outperform boundary arrays [6]. Moreover, most prior studies on aperture ANC systems employed a grid-based array. Figure 3-2 shows the schematic positioning of the loudspeakers for both arrays. Again, calibration of the loudspeakers is out of scope, and we assume that the loudspeakers have a flat frequency response.

For the reference algorithm, the multiple-error Least Mean Squares (LMS), error microphones must be positioned in the room. Their purpose is to feedback the error signals such that the algorithm can update the filter weights, as discussed in Chapter 2. Similar to the reference microphone, we assume that the error microphones have flat frequency responses and do not need calibration.

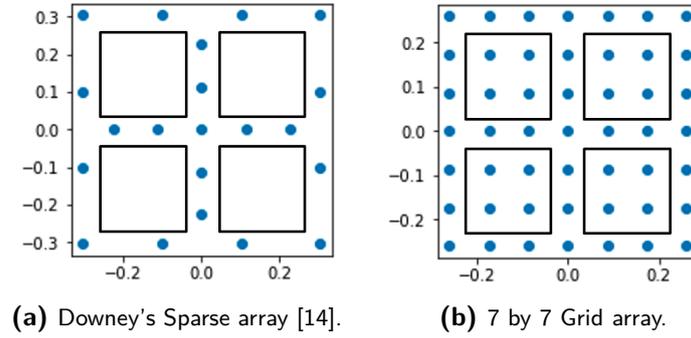


Figure 3-2: 3D loudspeaker array schematics.

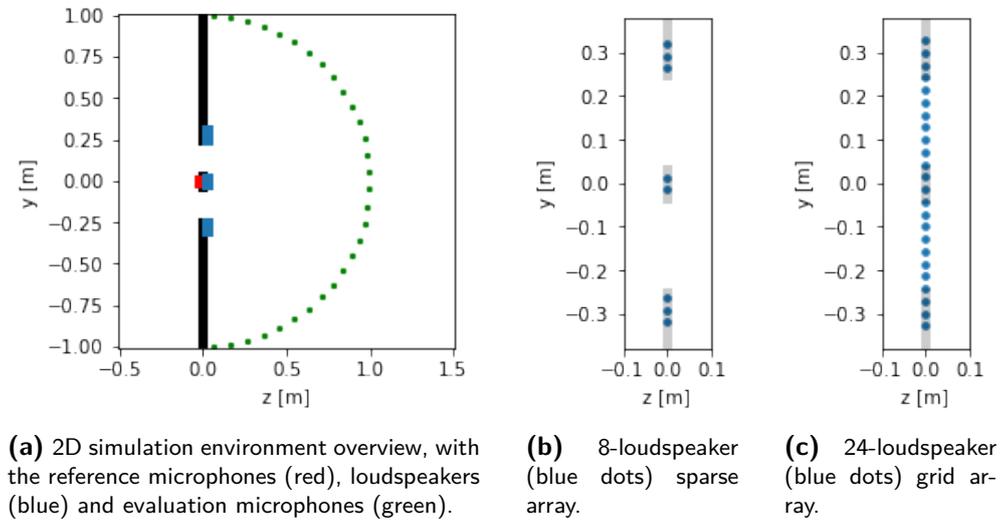


Figure 3-3: 2D simulation environment and array schematics.

Finally, we have the evaluation microphones. Their function is to measure the sound pressure from the aperture in all directions, both when a control algorithm is active as well as without control. In the virtual simulation environment, we evenly distributed these microphones over a hemisphere surrounding the aperture [41], such that we can measure sound energy in all directions from the aperture into the room.

3-1-2 2D Representation

Additional to the 3D simulation environment, a 2D simplification is developed. The computational effort of a 2D model is much lower compared to 3D. This gives the opportunity to quickly iterate and test algorithms before applying them in the 3D environment. Furthermore, 2D modeling is a common approach in wave-domain studies [23, 24].

The 2D representation is a cross-section of the 3D aperture. We remove the height and model only in (z, y) coordinates. The aperture entails a L_y wide opening, containing a crossbar in the middle, with width set as W_+ . A schematic overview is shown in Figure 3-3a. Similar to the 3D model, we position a reference microphone at the origin and assume perfect calibration.

Similar to the 3D tests, we execute the 2D simulations with two types of arrays. The sparse array contains 8 loudspeakers, divided over the boundaries and crossbar, as can be seen in Figure 3-3b. The grid array is modeled as a row of 24 loudspeakers over the whole width of the aperture, shown in Figure 3-3c. Finally, the evaluation microphones are positioned at an arc, clarified as the green dots in Figure 3-3a.

3-2 Acoustic Transfer Functions

The simulation environment employs multiple ATFs. These are used in parallel to describe what happens when a wave propagates from outside through the aperture into the room, as well as the waves from the loudspeakers. Basic plane wave modeling is derived in Appendix B. In this section, we focus on the aperture ATF and loudspeaker ATF.

3-2-1 Aperture ATF

To model the aperture, we seek an ATF that relates the pressure of the plane wave signal in the aperture with the pressure at an arbitrary evaluation position in the room. For this, we follow the studies of Miller [13], and Downey [14]. In these studies, the aperture is modeled as a vibrating plate in an infinitely large wall [27]. The ATF of a single square vibrating plate is given as [13, 27]:

$$H^{ap}(\mathbf{x}, k, \theta_0, \phi_0) = \frac{jck\rho_0}{2\pi} \dot{\omega}_0 L_x L_y \frac{e^{-jkr}}{r} \operatorname{sinc}\left(\frac{L_x k(\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)}{2}\right) \operatorname{sinc}\left(\frac{L_y k(\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)}{2}\right), \quad (3-1)$$

where j is the imaginary number, c is the speed of sound, k is the wavenumber, ρ_0 is the density of air, $\dot{\omega}_0$ is a gain constant, L_x and L_y denote the aperture dimensions, $\mathbf{x} = (r, \theta, \phi) = (x, y, z)$ describes the position at which we calculate the pressure and θ_0 and ϕ_0 indicate the incident angle of the primary noise. The derivation of this equation, derived from Rayleigh's first Integral equation, can be found in the accompanying literature study [12]. In order to model the cross-bar in the aperture, [13] extends Eq. (3-1) as a stack of four vibrating plates with the a single origin. This gives:

$$H^{ap}(\mathbf{x}, k, \theta_0, \phi_0) = \frac{jck\rho_0}{2\pi} \dot{\omega}_0 \frac{e^{-jkr}}{r} \left(L_x L_y \operatorname{sinc}\left(\frac{L_x k(\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)}{2}\right) \operatorname{sinc}\left(\frac{L_y k(\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)}{2}\right) - L_x W_+ \operatorname{sinc}\left(\frac{L_x k(\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)}{2}\right) \operatorname{sinc}\left(\frac{W_+ k(\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)}{2}\right) - L_y W_+ \operatorname{sinc}\left(\frac{L_y k(\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)}{2}\right) \operatorname{sinc}\left(\frac{W_+ k(\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)}{2}\right) + W_+^2 \operatorname{sinc}\left(\frac{W_+ k(\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)}{2}\right) \operatorname{sinc}\left(\frac{W_+ k(\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)}{2}\right) \right), \quad (3-2)$$

where W_+ is the crossbar width. This equation is valid in the far-field. However, if we have aperture dimensions of, e.g. $L_x = L_y = 0.5$ m, the far-field starts at $r \gg ka^2 = 2\pi fa^2/c = 2\pi \cdot 2048 \cdot 0.354^2/343 = 4.70$ m [42]. This is too far from the aperture for our application. We seek an approach that validly describes the wave from approximately 1 m from the aperture onwards. Hence, we elaborate further and develop the following aperture ATF.

The method is extended by summing a multitude of smaller vibrating plates. With this approach we model more specifically what happens when a wave propagates through an aperture. It describes the soundfield by an aperture with a crossbar more accurately at closer distances. This allows us to use it for our algorithm. So, we express the pressure at evaluation position ($\mathbf{x} = (x_e, y_e, z_e)$) as a sum of the pressures by \hat{P} square vibrating plates. The equation for 3D modeling is then derived as:

$$H^{ap}(\mathbf{x}, k, \theta_0, \phi_0) = \frac{jck\rho_0}{2\pi} \dot{\omega}_0 \Delta L_x \Delta L_y \sum_{i=1}^{\hat{P}} D_i, \quad (3-3)$$

where ΔL_x and ΔL_y are aperture section dimensions. D_i is the directivity, of each plate, defined as:

$$D_i = \frac{e^{-jk(r_i+\tau_i)}}{r_i} \text{sinc}\left(\frac{\Delta L_x k(\sin \theta_i \cos \phi_i - \sin \theta_0 \cos \phi_0)}{2}\right) \text{sinc}\left(\frac{\Delta L_y k(\sin \theta_i \sin \phi_i - \sin \theta_0 \sin \phi_0)}{2}\right), \quad (3-4)$$

where, for section i , r_i , θ_i and ϕ_i are the adjusted spherical coordinates and τ_i is a delay term due to the incident angle of the plane wave. We define the coordinates as:

$$r_i = \sqrt{(x_e - x_i)^2 + (y_e - y_i)^2 + z_i^2}, \quad (3-5)$$

$$\theta_i = \arccos(z_i/r_i), \quad (3-6)$$

$$\phi_i = \text{atan2}((y_e - y_i), (x_e - x_i)), \quad (3-7)$$

where (x_i, y_i, z_i) denotes the origin of section i . Furthermore, the delay term is calculated as the perpendicular distance between the plane of the plane wave in the origin of the aperture, and the origin of section i . It is defined as:

$$\tau_i = \frac{\sin(\theta_0) \cos(\phi_0)x_i + \sin(\theta_0) \sin(\phi_0)y_i}{\sqrt{(\sin(\theta_0) \cos(\phi_0))^2 + (\sin(\theta_0) \sin(\phi_0))^2 + (\cos(\theta_0))^2}}, \quad (3-8)$$

and it makes sure that section i has the correct phase shift resulting from the incident angle of the incoming noise.

Modeling in 2D is done by removing the height ΔL_x and emitting the sinc function of the x direction. Essentially, this describes an infinitely thin window. The transfer function of 3D, in Eq. (3-3), reduces to:

$$H^{ap}(\mathbf{x}, k, \theta_0, \phi_0) = \frac{jc k \rho_0}{2\pi} \dot{\omega}_0 \Delta L_y \sum_{i=1}^{\hat{P}} D_i, \quad (3-9)$$

and the directivity from Eq. (3-4) is downsized to:

$$D_i = \frac{e^{-jk(r_i + \tau_i)}}{r_i} \text{sinc}\left(\frac{\Delta L_y k (\sin \theta_i \sin \phi_i - \sin \theta_0 \sin \phi_0)}{2}\right), \quad (3-10)$$

and we have the adjusted coordinates as:

$$r_i = \sqrt{(y_e - y_i)^2 + z_i^2}, \quad (3-11)$$

$$\theta_i = \arccos(z_i / r_i), \quad (3-12)$$

$$\phi_i = \text{atan2}((y_e - y_i), 0), \quad (3-13)$$

and the delay-term ends up being:

$$\tau_i = \frac{\sin(\theta_0) \sin(\phi_0) y_i}{\sqrt{(\sin(\theta_0) \sin(\phi_0))^2 + (\cos(\theta_0))^2}}. \quad (3-14)$$

3-2-2 Loudspeaker ATF

Similar to the aperture ATF, we need an ATF that relates the sound pressure at an evaluation position to the loudspeaker signal. This is achieved by modeling the loudspeaker ATF as a monopole [27, 43]. Accordingly, the pressure at position \mathbf{x} from the loudspeaker array is a sum of each individual loudspeaker. A monopole is modeled as:

$$H_q^{ls}(\mathbf{x}, k) = \frac{jc k \rho_0}{4\pi} A_q \frac{e^{-jkr_q}}{r_q}, \quad (3-15)$$

in which $A_q = 4\pi a^2 u_0$ is the monopole amplitude, with u_0 a surface velocity gain constant and a the radius of the monopole. Furthermore, r_q is the adjusted spherical radius from the monopole to a position \mathbf{x} in the room, defined as:

$$r_q = \sqrt{(x_e - x_q)^2 + (y_e - y_q)^2 + z_q^2}, \quad (3-16)$$

where (x_q, y_q, z_q) denotes the position of the loudspeaker. This ATF holds in 3D, and for 2D we set $x_q = 0$.

3-3 Block Processing

This section briefly covers the block processing concept that is deployed in this work. This enables us to describe any transformation by use of an ATF, such as the ones discussed in the prior section. An arbitrary input signal ($x(n)$) is transformed to the wave-domain with the Short-time Fourier Transform (STFT). First, we cover implementation of the block processing. Thereafter, we discuss the time-delay wrapping, as time shifts in the STFT cause for issues. Finally, we present a method to quantify the limitation of a fixed window-size N .

3-3-1 Implementation

The signal is broken into M blocks ($x_m(n)$) using an analysis window function $w(n)$, of length N samples and we apply the Discrete Fourier Transform (DFT) to each block. The window-function, $w(n)$ is chosen to fulfill $\sum_{m \in \mathcal{Z}} w(n - mH)^2 = 1$, such that we have perfect reconstruction, as discussed in the Section 2-2-5. Lets denote the coefficient vector containing frequency information of the m -th block as:

$$\mathbf{X}_m(k) = \text{STFT}(x_m(n)). \quad (3-17)$$

Thereafter, we do an element-wise multiplication of the coefficient vector with an ATF $H(k)$:

$$\hat{\mathbf{X}}_m(k) = \mathbf{X}_m(k) \odot H(k). \quad (3-18)$$

Finally, the transformed signal $\hat{x}(n)$ can be obtained with the Inverse Short-time Fourier Transform (I-STFT):

$$\hat{x}(n) = \text{I-STFT}(\hat{\mathbf{X}}_m(k)). \quad (3-19)$$

3-3-2 Time-delay Wrapping

The block-processing, elaborated in the prior section, has a limiting artifact. When phase-shifts by ATFs become significant compared to the window length, the circularity property of the STFT, assuming that $x_m(n)$ is periodic, causes for wrapping of the signals. That means that a positive time-delay shifts the signal such that the last part (in time), appears at the beginning of the block. This issue causes the block processing approach to induce errors in the transformed signals. An illustration is shown in Figure 3-4. Deploying zero-padding can reduce this issue. However, this emits the shifted signal content that would otherwise appear at the beginning of the block. Omitting this signal part leads to a loss of signal, limiting the accuracy of the block processing. In this section, we briefly cover the method we developed to reduce this issue significantly.

In some ATFs, like Eq. (3-3), Eq. (3-9) and Eq. (3-15), the time-delay is encapsulated in the e^{-jkr} term, where r is a distance. The wave propagates over this distance with the speed of sound c , leading to a time-delay. To overcome the issue of wrapping, we apply the most

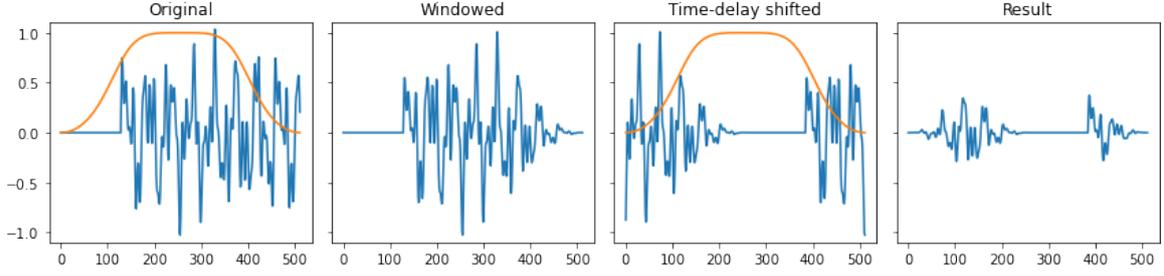


Figure 3-4: The time delay wrapping issue that occurs when long delays are implemented with short STFT blocks. Due to the periodicity assumption of the Fourier transform, the time-delay shift causes the end of the signal block to wrap to the beginning of the block, visible when taking the following steps. The original signal (1) is windowed to obtain a windowed signal (2). Then, the signal is transformed to the frequency-domain, a time-delay is applied and transformed back to the time-domain (3). Finally, the window is applied again, resulting in a wrapped signal (4).

significant part of the time-delay in the time-domain. Let us define the procedure for a simplified ATF, defined as:

$$H = Ae^{-jkr_{delay}}. \quad (3-20)$$

where A is any other part of the ATF that does not include the phase-shift and k is the wavenumber. We calculate the total delay in samples:

$$T_{total} = \frac{f_s r_{delay}}{c}, \quad (3-21)$$

where f_s is the sample rate and c the speed of sound. However, T_{total} is often not an integer, and in the discrete time-domain, we can only shift signals by integer steps. Hence, we divide the total delay between an integer and a decimal term:

$$T_{total} = \hat{T}_{int} + \tilde{T}_{dec}, \quad (3-22)$$

where the integer term is defined as:

$$\hat{T}_{int} = \lfloor T_{total} \rfloor, \quad (3-23)$$

where $\lfloor \cdot \rfloor$ is rounding to the next integer. We retrieve the adjusted ATF as:

$$\hat{H}(k) = e^{-jkc\tilde{T}_{dec}/f_s}, \quad (3-24)$$

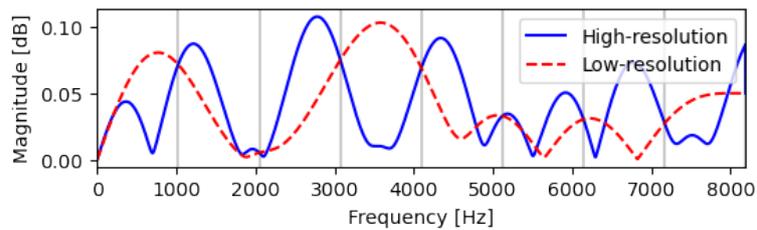
and plug this, together with the integer time shift, in Eq. (3-18). Then, we achieve a non-wrapping, time-shifted block processing procedure for this ATF as:

$$\hat{x}(n + \hat{T}_{int}) = \text{I-STFT}(\mathbf{X}_m(k) \odot e^{-jkc\tilde{T}_{dec}/f_s}). \quad (3-25)$$

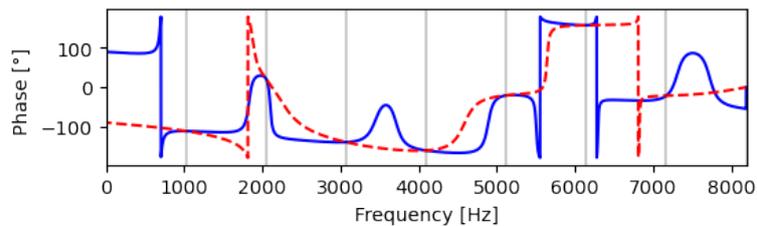
3-3-3 Window-size and Frequency Resolution

Aside from the time-delay wrapping that influences the accuracy of block-processing with the STFT, another limitation arises due to the blockwise processing. As the STFT uses the DFT, we work with a sampled frequency response. That means that we sample the continuous ATFs given in Eq. (3-3) and Eq. (3-15). When sampling, aliasing can occur. We elaborate on this issue in this section and develop a tool to measure the significance.

The application of the ATF in the discrete wave-domain is the root of the problem. The ATF is a continuous function. However, we apply it in a discrete sense. This means that we sample the frequency response of the ATF. Similar to the sampling of signals in the time-domain, aliasing occurs when we sample in the wave-domain. More specifically, when sampling, part of the behavior that happens ‘in between’ the sampled points is disregarded: the sample is the average of that measured section of the signal. With a shorter STFT window-size N , we have less discrete frequency bins, leading to a lower frequency resolution. Similar to sampling in the time-domain, sampling in frequency with less frequency bins means that only smooth behavior of the frequency response is captured. Let us look at an example of the frequency response of the aperture ATF, evaluated at a point in the room. The frequency response with high frequency resolution, with $N = f_s$ (close to the continuous case) is compared with the low frequency resolution version, with $N = 16$ samples.



(a) Magnitude plot.



(b) Phase plot.

Figure 3-5: Aperture ATF frequency responses for the high (blue) and low-resolution (red) scenario. The low-resolution line shows a smoothed version and corresponds with the high-resolution version at the grey lines, that indicate the frequency bins.

Figure 3-5 shows the relatively non-smooth frequency response in solid blue. The grey lines indicate the frequency bins that correspond to the forward-STFT for a block size of $N = 16$ samples. The red dashed line corresponds to the low-resolution frequency response of the aperture transfer function. In this small window-size case, the impulse response of the high-resolution is windowed drastically. The red dashed line in Figure 3-6 shows the windowed impulse response. It becomes clear that the low resolution results in an error, as the two impulse responses do not overlap.

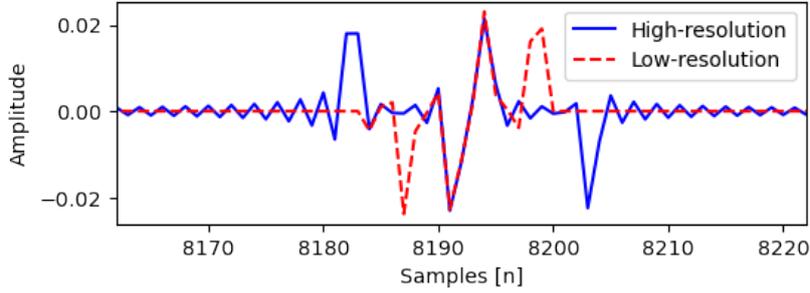


Figure 3-6: The impulse response of the aperture ATF. The solid blue line shows the original impulse response. The dashed red line shows the filtered case, where a rectangular window is applied due to the low frequency resolution. Only the first part of the middle part of the impulse response remains, which directly shows the issue of a short block length.

Let us derive an analytical method to calculate the error caused by approximating an ATF with low resolution. Let us denote the wave-domain variables in bold and time-domain variables in normal font. We start with a frequency weighting, which weights certain frequency content based on the primary noise signals. We denote this by: $\mathbf{s}(k) : \mathbb{R} \rightarrow \mathbb{R}$. This frequency weighting is the average power spectral density of a certain audio set. We use a perfectly flat frequency response weighting, so $\mathbf{s}(k) = 1 \forall k$. Furthermore, $\mathbf{y}(k)$ and $\hat{\mathbf{y}}(k)$ denote an weighted frequency response of ATF and it's approximated (lower frequency resolution) version, respectively. The arbitrary transfer function is denoted as $\mathbf{h}(k)$. Finally, the low 'time' filter, corresponding to the window-size is defined in the time-domain as a rectangular window:

$$w(t) = \begin{cases} 1, & -\frac{N}{2} \leq t \leq \frac{N}{2}, \\ 0, & \text{elsewhere,} \end{cases}, \quad (3-26)$$

where N is the window-size and its wave-domain equivalent is defined as $\mathbf{w}(k) = N\text{sinc}(ckN)$, as the fourier transform of a rectangular window is a sinc-function [27].

The error in wave-domain $\mathbf{e}(k)$ is derived as follows. We begin with the weighted frequency response:

$$\mathbf{y}(k) = \mathbf{h}(k)\mathbf{s}(k), \quad (3-27)$$

and use it to describe the filtered frequency response:

$$\hat{\mathbf{y}}(k) = \mathbf{w}(k) * \mathbf{y}(k) = \mathbf{w}(k) * (\mathbf{h}(k)\mathbf{s}(k)). \quad (3-28)$$

The filtering in the time-domain, a multiplication of the weighted impulse response with the filter, corresponds to a linear convolution between the weighted frequency response and the frequency transformation of the filter in the frequency domain. $\mathbf{y}(k)$ and $\hat{\mathbf{y}}(k)$ are used as ATFs the simulation model. Finally, the frequency response error is the difference between the two frequency responses:

$$\begin{aligned}
\mathbf{e}(k) &= \mathbf{y}(k) - \hat{\mathbf{y}}(k) \\
&= \mathbf{h}(k)\mathbf{s}(k) - \mathbf{w}(k) * (\mathbf{h}(k)\mathbf{s}(k)) \\
&= (1 - \mathbf{w}(k)) * (\mathbf{h}(k)\mathbf{s}(k)).
\end{aligned}
\tag{3-29}$$

The method is summarized with a block diagram in Figure 3-7.

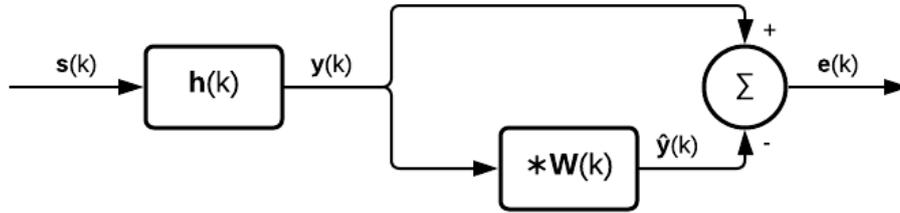


Figure 3-7: The schematic overview of the error analysis procedure with frequency weighting $\mathbf{h}(k)$, weighted frequency response $\mathbf{y}(k)$, its low frequency resolution version $\hat{\mathbf{y}}(k)$ and the error $\mathbf{e}(k)$. The * denotes convolution.

From this, we can calculate the Signal-to-Noise Ratio (SNR) between the frequency response error and the weighted frequency response, or, equivalently by Parseval's theorem [44], the ratio between the weighted impulse response and the error impulse response:

$$SNR = 10\log_{10}\left(\frac{\sum(\mathbf{y}(k))^2}{\sum(\mathbf{e}(k))^2}\right) = 10\log_{10}\left(\frac{\sum(y(t))^2}{\sum(e(t))^2}\right),
\tag{3-30}$$

in dB. The resulting SNR describes how well the approximated weighted frequency response $\hat{\mathbf{y}}(k)$ described the actual weighted frequency response $\mathbf{y}(k)$. This gives us the fundamental performance limit of using the approximated weighted frequency response for frequency weight calculation. In the result section, we will explore the limits of the aperture ATF for various window-sizes, used by the wave-domain algorithm.

3-4 Schematic Overview

The simulation model deploys the discussed ATFs, block processing, and the method to prevent time-delay wrapping. Figure 3-8 is a schematic overview of the implementation for a single evaluation position in the room. The primary noise ($\xi(n)$) takes the primary path via the aperture ATF $H^{ap}(k)$. Here, a STFT with a very large window-size ($\tilde{N} = f_s$) is used, for high frequency resolution. After transforming back to time, we implement the time-delay T^{ap} that we split from the frequency implementation. Eventually, we have the primary noise signal at the evaluation position ($d(n)$). In the secondary path, the primary noise signal is measured by reference microphone R . The measured signal is transformed to the wave-domain with

a STFT with window-size N . Then, for each loudspeaker q , the signal is transformed with its corresponding filter weight $W_q(k)$. The calculation of this weight is discussed in Chapter 4. Next, each adjusted loudspeaker signal is multiplied with the corresponding ATF of the loudspeaker $H_q^{ls}(k)$ and transformed back to time-domain with an I-STFT. We implement the time-delay that we omitted from the loudspeaker ATF. In the end, we sum the signals of the aperture ($d(n)$) and from all loudspeakers ($y_q(n)$) and obtain the error in the evaluation position $e(n)$.

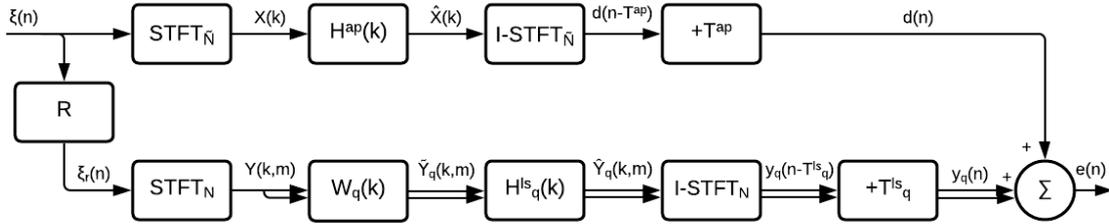


Figure 3-8: Block diagram of the simulation environment.

3-5 Summary

In this chapter, we investigated the simulation environment. We focused on the necessary ATFs for wave propagation modeling and the wave-domain algorithm and covered various aspects of the implementation with block-processing. The most important topics are summarized below:

- We established a simulation environment by using ATFs that generate the soundfields from the aperture and loudspeaker arrays. Block processing is used for implementation.
- A 3D representation of the environment was given, as well as a 2D simplification.
- The aperture is modeled as a vibrating plate. We extended the existing theory by deriving equations that describe a sum of multiple vibrating plates. This allows us to use the ATF at 0.88 m from the aperture.
- Loudspeakers are represented by monopoles in the simulation environment.
- The use of block processing induces an issue we call time-delay wrapping. Here, relatively large phase-shifts in the ATF cause issues. We presented a solution that removes the significant phase-shifts to a time-delay in the time-domain.
- In the block processing approach, the window-size of the STFT results in wave-domain aliasing. This window-size determines the number of frequency bins for which we obtain the frequency response of the ATF. A small window-size leads to a loss of frequency response information. Hence, this fundamentally limits the performance of a wave-domain algorithm. We introduced an error analysis method that can give us insight into these fundamental limits.

Wave-Domain Algorithm Design

In the previous chapter, we developed a simulation environment and defined Acoustic Transfer Functions (ATFs) for aperture and loudspeaker soundfields. We continue with this information in this chapter as we design the wave-domain algorithm. The main focus is to come up with the necessary equations to calculate loudspeaker filter-weights that minimize the soundfield of the aperture. For this, we first discuss the control region, the spatial region in which we want to minimize the sound energy in Section 4-1. Thereafter, the algorithm is developed for which a soundfield basis expansion forms the fundament. We will include a brief discussion on algorithmic delay compensation. Both are given in Section 4-2. An in-depth discussion of two parameters of the algorithm is covered in Section 4-3. Finally, Section 4-4 concludes this chapter.

4-1 Control Region

Any wave-domain algorithm rests on the principle of minimizing the sum of soundfields in a spatial control region. We denote this region with \mathbb{D} . For proper Active Noise Control (ANC) for apertures, we ensure global control by specifying this control region in all directions from the aperture into the room. Hence, in the 2D simulations, we denote the control region as an arc with finite thickness [6, 11, 45]:

$$\mathbb{D}_{2D} = \begin{cases} r_{min} \leq r \leq r_{max}, \\ 0 \leq \theta \leq \frac{\pi}{2}, \\ \phi = \frac{\pi}{2} \ \& \ \phi = -\frac{\pi}{2}, \end{cases} \quad (4-1)$$

where r_{min} and r_{max} determine the thickness of the arc. This is visualized in Figure 4-1. Moreover, we specify the 3D control region as a half spherical shell with finite thickness [46], and extend Eq. (4-1) to:

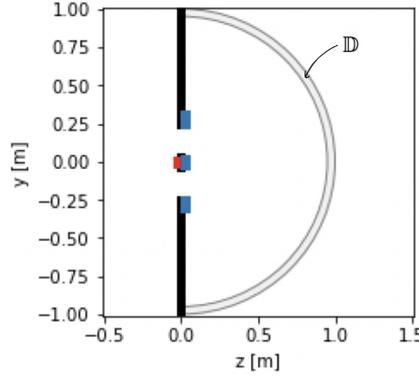


Figure 4-1: 2D cross-section of the environment with control region \mathbb{D} , a hemisphere in the far-field, between r_{min} and r_{max} from the aperture.

$$\mathbb{D}_{3D} = \begin{cases} r_{min} \leq r \leq r_{max}, \\ 0 \leq \theta \leq \frac{\pi}{2}, \\ 0 \leq \phi \leq 2\pi. \end{cases} \quad (4-2)$$

A finite thickness ensures that we obtain global control in all directions. We create a new wavefront, based on the current wavefront with reduced sound energy in the control region. Consequently, the new wavefront behind the control region has reduced sound energy, according to the Huygens' principle discussed in 2-3.

4-2 Algorithm development

This section discusses the general open-loop wave-domain algorithm, applicable to both the 2D and 3D situations. The theory closely follows the steps in [23,24]. The algorithm employs a soundfield basis expansions. Hence, we explore that theory before deriving the actual loudspeaker weight calculations.

The following notation is used: matrices and vectors are denoted with upper and lower boldface respectively: \mathbf{C} and \mathbf{y} . $\mathbf{x} \in \mathbb{R}^3$ is an arbitrary spatial observation point. The number of loudspeakers is Q .

4-2-1 Soundfield Basis Expansion

Any soundfield function can be written as a sum of weighted basis functions, where the basis function set is an orthonormal set of solutions to the Helmholtz equation [22, 23, 27]. The Fourier transform of the time-domain wave equation gives the Helmholtz equation, defined as:

$$\nabla^2 p + k^2 p = 0, \quad (4-3)$$

with function $p(x, y, z, \omega)$ and wavenumber k . Figure 4-2 illustrates the concept of soundfield basis expansion, where a finite sum of simple waves can be used to describe an arbitrary soundfield in an observation region.

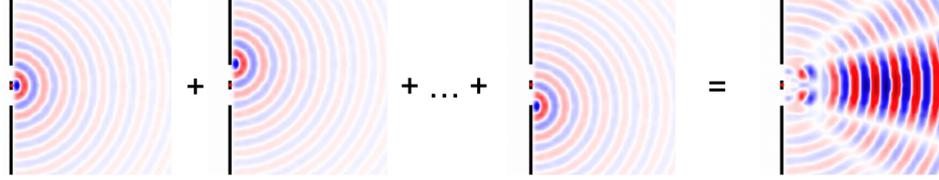


Figure 4-2: Any soundfield can be described by a finite weighted sum of simple waves.

Let us derive this in equations. The soundfield over the observation region at single wavenumber k , denoted $S(\mathbf{x}, k) : \mathbb{D} \times \mathbb{R} \rightarrow \mathbb{C}$ is written as a weighted series of basis functions $\{U_g\}_{g \in \mathcal{G}}$:

$$S(\mathbf{x}, k) = \sum_g E_g U_g(\mathbf{x}, k), \quad (4-4)$$

where $S(\mathbf{x}, k)$ is the soundfield, E_g are G coefficients and $U_g(\mathbf{x}, k)$ is a $G \times 1$ vector. We assume that all feasible solutions on \mathbb{D} fall in the Hilbert space spanned by the orthonormal set $\{U_g\}_{g \in \mathcal{G}}$. The inner-product is defined as:

$$\langle Y_1, Y_2 \rangle = \int_{\mathbb{D}} Y_1(\mathbf{x}) Y_2^H(\mathbf{x}) d\mathbf{x}, \quad (4-5)$$

where Y_1 and Y_2 are functions of the form $Y_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $Y_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$. The integration is conducted in the domain of \mathbb{D}^3 . The orthonormal set $U_g(\mathbf{x}, k)$ has the property $\langle U_i(\mathbf{x}, k), U_j(\mathbf{x}, k) \rangle = \delta_{ij}$. For a given $S(\mathbf{x}, k)$ and $U_g(\mathbf{x}, k)$, the coefficients E_g are obtained with $E_g = \langle S(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle$.

4-2-2 Orthonormalization of a Set of Basis Functions

Next, we discuss how to construct the orthonormal set of basis functions. We denote this orthonormal set of basis functions as a vector:

$$U = [U_1 \ U_2 \ \cdots \ U_G]^T, \quad (4-6)$$

To find this set, we start with a set of non-orthogonal functions that solve the wave-equation. A simple set of solutions is are plane waves. We set $f_g(\mathbf{x}, k) : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{C}$ that represent G plane waves in G directions, defined as:

$$f_g(\mathbf{x}, k) = e^{jk\mathbf{x} \cdot \hat{\beta}_g}, \quad (4-7)$$

where $\hat{\beta}_g$ is the unit vector in the direction of the g -th plane wave. Let's first derive the 2D case. Here, we $\beta'_g = (g-1)\Delta\beta$, $g = 1, \dots, G$ with $\Delta\beta = 2\pi/G$ and finally $\hat{\beta}_g \equiv (1, \beta'_g)$, such that we have the directions evenly distributed over a 2D plane. For the 3D case, we

use a dataset of evenly distributed directions in a sphere [41] and set $\hat{\beta}_{\mathbf{g}} \equiv (1, \theta_g, \phi_g)$. We normalize each basis function with Eq. (4-5) to obtain $\hat{f}_g(\mathbf{x}, k) = \frac{f_g}{\|f_g\|}$ and combine the set of normalized plane waves in a vector:

$$\hat{f} = [\hat{f}_1 \hat{f}_2 \cdots \hat{f}_G]^T, \quad (4-8)$$

Next, we find a lower triangular matrix R such that $U = R\hat{f}$, where U is the vector containing G orthonormal basis functions. We define a matrix containing inner-products of Eq. (4-8) with itself for all angles:

$$F = \hat{f}\hat{f}^T = \begin{bmatrix} F_{(1,1)} & F_{(1,2)} & \cdots & F_{(1,G)} \\ F_{(2,1)} & F_{(2,2)} & & \vdots \\ \vdots & & \ddots & \\ F_{(G,1)} & \cdots & \cdots & F_{(G,G)} \end{bmatrix}, \quad (4-9)$$

where we assume that F is positive definite: $\mathbf{x}^H F \mathbf{x} > 0 \forall \mathbf{x} \in \mathbb{C}^n$. The required matrix R is defined as lower triangular, leading to:

$$\begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_G \end{bmatrix} = \begin{bmatrix} R_{(1,1)} & 0 & \cdots & 0 \\ R_{(2,1)} & R_{(2,2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ R_{(G,1)} & R_{(G,2)} & \cdots & R_{(G,G)} \end{bmatrix} \begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \vdots \\ \hat{f}_G \end{bmatrix}, \quad (4-10)$$

Next, we define $V = R^{-1}$, also a lower triangular matrix, and we take the following steps:

$$I = UU^T = R\hat{f}\hat{f}^T R^T = RFR^T, \quad (4-11)$$

and multiply both sides with P leads to

$$VIP^T = VRFR^T V^T, \quad (4-12)$$

which, with $V = R^{-1}$, is equal to the Choleski decomposition [47]:

$$VV^T = F. \quad (4-13)$$

Finally, we obtain the orthonormal set of basis functions as $U = R\hat{f} = V^{-1}\hat{f}$, where the inverse exists because P is square and positive definite.

Numerical Stability The inner product between two plane waves in a perfect opposite direction results in 0. However, the Choleski decomposition requires a positive-definite matrix. Therefore, the Choleski decomposition is implemented with an adjusted F matrix. We define:

$$\tilde{F} = F + \nu I, \quad (4-14)$$

where $\nu = 10^{-20}$ and I is the identity matrix. This further ensures numerical stability and prevents rounding errors due to numerical integration.

4-2-3 Soundfield Expressions

In this section, we discuss the procedure to obtain filter weights $l_q(k)$ for all loudspeakers q at wavenumber k . The following procedure is repeated for wavenumbers k frequency bins corresponding to up to 2kHz.

First, we write the soundfields of the aperture as a sum of orthonormal basis functions:

$$S^{ap}(\mathbf{x}, k) = \sum_{g=1}^G A_g U_g(\mathbf{x}, k). \quad (4-15)$$

Weights A_g are obtained with the inner product:

$$A_g = \langle H^{ap}(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle, \quad (4-16)$$

where H^{ap} is from Eq. (3-9) (for 2D) or Eq. (3-3) (for 3D). Note that this is the low resolution frequency response, depending on the window-size N . This limits the accuracy of the algorithm, as discussed in Section 3-3-3. Next, we create a coefficient vector:

$$\mathbf{a} = [A_1 \ A_2 \ \cdots \ A_G]^T, \quad (4-17)$$

and a vector containing inner products between the ATF and the normalized basis functions denoted as:

$$\mathbf{H}_{\hat{f}}^{ap} = [\langle H^{ap}, \hat{f}_1 \rangle \ \langle H^{ap}, \hat{f}_2 \rangle \ \cdots \ \langle H^{ap}, \hat{f}_G \rangle]^T. \quad (4-18)$$

Plugging in $U = R\hat{f}$ gives $\mathbf{a} = R\mathbf{H}_{\hat{f}}^{ap}$. Consequently, we have a vector \mathbf{a} containing the coefficients to describe the soundfield of the aperture as a sum of plane waves from Eq. (4-7). Note that we produce this final equation such that it depends on R . By doing this, the complexity of the evaluated integrals is limited. Instead of having to evaluate the inner-products between the orthonormal basis functions in U and H^{ap} , we only need to compute the less complex inner-products between \hat{f} and H^{ap} .

Next, we apply a similar procedure to the soundfield from the loudspeaker array. We write the soundfield from a single loudspeaker as:

$$H_q^{ls}(\mathbf{x}, k) = \sum_{g=1}^G C_g^q U_g(\mathbf{x}, k), \quad (4-19)$$

with H_q^{ls} from Eq. (3-15) and coefficients C_g^q . The soundfield of the complete array is expanded as:

$$S^{ar}(\mathbf{x}, k) = \sum_{g=1}^G B_g U_g(\mathbf{x}, k), \quad (4-20)$$

with coefficients B_g . The soundfield from the array can also be expressed the sum of the soundfields from all individual loudspeakers, multiplied by their filter weights, giving:

$$S^{ar}(\mathbf{x}, k) = \sum_{q=1}^Q l_q(k) H_q^{ls}(\mathbf{x}, k). \quad (4-21)$$

Substituting Eq. (4-20) and Eq. (4-19) in Eq. (4-21), generates coefficients B_g as:

$$B_g = \sum_{q=1}^Q l_q(k) C_g^q, \quad (4-22)$$

where the coefficients are calculated using $C_g^q = \langle H_q^{ls}(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle$. In matrix form we have $\mathbf{C} = R \mathbf{H}_{\hat{\mathbf{f}}}^{ls}$ defined as:

$$\begin{bmatrix} C_1^1 & C_1^2 & \dots & C_1^G \\ C_2^1 & C_2^2 & & \vdots \\ \vdots & & \ddots & \\ C_G^1 & \dots & \dots & C_G^G \end{bmatrix} = R \begin{bmatrix} \langle H_1^{ls}, \hat{f}_1 \rangle & \langle H_2^{ls}, \hat{f}_1 \rangle & \dots & \langle H_Q^{ls}, \hat{f}_1 \rangle \\ \langle H_1^{ls}, \hat{f}_2 \rangle & \langle H_2^{ls}, \hat{f}_2 \rangle & & \vdots \\ \vdots & & \ddots & \\ \langle H_1^{ls}, \hat{f}_G \rangle & \dots & \dots & \langle H_Q^{ls}, \hat{f}_G \rangle \end{bmatrix}. \quad (4-23)$$

Here $\mathbf{H}_{\hat{\mathbf{f}}}^{ls}$ is filled with the inner products between the basis functions and the loudspeaker ATFs. Finally, we have the matrix \mathbf{C} that contains the coefficients to describe the soundfield from the loudspeaker array as a sum of plane waves from Eq. (4-7). Again, note that we use R in this final notation to limit the complexity of the integrals.

4-2-4 Filter-Weight Calculation

The final step is to calculate the loudspeaker weights such that the sum of the soundfields is minimized. We set the control problem as $J(l_q) = S^{ap}(\mathbf{x}, k) + S^{ar}(\mathbf{x}, k)$ and $\eta = \|J(l_q)\|^2$ and minimize in least mean square sense: $\min_{l_q} \|J(l_q)\|^2$, obtaining:

$$\eta = \|S^{ap}(\mathbf{x}, k) + S^{ar}(\mathbf{x}, k)\|^2. \quad (4-24)$$

With the orthonormality property ($\langle U_i(\mathbf{x}, k), U_j(\mathbf{x}, k) \rangle = \delta_{ij}$) and by plugging in Eq. (4-15) and Eq. (4-20) we reduce Eq. (4-24) to

$$\eta = \left\| \sum_{g=1}^G A_g U_g(\mathbf{x}, k) + \sum_{g=1}^G B_g U_g(\mathbf{x}, k) \right\|^2 = \sum_g \|A_g + B_g\|^2. \quad (4-25)$$

With the knowledge that $\langle U_i, U_j \rangle = 0$, we can rewrite in matrix form, similar to [23]. We denote $\mathbf{b} = \mathbf{C}\mathbf{l}$, where $\mathbf{l} = [l_1 \ l_2 \ \dots \ l_Q]^T$ and omit k for notation purposes. Furthermore, we add the regularization term $\tau\mathbf{l}$ with $\tau > 0$, to constrain the loudspeaker effort to preventing distortion and ensure a robust solution:

$$\eta = \|\mathbf{b} + \mathbf{a}\|^2 + \|\tau\mathbf{l}\|^2 = (\mathbf{b} + \mathbf{a})^H(\mathbf{b} + \mathbf{a}) + \tau\|\mathbf{l}\|^2. \quad (4-26)$$

We rewrite Eq. (4-26) to get:

$$\begin{aligned} \eta &= (\mathbf{C}\mathbf{l} + \mathbf{a})^H(\mathbf{C}\mathbf{l} + \mathbf{a}) + \tau\|\mathbf{l}\|^2 \\ \eta &= (\mathbf{C}\mathbf{l})^H\mathbf{C}\mathbf{l} + (\mathbf{C}\mathbf{l})^H\mathbf{a} + \mathbf{a}^H\mathbf{C}\mathbf{l} + \mathbf{a}^H\mathbf{a} + \tau\|\mathbf{l}\|^2, \end{aligned} \quad (4-27)$$

and take the derivative:

$$\begin{aligned} \frac{\partial}{\partial \eta} &= 2\mathbf{C}^H\mathbf{C}\mathbf{l} + 2\mathbf{C}^H\mathbf{a} + 2\tau\|\mathbf{l}\| = 0 \\ (\mathbf{C}^H\mathbf{C} + \tau I)\mathbf{l} &= -\mathbf{C}^H\mathbf{a}, \end{aligned} \quad (4-28)$$

to obtain the final equation for the filter weights at single wavenumber k :

$$\mathbf{l} = -(\mathbf{C}^H\mathbf{C} + \tau I)^{-1}\mathbf{C}^H\mathbf{a}. \quad (4-29)$$

with $\mathbf{C} = R\mathbf{H}_{\hat{f}}^{ls}$ and $\mathbf{a} = R\mathbf{H}_{\hat{f}}^{ap}$.

4-2-5 Algorithmic Delay Compensation

The block processing with Short-time Fourier Transform (STFT) in the wave-domain algorithm induces an algorithmic delay. More specifically, the window-size N of the STFT sets the length of the delay. The influence of varying the window-size is discussed in Section 3-3-3. Algorithmic delay compensation can be done in various ways. In this study, we compare two options: reference microphone placement and signal prediction.

Reference Microphone Placement for Algorithmic Delay Compensation

As mentioned, the algorithmic delay is equal to the length of the STFT block set by the window-size N . One method to compensate for the algorithmic delay is by positioning the reference microphone at a certain distance from the aperture. This is a feasible solution for certain physical setups where the noise source is far from the aperture. However, this distance can not be too long to keep the setup practical. The time the wave travels from the microphone to the aperture is the time for which we can compensate. We have the simple equation:

$$r_{ref} = \frac{cN}{f_s}, \quad (4-30)$$

where r_{ref} is the distance in m from the reference microphone to the middle of the aperture, c is the speed of sound, N is the window-size, and f_s is the sample rate. For example, a window-size of $N = 32$ samples would lead to $r_{ref} \approx 1.4$ m, which is a feasible distance.

Signal Predictor for Algorithmic Delay Compensation

The second compensation method we explore is a signal predicting algorithm. Here, the concept is to predict, each hop m , N samples in the future, with the measured signals up to that point. We built an Autoregressive (AR) model of order p :

$$v(n) = \sum_{i=1}^p \alpha_i v(n-i), \quad (4-31)$$

that has no trends or seasonality. We employ the *Yule-Walker Equations* to calculate its coefficients (α_i), fitting the model on the last W samples. This predictor is implemented such that the predicted signal is the input of the STFT in the block processing. Expressed in equations, for each hop m , we repeat the following process. The input vector is:

$$x_m = [x_m(n-1), x_m(n-2), \dots, x_m(n-W)], \quad (4-32)$$

where W is the number of input samples. x_m is then used in the *Yule-Walker equations* to obtain $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_p]$, the AR model with p parameters. Then, we obtain the predicted signal:

$$v_m = [v_m(n), v_m(n+1), \dots, v_m(n+N-1)], \quad (4-33)$$

by deploying Eq. (4-31) over the prediction horizon N . Finally, v_m is the input of the STFT-hop m in Eq. (3-17) in the simulation model. This process is repeated each hop m .

4-3 Parameters

For the implementation of the wave-domain algorithm, two main parameters influence the performance. We discuss the influence of the window-size and the number of basis functions (G) in this section.

4-3-1 Number of Basis Functions

The soundfield basis function expansion rests on the fact that we use a finite number of basis functions to describe any soundfield within a defined region. A study determined how many basis functions were necessary for 2D and 3D regions to describe the soundfield within a reasonable error [48]. Prior implemented soundfield reproduction [23] and real-time wave-domain algorithms [24] followed the results by Kennedy et al. where a relative truncation error of 16.1% is found as limit. The size of the defined region and the wavenumber influence the number of basis functions necessary. For 2D disc-shaped spatial regions of radius r , a minimum of $G_{2D} = (\lceil 2kr \rceil + 1)$. In the case of a spherical 3D spatial region, we need at least $G_{3D} = \lceil ekr/2 + 1 \rceil^2$ basis functions.

The number of basis functions directly influences the number of calculations necessary in the algorithm, as the shape of \mathbf{C} and \mathbf{a} in Eq. (4-29) depend on it. More basis functions result in

a higher computational effort. Moreover, in our control problem, we defined the 2D control region not as a disc but as a thick arc in 2D (Eq. (4-1)). In 3D, we only use a half-spherical thick shell and not a full sphere (Eq. (4-2)). As our regions are smaller than the regions in the study by Kennedy et al., a lower number of basis functions is expected to result in similar performance. Computational wins for the 2D simulations are negligible, but reducing G in 3D calculations can make a substantial difference. In summary, we set

$$G_{2D} = (\lceil 2kr_{max} \rceil + 1), \quad (4-34)$$

and determine G_{3D} by calculating the attenuation performance for a various number of basis functions, where we use

$$G_{3D} = \lceil \beta ekr_{max}/2 + 1 \rceil^2, \quad (4-35)$$

and compare for various scaling factors of $\beta = 1/32, 1/16, 1/8, 1/4$.

4-3-2 Window-size

As we have discussed in Section 3-3-3, the window-size influences how well the low resolution aperture Acoustic Transfer Function (ATF) in Eq. (4-16) mimics the full resolution ATF. From this, we expect that a higher window-size will result in a better result. However, we have limitations on the window-size for the algorithm with algorithmic delay compensation methods. The performance is directly linked to the window-size. Due to the opposite effect of better algorithm performance and worse prediction performance, for larger window-sizes, we will see an optimal window-size for the predictor case. We will explore the results for various window-sizes to see the eventual impact of the window-size.

4-4 Summary

The main purpose of this chapter was to develop the wave-domain algorithm. We derived the necessary equations and discussed the algorithmic delay compensation. The following main points were discussed:

- To ensure global attenuation, the control region (in 3D) is defined as a half-spherical shell with finite thickness. This region makes sure we minimize sound energy in all directions from the aperture into the room.
- Any soundfield can be written as a sum of weighted basis functions. This principle formed the basis of our wave-domain algorithm.
- We presented a new method to orthonormalize a set of basis functions by applying the Choleski decomposition on a inner-product matrix of the normalized basis functions.
- A procedure in matrix notation was developed. By using matrix R , we ensure that we have to evaluate less complex integrals, reducing the computational complexity of the algorithm.

- The derived algorithm resulted in a single expression for the filter-weights. This expression calculates the filter-weights for all loudspeakers, for a single wavenumber k , and is repeated over each wavenumber that we want to control.
- We proposed two methods for algorithmic delay compensation. The first method entails positioning a reference microphone further in front of the aperture. The second method deploys a signal predictor.
- Both the number of basis functions in the basis function expansion and window-size will influence the performance of the wave-domain algorithm.

Simulation Results

In the prior chapters, we discussed the reference algorithm, explained the simulation environment, and developed the wave-domain algorithm. In this chapter, we explore the results of how the simulation environment in itself performs, as well as how the multiple-error Normalized Least Mean Squares (NLMS) compares to the proposed wave-domain algorithm. We begin by explaining the experimental setup (Section 5-1) and discussing evaluation methods (Section 5-2). Thereafter, we present the results of the simulation environment modeling using Acoustic Transfer Functions (ATFs) in Section 5-3. This is followed Section 5-4 in which we cover the results of the 3D controller simulations. We compare various scenarios, parameters, and algorithmic delay compensation methods. Finally, we give a summary table containing the main results in Section 5-5.

5-1 Experimental Setup

This section briefly covers the information explaining how we implemented the algorithms and what we used to determine the performance of both algorithms.

5-1-1 Implementation

Most of the project, such as the simulation model and animations, NLMS and WDC algorithms and fundamental error analysis was implemented in *Python 3.8* [49] within *Visual Studio Code* [50], using various packages like *SciPy*, *NumPy* and more. Additionally, the calculations of the matrices R , \mathbf{a} , \mathbf{C} , \mathbf{H}_f^{ap} and \mathbf{H}_f^{ls} for the wave-domain algorithm were done on the Delft Center for Systems and Control (DCSC) hpc-cluster using *MATLAB R2020a* [51]. This helped to speed up the heavy calculations by reducing calculation time by more 20 times.

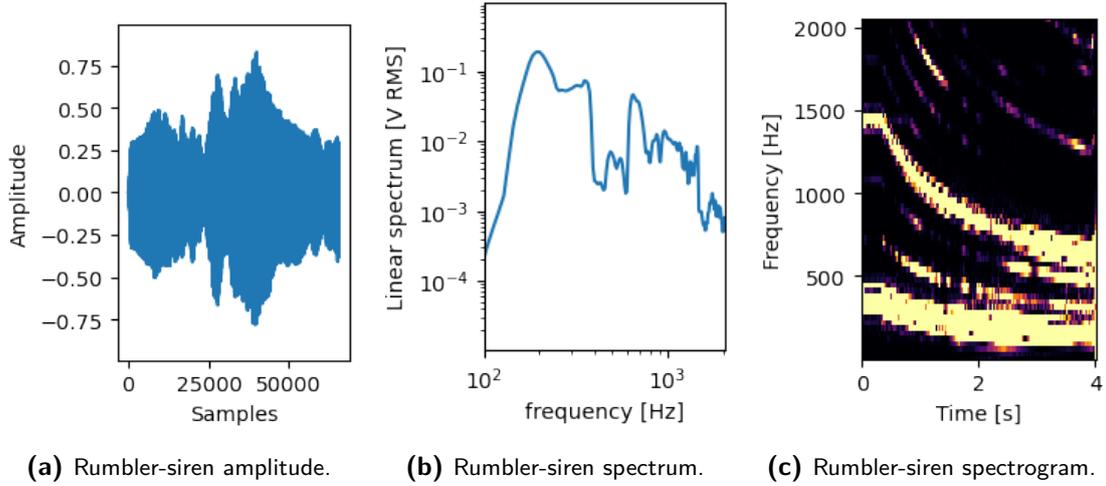


Figure 5-1: Rumbler-siren characteristics, one of the signals that is used in the simulations.

5-1-2 General Parameters

The relevant frequency range is 128 to 2048 Hz, which we denote with 'up to 2 kHz' from now on. This includes the frequency range for which the physical setup built at PennState University [13, 14] is designed. We set the sampling frequency at $f_s = 2^{14} = 16.384$ kHz. For the simulations, we assume that the density ($\rho_0 = 1.2041$ kg/m³) and temperature of the air are constant, setting a speed of sound at $c = 343$ m/s.

Moreover, the equations used in the simulations are valid in far-field circumstances. The minimal distance for a valid far-field varies per type of equation. For a monopole, far-field is defined as $kr \gg 1$ [42], where k is the wavenumber and r the distance between the monopole and evaluation point. Far-field, at the limit of 128 Hz starts at $r \gg 1/k = 1/(2\pi f/c) = 1/(2\pi \cdot 128/343) = 0.43$ m. For the aperture, we use the formula for a circular piston: $r/a \gg ka$, where a is the radius of the piston. The modeled aperture contains four squares, so we conservatively take a as the distance between the middle and the corner of a square, giving $a = 0.153$ m. Consequently, we need $r \gg ka^2 = 2\pi fa^2/c = 2\pi \cdot 2048 \cdot 0.153^2/343 = 0.88$ m. As such, we position the control region in the far-field with $r_{min} = 0.95$ m and $r_{max} = 1$ m.

To measure performance, we evenly distributed 114 evaluation microphones on a hemisphere at $r = 1$ m inside the room, using a spherical Fibonacci mapping [52].

5-1-3 Primary Noise Signals

The various algorithms were compared and tested using three types of noise: white noise, airplane noise, and a rumbler-siren signal. Each signal was four seconds long. We used the white noise to explore the performance of the optimal wave-domain and multiple-error NLMS algorithms. The rumbler-siren signal was used to explore the algorithmic delay compensation methods as it contains predictable signals. Additionally, the airplane noise was used as it is a real-world noise with emphasis on low frequency. An illustration of the rumbler-siren signal over time is shown in Figure 5-1a. A rumbler-siren is well suited for the wave-domain algorithm with a predictor. It covers both a higher-pitched (0.5-1.5 kHz) and a lower-pitched

(0.1-0.5 kHz) siren that sweeps, shown in Figure 5-1b and Figure 5-1c. The signal's non-stationary character resembles the behavior of real-world sounds. Moreover, a signal that is (partially) deterministic is necessary to illustrate the performance of a predictor-based algorithm. With an auto-regressive model, a random white noise signal can not be predicted, while a pure sinusoidal signal is perfectly predictable. The usage of the rumbler-siren signal enables us to showcase the trade-off between the algorithm and prediction. This is discussed in more detail in Chapter 6.

5-2 Evaluation Methods

To evaluate the attenuation performance, we deploy two procedures. These are described in this section.

5-2-1 Segmental SNR

Performance was evaluated on the boundary of the control regions in all directions. We define the segmental SNR in dB, summed over all evaluation microphones e as:

$$SEG_f(k, m) = 10 \log_{10} \frac{\sum_e^E |d_e(k, m)|^2}{\sum_e^E |d_e(k, m) + y_e(k, m)|^2}, \quad (5-1)$$

with $d_e(k, m)$ the Short-time Fourier Transform (STFT) block primary noise from the aperture and $y_e(k, m)$ the STFT block of the loudspeaker array, at each evaluation microphone e . The mean of $SEG_f(k, m)$ up to 2 kHz gives the performance over time as:

$$SEG_t(m) = \frac{\sum_k^K SEG_f(k, m)}{K}, \quad (5-2)$$

where K denotes the frequency bin number at 2 kHz. Finally, we obtain a single value for the Signal-to-Noise Ratio:

$$SNR = \frac{\sum_m^M SEG_t(m)}{M}, \quad (5-3)$$

where M is the total number of hops of the signal. The segmental SNR was calculated over signal segments with an 8 ms STFT with 50% overlap ($H = N/2$), and a confidence interval of 99% was presented in some cases. The relatively short 8 ms window-size for evaluation was chosen because of the nature of the used primary noise signal. Performance over a non-stationary noise changes faster over time. Therefore, shorter segments indicate more accurately what the performance over time is, with the downside of a lower frequency resolution.

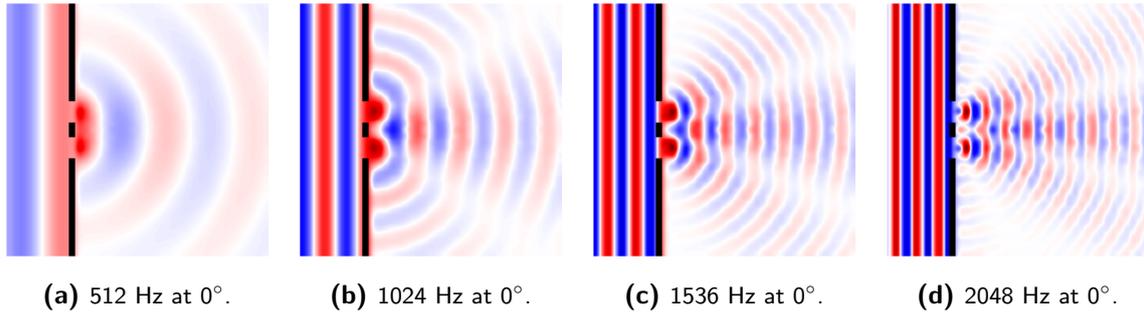


Figure 5-2: Aperture soundfields for various frequencies at 0° incident angle.

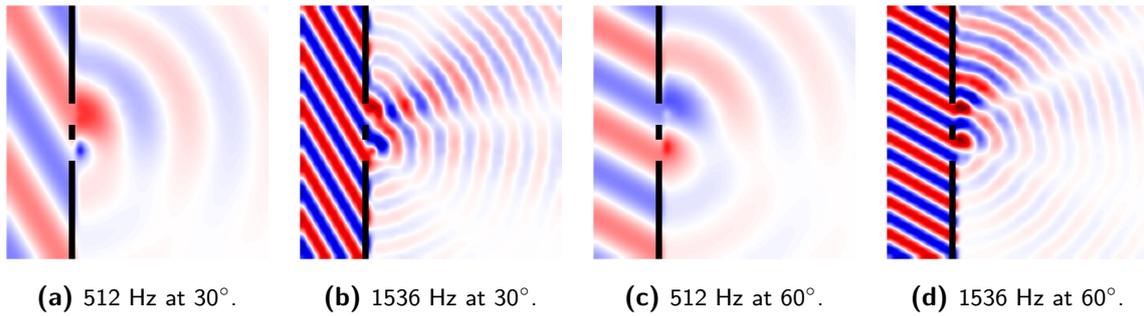


Figure 5-3: Aperture soundfields for various frequencies at 30° and 60° incident angle.

5-2-2 Animations

Next to the quantitative Segmental SNR values, a qualitative measure in the form of animations was implemented. These animations were mostly consulted during the development phase of the thesis work. They indicated what the calculated soundfields look like and how the loudspeaker soundfield aligns (or not) with the soundfield of the aperture. The measure indicates the emergence of sidelobes due to diffraction. Animations contained three parts: (1) a soundfield generated by the primary noise propagating through the aperture, (2) a soundfield from the loudspeaker array, and (3) the sum of both soundfields, resulting in a controlled soundfield.

5-3 Soundfield Modeling Results

We briefly cover the modeling of the aperture behavior in this section. In Chapter 3, we derived the aperture Acoustic Transfer Function (ATF) based on a sum of vibrating plates. We used $\hat{P} = 4$, to describe the four squares of the aperture with crossbar. The surface velocity gain parameter was set at $\dot{w}_0 = 0.0025$. For the loudspeaker ATF, we used monopole radius $a = 0.022225$ m and surface velocity gain $u_0 = 0.0025$.

In Figure 5-2 and Figure 5-3, soundfields are shown for pure sinusoidal signals at various frequencies and incident angles. The following is noticed. A higher frequency resulted in more diffraction and a more directed beam in the direction of the incident angle. Furthermore, the soundfield degraded in strength when moving away from the aperture. When the incident

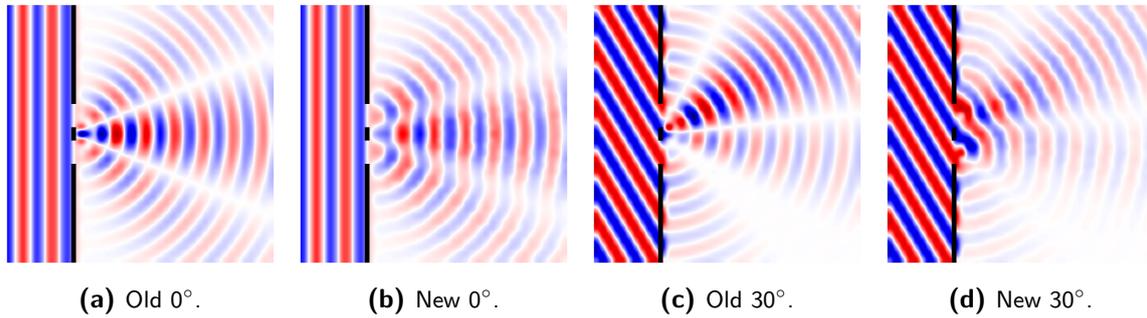


Figure 5-4: Comparison between the soundfields generated with the old and new Aperture Acoustic Transfer Function. The soundfields by the old version show clearly that the sound originates from the aperture. Because of that, there is a clear line between the side and main lobes visible. The new soundfield shows that the waves originate from the aperture sections and has more smeared out overlap between the main and sidelobes. The latter is a more realistic representation.

angle was increased, the sidelobes moved along with the main lobe. Lastly, some form of misalignment is visible in the soundfields of the plane wave and the aperture.

Let us compare the results of the new Aperture ATF with the old method, developed by [13]. The soundfields, generated with the original equation are shown in Figure 5-4a and Figure 5-4c. Sharp lines are visible between the lobes. Furthermore, the soundfield clearly originates from the middle of the aperture. In contrast, the new method (Figure 5-4b and Figure 5-4d) generates soundfields that originate from the individual aperture parts, clearly visible in Figure 5-4b. Consequently, this causes the sidelobes to be smeared out over the region resulting in less sharp lines. Hence, we create a more realistic aperture soundfield with the new aperture ATF.

5-4 3D Algorithm Results

This section covers the results of the 3D simulations. The 2D results were very similar and are, therefore, left out in this chapter and presented in Appendix C. We present the results in four sections. First, we briefly show graphical illustrations of a cancelled soundfield. Secondly, we investigate how the performance of the wave-domain algorithm depends on the window-size. Thirdly, the optimal wave-domain algorithm is compared to the multiple-error NLMS algorithm. Finally, we look into two methods of algorithmic delay compensation, by microphone placement (Wave-Domain Algorithm with Reference Microphone Clacement (WDC-M)) and prediction (Wave-Domain Algorithm with Predictor (WDC-P)). Hence, in total, four types of algorithms were implemented. After a parameter optimization, the following parameters were found for each algorithm.

NLMS For the NLMS algorithm we set the learning rate, $\tilde{\mu} = 1/2$, used an impulse response length of $J = 128$ samples and an filter-order of $F = 64$ samples. The algorithm used $R = 1$ reference microphone, $L_{sparse} = 21$ and $L_{grid} = 49$ loudspeaker for the sparse and grid array, respectively, and $E = 128$ error microphones. The impulse response of the aperture

is set to the impulse response of the ATF in Eq. (3-3): $P_e = \text{I-FFT}(H^{ap}(k))$. Similarly, we set the impulse response of each loudspeaker to the transform of the ATF from Eq. (3-15): $S_{te} = \text{I-FFT}(H_q^{ls}(k))$.

WDC-O We limited the control effort as follows for the Optimal Wave-Domain Algorithm (WDC-O) as follows. We use $\tau = 10^{-7}$ for the 3D sparse and $\tau = 10^{-12}$ for the 3D grid array. Furthermore, the window-size was set at the largest option to obtain the best performance, so $N=125$ ms.

WDC-M The WDC-M represents the case of microphone placement at $r_{ref} = 1.4$ m, a reasonable distance for microphone placement in far-field sound cancellation. This distance means that a fixed window-size of $N \approx 3.9$ ms is selected for this algorithm. The values for τ are similar to WDC-O.

WDC-P Lastly, the WDC-P had a variable window-size, for which we optimize. The results for this are discussed in a later section. The limitation to control effort (τ) was set equal to the case for the WDC-O. Finally, we used an auto-regressive model of order $p = 16$ and an input-signal length of $W = 256$ samples.

5-4-1 Graphical Illustrations

To obtain a feeling of the generated results, we animated the propagation of a pure sine wave primary noise with an incident angle of 30° and show the attenuation. Results are shown in Figure 5-5. In Figure 5-5a, we see clear diffraction occurring as the wavelength ($\lambda = c/f = 343/1000 = 0.343$ m) is smaller than the aperture ($L_y = 0.5$ m). Similarly, the created soundfield by the loudspeaker array, shown in Figure 5-5b, resembles the inverse of the aperture soundfield. The controlled soundfield in Figure 5-5c shows the result of the attenuation. It clearly shows the obtained attenuation.

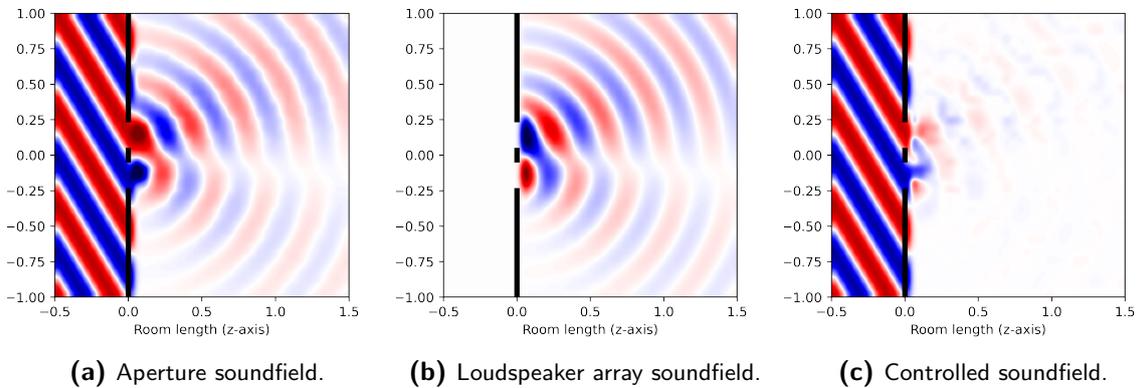


Figure 5-5: 2D animation for a pure 1kHz sine signal controlled with the WDC-O and a grid array. The algorithm is able to cancel the soundfield as we see less amplitude in graph Figure 5-5c.

5-4-2 Number of Basis Functions

The wave-domain algorithms use a finite number of basis functions to express the expansion. To limit computational load, we used a low number of basis functions that described the soundfields without a large truncation error. We set $G_{3D} = (\beta ekr_{max}/2 + 1)^2$, and compare for $\beta = 1/32, 1/16, 1/8, 1/4$. Results are shown in Figure 5-6. Clearly, results were close. This indicates that the scaling we did was reasonable. Results with $\beta = 1/4$, $\beta = 1/8$ and $\beta = 1/16$ were very similar, while $\beta = 1/32$ showed less performance at mid-frequency range. Therefore, we chose to work with $\beta = 1/16$.

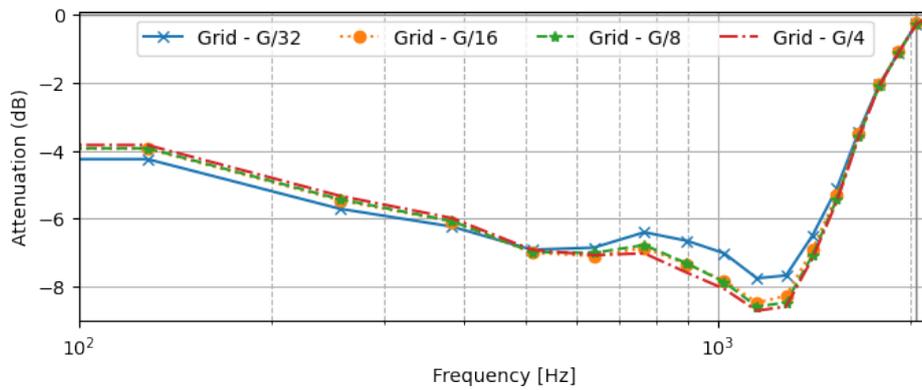


Figure 5-6: 3D results for various scaling of G , using white noise, 0° incident angle, the grid array and a window-size of $N=3.91$ ms. The attenuation performance is similar for all scaling factors. a smaller scaling factor is beneficial for the computational load. Also, $\beta = 1/32$ showed less performance at the mid-frequency range. Hence, we chose $\beta = 1/16$.

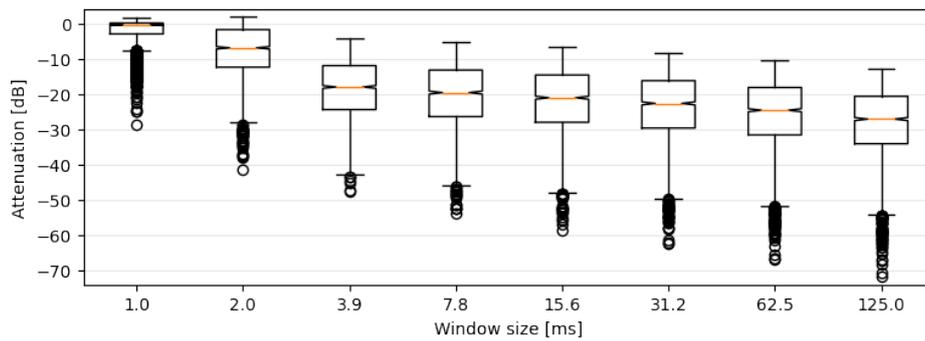


Figure 5-7: 3D Fundamental error analysis, with the Signal-to-Noise Ratio (SNR) over the evaluation microphones. We used a flat frequency weighting $s(k)$, and evaluated for 20 random incident angle scenarios, leading to 20 different $\mathbf{h}(k)$. We can conclude that increasing the window-size leads to less fundamental error in the aperture ATF. There is major difference between $N = 2.0$ ms and $N = 3.9$ ms, indicating that the ATF impulse response for a window-size smaller than $N = 3.9$ ms is cut off drastically, leading to major errors.

5-4-3 Wave-domain algorithm with various Window-sizes

First, we look into the fundamental limiting results, with the error analysis algorithm developed in Section 3-3-3. The distribution of SNR over the evaluation microphones, for a flat frequency weighting, $s(k) = 1 \forall k$, and 20 random incident angle scenarios (depicted in $\mathbf{h}(k)$) is shown in Figure 5-7.

Clearly, a larger window-size resulted in better performance, where the median attenuation was almost -30 dB at $N = 125.0$ ms. Noteworthy is the spread of the attenuation. It varied between +3 and -30 dB for $N = 1.0$ ms and even between -12 and -73 dB for $N = 125.0$ ms. As we have seen, the window-size determines the length of the impulse response. The spread in results could be caused by following two things or a combination of the two. First, the evaluation point in the room, either close to the wall or in the middle, could have largely influenced the attenuation. The relative distance between the borders of the aperture towards the evaluation point is much larger if the evaluation point is close to the wall. This leads to a longer impulse response. Secondly, the incident angle of the primary noise, could have caused a large impact on the fundamental error, as this has a similar impact on the length of the impulse response. We see a major jump in results between $N = 2.0$ ms and $N = 3.9$ ms, indicating that $N = 3.9$ ms is long enough to include the total impulse response length. We now follow with results for varying window-sizes to see the impact in the simulations. We look into incident angle differences in Section 5-4-4 later in this chapter.

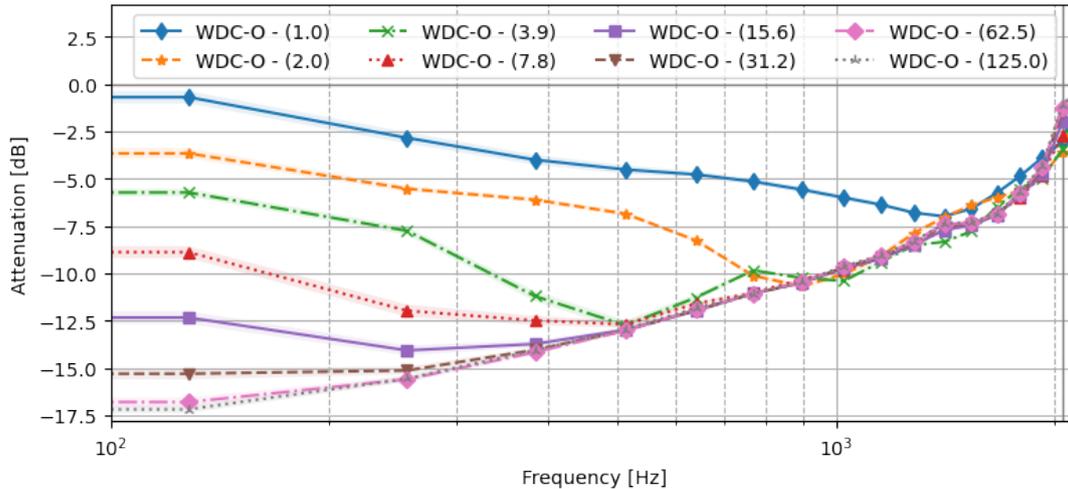


Figure 5-8: 3D Grid array SEG_f results with narrow 99% confidence interval for various window-sizes. Increasing the window-size increases performance. The major difference is visible at the lower frequencies.

The influence of the window-size on performance in simulation plotted over frequency is shown in Figure 5-8 with the 3D grid array case and the WDC-O for white noise and 0° incident angle. In general, an increase in window-size resulted in more attenuation, but even very small window-sizes generated attenuation. There is a clear trend visible: the smaller the window-size, the less performance at low frequencies. In contrast, from 1.5 kHz onwards, the window-size did not influence attenuation anymore.

In Figure 5-9, the frequency performance for various window-sizes is shown for the sparse

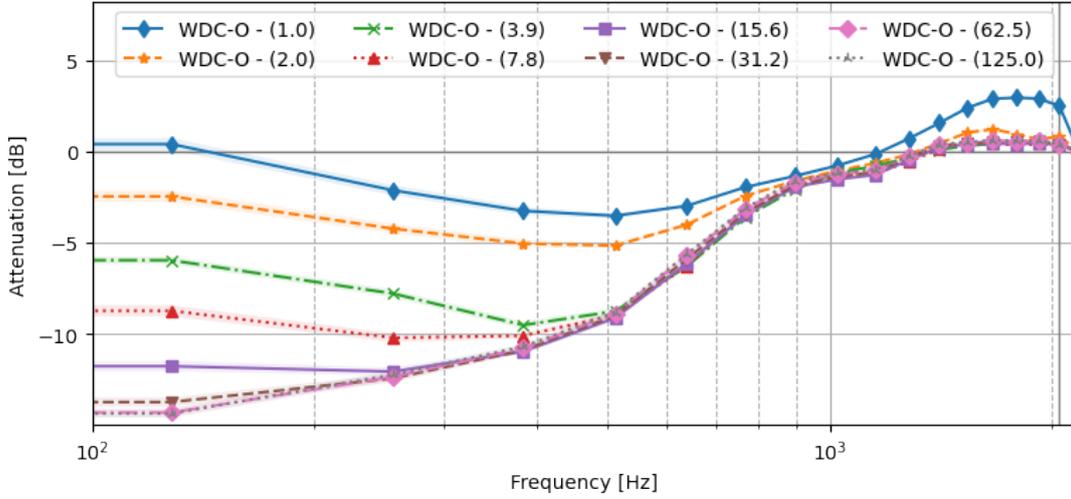


Figure 5-9: 3D Sparse array SEG_f results with narrow 99% confidence interval for various window-sizes. A larger window-size resulted in better performance. The major difference is visible at low frequencies.

array. We see similar trends as with the grid array. However, from 1.2 kHz upwards, there is no attenuation (and even amplification for $N = 1.0$ ms). Again, by increasing the window-size, increased performance was obtained. In addition, we see a two-split between $N \leq 2$ ms compared to $N \geq 3.9$ ms, where the middle frequency performance was reasonable for the 3.9 ms window-size algorithm. For larger window-sizes, performance degraded with increasing frequency. For smaller window-sizes, the results become more valley-shaped.

From this, we can conclude that using a lower window-size mostly affects the performance at low frequencies. Moreover, to prevent amplification, we need at least a window-size of 2.0 ms. Because we want to find the optimal wave-domain performance, we continue with an optimal wave-domain algorithm with $N = 125$ ms.

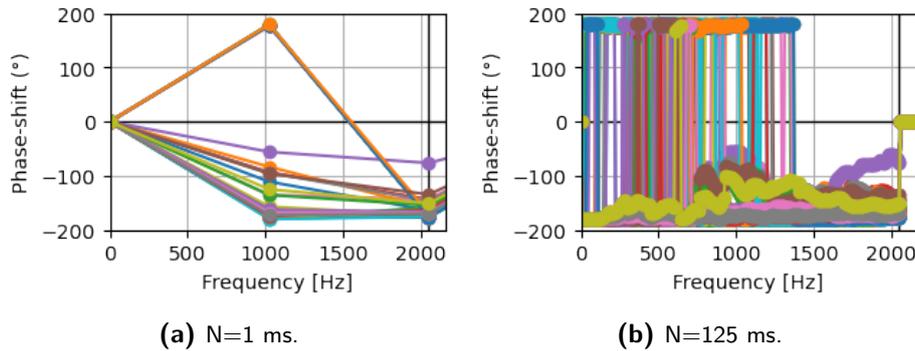
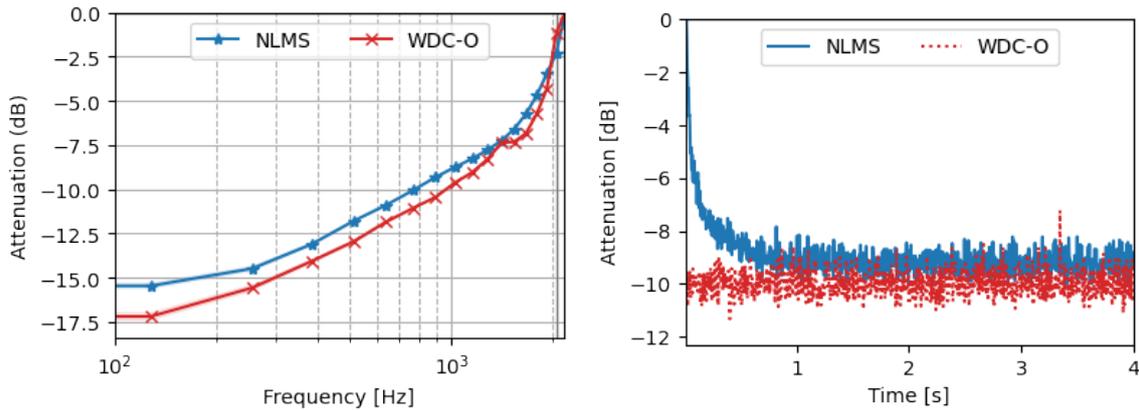


Figure 5-10: 3D phase-shift values in $W_q(k)$, for all speakers for 0° incident angle. As expected, they are concentrated around 180° , which leads to large phase-shifts in the block processing.

As we have seen, a smaller window-size mostly affected the performance at lower frequencies. We expect that this is caused by the block-processing approach, alongside the low resolution issue indicated by the fundamental error limitation. The time-delay wrapping, discussed in

Section 3-3-2, does not only occur when transforming a signal with an ATF. A similar issue could arise when transforming the signal with filter-weights $W_q k$. Therefore, we give some insights into the phase-shifts by the filter-weights in this section. Figure 5-10 shows the phase-shift of the filter weight $W_q(k)$ for all speakers for 0° incident angle. As we can see, the phase-shifts slightly differ per speaker but are mostly concentrated around -180° . This can be expected if one tries to generate a loudspeaker soundfield which is the opposite of the aperture soundfield. Therefore, these phase-shifts occur both for the lower-limit window-size of 1 ms and the upper-limit window-size of 125 ms.



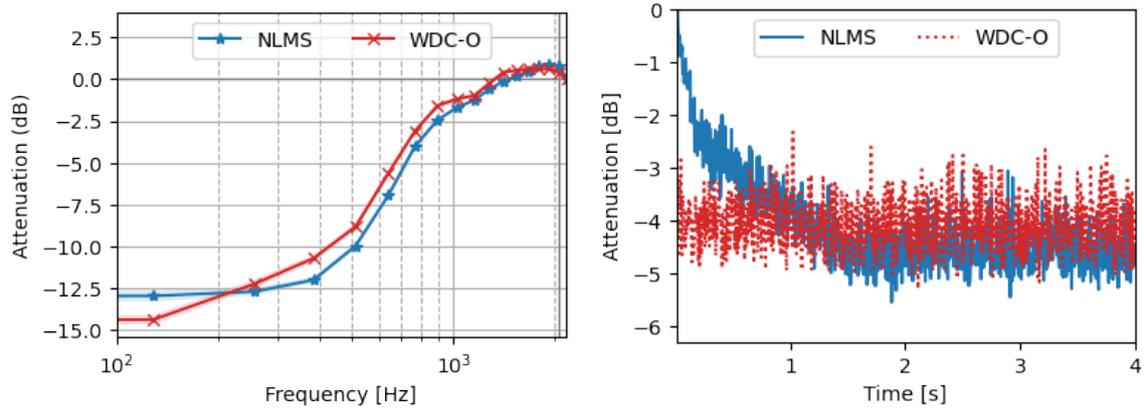
(a) 3D grid SEG_f with narrow 99% confidence interval (b) 3D grid SEG_t over time where NLMS shows slow convergence while WDC-O does not.

Figure 5-11: 3D Grid results for the NLMS and WDC-O algorithms. Performance is comparable. However, WDC-O outperforms in terms of convergence time, as it converges instantaneously. In contrast, the NLMS needs 1 s to converge.

5-4-4 Multiple-error NLMS algorithm versus Optimal Wave-domain Algorithm

We have derived the optimal wave-domain algorithm and compare its white noise cancelling performance to the NLMS algorithm in this section. Let us focus on the grid array first. Figure 5-11a shows the attenuation over frequency. Clearly, the WDC-O outperformed the NLMS and showed major attenuation almost up to 2 kHz. The NLMS algorithm had similar performance, albeit performing slightly worse for all frequencies except at 2 kHz. The narrow confidence intervals indicate that we had steady performance over time. Another difference between the NLMS and wave-domain based approach occurs when performance is shown over time. This is visible in Figure 5-11b. The NLMS algorithm showed a long convergence time (1.0 s), while the wave-domain algorithm did not. In addition, in Figure 5-11b we see that the performance of both algorithms fluctuates slightly over time, but this is not significant (within +1 and -1 dB from the mean).

Additionally, we tested the sparse array. Results are shown in Figure 5-12. The NLMS performed slightly better than the wave-domain algorithm. Performance is more than -5 dB attenuation from 0 to 0.7 kHz. However, from 0.7 kHz onwards, performance degraded for both algorithms. Above 1.4 kHz, no performance was obtained with either algorithm, indicating that the sparse array only performs well at lower frequencies. The sparse array



(a) 3D sparse SEG_f with narrow 99% confidence interval per frequency bin. (b) 3D sparse SEG_t over time where NLMS shows slow convergence while WDC-O does not.

Figure 5-12: 3D Sparse results for the NLMS and WDC-O algorithms. Performance is comparable. However, WDC-O outperforms in terms of convergence time, as it converges instantaneously. In contrast, the NLMS needs almost 2 s to converge.

performance over time, in Figure 5-12b shows similar trends as the grid array. Again, a long convergence time (almost 2.0 s) was shown by the NLMS algorithm, while the wave-domain algorithm converged instantaneously. If we compare the performance of the grid and sparse arrays (Figure 5-11a and Figure 5-12a), we see a significant difference occurring from around 0.7 kHz upwards. This indicates that the sparse array setup was only partially able to recreate the soundfield of the aperture while the grid array did a much better job.

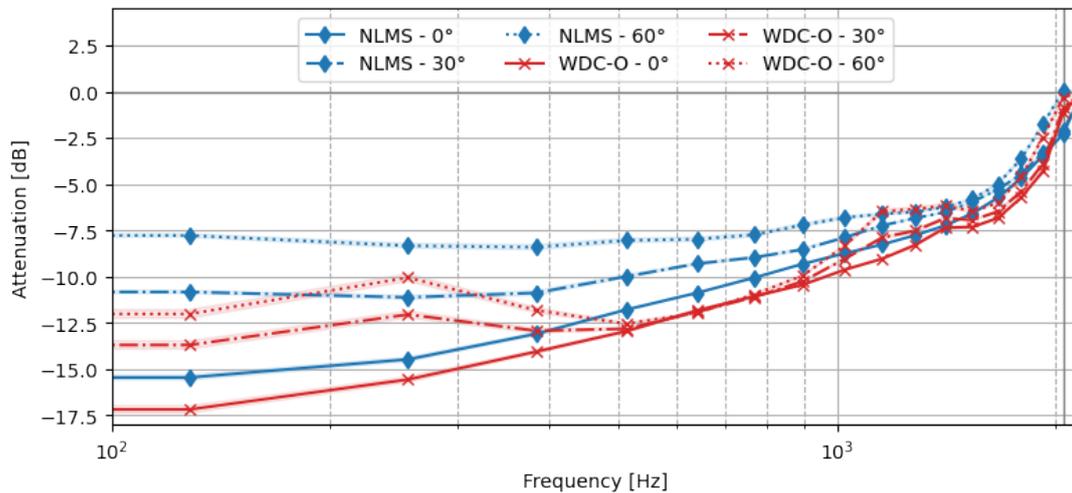


Figure 5-13: 3D grid array SEG_f results with narrow 99% confidence interval for WDC-O and NLMS at three incident angles. In general, a larger incident angle led to less attenuation. The performance of the NLMS is more sensitive to the incident angle compared to the WDC-O results.

Next, we investigate the performance when the noise comes from three incident angles: 0° , 30° and 60° . Again, we use white noise as the signal, because of its flat frequency spectrum. The grid array performance is shown in Figure 5-13. Clearly, attenuation was obtained by

both the NLMS algorithm and WDC algorithm. However, there are some differences. For both algorithms, it holds that attenuation degrades when the incident angle is increased, leading to the best performance at 0° , and lowest attenuation at 60° . In addition, the NLMS algorithm seems to have more difficulty to perform with an increased angle compared to the optimal wave-domain algorithm. Where the performance of WDC-O degraded mostly at low frequencies, the NLMS showed less performance for larger incident angles up to 1.5 kHz.

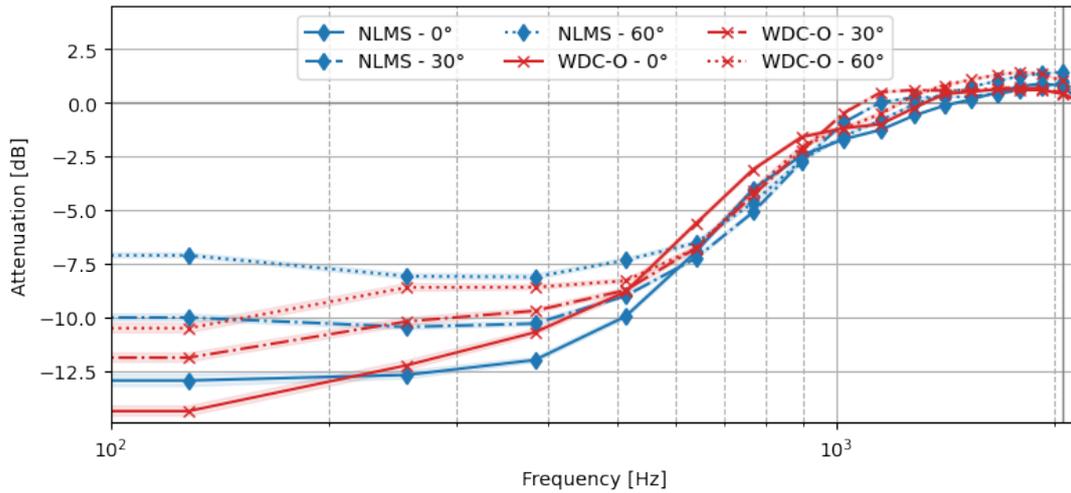


Figure 5-14: 3D sparse array SEG_f results with narrow 99% confidence interval for WDC-O and NLMS at three incident angles. In general, a larger incident angle led to less attenuation. The performance of the NLMS is more sensitive to the incident angle compared to the WDC-O results. The major difference is visible at lower frequencies.

The sparse array obtained similar results. It attenuated noise up to approximately 1 kHz for all three incident angles. Again, we see that the NLMS algorithm was more sensitive to the change in angle, while the results of the WDC-O were closer to one another.

To conclude, the NLMS algorithm performed slightly better with the sparse array, while the WDC-O came out on top for the grid array. Moreover, the wave-domain algorithm does not need time to convergence, while the NLMS algorithm needs significant time (> 1.0 s) to obtain proper filter weights. In addition, we saw that the performance of the NLMS algorithm degraded more for larger incident angles, while the WDC method maintained less diverge results. While larger incident angles led to less attenuation, both algorithms controlled white noise with 30° and 60° incident angles properly.

5-4-5 Grid and Sparse Array Comparison

As we have seen in the prior sections, the results for the grid and sparse array are not comparable. In fact, the grid array outperforms the sparse array, especially for higher frequencies. This holds for both the NLMS algorithm and WDC-O. In Figure 5-15, attenuation performance of the two arrays is shown. Here, we simulated a white noise signal with 0° angle and used the WDC-O. While the sparse array's performance is zero from approximately 1.5 kHz, the grid array achieved proper attenuation up to almost 2 kHz.

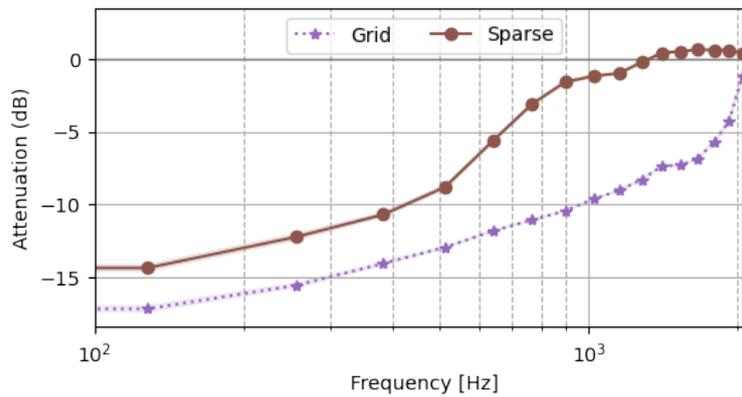
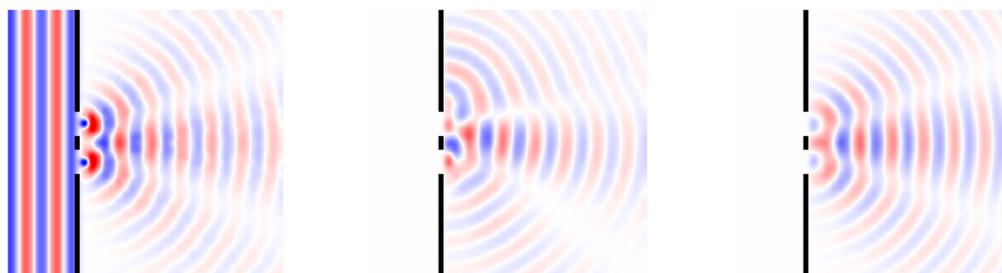


Figure 5-15: 3D grid and sparse array SEG_f results with narrow 99% confidence interval for WDC-O, white noise, 0° incident angle. Performance of the grid array was better for all frequencies. The grid array was able to cancel noise up to 2 kHz, while the sparse array only performed up to 1.2 kHz.

Apparently, the sparse array was not able to recreate the soundfield of the aperture. The aperture soundfield for a sine wave at 1.5 kHz is shown in Figure 5-16a. Diffraction, leading to a mainlobe and two sidelobes is visible. The wave-domain algorithm with sparse array generated the soundfield in Figure 5-16b, where the main- and sidelobes are not corresponding to the aperture soundfield. Also, we see sound energy close to the wall, while the aperture soundfield does not have much sound energy going there due to less diffraction at higher frequencies. In contrast, the soundfield generated by the grid array, visible in Figure 5-16c, does resemble the soundfield of the aperture, that would lead to a performance of approximately -7.5 dB, given in Figure 5-15.



(a) Aperture soundfield.

(b) Sparse array soundfield.

(c) Grid array soundfield.

Figure 5-16: Animation of aperture, grid and sparse array soundfields for 1.5 kHz sine wave. The sparse array soundfield does not correspond to the soundfield of the aperture array, while the grid array soundfield is very similar. This shows the major difference in performance between the two arrays.

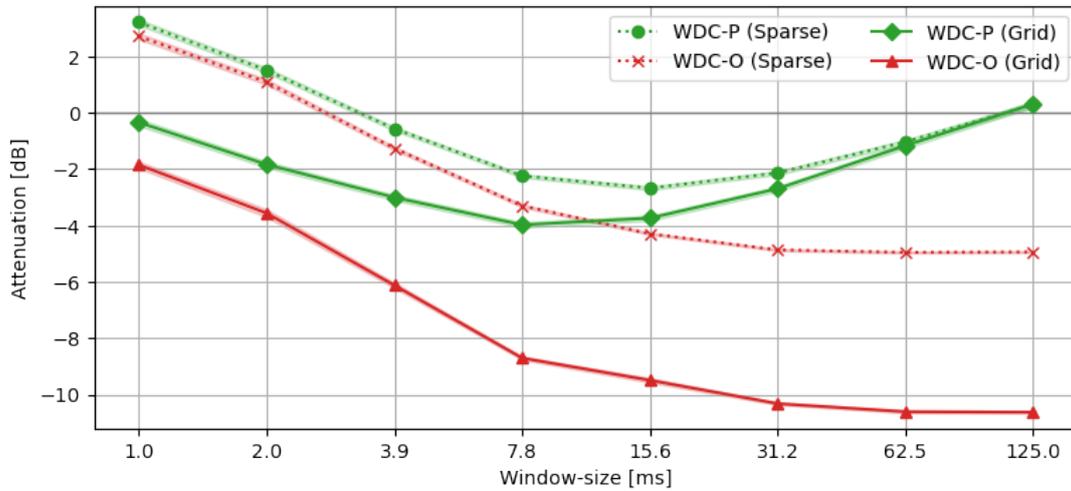


Figure 5-17: 3D SNR value with 99% confidence interval for both algorithms cancelling rumbler-siren noise to find the optimal window-size. The optimal point for the WDC-P (Sparse) lies at 15.6 ms, while the optimal window-size for the grid array is 7.8 ms.

5-4-6 Algorithmic Delay Compensation

The wave-domain algorithm showed promising results. However, we need to compensate for the algorithmic delay. In this section, we cover the results obtained for algorithmic delay compensation with microphone placement (WDC-M) and predictor (WDC-P). We fix the simulations to the 0° incident angle scenario and use rumbler-siren noise. Furthermore, the WDC-O and NLMS performances are included for comparison.

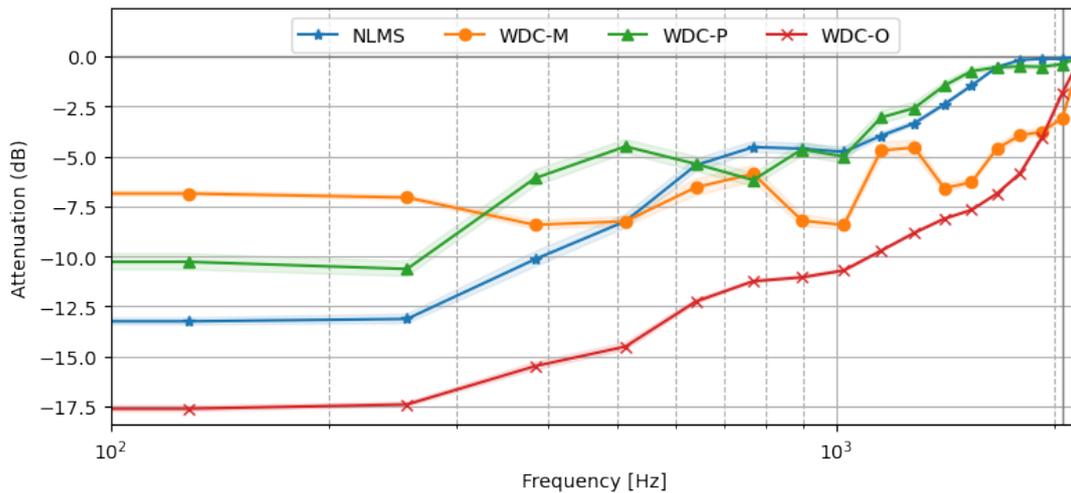


Figure 5-18: 3D grid array SEG_f results with narrow 99% confidence interval, for all algorithms, 0° angle and rumbler-siren noise. The best performing algorithm is the WDC-O. Only the WDC-M and WDC-O are able to attenuation noise over the full frequency range up to 2 kHz.

For the WDC-P algorithm, the optimal window-size was found by comparing performance for various window-sizes. Figure 5-17 indicates that increasing window-size improves performance

for the optimal wave-domain algorithm. However, when a predictor was used to compensate for the algorithmic delay, a too large window-size led to lower attenuation again. We expected this, as the prediction horizon is equal to the window-size. Increasing the prediction horizon generally leads to lower performance of the predictor. In contrast, increasing the window-size does lead to better algorithm performance. Hence, an optimal point can be found. It is important to note that the predictor performance highly depends on the type of signal. In this simulation, the rumbler-siren is reasonably predictable. However, a signal that is less deterministic, could give an optimal point located at a smaller window-size, because predictability is lower. This was the case with the airplane and white noise signals, for which the WDC-P algorithm has no performance at all. Similarly, if the signal is better predictable, the optimal point shifts towards the right as the algorithm performs better with a larger window-size. The optimal point for rumbler-siren noise with the sparse array was located at $N = 15.6$ ms, and with the grid array, $N = 7.8$ ms gave the best performance. We continue from here with the grid array only.

Figure 5-18 shows the attenuation over frequency for all four algorithms. As expected, the WDC-O outperformed all the others and showed major attenuation up to 1.8 kHz. The NLMS and WDC-P had performance up to 1.5 kHz. While the WDC-M did perform up to 2 kHz, it was more inconsistent above 1.0 kHz. As indicated in the prior section, the smaller window-size of the WDC-M shows a relatively flat spectrum while the other algorithms all show an upward (lower attenuation) trend over frequency. We can conclude that algorithmic delay compensation by microphone placement leads to better attenuation for higher frequencies. However, the predictor predicted lower frequencies well, giving it a slight advantage at frequencies below 0.3 kHz. Another observation is the relatively larger confidence interval for the WDC-P, indicating that its performance was not as consistent over time as the performance of the other algorithms.

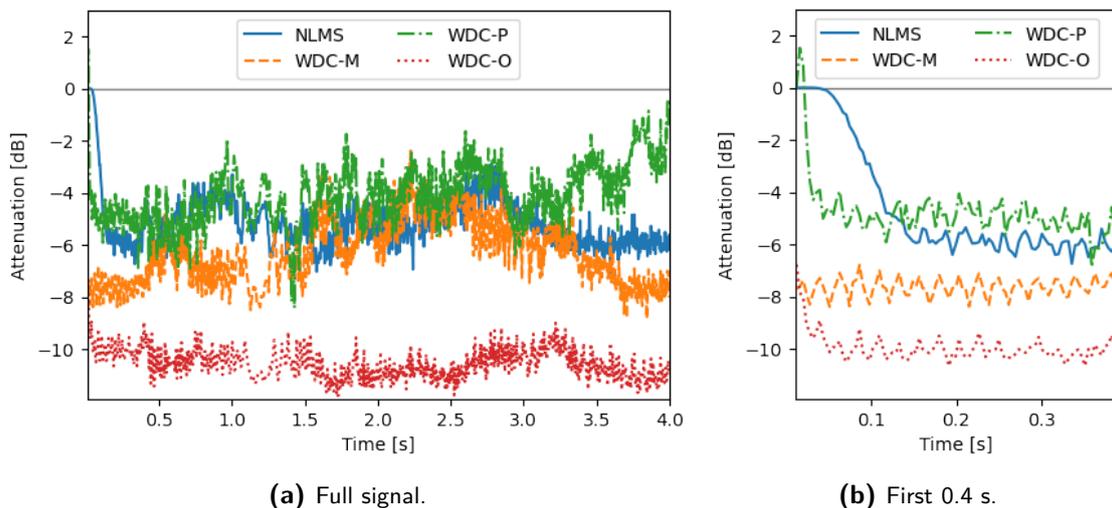


Figure 5-19: 3D grid SEG_t results for all algorithms, 0° incident angle and rumbler-siren noise. The WDC-O was clearly the best performing algorithm as the attenuation was kept steady at -10 dB. Attenuation of the other methods fluctuated over time, depending on the signal content. Moreover, the long convergence time of the NLMS algorithm is visible in Figure 5-19b, while the other had instantaneously or very fast convergence.

Table 5-1: 3D Results for rumbler-siren noise at 0° angle, in decibels. Given over four frequency ranges, for all algorithms. In general, less attenuation was obtained for higher frequencies. Results are similar between the sparse and grid array, but the grid array outperformed the sparse for all algorithms at higher frequencies. WDC-O performed best. Also, the WDC-M and WDC-O with grid array were able to obtain decent performance above 1.5 kHz while the other algorithms did not.

<i>Array type</i>	Grid					Sparse					
	<i>kHz</i>	0-0.5	0.5-1	1-1.5	1.5-2	<i>Total</i>	0-0.5	0.5-1	1-1.5	1.5-2	<i>Total</i>
NLMS		-11.5	-4.8	-2.8	-0.2	-5.2	-10.3	-2.0	-0.2	0.0	-3.5
WDC-M		-7.5	-7.3	-5.5	-3.9	-6.1	-7.1	-2.0	3.1	2.5	-1.3
WDC-P		-7.3	-5.3	-2.0	-0.5	-4.0	-6.6	-1.8	0.5	0.2	-2.2
WDC-O		-16.5	-11.3	-8.6	-4.7	-10.6	-13.3	-4.2	-0.6	0.4	-4.9

The attenuation of the rumbler-siren noise over time by all algorithms is shown in Figure 5-19. Again, the NLMS algorithm showed an convergence time (0.15 s), while the wave-domain algorithms did not, as indicated by Figure 5-19b. One exception is the WDC-P, where some signal data is necessary for the predictor to start predicting accurately. In addition, in Figure 5-19a we see that the performance of all algorithms fluctuated over time. The changing character of the rumbler-siren signal and low convergence speed of the NLMS limit the performance of this algorithm for such signals. In contrast, the optimal wave-domain algorithm performed consistently over time. Both WDC-M and WDC-P showed fluctuations. The prediction of the signal changes per segment, leading to diverging results. Similarly, the WDC-M method showed that this change in frequency content leads to varying results. We can conclude that the microphone placement outperformed the predictor system for the rumbler-siren signal at any time, except for a peak around 1.5 ms. A summary of the results for various frequency bands is shown in Table 5-1.

5-5 Summary Table

The 3D results for all types of noise and both arrays are summarized in Table 5-2. In general, the grid array outperformed the sparse array for all algorithms. Finally, in all simulations, the WDC-O outperformed the NLMS algorithms, especially for real-world sounds (rumbler-siren and airplane noise). The performance of the optimal wave-domain algorithm was consistent for all types of noise. In terms of algorithmic delay compensation, a predictor only worked when the primary noise signal had deterministic components. The microphone placement worked in all cases but did give up performance at low frequencies compared to the optimal algorithm.

Table 5-2: 3D results for all algorithms, for both grid (G) and sparse (S) arrays. Attenuation is given in decibel (dB). WDC-P did not attenuate airplane and white noise as it could not predict the signal. The grid array outperformed the sparse array in any scenario. The WDC-O algorithm is the clear winner. NLMS and WDC-M performed fairly similar.

<i>Noise type</i>	Rumbler siren		Airplane Noise		White Noise		
	<i>Array type</i>	G	S	G	S	G	S
NLMS		-5.2	-3.5	-7.6	-4.5	-9.2	-4.3
WDC-M		-6.1	-1.3	-6.6	-2.0	-8.1	-2.8
WDC-P		-4.0	-2.2	-	-	-	-
WDC-O		-10.6	-4.9	-10.6	-4.7	-10.0	-4.1

Chapter 6

Discussion

The main objective of this study was to develop an algorithm for filter-weight calculation in the wave-domain that actively cancels undesired sounds propagating through an aperture and assess its performance. As we have seen in the results in Chapter 5, we have successfully developed and implemented the algorithm and obtained substantial attenuation performance in the simulations. In this chapter, we will discuss the results in more detail and explore more practical implications. We focus on three main parts. First, in Section 6-1, we cover the results of the simulation environment. We indicate limitations as well as give future work suggestions to improve the simulation environment and soundfield calculations. Secondly, in Section 6-2 we discuss the performance of the algorithms and try to understand why the obtained performance occurred. We mention limitations and suggest future work towards a fully implementable real-time wave-domain algorithm. Finally, in Section 6-3 we briefly cover the algorithmic delay compensation topic, its limitations and present ideas for further research.

6-1 Simulation Environment and Acoustic Transfer Functions

The first objective was to develop a simulation environment in which we could test Active Noise Control (ANC) systems for apertures. The illustrations in Figure 5-2, Figure 5-3 (and more extensively in Appendix C) show that this has been achieved. The model allowed us to test the four types of algorithms. The soundfields generated by the new aperture Acoustic Transfer Function (ATF) (Eq. (3-3)) prove to be an improvement over the version by Miller [13] (Eq. (3-1)). It is more realistic to have the wave originate in the apertures, instead of the origin of the aperture. The use of multiple vibrating plates with different origins also resulted in less distinction between the main- and sidelobes. We expect that the aperture soundfield with the new ATF, which is less artificial, is easier to replicate with a set of loudspeakers, especially because of the less sharp lines between the lobes. Following these results, four specific points are addressed.

Validity of Aperture ATF We contributed to the existing literature by extending the aperture ATF, as in Eq. (3-3), which allowed us to have valid equations that describe soundfields

at a reasonable distance from the aperture. However, the use of this new aperture ATF is not verified with real-world experiments or measurements. Therefore, we have not validated the simulation environment and can not guarantee that the wave propagation through an aperture is properly described by Eq. (3-3) and Eq. (3-4). With the vibrating plate assumption, we model an infinitely thin aperture, while in reality, the border of an aperture could actually distort the sound waves propagating through it. In an open-loop algorithm, which is model-dependent, it is important to have valid Acoustic Transfer Functions (ATFs) such that performance is guaranteed. Validation of the aperture ATF is suggested as future work. This does not influence the algorithm itself, so the same wave-domain algorithm can be used with adjusted or validated ATF.

Time-delay Wrapping The method to prevent time-delay wrapping is an addition to existing modeling approaches. It allowed modeling in the wave-domain without having issues with the time-delay persistent in ATFs. This approach could be further explored in other wave-domain models where time-shifts limit the use of small window-sizes.

Conservation of Energy An advantage of modeling with ATFs, and only having to evaluate sound pressures on selective positions is that it keeps calculations relatively fast. It enabled us to run simulations on a normal laptop and quickly iterate and improve while working on the algorithms. However, some limitations to the model itself came to light. In our model, we measured the sound energy over an equally distributed array of evaluation microphones in the far-field. When we follow the Law of Conservation of Energy, attenuation of sound energy in all directions from the aperture into the room means that this sound energy has to go somewhere else. The total amount of energy in the system has to be the same. Our simulation model does not show where this energy goes, as we model the primary noise plane wave, aperture soundfield, and loudspeaker soundfields separately. The hypothesis is that a part of the energy, which normally would propagate through the aperture, is now reflected in the aperture opening due to the higher sound pressures at that location, while the other part is send back into the loudspeakers. Global attenuation can only be achieved if the total sound energy in the room is lower when the array is active. We suggest a theoretical study to answer the question where the energy is going to. The use of Finite Element Methods (FEMs), Finite Difference Methods (FDMs) or a software package like I-Simpa [53] most likely can hep to theoretically validate the ATF and study the total energy conservation.

Far-field Limitation More complex soundfield simulation software could also shed light on the properties of the soundfield closer to the aperture. Currently, the simulation model is valid in the far-field (> 0.88 m), limiting the system to implementations where we can only guarantee global attenuation outside this near-field region. If the far-field distance of the aperture ATF could be reduced, it could further reduce the computational load of the wave-domain algorithm. With valid equations closer to the aperture, the control region could be moved closer. This means a reduced r_{max} of the spherical control region. Consequently, this leads to a lower number of basis functions necessary for accurate filter-weight calculations, see Eq. (4-35).

Despite the above mentioned limitations, the simulation environment allowed us to develop the wave-domain algorithm, which was the main objective of this study. Keeping the simulation

model relatively straightforward improved the workflow and kept computational load low. With the new aperture ATF, we are one step closer to modeling the real propagation of waves through apertures. Finally, the simulation environment enabled us to generate insightful animations that helped during the development of the wave-domain algorithm.

6-2 Reference Algorithm and Optimal Wave-domain Algorithm

The goal of modeling a simulation environment was to enable us to develop a wave-domain based algorithm that can actively control sound propagating through apertures. We could compare the results of our new algorithm with existing algorithms used in ANC systems for apertures. We implemented an existing reference algorithm, the multiple-error Normalized Least Mean Squares (NLMS) algorithm, as well as the new wave-domain algorithm. The new algorithm we developed contains a new method of basis function orthonormalization and matrix formulation. Both concepts build on methods that were explored in related fields like wave-field synthesis [54] and wave-domain control [55]. Moreover, the attenuation performance of the wave-domain algorithm and reference algorithm were analyzed in the simulation environment with a sparse array [14] and grid array [6], for three distinct incident angles [20] while cancelling three types of noise. The results indicate that our open-loop wave-domain algorithm has the potential to attenuate more noise than existing closed-loop algorithms, while the components of no convergence time, low computational effort, and no need for error microphones ensure its feasibility in noise pollution problems of modern society. We discuss the results in more detail in this section.

First, we begin by comparing the optimal wave-domain algorithm to the reference algorithm. The error microphones of the NLMS based system were placed in the control region of the wave-domain algorithm. When enough error microphones are chosen for the NLMS, it essentially tries to minimize the soundfield in the spatial region similar to the wave-domain algorithm. For white noise, the results of both the reference algorithm and wave-domain algorithm are very similar. The attenuation performance for various frequency bins (Figure 5-11a) is in line with results found in literature [45]. Performance is best at low frequencies, but when moving to the higher frequencies, the attenuation degrades. Similar trends are seen in Figure 5-18 where rumbler-siren sound attenuation is shown. We can conclude that the reference algorithm performed less for real-world noise signal compared to white noise performance, while the wave-domain algorithm attenuation was almost the same for the two signals. The LMS-based algorithm uses the signal to calculate the filter-weights, and because of the changing character of the signal, the most optimal filter-weights are not found and change over time. Pursuant to existing findings, the closed-loop reference algorithm takes time to convergence to a steady-state [10]. This disadvantage is not persistent for the wave-domain algorithm, as offline filter-weight calculation directly generates the optimal filter, independent of the type of signal. Furthermore, by calculating the weights offline, the method does not have to cope with real-time closed-loop computational limitations, something that is limiting the feasibility of LMS-based algorithms [10,11]. According to prior implementations of ANC for apertures, the positioning of loudspeakers distributed over the aperture outperformed the boundary placed loudspeaker arrays [6,56]. The results found in this study support this. Similarly, the performance for attenuating noise that arrives with an incident angle was slightly lower than the on-axis scenario, which is also noted previously [20]. Noteworthy is that the

newly developed algorithm is less sensitive for incident angle increase. In summary, the results imply that the wave-domain algorithm is a promising approach to active noise control for apertures, where the optimal wave-domain algorithm is able to cancel undesired noise up to -10 dB, in the range of 0 to 2 kHz. The following six points need to be mentioned since they limit the algorithm, and could lead to further improvement of our work.

Number of Basis Functions Prior studies indicated that the number of basis functions depends on both the wavenumber and size of the region [48], where they focused on two spatial regions, a disc and a ball. We followed their suggestions on the number of basis functions for our regions. However, as we have smaller sections of their proposed control region shapes, we expected to need less basis functions for accurate results. After an informal comparison, we chose to scale the number of 3D basis functions with $\beta = 1/16$ and obtained proper results. Nevertheless, we can not state that, with our basis function expansion, the truncation error is at least lower than 16.1%. Therefore, a suggestion for future work is to derive a mathematical proof, as well as a value for the relative truncation error, specifically for thick hemisphere control regions. A similar approach as in the work of Kennedy et al. [48] is recommended.

Impact of Window-size on Performance In this study, we explored the performance of the wave-domain algorithm for various window-sizes, ranging from 1 ms to 125 ms. As we expected, the performance is influenced by the chosen window-size. We supplemented the results with two indicators that shed light on this influence.

Firstly, we developed a fundamental error analysis method, presented in Section 3-3-3. With this method, we showed that the use of a small window-size leads to aliasing in the frequency domain. This coerced a frequency response that was a smoothed version of the high-resolution frequency response of the aperture ATF. Because of this, matrix \mathbf{a} , in the wave-domain algorithm, contains an error. Consequently, the filter-weights are calculated with this error, leading to less attenuation for smaller window-sizes, as we have seen in Figure 5-7. This error is larger for positions closer to the wall, as the impulse response is longer in such cases while the window-size remains short. Moreover, the implemented error analysis assumes that every evaluation microphone is weighted equally. However, we see in Figure 5-2 that the sound energy differs significantly for each position in the room. Therefore, the results obtained with the error analysis method do not resemble the simulations perfectly. Nevertheless, they are an indication that the window-size is a limiting factor in the wave-domain algorithm.

Secondly, we looked into the size of the phase-shifts that are present in the filter-weights $W_q(k)$. From Figure 5-8 and Figure 5-9 we concluded that the impact of the window-size mostly affected the performance at low frequencies. At lower frequencies, a phase-shift means a larger shift in samples. In our time-delay wrapping avoidance method in Section 3-3-2, we prevented the issue of large phase shifts wrapping due to the circularity assumption of the Short-time Fourier Transform (STFT). However, this issue is not solved in the multiplication with filter-weights $W_q(k)$. With a brief calculation, we can indicate that the phase-shifts are indeed a factor that needs to be considered. Figure 5-10 shows that the phase-shifts in $W_q(k)$ are concentrated around -180° . This makes sense as we try to generate a loudspeaker soundfield that is the inverse of the aperture soundfield. In the lower limit, we have a window-size of $N = 1$ ms. Application of the phase-shifts shown in Figure 5-10a to a signal at

$f = 128$ Hz, with a period of $T_{\text{phase-shift}} = 1/f = 1/128 \approx 7.8$ ms, leads to a shift of $T_{\text{phase-shift}}/2 = 3.9$ ms. Clearly, $T_{\text{phase-shift}}/2 = 3.9 \gg N = 1$ ms, so the signal is significantly wrapped and leads to a large error at that frequency. As seen in Figure 5-8, the influence becomes negligible when $N \ll T_{\text{phase-shift}}/2$, here in case of $N \geq 61.5$.

As a consequence, the window-size significantly limits the performance of the wave-domain algorithm. To obtain decent performance at smaller window-sizes, we recommend to investigate the implementation of the filter-weights $W_q(k)$ in wave-domain. A method that prevents the wrapping, perhaps similar to our time-delay wrapping solution in Section 3-3-2, could prevent this issue. Furthermore, a time-domain implementation with impulse responses and linear convolution could be explored. If this is possible, the algorithmic delay problem becomes less severe and the wave-domain algorithm becomes even more feasible for real-world application.

Grid versus Sparse Array Our literature study indicated that distributed loudspeaker arrays [6] outperform boundary arrays [5]. The simulation results we obtained are in line with those findings. Figure 5-15 shows that both arrays perform decently at low frequencies. However, for noise at frequencies of 1.5 kHz and higher, the sparse array is not able to replicate the soundfield from the aperture, as we see in Figure 5-16. One reason could be a lower number of loudspeakers in the sparse array (21), compared to the grid array (49). In that trend, the spacing between the loudspeakers could be too large, leading to issues like spatial aliasing. During the study, we briefly examined the performance of other sparse arrays, for which the results can be found in Appendix D. These results indicate that increasing the number of loudspeakers has a beneficial effect, but the performance limit still was around 1.5 kHz. Comparably, adding four loudspeakers around the aperture center led to similar improvements. In addition to these results, we have seen that the sparse array needs a larger τ . Constraining the loudspeaker effort resulted in better performance for the sparse array. This suggests that investigating optimal τ for different wavenumbers could further improve the performance of the wave-domain algorithm. It is recommended to look further into the specifications of the loudspeaker array, as sparse or boundary approaches are much more feasible than a distributed grid array. Grid arrays block the aperture, limiting its properties of light and air transmission, while the aperture functionalities remain intact with sparse arrays. While the performance of the sparse array was lower than the grid array, we have seen that proper attenuation at lower frequencies is achievable with the sparse array. To conclude, with our developed simulation model, comparing various array types is easily achievable and we propose to use it when designing a loudspeaker array.

Accuracy of ATFs The wave-domain algorithm deploys loudspeaker and aperture ATFs to describe the soundfield by the array and aperture, respectively. The loudspeakers are modeled as monopoles. However, when implementing the wave-domain algorithm in a physical setup, it is important to have a correct ATF of the used loudspeakers. In our simulations, any misalignment between the real loudspeaker behavior and the representing equation is not modeled. Similarly, the validity of the aperture ATF, as discussed in the prior section, is important here as well. The real-world implementation of the wave-domain algorithm highly depends on the accuracy of both ATFs, as the method is a model-based approach. To that extent, the use of a perfect measurement by the reference microphone directly influences the

performance as well. Therefore, the calibration of the loudspeakers and reference microphones is of much importance. In contrast, LMS-based algorithms have an advantage here. Closed-loop methods intrinsically account for the non-perfect frequency responses of the loudspeakers and microphones and do not use any aperture ATF. Perhaps, to account for this lack of feedback in the wave-domain algorithm, a combination between an open-loop and closed-loop algorithm could be developed. An idea could be to position a few error microphones in the room that solve the issue of calibration or misalignment between the used ATF equations and real-world ATFs of the loudspeakers and microphones.

Biased Algorithms In our simulations, both implementations of the algorithm were biased. On the one hand we induced bias in the Least Mean Squares (LMS)-based algorithm. As we denoted in the theory of the reference system, the method uses an estimate of the secondary path to feed a filtered signal of the input signal to the LMS algorithm: \hat{S}_{rle} . However, in our simulation environment, we implemented the predefined loudspeaker ATF both as the estimate and actual secondary path, so $\hat{S}_{rle} = S_{rle}$. This means that the LMS-based algorithm has a perfect estimate of the secondary path, something which is unrealistic. Therefore, the performance of the multiple-error NLMS algorithm could be lower when implemented in the real-world. The significance of this is not deductible from our simulation model. There is extensive literature on (online) secondary path modeling available [7], which can be used to overcome this issue. On the other hand, we invoked a bias in the wave-domain algorithm. We employed the aperture ATF, discussed in Section 3-2-1, both in the simulation environment and filter-weight calculation algorithm. Hence, the filter-weights are calculated in a way that they will exactly cancel the soundfield by the aperture in the simulation environment. The only limiting factors here are the wave-domain aliasing and wrapping due to the large phase-shifts. This rather ideal case is unlikely in reality, as we have not validated our aperture ATF. Similar to the multiple-error NLMS algorithm, the performance of the wave-domain algorithm could be lower when it is implemented in the real-world. The effect, however, is not measurable in our simulation environment. As mentioned in the prior section, validation of the aperture ATF is a useful continuation of this work to indicate the precision of the aperture ATF. In summary, both algorithms are realized with a slight bias towards better results, all the more reason to further develop the work and implement it in a real-world setup.

Moving Primary Noise Source A major drawback of closed-loop systems used for active noise control for apertures is the convergence time. As indicated in Figure 5-19, the multiple-error NLMS algorithm requires 0.1 s to convergence when a rumbler-siren sound is cancelled, and up to 1.0 s for white noise. In contrast, the wave-domain approach does not need any time to convergence, as offline weight calculation is employed. Currently, the simulations show results for fixed positions of the primary noise source. The next step would be to cancel noise coming from a moving primary noise source. To facilitate this, the LMS-based algorithm does not need any conceptual changes. However, as the convergence time of such algorithms is long, we expect that a moving source could cause the NLMS algorithm, in its current setup, not to converge fast enough. If that occurs, the performance will be limited, depending on the speed of movement and type of noise. On the contrary, the wave-domain algorithm currently operates under various assumptions. One is the fact that it knows the incident angle of the primary noise source. However, when implemented in the real-world with a primary noise source that is moving, the algorithm must include a method to determine the incident angle

in real-time. Existing sound source positioning algorithms already exist and could be used for this [39]. Moreover, the filter weights are calculated offline for distinct incident angle cases. When the incident angle changes over time, the filter weights also have to be adjusted, as matrix \mathbf{a} depends on the incident angle values θ_0, ϕ_0 . We propose to calculate, offline, a multitude of \mathbf{a} matrices for a large, discretized number of incident angles and use interpolation between the matrices to create a large database with \mathbf{a} matrices. When an incident angle calculation algorithm is in place, a complementarity set of filter-weights can be calculated with Eq. (4-29). This can be used to update the soundfield of the loudspeaker array to keep it in line with the soundfield of the aperture from the moving primary noise source.

In summary, the wave-domain algorithm has many advantages over current implementation ANC for apertures. The results indicate a continuous average attenuation of approximately -10 dB between 0 and 2 kHz, independent of the type of noise. In this section, we presented initial findings and suggestions for future work on the optimal wave-domain algorithm. As we noted, an artifact of the wave-domain algorithm is the algorithmic delay induced by the block-processing. In this study, we briefly investigated methods to compensate for this. The next section will cover a concise discussion on this subject.

6-3 Algorithmic Delay Compensation

In the search for an open-loop wave-domain algorithm, the compensation for algorithmic delay is an important factor. We limited the exploration to two practical methods: re-positioning of the reference microphone and including a prediction algorithm. Initial results were presented in this study, and we compared the two methods reciprocally as well as with the reference algorithm. Let us focus on Table 5-2. In comparison, the NLMS algorithms performance exceeded the Wave-Domain Algorithm with Reference Microphone Clacement (WDC-M) and the Wave-Domain Algorithm with Predictor (WDC-P) (no performance) when a more stationary primary noise, such as airplane and white noise, had to be cancelled. However, for more deterministic and non-stationary rumbler-siren noise, the microphone placement method came out on top with a Signal-to-Noise Ratio (SNR) of -6.1 dB compared to NLMS and WDC-P. The predictor-based method only generated results for the rumbler-siren signal. This is expected, as this type of signal is reasonably predictable. Therefore, the WDC-P method is only a useful method when primary noise signals are somewhat deterministic, such as sirens or speech [24]. Let us discuss some aspects that could help improve the performance of these two methods further.

For the first algorithmic delay compensation, the WDC-M, the main limiting factor is, in fact, the window-size. As we have discussed in the prior section, the performance for smaller window-sizes is limited by the low-frequency resolution and large phase-shifts. In order to improve the performance of this algorithmic delay compensation method, performance with small window-sizes should be increased. As suggested, looking into omitting large phase-shifts in the wave-domain, as well as full time-domain implementations, could improve the approach of positioning the reference microphone at a distance from the aperture. It is imaginable that, depending on the physical setup, the reference microphone could be placed even further from the aperture, which will increase performance. This is feasible when the primary noise source is located very far away from the aperture and remains in a fixed position.

In this study, we showed that the WDC-P approach only worked for the case of a rumbler-siren sound. It could not cancel airplane noise and white noise because the autoregressive model did not predict the signals well. Suggestions for future work are to look into different predictors. As we explored in our literature study [12], the use of machine learning models such as a WaveNet [57], a Recurrent Neural Network (RNN) or a Neural Ordinary Differential Equations (NODE) model are expected to have a better performance than Autoregressive (AR) models.

Finally, we propose a combination of the predictor and microphone placement, as this approach has the potential to perform better than the standalone methods. However, it is advised to implement a method that detects whether the prediction of the algorithm is valuable, such that the predictor could be turned off when it is not adding value or reducing performance.

The purpose of this chapter was to discuss the results obtained during the simulations in more detail. Additionally, we reviewed some limitations of this study and gave suggestions for future research. Altogether, the optimal wave-domain algorithm cancelled all three types of noise properly, leading to a global reduction of -10 dB. The performance was only slightly lower when noise entered the room with an incident angle. The low computational effort due to offline filter-weight calculation, as well as the absence of convergence time and error microphones proof that the wave-domain approach to ANC for apertures is a feasible method. Algorithmic delay due to the window-size in block processing limits the performance. However, there is space for improvement of the algorithmic delay compensation methods. All things considered, wave-domain algorithms are the best performing solution to actively control undesired noise through apertures.

Chapter 7

Conclusion

Society is threatened by noise pollution in homes. Researchers have focused on applying active noise control systems to attenuate noise propagating through open windows. With a loudspeaker array in the aperture, such systems aim to reduce noise in homes while maintaining ventilation and sight. Existing studies focused on implementing closed-loop algorithms that minimize undesired sounds at error microphones locations in the room. These Least Mean Squares (LMS) based approaches have the disadvantages of long convergence times and computational efforts when the systems include a large number of speakers and microphones, essential for global control. Furthermore, closed-loop setups are unfeasible due to the necessity of a large array of error microphones in the room, blocking the view. Also, these error microphones are located inside the room, and the measured error signals are affected by sound originating from indoor. This reduces the performance of closed-loop algorithms. To overcome these disadvantages, we propose an open-loop approach. This approach has instant convergence, deploys offline filter-weight calculation, and does not require error microphones. In this study, we developed and tested this new algorithm by taking the following steps. Firstly, we came up with an Acoustic Transfer Function (ATF) to describe the soundfields by an aperture. Thereafter, a simulation environment was created, in which we implemented both the newly developed algorithm as well as a multiple-error Normalized Least Mean Squares (NLMS) algorithm, used as reference. Within the wave-domain algorithm, we came up with a new method to orthonormalize basis functions and presented a matrix formulation that reduces calculation efforts. The algorithmic delay, induced by the window-size in the block-processing of wave-domain algorithms, must be compensated at the cost of attenuation performance loss. We compared two methods for compensation. On the one hand, we tested the approach of positioning the reference microphone at a distance in front of the aperture. On the other hand, we investigated compensation by predicting the undesired noise with an autoregressive model. Finally, we compared the performance of the reference algorithm and wave-domain algorithm for various types of noise and three fixed incident angles. These steps helped us to determine the attenuation performance of open-loop wave-domain algorithms for active cancellation of noise propagating through open windows.

By analyzing quantitative attenuation results up to 2 kHz, we conclude that the wave-domain algorithm achieved superior attenuation compared to the multiple-error NLMS algorithm.

The wave-domain algorithm achieved -10.6 dB attenuation for the airplane noise and rumbler-siren sound. In comparison, the reference algorithm obtained attenuation results of -7.6 and -5.2 dB for the airplane noise and rumbler-siren, respectively. Moreover, the wave-domain algorithm outperforms the closed-loop algorithm when the primary noise originates from an off-axis direction. In line with existing literature, we showed that closed-loop algorithms have the disadvantage of convergence time, while the wave-domain algorithm starts attenuating instantaneously. The absence of error microphones in the wave-domain approach increases its feasibility because the aperture is not blocked by a large array of microphones. Also, the proposed algorithm is not affected by sound originating from indoor. During our simulations, grid arrays outperformed sparse arrays, supporting previous studies. Furthermore, compensation for algorithmic delay in the wave-domain algorithm by microphone placement is preferred over the predictor approach, as the latter did not work for all signals. More specifically, the trade-off between prediction performance and algorithm performance shows that the optimal window-size can not be set beforehand and depends on the predictability of the signal. Hence, this method is only feasible with predictable signals that are known beforehand. Additionally, we compared our soundfield Acoustic Transfer Functions (ATFs) with previously used ATFs. Based on a qualitative assessment of soundfield images, we indicate that our newly developed aperture ATF is an improved version of the existing theory. Moreover, it allowed us to design the wave-domain algorithm with a control region at 1 m distance from the aperture, while previous ATFs could not. However, the newly obtained ATF needs to be validated.

With this work, we contributed to the existing theory on active noise control for apertures in various ways. We built a simulation environment that can be used for designing and testing loudspeaker arrays for any arbitrary aperture size. Moreover, the extended aperture ATF gives the opportunity to develop wave-domain algorithms that ensure attenuation close to the aperture. As such, we took this opportunity and developed a wave-domain algorithm, for which we found a new method of basis function orthonormalization, as well as a matrix formulation that reduces calculation efforts. We addressed the issue of time-delay wrapping in the Short-time Fourier Transform (STFT) block-processing. Finally, it is desired to limit the algorithmic delay by reducing the window-size, so we derived a method to analyze the fundamental attenuation limit of the wave-domain algorithm.

During this study, we found that the performance of the wave-domain algorithm is mostly limited by the window-size. Hence, we suggest investigating the omitting of large phase-shifts from the Short-time Fourier Transform for further improvement of the algorithm. Future work could include comparing more advanced predictors and combining microphone placement and predictors into one method for algorithmic delay compensation. Furthermore, extending the simulation environment and wave-domain algorithm to include a moving primary noise source is a natural continuation of this work.

With this study, we are a step closer towards silent open windows due to the insights of the proposed wave-domain algorithm obtained in the simulation environment. The proposed algorithm achieved superior performance, compared to state-of-the-art algorithms. To conclude, the use of wave-domain Active Noise Control (ANC) algorithms for apertures is natural in addressing the noise pollution problems of modern society, especially because the proposed method requires neither time to converge nor error microphones and operates at a low computational effort due to offline weight calculation.

Appendix A

Conference Paper

The conference paper submitted to IEEE ICASSP 2022 [58] is copied below.

WAVE-DOMAIN APPROACH FOR CANCELLING NOISE ENTERING OPEN WINDOWS

Daan Ratering^{*}, W. Bastiaan Kleijn^{†‡}, Jean Gonzalez Silva^{*}, and Riccardo M.G. Ferrari^{*}

^{*} Delft Center for Systems and Control and [‡] Dept. Microelect., Delft Univ. of Technology, Netherlands

[†] School of Engineering and Computer Science, Victoria University of Wellington, New Zealand

ABSTRACT

Active control of noise propagating through apertures is commonly realized with closed-loop LMS algorithms. However, these algorithms require a large number of error microphones and provide only local attenuation. Slow convergence and high computational effort are additional disadvantages. We propose a wave-domain approach that converges instantaneously, operates with low computational effort and does not require error microphones. It inherently controls sound in all directions in the far-field. The soundfield from the aperture is matched in a least squares sense with the generated soundfield from the loudspeaker array using orthonormal basis functions. Compensation for algorithmic delay, induced by blockwise processing, can be based on microphone placement or signal prediction, at the cost of a loss in attenuation performance. Our simulation results indicate that wave-domain processing has the potential to outperform LMS-based methods in practical active noise control for apertures.

Index Terms— open window, aperture, active noise control, wave-domain algorithm, multiple-error least mean squares algorithm, algorithmic delay compensation

1. INTRODUCTION

Noise pollution is a major health threat to society [1]. Active Noise Control (ANC) systems that attenuate noise propagating through open windows (apertures) have the potential to create quieter homes while maintaining ventilation and sight through the apertures [2]. ANC systems employ loudspeakers to produce anti-noise soundfields that reduce the sound energy in noise-cancelling headphones [3] or over large regions such as airplane cabins [4]. Actively controlling sound propagating through open windows is being studied [5, 6]. The objective for these systems is to reduce the sound energy in all directions from the aperture into the room. Current methods employ closed-loop algorithms, leading to long convergence times [7], heavy computational load [8] and the need for a large number of error microphones being positioned in the room. These drawbacks limit the feasibility of such systems [9].

Most ANC systems for apertures utilize closed-loop Least Mean Squares (LMS) algorithms, such as the well-known Filtered-x LMS (FxLMS) algorithm [7], or its multi-channel equivalent, the multiple-error LMS [7, 10]. These closed-loop algorithms aim to minimize error signals at error microphones placed in the room by adapting signals generated by loudspeakers in the aperture. Current studies indicate that LMS based ANC for apertures can obtain a 10 dB global reduction of white noise between 0.4 and 1 kHz [11] and 0.5 and 2 kHz [12]. To set a baseline, the multiple-error normalized LMS algorithm will be used as a reference system.

We note that the control problem has similarities with wave-domain spatial control of the sound produced by multi-speaker sound systems, e.g. [13–15] and active cancellation of wall reflection of sound of a talking person [16], as well as wave-field synthesis [17]. Such a wave-domain algorithm uses a temporal frequency domain basis function expansion over a control region. The soundfield from the aperture and loudspeaker array can be expressed in these basis functions and their sum can be minimized in least squares sense [13–15].

In general, ANC systems for open windows with loudspeakers distributed over the aperture [6] outperform those with loudspeakers placed on the boundary of the aperture [5]. A compromise between both setups is a sparse array, that consists of a window with a cross-bar containing the speakers [18, 19].

We propose a wave-domain approach to ANC for apertures that addresses the shortcomings of the closed-loop LMS approach. It intrinsically ensures global control, because it cancels noise in all directions from the aperture, and does not require microphones positioned in the room. To our knowledge, our work is the first to use the wave-domain approach for ANC. We calculate the optimal filter-weights that minimize far-field sound energy for each frequency. We extend known Acoustic Transfer Functions (ATFs) that describe the sound propagation through apertures [18] and from loudspeakers [20]. The wave-domain algorithm operates in the temporal frequency domain. Hence it is necessary to transform signals with the Short-time Fourier Transform (STFT). This operation induces an filter-delay equal to the window-size of the STFT. The delay can be compensated for by signal prediction or microphone placement. With prediction, optimal window-size results from a trade-off between prediction performance and the algorithm performance [14].

Our contribution is the development of a wave-domain ANC for apertures that has the potential to outperform current LMS systems, and its implementation in a simulation environment. Moreover, the development of an extension to aperture ATFs [18], a procedure for basis function orthonormalization with Cholesky decomposition, and the matrix implementation of filter-weight calculation, extend current literature. An advantage of our wave-domain control system over existing LMS-based systems is that the filter weights are calculated off-line, leading to a lower computational effort. Furthermore, these coefficients are computed independent of the incoming noise. Therefore, the wave-domain approach itself requires no time to convergence. Its performance is affected by the algorithmic delay compensation method, the accuracy with which the aperture is represented and the physical characteristics of the microphone and loudspeaker arrays. In this study, both sparse and grid arrays will be used, and their performance is compared.

Section 2 formulates the problem and covers necessary ATFs. The open-loop wave-domain algorithm is developed in Section 3. Section 4 describes the reference system and evaluation methods and discusses the results. Finally, a conclusion is given in Section 5.

We thank the GN Group for funding this research.

2. PROBLEM FORMULATION AND NOTATION

Our objective is to develop an open-loop wave-domain control algorithm that ensures global attenuation of noise propagating through an aperture. The algorithm is designed to achieve cancellation in the far-field ($r > 0.8$ m). Following Huygens' principle, the energy behind a finite control region is minimized if a wavefront, with minimized sound energy, is created in that region [21]. The aim of the algorithm is to generate such a wavefront in the control region.

Throughout the paper, $k = 2\pi f/c$ is the wavenumber, $j = \sqrt{-1}$ is the imaginary number, the unnormalized sinc function is used [20] and $[\cdot]^H$ and $\|\cdot\|$ are the conjugate transpose and the Euclidean norm, respectively. We use spherical coordinates with radius r , inclination θ and azimuth ϕ and corresponding Cartesian coordinates $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$ and $z = r \cos \theta$.

The noise is assumed to be a plane wave [12], with fixed incident angle (θ_0, ϕ_0) . Then, we constrain the problem by modeling the aperture as a sum of square baffled pistons in an infinitely large wall with an ATF [18]. Such an ATF relates the pressure of the plane wave with the pressure of the soundfield at position \mathbf{x} in the room. The equation, for 3D modeling, is derived from [20] as:

$$H^{ap}(\mathbf{x}, k, \theta_0, \phi_0) = \frac{jck\rho_0}{2\pi} \dot{\omega}_0 \Delta L_x \Delta L_y \sum_{i=1}^{\hat{P}} D_i, \quad (1)$$

where c is the speed of sound, $\dot{\omega}_0$ is a gain constant, ΔL_x and ΔL_y are aperture section dimensions and \hat{P} is the number of aperture sections. D_i is the directivity, of each piston, defined as:

$$D_i = \frac{e^{-jk(r_i + \tau_i)}}{r_i} \text{sinc}\left(\frac{\Delta L_x k(\sin \theta_i \cos \phi_i - \sin \theta_0 \cos \phi_0)}{2}\right) \text{sinc}\left(\frac{\Delta L_y k(\sin \theta_i \sin \phi_i - \sin \theta_0 \sin \phi_0)}{2}\right), \quad (2)$$

where, for section i , r_i , θ_i and ϕ_i are the adjusted spherical coordinates and τ_i is a delay term due to the incident angle of the plane wave. Modeling in 2D is done by removing the height ΔL_x , omitting the sinc function in the x direction, and setting $x = 0$.

Furthermore, we constrain the system by modeling the ATFs of Q number of loudspeakers as monopoles [20, 22]:

$$H_q^{ls}(\mathbf{x}, k) = \frac{jck\rho_0}{4\pi} A_q \frac{e^{-jkr_q}}{r_q}, \quad (3)$$

in which $A_q = 4\pi a_{point}^2 u_0$ is each monopole's amplitude, with u_0 a surface velocity gain constant, a_{point} the radius of the point source and r_q the adjusted spherical radius from a monopole to a position \mathbf{x} in the room. The soundfield from the loudspeaker array is the sum of multiple loudspeaker soundfields. The loudspeaker ATF in (3), holds in 2D and 3D.

3. WAVE-DOMAIN ALGORITHM

3.1. Basis functions and control region

We define the control region over which we want to minimize the sound pressures as a spherical shell with finite thickness in 3D: $\mathbb{D}_{3D} = \{r_{min} \leq r \leq r_{max}, 0 \leq \theta \leq \frac{\pi}{2}, 0 \leq \phi \leq 2\pi\}$ and an arc with finite thickness in 2D: $\mathbb{D}_{2D} = \{r_{min} \leq r \leq r_{max}, 0 \leq \theta \leq \frac{\pi}{2}, \phi = \frac{\pi}{2} \wedge \phi = -\frac{\pi}{2}\}$.

The soundfield over the control region at a single wavenumber k , denoted $S(\mathbf{x}, k) : \mathbb{D} \times \mathbb{R} \rightarrow \mathbb{C}$ is written as a weighted series of basis functions $\{U_g\}_{g \in \mathcal{G}}$ [15, 20]:

$$S(\mathbf{x}, k) = \sum_g E_g U_g(\mathbf{x}, k), \quad (4)$$

where $S(\mathbf{x}, k)$ is the soundfield, E_g are G coefficients and $U_g(\mathbf{x}, k)$ is a $G \times 1$ vector containing G basis functions. We assume that all feasible solutions on \mathbb{D} fall in the Hilbert space spanned by the orthonormal set $\{U_g\}_{g \in \mathcal{G}}$. The inner-product is defined as $\langle Y_1, Y_2 \rangle = \int_{\mathbb{D}} Y_1(\mathbf{x}) Y_2^H(\mathbf{x}) d\mathbf{x}$, where Y_1 and Y_2 are functions of the form $Y_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $Y_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$. For a given $S(\mathbf{x}, k)$ and $U_g(\mathbf{x}, k)$, the coefficients E_g are obtained with $E_g = \langle S(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle$.

The orthonormal set of basis functions is obtained by starting with a set of non-orthogonal functions $f_g(\mathbf{x}, k) : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{C}$ that represent G plane waves with directions evenly distributed in 3D [23]. The non-orthogonal functions are defined as plane waves: $f_g(\mathbf{x}, k) = e^{jk\mathbf{x} \cdot \hat{\beta}_g}$, where $\hat{\beta}_g \equiv (1, \theta'_g, \phi'_g)$ is the unit vector in the direction of the g th plane wave. We normalize each function to obtain $\hat{f}_g(\mathbf{x}, k) = \frac{f_g}{\|f_g\|}$ and create vector $\hat{f} = [\hat{f}_1 \hat{f}_2 \dots \hat{f}_G]^T$.

Next, we find a lower triangular matrix R such that $U = R\hat{f}$, where U is a vector containing G orthonormal basis functions. We define:

$$F = \hat{f}\hat{f}^T = \begin{bmatrix} F_{(1,1)} & F_{(1,2)} & \dots & F_{(1,G)} \\ F_{(2,1)} & F_{(2,2)} & & \vdots \\ \vdots & & \ddots & \\ F_{(G,1)} & \dots & \dots & F_{(G,G)} \end{bmatrix}, \quad (5)$$

where F is symmetric and positive definite: $\mathbf{x}^H F \mathbf{x} > 0 \forall \mathbf{x} \in \mathbb{C}^n$. Then, we set $V = R^{-1}$ and obtain V by applying the Cholesky decomposition on F , where $F = VV^T$ [24]. Finally, we obtain the orthonormal set of basis functions by inverting V , where $U = V^{-1}\hat{f} = R\hat{f}$.

In literature, $G_{3D} = (ekr_{max}/2+1)^2$, and $G_{2D} = ([2kr] + 1)$ basis functions are suggested for the expansion over a full ball and full disc, respectively [25]. This ensures a relative truncation error, in the generated soundfield, of less than 16%. However, our 3D spherical shell control region is not a full ball. An informal parameter optimization shows that we need $G_{3D}/16$ basis functions for equivalent results, resulting computational benefits. Finally, we expect that the 16% truncation error leads to performance loss.

3.2. Weight calculation

In this subsection, we discuss the procedure to obtain filter weights $l_q(k)$ for all loudspeakers q at wavenumber k .

First, we write the soundfields of the aperture as a sum of orthonormal basis functions:

$$S^{ap}(\mathbf{x}, k) = \sum_{g=1}^G A_g U_g(\mathbf{x}, k), \quad (6)$$

where H^{ap} is from (1). Weights A_g are obtained with the inner product: $A_g = \langle H^{ap}(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle$, written in matrix form with coefficient vector $\mathbf{a} = [A_1 A_2 \dots A_G]^T$ and a vector containing inner products between the ATF and the normalized basis functions denoted as $\mathbf{H}_{\hat{f}}^{ap} = [\langle H^{ap}, \hat{f}_1 \rangle \langle H^{ap}, \hat{f}_2 \rangle \dots \langle H^{ap}, \hat{f}_G \rangle]^T$. Plugging in $U = R\hat{f}$ gives: $\mathbf{a} = R\mathbf{H}_{\hat{f}}^{ap}$.

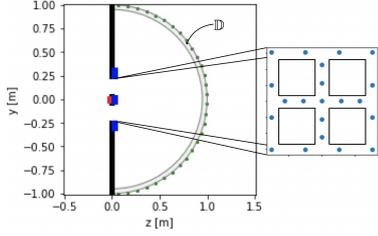


Fig. 1: 2D cross-section, with reference microphone (red), sparse loudspeaker array (blue), control region \mathbb{D} and evaluation microphones (green).

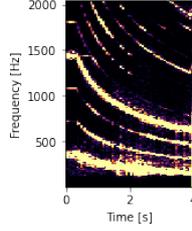


Fig. 2a: Spectrogram of rumbler signal with 32 ms window-size.

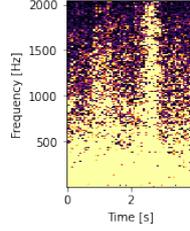


Fig. 2b: Spectrogram of airplane signal with 32 ms window-size.

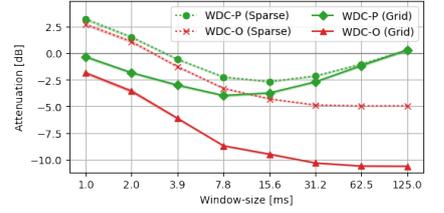


Fig. 4: 3D SNR results with optimal (WDC-O) and predictor (WDC-P) wave-domain controllers for rumbler-siren signals.

Similarly, we write the soundfield from a single loudspeaker as:

$$H_q^{ls}(\mathbf{x}, k) = \sum_{g=1}^G C_g^q U_g(\mathbf{x}, k), \quad (7)$$

with H_q^{ls} from (3) and coefficients C_g^q . The soundfield of the complete array is expanded with coefficients B_g , as well as a sum of the soundfields from all loudspeakers, multiplied by their filter weights:

$$S^{ar}(\mathbf{x}, k) = \sum_{g=1}^G B_g U_g(\mathbf{x}, k) = \sum_{q=1}^Q l_q(k) H_q^{ls}(\mathbf{x}, k), \quad (8)$$

from which we generate coefficients B_g by (8) and (7), leading to:

$$B_g = \sum_{q=1}^Q l_q(k) C_g^q, \quad (9)$$

where $C_g^q = \langle H_q^{ls}(\mathbf{x}, k), U_g(\mathbf{x}, k) \rangle$. In matrix form we obtain:

$$\mathbf{C} = \mathbf{R} \mathbf{H}_{\hat{\mathbf{f}}}^{ls} = \mathbf{R} \begin{bmatrix} \langle H_1^{ls}, \hat{f}_1 \rangle & \langle H_2^{ls}, \hat{f}_1 \rangle & \cdots & \langle H_Q^{ls}, \hat{f}_1 \rangle \\ \langle H_1^{ls}, \hat{f}_2 \rangle & \langle H_2^{ls}, \hat{f}_2 \rangle & & \vdots \\ \vdots & & \ddots & \\ \langle H_1^{ls}, \hat{f}_G \rangle & \cdots & \cdots & \langle H_Q^{ls}, \hat{f}_G \rangle \end{bmatrix} \quad (10)$$

where $\mathbf{H}_{\hat{\mathbf{f}}}^{ls}$ contains the inner products between the basis functions and the loudspeaker ATFs.

Finally, we set the control problem as the sum of the soundfields, $J(l_q) = S^{ap} + S^{ar}$, with $\eta = \|J(l_q)\|^2$ and minimize in least mean square sense: $\min_{l_q} \|J(l_q)\|^2$. With (6) and (8) gives

$$\eta = \left\| \sum_{g=1}^G A_g U_g(\mathbf{x}, k) + \sum_{g=1}^G B_g U_g(\mathbf{x}, k) \right\|^2 = \sum_g |A_g + B_g|^2. \quad (11)$$

With the knowledge that $\langle U_i, U_j \rangle = 0$, we can rewrite in matrix form. We denote $\mathbf{b} = \mathbf{C} \mathbf{l}$, where $\mathbf{l} = [l_1 \ l_2 \ \cdots \ l_Q]^T$. Furthermore, we add the regularization term $\tau \mathbf{l}$ with $\tau > 0$, to constrain the loudspeaker effort and ensure a robust solution [15]:

$$\eta = (\mathbf{a} + \mathbf{b})^H (\mathbf{a} + \mathbf{b}) + \tau \|\mathbf{l}\|^2, \quad (12)$$

where \mathbf{l} can be found by:

$$\mathbf{l} = -(\mathbf{C}^H \mathbf{C} + \tau \mathbf{I})^{-1} \mathbf{C}^H \mathbf{a}, \quad (13)$$

in which \mathbf{I} is an identity matrix.

3.3. Block-processing

An element-wise multiplication of the ATF with a STFT block is employed to transform signals, from the aperture and loudspeakers, to any position in the room. For the STFT, the window-function, $w(n)$ of length N is chosen to fulfill $\sum_{m \in \mathcal{Z}} w(n - mH)^2 = 1$ where n is the discrete time index, m is hop-number and H is the hop-size. This ensures that we have a tight frame with perfect reconstruction [26].

The circularity property of the STFT leads to wrapping of the signals, if phase-shifts by ATFs become significant compared to the window-size. Employing zero-padding can reduce this issue, however, it emits the shifted signal content. We overcome this problem by removing the major time shift from the wave-domain multiplication and implementing it in the time-domain.

The block-processing with STFT in the wave-domain approach induces an algorithmic delay. The window-size N determines the length of the delay. Compensating for this can either be done by placing the reference microphone at a distance of at least cN/f_s in front of the aperture, or, by predicting the noise signal. We compare both methods. For prediction, an auto regressive model with *Yule-Walker estimators* predicts the incoming noise signal each STFT hop m over a prediction horizon equal to N [27].

4. RESULTS AND DISCUSSION

4.1. Reference system

The performance of the proposed open-loop wave-domain algorithm was compared with a multiple-error Normalized LMS (NLMS) algorithm [6, 10]. Its error microphones were, for fair comparison, positioned in two rows within the control region of the wave-domain controller such that soundfield minimization was done over the same region. For the algorithm, the reader is referred to [28] and [29].

4.2. Experimental setup

The 3D simulation environment represents a physical setup, entailing a window with crossbar [19]. A grid 49-loudspeaker array and a sparse 21-loudspeaker array [19] were compared. We assumed that, by measuring the performance in all directions, any reflection is irrelevant. Therefore, no walls were modeled. The cross-section ($x = 0$) top-view of the environment is depicted in Figure 1, with coordinates (x, y, z) pointing into the paper, upwards and to the right. The red dot is a reference microphone, the blue dots are loudspeakers and the green dots represent evaluation microphones. In 3D, the aperture was a $L_x = 0.5$ m by $L_y = 0.5$ m window, with a crossbar of width $W_+ = 0.065$ m. Hence, the aperture consisted of four squares ($\hat{P} = 4$) with $\Delta L_x = (L_x - W_+)/2 = \Delta L_y$. The 2D model was a L_y -wide aperture with a crossbar of width W_+ and $\hat{P} = 2$.

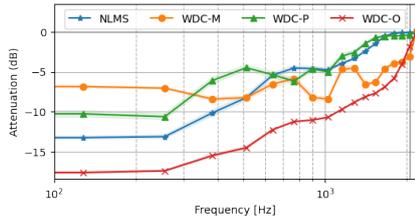


Fig. 4: $\overline{\text{SEG}}_f(k)$, with very narrow 99% confidence interval for grid array simulations in 3D, canceling rumbler-siren noise.

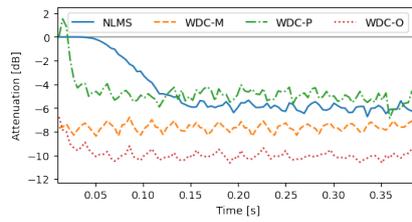


Fig. 5: First 0.4 s of $\overline{\text{SEG}}_f(k)$ 3D grid array results for all controllers, showing long convergence time of NLMS controller.

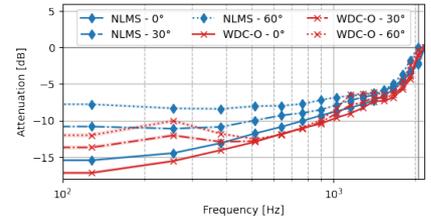


Fig. 6: $\overline{\text{SEG}}_f(k)$ for NLMS and WDC-O, with very narrow 99% confidence interval. 3D results for white noise signal, at 0°, 30° and 60° angle.

All controllers used one reference microphone, in the aperture origin and were implemented with the sparse and grid array. The NLMS was tested 32 (2D) and 128 (3D) error microphones in the control region. The optimal wave-domain controller (WDC-O) used a window-size of 125 ms. Additionally, algorithmic delay compensation was modeled by two approaches. One controller with the reference microphone positioned at 1.4 m in front of the aperture, implemented with a window-size of 3.9 ms (WDC-M) and the other as a wave-domain controller with auto regressive predictor (WDC-P). The wave-domain algorithms used a 75% STFT overlap.

Sample rate was set at $f_s = 2^{14}$ Hz. A fixed air temperature and density (ρ_0) were used, setting constant speed of sound at $c = 343$ m/s. To measure the performance of the controllers over time with a changing frequency spectrum, a rumbler-siren signal (Figure 2a) of 4 s was used as noise. Additionally, white noise and airplane noise (Figure 2b) were tested. Following existing studies [12], we evaluate the performance up to 2 kHz and for three incident angles: 0°, 30° and 60°. The performance was evaluated on the boundary of control regions \mathbb{D}_{2D} and \mathbb{D}_{3D} at 30 and 128 evenly distributed evaluation microphones, respectively. We define the segmental SNR in dB, summed over all evaluation microphones e as:

$$\text{SEG}_f(k, m) = 10 \log_{10} \frac{\sum_e |d_e(k, m)|^2}{\sum_e |d_e(k, m) + y_e(k, m)|^2} \quad (14)$$

where d_e is the noise signal and y_e is the loudspeaker array signal. We average $\text{SEG}_f(k, m)$ over frequency and time, to get insights per frequency bin ($\overline{\text{SEG}}_f(k)$), per hop ($\overline{\text{SEG}}_t(m)$) and in total (SNR). Performance was calculated over signal blocks with an 8 ms STFT with 50% overlap.

4.3. Attenuation performance

Table 1 shows the performance for all signals at 0° incident angle, where the grid outperformed the sparse array. WDC-O generated more attenuation than NLMS, when canceling rumbler-siren noise, especially at higher frequencies as shown in Figure 4. Additionally,

Noise type	Rumbler siren		Airplane Noise		White Noise	
	G	S	G	S	G	S
NLMS	-5.2	-3.5	-7.6	-4.5	-9.2	-4.3
WDC-M	-6.1	-1.3	-6.6	-2.0	-8.1	-2.8
WDC-P	-4.0	-2.2	-	-	-	-
WDC-O	-10.6	-4.9	-10.6	-4.7	-10.0	-4.1

Table 1: 3D, 0° angle, SNR for all controllers, with grid (G) and sparse (S) arrays. Attenuation is given in decibel (dB).

Figure 5 shows the slow convergence of NLMS, fast convergence of WDC-P, and instant convergence of WDC-O and WDC-M. Following Figure 6, WDC-O outperformed NLMS with better attenuation for each incident angle. When comparing algorithmic delay compensation methods, WDC-M slightly outperformed the WDC-P, with a grid array setup. Moreover, for WDC-P, a trade-off between prediction accuracy and algorithm performance was apparent so an optimal window-size is found, shown in Figure 3. However this optimum highly depends on the type of signal. For signals that are better predictable, the optimal window-size is larger. Finally, all controllers perform better at lower frequencies, except for WDC-M. For the latter, phase-shifts in the blockwise signal processing result in STFT wrapping.

The grid array outperformed the sparse array, confirming prior studies [30]. Besides, both the performance of white noise cancelling, and occurrence of long convergence time of the NLMS controller is in line with existing literature [12, 31]. For a stationary noise source, slow convergence is not a major issue. However, we expect that it limits the NLMS performance for moving noise source. In contrast, with instant convergence, the wave-domain controller is expected to perform better. Further investigation into performance for moving sources is necessary. Additionally, the trade-off apparent for WDC-P (Figure 3) follows prior results [14]. Improving algorithmic delay compensation methods, with a combination of microphone placement and prediction, or employing machine-learning, is a worthwhile continuation of this work. Offline calculation of filter-weights in WDC-O is a major advantage over closed-loop algorithms. Finally, we suggest to investigate STFT wrapping in wave-domain control algorithms, to improve performance at shorter window-sizes.

5. CONCLUSION

We proposed a new approach to ANC for apertures based on wave-domain processing with algorithmic delay compensation. We compared the new algorithm with a multiple-error Normalized LMS system for both sparse and grid arrays. Our 3D results showed that an ideal wave-domain algorithm obtains an average global attenuation of -10.6 dB up to 2 kHz for a 0.5 m window, compared to -5.2 dB for the LMS based algorithm when a grid array is used for attenuation rumbler-siren noise. Algorithmic delay compensation by microphone placement is preferred over predictor methods. Improving on delay solutions is a logical continuation of our work. Our results show that the use of wave-domain ANC algorithms for open windows is natural for addressing the noise pollution problems of modern society as the new method requires neither time to converge nor error microphones, and operates at a low computational effort due to off-line weight calculation.

6. REFERENCES

- [1] W. H. Organization, *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*, 2011.
- [2] L. Bhan and G. Woon-Seng, "Active acoustic windows: Towards a quieter home," *IEEE Potentials*, vol. 35, no. 1, pp. 11–18, 2016.
- [3] S. M. Kuo, S. Mitra, and W.-S. Gan, "Active noise control system for headphone applications," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 2, pp. 331–335, 2006.
- [4] S. J. Elliott, "Down with noise (active noise control)," *IEEE spectrum*, vol. 36, no. 6, pp. 54–61, 1999.
- [5] S. Wang, J. Yu, X. Qiu, M. Pawelczyk, A. Shaid, and L. Wang, "Active sound radiation control with secondary sources at the edge of the opening," *Applied Acoustics*, vol. 117, pp. 173–179, 2017.
- [6] B. Lam, S. Elliott, J. Cheer, and W.-S. Gan, "Physical limits on the performance of active noise control through open windows," *Applied Acoustics*, vol. 137, pp. 9–17, 2018.
- [7] S. M. Kuo and D. R. Morgan, "Active noise control: a tutorial review," *Proceedings of the IEEE*, vol. 87, no. 6, pp. 943–973, 1999.
- [8] L. Bhan, T. Murao, C. Shi, W.-S. Gan, and S. Elliott, "Feasibility of the full-rank fixed-filter approach in the active control of noise through open windows," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 253, no. 5. Institute of Noise Control Engineering, 2016, pp. 3548–3555.
- [9] C. Shi, T. Murao, D. Shi, B. Lam, and W.-S. Gan, "Open loop active control of noise through open windows," in *Proceedings of Meetings on Acoustics 172ASA*, vol. 29, no. 1. Acoustical Society of America, 2016, p. 030007.
- [10] C. Shi, N. Jiang, H. Li, D. Shi, and W.-S. Gan, "On algorithms and implementations of a 4-channel active noise canceling window," in *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. IEEE, 2017, pp. 217–221.
- [11] B. Kwon and Y. Park, "Interior noise control with an active window system," *Applied Acoustics*, vol. 74, no. 5, pp. 647–652, 2013.
- [12] T. Murao and M. Nishimura, "Basic study on active acoustic shielding," *Journal of Environment and Engineering*, vol. 7, no. 1, pp. 76–91, 2012.
- [13] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 1, pp. 107–116, 2008.
- [14] J. Donley, C. Ritz, and W. B. Kleijn, "Active speech control using wave-domain processing with a linear wall of dipole secondary sources," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 456–460.
- [15] W. Jin and W. B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2343–2355, 2015.
- [16] J. Donley, C. Ritz, and W. B. Kleijn, "On the comparison of two room compensation/dereverberation methods employing active acoustic boundary absorption," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 221–225.
- [17] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial sound with loudspeakers and its perception: A review of the current state," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 1920–1938, 2013.
- [18] L. P. Miller, "An analysis of acoustic beam-forming with sparse transducer arrays for active control," 2018, mSc. Thesis, The Pennsylvania State University, The Graduate School, College of Engineering.
- [19] K. Downey, "Loudspeaker array and testing facilities for performing large volume active noise cancelling measurements," 2020, mSc. Thesis, The Pennsylvania State University, The Graduate School, College of Engineering.
- [20] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*. Elsevier, 1999.
- [21] F. J. Fahy, *Foundations of engineering acoustics*. Elsevier, 2000.
- [22] A. D. Pierce, *Acoustics: an introduction to its physical principles and applications*. Springer, 2019.
- [23] J. Fliege and U. Maier, "The distribution of points on the sphere and corresponding cubature formulae," *IMA Journal of Numerical Analysis*, vol. 19, no. 2, pp. 317–334, 1999.
- [24] G. H. Golub and C. F. van Loan, *Matrix Computation*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.
- [25] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Transactions on Signal processing*, vol. 55, no. 6, pp. 2542–2556, 2007.
- [26] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing*. Cambridge University Press, 2014.
- [27] R. H. Shumway, D. S. Stoffer, and D. S. Stoffer, *Time series analysis and its applications*. Springer, 2000, vol. 3.
- [28] S. J. Elliott and P. A. Nelson, "Active noise control," *IEEE signal processing magazine*, vol. 10, no. 4, pp. 12–35, 1993.
- [29] C. H. Hansen, S. D. Snyder, X. Qiu, L. A. Brooks, and D. J. Moreau, *Active control of noise and vibration*. Spon London, 1997.
- [30] B. Lam, C. Shi, D. Shi, and W.-S. Gan, "Active control of sound through full-sized open windows," *Building and Environment*, vol. 141, pp. 16–27, 2018.
- [31] S. Elliott, I. Stothers, and P. Nelson, "A multiple error lms algorithm and its application to the active control of sound and vibration," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1423–1434, 1987.

Appendix B

Incoming Plane Wave Modeling

The first part of the system is the outside soundfield created by the primary source. For making proper animations, we model this. In this section, we cover the Acoustic Transfer Function (ATF) for this. As discussed, we assume that the noise source is far away, and, therefore, result in a plane wave on the aperture surface. Hence, we model this wave as a simple plane wave. A plane wave is simply a time delay, due to the distance between the plane of the primary source and the measurement location. The corresponding frequency response transfer function is:

$$H_{plane}(k, r) = e^{-jkr}, \quad (\text{B-1})$$

depending on acoustic wavenumber $k = \omega/c$, where $\omega = 2\pi f$ is the angular frequency and c is the speed of sound. Furthermore we have j the imaginary number and r the distance between the plane of the primary source (S) and the measurement location (M). In the model, we specify the primary source location as a single point. However, because we model a plane wave, we need a geometric relation to determine the perpendicular distance r , as shown in Figure B-1. We derive two equations for b . The first is:

$$\begin{aligned} a^2 + b^2 &= p^2 \\ (l - r)^2 + b^2 &= p^2 \\ b^2 &= p^2 - (l - r)^2 \end{aligned} \quad (\text{B-2})$$

And the second is:

$$\begin{aligned} r^2 + b^2 &= q^2 \\ b^2 &= q^2 - r^2 \end{aligned} \quad (\text{B-3})$$

combining gives the following expression for r :

$$\begin{aligned}
 p^2 - (l - r)^2 &= q^2 - r^2 \\
 p^2 - l^2 + 2lr - r^2 &= q^2 - r^2 \\
 p^2 - l^2 + 2lr &= q^2 \\
 2lr &= l^2 + q^2 - p^2 \\
 r &= \frac{l^2 + q^2 - p^2}{2l}
 \end{aligned}$$

(B-4)

where l is the distance between the primary noise source and the origin, q is the distance between the primary noise source and the measure location and p is the distance between the measure location and the origin.

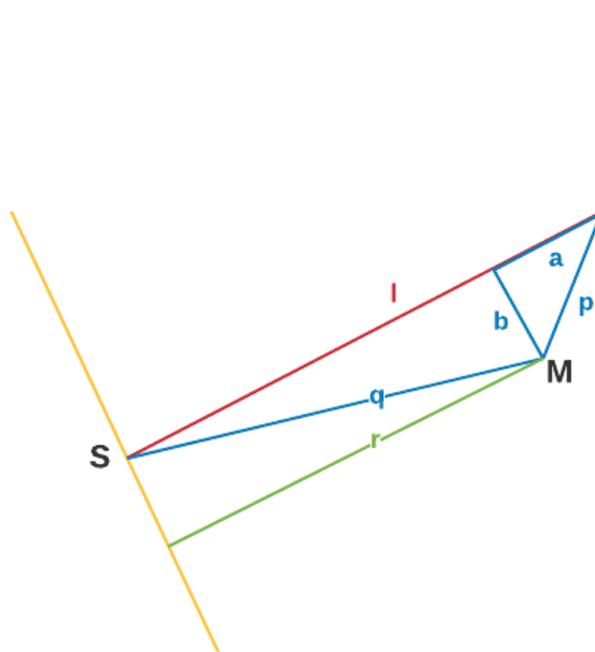


Figure B-1: The perpendicular distance between the measure location and the primary source location plane is denoted with r and can be calculated by $r = \frac{l^2 + q^2 - p^2}{2l}$.

Extended Soundfield Modeling Results

This Appendix contains additional plots of aperture soundfields for various frequencies and incident angles.

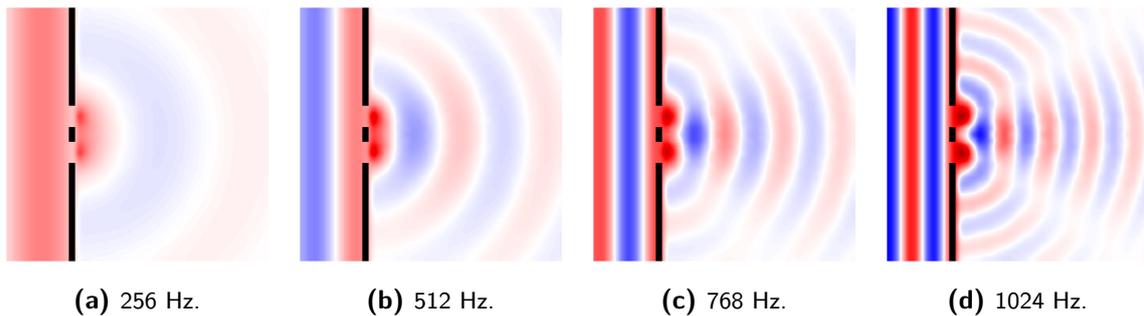


Figure C-1: Aperture soundfields for various frequencies at 0° incident angle.

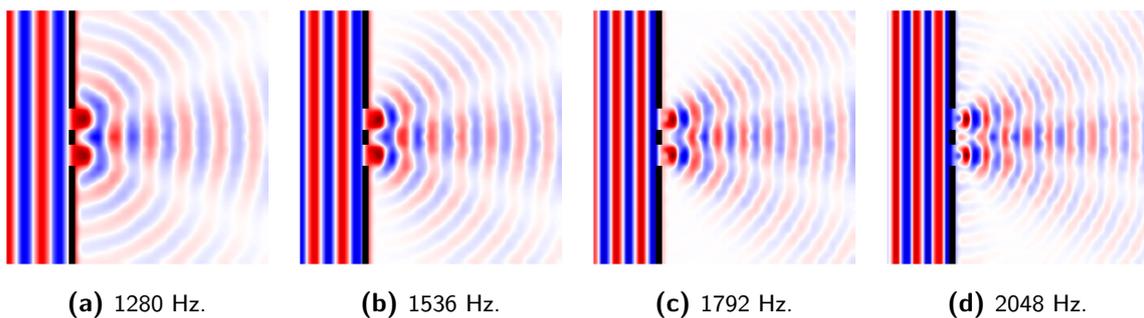


Figure C-2: Aperture soundfields for various frequencies at 0° incident angle.

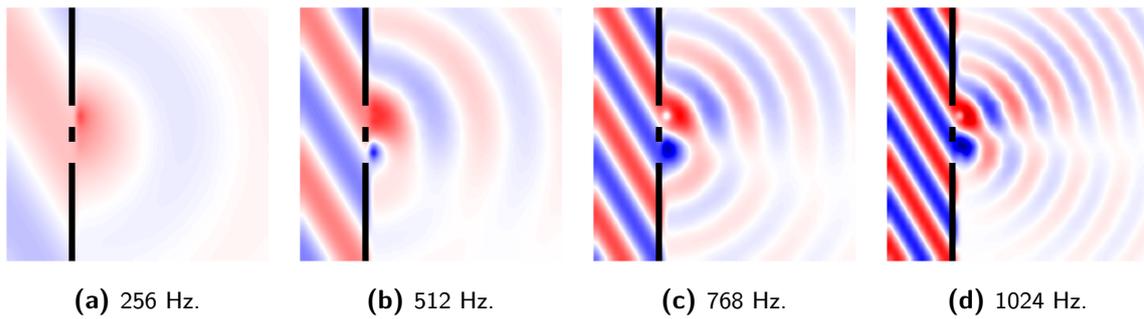


Figure C-3: Aperture soundfields for various frequencies at 30° incident angle.

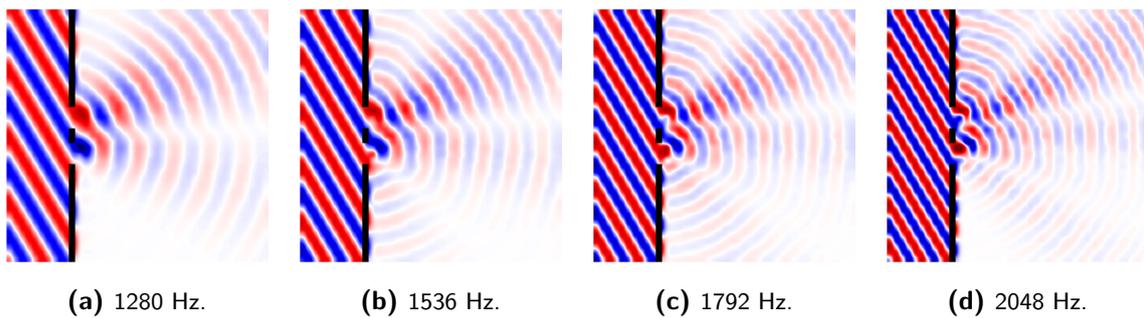


Figure C-4: Aperture soundfields for various frequencies at 30° incident angle.

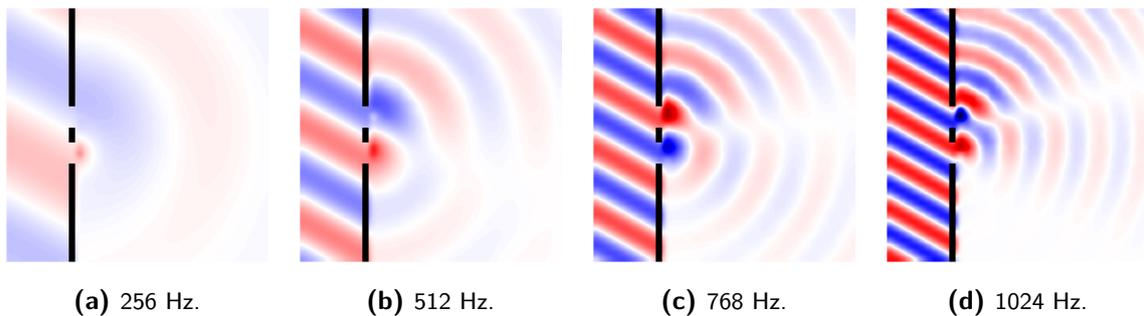


Figure C-5: Aperture soundfields for various frequencies at 60° incident angle.

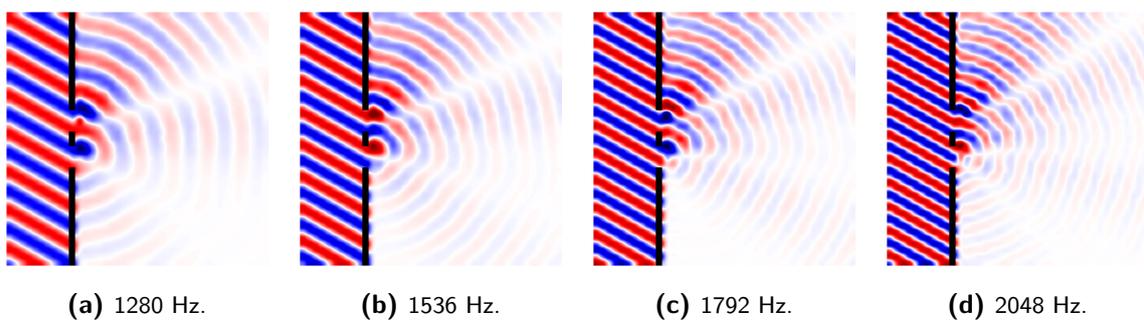


Figure C-6: Aperture soundfields for various frequencies at 60° incident angle.

Additional Simulation Results

This Appendix contains some additional results.

D-1 Other 3D Array Types

We briefly looked into other array types to measure performance. Figure D-1 shows the schematics of the arrays. The results of the simulation are shown in Figure D-2, where the array performances are compared for a white noise signal, with 0° angle and a window-size of $N=7.8$ ms. Clearly, the grid outperformed the other arrays, in line with results in literature. The Miller sparse array [13], with more speakers around the boundary does outperformed the Sparse array by Downey [14]. Especially at higher frequencies (0.6-1.5 kHz), more loudspeakers led to more attenuation. Finally, adding four loudspeakers in a circle did a slightly better job than the Circle Downey array.

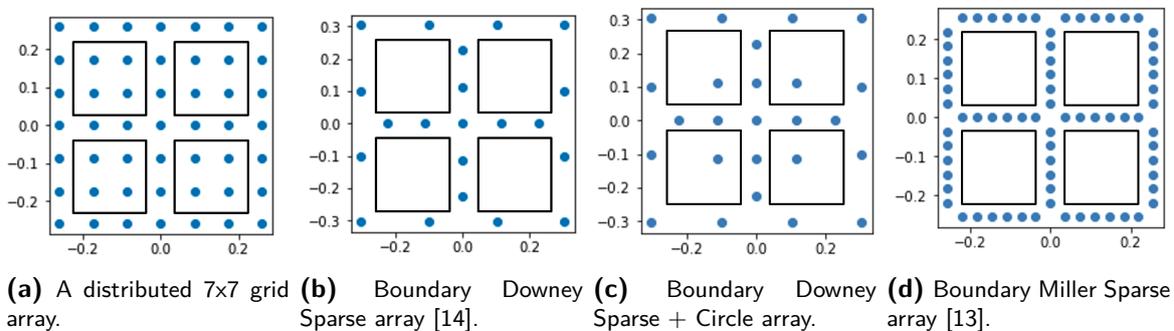


Figure D-1: Four arrays tested in 3D simulations.

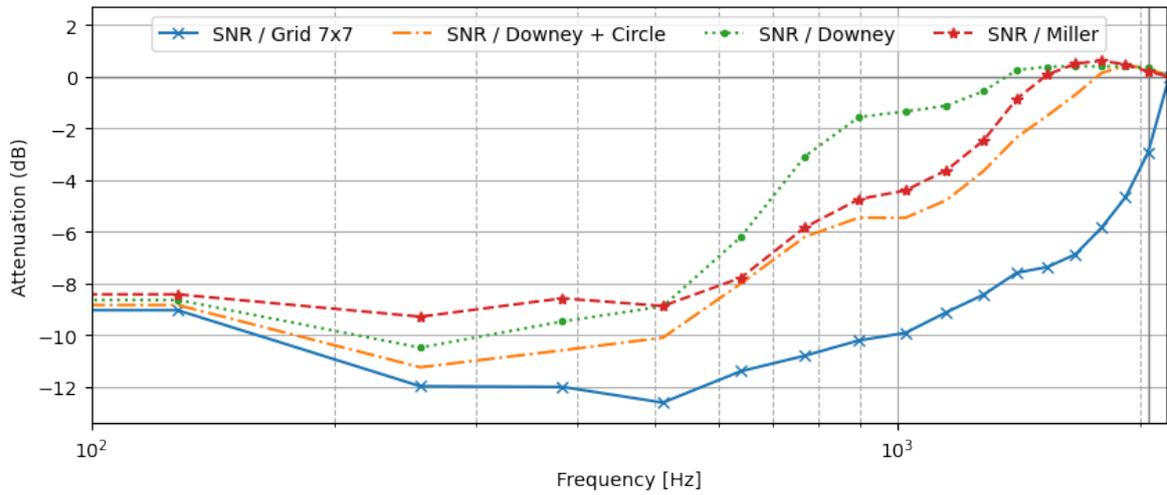


Figure D-2: 3D results for 4 types of arrays, the sparse array, sparse array by Downey with 4 additional speakers, sparse array by Miller, and the 7x7 grid array. Simulations with white noise, 0° incident angle and a window-size of N=7.8 ms.

D-2 2D Results

This section contains the results for 2D simulations in a similar fashion as the 3D results in Chapter 5.

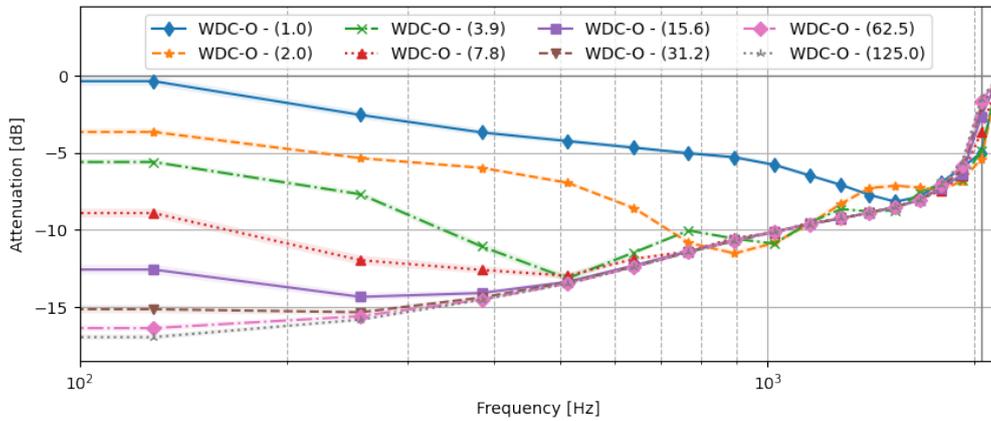


Figure D-3: 2D Grid array SEG_f results for various window-sizes.

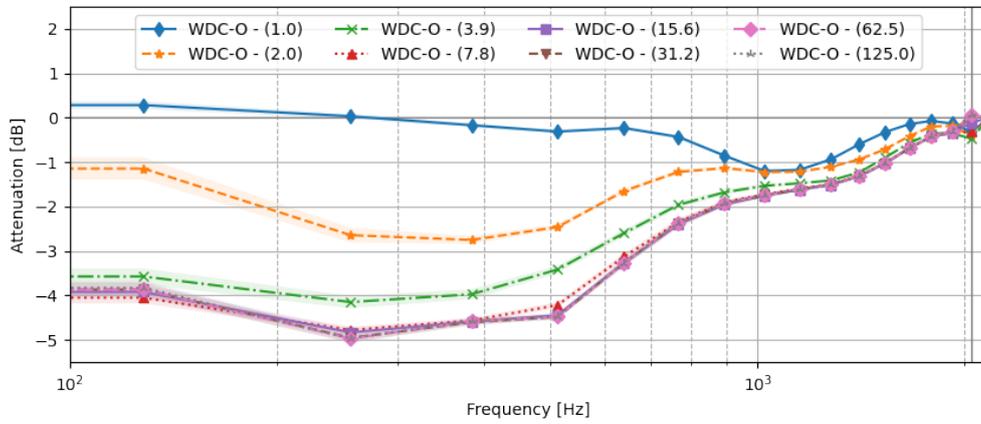


Figure D-4: 2D Sparse array SEG_f results for various window-sizes.

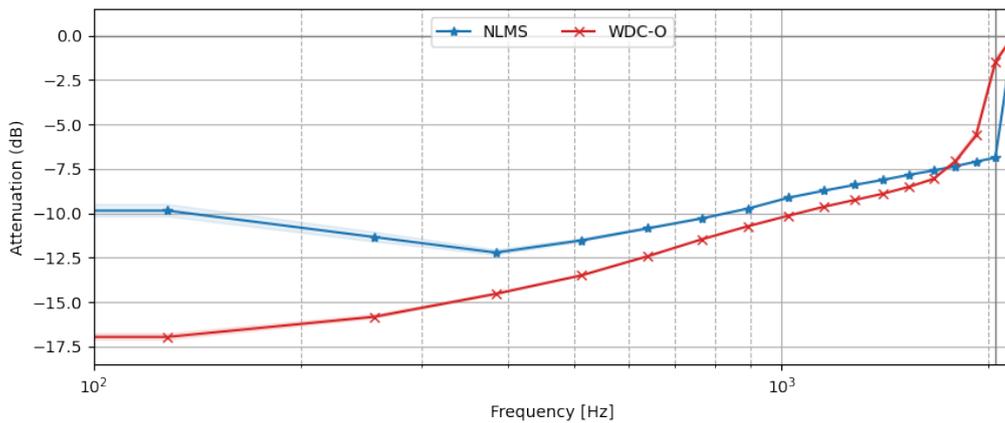


Figure D-5: 2D Grid array SEG_f results for the Normalized Least Mean Squares (NLMS) and Optimal Wave-Domain Algorithm (WDC-O) algorithms.

Table D-1: 2D Results for rumbler-siren noise at 0° angle, in decibels. Given over four frequency ranges, for all controllers.

Array type	Grid					Sparse				
	0-0.5	0.5-1	1-1.5	1.5-2	Total	0-0.5	0.5-1	1-1.5	1.5-2	Total
NLMS	-11.5	-5.2	-4.1	-0.4	-5.7	-8.4	-1.8	-0.2	0.0	-2.9
WDC-M	-7.4	-7.6	-6.3	-5.8	-6.8	-3.6	-0.3	0.8	0.1	-0.9
WDC-P	-7.4	-5.4	-2.0	-0.6	-4.0	-3.3	-1.4	-0.3	0.0	-1.4
WDC-O	-16.3	-11.6	-9.4	-5.7	-11.1	-5.0	-2.6	-1.5	-0.4	-2.5

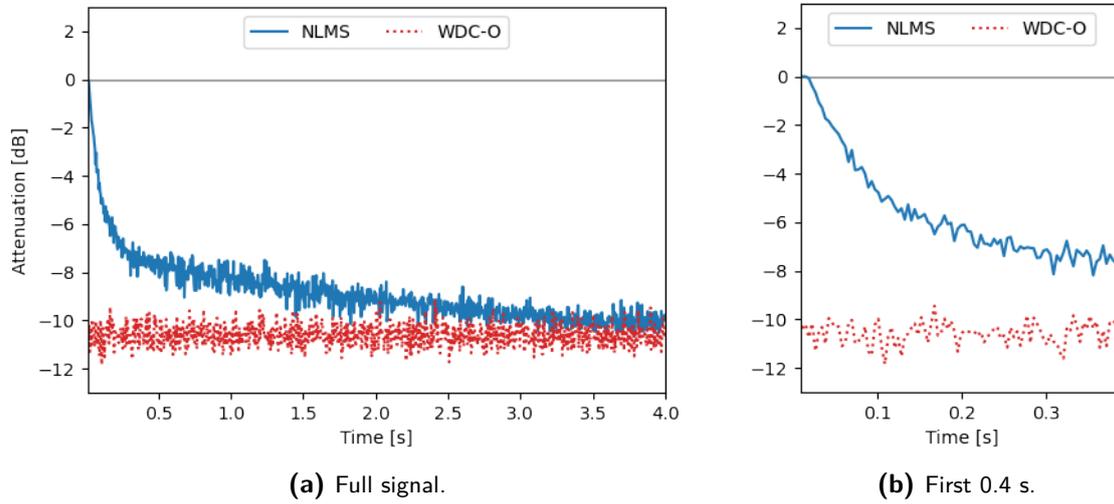


Figure D-6: 2D Grid time results for the NLMS and WDC-O algorithms.

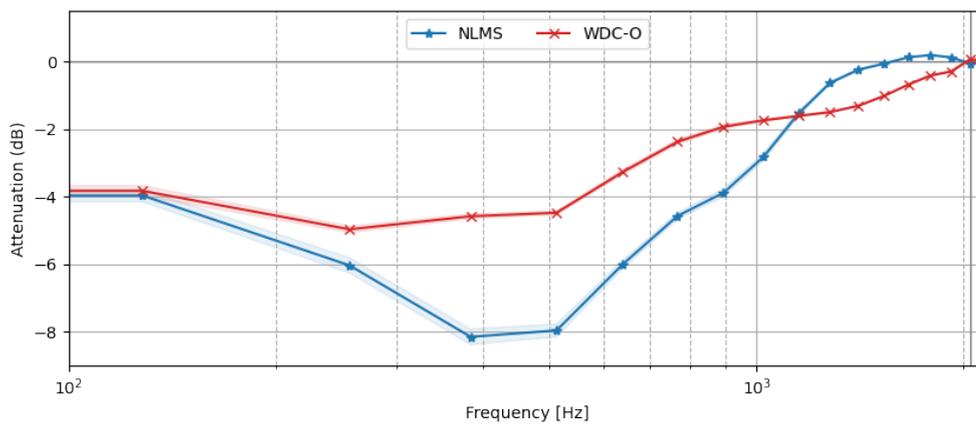


Figure D-7: 2D Sparse array SEG_f results for the NLMS and WDC-O algorithms.

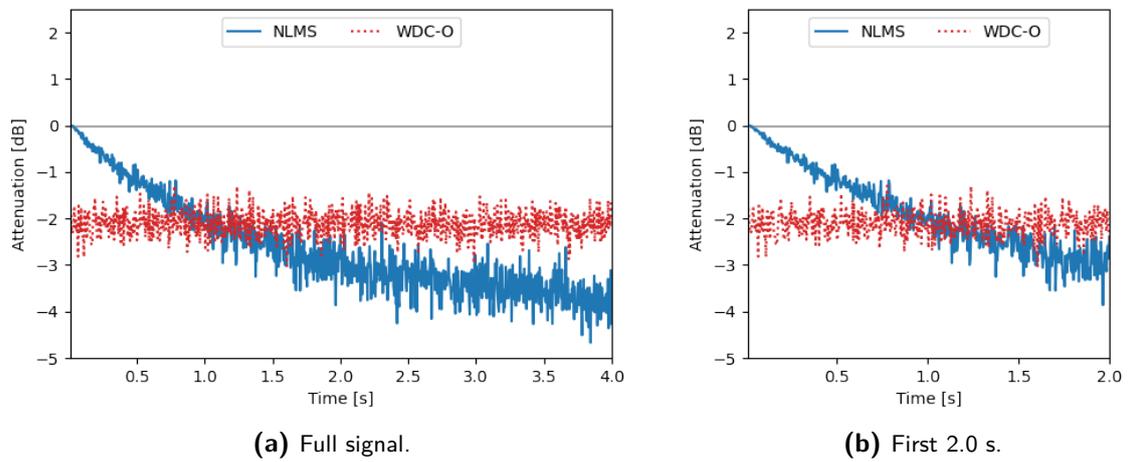


Figure D-8: 2D Sparse time results for the NLMS and WDC-O algorithms.

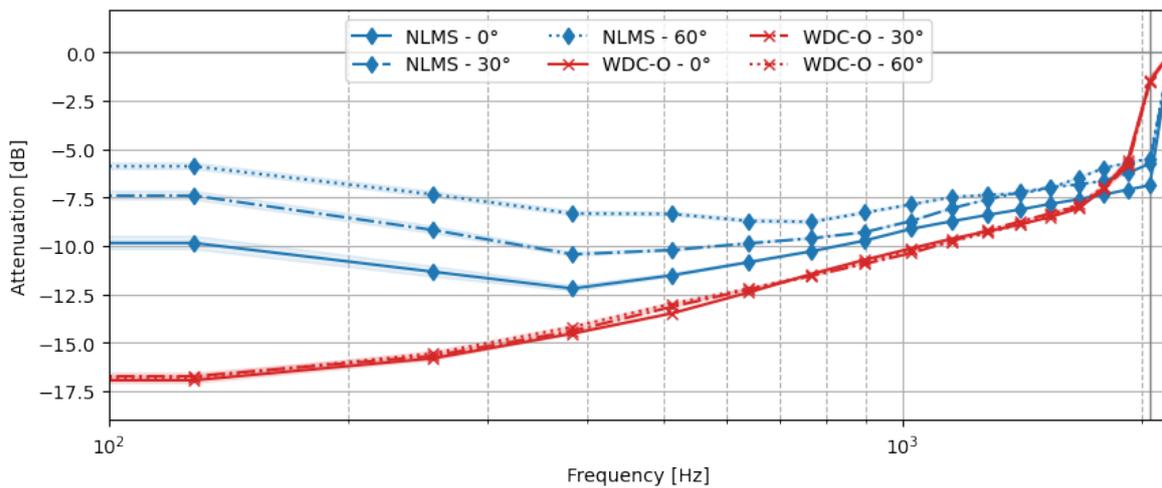


Figure D-9: 2D Grid array SEG_f results for WDC-O and NLMS for three different incident angles.

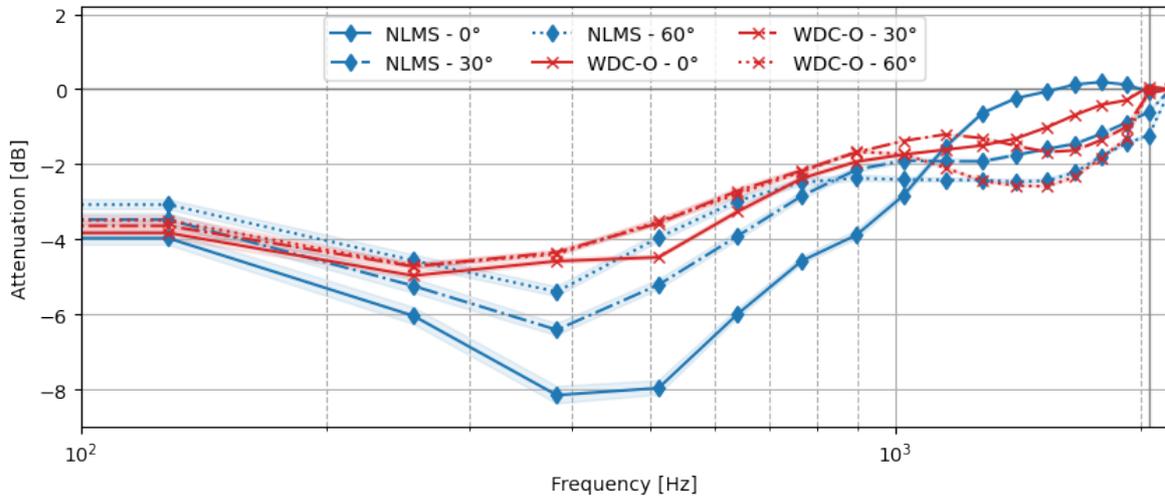


Figure D-10: 2D Sparse array SEG_f results for WDC-O and NLMS for three different incident angles.

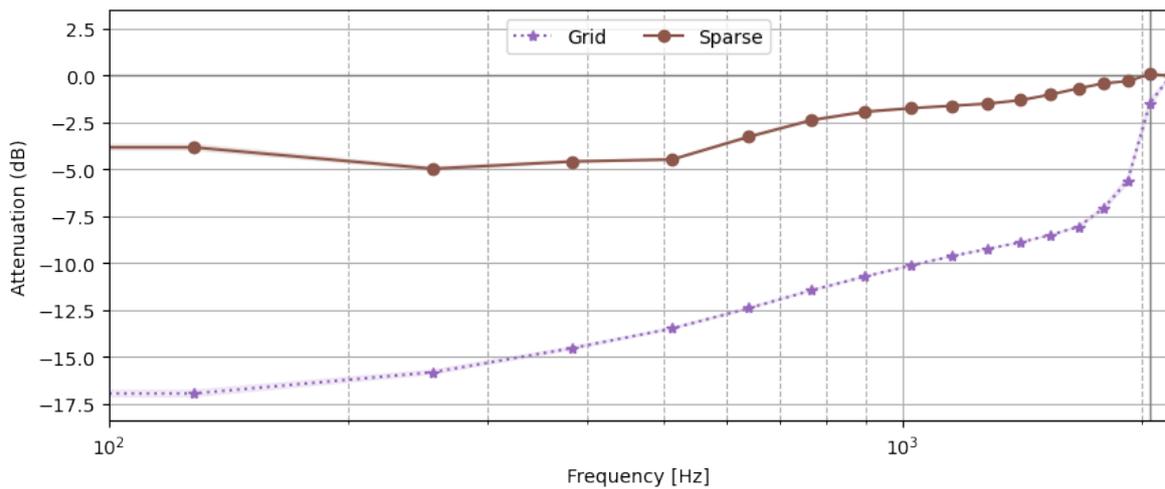


Figure D-11: 2D Grid and sparse array SEG_f results for WDC-O, white noise, 0° incident angle.

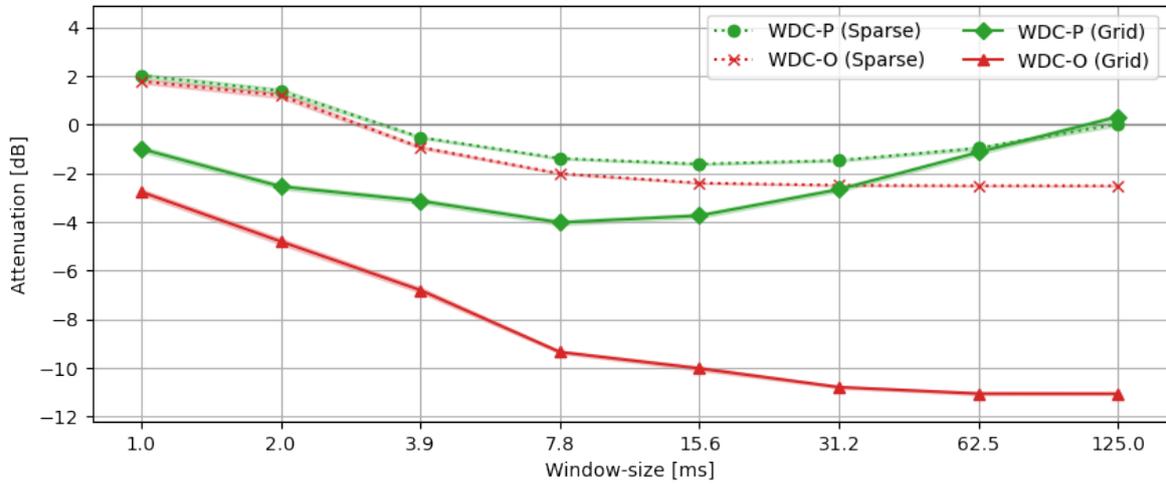


Figure D-12: Optimal window-size for 2D setup. The optimal point for the Wave-Domain Algorithm with Predictor (WDC-P) (Sparse) lies at 15.6 ms, while the optimal window-size for the grid array is 7.8 ms.

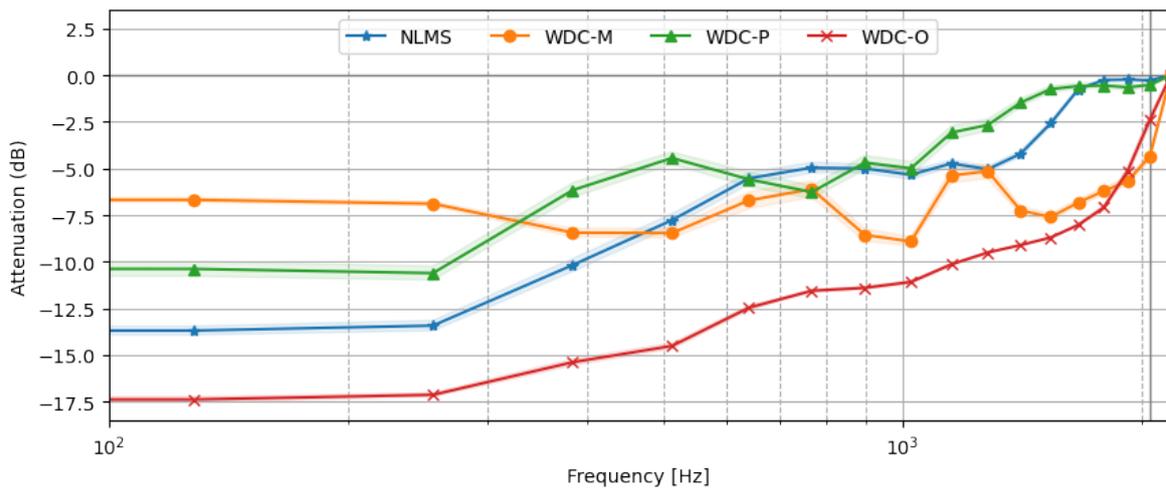


Figure D-13: 2D Grid array SEG_f results for all controllers, 0° angle and rumbler-siren noise.

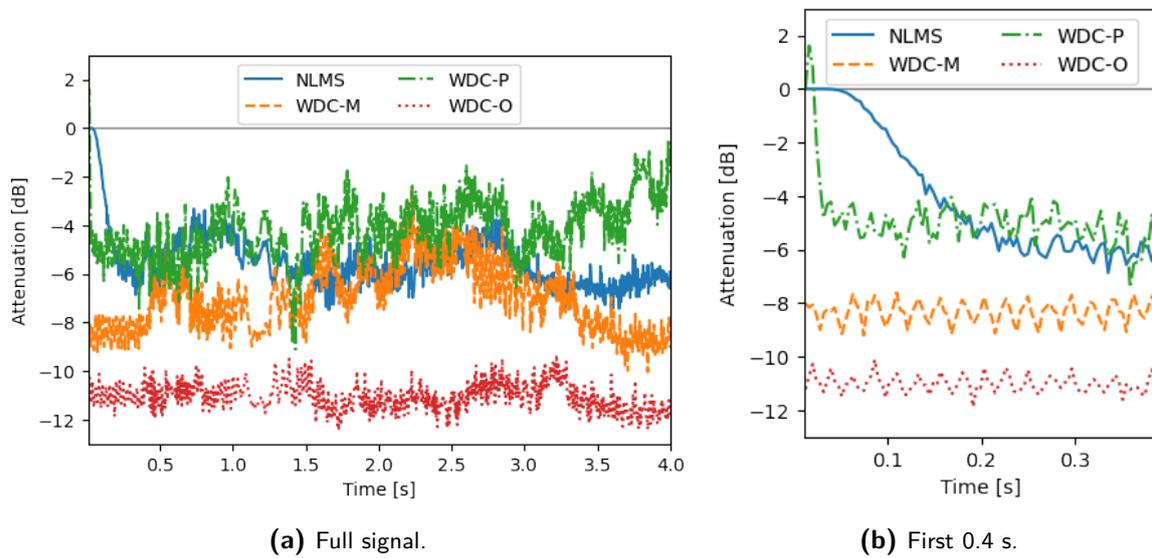


Figure D-14: 2D Grid time results for all controllers, 0° angle and rumbler-siren noise.

Bibliography

- [1] W. H. Organization *et al.*, *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*. World Health Organization. Regional Office for Europe, 2011.
- [2] L. Bhan and G. Woon-Seng, “Active acoustic windows: Towards a quieter home,” *IEEE potentials*, vol. 35, no. 1, pp. 11–18, 2016.
- [3] S. M. Kuo, S. Mitra, and W.-S. Gan, “Active noise control system for headphone applications,” *IEEE Transactions on Control Systems Technology*, vol. 14, no. 2, pp. 331–335, 2006.
- [4] S. J. Elliott, “Down with noise (active noise control),” *IEEE spectrum*, vol. 36, no. 6, pp. 54–61, 1999.
- [5] S. Wang, J. Yu, X. Qiu, M. Pawelczyk, A. Shaid, and L. Wang, “Active sound radiation control with secondary sources at the edge of the opening,” *Applied Acoustics*, vol. 117, pp. 173–179, 2017.
- [6] B. Lam, S. Elliott, J. Cheer, and W.-S. Gan, “Physical limits on the performance of active noise control through open windows,” *Applied Acoustics*, vol. 137, pp. 9–17, 2018.
- [7] S. M. Kuo and D. R. Morgan, “Active noise control: a tutorial review,” *Proceedings of the IEEE*, vol. 87, no. 6, pp. 943–973, 1999.
- [8] L. Bhan, T. Murao, C. Shi, W.-S. Gan, and S. Elliott, “Feasibility of the full-rank fixed-filter approach in the active control of noise through open windows,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 253, no. 5. Institute of Noise Control Engineering, 2016, pp. 3548–3555.
- [9] C. Shi, T. Murao, D. Shi, B. Lam, and W.-S. Gan, “Open loop active control of noise through open windows,” in *Proceedings of Meetings on Acoustics 172ASA*, vol. 29, no. 1. Acoustical Society of America, 2016, p. 030007.

- [10] C. Shi, N. Jiang, H. Li, D. Shi, and W.-S. Gan, "On algorithms and implementations of a 4-channel active noise canceling window," in *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. IEEE, 2017, pp. 217–221.
- [11] B. Lam, C. Shi, and W. S. Gan, "Active noise control systems for open windows: Current updates and future perspectives," in *Proceedings of the 24th International Congress on Sound and Vibration*, 2017, pp. 1–7.
- [12] D. Ratering, "Current research on active noise control for open windows," 2021, MSc. Thesis, Delft University of Technology.
- [13] L. P. Miller, "An analysis of acoustic beam-forming with sparse transducer arrays for active control," 2018, MSc. Thesis, The Pennsylvania State University, The Graduate School, College of Engineering.
- [14] K. Downey, "Loudspeaker array and testing facilities for performing large volume active noise cancelling measurements," 2020, MSc. Thesis, The Pennsylvania State University, The Graduate School, College of Engineering.
- [15] C. N. Hansen, *Understanding active noise cancellation*. CRC Press, 2002.
- [16] L. Paul, "Process of silencing sound oscillations," Jun. 9 1936, US Patent 2,043,416.
- [17] S. J. Elliott and P. A. Nelson, "Active noise control," *IEEE signal processing magazine*, vol. 10, no. 4, pp. 12–35, 1993.
- [18] Gracey and Associates, "Sound fields and related terms," <http://www.acoustic-glossary.co.uk/sound-fields.htm>, 2020.
- [19] J. Zhang, T. D. Abhayapala, W. Zhang, P. N. Samarasinghe, and S. Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 26, no. 4, pp. 774–786, 2018.
- [20] T. Murao and M. Nishimura, "Basic study on active acoustic shielding," *Journal of Environment and Engineering*, vol. 7, no. 1, pp. 76–91, 2012.
- [21] B. Kwon and Y. Park, "Interior noise control with an active window system," *Applied Acoustics*, vol. 74, no. 5, pp. 647–652, 2013.
- [22] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 1, pp. 107–116, 2008.
- [23] W. Jin and W. B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2343–2355, 2015.
- [24] J. Donley, C. Ritz, and W. B. Kleijn, "Active speech control using wave-domain processing with a linear wall of dipole secondary sources," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 456–460.

-
- [25] —, “On the comparison of two room compensation/dereverberation methods employing active acoustic boundary absorption,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 221–225.
- [26] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, “Spatial sound with loudspeakers and its perception: A review of the current state,” *Proceedings of the IEEE*, vol. 101, no. 9, pp. 1920–1938, 2013.
- [27] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*. Elsevier, 1999.
- [28] W. Kleijn, “A basis for source coding, 2013,” *KTH lecture notes*.
- [29] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing*. Cambridge University Press, 2014.
- [30] M. Verhaegen and V. Verdult, *Filtering and system identification: a least squares approach*. Cambridge university press, 2007.
- [31] R. N. Bracewell and R. N. Bracewell, *The Fourier transform and its applications*. McGraw-Hill New York, 1986, vol. 31999.
- [32] M. Müller, *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer, 2015.
- [33] F. J. Harris, “On the use of windows for harmonic analysis with the discrete fourier transform,” *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.
- [34] I. Bayram, “The short time fourier transform.” [Online]. Available: https://ilkerbayram.github.io/STFT_notes.html
- [35] F. J. Fahy, *Foundations of engineering acoustics*. Elsevier, 2000.
- [36] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.
- [37] S. Elliott, *Signal processing for active control*. Elsevier, 2000.
- [38] S. Elliott, I. Stothers, and P. Nelson, “A multiple error lms algorithm and its application to the active control of sound and vibration,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1423–1434, 1987.
- [39] H. Liu and E. Milios, “Acoustic positioning using multiple microphone arrays,” *The Journal of the Acoustical Society of America*, vol. 117, no. 5, pp. 2772–2782, 2005.
- [40] T. Nishiura, T. Yamada, S. Nakamura, and K. Shikano, “Localization of multiple sound sources based on a csp analysis with a microphone array,” in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*, vol. 2. IEEE, 2000, pp. II1053–II1056.
- [41] J. Fliege and U. Maier, “The distribution of points on the sphere and corresponding cubature formulae,” *IMA Journal of Numerical Analysis*, vol. 19, no. 2, pp. 317–334, 1999.

- [42] Y.-H. Kim, *Sound propagation: an impedance based approach*. John Wiley & Sons, 2010.
- [43] A. D. Pierce, *Acoustics: an introduction to its physical principles and applications*. Springer, 2019.
- [44] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.
- [45] B. Lam, C. Shi, D. Shi, and W.-S. Gan, “Active control of sound through full-sized open windows,” *Building and Environment*, vol. 141, pp. 16–27, 2018.
- [46] J. Tao, S. Wang, X. Qiu, and J. Pan, “Performance of an independent planar virtual sound barrier at the opening of a rectangular enclosure,” *Applied Acoustics*, vol. 105, pp. 215–223, 2016.
- [47] G. H. Golub and C. F. van Loan, *Matrix Computation*. Blatimore, MD, USA: Johns Hopkins Univ. Press, 1996.
- [48] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, “Intrinsic limits of dimensionality and richness in random multipath fields,” *IEEE Transactions on Signal processing*, vol. 55, no. 6, pp. 2542–2556, 2007.
- [49] G. Van Rossum and F. L. Drake Jr, *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.
- [50] Microsoft Corporation, “Visual studio code.” [Online]. Available: <https://code.visualstudio.com>
- [51] MATLAB, *version 9.8.0 (R2020a)*. Natick, Massachusetts: The MathWorks Inc., 2020.
- [52] B. Keinert, M. Innmann, M. Sanger, and M. Stamminger, “Spherical fibonacci mapping,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–7, 2015.
- [53] Universite Gustave Eiffel, Environmental Acoustics Research Unit (UMRAE), “I-simpa.” [Online]. Available: <https://i-simpa.univ-gustave-eiffel.fr>
- [54] S. Spors, R. Rabenstein, and J. Ahrens, “The theory of wave field synthesis revisited,” in *124th AES convention*. Citeseer, 2008, pp. 17–20.
- [55] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2764–2778, 1993.
- [56] S. Wang, J. Tao, and X. Qiu, “Controlling sound radiation through an opening with secondary loudspeakers along its boundaries,” *Scientific reports*, vol. 7, no. 1, pp. 1–6, 2017.
- [57] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, “Wavenet: A generative model for raw audio,” *arXiv preprint arXiv:1609.03499*, 2016.
- [58] D. Ratering, W. B. Kleijn, J. Gonzalez-Silva, and R. Ferrari, “Wave-domain approach for cancelling noise entering open windows,” in *Submitted to: 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022.

Glossary

List of Acronyms

DCSC	Delft Center for Systems and Control
ANC	Active Noise Control
FXLMS	Filtered-x Least Mean Squares
LMS	Least Mean Squares
NLMS	Normalized Least Mean Squares
STFT	Short-time Fourier Transform
I-STFT	Inverse Short-time Fourier Transform
AR	Autoregressive
RNN	Recurrent Neural Network
NODE	Neural Ordinary Differential Equations
DFT	Discrete Fourier Transform
I-DFT	Inverse Discrete Fourier Transform
FFT	Fast Fourier Transform
SNR	Signal-to-Noise Ratio
SISO	Single-input Single-output
MIMO	Multiple-input Multiple-output
ATF	Acoustic Transfer Function
ATFs	Acoustic Transfer Functions
FEMs	Finite Element Methods
FDMs	Finite Difference Methods
WDC-P	Wave-Domain Algorithm with Predictor
WDC-O	Optimal Wave-Domain Algorithm
WDC-M	Wave-Domain Algorithm with Reference Microphone Placement

