Background Noise Reduction in Wind Tunnels using Adaptive Noise Cancellation and Cepstral Echo Removal Techniques for Microphone Array Applications

Taylor B. Spalt

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science
in Mechanical Engineering

Christopher R. Fuller, Chairman
Thomas F. Brooks
Alfred L. Wicks

June 29, 2010
Hampton, Virginia

Keywords: Linear Array, Conventional Beamforming, Adaptive Noise Cancellation, Cepstrum, Wind Tunnel, Noise Sources, Background Noise Reduction
Background Noise Reduction in Wind Tunnels using Adaptive Noise Cancellation and Cepstral Echo Removal Techniques for Microphone Array Applications

Taylor B. Spalt

ABSTRACT

Two experiments were conducted to investigate Adaptive Noise Cancelling and Cepstrum echo removal post-processing techniques on acoustic data from a linear microphone array in an anechoic chamber. A point source speaker driven with white noise was used as the primary signal. The first experiment included a background speaker to provide interference noise at three different Signal-to-Noise Ratios to simulate noise propagating down a wind tunnel circuit. The second experiment contained only the primary source and the wedges were removed from the floor to simulate reflections found in a wind tunnel environment.

The techniques were applicable to both signal microphone and array analysis. The Adaptive Noise Cancellation proved successful in its task of removing the background noise from the microphone signals at SNRs as low as -20 dB. The recovered signals were then used for array processing. A simulation reflection case was analyzed with the Cepstral technique. Accurate removal of the reflection effects was achieved in recovering both magnitude and phase of the direct signal. Experimental data resulted in Cepstral features that caused errors in phase accuracy. A simple phase correction procedure was proposed for this data, but in general it appears that the Cepstral technique is and would be not well suited for all experimental data.
Acknowledgements

I would like to thank first Dr. Fuller for his support and encouragement during the thesis research and Dr. Brooks for all his invaluable aid in familiarizing me with the field of aeroacoustics. Without these two this work would not be possible. I would like to thank Dr. Wicks for participating as a committee member, but more importantly for making the subject of DSP interesting for a ME major.

I go on to thank others at NASA LaRC that enabled the work through assistance/guidance, were friendly to a student new to the field and area, or both: Mr. Humphreys, Mr. Smith, Mr. Plassman, Mr. Medley, Dr. Cabell, Mr. Klos, Dr. Loubeau, Dr. Farassat, Dr. Hutcheson, Mr. Marcolini, Dr. Tinetti, Mr. Kilzer, Mr. Swartzbaugh, Mr. Stead, Mr. Kuchta, Ms. Burley, Mr. Becker, Mr. Moen, and Mr. Reeves. I’d like to especially thank Dr. Schiller for his unending willingness to help; whether with theory, lab issues, a book, or career guidance, he made all the pieces come together.

Three people in particular stand out when I think of my experience in Blacksburg: Brian Goode, Gail Coe, and Cathy Hopkins. Thank you.

Although I believe everyone at NIA would lend a hand given the chance, the following people lent theirs to me: Lisa Burke, Rita Aguillard, Larry Battle, Dr. Grossman, and Lara Hawthorne.

My sincere thanks go out to Bart for being a role model.

Thanks to all of my friends in Hampton Roads: Scott, Alex, Duncan, Steve, Rob, Kerri, Mike, Charlie, Ryan, Steve, Herber, and Paolo; thank you for making graduate student life that much better.

Thank you Aimy for being selfless.

Thank you Brendan for being a brother.

Much appreciation goes to my family for supporting my every move with love and beyond.

Lastly, I recognize my mother for being more than a son could ever wish for.
Table of Contents

Acknowledgements ............................................................................................................. iii
Index of Figures ................................................................................................................ vi
Index of Tables .................................................................................................................. x
Nomenclature ................................................................................................................... xi

1 INTRODUCTION ............................................................................................................. 1
  1.1 LIMITATIONS OF CURRENT AEROCOUSTIC EXPERIMENTS .................................................. 3
  1.2 RESEARCH OBJECTIVES .................................................................................................. 5
  1.3 SUMMARY OF TECHNIQUES IN WIND TUNNEL BACKGROUND NOISE REDUCTION ................. 6
     Conventional Beamforming .......................................................................................... 6
     Adaptive Noise Cancellation ....................................................................................... 6
     Cepstral Echo Removal .............................................................................................. 7

2 LITERATURE REVIEW .................................................................................................... 8
  2.1 BEAMFORMING ........................................................................................................... 8
  2.2 ADAPTIVE NOISE CANCELLATION .............................................................................. 12
  2.3 CEPSRUM ECHO REMOVAL TECHNIQUES ................................................................ 14

3 THEORY .......................................................................................................................... 17
  3.1 DIRECTIONAL MICROPHONE ARRAY ....................................................................... 17
     Sound Source Concept ............................................................................................... 18
     Array Response .......................................................................................................... 18
     Beamforming .............................................................................................................. 20
  3.2 ADAPTIVE NOISE CANCELLATION ............................................................................ 22
  3.3 CEPSRUM REFLECTION REMOVAL .......................................................................... 26
     Power Cepstrum ........................................................................................................... 27
     Complex Cepstrum ....................................................................................................... 29
     Distortion in the Complex Cepstrum ......................................................................... 30
     Noise in the Complex Cepstrum .................................................................................. 31
     Distortion and Noise in the Complex Cepstrum .......................................................... 31
     Appending Zeros to the Time History ....................................................................... 32
     Data Averaging ............................................................................................................. 32
     Test Signal Considerations ....................................................................................... 33
     Filtering in the Cepstrum Domain ............................................................................. 33

4 EXPERIMENTAL SETUP .................................................................................................. 35
  4.1 FACILITY ..................................................................................................................... 35
  4.2 INSTRUMENTATION AND DATA ACQUISITION .............................................................. 36
     Instrumentation ............................................................................................................ 36
     Microphone Array ........................................................................................................ 37
     Data Acquisition .......................................................................................................... 39
  4.3 EXPERIMENTAL PROCEDURE ..................................................................................... 40
     Adaptive Noise Cancellation (ANC) ............................................................................ 40
     Cepstrum Reflection Removal .................................................................................... 45
5 EXPERIMENTAL RESULTS ................................................................. 49

5.1 ADAPTIVE NOISE CANCELLATION (ANC) ......................................................... 49

  SNR -6.5 dB .................................................................................. 57
  SNR -11.4 dB .................................................................................. 59
  SNR -19.9 dB .................................................................................. 61

  SNR -19.9 dB – Variations ................................................................. 62

5.2 CEPSTRUM REFLECTION REMOVAL ............................................................... 69

  Simulation Results ............................................................................ 69

  Experimental Results ........................................................................ 76

    First Liftering Scheme: Block Lifting, One Point Zeroed ....................... 83
    Second Liftering Scheme: Block Lifting, Four Points Zeroed .................... 86
    Third Liftering Scheme: Block Lifting, One Hundred Points Zeroed ........... 89
    Fourth Liftering Scheme: Interpolation ............................................... 92

6 CONCLUSIONS AND FUTURE WORK .......................................................... 97

  6.1 CONCLUSIONS ........................................................................ 97
  6.2 FUTURE WORK ..................................................................... 99

APPENDIX A: DAMAS METHODOLOGY ......................................................... 100

APPENDIX B: EXPERIMENT HARDWARE ..................................................... 102

BIBLIOGRAPHY .................................................................................. 103
Index of Figures

Figure 1.1 - Plane flying low over a neighborhood near Heathrow Airport [1]..............................1
Figure 1.2 - Noise source locations superimposed on a flyover shot of a turbojet [3]. 1) Front landing gear noise, 2) Trailing-edge flap noise, 3) Rear landing gear noise, and 4) engine noise. .................................................................2
Figure 1.3 - X-48C test in NASA Langley’s Full Scale Wind Tunnel [4]. ........................................2
Figure 1.4 – Microphone array examples. (a) Linear microphone array (University of Kentucky Audio Systems Lab) [6], and (b) Phased array (Virginia Tech Vibrations and Acoustics Lab) [7]. Microphones are flush mounted in a spiral pattern to an aluminum plate. ..........................................................................................3
Figure 1.5 – KAIST Aeroacoustic Lab’s anechoic wind tunnel [9]. The test section (center) is open and the walls surrounding it have been acoustically treated with anechoic wedges to absorb sound. ........................................................................5
Figure 2.1 - Test setup for rotor noise investigation [22]. ...............................................................9
Figure 2.2 - Beamforming plot for 6% scale model of a main element NACA 63:215 wing section with a 30% chord half-span flap; mach 0.17, angle-of-attack 16°, flap at 39°, and frequency 12.5 kHz. Lines seen around center point denote the flap-edge. Taken from [25]. .................................................................................................................10
Figure 2.3 – Beamform results for a flat-edge flap; Mach 0.11, one-third octave frequency 40 kHz. Integration area shown with dotted line. Taken from [17].................................................................11
Figure 2.4 - Example figures from [26]. (a) Synthetic point source simulation; frequency 20 kHz, source “placed” 5’ from array, source strength set to 100 dB, (b) Same as (a) with multiple synthetic sources placed to form the acronym NASA; frequency 30 kHz, and (c) NACA 63-216 airfoil in Leading and Trailing Edge test (16” chord, 36” span, -1.2° angle-of-attack to vertical flow); one-third octave frequency 20 kHz.....12
Figure 2.5 - Figures from [10]. (a) Adaptive noise canceller concept, and (b) ANC principle illustrated in the cancellation of electrocardiographic noise. .................................................................13
Figure 3.1 - Directional array principle [25]. ..................................................................................17
Figure 3.2 - Array/Source geometry relating sound propagation to array microphones from a monopole, omnidirectional source radiating spherically [40]. ......................................................18
Figure 3.3 - Ideal array response over a 10’ line ............................................................................19
Figure 3.4 - Example user-defined scanning grid (adapted from [40]). .........................................21
Figure 3.5 - Illustration of an example grid line for beamforming steering locations superimposed on a sound source. ...................................................................................................................21
Figure 3.6 - Adaptive noise canceller concept (adapted from [10]). .............................................23
Figure 3.7 - Computer rendering of hyperparaboloidal surface [42].............................................24
Figure 3.8 - Adaptive noise canceller concept with LMS algorithm and FIR Filter block diagrams included. Adapted from [43]. ..................................................................................................25
Figure 3.9 - FIR Filter block diagram. ............................................................................................26
Figure 3.10 - Power Cepstrum sequence [44]. .................................................................27
Figure 3.11 - Power Cepstrum for an impulse function and its echo [44]. ....................28
Figure 3.12 - Complex Cepstrum sequence [44]. ..........................................................29
Figure 3.13 - Complex Cepstrum for an impulse function and its echo [44]. .................30
Figure 3.14 - Example of lifting in the Power Cepstrum [35]. (a) Power Cepstrum, and (b) 
    Power Cepstrum with delta at quefrency = 0.07 seconds liftered..........................33
Figure 4.1 - Component interconnection........................................................................37
Figure 4.2 – (a) Horizontal microphone array (ANC test), (b) spacing dimensions for ANC array. 
                                                                                           ...........................................................................................................37
Figure 4.3 - Example beamform with beamwidth illustrated...........................................38
Figure 4.4 – Array sensitivities vs. frequency: (a) 1 kHz, B = 9.4', (b) 5 kHz, B = 1.4', (c) 10 
                              kHz, B = 0.8'. ........................................................................................................39
Figure 4.5 - ANC test setup with relevant dimensions. .................................................41
Figure 4.6 - ANC setup picture with background speaker and reference microphone in the near-
                              field and array in the far-field. ........................................................................41
Figure 4.7 - Source speaker with copper extension tube installed.................................42
Figure 4.8 - Block diagram of Signal-to-Noise Ratio calculation.....................................43
Figure 4.9 – Source speaker array averaged autospectrum (pink noise, BW = 0.98 Hz). ......44
Figure 4.10 – Background speaker array averaged autospectrum (white noise, BW = 0.98 Hz). 44
Figure 4.11 - (a) Vertical microphone array (Cepstrum test) and, (b) Microphone spacing for the 
    vertical array. ..............................................................................................................45
Figure 4.12 - Source speaker mounted; Cepstrum test. ....................................................46
Figure 4.13 - Cepstrum test setup diagram....................................................................46
Figure 4.14 – No reflection case, center array microphone autospectrum (white noise, BW = 
                               19.53 Hz). ........................................................................................................47
Figure 4.15 – Reflection case, center array microphone autospectrum (white noise, BW = 19.53 
                           Hz). ........................................................................................................48
Figure 5.1 – Array averaged autospectra: (a) SNR = -6.5 dB, (b) SNR = -11.4 dB, (c) SNR = - 
                          19.9 dB.................................................................51
Figure 5.2 - Center array and reference microphone autospectra; source speaker on only. .....52
Figure 5.3 - B&K free-field microphone polar plot [51]. .................................................52
Figure 5.4 - Center array and reference microphone autospectra; background speaker on only. 53
Figure 5.5 - Coherence between reference and center array microphone with background speaker 
    on. .........................................................................................................................54
Figure 5.6 - Coherence between reference and center array microphone with background and 
    source speaker both on for three SNRs. ..................................................................55
Figure 5.7 - Matlab LMS filter block. ...........................................................................56
Figure 5.8 – Recovered signal autospectrum (SNR = -6.5 dB, MSE = 2.66, w = 1024, µ = 0.1). 58
Figure 5.9 – Recovered signal beamform (5.95 kHz, SNR = -6.5 dB, MSE = 1.02, w = 1024, µ = 
                          0.1).................................................................................................59
Figure 5.10 – Recovered signal autospectrum (SNR = -11.4 dB, MSE = 3.64, w = 1024, µ = 0.1). ........................................................................................................................................60
Figure 5.11 – Recovered signal beamform (5.95 kHz, SNR = -11.4 dB, MSE = 1.50, w = 1024, µ = 0.1). ........................................................................................................................................60
Figure 5.12 – Recovered signal autospectrum (SNR = -19.9 dB, MSE = 6.61, w = 1024, µ = 0.1). ........................................................................................................................................61
Figure 5.13 – Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 4.43, w = 1024, µ = 0.1). ........................................................................................................................................62
Figure 5.14 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 16.90, w = 2048, µ = 0.1). ........................................................................................................................................63
Figure 5.15 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 7.70, w = 2048, µ = 0.1). ........................................................................................................................................64
Figure 5.16 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 23.25, w = 128, µ = 0.1). ........................................................................................................................................64
Figure 5.17 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 15.31, w = 128, µ = 0.1). ........................................................................................................................................65
Figure 5.18 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 33.41, w = 1024, µ = 0.01). ........................................................................................................................................66
Figure 5.19 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 34.19, w = 1024, µ = 0.01). ........................................................................................................................................66
Figure 5.20 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 52.59, w = 1024, µ = 0.3). ........................................................................................................................................67
Figure 5.21 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 25.37, w = 1024, µ = 0.3). ........................................................................................................................................68
Figure 5.22 - Matlab Power Cepstrum procedure ........................................................................................................................................69
Figure 5.23 - Matlab Complex Cepstrum procedure that leads to signal recovery with echo effects removed ........................................................................................................................................70
Figure 5.24 – Simulink simulation block ........................................................................................................................................70
Figure 5.25 - Simulation time histories. (a) 'Scope' signal simulating one direct pulse, and (b) 'Scope1' signal simulating one direct pulse and one reflection pulse at a lower effects removed. ........................................................................................................................................71
Figure 5.26 – (a) Magnitude of simulated non-reflection and reflection signals, and (b) Phase ..................................................72
Figure 5.27 – Power Cepstrum of simulated reflection data. (a) Full length Cepstrum, and (b) Zoomed version of (a) to emphasize the rahmonic content with direct signal, reflection, and first rahmonic in view. ........................................................................................................................................73
Figure 5.28 – Complex Cepstrum of simulated signal. (a) Full view of entire quefency scale, and (b) Zoomed version of (a) to enlarge low quefency end ........................................................................................................................................74
Figure 5.29 – (a) Magnitude of non-reflection, reflection, and recovered signals, and (b) Phase ........................................................................................................................................74
Figure 5.30 – Center array microphone autospectrum (no reflection vs. reflection present) ........................................................................................................................................77
Figure 5.31 – Average coherence between array microphones (no reflection) ........................................................................................................................................78
Figure 5.32 - Reflection coherence for microphones 1 & 5

Figure 5.33 - Power Cepstrum of microphone 5

Figure 5.34 - Power Cepstrum of microphone 5: (a) Zoomed into low frequency end, and (b) Shifted to provide rahmonic detail

Figure 5.35 - Complex Cepstrum of microphone 5, zoomed in to low frequency end

Figure 5.36 - Microphone 5 autospectrum: no reflection, reflection, and recovered (1 point zeroed)

Figure 5.37 – Beamform: recovered data (1 point zeroed, 8682 Hz)

Figure 5.38 - Phase correction procedure

Figure 5.39 - Beamform: recovered data, phase corrected (1 point zeroed, 8682 Hz)

Figure 5.40 - Microphone 5 autospectrum: no reflection, reflection, and recovered (4 points zeroed)

Figure 5.41 - Beamform: recovered data (4 points zeroed, 8242 Hz)

Figure 5.42 - Beamform: recovered data, phase corrected (4 points zeroed, 8242 Hz)

Figure 5.43 - Microphone 5 autospectrum: no reflection, reflection, and recovered (100 points zeroed)

Figure 5.44 - Beamform: recovered data (100 points zeroed, 8740 Hz)

Figure 5.45 - Beamform: recovered data, phase corrected (100 points zeroed, 8740 Hz)

Figure 5.46 - Microphone 5 autospectrum: no reflection, reflection, and recovered (interpolation)

Figure 5.47 - Beamform: recovered data (interpolation, 8242 Hz)

Figure 5.48 - Beamform: recovered data, phase corrected (interpolation, 8242 Hz)
Index of Tables

Table 4.1 - Microphone array beamwidths vs. frequency. .................................................................38
Table 5.1 – SNR values vs. coherence in frequency band 0.5-9.5 kHz for background and source
speaker both turned on.........................................................................................................................55
Table 5.2 - Matlab LMS algorithm variables. .......................................................................................56
Table 5.3 - ANC key results. ..................................................................................................................68
Table 5.4 - Cepstrum recovery key results. .........................................................................................95
Nomenclature

Roman

\(a\)  
Echo amplitude

\(b\)  
Grid point

\(B\)  
Grid points, total

\(c\)  
Speed of sound

\(c_n\)  
Filter weight

\(C\)  
Cross Spectral Matrix element

\(\hat{C}\)  
Cross Spectral Matrix

\(C_{xy}\)  
Coherence

\(d_n\)  
Primary input

\(D\)  
Characteristic dimension

\(e\)  
Mathematical constant

\(e_n\)  
Error

\(e_n^2\)  
Mean-square error

\(E\)  
Steering vector

\(\hat{E}\)  
Steering vector matrix

\(f\)  
Frequency

\(f_n\)  
Resonance frequency at nth harmonic

\(F\)  
Fourier transform

\(F^{-1}\)  
Inverse Fourier transform

\(F_s\)  
Sampling frequency

\(G_{xx}\)  
Auto-Spectral Density, one-sided

\(G_{xy}\)  
Cross-Spectral Density, one-sided

\(h\)  
Transfer function (time domain)

\(H\)  
Transfer function, FFT of

\(j\)  
Imaginary number

\(k\)  
Block average

\(K\)  
Block averages, total

\(L\)  
Length
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>Data block sample</td>
</tr>
<tr>
<td>$M$</td>
<td>Data block samples, total</td>
</tr>
<tr>
<td>$n$</td>
<td>Array microphone; Noise (time domain); Harmonic</td>
</tr>
<tr>
<td>$\hat{n}$</td>
<td>Noise, Complex Cepstrum of</td>
</tr>
<tr>
<td>$n_0$</td>
<td>Noise present in primary channel</td>
</tr>
<tr>
<td>$n_1$</td>
<td>Noise, reference</td>
</tr>
<tr>
<td>$N$</td>
<td>Noise, FFT of</td>
</tr>
<tr>
<td>$p$</td>
<td>Pressure (time domain)</td>
</tr>
<tr>
<td>$p_0$</td>
<td>Pressure, source speaker</td>
</tr>
<tr>
<td>$P$</td>
<td>Pressure, FFT of</td>
</tr>
<tr>
<td>$r$</td>
<td>Radial distance</td>
</tr>
<tr>
<td>$R$</td>
<td>Array response</td>
</tr>
<tr>
<td>$s$</td>
<td>Sound source, primary (time domain)</td>
</tr>
<tr>
<td>$\hat{s}$</td>
<td>Direct signal, Complex Cepstrum of</td>
</tr>
<tr>
<td>$S$</td>
<td>Array response, output power spectrum of; Primary sound source, FFT of</td>
</tr>
<tr>
<td>$t$</td>
<td>Time</td>
</tr>
<tr>
<td>$t_0$</td>
<td>Time shift</td>
</tr>
<tr>
<td>$T$</td>
<td>Acquisition time</td>
</tr>
<tr>
<td>$u$</td>
<td>Input sample vector</td>
</tr>
<tr>
<td>$w$</td>
<td>Window function</td>
</tr>
<tr>
<td>$w$</td>
<td>Weight vector</td>
</tr>
<tr>
<td>$w_s$</td>
<td>Weighting constant</td>
</tr>
<tr>
<td>$x$</td>
<td>Input signal (time domain)</td>
</tr>
<tr>
<td>$\hat{x}$</td>
<td>Location in space</td>
</tr>
<tr>
<td>$\hat{x}$</td>
<td>Composite signal, Complex Cepstrum of</td>
</tr>
<tr>
<td>$x_n$</td>
<td>Reference input</td>
</tr>
<tr>
<td>$x_p$</td>
<td>Power Cepstrum</td>
</tr>
<tr>
<td>$X$</td>
<td>Input signal, FFT of</td>
</tr>
<tr>
<td>$y_n$</td>
<td>Adaptive filter output</td>
</tr>
<tr>
<td>$z$</td>
<td>ANC system output</td>
</tr>
<tr>
<td>$\nabla$</td>
<td>Error gradient</td>
</tr>
</tbody>
</table>
Greek

δ  Delta function
ε_n  ANC error
θ  Angle
λ  Eigenvalue
μ  Step size
τ  Time shift
ψ  Variable of integration
ω  Angular frequency

Subscripts

construct  Constructive interference
destruct  Destructive interference
n  Array microphone; Step; Filter weight
n'  Array microphone (different from n)
k  Data block
N  Array microphones, total; Filter weights, total
c  Array microphone, center
opt  Optimal
xx  Signal x with respect to itself
yy  Signal y with respect to itself
xy  Signal x with respect to signal y
max  Maximum

Superscripts

°  Degrees
"  Inches
'  Feet
s  Sound source
T  Complex transpose
*  Complex conjugate
### Abbreviations/Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANC</td>
<td>Adaptive Noise Cancellation</td>
</tr>
<tr>
<td>B</td>
<td>Beamwidth</td>
</tr>
<tr>
<td>BW</td>
<td>Bandwidth</td>
</tr>
<tr>
<td>Const.</td>
<td>Constant</td>
</tr>
<tr>
<td>COP</td>
<td>Coherent Output Power</td>
</tr>
<tr>
<td>CSM</td>
<td>Cross Spectral Matrix</td>
</tr>
<tr>
<td>dB</td>
<td>Decibel</td>
</tr>
<tr>
<td>DAMAS</td>
<td>Deconvolution Approach for the Mapping of Acoustic Sources</td>
</tr>
<tr>
<td>FIR</td>
<td>Finite Impulse Response</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>ID</td>
<td>Inner Diameter</td>
</tr>
<tr>
<td>LaRC</td>
<td>Langley Research Center</td>
</tr>
<tr>
<td>LADA</td>
<td>Large Aperture Directional Array</td>
</tr>
<tr>
<td>LMS</td>
<td>Least Mean Squares</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>NACA</td>
<td>National Advisory Committee for Aeronautics</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>Pa</td>
<td>Pascal</td>
</tr>
<tr>
<td>sec</td>
<td>Second</td>
</tr>
<tr>
<td>SADA</td>
<td>Small Aperture Directional Array</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>SPL</td>
<td>Sound Pressure Level</td>
</tr>
<tr>
<td>V</td>
<td>Volt</td>
</tr>
</tbody>
</table>
1 Introduction

Noise pollution from aircraft is one of the most negative impacts created by the aviation industry. The reduction in noise generated began in the 1950s, which saw the debut of the commercial turbojet engine. As unrest due to this new pollution in the form of pressure waves known as sound grew, so did the response to combat it. England and the United States were the first into action as London Heathrow and New York JFK airports imposed regulations that reduced noise pollution durations. Later on a national level, the Federal Aviation Administration standardized regulations across the country with the goal of quiet aircraft on the technological side and adjusted flying procedures on the strategic end.

Figure 1.1 - Plane flying low over a neighborhood near Heathrow Airport [1].

Today, urban areas are becoming more and more populated and the demand for air travel is increasing, which means the problem posed by noise pollution is compounding. Concerns regarding the noise restrict expansion of the industry (number/frequency of flights, airport size etc.) thus generating an economic conflict which further increases the importance of the field of noise reduction, past just that of a public health concern. And this is not even to mention the effects airports play in the real-estate markets surrounding them. Until adequately reduced, noise pollution from aircraft will remain at the forefront of issues the industry faces.

In efforts to reduce noise emitted by aircraft the specific noise sources must first be identified. With aircraft, structure-fluid interaction is a major source of sound emission. Figure 1.2 provides a shot of the underside of a jet with a noise intensity map superimposed over the sources. The field of aeroacoustics is defined as the study of noise generation via turbulent fluid
motion or aerodynamic forces interacting with surfaces [2]. Flow in aeroacoustic experiments is essential to provide realistic modeling of the noise source. For a full grasp of the noise phenomena a thorough understanding of that flow is necessary.

![Figure 1.2](image1.jpg)

**Figure 1.2** - Noise source locations superimposed on a flyover shot of a turbojet [3]. 1) Front landing gear noise, 2) Trailing-edge flap noise, 3) Rear landing gear noise, and 4) engine noise.

Aeroacoustic experiments are usually carried out in wind tunnels which have two primary benefits: a scaled model can be tested reducing by orders of magnitude the associated cost, and it is a controlled experiment and thus the flow is characterized alleviating any environmental unknowns present in flight testing. Figure 1.3 gives an example photograph of an aircraft wind tunnel test.

![Figure 1.3](image2.jpg)

**Figure 1.3** - X-48C test in NASA Langley's Full Scale Wind Tunnel [4].
1.1 Limitations of Current Aeroacoustic Experiments

Wind tunnels, as normally designed, are not optimal for acoustics testing because they are designed with fluid flow considerations in mind. The acoustic measurements of components placed for testing in the tunnels are often masked by two main factors [5]: background noise and reverberations within the test section. Noise generated by the wind tunnel flow mechanism or the flow itself constitutes background noise. It is undesirable as it introduces another source of noise in addition to the desired sound to be studied, namely that emitted from the model. The reverberant noise originates from reflective surfaces within the test section. This can be especially of concern if microphone arrays (examples given in Figure 1.4) are used as they rely on a single wave front crossing the array. Hence, additional reflective waves mixing with primary sound waves contaminate the data acquired. Also, test section reverberation is intertwined with the background noise [5]. This is a serious problem as it affects and limits noise source testing, particularly at high speeds because aeroacoustic noise is proportional to velocity. In summary, the current wind tunnel testing situation limits the accurate measurement, identification, and quantification of noise sources.

Figure 1.4 – Microphone array examples. (a) Linear microphone array (University of Kentucky Audio Systems Lab) [6], and (b) Phased array (Virginia Tech Vibrations and Acoustics Lab) [7]. Microphones are flush mounted in a spiral pattern to an aluminum plate.

In response to the present wind tunnel testing conditions being hindered by background noise contamination, various methods are currently employed to improve experimentation procedures:

- First, the physical design of the wind tunnel when correct can significantly reduce the background noise. An example is the appropriate choice of fan rotational speed and
number of blades to ensure the fundamental fan mode does not propagate down the wind tunnel as an acoustic mode.

- Second, the positioning of the microphones used during testing can be tailored to the specific test. There are two main configurations considered when trying to solve the microphone placement problem. The first solution suggests placing the microphones near the model and thus in-flow so that the background noise is less relevant as the model noise at close distances will tend to be sufficiently greater than the background noise. However, due to the flow microphone self noise will be introduced into the measurements. The second configuration suggests placing the microphones out-of-flow which implies that the self noise issue disappears, but depending on the distance from the model, the Signal-to-Noise Ratio (SNR) might become an issue.

- Third, the use of microphone arrays can alleviate the aforementioned problems. When array data is processed accordingly, the noise source can be focused on and a contour noise map in just the area of interest can be obtained [8], ideally devoid of background noise. Problems can still arise with correlated noise over the width of the array and low frequency (< 300 Hz) reflections from the tunnel walls however.

- Lastly, the test section of the wind tunnel can be modified to increase its sound absorption. Methods available include acoustically treating the wind tunnel walls when using an open test section (as seen in Figure 1.5) and placing the array in an anechoic chamber within the wall, sealed with an acoustic transparent material. The former involves covering the wall with a non-smooth thus increasing the wall boundary layer background noise. The latter involves trade-offs between the strength of the chamber-sealing material and its acoustic transparency, which will also create additional boundary layer noise due to its flexibility.
1.2 Research Objectives

The overall objective of this work is to test post-processing techniques for improving wind tunnel acoustic data quality. Specifically, Adaptive Noise Cancellation and Cepstral echo removal techniques will be investigated. These techniques aim to improve the SNR of the acoustic signals at the microphones enabling a more accurate measurement of low level aeroacoustic noise sources in the presence of flow.

Specific objectives of the research include:

- Accurately recover the magnitude and phase of the source signal at the minimum SNR (~ -20 dB) using ANC post-processing techniques.
- Create an ideal signal simulation of a sound source with a reflecting surface present.
- Verify, through accurate magnitude and phase recovery of the non-reflection data, the Cepstral processing technique with aforementioned simulated signal.
- Verify the Cepstral echo estimations through physical sound path delay calculations and spectral characteristics.
- Accurately recover the magnitude and phase of the source signal without reflections using Cepstral echo removal techniques.
1.3 Summary of Techniques in Wind Tunnel Background Noise Reduction

Conventional Beamforming

Conventional beamforming is a microphone array post-processing technique that allows the user to focus on the desired source while suppressing sound from other directions. The array sensitivity is steered/focused to locations in space using phase delays applied to the input microphone signals. Using the assumption that the sound is emitted from monopole sources and spreads spherically, a delay-and-sum algorithm adds the output of all array microphones using the phase correction to achieve the steering/focusing.

Adaptive Noise Cancellation

Adaptive Noise Cancellation (ANC) requires a reference signal input which is highly correlated with the noise present in the primary signal that one wishes to cancel [10]. The reference signal is filtered with a Finite Impulse Response (FIR) filter that utilizes a Least Mean Squares (LMS) algorithm to adapt the filter weights in order to replicate the noise present in the primary signal. In an iterative process it is then combined with the primary signal and the output of the system converges on the primary input minus the noise.

Other methods of background noise removal do exist and have proven successful. One is the Coherent Output Power (COP) method, used with microphone pairs. This uses a cross-spectral approach between microphones and only correlated noise between them is retained in the output power spectrum. Extraneous and uncorrelated noises received at different microphones are mutually incoherent and therefore excluded (see [11-15] for applications and [16] for theory). Another method is spectral subtraction. Here, an acquisition is taken without the model under testing present. For a single microphone, the resulting spectrum is subtracted from (on a pressure-squared basis) that produced from data taken under test conditions with the model present, ideally giving the output of just the noise due to the model. For array processing this can be performed by the subtraction of the background Cross Spectral Matrix (CSM) from that obtained with the source present (see [17-19] for applications).
The ANC method presents a distinction to both of these as it removes noise from the time-domain signal. Thus upon converting to the frequency domain, both the magnitude and phase of the signal will be corrected. This method presents a niche for the case where the background noise is localized and a “background” acquisition needed for spectral subtraction is not possible.

**Cepstral Echo Removal**

Cepstral echo removal is a non-linear signal processing technique in which echoes in an electrical signal are separated from the direct source. Once separated, the signal can be filtered to remove the echoes thereby leaving only the primary signal.

The method behind the technique is the transformation of the convolved direct signal and echo into a sum through a logarithmic process. The echo is then filtered out and the process reversed to recover the primary signal only.

Background literature is given in Section 2.3 and in-depth theory in 3.3.
2 Literature Review

The following chapter presents example literature on three topics: beamforming, Adaptive Noise Cancellation, and Cepstral echo removal techniques. The respective sections are organized to start the reader at the relevant origin of the research and lead him or her to the present day with references pertaining to the research carried out in this work.

2.1 Beamforming

Research in acoustics has included beamforming as a primary analytical tool since the late 1980s. Two examples leading to that time are presented first.

In 1975, Soderman and Noble used an end-fire microphone array with digital time delays to directionally scan the array to focus on noise sources with the intentions of rejecting background noise and reverberations in the NASA Ames 40x80 ft Wind Tunnel [20]. Array performance depended on the microphone spacing to wavelength ratio and orientation in the wind tunnel. For an eight element array with time delays implemented based on the speed of sound and microphone spacing, 9 dB of microphone wind noise (at 3 kHz) and 11 dB of reverberations (at 1 kHz) were rejected in the main array lobe directed at the source.

One year later, 1976, Billingsley and Kinns published their work on “The Acoustic Telescope” [21]. Their telescope incorporated a microphone array with a digital computer that processed the signals and output source distributions with respect to position and frequency. Theory for a line source of arbitrarily correlated monopoles in frequency domain analysis was developed and its application in statistical property estimation of the sound source region was detailed. Jet stream results from an Olympus engine were presented and at high power the significant sound generation region occurred at 3 to 9 nozzle diameters downstream from the nozzle exit.

Brooks et al. investigated rotor noise with a 12 microphone, four-element, symmetrical square directional array of microphones [22] in 1987. The theory behind the frequency domain beamforming used was given. The array was placed out-of-flow with its main directional lobe positioned on the helicopter rotor sweep area as seen in Figure 2.1. A design goal was achieved
in that the sensing area of the main directional lobe was controlled independent of look angle (related to array focal point) and frequency over a “practical” range. This was desired because array side lobes were not of concern due to the anechoic environment the test was performed in. An array blending concept was proposed that divided the array into symmetrical clusters which were then “weighted” depending on what frequency was to be analyzed in order to control resolution and side lobe size. The blending was deemed useful for the array system because the main lobe resolution was freed from frequency dependency.

![Diagram of a test setup for rotor noise investigation](image)

**Figure 2.1** - Test setup for rotor noise investigation [22].

In 1995, Gramann and Mocio investigated the use of adaptive beamforming vs. conventional beamforming for aeroacoustic measurements in wind tunnels [23]. Conventional beamforming, also known as “delay-and-sum”, was achieved by multiplying the output of individual array sensors by time-invariant weights determined from array geometry, speed of sound propagation, and desired focal point. Adaptive beamforming minimized the beamforming response while holding the output at the desired focal point constant. This was achieved through weight adaptation with information taken from the Cross Spectral Matrix (refer to Chapter 3). A linear microphone array was used to map a speaker within a wind tunnel test section and both narrow and broad band inputs were employed. Results were shown to give adaptive beamforming levels with corrections applied to within 1-3 dB of semi-anechoic levels for a broadband source with significant side lobe attenuation, as compared to a 7 dB bias seen in the conventional beamforming results.
Mosher [24] and Humphreys Jr. et al. [25] gave very complete summaries on the theory, processing, array design, hardware, and testing consideration for the use of the technology in wind tunnels at that time, published in 1996 and 1998 respectively. Humphreys Jr. et al. addressed shear layer refraction correction, calibration, and shading algorithms for two different aperture sized arrays, named Small/Large Aperture Directional Array (SADA/LADA). The LADA was created with 35 flush-mounted microphones, providing greater resolution than the SADA. It had a nominal diagonal aperture size of 36 inches and a logarithmic spiral design for the microphone placement was chosen to provide a target frequency range of 2-30 kHz. Example beamforming results obtained for a 6% scale model of a main element NACA 632-215 wing section with a 30% chord half-span flap are given in Figure 2.3.

![Figure 2.2](image)

**Figure 2.2** - Beamforming plot for 6% scale model of a main element NACA 632-215 wing section with a 30% chord half-span flap; mach 0.17, angle-of-attack 16°, flap at 39°, and frequency 12.5 kHz. Lines seen around center point denote the flap-edge. Taken from [25].

Brooks and Humphreys [17] studied the relationship between directional microphone array size and noise measurement resolution and accuracy in 1999. Two arrays, one with a solid measurement angle of 31.6° (LADA) and another of 7.2° (SADA) of source directivity, were used in an anechoic, open-jet configuration outside of the flow. In general, the larger the array the better the resolution achieved yet the lower the sound level measured. To address these
findings an analytical model was created and noise scattering from shear layer turbulence was investigated. A technique called “source region integration” was developed for the analytical model to provide a total spectral sound output for distributed sources. An example beamform comparing data gathered from the SADA and LADA is shown in Figure 2.3. The combination of the analytical model and information indentified from the noise scattering accounted for the difference in source level between arrays of different sizes. Implications for open vs. closed-section wind tunnels such as Signal-to-Noise Ratio, scattering effects, and noise floor were considered. A benchmark was achieved in that criteria were established to guarantee noise source levels independent of array size when source region integration was performed.

![Figure 2.3](image)

**Figure 2.3** – Beamform results for a flat-edge flap; Mach 0.11, one-third octave frequency 40 kHz. Integration area shown with dotted line. Taken from [17].

In 2005, Brooks and Humphreys published a method that represented a step change in the noise measurement community [26]. Their methodology, called the Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS), deconvolves noise source results from array beamform response functions. Its success relies upon a positivity constraint, allowed by the independent source assumption, which renders a defined linear system of equations sufficiently deterministic. Figure 2.4 gives three example plots displaying the advance in beamforming resolution and accuracy achieved with the DAMAS methodology. Cross spectral subtraction (ref. Chapter 1 and [17-19]) was used to reduce background noise from the beamforming results, a precursor to DAMAS implementation. The objectives of the present research are to improve data quality through improved SNR by background noise reduction and acoustic reflection removal.
In doing so, more accurate beamforming will result. This directly ties into the DAMAS processing as inaccurate beamforms will cause the deconvolved solution to diverge. See Appendix A.1 for equations governing DAMAS methodology.

Figure 2.4 - Example figures from [26]. (a) Synthetic point source simulation; frequency 20 kHz, source "placed" 5" from array, source strength set to 100 dB, (b) Same as (a) with multiple synthetic sources placed to form the acronym NASA; frequency 30 kHz, and (c) NACA 63-216 airfoil in Leading and Trailing Edge test (16" chord, 36" span, -1.2° angle-of-attack to vertical flow); one-third octave frequency 20 kHz.

2.2 Adaptive Noise Cancellation

Pioneering work in the field of adaptive noise cancellation began at Stanford University in 1960 with work done by Widrow and Hoff on adaptive switching circuits [27]. This was
accompanied by work from Koford and Groner [28] and lead to the development of the Least-Mean-Square (LMS) adaptive algorithm and pattern recognition scheme (Adaline) respectively.

In 1975 Widrow et al. published a work that detailed the principles of adaptive noise cancellation to date and provided background theory development as well as examples of applications of the technology [10]. It reviewed optimal filtering and early progress made in adaptive filtering. The adaptive filter theory was presented and the concept is depicted in Figure 2.5. For successful noise cancellation a suitable reference input was crucial and allowed the processing of signals whose properties were unknown a priori. Weiner solutions to statistical noise cancelling problems were derived to demonstrate analytically the increase in SNR. The effect of primary signal in the reference input was investigated and it was determined that a small amount of signal in the reference input did not render the noise canceller useless (i.e. for a SNR at the filter output of 20, a 5% signal distortion was introduced). The adaptive noise canceller was presented as a notch filter due to its specific advantages i.e. bandwidth control, infinite null, adaptive frequency tracking, and was easily transformed into a high-pass filter by setting the notch to zero frequency.

![Diagram](image)

**Figure 2.5** - Figures from [10]. (a) Adaptive noise canceller concept, and (b) ANC principle illustrated in the cancellation of electrocardiographic noise.

The paper went on to provide example applications. The first was the cancelling of 60-Hz interference in electrocardiography with example results shown in Figure 2.5. Other medical
applications were presented including the cancellation of the donor’s (old) heart signal in heart transplant procedures and cancelling the mother’s heart signal while listening for the fetus’. Engineering applications were presented next. The authors were able to cancel background noise in a speech signal by 20-25 dB, which before cancellation made the speech indecipherable, on the order of seconds. Closer to topics addressed in the research presented in this thesis, antenna sidelobe cancellation was performed as an alternative to beamforming. The paper finished by defining the algorithms used in mathematical detail. The basis of the ANC post-processing method used in this research stems from the concept given in this paper.

2.3 Cepstral Echo Removal Techniques

The Cepstrum theory was first presented by Bogert et al. [29] in 1963. The objective of the technique was to determine the echo arrival times and relative strength in a signal composed of direct and reflected components from time domain data. The method is based on the conversion of a convolution into an addition through a non-linear logarithmic process.

In 1972, Kemerait and Childers [30] looked at the use of both the Power and Complex Cepstra in the decomposition of a signal with multiple wavelets. Data processing concerns were addressed such as the type of filtering to be employed in the Cepstrum domain and data window smoothing. Specific cases were investigated including noise present, multiple echoes, and echo distortion. The Power Cepstrum was deemed more robust than the Complex in the determination of echo arrival times. The wavelet recovery by the Complex Cepstrum was less distorted by comb filtering than short pass (block) filtering. The Mean-Square Error was increased as the echo magnitude approached that of the wavelet (direct signal) and decreased by appending zeroes to the data record length.

In 1975, Hassab and Boucher [31] looked at the Power, Complex Cepstrum, and the autocorrelation function regarding their performance in relation to distortion and noise present in the channel. Analytical solutions were checked with computer simulations. Distortion was determined to increase the width of the echo and decrease its peak height. Criteria were established for echo detection with the Power, Complex Cepstrum, and autocorrelation in simulations performed varying the signal-to-noise ratio and signal-to-noise bandwidth. It was shown that the relative bandwidth was more crucial to echo detection than signal-to-noise ratio.
The ideal condition proposed by the results was a relative bandwidth greater than unity and a signal-to-noise ratio greater than 10 dB.

Childers et al. [32] in 1977 released a paper intended to serve as a Cepstrum processing aide in response to the surge in the use of the Cepstrum in varied fields. Problems encountered in the processing linked to phase unwrapping, spectrum notching, aliasing, oversampling, and appending zeroes were addressed. Phase unwrapping exhibited errors when linear phase components with large slope were present. If phase changes between the samples were greater than $\pi$ the authors proposed that the problem could be solved by extending the data record length with zeros. The solution to aliasing in the Cepstrum was purported to be insuring the data record be as long as possible (experiment dependent) and then extending the record length with zeros. The authors go on to recommend that zeros should always be appended to the data record to avoid the two aforementioned processing issues. The advantages/disadvantages, data dependent not general, of windowing were presented. In speech analysis, for example, data windowing is desirable and a specific window [33] was suggested that preserved the separability of the wavelet and echo series. This window was further improved upon to include changing of the echo phase relation [34]. Both windows were issued with cautions due to presence of the wavelet outside of the windowing region. Methods of pre-whitening, trend removal, and filtering in the Cepstrum domain were analyzed with results. Finally, cautions were issued due to the data dependency of the processing techniques.

Syed et al. in 1980 [35] tested Cepstrum reflection removal from model jet, spinning rig, and jet engine data. The Rolls-Royce Viper 11 engine was tested in an open air environment with positive results not affected by wind or turbulence. They concluded that source size and directivity have little effect on the Cepstrum processing and that if test rigs are situated identically for every test, the reflection removal processing can be automated.

Two years later, Fuller et al. [36] used simulations of practical applications in acoustic source location to investigate the performance of the Complex Cepstrum in wavelet recovery with distorted echoes and noise present. Their phase unwrapping was performed with an algorithm first presented by Tribolet [37]. Distortions in the echo were simulated with a transfer function that was varied to represent real world surfaces (concrete, grass). The Complex Cepstrum process, which used block liftering to remove echo components, performed well with
grass as the simulated reflective surface ("relatively absorptive") for two decaying signals, an exponential and swept-sine.

In that same year, Martin and Burley at NASA LaRC [38] investigated the Power Cepstrum application to helicopter-rotor acoustic data. The data consisted of low-frequency tonal content which differs from the broadband or random character of jet noise. The mathematical theory behind the Power Cepstrum was presented at length for the single and two ideal reflection cases. The effects of additive noise and echo distortion were studied and numerical examples provided. Echo transfer function correction was presented as an alternate method of removing contamination in the Power Spectrum due to a distorted echo. An advantage was that case-by-case editing in the Cepstrum domain was not required; however, the echo transfer function must be known. A theoretical conclusion was made that the analysis bandwidth should be less than one-half the echo ripple frequency. As for the experimental rotor results, it was determined that Cepstrum editing improved the free-field spectrum by removing some of the contamination due to acoustic reflections.

Lastly, in 1988, Fuller et al. [39] used the Complex Cepstrum to correct directivity estimations of acoustic sources when a reflective surface was present. Block liftering of the Complex Cepstrum was employed to remove the echo deltas present in an effort to avoid user discernment involved in interpolation liftering and thus automate the task. As block liftering (setting all values within the block to zero) inevitably removed some information from the direct signal, a second step in the analysis was necessary. This step was to implement a coherence limit of 0.98 between the two recovered signals; any frequency under this was discarded. Through this bearing estimation and the echo time delay, the source distance was measured with only 6% error.

Many of the methods implemented and cautions issued in the work reviewed were considered in the Cepstrum processing done in the present research and are noted as such when applicable.
3 Theory

The following chapter presents the theory behind the three main concepts used in this research: directional microphone arrays, adaptive noise cancellation, and Cepstral echo removal techniques. Detailed mathematical justifications are given as well as diagrams and pictures where necessary to further aid explanation.

3.1 Directional Microphone Array

Conventional beamforming is used in order to focus on a noise source of interest. This is achieved in the frequency domain by phase shifting the outputs of each array microphone by an amount corresponding to the sound propagation delay due to array/source geometry then summing them together yielding a single output signal for the array as seen in Figure 3.1. Note that a time delay, \( \tau \), corresponds to a phase shift in the frequency domain, as illustrated by Equation (3.1), where \( p \) represents pressure in the time domain and \( P \) the frequency domain.

\[
p(t - \tau) \leftrightarrow P(\omega)e^{-j\omega \tau}
\]  

(3.1)

![Figure 3.1 - Directional array principle [25].](image)
**Sound Source Concept**

The principle behind the directional microphone array (adapted from [25]) starts by assuming that a monopole acoustic source exists in space at location \( \mathbf{x} \). An omnidirectional pressure wave propagating radially from the source can be defined as

\[
p(r, t) = \frac{C}{r} e^{j\omega(t - \frac{r}{c})}
\]  

(3.2)

where \( r \) is the radial distance from the source origin, \( t \) is time, \( C \) is a constant, \( \omega \) is the angular frequency of the wave, and \( c \) is the speed of sound.

A microphone array is placed a finite distance from the source. Each microphone in the array will receive a slightly different signal depending on its distance from the sound source as seen in Figure 3.2.

![Figure 3.2 - Array/Source geometry relating sound propagation to array microphones from a monopole, omnidirectional source radiating spherically [40].](image)

The pressure at the \( n^{th} \) microphone is then

\[
p_n(t) = \frac{\text{const.}}{r_n} e^{j\omega(t - \frac{r_n}{c})}
\]

(3.3)

where \( r_n \) is the distance from the \( n^{th} \) microphone to the source center. The term in parentheses is the delayed time from the source to microphone \( n \).

**Array Response**

The ideal array response for a simple source is given next. Let the phase center of the array be defined as
With this, the array response becomes

\[ R(\omega, \tilde{x}, \tilde{x}^s) = \sum_{n=1}^{N} \frac{r_s}{r_n} e^{j\omega \left( (r_s - r_c) - (r_n - r_n) \right)} \]  

(3.5)

where \( \tilde{x} \) represents the location in space to which the array is electronically steered, \( \tilde{x}^s \) the source location, \( r_c^s \) the distance from the source to the center array microphone, \( r_n^s \) the distance from the source to microphone \( n \), \( r_c \) the distance from the steering location to the center microphone, and \( r_n \) is the distance from the steering location to microphone \( n \). The response is plotted at a number of steering locations which form a line a finite distance away from the array. For the research performed here that line runs through the source as seen in Figure 3.5. The response represents a spatial filtering of the sound present at the steering locations due to the frequency and array/source geometry. It can be expressed in decibels referenced to the level obtained at the source, \( \tilde{x}^s \)

\[ dB(\tilde{x}) = 20 \log_{10} \left[ \frac{|R(\omega, \tilde{x}, \tilde{x}^s)|}{|R(\omega, \tilde{x}, \tilde{x}^s)|} \right] \]  

(3.6)

An example array response plotted with Eq. (3.6) is given for a nine microphone array, a distance 67” from the source, at 10 kHz, over a 10’ line in Figure 3.3.

![Figure 3.3 - Ideal array response over a 10’ line.]
Beamforming

The first step in beamforming is to compute the Cross Spectral Matrix (CSM) for the data set to be analyzed [26]. The CSM is formed from the Fast Fourier Transforms (FFTs) of the data set (in volts vs. time for instance) [41]. The FFTs of two microphones, \( n \) and \( n' \), are denoted \( P_n(f, m) \) and \( P_{n'}(f, m) \) and are formed from the microphones’ time records \( p_n(t) \) and \( p_{n'}(t) \) (Eq. (3.3)) after they are converted to engineering units, namely pressure,

\[
P_n(f) = \frac{2}{T} \int_0^T w(t) p_n(t) e^{-2j\pi ft} dt
\]

where \( T \) is the acquisition time and \( w(t) \) a chosen window function. The frequencies, \( f \), at which the transforms are defined are determined by the bandwidth, \( \Delta f = F_s/M \) (Hz), where \( F_s \) is the sampling frequency in Hz and \( M \) is the data block length, in number of samples. A CSM element, as a function of frequency, is defined

\[
c_{nn'}(f) = \frac{2}{k w_M} \sum_{k=1}^{K} \left[ P_{nk}^*(f, m) P_{n'k}(f, m) \right]
\]

The total record length is \( R = MK \). The summation is multiplied by two because it is a one-sided cross-spectrum. It is then divided by the data block length \( M \) to normalize the FFT output and the number of block averages \( K \) to get a mean value across the data blocks. The term \( w_s \) is a weighting constant used when a weighted window (e.g. Hanning) is implemented. The star on the first transform, \( P_{nk}^*(f, n) \), denotes the complex conjugate. The CSM element is a complex spectrum. The full CSM for \( M \) array microphones is

\[
\hat{C} = \begin{bmatrix}
C_{11} & C_{12} & \cdots & C_{1N} \\
C_{21} & C_{22} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
C_{N1} & \cdots & \cdots & C_{NN}
\end{bmatrix}
\]

The beamforming uses the CSM to electronically “steer” to positions in space defined by the user. For instance, these spatial positions would correspond to grid points on a plane or on a line as in Figure 3.4.
For the work done here, the array was one dimensional and the grid points were located along a line that bisected the sound source as shown in Figure 3.5.

Each position to be steered to is given a number, $b$, on the grid line. In order to steer to grid points on the line, vectors must be calculated between each array microphone and the grid point being steered to. The vector for microphone $n$ is

$$E_n = \left(\frac{r_n}{r_c}\right) e^{j2\pi f \tau_n}$$

where $r_n$ is the straight line distance from microphone $n$ to the grid point, $r_c$ denotes the straight line distance from the center microphone to the grid point, and $\tau_n$ is the propagation time for sound to travel between microphone $n$ and the grid point. The term $\left(\frac{r_n}{r_c}\right)$ is included to normalize the distance dependent amplitude of the steering vector to that of the center array microphone. The steering vector matrix, size $N \times 1$, is

$$\hat{E} = col[E_1 \ E_2 \ \cdots \ E_N]$$

The steering vectors, being complex, will phase shift the contributions from each microphone in the array to allow for constructive summing at the chosen grid point.

Finally, the array’s response is given as an output power spectrum.
\[ S(\hat{E}) = \frac{\hat{E}^T \hat{E}}{N^2} \] (3.12)

The response has units of mean-pressure-squared vs. frequency bandwidth i.e. Pascals\(^2\) vs. Hz. Dividing by the total number of microphones squared normalizes the response to that of a single microphone level. The superscript \(T\) is taking the complex transpose of the steering vector matrix. It should be noted that modifications to the beamforming presented here exist including shading algorithms to suppress different microphones in order to change the output response pattern and a Diagonal Removal process useful in low signal-to-noise situations.

Beamforming as summarized will be used for both the ANC and Cepstrum processing, providing a useful tool in the array analysis of results obtained.

### 3.2 Adaptive Noise Cancellation

Adaptive noise cancellation utilizes a reference input, ideally containing just noise, which is passed through an adaptive filter and later subtracted from a primary input containing both the desired signal and components of the noise present in the reference input. The output becomes the primary signal with the noise attenuated or cancelled altogether. Adaptive filters are those with the ability to adjust their own parameters automatically with little or no previous information about the signal to be cleaned or noise to be cancelled. The correlation between the noise present in the reference channel and the primary input is important: the higher it is, the better the cancellation. A diagram of the concept is pictured in Figure 3.6.
In Figure 3.6, a primary sound source, \( s \), that contains uncorrelated noise, \( n_0 \), is transmitted to the upper left channel. The bottom left channel receives noise, \( n_1 \), correlated with \( n_0 \) in some unknown manner. This is the reference input. It is fed into the adaptive filter with the goal of replicating \( n_0 \). As the noise characteristics are assumed to be unknown, an adaptive filter is necessary. The output of the adaptive filter is subtracted from the primary input, source plus \( n_0 \), to produce \( z \), which also serves as the error signal to the filter. Proof of the concept follows, from [10].

The output of the system, \( z \), is defined as

\[
    z = s + n_0 - y \tag{3.13}
\]

Squaring Eq. (3.13) and simplifying

\[
    z^2 = s^2 + (n_0 - y)^2 + 2s(n_0 - y) \tag{3.14}
\]

Again, the sound source is uncorrelated with the noise. Taking the expectation of Eq. (3.14) and simplifying

\[
    E[z^2] = E[s^2] + E[(n_0 - y)^2] + 2E[s(n_0 - y)] = E[s^2] + E[(n_0 - y)^2] \tag{3.15}
\]

As will be shown, noise cancelling requires that the adaptive filter minimize the total system output power. As the filter is trying to replicate \( n_0 \) with \( n_1 \), the signal power \( E[s^2] \) will be unaffected as the output power, \( E[z^2] \), is minimized. Thus, the minimum output power becomes

\[
    \min E[z^2] = E[s^2] + \min E[(n_0 - y)^2] \tag{3.16}
\]

As the filter adapts to minimize the output, \( E[(n_0 - y)^2] \) is also minimized, making the filter output an estimate of the primary noise, \( n_0 \). Also, from Eq. (3.13), \( (z-s) \approx (n_0 - y) \); thus when \( E[(n_0 - y)^2] \) is
minimized, $E[(z-s)^2]$ is also minimized. This infers that the output power will be an estimate of the source signal. The best possible minimization leads to $E[z^2] = E[s^2]$. Thus, $z=s$ and the output will be entirely noise free.

In order to produce an output that is as noise free as possible, a fit to the source signal, the adaptive filter uses a Least Mean Squares (LMS) algorithm. The purpose of the LMS algorithm, which operates in the time domain, is to adjust the filter weights to minimize the mean-square error. From [10], the mean-square error can be expressed as a function of the weight vector, $w$,

$$E[e_n^2] = E[(s + n_0)^2] - 2P^T w + w^T R w$$  \hspace{1cm} (3.17)

Where

$$P = E[(s + n_0)_n (n_1)_n] \quad \text{and} \quad R = E[(n_1)_n (n_1)_n^T]$$  \hspace{1cm} (3.18)

The error is a quadratic function of the weight vector, $w$, which can be pictured as a concave hyperparaboloidal surface, as illustrated by Figure 3.7.

![Figure 3.7 - Computer rendering of hyperparaboloidal surface [42].](image)

To minimize the error, the gradient of Eq. (3.17), $\nabla$, must be taken

$$\nabla = \frac{\partial E[e_n^2]}{\partial w_n} = -2P + 2Rw$$  \hspace{1cm} (3.19)

Setting Eq. (3.19) to zero, the optimal weight vector is found

$$w_{opt} = R^{-1}P$$  \hspace{1cm} (3.20)

The LMS algorithm [10] approximates Eq. (3.20) in real time using the steepest decent method. Accordingly, the current weight vector is equal to the previous weight vector plus a change proportional to the previous negative gradient multiplied by the user-defined step size, $\mu$,

$$w_n = w_{n-1} - \mu \nabla_{n-1}$$  \hspace{1cm} (3.21)
The Widrow-Hoff LMS algorithm uses an estimate in place of the true gradient, hence Eq. (3.21) becomes

\[ w_n = w_{n-1} + 2\mu e_{n-1}(n_1)_{n-1} \quad (3.22) \]

Two parameters are adjustable in the adaptation algorithm: the step size and the weight vector length. The step size controls the stability and rate of convergence. It has stability limits of

\[ 0 < \mu < \frac{1}{\lambda_{max}} \quad (3.23) \]

where \( \lambda_{max} \) represents the largest eigenvalue of the matrix \( R \). Within these limits, the larger the step size the faster the adaptation. The weight vector length (i.e. number of weights or “taps”) has a minimum that equals twice the ratio of the total signal bandwidth to the frequency resolution of the filter. Above this minimum, the more weights included (thus longer the vector length) the more accurate the response. This increase also comes with increased processing time, however.

The interior of the dashed box from Figure 3.6 is given in 3.8. The same terms and color coding are used in each to provide continuity.

![Figure 3.8 - Adaptive noise canceller concept with LMS algorithm and FIR Filter block diagrams included. Adapted from [43].](image)

As the signals are not shifted in time, success of the adaptation is dependent on the causality of the system. In the experimental setup (ref. Fig. 4.5) the reference microphone receives the background noise before the any of the array microphones do. From Figure 3.8, the reference signal (which ideally consists of only the undesired noise) is labeled \( x_n \) and is fed to the
FIR Filter at the start of the processing. The initial filter coefficients \((c_n)\) and number of weights are provided by the user. The output of the filter \((y_n)\) is subtracted from the desired signal \((d_n; \text{an array microphone})\) which contains a correlated version of the noise present in the reference signal from an earlier time. The error \((e_n)\) is then fed into the LMS algorithm and multiplied with the reference input \((x_n)\) and the step size \((\mu)\) (Eq. (3.22)) to produce the next set of filter coefficients \((c_{n+1})\). The order of the FIR Filter is the number of weights or “taps” used. FIR Filters are always stable and have linear phase response given that the coefficients are symmetrical.

The difference equation defining the output of the filter is

\[
y_n = c_0x_n + c_1x_{n-1} + \cdots + c_Nx_{n-N}
\]

where \(N\) is the number of filter weights specified. In block diagram form:

![Figure 3.9 - FIR Filter block diagram.](image)

As the noise to be cancelled is present in the reference signal before it is in the desired signal, the system is deemed causal and successful adaptation will be achieved. This is seen in Eq. (3.24). As long as the number of filter weights is sufficient to encompass the time delay between the reference and array microphone (in number of samples), the noise present in \(d_n\) will be related to that seen in an earlier sample still present in \(y_n\).

### 3.3 Cepstrum Reflection Removal

Cepstral processing is a signal processing technique with the goal of determining echo arrival times and magnitudes in an input signal where direct and reflected components are present. Once identified, the energy due to the echoes can be removed and process reversed to recover the original signal without reflection contamination. The basis of the algorithm relies on
the conversion of a convolution into an addition through a non-linear logarithmic process. The following paraphrased terms belonging to the Cepstrum domain are used in this work [29]:

- **Frequency** .............. **Quefrency**
- **Spectrum** .............. **Cepstrum**
- **Amplitude** .............. **Gamnitude**
- **Filtering** .............. **Liftering**
- **Harmonic** .............. **Rahmonic**

**Power Cepstrum**

The Power Cepstrum is defined as the magnitude of the inverse Fourier transform of the natural logarithm of the magnitude squared of the Fourier transform of the time domain data \( x(t) \) [44]. Figure 3.10 gives a block diagram of the Power Cepstrum.

![Block diagram of Power Cepstrum](image)

**Figure 3.10** - Power Cepstrum sequence [44].

In equation form,

\[
x_p(t) = \left| \int_{-\infty}^{\infty} \ln|X(\omega)|^2 e^{j\omega t} d\omega \right|
\]  
(3.25)

where \( X(\omega) \) represents the Fourier transform of \( x(t) \).

To analyze the operations that allow the Power Cepstrum to identify echoes in a signal, take an example signal, \( s(t) \), and one echo, amplitude \( a \), at \( t = t_0 \), \( as(t - t_0) \). The composite signal is

\[
x(t) = s(t) + as(t - t_0)
\]  
(3.26)

The Fourier and inverse Fourier transforms are defined as

\[
F = X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt
\]  
(3.27)

\[
F^{-1} = x(t) = \int_{-\infty}^{\infty} X(\omega)e^{j\omega t} d\omega
\]  
(3.28)

Taking the magnitude squared of the Fourier transform (block 1 then 2 in Figure 3.10) of Eq. (3.26), where \( S(\omega) \) is the Fourier transform of \( s(t) \),

\[
|X(\omega)|^2 = |S(\omega)|^2[1 + a^2 + 2a \cos(\omega t_0)]
\]  
(3.29)
The separation of the echo, $a$, from the wavelet, $S$, is realized through a non-linear process, implemented by taking the natural logarithm (block 3 in Fig. 3.10) of Eq. (3.29)

$$ln|X(\omega)|^2 = ln|S(\omega)|^2 + ln[1 + a^2 + 2a \cos(\omega t_0)]$$

(3.30)

By bounding $a$ to $-1 < a < 1$ and taking the inverse Fourier transform (block 4 in Fig. 3.10) of Eq. (3.30), the Power Cepstrum of $x(t)$ can be calculated [44]

$$x_p(t) = F^{-1}(ln|X(\omega)|^2) = F^{-1}(ln|S(\omega)|^2) + \sum_{n=1}^{\infty} A_n \pi[\delta(t - nt_0) + \delta(t + nt_0)]$$

(3.31)

where

$$A_n = (-1)^{n+1} \frac{2a^n}{n}$$

(3.32)

As seen in Eq. (3.31), the Power Cepstrum of the wavelet will have a series of delta functions that decay, due to the nature of $A_n$, superimposed onto it, yet the first delta will be separated from the wavelet by $t_0$, thus allowing the reflection components to be easily identified. Figure 3.11 gives an example of the Power Cepstrum for an impulse function and its echo. The impulse (at Quefrency=0) is not shown to allow for better resolution of the reflection deltas. The first echo is seen just after 0.01 on the Quefrency scale (x-axis). From there, decaying deltas can be seen that are the rahmonics (harmonics) of the first echo. Note that all deltas have corresponding high Quefrency components as the Power Cepstrum is an even function.

![Figure 3.11 - Power Cepstrum for an impulse function and its echo [44].](image-url)
Complex Cepstrum

The Complex Cepstrum is defined as the inverse Fourier transform of the complex logarithm of the Fourier transform of the time domain data \( x(t) \) [44]. Figure 3.12 gives a block diagram of the Complex Cepstrum.

![Figure 3.12 - Complex Cepstrum sequence [44].](image)

In equation form,

\[
\hat{x}(t) = \int_{-\infty}^{\infty} \ln[X(\omega)]e^{j\omega t} d\omega
\]  

(3.33)

It is a real number and defined for continuous valued functions. The term “Complex” is used to distinguish it from the Power Cepstrum as it uses the complex logarithm in order to preserve phase information which in turn allows for original time domain data recovery.

Again, a wavelet and single echo will be used for simplicity to analyze the Complex Cepstrum procedure. Taking the Fourier transform then complex logarithm (blocks 1 and 2 in Fig. 3.12) of Eq. (3.26)

\[
\ln[X(\omega)] = \ln[S(\omega)] + \ln[1 + ae^{-j\omega t_0}]
\]  

(3.34)

Logarithmic properties dictate that

\[
\ln(1 + y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \cdots \quad \text{for} \quad |y| < 1
\]  

(3.35)

Applying this, Eq. (3.34) becomes

\[
\ln[X(\omega)] = \ln[S(\omega)] + ae^{-j\omega t_0} - \frac{a^2}{2}e^{-2j\omega t_0} + \frac{a^3}{3}e^{-3j\omega t_0} - \cdots
\]  

(3.36)

Applying the inverse Fourier transform (block 3 in Fig. 3.12) to Eq. (3.36) the Complex Cepstrum is obtained, where \( \hat{x}(t) \) and \( \hat{s}(t) \) are the Complex Cepstra of the composite signal and the wavelet, respectively,

\[
\hat{x}(t) = \hat{s}(t) + a\delta(t - t_0) - \frac{a^2}{2}\delta(t - 2t_0) + \frac{a^3}{3}\delta(t - 3t_0) - \cdots \quad \text{for} \quad a < 1
\]  

(3.37)

From Eq. (3.37), it is shown that the Complex Cepstrum of a wavelet and one echo consists of deltas that alternate in sign and decay in amplitude superimposed on the wavelet’s Complex Cepstrum. Again, the first delta will occur at time \( t_0 \) from the wavelet. Figure 3.13
shows the Complex Cepstrum of an impulse function and its echo. As in Figure 3.11, the impulse is not shown to provide better resolution of the reflection deltas. The first echo appears just after Quefrency=0.01 and alternating sign decaying deltas are seen thereafter.

![Complex Cepstrum for an impulse function and its echo](image)

**Figure 3.13** - Complex Cepstrum for an impulse function and its echo [44].

**Distortion in the Complex Cepstrum**

Distortion in the Complex Cepstrum can be modeled as a transfer function that is convolved with the echo [31]. Implementing the convolution with Eq. (3.24)

$$x(t) = s(t) + ah(t - t_0) * s(t)$$  \hspace{1cm} (3.38)

with

$$h(t - t_0) * s(t) = \int_0^T h(t - t_0 - \psi)y(\psi)d\psi$$  \hspace{1cm} (3.39)

where $T$ is the acquisition time (large compared to $t_0$) and $\psi$ is a variable of integration.

Following steps previously taken to analyze the Complex Cepstrum, the Fourier transform then logarithm of Eq. (3.38) becomes

$$\ln X(\omega) = \ln S(\omega) + \ln [1 + aH(\omega)e^{-j\omega t_0}]$$  \hspace{1cm} (3.40)

When $|aH(\omega)e^{-j\omega t_0}|^2 < 1$ (for a reflection amplitude $a < 1$) Eq. (3.40) becomes

$$\ln X(\omega) = \ln S(\omega) + aH(\omega)e^{-j\omega t_0} - \frac{a^2}{2} H^2(\omega)e^{-j2\omega t_0} + \frac{a^3}{3} H^3(\omega)e^{-j3\omega t_0} - \cdots$$  \hspace{1cm} (3.41)

Taking the inverse Fourier transform
\[
\hat{x}(t) = \hat{s}(t) + ah(t - t_0) - \frac{a^2}{2} h(t - t_0) * h(t - t_0) + \frac{a^3}{3} h(t - t_0) * h(t - t_0) * h(t - t_0) - \ldots 
\]

(3.42)

Eq. (3.42) is a convergent series with a peak away from the origin determined by \(ah(t - t_0)\). Note the comparison between Eqs. (3.37) and (3.42). For a distortionless echo the transfer function is a unit impulse at \(t_0\). The distortion spreads the energy out around \(t_0\) per the transfer function, resulting in a lower echo amplitude and wider bandwidth in the Cepstrum domain.

**Noise in the Complex Cepstrum**

Noise is presented as an additive term [38]. Eq. (3.26) with noise becomes

\[
x(t) = s(t) + as(t - t_0) + n(t)
\]

(3.43)

Thus Eq. (3.37) with noise present becomes

\[
\hat{x}(t) = \hat{s}(t) + \hat{n}(t) + a\delta(t - t_0) - \frac{a^2}{2}\delta(t - 2t_0) + \frac{a^3}{3}\delta(t - 3t_0) - \ldots
\]

(3.44)

It can be seen from Eq. (3.44) that as long as the SNR is high the noise will not present an issue in echo delta location in the Cepstrum domain.

**Distortion and Noise in the Complex Cepstrum**

Distortion and noise in the signal combine Eqs. (3.38 and 43) to give

\[
x(t) = s(t) + ah(t - t_0) * s(t) + n(t)
\]

(3.45)

For \(a < 1\), a SNR greater than 0 dB point wise across the spectrum, the noise not correlated with the signal, the Fourier transform then complex logarithm of (3.45) give [31]

\[
\ln[X(\omega)] = \ln[S(\omega)] + L_1(\omega) + aL_2(\omega)H(\omega)e^{-j\omega t_0} - \frac{a^2}{2} L_2^2(\omega)H^2(\omega)e^{-j2\omega t_0} + \frac{a^3}{3} L_2^3(\omega)H^3(\omega)e^{-j3\omega t_0} - \ldots
\]

(3.46)

where

\[
L_1(\omega) = Z_2(\omega) - \frac{1}{2} Z_2^2(\omega) + \frac{1}{3} Z_2^3(\omega) - \ldots
\]

(3.47)

\[
Z_2(\omega) = \frac{N(\omega)}{S(\omega)}
\]

(3.48)

\[
L_2(\omega) = 1 - Z_2(\omega) + Z_2^2(\omega) - Z_2^3(\omega) + \ldots
\]

(3.49)
Performing the inverse Fourier transform on (3.46) yields the complex Cepstrum with distortion and noise present

\[ \hat{x}(t) = \hat{s}(t) + l_1(t) + ah_p(t - t_0) - \frac{a^2}{2} h_p(t - t_0) * h_p(t - t_0) + \frac{a^3}{3} h_p(t - t_0) * h_p(t - t_0) * h_p(t - t_0) - \cdots \]  

(3.50)

where

\[ l_1(t) = F^{-1}(L_1(\omega)) \]  

(3.51)

\[ h_p(t) = F^{-1}(L_2(\omega)H(\omega)) \]  

(3.52)

Note that \( \hat{s}(t) \) is the Complex Cepstrum of direct signal, \( N(\omega) \) is the FFT of the noise, \( S(\omega) \) is the FFT of the direct signal, and \( H(\omega) \) is the FFT of an impulse response function representing the distortion.

Noise would result in the low gamnitude perturbations, seen in \( l_1(t) \) and \( h_p(t) \). It will superpose and weaken the echo at delta time \( \tau \) which can be seen in \( L_2(\omega) \). Proper direct signal recovery will depend on the extent of overlap between the direct signal components (\( \hat{s}(t) \)) and the perturbations due to the noise (\( l_1(t) \)). The distortion present is also perturbed by the noise due to the presence of \( L_2(\omega) \) (Eq. (3.46)). And distortion will modulate the reflection’s strength as seen in Eq. (3.50).

**Appending Zeroes to the Time History**

Appending zeroes to the time history of the signal increases the frequency resolution of its Discrete Fourier Transform (DFT) without having to adjust the sampling rate. This in turn means that the phase is sampled more frequently leading to fewer phase unwrapping errors. A mathematical proof is given in [32].

**Data Averaging**

Two types of averaging can be employed in order to improve the accuracy of the results from a statistical perspective [45]: time and frequency.

Time averaging (the averaging of time history acquisitions) results in better signal coherence as it tends to increase the SNR. Frequency averaging is that of the Fourier transform of the signal. This reduces uncorrelated noise in the frequency domain and increases the resolution bandwidth.
Test Signal Considerations

Certain signal characteristics are more desirable than others when working with Cepstral processing. Similar to an infinitely large impulse in time resulting in an infinitely long and flat spectrum in the frequency domain, the flatter the signal’s spectrum, the sharper the deltas in the Cepstrum domain will be. Any sharp peaks in the signal’s spectrum will cause decaying oscillations in the Cepstrum domain. Also, the more broadband/random the signal content is the more the Cepstral information will be concentrated towards low quefrequencies [39]. To achieve both of these characteristics, the ideal setup would be a speaker with a flat frequency response over a frequency range larger than that of interest to the analysis, outputting pseudo-random white noise.

Filtering in the Cepstrum Domain

Filtering in the Cepstrum domain, called “liftering” [29], is used to remove the deltas seen in Eq. (3.37) that correspond to the reflection, as seen in Figure 3.13. Then the Cepstrum process can be reversed and the signal recovered without the effects of the reflection. For an ideal signal with calculated reflection times that could be verified, the liftering process could be automated. However, due to the non-ideal nature of experimental data the process is manual to ensure optimal signal recovery. Thus, with experimental data the Cepstral post-processing technique should be treated as a diagnostic.

Figure 3.14 - Example of liftering in the Power Cepstrum [35]. (a) Power Cepstrum, and (b) Power Cepstrum with delta at quefrecny = 0.07 seconds liftered.
As the deltas occur at a time lag of $t_0$ away from the information present due to the original signal, two types of liftering can be implemented: comb and block. Comb liftering is a more tedious yet accurate process that interpolates the Complex Cepstrum data at the reflection deltas as the average of the previous and subsequent points. Block liftering zeroes a chosen block of data in the Cepstrum domain. Comb and block liftering will be used in the Cepstrum processing to demonstrate the signal recovery effects each has when implemented.

Both the Power and Complex Cepstra will be used in the post-processing. Considerations and exceptions to ideal signals considered in the theory presented will also be considered where applicable.
4 Experimental Setup

The experimental research consists of two tests. The purpose of the first test (ANC) is to gather acoustic data from a linear microphone array under three conditions: a) a source speaker turned on only, b) a background noise speaker turned on only and, c) both speakers turned on. These acquisitions will allow for investigation of the removal of the background noise from the microphone signals in post-processing. The second test (Cepstrum) has the purpose of collecting data, using the same array rotated vertically, under two conditions: a) the source speaker turned on with fully anechoic surroundings and, b) the source speaker turned on with a reflective surface present.

4.1 Facility

A small anechoic chamber housed in the Physics Lab of the Acoustics Building at NASA Langley Research Center (LaRC) was used for the test area. Approximate dimensions of the room are 97x77x81 inches (length/width/height), wedge tip-to-tip. The anechoic wedges that cover the walls, the ceiling and floor absorb sound to minimize reflection/reverberation in the room. Two different wedge types were installed in the room, both polyurethane. The back wall had wedges with dimensions 12x12x36 inches (length/width/height) and the rest of the room was covered in wedges measuring 18x18x18 inches.

While the room is small, there is still sufficient distance to allow for the array to be located in the far-field. According to [46], a general rule for far-field sound radiation is that the radial distance be sufficiently greater than the characteristic dimension of the sound source,

\[ r \gg D \]

If sufficiently greater is taken to be an order of magnitude, the far-field distances for each sound source used in the testing would be the following, given that the source speaker tube has a radius of 1” and the diameter of the background speaker is 4”

\[ r_{\text{far-field, source speaker}} = 10 \times 1" = 10" \]
\[ r_{\text{far-field, background speaker}} = 10 \times 4" = 40" \]
For the ANC setup (Fig. 4.5), the source and background speaker are 67” and 43” away from the closest microphone in the array, respectively. For the Cepstrum setup (Fig. 4.13) the distance is 75”. Both provide sufficient distance and justification for far-field assumptions.

Three factors play a role in reducing reflected sound in an enclosure [47]: distance between source and receiver, surface area of the enclosure, and average energy absorption of the bounding surface. Although the wedges used were sufficient in their function for the present task (sound attenuation in the frequency range of interest), a larger room would greatly reduce potential reflection in that both the surface area and source/receiver distance would increase.

The cutoff frequency of the room was calculated from the smallest wedge height of 12” [48]:

\[
room\,\text{cutoff\,frequency} = \frac{\text{speed \,of \,sound}}{4 \times \text{angled \,wedge \,height}}
\]

\[
\frac{13397.24 \,\text{in}}{\text{sec}} \div \frac{4 \times 12 \,\text{in}}{\text{sec}} = 279.12 \,\text{Hz}
\]

This frequency is below the frequency range of interest used in this study, 0.5-9.5 kHz.

### 4.2 Instrumentation and Data Acquisition

**Instrumentation**

The instrumentation and data acquisition hardware remained the same for both tests with a few exceptions. Refer to Appendix B for a list of components.

A simplified block diagram displaying the connection between components for the Cepstrum test is shown in Figure 4.1. Note that for the ANC test setup two extra components were present: a background noise speaker and reference microphone (ref. Figure 4.5). Both the background speaker and reference microphone were connected in the same manner as the source speaker and array (respectively) are shown in Figure 4.1.
Figure 4.1 - Component interconnection.

Microphone Array

Nine B&K quarter-inch free-field microphones were used to construct a linear array using metal support rods with base stands, as shown in Figure 4.2a. The desired microphone spacing was 1” between diaphragm centers; actual spacing differed slightly due to equipment limitations (see Fig. 4.2b for exact dimensions). Points of contact between the microphones and support rod were insulated with duct tape to prevent conductivity. All areas around mics were wrapped with acoustic foam to reduce near-field reflections, as shown in Figure 4.2a.

Figure 4.2 – (a) Horizontal microphone array (ANC test), (b) spacing dimensions for ANC array.

Given the array’s geometry, its in-plane sensitivity can be calculated at a given frequency and perpendicular distance away (ref. Eq. (3.6), [25]). In the case of a linear array it describes the array’s theoretical beamformed response to a sound source in a user-defined line. The sensitivity
values in dB across the line are referenced to the source location on the line. As such, the source level is given a value of 0 dB with all other locations on the line in negative dB values.

The beamwidth (B) is the region of the beamforming map that lies within 3 decibels (dB) of the peak value. An example of the beamform’s beamwidth is given in Figure 4.3. For this array, Table 4.1 and Figure 4.4 show the calculated beamwidths as well as array sensitivities at 1, 5, and 10 kHz for a source 67” away from the array. Note that any asymmetries in the array sensitivity patterns are due to the asymmetry of the microphone placement in the array.

![Figure 4.3 - Example beamform with beamwidth illustrated.](image)

**Table 4.1 - Microphone array beamwidths vs. frequency.**

<table>
<thead>
<tr>
<th>Frequency (kHz)</th>
<th>B (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.4</td>
</tr>
<tr>
<td>5</td>
<td>1.4</td>
</tr>
<tr>
<td>10</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Data Acquisition

Data acquisition was facilitated through a custom Lab View program which interacted with the Data Acquisition System. Several output formats were available. A Matlab compatible file with volts versus time output was chosen.

The power supply was set to a 2 Ampere ceiling, based on an observed steady-state current draw of ~1.7 A by the microphones. The output voltage of 13 Volts was chosen to fall between the specified limits of the power supply (8-16 V). Zero gain was set on the microphone supplies.

Filters were set to high-pass the signal out to the speaker(s) at 400 Hz to avoid any low-frequency reflections present in the chamber (ref. Eq. 4.1).

The microphones used were calibrated using the handheld pistonphone and a custom calibration function of the same Lab View program used to acquire data. The pistonphone itself was in-house calibrated at the time of use. The nominal microphone sensitivity provided by the manufacturer was ~4.5 mV/Pa. The average across all 10 mics was 4.48 mV/Pa.
4.3 Experimental Procedure

Adaptive Noise Cancellation (ANC)

Since the microphone array is linear, its sensing ability lies only along a line parallel with the array at the same vertical height. Thus, the array, source speaker, and background speaker were all placed in the same plane parallel to the floor, at a distance of 46.5” above the floor to ensure that the speaker was high enough above the wedges to avoid low frequency near-field reflections. Refer to Figure 4.5 for a diagram of the setup. The source speaker was perpendicular to the array at a distance of 67”. The background speaker was at an angle of 59° (top view) to the array and a distance of 40” from microphone 1. A reference microphone, labeled 10, was situated a distance of 5” away from the diaphragm center of the background speaker with its rear facing the source speaker. The 5” distance was chosen to avoid near-field transient effects of the background speaker. Two types of near-field needed to be taken into account [49]:

- Acoustic: equal to the distance from the source to one wavelength of sound away (frequency dependent).
- Geometrical: equal to the largest diameter of the source, in this case 4”.

The back of the reference microphone was pointed at the source speaker to limit the source speaker’s influence on the reference microphone’s signal. As with the array, the reference microphone stand was wrapped in acoustic foam. Figures 4.5 and 4.6 show a diagram of the ANC setup and a picture of the background speaker, reference microphone, and array, respectively.
A Pioneer 5.25” Cup Midrange background speaker was chosen based on its size, flat frequency response, and output power. A custom source speaker with a Selenium Compression Driver was chosen due to its relatively flat frequency response and high output capability. As
shown in Figure 4.7, the speaker was fitted with a copper extension tube to further simulate an ideal point source and foam (green) was placed in the end of the tube to reduce high frequency reflections. Note that the copper tube will modulate resonances of the driver. The primary can be calculated from assuming the tube acts as an open ended pipe [44]

\[ f_n = \left( \frac{n}{2} \right) \frac{c}{L + \frac{8(ID)}{3\pi}} \]  

(4.2)

where \( f_n \) is the resonance frequency for the \( n \)th harmonic, \( c \) is the speed of sound, \( L \) is the tube length, and \( ID \) is the inside tube diameter. For the first harmonic the tube’s resonance frequency would be \( \sim 1154 \) Hz. As any narrow band analysis performed was above 6 kHz, this resonant frequency did not present a problem.

**Figure 4.7 -** Source speaker with copper extension tube installed.

Three acquisition groups were necessary to characterize the test:

1) Only the background speaker turned on (at different levels corresponding to Signal-to-Noise Ratios (SNRs) of -6.5, -11.4, and -19.9 dB, calculated over the frequency band 0.5-9.5 kHz)

2) Only the source speaker turned on (at different levels corresponding to SNRs of -6.5, -11.4, and -19.9 dB)

3) Source and background speaker both turned on (at different levels corresponding to SNRs of -6.5, -11.4, and -19.9 dB)

The Signal-to-Noise Ratio (SNR) was calculated following the steps in Figure 4.8.
Referencing Figure 4.8, the process started out by obtaining an acquisition of only the background speaker turned on at maximum level. Then acquisitions were taken of only the source speaker turned on at varied levels. For each source speaker level a SNR was calculated until the three aforementioned SNRs were obtained. The ‘Remove Unwanted Frequencies’ block refers to those outside of the frequency band of interest, 0.5-9.5 kHz. These would correspond to 0-0.5 kHz and 9.5-10 kHz (the Nyquist frequency). These frequency limits were chosen due to the HP and anti-aliasing filters respectively.

Once the desired individual SNRs were obtained the hardware settings were documented. In order to simulate two distinct yet broadband noise sources, two input noise types were used. A pink noise generator was used to drive the source speaker and a white noise generator was used to power the background speaker. This was also done to give the recovery data (emulating the source speaker) a distinct shape from the noisy signal (background and source speaker) from which it is processed from. Both were High Pass filtered at 400 Hz to avoid the chamber cutoff frequency (Eq. (4.1)). Example array averaged autospectra are shown in Figures 4.9 and 4.10.
The autospectra were calculated with the `pwelch` function in Matlab. `Pwelch` calculates the Power Spectral Density of the input signal vector using Welch’s averaged modified periodogram method of spectral estimation [50]. However, all autospectra presented are Power Spectra i.e. narrowband and thus bandwidth dependent. Bandwidth is calculated as

\[
Bandwidth = \frac{\text{Sampling Frequency}}{\text{FFT Block Size}}
\]  

(4.3)
The autospectra shown previously provide information of each speaker’s acoustic magnitude characteristics over a chosen frequency band. When only one or the other is turned on, information is gathered and the specific speaker is characterized. Once the settings for different SNRs were obtained, data was taken with both speakers turned on at those settings. Because each speaker was characterized separately, these autospectra will serve as benchmarks in quantifying the accuracy of the ANC technique in magnitude recovery. The aim will be to remove the noise from the combined acquisition in order to recover the noise free signal.

**Cepstrum Reflection Removal**

A second test was setup up in order to validate the performance of Cepstral post-processing on data from the microphone array. The array was configured vertically to be able to sense reflections coming from the floor (anechoic wedges removed, see Figure 4.11a). With the wedges removed, the floor of the anechoic chamber was used as the reflecting surface. The only alteration to the array was to turn it vertically. The microphone spacing, ordering, and numbering were left identical to the ANC test, as shown in Figure 4.11.

![Figure 4.11](image)

**Figure 4.11** - (a) Vertical microphone array (Cepstrum test) and, (b) Microphone spacing for the vertical array.

The speaker used as the background speaker in the ANC test was used as the source speaker to provide stronger interference of the reflective signal at the array. Both support rods
for the array and speaker were wrapped in foam to inhibit reflections. Figures 4.12 and 4.13 show the source speaker mounted and the Cepstrum test setup respectively.

**Figure 4.12** - Source speaker mounted; Cepstrum test.

**Figure 4.13** - Cepstrum test setup diagram.

Two acquisition groups were needed to characterize the Cepstrum test:

1) Source speaker turned on with floor wedges in place (no reflection)
2) Source speaker turned on with floor wedges removed (reflection)

The source speaker was driven with white noise at a high setting on the noise generator to ensure that the output was above the noise floor of the room. The reflecting surface provided a second “source”, the reflection, which interfered with the direct signal constructively and destructively. Example autospectra of both a non-reflection and reflection case are shown in Figures 4.14 and 4.15.

![Figure 4.14](image)

**Figure 4.14** – No reflection case, center array microphone autospectrum (white noise, BW = 19.53 Hz).
The source speaker was characterized at a set output level and then data was taken with and without floor wedges providing non-reflection and reflection acquisitions respectively. The autospectra shown in Figures 4.14 and 4.15 give examples of benchmarks to be used in the quantification of the Cepstral reflection post-processing in magnitude recovery; the reflection data will be processed and compared to the non-reflection case.
5 Experimental Results

This chapter details the results obtained from the processing of the data collected as outlined in Chapter 4.

The chapter is divided into two sections. The first section discusses the results obtained from the Adaptive Noise Cancellation (ANC) procedure. The speaker and microphone characteristics are discussed in relation to the experiment. Then, the coherence, SNR, and processing method used are defined and the results presented as function of the SNR. The second section details the Cepstrum test results. A simulation is used to provide ideal data for processing and illustrate Cepstral features. Then, the experimental data is presented with the implementation of four post-processing schemes. The results are shown as a function of the processing scheme and Mean-Square Error calculations.

5.1 Adaptive Noise Cancellation (ANC)

The first step in choosing a noise source to be cancelled from collected data was to ensure that the speaker chosen could:

1) Provide a strong output to achieve the desired SNRs (as low as -20 dB) at a level such that the lower power signal was still at least 10 dB up from the noise floor of the chamber.

2) Maintain the SNR over the frequency range of interest (0.5-9.5 kHz).

As described in the Chapter 4 (Fig. 4.8), the SNRs were calculated and autospectra from the center array microphone are given for three SNRs in Figure 5.1. The acquisition parameters used were: a sampling frequency of 20,000 samples/second, 50 blocks, and an FFT block size of 20480 samples. This gives a bandwidth for all spectra presented of 0.98 Hz (ref. Eq. 4.3).
As seen in Figure 5.1, all three SNRs are achieved with source signals that do not cross the 10 dB threshold above the noise floor at any point in the frequency range of interest, even at the lowest SNR of -19.9 dB (Figure 5.1c). Also note that using a pink noise input for the source speaker (red line, Figure 5.1) provides a spectrum that is notably distinct in shape from the background speaker driven with white noise (green line). This is desirable because it provides clear spectral features (peaks and valleys) for the recovered signal to match and a accurately recovered signal will look noticeably different from the combined signal it was recovered from.

As described in Chapter 4 (Fig. 4.5), a microphone was placed near the background speaker with its back facing the source speaker to reduce the source speaker signal’s influence on the reference microphone’s signal. Figure 5.2 shows the center array microphone’s signal vs. that of the reference microphone with only the source speaker turned on.
The reference microphone and center array microphone are roughly equidistant from the source speaker (60” and 67” respectively). Note that their autospectra are roughly the same magnitude for frequencies below 5 kHz. As the source becomes more directional at higher frequencies, the two microphones’ signals change slightly in form and the center array microphone’s magnitude is noticeably greater (above ~4 kHz). This is attributed to the microphone’s polar plot i.e. sensitivity variation as a function of angle of incidence, as shown in Figure 5.3.
Figure 5.4 shows the center array microphone’s signal vs. that of the reference microphone with only the background speaker on (ref. Fig. 4.5 for geometry).

With the background speaker on only it is seen that due to the difference in proximity the reference microphone’s level is ~20 dB greater than that of the center array microphone. Using “one-over-r” calculations this difference in sound power shows that the chamber upholds anechoic conditions [47]:

Reference Mic radial distance from background speaker = 5"
Center Array Mic radial distance from background speaker = 47"

\[ \text{Pressure}^2 \sim \frac{1}{\text{Radius}^2} \quad (5.1) \]

\[ \text{Sound Pressure Level (dB)} = 10 \times \log_{10} \left( \frac{\text{pressure}_1^2}{\text{pressure}_2^2} \right) \quad (5.2) \]

\[ 10 \times \log_{10} \left( \frac{\frac{1}{5^2}}{\frac{1}{47^2}} \right) = 19.46 \text{ dB} \]

Coherence between the signal received at the reference microphone and that received at the array is a critical parameter to ensure good noise cancellation [10]. The coherence between two signals, \( x(t) \) and \( y(t) \), is defined mathematically as [45]
\[
C_{xy}(\omega) = \frac{|G_{xy}(\omega)|^2}{G_{xx}(\omega)G_{yy}(\omega)}
\]  
(5.3)

where \(G_{xx}\) represents the one-sided auto-spectral density function for signal \(x\) and \(G_{xy}\) the one-sided cross-spectral density between signals \(x\) and \(y\) [52].

The coherence between the two signals was computed using the Matlab function `mscohere`, which uses Welch’s modified, averaged periodogram method to calculate the magnitude squared coherence of the inputs as a function of frequency [53]. Values are between 0 and 1 and indicate how well the signals correspond at each frequency, 1 being perfectly and 0 not at all.

Figure 5.5 shows the coherence between the reference microphone and the center array microphone when only the background speaker is turned on.

![Figure 5.5 - Coherence between reference and center array microphone with background speaker on.](image)

The average coherence in the frequency band 0.5-9.5 kHz is 0.99. This is expected as there is no other noise interfering with the background speaker output. Next, the coherence is shown for the three SNRs mentioned previously.
Figure 5.6 - Coherence between reference and center array microphone with background and source speaker both on for three SNRs.

Table 5.1 shows the mean coherence between 0.5-9.5 kHz for each SNR shown in Figure 5.6.

Table 5.1 – SNR values vs. coherence in frequency band 0.5-9.5 kHz for background and source speaker both turned on.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>Mean Coherence (0.5-9.5 kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6.5</td>
<td>0.80</td>
</tr>
<tr>
<td>-11.4</td>
<td>0.91</td>
</tr>
<tr>
<td>-19.9</td>
<td>0.99</td>
</tr>
</tbody>
</table>

From Figure 5.6 and Table 5.1, as the source speaker’s output level drops and the SNR decreases, the coherence between the reference and middle array microphone increases. This result is expected as the source speaker’s influence on the array mics will decrease as its output level is lowered.

The Matlab Least Mean Squares (LMS) Filter block was chosen to remove the noise from the measurements in post-processing. The block implements an adaptive Finite Impulse Response (FIR) filter using an LMS algorithm (ref. Fig. 3.6-9 and Eqs. (3.13-24)). The block estimates the filter weights needed to minimize the error between the output and desired signal. The error port outputs the result of subtracting the output signal (filtered input, estimate of
desired signal) from the desired signal. Figure 5.7 shows the Matlab block diagram. For reference see [10] and [54].

![Matlab LMS Filter Block Diagram](image)

**Figure 5.7** - Matlab LMS filter block.

Matlab’s LMS algorithm is defined as

\[
    y_n = \mathbf{w}_{n-1}^T \mathbf{u}_n
\]

(5.4)

\[
    e_n = d_n - y_n
\]

(5.5)

\[
    \mathbf{w}_n = \mathbf{w}_{n-1} + f(\mathbf{u}_n, e_n, \mu)
\]

(5.6)

where the weight update function, \( f \), is defined as

\[
    f(\mathbf{u}_n, e_n, \mu) = \mu e_n \mathbf{u}_n^* \]

(5.7)

which is a slightly modified version of Eq. (3.22).

Table 5.2 explains the variables used in the LMS algorithm.

<table>
<thead>
<tr>
<th>( n )</th>
<th>The current time index</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{u}_n )</td>
<td>The vector of buffered input samples at step ( n )</td>
</tr>
<tr>
<td>( \mathbf{u}_n^* )</td>
<td>The complex conjugate of the vector of buffered input samples at step ( n )</td>
</tr>
<tr>
<td>( \mathbf{w}_n )</td>
<td>The vector of filter weight estimates at step ( n )</td>
</tr>
<tr>
<td>( y_n )</td>
<td>The filtered output at step ( n )</td>
</tr>
<tr>
<td>( e_n )</td>
<td>The estimation error at step ( n )</td>
</tr>
<tr>
<td>( d_n )</td>
<td>The desired response at step ( n )</td>
</tr>
<tr>
<td>( \mu )</td>
<td>The adaptation step size</td>
</tr>
</tbody>
</table>

Two parameters affect Matlab’s LMS filter output: the filter weight vector length, \( \mathbf{w} \), and the adaptation step size, \( \mu \). The weight vector length affects the accuracy of the output; the longer it is (i.e. more weights), the more accurate. The adaptation step size affects the speed at which the filter converges; the larger it is, the quicker it converges. It should be noted that these
parameters are not limitless i.e. they cannot be arbitrarily maxed to ensure both high accuracy and rapid convergence (the ideal situation). Beyond the filter limits (Eq. 3.23) the filter response diverges and no solution is obtained. Depending on the data, an optimum combination exists such that while increasing $\mu$ will shorten the convergence time it will decrease accuracy and increasing $w$ will not lead to more accurate results at a fixed $\mu$.

The key result for judging the performance of the processing method (accuracy wise) is the Mean-Square Error (MSE) between the target and recovered data. It is defined as [45]

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (x_n - x'_n)^2 \quad (5.8)$$

where $x$ represents the original signal and $x'$ represents the signal being compared to it. In the ANC analysis the target data (desired) is considered that without the background noise speaker turned on. As the SNR between the signals decreases, recovery of the source signal (desired) becomes more difficult because it becomes further “buried” in the background noise.

The investigation of the robustness of the method is presented as a succession of the decreasing SNRs with an optimum $w$ and $\mu$ combination fixed. Afterward, variations of the $w$ and $\mu$ combination are shown.

**SNR -6.5 dB**

The first step in the recovery process is to run each array microphone’s data through the LMS filter with the reference microphone as the input (ref. Fig. 5.7). Both signals come from the same acquisition with both the background and source speaker turned on at one of the three SNRs. Because the data acquisition system had a built in anti-aliasing filter, the data was Low Pass filtered at 9.5 kHz before being fed into the LMS filter to ensure that signal information from the frequency range that was filtered by the anti-aliasing filter would not skew the ANC processing.

To strike an optimum balance between performance and accuracy, a filter length of 1024 and an adaptation step size of 0.1 were chosen through trial by iteration.

Figure 5.8 shows a graph of four autospectra. The signal plotted in green is the one corresponding to the background and source speakers turned on, the signal in red is the source speaker only turned on, the blue line is the recovered signal, and the black is the noise floor of the chamber, included for reference. The spectra are averaged over the 9 array microphones.
In the frequency band of interest, 0.5-9.5 kHz, the ‘Recovered’ signal’s MSE is 2.66 as compared to the ‘Source Only’ signal. The peaks of the ‘Source Only’ line are matched almost perfectly by the ‘Recovered’ signal and the valleys are where it has the most difficulty in recovering due to the lower SNR.

Next, the recovered signals at each microphone were put into a beamform code (adapted from [26]) to produce Figure 5.9. The frequency chosen (5.95 kHz) corresponds to a peak in the ‘Source Only’ signal where accurate recovery was seen. The x-axis of the beamform plot represents an 8’ horizontal line running through the source speaker with its center (0 ft) at the speaker.
Figure 5.9 – Recovered signal beamform (5.95 kHz, SNR = -6.5 dB, MSE = 1.02, \(w = 1024\), \(\mu = 0.1\)).

Figure 5.9 displays accurate beamform recovery: a MSE of 1.02 over 8’ line. The large lobe in the ‘Background + Source’ signal present in the far right of the plot (~3.8’) corresponds to the direction of the background speaker (ref. Fig. 4.5). The ‘Recovered’ signal achieves an 11.1 dB reduction in sound from that direction. Also note that the lobe structure and magnitude of the ‘Background + Source’ signal has been almost perfectly corrected by the ‘Recovered’ signal with the exception of the null near -3’ and the far right lobe.

**SNR -11.4 dB**

The process detailed for the first SNR of -6.5 dB was repeated for the SNR of -11.4 dB. The same parameters were kept for weight vector length and adaptation step size. The autospectrum with recovered results is shown in Figure 5.10.
Figure 5.10 – Recovered signal autospectrum (SNR = -11.4 dB, MSE = 3.64, w = 1024, µ = 0.1).

Figure 5.10 displays accurate recovery with some difficulty at the valleys and more success at the peaks. Compared to the higher SNR seen previously, the MSE over the frequency band 0.5–9.5 kHz has increased to 3.64.

Figure 5.11 gives the corresponding beamform. The frequency chosen is the same as the previous SNR case, 5.95 kHz, and remains the same throughout the ANC analysis.

Figure 5.11 – Recovered signal beamform (5.95 kHz, SNR = -11.4 dB, MSE = 1.50, w = 1024, µ = 0.1).
Figure 5.1 displays accurate recovery again, although there is a noticeable decrease in accuracy compared to the previous SNR’s beamform (Fig. 5.9) as proven with a slight increase in MSE to 1.50. The dB reduction has increased to 16.7 in the far right lobe, which is expected because the SNR decrease is due to decreasing the source speaker’s output. Thus, for a constant output from the background speaker (see far right lobe levels in Figures 5.9, 11, 13) the ‘Source Only’ level drops as the SNR decreases. If the ‘Recovered’ signal follows the trend (i.e. sustains accuracy), the far right side lobe dB reduction will increase with decreasing SNR.

**SNR -19.9 dB**

The lowest SNR is presented last, following the same procedure used for the two higher SNRs seen previously.

Figure 5.12 – Recovered signal autospectrum (SNR = -19.9 dB, MSE = 6.61, \(w = 1024, \mu = 0.1\)).

Figure 5.12 shows that the recovery of the lowest SNR suffers slightly compared to the higher SNRs. The trend remains in that better recovery is seen at the peaks as compared to the valleys. The MSE has increased to 6.61, the highest of all three SNR autospectra cases presented. The corresponding beamform is given in Figure 5.13.
The lowest SNR shows a much distorted beamform for the ‘Background + Source’ signal. This is due to the fact that the source level was so overwhelmed (i.e. overpowered) by the background noise. As such, no clear main lobe appears at X = 0’ (source speaker location). A 25.1 dB reduction is seen in the far right lobe, the greatest of all the SNRs investigated. Although the lobe structure is still well maintained at this SNR, the magnitude suffers slightly which is seen in the calculated MSE of 4.43, also the largest of the three SNR beamform cases investigated.

**SNR -19.9 dB – Variations**

The following figures illustrate the effect of varying the LMS filter parameters on the processing output. The lowest SNR was chosen arbitrarily and the filter length, w, and step size, μ, were adjusted in combinations above and below the optimum (w = 1024, μ = 0.1) by an order of magnitude (when algorithm convergence permitted). The figures are presented in order of descending accuracy, from lowest to highest MSE between the ‘Source Only’ and ‘Recovered’ signals.

The first combination is w = 2048 and μ = 0.1. An order-of-magnitude filter increase was attempted but with the step size of 0.1 the filter output diverged. Thus, a doubling of the weight vector was chosen, which still satisfied the condition that the vector length be a power of 2. Note
that characteristically this increase in weight vector length should result in a more accurate output. Figure 5.14 displays the resulting array averaged autospectra.

![Figure 5.14 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 16.90, w = 2048, µ = 0.1).](image)

The ‘Recovered’ signal seen in Figure 5.14 is the most inaccurate yet. The MSE confirms this: 16.90, more than double that achieved with the optimum filter settings at this SNR.

The corresponding beamform is presented in Figure 5.15.
Figure 5.15 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 7.70, \( w = 2048, \mu = 0.1 \)).

The ‘Recovered’ beamform results seen in Figure 5.15 display a loss in accuracy (MSE = 7.70) compared to Figure 5.13. This is also reflected in the increase in the far right lobe attenuation to 23.7 dB, a 1.3 dB increase compared with Figure 5.13.

The second combination is a decrease of the weight vector compared to optimum while holding the step size fixed: \( w = 128 \) and \( \mu = 0.1 \). Figure 5.16 displays the autospectra.

Figure 5.16 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 23.25, \( w = 128, \mu = 0.1 \)).
The ‘Recovered’ signal’s spectrum seen in Figure 5.16 is the least accurate thus far (MSE = 23.25). While peaks in the ‘Source Only’ autospectrum are recovered well, the valleys are not well matched in magnitude or shape. Also note the rippled distortion that exists below ~1.5 kHz.

The beamforming results for this combination are presented in Figure 5.17.

**Figure 5.17** - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 15.31, \( w = 128, \mu = 0.1 \)).

The beamform recovery continues to become less accurate. The MSE, 15.31, has more than doubled from that seen in the previous combination (Fig. 5.15). The side lobe magnitude and shape recovery is the most inaccurate yet. This is reflected in the slight increase of the far right side lobe reduction to a level of 22.7 dB.

The third combination fixes the weight vector length to that of the optimum and decreases the step size, corresponding to a slower convergence rate: \( w = 1024 \) and \( \mu = 0.01 \). The autospectra are presented in Figure 5.18.
Figure 5.18 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 33.41, \( w = 1024, \mu = 0.01 \)).

Figure 5.18, while not displaying the rippled distortion seen in Figure 5.16, has a ‘Recovered’ signal with a very large dB variation as seen by the thick blue line. This results in the highest autospectrum MSE seen up to this point, 33.41.

The beamform is presented in Figure 5.19.

Figure 5.19 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 34.19, \( w = 1024, \mu = 0.01 \)).
The ‘Recovered’ beamform of Figure 5.19 shows a more accurate side lobe structure than Figure 5.17 with the exception of the background speaker direction (positive x-axis). The MSE has increased by more than double again to 34.19. The far right lobe attenuation has decreased again to 19.2 dB.

The last combination holds the weight vector length fixed but increases the step size: \( w = 1024 \) and \( \mu = 0.3 \). The autospectrum is provided in Figure 5.20.

![Figure 5.20 - Recovered signal autospectrum (SNR = -19.9 dB, MSE = 52.59, \( w = 1024 \), \( \mu = 0.3 \)).](image)

The ‘Recovered’ spectrum is seen to replicate the shape of the ‘Source Only’ signal quite well. However, its magnitude is grossly off, as indicated by highest MSE of 52.59.

The beamform follows in Figure 5.21.
Figure 5.21 - Recovered signal beamform (5.95 kHz, SNR = -19.9 dB, MSE = 25.37, w = 1024, µ = 0.3).

As is the case with its autospectrum, the ‘Recovered’ signal’s beamform accurately replicates the ‘Source Only’ signal’s shape but does not match its magnitude. The MSE is high, second only to that seen in Figure 5.19, 25.37. Its improved MSE also leads to an improved far right lobe attenuation with respect to Figure 5.19, 21.3 dB.

Table 5.3 provides comparative and summarizing parameters of the ANC investigation.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>w</th>
<th>µ</th>
<th>MSE Autospectrum</th>
<th>MSE Beamform</th>
<th>Reduction from Reflection Direction (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6.5</td>
<td>1024</td>
<td>0.10</td>
<td>2.66</td>
<td>1.02</td>
<td>11.1</td>
</tr>
<tr>
<td>-11.4</td>
<td>1024</td>
<td>0.10</td>
<td>3.64</td>
<td>1.50</td>
<td>16.7</td>
</tr>
<tr>
<td>-19.9</td>
<td>1024</td>
<td>0.10</td>
<td>6.61</td>
<td>4.43</td>
<td>25.1</td>
</tr>
<tr>
<td>-19.9</td>
<td>2048</td>
<td>0.10</td>
<td>16.90</td>
<td>7.70</td>
<td>23.7</td>
</tr>
<tr>
<td>-19.9</td>
<td>128</td>
<td>0.10</td>
<td>23.25</td>
<td>15.31</td>
<td>22.7</td>
</tr>
<tr>
<td>-19.9</td>
<td>1024</td>
<td>0.01</td>
<td>33.41</td>
<td>34.19</td>
<td>19.2</td>
</tr>
<tr>
<td>-19.9</td>
<td>1024</td>
<td>0.30</td>
<td>52.59</td>
<td>25.37</td>
<td>21.3</td>
</tr>
</tbody>
</table>
Table 5.3 illustrates that although recovery was successful (due to the MSEs seen) at each SNR tested, accuracy of the recovered signal decreased with SNR. This is expected as the signal to be recovered is being overpowered by the background noise to a greater degree at each decrease of the SNR. The reduction in dB of the far right lobe in the beamforming plots, corresponding to the direction in which the background noise was sensed by the array, increased with decreasing SNR due to the fact that the lower SNRs were achieved by lowering the source speaker’s output. The variations at the SNR of -19.9 dB supported the fact that the chosen weight vector length/adaptations step size combination was indeed optimum. In terms of average MSE between the autospectrum and beamform recovery, the most accurate (MSE wise) variation was a doubling of the weight vector length while holding the step size fixed to that of the optimum, and the least accurate was increasing the step size by a factor of three with the weight vector length held at optimum. Finally, it was shown that an increase in the beamform MSE lead to an increase in the far right side lobe attenuation.

5.2 Cepstrum Reflection Removal

This section is divided into two parts. The first discusses the results of a simulation performed to verify the algorithms used to process the experimental results. The second details the processing and results of the experimental data gathered in the Cepstrum test.

Simulation Results

The Cepstrum processing was carried out with three Matlab functions: rceps for the Power Cepstrum, cceps for the Complex Cepstrum, and icceps for the Inverse Complex Cepstrum. The block diagrams showing the process used in the analysis for the Power Cepstrum is seen in Figure 5.22. The Complex Cepstrum as well as recovery is given in Figure 5.23.
A simulation was created in order to provide a simple processing case and to display ideal Cepstrum features. A Simulink block was used to produce a simulated direct signal spike and one reflection on a signal channel. The block is pictured in Figure 5.24.

The direct signal pulse (recorded on ‘Scope’ in Fig. 5.24) was set with a magnitude of 0.95. This corresponded to simulating the fully anechoic condition in the chamber. The reflection pulse, which had a magnitude of 0.08 and delay of 150 samples after the direct pulse, was added to the direct pulse and recorded on ‘Scope1’, recording the signal simulating the anechoic wedges being removed from the chamber floor. This signal corresponds to Eq. (3.26). The simulation sample rate was 40,000 samples/second and a sample time of 0.1024 seconds.

The time histories of the two Scopes are shown in Figure 5.25.
Figure 5.25 - Simulation time histories. (a) 'Scope' signal simulating one direct pulse, and (b) 'Scope1' signal simulating one direct pulse and one reflection pulse at a lower amplitude.

Figure 5.26 shows the magnitude and phase plots of both the non-reflection (‘Scope’) and reflection (‘Scope1’) signals.
Figure 5.26 – (a) Magnitude of simulated non-reflection and reflection signals, and (b) Phase.

The signals shown in Figure 5.26 are ideal. The non-reflection data (red) has a flat spectrum across the frequency band. The reflected signal (green) shows consistent constructive/destructive interference across the samples taken.

The main purpose behind simulating a signal is to produce a “perfect” signal i.e. without interference of any kind (ref. Eq. (3.26)). Example interference could be noise in the acquisition channels or a distorted reflection brought about by a non-rigid and/or porous reflecting surface. Without interference, the Power Cepstrum identifies only the direct signal, the reflection, and its
rahmonics, as per Eq. (3.31) and seen in Figure 5.27. At all other times on the quefrency scale, the power of the signal is zero. These are ideal characteristics of a simulated signal.

Figure 5.27 – Power Cepstrum of simulated reflection data. (a) Full length Cepstrum, and (b) Zoomed version of (a) to emphasize the rahmonic content with direct signal, reflection, and first rahmonic in view.

From Figure 5.27b, the echo delta is seen to occur at quefrency bin 151 (3.775e-3 seconds). The rahmonics of the echo appear at 150 quefrency bin intervals. Each rahmonic in the low quefrency end of the Cepstrum has an equivalent high quefrency component due to the fact
that the Power Cepstrum is an even function [44]. Note that as the echo is perfect, and thus not distorted or convolved with noise, the echo delta and rahmonics are all confined to one quefrency bin i.e. their energy is not spread out over multiple bins and the deltas decay as per Eq. (3.31).

Next, the Complex Cepstrum is performed on the simulated reflection signal (Eqs. (3.26, 34-37)).

**Figure 5.28** – Complex Cepstrum of simulated signal. (a) Full view of entire quefrency scale, and (b) Zoomed version of (a) to enlarge low quefrency end.
Figure 5.28 displays two views of the Complex Cepstrum of the simulated reflection data. The direct signal, reflection, and first rahmonic are clearly seen in 5.28b. Note the decaying nature, opposite sign, and one quefrency bin confinement of the deltas as per Eq. (3.37). Again the echo delta is seen at quefrency bin 151 and the rahmonics appear in intervals of 150 quefrency bins, in agreement with the Power Cepstrum.

With the locations of the reflection and its rahmonics identified, the Complex Cepstrum was liftered (ref. Chap. 3). The liftering was accomplished by replacing the value of the Complex Cepstrum at a reflection and its rahmonics (including high quefrency components) with the average of the preceding and subsequent values, know as interpolation. After liftering, the signal was run through the Inverse Complex Cepstrum (block 3, Fig. 5.23) and plotted with the original reflection and non-reflection magnitudes and phases as shown in Figure 5.29.
Figure 5.29 – (a) Magnitude of non-reflection, reflection, and recovered signals, and (b) Phase.

Figure 5.29 demonstrates that the Matlab procedure, with proper user-performed liftering, of recovering a signal by removing reflections through Cepstral processing provides perfect magnitude and phase recovery. This is seen in that the blue line (‘Recovered’) exactly masks the red (‘No Reflection’) in both plots. This magnitude and phase recovery is necessary for accurate array techniques such as beamforming which depends on the proper recovery of both.

**Experimental Results**

The data and results processed in this section were gathered in the test setup shown in Figure 4.13. The acquisition parameters were: sampling frequency of 20,000 samples per second, 30 blocks, and block size of 2048 samples. The bandwidth for all Cepstrum data presented is 9.77 Hz.

The reflections in Cepstral processing are identified by their arrival time relative to the direct sound’s arrival from the source. To further increase confidence in reflection identification, the echo arrival time can be calculated based on the measured dimensions of the test setup shown in Figure 4.13. The calculated echo arrival time (delay) is

\[
delta \text{echo arrival time} = \frac{\text{reflected–direct path}}{\text{speed of sound}} = \frac{107.24^\circ - 76^\circ}{\frac{13397.24 \text{ in}}{\text{sec}}} = 2.13 \text{ msec} \quad (5.9)
\]
First, an autospectrum of signals taken at the center array microphone (5) is shown in Figure 5.30 in order to display the constructive/destructive interference present due to the floor when the floor wedges were removed. White noise was used as the signal driving the speaker.

![Autospectrum](image)

**Figure 5.30** – Center array microphone autospectrum (no reflection vs. reflection present).

The frequency gap in subsequent valleys in the reflection’s autospectrum is related to the echo arrival time [44]

$$\Delta \text{freq.} = \frac{1}{\text{Echo Arrival Time}}$$  \hspace{1cm} (5.10)

Figure 5.30 confirms the arrival time delay calculated previously from the test set up

$$\text{Echo Arrival Time} = \frac{1}{469 \text{ Hz}} = 2.13 \text{ msec}$$

For a perfect reflection with the same radial distance to the microphone, the valleys would sink to negative infinity and the peaks would be 6 dB above the non-reflection data. Due to the difference in distance and non-perfect reflecting surface, this will change. Thus, if the pressure of the source speaker were $p_0$, a perfect reflection would represent another speaker of the same power, at a distance slightly farther away. Not taking into account any distortion, the dB output of the speaker without reflection would be

$$dB_1 = 20 \times \log_{10} \left( \frac{p_0}{78} \right)$$  \hspace{1cm} (5.11)

When their signals interfere destructively
When their signals interfere constructively,

\[ dB_{construct} = 20 \times \log_{10} \left( \frac{p_0}{78} + \frac{p_0}{106.5} \right) = dB_1 + 4.8 \]  

These maximum values are seen to within one dB at the low frequency peaks and valleys in Figure 5.30. The level of amplification/attenuation at the higher frequency peaks/valleys decreases due to the more directional nature of sound at higher frequencies.

Next, the average coherence between the array microphones for the non-reflection case is shown in Figure 5.31.

![Figure 5.31](image)

**Figure 5.31** – Average coherence between array microphones (no reflection).

Figure 5.31 illustrates that the regions before 0.5 kHz and after 9.5 kHz should be avoided in processing because the coherence drops off significantly, due to the High Pass (HP) and anti-aliasing filters respectively.

Now, the coherence between array microphones 1 and 5 is displayed in Figure 5.32, this time with the reflection due to the removal of the floor wedges present.
In Figure 5.32, the coherence suffers at the same frequencies that the reflected signal interferes destructively (valleys) with the direct signal.

Next, the Power Cepstrum of the center array microphone signal is shown in Figure 5.33. The Power Cepstrum is more useful than the Complex Cepstrum in detecting the echo arrival times and its rahmonics, as it squares its output which further separates the magnitude of deltas due to reflection(s) from signal noise in the Cepstrum domain (ref. Eq. 3.30 vs. 34).
Figure 5.33 - Power Cepstrum of microphone 5.

In Figure 5.33 the reflection is seen as the first spike after the initial group of spikes against the left side of the x-axis. Note that all the deltas in the Cepstrum have corresponding high quefrency components that are reflections about the midpoint.

The following figure, 5.34, is the Power Cepstrum presented in 5.33 zoomed in on the low quefrency end to provide details of the direct signal, the reflection, and its first two rahmonics. Note that the reflection is not contained in one quefrency bin and has leaked into three others. This leakage suggests a distorted echo and perturbations due to noise in the signal can be seen in the low level quefrency bins after the reflection [31]. Also, the rahmonics are barely visible as their leakage into adjacent quefrency bins has reduced their amplitude, as shown in Figure 5.34b. These characteristics suggest the signal is of the form given by Eq. (3.45). Note the clear differences between the ideal signal’s Power Cepstrum (Fig. 5.27b) and that of the experimental data.
The deltas seen in the lower quefrency region (bins 0-13) in Figure 5.34a are related to the direct signal. The largest echo component occurs at quefrency bin 42 and the second greatest amplitude is seen in bin 43. The echo arrival time can be calculated from the average of these bins

$$\frac{(42+43)}{2} \times (5e - 5) = 2.125 \text{ msec}$$  \hspace{1cm} (5.14)

This confirms the estimates from Eqs. (5.9, 10).
Figure 5.35 shows the Complex Cepstrum of the center array microphone signal zoomed in on the low quefrency section. The Complex Cepstrum will be used to recover the signal with the reflection effects removed.

![Complex Cepstrum of microphone 5, zoomed in to low quefrency end.](image)

Note that no clear reflection peak is seen at quefrency bin 42. Referring to Eq. 3.50, the Complex Cepstrum of the experimental data gives evidence that both echo distortion and noise are present in the signal. The distortion is seen in Figure 5.34 due to the echo energy being spread over four quefrency bins and in Figure 5.35 due to the reflection barely emerging above the noise deltas in the signal (compare Fig. 5.35 to the ideal Complex Cepstrum of Fig. 5.28b). The low gamnitude perturbations seen in the quefrencies after the direct signal deltas and in between the reflection and its rahmonics represent noise which will weaken the echo as seen in Figure 5.35. This noise is random in nature, may overlap the direct signal components in the Cepstrum domain, and is convolved with the reflection deltas (Eq. 3.46). This produces a reflection signal that differs from a non-reflection signal at more quefrency bins than just those containing energy due to the reflection and its rahmonics. Thus proper removal of the reflection through liftering does not ensure recovery of the non-reflection signal. This is an important aspect to be mindful of when working with data exhibiting similar features in the Cepstral domain.
According to the echo times given by the Power Cepstrum, the Complex Cepstrum was liftered. Two types of liftering were used to compare the effectiveness in signal recovery: block liftering and interpolation (see Chap. 3). Block liftering is attractive because it ensures that all reflection content will be removed (relying on correctly chosen quefreny blocks) and it can be automated to reduce processing time. Interpolation in theory is more accurate as less of the data is being modified, reducing the chances that information belonging to the direct signal in the Cepstrum domain will be altered.

The success of the beamforming, similar to the ANC results, will be judged on three factors: the average coherence of the recovered signal between the array microphones, the MSE of the recovered autospectrum and beamform, and the dB attenuation achieved by the recovered beamform at the reflection location.

**First Liftering Scheme: Block Liftering, One Point Zeroed**

The first three liftering schemes presented are block liftering. The first is the most minimal; it only zeroed the highest reflection delta and its corresponding high quefreny component. The second zeroed the highest reflection delta and the three other points around it where the reflection had leaked. The third zeroed from the start of the reflection until 100 quefreny bins later, roughly corresponding to the 2\textsuperscript{nd} rahmonic.

The autospectrum of the center array microphone is shown in Figure 5.36 displaying the ‘Reflection’ signal (green), No Reflection’ signal (red), ‘Recovered’ signal (blue), and ‘Noise Floor’ (black) for reference.
Figure 5.36 - Microphone 5 autospectrum: no reflection, reflection, and recovered (1 point zeroed).

The calculated MSE over the desired frequency range, 0.5-9.5 kHz, between the ‘Recovered’ and the ‘No Reflection’ data is 5.71 as compared to the ‘Reflection’ vs. ‘No Reflection’ which is 7.77; an improvement is realized, but not a significant one. The processing had little effect in removing the ‘Reflection’ peaks and valleys. The coherence between array mics of the recovered data was calculated to be 0.75. This is a non-accurate recovery, referencing that the coherence for the reflection signal is 0.97 and that of the non-reflection signal 0.99.

A frequency at which accurate magnitude recovery was seen was chosen to perform the beamforming. The recovered data beamform at this frequency is presented in Figure 5.37.
In Figure 5.37, it is seen that the ‘Recovered’ beamform doesn’t match the levels of the ‘No Reflection’ beamform at every valley and half of the peaks. At the reflection location a 1.63 dB reduction is seen in the recovered data. The MSE between the ‘Recovered’ and ‘No Reflection’ data is 27.08, reflecting the recovery inaccuracy.

In order to identify the importance of accurate phase recovery for the beamform processing, a phase correction procedure was devised as pictured in Figure 5.38.

The phase corrected recovered data at the same frequency as Figure 5.37 is given in Figure 5.39.
Figure 5.39 - Beamform: recovered data, phase corrected (1 point zeroed, 8682 Hz).

Figure 5.39 shows improvement in lobe structure over 5.38 due to the phase correction. A 6.33 dB reduction is seen at the reflection location, up from 1.63, and the MSE was reduced to 10.22 from 27.08, a 62% improvement. These improvements in reflection location dB attenuation and MSE demonstrate the significance of accurate phase recovery for array processing such as beamforming.

**Second Lifting Scheme: Block Lifting, Four Points Zeroed**

The next set of results employed a four point block lifter on the Complex Cepstrum data to obtain a recovered signal. The autospectrum of the center array microphone is shown with the recovered data in Figure 5.40.
The MSE between the ‘Recovered’ and the ‘No Reflection’ data is 3.46, more accurate than the 1 point block liftering scheme. Comparing the ‘Reflection’ spectrum to the ‘Recovered’, the four points zeroed liftering processing had more success in removing the reflections than the one point zeroed results, but improvement can be made at the destructive interference frequencies. The coherence between array microphones of the ‘Recovered’ data was calculated to be 0.56, lower than that seen with only one point zeroed.

Referencing Figure 5.38, the same process was used to correct the phase of the recovered signal. The same guidelines were used to choose an appropriate beamforming frequency. Figure 5.41 shows a beamform plot of the recovered data without phase correction.
In Figure 5.41, the ‘Recovered’ beamform hardly resembles the ‘No Reflection’ beamform. Only the lobe to the right of the main lobe is matched in magnitude to the ‘No Reflection’ data and no matching lobe structure is seen. The source strength (Y=0 on x-axis) is estimated more inaccurately than what is mapped by the ‘Reflection’ data. And only a 1.1 dB reduction is seen at the reflection location. The MSE between the ‘Recovered’ and ‘No Reflection’ data is 44.77, a large increase compared to what was calculated for the 1 point zero lifting scheme.

Figure 5.42 presents the same results with the phase correction applied.
Figure 5.42 shows the ‘Recovered’ signal to have an almost equal lobe structure to that of the ‘No Reflection’ signal. Although the source level is not matched, slight improvement is seen over the reflection data, and certainly over the non-phase-corrected recovery seen in Figure 5.41. A 4.6 dB reduction is seen at the reflection location and the MSE was reduced, as compared to the previous lifting scheme (Fig. 5.39), to 2.86.

**Third Lifting Scheme: Block Lifting, One Hundred Points Zeroed**

The final block lifting scheme was the most intrusive: 100 points on each end of the quefreny axis set to zero. The autospectrum of the center array microphone is shown with the recovered data in Figure 5.43.
The MSE in Figure 5.43 over 0.5-9.5 kHz between the ‘Recovered’ and the ‘No Reflection’ data is 2.69, the lowest of the three block lifter recovery schemes. Some of the valleys could be more accurate but all peaks of the ‘Recovered’ signal have been matched to the ‘No Reflection’ values. The coherence between array microphones of the recovered data was calculated to be 0.31, the lowest seen among the schemes, due to more information being removed in the liftering.
The ‘Recovered’ beamform doesn’t resemble the ‘No Reflection’ beamform in any way. No lobe level is matched. At the reflection location an increase of 0.5 dB exists. The MSE sums up the inaccuracy of the recovered beamform, 73.20, the highest yet.

Figure 5.45 presents the phase corrected recovered data at the same frequency.
Figure 5.45 shows an equal lobe structure to that of the non-reflection signal, with some error in dB levels present. The source level is not matched but a large improvement is seen over the non-phase-corrected data (Fig. 5.44). An 11.1 dB reduction is seen at the reflection location (an overshoot of 1.6 dB) and the calculated MSE is 17.40.

**Fourth Liftering Scheme: Interpolation**

As defined in Chapter 3, interpolation was carried out as the last liftering scheme. Ideally, only one bin for each delta must be interpolated (see simulation results). However, referencing Figure 5.34, the echo delta was spread into 4 quefrency bins. Thus, a linear interpolation was carried out over the four bins, using the previous and subsequent Complex Cepstrum values as start and end points. Note that only these four points belonging to the reflection and their high quefrency components were interpolated in each microphone signal, as rahmonic content was indistinguishable from the noise present in the Cepstrum domain (ref. Fig. 5.34). Figure 5.46 gives the recovered autospectrum from microphone 5.

\[\text{Figure 5.46 - Microphone 5 autospectrum: no reflection, reflection, and recovered (interpolation).}\]

The ‘Recovered’ data does not match the ‘No Reflection’ data well at the valleys (destructive interference). Although The MSE of the ‘Recovered’ data is 3.47 compared to that of the ‘Reflection’ data, 7. The array coherence was calculated to be 0.45.
Figure 5.47 presents the beamform plot with the reflection, non-reflection, and recovered signals present.

![Beamform plot](image)

**Figure 5.47 - Beamform: recovered data (interpolation, 8242 Hz).**

Figure 5.47 gives similar results as the block liftering beamforms have: the ‘Recovered’ beamform structure does not mimic that of the ‘No Reflection’ data in shape or magnitude. A small reduction at the reflection location is realized: 1.2 dB. The MSE is 42.10.

Figure 5.48 gives the phase corrected version.
Again, Figure 5.48 is similar to what was seen for the phase corrected block lifting results. The phase corrected ‘Recovered’ beamform replicates the ‘No Reflection’ data’s well and slight magnitude error is seen. The MSE has been reduced as compared to Figure 5.47 to 3.05. A 3.95 dB reduction is seen at the source location.

Table 5.4 provides the key parameters as a means to summarize the recovery method investigation.
Table 5.4 - Cepstrum recovery key results.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Array Coherence (Recovered Signal)</th>
<th>MSE Array Autospectrum</th>
<th>MSE Center Microphone Autospectrum</th>
<th>MSE Beamform</th>
<th>Reduction @ Reflection Location (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 point zeroed</td>
<td>0.75</td>
<td>5.71</td>
<td>27.08</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>4 points zeroed</td>
<td>0.56</td>
<td>3.46</td>
<td>44.77</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>100 points zeroed</td>
<td>0.31</td>
<td>2.69</td>
<td>73.20</td>
<td>+0.50</td>
<td></td>
</tr>
<tr>
<td>Interpolation</td>
<td>0.45</td>
<td>3.47</td>
<td>42.10</td>
<td>1.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>10.22</strong></td>
<td><strong>6.33</strong></td>
<td></td>
</tr>
<tr>
<td>1 point zeroed (phase corrected)</td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>2.86</strong></td>
<td><strong>4.60</strong></td>
<td></td>
</tr>
<tr>
<td>4 points zeroed (phase corrected)</td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>17.40</strong></td>
<td><strong>11.10</strong></td>
<td></td>
</tr>
<tr>
<td>100 points zeroed (phase corrected)</td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>3.37</strong></td>
<td><strong>3.95</strong></td>
<td></td>
</tr>
<tr>
<td>Interpolation (phase corrected)</td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td><strong>--</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4 demonstrates trends for the block liftered recovered data. It is seen that as the number of points zeroed in the Cepstrum domain increases, the array coherence, autospectrum MSE, and dB attenuation at the reflection location all decrease, and the beamform MSE increases. The interpolation data falls in the middle of the three block liftering schemes in every aforementioned category. These results suggest that the type of liftering to be implemented should depend on the desired processing. If only an autospectrum is needed, using block liftering with a large data block to be zeroed would be recommended. If array processing is necessary, then applying a 1 point zeroed block lifter would be the route of choice.
That said, it must be stated that the beamform MSEs of all the liftering techniques suggest that none of them are effective in recovering the necessary information to render them effective tools for beamforming. This is tied to the noise present that is convolved with the reflection energy in the Cepstrum domain. Lifting is useful when energy due to reflections can be replaced with the equivalent values of the non-reflection signal. When reflection energy is spread and modulated due to distortion and noise and/or when the noise present is random, the values in the Cepstrum domain used for interpolation will be inaccurate replications of those values the non-reflection signal exhibits, leading to inaccurate recovery. The phase correction procedure was used to illustrate the importance of the phase in accurate beamform recovery.

The phase correction results, while displaying improvement in accuracy, have two inherent problems. First, as the magnitude of the recovered signal is matched with a phase that does not come from its own signal, on inversing the Fourier transform the signal is skewed. Thus recovery is only viable at certain frequencies where the recovered and non-reflection spectra are close in magnitude. This also renders the beamform results to be non-predictive in their accuracy, as seen by the non-linear trend in both the beamform MSE and dB attenuation columns for the phase corrected data in Table 5.4. The second problem is that “non-reflection” data might not be present in actual experimental testing. The phase of this data was used here as it is known that the non-reflection signal is the goal of recovery. Without a signal that is known to be ideal under the given test circumstances, this phase information would not be available to the engineer.
6 Conclusions and Future Work

The main conclusions are that Adaptive Noise Cancellation was successful in accurate magnitude and phase recovery of the noise-free signal and that Cepstral echo removal was only successful in removing the effects of acoustic reflections from the magnitude of the signal. Future work involves comparison of the methods studied with those being actively used in industry, increasing the complexity of the experiment, and in situ wind tunnel testing.

6.1 Conclusions

Two experiments were setup to investigate the robustness of Adaptive Noise Cancelling and Cepstrum echo removal techniques on acoustic data from a linear microphone array. The first experiment introduced a background noise source into the anechoic chamber in order to overwhelm the desired source’s signal at the microphone array. The second introduced a reflective surface in order to produce constructive/destructive interference of the signals received by the array.

Autospectra, a single microphone analysis tool, were used to provide magnitude information of the processed signals. Beamforming was used as an array processing technique that provided spatial information not available from autospectra. Both the magnitude and phase of a signal are used in beamforming, giving a more complete description of the processed signal.

Adaptive Noise Cancellation was shown to be successful in recovering the magnitude of the autospectrum peaks of the noise-free data at all three SNRs investigated. The beamform magnitude and shape was also recovered accurately at all three SNRs. The recovery accuracy of both worsened as the SNR decreased. The success of both accurate autospectrum and beamform recovery indicates that this technique is applicable to both single microphone and array processing.

A simulation was created in order to investigate the algorithms and theory chosen to perform the Cepstral processing. A pulse and one reflection were used for the input, providing an ideal signal, which in turn gave ideal Cepstral features. The influence of the reflection on the
The experimental reflection data were processed using the outlined Cepstral techniques. Four schemes of data manipulation in the form of liftering were described. As the amount of data liftered increased, the magnitude recovery became closer to that of the non-reflection data (seen in the autospectra) but the phase recovery became less accurate (seen in the beamforms). Only the autospectrum of the most liftered data showed accurate magnitude recovery of the non-reflection signal. A phase correction procedure was implemented to demonstrate the importance of phase recovery for beamforming. The results demonstrated that the magnitude of the recovered signal and the phase of the non-reflection data could be combined to achieve accurate matches in the beamforming. With the phase correction applied, the MSE accuracy of the beamforms improved by ~77.4%, however no linear trend was observed.

Addressing the specific research objectives of Chapter 1,

- The source signal was accurately recovered at the minimum SNR (Figs. 5.12 and 13) using ANC post-processing techniques.
- An ideal signal, consisting of a wavelet and one echo, was used to illustrate Cepstral echo removal techniques (Fig. 5.25). Accurate magnitude and phase recovery were achieved (Fig. 5.29).
- The Cepstral echo arrival times produced (Eq. (5.14)) were successfully verified through setup geometry (Eq. (5.9)) and spectral characteristics (Eq. (5.10)).
- The Cepstral echo removal processing results using varied liftering schemes were investigated and quantified successfully (Table 5.4) with implementation of a phase correction procedure (Fig. 5.38). The magnitude was accurately recovered (Fig. 5.43) but the phase was unable to be recovered without the phase correction (Table 5.4, ‘MSE Beamform’ column).

Both background noise and surface reflections are major sources of acoustic data contamination in wind tunnels. Automated, in situ incorporation of the methods described could greatly increase fidelity of real time acquisitions. This in turn would lead to improved noise source identification and quantification of aircraft components tested in wind tunnels.
6.2 Future Work

A necessary step in identifying where the techniques studied fit in to methods/procedures currently used in industry is to quantitatively compare them. This would be necessary for either of the techniques to become commonly used.

Other improvements on the methods presented are related to complexity. The array could be made two dimensional, thus allowing for a mapped plane of acoustic data instead of a line. This would be necessary in the case of a larger, more complex source. Also, the speaker used could be substituted for something more complex than a point source. This would test the resolution capabilities of the processing by forcing the beamforming to distinguish between close sources or a radiation gradient across the source.

With success achieved in the laboratory, actual wind tunnel testing is required to validate experimental results. The wind tunnel environment poses challenges not present in the laboratory. The first problem is that the test section may not be anechoic. This signifies that reflection paths are possible from many directions. Second, the background noise present may not be localizable. This raises concerns on the obtainable upper limit of coherence that could be achieved between the array and a reference microphone. Third, when the tunnel is running a shear flow layer exists that distorts sound travelling through it. As the array would be placed out-of-flow, this would have to be taken into account. Last, the tunnel conditions mentioned all exist simultaneously, and would be distinct for each tunnel. These challenges present a bulk of work still left to accomplish.
Appendix A: DAMAS Methodology

After standard beamforming has been performed (refer to Chapter 3), the DAMAS processing may begin [26]. The general purpose of DAMAS is to separate source strength information from array beamforming characteristics, and take only the former as it is the desired and true sound information.

**Inverse Problem Definition**

The pressure transform $P_m$ (Eq. (3.7)) of microphone $m$ must be related to a model source located at a position $n$ in the source field by a new term, $Q_n$.

$$P_{mn} = Q_n e_{mn}^{-1}$$  \hspace{1cm} (A.1)

$Q_n$ represents $P_{mn}$ (originally $P_m$) if flow convection and shear layer refraction were not present and if the distance from $m$ to $n$ were $r_c$ instead of $r_m$. The term $e_{mn}^{-1}$ contains the parts of Eq. (3.9) that affect the signal in the actual pressure transmission to render $P_m$. The pressure-transform product from Eq. (3.7) becomes

$$P_m^* P_{mn} = (Q_n e_{mn}^{-1})^* (Q_n e_{m'n}^{-1}) = Q_n^* Q_n (e_{mn}^{-1})^* e_{m'n}^{-1}$$  \hspace{1cm} (A.2)

Subbing this back into Eq. (3.7) and forming (3.8) the modeled CSM for a single at $n$

$$\hat{G}_{n_{mod}} = X_n \begin{bmatrix} (e_1^{-1})^* e_1^{-1} & (e_2^{-1})^* e_2^{-1} & \cdots & (e_M^{-1})^* e_M^{-1} \\ (e_1^{-1})^* e_1^{-1} & (e_2^{-1})^* e_2^{-1} & & \\ \vdots & & \ddots & \\ (e_1^{-1})^* e_1^{-1} & & & (e_M^{-1})^* e_M^{-1} \end{bmatrix}$$  \hspace{1cm} (A.3)

Here, $X_n$ represents the mean-square-pressure per bandwidth at each microphone normalized at the level of $r_c$. The assumption is now made that there are $N$ statistically independent sources each at distinct $n$ positions. The total modeled CSM is

$$\hat{G}_{mod} = \sum_n \hat{G}_{n_{mod}}$$  \hspace{1cm} (A.4)

Subbing this into the array response, Eq. (3.11)

$$S_{n_{mod}} (\hat{e}) = \left[ \frac{\hat{e}^T e_{mod} \hat{e}}{M^2} \right]_n = \frac{\hat{e}^T \sum_n X_n^l \eta_{nl} \hat{e}_n}{M^2} = \sum_{n'} \frac{\hat{e}_{n'}^T \eta_{n'l} \hat{e}_{n'}}{M^2} X_{n'} = \hat{A} X_n$$  \hspace{1cm} (A.5)

Here, the term in brackets is Eq. (A.3). And matrix $\hat{A}$ components are

$$\hat{A}_{n'n'} = \frac{\hat{e}_{n'}^T \eta_{n'l} \hat{e}_{n'}}{M^2}$$  \hspace{1cm} (A.6)
Now, if the measured data are equated with the model data, i.e. Eq. (3.11) with (A.5), the result is

$$\hat{A}\hat{X} = \hat{Y}$$  \hspace{1cm} (A.7)

and it should be noted that the diagonal terms of $\hat{A}$ are equal to one.

Equation (A.7) is a system of linear equations relating a spatial field of point locations having array beamform output responses $Y_n$, to source distributions $X_n$ at those locations. It is the DAMAS inverse problem.

**Inverse Problem Solution**

The task of DAMAS is to solve for $X$. The solution is found by enforcing a positivity constraint on $X$, making the problem deterministic. For justification of this solution method, see [26].

A single linear equation component of Eq. (A.7) is

$$A_{n1}X_1 + A_{n2}X_2 + \cdots + A_{nn}X_n + \cdots + A_{nN}X_N = Y_n$$ \hspace{1cm} (A.8)

As $A_{nn} = 1$, this becomes

$$X_n = Y_n - \left[\sum_{n' = 1}^{n-1} A_{nn'}X_{n'} + \sum_{n' = n+1}^{N} A_{nn'}X_{n'}\right]$$ \hspace{1cm} (A.9)

Using the following constraints

$$X_1^i = Y_1 - \left[0 + \sum_{n' = 1+1}^{N} A_{1n'}X_{n'}^{i-1}\right]$$

$$X_n^i = Y_n - \left[\sum_{n' = 1}^{n-1} A_{nn'}X_{n'}^{i-1} + \sum_{n' = n+1}^{N} A_{nn'}X_{n'}^{i-1}\right]$$ \hspace{1cm} (A.10)

$$X_N^i = Y_N - \left[\sum_{n' = 1}^{N-1} A_{Nn'}X_{n'}^{i-1} + 0\right]$$

Eq. (A.9) is used iteratively to determine $X_n$ for all $n$ from 1 to $N$. For the first iteration ($i = 1$), $X_n$ can be made to equal zero or $Y_n$. If an $X_n$ calculation results in a negative number, it is set to zero. Equation (A.10) is the DAMAS inverse problem solution. For application criteria and experimental notes see [26].
# Appendix B: Experiment Hardware

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Make, Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microphones</td>
<td>Brüel &amp; Kjær (B&amp;K) quarter-inch free-field, type 4939</td>
</tr>
<tr>
<td>Power supply</td>
<td>Agilent DC, E363xA series</td>
</tr>
<tr>
<td>Pre-Amplifier</td>
<td>Carver, model TFM-42</td>
</tr>
<tr>
<td>Microphone Supply</td>
<td>B&amp;K Dual, type 5935L</td>
</tr>
<tr>
<td>Pistonphone</td>
<td>B&amp;K, type 4228</td>
</tr>
<tr>
<td>Microphone Preamplifier</td>
<td>Falcon Range half-inch, type 2669</td>
</tr>
<tr>
<td>Filter</td>
<td>Krohn-Hite, model 3202</td>
</tr>
<tr>
<td>Noise Generator</td>
<td>General Radio, type 1382</td>
</tr>
<tr>
<td>Data Acquisition System</td>
<td>National instruments, model PXI-1044</td>
</tr>
<tr>
<td>Speaker</td>
<td>Pioneer 5.25” Cup Midrange, model B11EC80-02F</td>
</tr>
<tr>
<td>CPU</td>
<td>Asus, Intel Core 2 Quad, 2.41 GHz, 2 GB Ram</td>
</tr>
<tr>
<td>Microphone Cabling</td>
<td>LEMO 30m extension, type AO-0414-D-300</td>
</tr>
<tr>
<td>Horn Driver</td>
<td>Selenium Compression Driver with Titanium Diaphragm, model DH200</td>
</tr>
</tbody>
</table>
Bibliography


