Multiarray Passive Acoustic Localization and Tracking

Daniel J. Mennitt

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mechanical Engineering

Dr. Marty Johnson, Co-Chairman
Dr. James Carneal, Co-Chairman
Dr. Michael Roan
Dr. Craig Woolsey
Dr. Andrew Kurdila

November 4, 2008
Blacksburg, VA

Keywords: Acoustic Localization, Sensor Networks, Numerical Modeling, Signal Classification, Matched Field Processing

Multiarray Passive Acoustic Localization and Tracking

Daniel J. Mennitt

ABSTRACT

Wireless sensor networks and data fusion has received increasing attention in recent years, due to the ever increasing computational power, battery and wireless technology, and proliferation of sensor modalities. Notably, the application of acoustic sensors and arrays of sensors has expanded to encompass surveillance, teleconferencing, and sound source localization in adverse environments. The ability to passively locate and track acoustic sources, be they gunfire, animals, or geological events, is crucial to a wide range of applications. The challenge addressed herein is how to best utilize the massive amount of data collected from spatially distributed sensors. Localization in two acoustic propagation scenarios is addressed: the free-field assumption and the general case. In both cases, it is found that performance is highly dependent on the array-source geometry which in turn drives the design of localization strategies.

First, the general surveillance problem including signal detection, classification, data association, localization and tracking is studied. Signal detectors are designed with a focus on robustness and capacity for real time implementation. Specifics of the data association problem relevant to acoustic measurements are addressed. Assuming free-field propagation, a localization algorithm is developed to harness some of the vast potential and robust nature of a sensor networks. In addition, a prototypical sensor network has been constructed to accompany the theoretical development, address real world situations, and demonstrate applicability. Experimental results obtained confirm the practicality of theoretical models, support numerical results, and illustrate the effectiveness of the proposed strategies and the system as a whole.

In many situations of interest, obstacles to wave propagation such as terrain or buildings exist that provide unique challenges to localization. These obstacles introduce multiple paths, diffraction, and scattering into the propagation. The second part of this dissertation investigates localization in the general propagation scenario of a multi-wave, semi-reverberant environment characteristic of urban areas. Matched field processing is introduced as a feasible method and found to offer superior performance and flexibility over time reversal techniques. The effects of uncertainty in model parameters are studied in an urban setting. Multiarray processing methods are developed and strategies to mitigate the effects of model mismatch are established.

Acknowledgements

First, I would like to thank Marty Johnson and Jamie Carneal for providing me with the opportunity to be part of several interesting projects in acoustics, for the constant guidance in my work, and for the invaluable and critical feedback on my research and presentation skills. I would also like to extend my appreciation to Dr. Craig Woolsey, Dr. Michael Roan, and Dr. Andrew Kurdila for serving on my advisory committee. This work would not have been possible without funding from AuSIM, Inc., Sennheiser and the Office of Naval Research.

Next, I would like to thank my fellow researchers at VAL, especially Philip Gillett, Kamal Idrisi, and Brian Goode. It has been a pleasure working with them and I wish them all great success in their future academic and professional lives. I would especially like to thank Cory Pappenfuss, engineering guru, for help and invaluable knowledge of all things electrical. Of course, I can't leave out all the excellent administrative assistants who have taken care of me throughout my years here: Lynda King, Mellissa Williams, and Gail Coe

A thank goes out to all my friends both here at Virginia Tech and elsewhere for reminding me there is life outside the anechoic chamber. The final word of acknowledgement is reserved for my family. I owe everything I achieve in life to their love, patience and support.

Table of Contents

List of Tables ..... *viii*

List of Figures ..... *ix*

Acronyms ..... *xiii*

Nomenclature ..... *xiv*

CHAPTER 1: INTRODUCTION **1**

1.1 Problem and Motivation *1*

1.2 System Overview *3*

1.3 Organization and Research Topics *4*

1.4 Contributions *7*

CHAPTER 2: BACKGROUND **9**

2.1 Sensor Networks *9*

2.2 Tracking *11*
2.2.1 Data Association 11
2.2.2 Target State Estimation 12

2.3 Acoustic Localization *13*
2.3.1 Free field methods 14
2.3.2 Full field methods 16

2.4 Numerical Modeling Methods for Acoustic Propagation *18*
2.4.1 Wave Based Methods 19
2.4.2 Statistical Methods 20
2.4.3 Ray Based Methods 21

CHAPTER 3: DETECTION AND CLASSIFICATION **23**

3.1 Coarse Classification *23*
3.1.1 Impulsiveness Metric 25
3.1.2 Tonality Metric 25

3.2 Elementary Signals *26*
3.2.1 Impulsiveness of white noise 27
3.2.2 Tonality of white noise 30

3.3 Binary Hypothesis Testing *32*

3.4 Impulse Detector 34
3.4.1 Approach 34
3.4.2 Formulation 35
3.4.3 Discussion 38

3.5 Tone Detector 39

3.6 Broadband Detector 43

3.7 Impulsiveness - Tonality Feature Space 43

CHAPTER 4: ASSOCIATION, LOCALIZATION, AND TRACKING IN A FREE FIELD 50

4.1 Detection Association 51
4.1.1 Network Information and Communication 52
4.1.2 Influence field of acoustic sensors 53
4.1.3 Criteria for Detection Association 54

4.2 Localization by Bearing Fusion 55
4.2.1 Linear Least Squares Localization 57
4.2.2 Pairwise Localization and Fusion 58

4.3 Localization Simulation 63
4.3.1 Simulation Method 64
4.3.2 Simulation Results 65

4.4 Tracking 68
4.4.1 The Kalman Filter 68
4.4.2 Event to Track Association 71

4.5 Experimental Validation 72
4.5.1 Experimental Setup 72
4.5.2 Experimental Results 74

CHAPTER 5: LOCALIZATION IN SCATTERING ENVIRONMENTS 76

5.1 Theory 78
5.1.1 Signal Model 78
5.1.2 Matched Field Processing 79
5.1.3 Multiarray Processing 82
5.1.4 Broadband Processing 84

5.2 Free-field beamforming 84

5.3 Finite Difference Time Domain Modeling 85
5.3.1 Formulation 86

5.4 Numerical Example: modeling and MFP in a scattering environment 89
5.4.1 FDTD Model 89

5.4.2 Impulse response 91
5.4.3 Matched Field Processing 93

CHAPTER 6: LOCALIZATION IN AN URBAN ENVIRONMENT WITH UNCERTAINTY 100

6.1 Test Environment 101

6.2 Forms of Uncertainty 104
6.2.1 Uncertainty due to sensor noise 105
6.2.2 Uncertainty due to model mismatch: calibration error 107
6.2.3 Uncertainty due to model mismatch: temperature 110
6.2.4 Uncertainty affecting inter-array coherence only 111

6.3 Spatial Variation of Performance 112

6.4 Time Reversal Refocusing 116
6.4.1 Method 116
6.4.2 Practical limitations 118

6.5 MFP Using Free-Field Steering Vectors 120
6.5.1 Beampatterns and performance 120
6.5.2 Parametric performance study 123

CHAPTER 7: SUMMARY & FUTURE WORK 126

7.1 Signal Detection 126

7.2 Detection Association 127

7.3 Bearings-only Localization 127

7.4 Matched Field Processing 128

BIBLIOGRAPHY 131

List of Tables

Table 3.1: Summary of impulsiveness and tonality of elementary signals. 31
Table 4.1: Mean RMS error (meters) of the three fusion strategies for the simulation with added noise only. 66
Table 4.2: Maximum RMS error (meters) of the three fusion strategies for the simulation with added noise only. 66
Table 4.3: Mean RMS error of the three fusion strategies for the simulation with additive noise and one biased node. 67
Table 4.4: Maximum RMS error of the three fusion strategies for the simulation with additive noise and one biased node. 67
Table 4.5: Summary of RMS localization error (meters). 75
Table 5.1: Parameters for the finite difference time domain model. 91
Table 6.1: Original and perturbed array coordinates in the x,y plane (meters). 108
Table 6.2: Mean and median error (meters) for various processing approaches under the influence of uncertainty in various parameters: medium temperature (20º perturbation) and array location (2m perturbation). 115

List of Figures

Figure 1.1: Schematic of system architecture. 3
Figure 1.2: Acoustic node featuring a diffracting microphone array. 4
Figure 2.1: Schematic representation of a target (red) moving through a network of multiple nodes (orange). Solid dashed lines represent the extent of the sensing and communication range respectively. 10
Figure 2.2: The Kalman filter is a recursive algorithm for optimally estimating the states, $\boldsymbol{\xi}$, of a linear dynamic system using a form of feedback control. Future states are predicted using a model, $\boldsymbol{\Lambda}$, and then corrected using noisy measurements (localization results). 12
Figure 2.3: Time and level difference of the output from two spatially separated microphones in response to an acoustic source. 14
Figure 2.4: Location finding by hyperbolic intersection [35]. 16
Figure 2.5: Examples of sound propagation in typical urban spaces. The full and broken lines represent the reflected and diffuse scattering fields respectively using a ray interpretation [50]. 19
Figure 3.1: Time and Frequency domain plots of elementary signals. 27
Figure 3.2: Cumulative density functions for a Gaussian distribution and the absolute value of the Gaussian distribution (unity variance). The median value occurs at a cumulative probability of 0.5. 29
Figure 3.3: Maximum expected value of a block of white Gaussian noise. 30
Figure 3.4: Neyman and Pearson binary hypothesis testing considering location testing with Gaussian error. 34
Figure 3.5: Receiver operating characteristic for the impulse detector. 37
Figure 3.6: Two examples of an impulse in noise (top left and top right). The SNR is 20dB in both cases. The impulsiveness of each signal is identical (bottom). 39
Figure 3.7: Receiver operating characteristic curves for the tonality detector given that $K = 512$ 42
Figure 3.8: Receiver operating characteristic curves for the tonality detector given that $SNR = -20\text{dB}$ 42
Figure 3.9: Spectrogram (left) and the impulsiveness-tonality feature space (right) for a white Gaussian noise signal. 44
Figure 3.10: Time domain (left) and frequency domain (right) of the Gaussian white noise signal showing divisions for impulsiveness and tonality estimation. 45
Figure 3.11: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of wind. 45
Figure 3.12: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of distant traffic from a 4-lane 45mph road. 46
Figure 3.13: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a gasoline generator 46
Figure 3.14: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a car driving past the microphone. 46
Figure 3.15: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a tracked vehicle (Sherman tank). 47

Figure 3.16: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a gunshots (three pistol shots and one shotgun).....47
Figure 3.17: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of speech (reading aloud in an anechoic chamber) 47
Figure 3.18: Signal data in the impulsiveness-tonality feature space. 48
Figure 4.1: Summary of the tracking process. Over time, node detections are associated to localize events and events are associated to update the target track. 50
Figure 4.2: Two dimensional geometry of source detection and localization problem using multiple nodes.....52
Figure 4.3: Schematic showing the most conservative estimate of propagation distance from the detecting node. 54
Figure 4.4: Favorable and unfavorable geometries illustrating geometric dilution of precision. The dotted lines indicate the range of bearing angle (all arrays have identical variance). 56
Figure 4.5: Demonstration of the pairwise fusion using the MLE estimator. 62
Figure 4.6: One sample of a scenario showing five nodes (circles) and LOBs to the source (asterisk). 1 node (solid black) is biased creating three skewed observations and one outlier (intersection is beyond the range of the plot).....64
Figure 4.7: Convergence of the mean RMS error during the Monte Carlo simulations for the noise-only case (left) and biased case (right).....65
Figure 4.8: Histograms showing the distribution of RMS error for the 6-node, additive noise only case for the LLS estimator (top), MLE estimator (middle), and MED estimator (bottom). 66
Figure 4.9: Histograms showing the distribution of RMS error for the 6-node, additive noise plus bias case for the LLS estimator (top), MLE estimator (middle), and MED estimator (bottom).....68
Figure 4.10: Photo of acoustic arrays and remote control helicopter.....73
Figure 4.11: Close up photo of drive-by-wire acoustic node and key components. 73
Figure 4.12. Real-time output of heterogeneous fusion center superimposed on satellite photograph of test area. Helicopter GPS track was added from post-test data processing to show the precision of the system. 74
Figure 4.13: Tracking performance of target position in the Eastern (left) and Northern (right) coordinate. Plots compare estimates of noise source position (KF track) with the known GPS track of the helicopter. 75
Figure 5.1: Representation of an urban setting illustrating spatial source – receiver relationships. 77
Figure 5.2: Cartesian grid arrangement for the leapfrog method.....88
Figure 5.3: Geometry of the FDTD model showing the relative position of the arrays and scattering object. 89
Figure 5.4: The source signal (Ricker wavelet) in the time (left) and frequency domain (right).....90
Figure 5.6: Snapshots of the 2D acoustic pressure field as calculated by the FDTD method. 92
Figure 5.5: Impulse response recorded at one sensor from node 1 (left) and node 2 (right). $\boldsymbol{ln}1 = 52, 62$, $\boldsymbol{ln}2 = 80, 15$ and $\boldsymbol{ls} = 125, 73$.....93

Figure 5.7: Transmission loss (dB) at 100Hz for array 1 (upper left), array 2 (upper right), and the coherent array network (bottom). 94
Figure 5.8: Matched-field power response function at 100 Hz for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{ls} = 125, 73.$ 95
Figure 5.9: Broadband matched-field power response function for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{ls} = 125, 73.$ 96
Figure 5.10: Matched-field power response function using the plane-wave model at 100 Hz for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{ls} = 125, 73.$ 97
Figure 5.11: Matched-field power response function of the coherent array network at 100 Hz in free-field propagation conditions using the free-field model. The source is located at $\boldsymbol{ls} = 125, 73$ 97
Figure 5.12: Matched-field power response function of the coherent array network at 100 Hz using the phase-only free-field model. The source is located at $\boldsymbol{ls} = 125, 73.$ 98
Figure 6.1: Overhead satellite image of the McKenna MOUT site at Fort Benning and individual building details (below) [94]. 101
Figure 6.2: 2D FDTD model of artificial training village used for the study. 102
Figure 6.3: Transmission loss (dB) of node 1 at 100Hz. Primary paths for propagation through the obstacles is indicated by high transmission loss. 103
Figure 6.4: Broadband matched-field power response function for array 1 (top left), array 2 (top right), the incoherent array network of 5 arrays (bottom left) and the coherent array network of 5 arrays (bottom right). The source is located at $\boldsymbol{ls} = 65, 73$ 104
Figure 6.5: Broadband noise-only matched field power response function for node 3 using weights derived from the plane wave model. 106
Figure 6.6: Broadband noise-only matched field power response function for node 1 (left) and node 3 (right) using weights derived from the a priori model. 107
Figure 6.7: Geometry of the simulation. Arrows indicate the direction in which the node location was perturbed for the mismatch study. 108
Figure 6.8: Broadband matched-field power response function with a 2m uncertainty in node location for node 1 (top left), node 2 (top right), the incoherent network of 5 arrays (bottom left) and the coherent network of 5 arrays (bottom right). The source is located at $\boldsymbol{ls} = 65, 73.$ 109
Figure 6.9: Broadband matched-field power response function with a 20ºC uncertainty in temperature for node 1 (upper left), node 2 (upper right), the incoherent array network of 5 nodes (lower left) and the coherent array network of 5 nodes (lower right). The source is located at $\boldsymbol{ls} = 65, 73.$ 110
Figure 6.10: Broadband matched-field power response function for the coherent network of 5 arrays with 0 inter-array coherence to be compared with Figure 6.4 (no mismatch, left) and Figure 6.8 (2m location perturbation, right). 112
Figure 6.11: Error (meters) of matched field processor due to a 2 meter uncertainty in array location for individual arrays (a-e), incoherent processor (f), coherent processor (g), and coherent processor with additional inter-array incoherence (h). 114
Figure 6.12: Received signal, rt, at microphone 1 on node 2 (left) and transmitted signal, $r - t$, from microphone 1 on node 2 (right). 117

Figure 6.13: Pressure field ($Nm2$ dB) at the time step coinciding with the peak of the source emission due to time reversal refocusing of 5 microphone signals. 118
Figure 6.14: Localization accuracy over space. Error (meters) due to a 20º C perturbation in temperature of the medium. Results are shown for the time reversal refocusing method (left) and the coherent matched field processor (right). 119
Figure 6.15: Broadband matched-field power response function using free-field weights for array 1 (upper left), array 2 (upper right), the incoherent array network of 5 arrays (lower left), the coherent array network (lower right). The source is located at $\boldsymbol{ls} = 65, 73$. 121
Figure 6.16: Error (meters) of matched field processor using free-field weights. Results are shown for individual arrays 1-5 (a-e), incoherent processor (f), coherent processor (g), and coherent processor with additional inter-array incoherence (h). 122
Figure 6.17: Percent coverage (error < 20m) of full-field weighted MFP considering increasing levels of uncertainty and free-field weighted MFP. 124

Acronyms

AOA	Angle Of Arrival
CDF	Cumulative Density Function
CSD	Cross Spectral Density
DI	Directivity Index
EDW	Event Detection Window
FDTD	Finite-Difference Time-Domain
FV	Feature Vector
GDOP	Geometric Dilution of Precision
GNN	Global Nearest Neighbor
GPS	Global Positioning System
ILD	Interaural Level Difference
LLS	Linear Least Squares
LOB	Line Of Bearing
LOS	Line of Sight
MFP	Matched Field Processing
MLE	Maximum Likelihood Estimator
MVDR	Minimum Variance Distortionless Reponse
PDF	Probability Density Function
RMS	Root Mean Square
SEA	Statistical Energy Analysis
SNR	Signal to Noise Ratio
TDOA	Time Difference of Arrival
WNG	White Noise Gain

Nomenclature

A	signal amplitude factor
$\boldsymbol{A}$	matrix in a least squares problem
$\boldsymbol{b}$	matched field power response function, vector in a least squares problem
c	speed of sound
C	cumulative density function
$\boldsymbol{d}, \boldsymbol{D}$	steering vector, steering vector matrix
e	kalman filter or FDTD step index
$E\{\cdot\}$	expectation
erf	Gaussian error function
f, F	frequency index, number of frequencies
g	number of spatial dimentions
G	subscript indicates Gaussian distribution
h	number of kalman filter states
H	hermitian transpose, hypothesis
i	imaginary number
I	subscript indicates impulsiveness domain
$\boldsymbol{I}$	identity matrix
$\boldsymbol{imp}$	impulsiveness vector
j, J	source index, number of sources
k, K	discrete sample index, number of samples (size of data block)
$\boldsymbol{l}$	location vector of spatial coordinates
L	likelihood ratio
m, M	microphone index, number of microphones
n, N	node index, number of nodes, Normal distribution
o, O	observation index, number of observations
p	probability density function, acoustic pressure
P	probability distribution
P_D	probability of detection
P_F	probability of false alarm
P_M	probability of miss
Pr	probability of
q, Q	subscript indicates network
$\boldsymbol{r}$	signal vector received at sensor
R	indicates signal matrix, Rayleigh distribution
$\boldsymbol{s}$	source signal vector, subscript indicates source
t	time
T	superscript indicates vector or matrix transpose, subscript indicates tonality domain
ton	tonality vector
$\boldsymbol{u}$	kalman filter process noise control input
$\boldsymbol{v}$	additive noise process, subscript indicates noise, acoustic particle velocity

$\boldsymbol{w}, \boldsymbol{W}$	weight vector,weight matrix
y, Y	observation (detection theory)
x, y, z	dimensional coordinates
X, Y	dimensional coordinate indices
α	beamformer design factor
β	phase factor parameter
γ	coherence
Γ	domain
δ	dirac delta function, decision rule
ε	error
ζ	phase factor
η	decision threshold
θ	node bearing or angle of arrival
$\boldsymbol{\Theta}$	kalman filter measurement matrix
κ	standard deviation factor
$\boldsymbol{\lambda}$	multivariate observation vector
$\boldsymbol{\Lambda}$	state transition matrix
μ	mean of a random process
$\boldsymbol{\xi}$	kalman filter state vector
$\boldsymbol{\Xi}$	kalman filter measurement vector
π	pi
$\boldsymbol{\Pi}$	kalman gain matrix
ρ	density of medium
σ	standard deviation of a random process
τ	time quantity
υ	uniformly distributed random variable
$\boldsymbol{\Upsilon}$	control matrix
φ	cumulative density function of the standard normal
ϕ	phase angle
$\boldsymbol{\Phi}$	covariance matrix
$\boldsymbol{\chi}$	unit normal noise vector
$\boldsymbol{\psi}$	transmission loss vector
Ψ	control input
ω	radian frequency
$\boldsymbol{\Omega}$	frequency bandwidth

Chapter 1:
Introduction

1.1 Problem and Motivation

All forms of life have evolved sensory mechanisms to thrive in a world where it is crucial to be knowledgeable of the surroundings. Paralleling this is development of cognitive facilities to utilize all the valuable information gathered. For some, this process is perpetual. As much as the dominant civilization on this planet is ever creating new problems, we are equally relentless in our effort to find solutions.

Wireless sensor networks have been identified as one of the most important technologies for the 21st century. Cheap, intelligent devices with onboard sensors networked through wireless links can be deployed in large numbers and provide extraordinary opportunities for instrumenting, monitoring, and controlling homes, cities and the environment. They can be deployed in air, under water, on vehicles, inside buildings, and in bodies. Current applications include surveillance, traffic control, industrial and manufacturing automation, and environmental and health monitoring [1].

The rising attention is due to the ever increasing computational power, battery and wireless technology, and proliferation of sensor modalities, each of which has its own unique set of characteristics [2]. Sensor types include magnetic, radar, thermal, acoustic, chemical, electric, seismic, and optical. Regardless of the sensor modality, the information that is commonly of most importance is the position of an object of interest.

Strengths of acoustic sensors include long range, high-fidelity, no-line-of-sight to target, and passive nature. Furthermore, many objects of interest have distinct acoustic signatures and the ability to detect, locate and identify acoustic sources, be they gunfire, animals, or geological events, is advantageous to a wide range of monitoring and

surveillance applications. Challenges of acoustic sensors include susceptibility to attenuation and time delay of the propagating medium, leading to distorted and out of sequence measurements. In addition, obstructions to propagation lead to scattering, multipath, and signal fading.

The most prominent application of acoustic localization capability is surveillance. Specifically, networked surveillance has potential in modern warfare as the ability to correctly detect, identify, and locate events on the battlefield will enhance situational awareness, leading to decreased casualties. Additionally, acoustic localization can aid residential security [3], traffic control [1], industrial processes, and structural monitoring of buildings [4]. Underwater acoustic arrays have been used to track and study the habits of marine animals [5], [6] such that human interference into animal habitats can be minimized.

While active modalities (e.g. radar) are capable of obtaining both bearing and range estimates, passive modalities are generally able to provide an angular position relative to the sensor only. However, a network of distributed sensors sharing information on an acoustic event can cooperatively localize a source absolutely. The use of distributed sensor networks has other desirable advantages, such as robustness to individual node failure and higher quality estimation ability due to the quantity and diversity of information. This is the subject of the first part of the dissertation.

In an environment characterized by complex acoustic propagation, such as the ocean and urbanized areas, a single passive sensor node can localize a source absolutely [47]. The physics of propagation lead to unique acoustic signatures from distinct points in space that can be exploited using prior knowledge of the environment. However, it will be shown that localization performance is dependent on the source-receiver geometry, among other factors. In addition to the aforementioned benefits of sensor networks, the disparate positions of multiple acoustic sensors can greatly aid localization efforts in a complex scattering environment. The second part of the dissertation investigates this situation with a focus on how to utilize information from spatially distributed sensors.

1.2 System Overview

A prototypical sensor network has been constructed to accompany the theoretical development, address real world situations, and demonstrate applicability. Figure 1.1 shows an overview of the system architecture and key connections between systems. The middle pane shows the flow of information between an acoustic event, nodes, and network. Any perturbation of the medium which can be attributed to a source at some discrete spatial location is termed an event, as opposed to acoustic energy from wind for example. Multiple acoustic sensors are attached to a node (left pane), additionally outfitted with local computational processing ability, GPS for position, a magnetic compass for orientation, and wireless networking capability.

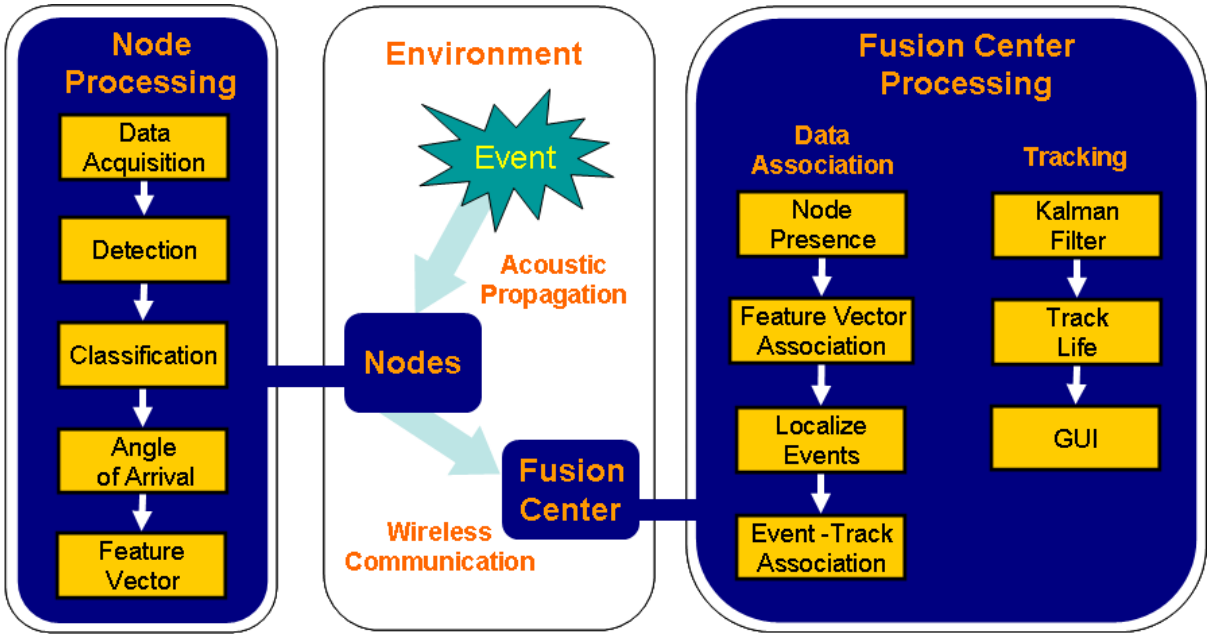

Figure 1.1: Schematic of system architecture.

Figure 1.2 shows a typical acoustic node used in experiments. Each spatially distributed node receives an environment filtered version of the source signal, and, upon successful detection estimates crucial parameters concerning the nature and origin of the source. This distilled information set, or feature vector, is communicated to a centralized

fusion center, typically another networked computer. The fusion center associates data from disparate nodes and times, localizes events, and tracks sources over time (right pane of Figure 1.1). Most elements of the system will be developed in more detail in chapters to follow.

Figure 1.2: Acoustic node featuring a diffracting microphone array.

1.3 Organization and Research Topics

Beyond the preliminaries, this dissertation is divided into two parts: the free field situation and the general situation. This introduction and background information makes up the preliminary material. Free field propagation is a common assumption and gross subset of the general situation. In addition to localization techniques, advanced applications of detection, data association, information fusion, and tracking in a multisource environment are investigated. In the second part, the more complicated situation of an arbitrary propagation environment is examined. The focus is on a field populated by multiple discrete scattering objects, representative of an outdoor urban environment. Although a specific scenario, the localization techniques developed herein

are applicable to other environments with heterogeneous medium. Throughout the dissertation, the theme is utilizing information from spatially distributed nodes.

Note that, due to the wide range of topics covered and the fact that various quantities commonly use the same symbols (such as p for pressure and probability), some crossover of variables occurs and symbols are redefined when necessary. A compromise between readability and distinction has been made and although some odd symbols are used, the correct identity of a variable should be clear from the context. In general, vector and matrix quantities are indicated by lowercase and capital letters respectively, and both use boldface type. Non-bold lowercase and capital letter pairs are commonly used to specify an index and quantity of an indexed amount, as in the k^{th} sample in a block of data containing K samples.

The dissertation is organized as follows; chapters 1 and 2 discuss the necessary preliminaries in regards to the research to follow. Chapter 2 discusses some of the relevant background information for the four major areas brought together in this proposal viz. sensor networks, tracking, acoustic localization, and numerical modeling methods for acoustic propagation. The research covered here is the fuel for many of the decisions regarding directions to pursue in later chapters. First, some theoretical background and advantages of distributed sensor networks is presented, followed by an explanation and literature review of the tracking problem. Next, passive acoustic localization techniques are discussed. These consist of methods that either assume the pressure field can be described by a single plane wave or that the field is composed of many waves. A review of many of the localization methods reported in the literature is provided, many of which cross over between acoustics, electromagnetics, seismology, and other wave phenomena. Finally, a survey of numerical methods for acoustic propagation is provided. Models of the wave field are necessary for localization methods that take advantage of a priori information when plane wave propagation cannot be assumed.

In chapter 3, the coupled problems of signal detection and classification are investigated. This is a vital step in the multitarget scenario. Although there is much prior work in the literature, a coarse classification hierarchy is established that emphasizes simplicity, robustness, and universal applicability. Metrics corresponding to the classes

are developed and the concept of elementary signals is introduced to demonstrate the properties of the classification metrics. Generic binary hypothesis testing and decision theory is reviewed and then applied to the classification metrics to create two new signal detectors.

Information fusion is the subject of chapter 4 as the evolution from node detections, to source events, to source tracks is covered. First, an ad-hoc methodology is presented to associate multiple node detections based on the classification results and other available information. A novel localization scheme that takes advantage of multiple node locations is presented. Variations of the method are compared with a conventional technique via simulation and experiment. Some elements of the tracking problem are discussed as well. The previous work on this topic is vast and fairly basic tools, such as the Kalman filter, are implemented here. The chapter closes with the description and results from an outdoor field test. The experiment supports the numerical results of the localization methods comparison study.

Chapter 5 marks the beginning of the second part of the dissertation, localization in scattering environments. Based on the literature review (chapter 2), matched field processing (MFP) has been chosen as the most appropriate tool to apply to this problem. The method is explained in detail and means of applying it to the distributed sensor network scenario are developed. MFP requires prior knowledge of the propagation-dependant qualities of the environment. Finite difference time domain (FDTD) models are used to supply this information. This numerical modeling technique is explained and a relatively simple example is presented to demonstrate both the FDTD method and MFP for atmospheric acoustic localization in a scattering environment.

Chapter 6 continues the discussion of localization in scattering environments with a more realistic urban scenario and the addition of uncertainty. The true conditions of a real-world scenario can never be known exactly and uncertainty in multiple parameters is possible. The influence of such uncertainty on the performance of individual arrays of varying location and multiarray fusion approaches is investigated. These approaches include incoherent and coherent MFP, time reversal refocusing, and free field triangulation. Consideration to other factors beyond accuracy, such as flexibility of implementation, is also discussed.

Chapter 7 summarizes the work presented in this dissertation. The research presented here is innovative, drawing on fresh technologies and offering novel approaches to many aspects of acoustic signal processing. Consequently work can be continued in several different areas of concentration. These include sensor networks, information fusion, acoustic propagation modeling, and performance bounds. Additional applications can be imagined and limits of the technology are to be found. Some of these avenues and possible applications are presented towards the end of final chapter.

1.4 Contributions

The key contributions of this work include:

1. ***Signal detectors***: Two signal detectors, for impulsive and tonal signals, were designed and characterized. In addition to being simple and robust, they serve as efficient features for classification of a very broad range of signals.

2. ***Detection association scheme for an acoustic sensor network***: There are data association problems unique to an acoustic sensor network. In order to handle the inevitability of missing, delayed, and out of sequence measurements, the concept of an event detection window is introduced. The association test includes a multiple hypothesis methodology to circumvent the range deficiency of passive localization by considering node pairs.

3. ***Algorithm for pairwise bearings-only target localization***: Pair-wise fusion strategies are developed to take exploit the magnitude and disparate nature of information collected by a spatially distributed sensor network. The algorithm is shown to be more accurate and robust than the conventional method.

4. ***Effective method of source localization in an urban environment:*** Matched field processing is combined with finite-difference time-domain modeling and applied to passive acoustic localization in a scattering environment. Multiarray post processing techniques are developed and evaluated through perturbation analysis of uncertain model parameters. These techniques are shown to be more accurate than free field methods and more powerful and flexible than existing time-reversal methods. Additionally, the mechanism through which uncertainty manifests itself in the beamformer output suggests strategies for array placement.

Chapter 2:
Background

In this chapter, a brief literature review of the some of the research topics is provided. The background and definitions provided will motivate some decisions concerning which methods were employed, as well as facilitate further development in later chapters

2.1 Sensor Networks

The availability of low-power sensors, actuators, embedded processors, and radios is enabling wireless distributed sensing to be used for a wide range of applications: physiological and environmental network monitoring, smart spaces, military, and precision agriculture [7].

Although a plethora of configurations are possible [8], the prevalent constituent is the node, a platform that includes some combination of sensors, computational power, and communication ability. Figure 2.1 is a schematic of a typical sensor network showing possible limitations of sensor and communication range. In a centralized configuration, nodes act as data collectors and a single fusion center is designated to perform all composite processing, whereas in distributed processing all nodes act as fusion centers. In contrast to platform-centric systems where all the sensors are located at a common site, sensors may be located at significant distances from one another in network-centric systems, as is the case of radar data from multiple ships and aircraft. While wireless network-centric systems introduce new challenges, increased cost and complexity, the distribution of spatial location and processing offset these drawbacks with numerous advantages.

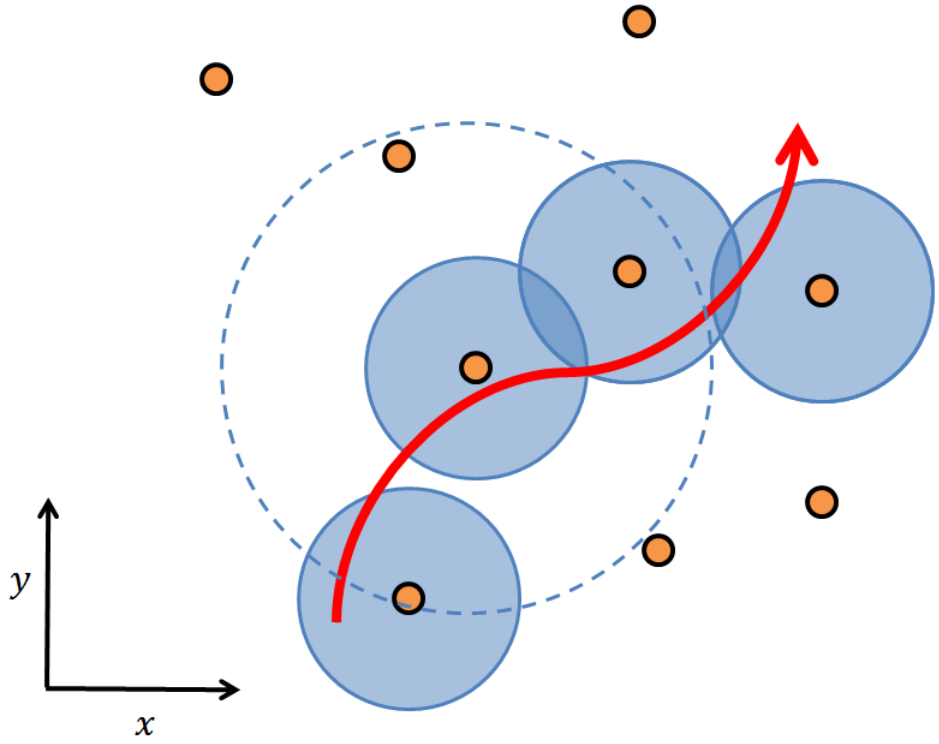

Figure 2.1: Schematic representation of a target (red) moving through a network of multiple nodes (orange). Solid dashed lines represent the extent of the sensing and communication range respectively.

A major focus of this work is *absolute localization by passive modalities*, which is made possible by a network of spatially distributed nodes. Furthermore, distributed sensing allows higher signal to noise ratios (SNR) as the proliferation of vantage points grants robustness to environmental obstacles, especially when the location of a signal of interest in unknown (see Figure 2.1). Although deploying one very large, sensitive sensor can increase SNR, line-of-sight opportunities and obstructions cannot be addressed by a single sensor regardless of its sensitivity. A single sensor is also less robust to component failure. It should be noted utilizing input from several sensors can be complicated by coordinate system transformations, know as gridlock [8].

While wired networking allows for advantages such as access to communication and energy channels, many environments do not have the necessary infrastructure. Wireless ability allows untethered freedom of movement while introducing the

complexity of relying on local, finite energy sources. Radio communications is often the prime energy consumer, making energy requirements the major design constraint in untended sensors. Communication also implies the necessity of a link, more equipment, and latency of transmitted information. Communication bandwidth requirements can be offset by intermediate processing of sensor data within the network at the node level. The reduction in transmitted data must be balanced with resource limitations of sensor nodes (processing, memory, bandwidth and energy).

2.2 Tracking

Multitarget-multisensor tracking has been extensively studied for several decades [9] and is essentially two coupled problems. Data association is the process of inferring which target, if any, a given measurement originated from. Target state estimation is the subsequent filtering of those measurements with a process model. These two problems are most commonly approached with statistical techniques designed to minimize the error due to noisy measurements, uncertainty in decisions and the approximations of deterministic models.

2.2.1 Data Association

A number of data association techniques have been proposed to deal with the measurement-track origin uncertainty, based primarily on kinematic quantities. The simplest, most commonly applied approach is the global nearest neighbor (GNN) [10], or single hypothesis tracking, in which an observation is assigned to at most one track. The joint probabilistic data association method forms and combines multiple hypotheses after each scan of data [11]. The multiple hypotheses tracking method is a deferred decision logic that forms multiple data associations and propagates their condition from scan to scan [12]. Numerous studies have concluded the performance advantages of multiple hypotheses tracking over joint probabilistic data association and GNN [13], [14].

2.2.2 Target State Estimation

Tracking is an observer design problem where the states and process model are specified. The Kalman filter [15], [16] is a recursive algorithm for optimally estimating the multidimensional state of a linear dynamic system. The Kalman filter estimates a process by using feedback control in the form of noisy measurements linearly related to the states (Figure 2.2). The Kalman filter is appropriate for predicting track updates if the underlying target dynamics and measurement noise can be assumed Gaussian. Otherwise, more accurate results can be obtained by propagating the probability density function in target space [17].

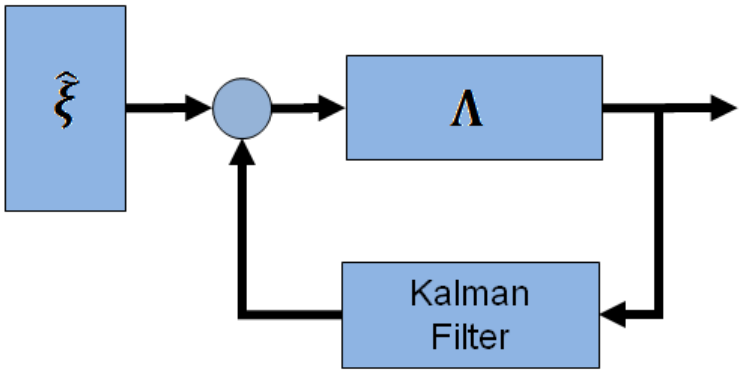

Figure 2.2: The Kalman filter is a recursive algorithm for optimally estimating the states, $\hat{\xi}$, of a linear dynamic system using a form of feedback control. Future states are predicted using a model, Λ, and then corrected using noisy measurements (localization results).

The majority of research in target motion considers dynamic models of a point target which describes the trajectory of the target state with respect to time [18]. Target motion, and models thereof, is usually classified as maneuver or nonmaneuver. Most two and three dimensional maneuver models attempt to capture the kinematics of various turning motions constrained in some way [19]. Although target motion is typically coupled across multiple coordinates in space, many models neglect this coupling, instead modeling the unknown acceleration as a random process [20]. Modern tracking systems

utilize many Kalman filters running in parallel. Model switching was originally performed using decision based methods [21] however much greater performance is achieved using the interacting multiple model method, currently the most widely accepted method for maneuvering targets [9].

Target motion models are best described in Cartesian coordinates while measurements of the target state are available in the original sensor coordinates. A variety of measurement models exist in which the target state, process noise, measurement, and sensor noise are in mixed coordinate systems or converted to a common coordinate system [22]. The weight of a particular measurement model's advantages and disadvantages are dependent on the overall system architecture.

2.3 Acoustic Localization

Acoustic localization is the process of discerning the location of an acoustic source via acoustic transducers. Passive methods entail listening only, whereas active localization additionally uses a known signal to excite the environment and is not considered herein. Generally, differences in the level and arrival time of a propagating wavefront at multiple microphones coupled with knowledge of acoustic propagation is manipulated to yield some information on the direction or absolute location of a source (Figure 2.3).

Most of the localization theory has focused on developing plane-wave methods based on free-field acoustic propagation. This assumption does not always hold, e.g. reverberant spaces. More recently, localization methods have been developed to take advantage of the full acoustic field which is often composed of multiple waves.

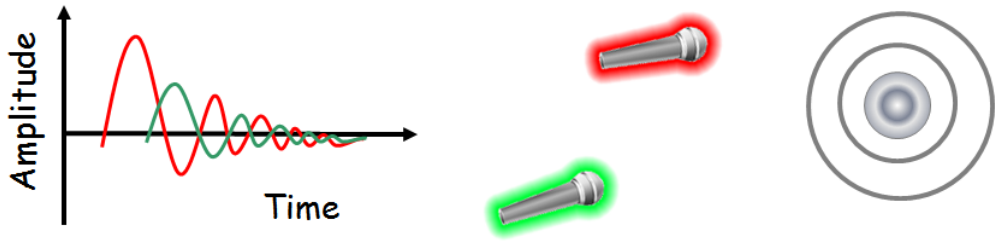

Figure 2.3: Time and level difference of the output from two spatially separated microphones in response to an acoustic source.

2.3.1 Free field methods

Differences in acoustic power between microphones, known as interaural level difference, has received little attention in the signal processing community as a localization cue [23]. Due to spherical spreading in free field acoustic propagation, energy decreases as the inverse of the square of the distance to the source. The level difference between two microphones defines a spherical locus of points upon which the source could lie, so multiple pairs of microphones are typically employed to localize the source absolutely. The accuracy of the method is sensitive to reverberation and the relative positions of source and microphones.

Difference in the time of arrival (TDOA) among microphones is a commonly exploited cue used in two fundamentally different ways: beamforming and geolocation by TDOA. When spacing between microphones is on the order of one wavelength (i.e. the microphones are configured as an array), the arrival time is discernable as a phase difference and coherent signal processing can be applied. In this case, the array can be electronically steered (beamforming) to obtain the angle of arrival (AOA) of the sound wave relative to the array [24], [25]. When the TDOA is weighted by the speed of sound to obtain an equivalent range difference, the range difference estimates can be used to localize a source absolutely [26].

In the area of acoustic array designs and geometries, diffracting arrays have been shown to offer superior low frequency performance over free field arrays of the same shape [27]. Algorithms based on experimental characterization of an array provide

accurate AOAs while maintaining high white noise gain. The diffraction around a rigid body enhances both the phase delay and level difference between sensors. Passive target localization using bearings-only measurements is a classic estimation problem supported by previous research [28], [29]. In this dissertation, new methods of combining multiple AOAs relative to spatially distributed nodes to absolutely localize a source in a global coordinate system are introduced, called pairwise bearing fusion. These methods are designed to harness the potential of a sensor network.

Geolocation by TDOA analysis is widely used in a variety of applications from locating lightning strikes to stop brush fires to location of tactile input in human computer interfaces [30]. Localization by using the TDOA between distinct microphones is usually implemented in two steps: time delay estimation and then incorporation of the TDOA estimates into a geometrical framework resulting in location estimates. The quality of the resulting location estimate is of course directly dependant on the quality of the time delay estimation process. There exists much work in the literature on TDOA estimation techniques, differing mainly on how to deal with background noise, reverberation, and movement of the sources [31]-[34]. Similar to signal intensity based target localization, TDOA methods for impulsive signals are attractive in regard to the amount of data transferred. However, finding the TDOA for non impulsive signal types using coherent signal processing techniques (cross correlation) requires phase difference information of signals measured at multiple sensors and thus involves high-precision time synchronization and large amount of data exchange among collaborating nodes. Even with available data, the number of periods can become ambiguous (i.e. multiple peaks). Therefore, coherent signal processing is generally limited to microphones collocated with the processing unit, i.e. an array on a node only, although there have been methods proposed to circumvent extensive data transfer for continuous signals [35].

Regardless of the time delay estimation methodology, there are a variety of additional methods to exploit the TDOA information, e.g. [30]. The time difference of arrival is equivalently a range difference of arrival assuming a homogenous propagation medium. Geometrically, a hyperbola defines a locus of points where the distance to two sensors (the foci) is constant. These curves can be intersected to identify the source location; one more sensor than the number of dimensions is required. Numerically, a set

of nonlinear coupled equations can be written describing the relationship between the emitter and receiver locations. No closed form solution exists, however, an iterative Taylor series method or Gaus-Newton implementation of the maximum likelihood estimator can be used [36].

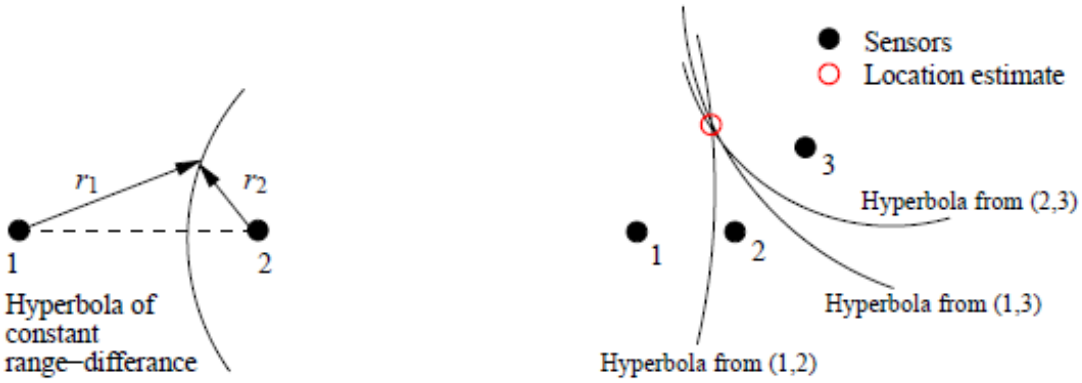

Figure 2.4: Location finding by hyperbolic intersection [35].

2.3.2 Full field methods

The methods discussed thus far are formulated assuming free field, plane wave propagation, and performance suffers in more complicated environments such as urban areas. The major difference is the presence of terrain or buildings that act as obstacles to acoustic wave propagation. These obstacles obstruct LOS and introduce multiple propagation paths, reflections, diffractions, and scattering into the propagation. Specifically, reverberant energy distorts loudness cues and indirect paths distort phase cues and time of arrival. Localization by bearings can tolerate some cases of non-LOS, such as diffraction over a hill. However, performance of this method will be degraded by image sources created by specular reflections. Returns from the environment that stand prominently above the diffuse and temporally decaying reverberation background and can be confused with or camouflage returns from an intended target and are known as clutter.

The problem of multipath propagation is well known and many statistical approaches have been proposed to mitigate the effect of clutter [37], [38]. In a known environment, a powerful method for acoustic source localization is time reversal refocusing [39], [40] in which sensor measured signals are back-propagated in a virtual environment using finite-difference time-domain (FDTD) models that incorporate the known geometry and propagation physics. Because of the reciprocity of the wave equation, inverted signals emitted from the sensor locations converge on the original source location. Time reversal acoustics has been developed recently in many fields [41], [42] especially medical ultrasound, and focusing ability has been found to degrade when there is a loss of information [43]. Specifically, dissipation in the medium breaks the invariance of wave equation and, in practice, the time reversal operation is often achievable only on a limited aperture time reversal mirror [44] that causes focusing quality to decrease in strongly diffracting media. Measurement of a propagation matrix, the collection of transfer functions between multiple source-receiver pairs, to create a matched filter has been shown to give superior performance at the cost of more measurements [45], [46].

In this dissertation, matched field processing (MFP) is presented as a method to locate acoustic sources in a scattering environment using microphone arrays. Whereas the conventional plane wave beamformer [25] steers an array in a particular direction, a matched field processor exploits the full-field acoustic propagation to predict the pressure in multiple dimensions, i.e. if the field is complex enough such that it cannot be described by a single plane wave, the fields are distinct enough such that muli-dimension resolution is possible. The process entails matching the sensor measured acoustic field with replicas of the expected field at multiple locations, essentially an inverse problem.

MFP has been largely developed in ocean acoustics [47] for source localization and tomography. The replica fields are obtained by applying the acoustic wave equation to waveguide propagation models. In contrast to the typical underwater application, there is no closed form solution for the array manifold vectors in scattering environments due to the complex nature of propagation. The array manifold is a matrix where each column describes how acoustic propagation alters the source signal received at the sensors from

each location. In the beamforming literature, each column is known as a steering vector, as it is used to "steer" the focus of the array in a particular direction.

In the case of plane-wave propagation, each steering vector accounts for the relative delays and attenuation between array elements and is dependent on the geometry of the array only. To represent the field in general without any simplifying assumptions, the steering vector - essentially a transfer function - must contain information concerning reflection, refraction, diffractions, transmission, absorption, and diffraction as appropriate. The array manifold can be obtained by measurement [48], [49] or numerical modeling. The widespread availability of urban geographic information and pseudo-static nature of the geometry makes this option attractive. Regardless of the employed approach, model mismatch is a concern. This is addressed in chapter 6.

2.4 Numerical Modeling Methods for Acoustic Propagation

Because of its major importance, noise propagation in urban spaces has been studied for a long time. A history of various approaches can be found by Picaut et al. [50]. Urban propagation is a very complicated problem because of the presence and often irregular structure of many intersecting streets amid variously shaped buildings (Figure 2.5).

Obstacles give rise to an often distorted direct path and reverberation. The complexity makes an exact analytical description of sound propagation unrealistic, and accordingly, much research has turned attention towards probabilistic models. However, time reversal refocusing and matched field methods require a precise representation of the environment. It is required that the representation be able to correctly incorporate the physics of acoustic wave propagation and boundary interaction such as that which is present in the actual space. Fortunately, advances in computational power enable numerical models of sufficient complexity to be calculated. An extensive survey of available techniques has been conducted and a brief summary is presented here.

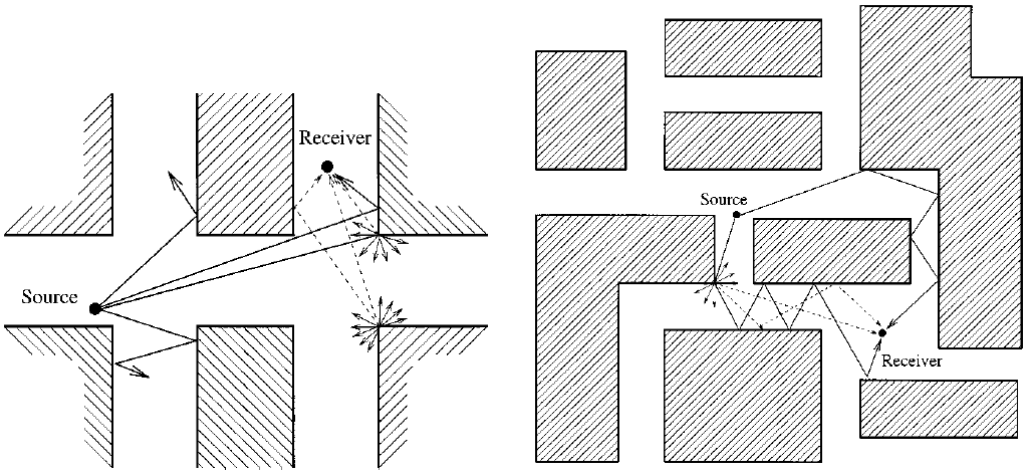

Figure 2.5: Examples of sound propagation in typical urban spaces. The full and broken lines represent the reflected and diffuse scattering fields respectively using a ray interpretation [50].

2.4.1 Wave Based Methods

The foundation of the various wave based methods is solving the acoustic wave equation. These are made up of difference methods, finite element methods and boundary element methods.

The finite difference time domain (FDTD) method approximates the relevant strong form of the differential equation as a finite difference which is then solved stepwise in time and space. Because they solve the approximated equations at each discretized point and each time step, they are flexible with regard to changing parameters spatially or temporally. As such these techniques allow complex phenomena such as acoustic propagation in wind and acoustic propagation through temperature gradients to be accurately calculated. Unfortunately, this method suffers from excessive computational time and memory requirements as wavelength of interest decreases. Regardless, due to the superior flexibility, accuracy, and ease of implementation, the FDTD is the method of choice for this work; the FDTD method for acoustic propagation will be covered in more detail in Chapter 5.

Whereas FDTD methods approximate the strong form of a partial differential equation, the finite element method is an approximation of the weak form [51]. Continuous structures are discretized into finite elements. The resulting sparse matrix is then inverted to obtain the solution to an $\boldsymbol{Ax} = \boldsymbol{b}$ type problem. Most finite element methods have similar discretization requirements as the FDTD (although not quite as demanding) and therefore are not that much faster. While finite element methods are able to handle complex geometries and boundaries with relative ease and irregular surfaces between different media can be defined with great accuracy, urban areas are typically characterized by simple rectangular geometries. The spectral element method is essentially a higher order finite element method where the behavior of each element is substantially more complex [52]. Finite element methods are not well suited for free-field radiation problems because it is difficult to include infinite elements.

Boundary element methods include the equivalent source method [53]. The solution to the free field wave equation is used to construct fields based on multiple sources. As opposed to the finite element method, only the boundaries are discretized. This leads to small, dense matrices as opposed to large, sparse matrices. The problem with these techniques is that the solution to the wave equation may not be readily calculated if the medium itself is changing due to spatial variations in acoustic properties such as speed of sound or wind. In addition, these techniques still require a fine 2-D grid and hence are computationally intensive with large memory requirements.

2.4.2 Statistical Methods

Statistical methods are not based on a deterministic model that assumes perfect knowledge of the boundary conditions, forcing functions, etc., and so may robust against mismatch between the assumed model and actual conditions. Statistical Energy Analysis (SEA) is the most commonly used statistical method. This is a very efficient technique for calculating the energy in systems at high frequencies. However it is based on the premise that a sub-system can be viewed as containing a large number of modes that transfer energy to other sub-systems in a statistical way. In the example of outdoor propagation of sound, the systems are largely "open" without any significant

reverberation or modal characteristics. Therefore Statistical Energy Analysis is not directly applicable in its current form.

The Energy Finite Element Analysis Method is used for similar applications to SEA but uses a set of discrete elements which pass energy between each other. The advantage of this method over SEA is that it can give the distribution of energy within a subsystem whereas SEA assumes the same energy level everywhere within a given sub-system. It has been shown that under the assumptions of plane wave propagation and a diffuse field, the equations of motion reduce to the equations for heat flow [54]. This vastly simplifies the equations and reduces the grid spacing required. It also allows the properties of the system (such as absorption) to be varied spatially. However, while the assumption of a purely diffuse environment may hold in some indoor spaces, it does not accurately represent wave propagation most environments.

There is also an Energy Boundary Element Technique [55] used for modeling interior acoustic problems but this suffers from the assumption of high frequency and spatial consistency of the properties of the medium.

Another statistical reduction method, the use of phonons, has been applied to sound propagation in a highly urban environment, such as a city block with tall buildings. In this case the sound is effectively restricted to travel along specific waveguides created by the buildings. The method developed by Bullen and Fricke [56] assumes the energy to be diffusely reflected from one building surface (or wall) to the next and the overall sound level throughout the environment can be rapidly estimated using very simple geometric relationships running through a simple probabilistic Markov process. However, this is a boundary technique that does not allow the properties of the medium to be varied spatially.

2.4.3 Ray Based Methods

Ray tracing methods find propagation paths between a source and receiver by generating rays emanating from the source position and following them individually as they propagate through the environment. They have been widely used to model the acoustics in large complex interior spaces such as concert halls and long range

outdoor/underwater propagation [57]. Versions of this technique include the image source method for simple geometries, path tracing and beam tracing. These methods are very computationally expensive and only valid when the wavelength in small compared to surface area and roughness of an obstacle.

Chapter 3:
Detection and Classification

The general estimation problem is characterized by the processing of signals that are distorted or corrupted in some unknown manner for the purpose of extracting information. Detection can be regarded as a subset of the estimation problem where the solution space is restricted to a binary decision: discrimination between a target's absence or presence. A successful detection process detects the presence of targets while avoiding false detections when no targets are present. Detection and classification are coupled processes in that the appropriate detector design is a consequence of the signal of interest. In this work, we are interested in the acoustic domain which influences the transducer used, frequency range of interest, and nature of interference to name a few. Due to electrical system noise, atmospheric disturbances, spurious reflections, and other signal distortions, it is not possible to determine with absolute certainty whether or not a target is present. Thus one must infer from the imperfect observation whether or not a target is present, and detection theory provides a means for choosing a good technique for doing so [58].

3.1 Coarse Classification

The goal of classification is to correctly label an input signal as belonging to one of several classes that define some library. The input signals, in general, may be one dimensional as in acoustic localization or two dimensional as in vision applications. Beyond target identification, classification aids data association processes immensely and allows detection of dissimilar sources that are not spatially separated. Sound sources can be classified in a hierarchy of categories varying in the level of detail. For example, the

acoustic radiation emitting from a Porsche 911 2.0 liter A series could be classified thusly, or progressively coarser as a car, a machine, or simply a tonal source. Because detectors are designed to respond to a specific signal model, the detection process acts as a coarse classification process as well. Coarse classification reduces the amount of data transferred within the network while capturing the main noise characteristics.

The acoustic energy present in the environment can be parsed into four groups: noise and three categories of signal type: impulsive, tonal, and broadband. These categories each vary greatly in their temporal and spectral characteristics. By identifying gross differences between the categories, features can be exploited for separation of sound types. The acoustic energy of impulsive signals, such as gunshots, is transient in time and occupies a wide bandwidth in the frequency domain. The acoustic energy from tonal signals, such as machinery, is continuous in time and concentrated in the frequency domain. Broadband signals are continuous in time and spread in frequency, such as automobile tire noise.

Due to the wide range of signals of interest, a separate detector has been designed to detect each of the three signal types in the presence of noise. As we are interested in locating point sources in space, noise is defined as acoustic energy that does not originate from a specific direction relative to the sensor array. Knowledge of characteristics of the noise is crucial, as the design of appropriate detectors is heavily influenced by not only the signal of interest but also the noise to reject. Sensor (thermal) noise, and wind are examples of noise that is spectrally similar to broadband noise. Beyond using the statistics and features of measured data, noise can also be rejected due to the lack of coherence between multiple sensors.

Finally, some consideration of the environment and propagating medium is necessary. Air changes the spectral characteristics of sound passing through it by absorbing acoustic energy in a frequency dependant manner. Although most parametric detectors are designed based on the null hypothesis of white noise, the ambient noise present in the environment is not white. However, received signals can be inverse filtered, or whitened, allowing the use of an optimal white noise detection scheme [58]. Various techniques differ in terms of performance, complexity, and effectiveness.

Two classification metrics, impulsiveness and tonality, have been devised and their properties are demonstrated here using elementary signals. In the following section it is shown how they can be used in conjunction with simple hypothesis tests to detect impulsive and tonal signals.

3.1.1 Impulsiveness Metric

Consider the discrete time (microphone) signal of length *K* samples, $\boldsymbol{r} = [r_k, r_{k+1}, \ldots, r_K]$. The corresponding impulsiveness vector is calculated by the transformation:

$$imp(k) = \frac{|r(k)|}{\text{median}(|\boldsymbol{r}|)} \tag{3.1}$$

The denominator of Eq. (3.1) is a measure of the power in the signal and so the transformation is analogous to an instantaneous signal to noise ratio. This scaling effect is convenient for comparing signals of various amplitudes as will be shown in the following section. Because most of the energy of an impulsive signal is concentrated into a few samples and the median ignores outliers, the impulsiveness will be very high. On the other hand, the consistent energy of a pure tone or Gaussian noise will counterbalance large amplitudes in the numerator.

Crest factor, or peak to average ratio, is a similar concept that scales a signal to create a dimensionless quantity. However, the average, or mean value has zero breakdown resistance [73] and will be ultimately corrupted by a single outlier. This severely limits crest factor usefulness as an impulsiveness metric.

3.1.2 Tonality Metric

The tonality metric is a frequency domain version of impulsiveness. Let $\boldsymbol{r}(\omega)$ be the complex sequence of numbers resulting from a discrete Fourier transform of $\boldsymbol{r}(t)$.

The corresponding tonality vector is achieved by the transformation:

$$\boldsymbol{ton}(\omega) = \frac{|\boldsymbol{r}(\omega)|}{\text{median}(|\boldsymbol{r}(\boldsymbol{\Omega})|)} \tag{3.2}$$

where $\boldsymbol{\Omega} = [\omega_f, \omega_{f+1}, \ldots, \omega_F]$ is a vector of frequency bins corresponding to the desired bandwidth. As in the impulsive case, a vector with concentrated energy in few elements will yield high values of tonality, as the metrics tend to accentuate "peakiness".

3.2 Elementary Signals

It is useful to explore the impulsiveness and tonality of some common signals, especially Gaussian white noise for the role it plays it the design of detectors. Elementary signals are the simplest forms of a signal that retain the distinguishing characteristics. Consider an elementary impulsive, tonal, and broadband signal as defined by

$$\begin{aligned} &\text{Impulse: } \boldsymbol{r}(t) = \delta(t-\tau) \\ &\text{Tone: } \boldsymbol{r}(t) = A\sin(\omega t + \phi) \\ &\text{Broadband: } \boldsymbol{r}(t) \sim N(\mathbf{0}, {\boldsymbol{\sigma}_G}^2) \end{aligned} \tag{3.3}$$

Time and frequency domain versions of these signals are shown in Figure 3.1. The tonality and impulsiveness of a tone and impulse are easily calculated by inspection. On the other hand, the tonality and impulsiveness of the broadband signal is more complicated. The maximum expected tonality and impulsiveness of a broadband signal is derived in the following subsections.

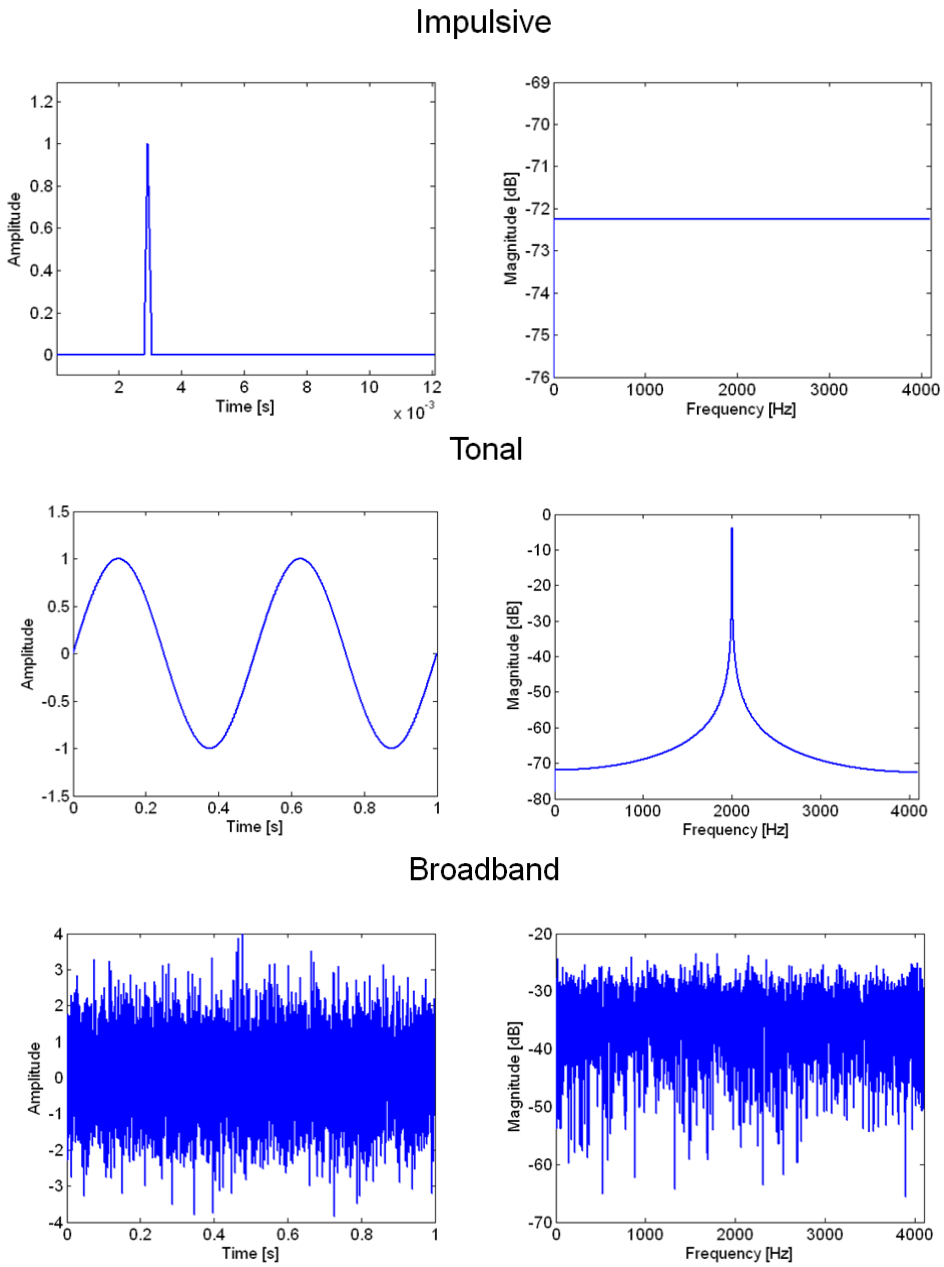

Figure 3.1: Time and Frequency domain plots of elementary signals.

3.2.1 Impulsiveness of white noise

Assume the broadband signal is a zero mean normally distributed random sequence with probability density function (PDF)

$$p_G(z) = \frac{1}{\sigma_G\sqrt{2\pi}} e^{\frac{-z^2}{2{\sigma_G}^2}} \tag{3.4}$$

and cumulative density function (CDF)

$$C_G(z) = \frac{1}{2}\left(1 + erf\frac{z}{\sigma_G\sqrt{2}}\right) \tag{3.5}$$

where erf is the error function

$$erf(z) = \frac{2}{\pi}\int_0^z e^{-x^2}\,dx \tag{3.6}$$

The absolute value operator of the impulsiveness metric causes the probability of negative values to vanish. If the random variable is zero mean, the probability of positive values double and the cumulative density function becomes

$$C_G(|z|) = erf\frac{|z|}{\sigma_G\sqrt{2}} \tag{3.7}$$

A comparison between the standard CDF and the absolute value CDF are shown in Figure 3.2.

The median of a probability distribution is the value at which half of the cumulative density is on either side. Therefore, the median value can be solved for to be

$$z_{median} = \sigma_G\sqrt{2}erf^{-1}\ 0.5\ = 0.6745\sigma_G \tag{3.8}$$

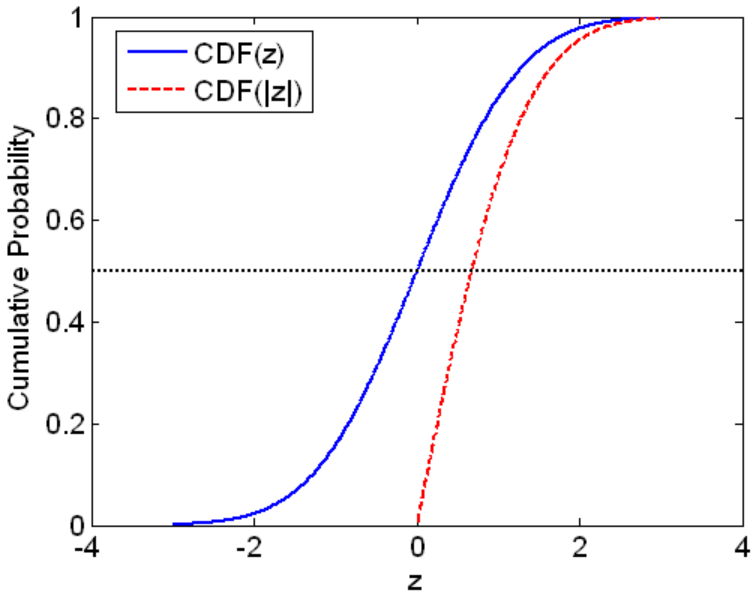

Figure 3.2: Cumulative density functions for a Gaussian distribution and the absolute value of the Gaussian distribution (unity variance). The median value occurs at a cumulative probability of 0.5.

To solve for the maximum expected value, first consider the probability of one occurrence of a high value

$$Pr_{high}(\kappa\sigma_G < \boldsymbol{r} < \infty) = 2\int_{\kappa\sigma_G}^{\infty} f_G(z) \tag{3.9}$$

where the magnitude of the high value, $\kappa\sigma_G$, is given as a constant multiple of the standard deviation. The probability of the high value occurring in a block of data of length *K* samples follows a binomial distribution

$$Pr_K = 1 - \left(1 - Pr_{high}\right)^K \tag{3.10}$$

Therefore, the maximum expected impulsiveness of white noise increases with block size. Figure 3.3 shows this relationship for a range of κ values. As an example, if we let $K = 8192$ and $Pr_{\text{high}} = 0.95$, $\kappa \cong 3.5$.

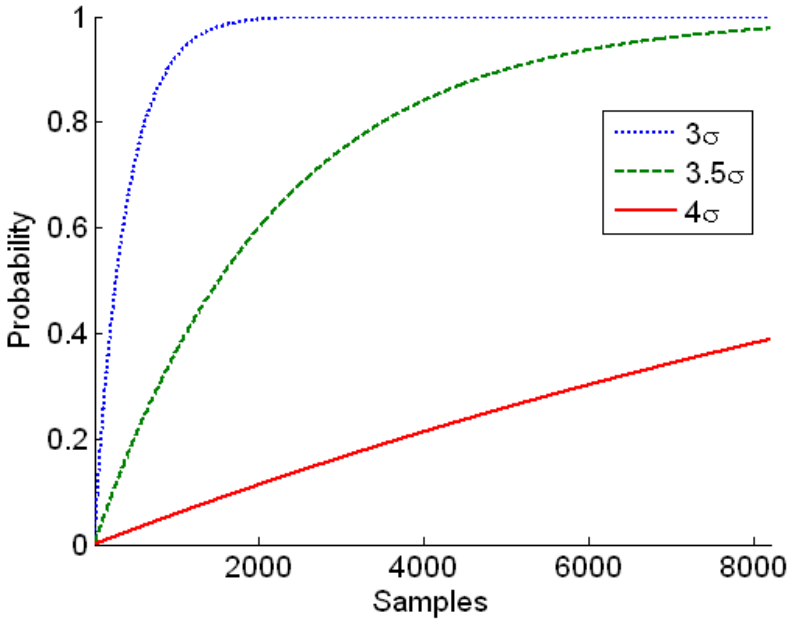

Figure 3.3: Maximum expected value of a block of white Gaussian noise.

3.2.2 Tonality of white noise

The tonality of Gaussian white noise is similarly of interest. The Fourier transform of a zero-mean Gaussian sequence, $\boldsymbol{r}(t) \sim N(\mathbf{0}, {\sigma_G}^2)$, yields a complex spectrum; the real and imaginary parts of which are independent zero mean Gaussian random sequences with variance σ_R. The Fourier transform preserves the power of the time domain signal, and for the case of white noise, divides it equally into the available frequency bins. Assuming a N-point transform where $N = K$, the number of frequency bins is also K with a 1 Hz resolution. Half of these bins correspond to negative frequency, the value of which is the complex conjugate of the positive frequency result. In most applications, it is more meaningful to ignore the negative side and double the amplitude of the positive frequency half. Based on these assumptions, the variance of the real and imaginary components of each bin is

$$ {\sigma_R}^2 = 2\frac{{\sigma_G}^2}{K} \tag{3.11}$$

The real and imaginary parts can be combined to obtain the magnitude of spectrum, i.e.

$$|R(\omega)| = \sqrt{\mathbb{R}\{R(\omega)\}^2 + \Im\{R(\omega)\}^2} \tag{3.12}$$

The amplitude of each frequency bin is then distributed as a Rayleigh random variable with PDF

$$p_R(z) = \frac{z}{{\sigma_R}^2} e^{\frac{-z^2}{2{\sigma_R}^2}} \tag{3.13}$$

where σ_R is the standard deviation of the normally distributed real and imaginary components, not of the Rayleigh distributed magnitude. Incidentally, the variance of a Rayleigh random variable is $\frac{4-\pi}{2}{\sigma_R}^2$. The median of the Rayleigh distribution can be expressed simply as $z_{\text{median}} = \sigma_R\sqrt{\log 4}$. The maximum expected value of the distribution can be determined similarly to the time domain case using eq. (3.9) and eq. (3.10).

Table 3.1 shows a summary of the maximum expected impulsiveness and tonality for the three elementary signals of Eq. (3.3). Values for the random noise case are based on a block size of $K = 8192$ samples and 95% confidence level.

	Impulse	Tone	Broadband Noise
	$r(t) = \delta(t-\tau)$	$r(t) = A\sin(\omega t + \phi)$	$r(t) \sim N(0, \sigma_G)$
$E\{\max(\boldsymbol{imp})\}$	∞	$\sqrt{2}$	5.2
$E\{\max(\boldsymbol{ton})\}$	1	∞	3.23

Table 3.1: Summary of impulsiveness and tonality of elementary signals.

3.3 Binary Hypothesis Testing

A method of identifying when a signal source is present in the influence field of the sensors is desired. For a set of observational data and probabilistic models which may have produced the data, statistical hypothesis testing provides a framework for making decisions in some optimal sense and quantifying the confidence of this choice. In the hypothesis testing problem, we assume that there are two possible states of nature, H_0 and H_1, corresponding to two possible probability distributions P_0 and P_1, on the observation space Γ. This can be written as

$$H_0: y \sim P_0 \quad \text{versus} \quad H_1: y \sim P_1$$

Selection between two hypotheses can be made with a binary hypothesis test, which formulates a decision rule δ. Two classical techniques of detection theory are Bayesian and Neyman-Pearson. The Bayesian methodology makes an optimal decision in terms of minimizing the overall expected cost, or average risk, which requires knowledge of the prior probabilities of hypotheses and the cost of wrong decisions. In many problems, such as the one of target detection considered here, the imposition of a specific cost structure on the decisions is not possible or desirable.

The Neyman-Pearson criterion offers an alternative by setting the probability of false alarm (P_F) to an appropriate constant value and then maximizing the probability of detection (P_D). A trade off exists between these two quantities which is described by the detector's receiver operating characteristics. P_F is known as a type 1 error and is defined as the probability that a source is identified as present when it is not (H_1 selected when H_0 is true). The probability of a miss (P_M) is known as a type 2 error and is defined as the probability that a source is identified as absent when it is present (H_0 selected when H_1 is true). $P_D(\delta) = 1 - P_M(\delta)$ which is also known as the power of the test. Let $p_0(y)$ and $p_1(y)$ be the probability density functions for an observation set y under hypotheses H_0 and H_1 respectively. The quantity

$$L(y) = \frac{p_1(y)}{p_0(y)}, \quad y \in \Gamma \tag{3.14}$$

is defined as the likelihood ratio between H_0 and H_1. A theorem attributed to Neyman and Pearson [58] shows that the most powerful test with a given false-alarm constraint is found from the likelihood-ratio test and given by the following decision rule:

$$\delta_{NP} = \begin{cases} 1 \ if \ L \ \ y \ \ \geq \eta \\ 0 \ if \ L \ \ y \ \ < \eta \end{cases} \tag{3.15}$$

where

$$Pr(L(y) \geq \eta | H_0) = \int_{\eta}^{\infty} p_0(y)\, dy \tag{3.16}$$

and η is the threshold. The maximum probability of detection obtained by this optimal test is

$$P_D = \Pr \ \ L \ \ y \ \ \geq \eta \,|\, H_1 \ \ = \int_{\eta}^{\infty} p_1 \ \ y \ \ dy \tag{3.17}$$

A summary of these quantities appears in Figure 3.4 based on the example of location testing with Gaussian error. p_1 and p_2 are normally distributed with identical variance but different mean values μ_1 and μ_2.

In the following subsections, binary hypothesis tests are developed for each of the three coarsely classified signal types. As discussed, if a signal is detected using a particular detector, the signal is classified as that signal type. It is entirely possible that a received signal may satisfy the true hypothesis for both the impulsive and tonal detectors, for example. This is indicative of a complex waveform, such as speech, and requires more advanced classification tools to sufficiently isolate.

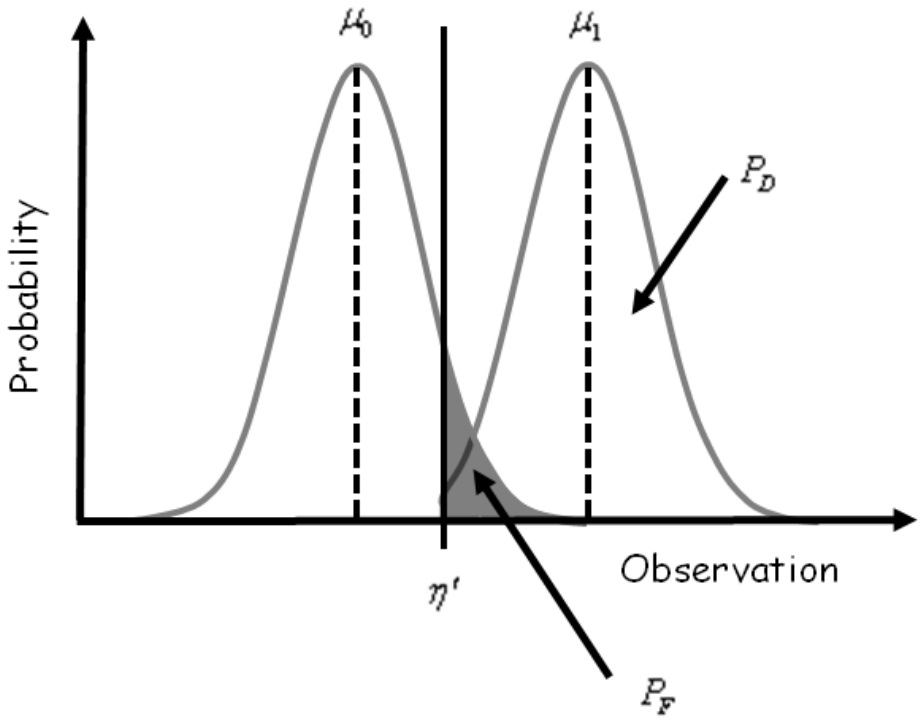

Figure 3.4: Neyman and Pearson binary hypothesis testing considering location testing with Gaussian error.

3.4 Impulse Detector

3.4.1 Approach

The defining feature of an impulsive signal is relatively high amplitude over a short time duration. The problem of detecting transient signals appears in many engineering applications, such as seismology and sonar, and a variety of approaches exist. One of the most straightforward is short time energy detection (see section 3.6), however, this method will be triggered by all signal types, limiting classification ability. On the other end of the complexity spectrum are methods that propose specific models of the impulse and seek to estimate many properties thereof, such as waveform, arrival time and duration [59], [60]. Unfortunately, the added complexity limits the applicability of

methods in situations where the anticipated signal is arbitrary and algorithm efficiency for real time processing is desired.

Instead of attempting to model the impulsive waveform, the approach presented here is to locate isolated peaks assumed to correspond to the beginning of a transient signal, i.e. the nature of the impulse tail is ignored. This is accomplished using a simple scalar test based on the well known example of location testing with Gaussian error [58] in conjunction with the impulsiveness metric. As the basic test is triggered simply by high amplitudes, the observations are first transformed into the impulsive domain which serves multiple purposes. First, high amplitude samples in the time domain corresponding to non-impulsive signals will be reduced in the impulsive domain. Secondly, the normalization of signal power achieved by the impulsiveness transformation eliminates the necessity to estimate the noise power level. The threshold of the location test required for a desired P_F is dependent on the noise power of the null hypothesis. By first transforming the data into the impulsiveness domain, the power becomes a known constant. Although the noise level can be estimated using a moving average scheme, this has its drawbacks. Typically, the noise power is estimated using a moving average during segments of no activity, requiring additional design parameters. The moving average cannot be updated when there is activity and is always lagging the true value, as it is dependent on a mixture of the signal history and current data.

3.4.2 Formulation

We seek to detect a sample of amplitude μ in zero mean white Gaussian noise with unknown variance ${\sigma_G}^2$. For the proposed test, the observation *y* is one sample of discrete time domain data from a single microphone channel that has been scaled by eq. (3.1). To be consistent with earlier examples, let *K = 8192*. For the following derivation, it is assumed that the sample corresponding to the impulse has a negligible influence on the median value of the block; i.e. the median of the signal plus noise data is equivalent to the noise data.

The time domain and impulsiveness domain hypotheses are as follows. A crucial aspect of the transformation from raw time data to impulsiveness for detection is that the random signal now has a known variance

$$\begin{aligned} H_0: Y \sim N(0, {\sigma_G}^2) &\rightarrow H_0: Y_I \sim |N(0, {\sigma_I}^2)| \\ H_1: Y \sim N(\mu, {\sigma_G}^2) &\rightarrow H_1: Y_I \sim |N(\mu_I, {\sigma_I}^2)| \end{aligned} \tag{3.18}$$

where $\mu_I = \frac{\mu}{0.6754\sigma}$ and $\sigma_I = 1.484$. The subscript 'I' refers to an impulsive domain quantity. The likelihood ratio for the two hypotheses is given by

$$L\ y\ = \frac{\frac{1}{\sigma_I\sqrt{2\pi}} e^{-\ y-\mu_I\ ^2/2{\sigma_I}^2}}{\frac{1}{\sigma_I\sqrt{2\pi}} e^{-y^2/2{\sigma_I}^2}} = \exp\left\{ \frac{\mu_I}{{\sigma_I}^2}\left(y - \frac{\mu_I}{2} \right) \right\} \quad for\ y \geq 0 \tag{3.19}$$

by integrating the probability of the null hypothesis in the rejection region ($y > \eta$), we obtain

$$P_F = P_0(L(Y) > \eta) = 1 - \varphi\left(\frac{\eta'}{\sigma_I}\right) \tag{3.20}$$

where

$$\eta' = \frac{{\sigma_I}^2 \log\ \eta}{\mu_I} + \frac{\mu_I}{2} \tag{3.21}$$

and φ denotes the CDF of a $N(0,1)$ random variable. Note that the probability of false alarm is a function of the threshold only.

The probability of detection is obtained in a similar fashion by integrating the probability of the alternative hypothesis in the acceptance region($y > \eta$):

$$P_D = P_1(L(Y) > \eta) = 1 - \varphi\left(\frac{\eta' - \mu_I}{\sigma_I}\right) \tag{3.22}$$

A parametric plot of the relationship between the probability of false alarm and the probability of detection considering various SNRs appears in Figure 3.5. This plot is known as a ROC (receiver operating characteristic) curve. SNR is calculated according to the definition according to the time domain quantities:

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \frac{\mu^2}{{\sigma_G}^2} \tag{3.23}$$

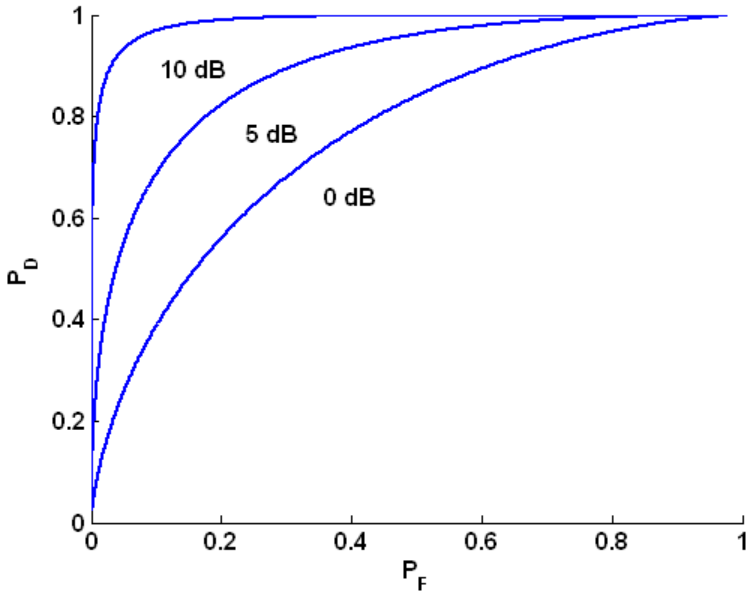

Figure 3.5: Receiver operating characteristic for the impulse detector.

The threshold is based on the desired false alarm rate and can be achieved exactly by rearranging Eq. (3.20)

$$\eta_0' = \sigma_I \varphi^{-1}(1 - P_F) \tag{3.24}$$

where φ^{-1} is the inverse function of φ. The Neyman-Pearson test for this case is then given by

$$\delta_{NP}\ y = \begin{cases} 1\ if\ \ y \geq \eta' \\ 0\ if\ \ y < \eta' \end{cases} \tag{3.25}$$

3.4.3 Discussion

This test does neglect some realistic qualities of a typical impulsive signal, such as reverberation. Reverberation will increase the median of the data block, the extent of which is dependent on the damping level of the environment. The test can be suspended for a short time after detection so that the reverb tail does not trigger additional false detections.

The usefulness of the impulsiveness metric is as classifier, rejecting loosely time invariant signals such as noise and tonal signals as will be shown explicitly in section 3.7. The detector itself is essentially equivalent to the detector for basic location testing with Gaussian error. In this case however, the noise level is a assumed constant value. Signal free estimates of the noise floor can be difficult to obtain in practice. Furthermore, the estimates lag behind the detection process. This is circumvented by using the median as an instantaneous estimate of the noise power (which can include any non-impulsive masking interference).

This leads to an important quality of the impulsiveness detector: the threshold defines a minimum signal to noise ratio. Because the mean and variance of the null hypothesis is fixed, the signal to noise ratio is dependent on the signal strength only. Consider the following two examples of the signal in noise hypothesis: in the first case, the impulse is located at $k = 35$ and the amplitude is 10. The noise is such that SNR = 20dB. In the second case, the impulse is located at $k = 65$ and the amplitude is 65. The noise is also such that SNR = 20dB. In both cases, $K = 100$ for ease of visibility.

Although these signals have different amplitudes, they have the same impulsiveness as shown in Figure 3.6.

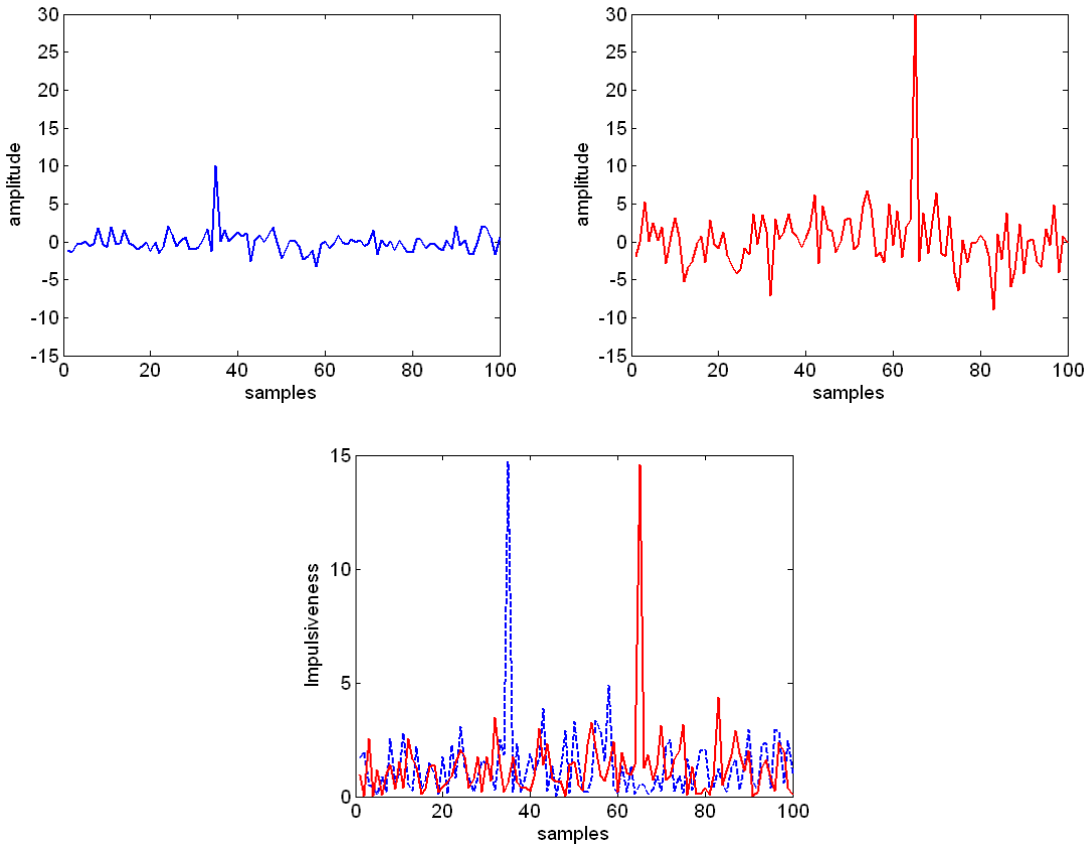

Figure 3.6: Two examples of an impulse in noise (top left and top right). The SNR is 20dB in both cases. The impulsiveness of each signal is identical (bottom).

Although presented here as a constant false alarm rate detector, the impulsiveness metric can be used to formulate a constant SNR detector. Because the median is related to the power in the signal, this detector would be useful for other null hypotheses beyond noise, for example, detecting an impulse embedded in a sine wave.

3.5 Tone Detector

The detection of tonal sources is pursued in a fashion similar to the impulse detector. Significant peaks in the frequency domain are sought using a version of the

location test described above. Consider the detection of a pure tone $r(t) = A\sin(\omega t + \phi)$ in additive zero mean white Gaussian noise with unknown variance ${\sigma_G}^2$. A FFT is performed on the time-domain data and the magnitude of the spectrum is scaled as per Eq. (3.2). It is assumed that ω is a multiple of the fundamental frequency of transform, where $f_0 = f_s/K$ and f_s is the sampling frequency.

The hypothesis test is performed on each bin in the tonality domain; the two hypotheses are

$$\begin{aligned} H_0 &: Y \sim R\ \ {\sigma_T}^2 \\ H_1 &: Y \sim N\ \ \mu_S, {\sigma_S}^2 \end{aligned} \tag{3.26}$$

where

$$\sigma_T = \frac{1}{\sqrt{\ln(4)}} \tag{3.27}$$

and

$$\mu_s = \frac{A}{\sigma_R\sqrt{\ln(4)}} \tag{3.28}$$

The subscript 'S' refers to a quantity of the source and 'T' refers to the tonality domain (note that the signal quantities are also in the tonality domain). The likelihood ratio for (3.26) is given by

$$L\ \ y\ = \frac{\frac{1}{\sigma_S\sqrt{2\pi}}e^{-y^2/2{\sigma_S}^2}}{\frac{ye^{-y^2/2{\sigma_T}^2}}{{\sigma_T}^2}} = \frac{{\sigma_T}^2}{y\sigma_S\sqrt{2\pi}}e^{-y^2/2{\sigma_S}^2+y^2/2\sigma_T} \tag{3.29}$$

by integrating the probability of the null hypothesis in the rejection region ($y > \eta$), we obtain

$$P_F = P_0 \ \ L \ \ Y \ \ > \eta \ \ = \exp\left(-\frac{\eta^2}{2{\sigma_T}^2} \right) \tag{3.30}$$

Similarly, the probability of detection is obtained by integrating the probability of the alternative hypothesis in the acceptance region($y > \eta$):

$$P_D = P_1(L(Y) > \eta) = 1 - \varphi\left(\frac{\eta - \mu_s}{\sigma_s}\right) \tag{3.31}$$

where $\sigma_s \cong \sigma_T$. ROC curves for the tonality detector are given in Figure 3.7 and Figure 3.8. Due to the dependence of ${\sigma_R}^2$ on K as given by eq. (3.11), the performance of the tonality detector increases with sample size. Note that the center curve in both plots is derived using identical parameters so that the relative effects of SNR and sample size can be compared. The SNR is defined in terms of the time domain quantities:

$$SNR_{dB} = 10\log_{10}\frac{\left(\frac{A}{\sqrt{2}}\right)^2}{{\sigma_G}^2} \tag{3.32}$$

The desired false alarm rate can be achieved by exactly by rearranging eq. (3.30)

$$\eta = \sigma_T\sqrt{-2\ln(P_F)} \tag{3.33}$$

The Neyman-Pearson test for this case is then given by

$$\delta_{NP} \ \ y \ \ = \begin{cases} 1 \ if \ \ y \geq \eta \\ 0 \ if \ \ y < \eta \end{cases} \tag{3.34}$$

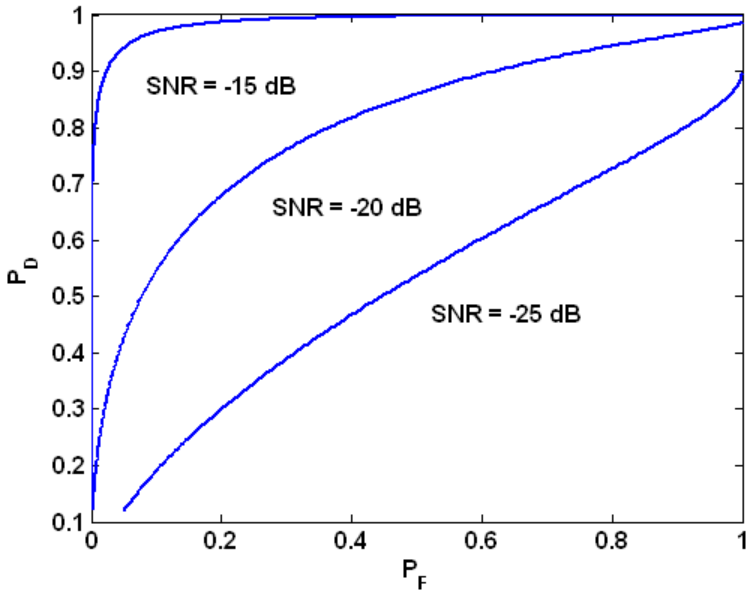

Figure 3.7: Receiver operating characteristic curves for the tonality detector given that $K = 512$.

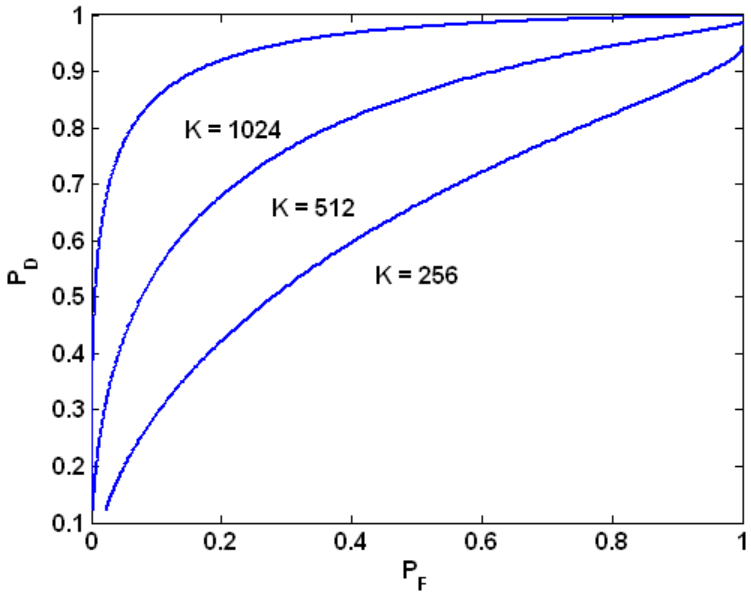

Figure 3.8: Receiver operating characteristic curves for the tonality detector given that $SNR = -20\text{dB}$.

3.6 Broadband Detector

In some cases, signals arise that are best modeled as being purely random or stochastic, such as signals with high modal density or those perturbed by propagation through turbulent media or along multiple paths. These signals may possess much acoustic energy, but are not very impulsive or tonal and therefore undetectable by the previous tests. An important special case of this problem is where both signals and noise are Gaussian random vectors or

$$\begin{aligned} H_0&: \boldsymbol{Y} \sim N(\boldsymbol{\mu}_0, {\boldsymbol{\sigma}_0}^2) \\ H_1&: \boldsymbol{Y} \sim N(\boldsymbol{\mu}_1, {\boldsymbol{\sigma}_1}^2) \end{aligned} \tag{3.35}$$

The optimum detector in this case compares the quantity $\sum_{k=1}^{K} y_k^2$ to a threshold. Since $\left(\frac{1}{K}\right)\sum_{k=1}^{K} y_k^2$ is the average energy in the observed waveform, the resulting detector structure is sometimes known as an energy detector, or radiometer. This is a common detector and thoroughly covered, for example in [58].

3.7 Impulsiveness - Tonality Feature Space

In this section, the impulsiveness and tonality metrics are applied to a range of realistic signals such as vehicles, gunshots, and speech to show the potential of these metrics as features for classification. Most of the signals acquired have been measured with an acoustic node during various field tests. Although the signals are a mixture including sensor and environmental noise, care was taken to choose examples in which the specified source of interest is dominant.

To bridge from the previous sections, consider a signal that consists of white Gaussian noise only. A spectrogram of this signal and the calculated impulsiveness and tonality appears in Figure 3.9.

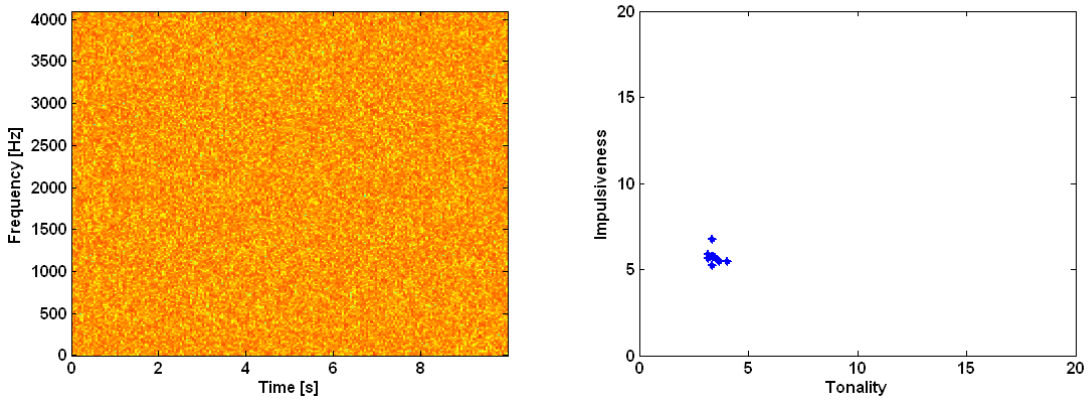

Figure 3.9: Spectrogram (left) and the impulsiveness-tonality feature space (right) for a white Gaussian noise signal.

The processing method to obtain the impulsiveness-tonality feature space is as follows. First, the discrete time domain signal is divided into one second blocks (Figure 3.10) of length $K = f_s$. For the Gaussian white noise example, $f_s = 8192$. The impulsiveness is calculated for each block and stored. Each one second block is then subject to the tonality metric. In practice, the metric is evaluated piecewise by octave bands (Figure 3.10). This serves to limit the impact of masking interference, for example, atmospheric noise typically increases with inverse frequency and would otherwise dominate the median of a spectrum that may contain high frequency tones. The choice of octave bands does introduce some variability in the results; the center frequencies are consistent for all example considered here. The maximum tonality over all the frequency bins in each octave band is compared over all octave bands and the maximum is stored. This results in one value for impulsiveness and one value for tonality per second of data. Notice how the impulsiveness and tonality values in Figure 3.9 correspond to expected maximum values calculated in section 3.2.

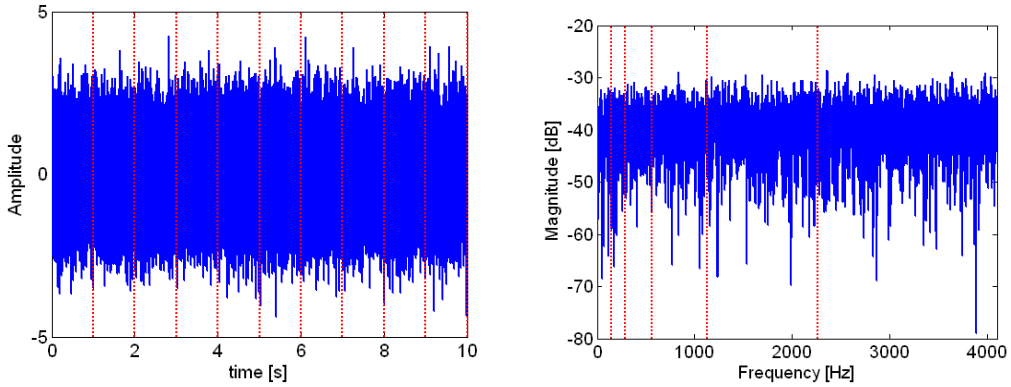

Figure 3.10: Time domain (left) and frequency domain (right) of the Gaussian white noise signal showing divisions for impulsiveness and tonality estimation.

This processing scheme was applied to other realistic signals of interest. The examples available are of varying sample rates and durations and include: wind, distant traffic, a gasoline generator, a car (Volkswagen) driving by, a tracked vehicle (tank), gunshots (pistol and shotgun), and read speech. Spectrograms and impulsiveness-tonality maps of each signal appear in Figure 3.11 to Figure 3.17. Note the changing scales of the impulsiveness and tonality axes to for clarity.

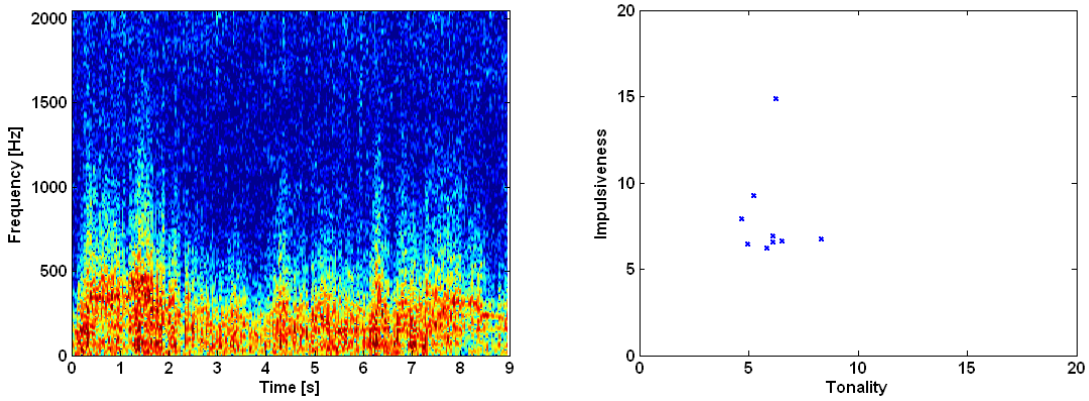

Figure 3.11: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of wind.

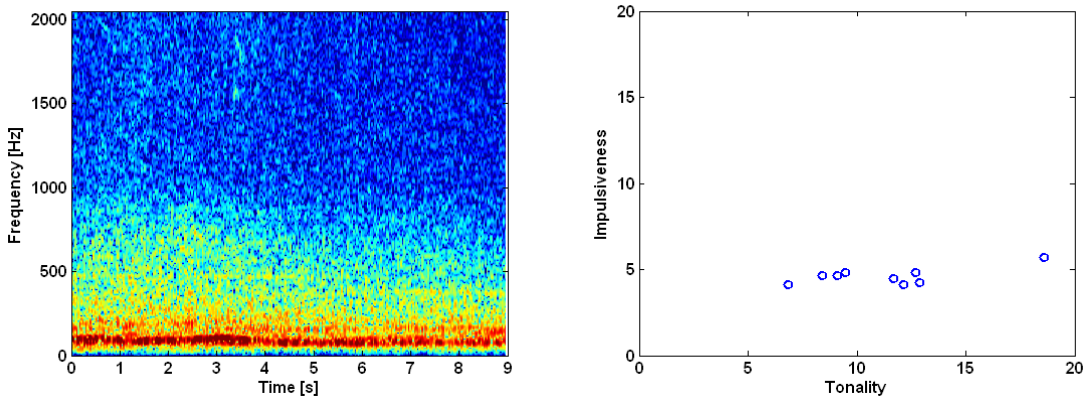

Figure 3.12: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of distant traffic from a 4-lane 45mph road.

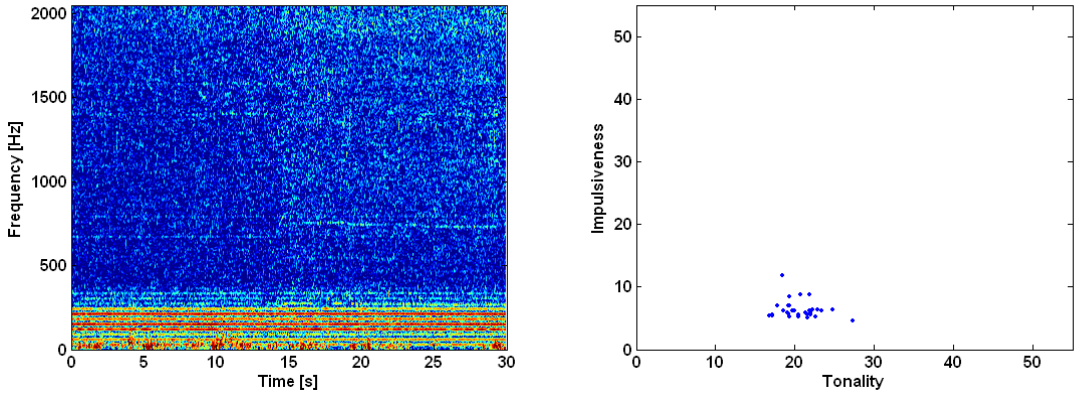

Figure 3.13: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a gasoline generator.

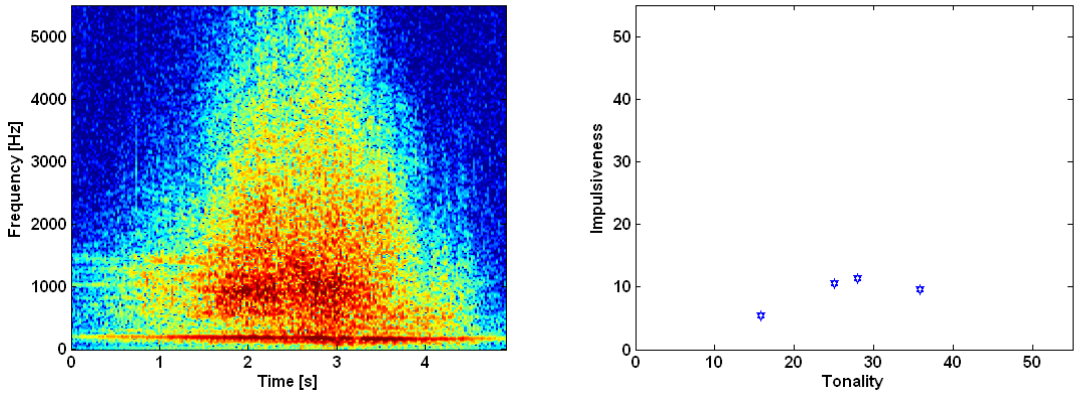

Figure 3.14: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a car driving past the microphone.

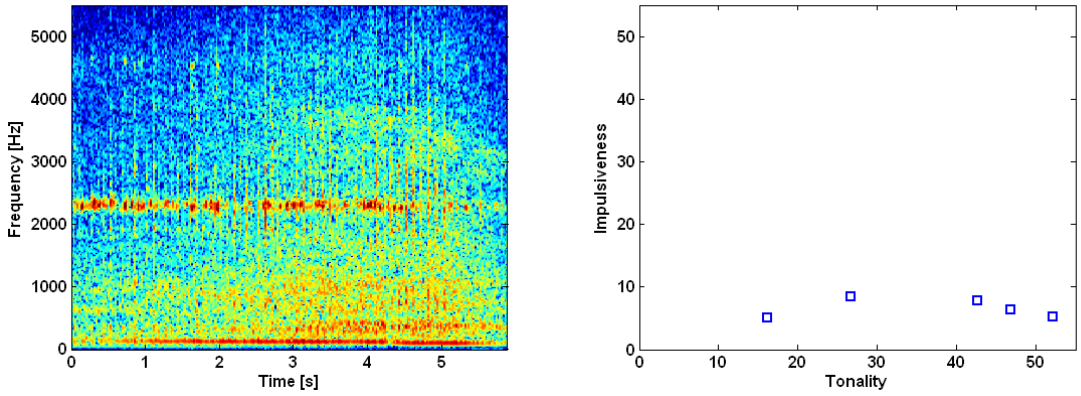

Figure 3.15: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a tracked vehicle (Sherman tank).

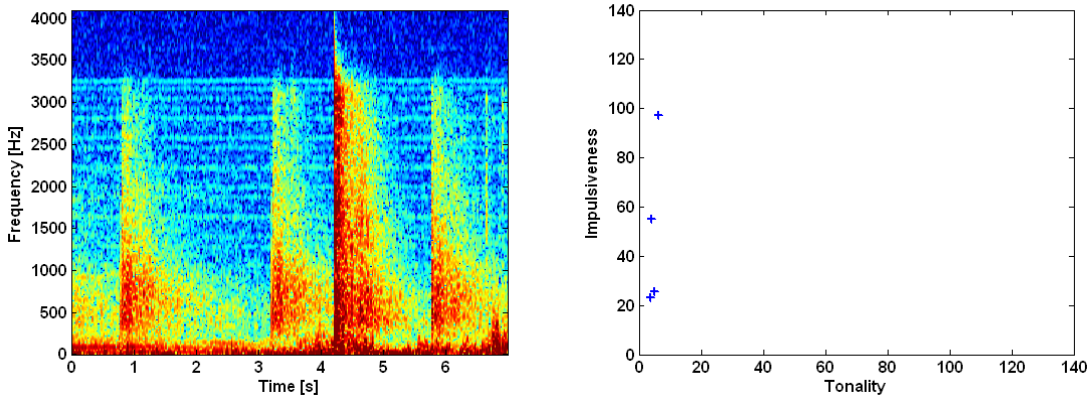

Figure 3.16: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of a gunshots (three pistol shots and one shotgun).

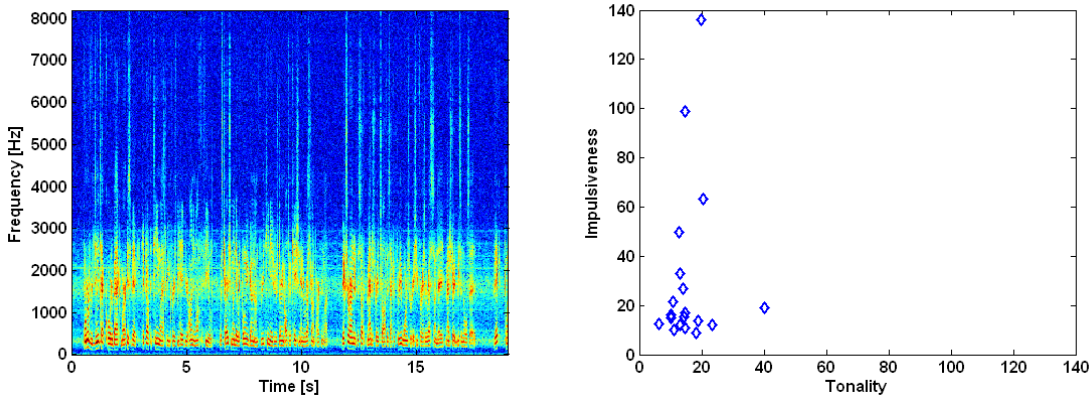

Figure 3.17: Spectrogram (left) and impulsiveness-tonality feature space (right) of a signal consisting primarily of speech (reading aloud in an anechoic chamber)

For clarity, a relative comparison of the impulsiveness and tonality of the eight signals are overplotted in Figure 3.18. It must be emphasized that many variables (window size, relationship of tone frequencies to octave bands, SNR, etc.) can affect the actual impulsiveness and tonality results as previously derived. All of these quantities are not uniform for each example signal, however, this is the reality of field measurements. For example, the car signal example goes through a range of SNRs as the vehicle approaches and then passes the microphone.

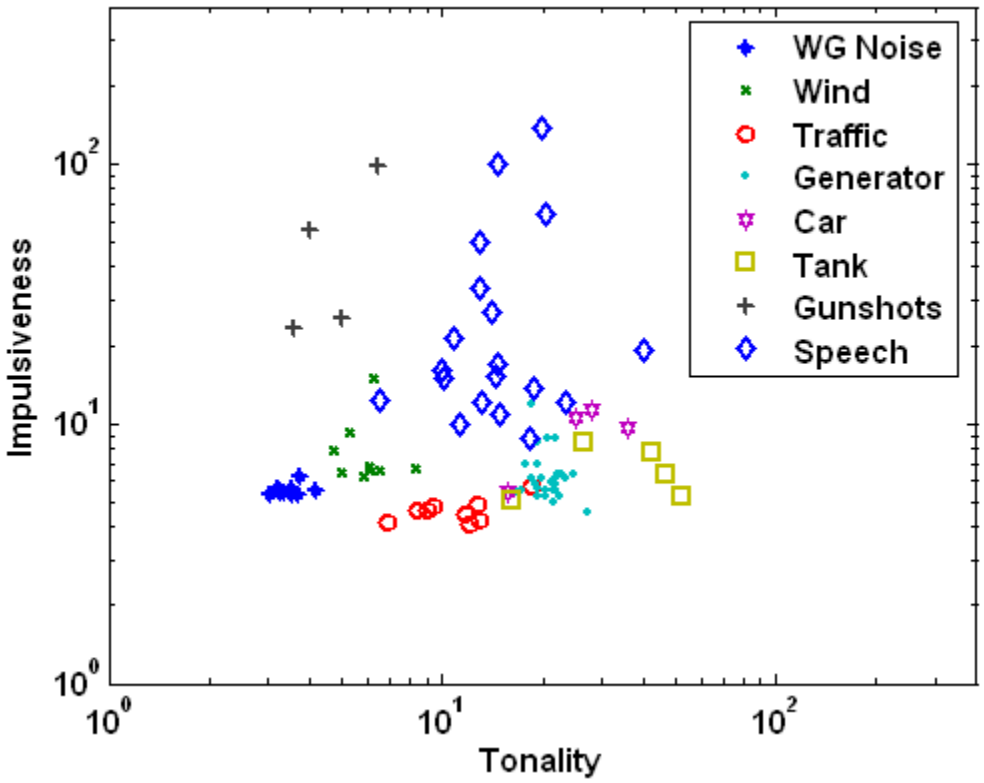

Figure 3.18: Signal data in the impulsiveness-tonality feature space.

Many signals are either predominantly impulsive (like the gunshots) or predominantly tonal (such as the machines). Signals that are not very tonal or impulsive (like the wind) carry broadband energy any can be detected as such. Perhaps the most interesting example is the speech, at times tonal, at times impulsive depending on which words were spoken and how the syllables lined up with the data block boundaries. Due

to the power of the impulsiveness and tonality features and the coarse classification desired, many signals can be classified trivially, i.e. if the signal is detected using the impulse detector, it is an impulsive signal. The ability to precisely classify complex signals on a fine scale, such as speech, requires more features and pattern-classification methods.

Chapter 4:
Association, Localization, and Tracking in a Free Field

An acoustic map of the environment can be made up of tracks of acoustic sources. Tracks are the spatial history of sources and provide a means to predict future target states. The inputs to the tracker are position fixes of acoustic events at discrete moments in time. In turn, inputs to the localization process are individual detections from the environment. This chapter develops the evolution from detections to events to tracks. Figure 4.1 shows an overview of the tracking process. Each successive step is a fusion of information that requires an accompanying association process.

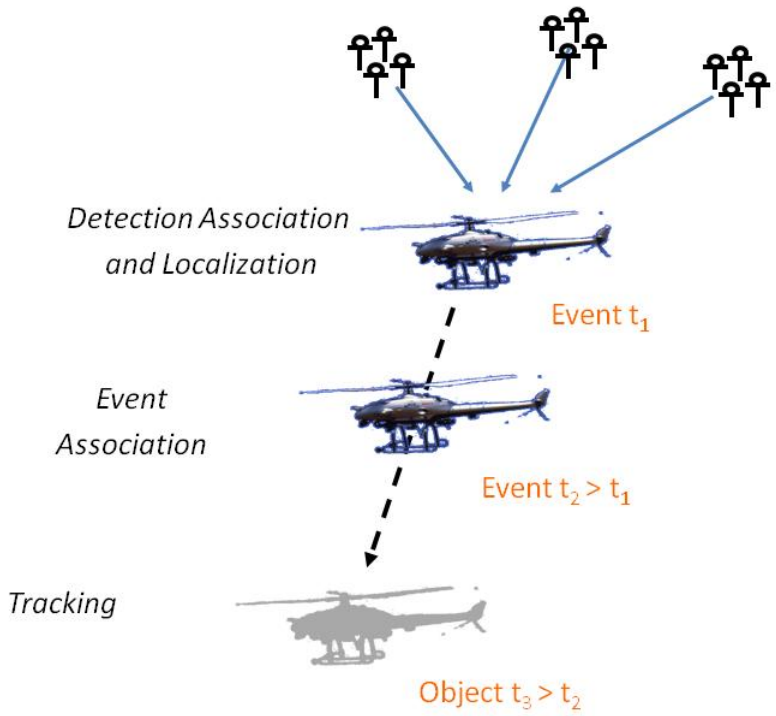

Figure 4.1: Summary of the tracking process. Over time, node detections are associated to localize events and events are associated to update the target track.

4.1 Detection Association

While active modalities (radar, sonar) are capable of obtaining both bearing and range estimates, passive modalities are generally able to estimate 1D or 2D angular position relative to the sensor only (assuming plane-wave propagation and beamforming). Angle-only tracking using techniques such as the Modified Spherical Coordinates filter is popular as it is able to provide stable accurate estimates of the observable states without a sensor platform maneuver [61]. However, passive ranging with a single sensor is possible only using an ownship maneuver under the assumption of no target maneuver. It will be shown that a network of distributed nodes (individual sensor platforms) sharing individual AOA estimates can collectively localize a source absolutely, including both maneuvering and transient sources. In general, we assume small compact nodes each equipped with a microphone array, limited onboard computational ability and a communication link to the network

Without loss of generality, consider N nodes coplanar with one far-field noise event as described by the geometry of Figure 4.2. The sound from the source at location $\boldsymbol{l}_s = [x_s, y_s]^T$ reaches the n^{th} node at location $\boldsymbol{l}_n = [x_n, y_n]^T$ with AOA estimate modeled as the true bearing angle embedded in additive Gaussian noise

$$\hat{\theta}_n = \theta_n + v_n \tag{4.1}$$

where $v_n \sim N(0, \sigma_n)$. It is possible that systematic error is present as well, and this condition is explored via simulation in section 4.3.

Absolute localization (i.e. bearing and range) using passive acoustic sensors requires the collaboration of multiple nodes. This section formulates an association process of relative bearings from two or more displaced nodes. In a multitarget environment, a strict association stage is necessary to prevent erroneous localization estimates from false intersections or 'ghosts' [62]. Even in a single source scenario, detection association is necessary to minimize track distortion due to clutter and resolve the delayed and out of sequence measurement ambiguities acoustic measurements are

prone to. The methodology presented herein uses a combination of course classification and spatial information.

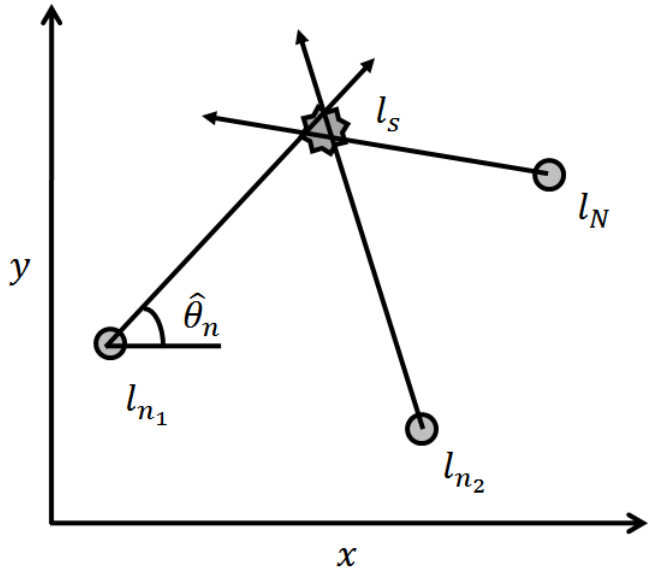

Figure 4.2: Two dimensional geometry of source detection and localization problem using multiple nodes.

4.1.1 Network Information and Communication

The network devised for the present application uses a combination of distributed and centralized processing to minimize network traffic. As the prototype network considered in this work is small enough that hopping (intermediate relaying) is unnecessary, the composite tracker system architecture is a convenient compromise between power expenditure and information transferred.

Nodes are enabled with sufficient computing power to distill the raw, high sample rate, multichannel acoustic data to reduced information sets called feature vectors (FV). Each FV contains the key information pertaining to a single detection as well as the current node status. Specifically, a FV contains the node name, node location (3D spatial coordinates), time of detection (t_d), AOA, AOA variance, impulsiveness, tonality, frequency (if tonal), octave band power levels. The AOA is derived from a beamforming algorithm local to the node. The uncertainty of the AOA measurement is based on the SNR and other characteristics individual to the node's acoustic array. Impulsiveness,

tonality, and broadband power levels are calculated according to the algorithms described in Chapter 3.

After detection and local processing, a node transmits the feature vector to the single, centralized fusion center node. Additionally, nodes can be instructed to regularly transmit a FV regardless of detection activity. This "heartbeat" provides a consistent confirmation of node presence and location updates.

4.1.2 Influence field of acoustic sensors

A. Aurora et al. [2] use the notion of an influence field, the spatial region surrounding an object in which the object causes fluctuations in one or more energy domains. In the acoustic domain, the shape and extent of an object's influence field is dependent on source directivity, amplitude, frequency content, atmosphere and terrain. Every noise event can only be sensed by nodes in the object's influence field and there is a finite time delay from event to detection. Assuming free field propagation in a homogenous medium, the delay is equal to the separation distance divided by wave speed. The varying delays in a multisource environment are dependent on source-receiver geometry and introduce measurement origin uncertainty.

In order to handle the inevitability of time delayed and out of sequence measurements, the concept of an event detection window (EDW) is introduced. Each window exclusively corresponds to an acoustic event, defined as a single impulsive event or a momentary scan of a continuous source. Upon receiving a FV, the fusion center attempts to associate it with an open EDW. If this association fails, a previously undetected event is assumed and a new window is opened at t_d. EDWs remain open for a time period based on the speed of propagating medium, node geometry, and network performance or until the window contains a detection from all available nodes. More succinctly, the closing time is given as

$$t_{EDW} = t_d + \frac{\max_N \Delta \boldsymbol{l}_{n_d n}}{c} + t_q \tag{4.2}$$

where $\Delta \boldsymbol{l}_{n_d n}$ are the distances from the detecting node to all the other nodes, c is the speed of sound, and t_q is the network communication delay. The second term allots time for the propagating wavefront to reach all nodes, by assuming that the source bearing is along the line connecting the detecting node to the farthest node (the most extreme possibility resulting in the largest time). This is illustrated by Figure 4.3.

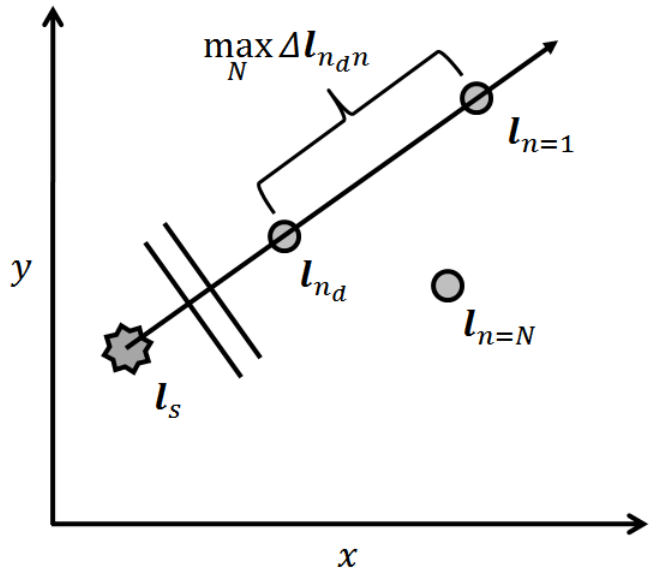

Figure 4.3: Schematic showing the most conservative estimate of propagation distance from the detecting node.

Eq. (4.2) relies on many assumptions: free field propagation, individual node processing times are equal, and consistency of the network communication delay and speed of sound. It also ignores the measured AOA. The window time can be tweaked to balance the tradeoff between acquiring more measurements for better localization versus rejecting unrelated detections and reporting events in real time.

4.1.3 Criteria for Detection Association

A three stage test is implemented to associate FVs in a given EDW. First, FVs classified as different signal types are rejected. Second, the LOBs must intersect, which

is dependent on the node geometry at the times of detection. An algorithm assuming coplanar node and source locations has been written. The third test is for agreement between hypothetical range estimates and the individual FV detection times, circumventing the typical bearing only limitation of passive sensors. For this test, the FVs are assumed to correspond to the same event and then triangulated to yield a temporary position fix. As a result, two range estimates are acquired to a possible source location estimate. Using the proposed position fix calculated, the time of event, t_E, is calculated by back propagating the acoustic wavefront:

$$t_{E,n} = t_{d,n} - \frac{\Delta \boldsymbol{l}_{n,s}}{c} \tag{4.3}$$

where $\Delta \boldsymbol{l}_{n,s}$ is the range from n^{th} node to the hypothetical event.

The range test is made by comparing the time of event differences to a threshold, η_r. This threshold is dependent on the uncertainty of the position fix, detection time, and speed of sound. A discussion on the uncertainty of the position fix is given in 4.2.2. The detection time quality is a function of the network time synchronization between nodes and also the sound type. For impulsive events, the time of detection can be pinpointed down to a precise instant on the scale of the sampling frequency. If the event is tonal or broadband (i.e. continuous type), the event time is a coarse estimate, dependent on the data block size at the node processing level. Finer resolution, such as that obtained by cross correlation of the receiver signals, requires substantial coherence between nodes. This is often rare in practice and is explored further in Chapter 6.

4.2 Localization by Bearing Fusion

A prime objective of a distributed sensor network engaged in target tracking is determination of the location of an object of interest. Typically, passive sensors are capable of providing a relative bearing to the target only. This section concentrates on exploring various fusion strategies to combine multiple bearings from spatially separated

nodes with a focus on accuracy and robustness to measurement bias. Pair-wise fusion strategies are developed and shown to be more accurate than the conventional linear least squares technique by minimizing the error magnification of the geometry. This is supported by two-dimensional, free-field Monte Carlo simulations and an outdoor experiment utilizing a network of passive acoustic arrays.

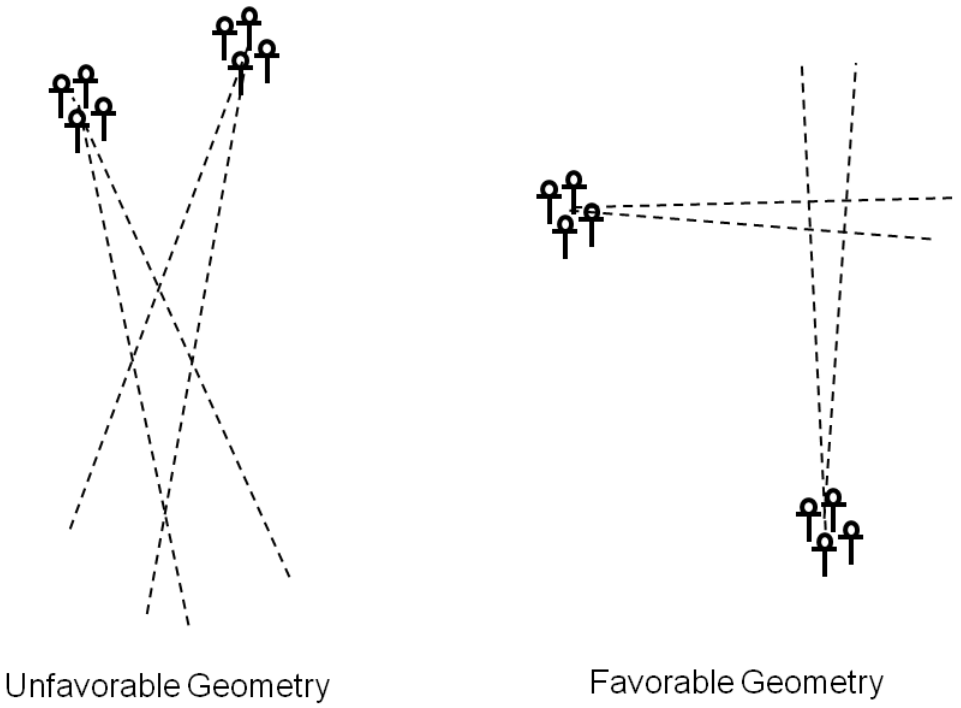

Figure 4.4: Favorable and unfavorable geometries illustrating geometric dilution of precision. The dotted lines indicate the range of bearing angle (all arrays have identical variance).

Passive target localization using bearings-only measurements is a classic estimation problem supported by previous research [63], [64] and also occurs outside of the acoustic modality, such as video in mobile robot position estimation [65]. However, with few exceptions [66], [67], the literature fails to address geometric dilution of precision (GDOP), a significant source of error in triangulation. The farther the source is from the baseline of sensors, the less accurate the position fix will be as the lines of bearing (LOBs) become increasingly parallel (see Figure 4.4). In the extreme case of collinear LOBs, the position fix coordinate along the sensor axis is indeterminate.

Alternatively, short LOBs at right angles to one another are the least sensitive to error. In [66], a closed-form expression for the gross magnification factor which transforms sensor error magnitudes into position error magnitudes is computed. This expression also produces the conditions under which triangulation becomes ill-conditioned.

To evaluate performance, pair-wise observation fusion methods are compared with a conventional least squares technique via simulation and experiment. Although the fusion strategies introduced here can be applied to bearings supplied by any sensor modality, the developed network of passive acoustic arrays is used to demonstrate the method experimentally. Experimental ground truth is provided by GPS. Although the GPS measurement itself is subject to uncertainty, it can be safely assumed to be orders of magnitude lower.

Next, several methods of combining LOBs to obtain an estimated position fix of the source in 2D Cartesian coordinates are presented. To focus on the localization process, it is assumed that all measurements are available to a fusion center and a data association process has associated bearings that originate from the same acoustic source and time.

4.2.1 Linear Least Squares Localization

For each node *n*, the linear least squares method (LLS) assumes that the source is located on a line described by the node's location and bearing estimate [63]

$$\tan(\hat{\theta}_n)x + y = l_{n,x}\tan(\hat{\theta}_n) + l_{n,y} \tag{4.4}$$

Accounting for all nodes leads to *n* linear equations with two unknowns. When the number of nodes exceeds two, the system of equations becomes over determined and multiple solutions may exist. A pseudo inverse can be used to minimize the error in a least-squares sense:

$$\hat{\boldsymbol{l}}_{LS} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \tag{4.5}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \tan(\hat{\theta}_{n_1}) & 1 \\ \vdots & \vdots \\ \tan(\hat{\theta}_{n_N}) & 1 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} l_{n_1,x}\tan(\hat{\theta}_{n_1}) + l_{n_1,y} \\ \vdots \\ l_{n_N,x}\tan(\hat{\theta}_{n_N}) + l_{n_N,y} \end{bmatrix} \tag{4.6}$$

4.2.2 Pairwise Localization and Fusion

Pair-wise localization and fusion is the concept of separately triangulating each pair of bearings into observations of target position and then fusing those observations. Assuming a constant bearing variance across all nodes, the accuracy of each pair varies due to GDOP and the extent can be gauged on a multidimensional basis by calculating a covariance matrix for each observation. A maximum likelihood estimator (MLE) is formulated using a priori knowledge to minimize the error due to GDOP. However, in the presence of unaccounted-for error, the performance of the MLE will deteriorate. This includes noise and/or bias in the AOAs due to inaccurate knowledge of the true sensor characteristics and consequences of the environment. Therefore, a robust estimator that is resistant to contamination while maintaining accuracy is desired. In this section, pair-wise localization is introduced and two fusion strategies are developed.

Consider the o^{th} pair of bearings consisting of the n_1 and n_2 nodes. They can be triangulated to form an observation $\hat{\boldsymbol{l}}_o$ according to

$$\hat{\boldsymbol{l}}_o = \begin{bmatrix} \hat{l}_x \\ \hat{l}_y \end{bmatrix}_o = \begin{bmatrix} \dfrac{(l_{n_1,y} - l_{n_2,y}) - (l_{n_1,x} - l_{n_2,x})\tan(\hat{\theta}_{n_2})}{\tan(\hat{\theta}_{n_1}) - \tan(\hat{\theta}_{n_2})} \\ (\hat{l}_{o,x} - l_{n_1,x})\tan(\hat{\theta}_{n_1}) + l_{n_1,y} \end{bmatrix}_o \tag{4.7}$$

The deviation between the observation and true value is defined as $\Delta \boldsymbol{l}_o = \boldsymbol{l}_o - \hat{\boldsymbol{l}}_o$. Using the law of propagation of error [70], the individual uncertainties in the bearing measurements can be translated to a coupled Cartesian covariance matrix. This involves finding the sensitivity of the source location to a perturbation of the bearing and taking a first order Taylor series approximation, which yields a contribution to the total differential:

$$\Delta \boldsymbol{l}_o = \begin{bmatrix} \Delta l_x \\ \Delta l_y \end{bmatrix}_o = \begin{bmatrix} \dfrac{\Delta l_x}{\partial \theta_{n_1}} & \dfrac{\Delta l_x}{\partial \theta_{n_2}} \\ \dfrac{\Delta l_y}{\partial \theta_{n_1}} & \dfrac{\Delta l_y}{\partial \theta_{n_2}} \end{bmatrix}_o \begin{bmatrix} \Delta \theta_{n_1} \\ \Delta \theta_{n_2} \end{bmatrix} \tag{4.8}$$

The total differential is squared and expectations are taken with the assumption that the measurement noise from each pair of nodes is uncorrelated. This leads to the fully coupled covariance matrix:

$$\boldsymbol{\Phi}_{oo} = E\{[\Delta \boldsymbol{l}_o \quad \Delta \boldsymbol{l}_o{}^T]\} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix}_o \tag{4.9}$$

$\boldsymbol{\Phi}_{oo}$ represents an error ellipse of the o^{th} observation in 2D space. For a multivariate Gaussian distribution, the percentage of source location estimates that fall inside the ellipse is given by [71]

$$p_{in} = 1 - e^{\frac{-\kappa^2}{2}} \tag{4.10}$$

where κ is the number of standard deviations used to form the ellipse. For example, 39.6% of estimates will fall inside a 1 standard deviation ellipse. Due to the nonlinearities in equation (4.7), the error distribution of an observation is non-Gaussian. However, simulations show that approximating the error as Gaussian is reasonable as observations fall into the error ellipse described by $\boldsymbol{\Phi}_{oo}$ at the same rate as a Gaussian distribution.

4.2.2.1 Pairwise Fusion: MLE Estimator

Using the covariance of the observation, the MLE can be derived based on linear least squares theory [72]. Let g denote the number of spatial dimensions of interest. The o^{th} pair of nodes whose detections have been associated yields a multivariate normal observation, $\boldsymbol{\lambda}_o$, with $g \times 1$ mean vector $\hat{\boldsymbol{l}}_o$ and $g \times g$ covariance matrix $\boldsymbol{\Phi}_{oo}$

$$\boldsymbol{\lambda}_o \sim N_g\left(\hat{\boldsymbol{l}}_o, \boldsymbol{\Phi}_{oo}\right) \tag{4.11}$$

In this case, $g = 2$ as all nodes and source are considered coplanar. Using the inverse square root covariance matrix, the observation is transformed to a variable, $\tilde{\boldsymbol{\lambda}}_o$, with unity variance:

$$\tilde{\boldsymbol{\lambda}}_o = \boldsymbol{\Phi}_{oo}{}^{-\frac{1}{2}}\boldsymbol{\lambda}_o \sim N_g\left(\boldsymbol{\Phi}_{oo}{}^{-\frac{1}{2}}\hat{\boldsymbol{l}}_o, I\right) \tag{4.12}$$

Next, form the following system of equations where N nodes have resulted in O observations:

$$\tilde{\boldsymbol{\lambda}}_{oO} = \tilde{\boldsymbol{\Phi}}_{oO}\boldsymbol{l}_s + \boldsymbol{\chi} \tag{4.13}$$

where $\tilde{\boldsymbol{\lambda}}_{oO}$ is a concatenated $Og \times 1$ vector of estimates and $\boldsymbol{\chi}$ is an $Og \times 1$ vector of normally distributed noise with zero mean and covariance equal to the identity matrix. Similarly, the concatenated covariance matrix is

$$\tilde{\boldsymbol{\Phi}}_{oO} = \begin{bmatrix} \boldsymbol{\Phi}_{11}{}^{-\frac{1}{2}} \\ \vdots \\ \boldsymbol{\Phi}_{OO}{}^{-\frac{1}{2}} \end{bmatrix} \tag{4.14}$$

The solution of the linear model is the MLE

$$\hat{\boldsymbol{l}}_{MLE} = \left(\widetilde{\boldsymbol{\Phi}}_{oO}{}^{T}\widetilde{\boldsymbol{\Phi}}_{oO}\right)^{-1}\widetilde{\boldsymbol{\Phi}}_{oO}{}^{T}\tilde{\boldsymbol{\lambda}}_{oO} = \sum_{o=1}^{O}\boldsymbol{W}_o\hat{\boldsymbol{l}}_o \tag{4.15}$$

where $\boldsymbol{W}_o$ is the optimal weight for the o^{th} observation given by

$$\boldsymbol{W}_o = \left(\sum_{o=1}^{O}\boldsymbol{\Phi}_{oo}\right)^{-1}\boldsymbol{\Phi}_{oo}{}^{-1} \tag{4.16}$$

and the covariance matrix of the MLE is given by

$$\boldsymbol{\Phi}_{MLE} = \sum_{o=1}^{O}\boldsymbol{W}_o\boldsymbol{\Phi}_{oo}\boldsymbol{W}_o{}^{T} \tag{4.17}$$

Figure 4.5 is a schematic demonstrating pairwise fusion using the MLE in a scenario where $n = 3$. The first 3 subplots show the LOBs and covariance matrix of the available observations. Calculation of $\boldsymbol{\Phi}_{oo}$ for each pair assumes all nodes have equivalent bearing uncertainty. One bearing pair (upper right) is the ideal node-source geometry, while the other two pairs exhibit mild GDOP. The final subplot (lower right) shows the tighter fused covariance matrix using all three observations.

Equation (4.18) is valid only if independent observations are used, such as using the nodal pair (1,2) as one observation and (3,4) as a second when $n = 4$. Forming observations from all possible pair combinations results in dependant observations. However, this improves the location estimate significantly and is necessary to utilize all nodes when an odd amount is available. The maximum number of realizable observations O can be calculated as a binomial coefficient:

$$\binom{N}{2} = \frac{N!}{2!(N-2)!} = O_{max} \tag{4.18}$$

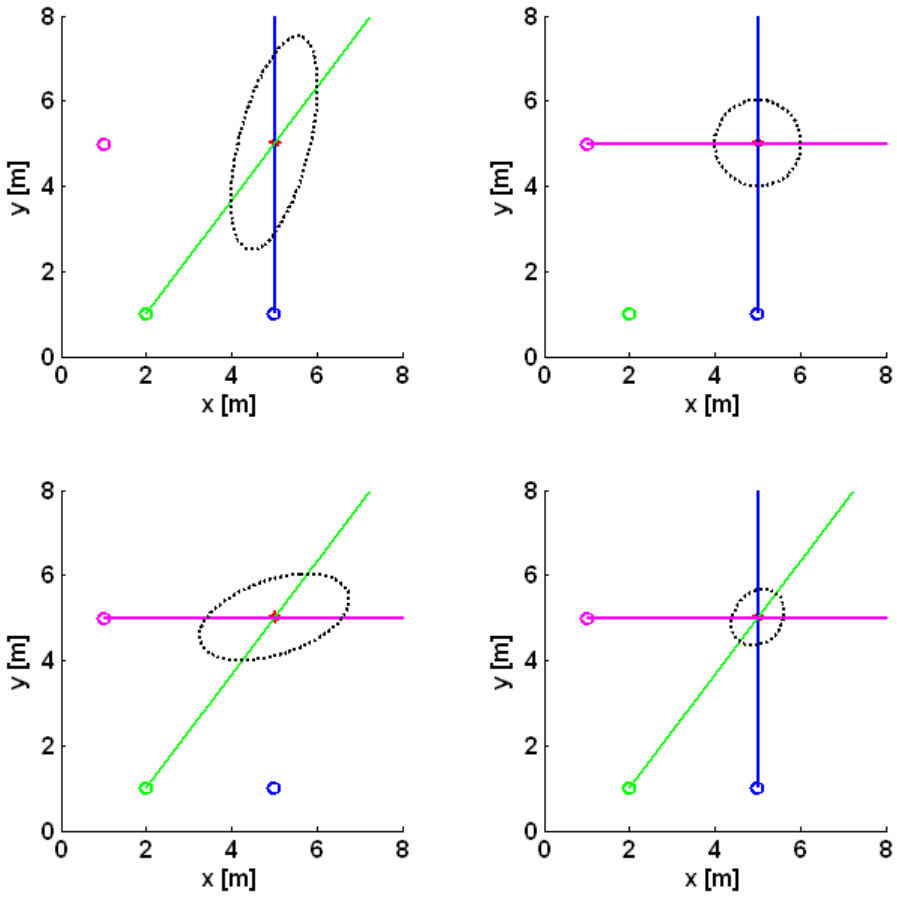

Figure 4.5: Demonstration of the pairwise fusion using the MLE estimator.

If all pairs contribute to the localization estimate, some dependencies will exist. For $n = 3$, the three pairs (1,2) (1,3) and (2,3) each share a common node. The covariance matrix must be augmented with additional terms, otherwise it is excessively optimistic (a consequence of transforming less angular degrees of freedom into more Cartesian degrees of freedom). $\boldsymbol{\Phi}_{io}$ for $i \neq o$ represents a measure of dependency between two observations. Accounting for dependent observations, the covariance of the MLE when using all observations is

$$\boldsymbol{\Phi}_{MLE} = \sum_{o=1}^{O} \sum_{i=1}^{O} \boldsymbol{W}_o \boldsymbol{\Phi}_{oi} \boldsymbol{W}_i^T \tag{4.19}$$

The validity of (4.19) has been confirmed by eq. (4.10) and numerical simulations.

The MLE is conditioned only against error described by the covariance matrix used to weight the observations. If the bearing error is not normally distributed or some contamination is present, e.g. bias due to array orientation, faulty sensors, or other real world conditions, the accuracy of the MLE will suffer; it is no longer the maximum likelihood estimator if the distribution of error changes.

4.2.2.2 Pairwise Fusion: Median Estimator

The median minimizes the L1 norm of the error and is a robust estimator that achieves the maximum breakdown point of 0.5 [73]. The MED estimator, $\hat{\boldsymbol{l}}_{MED}$, is defined as the coordinate-wise median of all the available observations:

$$\hat{\boldsymbol{l}}_{MED} = \begin{bmatrix} \text{median}\left(\hat{l}_{x,o}, \hat{l}_{x,o+1}, \ldots, \hat{l}_{x,O}\right) \\ \text{median}\left(\hat{l}_{y,o}, \hat{l}_{y,o+1}, \ldots, \hat{l}_{y,O}\right) \end{bmatrix} \tag{4.20}$$

This estimator is a robust way of fusing observations without regard to the source of uncertainty.

4.3 Localization Simulation

A Monte Carlo simulation was designed to test the fusion strategies under a variety of geometrical and additive noise conditions. The simulation is designed to estimate an unknown source location given that the measurement association step has been performed flawlessly, and the node locations and bearing uncertainty are known exactly.

4.3.1 Simulation Method

A scenario is defined as a spatial arrangement of source and nodes, with random locations on a 100×100 meter grid according to a uniform distribution. For each scenario, the AOA from the source to each node is calculated and corrupted with instances of Gaussian noise as in eq. (4.1), thereby creating many trials within one scenario. Sporadically, scenarios were discarded if the additive bearing noise and an unforgiving geometry repeatedly prohibited all of the LOBs to intersect. The standard deviation of the bearing error was 5° for all nodes. A separate case in which one node is given a 30 degree bias is also examined. Scenarios with 3, 4, 5, and 6 nodes were examined, resulting in 3, 6, 10, and 15 observations to be fused respectively.

A graphical example of 1 sample of 1 scenario from the bias case is shown in Figure 4.6. The 10 observations are at the intersections of each bearing pair, 4 of which are affected by the biased LOB.

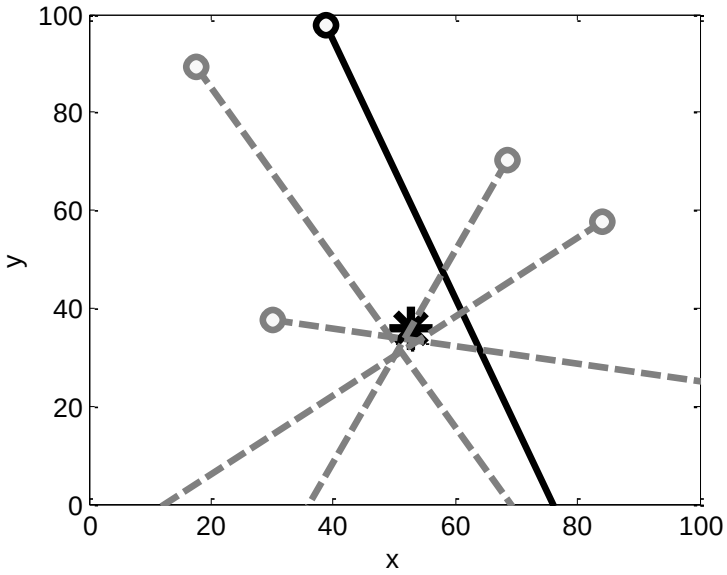

Figure 4.6: One sample of a scenario showing five nodes (circles) and LOBs to the source (asterisk). 1 node (solid black) is biased creating three skewed observations and one outlier (intersection is beyond the range of the plot).

4.3.2 Simulation Results

Monte Carlo simulations were run until the mean RMS error converged over the many scenarios. Due to the increased complexity, this required many more scenarios for the biased case as exemplified by the 3 node case in
Figure 4.7.

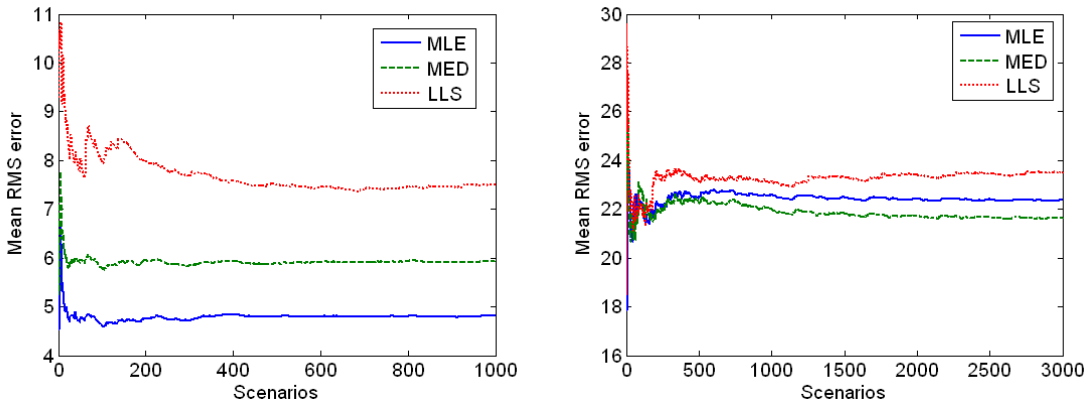

Figure 4.7: Convergence of the mean RMS error during the Monte Carlo simulations for the noise-only case (left) and biased case (right).

Table 4.1 and Table 4.2 show the mean and max RMS error for the case of bearings with added noise only. The pair-wise MLE consistently gives the closest estimate to the true value. The pair-wise MED yields maximum values similar to the MLE, as observation outliers are caused by strong GDOP. Overall, it is not as accurate as the MLE due to trials when the GDOP is weak. While all estimators generally improve with the number of observations, the LLS estimator does not take advantage of the quantity of measurements as efficiently as the other estimators. The LLS estimator suffers greatly from GDOP, evident by the long tailed error distribution shown in Figure 4.8.

	Nodes			
Fusion	3	4	5	6
LLS	8.64	7.5	7.01	6.75
MLE	6.69	4.83	3.97	3.5
MED	7.72	5.94	4.87	4.36

Table 4.1: Mean RMS error (meters) of the three fusion strategies for the simulation with added noise only.

	Nodes			
Fusion	3	4	5	6
LLS	80.21	52.43	40.71	61.49
MLE	70.34	12.45	8.85	6.89
MED	96.1	17.15	9.58	8

Table 4.2: Maximum RMS error (meters) of the three fusion strategies for the simulation with added noise only.

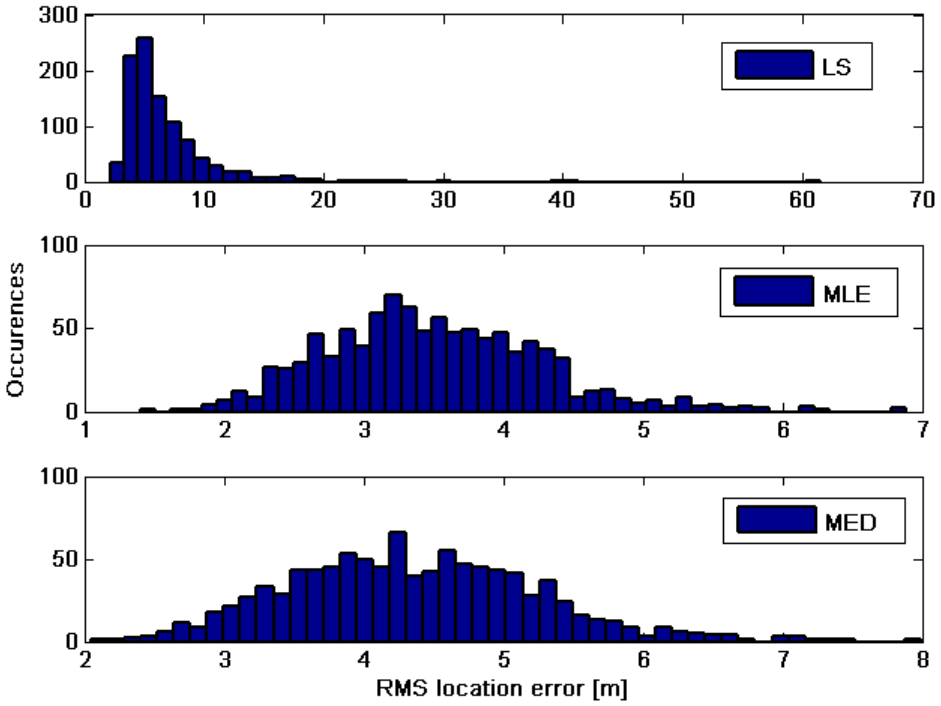

Figure 4.8: Histograms showing the distribution of RMS error for the 6-node, additive noise only case for the LLS estimator (top), MLE estimator (middle), and MED estimator (bottom).

Turning attention to the biased case, the accuracy of all of the estimators has suffered and the MED estimator gives the best performance. The mean and maximum error values are shown in Table 4.3 and Table 4.4. Because the observations are dependant, one biased node has a large influence over small sample sizes. The MED estimator becomes very accurate when the ratio of unbiased observations is greater than 50% as in the five and six node cases. The MLE, being simply a weighted average, has no resistance to the biased node and in fact weights biased observations heavily if the geometry is favorable. A detailed example of a distribution for the biased case appears in Figure 4.9.

	Nodes			
Fusion	3	4	5	6
LLS	23.5	17.18	14.87	12.58
MLE	22.39	17.3	15.48	14.83
MED	21.65	13.3	9.08	7.11

Table 4.3: Mean RMS error of the three fusion strategies for the simulation with additive noise and one biased node.

	Nodes			
Fusion	3	4	5	6
LLS	269.1	288.62	1005.78	82.05
MLE	84.3	66.64	67.46	64.01
MED	124.7	41.64	24.08	18.18

Table 4.4: Maximum RMS error of the three fusion strategies for the simulation with additive noise and one biased node.

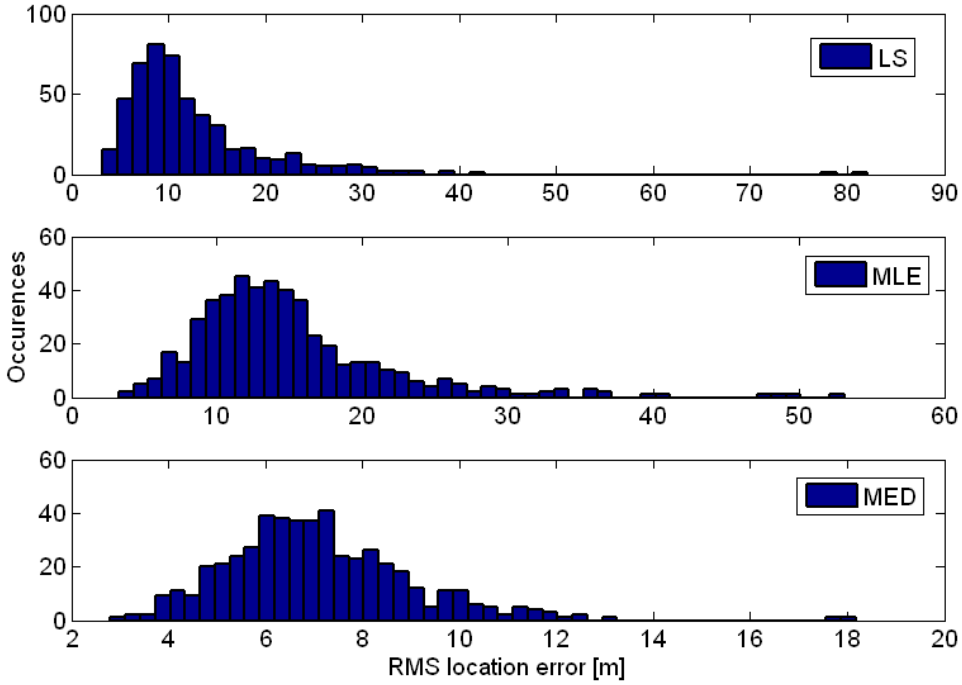

Figure 4.9: Histograms showing the distribution of RMS error for the 6-node, additive noise plus bias case for the LLS estimator (top), MLE estimator (middle), and MED estimator (bottom).

4.4 Tracking

4.4.1 The Kalman Filter

A four state linear coupled Kalman filter is employed for tracking objects. Although AOA measurements are inherently nonlinear, the measurements are linearized during the bearing fusion process allowing use of the standard filter. The performance gains over a linear uncoupled filter are similar to that of the extended filter, with the additional benefit of robustness to initialization errors [16] . The Kalman filter works in two stages, prediction and correction, to estimate the state of a dynamic system. Prior to incorporation of measurements, the state vector, $\xi \in \mathbb{R}^h$ is projected ahead according to a discrete-time difference equation of the following form:

$$\hat{\boldsymbol{\xi}}_e^- = \boldsymbol{\Lambda}\hat{\boldsymbol{\xi}}_{e-1} + \boldsymbol{\Upsilon}\Psi_e + u_{e-1} \tag{4.21}$$

where the hat denotes an estimate, the subscript e refers to the filter time step index, and the superscript minus refers to a predicted estimate before incorporation of measurements. $\boldsymbol{\Lambda}$ is the $h \times h$ state transition matrix which relates $\hat{\boldsymbol{\xi}}_{e-1}$ to $\hat{\boldsymbol{\xi}}_e^-$ according to some deterministic motion model. The matrix $\boldsymbol{\Upsilon}$ relates the control input Ψ_e to the state vector. $\boldsymbol{u}$ is a Gaussian white noise sequence that accounts for the uncertainty in the motion model.

For simplicity, the motion or process model used can be limited to the constant velocity model only due to the wide scope of targets it describes. Also, while a lower order model is subject to divergence during (e.g. constant acceleration) maneuvers, it allows for better position estimates with fewer measurements [16]. The motion model defines the state transition matrix

$$\underbrace{\begin{bmatrix} x_s(e) \\ y_s(e) \\ \dot{x}_s(e) \\ \dot{y}_s(e) \end{bmatrix}}_{\xi_e} = \underbrace{\begin{bmatrix} 1 & 0 & \Delta t_{KF} & 0 \\ 0 & 1 & 0 & \Delta t_{KF} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} x_s(e-1) \\ y_s(e-1) \\ \dot{x}_s(e-1) \\ \dot{y}_s(e-1) \end{bmatrix}}_{\xi_{e-1}} \tag{4.22}$$

where Δt_{KF} is the time step duration of the tracker. Process noise is added to the velocity estimate to account for dynamics that are not accounted for by the model. The discrete process noise matrix, $\boldsymbol{\Phi}_{uu}$, can be calculated from the continuous process noise matrix, $\breve{\boldsymbol{\Phi}}_{uu}$:

$$\boldsymbol{\Phi}_{uu} = \int_0^{\Delta t_{KF}} \boldsymbol{\Lambda}(\tau)\breve{\boldsymbol{\Phi}}_{uu}\boldsymbol{\Lambda}(\tau)d\tau \tag{4.23}$$

where

$$\boldsymbol{\Phi}_{uu} = E\{\boldsymbol{u}\boldsymbol{u}^T\} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & \sigma_u^2 \end{bmatrix} \tag{4.24}$$

and σ_u^2 is the continuous process noise spectral density.

The measurement or observation model is given by:

$$\boldsymbol{\Xi}_e = \boldsymbol{\Theta}\hat{\boldsymbol{\xi}}_e + \boldsymbol{v}_e \tag{4.25}$$

$\boldsymbol{v}_e$ is a Gaussian white noise sequence that accounts for the uncertainty in the measurement and is assumed to be independent of $\boldsymbol{u}$. The measurement noise covariance matrix $\boldsymbol{\Phi}_{vv}$ is determined from the node bearing fusion explained previously. For example, considering a fused measurement using the MLE $\boldsymbol{\Phi}_{vv} = \boldsymbol{\Phi}_{MLE}$. The $g \times h$ matrix $\boldsymbol{\Theta}$ relates the process state to the measurement $\boldsymbol{\Xi}_e = \hat{\boldsymbol{l}}$. In the present application, the matrix is simple as the measurements are in the same form as the states:

$$\boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{4.26}$$

The prediction stage, occurring at every time step, includes both predicting the state ahead and projecting the a priori estimate error covariance, $\boldsymbol{\Phi}_{\bar{\xi}\bar{\xi},e}$, given by

$$\boldsymbol{\Phi}_{\bar{\xi}\bar{\xi},e} = \boldsymbol{\Lambda}\boldsymbol{\Phi}_{\xi\xi,e-1}\boldsymbol{\Lambda}^T + \boldsymbol{\Phi}_{uu} \tag{4.27}$$

where $\boldsymbol{\Phi}_{\xi\xi,e-1}$ is the a posterior error covariance of the estimate. When a measurement is available, the update equations are used to correct the estimate. First, the Kalman gain is computed:

$$\boldsymbol{\Pi}_e = \boldsymbol{\Phi}_{\bar{\xi}\bar{\xi},e}\boldsymbol{\Theta}^T\left(\boldsymbol{\Theta}\boldsymbol{\Phi}_{\bar{\xi}\bar{\xi},e}\boldsymbol{\Theta}^T + \boldsymbol{\Phi}_{vv}\right) \tag{4.28}$$

The Kalman gain is a weighting based on the uncertainty of the new measurement versus how well the filter has been performing. The a priori estimate is updated with the measurement weighted by the Kalman gain

$$\hat{\boldsymbol{\xi}}_e = \hat{\boldsymbol{\xi}}_e^- + \boldsymbol{\Pi}_e\left(\boldsymbol{\Xi}_e - \boldsymbol{\Theta}\hat{\boldsymbol{\xi}}_e^-\right) \tag{4.29}$$

Lastly, the estimated error covariance matrix is updated

$$\boldsymbol{\Phi}_{\xi\xi,e} = (\boldsymbol{I} - \boldsymbol{\Pi}_e\boldsymbol{\Theta})\boldsymbol{\Phi}_{\xi\xi,e}^- \tag{4.30}$$

The posterior error covariance matrix provides a measure of how well the filter is tracking the states. Consistently acquiring measurements that agree with the predicted states will result in a small covariance, whereas dissentious and/or absent measurements will cause the covariance matrix to grow.

4.4.2 Event to Track Association

Measurements are provided to the Kalman filter via an event to track association process. The goal is to decide, which target, if any, a given event originated from. This is accomplished by matching attributes - such as class type and location - between the event and current tracks at the time of event. Similar to the detection association, event-track pairs with dissimilar class types are discarded. The remaining pairs are subject to the global nearest neighbor strategy [10]. The correlation decision is based on a Chi-squared test of the following form:

$$\left[\hat{\boldsymbol{\xi}}_{xy} - \boldsymbol{\Xi}\right]^T\left[\boldsymbol{\Phi}_{\xi\xi,xy} + \boldsymbol{\Phi}_{vv}\right]^{-1}\left[\hat{\boldsymbol{\xi}}_{xy} - \boldsymbol{\Xi}\right] < \eta_{NN} \tag{4.31}$$

where $\hat{\boldsymbol{\xi}}_{xy}$ denotes the position components of the object at the time of event and $\boldsymbol{\Phi}_{\xi\xi,xy}$ is the position components of the covariance matrix relating to the uncertainty in the current position. The decision criterion, or "gate" is denoted by η_{NN}. A track predicted

to be close to the event will fit in the gate, and uncertainty in the track or event location widens the gate. If no existing objects are within the gate, a new track is initiated.

Of course, it is possible the measurement satisfies the gate of multiple tracks or multiple measurements satisfy the fall within the gate of a single target track. This is known as the assignment problem and there are multiple approaches to conflict resolution [74]. The nearest neighbor strategy used here allows at most one measurement to update a given track.

4.5 Experimental Validation

Numerous tests have been performed over the course of this research. This section presents one such test in detail.

4.5.1 Experimental Setup

The system has been implemented experimentally to determine the effectiveness and accuracy at tracking acoustic noise sources. In collaboration with other Virginia Tech programs, The Autonomous Aerial Vehicle Team (AAVT) and The Joint Unmanned Systems Test, Experimentation and Research (JOUSTER), the system was recently deployed to track a small remote control helicopter. Three acoustic nodes and one fusion center were deployed at Kentland Farms facility at Virginia Tech. Two of the acoustic nodes were deployed on the ground, while the third was deployed on a remote control ATV as seen in Figure 4.10 and Figure 4.11. Each acoustic node was equipped with a 6 microphone acoustic array, 6 channel simultaneous sampling data acquisition system, and GPS and magnetometer for position and orientation. All equipment was mounted inside or on the sphere of the diffracting acoustic array. A laptop computer provided the node processing and wireless link to complete the networked system. The fusion center was a PC desktop computer with a wireless link. Common wireless internet data protocols (universal datagram protocol UDP, network time protocol NTP [75]) were

used for communication. The remote control helicopter flew around the roughly 100 square meter area between the three nodes and was equipped with a GPS unit so that the ground truth of the helicopter would be available.

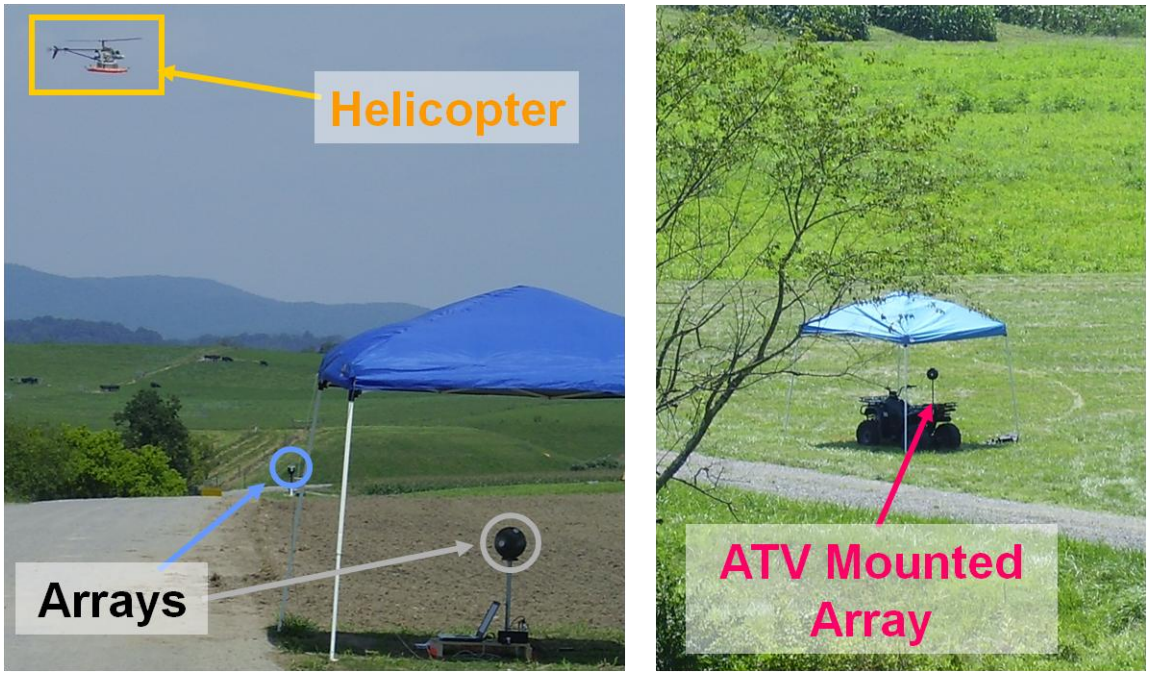

Figure 4.10: Photo of acoustic arrays and remote control helicopter.

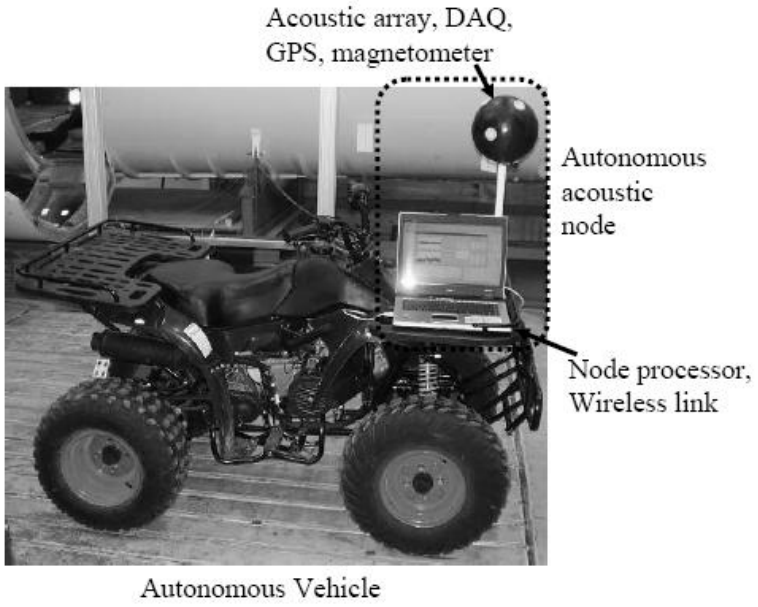

Figure 4.11: Close up photo of drive-by-wire acoustic node and key components.

4.5.2 Experimental Results

Figure 4.12 presents the real-time output of the system fusion center superimposed upon a satellite image of the test area. The plot shows the static node locations, node LOBs along with a fused observation, and error ellipse (covariance of the fused event using the MLE estimator). Additionally, the helicopter GPS track was added from post-test data processing to show the precision of system. The helicopter signature was predominantly tonal, and the first three harmonics (approximately 160, 320 and 480Hz) were used to provide the bearing estimates.

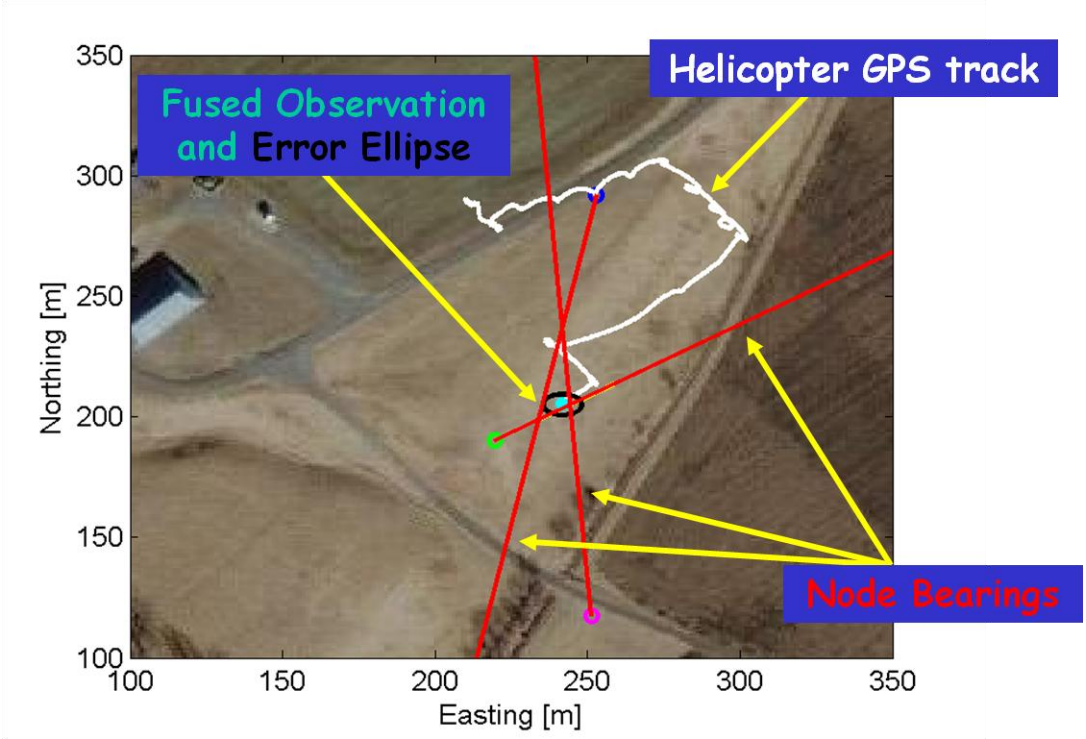

Figure 4.12. Real-time output of heterogeneous fusion center superimposed on satellite photograph of test area. Helicopter GPS track was added from post-test data processing to show the precision of the system.

The RMS error of the fused localization versus the helicopter GPS ground truth using the three fusion strategies was calculated and is shown in Table 4.5. Only localizations with contributions by all three nodes (sample size 363) were considered. Results support the unbiased, additive noise simulation. The MLE consistently provides the most accurate location estimates. Large error values for the LLS technique occurred

during particular times when the GDOP was severe. Again, a similar ceiling can be seen on the maximum error of the MLE and MED estimators, supporting the hypothesis that GDOP was the major influence on bearing error. It can be concluded that a pairwise approach utilizes the quantity and diversity of information available from a network of spatially distributed sensors.

Fusion	Mean	Min	Max	Median
LLS	19.04	0.28	294.95	14.04
MLE	8.52	0.56	58.96	6.78
MED	10.86	0.84	58.51	8.58

Table 4.5: Summary of RMS localization error (meters).

An example of the system tracking performance is shown in Figure 4.13. These plots compare estimates of the noise source position (KF track) with the known GPS track of the helicopter for the easting and northing coordinates, respectively. The uncertainty of the GPS ground truth is expected to be less than 1m. Individual acoustic events are denoted by an 'x'. As can be seen, the KF track is within 10 meters of the helicopter "ground truth" over a 100 square meter area of coverage. The interplay between the track following the motion model versus new measurements is governed by the relative magnitudes of the $\boldsymbol{\Phi}_{vv}$ and $\boldsymbol{\Phi}_{\bar{\xi}\bar{\xi},e}$ matrices.

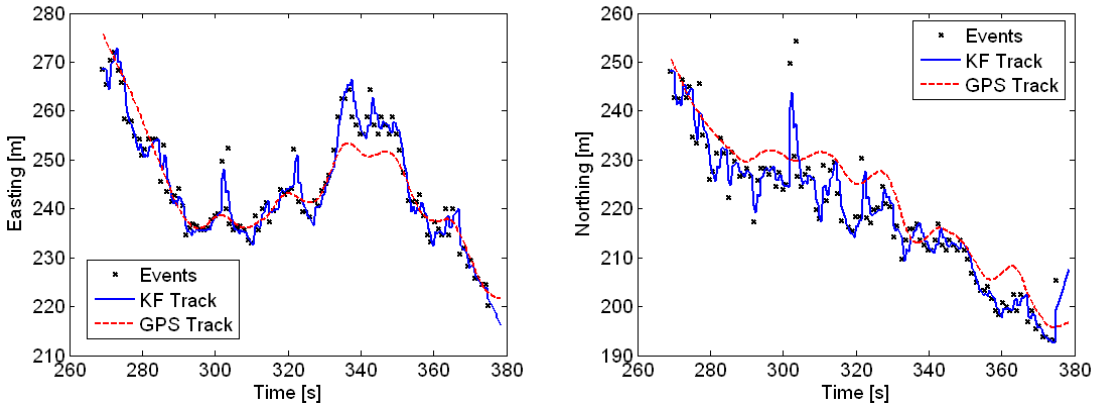

Figure 4.13: Tracking performance of target position in the Eastern (left) and Northern (right) coordinate. Plots compare estimates of noise source position (KF track) with the known GPS track of the helicopter.

Chapter 5:

Localization in Scattering Environments

In this chapter, matched field processing is presented as a method to locate acoustic sources in a scattering environment using microphone arrays. The detrimental effects of a reverberant environment on plane-wave localization techniques have been discussed in Chapter 2, and will be shown explicitly here. MFP exploits the full-field acoustic propagation to predict an expected pressure field by matching the sensor measurements with a replica or known field.

One of the great strengths of sensor networks is the multitude of information from disparate vantage points. This is especially apparent in urban environments, where large obstacles can obstruct LOS and cause heavy shadowing at some receiver locations. A schematic of the 2D node-source geometry appears in Figure 5.1. Solid black objects represent buildings which obstruct LOS and introduce multiple propagation paths, reflections, diffractions, and scattering into the propagation. All sensors and sources are assumed stationary.

In general, we assume small compact nodes each equipped with a microphone array, limited onboard computational ability and a communication link to the network. Individual microphone arrays operating outdoors are subject to atmospheric turbulence [76] especially between widely separated arrays. Furthermore, node data is referenced to a master network clock, which is subject to dissynchronization. Both of these factors compromise inter-array signal coherence. Model mismatch can also be interpreted as a loss of coherence between sensors and/or arrays. Source localization with distributed sensor arrays and partial spatial coherence has been studied in the context of plane wave

propagation [77]. In this work, coherent and incoherent matched field array processing is studied in the face of multiple forms of uncertainty.

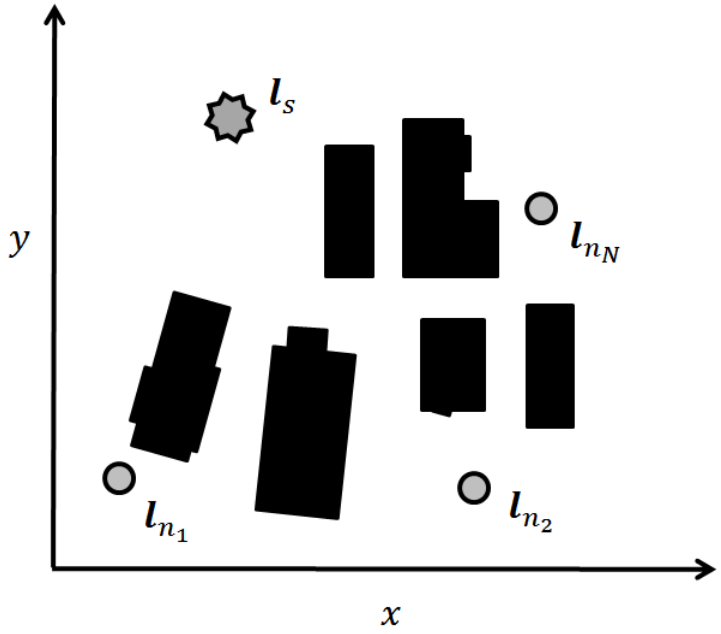

Figure 5.1: Representation of an urban setting illustrating spatial source – receiver relationships.

In the first section of this chapter, the signal processing methodology is introduced. Matched field processing is a generalization of plane wave beamforming and many of the algorithms developed in this context have been adapted to MFP. Multiarray processing methods are introduced and demonstrated with a simple example of a single scattering object in an otherwise free field. FDTD modeling is used to acquire signals at the sensor locations. Depending on the distance to the receiving array, proximity of geometric features, and nature of the medium, acoustic radiation is affected by reflections, absorption, diffraction, and other wavefield physics to a variable extent. In contrast to an acoustic free field, the distinct reverberant field strengthens matched field processing.

5.1 Theory

5.1.1 Signal Model

Consider N nodes each with M sensors and a single stationary source radiating at one of J discrete locations, $\boldsymbol{l}_n = \{x_s, y_s\}$. The sampled signal at the m^{th} microphone, $r(t)$ consists of the j^{th} source signal $s(t)$ filtered by the environment and a noise component, $v(t)$:

$$r_{n,m,j}(t) = d_{n,m,j}(t) * s_j(t) + v_{n,m}(t) \tag{5.1}$$

where $d_{n,m,j}(t)$ is the impulse response between source and receiver which captures the effects of propagation in a complex environment as well as the acousto-electric response of the receiver.

In the frequency domain, let $\boldsymbol{r}_n(\omega)$ be the vector of complex sound pressures measured by the M microphones of the n^{th} array and $\boldsymbol{s}(\omega)$ be a vector of complex source signals radiating from J locations at frequency ω. $\boldsymbol{r}_n(\omega)$ and $\boldsymbol{s}(\omega)$ are related by $\boldsymbol{D}_n(\omega)$, an $M \times J$ complex matrix of transfer functions, where $d_{n,m,j}(\omega)$ is the transfer function between the m^{th} microphone and a sound source radiating from the j^{th} location. Therefore, we have

$$\begin{bmatrix} r_1(\omega) \\ \vdots \\ r_M(\omega) \end{bmatrix}_n = \begin{bmatrix} d_{1,1}(\omega) & \cdots & d_{1,J}(\omega) \\ \vdots & \ddots & \vdots \\ d_{M,1}(\omega) & \cdots & d_{M,J}(\omega) \end{bmatrix}_n \begin{bmatrix} s_1(\omega) \\ \vdots \\ s_J(\omega) \end{bmatrix} + \begin{bmatrix} v_1(\omega) \\ \vdots \\ v_M(\omega) \end{bmatrix}_n \tag{5.2}$$

$$\boldsymbol{r}_n(\omega) = \boldsymbol{D}_n(\omega)\boldsymbol{s}(\omega) + \boldsymbol{v}_n(\omega)$$

In general, notation indicating time or frequency dependence will be suppressed when it is clear from the context.

The j^{th} column of the propagation matrix, $\boldsymbol{D}_n$, is the array steering vector for the j^{th} location. The transmission loss, a measure of the array's sensitivity to each source location, is the inner product of the steering vector (and hence real).

$$\boldsymbol{\psi}_n(\omega) = \left[\boldsymbol{d}_1^H \boldsymbol{d}_1, \boldsymbol{d}_2^H \boldsymbol{d}_2, \ldots, \boldsymbol{d}_j^H \boldsymbol{d}_j\right]_n \tag{5.3}$$

Considering the n^{th} array, the cross spectral density (CSD) of the vector of microphone outputs at frequency ω is

$$\boldsymbol{\Phi}_{RR}(\omega) = E\{\boldsymbol{r}\boldsymbol{r}^H\} \tag{5.4}$$

where E denotes expectation and H the Hermitian (complex conjugate) transpose. The diagonal elements $\boldsymbol{\Phi}_{RR,mm}$ are the power spectral density functions of the microphones $r_m(\omega)$ and convey the average received signal power with frequency. Assuming a narrowband sound source radiating from the j^{th} location, $\boldsymbol{\Phi}_{RR}$ can be decomposed into signal and noise components as follows:

$$\boldsymbol{\Phi}_{RR} = \sigma_s \boldsymbol{d}_j \boldsymbol{d}_j{}^H + \sigma_v \boldsymbol{\Phi}_{VV} \tag{5.5}$$

At any given frequency, the source radiates with power σ_s and some unknown phase. $\boldsymbol{\Phi}_{VV}$ is particular to a situation, although two special cases are of interest. When the noise is spatially white or uncorrelated from sensor to sensor, $\boldsymbol{\Phi}_{VV}$ becomes the identity matrix. Under the assumption of spatially diffuse noise, such that the signal is equally likely to have originated from any location, $\boldsymbol{\Phi}_{VV} = \boldsymbol{D}\boldsymbol{D}^H$.

5.1.2 Matched Field Processing

Matched field processing provides an estimate of the field distribution given the sample covariance matrix. The matched-field power response function is the beamformer output

$$b_j(\omega) = \boldsymbol{w}_j{}^H \boldsymbol{\Phi}_{RR} \boldsymbol{w}_j \tag{5.6}$$

where $\boldsymbol{w}_j$ is a $M \times 1$ vector of complex frequency domain sensor coefficients for the j^{th} source. The real vector $\boldsymbol{b}$ is the concatenation of j beamformer outputs and an estimate of the squared source location vector $\boldsymbol{s}^2$. The source location estimate is chosen as the argument that maximizes $\boldsymbol{b}$. There are many different choices for $\boldsymbol{w}$ based on the statistics of the data received at the array. The goal is to optimize the beamformer response so the output contains minimal contributions due to noise and signals arriving from directions other than the desired signal directions.

Matched field processing is a generalization of plane wave beamforming and many algorithms developed in this context have been adapted to MFP. The minimum variance distortionless response (MVDR) beamformer has been developed in the context of plane wave beamformers [25] and has been introduced to MFP [47], sometimes referred to as the Maximum Likelihood Method. It has a number of attributes suitable to MFP; these include good sidelobe suppression and a modest tolerance to mismatch when compared to other adaptive methods. The optimal beamformer minimizes the power of the output signal of the array. The minimization is constrained to give an undistorted signal response in the look direction in order to avoid the trivial solution of $w = 0$. The constrained minimization problem can be written:

$$min_{\boldsymbol{w}}\, \boldsymbol{w}^H \boldsymbol{\Phi_{VV}} \boldsymbol{w} \quad \text{subject to} \quad \boldsymbol{w}^H \boldsymbol{d} = 1 \tag{5.7}$$

The weights are given by

$$\boldsymbol{w}_j(\omega) = \frac{\boldsymbol{\Phi}_{VV}{}^{-1} \boldsymbol{d}_j}{\sqrt{\boldsymbol{d}_j{}^H \boldsymbol{\Phi}_{VV}{}^{-1} \boldsymbol{d}_j}} \tag{5.8}$$

The noise-only covariance matrix, $\boldsymbol{\Phi_{VV}}$, is used as inclusion of the signal in the matrix can increase sensitivity to mismatch [78] which is a concern in MFP.

Practically, it can be difficult to obtain signal free estimates of the noise cross spectral matrix. The beamformer can be designed using a theoretical noise field.

Assuming spatially uncorrelated noise only is present, $\boldsymbol{\Phi}_{VV} = I$. Substitution into eq. (5.8) results in the delay and sum beamformer. It is the optimal beamformer for optimizing the white noise gain, a quantity that indicates the ability of the array to suppress spatially uncorrelated noise:

$$WNG(\omega) = \frac{|\boldsymbol{w}^H \boldsymbol{d}|^2}{\boldsymbol{w}^H \boldsymbol{w}} \tag{5.9}$$

Another important quantity is the directivity index, which describes the ability of the array to suppress a diffuse noise field:

$$DI(\omega) = 10\, log_{10} \frac{|\boldsymbol{w}^H \boldsymbol{d}|^2}{\boldsymbol{w}^H \boldsymbol{D}\boldsymbol{D}^H \boldsymbol{w}} \tag{5.10}$$

Assuming a diffuse noise field is present, $\boldsymbol{\Phi}_{VV} = \boldsymbol{D}\boldsymbol{D}^H$. Substitution into eq. (5.8) results in a superdirective beamformer that optimizes the directivity index. This design rejects correlated noise, however, uncorrelated noise is amplified at low frequencies. In order to overcome this problem, the MVDR weight problem can be solved under a white noise constraint. The method adds a small scalar to the main diagonal of the coherence matrix:

$$\boldsymbol{\Phi}_{VV}(\omega) = \boldsymbol{D}\boldsymbol{D}^H + \alpha \frac{\|\boldsymbol{D}\boldsymbol{D}^H\|}{M} I \tag{5.11}$$

The factor α can range from 0 to infinity, which results in the unconstrained superdirective or delay and sum beamformer respectively. Although the noise field is unknown a priori, a conservative balance between directivity index and white noise gain can be made by assuming $\alpha = 0.01$, which retains most of the properties of the superdirective beamformer while maintaining sufficient white noise gain using modest equipment. This realization of $\boldsymbol{\Phi}_{VV}$ is used for all the results to follow.

5.1.3 Multiarray Processing

Incoherent network processing implies that the individual beamformer outputs of each array are combined in a magnitude only sense; no phase information is shared. For example, the outputs could be averaged such that

$$b_{inc} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{b}_n \tag{5.12}$$

Any correlation that may exist between the arrays is ignored. Although the potential for this information is lost, the processor is ultimately robust against any corruption of said information.

Coherent processing exploits the phase and magnitude relationships between microphone signals received at disparate arrays. To process the data from all arrays coherently, sensor outputs $\boldsymbol{r}_n$ from each of the N arrays are aggregated into one large vector:

$$\boldsymbol{r}_q = \begin{bmatrix} \boldsymbol{r}_{n=1} \\ \boldsymbol{r}_{n=2} \\ \vdots \\ \boldsymbol{r}_{n=N} \end{bmatrix} \tag{5.13}$$

where the subscript q indicates network. While the weight vectors $\boldsymbol{w}_{n,j}$ can be similarly aggregated, this is a suboptimal approach that leads to increased sidelobe levels; the network has a significantly different transmission loss function than the individual arrays. Instead, the steering vectors are aggregated and substitution into eq. (5.8) yields the coherent network weights.

The cross spectral density of the vector of network array outputs at frequency ω is

$$\boldsymbol{\Phi}_{QQ} = E\{\boldsymbol{r}_q \boldsymbol{r}_q{}^H\} \tag{5.14}$$

The block diagonal elements of $\boldsymbol{\Phi}_{QQ}$ are therefore $\boldsymbol{\Phi}_{RR}$**;** the off-block diagonal terms are CSD functions for the microphone outputs on distinct arrays. The CSD of any two microphones m_1 and m_2 is described by the corresponding autocorrelation functions and spectral coherence function,

$$\gamma^2_{m_1,m_2}(\omega) = \frac{\left|\boldsymbol{\Phi}_{QQ,m_1,m_2}\right|^2}{\boldsymbol{\Phi}_{QQ,m_1,m_1}\boldsymbol{\Phi}_{QQ,m_2,m_2}} \tag{5.15}$$

The coherence measures the extent to which one signal can be predicted from another using a linear relationship and is within the range $0 \leq \gamma^2_{m_1,m_2} \leq 1$. A coherence of $\gamma^2_{m_1,m_2} < 1$ implies some lack of correlation between microphones such as that due to sensor noise. Network time synchronization errors will not affect intra-array coherence but will adversely affect the coherence between microphones on distinct arrays. This has no effect on incoherent array processing but will degrade the coherent array beamformer performance.

Once the coherent network CSD signal matrix and weights have been acquired, matched field processing is performed identically to the single array case

$$b_{q,j}(\omega) = \boldsymbol{w}_{q,j}{}^H \boldsymbol{\Phi}_{QQ} \boldsymbol{w}_{q,j} \tag{5.16}$$

The output of the beamformer is a coherent processing of the multiarray system; because the array signals are combined in a complex sense, the relative magnitude and phase information is retained. Alternatively, incoherent processing discards this information and the criterion for any given optimal weights is not met. Assuming perfect coherence between all sensors in the network, the optimal coherent method will outperform the incoherent method. However, the incoherent processor is robust to a lack of inter-array coherence. This will be explored in more detail in the following chapter.

5.1.4 Broadband Processing

If the source is broadband such that there is energy distributed across some frequency bandwidth, the beamformer output can be combined across F frequencies. The use of increased amounts of data stabilizes the estimation process, and has been found to be effective to reduce ambiguous beamformer sidelobes [79].

There are a number of different ways to combine the beamformer outputs across frequency, these can be classified broadly into two classes: conventional incoherent methods and coherent broadband methods [80]. Depending on the propagating environment, the beamformer outputs may contain useful cross frequency information if spatially correlated noise is present. However, as the focus of this work is on multiarray systems, the conventional incoherent broadband approach is used. Broadband incoherent processing ignores any correlation across frequency and is optimal only in the unknown flat source spectrum case:

$$b_{BB,j} = \frac{1}{F}\sum_{f=1}^{F} \frac{b_j(\omega_f)}{\|\boldsymbol{b}(\omega_f)\|} \tag{5.17}$$

where the matched-field power response function $b_j(\omega_f)$ can be the result of incoherent or coherent array processing. Frequency averaging is rooted in the idea of constructive mainlobe interference and destructive side lobe interference as sidelobe position varies frequency to frequency. The normalization factor in the denominator ensures that sidelobes are in the same numerical range such that they will average out.

5.2 Free-field beamforming

Although the single-wave assumption no longer holds in a reverberant medium, exactly how this model responds and the limit of its usefulness is of interest. In lightly reverberant environments, a single wave may dominate the response of some source-

receiver geometries. However, it remains that this wave may originate from an image source (due to a reflection) or has been distorted via diffraction.

In a medium that is infinite, homogeneous, and isotropic, a simple source will produce an outgoing spherical wave

$$p(\omega, \Delta \boldsymbol{l}_{jm}) = \frac{A}{\Delta \boldsymbol{l}_{jm}} e^{-i\frac{\omega}{c}\Delta \boldsymbol{l}_{jm}} \tag{5.18}$$

where $\Delta \boldsymbol{l}_{jm}$ is the distance from the source to the receiver and A is a constant determined by the strength of the source, assumed unity. Note that the phase and amplitude of the pressure is dependent only on the range and frequency of the source. Assuming this model, the steering vector for the j^{th} source is given by

$$\boldsymbol{d}_j = \left[p(\omega, \Delta \boldsymbol{l}_{j,m=1}), p(\omega, \Delta \boldsymbol{l}_{j,m=2}), \ldots, p(\omega, \Delta \boldsymbol{l}_{jM})\right]^T \tag{5.19}$$

which accounts for the relative delays between sensors and attenuation due to spherical spreading. These steering vectors will be referred to as "free-field" as opposed to "plane-wave". Because the time delays are individually measured for each receiver, the wave does not necessarily have to be planar with respect to the array (nearfield versus farfield assumption).

5.3 Finite Difference Time Domain Modeling

Generally, source localization is very challenging in scattering environments due to multipath propagation. However, the MFP performance actually increases over the free field case as complex diffraction and scattering lends spatial distinction to the acoustic field, known as hyper resolution [81]. Aubry et al. [82] has shown the complexity of a scattering medium to increase the global number of degrees of freedom of the propagation matrix. In such environments, closed form expressions for acoustic

pressure between arbitrary source and receiver locations do not exist. In this section, FDTD modeling is introduced as a method for obtaining the propagation matrix. A relatively simple example of a reverberant field with 1 scatterer is presented, a situation that is still the interest of current research [83]. Visual results will illustrate some elements of multiarray MFP.

In practice the propagation matrix, $\boldsymbol{D}$, can be experimentally measured or acquired through numerical methods, the latter of which is utilized here. A FDTD model [84] is employed to calculate the acoustic pressure at all points in the field, including the sensor locations. The FDTD method solves a first-order approximation to the wave equation stepwise at discrete points in time and space. It is possible to create very high fidelity models which can simulate the acoustic field in the presence of complicated phenomena such as scattering obstacles, transient sources, nonlinear effects, and a moving inhomogeneous medium [85] .

5.3.1 Formulation

Finite difference methods approximate a partial differential equation and solve the result explicitly. The following formulation follows from Kinsler [86] and Tolan [87]. Assuming the ideal case of a constant sound velocity, relatively small wave amplitude, and a lossless medium, sound propagation obeys the linear wave equation:

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{5.20}$$

where p is the acoustic pressure. The equation of state for a perfect gas can be used to relate the linear Euler's equation to the linear continuity equation to yield the governing differential equations, here in 2 dimensions:

$$\frac{\partial p}{\partial t} = -\rho c^2 \nabla \cdot \boldsymbol{v} = -\rho c^2 \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right) \tag{5.21}$$

$$\frac{\partial v}{\partial t} = -\frac{1}{\rho}\nabla p \rightarrow \begin{cases} \frac{\partial v_x}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial x} \\ \frac{\partial v_y}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial y} \end{cases}$$

where $\boldsymbol{v} = v_x + v_y$ is the velocity vector. The density, ρ, and the speed of sound, c, may both be inhomogeneous. A substantial reduction in computational requirements is achieved by using a two dimensional model as an approximation of the three dimensional world. This involves a number of approximations and limitations to the modeling capability [88], for example, buildings are essentially infinitely tall and vertical propagation is ignored. However, the kinematics of the computations remains accurate and the 2D model is sufficient to describe all phenomena of interest for the purpose of this work.

The basis of the FDTD method is the replacement of derivatives with finite differences. Figure 5.2 shows a Cartesian grid arrangement of nodes that can be used for this replacement. The widely accepted staggered difference algorithm [89], or leapfrog method, is used in which velocity and pressure calculations are staggered in both time and space. This approach is robust, fast, and has several properties that give it an advantage over other numerical finite-difference methods [88]. The velocity and pressure nodes are known only at fixed locations in space and only at discrete times. The spatial location of the nodes is indicated by the indices (X, Y) such that the physical location is $\boldsymbol{l}_{FD} = (X\Delta x, Y\Delta y)$ where Δx and Δy are the step sizes in the x and y directions respectively. Furthermore, there is an additional offset of the nodes as indicated in Figure 5.2 so that each of the velocity components is offset one half spatial step from the pressure node with the same indices. These offsets are not explicitly included in the indices as they are dropped when implementing the algorithm in a computer. Instead, the offsets are implied by the node component itself to facilitate notational convenience and transformation into code.

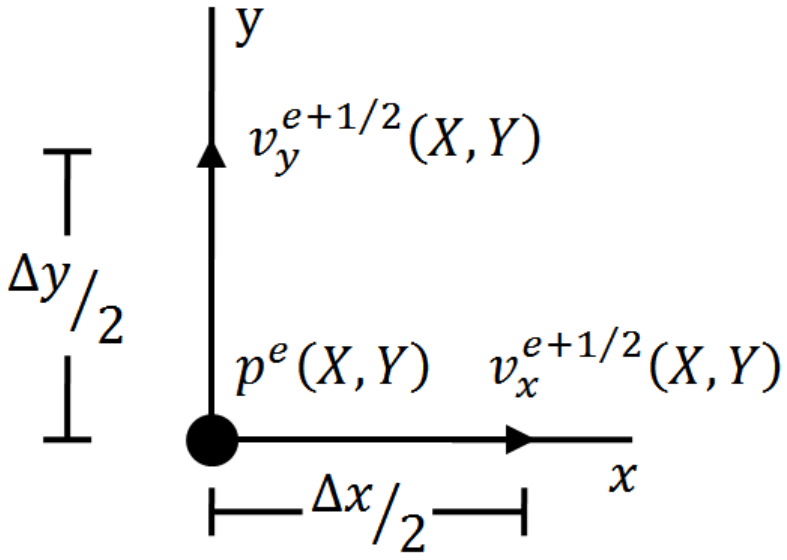

Figure 5.2: Cartesian grid arrangement for the leapfrog method

Velocity nodes are offset from pressure nodes by a half temporal step, but all velocity components are evaluated at the same instant of time, i.e., pressure nodes are defined to exist at integer multiples of the temporal step Δt_{FD} while the velocity nodes v_x and v_y exist at times half-way between the times at which the pressures are defined. The temporal index is e. Offset in temporal locations of the fields is explicitly indicated.

This arrangement of offset nodes allows the calculation of pressures from velocities at a previous time, and the calculation of velocities from these updated pressures. This alternating calculation is repeated for each successive time step as the simulation is run from an initial time to a final time. By replacing the derivatives with finite differences according to the node arrangement shown in Figure 5.2, the following update equations can be obtained from the governing differential equation (5.23):

$$p^{e+1}(X,Y) = p^e(X,Y) - \Delta t_{FD}\rho c^2\left(\frac{v_x^{e+1/2}(X,Y) - v_x^{e+1/2}(X-1,Yj)}{\Delta x} + \frac{v_y^{e+1/2}(X,Y) - v_y^{e+1/2}(X,Y-1)}{\Delta y}\right) \quad (5.22)$$

$$v_x^{k+1/2}(i,j) = v_x^{k-1/2}(i,j) - \frac{\Delta t}{\rho}\left(\frac{p^k(i+1,j) - p^k(i,j)}{\Delta x}\right)$$

$$v_y^{k+1/2}(i,j) = v_y^{k-1/2}(i,j) - \frac{\Delta t}{\rho}\left(\frac{p^k(i,j+1) - p^k(i,j)}{\Delta y}\right) \quad (5.23)$$

where the superscripts represent the temporal index, and the arguments the spatial indices. Eq. (5.22) is concerned with conservation of mass, i.e. the pressure increases when there is net flow into a node. Eq. (5.23) describes the acceleration of the fluid due to a pressure gradient. These update equations require that material properties change only at discrete locations – properties can change on the scale of the spatial step sizes.

5.4 Numerical Example: modeling and MFP in a scattering environment

5.4.1 FDTD Model

Consider the following scenario to demonstrate FDTD modeling and MFP. Figure 5.3 shows the geometrical relationship between two four-microphone arrays and a rectangular shaped scattering object over a 160×110 m^2 area.

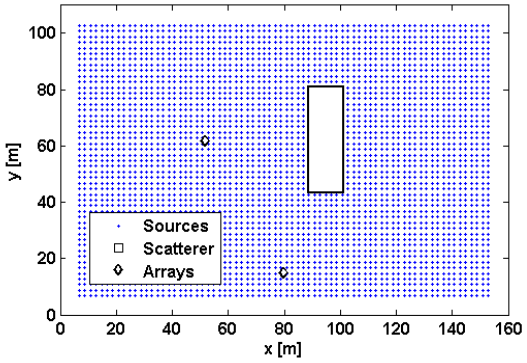

Figure 5.3: Geometry of the FDTD model showing the relative position of the arrays and scattering object.

The scattering object is treated as rigid (infinite impedance). Microphones are arranged in each array with radius 0.2 meters and centers at $\boldsymbol{l}_{n_1} = \{52, 62\}$ and $\boldsymbol{l}_{n_2} = \{80, 15\}$. The FDTD grid cells are $\{\Delta x, \Delta y\} = \{0.2, 0.2\}$ meters yielding 440,000 elements overall. This allows for roughly 17 grid points per primary wavelength. 3626 source locations are specified at 2 meter spacing in a rectangular grid. The model uses gradual absorption at the boundaries to simulate an infinite space satisfying the Sommerfield radiation condition.

The source signal is a Ricker wavelet [90] with a center frequency of $f_c = 100Hz$. The mathematical formula for a Ricker wavelet can be uniquely specified with the center frequency:

$$s(t) = (1 - 2\pi^2 f_c^2 t^2) e^{-\pi^2 f_c^2 t^2} \tag{5.24}$$

This a zero-phase wavelet commonly used in seismology to represent impulsive sources, identical to the second derivative of the Gaussian PDF. The zero-phase property allows the arrival of the source energy at a receiver to be associated with a specific time. The energy of this waveform decays rapidly in both the time and frequency domains, reducing computational effort and numerical distortion (see Figure 5.4).

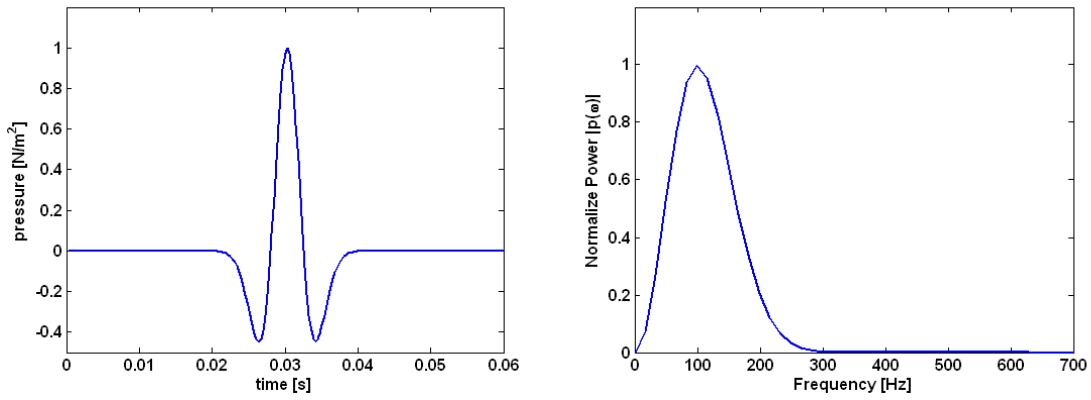

Figure 5.4: The source signal (Ricker wavelet) in the time (left) and frequency domain (right).

To ensure numerical stability of the FDTD calculation, the time step is chosen to satisfy the theoretical courant condition [91],

$$\Delta t \leq \frac{1}{c\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}}} \tag{5.25}$$

The time step is chosen as 0.3 ms and the simulation is run for 2500 time steps or 750 ms to allow the pulse excitation to sufficiently propagate through the environment. The model parameters are summarized in Table 5.1.

Area	160 × 110 m^2
Grid Cell	0.2 × 0.2 m^2
Elements	440,000; 17 per primary λ
Excitation	Ricker Wavelet, f_c = 100Hz
Time step	0.3 ms
Time run	750 ms
Source Locations	3626
Location Resolution	2 × 2 m^2
Array radius	0.2 m

Table 5.1: Parameters for the finite difference time domain model.

To build the steering vector matrix, $\boldsymbol{D}$, the FDTD simulation is run once for each source location. The pressure history at the sensor locations is stored and a fast Fourier transform is performed on each impulse response.

5.4.2 Impulse response

To demonstrate the acquisition of impulse responses at the sensor locations, consider a source at $\boldsymbol{l}_s = \{125, 73\}$. Snapshots of the propagation at 60ms intervals appear in Figure 5.5. The complex effects of a single scattering object on the propagating wavefront are made clear by these images.

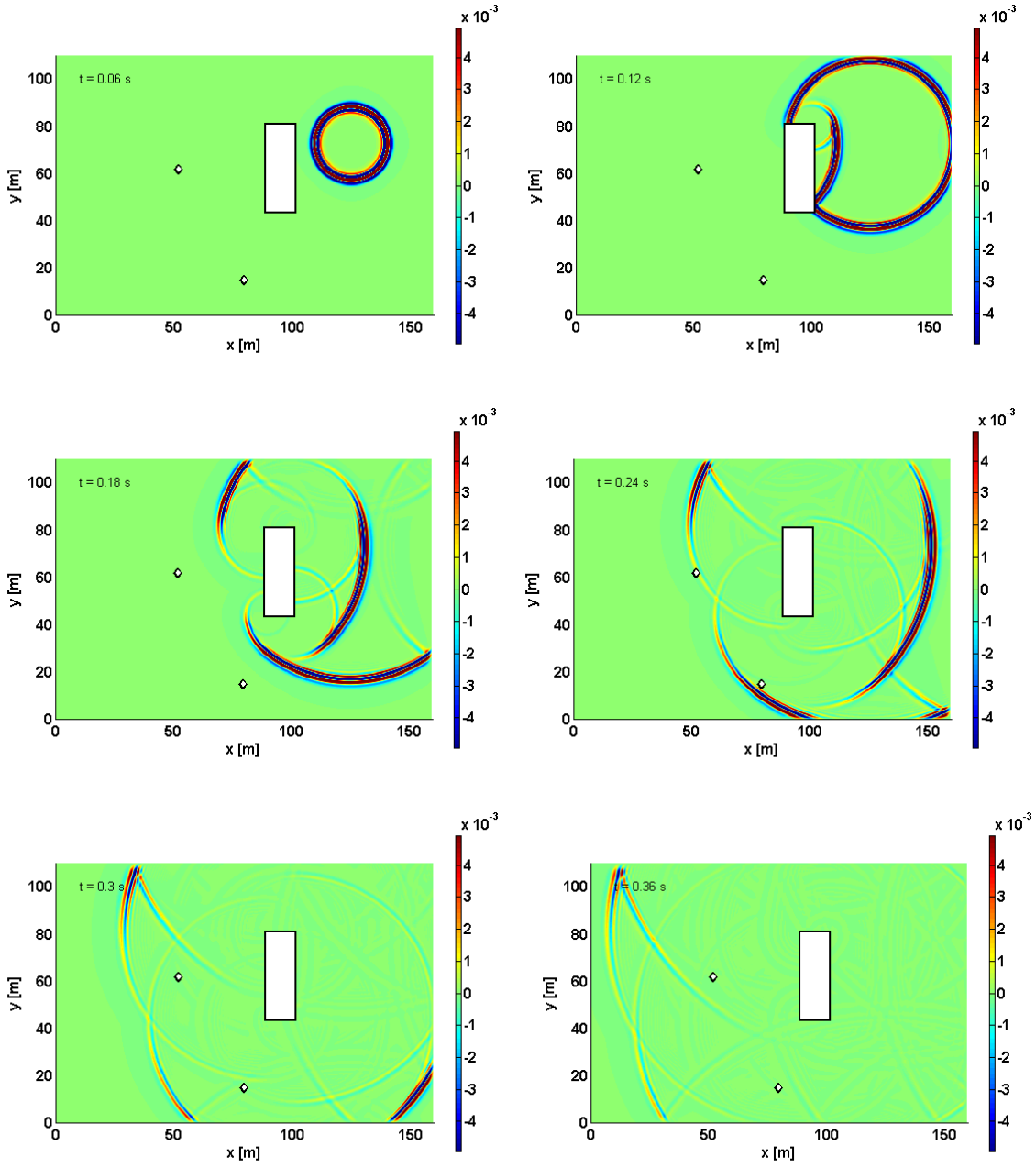

Figure 5.5: Snapshots of the 2D acoustic pressure field as calculated by the FDTD method.

The recorded impulse response at one sensor on each of the two arrays appears in Figure 5.6. All the sensors on an array receive delayed versions of the same waveform. Due to shadowing, the amplitude of the waveform received at node 1 is much lower than that received at node 2, which had a line of sight to the source location. In fact, the

source-receiver location relationship for node 2 resulted in a nearly free field-like response. Node 1 received multiple reflected waves and a much more distinctive response overall.

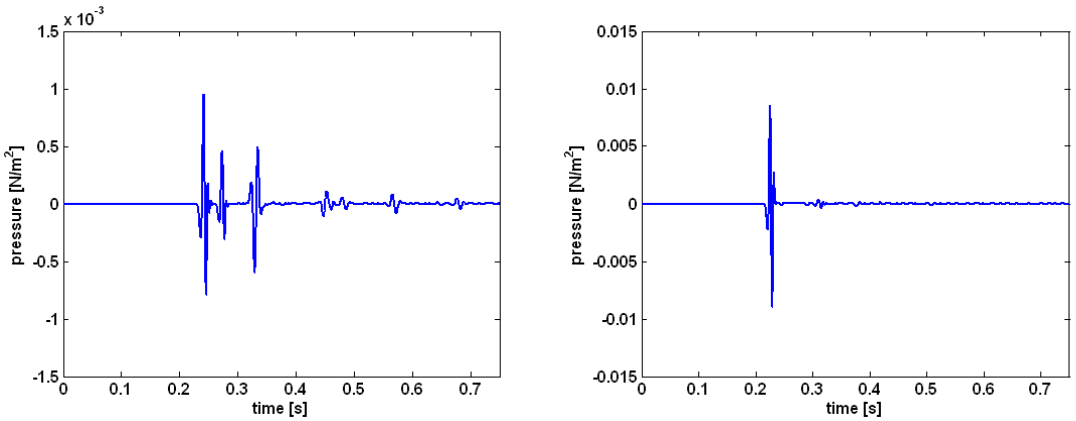

Figure 5.6: Impulse response recorded at one sensor from node 1 (left) and node 2 (right). $\boldsymbol{l}_{n_1} = \{52, 62\}$, $\boldsymbol{l}_{n_2} = \{80, 15\}$ and $\boldsymbol{l}_s = \{125, 73\}$.

5.4.3 Matched Field Processing

The transmission loss $\boldsymbol{\psi}$, of both arrays and the coherent array network has been reshaped into a two-dimensional matrix and laid onto the proper coordinates in relation to the scattering object (Figure 5.7). In addition to spherical spreading, the sensitivity is heavily influenced by reflection, diffraction, and shadowing introduced by the scattering object. The transmission loss of the coherent network is distinct from either array, although it shares features of both. Contrary to an average, $\boldsymbol{\psi}_q$ takes the better of the two, for example a location has low sensitivity only if both individual arrays have low sensitivity.

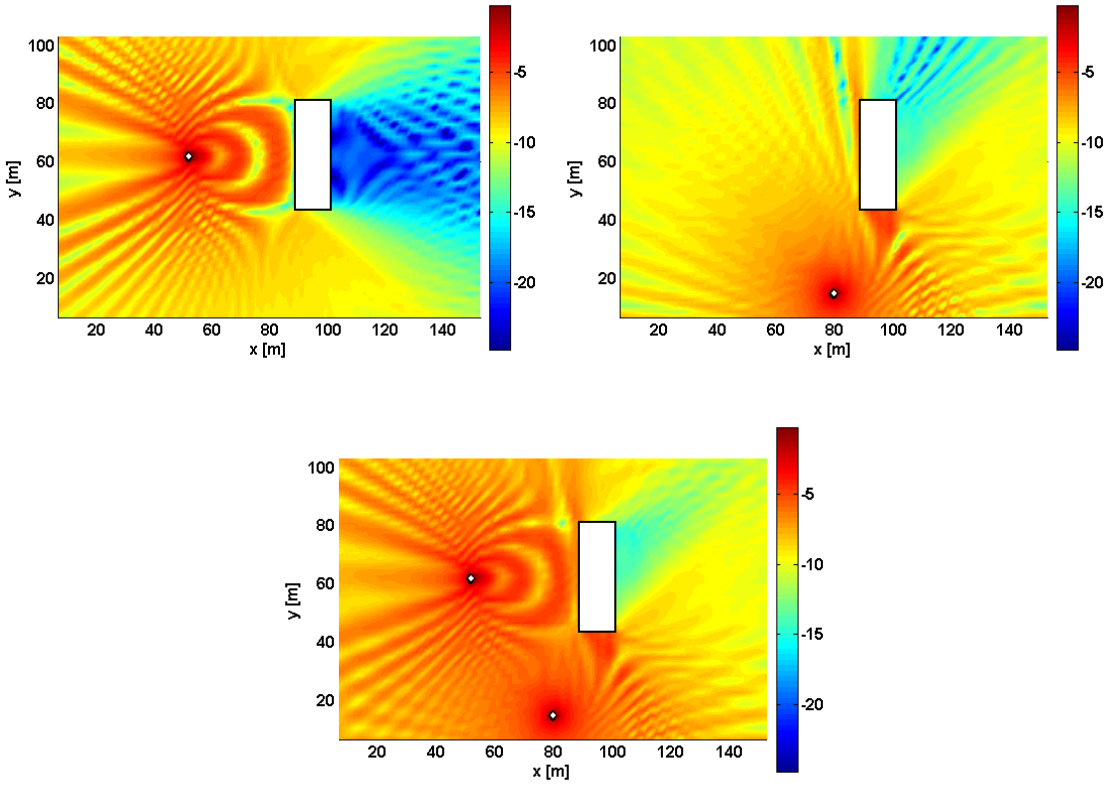

Figure 5.7: Transmission loss (dB) at 100Hz for array 1 (upper left), array 2 (upper right), and the coherent array network (bottom).

In Figure 5.8, the beamformer outputs for the source location $\boldsymbol{l}_s = \{125, 73\}$ are presented. In all cases, the maximum of the matched-field power response function is at $\boldsymbol{l}_s$, however the uniqueness of the solution is marginalized more in some cases than others. At this frequency, reflections from the obstacle create a fine interference pattern at the scale of a wavelength. Regions in space that share a similar amplitude and phase response to the true source location are essentially sidelobes as in directional beamforming. However, in MFP the characteristics of the environment have a far greater influence on beamformer sidelobes than microphone spacing. The upper half of Figure 5.8 shows the beamformer output for arrays 1 and 2 at 100Hz. For the current source location of interest, the obstacle has a much greater effect on array 1, and there is some confusion as to whether the non-LOS source is in front or behind the obstacle. The scattering object lends less distinction to the response of array 2, which is similar to the free field case where direction is resolved and an actual beam is formed but target range

is ambiguous. Multiarray approaches appear in the lower half of Figure 5.8. Again, the incoherent approach is an average so while array 2's contribution eliminates some of the ambiguity of array 1, it also adds some of its own. Coherent processing results in sidelobes that are significantly reduced.

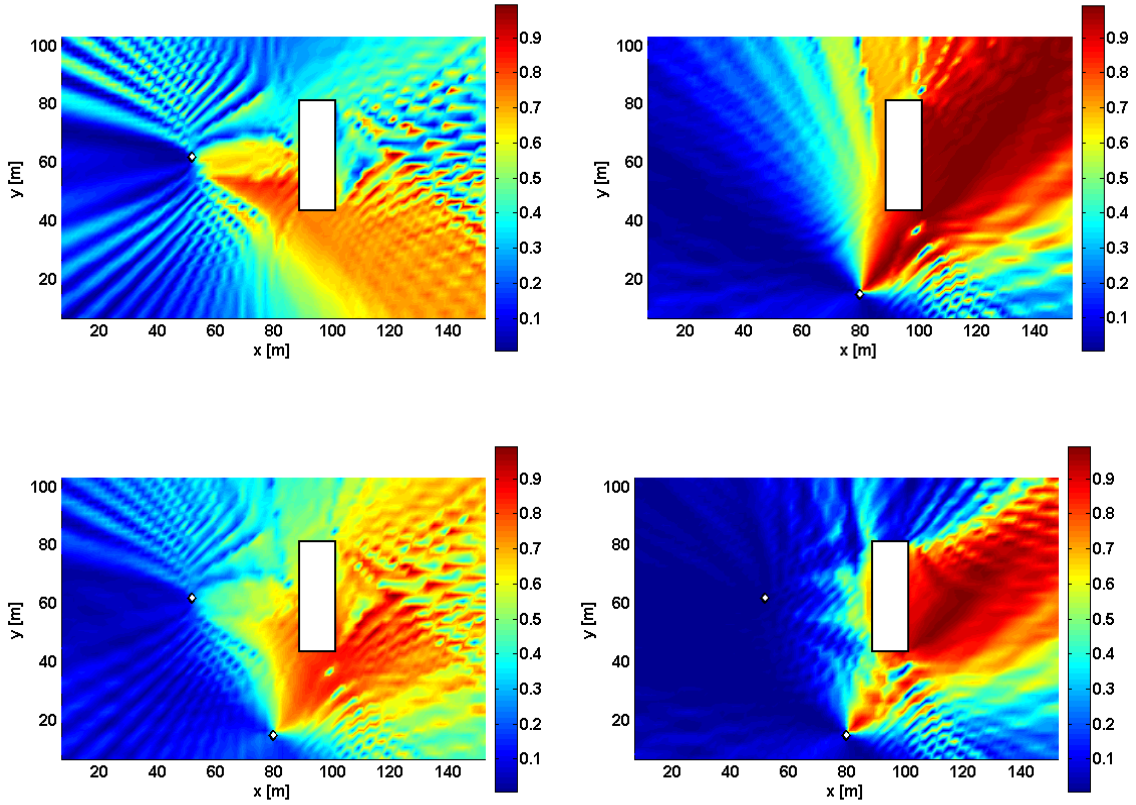

Figure 5.8: Matched-field power response function at 100 Hz for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{l}_s = \{125, 73\}$.

The result of incoherent broadband processing, eq. (5.17), is shown in Figure 5.9. 25 discrete frequencies have been averaged in the band 90-120 Hz. As expected, averaging the beampatterns across frequency reinforces the mainlobe, smoothes the output in the region near the source, and eliminates spurious sidelobes.

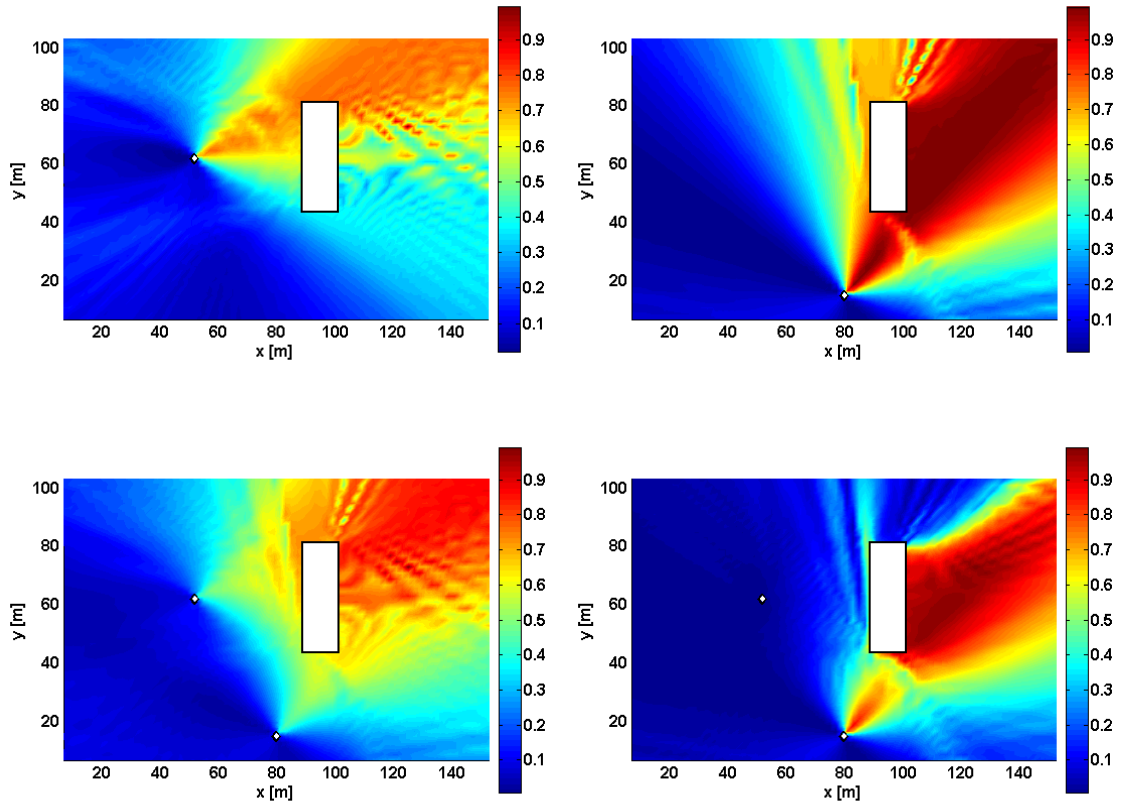

Figure 5.9: Broadband matched-field power response function for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{l}_s = \{125, 73\}$.

In Figure 5.10, the beamformer outputs using the free-field model are presented. Contrary to the a priori model results, the maximum of the matched-field power response function using the free-field model is no longer at $\boldsymbol{l}_s$. Array 1 receives the first and most powerful source signal due to diffraction around the object and therefore beams in that direction (towards the edge of the object). Array 2 has line of sight to the source and its beampattern, although ambiguous in range, has the correct direction. The incoherent averaging of the two nodes, essentially triangulation, yields a position estimate almost 40 meters off in this example.

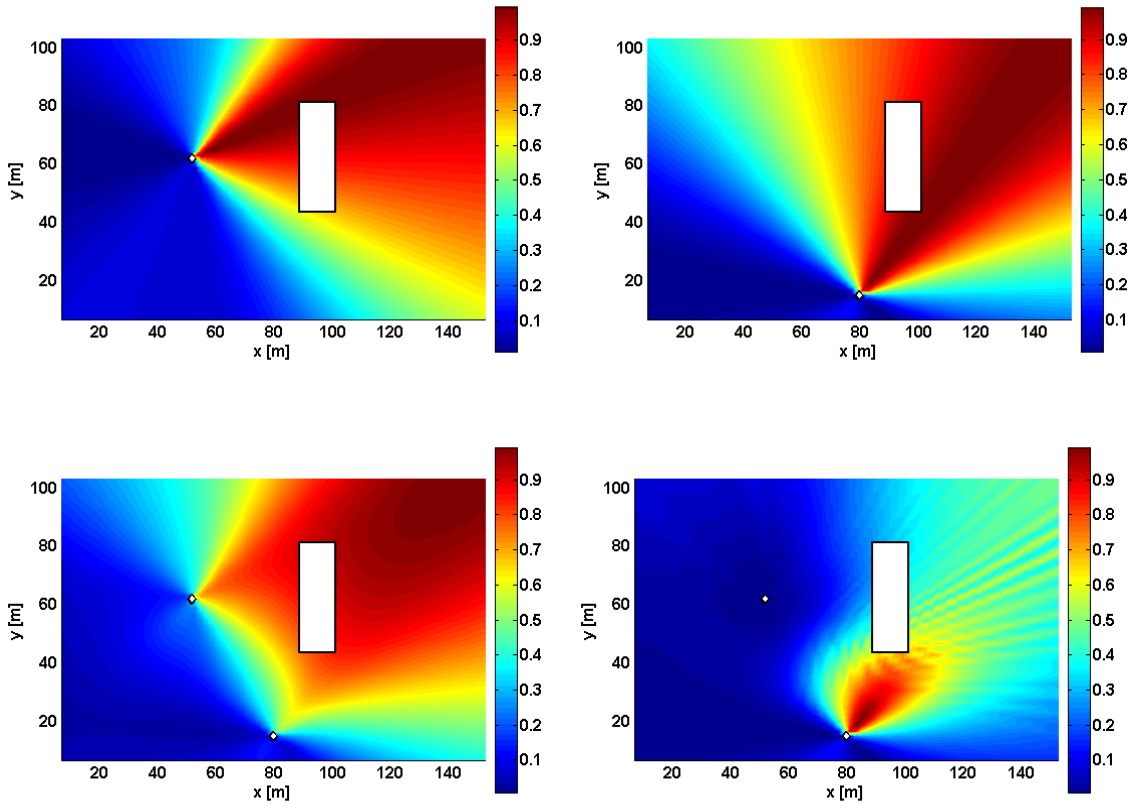

Figure 5.10: Matched-field power response function using the plane-wave model at 100 Hz for array 1 (upper left), array 2 (upper right), the incoherent array network (lower left) and the coherent array network (lower right). The source is located at $\boldsymbol{l}_s = \{125, 73\}$.

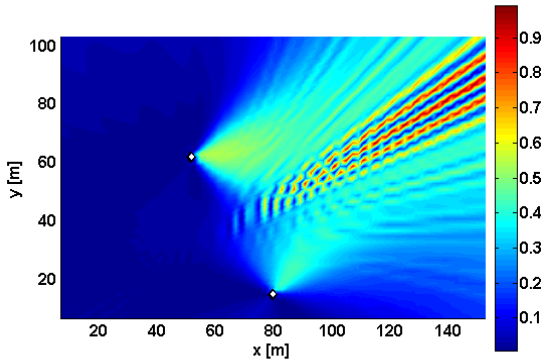

Figure 5.11: Matched-field power response function of the coherent array network at 100 Hz in free-field propagation conditions using the free-field model. The source is located at $\boldsymbol{l}_s = \{125, 73\}$.

The coherent estimator in Figure 5.10 is confused by the received signals. For comparison, the coherent beampattern that would have resulted had the scatterer not been present is shown in Figure 5.11. Although some qualities of the free-field coherent beampattern remain in the reverberant case (Figure 5.10), the response is dominated by a concentration near array 2. This is due to the unexpected shadowing at array 1 that diminishes its contribution to the networked array.

In this scenario, the shadowing at node 1 disrupted the coherent processor. This can be circumvented by removing the amplitude information from the weights, i.e., let the steering vectors be governed by the phase only:

$$\boldsymbol{d}_j = \left[e^{-i\frac{\omega}{c}\Delta \boldsymbol{l}_{jm_1}}, e^{-i\frac{\omega}{c}\Delta \boldsymbol{l}_{jm_2}}, \ldots, e^{-i\frac{\omega}{c}\Delta \boldsymbol{l}_{jm_M}}\right]^T \tag{5.26}$$

The resulting free-field phase-only coherent beampattern is given in Figure 5.12. Although in this scenario more accurate results are conferred by neglecting the amplitude information, high amplitude is typically related to line of sight and usually will act to positively weight a node's contribution. In general, neglecting the amplitude information results in poorly performing beamformer and will not be considered further.

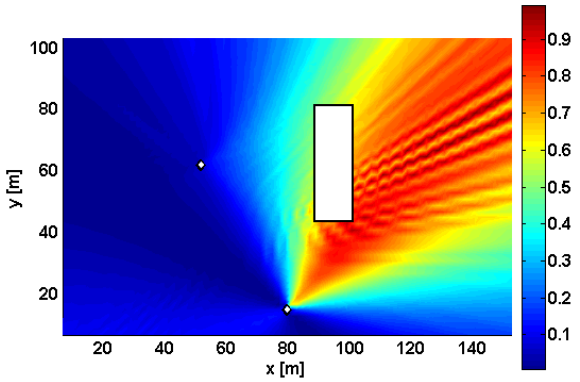

Figure 5.12: Matched-field power response function of the coherent array network at 100 Hz using the phase-only free-field model. The source is located at $\boldsymbol{l}_s = \{125, 73\}$.

In conclusion, it has been shown that MFP using an appropriate multiwave model is required if a beamformer is to operate in a cluttered environment. The simple model of a single freely propagating wave is violated by the presence of a single obstacle; a

beamformer designed under this assumption will be confused. Instead, a model of the environment can provide the steering vectors to be used in the design of a beamformer that accounts for complex propagation.

Chapter 6:

Localization in an Urban Environment with Uncertainty

MFP accounts for the unique signature of the reverberant acoustic field at a receiver such that the location of a propagating source can be determined anywhere in the known environment. Naturally, localization performance is correlated with the quality of a priori information. There is always some uncertainty in the precise location and even presence of scatterers, and fluctuation in the atmospheric conditions, temperature, and density of the medium are unavoidable. Mismatch has been studied extensively in the context of ocean acoustics [92], however, urban environments offer unique challenges. Some of the effects of model uncertainty have been studied in the context of time reversal refocusing. Liu et al. [40] found that localization was generally tolerant of minor changes in scatterer position, although the addition of an obstacle close to a receiver caused substantial degradation. On a much smaller scale, Yon et al. [93] reported robustness to perturbations such as moving persons within a room. In a second experiment, small variations in temperature were shown to increase the sidelobe level during focusing.

In this chapter, some aspects of practical implementation are studied in a complex urban environment. Previous work has demonstrated the viability of matched field processing. Here, an understanding of how and to what extent uncertainty affects localization is sought. Beamformer performance in the presence of uncertainty such as noise and model perturbations is examined. Fusion of information from multiple spatially distributed acoustic arrays is proposed to mitigate the effects of uncertain environmental parameters. The tradeoffs and ultimate performance of coherent and

incoherent mulitarray processing techniques are investigated. The influence of the quantity and position of available nodes is also considered.

6.1 Test Environment

A more complicated FDTD model has been constructed for the following numerical experiments. The scatterer locations are based on an artificial training village of 15 closely spaced concrete buildings in a flat area, similar to that used for previous time reversal experiments [39]. The artificial village, or MOUT (military operations on urban terrain) site exists at the US Army installation Fort Benning in Georgia. An overhead image of the site and close-up images of individual buildings appears in

Figure 6.1.

Figure 6.1: Overhead satellite image of the McKenna MOUT site at Fort Benning and individual building details (below) [94].

A general consistency between the sound pressure level given by a 2-dimensional FDTD model and measured microphone data at multiple locations has been found in previous work by Liu and Albert [95]. Explosive charges were detonated to produce acoustic pulses and the resulting signals were recorded at multiple receivers distributed throughout the site. Sound pressure levels were also predicted by a 2-dimensional FDTD model using uniform physical parameters and simple geometries. The measured sound pressure levels were in agreement with the predicted sound pressure levels at 13 sensor locations, while one location showed a discrepancy. High (roughly 1 order of magnitude) observed sound pressure levels suggest that a significant amount of energy was transmitted through a nearby building.

Given the results of the experimental validation by Liu et. al., a second two-dimensional FDTD model has been constructed using rigid scatterers. The FDTD model parameters for the test environment are identical to the single scatterer model in the previous section. There are five, 4-microphone arrays distributed throughout the source domain (Figure 6.2).

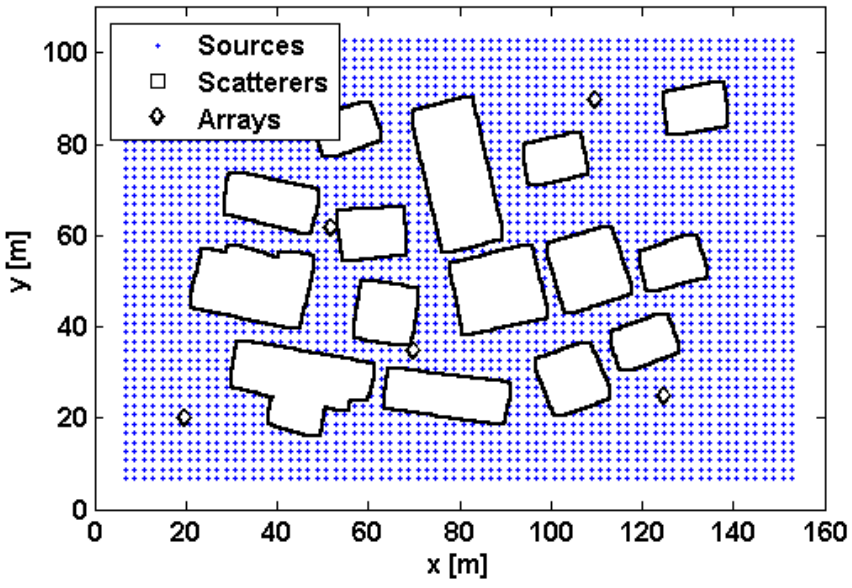

Figure 6.2: 2D FDTD model of artificial training village used for the study.

The transmission loss, eq. (5.3), of node 1 is plotted in Figure 6.3. Multiple paths through the obstacles are apparent, due to a multiple reflections and diffraction around

building corners. Still, the strongest path for acoustic energy is around the cluster of buildings and shadow zones form behind some buildings.

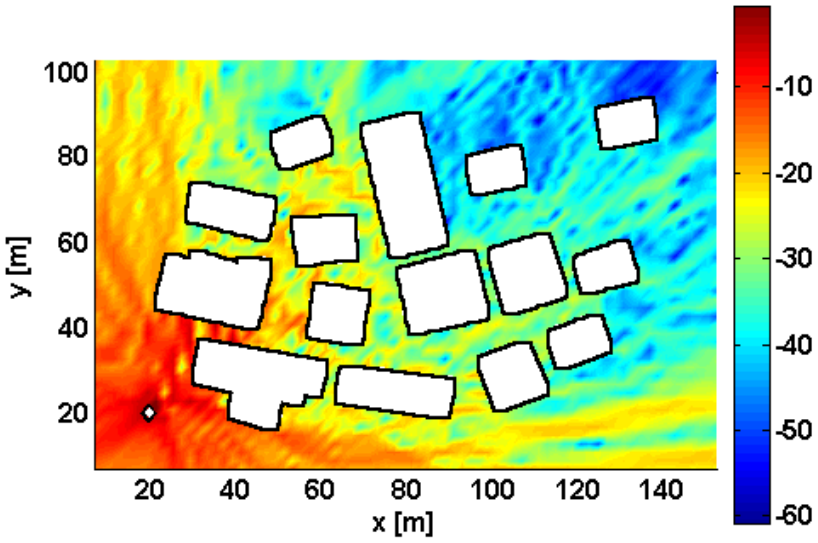

Figure 6.3: Transmission loss (dB) of node 1 at 100Hz. Primary paths for propagation through the obstacles is indicated by high transmission loss.

Considering two single arrays (node 1 and node 2) and two multiarray processing strategies (coherent and incoherent processing utilizing all 5 nodes), the broadband beamformer matched field power response functions corresponding to a source at $\boldsymbol{l}_s = \{65, 73\}$ have been calculated and are presented in Figure 6.4. These are also known as ambiguity functions because they show the relative uniqueness between the beamformer output at each location. 25 discrete frequencies have been averaged in the band 90-120 Hz. The frequency averaging has reduced sidelobes such that the source location is very distinct. Node 2 receives more direct energy and is in closer proximity than node 1 which receives propagation through many buildings. The complex geometry creates a distinct acoustic field, however, in a large domain the distinction is not necessarily unique at every point as evident by sidelobes. Regardless, the hyper resolution mentioned in section 5.3 is readily apparent compared to the single scatterer case. While both the incoherent and coherent network clearly resolves the source, sidelobes are more efficiently reduced by the coherent processing.

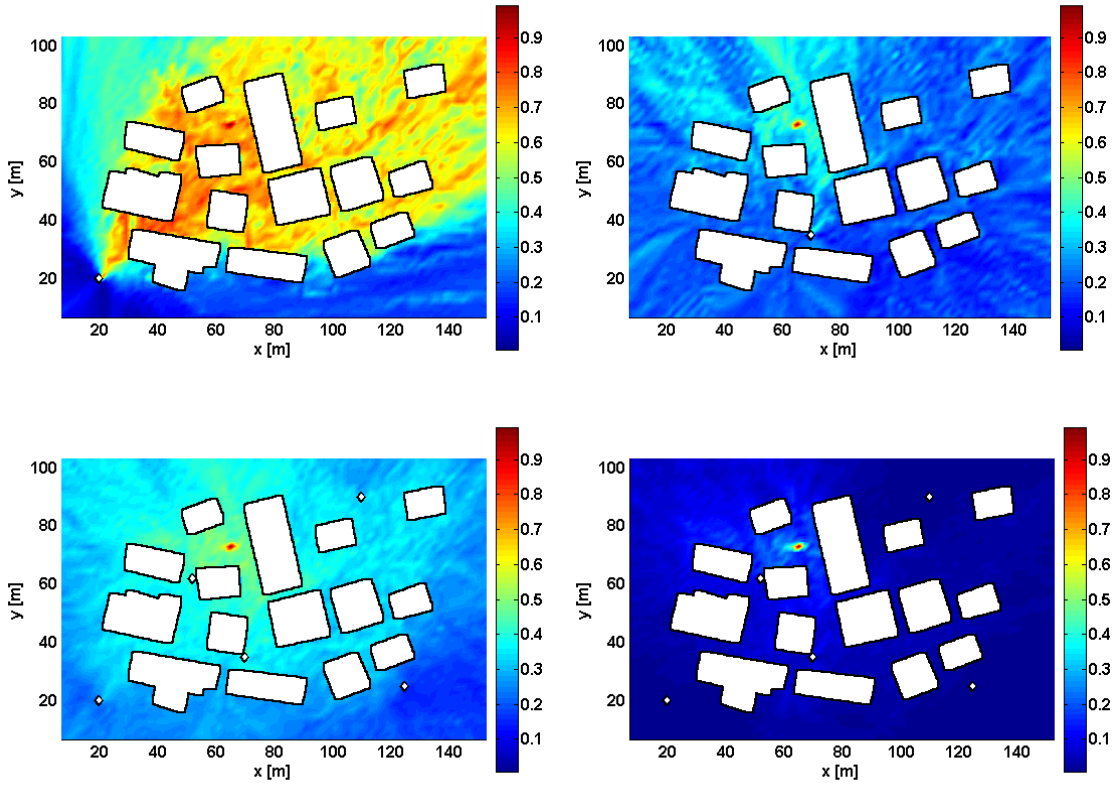

Figure 6.4: Broadband matched-field power response function for array 1 (top left), array 2 (top right), the incoherent array network of 5 arrays (bottom left) and the coherent array network of 5 arrays (bottom right). The source is located at $\boldsymbol{l}_s = \{65, 73\}$.

6.2 Forms of Uncertainty

The ability of the true source location to dominate the magnitude of the beamformer output will degrade in the presence of uncertainty. This uncertainty can manifest itself in a variety of ways, such as thermal noise originating from the sensors and associated hardware, noise from wind, and acoustic interferers. It is also likely that the numerical model or experimental situation used to obtain the steering vector matrix differs from the situation in which signals are measured, such as discrepancies in the geometry and impedance characteristics of buildings. Fine features at a small scale are computationally expensive to model and unlikely to be known at high levels of detail. Regardless, even an experimentally characterized array will suffer from the ever

changing atmospheric conditions, temperature, and density of the propagating medium. Numerical experiments offer fine control over individual parameters. This section outlines some forms of uncertainty and how they are simulated.

6.2.1 Uncertainty due to sensor noise

It can be assumed that the sensor noise manifests itself as normally distributed zero–mean white Gaussian noise, incoherent between sensors. Referring to Eq. (5.1),

$$v_m(k) \sim N\left(0, \sigma_g^2\right) \tag{6.1}$$

The noise level is set based on the desired signal to noise ratio (SNR):

$$SNR_{dB} = 10 \log_{10} \frac{\sum_{i=1}^{I} \sigma_s^2(\omega_i) \boldsymbol{d}_n(\omega_i) \boldsymbol{d}_n^H(\omega_i)}{\sum_{i=1}^{I} \sigma_g^2(\omega_i)} \tag{6.2}$$

where the noise power is calculated according to the following definition:

$$\sigma_g^2(\omega_i) = E\{\boldsymbol{v}(\omega_i) \boldsymbol{v}(\omega_i)^*\} \tag{6.3}$$

Additive noise can also been interpreted as a drop in intra array coherence and the SNR is proportional to the off-diagonal terms of $\boldsymbol{\Phi}_{RR}$. Like any superposition of independent signals, the beampattern resulting from a signal in additive noise is identical to the sum of the individual signal and noise beampatterns. Therefore, consider the noise only contribution to $\boldsymbol{\Phi}_{RR}$ as given by eq. (5.5). Assuming many averages are available to form the covariance matrix, $\boldsymbol{\Phi}_{RR} = I$. The shape of the noise-only beampattern is then given by:

$$b(\omega) = \boldsymbol{w}^H \boldsymbol{w} = \frac{\boldsymbol{d}^H \left(\boldsymbol{\Phi}_{VV}{}^{-1}\right)^H \boldsymbol{\Phi}_{VV}{}^{-1} \boldsymbol{d}}{\boldsymbol{d}^H \boldsymbol{\Phi}_{VV}{}^{-1} \boldsymbol{d}} \tag{6.4}$$

If the delay and sum beamformer is used, the response to incoherent noise is unity output at every location. However, the output of the superdirective beamformer is governed by the interaction between the assumed diffuse field and the strength of the steering vectors at each location.

In a free field, the noise-only beampattern takes on the form of mainlobes and notches derived from the geometry of the array relative to the source grid. For example, see Figure 6.5 which is the broadband noise-only beampattern for node 3 using the superdirective free field weights defined by eq. (5.11) and eq. (5.19). The lack of symmetry is due to the discrete nature of the source grid, i.e. the array is slightly closer to sources in the positive y and x directions.

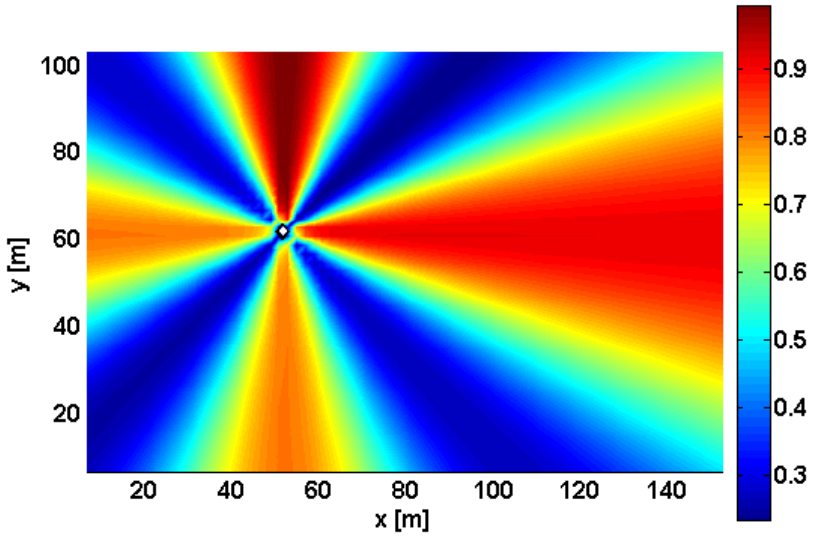

Figure 6.5: Broadband noise-only matched field power response function for node 3 using weights derived from the plane wave model.

In a scattering environment, the noise-only beampattern is affected by the complex nature of the beamformer weights. See Figure 6.6 where the noise only beampattern of nodes 1 and 3 are shown. Essentially half of node 1's field of view is on the boundary, which simulates free field propagation. The resemblance to the free-field pattern is evident. Meanwhile, node 3 is completely surrounded by scatterers and has a complex, random beampattern. While the noise power was identical at both arrays, the

overall resulting beamformer power output is much lower at node 3 as the diverse patterns destructively average out across frequency. In general, arrays with complex manifold vectors are more robust to additive white noise.

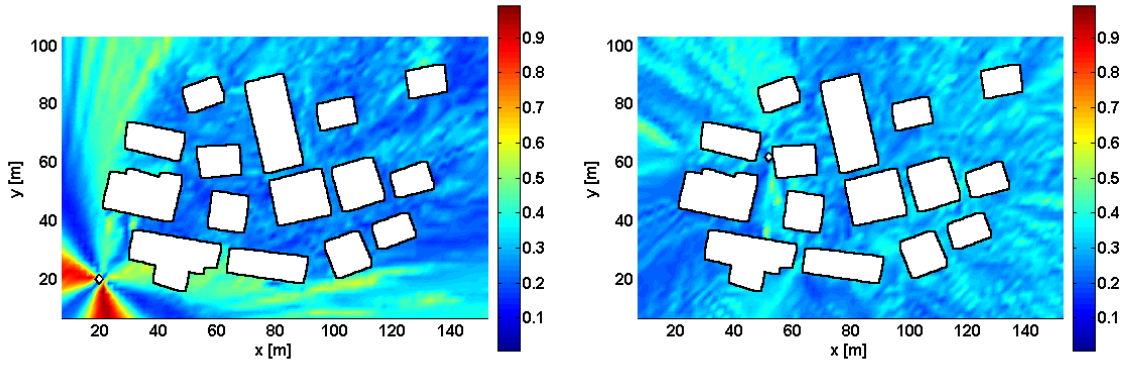

Figure 6.6: Broadband noise-only matched field power response function for node 1 (left) and node 3 (right) using weights derived from the a priori model.

6.2.2 Uncertainty due to model mismatch: calibration error

It is possible that the true location of a microphone array differs from the numerical model. Even precise location measuring devices such as those utilizing GPS can exhibit significant variance and drift. This is known as system mismatch and various forms, such as array tilt, have been studied in ocean acoustics [47].

It is informative to view these uncertainties in terms of the phase of a propagating or reflected wave. For example, at 100Hz an uncertainty of 2m represents 210º of phase. Therefore, near a reflecting boundary with an uncertainty of 2 meters we cannot be certain whether the reflected wave will be in phase or out of phase with the incoming wave at any given position. To investigate this case, a perturbed model was built in which the location of each sensor array is shifted 2 meters in one direction (Figure 6.7). The original and perturbed node locations are listed in Table 6.1. Data from the perturbed model simulates signals actually acquired while the data from the original model is an estimate, $\widehat{\boldsymbol{D}}$, of the true steering vector matrix.

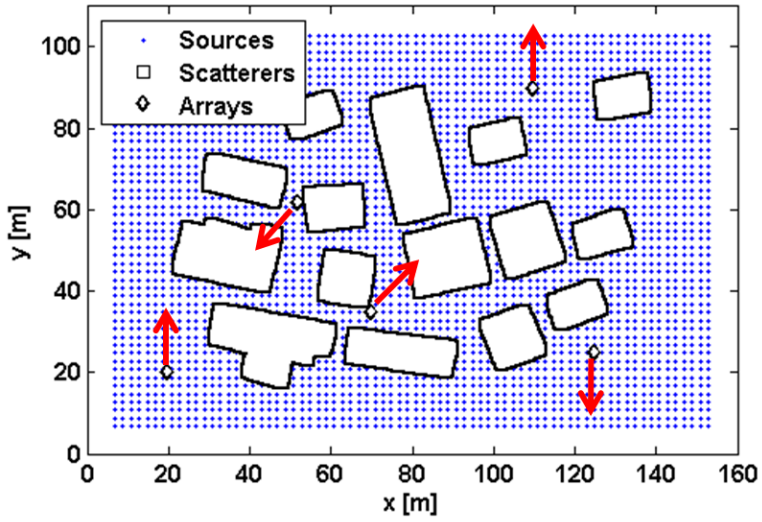

Figure 6.7: Geometry of the simulation. Arrows indicate the direction in which the node location was perturbed for the mismatch study.

Array	Original	Perturbed
1	20,20	20,22
2	70,35	71.4,36,4
3	52, 62	50.6,60.6
4	110,90	110,92
5	125,25	125,23

Table 6.1: Original and perturbed array coordinates in the x,y plane (meters).

Broadband matched field power response functions for a 2 meter uncertainty in array location are presented in Figure 6.8. In the case of node 1, the uncertainty results in increased sidelobes and a shift in the peak value. While node 2 is in close proximity to the source, it suffers from much greater sidelobe increase and main lobe decrease. The effect of model mismatch, while clearly dependent on the array-source geometry and model, is very complex. Both multi-array techniques are able to maintain a peak near the true source location. Note that, while a comparison with Figure 6.4 gives the impression that in general the sidelobes increase, in truth some values increase and others decrease.

However, the mainlobe is almost always reduced and the normalization to the maximum value of the plot brings up the sidelobes.

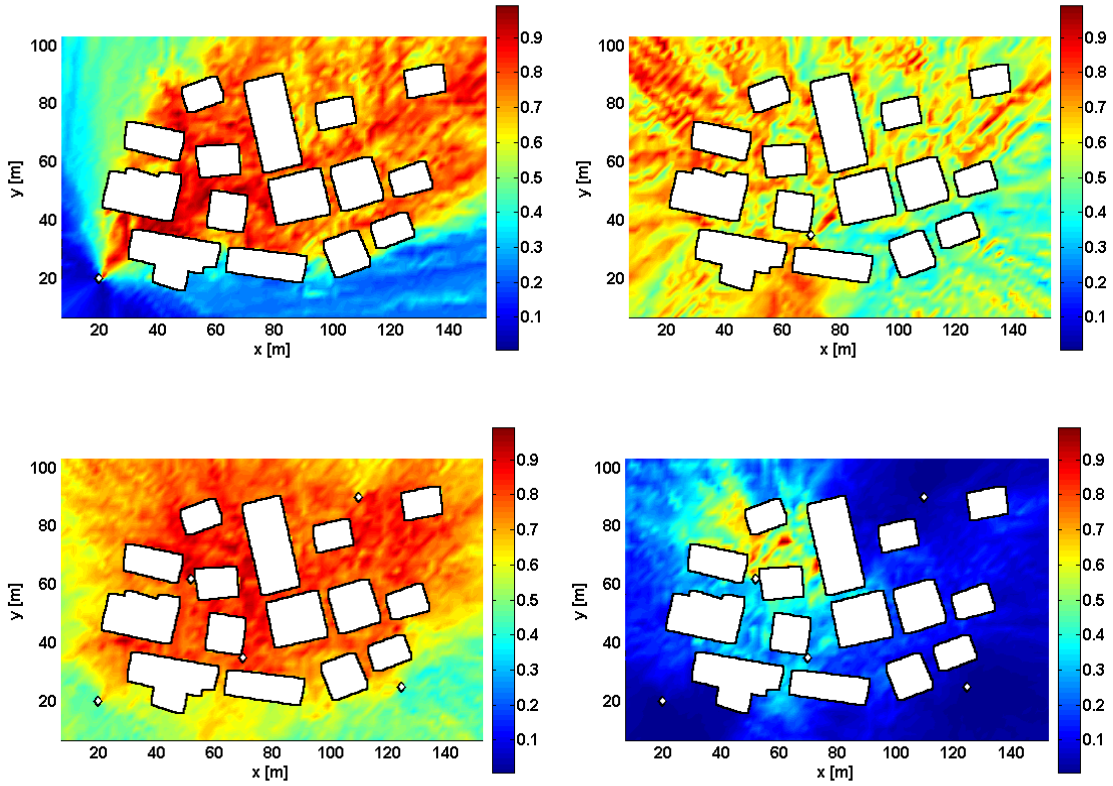

Figure 6.8: Broadband matched-field power response function with a 2m uncertainty in node location for node 1 (top left), node 2 (top right), the incoherent network of 5 arrays (bottom left) and the coherent network of 5 arrays (bottom right). The source is located at $\boldsymbol{l}_s = \{65, 73\}$.

In the case of mismatch, the signal model is no longer described by eq. (*5.2*). Although the received waves do not match any catalogued steering vectors, they are coherent across the array. Uncertainty in model parameters can be interpreted as corrupting the steering vectors with spatially correlated noise. Note that, from a MFP beamformer design standpoint this is a motivation for using the superdirective beamformer over the delay-and-sum beamformer (which optimally rejects incoherent noise).

6.2.3 Uncertainty due to model mismatch: temperature

The uncertainty in temperature of the propagating medium and hence speed of sound has a similar effect as uncertainty in sensor or scatterer location. Consider a 20^0C perturbation in temperature. There is a 3% increase in the speed of sound between 30^0C and 50^0C. With this change in temperature, a propagating wavefront at 100Hz will be completely out of phase with a wave at the original temperature after approximately 15 wavelengths or about 52m. To investigate this case, a perturbed model was built in which the temperature was increased by 20^0 C. Data from the perturbed model simulates signals actually acquired while the data from the original model is an estimate, $\widehat{\mathbf{D}}$, of the true steering vector matrix.

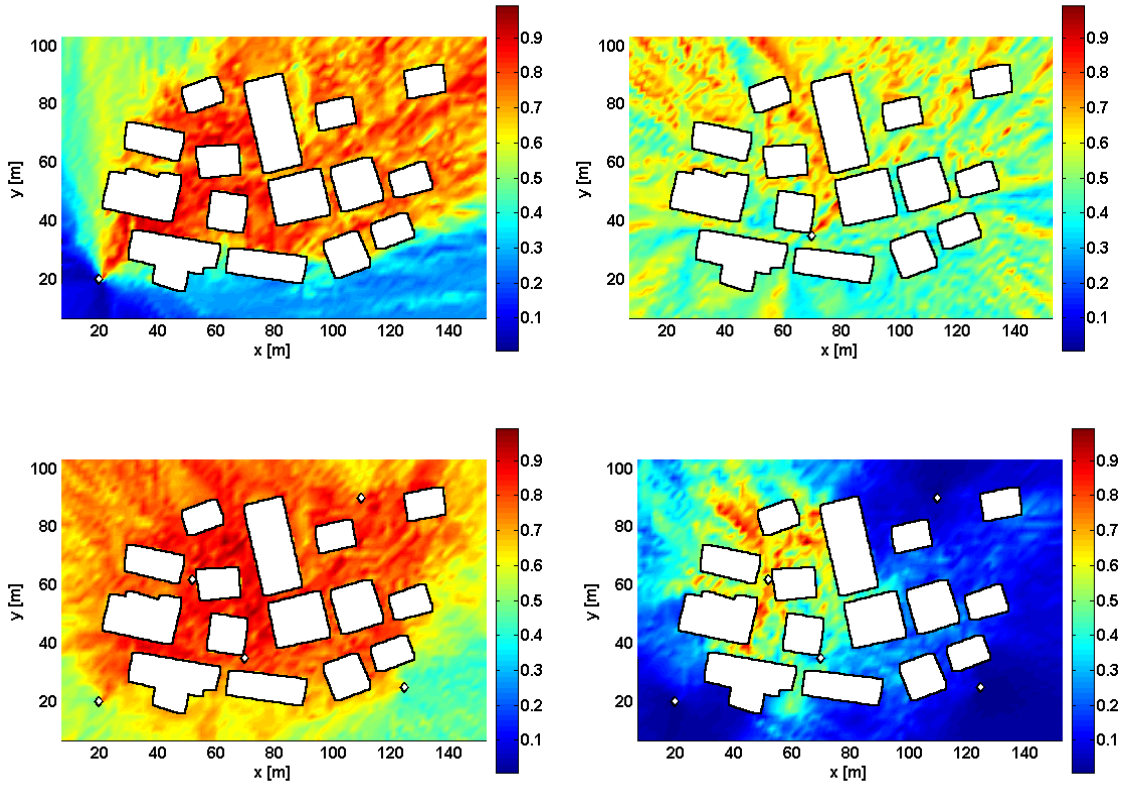

Figure 6.9: Broadband matched-field power response function with a 20ºC uncertainty in temperature for node 1 (upper left), node 2 (upper right), the incoherent array network of 5 nodes (lower left) and the coherent array network of 5 nodes (lower right). The source is located at $\boldsymbol{l}_s = \{65, 73\}$.

Broadband beampatterns due to a 20^0 C perturbation in temperature are presented in Figure 6.9. Compared to Figure 6.8, perturbations of temperature and node location have a similar effect.

6.2.4 Uncertainty affecting inter-array coherence only

Some uncertainty in parameters, such as that governing the wind speed of a moving medium, is likely to have a negligible effect on the closely spaced sensors of a single array but can cause substantial phase differences on a larger scale between microphones on separate nodes. Random propagation effects such as turbulence and wind speed uncertainty can be modeled as a drop in inter-array coherence. A loss of inter-array coherence will also occur as a consequence of time synchronization errors across separate sensor platforms. Each array is referenced to a separate clock which is subject to dissynchronization with the network. More robust than time domain methods, the frequency domain formulation limits phase error to one period, yet this is still substantial at low frequencies. To simulate perfect intra-array coherence while degrading inter-array coherence, we introduce a phase factor zeta

$$\zeta = e^{i\pi\beta\upsilon} \tag{6.5}$$

where $i = \sqrt{-1}$, $\upsilon \in \{0,1\}$ is a uniformly distributed random variable and $\beta \in \{0,2\}$ is a parameter governing the extent of phase distortion. The array signal vectors are concatenated for coherent multiarray processing. Prior to concatenation, $\boldsymbol{r}_n(\omega)\ \forall\ n$ is phase shifted by a realization of ζ such that the affected network signal vector is

$$\boldsymbol{r}_q = \begin{bmatrix} \zeta_1 \boldsymbol{r}_{n=1} \\ \zeta_2 \boldsymbol{r}_{n=2} \\ \vdots \\ \zeta_N \boldsymbol{r}_{n=N} \end{bmatrix} \tag{6.6}$$

Over many averages, the coherence matrix $\boldsymbol{\Phi}_{QQ}$ becomes block diagonal, with ones on the block diagonal and the reduced inter-array coherence elsewhere. Here, we confine our study to the limiting case of zero inter-array coherence induced by $\beta = 2$. Figure 6.10 displays two cases of a broadband coherent network matched-field power response function due to an absence of inter-array coherence. Coherent processing is still viable even though 0 coherence exists.

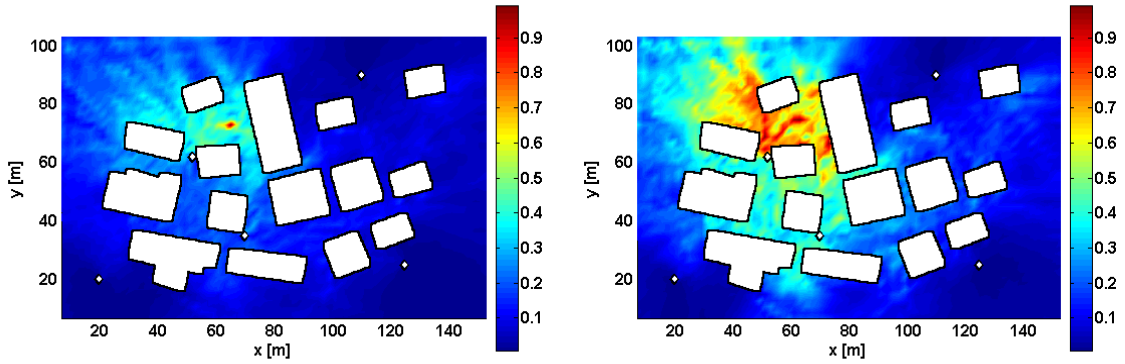

Figure 6.10: Broadband matched-field power response function for the coherent network of 5 arrays with 0 inter-array coherence to be compared with Figure 6.4 (no mismatch, left) and Figure 6.8 (2m location perturbation, right).

The first case, Figure 6.10 (left), considers the effect of zero inter-array coherence only. Compared to Figure 6.4, the result is a very similar response to the unperturbed coherent case with the addition of increased sidelobes. Figure 6.10 (right) is the coherent network response subject to both a 2m perturbation in phase and absence of inter-array coherence. Compared to Figure 6.8 (bottom right), again we see increased sidelobes and a peak that has shifted farther from the correct location.

6.3 Spatial Variation of Performance

Many studies have been done on the consequences of imperfect models [92]. While it is almost always assured that model uncertainty will degrade beamformer

performance, the extent of the degradation is both model specific and highly dependent on the array-source geometry as shown previously. To quantify the impact of mismatch on performance, it is necessary to search a large domain. In the past, this has been difficult as many time-consuming computations must be carried out. With the availability of ever increasing computational power, however, simulations of useful scope can be executed reasonably. The localization performance of each individual array and the entire network considering various forms of uncertainty was calculated over the entire domain.

The RMS localization error, ε_j, of the estimated source location versus the true source location is calculated as

$$\varepsilon_j = \sqrt{\left(\hat{\boldsymbol{l}}_j - \boldsymbol{l}_j\right)^T \left(\hat{\boldsymbol{l}}_j - \boldsymbol{l}_j\right)} \tag{6.7}$$

where $\hat{\boldsymbol{l}}_j$ is chosen as the location that maximizes the broadband matched-field power response function $\boldsymbol{b}_{BB,j}$. The spatial distribution of error considering a 2 meter uncertainty in array location appears in Figure 6.11 for individual arrays and the various network processing schemes.

Turning attention first to the individual array performance, there is a trend of increasing error with distance, especially visible for arrays outside the cluster of buildings. Possibly, it is more accurate to say the error increases with each interaction with a reflecting boundary where the phase mismatch causes constructive versus destructive interference or vice versa. Uncertainty in temperature shows similar trends (discussed later, an example is shown in Figure 6.16).

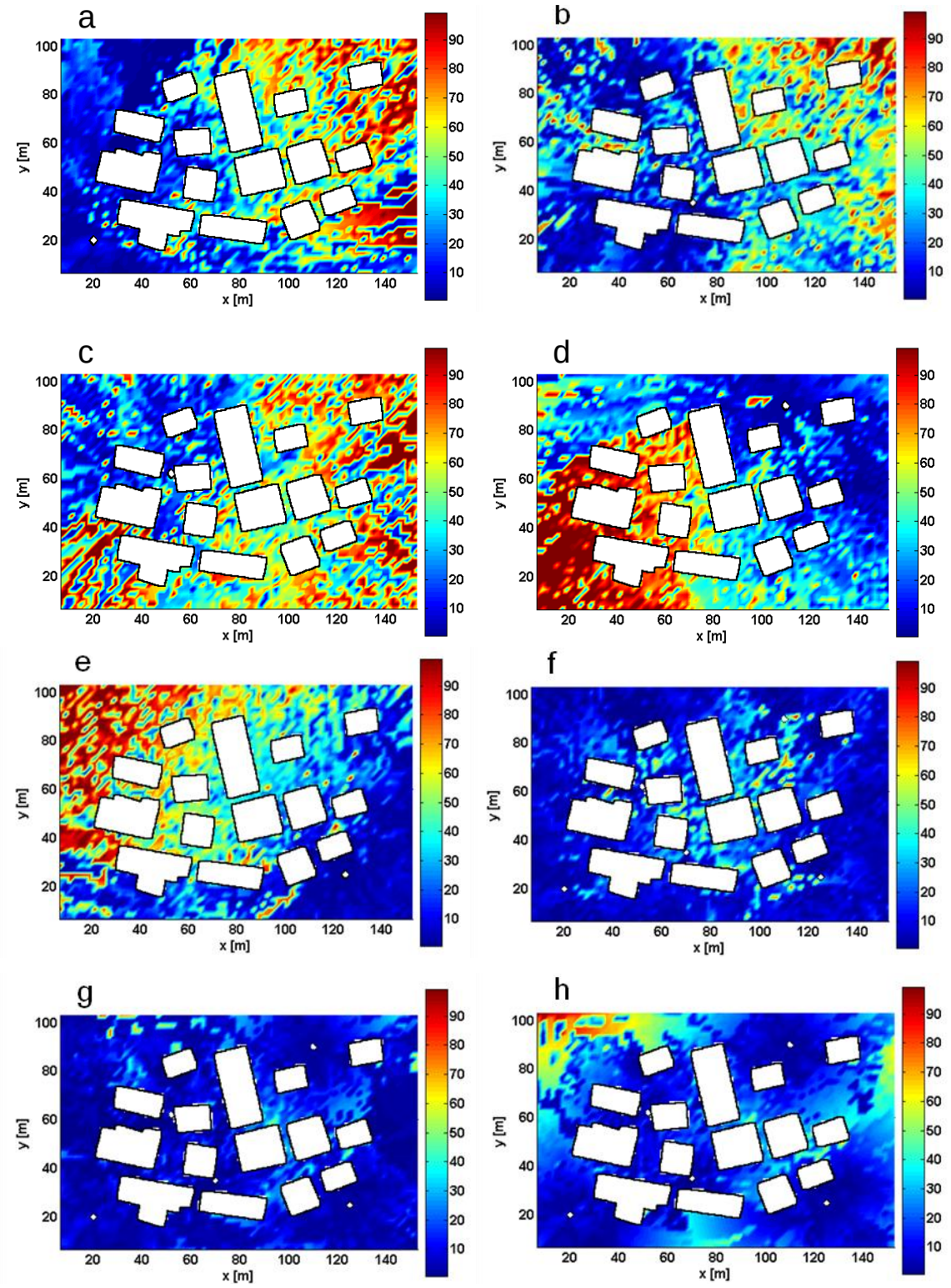

Figure 6.11: Error (meters) of matched field processor due to a 2 meter uncertainty in array location for individual arrays (a-e), incoherent processor (f), coherent processor (g), and coherent processor with additional inter-array incoherence (h).

Although individual arrays may perform well locally, clearly in this environment a network of spatially distributed sensors offers a dramatic increase in performance overall. The coherent processor preserves the effective range of each array and error builds up mostly on the edges. This trend is further pronounced in the coherent processor with additional inter-array incoherence. Even when the assumptions have been violated, coherent processing functions quite well. Meanwhile, the incoherent processor can be susceptible to the influence of one heavily degraded array. However, it is still able to utlize the multiarray information with some degree of effectiveness, on average outperforming any single array. While the maximum value of a single array's beamformer output may be way off mark, the incoherent processor accounts for any energy that is still present at the correct location.

Statistics of the performance considering other forms of uncertainty are collected in Table 6.2. In practical situations, an estimate that is wrong by 50m and an estimate that is wrong by 100m may be equally worthless. The median as well as the mean of the localization error is included as the median does not share the mean's susceptibility to influence by outliers.

	Location Uncertainty		Temperature Uncertainty	
	Mean	Median	Mean	Median
Array 1	30.2	17.9	30.5	16.5
Array 2	27.2	20.4	24.7	20
Array 3	40.1	32.1	35.2	30.5
Array 4	36.1	24.1	27.4	17.1
Array 5	31.4	20.9	28	16.5
Incoherent Q	10.3	5.7	8.1	4.5
Coherent Q	7.1	2.8	8.7	4.5
Coherent Q with ζ	15.6	10.2	15.7	10.2
Best Array	5.2	2	5.8	4
Worst Array	69.9	75.9	62.1	65.3

Table 6.2: Mean and median error (meters) for various processing approaches under the influence of uncertainty in various parameters: medium temperature (20^o perturbation) and array location (2m perturbation).

The incoherent processor actually outperforms the coherent processor in the perturbed temperature case. Because the coherent processor depends on more information than the incoherent, it is more vulnerable to mismatch. When inter-array coherence is completely eliminated (labeled *coherent Q with ζ*), the coherent processor degrades past the incoherent approach. The coherent processor may have less of a margin over the incoherent processor in the temperature case because the location perturbation affected all the sensors on each array similarly. Considering the improbability of coherent conditions, incoherent processing is a robust alternative with reasonable accuracy.

For each source location, the error of the best and worst-performing individual array was recorded (last two rows of Table 6.2). It is interesting to see that, at the majority of locations a single array is more accurate than the entire network. This suggests that there may be a more favorable way to fuse array information.

6.4 Time Reversal Refocusing

Application of time reversal refocusing to source localization has been of recent interest in the literature, indeed in this very same test environment [39], [40]. The method has previously been heavily developed in ultrasound applications, and was shown to yield inferior performance to a matched filter [45], [46]. However, for completeness and perspective on the current application of MFP in an urban environment, a brief example is presented here.

6.4.1 Method

The time reversal refocusing technique is rooted in the time-reversal invariance of the wave equation (5.20). Because this equation contains only second-order derivatives with respect to time, the equation is time-reversal invariant [96]. This means that for any acoustic energy originating from a source that can possibly be diffracted, reflected, or

otherwise transmitted through the propagation medium, there exists a signal that can effectively retrace the intricate path and reconverge on the source location. For example, Figure 6.12 shows the received signal at 1 microphone on node 2 due to a source located at $\boldsymbol{l}_s = \{65, 73\}$ in the test environment. Also shown is the time reversed signal that is virtually emitted from the microphone location using the FDTD model.

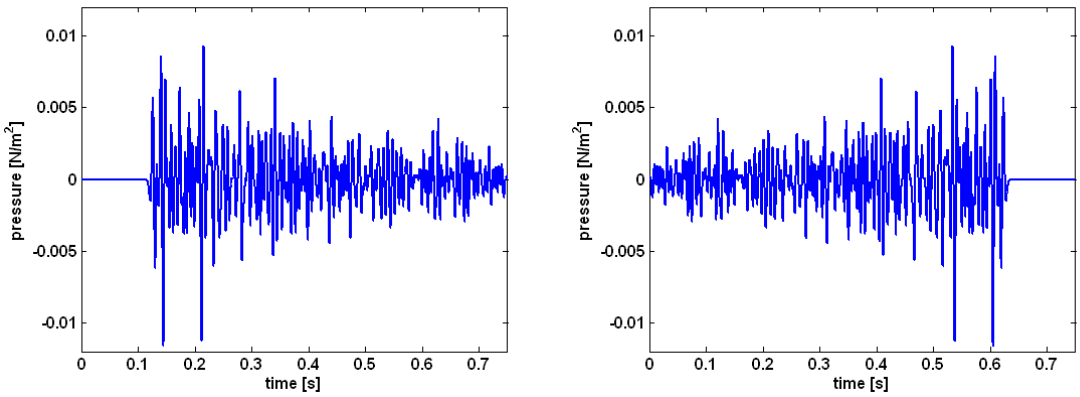

Figure 6.12: Received signal, $r(t)$, at microphone 1 on node 2 (left) and transmitted signal, $r(-t)$, from microphone 1 on node 2 (right).

Time reversal of the received signal can be similarly processed for microphones at the other nodes. A snapshot of the resulting pressure field at the time step coinciding with the peak of the source emission (see Figure 5.4) is shown in Figure 6.13 as calculating by the FDTD model. The 5 individual reversed signals have reconverged on the original source location resulting in a large pressure spike.

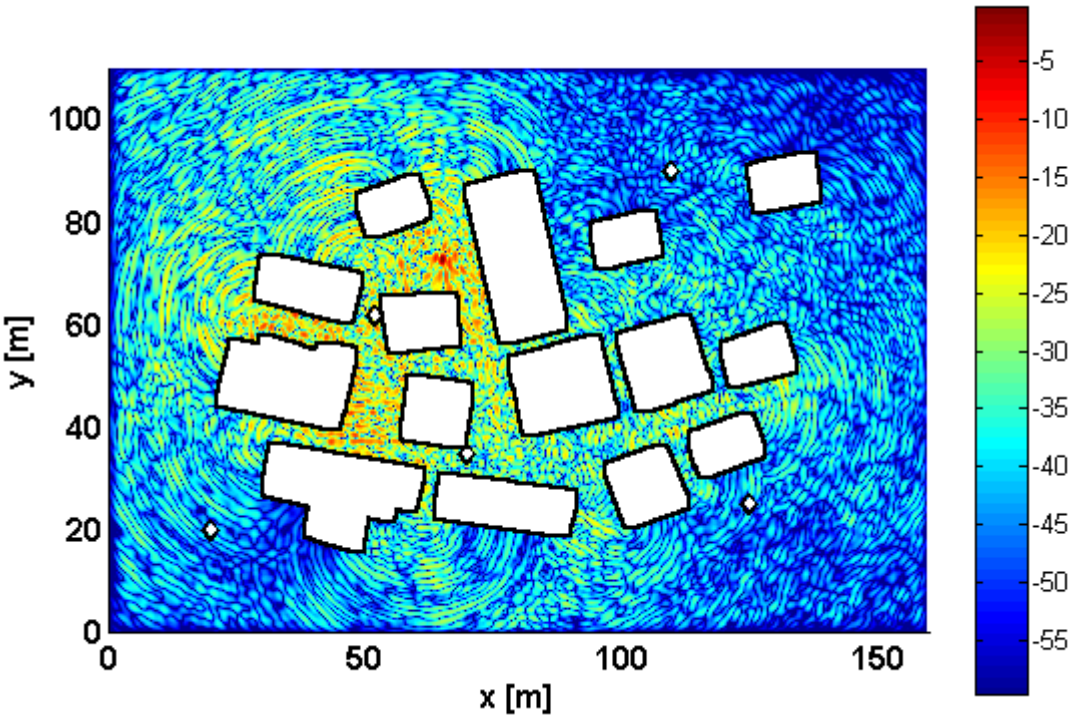

Figure 6.13: Pressure field (N/m^2 dB) at the time step coinciding with the peak of the source emission due to time reversal refocusing of 5 microphone signals.

6.4.2 Practical limitations

Time reversal is an idealized concept not precisely realized in practice (see Chapter 2.3.2 and references). Although very useful for many other applications, the potential for source localization suffers from multiple drawbacks. Beyond the limitations of a time reversal mirror as opposed to the idealized time reversal cavity, time reversal refocusing is subject to dissynchronization, model mismatch, and high computational requirements. Furthermore, the method is inflexible compared to MFP.

In practice, multiple sensors are required and prior to reversal, a common point in time corresponding to the end of the received source signal must be designated. Any network time dissynchronization of individual node clocks will disrupt the precise constructive and deconstructive addition of waves from disparate sensors and degrade the overall focusing ability. Sample accurate coherent processing is required; there is no flexibility in an incoherent processing alternative.

Furthermore, while the refocused pressure field due to a continuous source may be averaged over time to yield the source location estimate, the maximum refocused pressure of a transient source occurs at the time corresponding to the maximum of the original source impulse. This time is not generally know a priori and must somehow be estimated.

In addition, time reversal refocusing suffers from uncertainty in the model similar to MFP. Assuming no synchronization error, the spatial accuracy of the time reversal refocusing method has been calculated for the case of a 20º C perturbation in temperature. The results are shown in Figure 6.14 and compared to the coherent MFP processing results. Because MFP incorporates more information, the resulting estimation power is greater than time reversal refocusing.

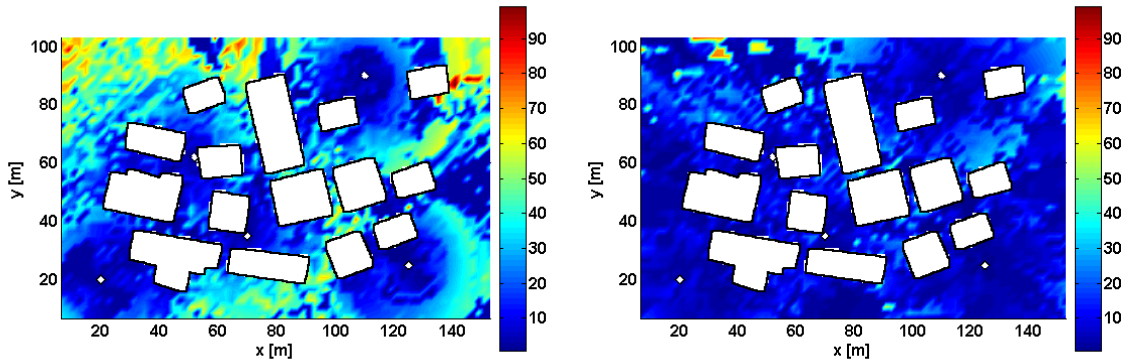

Figure 6.14: Localization accuracy over space. Error (meters) due to a 20º C perturbation in temperature of the medium. Results are shown for the time reversal refocusing method (left) and the coherent matched field processor (right).

Finally, because a steering vector matrix is not required, time reversal refocusing requires far less upfront computation than MFP. However, the time reversal process requires that the numerical model to be run after a detection to refocus the signals. This limits time reversal refocusing as a real-time source localization method. On the other hand, the steering vector matrix can measured a priori and then stored. Consequently, MFP has been boiled down to a single matrix multiplication at each frequency and can be calculated much more efficiently during implementation.

6.5 MFP Using Free-Field Steering Vectors

The motivation for MFP was the degradation of free-field localization methods in non free-field environments and the discussion has now come full circle. Free field techniques are attractive because they do not require a large computational or knowledge investment. Furthermore, simplicity goes hand in hand with robustness, as shown by the coherent versus incoherent processing experiment. The more information an algorithm employs, the more vulnerable it is to poor quality. The free-field weights are susceptible to temperature changes for example, but not to the degree a numerical model is. Moreover, they are completely independent of any information concerning the location of buildings.

While the performance of a matched field processor using free-field weights is independent of much information, the performance is most certainly affected by an acoustic field distorted by scattering objects. In chapter 5, it was shown how a matched field processor utilizing free-field weights can possibly fail, and also how a LOS node was able to retain good performance. In the first part of this subsection, the consequences of the free-field assumption in the very complicated test environment are explored.

The susceptibility of MFP to model mismatch raises many questions, for example: What is the relationship between uncertainty and performance? Do conditions exist in which it is advantageous to abandon the model and rely on free-field techniques? The extent of error due to mismatch can be addressed by performance bounds such as the Cramer-Rao Lower Bound. Although it is possible to apply the Cramer-Rao Bounds to MFP [97] in some situations, it requires closed form expressions for the pressure distribution of the acoustic field. Therefore, a single example is studied and although the results cannot be extrapolated with complete confidence, some insight is offered.

6.5.1 Beampatterns and performance

To begin, MFP using the weights derived from the free-field propagation model (see section 5.2) is applied to the familiar scenario of section 6.1 and 6.2. Example

beampatterns using individual nodes and the network processing approaches are shown below in Figure 6.15.

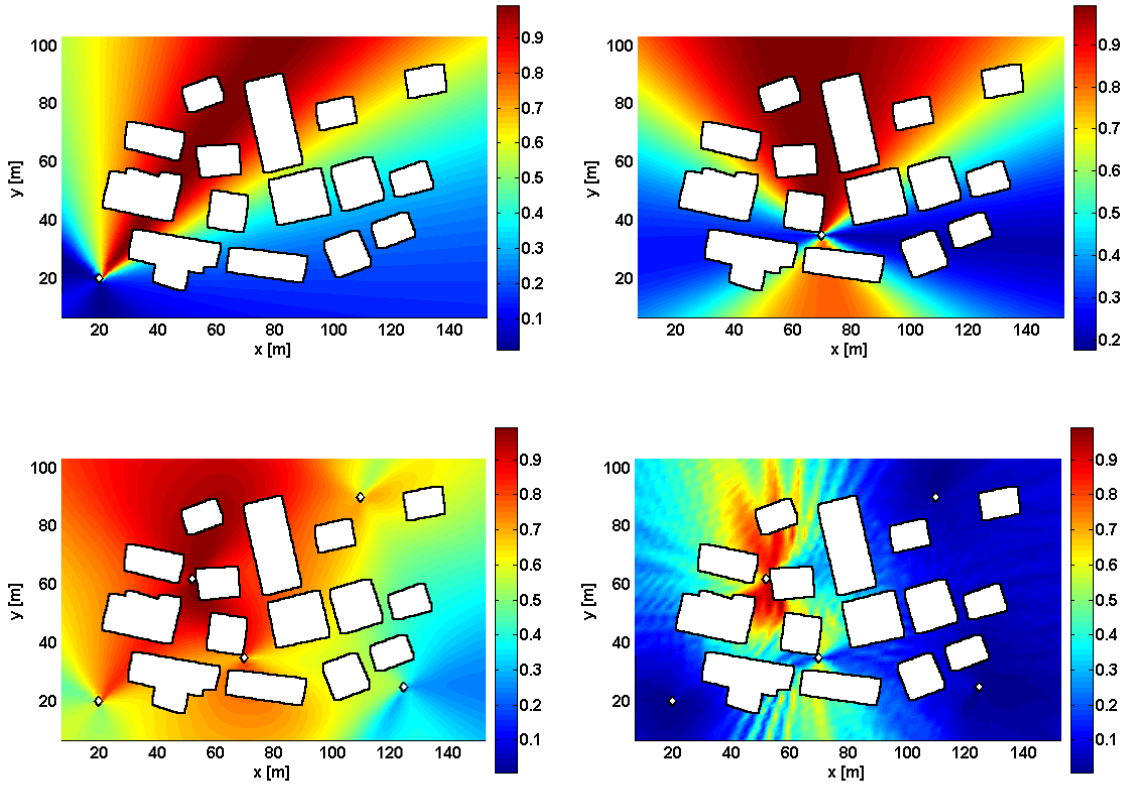

Figure 6.15: Broadband matched-field power response function using free-field weights for array 1 (upper left), array 2 (upper right), the incoherent array network of 5 arrays (lower left), the coherent array network (lower right). The source is located at $\boldsymbol{l}_s = \{65, 73\}$.

Figure 6.15(a) is an example of free-field based weights succeeding in a scattering environment for a non-LOS array. Note however that range is ambiguous. The beampattern of array 2 is a more common example of how a free-field based processor reacts in a multi-wave environment. Positively, the processor is somewhat flexible in its ability to recognize energy approaching from multiple directions.

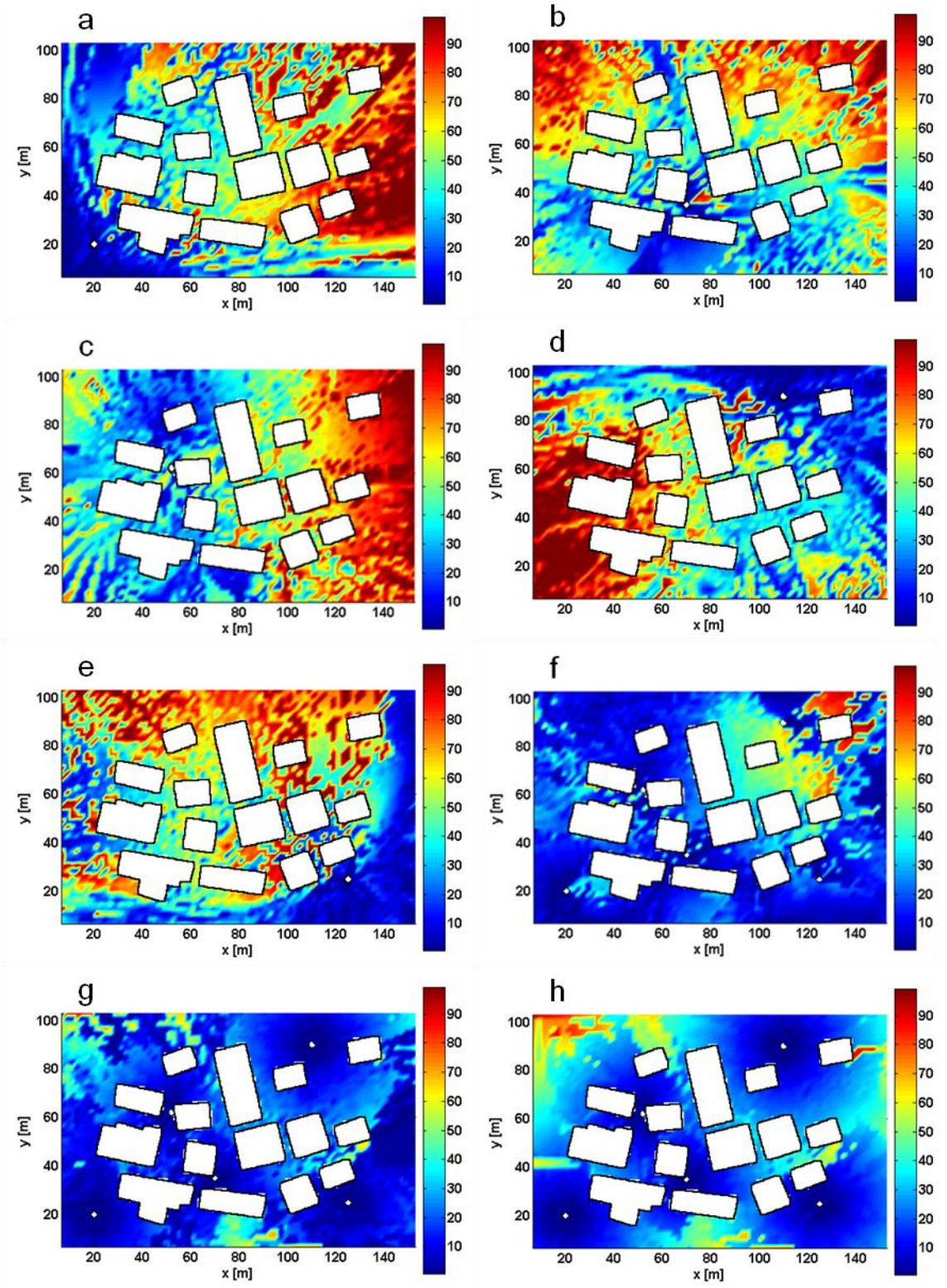

Figure 6.16: Error (meters) of matched field processor using free-field weights. Results are shown for individual arrays 1-5 (a-e), incoherent processor (f), coherent processor (g), and coherent processor with additional inter-array incoherence (h).

Similar to the model-based matched field processor, the coherent free-field processor has greatly reduced sidelobes compared to the incoherent case. None of the individual arrays or networked beamformer outputs has a maximum at the true source location however. Beyond the current example, this is true for the majority of source locations. Similar to section 6.3, the spatial performance of the MFP using free-field weights was calculated and is shown in Figure 6.16.

In that the individual node performance is correlated with distance, the results are similar to the perturbed location case. However, a more important condition for accurate MFP with free-field weights is LOS, most clearly indicated by the performance of node 2. The other strong resemblance to the full-field model mismatch results is how the coherent processor retains accuracy in the immediate vicinity of each node. This supports the intuition that the quantity of sensors is proportional to quality of information. Of course, in practice there is a limit to the number of sensors available.

6.5.2 Parametric performance study

In the limit of having infinite arrays, good performance is assured. Practically, the quantity of sensors is limited. In this subsection, the rate of change of performance with the number of available nodes is investigated. This study is carried across multiple scenarios: MFP using free-field weights and using full-field weights under variable degrees of uncertainty.

The various cases are evaluated by the metric of percent coverage. Previous results concentrated on the absolute deviation of the location estimate. Practically, at some point the magnitude of error is irrelevant, i.e. the estimate is useless whether it is far off or way far off. Percent coverage is defined here as the percentage of total locations J at which the error ε_j is less than some threshold. Figure 6.17 shows the percent coverage for 5 situations using a threshold of 20 meters: MFP using free-field weights, MFP using full-field weights considering a perturbation in array location of 1 meter, 2 meters, 5 meters, and 9 meters. Each array is perturbed progressively farther in one direction as indicated by Figure 6.7. Both the free-field and full-field situations utilize incoherent processing with 1, 2, 3, 4 or 5 nodes. The 1 node case uses data from node 1 only as

defined by Table 6.1. The 2 node case uses data from nodes 1 and 2. The 3 node case uses data from nodes 1, 2, and 3 and so on.

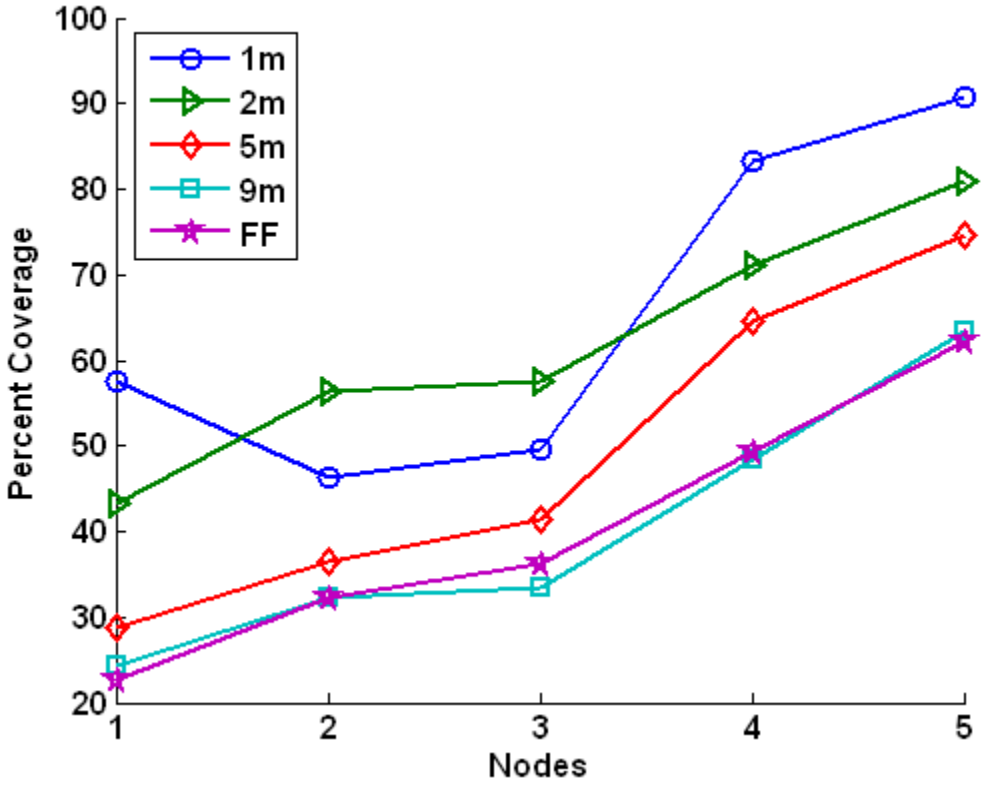

Figure 6.17: Percent coverage (error < 20m) of full-field weighted MFP considering increasing levels of uncertainty and free-field weighted MFP.

As presented explicitly early in this chapter, the effect of uncertainty is very complex over space. The results in Figure 6.17 consider only one spatial distribution of arrays in one arrangement and density of buildings and location perturbation along a single direction. Despite the noted limitations, the performance is evaluated at a reasonable sample size of over 3000 locations.

It should be noted that the free-field results for the single array represent a consequence of the simulation more than physical performance. Although on some numerical scale there are slight inter-microphone phase differences, passive ranging with a single, small aperture array is ambiguous.

Two intuitive trends are confirmed by Figure 6.17: performance decreases as the level of uncertainty increases and performance increases as the number of nodes (and therefore available information) increases. Furthermore, it is advantageous to use uncertain model data over a free-field model for MFP weights, at least in this test environment. Based on typical GPS equipment / geospatial data, uncertainty in the geometry of the boundaries (position and size of buildings and land contours) is approximately 1m. The movement of vehicles in an urban environment can be interpreted as an additional uncertainty of geometry on the order of 0-5m, depending on the proximity of the obstacle to the array [134].

Chapter 7:

Summary & Future Work

The problem of acoustic source localization using passive methods has been studied in a variety of environments. Supporting topics including signal detection, classification, and tracking as they pertain to acoustic sources were also addressed. The work in this dissertation makes multiple contributions to the fields of array processing, sensor networks, detection theory, and acoustic localization. This final chapter reviews these contributions in light of the work presented and offers some comments on future directions.

7.1 Signal Detection

In chapter 3, the coupled problems of signal detection and classification are investigated. Through the study of elementary signals, simple metrics that capture the underlying physics of many signals were presented. These metrics efficiently separate signals into coarsely defined categories, the top layer of a classification hierarchy. Additionally, the metrics were combined with detection theory to show how signal detectors could be formulated. The simplicity and use of an ordered statistic (the median) makes these detectors robust and useful. The metrics have two key properties: attenuation of unwanted signals and adaptive, instantaneous noise estimation.

The ability of the proposed signal detectors to be formulated non-parametrically makes them potentially very valuable for real-time applications. In a real-time situation, non stationary conditions may cause the noise distribution to change faster than it is

possible to update the threshold of a constant false alarm rate detector, resulting in sub optimal performance. The proposed detectors can be formulated as constant signal to noise ratio detectors. Future work implementing these detectors could include a performance study versus optimal detectors in adverse conditions.

7.2 Detection Association

Chapter 4 was concerned with many of the practical issues of source localization and tracking. A prototypical sensor network was designed and built to experimentally support the theoretical development. An original system of associating individual node detections and bearing estimates was presented. There are data association problems unique to an acoustic sensor network. In order to handle the inevitability of missing, delayed, and out of sequence measurements, event detection windows are created and managed. A multiple hypothesis methodology was formulated to circumvent the range deficiency of passive localization by considering node pairs. This was formulated using the available classification information and the assumption of free-field propagation, although the method is not limited to this condition. The detection association method is embedded within a centralized fusion center that incorporates multiple data association routines to perform localization and tracking. This method was successfully employed to associate hundreds of measurements during an outdoor experiment in a multisource environment.

7.3 Bearings-only Localization

Additionally in chapter 4, methods of obtaining a position fix using a distributed sensor network and bearings-only measurements were investigated. Novel pair-wise fusion strategies were introduced and found to provide more accurate location estimates

than the conventional linear least squares technique in both simulation and experimental settings.

Pair-wise fusion exploits the robust nature of sensor networks; the fused result can be conditioned against the error magnifying effects of GDOP and systematic error in individual measurements. Conditioning against GDOP is achieved by inversely weighting the observations by their respective covariance. Weighting by the multi-dimensional covariance matrix as opposed to a scalar measure of GDOP [66] allows measurements of accurate coordinates to be utilized while other dimensions are suppressed. Conditioning against strong GDOP as well as bias can achieved using the robust pair-wise MED estimator. By first triangulating pairs of bearings into spatial observations, a statistical population is created in which outliers are exposed in Cartesian space.

The uncertainty of the fused MLE was calculated and confirmed to be an accurate representation of the fused location error, making this fusion strategy suitable for further use in tracking applications such as the Kalman filter. The analytical formulation of the covariance matrix of observations developed herein can also be used for future applications, such as finding the optimal configuration of nodes for a given source location. Future work may also involve estimating bearing biases, expanding the MLE formulation to include uncertainty in node location, and determining the fused covariance matrix of the pairwise MED estimator.

7.4 Matched Field Processing

On a coarse scale, acoustic propagation in an urban environment is analogous to the reverberant field created by scattering objects and has been modeled thusly. The complex field gives rise to a unique relationship between any two points in space that can be exploited for source localization. FDTD models can be used to generate the steering vector matrix for MFP, which is an attractive alternative to time reversal refocusing.

In addition to superior performance, MFP offers flexibility of implementation and control over design parameters through established beamforming framework. Because beamforming can be formulated in the frequency domain, it is not necessary to transmit

the full time domain data and network time synchronization requirements can be relaxed for incoherent processing. Sensor data can be locally processed at each array and only the distilled information need be transmitted to be fused with other arrays. The time accuracy required is on the order of an acoustic event as opposed to a sample. A disadvantage of this inverse method is that many calculations are needed to build the propagation matrix, even for conservative resolution of the source location vector. However, these calculations can be made before implementation and stored for real-time source localization.

Realistically, there is always some uncertainty in various model parameters that in turn can adversely affect localization ability. Multiple forms of uncertainty were investigated: additive incoherent noise, perturbation of the model parameters array location and medium temperature, and network time dissynchronization. Uncertainty in model parameters can be interpreted as coherent noise. The uncertainty manifests itself as error in the position estimate and some insight into the mechanism was garnered by quantifying the error over space. It was found that the degree of performance degradation is dependent on the spatial source-receiver relationship. Specifically, error increases through interaction with the boundaries.

Incoherent as well as coherent multi-array fusion strategies were investigated and shown to mitigate the effects of model uncertainty. Because individual arrays are spatially separated and subject to clocking errors and large-scale medium inhomogeneities, the inter-array coherency is compromised. While the coherent processor is able to make use of more information, it is ultimately more vulnerable to mismatch. In severe cases, it performs worse than the incoherent processor over the entire space. More information is not always detrimental if applied robustly, e.g. incoherent MFP uses more information than time reversal refocusing. While it is only certain that each individual array is locally accurate, one random array could be identified to outperform the multiarray strategies over the majority of the domain. This suggests that there may be a more favorable way to fuse array information, and the future understanding of factors that influence array performance in such complex situations will yield superior array positioning and fusion strategies.

Finally, the feasibility of free-field weights in a full-field environment was investigated. This approach incurred large localization errors in the obstacle-dense test environment. However, when line of sight is available such as in a more open environment or if many nodes are available, reasonable performance is achievable. A final study showed that if uncertainty in model parameters is high, free-field weights are a viable alternative. It is postulated that in real propagation conditions, the margin between free-field weights and numerical model based weights may be narrower. Semi-porous obstacles with finite impedance and 3-dimensional propagation can facilitate line-of-sight while the increase in complexity heightens the potential for mismatch in the numerical model.

There is a plethora of ways in which future work can proceed in this research area. The research area is still very much in the nascent stage and several important questions need to be addressed and answered in a quantifiable manner due to the wide-ranging applicability of acoustic source localization. Future work should include experimental implementation and assess the necessity of more complicated models, such as those integrating 3D propagation and varying ground and building impedances for example. There is also opportunity for extension of the model information to other free-field localization techniques, similar to how MFP is an adaptation of plane-wave beamforming to reverberant environments. For example, localization using the relative loudness of spatially separated receivers in a free field is formulated assuming spherical spreading only. In a heterogeneous medium, the relative amplitude is governed by a sensor's transmission loss function. Likewise, TDOA methods assume free field propagation along a straight line. These methods can be adapted by acquiring the true distorted path length from a model as well.

Bibliography

[1] C. Chong and S. P. Kumar, "Sensor Networks: Evolution, Opportunities, and Challenges," Proceedings of the IEEE, Vol. 91, No. 8, pp. 1247-1256, August 2003.

[2] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda, Y. Choi, T. Herman, S. Kulkarni, U. Arumugam, M. Nesterenko, A. Vora, and M. Miyashita, "A line in the sand: a wireless sensor network for target detection, classification, and tracking," Computer Networks, Vol. 46, pp. 605-634, 2004.

[3] L.G. Mazerolle, J. Frank, D. Rogan, and C. Watkins, "Field Evaluation of the ShotSpotter Gunshot Location System: Final Report on the Rewood City Field Trial," submitted to the U.S. Department of Justice, Document No. 180112, January 2000.

[4] J. Carey, R. Govindan, E. Johnson, B. Krishnamachari, S. Masri, and G. Sukhatme. Networked Sensing for Structural Health Monitoring. In Proceedings of the Fourth International Workshop on Structural Monitoring and Control, June 2004.

[5] C. O. Tiemann, A. M. Thode, J. Straley, V. O'Connell, and K. Folkert, "Three-dimensional localization of sperm whales using a single hydrophone," Journal of the Acoustical Society of America, Vol. 120, No. 4, pp. 2355-2365, Oct. 2006.

[6] R. Szewczyk, A. Mainwaring, , J. Polastre, J. Anderson, and D. Culler. "An Analysis of a Large Scale Habitat Monitoring Application," SenSys '04, Baltimore, Maryland, USA, November 2004.

[7] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the World With Wireless Sensor Networks," Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing ICASSP, May 2001.

[8] J.R. Moore and W.D. Blair, “Practical Aspects of Multisensor Tracking,” Chapt.1, Multitarget, Multisensor tracking: Applications and Advances, Volume III, Y. Bar-Shalom and W.D. Blair, ed., Artech House, Norwood, MA, 1990.

[9] Y. Bar-Shalom, and W.D. Blair, “Multitarget-Multisensor Tracking: Applications and Advances, Volume 1-3,” Artech House, 2000.

[10] M.P Dana, “Registration: A Prerequisite for Multiple Sensor Tracking,” article in “Multitarget-Multisensor Tracking: Advanced Applications," Artech House, 1990.

[11] T. Fortman, “Sonar tracking of multiple targets using joint probabilistic data association, “ IEEE Journal of Oceanic Engineering, Vol. 8, No. 3, pp. 173-184, 1983.

[12] D. Reid, “An algorithm for tracking multiple targets, “IEEE Transactions on Automatic Control, Vol. 24, No. 6, pp. 843-854, 1979.

[13] D. S. K. Chan, and D. A. Langan, “Performance Results of the Bilevel MHT Tracking Algorithm for Two Crossing Targets in a High Clutter Environment,” Signal and Data Processing of Small Targets 1994, Proc. SPIE, Vol. 2235, pp. 406-416, April 1994.

[14] M. deFeo, et al., “IMM JPDA versus MHT and Kalman Filter with NN Correlation Performance Comparison,” IEEE Proc. Radar, Sonar, Navigation, Vol. 144, No. 2, pp. 49-56, Apr. 1997.

[15] R. E. Kalman, “A New Approach to Linear Filtering and Prediction Problems,” Journal of Basic Engineering, vol. 82, No. 1, pp. 35-46, 1960.

[16] P. Zarchan and H. Musoff, “Fundamentals of Kalman Filtering: A Practical Approach,” American Institute of Aeronautics and Astronautics, Inc. 2005.

[17] V. Cevher, R. Velmurugan, and J. H. McClellan, “A Range-Only Multiple Target Particle Filter Tracker,” Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing ICASSP, May 2006.

[18] X. R. Li and V. P. Jilkov, “A Survey of Maneuvering Target Tracking: Dynamic Models,” Signal and Data Processing of Small Targets 2000, Proc. SPIE, Vol. 4048, pp. 212-235, 2000.

[19] J. L. Gertz, "Multisensor Surveillance for Improved Aircraft Tracking," Lincoln Laboratory Journal, Vol. 2, No. 3, pp. 381-396, 1989.

[20] R. A. Singer, "Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets," IEEE Trans. Aerospace and Electronic Systems, AES, Vol. 6, pp. 473-483, July 1970.

[21] X. R. Li and V. P. Jilkov, "A Survey of Maneuvering Target Tracking-Part IV: Decision-Based Methods," Signal and Data Processing of Small Targets 2002, Proc. SPIE, Vol. 4728, pp. 511-534, 2002.

[22] X. R. Li and V. P. Jilkov, "A Survey of Maneuvering Target Tracking-Part III: Measurement Models," Signal and Data Processing of Small Targets 2001, Proc. SPIE, Vol. 4473, pp. 423-446, 2001.

[23] S. T. Birchfield and R. Gangishetty, "Acoustic Localization by Interaural Level Difference," In Proceedings of ICASSP, Vol. 4, pp. 1109-1112, 2005.

[24] J. L. Flanagan, "Computer-steered microphone arrays for sound transduction in large rooms, "Journal of the Acoustical Society of America, Vol. 78, no. 5, pp. 1508-1518, Nov 1985.

[25] M. Brandstein, and D. Ward, *Microphone Arrays,* Springer, 2001.

[26] S. P. Drake, K. Dogancay, "Geolocation By Time Difference of Arrival Using Hyperbolic Asymptotes," Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing ICASSP, 2004.

[27] M. Johnson, J. Carneal, and P. Gillett, "Comparison of a diffracting and a non-diffracting circular acoustic array," ICASSP 2006, Toulouse, France, vol. 4, pp. 1081-1084, 2006.

[28] Y. Oshman and P. Davidson, "Optimization of observer trajectories for bearings-only target localization," IEEE Trans. On Aerospace and Electronic Systems, vol.35, no.3, pp. 892-902, July 1999.

[29] L. M. Kaplan, Q. Le, and P., Molnár, "Maximum Likelihood Methods for Bearings-Only Target Localization," IEEE Proceedings on Acoustics, Speech, and Signal Processing, vol. 5, pp. 3001-3004, May 2001.

[30] G. De Sanctis, D. Rovetta, A. Sarti, “Localization of Tactile Interactions Through TDOA Analysis: Geometric Vs. Inversion-Based Method, “ XIV European Signal Processing Conference EUSIPCO, Florence, Italy, September 4 - 8, 2006.

[31] C. Knapp and G. Carter, “The generalized correlation method for estimation of the time-delay,” IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-24, pp. 320–327, September 1976.

[32] M. Brandstein and H. Silverman, “A practical methodology for speech source localization with microphone arrays,” Comput. Speech Lang., vol. 2, pp. 91–126, 1997.

[33] M. Omologo and P. Svaizer, “Use of the cross-power spectrum phase in acoustic event localization,” Acoust., Speech, Signal Process., vol. 5, pp. 288–292, 1976.

[34] A. Stephenne and B. Champagne, “Cepstral pre-filtering for time delay estimation in reverberant environments,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, ICASSP, pp. 3081–3084, 1999.

[35] P. W. Boettcher, J. A. Sherman, and G. A. Shaw, “Target localization using acoustic time-difference of arrival in distributed sensor networks, “Battlespace Digitization and Network-Centric Warfare II, Proceedings of SPIE Vol. 4741, 2002.

[36] S. P. Drake, K. Dogancay, “Geolocation by time difference of arrival using hyperbolic asymptotes, “Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing ICASSP, 2004.

[37] D. Carevica, “Tracking target in cluttered environment using multilateral time-delay measurements,” J. Acoust. Soc. Am. 115, pp. 1198-1206, 2004.

[38] M. K. Broadhead, L. A. Pflug, and R. L. Field, “Use of higher order statistics in source signature estimation,” J. Acoust. Soc. Am. 107, pp. 2576-2585, 2000.

[39] D. G. Albert, L. Liu, and M. L. Moran, “Time reversal processing for source location in an urban environment,” J. Acoust. Soc. Am. 118, pp. 616-619, 2005.

[40] L. Liu, D. G. Albert, and D. K. Wilson, “The effect of changing scatterer positions on acoustic time-reversal refocusing in a 2D urban environment at low frequencies,” Journal of Geophysical Engineering 4, pp. 276-284, 2007.

[41] M. Tanter, J.-F. Aubry, J. Gerber, J.-L. Thomas, and M. Fink, “Optimal focusing by spatio-temporal inverse filter. I. Basic principles,” J. Acoust. Soc. Am. 110, pp. 37–47, 2001.

[42] K. G. Sabra, P. Roux, H. C. Song, W. S. Hodgkiss, W. Kuperman, T. Akal, and M. Stevenson, “Experimental demonstration of iterative time-reversed reverberation focusing in a rough waveguide: Application to target detection,” J. Acoust. Soc. Am. 120, pp. 1305–1314, 2006.

[43] G. Montaldo, M. Tanter, and M. Fink, "Real time inverse filter through iterative time reversal," J. Acoust. Soc. Am. 115, 2004.

[44] D. Cassereau and M. Fink, "Time reversal of ultrasonic fields. III. Theory of the closed time reversal cavity," IEEE Trans. Ultrason. Ferroelectr. Freq. Control 39, 579, 1992.

[45] J.-F. Aubry , M. Tanter, J. Gerber, J.-L. Thomas, and M. Fink, “Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media,” J. Acoust. Soc. Am. 110, pp. 48–58, 2001.

[46] M. Tanter, J.-L. Thomas, and M. Fink, “Time-reversal and the inverse filter,” J. Acoust. Soc. Am. Vol. 108, pp. 223–234, 2000.

[47] A. B. Baggeroer, W.A. Kuperman, and P. N. Mikhalevsky, “An overview of matched field methods in ocean acoustics,” IEEE J. Ocean Eng. 18, pp. 401-424, 1993.

[48] J. Carneal, M. Johnson and P. Gillett, “Comparison of a Diffracting and a Non-Diffracting Circular Acoustic Array,” IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 1081-1084, 2006.

[49] L. Fialkowski, M. Collins, W. Kuperman, J. Perkins, L. Kelly, A. Larsson, J. Fawcett, and L. Hall, "Matched-field processing using measured replica fields," J. Acoust. Soc. Am. 107, pp. 739–746, 2000.

[50] J. Picaut, J. Hardy, and L. Simon, “Sound propagation in urban areas: A periodic disposition of buildings,” Physical Review E, Vol. 60, no. 4, pp. 4851-4859, 1999.

[51] T. J. R. Hughes, *The finite element method, linear static and dynamic finite element analysis*, Prentice-Hall International, Englewood Cliffs, NJ, 1987.

[52] G. Seriani, and E. Priolo, “Spectral element method for acoustic wave simulation in heterogeneous media,” Finite Elements in Analysis and Design, vol. 16, pp. 337-348, 1994.

[53] G. H. Koopmann, L. Song, and J. B. Fahnline, ‘‘A method for computing acoustic fields based on the principle of wave superposition,’’ J. Acoust.Soc. Am. 86, pp. 2433–2438, 1989.

[54] O. M. Bouthier and R. J. Bernhard, “Simple models of the energetics of transversely vibrating plates,” Journal of the Acoustical Society of America, Vol. 189, pp. 149-164, 1995.

[55] L. P. Franzoni, D. B. Bliss, and J. W. Rouse, “An acoustic boundary element method based on energy and intensity variables for prediction of high-frequency broadband sound fields,” Journal of the Acoustical Society of America, Vol. 110, no. 6, pp. 3071-3080, 2001.

[56] R. Bullen and F. Fricke, “Sound propagation at a street intersection in an urban environment,” Journal of Sound and Vibration, Vol. 54, pp. 123-129, 1997.

[57] J. S. Lamancusa and P.A. Daroux, “Ray tracing in a moving medium with two-dimensional sound-speed variation and application to sound propagation over terrain discontinuities,” Journal of the Acoustical Society of America, Vol. 93 (4), pp. 1716-1726, 1993.

[58] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer Press, 1994.

[59] B. Friedlander and B. Porat, “Detection of Transient Signals by the Gabor Representation,” IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 2, 1989.

[60] M. Frisch, and H. Messer, “Transient Signal Detection Using Prior Information in the Likelihood Ratio Test,” IEEE Transactions on Signal Processing, Vol. 41, No. 6, 1993.

[61] R. R. Allen and S. S. Blackman, “Implementation of an Angle-only Tracking Filter,” SPIE Signal and Data Processing of Small Targets, Vol. 1481, 1991.

[62] S.S. Blackman, "Association and Fusion of Multiple Sensor Data," Chapter 6, *Multi-target, Multi-sensor Tracking: Advanced Applications*, Y. Bar-Shalom, ed., Artech House, Norwood, MA, 1990.

[63] Y. Oshman and P. Davidson, "Optimization of observer trajectories for bearings-only target localization," IEEE Trans. On Aerospace and Electronic Systems, Vol. 35, No. 3, pp. 892-902, July 1999.

[64] L. M. Kaplan, Q. Le, and P., Molnár, "Maximum Likelihood Methods for Bearings-Only Target Localization," IEEE Proceedings on Acoustics, Speech, and Signal Processing, Vol. 5, pp. 3001-3004, May 2001.

[65] C. Cohen and F. Koss, "A Comprehensive Study of Three Object Triangulation", Proceedings of the 1993 SPIE Conference on Mobile Robots, Boston, MA, November 1993.

[66] A. Kelly, "Precision Dilution in Triangulation-Based Mobile Robot Position Estimation," Intelligent Autonomous Systems, Amsterdam, 2003.

[67] M. Betke and L. Gurvits, "Mobile Robot Localization Using Landmarks", IEEE Transactions on Robotics and Automation, Vol. 13, No. 2, April 1997.

[68] H. McCabe, and M. Al-Samara, "Application of a sensor fusion algorithm to nonlinear data containing nonstationary, multiplicative noise of unknown distribution," Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on, Vol. 2, Issue 14-16, pp. 1205–1206, December 1994.

[69] S. McLaughlin, V. Krishnamurthy, and S. Challa, "Managing data incest in a distributed sensor network," IEEE Proceedings on Acoustics, Speech, and Signal Processing, vol. 5, pp. 269-272, 2003.

[70] R. S. Figliola, and D.E. Beasley, *Theory and Design for Mechanical Measurements*, John Wiley & Sons, Inc. 2000.

[71] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Ch. 26, Dover Publications, New York, 1970.

[72] R. Christensen, "Linear Models for Multivariate, Time Series, and Spatial Data," Springer-Verlag, 1991.

[73] F. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, John Wiley and Sons, 1985.

[74] Blackman, Samuel S., “Multiple-Target Tracking with Radar Applications,” Artech House, Inc., 1986.

[75] D.L. Mills, “Improved Algorithms for Synchronizing Computer Network Clocks,” Proc. of ACM SIGCOMM 94 Symposium, September 1994.

[76] D. K. Wilson, "Performance bounds for acoustic direction-of-arrival arrays operating in atmospheric turbulence," J. Acoust. Soc. Am., Vol. 103, pp. 1306–1319, 1998.

[77] R. J. Kozick and B. M. Sadler, “Source localization with distributed sensor arrays and partial spatial coherence,” IEEE Transactions on Signal Processing, Vol. 52, pp. 601-616, 2004.

[78] Cox, H., “Resolving power and sensitivity to mismatch of optimum array processors,” J. Acoust. Soc. Am. Vol. 54, pp. 771-785, 1973.

[79] A. B. Baggeroer, W.A. Kuperman, and H. Schmidt, “Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem,” J. Acoust. Soc. Am. Vol. 83, pp. 571–587, 1988.

[80] C. Soares and S. M. Jesus, “Broadband matched-field processing: Coherent vs incoherent approaches,” J. Acoust. Soc. Am. Vol. 113, No. 5, pp. 2587–2598, 2003.

[81] A. Derode, A. Tourin, and M. Fink, “Time reversal in multiply scattering media,” Ultrasonics, Vol. 36, pp. 443–447, 1998.

[82] J. F. Aubry , M. Tanter, J. Gerber, J. L. Thomas, and M. Fink, “Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media,” J. Acoust. Soc. Am. Vol. 110, pp. 48–58, 2001.

[83] W. C. Kirkpatrick, J. M. Noble, and M. A. Coleman, “Sound propagation in the vicinity of an isolated building: An experimental investigation,” J. Acoust. Soc. Am. Vol. 124, pp. 733–742, 2008.

[84] D. Botteldooren, “Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid,” Journal of the Acoustical Society of America, Vol. 95, pp. 2313-2319, 1994.

[85] V. E. Ostashev, D.K. Wilson, L. Liu, D. F. Aldridge, N. P. Symons and D. Marlin, "Equations for finite-difference, time-domain simulation of sound propagation in moving inhomogeneous media and numerical implementation," Journal of the Acoustical Society of America, Vol. 117, pp. 503-517, 2005.

[86] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 4th ed., Wiley, New York, 2000.

[87] J. G. Tolan, J. B. Schneider, "Locally conformal method for acoustic finite-difference time-domain modeling of rigid surfaces," J. Acoust. Soc. Am., Vol. 114, pp. 2575–2581, 2003.

[88] L. Liu and D. G. Albert, "Acoustic pulse propagation near a right-angle wall," Journal of the Acoustical Society of America, Vol. 119, pp. 2073-2083, 2006.

[89] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," IEEE Transactions on Antennas Propagation, Vol. 14, pp. 302-307, 1966.

[90] R. E. Sheriff, *Encyclopedic dictionary of exploration geophysics*, 3rd Ed., Society of Exploration Geophysicists, pp. 376, 1994.

[91] L. R. Lines, R. Slawinski, and R. P. Bording, "A recipe for stability of finite-difference wave-equation computations," Geophysics, Vol. 64, pp. 967–969, 1999.

[92] D. F. Gingras, "Methods of predicting sensitivity of matched-field processors to mismatch," J. Acoust. Soc. Am., Vol. 86, pp. 1940–1949, 1989.

[93] S. Yon, M. Tanter, and M. Fink, "Sound focusing in rooms. II. The spatio-temporal inverse filter," J. Acoust. Soc. Am. Vol. 114, pp. 3044–3052, 2003.

[94] Available at: http://www.globalsecurity.org/military/facility/fort-benning.htm

[95] L. Liu, D. G. Albert, "Sound propagation in an urban environment. II: Preliminary FDTD model," J. Acoust. Soc. Am. Vol. 114, pp. 2442, 2003.

[96] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J. Thoms, and F. Wu, "Time-reversed Acoustics," Rep. Prog. Phys., Vol. 63, pp. 1933-1995, 2000.

[97] A. Sahin and E. L. Miller, "Performance Bounds for Matched Field Processing in Subsurface Object Detection Applications," SPIE Conference on Detection and

Remediation Technologies for Mines and Minelike Targets III, Vol. 3392, pp. 794-807, April 1998.