Microphone Array Beamforming with Near-field Correlated Sources

by

Jonathan Odom

Department of Electrical and Computer Engineering
Duke University

Date: April 21, 2009

Approved:

____
Jeffrey Krolik, Advisor

Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in the Department of Electrical and Computer Engineering in the Pratt School of Engineering of Duke University
2009

Abstract

Due to the traditionally high cost of large microphone arrays, the application of arrays for large room audio capture has not become widespread. However, arrays of many small omnidirectional microphones can now be efficiently developed and deployed. A single array could replace several wired and wireless microphones. The microphone would no longer have to be used for one person at a time or even be near a person. The large wavelength of speech places most applications in the near-field. Large room applications have loudspeakers, which create correlated sources and severely limit the use of optimum beamformers. Widrow and Kailath's work on correlation assumed far-field sources (often seen in sonar and radar) but little work can be directly applied to large room acoustics. A new beamforming method has been developed which incorporates near-field and correlation. The ability to control the loudspeaker signal is exploited. An estimate of the transfer function for the loudspeaker is used to form a new covariance matrix. A null can be placed in the array pattern even when signals are perfectly correlated.

Contents

List of Figures

List of Abbreviations and Symbols

Symbols

The symbols used through the document are listed below.

r_o	The range from the phase center of the array to the target.
r_s	The range from the phase center of the array to a steered location.
$\mathbf{x}$	The signal received vector.
$\mathbf{w}$	The steering vector.
$\mathbf{M}^{\mathrm{H}}$	The hermitian (conjugate transpose) of the matrix $\mathbf{M}$.
$\mathbf{y}$	The output of a beamformer.
$\mathbf{R}$	The covariance matrix.
$\mathbf{I}$	The identity matrix.

Abbreviations

MVDR	Minimum Variance Distionless Response defined by (2.9a).
iRv	Adaptive beamformer define by (2.10).
CC	Used in legends for cross correlation.

1
Introduction

1.1 Problem

Large room acoustics often require loudspeakers and a sound system to amplify a target, such as a person's voice. The common method for amplifying a target is to place a wireless microphone on a person's clothing or on a microphone stand. This causes problems with a wireless setup when batteries fail or clothing moves. Stand microphones can be wireless or wired; but, running wires creates another problem. Wires must be run through buildings, often during initial construction. Stand microphones have the extra problem that the distance between the microphone and person speaking can change. Lapel microphones can often be placed exactly by a trained technician. A technician does not have control over the exact location of a person talking and a stand microphone. People may stand far away enough that the system can no longer amplify a target without causing feedback. Stand microphones must also be moved with great caution since the microphone must be turned down first.

Most large and small room sound systems have several microphones which are treated as independent channels. Advanced systems have microphones placed in many locations throughout the room in order to electronically change the properties of the room in these systems [2]. Array processing is used to design loudspeaker systems. However, modern audio systems do not use array processing and have only recently being used in phone and laptop technology. Microphone arrays could be implemented to improve audio systems in large and small rooms.

A single microphone array could be placed meters away from several targets and beamforming would be used to collect selected signals. An example for a simple setup with a target and an interferer can be seen in Figure 1.1. The two sources would be spatially separated, but perfectly correlated. This is similar to a multipath problem where the multipath is stronger than the target. Optimum beamformers have severe performance degradation with correlated sources. Since speech has a wavelength on the order of meters, a large microphone array of a meter places all targets within the near-field.

1.2 Background

The use of microphone arrays is an up-and-coming field. *Microphone Arrays* by Brandstein and Ward explains the current development of microphone arrays for localization, noise reduction, and signal separation. They specifically mention the use of large arrays for large rooms, but believe that the design cost for such systems

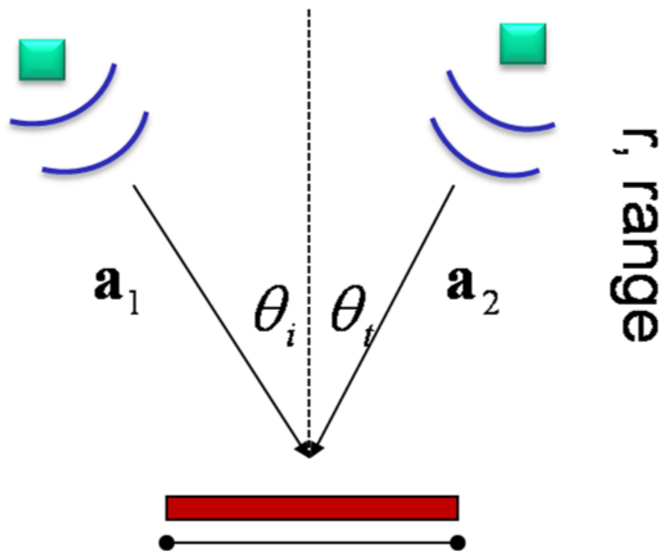

FIGURE 1.1: Overview of physical setup.

makes them impossible [3, p 393]. This research focuses on the ability to overcome the design cost problem with generalized signal processing.

Array processing as a field is well-developed for beamforming, which combines arrays of sensors into a single output. A summary of the most common beamforming methods was written by Van Veen and Buckley [14]. Research has been focused on statistically optimum beamformers. A high-resolution beamformer developed in 1969 by Capon, known as Minimum Variance Distorionless Response (MVDR), is discussed in Section 2.4. One of the largest problems with statistically optimum methods is performance degradations with correlated sources. Historically, the problem has been studied exclusively in the far-field. Only recently has the problem of correlated sources been analyzed in the near-field [7],[1].

In 1982, performance problems from correlated sources were described by Widrow as signal cancellation, which is a problem for adaptive processing in general. The problem arose from jammers that created perfectly correlated interference. The signal cancellation could be reduced by physically moving the array, called "spatial dithering." This is the basis for most of "decorrelation" research. [15]

Shan and Kailith, 1985, approached the problem with "spatial smoothing," which moves the array electronically instead of physically. The mathematical description is given in Section 2.5. Spatial invariance along the axis of the array for the far-field was used in the same way as Widrow; however, aperture length was reduced instead of mechanical movement [10]. The performance of the method for direction of arrival was also studied [11].

A few methods were developed which did not rely on spatial smoothing. Kesler *et al.*, 1985, used linear prediction in order to estimate the direction of arrival of the signals. The Burgs algorithm, Generalized Burgs algorithm, and iterative methods were all used for estimation and the results were compared [6]. Separately, a new beamformer was developed by Luthra, 1986. The signals are first pre-steered to the desired look direction. Then, the beamformer is steered to each of the of the source location and weights are calculated which null each of the sources except the desired source. The beamformer then performs beamspace beamforming using three beams in order to suppress the correlated signals further. The method provides nulls in the array pattern for correlated targets [8].

Spatial smoothing remained the most used method for separating correlated sources. Reedy *et al.*, 1987, studied the performance of MVDR and how spatial smoothing improved the method. A representation of the output power of MVDR and the specific effects of correlation were mathematically proven [9]. Spatial smoothing was also shown to be a toeplitzization of the covariance matrix by Takao *et. al.* [12]. The covariance matrix becomes more diagonal as signals are decorrelated. The work was extended to a more general form by Tasi in 1995. The array parameters relationship to the effect SINR was also formally studied. [13]. The consideration of mismatch sensitivity was included by Zoltowski, who suggested using total least squares to determine the weights [16]. Bresler *et al.*, 1988, provided a proof of spatial smoothing through signal subspace and parameter estimation. The proof allowed iterative methods to be used for estimating signal parameters. Three different optimum beamformers were used as examples to extend beyond MVDR: combiner, interference plus noise rejector, and interference notch [4].

Another beamforming method without spatial smoothing was developed by Godara, 1990. The beamformer has similarities to the one developed by Kesler. Two beamformers are used at once. One beamformer is dedicated to steering at a chosen target while another steers at interference signals. The output of the interference beamformer is then subtracted from the target beamformer. The SNR performance of this beamformer was unaffected by correlation.

2
Beamforming

Beamforming applies weights to a vector of data in order to seperate signals based on a property. All beamforming discussed here separates signals based on the spatial location of the source. Narrowband beamforming applies frequency-dependent weights to frequency filtered data. Here, signals are assumed to be at a single frequency and the dependency notation is suppressed for simplification.

2.1 Near-field Model

A near-field representation of a narrow-band signal is a spherical wave. A free-space wave can be represented by (2.1). The spherical wave is modeled by $r(\theta, d)$ as shown in (2.2). The range from a target to a sensor, given by r_o, is determined by the distance from a target to the phase sensor of the array, d, and the distance of the sensor from the phase sensor of the array, τ. A single range bin, which is a circle from the phase center of the array, can be selected by holding d constant and varying θ.

$$x(r_o) = e^{-jkr_o} \tag{2.1}$$

$$r_o = \sqrt{\left(d\sin(\frac{\pi}{2} - \theta)\right)^2 + \left(d\cos(\frac{\pi}{2} - \theta) - \tau\right)^2} \tag{2.2}$$

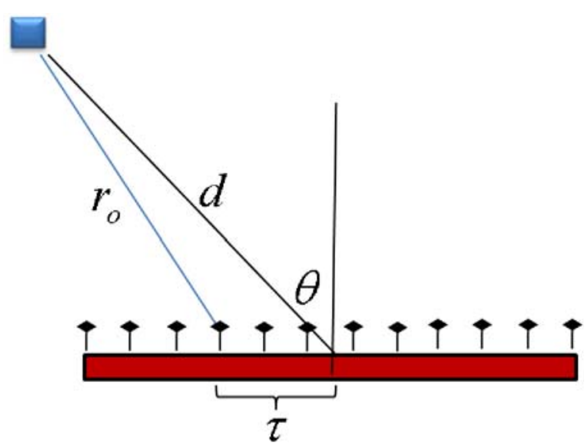

FIGURE 2.1: Geometry of Near-field

The signal which impinges on the array, $\mathbf{x}$, is shown by

$$\mathbf{x} = \begin{bmatrix} x_1(r_o) & x_2(r_o) & \cdots & x_{m-1}(r_o) & x_m(r_o) \end{bmatrix}^{\mathrm{T}} \tag{2.3}$$

2.2 Conventional Beamforming

The simplest form of beamforming is conventional, or delay-and-sum beamforming, which uses the propagation time from an assumed location and sums the channels up accordingly. The weights, $\mathbf{w}$, are unit vectors scaled by the number of sensors, m, for unity gain, which means a gain of 0 dB.

$$\mathbf{v} = \begin{bmatrix} x_1(r_s) & x_2(r_s) & \cdots & x_{m-1}(r_s) & x_m(r_s) \end{bmatrix}^{\mathrm{T}} \tag{2.4}$$

$$\mathbf{w}_{\text{conv}} = \frac{1}{m}\mathbf{v}, \text{ with } m \text{ sensors} \tag{2.5}$$

$$\boxed{\mathbf{y} = \mathbf{w}^{\mathrm{H}}\mathbf{x}} \tag{2.6}$$

Conventional and data-independent methods are not affected by correlation (discussed in Section 2.5). The array pattern is the output of a beamformer with fixed steering location, r_s, while varying $\mathbf{x}$. This describes how the output is affected by signals from different locations. A null in a direction means that the power is low for signals coming from that direction. A peak to zero in a direction means the signals from that direction are passed with unity gain. The spatial spectrum is the output of the beamformer by varying the steering location, r_s, with fixed $\mathbf{x}$. Conventional beamformers have -13 dB sidelobes, assuming rectangular windowing.

$$\text{Array Pattern: } \mathbf{y}(r_o) = \mathbf{w}^{\mathrm{H}} \cdot \mathbf{x}(r_o) \tag{2.7a}$$

$$\text{Spatial Spectrum: } \mathbf{y}(r_s) = \mathbf{w}^{\mathrm{H}}(r_s) \cdot \mathbf{x} \tag{2.7b}$$

2.3 Covariance Matrix

In order to understand the effects of multiple signals, a covariance matrix model is used. A two-signal model is formed in (2.8), where $\mathbf{v}_1$ refers to the steering vector to the 1st source. Optimum beamformers use the covariance matrix, $\mathbf{R}$, since it contains all the data needed for each source: power, direction and correlation. The eigenvectors of $\mathbf{R}$ are the steering vectors to targets, and the corresponding eigenvalues show the power of the signals. The statistical correlation between the two signals is p and increases the off-diagonal elements.

$$\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \tag{2.8a}$$

$$\mathbf{P} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 p \\ \sigma_1\sigma_2 p^* & \sigma_2^2 \end{bmatrix} \tag{2.8b}$$

$$\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I} = \mathrm{E}\left[\mathbf{x}\mathbf{x}^{\mathrm{H}}\right] \tag{2.8c}$$

2.4 Minimum Variance Distortionless Response

Statistically optimum beamformers are used to obtain specific constraints for a given problem. Minimum Variance Distortionless Response (MVDR) is one of the most common since it has unity gain in the direction of steering (2.9a). MVDR minimizes the power of the beamformer in directions outside the steered location. This places deep nulls at the location of other signals.

$$\min \mathbf{w}\mathbf{R}\mathbf{w}^{\mathrm{H}} \text{ subject to } \mathbf{w}^{\mathrm{H}}\mathbf{v} = 1 \tag{2.9a}$$

$$\mathbf{w}_{\mathrm{MVDR}} = \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{v}} \tag{2.9b}$$

The denominator of (2.9b), $\mathbf{v}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{v}$, is a normalization due to the constraint. The beamformer can also be used without the constraints, known as an adaptive beamformer. It is one of the simplest data-dependent beamformers and has many of the same properties of MVDR without having a unit constraint. Deep nulls are still placed in the direction of sources.

$$\mathbf{w}_{\mathrm{iRv}} = \mathbf{R}^{-1}\mathbf{v} \tag{2.10}$$

2.5 Correlation

As seen in Section 1.2, it has been well studied that correlation, p from (2.8b), causes many problems for beamformers based on covariance matrix inversion. In the uncorrelated case, the number of sources is the same as the number of dominant eigenvectors of the covariance matrix. Johnson and Dudgeon [5, p. 386] show that perfectly correlated signals, $|p| = 1$, of the same power, σ^2, result with $\mathbf{R}$ of the form shown in (2.11). The result is $m-1$ eigenvectors with eigenvalues of σ_n^2 due to noise and a single eigenvector with a eigenvalue of $m\sigma^2 + \sigma_n^2$ due to signal plus noise. The reduction of the eigenvector space means that the beamformer will not place nulls in the directions of the correlated signals. This is because the beamformer relies on the inverse of the eigenvectors to place nulls. When the eigenvector space collapses, all correlated signals are grouped into an indistinguishable signal. Therefore all signals will have a null if any of the signals are nulled, which is signal cancellation.

$$\mathbf{R} = \sigma_n^2\mathbf{I} + \sigma^2\left[\mathbf{v}_1 - \mathbf{v}_2\right]\left[\mathbf{v}_1 - \mathbf{v}_2\right]^{\mathrm{H}} \tag{2.11}$$

The power performance of MVDR was studied in depth by Reddy *et. al.* [9], who showed that correlated signals could not be separated by the beamformer. When the beamformer was steered at source 1, the power of the beamformer is given by (2.12), where $\beta = \left(\mathbf{v}_2^{\mathrm{H}}\mathbf{v}_1\right)$. This shows that there is a contribution of the correlated interferer determined by the interference pattern. It is interesting to notice that there will still be a contribution of the interferer even in the uncorrelated case.

$$\mathbf{P}_{\text{opt}} = \sigma_1^2 + \frac{1}{\sigma_2^2(m^2-\beta^*\beta)+m\sigma^2} \left(\sigma_1^2\sigma_2^2\left(\beta^*\beta - m^2\right) + \sigma_{\mathrm{N}}^2\left(\frac{\sigma_2^2\beta^*\beta}{m} + \sigma_1\sigma_2\left(\rho\beta+\rho^*\beta^*\right)\right)\right) + \frac{\sigma^2}{m} \quad (2.12)$$

The solution for the far-field case was spatial smoothing, first analyzed as spatial dithering. It averages along the diagonal of the covariance matrix as shown in Figure 2.2. The averaging assumes a plane-wave since each source will have the same direction of arrival for any section of the array. However, this does not hold in the near-field.

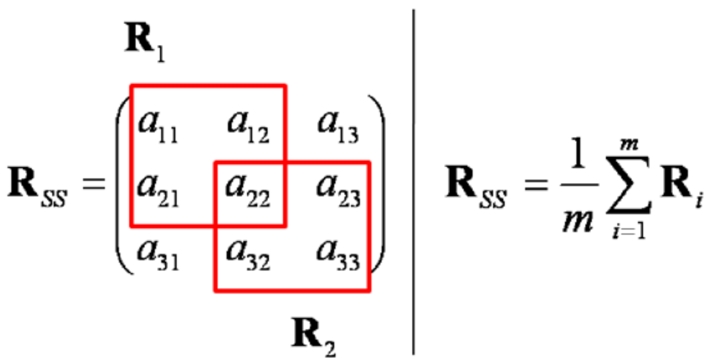

FIGURE 2.2: Spatial smoothing is a block diagonal averaging.

Specifically, averaging reduces the correlation by combining spatially separated parts of the array. The correlation of the sources appears different for different spatial sections. This can be seen in the p terms of the bth and dth sensor term of $\mathbf{R}_{b,d}$ in (2.13). Note that $r_b(\alpha_1^o)$ represents the distance from the bth sensor to the target α_1^o target location. The first two terms do not change from sensor to sensor in a uniform linear array, but the second two terms vary and are reduced by the averaging.

$$\mathbf{R}_{b,d} = \left(e^{-jk\left(r_b(\alpha_1^o)-r_d(\alpha_1^o)\right)} + e^{-jk\left(r_b(\alpha_2^o)-r_d(\alpha_2^o)\right)}\right) + \ldots \\ + p^* e^{-jk\left(r_b(\alpha_2^o)-r_d(\alpha_1^o)\right)} + p e^{-jk\left(r_b(\alpha_1^o)-r_d(\alpha_2^o)\right)} \quad (2.13)$$

In the near-field, the first two terms will also vary between sensors. This results in a spread of sources increased by the arc of the near-field. The sources are still decorrelated since the p terms still vary more in the near-field. The loss of aperture due to the spatial smoothing also causes loss of resolution.

3
Two Proposed Methods

The first proposed method fits the data to an uncorrelated model and then applies the weights accordingly. The performance degradation is caused by the rank reduction of the covariance matrix, so using a covariance matrix with uncorrelated data will not cause signal cancellation. The second method measures the impulse response directly instead of fitting to a model, in an effect to circumvent the problems of the model itself.

3.1 Model Based MVDR

It is assumed that there are a known number of strong signals. A beamformer can be used to estimate the steering vector of the signals. This can be done by beamforming over different locations and using the top p peaks, where p is the number of strong signals. These peaks correspond to steering vectors from the beamformer. A new model covariance matrix is formed by the p steering vectors. The model covariance matrix formed assumes the signals are uncorrelated. Adaptive beamforming is then used with the model covariance matrix on the correlated data and places nulls without the rank reduction.

3.1.1 Model Based MVDR simulation

This is demonstrated at a single range bin with two sources. Both sources are perfectly correlated and are narrowband with a frequency of 800 Hz. The array is a 64 element uniform linear array (ULA) with 0.02 m element spacing of length 1.26 m. The target appears at -0.5 rad with an SNR of 30 dB. The interferer appears at broadside with an SNR of 40 dB. Only the louder source needs to be estimated since only the loudspeaker or interferer needs to be nulled in this problem. A conventional beamformer is used to estimate the location of the source by sweeping through angles at a resolution of 1 degree. The maximum is the steering vector for the interfering source. When forming the model covariance matrix, an SNR of 60 dB is assumed. This is relevant to the diagonal loading of the model covariance matrix and has little impact on the location of the null. However, the depth of the null is related to the amount of diagonal loading. Any reasonable SNR level will give a deep null, larger than 50 dB.

The performance of the beamformer is shown by the array pattern in the second half of Figure 3.1. A deep null can be clearly seen at broadside, while the rest of the array pattern is similar to a conventional beamformer. The beamformer output, or spatial spectrum, shows the peak on the interferer. This is the usual behavior of

MVDR when steering at a source in the covariance matrix. A conventional pattern is found at the target at 30^o since it is not in the covariance matrix.

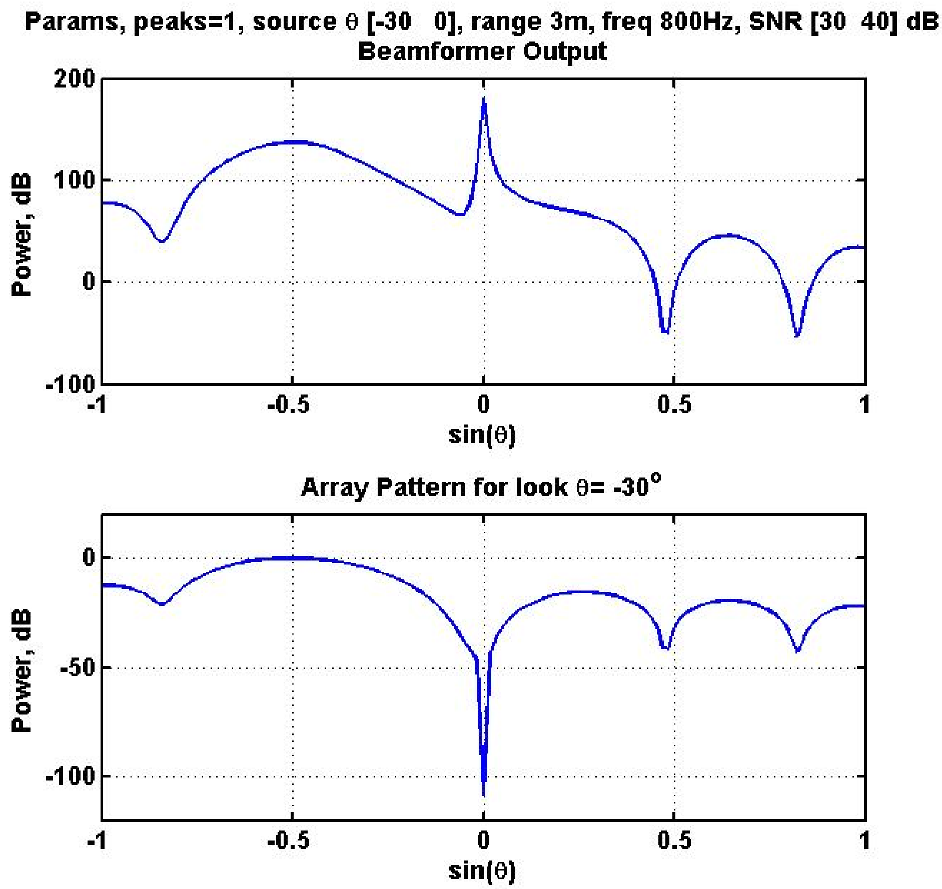

FIGURE 3.1: Model MVDR with two correlated sources

However, this beamforming technique is limited by the ability to model the signals. A free space model, unfortunately, is not sufficient for sound waves in a real room. Experiments show that conventional beamforming was robust enough to work in the lab room. However, MVDR or other optimum beamformers are not robust enough to work when multiple sources are present. The reflections from the walls and objects in the room are not given in the model and cause MVDR beamformers to fail. Given a situation when the impulse response for each source is known, a model-based approach would work. The simulation shows the correctness of the method. For a loudspeaker system, it is unrealistic to be able to model a room. Instead, a different approach is taken to determine the impulse response experimentally and adaptively.

3.2 Pseudo-Random Noise

In audio systems, the signal sent to the loudspeaker can be controlled directly. Instead of first decorrelating, or reducing ρ, a signal-free estimate could be made if the target power was reduced. This is the same as estimating the impulse response of the loudspeaker. This could be measured continuously using a known pseudo-random sequence and cross-correlation.

3.2.1 Theory

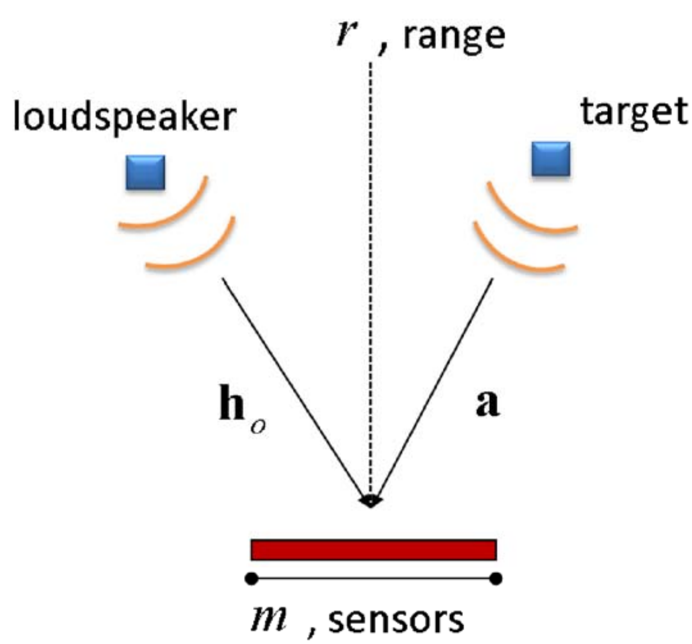

FIGURE 3.2: Explanation of physical setup.

The pseudo-random noise can be added to the output of any number of speakers. This sequence must be stored to be used later in the cross-correlation. Generally, white noise is Gaussian with a mean of 0 and the variance determines the power of the noise. The pseudo-random sequence will be a Gaussian so that the noise will appear to be part of the environment. However, the pseudo-random noise can be distinguished from other noises by the fact that it comes from a single location, the loudspeaker. The goal will be to limit the power of the pseudo-random signal so it will not be noticed by human listeners. The beginning of the problem is approached as usual. The target and loudspeaker signals without the pseudo-random noise will be modeled in the Cholesky decomposition of the covariance matrix.

$$\mathbf{R_x} = \mathrm{E}\left[\mathbf{x}\mathbf{x}^{\mathrm{H}}\right] \tag{3.1a}$$

$$\mathbf{R_x} = \mathbf{C}^{\mathrm{H}}\mathbf{C} \tag{3.1b}$$

The lth snapshot of the received signal, with the original signals and the pseudo-random sequence, are contained by $\tilde{\mathbf{x}}_l$. The snapshots can be cross-correlated with the original pseudo-random noise sequence, which in the frequency domain is multiplication by the complex conjugate.

The actual data samples with noise can be simulated by multiplying $\mathbf{C}^{\mathrm{H}}$ by a complex Gaussian sequence. Note that the magnitude of the variance must be 1 for the complex case. The pseudo-random sequence is the same, except the power of the signal is the variance, not 1. The full signal, $\tilde{\mathbf{x}}$, can be modeled by addition since the pseudo-random sequence is uncorrelated with the other signals.

$$\tilde{\mathbf{x}}_l = \mathbf{h}_o s_l + \mathbf{C}^{\mathrm{H}}\ \mathbf{e}_l \tag{3.2}$$

Now the cross-correlation is used to reduce signals which do not appear in both terms. With enough snapshots, the cross-correlation will reduce all signals below the noise level except for the pseudo-random noise.

$$\begin{aligned}
\mathbf{y} &= \frac{1}{L}\sum_{l}^{L}(\tilde{\mathbf{x}}_l)s_l^* \\
&= \frac{1}{L}\sum_{l}^{L}\left(\mathbf{h}_o s_l + \mathbf{C}^{\mathrm{H}}\mathbf{e}_l\right)s_l^* \\
&= \frac{1}{L}\sum_{l}^{L}(\mathbf{h}_o|s_l|) + \frac{1}{L}\sum_{l}^{L}\left(\mathbf{C}^{\mathrm{H}}\mathbf{e}_l s_l^*\right) \\
&= \mathbf{h}_o\frac{1}{L}\sum_{l}^{L}|s_l| + \mathbf{C}^{\mathrm{H}}\frac{1}{L}\sum_{l}^{L}\mathbf{e}_l s_l^* \\
&= \mathbf{h}_o\sigma_{\mathrm{PN}}^2
\end{aligned} \tag{3.3}$$

The cross-correlation will give an estimate of the impulse response of the interferer scaled by the power of the pseudo-random noise. The second half of (3.3) shows the sum of $\mathbf{e}_l s_l^*$, which is a vector product of Gaussian signals. Each element of the vector will have a mean of zero. This is what causes the reduction of the power of the interfering signals.

3.2.2 Simulation

The simplest sound system would have a single person talking and a single loudspeaker. Since the placement in a real situation would vary dramatically, a situation was chosen based on the size of the laboratory room for the experiment. Due to the expense of collecting real data from an auditorium, the levels of signals in a room were determined by computer modeling software. This is a typical practice in the audio industry [2, p. 216]. ODEON room acoustics software provides a model of an auditorium at Technical University of Denmark. The software allows for sources and receivers to be placed in the room and then for acoustic properties to be calculated. A model for a person talking and a typical Yamaha loudspeaker are included with the software. An omni-directional receiver is placed in the middle of the room. Typical large rooms can have a maximum noise level of 30 dB(A) before the noise level causes problems [2, p. 68]. It was determined for the given room that a worst-case SNR would be 30 dB for the loudspeaker and 20 dB for the person talking.

A setup from Figure 3.2 is used. The target, or person talking, is at 0.2 rad and the interferer is at -0.4 rad. The interferer location was chosen so that the null would be placed near the conventional sidelobe. Since it is not likely that the signals would arrive at the same time, due to processing of the sound system, a delay time of 1e-5 seconds is used to simulate the correlated data.

The data is simulated using the method previously described in Section 3.2.1, using the Cholesky decomposition of the covariance matrix. The number of snapshots is 1292, which is one minute of data collected at 22050 Hz using a non-overlapping FFT with length 1024. The array pattern is an average of 100 independent runs. The pseudo-random sequence is a Gaussian signal.

The array pattern is shown in Figure 3.2.2 with arrows pointing in the location of sources. The CC-iRv are weights from the cross-correlated data, $\mathbf{y}$, from (3.3). The MVDR weights are derived from the received signal, $\mathbf{x}$ without any pseudo-random noise added. Thus, the MVDR is the standard method of applying optimum weights with unit gain. It can be seen in Figure 3.2.2 that MVDR has a -10 dB gain in the direction of the interferer. The gain in the array pattern from the interferer for MVDR will be the interference to signal ratio. However, this is irrelevant for MVDR since the SNR and SINR go to zero, which was shown by Reddy et al [9] and Tasi et al [13]. The proposed method solves this problem by averaging out the correlation to below the noise level.

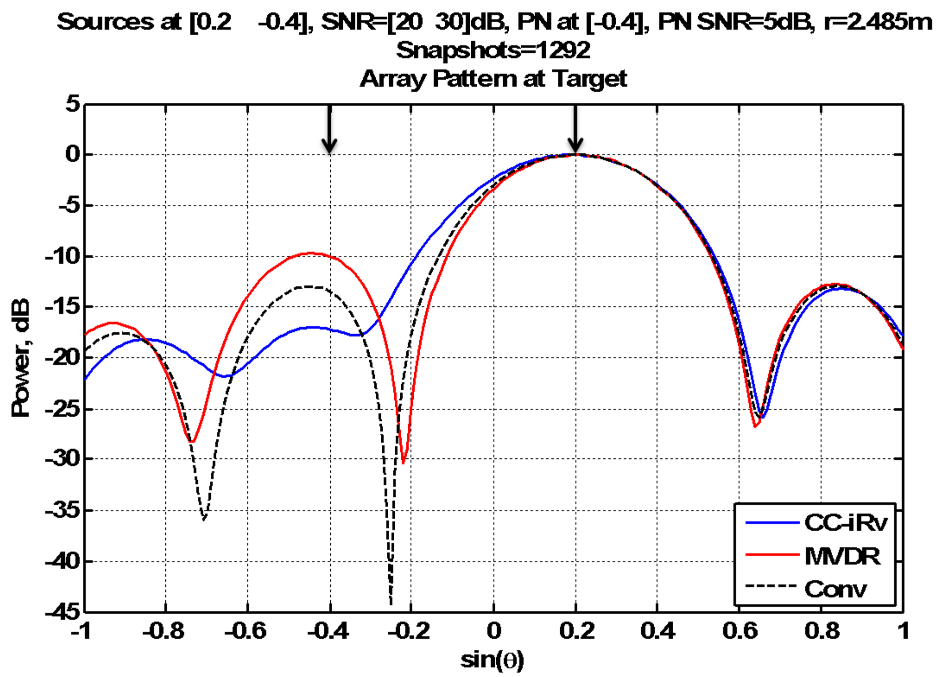

FIGURE 3.3: Simulation of array pattern for new method with perfectly correlated sources. Target at 0.2 rad with SNR 20dB, interferer at -0.4 rad with SNR 30 dB, pseudo-random noise at -0.4 rad with SNR 5 dB

4
Experiment

4.1 Microphone Array

An experiment was run in order to verify the simulations of Section 3.2.2. The NIST Mark-III Version 2 is used for all data collection, shown in Figure 4.1. Each board of the array is 8 microphones with on-board A/D converters synchronized by a single motherboard. The recording is sampled at 22050 Hz via wired ethernet connections through a TL-WR340GD router acting as a switch. Data is collected in one file, which is then processed after the experiment. The data collection computer is running Linux with the mk3 capture programs written by NIST. The data is recorded as a 24-bit Big Indian stream. All experiments have static sources. The microphone array is placed in the center of the room in order to reduce the effects of walls objects in the room.

The setup is chosen so that the room could be easily setup for sources at -20^o and 20^o; this places the targets at a range of 2.485m. This is well within the near-field of 800 Hz signal to be used. The relative SNR levels of the experiment are used from the simulation, based on the computer room modeling software. The SNR levels of the individual microphones elements ranged from 30 to 50 dB. The sources are Harman Kardon computer speakers using a SoundBlaster 7 channel sound card. Matlab is used to generate the sound sources. The computer speakers and receivers have a flat response around the 100-10,000 Hz range.

The signal plays continuously for one minute in the room. The transient response of the room at the start of the signal is ignored, since it would be averaged out in a real system.

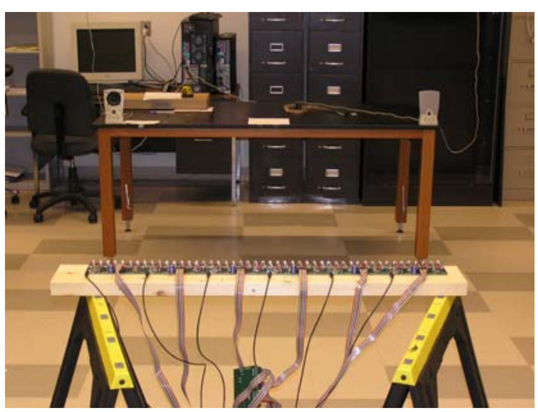

FIGURE 4.1: NIST Mark-III Version 2 microphone array setup for experiment

Array pattern steered at 20^o shows the ability to null a strong correlated source, as seen in Figure 4.2. Several experiments were run using the same pseudo random sequence and similar results were obtained, confirming that the results were similar to the simulations. The null does vary between experiments as with the simulations: results of -18 dB were obtained in a second run.

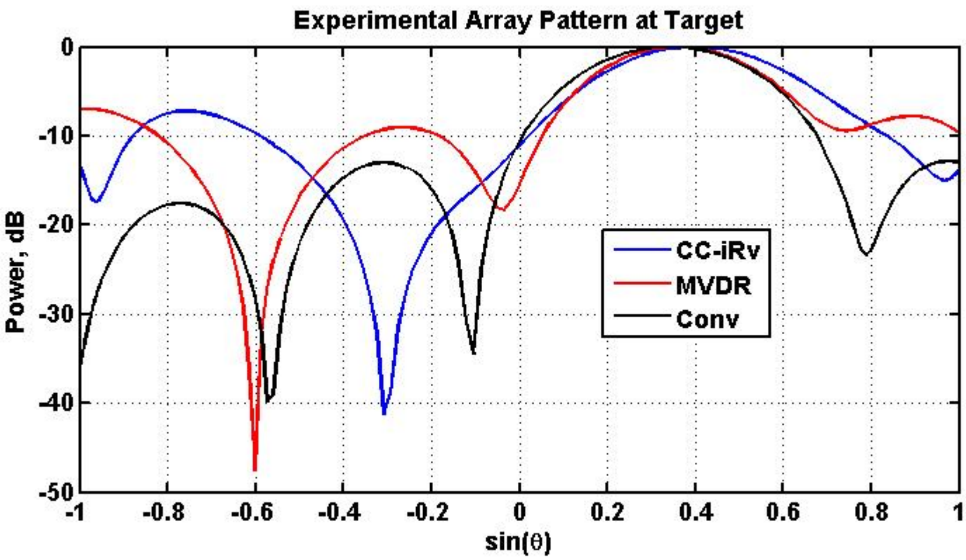

Figure 4.2: Array pattern for experimental results.

The deepest point of the null appears at $-.309$ rad, or -18^o. The source is placed at -0.342 rad, -20, which is contained within the null. This shows that the method works very well in real world conditions. The inability of the simulation model to capture the acoustic response of the room does not hinder the method. The impulse response of the loudspeaker is correctly estimated and then eliminated from the signal. Since the pseudo-random sequence is not narrow-band, it could applied to any frequency bin with no additional overhead other than applying the weights to the frequency bin.

5
Conclusions

The simulation shows that both proposed methods work. A model-based MVDR beamformer works in a known environment. This verifies that weights from an uncorrelated model can be applied to correlated data. The performance degradation from MVDR is not inherent in the data but only in the weight computation. Preprocessing methods such as spatial smoothing also rely on this property. However, previous methods assumed far-field sources and do not work well in the near-field. Also, previous methods had an aperture reduction trade-off and reduced target resolution. A model-based method would be limited by the computation cost of the model itself. The overhead involved in the simulation was conventional beamforming through the angular resolution of 1 degree. The real problem with a model-based approach is the need for an accurate model. Knowledge of the impulse response of a room can not yet be obtained adaptively. This limits the performance or capabilities of any model-based approach. For any room, people entering and leaving can dramatically change the room response. Large-room beamformer techniques should include methods of measuring the this response.

The second method attempts to solve the impulse response problem. Unlike many beamforming applications, sound systems have a large degree of control over the signals traveling in a room. Exploiting access to the signal sent to the loudspeaker by inserting pseudo-random noise is not a method that can be used in general. This approach only solves the specific problem for sound systems. However, it does provide a way of accurately and adaptively measuring the impulse response of the room. Current systems rely on one-time measurements at installation or specific equipment only used for measuring room response. The actual level of the pseudo-noise required for varying levels of target and loudspeaker SNR has not been fully developed. However, the simulation and experiment do show that there is a possibility of using low levels of pseudo-random noise. The ability to suppress the correlated signal creates many new possibilities for large arrays. The experimental results verify that the system could work without exceedingly sophisticated hardware or room design. It shows that dramatic signal reduction based on spatial separation can be obtained without modifying a room or even including a known model of the room. This proves that with more research, a large room microphone array could be developed without expensive or tedious room-specific design.

5.1 Future Work

The relation of pseudo-random noise SNR to the target and loudspeaker SNR needs to be analyzed. The final metrics for solving this problem should be the effective

SNR for a listener in the room and the maximum gain that can be applied to a sound system before feedback causes the system to fail. Ideally, the experiment would be conducted in a large room, such as an auditorium, using broadband signals. All research can be theoretically expanded to include broadband signals such as speech and music. Also, the final result would be a real-time system. The field of arrays in large rooms has not been fully developed because arrays in the past have been too expensive. With the introduction of arrays such as the one developed by NIST, research can be conducted in parallel with experiments.

Acknowledgements

I would like to thank my advisor, Dr. Jeffrey Krolik, for guiding my research and treating me as a graduate student for the last year and a half. Also, the members of the lab have shown me how research should be conducted and convinced me to stay on for another two degrees. I would like to thank my family for supporting my work even when I could not explain clearly the purpose of the work. To Ashley DeMass, I would like express the deepest thanks for constantly supporting me in whatever I do.

Bibliography

[1] Monika Agrawal, Richard Abrahamsson, and Per Ahgren. Optimum beamforming for a nearfield source in signal-correlated interferences. *Signal Processing*, 86(5):915 – 923, 2006.

[2] Glen Ballou. *Handbook for sound engineers*. Oxford : Focal, c2008.

[3] M. Brandstein and D. Ward, editors. *Microphone arrays : signal processing techniques and applications*. New York : Springer, c2001.

[4] Y. Bresler, V.U. Reddy, and T. Kailath. Optimum beamforming for coherent signal and interferences. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(6):833–843, Jun 1988.

[5] Don H. Johnson and Dan E. Dudgeon. *Array signal processing : concepts and techniques / Don H. Johnson, Dan E. Dudgeon*. P T R Prentice Hall, Englewood Cliffs, NJ :, 1993.

[6] S. Kesler, S. Boodaghians, and J. Kesler. Resolving uncorrelated and correlated sources by linear prediction. *Antennas and Propagation, IEEE Transactions on*, 33(11):1221–1227, Nov 1985.

[7] Ju-Hong Lee, Yih-Min Chen, and Chien-Chung Yeh. A covariance approximation method for near-field direction-finding using a uniform linear array. *Signal Processing, IEEE Transactions on*, 43(5):1293–1298, May 1995.

[8] A. Luthra. A solution to the adaptive nulling problem with a look-direction constraint in the presence of coherent jammers. *Antennas and Propagation, IEEE Transactions on*, 34(5):702–710, May 1986.

[9] V. Reddy, A. Paulraj, and T. Kailath. Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 35(7):927–936, Jul 1987.

[10] Tie-Jun Shan and T. Kailath. Adaptive beamforming for coherent signals and interference. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 33(3):527–536, Jun 1985.

[11] Tie-Jun Shan, M. Wax, and T. Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 33(4):806–811, Aug 1985.

[12] K. Takao, N. Kikuma, and T. Yano. Toeplitzization of correlation matrix in multipath environment. volume 11, pages 1873–1876, Apr 1986.

[13] Churng-Jou Tsai, Jar-Ferr Yang, and Tsung-Hau Shiu. Performance analyses of beamformers using effective sinr on array parameters. *Signal Processing, IEEE Transactions on*, 43(1):300–303, Jan 1995.

[14] B.D. Van Veen and K.M. Buckley. Beamforming: a versatile approach to spatial filtering. *ASSP Magazine, IEEE*, 5(2):4–24, April 1988.

[15] B. Widrow, K. Duvall, R. Gooch, and W. Newman. Signal cancellation phenomena in adaptive antennas: Causes and cures. *Antennas and Propagation, IEEE Transactions on*, 30(3):469–478, May 1982.

[16] M.D. Zoltowski. On the performance analysis of the mvdr beamformer in the presence of correlated interference. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(6):945–947, Jun 1988.