Approved for public release; distribution is unlimited

Acoustic Sensor Network Self-Localization: Experimental Results

J.N. Ash and R.L. Moses

Department of Electrical Engineering, The Ohio State University
2015 Neil Avenue, Columbus, OH 43210 USA

Abstract

In this paper we present experimental results on propagation, coherence, and time delay estimation (TDE) in an acoustic microphone array. The primary goal is to understand the achievable accuracy of acoustic TDE using low-cost, commercial off the shelf (COTS) speakers and microphones. In addition, the experiment seeks to provide an empirical understanding of the effects of center frequency, bandwidth, and signal duration on TDE effectiveness. As part of this study, limited empirical signal propagation characteristics were obtained using this COTS equipment.

1. Introduction

Sensor networks are emerging in a large number of civilian and military applications to sense and process information about their surroundings [1]. To make use of this information the network needs to know the locations of the individual sensors. However, in many scenarios manual assignment of sensor locations is impossible or impractical due to the volume of sensors deployed or the placement method. Therefore, the problem of self-localizing sensor networks becomes increasingly important.

There are a number of techniques for self-localization of sensors from time-of-arrival, time-difference-of-arrival, direction-of-arrival, or received signal signal strength measurements; see, e.g. [2]. Most of these techniques employ TOA or TDOA measurements, perhaps in conjunction with other measurements. The localization accuracy thus depends on the accuracy with which one can estimate TOA or TDOA. In this paper we explore, through statistical bounds and outdoor experimentation, the limits on the mean square error of TDOA estimates for an acoustic sensor network.

The remainder of this paper is organized as follows. Section 2 describes our experimental procedure, including hardware setup and signal generation. In Section 3 we present empirical results characterizing the attenuation and coherence loss of the acoustic channel. Section 4 contains the results of our empirical study of the effects of center frequency, bandwidth, and signal duration on

Appears in Proceedings of Military Sensing Symposia (MSS) Specialty Group on Battlefield Acoustic and Seismic Sensing, 2003

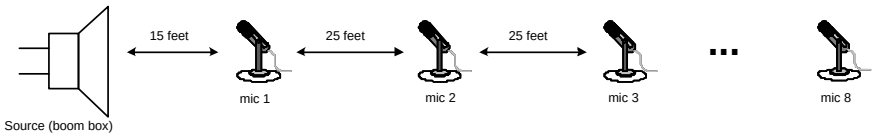

Figure 1: Linear array used in field measurements.

time delay estimation accuracy. In Section 5 we present the results of self-calibrating an outdoor sensor network using PN sequences. Finally, in Section 6 we conclude.

2. Experimental Procedure

As depicted in Figure 1 a linear array of eight Knowles BL-1994 microphones, each separated by 25 feet, was used in this experiment. The 100 feet of coaxial cable available for each microphone yielded maximum span of 200 feet with hardware positioned near the center of the linear array. Each microphone was also equipped with a spherical windscreen with 3" radius. The Knowles microphones were connected via the coaxial cables to an 8-channel National Instruments analog signal conditioner followed by an analog-to-digital converter. The signal conditioning consisted of amplification and anti-alias filtering the signals to 4kHz. The microphone signals were sampled at 12kHz per channel and 12 bits per sample.

The low-cost COTS sound source consisted of a portable stereo (boom box) playing a CD of prerecorded PN sequences. The boom box was co-linear with the array and located 15 feet from the first microphone as shown in Figure 1.

The PN sequences were generated as maximum-length shift-register sequences using an m-stage shift register with linear feedback[3]. The sequences, which have length $n = 2^m - 1$, were chosen for their nearly ideal periodic auto-correlation

$$R(k) = \begin{cases} n & (k = 0) \\ -1 & (1 \leq k \leq n-1). \end{cases} \quad (1)$$

Our application uses a non-periodic correlation, however the correlation values at non-zero lags are still small compared to the strong peak at zero lag. The sequences were designed to cover a wide range of center frequencies, bandwidths, and durations. The center frequencies ranged from 100Hz to 2000Hz, bandwidths ranged from 3Hz to 3200Hz, and total signal durations ranged from 0.2s to 10s. In total, 319 PN waveforms were used in this experiment. The PN sequences were generated in MATLAB, modulated to different center frequencies, exported as a .wav files, and then copied to an audio CD.

The microphone array was set up in a flat, grassy field on the aftenoon of June 16, 2003. The temperature was 74°F, the relative humidity was approximately 75%, and the wind was light. First, background noise was recorded to establish the noise spectrum of the microphone outputs. Then the PN sequences were played and recorded to study empirical propagation and time delay estimation. Finally, the array was reconfigured in a nonlinear fashion and a small subset of the PN

sequences were played from various locations to serve as calibration signals for a self-localization experiment.

We note that post analysis of the data indicated that measurements from microphones 2 and 8 were corrupt — likely due to hardware failures. Thus, these microphone signals have been omitted from the empirical results below.

3. Empirical Propagation

In this section we present the results of experiments that were designed to give us an empirical understanding of the acoustic channel over short distances and moderate bandwidths. The aim of this section is not to develop a precise model of acoustic propagation, but to empirically study the trends of parameters that are important to time delay estimation. In particular, we evaluate signal attenuation and signal coherence as a function of distance.

3.1 Noise and Attenuation

Figure 2 presents the observed power spectral density of the background noise observed at microphones 1, 3, and 7 (chosen arbitrarily for illustration). The nearly perfect overlap of the noise PSDs indicates similar noise levels at each position and similar frequency responses from each microphone. The drop at 4000Hz is due to the anti-aliasing filter in the signal conditioning hardware. The source of the peak at 2000Hz is unknown.

In Figure 3 we present the observed PSDs for received signals from a high bandwidth PN sequence ($F_c = 2000$Hz, $BW = 3200$Hz). Because the sound source was uncalibrated, the exact PSD of the transmitted signal was unknown. Thus, the attenuation loss to the first microphone could not be determined. There are, however, several trends that can be noted from the figure. As expected, the signal experiences greater attenuation as the source-receiver distance is increased. There also appear to be nulls in the responses that occur at different frequencies depending on the location of the microphone. Microphone 1 experiences a null at approximately 1800Hz, Microphone 2 at 1300Hz, and Microphone 3 at 900Hz. We hypothesize that these location-dependent nulls are due to destructive interference from ground reflections (see, e.g. [7]). Beyond each null, the response is relatively flat until the cutoff frequency of 4kHz.

By examining the difference in signal power and the average noise power in Figure 3 we can investigate the SNR as a function of frequency and distance. For example, at 2500Hz, microphone signal 7 (which is 165 feet from the sound source) has an SNR of approximately 15 dB in this experiment. All of the SNRs are observed to be greater than 10 dB except near the null of microphone 7, and near the 2000Hz noise spike for microphones 3 and 7.

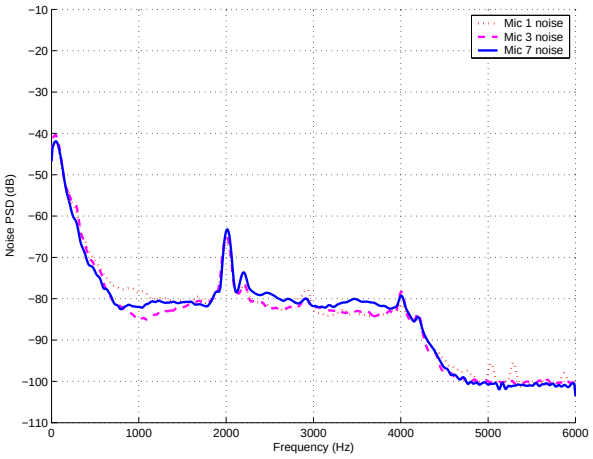

Figure 2: Power spectral density of background noise as observed at microphones 1, 3, and 7. The nearly perfect overlap indicates similar noise levels and microphone responses.

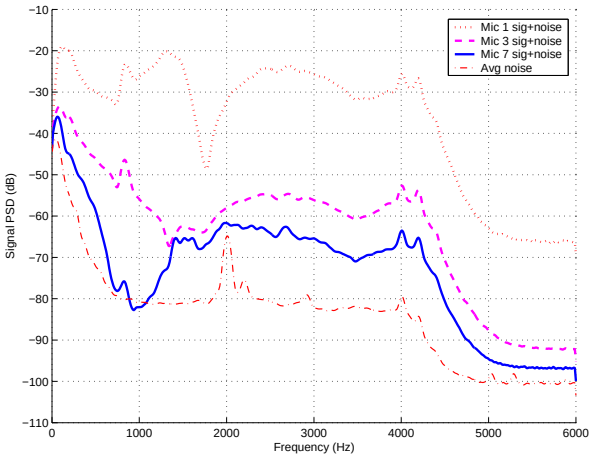

Figure 3: Observed PSD's of a high bandwidth PN sequence: $F_c = 2000Hz, BW = 3200Hz$. The average of the three noise floors of Figure 2 is plotted as well for comparison.

3.2 Coherence

Figure 4 presents our empirical observations of signal coherence as a function of microphone separation, center frequency, and bandwidth. For each received waveform, the coherence was estimated between the received signal at microphone 1 and microphones 3, 4, 5, 6, and 7. The coherence reported is the maximum value of the magnitude squared coherence (MSC) over the bandwidth of the signal. The expected trend of reduced coherence with distance is observed for all frequencies. However, we also observe a substantial decrease in coherence for $F_c = 1600$Hz and $F_c = 2000$Hz. The loss in coherence at $F_c = 2000$Hz may be caused in part by the noise spike at the same frequency. Therefore, from these observations alone, we cannot say for certain that the coherence above 1600Hz is poor due solely to loss of phase coherence.

The impact of coherence loss on time delay estimation will be illustrated in the following section.

4. Time Delay Estimation

4.1 Background

The time delay estimation (TDE) problem is to estimate the time difference of arrival between the received signals at two distant receivers. For a transmitted signal $s(t)$ we have the following received signal model:

$$\begin{aligned} r_1(t) &= h_1(t) * s(t) + n_1(t) \\ r_2(t) &= h_2(t) * s(t-d) + n_2(t), \end{aligned} \tag{2}$$

where, $*$ denotes convolution, d is the delay to be estimated, $h_i(t)$ is channel impulse response from source to node i, and $n_i(t)$ is additive noise that is assumed to be uncorrelated with $s(t)$ and other

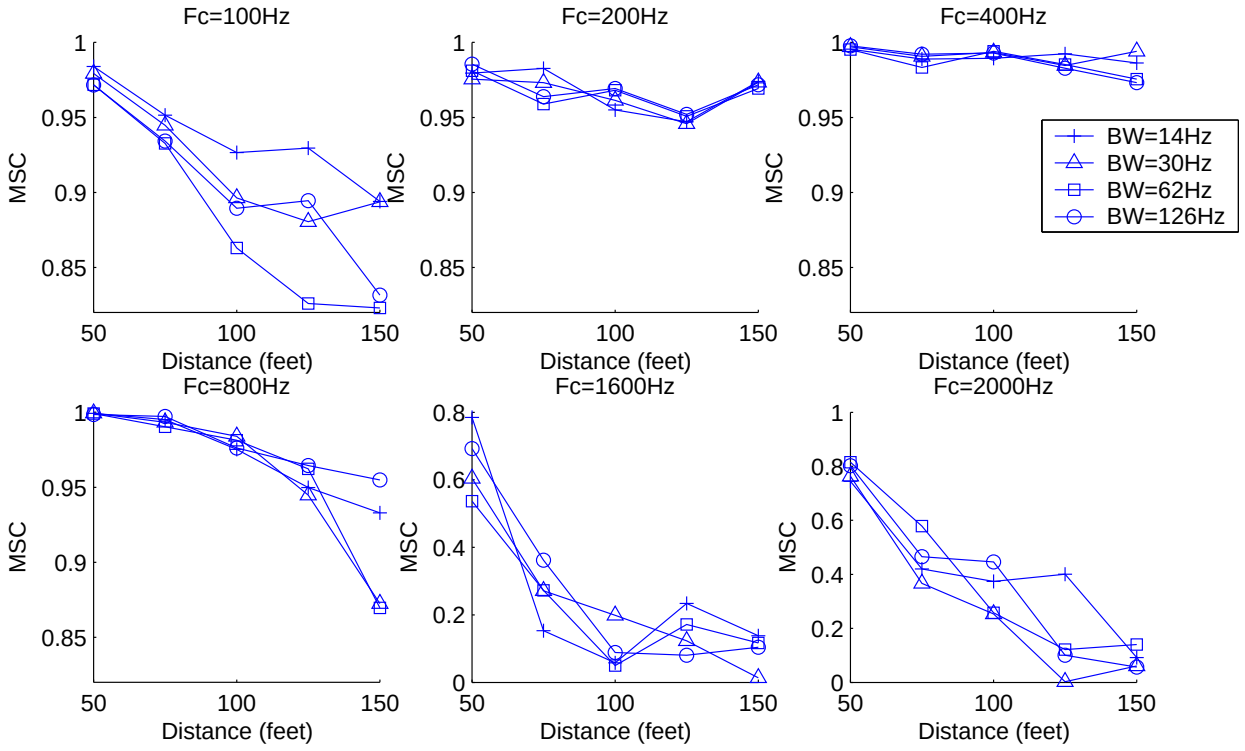

Figure 4: Mean squared coherence (MSC) as a function of microphone separation, center frequency, and bandwidth. The signal duration is $T = 1$s in all cases. Signal coherence is seen to be low for $F_c = 1600$Hz and $F_c = 2000$Hz.

with $n_j(t)$ for $j \neq i$. As the distance from source to receiver is increased, $h_i(t)$ is expected to apply greater attenuations to $s(t)$.

In Figure 5 we present the results of a simulation that illustrates key trends in the TDE problem. In implementating the simulation we used the signal model in (2) and imposed a random artificial delay. The channel responses were estimated from propagation models given in [7]. From Figure 5 we observe that as the sensor separation increases the SNR decreases accordingly. At a certain distance, about 34m in Figure 5, the SNR drops below the *threshold* signal to noise ratio, SNR_{th}, and the error dramatically increases. This is due to the well known threshold effect in time delay estimation [8, 11]. When the SNR falls below SNR_{th} the dominant peak of the cross-correlation function can no longer be reliably identified by the estimator. This peak ambiguity results in high estimation error.

Next we consider means of statistically bounding the TDE error for certain signal and noise parameters. The Cramér-Rao lower bound (CRLB) would be the usual tool of choice, however the CRLB is a local bound that is only tight for high SNR regions. An alternative bound that is tighter over all SNRs is the Weiss-Weinstein lower bound (WWLB) given in [9]. In [10] the WWLB was derived for the time delay estimation problem when the unknown delay is assumed to have a uniform prior distribution over $[-D/2, D/2]$. The variance of the time delay estimate is bounded by

$$\bar{\epsilon}^2 \geq \max_{0<h<D} J(h) \tag{3}$$

where $J(h)$ is given by

$$J(h) = \begin{cases} \frac{\frac{1}{2}h^2(1-h/D)^2 e^{-R(1-\rho(h))/2}}{1-h/D-(1-2h/D)e^{-R(1-\rho(2h))/4}}, & 0 \le h < D/2 \\ \frac{1}{2}h^2(1-h/D)e^{-R(1-\rho(h))/2}, & D/2 \le h < D, \end{cases} \tag{4}$$

and where $\rho(h)$ is the source auto-correlation function and $R = 2E/N_0$ the so called post-integration SNR. When SNR $>$ SNR$_{th}$ there is no peak ambiguity problem and the WWLB reduces to the CRLB [8].

Also from [8], the value of SNR$_{th}$ can be calculated from the source's time-bandwidth product WT and bandwidth to center frequency ratio W/ω_0:

$$\text{SNR}_{th} = \frac{6}{\pi^2(WT/2\pi)} \left(\frac{\omega_0}{W}\right)^2 \left[\phi^{-1}\left(\frac{W^2}{24\omega_0^2}\right)\right]^2, \tag{5}$$

where $\phi(y) = 1/\sqrt{2\pi}\int_y^\infty e^{-t^2/2}dt$.

The bound given by (3) is still an optimistic bound however, because it assumes the received signals are fully coherent. Similarly, even for high SNRs, the CRLB is overly optimistic if the received signals are only partially coherent. To account for this, in [5] Kozick and Sadler derive an effective SNR based on the coherent signal component of the of the received signal,

$$\begin{aligned} \text{SNR}_1 &= \frac{|\gamma_{r,12}(f_c)|G_{r,11}(f_c)}{G_n(f_c) + (1-|\gamma_{r,12}(f_c)|)G_{r,11}(f_c)} \\ \text{SNR}_2 &= \frac{|\gamma_{r,12}(f_c)|G_{r,22}(f_c)}{G_n(f_c) + (1-|\gamma_{r,12}(f_c)|)G_{r,22}(f_c)}, \end{aligned} \tag{6}$$

In these expressions $\gamma_{r,12}(f)$ is the coherence between $r_1(t)$ and $r_2(t)$. Also, $G_{r,11}(f)$, $G_{r,22}(f)$, and $G_n(f)$ are the PSDs of $r_1(t)$, $r_2(t)$, and $n(t)$ respectively. With these modified SNR values, the total SNR as given in [8] can be modified as

$$\text{SNR} = \frac{\text{SNR}_1\text{SNR}_2}{1 + \text{SNR}_1 + \text{SNR}_2}. \tag{7}$$

With this formulation, we can compare the SNR in (7) to the threshold SNR in (5) for any level of coherence.

4.2 Experimental results

In this section we present the results of field experiments designed to study the effects of source signal parameters on time delay estimation accuracy. A set of 319 different PN sequences was generated with varying durations, bandwidths, and center frequencies as described in Section 2. Each source waveform was then transmitted from an endfire position as shown in Figure 1.

The time delay estimates, $\hat{d}$, were obtained using a simple cross-correlator (SCC) that estimates the delay as the position of the peak in the cross-correlation between two received signals. The true time differences were calculated from the known geometry of the array and used to determine

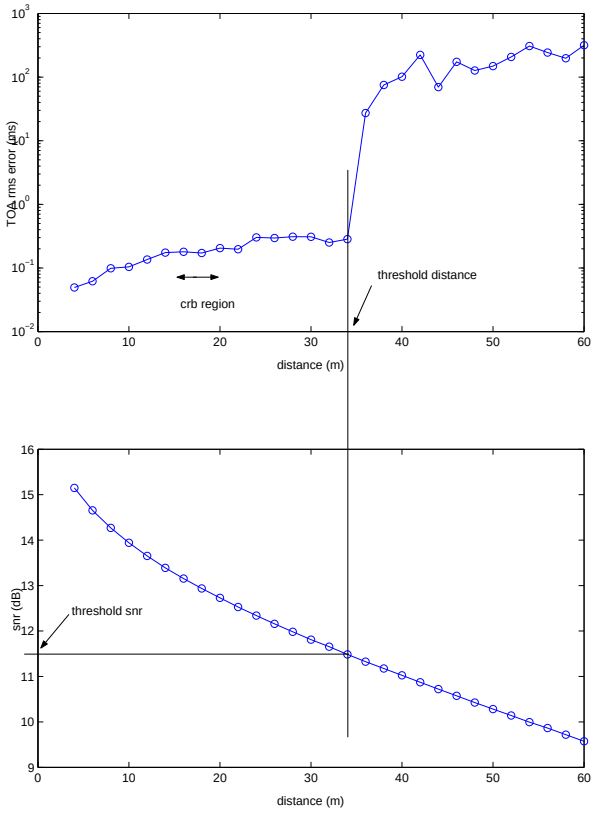

Figure 5: Simulation result illustrating the threshold phenomenon in time delay estimation. As the inter-sensor distance increases, the SNR falls below the threshold SNR and the estimation error rapidly increases. This is due to the peak ambiguity problem that arises when attempting to identify the maximum of the received signals' cross-correlation.

the error of the estimates. It is recognized that the generalized cross-correlator (GCC) is the maximum likelihood estimator in this problem [12], however we chose the SCC because of its ease in implementation and for its robustness. The GCC requires knowledge of the signal and noise power spectra and can give poor performance if the estimated spectra is mismatched from the true spectra (see, e.g. [6]).

4.2.1 Error versus Bandwidth and Signal Length

Figure 6 illustrates the observed TDE error of three different source signals. The center frequency and length were held constant at $100Hz$ and 10s, while only the bandwidth was varied. The time delay error as a function of distance was empirically determined by cross-correlating the received signal at microphone 1 with with the received signals at microphones 3, 4, 5, 6, and 7. As expected, the higher bandwidth signals had better delay estimation performance. Figure 7 was produced in the same way, except that signal duration was varied while center frequency and bandwidth were held constant. The low TDE errors of these two plots indicate that all the points are above the threshold SNR and that we are in the CRLB region.

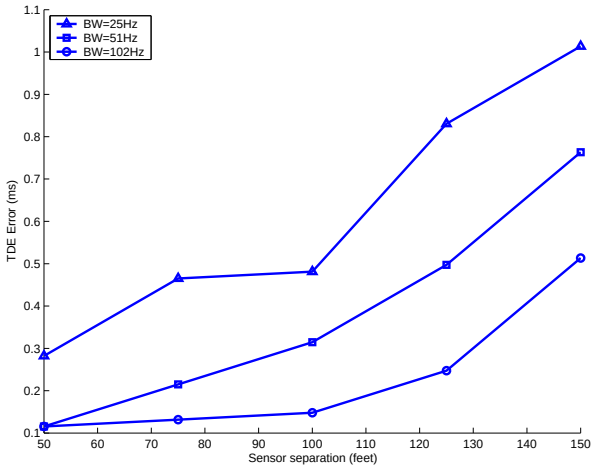

Figure 6: Time delay estimation error as a function of sensor separation and signal bandwidth. $F_c = 100$Hz, $T = 10$s in all cases.

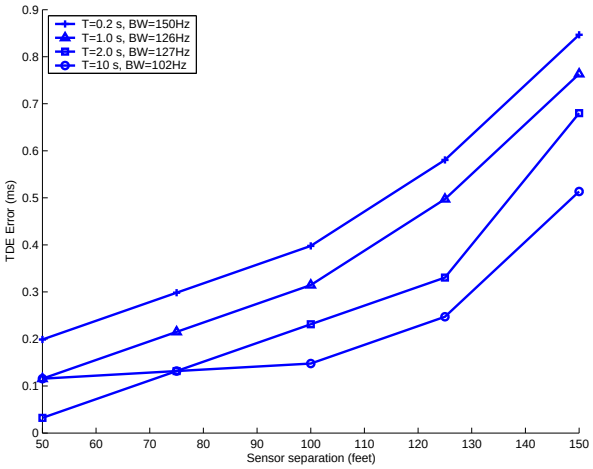

Figure 7: Time delay estimation error as a function of sensor separation and signal Length. $F_c = 100$Hz in all cases.

Figure 8 shows TDE errors for a case in which the SNR of the signals received at the microphones sometimes falls below the threshold SNR. Figure 8 is the same as Figure 6 except the signal length has been reduced to $T = 2$s and the center frequency has been raised to $F_c = 400$Hz. Because of the shorter duration, the energy in the signals is lower and the post integration SNR is reduced. From the Figure, the error of the two low-bandwidth signals is seen to dramatically increases with distance. In this case, the SNR is falling below the threshold somewhere between 50 and 150 feet. The threshold SNR of the 31Hz signal is lower than the 15Hz one (as predicted by equation (5)) causing it to diverge at a greater distance. The signals with bandwidth of 63Hz and 127Hz appear to be above the SNR threshold for all distances considered.

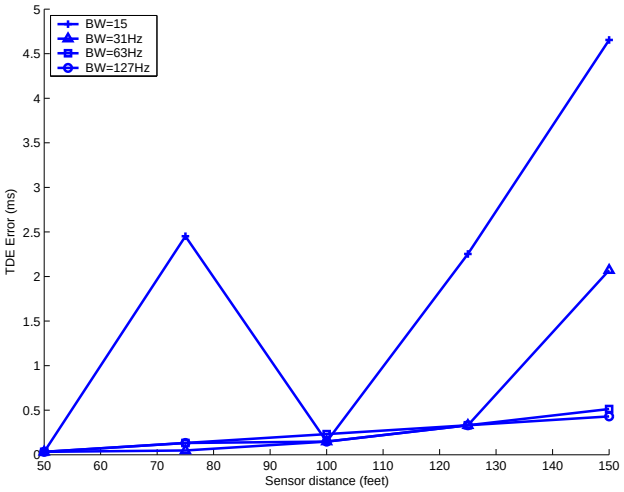

Figure 8: Example of signals falling below SNR_{th}. As sensor distance is increased, the two low-bandwidth signals fall below their SNR_{th} and their error dramatically increases. $F_c = 400$Hz, $T = 2$s.

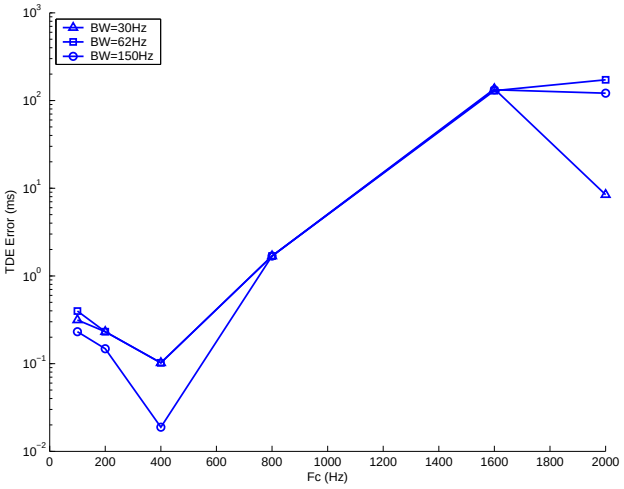

Figure 9: Experimental time delay estimation error versus center frequency. $T = 1$s.

4.2.2 Error versus Center Frequency

Figure 9 illustrates the observed increase in TDE error as the center frequency of the PN sequence is increased from 100Hz to 2000Hz for different PN sequence bandwidths. There are two major effects contributing to the increase in error. First, for a fixed bandwidth signal, the percent bandwidth, W/ω_0, decreases with center frequency and the threshold SNR increases according to (5). The second cause for the increase in TDE error is the loss of signal coherence. This connection is illustrated by comparing the observed coherences in Figure 4 to the errors in Figure 9. Center frequencies of 1600Hz and 2000Hz exhibit the worst coherence and these have the highest TDE errors. Nearly perfect coherence was observed at $F_c = 400$Hz and this has the smallest TDE error (0.02ms for 150Hz bandwidth) over the range of bandwidths considered. From equations (6) and (7) the coherence loss can be interpreted as a loss in the effective SNR.

Figure 10 illustrates how the SNR threshold changes with center frequency. The top row of plots illustrates the observed TDE error as a function of bandwidth for center frequencies of 200, 400, and 800Hz. Each plot in the second row corresponds to the one above it and gives the measured total SNR (see equation (7)) of each bandwidth signal. The solid lines in the second row of plots are the threshold SNRs as a function of bandwidth as predicted by (5). When the observed SNR is above SNR_{th}, we expect low TDE error; this is seen in the top row of plots. The match between the observed and theoretical bandwidth thresholds is not perfect, but the trends are evident. In Figure 11 we plot the TDE error predicted by the WWLB and observe the same trend seen empirically in the top row of Figure 10.

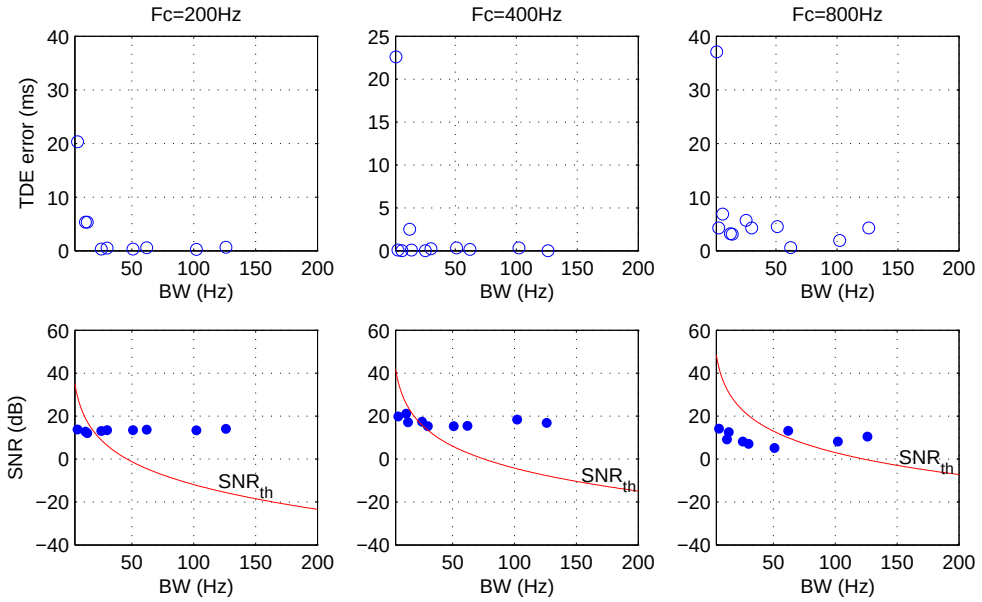

Figure 10: Illustration of threshold SNR at different center frequencies. The top row shows empirical TDE error versus bandwidth for three different PN signal center frequencies. Below each of these is a plot showing the empirical SNR for each bandwidth signal. The solid curves in the bottom row of figures show the threshold SNR values predicted by (5). We expect low TDE error when the signal SNR is above the threshold and high TDE error when the signal is below the threshold (see Figure 5).

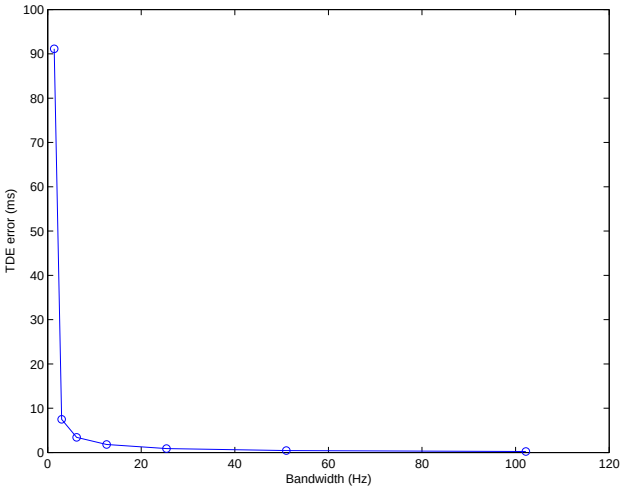

Figure 11: Time delay estimation error as a function of signal bandwidth as predicted by the WWLB, equation (3).

5. Localization Experiment

The self-localization scenario consists of a number of sources that are placed in a field of sensors with unknown locations. Each source transmits a source signal which is detected by a subset of the sensors and used to compute the TDOAs. The time measurements are then passed to a localization algorithm to determine the locations of the sensors. In this section we present results of self-localizing an acoustic sensor network using a subset of the PN sequences described earlier as the calibration source signals.

To study self-localization performance, we reconfigured the acoustic array into the nonlinear configuration depicted in Figure 12. The six sensors (microphones) are represented by the circles, and the four sources are represented by the $\times$'s. We emulated the four calibration sources by moving the boom box to these different positions. Because the emission times were unknown, the only information available to the localization algorithm were the TDOAs obtained from simple cross correlations of the received PN sequences as described in Section 4. With these time estimates, self-localization was then performed using the maximum likelihood algorithm given in [13]. The results from a 2s calibration signal with $F_c = 200$Hz and $BW = 127$Hz are presented in Figure 12. Sensor location estimates are shown by triangles, while estimates of the source locations are given by filled circles. The average localization error of the scene estimate was calculated as

$$\frac{1}{N}\sum_{n=1}^{N}\sqrt{(x_i-\hat{x_i})^2+(y_i-\hat{y_i})^2}, \tag{8}$$

where $N = 6$ is the number of sensors, x_i and $\hat{x_i}$ are the true and estimated x-coordinates of the i^{th} sensor respectively, and y_i and $\hat{y_i}$ are the true and estimated y-coordinates of the i^{th} sensor. For this source signal, the average error was 0.9 feet.

This experiment was repeated for several different PN sequences. The source signal parameters and their resulting localization errors are given in Table 1. The estimates in Figure 12 correspond to source signal 4 in this table. For most of the signals, we see good agreement between estimated and actual sensor locations. The exception is signal 7 which has the worst performance although it has the greatest bandwidth. The poor localization performance is due to poor time delay estimates which are caused by the degraded signal coherence at $F_c = 1600$Hz. Similarly, the smallest average localization errors correspond to low frequency source signals that were previously observed to have high signal coherence.

6. Conclusions

In this paper we have presented experimental results from an outdoor field experiment designed to study the effects of source signal bandwidth, center frequency, and duration on time delay estimation error. A set of 319 different PN source signal waveforms spanning these three signal parameters were emitted endfire upon a linear array of length 150 feet. We specifically investigated signal attenuation and coherence in the channel as function of distance. Our empirical findings

Signal #	F_c (Hz)	BW (Hz)	length (sec)	average localization error (feet)
1	200	14	1	9.6
2	200	127	2	0.7
3	400	30	1	3.18
4	400	127	2	0.9
5	800	127	2	18.3
6	800	254	1	5.1
7	1600	1023	2	103

Table 1: Source signals used in network self-calibration. Position estimates were made from time delay estimates of these signals and errors were calculated from the known true positions.

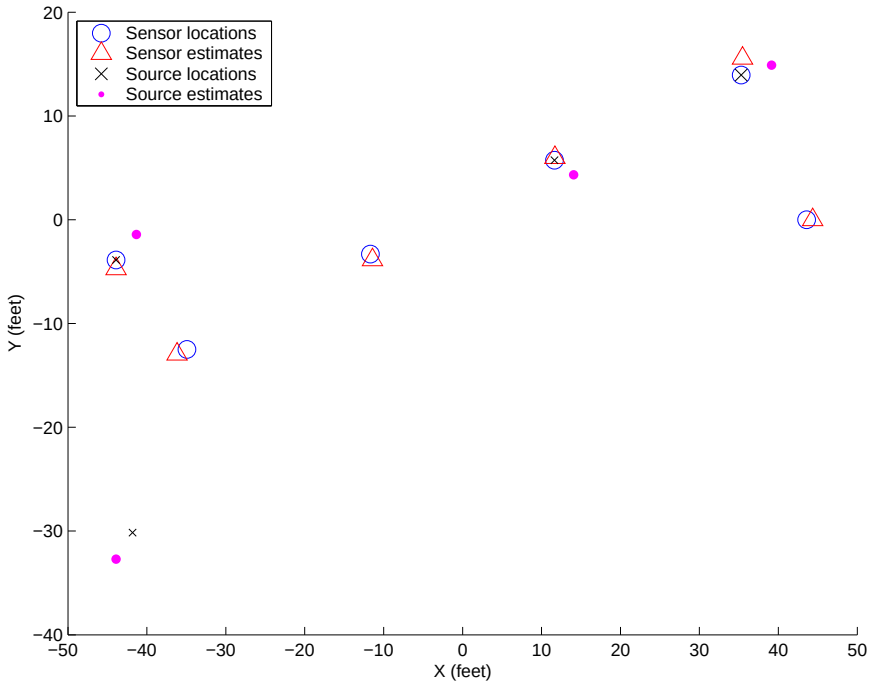

Figure 12: Sensor network used in self-localization. Shown estimates correspond calibration signal 4 of Table 1.

showed that, while raw SNR may be high, loss of signal coherence severely degraded the time delay estimates. We observed poor coherence values for frequencies above 800Hz. For center frequencies lower than this, we obtained time delay estimate errors on the order of 0.5ms at 150 feet. Finally, we presented results from an outdoor self-localization experiment in which we used PN sequences to obtain time difference of arrival estimates for 6 six randomly placed sensors. Using four calibration sources we obtained an average localization error of 0.7 feet in the best case.

References

[1] S.Kumar, F.Zhao, and D.Shepherd "Collaborative Signal and Information Processing in Microsensor Networks," IEEE Signal Processing Magazine, Special Issue on Collaborative Information Processing, Vol. 19, No. 2, March 2002.

[2] R.L. Moses, D. Krishnamurthy, and R. Patterson "A self-localization method for wireless sensor networks," Eurasip Journal on Applied Signal Processing, Special Issue on Sensor Networks, Vol. 2003, pp. 348-358, March 2003.

[3] J.G. Proakis, "Digital Communications", Boston, McGraw-Hill, 1995.

[4] International Standard ISO 9613-1:1993(E).

[5] R.J. Kozick and B.M. Sadler. "Algorithms for Localization and Tracking of Acoustic Sources with Widely Separated Sensors," *2000 Meeting of the MSS Specialty Group on Battlefield Acoustic and Seismic Sensing*, October 17-19, 2000.

[6] G.C.Carter "Coherence and Time Delay Estimation," Piscataway, NJ, IEEE Press, 1993.

[7] Erik M. Salomons "Computational atmospheric acoustics", Dordrecht/Boston/London, Kluwer Academic Publishers, 2001.

[8] A.J. Weiss, and E. Weinstein, "Fundamental Limitations in Passive Time Delay Estimation, Part I: Narrow-band Systems," IEEE Transactions on Acoustics, Speech, and Signal-Processing, Vol. 31, No. 2, pp. 472-485, April 1983.

[9] A.J. Weiss, and E. Weinstein, "Lower Bounds on the Mean Square Error in Random Parameter Estimation," IEEE Transactions on Information Theory, Vol. IT-31, No. 5, pp. 680-682, September 1985.

[10] A.J. Weiss, "Composite Bound on Arrival Time Estimation Errors," IEEE Transactions on Aerospace and Electronic Systems, Vol. 22, No. 6, pp. 751-756, November 1986.

[11] J.P. Ianniello "Time Delay Estimation Via cross-correlation in the Presence of Large Estimation Errors," IEEE Transactions on Acoustics, Speech, and Signal-Processing, Vol. 30, No. 6, pp. 998-1003, 1982.

[12] C.H. Knapp and G.C Carter "The Generalized Correlation Method for Estimation of Time Delay," IEEE Transactions on Acoustics, Speech, and Signal-Processing, Vol. 24, No. 4, pp. 320-327, 1976.

[13] R.L. Moses, D. Krishnamurthy, and R. Patterson "An Auto-calibration Method for Unattended Ground Sensors," IEEE International Conference on Acoustics, Speech, and Signal Processing. Vol. 3, pp. III-2941-III-2944, May 13-17, 2002.