I. INTRODUCTION

Sensor networks are emerging in a large number of civilian and military applications to sense and process information about their surroundings. To make use of this information the locations of the individual sensors need to be known. 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 based on measurements of time-of-arrival (TOA), time-difference-of-arrival (TDOA), direction-of-arrival (DOA), or received signal strength measurements. Several techniques employ hybrid schemes that combine multiple types of measurements, such as ranging information from TOA estimates and angular information from DOA estimates. In this paper we consider TDOA measurements from acoustic sources to localize elements of a microphone sensor array. The resultant localization accuracy is closely tied to the accuracy with which we can estimate TDOA. In this paper we explore, through experimentation the impact of acoustic source signal parameters on time-delay estimation (TDE) and source localization. The theory of time-delay estimation from propagating waves is relatively well understood, however very little work exists to validate this theory in the aeroacoustic regime. In this paper we compare theoretical expectations of TDE performance to experimental observations in an aeroacoustic environment. In particular, we evaluate TDE performance with respect to SNR and the source signal parameters of bandwidth, center frequency, and signal duration.

The remainder of this paper is organized as follows. Section II describes our experimental procedure, including hardware setup and signal generation. In Sec. III we present empirical results characterizing the attenuation and coherence loss of the acoustic channel. Section IV contains the results of our empirical study of the effects of center frequency, bandwidth, and signal duration on time-delay estimation accuracy. In Sec. V we present the results of self-calibrating an outdoor sensor network using broadband pseudo-noise (PN) sequences. Finally, in Sec. VI we conclude.

II. EXPERIMENTAL PROCEDURE

A. Equipment and environment

As depicted in Fig. 1, a linear array of eight Knowles BL-1994 microphones, each separated by 7.7 m, was used in this experiment. All microphones were positioned at the same height \( h_m = 18 \) cm above the grass surface. Each microphone was equipped with an 8-cm-radius spherical windscreen and was connected to a National Instruments data acquisition system by a coaxial cable. The data acquisition system consisted of an eight-channel analog signal condi-
tioner (performing amplification and anti-alias filtering to 4 kHz) followed by an analog-to-digital converter. Each channel was sampled at 12 k samples per second with 12 bits per sample and stored to disk for subsequent analysis and processing.

The low-cost COTS sound source consisted of a portable stereo (Sony CFDG55) playing prerecorded signals from a compact disc. The portable stereo was located collinear with the array 4.6 m from the first microphone and with a height of \( h_2 = 38 \) cm as shown in Fig. 1.

The microphone array was set up in a flat, mowed grass field, located in Glacier Ridge Metro Park of Union County, Ohio, on the afternoon of 16 June 2003. The temperature was 21 °C, the relative humidity was approximately 83%, and the wind was light (<2 m/s) with no dominant direction clearly discernable. First, background noise was recorded to establish the noise spectrum of the microphone outputs. Then, a series of PN sequences was 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 and recorded to study discernable. First, background noise was recorded to establish the noise spectrum of the microphone outputs. Then, a series of PN sequences was 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 described in Sec. V.

B. Source signal generation

A set of source signals, \( s(t) \), was generated spanning different time durations, center frequencies, and bandwidths. The source signals were based on PN sequences that were generated from maximum-length shift-register sequences using an \( m \)-stage shift register with linear feedback. The binary PN sequences \( \{b_i\}_{i=1}^n \) drawn from \( \{-1, +1\} \) have length \( n = 2^m - 1 \) and were chosen for their nearly ideal periodic autocorrelation

\[
R(k) = \begin{cases} 
n, & k = 0 \\ -1, & 1 \leq k \leq n-1. 
\end{cases}
\]

The baseband source signal is built up from the sequence values

\[
s_{bb}(t) = \sum_{i=1}^{n} b_i p(t-iT_c),
\]

where \( p(t) \) is a unit amplitude rectangular pulse of duration \( T_c \), and then modulated to the desired center frequency, \( F_c \), to obtain the desired source signal for transmission

\[
s(t) = \sin(2\pi F_c t) s_{bb}(t).
\]

The bandwidth of \( s(t) \) is controlled through \( T_c \), while the signal duration is equal to \( nT_c \). Figure 2 illustrates a typical PN-based source signal at baseband and passband in the time and autocorrelation domains.

In our experiment we chose center frequencies that ranged from 100 to 2000 Hz, bandwidths that ranged from 3 to 3200 Hz, and signal durations that ranged from 0.2 to 10 s. In total, 319 PN source signals were evaluated. The source signals were generated in MATLAB at a sampling rate of 44.1 kHz, exported as audio (.wav) files, and then copied to an audio CD that was played by the portable stereo system during the experiments.

C. Postprocessing

Figure 3 illustrates the normalized power spectral density (PSD) of a \( \{F_c=200 \text{ Hz}, B=63 \text{ Hz}, T=2 \text{ s}\} \) source signal as it was designed, \( G_s(f) \), and as it was received at the first microphone, \( G_{r1,r1}(f) \). In this paper, \( G_{r1,r1}(f) \) denotes the PSD of the signal received at microphone \( i \), and \( G_{rj,rf}(f) \) denotes the cross-spectral density of signals received at microphones \( i \) and \( j \). The power spectra in Fig. 3 and elsewhere are estimated using Welch’s method of averaging periodograms obtained from windowed segments of time series data. A 20-ms Hanning window and 50% window overlap are used throughout this paper.

Harmonic components in the portable stereo are clearly visible in the received signal of Fig. 3. The additional bandwidth provided by these components would aid in our subsequent attempts of time-delay estimation, however they interfere with our attempt to characterize TDE performance based on bandwidth. To provide a fair comparison of the

FIG. 1. Linear array used in field measurements (not to scale).

FIG. 2. Example of a baseband PN source signal, \( s_{bb}(t) \), and the modulated version, \( s(t) \), for transmission (left). Example normalized autocorrelation functions for the baseband (---) and passband (--) signals (right).
different source signals, we include a bandpass filter in the postprocessing phase to eliminate the harmonics from the portable stereo. The bandpass filter was centered at $F_c$ with bandwidth set to the designed bandwidth of the signal, which was taken as the null-to-null bandwidth given by $B=2T_c$.

We also note that postexperiment analysis of the data indicated that measurements from microphones 2 and 8 were corrupt due to hardware failures in those channels. Thus, results from these two microphone signals have been omitted from the empirical results below.

### III. EMPIRICAL PROPAGATION

In this section we present the results of experiments that were designed to give an empirical understanding of the acoustic channel over short distances (<75 m) and moderate bandwidths. The aim of this section is not to develop a precise model of acoustic propagation, but rather 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.

#### A. Noise and attenuation

Figure 4 presents the observed power spectral density (PSD) of the background noise observed at microphones 1, 3, and 7 (chosen arbitrarily for illustration). The close agreement of the noise PSDs indicates similar noise levels at each position and similar frequency responses from each microphone. The drop at 4000 Hz is due to the antialiasing filter in the signal conditioning hardware.

In Fig. 5 we present the observed PSDs for received signals from a high bandwidth PN sequence \( \{F_c = 2000 \text{ Hz}, B = 3200 \text{ Hz}, T = 10 \text{ s}\} \). Because the sound source was uncalibrated, the exact PSD of the transmitted signal was unknown and the attenuation 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. These losses also exceed those expected from spherical spreading and atmospheric absorption alone. For example, in going from microphone 1 to 3, the distance from the source approximately quadruples and a loss of 12.7 dB would be expected from spherical spreading,\(^{20}\) moreover, the expected atmospheric absorption is less than 1 dB for the distances and frequencies considered.\(^{21}\) However, for most frequencies the actual power loss between $G_{1,1}(f)$ and $G_{3,3}(f)$ exceeds 20 dB, implying the presence of other losses. Also apparent in the figure is a sequence of nulls whose position in frequency decreases with distance. This trend is the opposite of what would be expected from simple destructive interference from a single ground reflection. Although the details of turbulence models are outside the scope of this paper, we comment that basic models imposing random phase combining of direct and reflected rays at each microphone can explain this phenomenon. The interested reader is referred to Ref. 22 and Ref. 20, Appendix K.

By examining the difference in signal power and the average noise power in Fig. 5 we can investigate the SNR as a function of frequency and distance. For example, at 2000 Hz, microphone signal 7 (which is 50.8 m from the sound source) has an SNR of approximately 18 dB in this experiment. All of the SNRs are observed to be greater than 10 dB except near the null frequency for microphone 7.

---

**FIG. 3.** Power spectral density, $G_s(f)$, of a \( \{F_c = 200 \text{ Hz}, B = 63 \text{ Hz}, T = 2 \text{ s}\} \) source signal as designed (left) and as received at the first microphone $G_{1,1}(f)$ (right). Harmonic components, likely caused by nonlinearities in the sound system, are clearly visible in the received signal.

**FIG. 4.** Power spectral density of background noise as observed at microphones 1, 3, and 7. The close agreement indicates similar noise levels and microphone responses.
B. Coherence

Signal coherence, $\gamma_{i,r}(f)$, describes the degree of correlation between like frequency components in two signals $r_i(t)$ and $r_j(t)$. As we describe in Sec. IV, signal coherence plays an important role in time-delay estimation because it affects the quality of the cross correlation function that is used in estimating time delay. In this section we present empirical observations of signal coherence as a function of microphone separation and frequency. The coherence between signals received at microphones $i$ and $j$ is calculated from the PSD estimates as

$$\gamma_{i,r}(f) = \frac{G_{r,r}(f)}{\sqrt{G_{r,r}(f)G_{r,r}(f)}}. \tag{4}$$

In the subsequent figures we plot the magnitude-squared coherence (MSC): $|\gamma_{i,r}(f)|^2$ - a quantity bounded by 0 and 1, with endpoints representing no correlation and perfect correlation between like frequency components, respectively.$^{14}$

Figure 6 illustrates the coherence for the same 3200-Hz broadband signal whose PSD is shown in Fig. 5. The spatial coherences for pairs of received signals separated by 15.4, 30.8, and 46.2 m are presented from the evaluation of $\gamma_{1,3}(f)$, $\gamma_{1,5}(f)$, and $\gamma_{1,7}(f)$, respectively. This figure provides a general understanding of the coherence effects in the channel during this experiment. The region of severe coherence loss between 1000 and 2000 Hz is attributed to the power loss at the same frequencies. In this case, the background noise is a more significant portion of the received signal and the coherence is reduced accordingly. We observe this by noting that the points of lowest signal power correspond to minima in coherence as well. For example, microphones 1 and 3 experience power minima at 1300 and 1800 Hz, respectively (see Fig. 5), hence $|\gamma_{1,3}(f)|^2$ is significantly reduced at these frequencies in Fig. 6. The same holds true for the other microphone pairs. Also apparent in Fig. 6 is the general loss of coherence as distance increases. In addition to power loss, we believe coherence loss is partially a result of distortions in the signal wavefront due to atmospheric turbulence.$^{22}$ Coherence versus distance is further investigated below.

In Fig. 7 we explicitly plot the coherence as a function of distance for six low bandwidth signals with different center frequencies. Each line is a plot of $|\gamma_{i,r}(f)|^2$ for $i=3$, and the six lines are generated from center frequencies $F_c \in \{100, 200, 400, 800, 1600, 2000\}$. Here, $B=30$ Hz and

FIG. 5. Received PSDs of a high bandwidth modulated PN sequence $\{F_c = 2000 \text{ Hz}, B=3200 \text{ Hz}, T=10 \text{ s}\}$ at microphones 1, 3, and 7. The sharp cutoffs at 400 and 3600 Hz are due to the postprocessing bandpass filter (see Sec. II C). For comparison, the average noise floor after bandpass filtering, $G_d(f)$, is also plotted.

FIG. 6. Magnitude-squared coherence (MSC) as a function of frequency for microphone separations of 15.4, 30.8, and 46.2 m.

FIG. 7. Magnitude-squared coherence (MSC) versus distance for different center frequencies.
The impact of coherence loss on time-delay estimation is studied in the following section.

### IV. TIME-DELAY ESTIMATION

#### A. Background

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

\[
\begin{align*}
    r_i(t) &= h_i(t) * s(t) + n_i(t), \\
    r_j(t) &= h_j(t) * s(t - \tau) + n_j(t),
\end{align*}
\]

where \(\ast\) denotes convolution, \(\tau\) is the delay to be estimated, \(h_i(t)\) is the deterministic channel impulse response from source to node \(i\), and the \(n_i(t)\) are independent and identically distributed (i.i.d.) additive noise signals that are assumed to be uncorrelated with \(s(t)\) and with each other. As the distance from source to receiver is increased, \(h_i(t)\) is expected to apply greater attenuation to \(s(t)\).

We use the simple cross-correlator (SCC) to estimate the delay. The SCC delay estimate is the position of the peak in the sample cross correlation between two received signals, \(^{17}\)

\[
\hat{\tau} = \arg \max_{\tau} R_{r_i r_j}(\tau).
\]

It is recognized that the generalized cross correlator (GCC) is the maximum likelihood estimator in this problem, \(^{17}\) 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 are mismatched from the true spectra (see, e.g., Ref. 14). Moreover, the loss of statistical efficiency incurred with the SCC is modest for the signal and noise spectra observed and the bandwidths encountered in these experiments. The experiments described in this paper were performed in an open field to minimize the effects of multipath from buildings and other obstacles. For a consideration of identifying first-arriving signal components in a multipath environment see Ref. 23.

Figure 8 presents the results of a computer simulation that illustrates key aspects of the TDE problem. In this example a \(\{f_c = 200\text{ Hz}, B = 30\text{ Hz}, T = 1 \text{ s}\}\) point source signal is subjected to geometrical spreading, and the resulting SNR and TDE effects are examined. From the lower portion of Fig. 8 we observe that as the sensor separation increases the SNR decreases 6 dB per doubling of distance, as expected for spherical expansion and constant background noise. The rms time-delay estimation error is given in the upper half of the figure for the same distances. At a certain distance, about 22 m in Fig. 8, the SNR drops below a threshold signal-to-noise ratio, denoted \(\text{SNR}_{th}\), and the error dramatically increases. This is due to the well-known threshold effect in time-delay estimation\(^{15,24}\) and is caused when the estimator can no longer reliably identify the peak in the cross-correlation because of excess noise or signal decorrelation in the channel. This peak ambiguity results in high estimation error and occurs whenever \(\text{SNR} < \text{SNR}_{th}\).

Next we consider means of statistically bounding the TDE error for certain signal and noise parameters. The Cramér-Rao lower bound\(^{25}\) (CRLB) is the usual statistical tool of choice, however the CRLB is a local bound that is only tight under high SNR conditions. An alternative bound that is tighter over all SNRs is the Weiss-Weinstein lower bound\(^{26}\) (WWLB). The WWLB is derived for the time-delay estimation problem\(^{16}\) 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 then bounded by

\[
\sigma^2 = \max_{0 < h < D} J(h),
\]

where \(J(h)\) is given by

\[\text{CRLB region} \]

\[\text{Threshold distance} \]

\[\text{RMS TDE error (ms)} \]

\[\text{Distance (m)} \]

\[\text{SNR (dB)} \]

\[\text{Distance (m)} \]
where $R_s(h)$ is the source signal’s autocorrelation function and $\Gamma = 2E/N_0$ is the so-called postintegration SNR, with $E$ being the signal energy and $N_0$ the single-sided noise power. When $\text{SNR} > \text{SNR}_{th}$ there is no peak ambiguity in the signals’ cross-correlation and the bound in Eq. (7) reduces to the CRLB:\(^{15}\)

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

and where $R_s(h)$ is the signal portion at microphone $i$, the duration in seconds. From Ref. 15, the value of $\text{SNR}$ reduces to the CRLB:\(^{15}\)

$$
\sigma_i^2 \geq \frac{1}{8\pi^2 B T F_c^2 \text{SNR}} \quad (\text{SNR} > \text{SNR}_{th}),
$$

(9)

where $B$ is the signal bandwidth in Hz, and $T$ is the signal duration in seconds. From Ref. 15, the value of $\text{SNR}_{th}$ can be estimated from the source’s time-bandwidth product $BT$ and bandwidth to center frequency ratio $B/F_c$:

$$
\text{SNR}_{th} = \frac{6}{\pi^2 (BT)} \left( \frac{F_c}{B} \right)^2 \left( \frac{\phi^{-1}(\pi^2 B^2/24 F_c^2)}{B^2} \right)^2,
$$

(10)

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

Although the bound in Eq. (7) accounts for peak ambiguities under low SNR, it is still an optimistic bound because it assumes the received signals are fully coherent except for additive noise. In practice, the signals are not fully coherent as we observed in Sec. III. As such, this bound underestimates the error because the assumption of a linear, time-invariant, convolutional channel in Eq. (5) is violated in the acoustic setting. Under these conditions, the CRLB would also be overly optimistic even for high SNRs. To account for this, Kozick and Sadler\(^9\) propose an effective SNR that treats the coherent portion of a received waveform as “signal” and treats the incoherent portion as additional noise. With this designation, the effective SNRs of a signal received at microphones $i$ and $j$ are, respectively,

$$
\text{SNR}_i = \frac{|\gamma_{i,j}(F_c)| G_{i,i}(F_c)}{G_s(F_c) + (1 - |\gamma_{i,j}(F_c)|) G_{i,i}(F_c)},
$$

$$
\text{SNR}_j = \frac{|\gamma_{i,j}(F_c)| G_{j,j}(F_c)}{G_s(F_c) + (1 - |\gamma_{i,j}(F_c)|) G_{i,i}(F_c)},
$$

(11)

where $G_s(f)$ is the noise PSD that is assumed equal at each microphone, $\gamma_{i,j}$ is the coherence between the signal portion of the received waveforms (neglecting background noise) at microphones $i$ and $j$, and similarly $G_{i,i}$ is the PSD of only the signal portion at microphone $i$. Using these modified SNR expressions, the joint SNR as given in Ref. 15 is modified as

$$
\text{SNR} = \frac{\text{SNR}_i \text{SNR}_j}{1 + \text{SNR}_i + \text{SNR}_j},
$$

(12)

which can be compactly expressed in terms of the MSC of the received waveforms as

$$
\text{SNR} = \frac{|\gamma_{i,j}(F_c)|^2}{1 - |\gamma_{i,j}(F_c)|^2}.
$$

(13)

With this formulation, we can compare the SNR of Eq. (13) to the threshold SNR of Eq. (10) for any level of signal coherence.

### B. TDE accuracy: 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 source signals formed from modulated PN sequences was generated with varying durations, bandwidths, and center frequencies as described in Sec. II. Each source waveform was then transmitted from an endfire position as shown in Fig. 1. The time delays were estimated using the SCC in Eq. (6) and compared to the true time differences obtained from the known geometry of the array.

In computing the true time delays from the known geometry we used an estimate of the speed of sound derived from a least-squares fit to a subset of our measurements. This subset was carefully selected to only contain signals with high SNR and from adjacent microphones. The speed of sound so obtained was 343.6 m/s.

#### 1. Error versus bandwidth and signal length

Figure 9 illustrates the observed TDE error of three different source signals. The center frequency and signal duration were held constant at 100 Hz and 10 s, while only the bandwidth was varied. The time-delay estimate as a function of distance was empirically determined by cross correlating the received signal at microphone 1 with the received signals at microphones 3–7. The absolute error, when compared to the true delays, is plotted on the vertical axis. As expected, the higher bandwidth signals had better delay estimation performance. The low TDE errors suggest that all the points are above the threshold SNR and that we are in the CRLB region. From the CRLB in Eq. (9), we expect that doubling the
bandwidth, and thus the time-bandwidth product, will decrease the TDE error by a factor of $(1 - \sqrt{1/2}) = 29\%$. This expectation is generally confirmed in Fig. 9 where the error reduction, for doubling of bandwidth, ranges from 25\% to 59\% with a median value of 43\%.

Figure 10 was produced the same way as Fig. 9 except that signal duration was varied while center frequency and bandwidth were held constant. In this case the added signal length decreases the error as expected, but less than the amount predicted by the CRLB. For example, the time-length decreases the error as expected, but less than the bandwidth were held constant. In this case the added signal duration was varied while center frequency and bandwidth were held constant. Figure 10 was produced the same way as Fig. 9 except that signal duration was varied while center frequency and bandwidth were held constant.

It should also be noted that the errors reported in Figs. 9 and 10 are from a single realization of the experiment, whereas the CRLB represents the squared error averaged over multiple realizations.

Figure 11 illustrates TDE errors for a case where received signal SNRs fall below the threshold SNR in some cases. Figure 11 was produced the same way as Fig. 9 except that the signal length has been reduced to $T=2$ s and the center frequency has been raised to $F_c=400$ Hz. From Eq. (10), both of these changes necessarily raise the value of SNR_{th} which is given in the legend of the figure for the four different bandwidths. The coherence-corrected SNR values from Eq. (13) are also presented for selected points. In Fig. 11 we observe a rapid increase in error for the two lowest bandwidth signals as the distance increases. The increase in error is quantized to multiples of approximately 2.5 ms. This corresponds to the spacing $(1/F_c)$ between peaks of the correlation function and indicates that the estimator has selected the incorrect peak of the cross-correlation function (see Fig. 2).

The two highest bandwidth signals in Fig. 11 maintain low TDE error over the entire range of distances, and their SNRs are well above the predicted threshold. For example, at 46.2 m the 127-Hz signal is still 21.0 dB above its threshold. In contrast, the two lower bandwidth signals sometimes fall below their threshold SNR, giving rise to a dramatic increase in TDE error. The 31-Hz signal has a threshold SNR of 13.0 dB; measurements at 38.5 m and below are above this threshold, and low TDE errors are seen for these distances. At 46.2 m the estimator has clearly identified an incorrect peak in the cross correlation and the measured SNR (14.9 dB) is very close to the predicted threshold. Finally, for the 15-Hz signal, we observe that three of the five measurements exhibit large TDE errors and that these points all have SNR below the predicted 23.3 dB threshold. Thus, we see that the SNR threshold is a good predictor of whether the TDE error will be low or high in these experiments. As in Figs. 9 and 10, the experimental results in Fig. 11 are for single realizations, whereas the SNR threshold predicts average performance.

2. Error versus center frequency

Figure 12 illustrates the observed increase in TDE error as the center frequency of the PN sequence is increased from 100 to 2000 Hz 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, $B/F_c$, decreases with increasing center frequency and the threshold SNR increases according to Eq. (10). The second cause for the increase in TDE error is the loss of signal coherence as the center frequency increases—which from Eqs. (11) and (13) implies a loss in effective SNR. This connection is illustrated by comparing the observed coherences in Fig. 7 to the TDE errors in Fig. 12. Center frequencies of 1600 and 2000 Hz exhibit the lowest 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 over the range of bandwidths considered.

Figure 13 illustrates how the SNR threshold changes with center frequency and bandwidth. The top row of plots illustrates the observed TDE error as a function of bandwidth...
for center frequencies of 200, 400, and 800 Hz. Each plot in the second row corresponds to the one above it and gives the measured total SNR [from Eq. (13)] for each signal. The solid lines in the second row of plots are the threshold SNRs as a function of bandwidth as predicted by Eq. (10). 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 Fig. 14 we plot the TDE error bound predicted by the WWLB in Eq. (7) and compare it to empirical errors measured for the $F_c=400$ Hz case. The postintegration SNR for the plot was taken as $\Gamma=16$ dB by fitting Eq. (7) to the TDE measurements. The measured $\Gamma$ was approximately 35 dB. The tight agreement between the bound and the estimates verifies the validity of the WWLB, however the effective SNR (16 dB) is less than the observed SNR (35 dB).

V. LOCALIZATION EXPERIMENT

In this section we present an experiment in which we estimate both sensor and signal source locations from TDOA measurements obtained from time-delay estimates. The self-localization scenario consists of a number of signal sources placed in a field of sensors with unknown locations. The sources, which also have unknown positions, each transmit a calibration source signal that is detected by a subset of the sensors and used to compute the TDOAs. It is assumed that the emission times of the sources are unknown, but that the sensors all have a common time base. The time measurements are then passed to a localization algorithm to determine the locations of the sensors. Because TDOA measurements only provide information about the relative configuration of sensors, we only evaluate the performance of relative localization. Absolute localization requires prior
four signal sources are represented by the X’s. We emulated the four calibration sources by moving the portable stereo to different positions. Neither the source nor sensor locations are known to the localization algorithm; in addition, while three source locations are co-located with sensors, this information was not provided to the localization algorithm. Because the signal emission times were unknown, the only information available to the localization algorithm was the TDOAs obtained from cross correlations of the received PN sequences described earlier as the calibration source signals.

The self-localization experiment used the same equipment described in Sec. II, but altered the acoustic array into the nonlinear configuration depicted in Fig. 15. The six sensors (microphones) are represented by the circles, and the four signal sources are represented by the X’s. We emulated the four calibration sources by moving the portable stereo to these different positions. Neither the source nor sensor locations are known to the localization algorithm; in addition, while three source locations are co-located with sensors, this information was not provided to the localization algorithm. Because the signal emission times were unknown, the only information available to the localization algorithm was the TDOAs obtained from cross correlations of the received PN sequences as described in Sec. IV. With these time estimates, self-localization was then performed using the maximum likelihood algorithm described in Ref. 27 which simultaneously estimates the x and y locations of the sensors along with the locations and signal emission times of the sources. The unknown source locations and unknown emission times are considered nuisance parameters in the sensor localization problem. The results from a T = 2 s calibration signal with $F_c = 200$ Hz and $BW = 127$ Hz are presented in Fig. 15. Sensor location estimates are shown by triangles, while estimates of the source locations are given by solid dots. The average sensor localization error of the scene estimate was calculated as

$$\frac{1}{N} \sum_{i=1}^{N} \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2},$$

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 localization error was 0.28 m.

The preceding experiment was repeated for several different PN-based source signals. The source signal parameters and their resulting average localization errors are given in Table I. The estimates in Fig. 15 correspond to experiment 4 in this table. For most of the signals, we see good agreement between estimated and actual sensor locations. The exception is experiment 7, which has the worst performance even though 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.

![FIG. 15. Sensor network used in self-localization. Shown estimates correspond to calibration experiment 4 of Table I.](image)

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

<table>
<thead>
<tr>
<th>Experiment no.</th>
<th>$F_c$ (Hz)</th>
<th>$BW$ (Hz)</th>
<th>length (s)</th>
<th>Average localization error (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>14</td>
<td>1</td>
<td>2.95</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>127</td>
<td>2</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>400</td>
<td>30</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td>4</td>
<td>400</td>
<td>127</td>
<td>2</td>
<td>0.28</td>
</tr>
<tr>
<td>5</td>
<td>800</td>
<td>127</td>
<td>2</td>
<td>5.63</td>
</tr>
<tr>
<td>6</td>
<td>800</td>
<td>254</td>
<td>1</td>
<td>1.57</td>
</tr>
<tr>
<td>7</td>
<td>1600</td>
<td>1023</td>
<td>2</td>
<td>31.69</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS

We have presented experimental results from an outdoor field experiment designed to study the effects of source signal bandwidth, center frequency, and duration on both signal coherence and time-delay estimation error. One goal was to understand the achievable accuracy of acoustic TDE using low-cost commercial equipment and simple estimation algorithms.

Modulated pseudo-noise signals were used to estimate both signal coherence and TDE accuracy as functions of bandwidth, signal duration, and center frequency. A set of 319 different modulated PN source signals spanning these three signal parameters was transmitted toward a linear array of length 46.2 m, and received signals were used to assess signal coherence and TDE performance. A simple cross-correlation estimator was employed to examine time-delay estimation accuracy. In general, we found that the modulated PN source signals and the SCC estimator were effective in performing time-delay estimation using our uncalibrated COTS equipment.
Both signal coherence and TDE accuracy followed trends predicted by theory to within the limits of the low-cost hardware. As expected, signal coherence was generally higher at lower frequencies, but measured signal coherence dropped off slightly for frequencies below 400 Hz, as a result of greater noise levels and the reduced signal transmission power of the equipment at low frequencies. For center frequencies below 800 Hz, we obtained TDE errors on the order of 0.5 ms for a sensor separation of 46.2 m.

We also considered how well mathematical theory predicted TDE behavior in the aeroacoustic environment. The CRLB and WWLB provide a theoretical expectation that the mean-squared error of time-delay estimates will vary inversely with the source signal’s time-bandwidth product and the coherence-corrected effective SNR; however, higher bandwidths require higher center frequencies, and coherence decreases sharply with distance at higher center frequencies. Our experiments have confirmed these theoretically predicted trends in the aeroacoustic regime; however, in many cases the actual performance was lower than theoretical predictions based on measured SNR values. Besides predicting the TDE error, detecting highly erroneous time-delay estimates is an important element in network self-localization. TDE theory predicts an SNR threshold below which estimates are subject to peak ambiguities resulting in relatively high error. Using Eqs. (10) and (13) we were successful in calculating a coherence-corrected SNR that, when compared to the theoretical threshold, correctly predicted whether the estimate was in the peak ambiguous region or not.

Finally, we presented results from an outdoor self-localization experiment where we used the modulated PN sequences to obtain time-difference-of-arrival estimates for six randomly placed sensors. Our primary goal was to experimentally assess localization performance as a function of the center frequency, bandwidth, and duration of calibration signals used to obtain TDE estimates that form the basis for localization estimates. Using four calibration sources we obtained an average localization error of 0.22 m in the best case.

ACKNOWLEDGMENTS

This work was prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-02-0008. The authors gratefully acknowledge the assistance of D. Keith Wilson in the evaluation of turbulence effects.