Source Localization With Distributed Sensor Arrays and Partial Spatial Coherence

Richard J. Kozick, Member, IEEE, and Brian M. Sadler, Member, IEEE

Abstract—Multiple sensor arrays provide the means for highly accurate localization of the \((x, y)\) position of a source. In some applications, such as microphone arrays receiving aeroacoustic signals from ground vehicles, random fluctuations in the air lead to frequency-selective coherence losses in the signals that arrive at widely separated sensors. We present performance analysis for localization of a wideband source using multiple, distributed sensor arrays. The wavefronts are modeled with perfect spatial coherence over individual arrays and frequency-selective coherence between distinct arrays, and the sensor signals are modeled as wideband, Gaussian random processes. Analysis of the Cramér–Rao bound (CRB) on source localization accuracy reveals that a distributed processing scheme involving bearing estimation at the individual arrays and time-delay estimation (TDE) between sensors on different arrays performs nearly as well as the optimum scheme while requiring less communication bandwidth with a central processing node. We develop Ziv–Zakai bounds for TDE with partially coherent signals in order to study the achievability of the CRB. This analysis shows that a threshold value of coherence is required in order to achieve accurate time-delay estimates, and the threshold coherence value depends on the source signal bandwidth, the additive noise level, and the observation time. Results are included based on processing measured aeroacoustic data from ground vehicles to illustrate the frequency-dependent signal coherence and the TDE performance.

Index Terms—Acoustic arrays, array signal processing, decentralized signal processing, delay estimation, imperfect spatial coherence, sensors, source localization, statistical performance bounds.

I. INTRODUCTION

IN THIS PAPER, we are concerned with estimating the location of a source using a network of sensors. We assume that the sensors are placed in an “array of arrays” configuration containing several small-aperture arrays distributed over a wide area. Each array contains local processing capability and a communication link with a fusion center. A standard approach for estimating the source locations involves bearing estimation at the individual arrays, communication of the bearings to the fusion center, and triangulation of the bearing estimates at the fusion center (e.g., see [1]–[5]). This approach is characterized by low communication bandwidth and low complexity, but the localization accuracy will generally be inferior to the optimal solution in which the fusion center jointly processes all of the sensor data. The optimal solution requires high communication bandwidth and high processing complexity. The amount of improvement in localization accuracy that is enabled by greater communication bandwidth and processing complexity is dependent on the scenario, which we characterize in terms of the power spectra (and bandwidth) of the signals and noise at the sensors, the coherence between the source signals received at widely separated sensors, and the observation time (amount of data). We present a framework to identify scenarios that have the potential for improved localization accuracy relative to the standard bearings-only triangulation method. We propose an algorithm that is bandwidth efficient and nearly optimal that uses beamforming at small-aperture sensor arrays and time-delay estimation (TDE) between widely-separated sensors. The sensor signals are modeled as Gaussian random processes, allowing the inclusion of deterministic as well as random propagation effects.

Accurate TD estimates using widely separated sensors are required to achieve improved localization accuracy relative to bearings-only triangulation. We present results in this paper for the application of aeroacoustic tracking of ground vehicles using a collection of microphone arrays. In aeroacoustics, the signal coherence is known to degrade with increased sensor separation for low-frequency sounds (10–300 Hz) propagating through the air, e.g., [6], [7]. Thus, it is important to understand the fundamental limitations on TD estimation when the signals are partially coherent, and we provide a detailed study of this question in this paper. Our results quantify the scenarios in which TDE is feasible as a function of signal coherence, SNR per sensor, fractional bandwidth of the signal, and time-bandwidth product of the observed data. The basic result is that for a given SNR, fractional bandwidth, and time-bandwidth product, there exists a “threshold coherence” value that must be exceeded in order for TDE to achieve the CRB. The analysis is based on Ziv–Zakai bounds for TDE, using the results in [8] and [9]. Time synchronization is required between the arrays for TDE.

Previous work on source localization with aeroacoustic arrays has focused on angle of arrival estimation with a single array, e.g., [10]–[12]. The problem of imperfect spatial coherence in the context of narrowband angle-of-arrival estimation with a single array was studied in [13]–[19] and [6]. Pauraj and Kailath [13] presented a MUSIC algorithm that incorporates nonideal spatial coherence, assuming that the coherence losses are known. Song and Ritey [14] provided maximum-likelihood (ML) methods for estimating the angles of arrival and the parameters in a coherence model, and Wilson [6] incorporated physics-based models for the spatial coherence losses. Gershman et al. [15] provided a procedure to jointly estimate
the spatial coherence loss and the angles of arrival. In the series of papers [16]–[19], stochastic and deterministic models were studied for imperfect spatial coherence, and the performance of various bearing estimators was analyzed. The problem of decentralized array processing was studied in [20], [21]. Wax and Kailath [20] presented subspace algorithms for narrowband signals and distributed arrays, assuming perfect spatial coherence across each array but neglecting any spatial coherence that may exist between arrays. Stoica et al. [21] considered ML angle of arrival estimation with a large, perfectly coherent array that is partitioned into subarrays. Weinstein [22] presented performance analysis for pairwise processing of the wideband sensor signals from a single array, and he showed that pairwise processing is nearly optimal when the SNR is high. In [23], Moses et al. studied autocalibration of sensor arrays, and for aeroacoustic arrays, the loss of signal coherence at widely separated sensors impacts the performance of autocalibration.

The results in this paper are distinguished from [10]–[22] in that our primary focus is a performance analysis that explicitly models partial spatial coherence in the signals at different sensor arrays in an array of arrays configuration, along with an analysis of decentralized processing schemes for this model. The previous works have considered wideband processing of aeroacoustic signals using a single array with perfect spatial coherence [10]–[12], imperfect spatial coherence across a single array aperture [6], [13]–[19], and decentralized processing with either zero coherence between distributed arrays [20] or full coherence between all sensors [21], [22].

This paper is organized as follows. Section II contains the sensor data model and the CRB analysis of various distributed processing schemes. Performance bounds for TD estimation with partially coherent signals are developed in Section III. Numerical examples as well as results from measured data are included in Sections II and III. The paper concludes in Section IV with a summary and concluding remarks.

II. DATA MODEL AND CRBs

A model is formulated in this section for the discrete-time signals received by the sensors in an array of arrays configuration. To begin, suppose a single nonmoving source is located at coordinates \( (x_s, y_s) \) in the \((x, y)\) plane, and consider \( H \) arrays that are distributed in the same plane, as illustrated in Fig. 1. Each array \( h \in \{1, \ldots, H\} \) contains \( N_h \) sensors and has a reference sensor located at coordinates \( (x_h, y_h) \). The location of sensor \( n \in \{1, \ldots, N_h\} \) is at \( (x_h + \Delta x_{hn}, y_h + \Delta y_{hn}) \), where \( (\Delta x_{hn}, \Delta y_{hn}) \) is the relative location with respect to the reference sensor. If \( c \) is the speed of propagation, then the propagation time from the source to the reference sensor on array \( h \) is

\[
\tau_h = \frac{d_h}{c} = \frac{1}{c} \left[ (x_s - x_h)^2 + (y_s - y_h)^2 \right]^{1/2} \quad (1)
\]

where \( d_h \) is the distance from the source to array \( h \). We model the wavefronts over individual array apertures as perfectly coherent plane waves. Then, in the far-field approximation, the propagation time from the source to sensor \( n \) on array \( h \) is expressed by \( \tau_{hn} = \tau_h + \tau_{nm} \), where

\[
\tau_{nm} \approx \frac{1}{c} \left[ \frac{x_s - x_h}{d_h} \Delta x_{nm} + \frac{y_s - y_h}{d_h} \Delta y_{nm} \right] = \frac{1}{c} \left[ (\cos \phi_h) \Delta x_{hn} + (\sin \phi_h) \Delta y_{hn} \right] \quad (2)
\]

is the propagation time from the reference sensor on array \( h \) to sensor \( n \) on array \( h \), and \( \phi_h \) is the bearing of the source with respect to array \( h \). Note that while the far-field approximation (2) is reasonable over individual array apertures, the wavefront curvature that is inherent in (1) must be retained in order to model wide separations between arrays.

The time signal received at sensor \( n \) on array \( h \) due to the source will be denoted as \( s_h(t - \tau_{hn} - \tau_{nm}) \), where the vector \( s(t) = [s_1(t), \ldots, s_H(t)]^T \) contains the signals received at the reference sensors on the \( H \) arrays. The elements of \( s(t) \) are modeled as real-valued, continuous-time, zero-mean, jointly wide-sense stationary, Gaussian random processes with \(-\infty < t < \infty\). These processes are fully specified by the \( H \times H \) cross-correlation matrix

\[
R_s(\tau) = E\{s(t + \tau)s(t)^T\} \quad (3)
\]

where \( E \) denotes expectation, superscript \( T \) denotes transpose, and we will later use the notation superscript \( * \) and superscript \( \dagger \) to denote complex conjugate and conjugate transpose, respectively. The \((g, h)\) element in (3) is the cross-correlation function

\[
r_{g,h}(\tau) = E\{s_g(t + \tau)s_h(t)\} \quad (4)
\]

between the signals received at arrays \( g \) and \( h \). The correlation functions (3) and (4) are equivalently characterized by their
Fourier transforms, which are the cross-spectral density (CSD) functions in (5) and CSD matrix in (6):

\[
G_{s,gh}(\omega) = \mathcal{F}\{r_{s,gh}(\tau)\} = \int_{-\infty}^{\infty} r_{s,gh}(\tau) e^{-j\omega \tau} d\tau \tag{5}
\]

\[
G_{a}(\omega) = \mathcal{F}\{R_{s}(\tau)\}, \tag{6}
\]

The diagonal elements \(G_{s,gh}(\omega)\) of (6) are the power spectral density (PSD) functions of the signals \(s_h(t)\), and hence, they describe the distribution of average signal power with frequency. The model allows the PSD to vary from one array to another to reflect propagation and source aspect angle differences.

The off-diagonal elements of (6) \(G_{s,gh}(\omega)\) are the CSD functions for the signals \(s_g(t)\) and \(s_h(t)\) received at distinct arrays \(g \neq h\). In general, the CSD functions have the form

\[
G_{s,gh}(\omega) = \gamma_{s,gh}(\omega) [G_{s,gg}(\omega)G_{s,hh}(\omega)]^{1/2} \tag{7}
\]

where \(\gamma_{s,gh}(\omega)\) is the spectral coherence function for the signals, which has the property \(0 \leq |\gamma_{s,gh}(\omega)| \leq 1\). Coherence magnitude \(|\gamma_{s,gh}(\omega)| = 1\) corresponds to perfect correlation between the signals at sensors \(g\) and \(h\), whereas the partially coherent case \(|\gamma_{s,gh}(\omega)| < 1\) models random effects in the propagation paths from the source to sensors \(g\) and \(h\). Note that our assumption of perfect spatial coherence across individual arrays implies that the random propagation effects have negligible impact on the intra-array delays \(\tau_{hm}\) in (2) and the bearings \(\phi_1, \ldots, \phi_H\).

The signal received at sensor \(n\) on array \(h\) is the delayed source signal plus noise

\[
z_{hn}(t) = s_{h}(t - \tau_{h} - \tau_{hn}) + w_{hn}(t) \tag{8}
\]

where the noise signals \(w_{hn}(t)\) are modeled as real-valued, continuous-time, zero-mean, jointly wide-sense stationary, Gaussian random processes that are mutually uncorrelated at distinct sensors and are uncorrelated from the signals. That is, the noise correlation properties are

\[
E\{w_{gm}(t + \tau)w_{hn}(t)\} = r_{w}(\tau) \delta_{gn} \delta_{mn} \quad \text{and} \quad E\{w_{gm}(t + \tau)s_{h}(t)\} = 0 \tag{9}
\]

where \(r_{w}(\tau)\) is the noise autocorrelation function, and the noise PSD is \(G_{w}(\omega) = \mathcal{F}\{r_{w}(\tau)\}\). We then collect the observations at each array \(h\) into \(N_h \times 1\) vectors \(z_{h}(t) = [z_{h1}(t), \ldots, z_{hn}(t)]^T\) for \(h = 1, \ldots, H\), and we further collect the observations from the \(H\) arrays into a vector

\[
Z(t) = [z_{1}(t)^T \cdots z_{H}(t)^T]^T. \tag{10}
\]

The elements of \(Z(t)\) in (10) are zero-mean, jointly wide-sense stationary, Gaussian random processes. We can express the CSD matrix of \(Z(t)\) in a convenient form with the following definitions. The array manifold for array \(h\) at frequency \(\omega\) is

\[
a_h(\omega) = \begin{bmatrix}
\exp(-j\omega \tau_{h1}) \\
\vdots \\
\exp(-j\omega \tau_{hn}) \\
\exp\left[j\omega \left(\cos\phi_{h}\Delta x_{h1} + \sin\phi_{h}\Delta y_{h1}\right)\right] \\
\vdots \\
\exp\left[j\omega \left(\cos\phi_{h}\Delta x_{hn} + \sin\phi_{h}\Delta y_{hn}\right)\right]
\end{bmatrix} \tag{11}
\]

using \(\tau_{hn}\) from (2) and assuming that the sensors have omnidirectional response. Let us define the relative time delay of the signal at arrays \(g\) and \(h\) as

\[
D_{gh} = \tau_{g} - \tau_{h} \tag{12}
\]

where \(\tau_{h}\) is defined in (1). Then, the CSD matrix of \(Z(t)\) in (10) has the form of (13), shown at the bottom of the page. Recall that the source cross-spectral density functions \(G_{s,gh}(\omega)\) in (13) can be expressed in terms of the signal spectral coherence \(\gamma_{s,gh}(\omega)\) using (7).

Note that (13) depends on the source location parameters \((x_{s}, y_{s})\) through the bearings \(\phi_{h}\) in \(a_{h}(\omega)\) and the pairwise TD differences \(D_{gh}\). However, (13) points out that the observations are also characterized by the bearings \(\phi_{h}\) to the source from the individual arrays and the relative delays \(D_{gh}\) between pairs of arrays. Therefore, one way to estimate the source location \((x_{s}, y_{s})\) is first to estimate the bearings \(\phi_{1}, \ldots, \phi_{H}\) and the pairwise TDs \(D_{gh}\) and then to estimate the source location \((x_{s}, y_{s})\) by triangulation with the equations

\[
\cos(\phi_{h}) = \frac{x_{s} - x_{h}}{[(x_{s} - x_{h})^{2} + (y_{s} - y_{h})^{2}]^{1/2}}, \quad h = 1, \ldots, H \tag{14}
\]

\[
\sin(\phi_{h}) = \frac{y_{s} - y_{h}}{[(x_{s} - x_{h})^{2} + (y_{s} - y_{h})^{2}]^{1/2}}, \quad h = 1, \ldots, H \tag{15}
\]

\[
D_{gh} = \frac{1}{c} \left(\frac{(x_{s} - x_{g})^{2} + (y_{s} - y_{g})^{2}}{(x_{s} - x_{h})^{2} + (y_{s} - y_{h})^{2}}\right)^{1/2} - \frac{1}{c} \left(\frac{(x_{s} - x_{g})^{2} + (y_{s} - y_{g})^{2}}{(x_{s} - x_{h})^{2} + (y_{s} - y_{h})^{2}}\right)^{1/2} \quad h = 2, \ldots, H
\]

\[
g = 1, \ldots, h - 1. \tag{16}
\]

A. Cramér–Rao Bound (CRB)

The Cramér–Rao bound (CRB) provides a lower bound on the variance of any unbiased estimator. The problem

\[
G_{Z}(\omega) = \begin{bmatrix}
a_{1}(\omega)a_{1}(\omega)^{\dagger}G_{s,11}(\omega) & \cdots & a_{1}(\omega)a_{H}(\omega)^{\dagger}\exp(-j\omega D_{1H})G_{s,1H}(\omega) \\
\vdots & \ddots & \vdots \\
a_{H}(\omega)a_{1}(\omega)^{\dagger}\exp(+j\omega D_{1H})G_{s,1H}(\omega) & \cdots & a_{H}(\omega)a_{H}(\omega)^{\dagger}G_{s,HH}(\omega)
\end{bmatrix} + G_{tr}(\omega)L \tag{13}
\]
of interest is estimation of the source location parameter vector $\Theta = [x_s, y_s]^T$ using $T$ samples of the sensor signals $Z(0), Z(T_s), \ldots, Z(T(T-1) \cdot T_s)$, where $T_s$ is the sampling period. The total observation time is $T = T_s \cdot T$. Let us denote the sampling rate by $f_s = 1/T_s$ and $\omega_s = 2\pi f_s$. We will assume that the continuous-time random processes $Z(t)$ are bandlimited and that the sampling rate $f_s$ is greater than twice the bandwidth of the processes. Then, it has been shown [24], [25] that the Fisher information matrix (FIM) $J$ for the parameters $\Theta$ based on the samples $Z(0), Z(T_s), \ldots, Z(T(T-1) \cdot T_s)$ has elements

$$J_{ij} = \frac{T}{4\pi} \int_{-\omega_s/2}^{\omega_s/2} \left\{ \frac{\partial G_Z(\omega)}{\partial \theta_i} G_Z(\omega)^{-1} \frac{\partial G_Z(\omega)}{\partial \theta_j} G_Z(\omega)^{-1} \right\} d\omega,$$

where “$\omega$” denotes the trace of the matrix. The CRB matrix $C = J^{-1}$ then has the property that the covariance matrix of any unbiased estimator $\hat{\Theta}$ satisfies $\text{Cov}(\hat{\Theta}) - C \geq 0$, where $\geq 0$ means that $\text{Cov}(\hat{\Theta}) - C$ is positive semidefinite. Equation (17) provides a convenient way to compute the FIM for the array of arrays model as a function of the signal coherence between distributed arrays, the signal and noise bandwidth and power spectra, and the sensor placement geometry.

Let us consider the CRB for an acoustic source that has a narrowband power spectrum, i.e., the PSD $G_{\omega h}(\omega)$ of the signal at each array $h = 1, \ldots, H$ is nonzero only in a narrow band of frequencies $\omega_h - (\Delta \omega/2) \leq \omega \leq \omega_h + (\Delta \omega/2)$. If the bandwidth $\Delta \omega$ is chosen small enough so that the $\omega$-dependent quantities in (17) are well approximated by their value at $\omega_h$, then the narrowband approximation to the FIM (17) is

$$J_{ij} \approx \frac{T \Delta \omega / \omega_s}{\omega_s} \text{tr} \left\{ \frac{\partial G_Z(\omega_h)}{\partial \theta_i} G_Z(\omega_h)^{-1} \frac{\partial G_Z(\omega_h)}{\partial \theta_j} G_Z(\omega_h)^{-1} \right\}.$$

The quantity $T \Delta \omega / \omega_s$ multiplying the FIM in (18) is the time-bandwidth product of the observations. In narrowband array processing, the $T$ time samples per sensor are often segmented into $M$ blocks containing $T/M$ samples each. Then, the discrete Fourier transform (DFT) is applied to each block, and the complex coefficients at frequency $\omega_h$ (at each sensor) are used to form $M$ array “snapshots.” In this case, the quantity $T \Delta \omega / \omega_s$ is approximately equal to $M$, which is the number of snapshots. The narrowband approximation in (18) is most useful when the coherence is zero between all array pairs $\gamma_{\omega h}(\omega_h) = 0$. As we will show in Section III, the coherence is difficult to exploit when the signals are narrowband.

The CRBs presented in (17) and (18) provide a performance bound on source location estimation methods that jointly process all the data from all the sensors. Such processing provides the best attainable results but also requires significant communication bandwidth to transmit data from the individual arrays to the fusion center. Next, we develop approximate performance bounds on schemes that perform bearing estimation at the individual arrays in order to reduce the required communication bandwidth to the fusion center. These CRBs facilitate a study of the tradeoff between source location accuracy and communication bandwidth between the arrays and the fusion center. The methods that we consider are summarized as follows.

1) Each array estimates the source bearing and transmits the bearing estimate to the fusion center, and the fusion processor triangulates the bearings to estimate the source location. This approach does not exploit wavefront coherence between the distributed arrays, but it greatly reduces the communication bandwidth to the fusion center.

2) The raw data from all sensors is jointly processed to estimate the source location. This is the optimal approach that fully utilizes the coherence between distributed arrays, but it requires large communication bandwidth.

3) This is a combination of methods 1 and 2, where each array estimates the source bearing and transmits the bearing estimate to the fusion center. In addition, the raw data from one sensor in each array is transmitted to the fusion center. The fusion center estimates the propagation TD between pairs of distributed arrays and triangulates these TD estimates with the bearing estimates to localize the source.

Consider the simplest scheme (method 1) in which each array transmits only its bearing estimate to the fusion center. The fusion center then triangulates the bearings $\phi_1, \ldots, \phi_H$ to estimate the source location $(x_s, y_s)$ using (14) and (15). This scheme processes the data from each array separately to estimate the bearings, and it ignores coherence that may exist between the signals arriving at different arrays. It is difficult to incorporate ignored data in the CRB, so we proceed by considering the case in which there is no coherence between arrays: $\text{Cov}(\hat{\Theta} - C)$ is a lower bound on triangulation of bearings when $\gamma_{\omega h}(\omega_h) = 0$ for all $g < h$ and all $\omega$. Then, the CSD matrix in (13) is block diagonal, which we denote by $G_Z^{(B)}(\omega)$. Evaluating the CRB in (17) using $G_Z^{(B)}(\omega)$ provides a lower bound on triangulation of bearings\(^2\) when $\gamma_{\omega h}(\omega_h) = 0$. If the sensor data contains coherence between arrays, and if an algorithm exploits the coherence, then the variance may be lower than the CRB based on $G_Z^{(B)}(\omega)$. The “joint CRB” (method 2) provides a lower bound for algorithms that exploit coherence between the arrays. The CRB based on block diagonal CSD $G_Z^{(B)}(\omega)$ is useful because in many acoustic scenarios, the signal bandwidth and coherence are such that the block-diagonal CSD represents the best achievable performance. We use this CRB to bound the performance of triangulation with bearing estimates.

Next, consider method 3, in which each array transmits its bearing estimate and the $T$ samples from one sensor to the fusion center. We assume that the sensor whose samples are transmitted is located at the reference location $(x_h, y_h)$ for the array. In this case, the fusion center is able to exploit signal coherence at distributed arrays by estimating the time delays $D_{\omega h}$. However, coherence between arrays is not exploited in the estimation of the bearings.

\(^2\)The block diagonal CSD matrix $G_Z^{(B)}(\omega)$ corresponds to the case in which each array transmits its covariance matrix to the fusion center. While transmitting the covariance matrix requires slightly more communication bandwidth than transmitting the bearing, it is significantly less than transmitting the raw data from all sensors. The decentralized scheme in which covariance matrices are transmitted from each array was studied in [20], and this method is shown to be slightly more accurate than triangulation of bearings. The CRB based on the CSD matrix $G_Z^{(B)}(\omega)$ is a lower bound on triangulation of bearings.
We approximate the performance bound for this scheme as follows. To simplify the modeling, we assume the existence of an additional independent sensor that is collocated at the reference location \((x_{h}, y_{h})\) of each array. The samples from this independent sensor are transmitted to the fusion center, but they are not used for bearing estimation. Similar to (8), the observations at these additional sensors are modeled as

\[
\tilde{z}_{h}(t) = \tilde{\eta}_{h}(t - \tau_{h}) + \tilde{\nu}_{h}(t), \quad h = 1, \ldots, H
\]

where the noise \(\tilde{\nu}_{h}(t)\) is independent from the noise at all other sensors and shares the common noise PSD \(G_{\nu}(\omega)\). We define a vector \(\tilde{\mathbf{z}}(t) = [\tilde{z}_{1}(t), \ldots, \tilde{z}_{H}(t)]^{T}\) and a larger vector \(\tilde{\mathbf{Z}}(t) = [\mathbf{z}(t)^{T}, \tilde{\mathbf{z}}(t)^{T}]^{T}\) that collects all of the sensor signals in this model. In order to reflect the fact that the signal coherence is not exploited in the bearing estimation using \(\mathbf{Z}(t)\) while it is exploited in the estimation of the time delays \(D_{gh}\) using \(\tilde{z}_{h}(t)\), the cross-spectral density matrix of \(\tilde{\mathbf{Z}}(t)\) is modeled as

\[
G_{\tilde{\mathbf{Z}}}(\omega) = \begin{bmatrix}
G_{\mathbf{Z}}^{(1)}(\omega) & 0 \\
0 & G_{\mathbf{Z}}^{(11)}(\omega)
\end{bmatrix}
\]

(20)

where \(G_{\mathbf{Z}}^{(11)}(\omega)\) is formed from (13), assuming incoherent signals for bearing estimation

\[
G_{\mathbf{Z}}^{(11)}(\omega) = \begin{bmatrix}
a_{1}(\omega)a_{1}(\omega)^{T}G_{n,11}(\omega) & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & a_{H}(\omega)a_{H}(\omega)^{T}G_{s,HH}(\omega) + G_{w}(\omega)I
\end{bmatrix}
\]

(21)

and \(G_{\mathbf{Z}}^{(11)}(\omega)\) includes the signal coherence to allow time-delay (TD) estimation, as in (22), shown at the bottom of the page.

We obtain the FIM for estimation of the source location parameters \((x_{s}, y_{s})\) using this scheme by inserting \(G_{\tilde{\mathbf{Z}}}(\omega)\) in (20) into the general expression (17). In practice, the use of the same sensor for bearing estimation and TDE will have little effect on the estimation performance. Note that the model assumes that the fusion processor estimates the TDs \(D_{gh}\) for \(h = 2, \ldots, H, \quad g = 1, \ldots, H\), jointly based on the time samples \(\tilde{z}(1), \ldots, \tilde{z}(T)\). A practical TDE method is likely to estimate only the \(H - 1\) TDs \(D_{12}, \ldots, D_{1H}\) through independent, pairwise processing of the sensor samples. Such a pairwise processing scheme cannot perform better than the CRB based on (20). However, the results of Weinstein [22] regarding pairwise processing of sensor signals on a single array suggest that the performance degradation is negligible as long as the SNR is greater than 0 dB. It is possible to obtain an exact CRB for pairwise TDE using our model by following Weinstein’s approach [22]. However, the exact CRB is considerably more complicated and is valuable only for low SNR scenarios.

### B. Examples

Next, we present numerical examples that evaluate the CRB on localization accuracy in (17) for a narrowband and a wideband source. We also showed measured aeracoustic data from a ground vehicle and evaluate the spectral coherence of the source at widely separated sensors.

1) CRB Evaluation: We evaluate the CRBs for the three schemes described above for a narrowband source and a wideband source in this section. We will refer back to these examples in Section III after developing the threshold coherence analysis for TDE. The analysis in Section III will show that the CRBs for the narrowband source case are achievable when there is any appreciable loss of signal coherence between arrays, whereas the CRBs for the wideband source case are achievable when moderate coherence losses occur.

Consider \(H = 3\) identical arrays, each of which containing \(N_{1} = \cdots = N_{H} = 7\) sensors. Each array is circular with a 4-ft radius, six sensors are equally spaced around the perimeter, and one sensor is in the center. We first evaluate the CRB for a narrowband source with a 1-Hz bandwidth centered at 50 Hz, and SNR = 10 dB at each sensor. That is, \(G_{\sigma,hh}(\omega)/G_{\nu}(\omega) = 10\) for \(h = 1, \ldots, H\) and \(2\pi(49,5) < \omega < 2\pi(50,5)\) rad/s. The signal coherence \(\gamma_{s,gh}(\omega) = \gamma_{s}(\omega)\) is varied between 0 and 1. We assume that \(T = 4000\) time samples are obtained at each sensor with sampling rate \(f_{s} = 2000\) samples/s. The source localization performance is evaluated by computing the ellipse in \((x, y)\) coordinates that satisfies the expression

\[
[x \ y] \mathbf{J}^{-1} [x \ y] = 1
\]

where \(\mathbf{J}\) is the FIM in (17). If the errors in \((x, y)\) localization are jointly Gaussian distributed, then the ellipse (23) represents the contour at one standard deviation in root-mean-square (RMS) error. The error ellipse for any unbiased estimator of source location cannot be smaller than this ellipse derived from the FIM.

The \(H = 3\) arrays are located at coordinates \((x_{1}, y_{1}) = (0, 0), (x_{2}, y_{2}) = (400, 400),\) and \((x_{3}, y_{3}) = (100, 0)\), where the units are meters. One source is located at \((x_{s}, y_{s}) = (200, 300)\), as illustrated in Fig. 2(a). The RMS error ellipses for joint processing of all sensor data for coherence values \(\gamma_{s}(\omega) = 0, 0.5,\) and 1 are also shown in Fig. 2(a). The coherence between all pairs of arrays is assumed to be identical, i.e., \(\gamma_{s,gh}(\omega) = \gamma_{s}(\omega)\) for \((g, h) = (1, 2), (1, 3), (2, 3).\) The largest ellipse in Fig. 2(a) corresponds to incoherent signals, i.e., \(\gamma_{s}(\omega) = 0,\) and characterizes the performance of the simple method of triangulation using

\[
G_{\mathbf{Z}}^{(11)}(\omega) = \begin{bmatrix}
G_{s,11}(\omega) + G_{w}(\omega) \\
\vdots \\
e^{-j\omega D_{1H}\gamma_{s,1H}(\omega)} (G_{s,11}(\omega)G_{s,HH}(\omega))^{1/2} \\
\vdots \\
e^{-j\omega D_{2H}\gamma_{s,2H}(\omega)} (G_{s,11}(\omega)G_{s,HH}(\omega))^{1/2} \\
\vdots \\
G_{s,HH}(\omega) + G_{w}(\omega)
\end{bmatrix}.
\]

(22)
the bearing estimates from the three arrays. Fig. 2(b) shows the ellipse radius $r = \left( \frac{\text{(major axis)}^2 + \text{(minor axis)}^2}{2} \right)^{1/2}$ for various values of the signal coherence $\gamma_s(\omega)$. The ellipses for $\gamma_s(\omega) = 0.5$ and 1 are difficult to see in Fig. 2(a) because they fall on the lines of the $x$ that marks the source location, illustrating that signal coherence between the arrays significantly improves the CRB on source localization accuracy. Note also that for this scenario, the localization scheme based on bearing estimation with each array and TDE using one sensor from each array performs equivalently to the optimum, joint processing scheme. Fig. 2(c) shows a clearer view of the error ellipses for the scheme of bearing estimation plus TDE with one sensor from each array. The ellipses are identical to those in Fig. 2(a) for joint processing.

Fig. 2(d)–(f) present corresponding results for a wideband source with bandwidth 20 Hz centered at 50 Hz and SNR 16 dB. That is, $G_{s,sh}/G_w = 40$ for $2\pi(40) < \omega < 2\pi(60) \text{ rad/s}$, $h = 1, \ldots, \cdot H; T = 1000$ time samples are obtained at each sensor with sampling rate $f_s = 2000$ samples/s; therefore, the observation time is 1 s. As in the narrowband case in Fig. 2(a)–(c), joint processing reduces the CRB compared with bearings-only triangulation, and bearing plus TDE is nearly optimum.

The CRB provides a lower bound on the variance of unbiased estimates; therefore, an important question is whether an estimator can achieve the CRB. We show in Section III that the coherent processing CRBs for the narrowband scenario illustrated in Fig. 2(a)–(c) are achievable only when the coherence is perfect, i.e., $\gamma_s = 1$. Therefore, for that scenario, bearings-only triangulation is optimum in the presence of even small coherence losses. However, for the wideband scenario illustrated in Fig. 2(d)–(f), the coherent processing CRBs are achievable for coherence values $\gamma_s > 0.75$.

2) Coherence in Measured Data: Next, we present results from measured aeroacoustic data to illustrate typical values of signal coherence at distributed arrays. The experimental setup is illustrated in Fig. 3(a), which shows the path of a moving ground vehicle and the locations of four microphone arrays (labeled 1, 3, 4, 5). Each array is circular with $N = 7$ sensors and a 4-ft radius, as in the previous example. We focus on the 10-s interval indicated by the $\phi$'s in Fig. 3(a), and we process the data in 1-s segments to reduce the effects of the source motion. Fig. 3(b) shows the mean power spectral density (PSD) of the data measured at arrays 1 and 3. The mean PSD in Fig. 3(b) is computed over the 10-s interval by averaging the PSD's from each 1-s data segment. Note the dominant harmonic at 39 Hz. Fig. 3(c) shows the estimated coherence between arrays 1 and 3 during the 10-s segment. The coherence is approximately 0.85 at 40 Hz, which demonstrates the presence of significant coherence at widely separated microphones. Fig. 3(d) shows the estimated coherence between two sensors on array 1, spaced by 8 ft. Note that the coherence is close to unity for frequencies in the range from about 40 to 200 Hz; therefore, our model of perfect signal coherence over individual arrays seems reasonable. An anti-aliasing filter accounts for the coherence drop above $\approx 300$ Hz.

The Doppler effect due to source motion was compensated prior to the coherence estimates shown in Fig. 3(c). Without Doppler compensation, the coherence is significantly reduced, as shown in Fig. 3(e). The time-varying radial velocity of the source with respect to each array in Fig. 3(a) is plotted in the top panel of Fig. 3(f) for a 30-s interval that is centered on the $\phi$’s in Fig. 3(a). If $s(t)$ is the waveform emitted by the source that is moving with radial velocity $v$ with respect to the sensor, then the sensor receives a waveform with the form $\tilde{s}(\alpha t)$, where the scaling factor $\alpha$ is

$$\alpha = 1 - \frac{v}{c}$$

and $c$ is the speed of wave propagation. The scaling factor $\alpha$ is plotted in the bottom panel of Fig. 3(f). Note that for this data set, $0.98 < \alpha < 1.02$, which corresponds to a Doppler frequency shift of approximately $\pm 1$ Hz for an emitted tone at 50 Hz. We use a digital resampling algorithm to compensate for the Doppler effect.

Arrays 1 and 3 are separated by approximately 200 m in Fig. 3(a). We have performed a similar analysis for arrays 1 and 5, which are separated by 500 m, and the coherence is negligible in this case.

III. TIME DELAY ESTIMATION (TDE)

The CRB results presented in Section II-B1 indicate that TDE between widely spaced sensors is an effective way to improve the source localization accuracy with joint processing. Fundamental performance limits for passive time delay and Doppler estimation have been studied extensively for several decades, e.g., see the collection of papers in [26]. The fundamental limits are usually parameterized in terms of the signal-to-noise ratio (SNR) at each sensor, the spectral support of the signals (fractional bandwidth), and the time-bandwidth product of the observations. When a collection of microphone arrays is used for aeroacoustic tracking of ground vehicles, signal coherence degrades with increased spatial separation between the sensors due to random scattering caused by atmospheric turbulence [6], [7]. This coherence loss significantly affects the TDE accuracy.

In this section, we quantify the effect of partial signal coherence on TDE. We present Cramér–Rao and Ziv–Zakai bounds that are explicitly parameterized by the signal coherence, along with the traditional parameters of SNR, fractional bandwidth, and time-bandwidth product. This analysis of TDE is relevant to method 3 in Section II-B1. We focus on the case of $H = 2$ sensors, and we then outline the extension to $H > 2$ sensors.

Let us parameterize the model in (13) by the bearings $\theta_h$ and the time-delay differences $D_{sh}$, and consider first the case of $H = 2$ sensors. Then, the signals at the reference sensors are modeled as

$$z_1(t) = s_1(t) + w_1(t)$$
$$z_2(t) = s_2(t - D) + w_2(t)$$

where $D = D_{z1}$ is the differential time delay. Following (13), the CSD matrix of the sensor signals in (25) and (26) is as in (27), shown at the top of the page after the next page. The signal coherence function $\gamma_{s,2}(\omega)$ describes the degree of correlation
that remains in the signal emitted by the source at each frequency $\omega$ after propagating to sensors 1 and 2. Next, we develop an SNR-like expression for the two-sensor case that appears in all subsequent expressions for fundamental limits on TDE. We
begin with the magnitude-squared coherence (MSC) \([26]\) of the observed signals \(z_1(t), z_2(t)\) as a function of the signal coherence magnitude, \(\gamma_{s,12}(\omega)\), and other spectral density parameters

\[
\text{MSC}_z(\gamma_{s,12}(\omega)) = \frac{\text{CSD}[z_1(t), z_2(t)]^2}{\text{PSD}[z_1(t)] \cdot \text{PSD}[z_2(t)]} = \frac{\gamma_{s,12}(\omega)^2 G_{n,11}(\omega)G_{n,22}(\omega)}{[G_{s,11}(\omega) + G_w(\omega)][G_{s,22}(\omega) + G_w(\omega)]} \leq 1. \tag{28}
\]

Then, the following SNR-like expression, which we denote by \(\text{SNR}_{\text{TDE}}\), is well known to characterize the performance of TDE [26]:

\[
\text{SNR}_{\text{TDE}}(\gamma_{s,12}(\omega)) = \frac{\text{MSC}_z(\gamma_{s,12}(\omega))}{1 - \text{MSC}_z(\gamma_{s,12}(\omega))} \geq 1 - \frac{\gamma_{s,12}(\omega)^2}{1 - \gamma_{s,12}(\omega)^2}. \tag{29}
\]

The standard analysis of TDE is based on (28)–(30) with perfect signal coherence \(\gamma_{s,12} = 1\). Our formulation shows the effect of partial signal coherence on these quantities. The inequality (31) shows that signal coherence loss \(\gamma_{s,12}(\omega) < 1\) severely limits the \(\text{SNR}_{\text{TDE}}\) quantity that characterizes performance, even if the SNR per sensor \(G_{n,h}(\omega)/G_w(\omega)\) is very large.

**A. Bounds for TDE**

We can use (27) in (17) to find the \(\text{CRB}\) for TDE with \(H = 2\) sensors, yielding

\[
\text{CRB}(D) = 2\pi \int_{0}^{\omega_T} \omega^2 \text{SNR}_{\text{TDE}}(\gamma_{s,12}(\omega)) \, d\omega \tag{32}
\]

where \(\omega_T\) is the total observation time of the sensor data, and \(\text{SNR}_{\text{TDE}}(\gamma_{s,12}(\omega))\) is defined in (30). Let us consider the case in which the signal PSDs, the noise PSD, and the coherence are flat (constant) over a bandwidth \(\Delta \omega\) rad/s centered at \(\omega_0\) rad/s. If we omit the frequency dependence of \(G_{n,11}, G_{n,22}, G_w,\) and \(\gamma_{s,12}\), then the integral in (32) may be evaluated to yield the following \(\text{CRB}\) expression:

\[
\text{CRB}(D) = \frac{1}{2\omega_0^2 \frac{\Delta \omega}{2\pi}} \left[ 1 + \frac{1}{12} \left( \frac{2\Delta \omega}{\omega_0} \right)^2 \right] \text{SNR}_{\text{TDE}}(\gamma_{s,12}) \tag{33}
\]

\[
\geq \frac{1}{2\omega_0^2 \frac{\Delta \omega}{2\pi}} \left[ 1 + \frac{1}{12} \left( \frac{2\Delta \omega}{\omega_0} \right)^2 \right] \frac{1}{\gamma_{s,12}(\omega)} \leq 1 - \left( \frac{G_{s,11}}{G_w} \right)^{-1} - 1 \tag{34}
\]

The quantity \((\Delta \omega \cdot T / 2\pi)\) is the time-bandwidth product of the observations, \((\Delta \omega / \omega_0)\) is the fractional bandwidth of the signal, and \(G_{s,h}/G_w\) is the SNR at sensor \(h\). Note from the high-SNR limit in (34) that when the signals are partially coherent, so that \(\gamma_{s,12} < 1\), increased source power does not reduce the CRB. Improved TDE accuracy is obtained with partially coherent signals by increasing the observation time \(T\) or changing the spectral support of the signal, which is \((\omega_0 - \Delta \omega / 2, \omega_0 + \Delta \omega / 2)\).

The spectral support of the signal is not controllable in passive TDE applications; therefore, increased observation time is the only means for improving the TDE accuracy with partially coherent signals. Source motion becomes more important during long observation times, and we have extended the model to include source motion in [27].

The analysis in this paper is focused on passive TDE, but similar results are obtained in active systems, such as medical ultrasound with partially correlated speckle signals [28], [29]. Since the medical ultrasound systems are active, the designer has much more control over the SNR and bandwidth of the signals. Of course, in passive aeroacoustics, there is no control over the source.

With perfectly coherent signals, it is well known that the CRB on TDE is achievable only when the \(\text{SNR}_{\text{TDE}}\) expression in (30) (with \(\gamma_{s,12}(\omega) = 1\)) exceeds a threshold [8], [9]. Next, we show that for TDE with partially coherent signals, a similar threshold phenomenon occurs with respect to coherence. That is, the coherence must exceed a threshold in order to achieve the CRB (32) on TDE. We state the threshold coherence formula for the following simplified scenario. The signal and noise spectra are flat over a bandwidth of \(\Delta \omega\) rad/s centered at \(\omega_0\) rad/s, and the observation time is \(T\) s. Further, assume that the signal PSDs are identical at each sensor, and define the following constants for notational simplicity:

\[
G_{n,11}(\omega_0) = G_{n,22}(\omega_0) = G_n
\]
Fig. 3. (a) Path of ground vehicle and array locations for measured data. (b) Mean PSD at arrays 1 and 3 estimated over the 10-s segment in (a), where the top panel is $G_{\omega_{11}}(f)$ and the bottom panel is $G_{\omega_{33}}(f)$. (c) Mean short-time spectral coherence $\gamma_{\omega_{12}}(f)$ between arrays 1 and 3, with Doppler compensation. (d) Mean short-time spectral coherence for two sensors on array 1 that are spaced by 8 ft. (e) Spectral coherence as in (c), but Doppler is not compensated. (f) Radial velocity and Doppler scaling factor $\alpha$ in (24) for source in part (a).

\[ G_{\omega}(\omega_{0}) = G_{\omega}, \quad \text{and} \quad \gamma_{\omega_{12}}(\omega_{0}) = \gamma_{s}. \]  
(35)
Then, the SNR\(_{\text{TDD}}\) expression in (30) has the form

\[
\text{SNR}_{\text{TDD}}(|\gamma_s|) = \left( \frac{1}{|\gamma_s|^2} \left( 1 + \frac{1}{G_s/G_w} \right)^2 - 1 \right)^{-1}. \tag{36}
\]

The Ziv–Zakai bound developed by Weiss and Weinstein [8], [9] shows that the CRB is attainable only if SNR\(_{\text{TDD}}\) exceeds a threshold value that is a function of the time-bandwidth product \((\Delta \omega \cdot T/2\pi)\) and the fractional bandwidth \((\Delta \omega/\omega_0)\). The condition for CRB attainability and the threshold value is given by [8], [9]

\[
\text{SNR}_{\text{TDD}}(|\gamma_s|) \geq \text{SNR}_{\text{thresh}} = \frac{6}{\pi^2} \left( \frac{\omega_0}{\Delta \omega} \right)^2 \left[ \varphi^{-1} \left( \frac{1}{24} \left( \frac{\Delta \omega}{\omega_0} \right)^2 \right) \right]^2, \tag{37}
\]

where \(\varphi(y) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} \exp(-t^2/2) dt\). If we substitute (36) into (37) and rearrange so that \(|\gamma_s|^2\) is on the left side, then the following inequality defines a threshold coherence value that must be exceeded for CRB attainability:

\[
|\gamma_s|^2 \geq \frac{ \left( 1 + \frac{1}{G_s/G_w} \right)^2 }{ 1 + \text{SNR}_{\text{thresh}}} , \text{ so } \quad |\gamma_s|^2 \geq \frac{1}{1 + \text{SNR}_{\text{thresh}}} \text{ as } G_s/G_w \to \infty. \tag{38}
\]

Since \(|\gamma_s|^2 \leq 1\), (38) is useful only if \(G_s/G_w > \text{SNR}_{\text{thresh}}\). Note that the threshold coherence value in (38) is a function of \((\Delta \omega \cdot T/2\pi)\) and \((\Delta \omega/\omega_0)\) through the formula for \(\text{SNR}_{\text{thresh}}\) in (37).

Fig. 4(a) contains a plot of (38) for a particular case in which the signals are in a band centered at \(\omega_0 = 2\pi 50\) rad/s and the time duration is \(T = 2\) s. Fig. 4(a) shows the variation in threshold coherence as a function of signal bandwidth \(\Delta \omega\). Note that nearly perfect coherence is required when the signal bandwidth is less than 5 Hz (or 10% fractional bandwidth). The threshold coherence drops sharply for values of signal bandwidth greater than 10 Hz (20% fractional bandwidth). Thus, for sufficiently wideband signals, e.g., \(\Delta \omega > 2\pi 10\) rad/s, a certain amount of coherence loss can be tolerated while still allowing unambiguous TDE. Fig. 4(b) shows corresponding results for a case with twice the center frequency and half the observation time.

Fig. 5(a)–(c) contains plots of the threshold coherence in (38) as a function of the time-bandwidth product \((\Delta \omega \cdot T/2\pi)\), SNR \(G_s/G_w\), and fractional bandwidth \((\Delta \omega/\omega_0)\). Note that \(G_s/G_w = 10\) dB is nearly equivalent to \(G_s/G_w \to \infty\) and that very large time-bandwidth product is required to overcome coherence loss when the fractional bandwidth is small. For example, in Fig. 5(a) with fractional bandwidth 0.1, the time-bandwidth product must exceed 100 for coherence 0.9. The variation of threshold coherence with fractional bandwidth is illustrated in Fig. 5(d). For threshold coherence values in the range from about 0.1 to 0.9, each doubling of the fractional bandwidth reduces the required time-bandwidth product by a factor of 10.

Let us examine a scenario that is typical in acoustics, with center-frequency \(f_0 = \omega_0/(2\pi) = 50\) Hz and bandwidth \(\Delta f = \Delta \omega/(2\pi) = 5\) Hz, so that the fractional bandwidth is \(\Delta f/f_0 = 0.1\). From Fig. 5(a), signal coherence \(|\gamma_s| \approx 0.8\) requires time-bandwidth product \(\Delta f \cdot T > 200\); therefore, the necessary time duration \(T = 40\) s for TDE is impractical for moving sources.

Larger time-bandwidth products of the observed signals are required in order to make TDE feasible in environments with signal coherence loss. As discussed previously, only the observation time is controllable in passive applications, thus leading us to consider source motion models in [27] for use during long observation intervals.
We can evaluate the threshold coherence for the narrowband and wideband scenarios considered in Section II-B1 for the CRB examples in Fig. 2. The results are as follows, using (37) and (38).

- **Narrowband case**: \( G_s/G_w = 10, \omega_0 = 2\pi \text{ rad/s, } \Delta \omega = 2\pi \text{ rad/s, } T = 2 \text{ s} \)  
  \( \implies \) Threshold coherence \( \approx 1 \);
- **Wideband case**: \( G_s/G_w = 40, \omega_0 = 2\pi \text{ rad/s, } \Delta \omega = 2\pi \cdot 20 \text{ rad/s, } T = 1 \text{ s} \)  
  \( \implies \) Threshold coherence \( \approx 0.75 \).

Therefore, for the narrowband case, joint processing of the data from different arrays will not achieve the CRBs in Fig. 2(a)–(c) unless there is any loss in signal coherence. For the wideband case, joint processing can achieve the CRBs in Fig. 2(d)–(f) for coherence values \( \geq 0.75 \).

The remainder of this section contains examples of TDE with partial spatial coherence. A simulation example is presented in Section III-B that verifies the CRB and threshold coherence values. In Section III-C, we discuss the extension from \( H = 2 \) sensors to TDE with \( H > 2 \) sensors, and Section III-D contains examples of TDE with measured aeroacoustic data.

### B. TDE Simulation Examples

Consider TDE at \( H = 2 \) sensors with varying signal coherence \( \gamma_s \). Our first simulation example involves a signal with bandwidth \( \Delta f = 30 \text{ Hz} \) centered at \( f_0 = 100 \text{ Hz} \); therefore, the fractional bandwidth \( \Delta f/f_0 = 0.3 \). The signal, noise, and coherence are flat over the frequency band, with SNR \( G_s/G_w = 100 \) (20 dB). The signals and noise are bandpass Gaussian random processes. The sampling rate in the simulation is \( f_s = 10^4 \) samples/s, with \( T = 3 \times 10^4 \) samples; therefore, the time interval length is \( T = 3 \text{ s} \).

Fig. 6(a) displays the simulated RMS error on TDE for \( 0.2 \leq \gamma_s \leq 1.0 \), along with the corresponding CRB from (33). The simulated RMS error is based on 100 runs, and the TDE is estimated from the location of the maximum of the cross-correlation of the sensor signals. The threshold coherence for this case
is 0.41, from (38) and (37). Note in Fig. 6(a) that the simulated RMS error on TDE diverges sharply from the CRB very near the threshold coherence value of 0.41, illustrating the accuracy of the threshold coherence in (38).

Next, we consider TDE with a narrowband signal with $\Delta f = 2$ Hz centered at $f_0 = 40$ Hz. The signal, noise, and coherence are flat over the frequency band, with SNR $G_n/G_w = 100$ (20 dB). The signals and noise are bandpass Gaussian random processes. Fig. 6(b) displays the simulated RMS error on TDE (based on 1000 runs) for coherence values $0.7 \leq \gamma_s \leq 1.0$. As in the previous wideband signal example, the TDE is obtained by cross-correlation. The threshold coherence value is $\approx 1$ for this narrowband case. Fig. 6(b) illustrates the divergence of the simulated RMS error from the CRB, except at $\gamma_s = 1$.

C. TDE With $H > 2$ Sensors

We can extend the analysis of the $H = 2$ sensor case to TDE with $H > 2$ sensors following the approach of Weinstein [22], leading to the conclusion that pairwise TDE is essentially optimum for cases of interest with reasonable signal coherence between sensors. By pairwise TDE, we mean that one sensor, say $H$, is identified as the reference, and only the $H - 1$ time differences $D_{1H}, D_{2H}, \ldots, D_{H-1,H}$ are estimated. Under the conditions described below, these $H - 1$ estimates are nearly as accurate for source localization as forming all pairs of TDEs $D_{gh}$ for all $g < h$.

Extending (35) and (36) to $H > 2$ sensors, let us assume equal $G_{n,h}/G_w$ at all sensors $h = 1, \ldots, H$ and equal coherence $\gamma_s$ between all sensor pairs so that the SNR($\gamma_s$) in (36) is equal for all sensor pairs. Then, as long as $H \cdot \text{SNR}(\gamma_s) \gg 1$, it follows from [22, (67)] that forming all TDE pairs $D_{gh}$ potentially improves the source localization variance relative to pairwise processing by the factor

$$V = \frac{H \left(1 + 2 \cdot \frac{\gamma_s}{1 - \gamma_s}\right)}{2 \left(1 + H \cdot \frac{\gamma_s}{1 - \gamma_s}\right)}.$$  

Clearly, $V \rightarrow 1$ as $\gamma_s \rightarrow 1$, and $V < (3H)/(2(1 + H)) \approx 1.5$ for $\gamma_s > 0.5$. Therefore, the potential accuracy gain from processing all sensor pairs is negligible when the coherence exceeds the threshold values that are typically required for TDE.

This result suggests strategies with moderate communication bandwidth that potentially achieve nearly optimum localization performance. The reference sensor $H$ sends its raw data to all other sensors. Those sensors $h = 1, \ldots, H - 1$ locally estimate the time differences $D_{1H}, \ldots, D_{H-1,H}$, and these estimates are passed to the fusion center for localization processing with the bearing estimates $\phi_1, \ldots, \phi_H$. A modified scheme with more communication bandwidth and more centralized processing is for all $H$ sensors to communicate their data to the fusion center, with TDE performed at the fusion center.

D. TDE With Measured Data

First, we present an illustration based on processing the measured data for the source in Fig. 3(a) that was discussed in Section II-B2. Fig. 7 shows results of cross-correlation processing of the data for a 2-s segment. Fig. 7(a) is obtained by cross-correlating the signals received at one sensor from each of arrays 1 and 3, for which the coherence is appreciable only over a narrow band near 39 Hz [see Fig. 3(c)]. A peak in the cross-correlation is not evident, which is expected based on the preceding analysis, since nearly perfect coherence is needed for narrowband TDE in this scenario. Fig. 7(b) is obtained by cross-correlating the signals received at two sensors on array 1, where the coherence is large over a wide bandwidth [see Fig. 3(d)]. The peak is clearly evident in the cross-correlation in Fig. 7(b).

Next, we present a TDE example based on data that was measured by BAE systems using a synthetically generated, non-moving, wideband acoustic source. The PSD of the source is shown in Fig. 8(a), which indicates that the source bandwidth is about 50 Hz with center frequency 100 Hz. With reference to the sensor locations in Fig. 8(b), the source is at node 2, and the
two receiving sensors are at nodes 0 and 1. The source and sensors form a triangle, with dimensions as follows: The distance from the source (node 2) to sensors 0 and 1 is 233 ft and 329 ft, respectively, and the distance between sensors 0 and 1 is 233 ft. The PSD and coherence magnitude estimated from 1-s segments of data measured at sensors 0 and 1 is shown in Fig. 8(c). Note that the PSDs of the sensor signals do not have their maxima at 100 Hz due to the acoustic propagation conditions. However, the coherence magnitude is roughly 0.8 over a 50-Hz band centered at 100 Hz.

Fig. 4(b) shows the threshold coherence computed with (37) and (38) for the signal in Fig. 8(a) that is centered at $\omega_0/2\pi = 100$ Hz and $T = 1$ s observation time. For bandwidth $\Delta\omega/2\pi = 50$ Hz, the threshold coherence in Fig. 4(b) is approximately 0.5. The actual coherence of 0.8 in Fig. 8(c) significantly exceeds the threshold value; therefore, the TDE between sensors 0 and 1 should be feasible. Fig. 8(d) shows that the generalized cross-correlation has its peak at zero lag, which is the correct location because the sensor data is time aligned before processing. This example shows the feasibility of TDE with acoustic signals measured at widely separated sensors, provided that the SNR, fractional bandwidth, time-bandwidth product, and coherence meet the required thresholds.

Fig. 9 contains another example from the same data set using the sensor locations in Fig. 8(b) and the wideband source with spectrum in Fig. 8(a). In this example, the source is at node 0, and the receiving sensors are at nodes 1 and 3. Note the difference in the PSD shapes in Fig. 9(a), which is similar to our observation about the PSDs in Fig. 8(c). The signal coherence between nodes 1 and 3 is shown in Fig. 9(b), indicating high coherence over an appreciable bandwidth. The cross-correlation is shown in Fig. 9(c), and the peak is clearly evident at the correct location.

Fig. 9(d) and (e) contain a final example using this data. The source is at node 0 and measurements are recorded at nodes 1, 2, and 3, and the time delays are hyperbolically triangulated to estimate the location of the source. Fig. 9(d) shows the hyperbolas obtained from the three differential time delay estimates, and Fig. 9(e) shows an expanded view near the intersection point. The triangulated location is within 1 ft of the true source location, which is at (−3, 0) ft.

We conclude this section with an example based on a different set of aerodynamic data that was measured in an open field. The source in this data is a heavy-tracked vehicle that is moving at a range of approximately 140 m from a collection of 3 sensor arrays. The sensor arrays are located along a straight line with labels A, B, and C. Array A is in the center, the distance from array B to A is 15 m, and the distance from array C to A is 8 m. The source is moving parallel to the line connecting the three arrays. Fig. 10 shows the signal coherence between pairs of arrays in (a) and the cross-correlations in (b). The PSDs at each array are not shown in Fig. 10, but they exhibit strong harmonic components. The coherence in Fig. 10(a) is high over a rather large bandwidth; therefore, the cross-correlation functions in Fig. 10(b) have a clear peak at the correct location.

IV. SUMMARY AND CONCLUDING REMARKS

In this paper, we have presented an analysis of source localization with sensors arranged in an “array of arrays” configuration. We have paid particular attention to aerodynamic localization of ground vehicles, where the signals measured at widely separated sensors are not perfectly coherent due to random propagation effects. We analyzed an algorithm that combines bearing estimation from individual arrays with pairwise TDE between separate arrays. This scheme incorporates distributed processing and data compression so that the communication bandwidth with a fusion center is reduced, with little loss in localization accuracy versus optimal processing. We provided an analysis based on Ziv–Zakai bounds that quantifies the requirements on signal-to-noise ratio, signal bandwidth, signal coherence, and observation time so that joint (coherent)
processing of widely spaced sensor data provides improved localization accuracy. We presented computer simulations and results from processing measured data to illustrate and support the theoretical developments.

Many array processing algorithms have considered subarray processing, such as ESPRIT [30] and its extensions [31]–[33]. ESPRIT is not directly applicable to the model we have studied in this paper. The standard ESPRIT algorithm assumes narrow-band signals, small displacement (< half-wavelength) between identical subarrays, and perfect signal coherence at the subarrays. In this paper, the distinct arrays are widely separated; therefore, the source is near-field with respect to the overall "array of arrays." ESPRIT has been extended to near-field [34] and wideband [35] cases, but the partial signal coherence in our model complicates the application of ESPRIT. We have studied subspace processing for our partial coherence model in [27].
Fig. 9. (a) PSDs at nodes 1 and 3 when transmitter is at node 0. (b) Coherence between nodes 1 and 3. (c) Generalized cross-correlation between nodes 1 and 3. (d) Intersection of hyperbolas obtained from differential time delays estimated at nodes 1, 2, and 3. (e) Expanded view of part (d) near the point of intersection.
ACKNOWLEDGMENT

The authors thank BAE Systems for providing measured aeroacoustic data that was used for some of the examples in Section III-D and the referees for providing suggestions to improve the paper.

REFERENCES


Richard J. Kozick (M’92) received the B.S. degree from Bucknell University, Lewisburg, PA, in 1986, the M.S. degree from Stanford University, Stanford, CA, in 1988, and the Ph.D. degree from the University of Pennsylvania, Philadelphia, in 1992, all in electrical engineering.

From 1986 to 1989 and from 1992 to 1993, he was a Member of Technical Staff at AT&T Bell Laboratories, Murray Hill, NJ. Since 1993, he has been with the Electrical Engineering Department at Bucknell University, where he is currently an Associate Professor. His research interests are in the areas of statistical signal processing, communications, and sensor array processing.

Dr. Kozick serves on the editorial board of the *Journal of the Franklin Institute*. He received the Presidential Award for Teaching Excellence from Bucknell University in 1999 and is a member of the ASEE, Tau Beta Pi, and Sigma Xi.

Brian M. Sadler (M’90) received the B.S. and M.S. degrees from the University of Maryland, College Park, in 1981 and 1984, respectively, and the Ph.D. degree from the University of Virginia, Charlottesville, in 1993, all in electrical engineering.

He is a senior research scientist at the Army Research Laboratory (ARL), Adelphi, MD. He was a lecturer at the University of Maryland from 1985 to 1987 and has been lecturing on statistical signal processing and communications at Johns Hopkins University, Baltimore, MD, since 1994. His research interests generally include statistical signal processing with applications in communications, radar, and aeroacoustics.

Dr. Sadler was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and is on the editorial board for the EURASIP Journal on Wireless Communications and Networking. He is a member of the IEEE Technical Committee on Signal Processing for Communications and co-chaired the 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC-99).