ARMY RESEARCH LABORATORY

Tracking Moving Acoustic Sources With a Network of Sensors

by Richard J. Kozick and Brian M. Sadler

ARL-TR-2750 **October 2002**

Army Research Laboratory
Adelphi, MD 20783-1197

ARL-TR-2750 **October 2002**

Tracking Moving Acoustic Sources With a Network of Sensors

Richard J. Kozick
Bucknell University, Electrical Engineering Department

Brian M. Sadler
Computational and Information Sciences Directorate, ARL

Contents

Figures

Tables

1. Introduction

We are concerned in this report with tracking moving sources using a network of aeroacoustic 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 processing of the bearing estimates at the fusion center with a tracking algorithm [1–5]. This approach is characterized by low communication bandwidth and low complexity, but the localization accuracy may 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 source motion parameters, the power spectra (and bandwidth) of the signals and noise in the sensor data, 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 tracking 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. We have previously studied this model [6] for the case of a single source with fixed position (no motion). We extend the analysis in this report to moving sources that follow a parametric motion model.

Accurate time-delay (TD) estimates are required in order to achieve improved localization accuracy relative to bearings-only triangulation. We are particularly interested in aeroacoustic tracking of ground vehicles using an array of microphone arrays. Signal coherence is known to degrade with increased sensor separation for low frequency sounds (10–300 Hz) propagating through air (for examples, see Wilson [7, 8]). Thus it is important to understand the fundamental limitations of TD and Doppler estimation when the signals are *partially* coherent, and we provide a detailed study of this question in this report. Our results quantify the scenarios in which TDE is feasible as a function of signal coherence, signal-to-noise ratio (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 Cramér-Rao Bound (CRB). The analysis is based on Ziv-Zakai bounds for TDE, as described by Weiss and Weinstein [9,10].

The specific tasks in the statement of work for this contract are as follows.

a. Study alternative approaches to modeling source motion and the impact on localization and tracking performance. Two possible approaches are as follows:

   – Model the sources as nonmoving over a short time interval. The validity of this model depends on the source velocity and the length of the time interval.

   – Describe the source motion over a time interval with a parametric model. Evaluate the tradeoffs in processing complexity and tracking performance associated with each approach.

b. Investigate the performance of algorithms for TDE with partially coherent signals that are measured at widely separated aeroacoustic sensors.

c. Study the impact of compensating for the Doppler effect that is caused by the source motion. Evaluate the performance improvements and processing complexity that are associated with Doppler compensation.

d. Investigate performance bounds on localization accuracy (Cramér-Rao, Ziv-Zakai, and/or Barankin) that include the signal structure and the partial signal coherence at distributed arrays.

e. Investigate algorithms for joint tracking and identification that share information about source location, heading, orientation, and source signal attributes.

   This task was elaborated in the modification of July 2001 as follows:

   Classification of ground vehicles is currently performed using features derived solely from the signal waveforms measured at distributed sensors. Classification bounds will be studied that quantify the amount of additional information that is available if the following data are used in conjunction with the signal waveforms: target aspect angle, range, and heading/velocity (obtained from the tracking algorithm), and coherence level of the signals measured at widely separated sensors. Acoustic waveforms for ground vehicles as a function of range, aspect angle, and heading/velocity will be obtained from existing data files containing moving vehicles and distributed sensors. New measurements may also be required to obtain more complete data sets.

f. The following task was added in the modification of July 2001:

   The previous work has focused on localization of a single source using widely separated sensors with partial spatial coherence. Investigate extensions of the single-source analysis and techniques to multiple-source scenarios.

Task a is addressed in section 5 of this report. Tasks b–d are closely related and are addressed in sections 2–5. Task f is addressed in section 6, where we present an example with two moving sources. Task e relates to classification of ground vehicles and is not addressed in this report. We have developed a theoretical methodology for quantifying the amount of information provided by additional "features for classification [11], but we have not applied the methodology to acoustical signatures of ground vehicles.

This report is organized as follows. The sensor data model is presented in section 2 for the case of a nonmoving source. TDE with partially coherent signals is studied in section 3, and section 4 presents a subspace algorithm for narrowband source localization with distributed arrays and partially coherent signals. The sensor data model is extended to moving sources in section 5, and an algorithm is proposed for tracking with an array of arrays. Section 6 includes an example based on measured aeroacoustic data for two moving ground vehicles. The report concludes in section 7 with a summary and concluding remarks.

2. Data Model for a Nonmoving Source

A model is formulated in this section for the discrete-time signals received by the sensors in an array of arrays. 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 Figure 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

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

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 $\tau_h + \tau_{hn}$, where

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

is the propagation time from the reference sensor on array h to sensor n on array h, and ϕ_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 represented as $s_h(t - \tau_h - \tau_{hn})$, where the vector of signals $\mathbf{s}(t) = [s_1(t), \ldots, s_H(t)]^T$ received at the H arrays 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 function matrix

$$\mathbf{R}_s(\tau) = E\{\mathbf{s}(t+\tau)\,\mathbf{s}(t)^T\}, \tag{3}$$

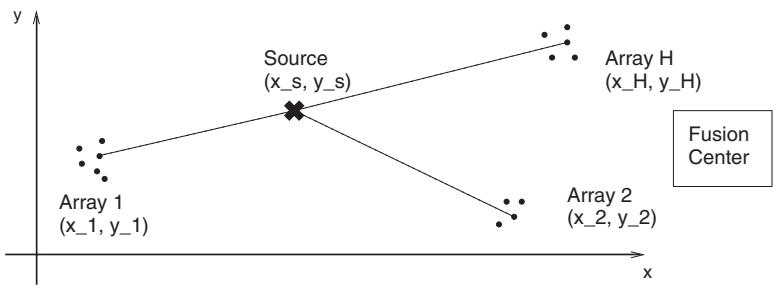


Figure 1. Geometry of nonmoving source location and an array of arrays. A communication link is available between each array and the fusion center.

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

$$r_{s,gh}(\tau) = E\{s_g(t+\tau)\, s_h(t)\} \tag{4}$$

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 functions

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

and the associated cross-spectral density matrix

$$\mathbf{G}_s(\omega) = \mathcal{F}\{\mathbf{R}_s(\tau)\}. \tag{6}$$

The diagonal elements $G_{s,hh}(\omega)$ 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 differences and source aspect angle differences.

The off-diagonal elements of (6), $G_{s,gh}(\omega)$, are the cross-spectral density (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

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

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, while 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 τ_{hn} in (2) and the bearings $\phi_1, \ldots \phi_H$.

The signal received at sensor n on array h is modelled as a sum of the delayed source signal and 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

$$E\{w_{gm}(t+\tau) w_{hn}(t)\} = r_w(\tau)\, \delta_{gh} \delta_{mn} \tag{9}$$

and

$$E\{w_{gm}(t+\tau) s_h(t)\} = 0, \tag{10}$$

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 $\mathbf{z}_h(t) = [z_{h1}(t), \ldots, z_{h,N_h}(t)]^T$ for $h = 1, \ldots, H$, and we further collect the observations from the H arrays into a $(N_1 + \cdots + N_H) \times 1$ vector

$$\mathbf{Z}(t) = \begin{bmatrix} \mathbf{z}_1(t) \\ \vdots \\ \mathbf{z}_H(t) \end{bmatrix}. \tag{11}$$

The elements of $\mathbf{Z}(t)$ in (11) are zero-mean, jointly wide-sense stationary, Gaussian random processes. We can express the CSD matrix of $\mathbf{Z}(t)$ in a convenient form with the following definitions. The array manifold for array h at frequency ω is

$$\mathbf{a}_h(\omega) = \begin{bmatrix} \exp(-j\omega\tau_{h1}) \\ \vdots \\ \exp(-j\omega\tau_{h,N_h}) \end{bmatrix} = \begin{bmatrix} \exp\left[j\frac{\omega}{c}\left((\cos\phi_h)\Delta x_{h1} + (\sin\phi_h)\Delta y_{h1}\right)\right] \\ \vdots \\ \exp\left[j\frac{\omega}{c}\left((\cos\phi_h)\Delta x_{h,N_h} + (\sin\phi_h)\Delta y_{h,N_h}\right)\right] \end{bmatrix}, \tag{12}$$

using τ_{hn} from (2) and assuming that the sensors have omnidirectional response to sources in the plane of the array. Let us define the relative TD of the signal at arrays g and h as

$$D_{gh} = \tau_g - \tau_h, \tag{13}$$

where τ_h is defined in (1). Then the cross-spectral density matrix of $\mathbf{Z}(t)$ in (11) has the form

$$\mathbf{G}_{\mathbf{Z}}(\omega) = \tag{14}$$

$$\begin{bmatrix} \mathbf{a}_1(\omega)\mathbf{a}_1(\omega)^H G_{s,11}(\omega) & \cdots & \mathbf{a}_1(\omega)\mathbf{a}_H(\omega)^H \exp(-j\omega D_{1H}) G_{s,1H}(\omega) \\ \vdots & \ddots & \vdots \\ \mathbf{a}_H(\omega)\mathbf{a}_1(\omega)^H \exp(+j\omega D_{1H}) G_{s,1H}(\omega)^* & \cdots & \mathbf{a}_H(\omega)\mathbf{a}_H(\omega)^H G_{s,HH}(\omega) \end{bmatrix} + G_w(\omega)\mathbf{I}.$$

The source CSD functions $G_{s,gh}(\omega)$ in (14) can be expressed in terms of the signal spectral coherence $\gamma_{s,gh}(\omega)$ using (7). Note that (14) depends on the source location parameters (x_s, y_s) through the bearings ϕ_h in $\mathbf{a}_h(\omega)$ and the pairwise TD differences D_{gh}.

2.1 Cramér-Rao Bound (CRB)

The CRB provides a lower bound on the variance of any unbiased estimator. The problem of interest is estimation of the source location parameter vector $\mathbf{\Theta} = [x_s, y_s]^T$ using T samples of the sensor signals $\mathbf{Z}(0), \mathbf{Z}(T_s), \ldots, \mathbf{Z}((T-1)\cdot T_s)$, where T_s is the sampling period. The total observation time is $\mathcal{T} = T \cdot T_s$. 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

$\mathbf{Z}(t)$ are band-limited, and that the sampling rate f_s is greater than twice the bandwidth of the processes. Then Friedlander and Whittle [12, 13] have shown that the Fisher information matrix (FIM) $\mathbf{J}$ for the parameters $\mathbf{\Theta}$ based on the samples $\mathbf{Z}(0), \mathbf{Z}(T_s), \ldots, \mathbf{Z}((T-1)\cdot T_s)$ has elements

$$J_{ij} = \frac{\mathcal{T}}{4\pi} \int_0^{\omega_s} \mathrm{tr} \left\{ \frac{\partial \mathbf{G}_{\mathbf{Z}}(\omega)}{\partial \theta_i} \mathbf{G}_{\mathbf{Z}}(\omega)^{-1} \frac{\partial \mathbf{G}_{\mathbf{Z}}(\omega)}{\partial \theta_j} \mathbf{G}_{\mathbf{Z}}(\omega)^{-1} \right\} d\omega, \quad i,j = 1,2, \tag{15}$$

where "tr" denotes the trace of the matrix. The CRB matrix $\mathbf{C} = \mathbf{J}^{-1}$ then has the property that the covariance matrix of any unbiased estimator $\hat{\mathbf{\Theta}}$ satisfies $\mathrm{Cov}(\hat{\mathbf{\Theta}}) - \mathbf{C} \geq \mathbf{0}$, where $\geq \mathbf{0}$ means that $\mathrm{Cov}(\hat{\mathbf{\Theta}}) - \mathbf{C}$ is positive semidefinite [14]. Equation (15) 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. The CRB is evaluated for various scenarios in our previous work [6], and additional examples are presented in section 2.2.1 of this report.

Let us consider the CRB for an acoustic source that has a *narrowband* power spectrum, i.e., the PSD $G_{s,hh}(\omega)$ of the signal at each array $h = 1, \ldots, H$ is nonzero only in a narrow band of frequencies $\omega_0 - (\Delta\omega/2) \leq \omega \leq \omega_0 + (\Delta\omega/2)$. If the bandwidth $\Delta\omega$ is chosen small enough so that the ω-dependent quantities in (15) are well approximated by their value at ω_0, then the narrowband approximation to the FIM (15) is

$$J_{ij} \approx \frac{T\Delta\omega}{\omega_s} \mathrm{tr} \left\{ \frac{\partial \mathbf{G}_{\mathbf{Z}}(\omega_0)}{\partial \theta_i} \mathbf{G}_{\mathbf{Z}}(\omega_0)^{-1} \frac{\partial \mathbf{G}_{\mathbf{Z}}(\omega_0)}{\partial \theta_j} \mathbf{G}_{\mathbf{Z}}(\omega_0)^{-1} \right\}. \tag{16}$$

The quantity $\frac{T\Delta\omega}{\omega_s}$ multiplying the FIM in (16) 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 sample each. Then the discrete Fourier transform (DFT) is applied to each block, and the complex coefficients at frequency ω_0 (at each sensor) are used to form M array "snapshots." In this case, the quantity $\frac{T\Delta\omega}{\omega_s}$ is approximately equal to M, the number of snapshots.

2.2 Examples

Next, we present numerical examples that evaluate the CRB on localization accuracy in (15) for a narrowband and a wideband source. We also show measured aeroacoustic data from a ground vehicle and evaluate the spectral coherence of the source at widely separated sensors.

2.2.1 CRB Evaluation

In our previous work [6], we developed CRBs on localization accuracy for the following three schemes.

1. Each array estimates the source bearing and transmits the bearing estimate to the fusion center. The fusion processor then triangulates the

bearings to estimate the source location. This approach does not exploit wavefront coherence between the distributed arrays, but it does minimize the communication bandwidth required to transmit data from the arrays to the fusion center.

2. The raw data from all sensors is jointly processed to estimate the source location. This is the optimum approach and it fully utilizes the coherence between distributed arrays. However, it requires a large communication bandwidth, since the data from all of the sensors must be transmitted to the fusion center.

3. Combination of methods 1 and 2: The objective is to perform some processing at the individual arrays to reduce the communication bandwidth requirement while still exploiting the coherence between distributed arrays. 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 then estimates the propagation time delay between pairs of distributed arrays, and triangulates these time delay estimates with the bearing estimates to localize the source.

The first scheme is "incoherent triangulation" of the bearing estimates, and it is obtained from the CRB in (15) by setting the coherence parameters to zero, $\gamma_{gh} = 0$, in (7) and (14). The second scheme is obtained from the CRB in (15) with nonzero coherence, γ_{gh}, and the CRB for the third scheme is described in [6]. Our objective in this section is to evaluate these CRBs for a narrowband source and a wideband source. We will refer back to these examples in section 3 after developing the threshold coherence analysis for TDE. The analysis in section 3 will show that the CRBs for the narrowband source case are unachievable when there is any loss of signal coherence between arrays, while the CRBs for the wideband source case are achievable when moderate coherence losses occur.

Consider $H = 3$ identical arrays, each of which contains $N_1 = \cdots = N_H = 7$ sensors. Each array is circular with 4-ft radius, with six sensors equally spaced around the perimeter and one sensor 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, $G_{s,hh}(\omega)/G_w(\omega) = 10$ 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

$$\begin{bmatrix} x & y \end{bmatrix} \mathbf{J} \begin{bmatrix} x \\ y \end{bmatrix} = 1, \tag{17}$$

where $\mathbf{J}$ is the FIM in (15). If the errors in (x, y) localization are jointly Gaussian distributed, then the ellipse (17) 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 $(x_3, y_3) = (100, 0)$. One source located at $(x_s, y_s) = (200, 300)$, where the units are meters, as illustrated in Figure 2a. 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 Figure 2a. 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 Figure 2a corresponds to incoherent signals $\gamma_s(\omega) = 0$ and characterizes the performance of the simple method of triangulation using the bearing estimates from the three arrays. Figure 2b shows the ellipse radius 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 Figure 2a because they fall on the lines of the $\times$ 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. Figure 2c shows a closer 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 Figure 2a for joint processing.

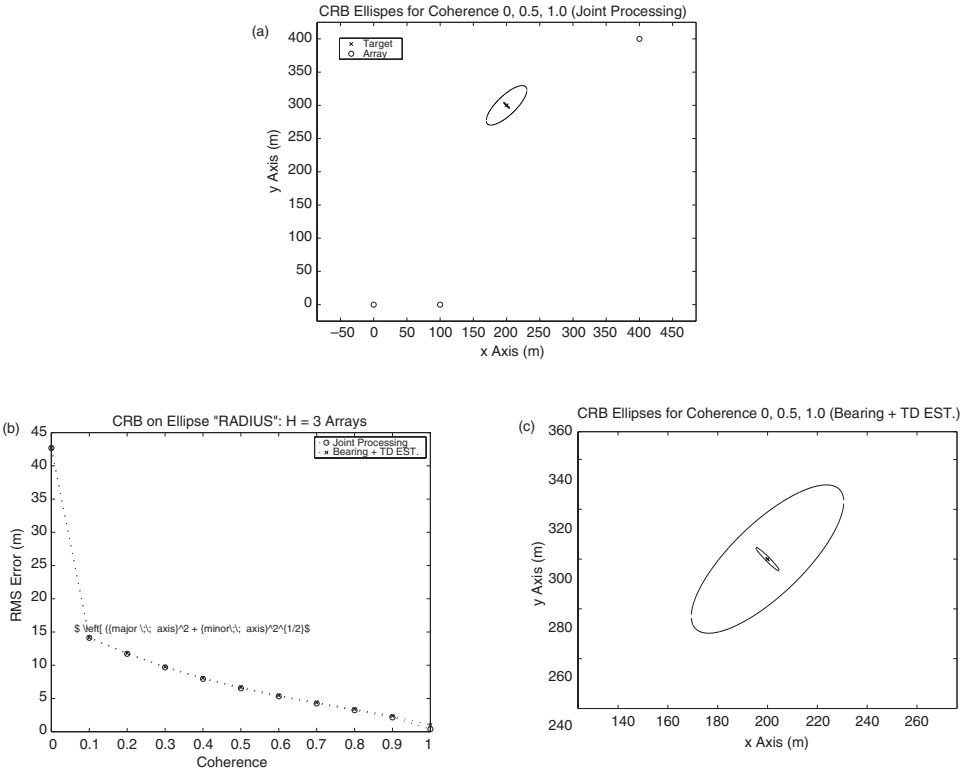


Figure 2. RMS source localization error ellipses based on the CRB for $H = 3$ arrays and one *narrowband* source. The array and source locations are shown in (a), along with the error ellipses for joint processing of all sensor data for coherence values $\gamma_s(\omega) = 0, 0.5$, and 1. Part (b) shows the error ellipse radius (major axis^2 + minor axis^2)$^{1/2}$ for a range of coherence values, comparing joint processing with the reduced-complexity scheme of bearing estimation plus TDE using data from one sensor per array. Part (c) is a closer view of the RMS error ellipses for the bearing plus TDE scheme.

Figure 3 presents corresponding results for a wideband source with bandwidth 20 Hz centered at 50 Hz and SNR 16 dB. That is, $G_{s,hh}/G_w = 40$ for $2\pi(40) < \omega < 2\pi(60)$ rad/s, $h = 1, \ldots, H$. $T = 2000$ time samples are obtained at each sensor with sampling rate $f_s = 2000$ samples/s, so the observation time is 1 s. As in the narrowband case in Figure 2, Figure 3 shows that joint processing reduces the CRB compared with bearings-only triangulation, and bearing plus TD estimation is nearly optimum.

The CRB provides a lower bound on the variance of unbiased estimates, so an important question is whether an estimator can *achieve* the CRB. We show in section 3 that the coherent processing CRBs for the narrowband scenario illustrated in Figure 2 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 Figure 2, the coherent processing CRBs are achievable for coherence values $\gamma_s > 0.75$, so the wideband source provides tolerance to moderate values of coherence loss.

2.2.2 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 Figure 4a, 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, 4-ft radius, and six sensors equally

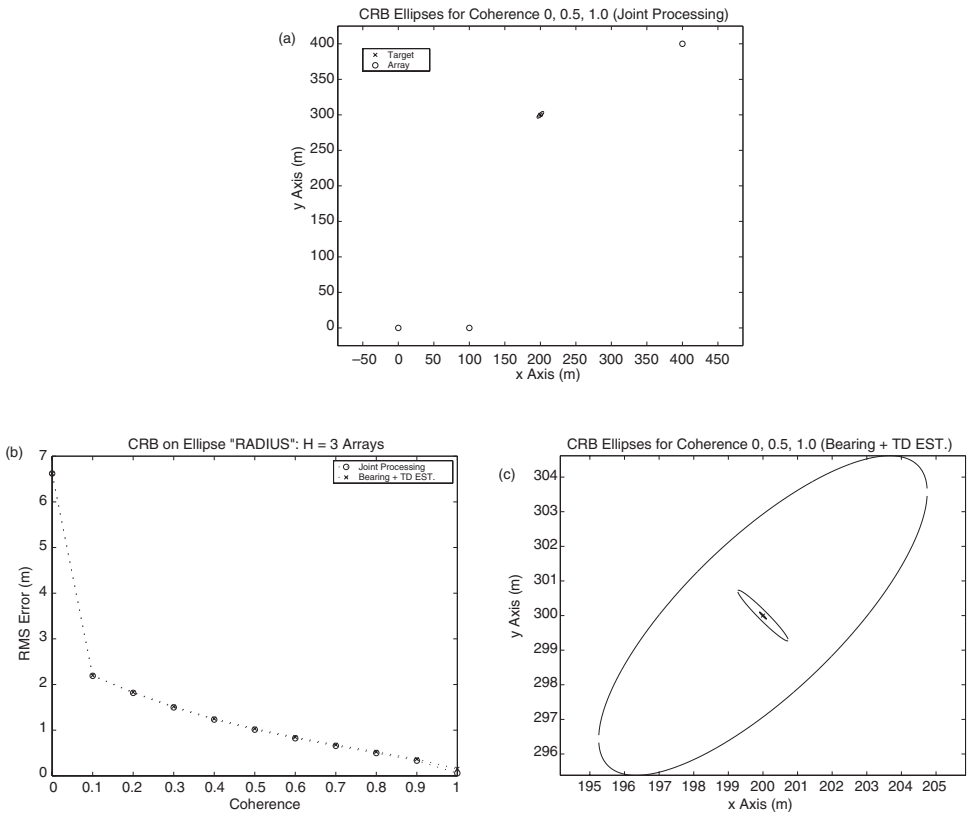


Figure 3. RMS source localization error ellipses based on the CRB for $H = 3$ arrays and one *wideband* source. The array and source locations are shown in (a), along with the error ellipses for joint processing of all sensor data for coherence values $\gamma_s(\omega) = 0, 0.5$, and 1. Part (b) shows the error ellipse radius, comparing joint processing with the reduced-complexity scheme of bearing estimation plus TDE using data from one sensor per array. Part (c) is a closer view of the RMS error ellipses for the bearing plus TDE scheme.

spaced around the perimeter with one sensor in the center. We focus on the 10-s segment indicated by the ◊'s in Figure 4a (which correspond to the time segment 340–350 s in the data). Figure 4b shows the power spectral density (PSD) of the data measured at arrays 1 and 3 during the 10-s segment. Note the dominant harmonic at 39 Hz. Figure 4c 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. Figure 4c 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, so our model of perfect signal coherence over individual arrays seems reasonable.

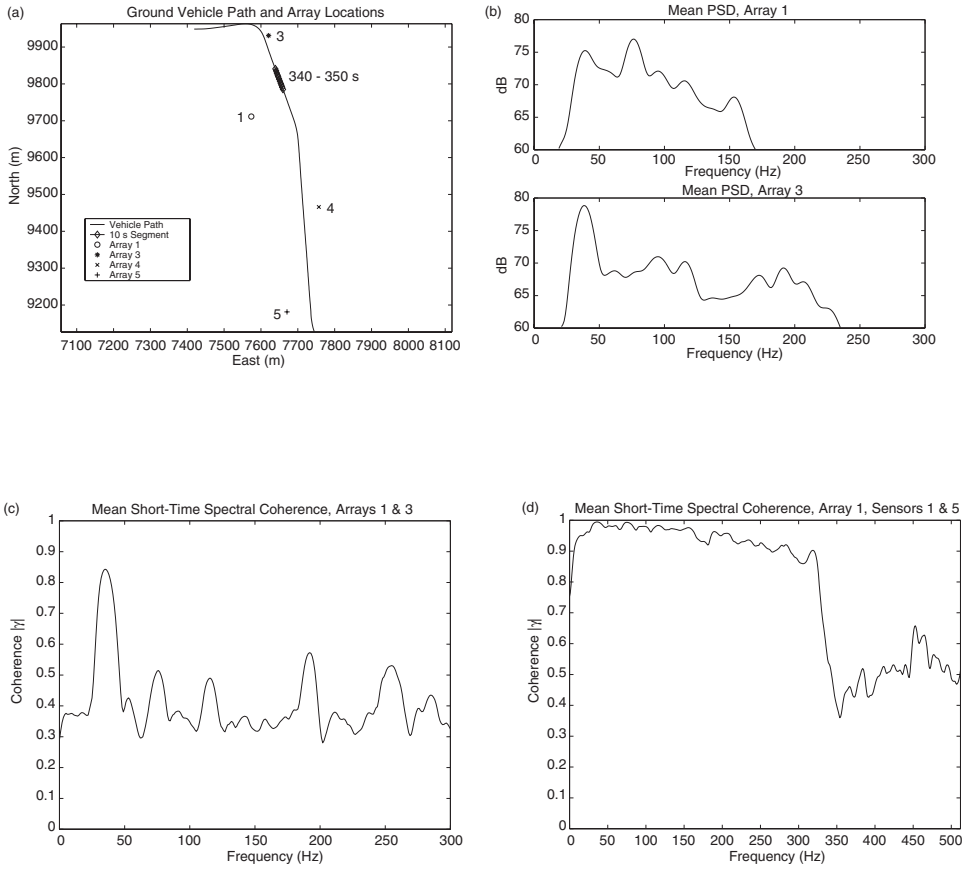


Figure 4. (a) Path of ground vehicle and array locations for measured data. (b) Mean PSD at arrays 1 and 3 estimated from measured data over the 10-s segment ◊ in (b). Top panel is $G_{s,11}(f)$, bottom panel is $G_{s,33}(f)$. (c) Mean spectral coherence $\gamma_{s,13}(f)$ between arrays 1 and 3 estimated over the 10-s segment. (d) Mean spectral coherence for two sensors on array 1, with sensor spacing 8 ft.

The Doppler effect due to source motion was compensated prior to the coherence estimate shown in Figure 4c. Without Doppler compensation, the coherence is significantly reduced, as shown in Figure 5a. The time-varying radial velocity of the source with respect to each array in Figure 4a is plotted in the top panel of Figure 5b. 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 $s(\alpha t)$, where the scaling factor α is

$$\alpha = 1 - \frac{v}{c} \tag{18}$$

and c is the speed of wave propagation. The scaling factor α is plotted in the bottom panel of Figure 5b. Note that for this data set, $0.98 < \alpha < 1.02$, which corresponds to a Doppler frequency shift of approximately ± 1 Hz for an emitted tone at 50 Hz. We use a digital resampling algorithm to compensate for the Doppler effect.

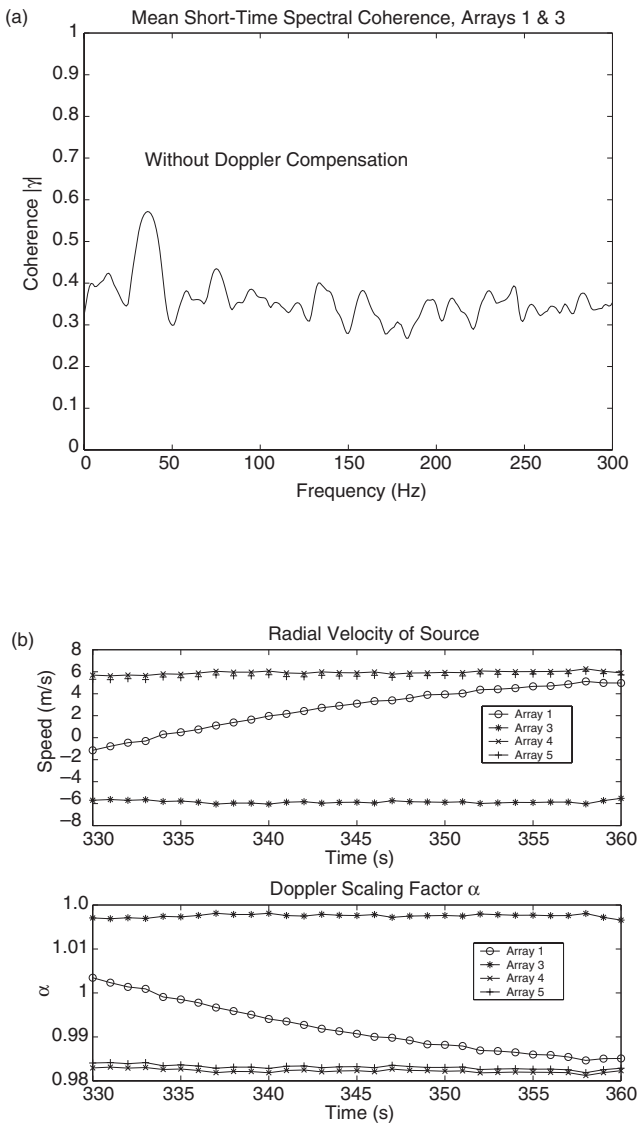


Figure 5. (a) Mean short-time spectral coherence between arrays 1 and 3 if Doppler is *not* compensated. (b) Radial velocity and Doppler scaling factor α in (18) for source in Figure 4a with respect to each array.

3. Time Delay Estimation (TDE)

The CRB results presented in section 2.2.1 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 edited by Carter [15]). The fundamental limits are usually parameterized in terms of the SNR at each sensor, the spectral support of the signals (fractional bandwidth), and the time-bandwidth product of the observations. We are particularly interested in the application of aeroacoustic tracking of ground vehicles using an array of microphone arrays. The coherence between the signals is known to degrade with increased spatial separation between the signals [7, 8] due to random motion of the air. This coherence loss significantly affects the time delay and Doppler estimation accuracy.

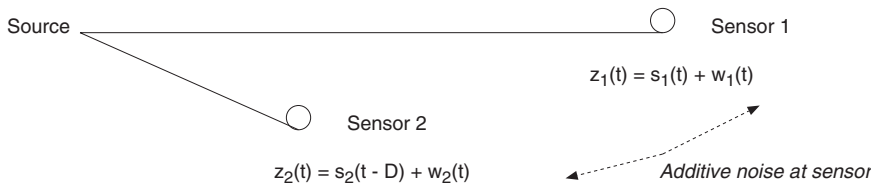
In order to quantify the effect of partial signal coherence on TD and Doppler estimation, we present Cramér-Rao and Ziv-Zakai bounds that are explicitly parameterized by the signal coherence, along with the traditional parameters of SNR, fractional bandwidth, and time-bandwidth product. The basic result 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, as in Weiss and Weinstein [9, 10].

Let us parameterize the model in (14) by the bearings ϕ_h and the time-delay differences D_{gh}, and consider first the case of $H = 2$ sensors. Figure 6 contains an illustration of the TDE problem, with the differential TD defined as $D = D_{21}$. It follows from (14) that the CSD matrix of the sensor data in Figure 6 is

$$\text{CSD}\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \mathbf{G}_{\mathbf{Z}}(\omega) = \tag{19}$$

$$\begin{bmatrix} G_{s,11}(\omega) + G_w(\omega) & e^{+j\omega D}\gamma_{s,12}(\omega)\left[G_{s,11}(\omega)G_{s,22}(\omega)\right]^{1/2} \\ e^{-j\omega D}\gamma_{s,12}(\omega)^*\left[G_{s,11}(\omega)G_{s,22}(\omega)\right]^{1/2} & G_{s,22}(\omega) + G_w(\omega) \end{bmatrix}.$$

Figure 6. TDE problem for a nonmoving source with $H = 2$ sensors.



The signal coherence function $\gamma_{s,12}(\omega)$ describes the degree of correlation that remains in the signal emitted by the source at each frequency ω 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 TD and Doppler estimation. We begin with the magnitude-squared coherence (MSC) [15] 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:

$$\begin{aligned}\mathrm{MSC}_z\left(|\gamma_{s,12}(\omega)|\right) &= \frac{|\mathrm{CSD}[z_1(t), z_2(t)]|^2}{\mathrm{PSD}[z_1(t)] \cdot \mathrm{PSD}[z_2(t)]} = \frac{|\gamma_{s,12}(\omega)|^2 \; G_{s,11}(\omega) G_{s,22}(\omega)}{[G_{s,11}(\omega) + G_w(\omega)] \; [G_{s,22}(\omega) + G_w(\omega)]} \\ &= \frac{|\gamma_{s,12}(\omega)|^2}{\left[1 + \left(\frac{G_{s,11}(\omega)}{G_w(\omega)}\right)^{-1}\right] \left[1 + \left(\frac{G_{s,22}(\omega)}{G_w(\omega)}\right)^{-1}\right]} \leq 1 \, . \end{aligned} \tag{20}$$

Then the following "SNR" expression appears in subsequent performance bounds:

$$\mathrm{SNR}\left(|\gamma_{s,12}(\omega)|\right) = \frac{\mathrm{MSC}_z\left(|\gamma_{s,12}(\omega)|\right)}{1 - \mathrm{MSC}_z\left(|\gamma_{s,12}(\omega)|\right)} \tag{21}$$

$$= \left\{ \frac{1}{|\gamma_{s,12}(\omega)|^2} \left[1 + \left(\frac{G_{s,11}(\omega)}{G_w(\omega)}\right)^{-1}\right] \left[1 + \left(\frac{G_{s,22}(\omega)}{G_w(\omega)}\right)^{-1}\right] - 1 \right\}^{-1} \tag{22}$$

$$\leq \frac{|\gamma_{s,12}(\omega)|^2}{1 - |\gamma_{s,12}(\omega)|^2}. \tag{23}$$

The inequality (23) shows that signal coherence loss ($|\gamma_{s,12}(\omega)| < 1$) severely limits the "SNR" quantity that characterizes performance, even if the SNR per sensor $G_{s,ii}(\omega)/G_w(\omega)$ is very large.

3.1 Bounds for TDE

We can use (19) in (15) to find the CRB for TDE with $H = 2$ sensors, yielding

$$\mathrm{CRB}(D) = \frac{4\pi}{\mathcal{T}} \left[\int_0^{\omega_s} \omega^2 \, \mathrm{SNR}\left(|\gamma_{s,12}(\omega)|\right) \, d\omega \right]^{-1}, \tag{24}$$

where $\mathcal{T}$ is the total observation time of the sensor data and $\mathrm{SNR}\left(|\gamma_{s,12}(\omega)|\right)$ is defined in (22). 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 ω_0 rad/s. If we omit the frequency dependence of $G_{s,11}, G_{s,22}, G_w$, and $\gamma_{s,12}$, then the integral in (24) may be evaluated to yield the following CRB expression:

$$\mathrm{CRB}(D) = \frac{1}{2\omega_0^2\left(\frac{\Delta\omega\,\mathcal{T}}{2\pi}\right)\left[1+\frac{1}{12}\left(\frac{\Delta\omega}{\omega_0}\right)^2\right]\mathrm{SNR}\left(|\gamma_{s,12}|\right)} \tag{25}$$

$$= \frac{1}{2\omega_0^2\left(\frac{\Delta\omega\,\mathcal{T}}{2\pi}\right)\left[1+\frac{1}{12}\left(\frac{\Delta\omega}{\omega_0}\right)^2\right]}\left\{\frac{1}{|\gamma_{s,12}|^2}\left[1+\left(\frac{G_{s,11}}{G_w}\right)^{-1}\right]\left[1+\left(\frac{G_{s,22}}{G_w}\right)^{-1}\right]-1\right\}$$

$$> \frac{1}{2\omega_0^2\left(\frac{\Delta\omega\,\mathcal{T}}{2\pi}\right)\left[1+\frac{1}{12}\left(\frac{\Delta\omega}{\omega_0}\right)^2\right]}\left[\frac{1}{|\gamma_{s,12}|^2}-1\right]. \tag{26}$$

The quantity $\left(\frac{\Delta\omega\cdot\mathcal{T}}{2\pi}\right)$ is the time-bandwidth product of the observations, $\left(\frac{\Delta\omega}{\omega_0}\right)$ is the fractional bandwidth of the signal, and $G_{s,hh}/G_w$ is the SNR at sensor h. Note from the high-SNR limit in (26) that when the signals are partially coherent $|\gamma_{s,12}| < 1$, increased source power does not reduce the CRB. Improved TDE accuracy is obtained with partially coherent signals by increasing the observation time $\mathcal{T}$ 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, so 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 in section 5 we extend the model to include source motion.

The CRB in (25) agrees with known results for perfectly coherent signals [15] and with results from the medical ultrasound literature [16, 17] for partially correlated speckle signals. The medical ultrasound application is distinguished from the aeroacoustic tracking of ground vehicles in that the former is typically an *active* system while the latter is *passive*. The medical ultrasound application therefore allows much more control over the SNR and bandwidth of the signals. In passive aeroacoustics, the received signals are emitted by a vehicle and are not controllable for the purposes of TD estimation.

With perfectly coherent signals, it is well known that the CRB on TDE is achievable only when the SNR expression in (22) (with $|\gamma_{s,12}(\omega)| = 1$) exceeds a threshold [9, 10]. 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 (24) 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 ω_0 rad/s, and the observation time is $\mathcal{T}$ seconds. Further, assume that the signal PSDs are identical at each sensor, and define the following constants for notational simplicity:

$$G_{s,11}(\omega_0) = G_{s,22}(\omega_0) = G_s,\ G_w(\omega_0) = G_w,\ \text{and}\ \gamma_{s,12}(\omega_0) = \gamma_s. \tag{27}$$

Then the SNR expression in (22) has the form

$$\mathrm{SNR}(|\gamma_s|) = \left[\frac{1}{|\gamma_s|^2} \left(1 + \frac{1}{(G_s/G_w)} \right)^2 - 1 \right]^{-1}. \tag{28}$$

The Ziv-Zakai bound developed by Weiss and Weinstein [9, 10] shows that the threshold SNR for CRB attainability is a function of the time-bandwidth product $\left(\frac{\Delta\omega \cdot \mathcal{T}}{2\pi}\right)$ and the fractional bandwidth $\left(\frac{\Delta\omega}{\omega_0}\right)$,

$$\mathrm{SNR}_{\mathrm{thresh}} = \frac{6}{\pi^2 \left(\frac{\Delta\omega \mathcal{T}}{2\pi}\right)} \left(\frac{\omega_0}{\Delta\omega}\right)^2 \left[\varphi^{-1} \left(\frac{1}{24} \left(\frac{\Delta\omega}{\omega_0} \right)^2 \right) \right]^2, \tag{29}$$

where $\varphi(y) = 1/\sqrt{2\pi} \int_y^\infty \exp(-t^2/2)\, dt$. It follows that the threshold coherence value is

$$|\gamma_s|^2 \geq \frac{\left(1 + \frac{1}{(G_s/G_w)}\right)^2}{1 + \frac{1}{\mathrm{SNR}_{\mathrm{thresh}}}}, \quad \text{so} \quad |\gamma_s|^2 \geq \frac{1}{1 + \frac{1}{\mathrm{SNR}_{\mathrm{thresh}}}} \ \text{ as } \ \frac{G_s}{G_w} \to \infty. \tag{30}$$

For a specific TDE scenario, the threshold SNR for CRB attainability is given by (29), and (30) provides a corresponding threshold coherence. Since $|\gamma_s|^2 \leq 1$, (30) is useful only if $G_s/G_w > \mathrm{SNR}_{\mathrm{thresh}}$.

Figure 7 contains a plot of (30) 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 $\mathcal{T} = 2$ s. Figure 7 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., $\Delta\omega \geq 2\pi 10$ rad/s, a certain amount of coherence loss can be tolerated while still allowing unambiguous TDE.

Figures 8a–c contain plots of the threshold coherence in (30) as a function of the time-bandwidth product $\left(\frac{\Delta\omega \cdot \mathcal{T}}{2\pi}\right)$, SNR $\frac{G_s}{G_w}$, and fractional bandwidth $\left(\frac{\Delta\omega}{\omega_0}\right)$. Note that $\frac{G_s}{G_w} = 10$ dB is nearly equivalent to $\frac{G_s}{G_w} \to \infty$. We note that very large time-bandwidth product is required to overcome coherence loss when the fractional bandwidth is small at 0.1. The variation of threshold coherence with fractional bandwidth is illustrated in Figure 8d. For a fixed threshold coherence value, such as 0.7, each doubling of the fractional bandwidth reduces the required time-bandwidth product by about a factor of 10.

Let us examine a narrowband signal scenario that is typical in aeroacoustics, with center frequency $f_0 = \omega_0/(2\pi) = 50$ Hz and bandwidth $\Delta f = \Delta\omega/(2\pi) = 5$ Hz, so the fractional bandwidth is $\Delta f / f_0 = 0.1$. From Figure 8a, signal coherence $|\gamma_s| = 0.8$ requires time-bandwidth product $\Delta f \cdot \mathcal{T} > 200$, so the necessary time duration $\mathcal{T} = 40$ s for TDE may be impractical for moving sources.

Figure 7. Threshold coherence vs. bandwidth based on (30) for $\omega_0 = 2\pi 50$ rad/s and $G_s/G_w = 0, 10$, and ∞ dB.

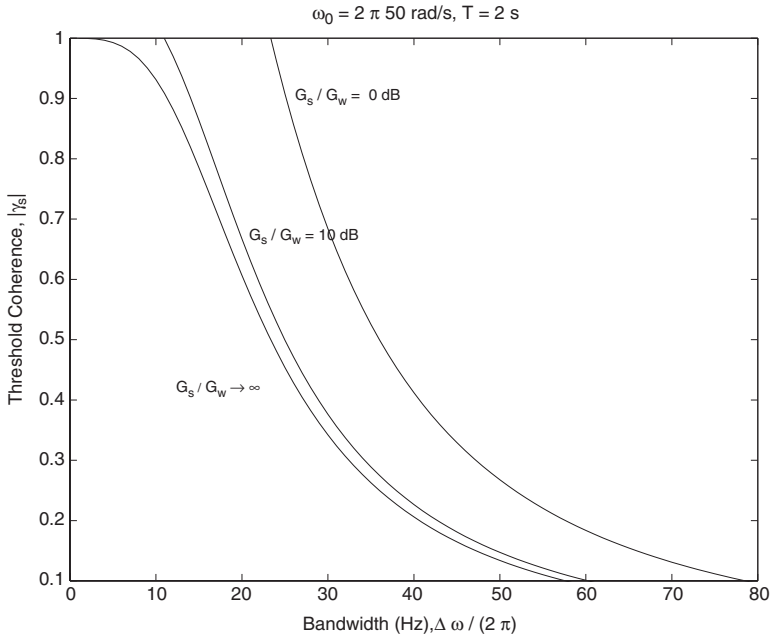
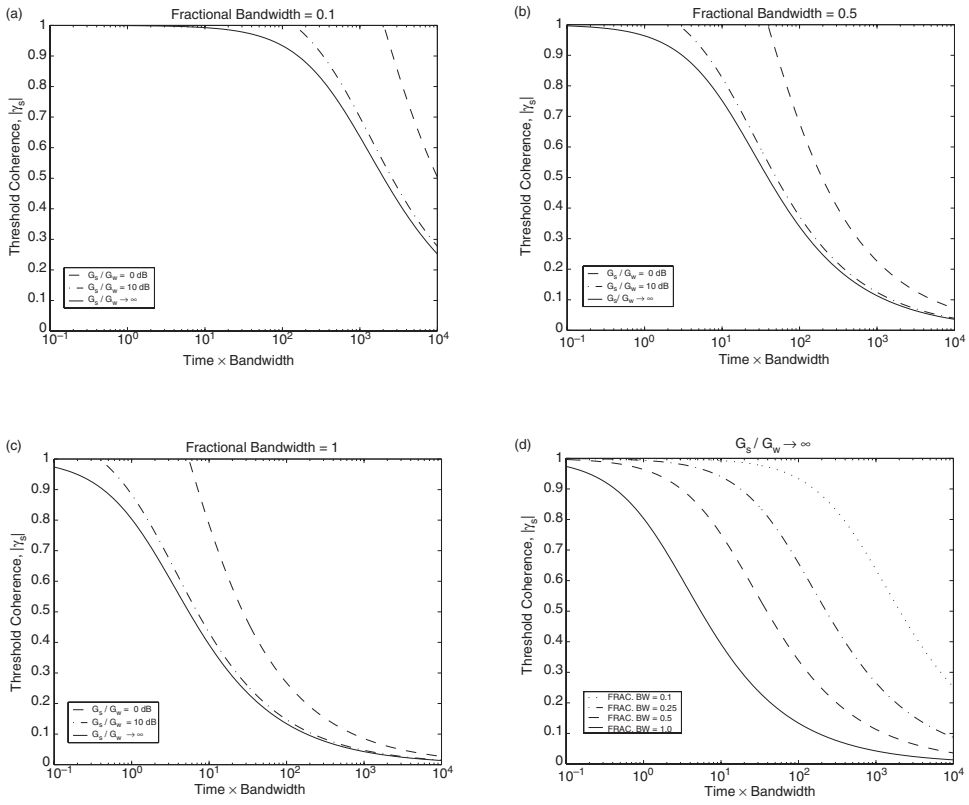


Figure 8. Threshold coherence value from (30) vs. time-bandwidth product $\left(\frac{\Delta\omega \cdot \mathcal{T}}{2\pi}\right)$ and SNR G_s/G_w for fractional bandwidth values $\left(\frac{\Delta\omega}{\omega_0}\right)$ (a) 0.1, (b) 0.5, and (c) 1.0. In (d), the high SNR curves $G_s/G_w \to \infty$ are superimposed for several values of fractional bandwidth.



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 section 5 for use during long observation intervals.

We can evaluate the threshold coherence for the narrowband and wideband scenarios considered in section 2.2.1 for the CRB examples. The results are as follows, using (29) and (30):

- Narrowband case: $G_s/G_w = 10$, $\omega_0 = 2\pi 50$ rad/s, $\Delta\omega = 2\pi$ rad/s, $\mathcal{T} = 2$ s
  $\Longrightarrow$ Threshold coherence ≈ 1.
- Wideband case: $G_s/G_w = 40$, $\omega_0 = 2\pi 50$ rad/s, $\Delta\omega = 2\pi \cdot 20$ rad/s, $\mathcal{T} = 1$ s
  $\Longrightarrow$ Threshold coherence ≈ 0.75.

Therefore, for the narrowband case, joint processing of the data from different arrays will not achieve the CRBs in Figure 2 when there is any loss in signal coherence. For the wideband case, joint processing can achieve the CRBs in Figure 3 for coherence values that exceed 0.75.

The remainder of this section continues to focus on nonmoving sources, with a simulation example presented in section 3.2 that verifies the CRB and threshold coherence values for TDE. In section 3.3, the $H = 2$ sensor case of this subsection is extended to TDE with $H > 2$ sensors. Section 3.4 contains examples of TDE with measured aeroacoustic data.

3.2 TDE Simulation Examples

Consider TDE at $H = 2$ sensors with varying signal coherence γ_s. Our first simulation example involves a signal with reasonably wide bandwidth, $\Delta f = 30$ Hz centered at $f_0 = 100$ Hz, so 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 band-pass Gaussian random processes. The sampling rate in the simulation is $f_s = 10^4$ samples/s, with $T = 3 \times 10^4$ samples, so the time interval length is $\mathcal{T} = 3$ s.

Figure 9a displays the simulated RMS error on TDE for $0.2 \leq \gamma_s \leq 1.0$, along with the corresponding CRB from (25). 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 (30) and (29). Note in Figure 9a that the simulated RMS error on TDE diverges sharply from the CRB very near to the threshold coherence value of 0.41, illustrating the accuracy of the analytical threshold coherence in (30).

Next we consider TDE with three different signals:

- A narrowband signal with $\Delta f = 2$ Hz centered at $f_0 = 40$ Hz. We refer to this as "1 harmonic" (it is the fundamental frequency of the signals defined next).

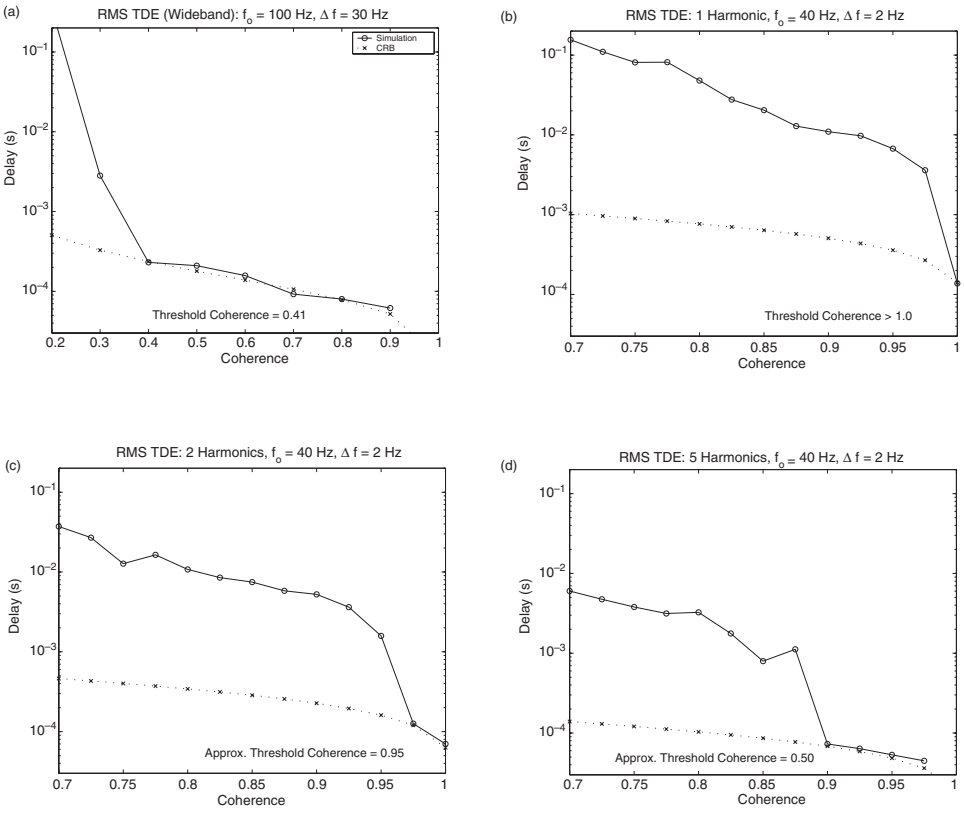


Figure 9. Comparison of simulated RMS error for TDE with CRBs and threshold coherence value. (a) Wideband signal with $\Delta f = 30$ Hz centered at $f_0 = 100$ Hz. (b–d) Narrowband signal with $\Delta f = 2$ Hz, fundamental frequency $f_0 = 40$ Hz and 1, 2, and 5 harmonic components, respectively.

- "2 harmonics" at 40 and 80 Hz, with bandwidth $\Delta f = 2$ Hz at each harmonic.
- "5 harmonics" at 40, 80, 120, 160, 200 Hz, with bandwidth $\Delta f = 2$ Hz at each harmonic.

The signal, noise, and coherence are flat over each frequency band, with SNR $G_s/G_w = 100$ (20 dB), and the signals and noise are band-pass Gaussian random processes. The sampling rate in the simulation is $f_s = 10^4$ samples/s, with $T = 2 \times 10^4$ samples, so the time interval length is $\mathcal{T} = 2$ s.

Figures 9b–d display 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 example, the TDE is obtained by cross correlation. The threshold coherence is defined only for the "1 harmonic" signal, and the threshold coherence value is ≈ 1. Figure 9b illustrates the divergence of the simulated RMS error from the CRB, except at $\gamma_s = 1$.

Figures 9c and d display the results for the signals with "2 harmonics" and "5 harmonics." The additional harmonics enable accurate TDE for lower coherence values, but we cannot use (30) to compute the analytical threshold coherence for the harmonic signals. The "approximate threshold coherence" values indicated in Figures 9c and d are computed as follows. For the K harmonics, suppose the *total bandwidth* of all harmonics, $\Delta f = K \cdot 2$ Hz, is centered at the fundamental frequency $f_0 = 40$ Hz. The approximate threshold coherence values, 0.95 in Figure 9c and 0.50 in Figure 9d, are

considerably lower than the actual points of divergence from the CRB. Not surprisingly, the total bandwidth that is "spread" across the harmonics is less useful for overcoming signal coherence loss than an equivalent bandwidth concentrated at the fundamental $f_0 = 40$ Hz. Narrowband and harmonic signals are generally difficult for TDE due to ambiguous peaks in the cross-correlation function.

3.3 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 [18], 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 next, these $H - 1$ estimates are nearly as accurate for source localization as forming *all pairs* of TDEs D_{gh} for all $g < h$. Weinstein's analysis [18] is valid for moving as well as nonmoving sources.

Extending (27) and (28) to $H > 2$ sensors, let us assume equal $G_{s,hh}/G_w$ at all sensors $h = 1, \ldots, H$ and equal coherence γ_s between all sensor pairs, so that the $\mathrm{SNR}(|\gamma_s|)$ in (28) is equal for all sensor pairs. Then as long as $H \cdot \mathrm{SNR}(|\gamma_s|) \gg 1$, improves the source localization variance relative to pairwise processing by the factor

$$V = \frac{H\left(1 + 2 \cdot \frac{\gamma_s}{1-\gamma_s}\right)}{2\left(1 + H \cdot \frac{\gamma_s}{1-\gamma_s}\right)}. \tag{31}$$

Clearly $V \to 1$ as $\gamma_s \to 1$, and $V < (3H)/[2(1 + H)] < 1.5$ 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 $D_{1,H}, \ldots, D_{H-1,H}$, 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 but more centralized processing is for all H sensors to communicate their data to the fusion center, with TDE performed at the fusion center.

3.4 TDE With Measured Data

First, we present an illustration based on processing the measured data for the source in Figure 4a that was discussed in section 2.2.2. Figure 10 shows results of cross-correlation processing of the data for a 2-s segment at time

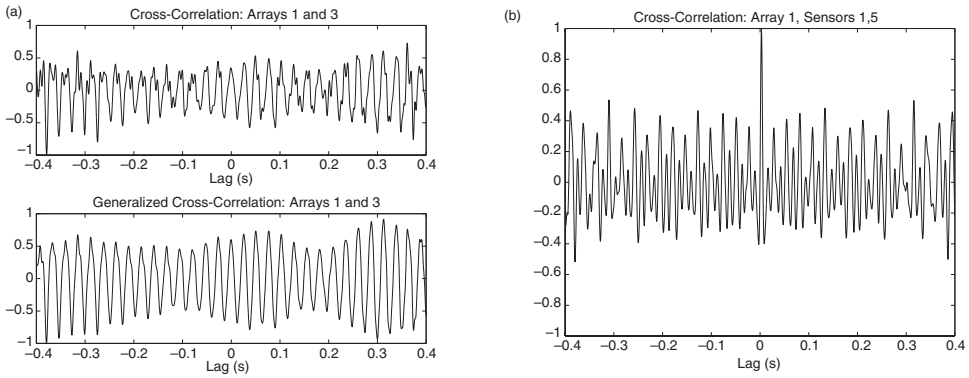


Figure 10. (a) Cross correlation of signals at arrays 1 and 3 in Figure 4a for source at time 342 s. (b) Cross correlation of signals at two sensors on array 1, separated by 8 ft.

342 s. Figure 10a is obtained by cross-correlating the signals received at arrays 1 and 3, for which the coherence is appreciable only over a narrow band near 39 Hz (see Figure 4c). 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 time delay estimation in this scenario. Figure 10b is obtained by cross correlating the signals received at two sensors on array 1, where the coherence is large over a wide bandwidth (see Figure 4d). The peak is clearly evident in the cross correlation in Figure 10b.

Next, we present a TDE example based on measured data from Sanders Corporation [19] using a synthetically generated, nonmoving, wideband acoustic source. The PSD of the source is shown in Figure 11, which in-

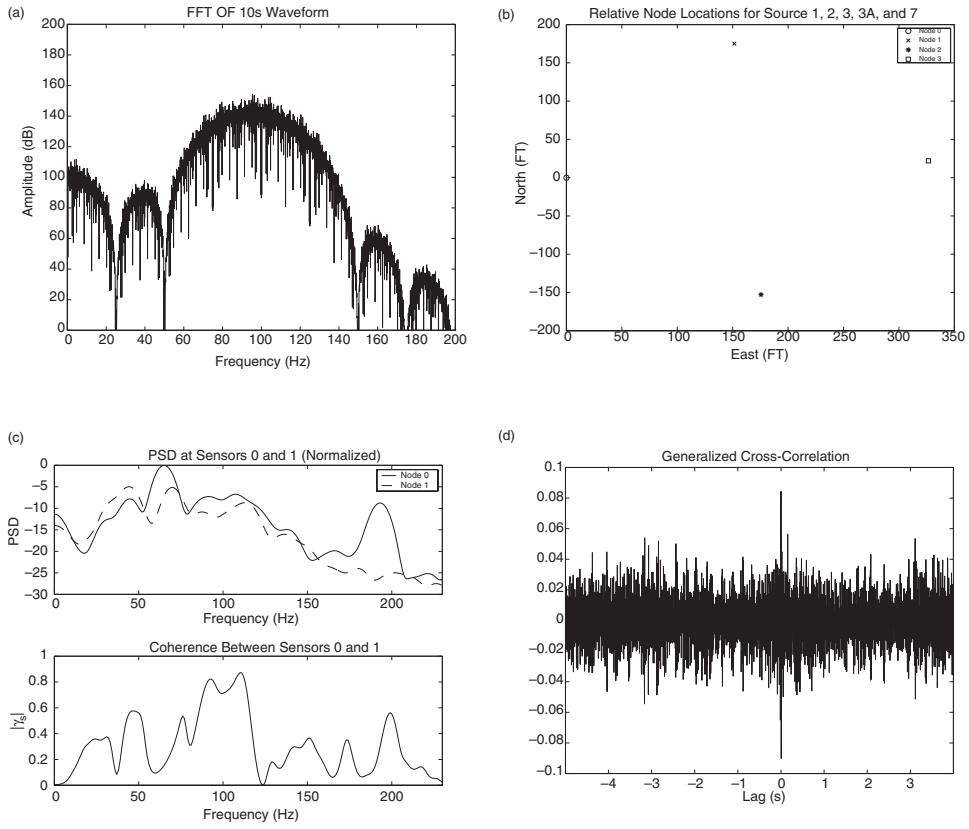


Figure 11. (a) Frequency spectrum of wideband signal. (b) Location of nodes where transmitters and receivers are placed. (c) PSD and coherence for synthetically generated wideband source located at node 2 and measured at nodes 0 and 1. (d) Generalized cross correlation of 1-s segments from node 0 with a 10-s segment from node 1. The data is time-aligned so the true peak location is zero lag.

dicates that the source bandwidth is about 50 Hz with center frequency 100 Hz. With reference to the sensor locations in Figure 11b, 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 Figure 11. Note that the PSDs of the sensor signals do *not* have their maxima at 100 Hz due to the acoustic propagation. However, the coherence magnitude is roughly 0.8 over a 50-Hz band centered at 100 Hz.

Figure 12 shows the threshold coherence computed with (29) and (30) for the signal in Figure 11a that is centered at $\omega_0/2\pi = 100$ Hz and $\mathcal{T} = 1$-s observation time. For bandwidth $\Delta\omega/2\pi = 50$ Hz, the threshold coherence in Figure 12 is approximately 0.5. The actual coherence of 0.8 in Figure 11c exceeds the threshold value, so TDE between sensors 0 and 1 should be feasible. Figure 11d 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.

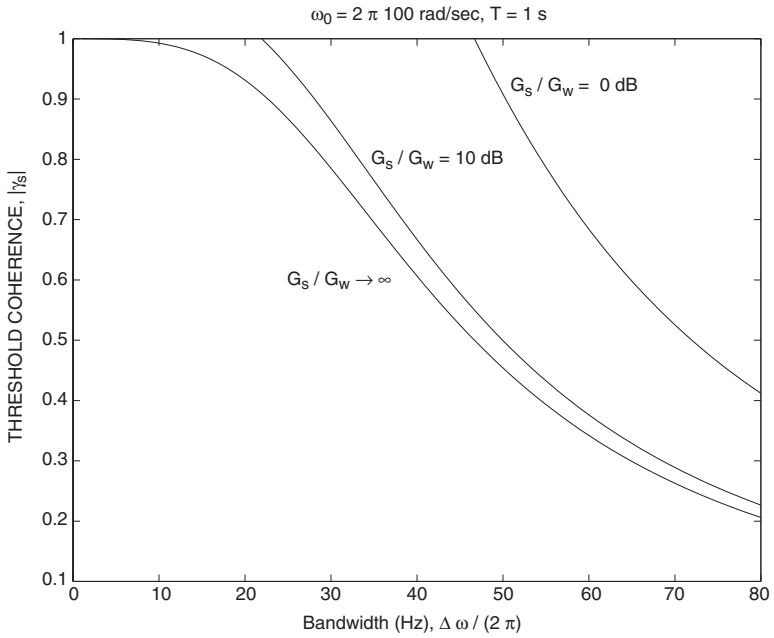


Figure 12. Threshold coherence for synthetic wideband source with PSD in Figure 11a, which has center frequency $\omega_0 = 2\pi 100$ rad/s and observation time is $\mathcal{T} = 1$ s.

Figure 13 contains another example from the Sanders data using the sensor locations in Figure 11b and the wideband source with spectrum in Figure 11a. 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 Figure 13a, which is similar to our observation about the PSDs in Figure 11c. The signal coherence between nodes 1 and 3 is shown in Figure 13b, indicating high coherence over an appreciable bandwidth. The cross correlation is shown in Figure 13, and the peak is clearly evident at the correct location.

Figure 14 contains a final example using the Sanders data in which the source is at node 0 and measurements are recorded at nodes 1, 2, and 3 (see Figure 11b for the node locations). Differential TDs are estimated using the signals measured at nodes 1, 2, and 3, and the TDs are hyperbolically triangulated to estimate the location of the source. Figure 14a shows the hyperbolas obtained from the three differential TD estimates, and Figure 14b 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 measured data in an open field. The source is an M1 tank that is moving at a range of approximately 140 m from a set of $H = 3$ sensor arrays. The sensor arrays are labeled A, B, and C, and they are located along a line with the following separations:

$$B \quad \leftarrow \quad 15\,\mathrm{m} \quad \rightarrow \quad A \quad \leftarrow \quad 8\,\mathrm{m} \quad \rightarrow \quad C.$$

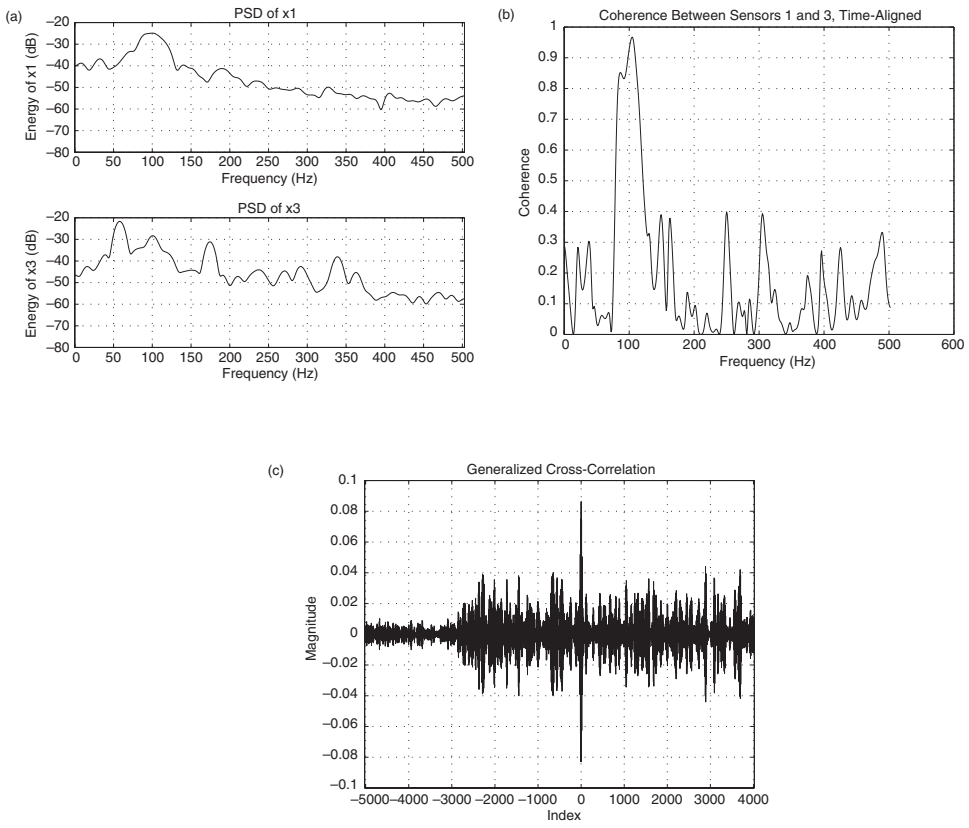
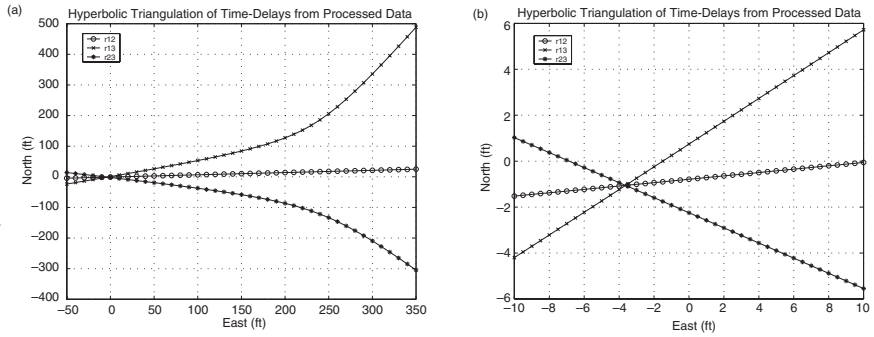


Figure 13. (a) PSDs at nodes 1 and 3 when transmitter is at node 0. (b) Coherence between nodes 1 and 3. (c) Generalized cross correlation between nodes 1 and 3.

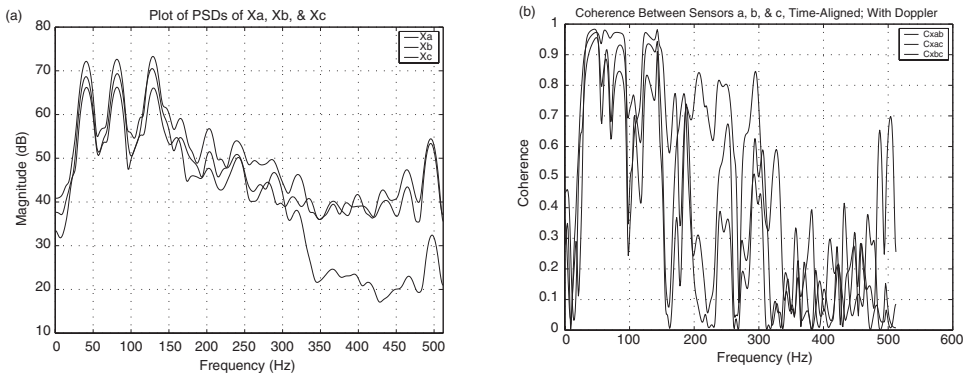
Figure 14.
(a) Intersection of hyperbolas obtained from differential TDs estimated at nodes 1, 2, and 3 in Figure 11b.
(b) Expanded view of part (a) near the point of intersection.

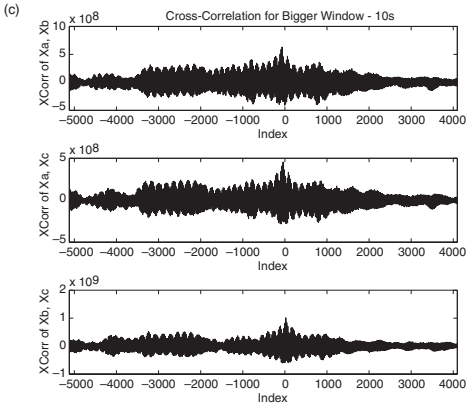


The source is moving parallel to the line connecting the three arrays. Figure 15 shows the PSDs at each array in (a), the signal coherence between pairs of arrays in (b), and the cross correlations in (c). The PSDs in Figure 15a exhibit strong harmonics, perhaps due to tread slap. The coherence in Figure 15b is high over a rather large bandwidth, so the cross-correlation functions in Figure 15c have a clear peak at the correct location.

The examples presented in this section have demonstrated success and failure of time delay estimation, depending on the coherence and bandwidth of the measured signals.

Figure 15. Results for M1 tank moving at range 140 m from three arrays A, B, and C separated by 8, 15, and 23 m. (a) Power spectral densities. (b) Signal coherence between pairs of arrays. (c) Cross correlations between pairs of arrays.





4. Subspace Processing

In this section, we begin with an eigenanalysis of the cross-spectral density matrix (14) for the case of $H = 2$ arrays containing $N_1 = N_2 = N$ sensors operating in a narrow band of frequencies centered at ω_0. This analysis leads to a MUSIC-like subspace algorithm for source localization with distributed arrays and partial signal coherence. We illustrate with computer simulations and measured aeroacoustic data that the new subspace algorithm is limited in performance due to source location ambiguities that arise from the large separation between arrays and the narrowband signals. These ambiguities are identical to those that we studied with the Ziv-Zakai bounds for TDE in section 3.1.

4.1 Eigenanalysis and Algorithm

To simplify notation, we will let $\mathbf{a}_h$ represent the array manifold $a_h(\omega_0), \sigma_h^2$ represent the average signal power at array h in the frequency band of interest, σ_w^2 is the average noise power, and γ is the coherence $\gamma_{s,12}(\omega_0)$. We allow γ to be complex-valued with $|\gamma| \leq 1$. Then (14) can be expressed as the correlation matrix

$$\mathbf{R}^{(12)} = \left(\begin{bmatrix} \sigma_1^2 & \gamma\sigma_1\sigma_2 \\ \gamma^*\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \otimes \mathbf{1}_{N\times N} \right) \circ \begin{bmatrix} \mathbf{a}_1\,\mathbf{a}_1^H & \mathbf{a}_1\,\mathbf{a}_2^H\,\exp(-j\omega_0 D_{12}) \\ \mathbf{a}_2\,\mathbf{a}_1^H\,\exp(+j\omega_0 D_{12}) & \mathbf{a}_2\,\mathbf{a}_2^H \end{bmatrix} + \sigma_w^2\mathbf{I}, \tag{32}$$

where $\otimes$ denotes Kronecker product, $\circ$ denotes elementwise product, and $\mathbf{1}_{N\times N}$ is an $N \times N$ matrix of 1's. The superscript (12) in (32) indicates that the quantity is the combined correlation matrix of all sensors in arrays 1 and 2. The correlation matrices for the individual arrays 1 and 2 are

$$\mathbf{R}^{(\mathrm{h})} = \sigma_h^2\,\mathbf{a}_h\,\mathbf{a}_h^H + \sigma_w^2\mathbf{I}, \quad h = 1, 2. \tag{33}$$

It is well known that the individual array correlation matrices $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$ in (33) have a rank-1 signal subspace, and that the signal subspace eigenvalue is

$$\lambda_1^{(h)} = N\sigma_h^2 + \sigma_w^2, \quad h = 1, 2, \tag{34}$$

with corresponding eigenvector

$$\mathbf{e}_1^{(h)} = \mathbf{a}_h, \quad h = 1, 2. \tag{35}$$

The remaining eigenvalues $\lambda_n^{(h)}$, $n = 2, \ldots, N$ are equal to σ_w^2.

Now, we present the eigenanalysis of the combined correlation matrix $\mathbf{R}^{(12)}$ in (32). The derivation of these results is similar to the reasoning in the appendix of [20]. We assume that the arrays are labeled so that $\sigma_1^2 \geq \sigma_2^2$, and

that the sensors are omnidirectional so that the array manifold elements are of the form $[\mathbf{a}_h]_n = \exp(-j\omega_0\tau_{hn})$. The rank of the signal subspace component of $\mathbf{R}^{(12)}$ is at most two, and the two largest eigenvalues are

$$\lambda_1^{(12)} = \frac{N}{2}\left(\sigma_1^2+\sigma_2^2\right)\left[1+\sqrt{1-(1-|\gamma|^2)\left(\frac{2}{\frac{\sigma_1}{\sigma_2}+\frac{\sigma_2}{\sigma_1}}\right)^2}\right]+\sigma_w^2 \quad (36)$$

$$\lambda_2^{(12)} = \frac{N}{2}\left(\sigma_1^2+\sigma_2^2\right)\left[1-\sqrt{1-(1-|\gamma|^2)\left(\frac{2}{\frac{\sigma_1}{\sigma_2}+\frac{\sigma_2}{\sigma_1}}\right)^2}\right]+\sigma_w^2. \quad (37)$$

The corresponding eigenvectors are

$$\mathbf{e}_1^{(12)} = \begin{bmatrix} \frac{1}{2}\left(\frac{\sigma_1}{\sigma_2}-\frac{\sigma_2}{\sigma_1}\right)\left[1+\sqrt{1+\left(\frac{2|\gamma|}{\frac{\sigma_1}{\sigma_2}-\frac{\sigma_2}{\sigma_1}}\right)^2}\right]\mathbf{a}_1 \\ \gamma^*\exp(j\omega_0 D_{12})\,\mathbf{a}_2 \end{bmatrix} \quad (38)$$

$$\mathbf{e}_2^{(12)} = \begin{bmatrix} -\gamma\,\mathbf{a}_1 \\ \frac{1}{2}\left(\frac{\sigma_1}{\sigma_2}-\frac{\sigma_2}{\sigma_1}\right)\left[1+\sqrt{1+\left(\frac{2|\gamma|}{\frac{\sigma_1}{\sigma_2}-\frac{\sigma_2}{\sigma_1}}\right)^2}\right]\exp(j\omega_0 D_{12})\,\mathbf{a}_2 \end{bmatrix}, \quad (39)$$

provided $\sigma_1 \neq \sigma_2$. Some observations are as follows.

- In the case that $\sigma_1 = \sigma_2 = \sigma_s$, (36)–(39) reduce to

$$\lambda_1^{(12)} = N\sigma_s^2(1+|\gamma|)+\sigma_w^2, \qquad \lambda_2^{(12)} = N\sigma_s^2(1-|\gamma|)+\sigma_w^2 \quad (40)$$

$$\mathbf{e}_1^{(12)} = \begin{bmatrix} \mathbf{a}_1 \\ (\gamma^*/|\gamma|)\,\exp(j\omega_0 D_{12})\,\mathbf{a}_2 \end{bmatrix}, \quad \mathbf{e}_2^{(12)} = \begin{bmatrix} -(\gamma/|\gamma|)\,\mathbf{a}_1 \\ \exp(j\omega_0 D_{12})\,\mathbf{a}_2 \end{bmatrix}. \quad (41)$$

  Note that (41) is valid as long as $\gamma \neq 0$, and in this case the eigenvectors are independent of the coherence magnitude. In the case of high SNR so that $\sigma_w^2 \ll N\sigma_s^2(1-|\gamma|)$, the ratio of signal eigenvalues is approximately

$$\frac{\lambda_1^{(12)}}{\lambda_2^{(12)}} \approx 1+2\frac{|\gamma|}{1-|\gamma|}. \quad (42)$$

  From (42), in order for $\lambda_1^{(12)} > 10\lambda_2^{(12)}$ so that $\lambda_1^{(12)}$ is the dominant eigenvalue, the coherence must satisfy $|\gamma| > 0.8181$.

- In the case of perfect coherence $|\gamma| = 1, \mathbf{R}^{(12)}$ has a rank-1 signal subspace, with signal eigenvalue and eigenvector

$$\lambda_1^{(12,COH)} = N(\sigma_1^2+\sigma_2^2)+\sigma_w^2, \quad \mathbf{e}_1^{(12,\mathrm{COH})} = \begin{bmatrix} \sigma_1\,\mathbf{a}_1 \\ \sigma_2\gamma^*\exp(j\omega_0 D_{12})\,\mathbf{a}_2 \end{bmatrix}. \quad (43)$$

The expressions in (43) are the natural extension of (34) and (35) applied to the "superarray" of all sensors, allowing for different signal power levels at the two arrays.

- In the case of partial coherence $|\gamma| < 1$, $\mathbf{R}^{(12)}$ has a rank-2 signal subspace. The signal eigenvalues are ordered $\lambda_1^{(12)} \geq \lambda_2^{(12)}$, with equality if and only if $\gamma = 0$ and $\sigma_1 = \sigma_2$. In the incoherent case $\gamma = 0$, the signal eigenvectors are

$$\mathbf{e}_1^{(12,\mathrm{INC})} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{e}_2^{(12,\mathrm{INC})} = \begin{bmatrix} \mathbf{0} \\ \mathbf{a}_2 \end{bmatrix}. \tag{44}$$

  Thus, in the incoherent case, the signal eigenvectors (44) of $\mathbf{R}^{(12)}$ are completely determined by the signal eigenvectors (35) of individual array correlation matrices $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$.

- Note that the span of the $\mathbf{R}^{(12)}$ signal subspace eigenvectors in (38) and (39) is *equal* to the span of the incoherent eigenvectors in (44). Therefore, for $|\gamma| < 1$, the signal subspace of $\mathbf{R}^{(12)}$ is spanned by the signal subspace eigenvectors of the *individual* array correlation matrices $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$. It follows that $\mathbf{R}^{(12)}$ and the block diagonal matrix $\begin{bmatrix} \mathbf{R}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{(2)} \end{bmatrix}$ have identical signal and noise subspaces. Thus a subspace algorithm that exploits *joint* processing of the two arrays must exploit the structure of the dominant signal subspace eigenvector $\mathbf{e}_1^{(12)}$ in (38).

The following subspace procedure may be used to estimate the source location (x_s, y_s) using the correlation matrices $\mathbf{R}^{(1)}$, $\mathbf{R}^{(2)}$, and $\mathbf{R}^{(12)}$.

1. Form the correlation matrices $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$ using the observations from the individual arrays. Estimate the noise power σ_w^2 from the noise eigenvalues of $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$. Then estimate the signal power at each array using the largest eigenvalue of $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$ with (34):

$$\sigma_h^2 = \frac{1}{N}\left(\lambda_1^{(h)} - \sigma_w^2\right), \quad h = 1, 2. \tag{45}$$

2. Form the joint correlation matrix $\mathbf{R}^{(12)}$, and estimate the coherence magnitude $|\gamma|$ using the two largest eigenvalues $\lambda_1^{(12)}, \lambda_2^{(12)}$ of $\mathbf{R}^{(12)}$ with (36) and (37):

$$|\gamma| = \left[1 - \frac{1 - \left(\frac{\lambda_1^{(12)} - \lambda_2^{(12)}}{N(\sigma_1^2 + \sigma_2^2)} \right)^2}{\left(\frac{2}{\frac{\sigma_1}{\sigma_2} + \frac{\sigma_2}{\sigma_1}} \right)^2} \right]^{1/2}, \tag{46}$$

   as long as $\sigma_1 \neq \sigma_2$. Note that if $\sigma_1 = \sigma_2$, then it is not necessary to Estimate the coherence magnitude (see (41) and step 3). We will assume that γ is real-valued and positive, which is motivated by physical considerations as discussed earlier and in Wilson [7, 8].

3. Use the estimates for σ_1, σ_2, and $|\gamma|$ in the signal eigenvector expression for $\mathbf{e}_1^{(12)}$ in (38) to create a "search vector" $\mathbf{g}$:

$$\mathbf{g}(x_s, y_s) = \begin{bmatrix} \frac{1}{2}\left(\frac{\sigma_1}{\sigma_2} - \frac{\sigma_2}{\sigma_1}\right)\left[1 + \sqrt{1 + \left(\frac{2|\gamma|}{\frac{\sigma_1}{\sigma_2} - \frac{\sigma_2}{\sigma_1}}\right)^2}\right] \mathbf{a}_1 \\ \gamma^* \exp(j\omega_0 D_{12})\, \mathbf{a}_2 \end{bmatrix} = \begin{bmatrix} \beta_1\, \mathbf{a}_1 \\ \gamma^* \exp(j\omega_0 D_{12})\, \mathbf{a}_2 \end{bmatrix}, \quad (47)$$

where $\mathbf{a}_1, \mathbf{a}_2$, and D_{12} are functions of the source location (x_s, y_s). (If $\sigma_1 = \sigma_2$, then the search vector is based on $\mathbf{e}_1^{(12)}$ in (41) instead.)

4. Let $\mathbf{e}_n^{(12)}$, $n = 1, \ldots, 2N$ denote the eigenvectors of $\mathbf{R}^{(12)}$, ordered such that the corresponding eigenvalues follow $\lambda_1^{(12)} \geq \cdots \geq \lambda_{2N}^{(12)}$. Then form the matrix

$$\mathbf{V} = \begin{bmatrix} \mathbf{e}_2^{(12)} & \cdots & \mathbf{e}_{2N}^{(12)} \end{bmatrix}. \quad (48)$$

A MUSIC-type spectrum may then be defined as

$$P_c(x_s, y_s) = \frac{1}{\mathbf{g}^H \mathbf{V}\mathbf{V}^H \mathbf{g}}, \quad (49)$$

and the value of (x_s, y_s) that maximizes (49) is an estimate of the source location.

5. Wax and Kailath [21] proposed an "incoherent" MUSIC-type source location estimator that ignores coherence between the arrays. The form of the Wax and Kailath estimator can be expressed as follows:

$$P_i(x_s, y_s) = \frac{1}{\mathbf{a}_1^H \mathbf{V}^{(1)} \mathbf{V}^{(1)H} \mathbf{a}_1 + \mathbf{a}_2^H \mathbf{V}^{(2)} \mathbf{V}^{(2)H} \mathbf{a}_2}, \quad (50)$$

where $\mathbf{V}^{(h)} = [\mathbf{e}_2^{(h)}, \ldots, \mathbf{e}_N^{(h)}]$ contains the noise eigenvectors of $\mathbf{R}^{(\mathrm{h})}$. This analysis extends to the case of H arrays with a single source. If the signals are partially coherent at all arrays, then the dimension of the signal subspace of the combined correlation matrix is H. If K signals are present, then the combined correlation matrix has signal subspace with rank KH.

4.2 Examples of Subspace Processing

We present examples of the "partially coherent" MUSIC algorithm (49) and the "incoherent" MUSIC algorithm (50) with computer simulations and measured data. First, we simulate a scenario with $H = 2$ arrays. The individual arrays are identical and contain $N_1 = N_2 = N = 7$ sensors. Each array is circular and has 4-ft radius, with six sensors equally spaced around the perimeter and one sensor in the center. Narrowband processing centered at 40 Hz is assumed, with an SNR of 40 dB/sensor at array 1 and 34 dB at array 2, i.e., $G_{s,11}(\omega)/G_w(\omega) = 10^4$, $G_{s,22}(\omega)/G_w(\omega) = 2500$. The coherence is set to two values, $\gamma_{s,12} = 0.9$ and 0.999. The arrays are located

at coordinates $(x_1, y_1) = (0, 150)$, $(x_2, y_2) = (0, 0)$, and one source is located at $(x_s, y_s) = (200, 300)$, where the units are meters. Source location estimation is performed using $M = 30$ snapshots, where each snapshots contains an estimate of the measured complex amplitude at 40 Hz at each sensor.

Results are presented in Figure 16, where part a shows an approximate CRB as a function of signal coherence.* Increasing coherence leads to modest improvement in the localization potential. Figures 16b and 16c contain representative spectra of incoherent and partially coherent MUSIC, respectively, for coherence $\gamma_{s,12} = 0.9$. Note that the incoherent MUSIC spectrum has a rather broad but unique maximum. The partially coherent MUSIC spectrum has a sharp ridge in one direction, with multiple peaks along the ridge. The multiple peaks arise due to the ambiguities in source location caused by the narrowband signals and large separation between arrays.

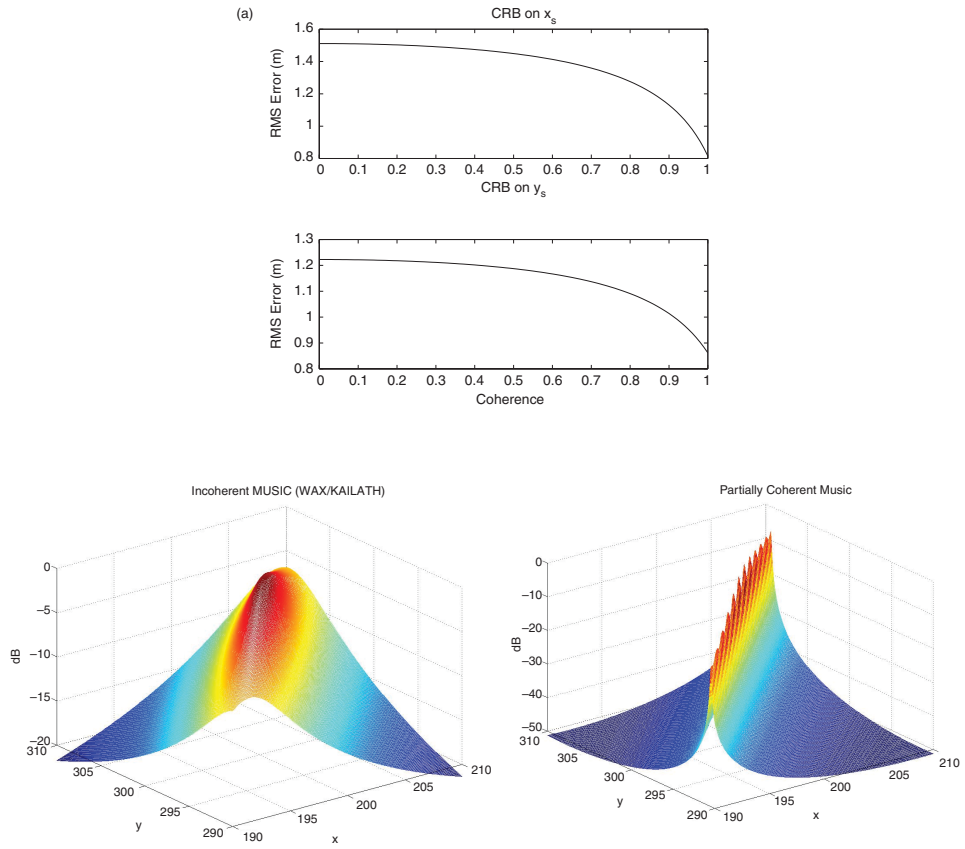


Figure 16. (a) CRBs on localization accuracy. (b) Incoherent MUSIC and (c) partially coherent MUSIC spectra.

Table 1 contains results of Monte Carlo simulations with 100 runs for both coherence values, $\gamma_{s,12} = 0.9$ and 0.999. The incoherent MUSIC performs close to the approximate CRB in Figure 16a corresponding to $\gamma_{s,12} = 0$. Partially coherent MUSIC performs *worse* than incoherent MUSIC because of the ambiguities: partially coherent MUSIC sometimes chooses the incorrect peak along the ridge in the spectrum shown in Figure 16c. Note that par-

*The CRB is approximate because the mapping of the scaling factor $\frac{T\Delta\omega}{\omega_s}$ in (16) to the number of snapshots M is approximate.

Table 1. RMS error simulation results for incoherent and partially coherent MUSIC.

Coherence	**Approx. CRB**		**Incoherent MUSIC**		**Partially Coherent MUSIC**	
	x_s	y_s	x_s	y_s	x_s	y_s
0.9	1.17	1.10	1.36	1.10	2.14	1.71
0.999	0.81	0.87	1.41	1.16	1.96	1.61
0.0	1.50	1.22				

tially coherent MUSIC is more accurate for $\gamma_s = 0.999$ than for $\gamma_s = 0.9$, suggesting that the ambiguities are less severe when the signal coherence is larger.

We have applied incoherent MUSIC and partially coherent MUSIC to the measured data scenario discussed in section 2.2.2 and in Figures 4 and 10. The data is processed in a narrow band around 39 Hz for the 2-s interval between 345 and 347 s. Both algorithms produce identical estimates for the source location in this case, but the spectra for partially coherent MUSIC are shown in Figure 17. The top panel clearly shows the ambiguities that limit the performance with narrowband signals. The contour plot in the bottom panel also shows the periodic structure of the ambiguities.

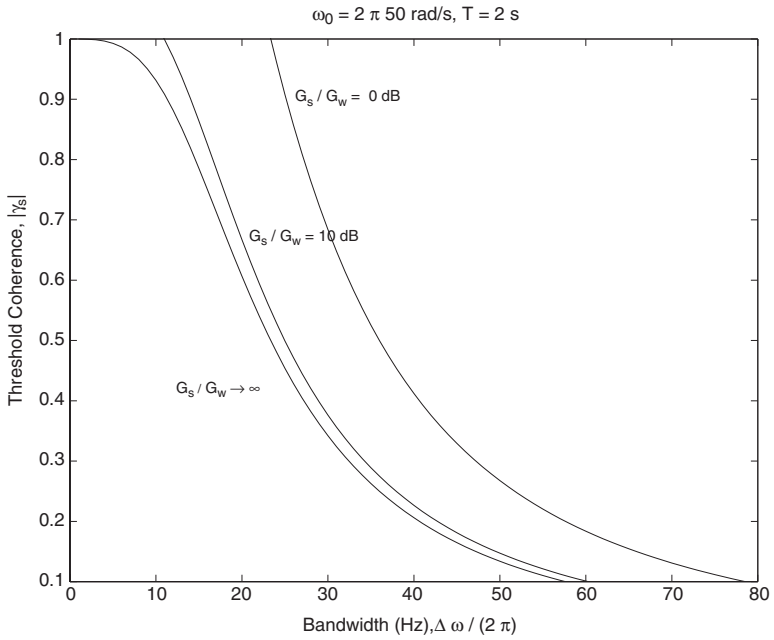


Figure 17. Partially coherent MUSIC spectrum for measured data.

5. Moving Sources

One of our objectives in this report is to quantify scenarios in which jointly processing data from widely spaced sensors has the potential for improved source localization accuracy, compared with incoherent triangulation/ tracking of bearing estimates. We established in section 2 that the potential for improved accuracy depends directly on TDE between the sensors. Then we showed in section 3 that TDE with partially coherent signals is feasible only with an increased time-bandwidth product of the sensor signals. This leads to a constraint on the minimum observation time, $\mathcal{T}$, in passive applications where the signal bandwidth is fixed. If the source is moving, then approximating it as nonmoving becomes poorer as $\mathcal{T}$ increases, so modelling the source motion becomes more important.

Approximate bounds are known [22, 23] that specify conditions of validity for nonmoving and moving source models. Let us consider $H = 2$ sensors with Doppler values $\alpha_2 > \alpha_1$ (see (64) for the definitions of α_1, α_2). If $f_{\max}$ (Hz) is the maximum signal frequency that is processed, then TDE estimation accuracy is not seriously affected by ignoring source motion, as long as the time interval $\mathcal{T}$ satisfies

$$\mathcal{T} \ll \frac{1}{f_{\max}\left(\frac{\alpha_2}{\alpha_1} - 1\right)}. \tag{51}$$

Taking typical parameters for ground vehicles in aeroacoustics, let us consider a vehicle moving at 5% the speed of sound (15 m/s), with radial motion that is in opposite directions at the two sensors. Then $\alpha_2/\alpha_1 - 1 \approx 0.1$ and (51) becomes $\mathcal{T} \ll 10/f_{\max}$. For $f_{\max} = 100$ Hz, the requirement is $\mathcal{T} \ll 0.1$ s, which, according to the analysis in section 2, yields insufficient time-bandwidth product for partially coherent signals that are typically encountered. Thus motion modeling and Doppler compensation are critical, even for aeroacoustic sources that move more slowly than in this example.

5.1 Data Model

In this section, we extend the nonmoving source model from section 2 using first-order motion models (see (52), (53), (67)). The first-order motion models are simple and accurate over larger time intervals $\mathcal{T}$ compared with the nonmoving source model. However, accurate modeling of more complex trajectories over longer time intervals requires higher-order polynomial models, with added complexity. The source position trajectory is modeled as a straight line with constant velocity over an interval of length $\mathcal{T}$,

$$x_s(t) = x_{s,0} + \dot{x}_s \cdot (t - t_0), \quad t_0 \le t \le t_0 + \mathcal{T} \tag{52}$$

$$y_s(t) = y_{s,0} + \dot{y}_s \cdot (t - t_0), \tag{53}$$

so $\dot{x}_s, \dot{y}_s$ are the velocity components. The source trajectory parameter vector is

$$\mathbf{\Theta} = [x_{s,0}, \dot{x}_s, y_{s,0}, \dot{y}_s]^T, \tag{54}$$

and the (time-varying) propagation time from the source to the sensors on array h follows from (1) and (2):

$$\begin{aligned} \tau_h(t) &= \frac{d_h(t)}{c} = \frac{1}{c}\left[(x_s(t)-x_h)^2 + (y_s(t)-y_h)^2\right]^{1/2} && (55) \\ \tau_{hn}(t) &\approx -\frac{1}{c}\left[\frac{x_s(t)-x_h}{d_h(t)}\Delta x_{hn} + \frac{y_s(t)-y_h}{d_h(t)}\Delta y_{hn}\right] \\ &= -\frac{1}{c}\left[(\cos\phi_h(t))\Delta x_{hn} + (\sin\phi_h(t))\Delta y_{hn}\right]. && (56) \end{aligned}$$

The bearing and bearing rate are related to the source motion parameters $\mathbf{\Theta}$ as

$$\begin{aligned} \phi_h(t) &= \tan^{-1}\left[\frac{y_s(t)-y_h}{x_s(t)-x_h}\right] = \tan^{-1}\left[\frac{y_{s,0}+\dot{y}_s\cdot(t-t_0)-y_h}{x_{s,0}+\dot{x}_s\cdot(t-t_0)-x_h}\right] && (57) \\ \dot{\phi}_h(t) &= \frac{\dot{y}_s\cos\phi_h(t) - \dot{x}_s\sin\phi_h(t)}{d_h(t)}. && (58) \end{aligned}$$

The radial velocity of the source with respect to array h is

$$v_{r,h}(t) = \dot{x}_s\cos\phi_h(t) + \dot{y}_s\sin\phi_h(t). \tag{59}$$

We can insert (52) and (53) into (55) to obtain the following approximation for the propagation time to array h:

$$\begin{aligned} \tau_h(t) &= \tau_h(t_0)\left[1 + \frac{2\cdot\cos\phi_h(t_0)\cdot\dot{x}_s\cdot(t-t_0)}{d_h(t_0)} + \frac{2\cdot\sin\phi_h(t_0)\cdot\dot{y}_s\cdot(t-t_0)}{d_h(t_0)}\right]^{1/2} && (60) \\ &\approx \tau_h(t_0) + \frac{v_{r,h}(t_0)}{c}\cdot(t-t_0), && (61) \end{aligned}$$

where $d_h(t_0)$ and $v_{r,h}(t_0)$ are the source distance and radial velocity at the start of time interval $t = t_0$. The approximation (61) is valid as long as the total motion during the time interval $\mathcal{T}$ is much less than the range, i.e., $|2\,\dot{x}_s\mathcal{T}| \ll d_h(t_0)$ and $|2\,\dot{y}_s\mathcal{T}| \ll d_h(t_0)$.

Next, we use the approximation (61) and model the received signal at the reference sensor on array h as

$$\begin{aligned} s_h\left(t-\tau_h(t)\right) &= s_h\left[\left(1-\frac{v_{r,h}(t_0)}{c}\right)t - \tau_h(t_0) + \frac{v_{r,h}(t_0)\,t_0}{c}\right] && (62) \\ &= s_h\left(\alpha_h\,t - \tau_h(t_0) + \frac{v_{r,h}(t_0)\,t_0}{c}\right), \quad t_0 \le t \le t_0+\mathcal{T}, && (63) \end{aligned}$$

where

$$\alpha_h = 1 - \frac{v_{r,h}(t_0)}{c} = 1 - \frac{1}{c}\left[\dot{x}_s\cos\phi_h(t_0) + \dot{y}_s\sin\phi_h(t_0)\right] \tag{64}$$

is the Doppler compression and

$$\tau_h(t_0) = \frac{d_h(t_0)}{c} = \frac{1}{c}\left[(x_{s,0} - x_h)^2 + (y_{s,0} - y_h)^2\right]^{1/2} \tag{65}$$

is the propagation delay at the initial time $t = t_0$. Without loss of generality, we set $t_0 = 0$, so the received signal at sensor n on array h is

$$s_h\left(\alpha_h\, t - \tau_h(0) - \tau_{hn}(t)\right), \tag{66}$$

which is the extension of the signal component of (8) to the moving source case. Note from (2) that $\tau_{hn}(t)$ depends on the source location only through the time-varying bearing $\phi_h(t)$, which we approximate with a first-order model

$$\phi_h(t) \approx \phi_h(t_0) + \dot{\phi}(t_0) \cdot (t - t_0), \quad t_0 \le t \le t_0 + \mathcal{T}. \tag{67}$$

For a single array h, the Doppler compression α_h and time delay $\tau_h(t_0)$ have negligible effect on estimation of the intra-array delays $\tau_{hn}(t)$, since α_h and $\tau_h(t_0)$ are identical for each $n = 1, \ldots, N_h$. Thus, each array can be processed separately to estimate the bearings $\phi_1(t_0), \ldots, \phi_H(t_0)$ and bearing rates $\dot{\phi}_1(t_0), \ldots, \dot{\phi}_H(t_0)$, and these can be "triangulated" via (57) and (58) to estimate the source motion parameters $\mathbf{\Theta}$ in (54). An algorithm [24] for estimating $\phi_h(t_0)$ and $\dot{\phi}_h(t_0)$ is described in section 5.2.

Let us consider the signals received at the reference sensors at each array, so $\tau_{hn}(t) = 0$ in (66):

$$s_1\left[\left(1 - \frac{v_{r,1}(t_0)}{c}\right) t - \tau_1(t_0)\right], \ldots, s_H\left[\left(1 - \frac{v_{r,H}(t_0)}{c}\right) t - \tau_H(t_0)\right]. \tag{68}$$

Our modeling assumptions imply that each signal $s_h\left[\left(1 - \frac{v_{r,h}(t_0)}{c}\right) t - \tau_h(t_0)\right]$ is a wide-sense stationary Gaussian random process. However, for two arrays g, h with *unequal* Doppler $v_{r,g}(t_0) \neq v_{r,h}(t_0)$, the signals at arrays g, h are *not* jointly wide-sense stationary [22, 25], complicating the analytical description and the CRB performance analysis. The jointly nonstationary sensor signals generally are not characterized by a cross-spectral density matrix, TDE with jointly nonstationary signals as in (68) was derived by Knapp and Carter [22]. The CRB analysis was rigorously justified by Schultheiss and Weinstein [26], and they extended the results to CRBs on differential Doppler. A clever transformation is used so that the jointly nonstationary signals in (68) are locally modeled by a CSD of the form (14), and Schultheiss and Weinstein show [25] that the representation is accurate for CRB analysis.

Formulating the results of Schultheiss and Weinstein [25] for the case of partially coherent signals,* we make the following observations for $H = 2$ arrays. The results are valid for large observation time ($\mathcal{T}$ much larger than the coherence time of the signals and noise). The TDE is $D_{12} = \tau_1(t_0) - \tau_2(t_0)$ and the differential Doppler is $\Delta v_{12} = v_{r,1}(t_0) - v_{r,2}(t_0)$.

*The signal coherence between the signals at arrays g and h in (68) is defined assuming perfect compensation of the Doppler compression α_g, α_h, thus yielding the definition in (7).

- Estimation of TDE and differential Doppler are decoupled, so the CRB on D_{12} is given by (24), which is identical to the nonmoving source case.
- The threshold coherence analysis for TDE in (30) and Figure 8 extends to the moving source case. In the best case that Doppler effects are perfectly estimated and compensated, the TDE problem that remains is identical to the nonmoving source case. Doppler estimation is less demanding in terms of time-bandwidth product compared with TDE. Indeed, Doppler estimation is possible with sinusoidal signals that have negligible bandwidth [25].

  The CRB on differential Doppler [25], modified for partially coherent signals, is

$$\mathrm{CRB}(\Delta v_{12}) = \frac{24\pi}{\mathcal{T}} \left(\frac{c}{\mathcal{T}}\right)^2 \left[2 \int_0^{\omega_s} \omega^2 \,\mathrm{SNR}\left(|\gamma_{s,12}(\omega)|\right)\, d\omega\right]^{-1}. \quad (69)$$

  Note that (69) is a scalar multiple of the CRB on TDE in (24). The CRB on differential Doppler may be achievable in scenarios where the time-bandwidth product is insufficient for TDE.

  Interestingly, differential Doppler provides sufficient information for source localization, even without TDE, as long as five or more sensors are available [26]. Thus the source motion may be exploited in scenarios where TDE is not feasible, such as narrowband signals.
- We discussed TDE with $H > 2$ sensors in section 2, concluding that pairwise processing of TDEs $D_{1H}, \ldots, D_{H-1,H}$ with a reference sensor H is nearly optimum for scenarios of interest (see (31)). A similar result holds for differential Doppler estimation [26], where pairwise estimation of $\Delta v_{1H}, \ldots, \Delta v_{H-1,H}$ is nearly as accurate as estimation of all pairs Δv_{gh}, as long as $H \cdot \mathrm{SNR}(|\gamma_s|) \gg 1$.

5.2 An Algorithm

The parameters that can be directly estimated from the sensor data are the bearings $\phi_1(t_0), \ldots, \phi_H(t_0)$, bearing rates $\dot{\phi}_1(t_0), \ldots, \dot{\phi}_H(t_0)$, pairwise time differences $D_{1H} = \tau_1(t_0) - \tau_H(t_0)$, ..., $D_{H-1,H} = \tau_{H-1}(t_0) - \tau_H(t_0)$, and differential Doppler $\Delta v_{1H} = v_{r,1}(t_0) - v_{r,H}(t_0), \ldots, \Delta v_{H-1,H} = v_{r,H-1}(t_0) - v_{r,H}(t_0)$. Equations (57), (58), (64), and (65) define the nonlinear relations that "triangulate" these parameters and relate them to the source motion parameters $\boldsymbol{\Theta} = [x_{s,0}, \dot{x}_s, y_{s,0}, \dot{y}_s]^T$.

A distributed processing algorithm is outlined next, and parts of the algorithm are illustrated with measured aeroacoustic data in section 6.

1. Use the local polynomial approximation (LPA) beamformer [24] at each array to estimate the bearings and bearing rates. The LPA beamformer in Katkovnik and Gershman [24] is formulated for narrowband processing, and it is a generalization of the classical beamformer

to moving sources. We extend it in a straightforward way to wideband signals by incoherently averaging the LPA beampatterns at different frequencies.

2. Solve (57) and (58) to obtain initial estimates of the source motion parameters $\boldsymbol{\Theta}$. These estimates correspond to incoherent triangulation of the bearings and bearing rates from individual arrays.

3. Estimate the Doppler compression factors $\alpha_1, \ldots, \alpha_H$, compensate for Doppler, and test whether the signals at distinct arrays have sufficient coherence, fractional bandwidth, and time-bandwidth product to enable TDE between arrays (see section 2 for the conditions).

4. If the conditions are *not* met, then incoherent triangulation of the bearings and bearing rates is nearly optimum, and further joint processing is not informative.

5. If the conditions are met, then identify a reference array H (the array with maximum SNR) and estimate the time differences $D_{1H}, \ldots, D_{H-1,H}$ and differential Dopplers $\Delta v_{1H}, \ldots, \Delta v_{H-1,H}$. The maximum likelihood solution involves wideband ambiguity function search over Doppler and TDE [22], but computationally simpler alternatives have been investigated [26].

6. A suboptimum procedure is to avoid the joint Doppler and TDE estimation in the preceding step, and instead use the *initial* Doppler estimates from steps 1 and 2 and perform TDE after approximate Doppler compensation. With this approach, triangulation of the TDEs via (65) will improve the estimates of $x_{s,0}$ and $y_{s,0}$ only (and not the source velocity $\dot{x}_s, \dot{y}_s$).

7. If multiple sources are present, then the LPA beamformer in step 1 may be used to separate the source signals at each array prior to Doppler/TDE estimation.

The LPA beamformer in steps 1 and 7 is illustrated in section 6 for a two-source scenario based on measured aeroacoustic data. Examples of TDE (step 5) with measured data were presented in section 3.4.

6. Example with Two Moving Sources

We present an application of the LPA beamformer [24] to measured aeroacoustic data with two ground vehicles: an M1 tank and an M3 tracked armored vehicle. The vehicle trajectories over a 10-s segment are shown in Figure 18a. Two arrays, labeled 8B and 8C, are separated by about 23 m. Each array is circular with $N = 7$ sensors, 4-ft radius, and six sensors equally spaced around the perimeter with one sensor in the center. We present results based on processing the data at array 8B, for which the bearing and bearing rate of the sources are shown in Figure 18b. The bearing of the M3 varies by more than 30° over the 10-s time interval. The range of the M3 is closer at about 100 m, with the M1 range at approximately 200 m.

The data from array 8B is processed over a wide bandwidth from 30 to 150 Hz. The beampattern for a classical beamformer based on a nonmoving source model is shown in Figure 18c. The beampattern is the incoherent sum of narrowband beampattern over the 30- to 150-Hz frequency band. The peak of the beampattern is located at approximately the mean bearing of the stronger source (M3) over the 10 s interval.

The beampattern of the LPA beamformer is shown in Figure 19a. The LPA beampattern exploits a first-order model for time-varying bearing, as in (67). The LPA beampattern is two-dimensional, with axes of initial bearing $\phi(t_0)$ and bearing rate $\dot{\phi}(t_0)$. The LPA beamformer in Katkovnik and Gershman [24] is formulated for narrowband processing, so we extend to the wideband case by incoherently adding the narrowband LPA beampatterns over the frequency range. The peak of the beampattern in Figure 19a is close to the true values of $\phi(t_0)$, $\dot{\phi}(t_0)$ for the M3 that are shown in Figure 18b. The location of the weaker M1 source is not evident in the LPA beampattern in Figure 19a, so we subtract an estimate of the stronger source from the data. The subtraction is performed based on the bearing and bearing rate estimates from Figure 19a. The subtraction is coherent over the processing bandwidth and includes the time-varying bearing. The LPA beampattern after subtraction is shown in Figure 19b, which indicates the bearing and bearing rate of the weaker M1 source. This example illustrates the gain in resolution that is achieved by exploiting the source motion in a beamformer.

Figure 18. (a) Array locations and trajectory of M1 and M3 vehicles over a 10-s time interval. (b) Bearing and bearing rate of sources with respect to array 8B. (c) Beampattern of classical beamformer based on nonmoving source model.

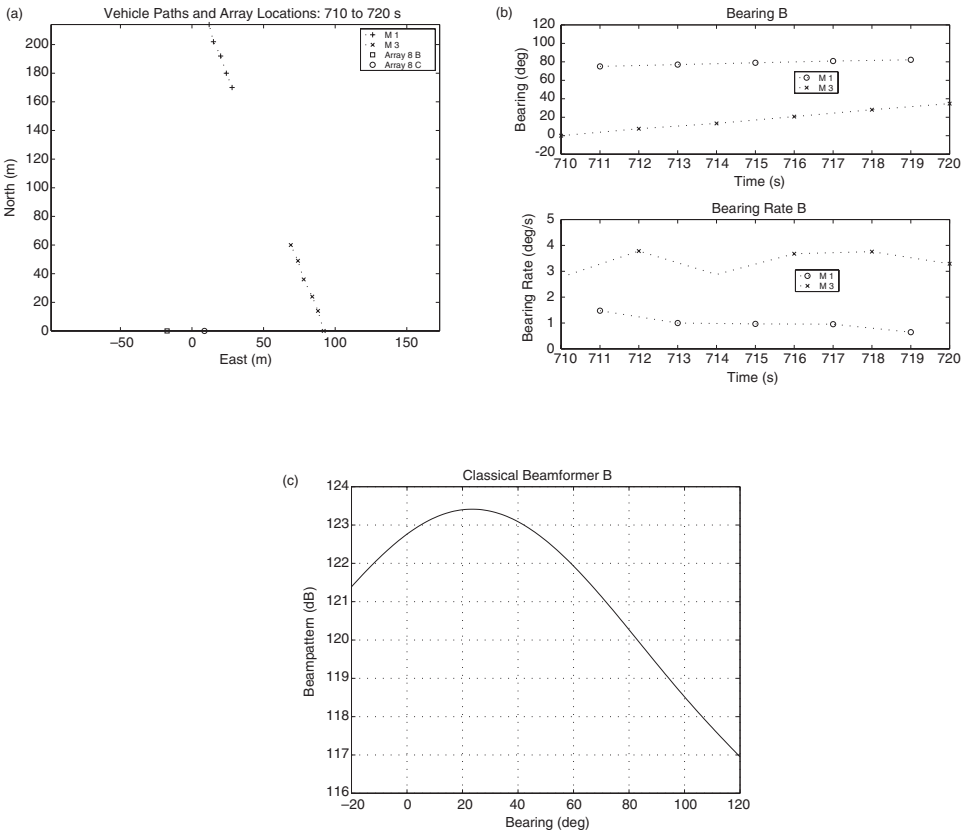
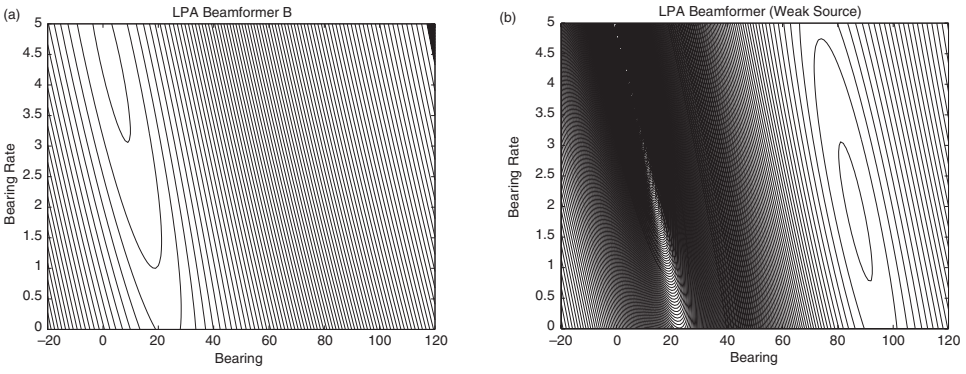


Figure 19.
(a) Beampattern of LPA beamformer, showing the location of the stronger source (M3). (b) Beampattern of LPA beamformer after subtracting an estimate of the strong source signal from the data, showing the location of the weaker source (M1).



7. Summary and Concluding Remarks

We have presented in this report an analysis of source localization with sensors arranged in an "array of arrays" configuration. We have paid particular attention to aeroacoustical 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 with pairwise TDE between separate arrays. We showed that 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. The algorithm was presented first for a single, nonmoving source, and then we extended to the case of moving sources. The algorithm for moving sources involves estimation of bearing-rate and differential Doppler between sensors. We provided an analysis based on Ziv-Zakai bounds that quantifies the requirements on SNR, 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 in the report.

We are continuing our research on aeroacoustic source localization and tracking in several ways. In particular, we are investigating algorithms that exploit the "sum of harmonics" structure that characterizes the acoustical signature of many ground vehicles. We are also investigating joint algorithms for localization, tracking, and classification that share information to improve performance.

References

1. Tenney, R. R., and J. R. Delaney. "A Distributed Aeroacoustic Tracking Algorithm." *Proc. American Control Conf.*, pp. 1440–1450, June 1984.

2. Bar-Shalom, Y., and X.-R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. YBS, Storrs, CT, 1995.

3. Farina, A. "Target Tracking with Bearings-Only Measurements." *Signal Processing*, vol. 78, pp. 61–78, 1999.

4. Ristic, B., S. Arulampalam, and C. Musso. "The Influence of Communication Bandwidth on Target Tracking With Angle Only Measurements From Two Platforms." *Signal Processing*, vol. 81, pp. 1801–1811, 2001.

5. Kaplan, L. M., P. Molnar, and Q. Le. "Bearings-Only Target Localization for an Acoustical Unattended Ground Sensor Network." *Proceedings SPIE AeroSense*, Orlando, FL, April 2001.

6. Kozick, R. J. "Distributed Sensor Array Processing of Wideband Aeroacoustic Signals With Frequency-Selective Spatial Coherence." Final report for U.S. Army Research Office, contract DAAH04-96-C-0086, 19 July 2000.

7. Wilson, D. K. "Performance Bounds for Acoustic Direction-of-Arrival Arrays Operating in Atmospheric Turbulence." *J. Acoust. Soc. Am.*, vol. 103, no. 3, pp. 1306–1319, March 1998.

8. Wilson, D. K. "Atmospheric Effects on Acoustic Arrays: A Broad Perspective From Models." 1999 Meeting of the IRIS Specialty Group on Battlefield Acoustics and Seismics. Laurel, MD, 13–15 September 1999.

9. Weiss, A. J., and E. Weinstein. "Fundamental Limitations in Passive Time Delay Estimation—Part 1: Narrowband Systems." *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-31, no. 2, pp. 472–485, April 1983.

10. Weinstein, E., and A. J. Weiss. "Fundamental Limitations in Passive Time Delay Estimation—Part 2: Wideband Systems." *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-32, no. 5, pp. 1064–1077, October 1984.

11. Kozick, R. J. "Classification of Pulsed Signals: Classification Limits and Performance Analysis." Final report for U.S. Army Research Office, contract DAAH04-96-C-0086, 27 November 2000.

12. Friedlander, B. "On the Cramer-Rao Bound for Time Delay and Doppler Estimation." *IEEE Trans. on Info. Theory*, vol. IT-30, no. 3, pp. 575–580, May 1984.

13. Whittle, P. "The Analysis of Multiple Stationary Time Series." *J. Royal Statist. Soc.*, vol. 15, pp. 125–139, 1953.

14. Kay, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, 1993.

15. Carter, G. C. (ed.). *Coherence and Time Delay Estimation* (Selected Reprint Volume). IEEE Press, New York, NY, 1993.

16. Walker, W. F., and G. E. Trahey. "A Fundamental Limit on Delay Estimation Using Partially Correlated Speckle Signals." *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 42, no. 2, pp. 301–308, March 1995.

17. Cespedes, I., J. Ophir, and S. K. Alam. "The Combined Effect of Signal Decorrelation and Random Noise on the Variance of Time Delay Estimation." *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 44, no. 1, pp. 220–225, January 1997.

18. Weinstein, E. "Decentralization of the Gaussian Maximum Likelihood Estimator and Its Applications to Passive Array Processing." *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-29, no. 5, pp. 945–951, October 1981.

19. Deadrick D., Sanders Corp., Nashua, NJ, and A. Filipov, U.S. Army Research Laboratory, Adelphi, MD. Personal communication with B. Sadler, May 2000.

20. Paulraj, A., and T. Kailath. "Direction of Arrival Estimation by Eigenstructure Methods With Imperfect Spatial Coherence of Wavefronts." *J. Acoust. Soc. Am.*, vol. 83, pp. 1034–1040, March 1988.

21. Wax, M., and T. Kailath. "Decentralized Processing in Sensor Arrays." *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 1123–1129, October 1985.

22. Knapp, C. H., and G. C. Carter. "Estimation of Time Delay in the Presence of Source or Receiver Motion." *J. Acoust. Soc. Am.*, vol. 61, no. 6, pp. 1545–1549, June 1977.

23. Adams, W. B., J. P. Kuhn, and W. P. Whyland. "Correlator Compensation Requirements for Passive Time-Delay Estimation With Moving Source or Receivers." *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 2, pp. 158–168, April 1980.

24. Katkovnik, V., and A. B. Gershman. "A Local Polynomial Approximation Based Beamforming for Source Localization and Tracking in Nonstationary Environments." *IEEE Signal Processing Letters*, vol. 7, no. 1, pp. 3–5, January 2000.

25. Schultheiss, P. M., and E. Weinstein. "Estimation of Differential Doppler Shifts." *J. Acoust. Soc. Am.*, vol. 66, no. 5, pp. 1412–1419, November 1979.

26. Betz, J. W. "Comparison of the Deskewed Short-Time Correlator and the Maximum Likelihood Correlator." *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 2, pp. 285–294, April 1984.

REPORT DOCUMENTATION PAGE

or A roved o

1. AGENCY USE ONLY *eave an*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2002	Final, October 2000–October 2001

4. TITLE AND SUBTITLE
Tracking Moving Acoustic Sources With a Network of Sensors

5. FUNDING NUMBERS
DA PR: AH48
PE: 61102A

6. AUTHOR(S)
Richard J. Kozick and Brian M. Sadler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
Attn: AMSRL- CI-CN
2800 Powder Mill Road
Adelphi, MD 20783-1197

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-2750

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
ARL PR: 2FEPC1
AMS code: 611102H4811

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *a i ords*

This report presents issues and algorithms for the problem of source tracking with a network of aeroacoustic sensors. We study fusion of data from sensors that are widely separated, and we give particular attention to the important issues of limited communication bandwidth between sensor nodes, effects of source motion, coherence loss between signals measured at different sensors, signal bandwidth, and noise. We compare the tracking performance of various schemes, including joint (coherent) processing of all sensor data, as well as data-reduction schemes that employ distributed computation and reduced communication bandwidth with a fusion center. Our analysis provides a quantification of the potential gain in source tracking accuracy that is achievable with greater communication bandwidth and joint processing of sensor data. We show that the potential gain in accuracy depends critically on the scenario, as determined by the source motion parameters, signal coherence between sensors, bandwidth of the source signals, and noise level. For scenarios that admit increased accuracy with joint processing, we present a bandwidth-efficient algorithm that involves beamforming at small-aperture sensor arrays combined with time-delay estimation between widely spaced sensor arrays. The algorithms and their performance are illustrated using measured aeroacoustic data from ground vehicles.

14. SUBJECT TERMS
sensor networks, acoustics, tracking

15. NUMBER OF PAGES
48

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL