

Acoustic Source Localization and Beamforming: Theory and Practice

Joe C. Chen

*Electrical Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594, USA
Email: jcchen@ee.ucla.edu*

Kung Yao

*Electrical Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594, USA
Email: yao@ee.ucla.edu*

Ralph E. Hudson

*Electrical Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594, USA
Email: ralph@ee.ucla.edu*

Received 17 February 2002 and in revised form 21 September 2002

We consider the theoretical and practical aspects of locating acoustic sources using an array of microphones. A maximum-likelihood (ML) direct localization is obtained when the sound source is near the array, while in the far-field case, we demonstrate the localization via the cross bearing from several widely separated arrays. In the case of multiple sources, an alternating projection procedure is applied to determine the ML estimate of the DOAs from the observed data. The ML estimator is shown to be effective in locating sound sources of various types, for example, vehicle, music, and even white noise. From the theoretical Cramér-Rao bound analysis, we find that better source location estimates can be obtained for high-frequency signals than low-frequency signals. In addition, large range estimation error results when the source signal is unknown, but such unknown parameter does not have much impact on angle estimation. Much experimentally measured acoustic data was used to verify the proposed algorithms.

Keywords and phrases: source localization, ML estimation, Cramér-Rao bound, beamforming.

1. INTRODUCTION

Acoustic source localization has been an active research area for many years. Applications include unattended ground sensor (UGS) network for military surveillance, reconnaissance, or around the perimeter of a plant for intrusion detection [1]. Many variations of algorithms using a microphone array for source localization in the near field as well as direction-of-arrival (DOA) estimation in the far field have been proposed [2]. Many of these techniques involve a relative time-delay-estimation step that is followed by a least squares (LS) fit to the source DOA, or in the near-field case, an LS fit to the source location [3, 4, 5, 6, 7].

In our previous paper [8], we derived the “optimal” parametric maximum likelihood (ML) solution to locate acoustic sources in the near field and provided computer simulations to show its superiority in performance over other methods. This paper is an extension of [8], where both the far- and the near-field cases are considered, and the theoretical analysis is provided by the Cramér-Rao bound

(CRB), which is useful for both performance comparison and basic understanding purposes. In addition, several experiments have been conducted to verify the usefulness of the proposed algorithm. These experiments include both indoor and outdoor scenarios with half a dozen microphones to locate one or two acoustic sources (sound generated by computer speaker(s)).

One major advantage that the proposed ML approach has is that it avoids the intermediate relative time-delay estimation. This is made possible by transforming the wideband data to the frequency domain, where the signal spectrum can be represented by the narrowband model for each frequency bin. This allows a direct optimization for the source location(s) under the assumption of Gaussian noise instead of the two-step optimization that involves the relative time-delay estimation. The difficulty in obtaining relative time delays in the case of multiple sources is well known, and by avoiding this step, the proposed approach can then estimate multiple source locations. However, in practice, when we apply the discrete Fourier transform (DFT), several artifacts

can result due to the finite length of data frame (see Section 2.1.1). As a result, there does not exist an exact ML solution for data of finite length. Instead, we ignore these finite effects and derive the solution which we refer to as the approximated ML (AML) solution. Note that a similar solution has been derived independently in [9] for the far-field case.

In practice, the number of sources may be determined independent of or together with the localization algorithm, but here we assume that it is known for the purpose of this paper. For the single-source case, we have shown that the AML formulation is equivalent to maximizing the sum of the weighted cross-correlation functions between time-shifted sensor data in [8]. The optimization using all sensor pairs mitigates the ambiguity problem that often arises in the relative time-delay estimation between two widely separated sensors for the two-step LS methods. In the case of multiple sources, we apply an efficient alternating projection (AP) procedure, which avoids the multidimensional search by sequentially estimating the location of one source while fixing the estimates of other source locations from the previous iteration. In this paper, we demonstrate the localization results using the AML method to the measured data, both in the near-field and far-field cases, and for various types of sound sources, for example, vehicle, music, and even white noise. The AML approach is shown to outperform the LS-type algorithms in the single-source case, and by applying AP, the proposed algorithm is able to locate two sound sources from the observed data.

The paper is organized as follows. In Section 2, the theoretical performances of DOA estimation and source localization with the CRB analysis are given. Then, we derive the AML solution for DOA estimation and source localization in Section 3. In Section 4, simulation examples and experimental results are given to demonstrate the usefulness of the proposed method. Finally, we give our conclusions.

2. THEORETICAL PERFORMANCE AND ANALYSIS

In this section, the theoretical performances of DOA estimation for the far-field case and of source localization for the near-field case are analyzed. First, we define the signal models for the far- and near-field cases. Then, the CRBs are derived and analyzed. The CRB is most often used as a theoretical lower bound for any unbiased estimator [10]. Most of the derivations of the CRB for wideband source localization found in the literature are in terms of relative time-delay estimation error. In the following, we derive a more general CRB directly from the signal model. By developing a theoretical lower bound in terms of signal characteristics and array geometry, we not only bypass the involvement of the intermediate time-delay estimator but also offer useful insights to the physical properties of the problem.

The DOA and source localization variances both depend on two separate parts, one that only depends on the signal and another that only depends on the array geometry. This suggests separate performance dependence on the signal and the geometry. Thus, for any given signal, the CRB can provide the theoretical performance of a particular ge-

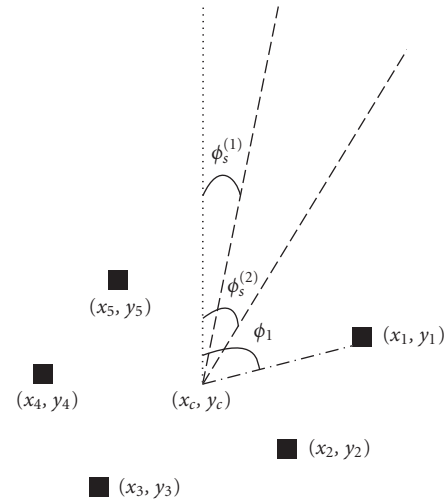


FIGURE 1: Far-field example with randomly distributed sensors.

ometry and helps the design of an array configuration for a particular scenario of interest. The signal dependence part shows that theoretically the DOA and source location root mean squares (RMS) error are linearly proportional to the noise level and the speed of propagation, and inversely proportional to the source spectrum and frequency. Thus, better DOA and source location estimates can be obtained for high-frequency signals than low-frequency signals. In further sensitivity analysis, large range estimation error is found when the source signal is unknown, but such unknown parameter does not affect the angle estimation.

The CRB analysis also shows that the uniformly spaced circular array provides an attractive geometry for good overall performance. When a circular array is used, the DOA variance bound is independent of the source direction, and it also does not degrade when the speed of propagation is unknown. An effective beamwidth for DOA estimation can also be given by the CRB. The beamwidth provides a measure of how dense the angles should be sampled for the AML metric evaluation, thus prevents unneeded iterations using numerical techniques.

Throughout this paper, we denote superscript T as the transpose, H as the complex conjugate transpose, and $*$ as the complex conjugate operation.

2.1. Signal model of the far- and near-field cases

2.1.1 The far-field case

When the source is in the far-field of the array, the wave front is assumed to be planar and only the angle information can be estimated. In this case, we use the array centroid as the reference point and define a signal model based on the relative time delays from this position. For simplicity, we assume a randomly distributed planar (2D) array of R sensors, each at position $\mathbf{r}_p = [x_p, y_p]^T$, as depicted in Figure 1. The centroid position is given by $\mathbf{r}_c = (1/R) \sum_{p=1}^R \mathbf{r}_p = [x_c, y_c]^T$. The sensors are assumed to be omnidirectional and have identical responses. On the same plane as the array, we assume that there are M sources ($M < R$), each at an angle $\phi_s^{(m)}$

from the array, for $m = 1, \dots, M$. The angle convention is such that north is 0 degree and east is 90 degrees. The relative time delay of the m th source is given by $t_{cp}^{(m)} = t_c^{(m)} - t_p^{(m)} = [(x_c - x_p) \sin \phi_s^{(m)} + (y_c - y_p) \cos \phi_s^{(m)}] / v$, where $t_c^{(m)}$ and $t_p^{(m)}$ are the absolute time delays from the m th source to the centroid and the p th sensor, respectively, and v is the speed of propagation in length unit per sample. The data collected by the p th sensor at time n can be given by

$$x_p(n) = \sum_{m=1}^M s_c^{(m)}(n - t_{cp}^{(m)}) + w_p(n), \quad (1)$$

for $n = 0, \dots, L - 1$, $p = 1, \dots, R$, and $m = 1, \dots, M$, where $s_c^{(m)}$ is the source signal arriving at the array centroid position, $t_{cp}^{(m)}$ is allowed to be any real-valued number, and w_p is the zero-mean white Gaussian noise with variance σ^2 .

For the ease of derivation and analysis, the wideband signal model should be given in the frequency domain, where a narrowband model can be given for each frequency bin. A block of L samples in each sensor data can be transformed to the frequency domain by a DFT of length N . It is well known that the DFT creates a circular time shift when applying a linear phase shift in the frequency domain. However, the time delay in the array data corresponds to a linear time shift, thus creating a mismatch in the signal model, which we refer to as an edge effect. When $N = L$, severe edge effect results for small L , but it becomes a good approximation for large L . We can apply zero padding for small L to remove such edge effect, that is, $N \geq L + \tau$, where τ is the maximum relative time delay among all sensor pairs. However, the zero padding removes the orthogonality of the noise component across frequency. In practice, the size of L is limited due to the nonstationarity of the source location. In the following, we assume that either L is large enough or the noise is almost uncorrelated across frequency. Note that the CRB derived based on this frequency-domain model is idealistic and does not take this edge effect into account.

In the frequency domain, the array signal model is given by

$$\mathbf{X}(k) = \mathbf{D}(k)\mathbf{S}_c(k) + \boldsymbol{\eta}(k), \quad (2)$$

for $k = 0, \dots, N - 1$, where the array data spectrum is given by $\mathbf{X}(k) = [X_1(k), \dots, X_R(k)]^T$, the steering matrix is given by $\mathbf{D}(k) = [\mathbf{d}^{(1)}(k), \dots, \mathbf{d}^{(M)}(k)]$, the steering vector is given by $\mathbf{d}^{(m)}(k) = [d_1^{(m)}(k), \dots, d_R^{(m)}(k)]^T$, $d_p^{(m)}(k) = e^{-j2\pi k t_{cp}^{(m)}/N}$, and the source spectrum is given by $\mathbf{S}_c(k) = [S_c^{(1)}(k), \dots, S_c^{(M)}(k)]^T$. The noise spectrum vector $\boldsymbol{\eta}(k)$ is zero-mean complex white Gaussian, distributed with variance $L\sigma^2$. Note that, due to the transformation of the frequency domain, $\boldsymbol{\eta}(k)$ asymptotically approaches a Gaussian distribution by the central limit theorem even if the actual time-domain noise has an arbitrary i.i.d. distribution (with bounded variance) other than Gaussian. This asymptotic property in the frequency domain provides a more reliable noise model than the time-domain model in some practical cases. For convenience of notation, we define $\mathbf{S}(k) =$

$\mathbf{D}(k)\mathbf{S}_c(k)$. By stacking up the $N/2$ positive frequency bins (zero frequency bin is not important and the negative frequency bins are merely mirror images) of the signal model in (2) into a single column, we can rewrite the sensor data into an $NR/2 \times 1$ space-temporal frequency vector as $\mathbf{X} = \mathbf{G}(\boldsymbol{\Theta}) + \boldsymbol{\xi}$, where $\mathbf{G}(\boldsymbol{\Theta}) = [\mathbf{S}(1)^T, \dots, \mathbf{S}(N/2)^T]^T$, and $\mathbf{R}_\xi = E[\boldsymbol{\xi}\boldsymbol{\xi}^H] = L\sigma^2\mathbf{I}_{NR/2}$.

2.1.2 The near-field case

In the near-field case, the range information can also be estimated in addition to the DOA. Denote \mathbf{r}_{s_m} as the location of the m th source, and in this case we use this as the reference point instead of the array centroid. Since we consider the near-field sources, the signal strength at each sensor can be different due to nonuniform spatial loss in the near-field geometry. The sensors are again assumed to be omnidirectional and have identical responses. In this case, the data collected by the p th sensor at time n can be given by

$$x_p(n) = \sum_{m=1}^M a_p^{(m)} s_0^{(m)}(n - t_p^{(m)}) + w_p(n), \quad (3)$$

for $n = 0, \dots, L - 1$, $p = 1, \dots, R$, and $m = 1, \dots, M$, where $a_p^{(m)}$ is the signal-gain level of the m th source at the p th sensor (assumed to be constant within the block of data), $s_0^{(m)}$ is the source signal, and $t_p^{(m)}$ is allowed to be any real-valued number. The time delay is defined by $t_p^{(m)} = \|\mathbf{r}_{s_m} - \mathbf{r}_p\| / v$, and the relative time delay between the p th and the q th sensors is defined by $t_{pq}^{(m)} = t_p^{(m)} - t_q^{(m)} = (\|\mathbf{r}_{s_m} - \mathbf{r}_p\| - \|\mathbf{r}_{s_m} - \mathbf{r}_q\|) / v$. With the same edge-effect problem mentioned above, the frequency-domain model for the near-field case is given by

$$\mathbf{X}(k) = \mathbf{D}(k)\mathbf{S}_0(k) + \boldsymbol{\eta}(k), \quad (4)$$

for $k = 0, \dots, N - 1$, where each element of the steering vector now becomes $d_p^{(m)}(k) = a_p^{(m)} e^{-j2\pi k t_p^{(m)}/N}$, and the source spectrum is given by $\mathbf{S}_0(k) = [S_0^{(1)}(k), \dots, S_0^{(M)}(k)]^T$.

2.2. Cramér-Rao bound for DOA estimation

In the following CRB derivation, we consider the single-source case ($M = 1$) under three conditions: known signal and known speed of propagation, known signal but unknown speed of propagation, and known speed of propagation but unknown signal. The comparison of the three conditions provides a sensitivity analysis of different parameters. Only the single-source case is considered since valuable analysis can be obtained using a single source while the analytic expression of the multiple-sources case becomes much more complicated. The far-field frequency-domain signal model for the single-source case is given by

$$\mathbf{X}(k) = S_c(k)\mathbf{d}(k) + \boldsymbol{\eta}(k), \quad (5)$$

for $k = 0, \dots, N - 1$, where $\mathbf{d}(k) = [d_1(k), \dots, d_R(k)]^T$, $d_p(k) = e^{-j2\pi k t_{cp}/N}$, and $S_c(k)$ is the source spectrum of this source.

After considering all the positive frequency bins, we can construct the Fisher information matrix [10] by

$$\mathbf{F} = 2 \operatorname{Re} [\mathbf{H}^H \mathbf{R}_\xi^{-1} \mathbf{H}] = (2/L\sigma^2) \operatorname{Re} [\mathbf{H}^H \mathbf{H}], \quad (6)$$

where $\mathbf{H} = \partial \mathbf{G} / \partial \phi_s$, for the case of known signal and known speed of propagation. In this case, the Fisher information matrix is indeed a scalar $F_{\phi_s} = \zeta \alpha$, where $\zeta = (2/L\sigma^2 v^2) \sum_{k=1}^{N/2} (2\pi k |S_c(k)|/N)^2$ is the scale factor that is proportional to the total power in the derivative of the source signal, and $\alpha = \sum_{p=1}^R b_p^2$ is the geometry factor that depends on the array and the source direction, where

$$b_p = (x_c - x_p) \cos \phi_s - (y_c - y_p) \sin \phi_s. \quad (7)$$

Hence, for any arbitrary array, the RMS error bound for DOA estimation is given by $\sigma_{\phi_s} \geq 1/\sqrt{\zeta \alpha}$. The geometry factor α provides a measure of geometric relations between the source and the sensor array. Poor array geometry may lead to a small α , which results in large estimation variance. It is clear from the scale factor ζ that the performance does not solely depend on the SNR but also the signal bandwidth and spectral density. Thus, source localization performance is better for signals with more energy in the high frequencies.

In the case of unknown source signal, the matrix $\mathbf{H} = [\partial \mathbf{G} / \partial \phi_s, \partial \mathbf{G} / \partial |S_c|^T, \partial \mathbf{G} / \partial \Phi_c^T]$, where $S_c = [S_c(1), \dots, S_c(N/2)]^T$, and $|S_c|$ and Φ_c are the magnitude and phase part of S_c , respectively. The resulting bound after applying the well-known block matrix inversion lemma (see [11, Appendix]) on \mathbf{F}_{ϕ_s, S_c} is given by $\sigma_{\phi_s} \geq 1/\sqrt{\zeta(\alpha - z_{S_c})}$, where $z_{S_c} = (1/R)[\sum_{p=1}^R b_p]^2$ is the penalty term due to the unknown source signal. It is known that the DOA performance does not degrade when the source signal is unknown; thus, we can show that z_{S_c} is indeed zero, that is, $\sum_{p=1}^R b_p = \cos \phi_s \sum_{p=1}^R (x_c - x_p) - \sin \phi_s \sum_{p=1}^R (y_c - y_p) = 0$ since $\sum_{p=1}^R (x_c - x_p) = R x_c - \sum_{p=1}^R x_p = 0$ and $\sum_{p=1}^R (y_c - y_p) = 0$. Note that the above analysis is valid for any arbitrary array. When the speed of propagation is unknown, the matrix $\mathbf{H} = [\partial \mathbf{G} / \partial \phi_s, \partial \mathbf{G} / \partial v]$, and the resulting bound after applying the matrix inversion lemma on $\mathbf{F}_{\phi_s, v}$ is given by $\sigma_{\phi_s} \geq 1/\sqrt{\zeta(\alpha - z_v)}$, where $z_v = (1/\sum_{p=1}^R t_{cp}^2)[\sum_{p=1}^R b_p t_{cp}]^2$ is the penalty term due to the unknown speed of propagation. This penalty term is not necessarily zero for any arbitrary array, but it becomes zero for a uniformly spaced circular array.

2.2.1 The circular-array case

In the following, we show the CRB for a uniformly spaced circular array. Not only a simple analytic form can be given but also the optimal geometry for DOA estimation. The variance of the DOA estimation is independent of the source direction, and also does not degrade when the speed of propagation is unknown. Without a loss of generality, we pick the array centroid as the origin, that is, $\mathbf{r}_c = [0, 0]^T$. The location of the p th sensor is given by $\mathbf{r}_p = [\rho \sin \phi_p, \rho \cos \phi_p]^T$, where ρ is the radius of the circular array, $\phi_p = 2\pi p/R + \phi_0$ is the

angle of the p th sensor with respect to north, and ϕ_0 is the angle that defines the orientation of the array. Then, $\alpha = \rho^2 R/2$. The DOA variance bound is given by $\sigma_{\phi_s}^2$ (circular array) $\geq 2/\zeta \rho^2 R$, which is independent of the source direction. It is useful to define the following terms for a better interpretation of the CRB. Define the *normalized root weighted mean squared (nrwms) source frequency* by

$$k_{\text{nrwms}} \equiv \frac{2}{N} \sqrt{\frac{\sum_{k=1}^{N/2} k^2 |S_c(k)|^2}{\sum_{k=1}^{N/2} |S_c(k)|^2}}, \quad (8)$$

and the *effective beamwidth* by

$$\phi_{\text{BW}} \equiv \frac{v}{\pi \rho k_{\text{nrwms}}}. \quad (9)$$

Then, the RMS error bound for DOA estimation can be given by

$$\sigma_{\phi_s}(\text{circular array}) \geq \frac{\phi_{\text{BW}}}{\sqrt{\text{SNR}_{\text{array}}}}, \quad (10)$$

where the effective SNR $\simeq \sum_{k=1}^{N/2} |S_c(k)|^2 / L\sigma^2$ and $\text{SNR}_{\text{array}} = R \cdot \text{SNR}$.

This shows that the effective beamwidth is proportional to the speed and propagation and inversely proportional to the circular array radius and the nrwms source frequency. For example, take $v = 345/1000 = 0.345$ m/sample, $N = 256$, $\rho = 0.1$ m, $k_{\text{nrwms}} = 0.78$, and $\phi_{\text{BW}} = 2.8$ degree. If we use a larger circular array where $\rho = 0.5$ m, $\phi_{\text{BW}} = 0.6$ degree. The effective beamwidth is useful to determine the angular sampling for the AML maximization. This avoids excessive sampling in the angular space and also prevents further iterations on the AML maximization. Based on the angular sampling by the effective beamwidth, a quadratic polynomial interpolation (concave function) of three points can yield the DOA estimate easily (see Appendix A). The explicit analytical form of the CRB for the circular array is also applicable to a randomly distributed 2D array. For instance, we can compute the RMS distance of the sensors from its centroid and use that as the radius ρ in the circular array formula to obtain the effective beamwidth to estimate the performance of a randomly distributed 2D array. For instance, for a randomly distributed array of 5 sensors at positions $\{(1, 1), (2, 0.8), (3, 1.4), (1.5, 3), (1, 2.5)\}$, the RMS distance of the array to its centroid is 1.14. Since we cannot obtain an explicit analytical form for this random array, we can simply use the circular array formula for $\rho = 1.14$ to obtain the effective beamwidth ϕ_{BW} . For some random arrays, the DOA variance depends highly on the source direction, and an elliptical model is better than the circular one (see Appendix B).

2.3. CRB for source localization

For the near-field case, we also consider the CRB for a single source under three different conditions. The source signal S_c and steering vector in the far-field case are replaced by S_0 and by the steering vector with signal-gain level a_p in

the signal component \mathbf{G} , respectively. For the first case, we can construct the Fisher information matrix by (6), where $\mathbf{H} = \partial\mathbf{G}/\partial\mathbf{r}_s^T$, assuming that \mathbf{r}_s is the only unknown. In this case, $\mathbf{F}_{\mathbf{r}_s} = \zeta\mathbf{A}$, where

$$\mathbf{A} = \sum_{p=1}^R a_p^2 \mathbf{u}_p \mathbf{u}_p^T \quad (11)$$

is the *array matrix* and $\mathbf{u}_p = (\mathbf{r}_s - \mathbf{r}_p)/\|\mathbf{r}_s - \mathbf{r}_p\|$. The \mathbf{A} matrix provides a measure of geometric relations between the source and the sensor array. Poor array geometry may lead to degeneration in the rank of matrix \mathbf{A} . Note that the near-field CRB has the same dependence ζ on the signal as the far-field case.

When the speed of propagation is also unknown, that is, $\Theta = [\mathbf{r}_s^T, \nu]^T$, the \mathbf{H} matrix is given by $\mathbf{H} = [\partial\mathbf{G}/\partial\mathbf{r}_s^T, \partial\mathbf{G}/\partial\nu]$. The Fisher information block matrix for this case is given by

$$\mathbf{F}_{\mathbf{r}_s, \nu} = \zeta \begin{bmatrix} \mathbf{A} & -\mathbf{U}\mathbf{A}_a\mathbf{t} \\ -\mathbf{t}^T\mathbf{A}_a\mathbf{U}^T & \mathbf{t}^T\mathbf{A}_a\mathbf{t} \end{bmatrix}, \quad (12)$$

where $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_R]$, $\mathbf{A}_a = \text{diag}([a_1^2, \dots, a_R^2])$, and $\mathbf{t} = [t_1, \dots, t_R]^T$. By applying the block matrix inversion lemma, the leading $D \times D$ submatrix of the inverse Fisher information block matrix can be given by

$$[\mathbf{F}_{\mathbf{r}_s, \nu}^{-1}]_{11:DD} = \frac{1}{\zeta} (\mathbf{A} - \mathbf{Z}_\nu)^{-1}, \quad (13)$$

where the *penalty matrix* due to the unknown speed of propagation is defined by $\mathbf{Z}_\nu = (1/\mathbf{t}^T\mathbf{A}_a\mathbf{t})\mathbf{U}\mathbf{A}_a\mathbf{t}\mathbf{t}^T\mathbf{A}_a\mathbf{U}^T$. The matrix \mathbf{Z}_ν is nonnegative definite; therefore, the source localization error of the unknown speed of propagation case is always larger than that of the known case.

When the source signal is also unknown, that is, $\Theta = [\mathbf{r}_s^T, |\mathbf{S}_0|^T, \Phi_0^T]^T$, the \mathbf{H} matrix is given by $\mathbf{H} = [\partial\mathbf{G}/\partial\mathbf{r}_s^T, \partial\mathbf{G}/\partial|\mathbf{S}_0|^T, \partial\mathbf{G}/\partial\Phi_0^T]$, where $\mathbf{S}_0 = [S_0(1), \dots, S_0(N/2)]^T$, and $|\mathbf{S}_0|$ and Φ_0 are the magnitude and phase part of \mathbf{S}_0 , respectively. The Fisher information matrix can then be explicitly given by

$$\mathbf{F}_{\mathbf{r}_s, \mathbf{S}_0} = \begin{bmatrix} \zeta\mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{D} \end{bmatrix}, \quad (14)$$

where \mathbf{B} and \mathbf{D} are not explicitly given since they are not needed in the final expression. By applying the block matrix inversion lemma, the leading $D \times D$ submatrix of the inverse Fisher information block matrix can be given by

$$[\mathbf{F}_{\mathbf{r}_s, \mathbf{S}_0}^{-1}]_{11:DD} = \frac{1}{\zeta} (\mathbf{A} - \mathbf{Z}_{\mathbf{S}_0})^{-1}, \quad (15)$$

where the penalty matrix due to the unknown source signal is defined by

$$\mathbf{Z}_{\mathbf{S}_0} = \frac{1}{\sum_{p=1}^R a_p^2} \left(\sum_{p=1}^R a_p^2 \mathbf{u}_p \right) \left(\sum_{p=1}^R a_p^2 \mathbf{u}_p \right)^T. \quad (16)$$

The CRB with the unknown source signal is always larger than that with the known source signal, as discussed below. It can be easily shown that since the penalty matrix $\mathbf{Z}_{\mathbf{S}_0}$ is non-negative definite. The $\mathbf{Z}_{\mathbf{S}_0}$ matrix acts as a penalty term since it is the average of the square of weighted \mathbf{u}_p vectors. The estimation variance is larger when the source is faraway since the \mathbf{u}_p vectors are similar in directions to generate a larger penalty matrix, that is, \mathbf{u}_p vectors add up. When the source is inside the convex hull of the sensor array, the estimation variance is smaller since $\mathbf{Z}_{\mathbf{S}_0}$ approaches zero, that is, \mathbf{u}_p vectors cancel each other. For the 2D case, the CRB for the distance error of the estimated location $[\hat{x}_s, \hat{y}_s]^T$ from the true source location can be given by

$$\sigma_d^2 = \sigma_{x_s}^2 + \sigma_{y_s}^2 \geq [\mathbf{F}_{\mathbf{r}_s, \mathbf{S}_0}^{-1}]_{11} + [\mathbf{F}_{\mathbf{r}_s, \mathbf{S}_0}^{-1}]_{22}, \quad (17)$$

where $d^2 = (\hat{x}_s - x_s)^2 + (\hat{y}_s - y_s)^2$. By further expanding the parameter space, the CRB for multiple source localization can also be derived, but its analytical expression is much more complicated and will not be considered here. The case of the unknown signal and the unknown speed of propagation is also not shown due to its complicated form but numerical similarity to the unknown signal case. Note that when both the source signal and sensor gains are unknown, it is possible to determine the values of the source signal and the sensor gains (they can only be estimated up to a scaled constant).

2.3.1 The circular-array case

In the following, we again consider the uniformly spaced circular array with radius ρ for the near-field CRB. Assume that the source is at distance r_s from the array centroid that is large enough so that the signal-gain levels are uniform, that is, $a_p = a$. Consider the 2D case of unknown source signal, and without loss of generality, let the line of sight (LOS) be the X -axis and let the cross line of sight (CLOS) be the Y -axis. Then, the error covariance matrix is given by

$$\begin{aligned} & [\mathbf{F}_{\mathbf{r}_s, \mathbf{S}_0}^{-1}]_{11:22}(\text{circular array}) \\ &= \frac{1}{\zeta} (\mathbf{A} - \mathbf{Z}_{\mathbf{S}_0})^{-1} \\ &= \begin{bmatrix} \sigma_{\text{LOS}}^2 & 0 \\ 0 & \sigma_{\text{CLOS}}^2 \end{bmatrix} \approx \frac{2r_s^2}{\zeta R a^2 \rho^2} \begin{bmatrix} O\left(\frac{r_s}{\rho}\right) & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \quad (18)$$

The intermediate approximations are given in Appendix C. The above result shows that as r_s increases, the LOS error increases much faster than the CLOS error. For any arbitrary source location, the LOS error is always uncorrelated with the CLOS error. The variance of the DOA estimation is given by $\sigma_{\phi_s}^2 = \sigma_{\text{CLOS}}^2/r_s^2 \approx 2/\zeta R a^2 \rho^2$, which is the same as the far-field case for $a = 1$. The ratio of the CLOS and LOS error can provide a quantitative measure to differentiate far-field from near-field. For example, define far-field as the case when the ratio $r_s/\rho > \gamma$. Then, for a given circular array, we can define far-field as the case when the source range exceeds the array radius γ times. The explicit analytical form of the circular array CRB in the near-field case is again useful for a randomly

distributed 2D array. In the near-field case, the location error bound can be represented by an ellipse, where its major axis represents the LOS error and its minor axis represents the CLOS error.

3. ML SOURCE LOCALIZATION AND DOA ESTIMATION

3.1. Derivation of the ML solution

The derivation of the AML solution for real-valued signals generated by wideband sources is an extension of the classical ML DOA estimator for narrowband signals. Due to the wideband nature of the signal, the AML metric results in a combination of each subband. In the following derivation, the near-field signal model is used for source localization, and the DOA estimation formulation is merely the result of a trivial substitution.

We assume initially that the unknown parameter space is $\Theta = [\tilde{\mathbf{r}}_s^T, \mathbf{S}_0^{(1)T}, \dots, \mathbf{S}_0^{(M)T}]^T$, where the source locations are denoted by $\tilde{\mathbf{r}}_s = [\mathbf{r}_{s_1}^T, \dots, \mathbf{r}_{s_M}^T]^T$ and the source signal spectrum is denoted by $\mathbf{S}_0^{(m)} = [S_0^{(m)}(1), \dots, S_0^{(m)}(N/2)]^T$. By stacking up the $N/2$ positive frequency bins of the signal model in (4) into a single column, we can rewrite the sensor data into an $NR/2 \times 1$ space-temporal frequency vector as $\mathbf{X} = \mathbf{G}(\Theta) + \xi$, where $\mathbf{G}(\Theta) = [\mathbf{S}(1)^T, \dots, \mathbf{S}(N/2)^T]^T$, $\mathbf{S}(k) = \mathbf{D}(k)\mathbf{S}_0(k)$, and $\mathbf{R}_\xi = E[\xi\xi^H] = L\sigma^2\mathbf{I}_{NR/2}$. The log-likelihood function of the complex Gaussian noise vector ξ , after ignoring irrelevant constant terms, is given by $\mathcal{L}(\Theta) = -\|\mathbf{X} - \mathbf{G}(\Theta)\|^2$. The ML estimation of the source locations and source signals is given by the following optimization criterion:

$$\max_{\Theta} \mathcal{L}(\Theta) = \min_{\Theta} \sum_{k=1}^{N/2} \|\mathbf{X}(k) - \mathbf{D}(k)\mathbf{S}_0(k)\|^2, \quad (19)$$

which is equivalent to finding $\min_{\tilde{\mathbf{r}}_s, \mathbf{S}_0(k)} f(k)$ for all k bins, where

$$f(k) = \|\mathbf{X}(k) - \mathbf{D}(k)\mathbf{S}_0(k)\|^2. \quad (20)$$

The minima of $f(k)$, with respect to the source signal vector $\mathbf{S}_0(k)$, must satisfy $\partial f(k)/\partial \mathbf{S}_0^H(k) = 0$, hence the estimate of the source signal vector which yields the minimum residual at any source location is given by

$$\hat{\mathbf{S}}_0(k) = \mathbf{D}^\dagger(k)\mathbf{X}(k), \quad (21)$$

where $\mathbf{D}^\dagger(k) = (\mathbf{D}(k)^H\mathbf{D}(k))^{-1}\mathbf{D}(k)^H$ is the pseudoinverse of the steering matrix $\mathbf{D}(k)$. Define the orthogonal projection $\mathbf{P}(k, \tilde{\mathbf{r}}_s) = \mathbf{D}(k)\mathbf{D}^\dagger(k)$ and the complement orthogonal projection $\mathbf{P}^\perp(k, \tilde{\mathbf{r}}_s) = \mathbf{I} - \mathbf{P}(k, \tilde{\mathbf{r}}_s)$. By substituting (21) into (20), the minimization function becomes $f(k) = \|\mathbf{P}^\perp(k, \tilde{\mathbf{r}}_s)\mathbf{X}(k)\|^2$. After substituting the estimate of $\mathbf{S}_0(k)$, the AML source locations estimate can be obtained by solving the following maximization problem:

$$\max_{\tilde{\mathbf{r}}_s} J(\tilde{\mathbf{r}}_s) = \max_{\tilde{\mathbf{r}}_s} \sum_{k=1}^{N/2} \|\mathbf{P}(k, \tilde{\mathbf{r}}_s)\mathbf{X}(k)\|^2. \quad (22)$$

Note that the AML metric $J(\tilde{\mathbf{r}}_s)$ has an implicit form for the estimation of $\mathbf{S}_0(k)$, whereas the metric $\mathcal{L}(\Theta)$ shows the explicit form. Once the AML estimate of $\tilde{\mathbf{r}}_s$ is obtained, the AML estimate of the source signals can be given by (21). Similarly, in the far-field case, the unknown parameter vector contains only the DOAs, that is, $\phi_s = [\phi_s^{(1)}, \dots, \phi_s^{(M)}]^T$. Thus, the AML DOA estimation can be obtained by $\arg \max_{\phi_s} \sum_{k=1}^{N/2} \|\mathbf{P}(k, \phi_s)\mathbf{X}(k)\|^2$. It is interesting that, when zero padding is applied, the covariance matrix \mathbf{R}_ξ is no longer diagonal and is indeed singular; thus, an exact ML solution cannot be derived without the inverse of \mathbf{R}_ξ . In the above formulation, we derive the AML solution using only a single block. A different AML solution using multiple blocks could also be formed with some possible computational advantages. When the speed of propagation is unknown, as in the case of seismic media, we may expand the unknown parameter space to include it, that is, $\Theta = [\tilde{\mathbf{r}}_s^T, v]^T$.

3.2. Single-source case

In the single-source case, the AML metric in (22) becomes $J(\mathbf{r}_s) = \sum_{k=1}^{N/2} |B(k, \mathbf{r}_s)|^2$, where $B(k, \mathbf{r}_s) = \bar{\mathbf{d}}(k, \mathbf{r}_s)^H \mathbf{X}(k)$ is the beam-steered beamformer output in the frequency domain [12], $\bar{\mathbf{d}} = \mathbf{d}/\sqrt{\sum_{p=1}^R a_p^2}$ is the normalized steering vector, and $\bar{a}_p = a_p/\sqrt{\sum_{p=1}^R a_p^2}$ is the normalized signal-gain level at the p th sensor. It is interesting to note that in the near-field case, the AML beamformer output is the result of forming a focused spot (or area) on the source location rather than a beam since the range is also considered. In the far-field case, the AML metric becomes $J(\phi_s)$. In [8], the AML criterion is shown to be equivalent to maximizing the weighted cross correlations between sensor data, which is commonly used for estimating relative time delays.

The source location can be estimated, based on where, $J(\mathbf{r}_s)$ is maximized for a given set of locations. Define the normalized metric

$$J_N(\mathbf{r}_s) \equiv \frac{\sum_{k=1}^{N/2} |B(k, \mathbf{r}_s)|^2}{J_{\max}} \leq 1, \quad (23)$$

where $J_{\max} = \sum_{k=1}^{N/2} [\sum_{p=1}^R \bar{a}_p |X_p(k)|]^2$, which is useful to verify estimated peak values. Without any prior information on possible region of the source location, the AML metric should be evaluated on a set of grid points. A nonuniform grid is suggested to reduce the number of grid points. For the 2D case, polar coordinates with nonuniform sampling of the range and uniform sampling of the angle can be transformed to Cartesian coordinates that are dense near the array and sparse away from the array. When the crude estimate of the source location is obtained from the grid-point search, iterative methods can be applied to reach the global maximum (without running into local maxima, given appropriate choice of grid points). In some cases, grid-point search is not necessary since a good initial location estimate is available from, for example, the estimate of the previous data frame for a slowly moving source. In this paper, we consider the Nelder-Mead direct search method [13] for the purpose of performance evaluation.

3.3. Multiple-sources case

For the multiple-sources case, the parameter estimation is a challenging task. Although iterative multidimensional parameter search methods such as the Nelder-Mead direct search method can be applied to avoid an exhaustive multidimensional grid search, finding the initial source location estimates is not trivial. Since iterative solutions for the single-source case are more robust and the initial estimate is easier to find, we extend the AP method in [14] to the near-field problem. The AP approach breaks the multidimensional parameter search into a sequence of single-source-parameter search, and yields fast convergence rate. The following describes the AP algorithm for the two-sources case, but it can be easily extended to the case of M sources. Let $\Theta = [\Theta_1^T, \Theta_2^T]^T$ be either the source locations in the near-field case or the DOAs in the far-field case.

AP algorithm 1.

Step 1. Estimate the location/DOA of the stronger source on a single-source grid

$$\Theta_1^{(0)} = \arg \max_{\Theta_1} J(\Theta_1). \quad (24)$$

Step 2. Estimate the location/DOA of the weaker source on a single-source grid under the assumption of a two-source model while keeping the first source location estimate from Step 1 constant

$$\Theta_2^{(0)} = \arg \max_{\Theta_2} J\left([\Theta_1^{(0)T}, \Theta_2^T]^T\right). \quad (25)$$

Step 3. Iterative AML parameter search (direct or gradient search) for the location/DOA of the first source while keeping the estimate of the second source location from the previous iteration constant

$$\Theta_1^{(i)} = \arg \max_{\Theta_1} J\left([\Theta_1^T, \Theta_2^{(i-1)T}]^T\right). \quad (26)$$

Step 4. Iterative AML parameter search (direct or gradient search) for the location/DOA of the second source while keeping the estimate of the first source location from Step 3 constant

$$\Theta_2^{(i)} = \arg \max_{\Theta_2} J\left([\Theta_1^{(i)T}, \Theta_2^T]^T\right). \quad (27)$$

For $i = 1, \dots$ (repeat Steps 3 and 4 until convergence).

4. SIMULATION EXAMPLES AND EXPERIMENTAL RESULTS

4.1. Cramér-Rao bound example

In the following simulation examples, we consider a prerecorded tracked vehicle signal with significant spectral content of about 50-Hz bandwidth centered about a dominant frequency at 100 Hz. The sampling frequency is set to

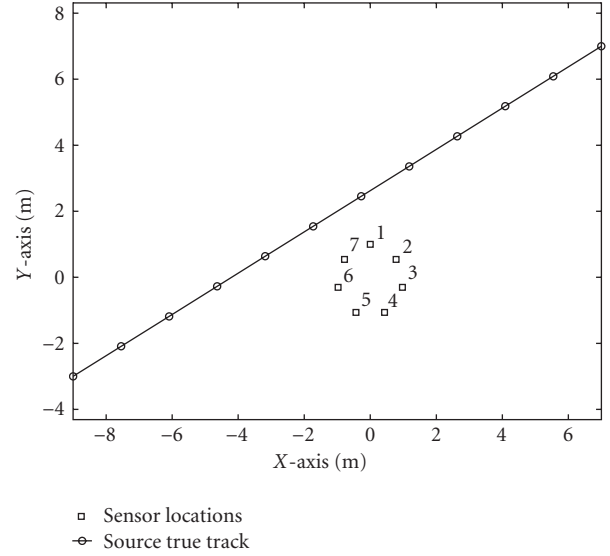


FIGURE 2: Single-traveling-source scenario. Uniformly spaced circular array of 7 elements.

be 1 kHz and the speed of propagation is 345 m/s. The data length $L = 200$ (which corresponds to 0.2 second), the DFT size $N = 256$ (zero padding), and all positive frequency bins are considered. We consider a single-traveling-source scenario for a circular array of seven elements (uniformly spaced on the circumference), as depicted in Figure 2. In this case, we consider the spatial loss that is a function of the distance from the source location to each sensor location, thus the gains a_p 's are not uniform. To compare the theoretical performance of source localization under different conditions, we compare the CRB for the known source signal and speed of propagation, for the unknown speed of propagation, and for the unknown source signal cases for this single-traveling-source scenario. As depicted in Figure 3, the unknown source signal is shown to be a much more significant parameter factor than the unknown speed of propagation in source location estimation. However, these parameters are not significant in the DOA estimations.

4.2. Single-source experimental results

Several acoustic experiments were conducted in Xerox PARC, Palo Alto, Calif, USA. The experimental data was collected indoor as well as outdoor by half to a dozen omnidirectional microphones. A semianechoic chamber with sound absorbing foams attached to the walls and ceiling (shown to have a few dominant reflections) was used for the indoor data collection. An omnidirectional loud speaker was used as the sound source. In one indoor experiment, the source is placed in the middle of the rectangular room of dimension 3×5 m surrounded by six microphones (convex hull configuration), as depicted in Figure 4. The sound of a moving light-wheeled vehicle is played through the speaker and collected by the microphone array. Under 12 dB SNR, the speaker location can be accurately estimated (for every 0.2 second of data)

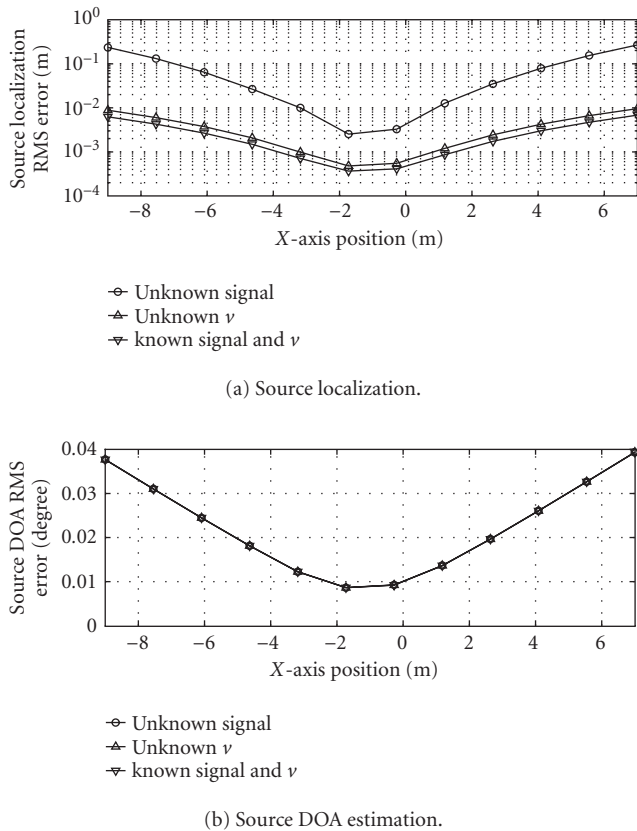


FIGURE 3: CRB comparison for the traveling-source scenario ($R = 7$): (a) localization bound, and (b) DOA bound.

with an RMS error of 73 cm using the near-field AML source localization algorithm. An RMS error of 127 cm is reported the same data using the two-step LS method. This shows that both methods are capable of locating the source despite some minor reverberation effects.

In the outdoor experiment (next to Xerox PARC building), three widely separated linear subarrays, each with four microphones (1 ft interelement spacing), are used. A stationary noise source (possibly air conditioning) is observed from an adjacent building. To demonstrate the effectiveness of the algorithms in handling wideband signals, a white Gaussian signal is played through the loud speaker placed at the two locations (from two independent runs) shown in Figure 5. In this case, each subarray estimates the DOA of the source independently using the AML method, and the bearing crossing (see Appendix D) from the three subarrays (labeled as A, B, and C in the figures) provides an estimate of the source location. The estimation is again performed for every 0.2 second of data. An RMS error of 32 cm is reported for the first location, and an RMS error of 97 cm is reported for the second location. Then, we apply the two-step LS DOA estimation to the same data, which involves relative time-delay estimation among the Gaussian signals. Poorer results are shown in Figure 6, where an RMS error of 152 cm is reported for the first location, and an RMS error of 472 cm is

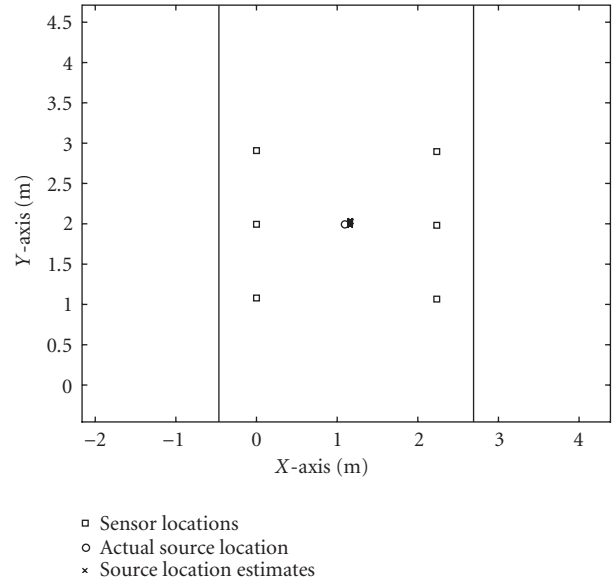


FIGURE 4: AML source localization of a vehicle sound in a semi-anechoic chamber.

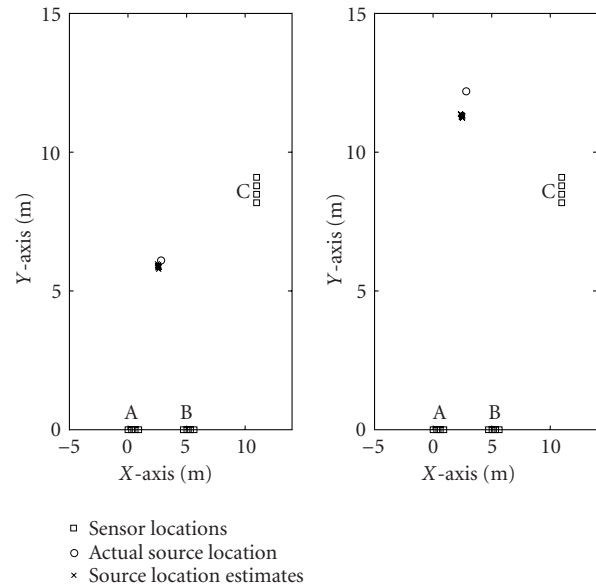


FIGURE 5: Source localization of white Gaussian signal using AML DOA cross bearing in an outdoor environment.

reported for the second location. This shows that when the source signal is truly wideband, the time-delay-based techniques can yield very poor results. In other outdoor runs, the AML method was also shown to yield good results for music signals.

Then, a moving source experiment is conducted by placing the loud speaker on a cart that moves on a straight line from the top to the bottom of Figure 7. The vehicle sound is again played through the speaker while the cart is moving. We assume that the source location is stationary within each

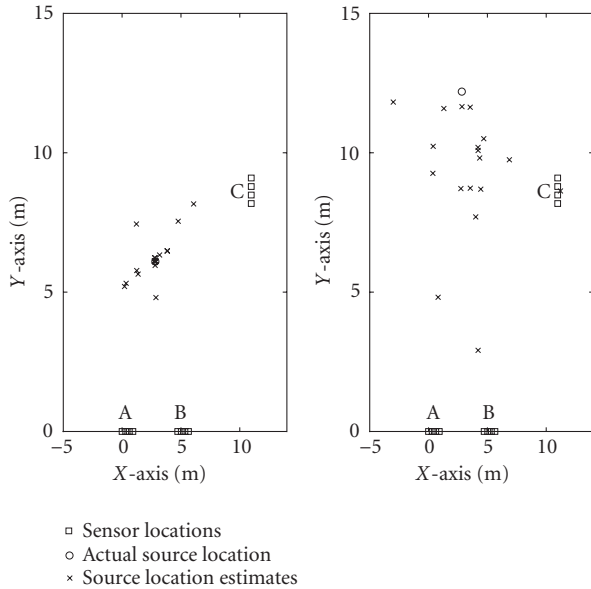


FIGURE 6: Source localization of white Gaussian signal using LS DOA cross bearing in an outdoor environment.

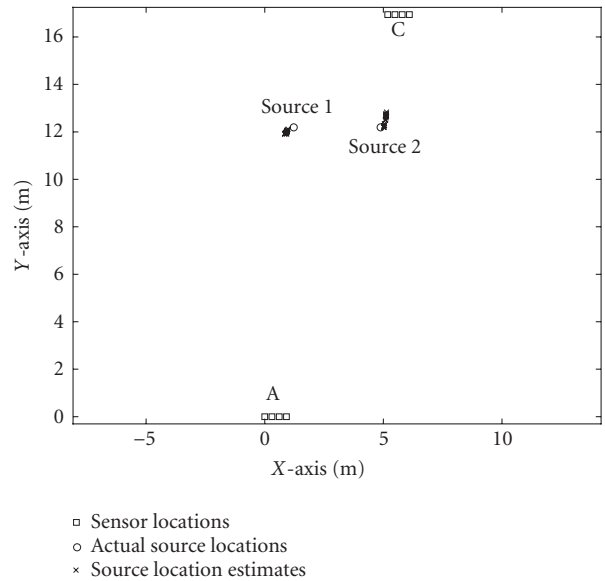


FIGURE 8: Two-source localization using AML DOA cross bearing with AP in an outdoor environment.

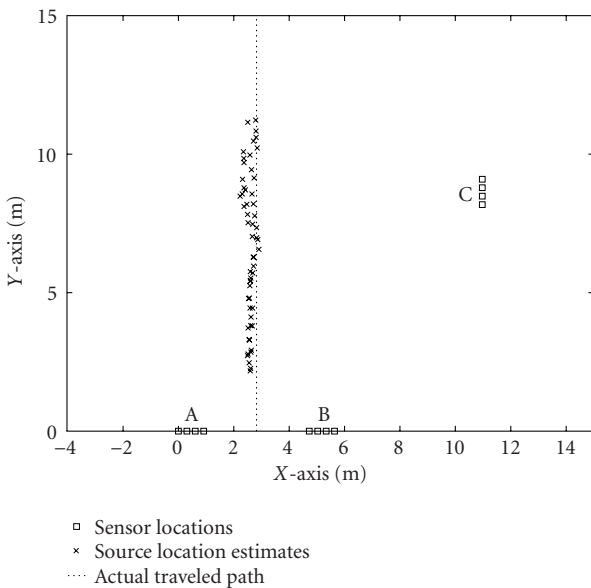


FIGURE 7: Source localization of a moving speaker (vehicle sound) using AML DOA cross bearing in an outdoor environment.

data frame of about 0.1 second, and the DOA is estimated for each frame using the AML method. The source location is again estimated by the cross bearing of the three DOAs. As shown in Figure 7, the source can be well estimated to be very close to the actual traveled path. The results using the LS method (not shown) are much worse when the source is faraway.

4.3. Two-source experimental results

In a different outdoor configuration, two linear subarrays (labeled as A and C), each consisting of four microphones, are placed at the opposite sides of the road and two omnidirectional loud speakers are placed between them, as depicted in Figure 8. The two loud speakers play two independent prerecorded sounds of light-wheeled vehicles of different kinds. By using the AP steps on the AML metric, the DOAs of the two sources are jointly estimated for each array under 11 dB SNR (with respect to the bottom array). Then, the cross bearing yields the location estimates of the two sources. The estimation is performed for every 0.2 second of data. An RMS error of 37 cm is observed for source 1 and an RMS error of 45 cm is observed for source 2. Note that the range estimate of the second source is slightly worse than that of the first source because the bearings from the two arrays are close to being collinear for the second source.

Another two-source localization experiment was also conducted inside the semianechoic chamber. In this setup, twelve microphones are placed in a linear manner near one of the walls. Two speakers are placed inside the room, as depicted in Figure 9. The microphones are then divided into three nonoverlapping groups (subarrays, labeled as A, B, and C), each with four elements. Each subarray performs the AML DOA estimation using AP. The cross bearing of the DOAs again provides the location estimate of the two sources. The estimation is again performed for every 0.2 second of data. An RMS error of 154 cm is observed for the first source, and an RMS error of 35 cm is observed for the second source. Since the bearing angles are not too different across the three subarrays, the source range estimate becomes poor, especially for source 1. This again suggests that

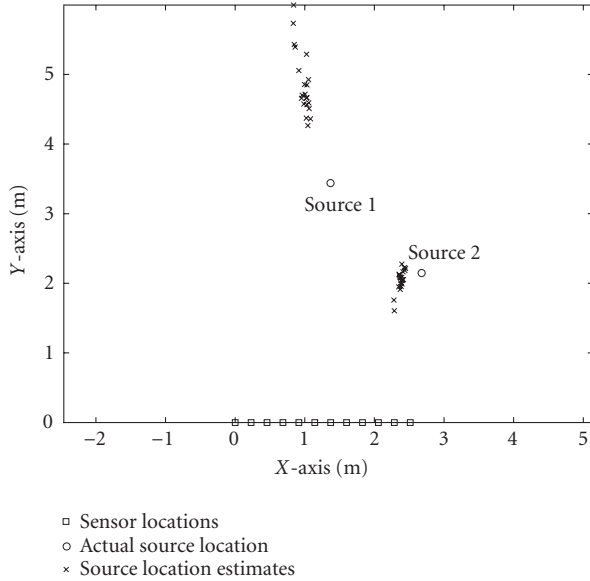


FIGURE 9: Two-source localization using AML DOA cross bearing with AP in a semianechoic chamber.

the geometry of the subarrays used in this experiment was far from ideal, and widely separated subarrays would have yielded better triangulation (cross bearing) results.

5. CONCLUSION

In this paper, the theoretical CRBs for source localization and DOA estimation are analyzed and the AML source localization and DOA estimation methods are shown to be effective as applied to measured data. For the single-source case, the AML performance is shown to be superior to that of the two-step LS method in various types of signals, especially for the truly wideband ones. The AML algorithm is also shown to be effective in locating two sources using AP. The CRB analysis suggests the uniformly spaced circular array as the preferred array geometry for most scenarios. When a circular array is used, the DOA variance bound is independent of the source direction, and it also does not degrade when the speed of propagation is unknown. The CRB also proves the physical observations which favor high energy in the higher-frequency components of a signal. The sensitivity of source localization to different unknown parameters has also been analyzed. It has been shown that unknown source signal results in a much larger error in range than that of unknown speed of propagation, but those parameters are not significant in DOA estimation.

APPENDICES

A. DOA ESTIMATION USING INTERPOLATION

Denote the three data points $\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ as the angular samples and their corresponding AML function

values, where y_2 is the overall maximum and the other two are the adjacent samples. By the Lagrange interpolation polynomial formula [15], we can obtain a quadratic polynomial that interpolates the three data points. The angle (or the DOA estimate) that yields the maximum value of the quadratic polynomial is given by

$$\hat{x} = \frac{c_1(x_2 + x_3) + c_2(x_1 + x_3) + c_3(x_1 + x_2)}{2(c_1 + c_2 + c_3)}, \quad (\text{A.1})$$

where $c_1 = y_1/(x_1 - x_2)/(x_1 - x_3)$, $c_2 = y_2/(x_2 - x_1)/(x_2 - x_3)$, and $c_3 = y_3/(x_3 - x_1)/(x_3 - x_2)$. The interpolation step avoids further iterations on the AML maximization.

B. THE ELLIPTICAL MODEL OF DOA VARIANCE

In Section 2.2.1, we show that we can conveniently define an effective beamwidth for a uniformly spaced circular array. This gives us one measure of the beamwidth that is independent of the source direction. When we have randomly distributed arrays, the circular CRB may be a reasonable approximation if the sensors are distributed uniformly in both the X and Y directions. However, in some cases, the sensors may span more in one direction than the other. In that case, we may model the effective beamwidth using an ellipse. The direction of the major axis indicates the best DOA performance, where a small beamwidth can be defined. The direction of the minor axis indicates the poorest DOA performance, and a large beamwidth is defined in that direction. This suggests the use of a variable beamwidth as a function of angle, which is useful for the AML metric evaluation.

First, we need to determine the orientation of the ellipse for an arbitrary 2D array. Without loss of generality, we define the origin at the array centroid $\mathbf{r}_c = [x_c, y_c]^T = [0, 0]^T$. Let there be a total of R sensors. The location of the p th sensor is denoted as $\mathbf{r}_p = [x_p, y_p]^T$ in the coordinate system. Our objective is to find a rotation angle ψ from the X-axis such that the cross terms of the new sensor locations are summed to zero. The major and minor axes will be the new X- and Y-axes. Denote $[x'_p, y'_p]^T$ as the new coordinate of the p th sensor in the rotated coordinate system. The new coordinate has the following relation with the old coordinate:

$$\begin{aligned} x'_p &= x_p \cos \psi + y_p \sin \psi, \\ y'_p &= -x_p \sin \psi + y_p \cos \psi. \end{aligned} \quad (\text{B.1})$$

The sum of the cross terms is then given by

$$\sum_{p=1}^R x'_p y'_p = c_1 \cos \psi \sin \psi + c_2 (1 - 2 \sin^2 \psi), \quad (\text{B.2})$$

where $c_1 = \sum_{p=1}^R (y_p^2 - x_p^2)$ and $c_2 = \sum_{p=1}^R x_p y_p$. After double angle substitutions and some algebraic manipulation to equate the above to zero, we obtain the solution

$$\psi = -\frac{1}{2} \tan^{-1} \left(\frac{2c_2}{c_1} \right) + \frac{\pi}{2} \ell, \quad (\text{B.3})$$

for $\ell = 0$ and 1, which means that the two solutions that are different by 90 degrees exist.

We have shown that, for a circular array, the DOA variance bound is given by $1/\zeta\alpha$, where $\alpha = \rho^2 R/2$. For an ellipse with the center at the origin, the corresponding $\alpha = \sum_{p=1}^R b_p^2 = \cos^2 \phi_s \sum_{p=1}^R (x'_p)^2 + \sin^2 \phi_s \sum_{p=1}^R (y'_p)^2$. Note that the cross terms become zero in this case. To put the above in a form similar to that of the circular array, we can write $\alpha = R[V_x \cos^2 \phi_s + V_y \sin^2 \phi_s]$, where $V_x = (1/R) \sum_{p=1}^R (x'_p)^2$ and $V_y = (1/R) \sum_{p=1}^R (y'_p)^2$. Note that at the major or the minor axis, the source angles are either 0 degree or 90 degrees. This means that $\alpha = RV_x$ or RV_y , depending on which axis is the major or minor axis. Define $\rho_x = \sqrt{2V_x}$ and $\rho_y = \sqrt{2V_y}$. These two values can be used to determine the largest and the smallest beamwidth for the ellipse, that is, $\phi_{BW,x} \equiv v/\pi\rho_x k_{nrwms}$ and $\phi_{BW,y} \equiv v/\pi\rho_y k_{nrwms}$.

C. CIRCULAR ARRAY CRB APPROXIMATIONS

The approximations used in the near-field circular array CRB involve several steps, including the approximations for \mathbf{A} and \mathbf{Z}_{S_0} . The array matrix \mathbf{A} defined in (11) can be given explicitly by

$$\begin{aligned} \mathbf{A} &\simeq a^2 \sum_{p=1}^R \mathbf{u}_p \mathbf{u}_p^T \\ &= \frac{Ra^2}{r_s^2} \left[\left(1 - \frac{\rho^2}{r_s^2} + O\left(\frac{\rho^3}{r_s^3}\right) \right) \mathbf{r}_s \mathbf{r}_s^T \right. \\ &\quad \left. + \frac{\rho^2}{2} \left(1 + \frac{2\rho^2}{r_s^2} + O\left(\frac{\rho^3}{r_s^3}\right) \right) \mathbf{I} \right], \end{aligned} \quad (\text{C.1})$$

where uniform gain a is assumed and power series expansion for $R > 3$, preserving only the second order, is used to obtain the final expression. Similarly, the penalty matrix \mathbf{Z}_{S_0} can be approximated by

$$\mathbf{Z}_{S_0} \simeq \frac{Ra^2}{r_s^2} \left(1 - \frac{\rho^2}{2r_s^2} + O\left(\frac{\rho^3}{r_s^3}\right) \right) \mathbf{r}_s \mathbf{r}_s^T, \quad (\text{C.2})$$

which also uses power series expansion preserving only the second order. After some simplifications, the difference matrix can be given by

$$\begin{aligned} \mathbf{A} - \mathbf{Z}_{S_0} &\simeq \frac{Ra^2}{r_s^2} \left(\frac{\rho^2}{2} \mathbf{I} - \frac{\rho^2}{2r_s^2} \mathbf{r}_s \mathbf{r}_s^T + O\left(\frac{\rho^3}{r_s^3}\right) \mathbf{r}_s \mathbf{r}_s^T \right) \\ &= \frac{Ra^2 \rho^2}{2r_s^2} \begin{bmatrix} O\left(\frac{\rho}{r_s}\right) & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \quad (\text{C.3})$$

where $\mathbf{r}_s \mathbf{r}_s^T / r_s^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ for this coordinate system. Hence, the final approximation for the inverse Fisher information matrix is given by

$$\frac{1}{\zeta} (\mathbf{A} - \mathbf{Z}_{S_0})^{-1} \simeq \frac{2r_s^2}{\zeta Ra^2 \rho^2} \begin{bmatrix} O\left(\frac{r_s}{\rho}\right) & 0 \\ 0 & 1 \end{bmatrix}. \quad (\text{C.4})$$

D. SOURCE LOCALIZATION VIA BEARING CROSSING

When two or more subarrays simultaneously detect the same source, the crossing of the bearing lines can be used to estimate the source location. This step is often called triangulation. Without loss of generality, let the centroid of the first subarray be the origin of the coordinate system. Denote $\mathbf{r}_{c_k} = [x_{c_k}, y_{c_k}]^T$ as the centroid position of the k th subarray, for $k = 1, \dots, K$. Denote ϕ_k as the DOA estimate (with respect to north) of the k th subarray. Then, the following system of linear equations can yield the bearing crossing solution

$$\begin{aligned} &\begin{bmatrix} \cos(\phi_1) & -\sin(\phi_1) \\ \vdots & \vdots \\ \cos(\phi_K) & -\sin(\phi_K) \end{bmatrix} \begin{bmatrix} x_s \\ y_s \end{bmatrix} \\ &= \begin{bmatrix} x_{c_1} \cos(\phi_1) - y_{c_1} \sin(\phi_1) \\ \vdots \\ x_{c_K} \cos(\phi_K) - y_{c_K} \sin(\phi_K) \end{bmatrix}. \end{aligned} \quad (\text{D.1})$$

Note that the source location $[x_s, y_s]^T$ is defined in the coordinate system with respect to the centroid of the first subarray.

ACKNOWLEDGMENTS

This work was partially supported by DARPA-ITO under Contract N66001-00-1-8937. The authors wish to thank J. Reich, P. Cheung, and F. Zhao of Xerox PARC for conducting and planning the experiments presented in this paper.

REFERENCES

- [1] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 30–39, 2002.
- [2] M. S. Brandstein and D. Ward, *Microphone Arrays: Techniques and Applications*, Springer-Verlag, Berlin, Germany, September 2001.
- [3] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 35, no. 12, pp. 1661–1669, 1987.
- [4] H. C. Schau and A. Z. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 35, no. 8, pp. 1223–1225, 1987.
- [5] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 1905–1915, 1994.
- [6] M. S. Brandstein, J. E. Adcock, and H. F. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 1, pp. 45–50, 1997.
- [7] K. Yao, R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1555–1567, 1998.
- [8] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp. 1843–1854, 2002.

- [9] P. J. Chung, M. L. Jost, and J. F. Böhme, "Estimation of seismic-wave parameters and signal detection using maximum-likelihood methods," *Computers and Geosciences*, vol. 27, no. 2, pp. 147–156, 2001.
- [10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory, Vol. 1*, Prentice-Hall, New Jersey, NJ, USA, 1993.
- [11] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, New Jersey, NJ, USA, 2000.
- [12] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*, Prentice-Hall, New Jersey, NJ, USA, 1993.
- [13] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.
- [14] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.
- [15] R. L. Burden and J. D. Faires, *Numerical Analysis*, PWS Publishing, Boston, Mass, USA, 5th edition, 1993.

Joe C. Chen was born in Taipei, Taiwan, in 1975. He received the B.S. (with honors), M.S., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), in 1997, 1998, and 2002, respectively. From 1997 to 2002, he was with the Sensors and Electronics Systems group of Raytheon Systems Company (formerly Hughes Aircraft), El Segundo, Calif. From 1998 to 2002, he was a Research Assistant at UCLA, and from 2001 to 2002, he was a Teacher Assistant at UCLA. Since 2002, he joined TRW Space & Electronics, Redondo Beach, Calif, as a Senior Member of the Technical Staff. His research interests include estimation theory and statistical signal processing as applied to sensor array systems, communication systems, and radar. Dr. Chen is a member of Tao Beta Pi and Eta Kappa Nu honor societies and the IEEE.



Kung Yao received the B.S.E., M.A., and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ. He has worked at the Princeton-Penn Accelerator, the Brookhaven National Lab, and the Bell Telephone Labs, Murray Hill, NJ. He was a NAS-NRC Postdoctoral Research Fellow at the University of California, Berkeley. He was a Visiting Assistant Professor at the Massachusetts Institute of Technology and a Visiting Associate Professor at the Eindhoven Technical University. In 1985–1988, he was an Assistant Dean of the School of Engineering and Applied Science at UCLA. Presently, he is a Professor in the Electrical Engineering Department at UCLA. His research interests include sensor array systems, digital communication theory and systems, wireless radio systems, chaos communications and system theory, and digital and array signal processing. He has published more than 250 papers. He received the IEEE Signal Processing Society's 1993 Senior Award in VLSI Signal Processing. He is the coeditor of *High Performance VLSI Signal Processing* (IEEE Press, 1997). He was on the IEEE Information Theory Society's Board of Governors and is a member of the Signal Processing System Technical Committee of the IEEE Signal Processing Society. He has been on the editorial boards of various IEEE Transactions, with the most recent being *IEEE Communications Letters*. He is a Fellow of the IEEE.



Ralph E. Hudson received his B.S. degree in electrical engineering from the University of California at Berkeley in 1960 and the Ph.D. degree from the US Naval Postgraduate School, Monterey, Calif, in 1969. In the US Navy, he attained the rank of Lieutenant Commander and served with the Office of Naval Research and the Naval Air Systems Command. From 1973 to 1993, he was with Hughes Aircraft Company, and since then he has been a Research Associate in the Electrical Engineering Department at the University of California at Los Angeles. His research interests include signal and acoustic and seismic array processing, wireless radio, and radar systems. He received the Legion of Merit and Air Medal, and the Hyland Patent Award in 1992.

