Maximum Likelihood Multiple-Source Localization Using Acoustic Energy Measurements with Wireless Sensor Networks

Xiaohong Sheng and Yu-Hen Hu, *Fellow, IEEE*

Abstract—A maximum likelihood (ML) acoustic source location estimation method is presented for the application in a wireless *ad hoc* sensor network. This method uses acoustic signal energy measurements taken at individual sensors of an *ad hoc* wireless sensor network to estimate the locations of multiple acoustic sources. Compared to the existing acoustic energy based source localization methods, this proposed ML method delivers more accurate results and offers the enhanced capability of multiple source localization. A multiresolution search algorithm and an expectation–maximization (EM) like iterative algorithm are proposed to expedite the computation of source locations. The Cramér–Rao Bound (CRB) of the ML source location estimate has been derived. The CRB is used to analyze the impacts of sensor placement to the accuracy of location estimates for single target scenario. Extensive simulations have been conducted. It is observed that the proposed ML method consistently outperforms existing acoustic energy based source localization methods. An example applying this method to track military vehicles using real world experiment data also demonstrates the performance advantage of this proposed method over a previously proposed acoustic energy source localization method.

Index Terms—Acoustic energy, acoustic intensity, maximum likelihood, nonlinear least square, sensor network, source localization.

I. INTRODUCTION

THE emergence of miniature low-power devices that integrate micro-sensing and actuation with on-board processing and wireless communication capabilities has stimulated great interests in developing wireless ad hoc sensor network [1], [2]. In an *ad hoc* sensor network, the sensors' locations do not correspond to a particular structure. However, the location of each sensor can be assumed to be known in advance.

A wireless *ad hoc* sensor network often performs monitoring tasks such as detection, classification, localization, and tracking of one or more targets in the sensor field. The sensors are typically battery powered and have limited wireless communication bandwidth. Therefore, efficient collaborative signal processing

algorithms that consume less energy for computation and less bandwidth for communication are needed [3].

In this paper, we focus on the task of source location estimation using passive and stationary acoustic sensors (microphones) in a wireless *ad hoc* sensor network. Source localization using acoustic sensors has found in numerous applications. In sonar signal processing, the focus is on locating underwater acoustic sources using an array of hydrophones [4]. In video conference and multimedia human computer interface applications, microphone arrays have been developed to locate and track speakers head position in a room environment [6]–[12]. Acoustic signatures have also been used to estimate vehicle locations in an open-field sensor network [13], [14].

Existing acoustic source localization methods make use of three types of physical measurements: time delay of arrival (TDOA), direction of arrival (DOA), and received signal strength or energy. DOA can be estimated by exploiting the phase difference measured at receiving sensors [15]–[19] and is applicable when the acoustic source emits a coherent, narrowband signal. TDOA is suitable for broadband acoustic source localization and has been extensively investigated [5], [6], [13], [14], [20], [21]. It requires accurate measurements of the relative time delay between sensor nodes. It is known that the intensity or equivalently the energy of acoustic signal attenuates as a function of distance from the source. Using this property, an energy-based acoustic source localization method has been reported recently [22] for locating single target in an open-terrain acoustic sensor field.

For wireless *ad hoc* sensor network applications, energy (intensity) based acoustic features is an appropriate choice since the acoustic power emitted by targets such as moving vehicles usually varies slowly with respect to time. As such, the acoustic energy time series can be sampled at a much lower rate compared to the raw acoustic time series. Furthermore, the positions of ground moving targets need not be updated too often. Therefore, few data will need to be transmitted to a data fusion center via the often congested wireless communication channels. This, in terms, will reduce the energy consumption for data transmission on individual sensor nodes and will reduce the demand of communication bandwidth over shared wireless channels. The coarse time interval for computing each new energy reading also greatly relieved the burden of accurate time synchronization among sensors using the wireless communication channel. Moreover, acoustic signal intensity

measurements can also be used to detect the presence of a target. Hence, there is no need to compute additional features for the task of source localization.

The method to be presented in this paper is based on a maximum likelihood estimation of both the source locations and corresponding acoustic energy readings. Compared to a previously reported method [22], this proposed method promises two significant advantages.

- The ML method can handle more than one targets within a sensor field. The previously reported method can handle only single target localization.
- The ML method yields higher accuracy in terms of source location estimates compared to the earlier methods.

In this paper, an acoustic signal energy attenuation model as a function of source-to-sensor distance is first established. Based on this model, the acoustic source localization problem is formulated as a maximum likelihood estimation problem. Two complementary methods are proposed to solve this nonlinear optimization problem. The first method is based on a projection formulation and uses multiresolution search to expedite computation. The second method is based on an Expectation–Maximization (EM) like iterative algorithm. This method has less computation complexity but may converge to a local minimum. A formulation of the Cramér–Rao Bound (CRB) of the location estimates is derived. When there is only a single target, the corresponding CRB can be used to analyze the impact of sensor deployment strategy on the localization accuracy. Specifically, the formula clearly verified the common sense that smaller estimation error is achieved if the sensors are densely and uniformly laid out over the sensor field.

The remaining of this paper is organized as follows: In Section II, the acoustic energy attenuation model is developed. In Section III, the ML source location estimation method is derived. Both the projection and the EM solution methods are also presented. In Section IV, the CRB is derived. In particular, in the single source case, CRB is employed to explore the impacts of different sensor deployment strategies on the localization accuracy. Experiments and simulations are provided in Section V. Extensive simulation have been conducted to compare the performance of the proposed ML method to three existing energy-based source localization methods. Then, time series data samples obtained from a field experiment conducted in California during November 2001 are used to demonstrate the feasibility of using this proposed method to solve real world sensor network source localization problem.

II. ACOUSTIC ENERGY ATTENUATION MODEL

Let N be the number of acoustic sensors and K be the number of acoustic sources in the sensor field. In theory [23], the intensity of an acoustic signal emitted omni-directionally from a point sound source and propagating through ground surface will attenuate at a rate that is inversely proportional to the distance from the source. In this paper, we will further assume that the acoustic intensities of the K sources will be linearly superimposed without any interaction among them. The acoustic signal received at the ith sensor $i = 1, 2, \ldots N$ during time interval n then can be expressed as

$$x_i(n) = s_i(n) + \nu_i(n) \qquad (1)$$

where

$$s_i(n) = \gamma_i \sum_{k=1}^{K} \frac{a_k(n - t_{ki})}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|} \qquad (2)$$

is the acoustic intensity measured at the ith sensor due to all K acoustic sources, and $\nu_i(n)$ is the background noise. In this paper, $\nu_i(n)$ is modeled as a zero-mean additive white Gaussian (AWGN) noise random variable with variance ς_i^2. $a_k(n - t_{ki})$ is the intensity of the kth acoustic source measured 1 m away from that source, and t_{ki} is the propagation delay of the acoustic signal from the kth source to the ith sensor. $a_k(n - t_{ki})$ will be modeled as a random variable uncorrelated with each other. $\boldsymbol{\rho}_k$ is an unknown $p \times 1$ vector denoting the position vector of the kth target; and $\mathbf{r}_i$ is a given $p \times 1$ vector denoting the position vector of the ith stationary sensor, where p is the dimension of the location coordinates. γ_i is sensor gain factor of the ith acoustic sensor.

We remark that this model does not account for the potential echoes and reverberation [24] due to the presence of obstacles such as man-made walls or rocky hills. It also does not account for the potential impacts of wind direction [25], [26] and dense vegetation [27] in the sensor field. Moreover, in practice, some of the assumptions made in deriving this model may not be always true and, hence, may affect its accuracy. For example, the engine sound of a vehicle may not be omni-directional and will be biased toward the side closer to the engine. The physical size of the acoustic source may be too large to be adequately modeled as a point source for sensors very close to the source. In an outdoor environment, strong background noise, including wind gusts, may be encountered during operation. In addition, the gain of individual microphones will need to be calibrated to yield consistent acoustic energy readings. However, we are optimistic that these simplifications will not hinder the application of this proposed localization algorithm for two reasons: First, the acoustic energy is measured over a rather long period of time (in the order of tenth of a second) compared to the time constants of some of these physical anomalies. As such, the impacts they may cause will be diminished due to time average. Second, the measurements are made on large number of sensors that spread over the sensor field. Not all sensors will suffer from the same type of physical anomaly.

Assume that $s_i(n)$ and $\nu_i(n)$ are uncorrelated such that $E\{s_i(n)\nu_i(n)\} = E\{s_i(n)\}E\{\nu_i(n)\} = 0$. Then, one may represent the acoustic energy as (setting $g_i = \gamma_i^2$ and $S_k(n - t_{ki}) = E\left[a_k^2(n - t_{ki})\right]$):

$$E\left[s_i^2(n)\right] = \gamma_i^2 \sum_{k=1}^{K} \frac{E\left[a_k^2(n - t_{ki})\right]}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|^2}$$

$$= g_i \sum_{k=1}^{K} \frac{S_k(n - t_{ki})}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|^2}.$$

In practice, the expectation is realized using the ensemble average over a time window $T = M/f_s$, where M is the number of sample points used for estimating the acoustic energy received by the sensor during this time interval, and f_s is the sampling frequency. Denoting this average energy measurements over the time window $[t - T/2, t + T/2]$ as $y_i(t)$ leads to

$$
\begin{aligned}
y_i(t) &= \frac{1}{f_s T} \sum_{n=(t-T/2)f_s}^{(t+T/2)f_s} x_i^2(n) \\
&= \underbrace{\frac{1}{f_s T} \sum_{n=(t-T/2)f_s}^{(t+T/2)f_s} s_i^2(n)}_{\text{signal energy, } y_{si}(t)} \\
&\quad + \underbrace{\frac{1}{f_s T} \sum_{n=(t-T/2)f_s}^{(t+T/2)f_s} \nu_i^2(n)}_{\text{noise energy, } \varepsilon_i(t)}.
\end{aligned}
\tag{3}
$$

In this work, we assume that the sources have been detected within a sensor field of size 100×100 m^2. Thus, the maximum possible propagation delay between any pair of sensors is less than 0.33 s. Therefore, we can assume that the sound intensity and hence energy emitted from each source will not change significantly over such a short time delay. During the 0.33 s, the vehicle movement is less than 6 m (assuming the vehicle moves at a speed of 72 km/h). Furthermore, when we average the energy and target locations over a time period, the net effects caused by the delay propagation are even less. Therefore, it is safe to neglect the propagation delay for the energy decay function, leading to a more concise acoustic energy decay model

$$
y_i(t) = y_{si}(t) + \varepsilon_i(t) = g_i \sum_{k=1}^{K} \frac{S_k(t)}{d_{ik}^2(t)} + \varepsilon_i(t)
\tag{4}
$$

where $d_{ik}(t) = \|\boldsymbol{\rho}_k(t) - \mathbf{r}_i\|$ is the Euclidean distance between the ith sensor and the kth source. The square of the background noise $\nu_i^2(n)$ will have a χ^2 distribution with mean equal to $E\left[\nu_i^2(n)\right] = \varsigma_i^2$ and variance equal to $2\varsigma_i^4/M$. If M is sufficiently large $(M \gg 30)$, according to the central limit theorem, ε_i can be approximated well with a normal distribution, namely, $\varepsilon_i \sim N\left(\varsigma_i^2, 2\varsigma_i^4/M\right)$. For convenience, we will denote $\mu_i = \varsigma_i^2$ and $\sigma_i^2 = 2\varsigma_i^4/M$ in later derivations.

The validity of this energy attenuation model has been verified with a simple experiment. Details can be found in [22].

III. MAXIMUM LIKELIHOOD LOCATION ESTIMATION

In this section, we will introduce the ML estimation with different solutions to estimate the source location. Note that the estimation is based on a single frame of energy readings from different individual sensors. Let us now define the following

matrix notations. The time index t is omitted in the interests of brevity.

$$
\begin{aligned}
\mathbf{Z} &= \begin{bmatrix} \frac{(y_1-\mu_1)}{\sigma_1} & \cdots & \frac{(y_N-\mu_N)}{\sigma_N} \end{bmatrix}^T \\
\mathbf{G} &= \mathrm{diag}\begin{bmatrix} \frac{g_1}{\sigma_1} & \frac{g_2}{\sigma_2} & \cdots & \frac{g_N}{\sigma_N} \end{bmatrix} \\
\mathbf{D} &= \begin{bmatrix}
\frac{1}{d_{11}^2} & \frac{1}{d_{12}^2} & \cdots & \frac{1}{d_{1K}^2} \\
\frac{1}{d_{21}^2} & \frac{1}{d_{22}^2} & \cdots & \frac{1}{d_{2K}^2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{d_{N1}^2} & \frac{1}{d_{N2}^2} & \cdots & \frac{1}{d_{NK}^2}
\end{bmatrix} \\
\mathbf{S} &= \begin{bmatrix} S_1 & S_2 & \cdots & S_K \end{bmatrix}^T \\
\mathbf{H} &= \mathbf{GD} \\
\boldsymbol{\xi} &= \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_N \end{bmatrix}^T
\end{aligned}
\tag{5}
$$

where $\xi_i = (\epsilon_i - \mu_i)/\sigma_i \sim N(0,1)$ are independent Gaussian random variables. Using these notations, (4) can be represented as

$$
\mathbf{Z} = \mathbf{GDS} + \boldsymbol{\xi} = \mathbf{HS} + \boldsymbol{\xi}.
\tag{6}
$$

The joint probability density function of $\mathbf{Z}$ then can be expressed as

$$
f(\mathbf{Z}|\boldsymbol{\theta}) = (2\pi)^{-(N/2)} \exp\left\{ -\frac{1}{2} (\mathbf{Z} - \mathbf{HS})^T (\mathbf{Z} - \mathbf{HS}) \right\}
\tag{7}
$$

where

$$
\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\rho}_1^T & \boldsymbol{\rho}_2^T & \cdots & \boldsymbol{\rho}_K^T & S_1 & S_2 & \cdots & S_K \end{bmatrix}^T
$$

is the vector of unknown parameters. $\boldsymbol{\rho}_k$ is the kth source location, and S_k is the kth source energy. The negative log-likelihood function is proportional to a quadratic form

$$
\ell(\boldsymbol{\theta}) = \|\mathbf{Z} - \mathbf{GDS}\|^2.
\tag{8}
$$

Thus, the maximum likelihood parameter estimation of $\boldsymbol{\theta}$ can be obtained by minimizing $\ell(\boldsymbol{\theta})$.

Equation (8) is a *nonlinear* least square cost function because the $\mathbf{D}$ matrix that contains NK elements is a nonlinear function of the pK unknown source location coordinates $\{\boldsymbol{\rho}_k; 1 \leq k \leq K\}$, p is the dimension of the location coordinates. The $\mathbf{S}$ vector also contains K unknown parameters. Since there are $K(p+1)$ unknown parameters, there must be at least $K(p+1)$ or more sensors reporting acoustic energy readings to yield a unique solution.

To minimize $\ell(\boldsymbol{\theta})$, the solution must lie on a stationary point where

$$
\frac{\partial \ell(\boldsymbol{\theta})}{\partial S_k} = 0, \text{ and } \nabla_{\boldsymbol{\rho}_k} \ell(\boldsymbol{\theta}) = 0
$$

for $k = 1, 2, \ldots K$. These conditions lead to the following set of relations:

$$
\mathbf{S} = \mathbf{H}^\dagger \mathbf{Z}
\tag{9}
$$

$$
\mathbf{B}_k^T \mathbf{G}(\mathbf{Z} - \mathbf{HS}) = 0 \quad \{k = 1, 2, \ldots, K\}
\tag{10}
$$

where $\mathbf{H}^{\dagger}$ is the *pseudo-inverse* of the matrix $\mathbf{H}$, and

$$\mathbf{B}_k^T \overset{def}{=} \frac{\partial(\mathbf{DS})^T}{\partial \boldsymbol{\rho}_k} = -2S_k \begin{bmatrix} \frac{\mathbf{b}_{1k}}{d_{1k}^3} & \cdots & \frac{\mathbf{b}_{Nk}}{d_{Nk}^3} \end{bmatrix} \quad (11)$$

$\mathbf{b}_{ik} = \partial d_{ik}/\partial \boldsymbol{\rho}_k = (\boldsymbol{\rho}_k - \mathbf{r}_i)/d_{ik}$ is a unit vector from the kth source to the ith sensor.

Equation (10) may be expressed alternatively as

$$2S_k \sum_{i=1}^{N} \frac{g_i}{\sigma_i} \left(\frac{\boldsymbol{\rho}_k - \mathbf{r}_i}{d_{ik}^4} \right) \left(z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^{K} \frac{S_m}{d_{im}^2} \right) = 0. \quad (12)$$

Solving the above equation, $\boldsymbol{\rho}_k$ can be represented as a linearly weighted sum of all sensor locations $\mathbf{r}_i$:

$$\boldsymbol{\rho}_k = \frac{\left(\sum_{i=1}^{N} \alpha_{ik} \mathbf{r}_i \right)}{\left(\sum_{i=1}^{N} \alpha_{ik} \right)} \quad (13)$$

where

$$\alpha_{ik} = \frac{g_i}{\sigma_i} \left(\frac{1}{d_{ik}^4} \right) \left(z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^{K} \frac{S_m}{d_{im}^2} \right).$$

It must be emphasized that (13) is *not* an explicit expression for $\boldsymbol{\rho}_k$ since the distance d_{im} is a function of $\boldsymbol{\rho}_m$ on the right-hand side of this equation.

Note that (9) gives explicit expression of $\mathbf{S}$ as a function of $\mathbf{H}$ and $\mathbf{Z}$. However, (10) has an implicit relation where the unknown $\{\boldsymbol{\rho}_k\}$ appears in both sides of the equation. These two equations motivated the development of two different approaches to solving for the maximum likelihood source localization problem as described below:

A. Multiresolution Projection Solution

Substituting (9) into (8), the variables $\{S_k\}_{k=1}^{K}$ can be eliminated, giving a modified negative log-likelihood function

$$\begin{aligned} \mathbf{L}'(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_K) &= \mathbf{Z}^{\mathbf{T}} (\mathbf{I} - \mathbf{P_H})^{\mathbf{T}} (\mathbf{I} - \mathbf{P_H}) \mathbf{Z} \\ &= \mathbf{Z}^{\mathbf{T}} (\mathbf{I} - \mathbf{P_H}) \mathbf{Z} = \mathbf{Z}^{\mathbf{T}} \mathbf{Z} - \mathbf{Z}^{\mathbf{T}} \mathbf{P_H} \mathbf{Z} \end{aligned}$$

where

$$\mathbf{P}_H = \mathbf{H}(\mathbf{H^T H})^{-1} \mathbf{H^T} = \mathbf{U_H U_H^T}$$

is a projection matrix, and $\mathbf{U}_H$ is the matrix of the left singular vectors of the $\mathbf{H}$ matrix. The properties $\mathbf{P}_H = \mathbf{P}_H^T$ and $\mathbf{P}_H \cdot \mathbf{P}_H = \mathbf{P}_H$ have been used during derivation. Since $\mathbf{Z^T Z}$ are known (normalized) energy measurements, **minimizing $\mathbf{L}'$** is equivalent to **maximizing**

$$\begin{aligned} \mathbf{L}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_K) &= \mathbf{Z^T P_H Z} = \mathbf{Z^T U_H U_H^T Z} \\ &= \left\| \mathbf{U_H^T Z} \right\|^2. \quad (14) \end{aligned}$$

From (14), we know that projection solution for *ML* estimation is trying to find the target locations so that its constructed subspace of $\mathbf{H}_{op}$ has the maximum projection energy for $\mathbf{Z}$. Equation (14) is the log-likelihood function with the constraint

of (9) imposed. As such, $\mathbf{L}$ contains K fewer unknown variables than $\ell(\boldsymbol{\theta})$.

We caution the reader that although $\mathbf{L}$ is expressed as a quadratic form, maximizing it still requires the solution of a *nonlinear* least square problem as $\mathbf{U}_H$ is still a nonlinear function of the source locations $\{\boldsymbol{\rho}_k\}_{k=1}^{K}$.

1) Multiresolution (MR) Search: A straightforward method to find a solution that maximizes $\mathbf{L}$ is exhaustive search. However, the computation cost is extremely high especially when there are multiple sources. For example, let there be K sources and q grid points to be searched in each dimension. Then the total number of search points with a p-dimensional sensor field will be equal to q^{pK}. While the computation complexity may be feasible for a desktop computer, it is likely to be excessive for low power sensor nodes with limited computing capabilities.

This exponentially growing computation complexity can be mitigated with the use of MR search method. Among several choices, a logarithmic MR search strategy will examine only w points in each dimension per iteration, where $q = w^m$, with m being a positive integer. Hence, in each iteration, only w^{pK} grid points needs to be searched. Then, another iteration of search will be confined in the neighborhood of the current best solution by subdividing the coarser mesh around the current solution into w subdivisions and performing search. After m iterations, the MR method will search at a grid size equal to that of exhaustive search. However, the total search points will be $m \cdot w^{pK}$ rather than $q^{pK} = w^{mpK}$. To appreciate the amount of saving, let $q = 128 = 2^7 = w^m$, and $p = 2$. Using exhaustive search, the total search points will be 2^{14K}. Using MR search, the total number of search points will be $7 \cdot 2^{2K}$. For $K = 1$, the difference is 4096 points versus 28 points. For $K = 2$, the difference is 268 435 456 points versus 112 points. Obviously, due to the coarser search grid at earlier iterations, the MR method may be trapped in a local minimum and yields an inferior solution.

B. E–M Iterative Solution

Equations (9) and (13) can be used together to yield an iterative solution similar to the E–M algorithm. With this solution, we assume the source intensity vector $\mathbf{S}$ is the *missed* data rather than unknown parameter. We initiate the K unknown source location ρ_k at the beginning. During the iteration, we estimate the *missed* data according to (9), and then, we maximize the log-likelihood function using (13) to get the updated estimate ρ_k. The iteration keeps on going until convergence. This EM algorithm has much less computation complexity compared to the projection solution. However, it might trap into a local minimum, depending on the initialization of parameter estimates. Hence, it may be applied to refine the MR search results.

C. Comparison with Other Energy-Based Source Localization Methods

In addition to the maximum likelihood method presented in this paper, there are other energy based acoustic source localization methods. We now summarize these algorithms and explore their relationship to the proposed ML method.

1) Closest Point Approach (CPA) Methods: The closest point approach (CPA) is a navigation term that describes the closest position of two objects moving along nonintersecting straight lines. Here, we borrow the term to refer to a nearest neighbor localization method: *Identify a sensor with largest acoustic energy measurement:*

$$i^* = \arg \max_i y_i.$$

Assign the source location to be the sensor location

$$\boldsymbol{\rho} = \mathbf{r}_{i*}. \tag{15}$$

In general, the CPA method searches the sensor with the maximum energy reading for single source situation. When there are multiple sources, the algorithm must first identify all local maxima among all sensor acoustic energy readings. As a matter of fact, the CPA algorithm can be deduced from (13) by setting $K = 1$. Specifically, let $d_i = d_{ik}$ when $K = 1$; then

$$\alpha_i = \frac{g_i}{\sigma_i} \left(\frac{1}{d_i^4} \right) \left(z_i - \frac{g_i}{\sigma_i} \frac{S}{d_i^2} \right) = \frac{g_i}{\sigma_i} \left(\frac{\xi_i}{d_i^4} \right).$$

Let $d_{i*} \ll d_j, \forall j \neq i^*$. If g_i, σ_i, and ξ_i have the same order for all sensors i, the ML solution (13) becomes the CPA solution (15). Another implementation detail is that the actual source location cannot overlap with the sensor location. Hence, it is often chosen as a location that is nearest to the sensor with largest energy reading.

D. Energy Ratios Source Localization, Nonlinear Least Square (ER-NLS), and Least Square (ER-LS) Formulations

When there is only a single source $(K = 1)$ within the sensor field, the $\mathbf{H}$ matrix become a vector

$$\mathbf{H} = \left[\frac{g_1}{\sigma_1 d_1^2}, \frac{g_2}{\sigma_2 d_2^2}, \ldots, \frac{g_N}{\sigma_N d_N^2} \right]^T.$$

As such, $\mathbf{U_H} = \mathbf{H}/\|\mathbf{H}\|$ will be a unit vector. If each entry of the $\mathbf{U}_H$ vector *were* an independent variable, then an obviously solution to maximize the modified cost function $\mathbf{L}$ in (14) would be $\mathbf{U_H} = \mathbf{c} \cdot \mathbf{Z}$, where c is a proportional constant. Equivalently, for $i = 1, \ldots, N$

$$\frac{y_i - \mu_i}{\sigma_i} = c \cdot \frac{g_i}{\sigma_i d_i^2} = c \cdot \frac{g_i}{\sigma_i \|\boldsymbol{\rho} - \mathbf{r}_i\|^2} \tag{16}$$

where $\boldsymbol{\rho}$ denotes the source location. Note that although there are N equalities, there are actually $p+1 < N$ unknowns, including the proportional constant c.

One way to solve this set of nonlinear equalities is to first eliminate the unknown constant c by computing the energy ratio φ_{ij} of the ith and the jth sensors as follows:

$$\varphi_{ij} = \left[\frac{\frac{(y_i - \mu_i)}{(y_j - \mu_j)}}{\frac{g_i}{g_j}} \right]^{-1/2} = \frac{\|\boldsymbol{\rho} - \mathbf{r}_i\|}{\|\boldsymbol{\rho} - \mathbf{r}_j\|}. \tag{17}$$

By sorting the calibrated energy readings $(y_i - \mu_i)/g_i$, for $0 < \varphi_{ij} \neq 1$, all the possible source coordinates $\boldsymbol{\rho}$ that satisfy

(17) reside on a p-dimensional hyper-sphere described by the equation:

$$\|\boldsymbol{\rho} - \mathbf{c}_{ij}\|^2 = \zeta_{ij}^2 \tag{18}$$

where the center $\mathbf{c}_{ij}$ and the radius ζ_{ij} of this hyper-sphere associated with sensor i and j are given by

$$\mathbf{c}_{ij} = \frac{\mathbf{r}_i - \varphi_{ij}^2 \mathbf{r}_j}{1 - \varphi_{ij}^2}, \quad \zeta_{ij} = \frac{\varphi_{ij} \|\mathbf{r}_i - \mathbf{r}_j\|^2}{1 - \varphi_{ij}^2}. \tag{19}$$

If $\varphi_{ij} = 1$, the solution of (17) forms a hyperplane between $\mathbf{r}_i$ and $\mathbf{r}_j$, i.e.,

$$\boldsymbol{\rho}(t)\boldsymbol{\iota}_{ij} = \tau_{ij} \tag{20}$$

where $\boldsymbol{\iota}_{ij} = \mathbf{r}_i - \mathbf{r}_j, \tau_{ij} = (|\mathbf{r}_i|^2 - |\mathbf{r}_j|^2)/2$. With all the energy ratios computed, the target location can be solved by minimizing a nonlinear least square cost function

$$J(\boldsymbol{\rho}) = \sum_{l_1=1}^{L_1} (|\boldsymbol{\rho} - \mathbf{c}_{l_1}| - \zeta_{l_1})^2 + \sum_{l=1_2}^{L_2} \left(\boldsymbol{\iota}_{l_2}^T \boldsymbol{\rho} - \tau_{l_2} \right)^2 \tag{21}$$

where L_1 and L_2 are the number of hyper-pheres and the number of hyperplanes, respectively. We call it the energy-ratio nonlinear least square (ER-NLS) method.

Since every pair of hyper-spheres (with double indices ij replaced by a single index m for the brevity of notations) $\|\boldsymbol{\rho} - \mathbf{c}_m\|^2 = \zeta_m^2$ and $\|\boldsymbol{\rho} - \mathbf{c}_n\|^2 = \zeta_n^2$, a hyper plane can be determined by eliminating the common terms:

$$(\mathbf{c}_m - \mathbf{c}_n)^T \boldsymbol{\rho} = \frac{\left((\|\mathbf{c}_m\|^2 - \|\mathbf{c}_n\|^2) - \left(\zeta_m^2 - \zeta_n^2 \right) \right)}{2}. \tag{22}$$

Combining (22) with (20), the source location can be solved using a least square solution without lengthy nonlinear optimization search. We call it the energy-ratio least square (ER-LS) method.

The ER-NLS and ER-LS methods have been reported previously in [22]. Here, we show the relationship between ML estimation and these two methods. We show that these two methods are the approximate solutions for ML estimation for single target source localization.

IV. CRAMÉR–RAO BOUNDS (CRB)

A. CRB Derivation

CRB is a theoretical lower bound of the variance of an unbiased parameter estimate. It is defined as the inverse of the Fisher information matrix:

$$\mathbf{J} = - E \left[\frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{Z}|\boldsymbol{\theta}) \right)^T \right]$$

$$= E \left\{ \left[\frac{\partial \ln f(\mathbf{Z}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \left[\frac{\partial \ln f(\mathbf{Z}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T \right\}. \tag{23}$$

Substituting (7) into (23), one has

$$\mathbf{J} = \frac{\partial (\mathbf{DS})^T}{\partial \boldsymbol{\theta}} \mathbf{G^T G} \frac{\partial (\mathbf{DS})}{\partial \boldsymbol{\theta}^T} \tag{24}$$

where $\partial(\mathbf{DS})^T/\partial\boldsymbol{\theta} = [\mathbf{B} \quad \mathbf{D}]$ since $\partial(\mathbf{DS})^T/\partial\mathbf{S} = \mathbf{D}$, and $\partial(\mathbf{DS})^T/\partial\boldsymbol{\rho}_j = \mathbf{B}_j$, as defined in (11). Defining $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_K]$, one has

$$\mathbf{J} = \begin{bmatrix} \mathbf{B^T} \\ \mathbf{D^T} \end{bmatrix} \mathbf{G^T G} \, [\mathbf{BD}]. \tag{25}$$

The lower bound of the variance of the source location estimates $\mathrm{var}\left(\widehat{\rho_{ij}}\right)$ can be expressed as

$$\mathrm{var}\left(\widehat{\rho_{ij}}\right) \geq (J^{-1})_{(i-1)p+j,(i-1)p+j}.$$

$1 \leq i \leq K$, and $1 \leq j \leq p$.

B. Impact of Sensor Placement to the Estimation Accuracy Using CRB Analysis for Single-Source Case

For convenience, we will focus on the single source $(K = 1)$ and two-dimensional (2-D) sensor field $(p = 2)$ situation. Furthermore, let us assume noise is prewhitened and that all sensor gains are the same. Then, the Fisher information matrix becomes

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & J_{22} \end{bmatrix} \tag{26}$$

where

$$\mathbf{J}_{11} = \frac{4S^2 g^2}{\sigma^2} \sum_{i=1}^{N} \frac{\mathbf{b}_i \mathbf{b}_i^T}{d_i^6}$$

$$\mathbf{J}_{21}^T = \mathbf{J}_{12} = -2S \frac{g^2}{\sigma^2} \sum_{i=1}^{N} \frac{\mathbf{b}_i}{d_i^5}$$

$$J_{22} = \frac{g^2}{\sigma^2} \sum_{i=1}^{N} \frac{1}{d_i^4}.$$

Using block inverse matrix theory, we get

$$\mathbf{J}^{-1} = \begin{bmatrix} \mathbf{E}^{-1} & L^{-1}\mathbf{F} \\ L^{-1}\mathbf{F}^T & L^{-1} \end{bmatrix} \tag{27}$$

where

$$\mathbf{E} = \mathbf{J}_{11} - \frac{1}{J_{22}}\mathbf{J}_{12}\mathbf{J}_{12}^T$$
$$\mathbf{F} = -\mathbf{J}_{11}^{-1}\mathbf{J}_{12}$$
$$L = J_{22} - \mathbf{J}_{12}^T\mathbf{J}_{11}^{-1}\mathbf{J}_{12}.$$

$\mathbf{E}$ can be factored as

$$\mathbf{E} = \kappa \begin{bmatrix} \mathbf{L_X^T} \\ \mathbf{L_X^T} \end{bmatrix} [\mathbf{L_X} \quad \mathbf{L_Y}]$$
$$= \kappa \begin{bmatrix} |\mathbf{L_X}|^2 & |\mathbf{L_X}| \cdot |\mathbf{L_Y}| \cos\beta \\ |\mathbf{L_X}| \cdot |\mathbf{L_Y}|\cos\beta & |\mathbf{L_Y}|^2 \end{bmatrix} \tag{28}$$

where $\kappa = 4S^2 g^2/\sigma^2$, $\mathbf{L_X}$, and $\mathbf{L_Y}$ are both $N \times 1$ vectors, and $\beta = \cos^{-1}\left(\left(\mathbf{L_X^T}(\mathbf{L_Y})\right)/(|\mathbf{L_X}||\mathbf{L_Y}|)\right)$. The jth column of the $2 \times N$ matrix $[\mathbf{L_X} \quad \mathbf{L_Y}]^T$ can be expressed as

$$[\mathbf{L_X} \quad \mathbf{L_Y}]_j^T = \frac{\mathbf{b}_j}{d_j^3} - \frac{\frac{1}{d_j^2}}{\sum_{i=1}^{N}\left(\frac{1}{d_i^2}\right)^2} \sum_{i=1}^{N}\left(\frac{1}{d_i^2}\right)\left(\frac{\mathbf{b}_i}{d_i^3}\right)$$

$$= \frac{\mathbf{b}_j}{d_j^3} - \frac{\overline{\mathbf{b}}}{d_j^2} \tag{29}$$

where $\overline{\mathbf{b}}$ is a weighted average of the individual (weighted) source-to-sensor direction vector.

The lower bound of the variance of the ML source location estimates can be found in terms of the CRB as

$$\mathrm{Var}(\hat{\rho}_x) \geq \{\mathbf{E}^{-1}\}_{11} = \frac{1}{\kappa} \frac{1}{|\mathbf{L_X}|^2(1 - \cos^2\beta)} \tag{30}$$

$$\mathrm{Var}(\hat{\rho}_y) \geq \{\mathbf{E}^{-1}\}_{22} = \frac{1}{\kappa} \frac{1}{|\mathbf{L_Y}|^2(1 - \cos^2\beta)}. \tag{31}$$

Due to the weight factors of $1/d_i^3$, sensors that are close to the source will have a much higher impact on the CRB than those far away. Specifically, the fact that the CRBs are the diagonal elements of $\mathbf{E}^{-1}$ implies that the $\mathbf{E}$ matrix must be strongly nonsingular. That in turn implies that at least two of the sensors must be close enough to the source so that the $[\mathbf{L_X} \quad \mathbf{L_Y}]^T$ matrix will have full column rank $(= p)$.

To reduce the CRB and hence to improve the accuracy of the source localization results, three approaches may be taken: i) to keep κ large, ii) to increase both $|\mathbf{L_X}|^2$ and $|\mathbf{L_Y}|^2$, and iii) to let $\beta \to 90°$. Since source energy emission level S and background noise level σ cannot be controlled, to increase κ, the sensor gain g needs to be increased. To increase the magnitude of both $|\mathbf{L_X}|^2$ and $|\mathbf{L_Y}|^2$, the set of sensor target distances $\{d_i; 1 \leq i \leq N\}$ must be reduced. This implies that there must be sufficient number of sensors close to the acoustic source whenever the source is within the sensor field. Since the target may be moving within the sensor field, this requirement translates to the requirement of dense sensor deployment within the sensor field. As for iii), to make $\beta \to 90°$, it would require the sensors within the sensor field be laid out in a uniform formation. Nonetheless, this requirement is not suitable for an *ad hoc* sensor network application where sensor deployment is assumed to have no specific formation or structure.

V. SIMULATION AND EXPERIMENTATION

A. Performance Comparison Using Simulation

Extensive simulation runs have been conducted to compare the performance of the proposed ML energy based source localization algorithm to other energy based source localization algorithms. We use (4) to generate the acoustic energy readings of a 2-D $(p = 2)$ sensor field of size 100 by 100 m. Since existing localization algorithms are all for single target situations, we set $K = 1$. The source location and the sensor locations are randomly chosen from within the sensor field in each run. The source energy is set at $S = 5000$, and the background noise level is set at $\sigma_i = 1$ for all sensors in the sensor field. Note that although the SNR is 37 dB at the source location, the actual SNR at different sensors depends very much on the sensor to source distance. For example, for a sensor that is 50 m away from the source, its SNR is merely $10 \times \log_{10}(5000/50^2) = 3$ dB. The energy variation $\epsilon_i(t)$ is modeled as a Gaussian random variable $N\left(\sigma_i^2, \sigma_i^4/M\right)$ with $M = 100$.

We conducted 2000 repeated trials. In each trial, all four energy-based acoustic source localization methods (ML, CPA, ER-NLS, and ER-LS) were used to locate the source location, and the error was recorded. For the CPA method, the target location is estimated as the location of the sensor that has

TABLE I
MEAN AND COVARIANCE MATRICES OF LOCATION ESTIMATION ERROR

|	4 sensor	10 sensors	25 sensors
CPA	$\begin{bmatrix} .26 & .76 \end{bmatrix}$ $\begin{bmatrix} 513 & -3 \\ -3 & 521 \end{bmatrix}$	$\begin{bmatrix} .27 & -.14 \end{bmatrix}$ $\begin{bmatrix} 215 & 7 \\ 7 & 191 \end{bmatrix}$	$\begin{bmatrix} .27 & -.14 \end{bmatrix}$ $\begin{bmatrix} 78 & -1 \\ -1 & 77 \end{bmatrix}$
ML	$\begin{bmatrix} .19 & -.39 \end{bmatrix}$ $\begin{bmatrix} 99 & -8 \\ -8 & 119 \end{bmatrix}$	$\begin{bmatrix} -.09 & .03 \end{bmatrix}$ $\begin{bmatrix} 15 & 0 \\ 0 & 15 \end{bmatrix}$	$\begin{bmatrix} .00 & .08 \end{bmatrix}$ $\begin{bmatrix} 9 & 0 \\ 0 & 8 \end{bmatrix}$
ER-NLS	$\begin{bmatrix} .58 & .28 \end{bmatrix}$ $\begin{bmatrix} 295 & -17 \\ -17 & 333 \end{bmatrix}$	$\begin{bmatrix} -.24 & -.55 \end{bmatrix}$ $\begin{bmatrix} 134 & -2 \\ -2 & 130 \end{bmatrix}$	$\begin{bmatrix} .15 & -.14 \end{bmatrix}$ $\begin{bmatrix} 111 & -3 \\ -3 & 112 \end{bmatrix}$
ER-LS	$\begin{bmatrix} 1.08 & -.02 \end{bmatrix}$ $\begin{bmatrix} 615 & -16 \\ -16 & 601 \end{bmatrix}$	$\begin{bmatrix} .01 & -.28 \end{bmatrix}$ $\begin{bmatrix} 98 & 2 \\ 2 & 89 \end{bmatrix}$	$\begin{bmatrix} .08 & .15 \end{bmatrix}$ $\begin{bmatrix} 87 & -3 \\ -3 & 105 \end{bmatrix}$

Fig. 1. Distribution of the magnitude of errors of the four localization algorithms for single target localization.

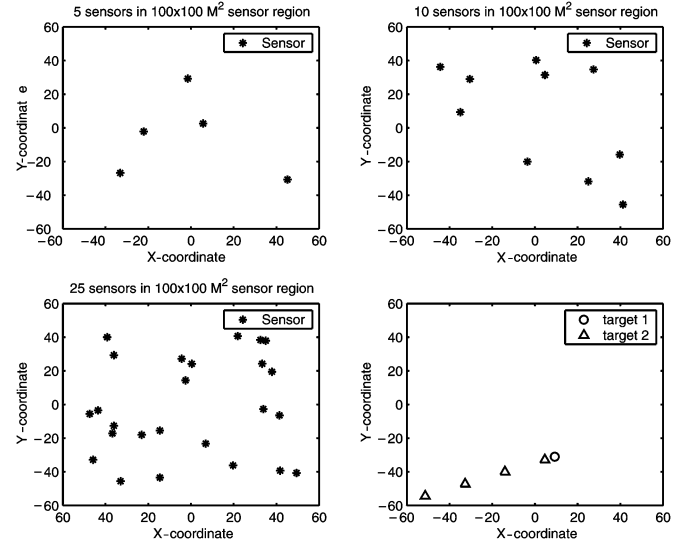Fig. 2. Sensor deployment for the two-target simulations and four different source separations.

maximum energy reading plus a Gaussian random variable with zero mean and unit variance. For the ML and ER-NLS methods, exhaustive search is used at a resolution of 5 by 5 m^2 per search grid. Three different sensor densities have been used: 4, 10, and 25. In the last case, there is approximately one sensor in every 20 by 20 m cells in the sensor field.

The mean and covariance matrices of location estimation error are listed in Table I.

From simulation results, we observe that the mean values of these methods do not show any statistically significant bias and, hence, yield unbiased estimates. Furthermore, the estimation error in different dimensions are also uncorrelated. In addition, we have the following.

- The ML method consistently outperforms all other three existing methods in terms of estimation variances and mean estimation error.
- The CPA method benefits most when the number of sensors increases.
- With both the ten-or 25-sensor cases, the least square formulation (ER-LS) outperforms the computation-wise more expensive nonlinear least square formulation (ER-NLS).

Another way to analyze the simulation results is to examine the distribution of the *magnitude* of location estimation error. The results are summarized in Fig. 1. In this figure, each row represents results obtained from a particular method. Each column represents results from a particular sensor density. The histograms of the magnitudes of the localization error are plotted using a bin with a 5-m increment. Since the histogram can be regarded as an approximation of the probability density function, the mean and standard deviation of the *magnitude* of the localization error are calculated and listed in each figure.

B. Performance Analysis: Two Sources Case

We have also performed simulations to examine the performance of the proposed algorithm when there are two targets present in the sensor field. Again, we assume a sensor field of 100 by 100 m^2. Fig. 2 shows the sensor field of the two target simulations. Four different source separations are used: 65, 45, 25, and 5 m. The background noise parameters are $\mu = 2$ and $\sigma = 2^4/100$. Both source intensities at 1 m distance are defined as 10 000 units. Three sensor network configurations are used: 5, 10, and 25 sensors. For each source location and each sensor configuration, 1000 simulations are performed. The mean and
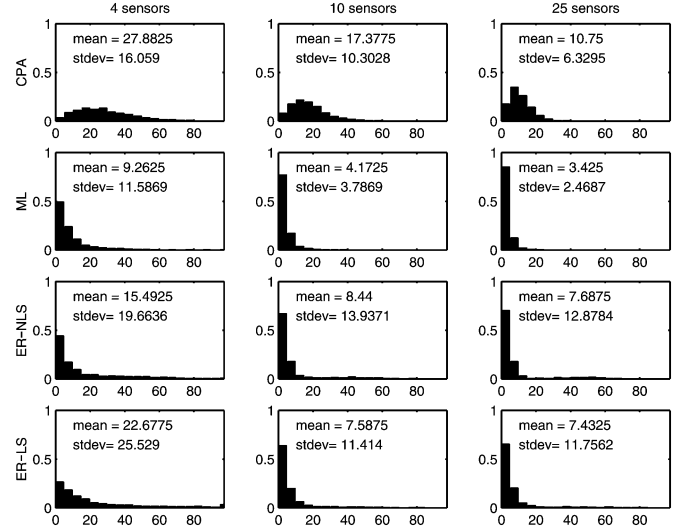
variance are computed. Since there are two targets, the localized results will return two positions. We assume the target association is done properly so that the location estimation error is the minimum of the two possible assignments. The histograms of the location estimation error of each source are listed in Figs. 3 and 4.

From these figures, it is noted that **when two targets are very close to each other**, the energy profile of the sensors has a single peak around both targets' locations. *If the sensors are placed densely enough*, the ML algorithm gives smaller localization error, despite the fact that it is unable to resolve these two targets. However, *if the sensors are deployed sparsely*, ML algorithm gives larger localization error because only one single sensor gets significant energy reading in the sparsely deployed sensor field when targets are very close to each other. From the CRB analysis described in the previous section, we know that it requires at least two significant energy readings for better source localization accuracy. When the two targets
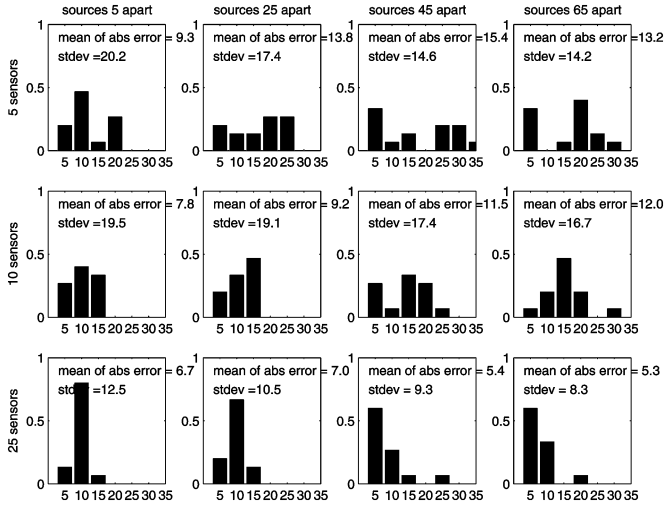
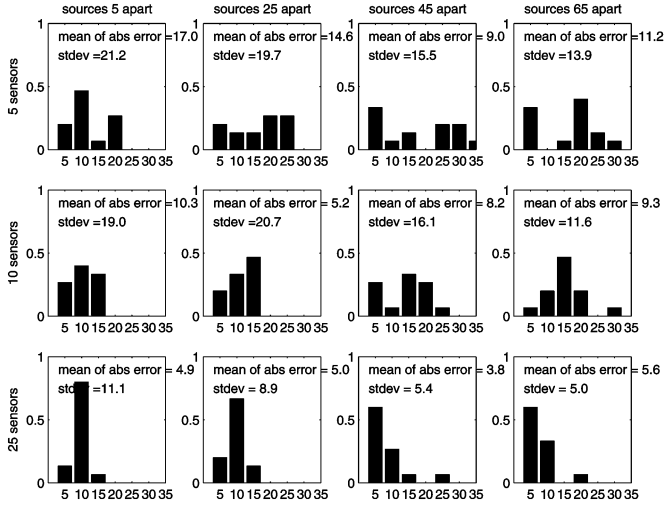Fig. 3. Localization error of first target of the two targets source localization.

Fig. 4. Localization error of second target of the two targets source localization.

are well separated, two or more sensors might have significant acoustic energy reading.

When the two targets are far apart, the estimation error is larger. This may be due to the fact that the cost function is sensitive to parameter perturbations. Further parameter sensitivity analysis is currently underway and will be reported in the near future. In addition, when targets are far apart, the sensors near one target are far from the other target. Therefore, the relative SNR for the target located far from the sensors is very small. The sensors located between the targets might be far from both of the targets, and therefore, the relative SNR is small for both targets.

Multitarget simulations using ML algorithm under different noise energy levels are also performed. They present the similar phenomena, as shown in Figs. 3 and 4 with the overall performance increasing as noise power decreasing. It has also been found that two targets can be treated as a single target when they are very close to each other without the degression of performance, and they can be treated as two single targets when they are far apart with better performance and reduced computation burden since the 4-D search is reduced to two 2-D searches.
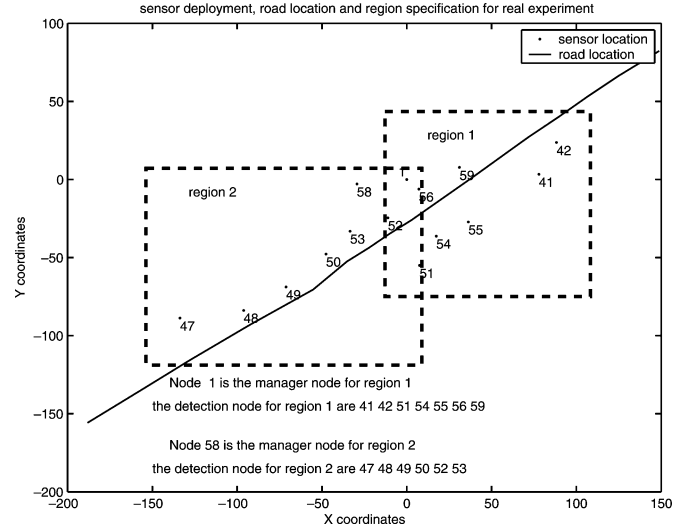

Fig. 5. Sensor deployment, road coordinate, and region specification for experiments.

Details of multitarget source localization will be reported by our follow-up paper.

C. Experiments: Application to Moving Vehicle Localization

An application of the proposed energy based source localization (EBSL) methods to locate moving vehicle using distributed microphone sensor nodes will be reported in this section. First, we describe the overall system. Then, the experiment results will be presented. Finally, the systematic bias analysis is given.

1) Sensor Network System: In November 2001, a field experiment sponsored by the DARPA ITO SensIT project has been carried out at Twentynine Palms, CA. Custom-made prototype sensor nodes were laid out along side a road. Each sensor node is equipped with acoustic, seismic, and polarized infrared sensors, a 16-bit microprocessor, and a radio transceiver and modem. It is powered by an external car battery. During the experiment, military vehicles such as an amphibious assault vehicle (AAV) and a dragon wagon (DW) were driven through the road, and sensors sampled the corresponding multimodal data. The acoustic signal is sampled at 4960 Hz at 16-bit resolution. The set of data segments reported below are taken from the acoustic signatures of a single AAV travelling from east to west along the road during a time period of approximate 2 min. The sensor field is partitioned into two regions with some details given in Fig. 5.

2) Acoustic Energy-Based Localization Experiment: The energy reading collected from all sensor nodes within the region at the same 0.75-s time interval were used for acoustic source localization at the manager node. Fig. 6 shows the AAV ground truth and the localization results based on the ML algorithm with MR projection solution and ER-NLS algorithm. The ground truth is obtained by interpolating an onboard GPS recording that recorded a position fix every 15 s. To use multiresolution search, three grid sizes of 4, 2, and 1 were applied. The localization results are summarized in Figs. 6 and 7. From Fig. 7, it is clear that the ML method outperforms the ER-NLS method.

3) Systematic Bias Analysis: The data sampled during this experiment is very noisy. During the experiments, a gusty wind
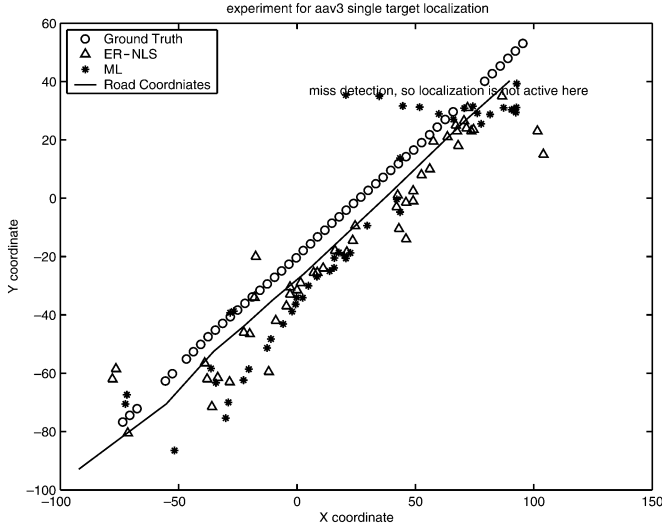
Fig. 6.   AAV ground truth and localization estimation results based on ML algorithm with projection solution and NLS algorithm (MR search is used, grid size is $4*4$, $2*2$, and $1*1$. Estimation results look biased from the ground-truth; refer to "systematic bias analysis" for more details).
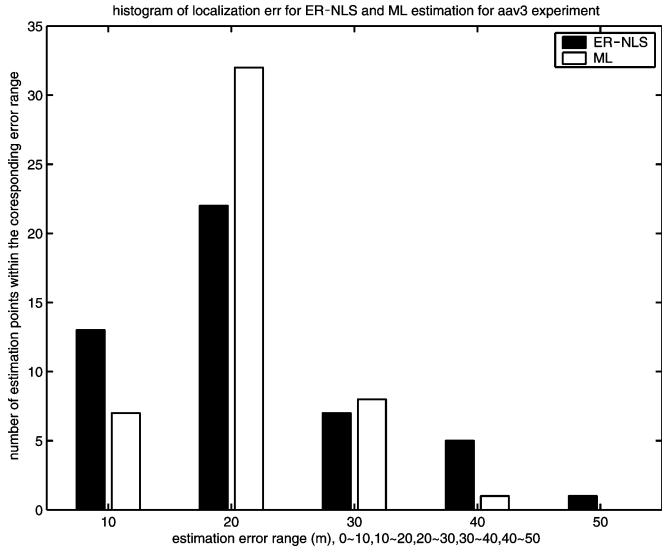

Fig. 7.   Estimation error histogram for AAV experiment data.

often blew directly into the microphone, creating isolated energy spikes in some of the sensors. Many microphones were not properly calibrated, and the gain factors g_i were estimated from the time series manually and can be grossly inaccurate. The ground truth is not necessarily correct either due to the lack of differential GPS settings or due to the lack of long term averaging at each position fix. It is evident from Fig. 6 that the GPS markings of the ground truth are consistently off the road. All these factors may cause the estimation bias. Yet, it does not affect the feasibility of this algorithm. In fact, it is found that some location estimations are even closer to the road than the GPS markings of the ground truth. Hence, the superior performance of the proposed ML source localization method in this experiment clearly demonstrates the feasibility of applying this method to handle real-world data.

To analyze the details of the effects of the systematic bias on the performance of the ML estimation and other existing

methods, we need to do more sensitivity analysis on these algorithms, which is under development and will be presented in the follow-up paper. According to our initial sensitivity analysis, the performance of both ML estimation and ER-NLS, ER-LS will not degrade significantly when the sensor locations from GPS measurement are biased. Obviously, if the bias of the sensor location is not severe, it also will not affect CPA estimation significantly. On the other hand, the proposed algorithms are sensitive to the incorrect gain factors.

VI. DISCUSSION

A. Number of Sources

In this proposed method, it is assumed that the number of acoustic sources is known in advance before the localization algorithm is applied. Indeed, the proposed method can be extended to the situation of unknown number of the targets, using a classical generalized likelihood ratio test (GLRT) or weighted subspace fitting method [28], [29]. Yet, these two methods will greatly increase the computational burden, and therefore, they are not suitable for sensor network application, whereas the power supply is limited. Instead, in sensor network applications, we use other methods to estimate the number of targets. If the sources are well-separated and sensors are densely deployed over the sensor field, the number of sources can be determined by finding the number of peaks of the energy profile. When several sources are closely positioned, the proposed method may be unable to resolve the individual positions of each source, even if the number of sources is correctly estimated.

On the other hand, using different sensing modalities, such as passive infrared sensors, it may be possible to separate sources that are closely positioned. However, it is often required that the sources travel on a predefined road, and sensors are placed along side of that road.

A number of the target in the sensor field can also be counted by the tracking algorithm. To save computation energy and communication bandwidth, source detection, localization, and tracking algorithms work cooperatively to detect, localize, and predict the targets in the sensor field [30]. Localization algorithm is performed only when the region detection announces the target presence. Localization results are fed into the tracking algorithm to predict the target locations at the next time slot. Therefore, using the tracking algorithm, we can count the number of the targets in the sensor field.

Incidentally, in [22], the localization algorithm is for a single target case only.

B. Accuracy of the Location Estimates

In the application considered in this work, a moving military vehicle is to be located within a sensor field. The sizes of the military vehicles vary from 2 to 10 m. The vehicle may be moving at a speed up to 20 ms/s (45 mi/h). Over the time interval of 0.75 s when a single energy reading is calculated, the target's location may have an uncertainty range of up to 25 m. Hence, 5 to 20 m location estimation error is quite reasonable. In fact, the estimation error shown in Fig. 6 is calculated by comparing the estimations with the GPS measurements of the ground truth. Yet, as shown in Fig. 6, GPS measurements are biased off the

road. Therefore, the real estimation errors are less than 15 m. If some imperfect factors of the experiments can be calibrated by the preprocessing, the estimation error can be even less.

VII. CONCLUSION

In this paper, a novel ML source localization method is presented. This method promises superior performance and multiple source localization and is easy to implement. In addition to the algorithm derivation, the CRB of this algorithm has also been reported. Extensive simulations show that the proposed algorithm consistently outperforms other existing energy-based localization methods.

Future works include parameter sensitivity analysis and sequential Bayesian estimation.

REFERENCES

[1] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the physical world with pervasive networks," *IEEE Pervasive Comput.*, vol. 1, no. 1, pp. 59–69, 2002.

[2] C. Savarese, J. M. Rabaey, and J. Reutel, "Localization in distributed Ad-hoc wireless sensor networks," in *Proc. ICASSP*, Salt Lake City, UT, 2001, pp. 2037–2040.

[3] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, pp. 17–29, 2002.

[4] A. Tolstory, *Matched-Field Processing for Underwater Acoustics*, Singapore: World Scientific, 1993.

[5] M. Brandstein and H. Silverman, "A localization-error-based method for microphone-array design," in *Proc. ICASSP*, Atlanta, GA, 1996, pp. 901–904.

[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 Audio Process.*, vol. 5, no. 1, pp. 45–50, Jan. 1997.

[7] J. Huang, N. Ohnishi, and N. Sugie, "Sound localization in reverberant enviornment based on model of the precedence effect," *IEEE. Trans. Instrum. Meas.*, vol. 46, pp. 842–846, Aug. 1997.

[8] M. Omologo and P. Svaizer, "Acoustic source location in noisy and reverberant environment using CSP analysis," in *Proc. ICASSP*, Atlanta, GA, 1996, pp. 921–924.

[9] H. Wang and P. Chu, "Voice source localization for automatic camera pointing system in videoconferencing," in *Proc. ICASSP*, vol. 1, 1997, pp. 187–190.

[10] K. Suyama, K. Takahashi, and R. Hirabayashi, "A robust technique for sound source localization in consideration of room capacity," in *Proc. IEEE Workshop Appl. Signal Processing Audio Acoust.*, New Paltz, NY, 2001, pp. 63–66.

[11] J. Weng and K. Y. Guentchev, "Three-dimensional sound localization form a compact noncoplanar array of microphones using three-based learning," *J. Acoust. Soc. Amer.*, vol. 110, pp. 311–323, 2001.

[12] S. M. Griebel and M. S. Brandstein, "Microphone array source localization using realizable delay vectors," in *Proc. IEEE Workshop Appl. Signal Processing Audio Acoust.*, New Paltz, NY, 2001, pp. 71–74.

[13] K. Yao, R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1555–1567, Oct. 1998.

[14] C. W. Reed, R. Hudson, and K. Yao, "Direct joint source localization and propagation speed estimation," in *Proc. ICASSP*, Phoenix, AZ, 1999, pp. 1169–1172.

[15] S. Haykin, *Array Signal Processing*. Englewood-Cliffs, NJ: Prentice-Hall, 1985.

[16] L. G. Taff, "Target localization from bearings-only observations," *IEEE Trans. Aerosp. Electron.*, vol. 3, no. 1, pp. 2–10, Jan. 1997.

[17] Y. Oshman and P. Davidson, "Optimization of observer trajectories for bearings-only target localization," *IEEE Trans. Aerosp. Electron.*, vol. 35, no. 3, pp. 892–902, Jul. 1999.

[18] K. M. Kaplan, Q. Le, and P. Molnar, "Maximum likelihood methods for bearings-only target localization," in *Proc. IEEE ICASSP*, vol. 5, 2001, pp. 3001–3004.

[19] G. C. Carter, *Coherence and Time Delay Estimation*. New York: IEEE Press, 1993.

[20] "Special issue on time-delay estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, 1981.

[21] J. O. Smith and J. S. Abel, "Closed form least square source location estimation from range difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 1661–1669, Dec. 1987.

[22] D. Li and Y. H. Hu, "Energy based collaborative source localization using acoustic micro-sensor array," *J. EUROSIP Applied Signal Process.*, vol. 4, pp. 321–337, 2003.

[23] L. E. Kinsler *et al.*, *Fundamentals of Acoustics*. New York: Wiley, 1982.

[24] K. Srodecki, "Evaluation of the reverberation decay quality in rooms using the autocorrelation function and the cepstrum analysis," *Acustica*, vol. 80, pp. 216–25, 1994.

[25] Y. L. Li, M. J. White, and S. J. Franke, "New fast field programs for anisotropic sound propagation through an atmosphere with a wind velocity profile," *J. Acoust. Soc. Amer.*, vol. 95, pp. 718–26, 1994.

[26] E. M. Salomons, "Downwind propagation of sound in an atmosphere with a realistic sound-speed profile: A semianalytical ray model," *J. Acoust. Soc. Amer.*, vol. 95, pp. 2425–36, 1994.

[27] T. Watanabe and S. Yamada, "Sound attenuation through absorption by vegetation," *J. Acoust. Soc. Japan*, vol. 17, pp. 175–82, 1996.

[28] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample ML techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, Haykin, Litva, and Shepherd, Eds. Berlin, Germany: Spriger-Verlag, 1993, pp. 99–151.

[29] M. Viberg and A. L. Swindlehurst, "A Bayesian approach to auto calibration for parametric array signal processing," *IEEE Trans. Signal Process.*, vol. 42, pp. 3495–3507, Dec. 1994.

[30] X. Sheng and Y. H. Hu, "Energy based source localization," *Inform. Processing Sensor Networks*, vol. IPSN2003, pp. 285–300, 2003.

Xiaohong Sheng received the B.Eng. degree from Zhejiang University, Hangzhou, China, and the M.Eng. degree from Nanyang Technological University, Singapore, in 1998. She received the M.S.E.E. degree from the Department of electrical and computer engineering, University of Wisconsin-Madison in 2001, where she is currently pursuing the Ph.D. degree.

From 1998 to 1999, she was with the Audio group at R&D, Creative Technology Ltd., Singapore. Her research interests include collaborative signal processing in wireless sensor networks, multimedia signal processing, and statistical signal processing in communications.

Yu-Hen Hu (M'83–SM'88–F'99) received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1976 and the M.S. and Ph.D. degrees, both in electrical engineering, from University of Southern California, Los Angeles, in 1980 and 1982 respectively.

Currently, he is a Professor with the Electrical and Computer Engineering Department, University of Wisconsin–Madison. Previously, he was with the Electrical Engineering Department, Southern Methodist University, Dallas, TX. His research interests include multimedia signal processing, design methodology and implementation of signal processing algorithms and systems, and neural network signal processing. He has published more than 200 journal and conference papers and edited two books: *Programmable Digital Signal Processors* and *Handbook of Neural Network Signal Processing*.

Dr. Hu served as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, *the Journal of VLSI Signal Processing*, and *the European Journal of Applied Signal Processing*. He served as Secretary of the IEEE Signal Processing Society, the Board of Governors of the IEEE Neural Networks Council, chair of IEEE Signal Processing Society, and the IEEE Neural Network Signal Processing Technical Committee.