

WHOI-89-32

Optimal Source Localization and Tracking Using Arrays with Uncertainties in Sensor Locations

by

Mordechai Segal and Ehud Weinstein

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

August 1989

Technical Report



Funding was provided by the Office of Naval Research
under contract Number N00014-85-K-0272.

Reproduction in whole or in part is permitted for any purpose of the
United States Government. This report should be cited as:
Woods Hole Oceanog. Inst. Tech. Rept., WHOI-89-32.

Approved for publication; distribution unlimited.

Approved for Distribution:

A handwritten signature in cursive that reads "Albert J. Williams 3rd".

Albert J. Williams 3rd, Chairman
Department of Applied Ocean Physics and Engineering

Optimal Source Localization and Tracking Using Arrays
with Uncertainties in Sensor Locations

by

Mordechai Segal and Ehud Weinstein*

Department of Electronic Systems

Faculty of Engineering

Tel-Aviv University

Ramat-Aviv, Tel-Aviv 69978

ISRAEL

* Ehud Weinstein is also an adjunct scientist with the Department of Ocean Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, U.S.A.

Abstract

We develop a computationally efficient iterative algorithm for source localization and tracking using active/passive arrays with uncertainties in sensor locations. We suppose that the available data consist of time delay, or differential time delay, measurements of the signal wavefront across the array. We consider a general scenario in which the array uncertainties may be correlated in time and in space. The proposed algorithm is optimal in the sense that it converges monotonically to the Maximum Likelihood (ML) estimate of the source trajectory parameters. In the case of multiple sources, the algorithm makes an essential use of the information available from all sources to reduce the array uncertainties (the so-called array calibration) and thus to improve the localization accuracy of each signal source. We also derive new expressions for the log-likelihood gradient, the Hessian, and the Fisher's information matrix, that may be used for efficient implementation of gradient based algorithms, and for assessing the mean square error of the resulting ML parameter estimates.

I. Introduction

The location of a radiating source can be determined by measurement of its signal wavefront using an array of spatially distributed sensors/receivers. In many situations of practical interest, the sensor positions are not precisely known. As demonstrated in [1] [2], uncertainties in the sensor positions may seriously degrade source localization accuracy. In such situations, it has been suggested to use auxiliary sources at known/unknown locations for the purpose of array calibration [2]-[5]. However, most of these studies have concentrated on analyzing the attainable mean square estimation errors without referring to the estimation algorithm that may achieve the indicated performance predictions.

In this report, we develop a computationally efficient iterative scheme for source localization and tracking, based on the Estimate-Maximize (EM) algorithm. We suppose that the available data consist of time delay, or differential time delay, measurements of the signal wavefront across the array. We consider a general scenario in which the array uncertainties may be correlated in time and in space. The proposed algorithm is optimal in the sense that it converges monotonically to the Maximum Likelihood (ML) estimate of the source trajectory parameters. In the case of multiple sources, the algorithm makes an essential use of the information available from all sources to reduce the array uncertainties (the so-called calibration effect) and thus to improve the localization accuracy of each signal source. We also derive new expressions for the log-likelihood gradient, the Hessian, and the Fisher's information matrix, that may be used for efficient implementation of gradient-based algorithms, and for assessing the mean square error (m.s.e.) of the resulting ML parameter estimates.

The organization of the paper is as follows: In section II we formulate the problem and indicate the difficulties associated with the direct ML approach. In section III we apply the EM algorithm for iteratively solving the ML problem in the single source case, and in section

IV we consider the multiple source case. Section V is devoted to gradient-based algorithms and performance evaluation, and finally in section VI we summarize the study.

II. Problem Formulation

Consider the problem of source localization and tracking using an array of M spatially distributed receivers as illustrated in Figure 1. Let the radius vector $\mathbf{p}(t; \theta)$ denote the source location at time t , where θ is the vector unknown trajectory parameters.

Similarly, let $\mathbf{q}_m(t)$ denote the location of the m^{th} receiver, which is composed of a nominally known track $\mathring{\mathbf{q}}_m(t)$, and a random displacement $\delta_m(t)$:

$$\mathbf{q}_m(t) = \mathring{\mathbf{q}}_m(t) + \delta_m(t) \quad (1)$$

Assuming perfect propagation conditions in the medium, the travel time of the signal from the source to the m^{th} receiver is:

$$\begin{aligned} \tau_m(t; \theta) &= \frac{1}{c} \|\mathbf{p}(t; \theta) - \mathbf{q}_m(t)\| \\ &= \frac{1}{c} \|\mathbf{p}(t; \theta) - \mathring{\mathbf{q}}_m(t) - \delta_m(t)\| \end{aligned} \quad (2)$$

where $\|\cdot\|$ denotes the magnitude (norm) of the bracketed vector, and c is the propagation velocity in the medium. Using Taylor series expansion about the nominal sensor locations:

$$\tau_m = \tau_m|_{\mathring{\mathbf{q}}_m} + \frac{\partial \tau_m}{\partial \mathbf{q}_m|_{\mathring{\mathbf{q}}_m}} \delta_m + \frac{1}{2} \delta_m^T \frac{\partial^2 \tau_m}{\partial \mathbf{q}_m^2|_{\mathring{\mathbf{q}}_m}} \delta_m + \dots \quad (3)$$

Differentiating (2),

$$\frac{\partial \tau_m}{\partial \mathbf{q}_m} = -\frac{1}{c} \frac{1}{\|\mathbf{p} - \mathbf{q}_m\|} (\mathbf{p} - \mathbf{q}_m)^T \quad (4)$$

Differentiating (4),

$$\frac{\partial^2 \tau_m}{\partial \mathbf{q}_m^2} = \frac{1}{c} \left[\frac{1}{\|\mathbf{p} - \mathbf{q}_m\|} I - \frac{1}{\|\mathbf{p} - \mathbf{q}_m\|^3} (\mathbf{p} - \mathbf{q}_m)(\mathbf{p} - \mathbf{q}_m)^T \right] \quad (5)$$

where I is the identity matrix. Substituting (4) and (5) into (3)

$$\tau_m = \frac{1}{c} \|\mathbf{p} - \mathring{\mathbf{q}}_m\| \left\{ 1 - \frac{(\mathbf{p} - \mathring{\mathbf{q}}_m)^T \delta_m}{\|\mathbf{p} - \mathring{\mathbf{q}}_m\|^2} + \frac{1}{2} \left[\frac{\|\delta_m\|}{\|\mathbf{p} - \mathring{\mathbf{q}}_m\|} \right]^2 - \frac{1}{2} \left[\frac{(\mathbf{p} - \mathring{\mathbf{q}}_m)^T \delta_m}{\|\mathbf{p} - \mathring{\mathbf{q}}_m\|^2} \right]^2 + \dots \right\} \quad (6)$$

We suppose that the sensor displacements are small compared with the spacing between sensors and with the distance to the source. Therefore, ignoring terms of the order of $(\|\delta_m\| / \|\mathbf{p} - \mathring{\mathbf{q}}_m\|)^2$ (that is, the ratio of sensor displacement to source range squared), τ_m is given to an excellent approximation by:

$$\tau_m(t; \theta) \simeq \mathring{\tau}_m(t; \theta) - \frac{1}{c} \mathbf{d}_m^T(t; \theta) \delta_m(t) \quad (7)$$

where $\mathring{\tau}_m$ is the nominal delay measured relative to the nominal sensor position:

$$\mathring{\tau}_m(t; \theta) = \frac{1}{c} \|\mathbf{p}(t; \theta) - \mathring{\mathbf{q}}_m(t)\|, \quad (8)$$

and \mathbf{d}_m is the unit vector along the line connecting the nominal sensor location with the source location (see Figure 2):

$$\mathbf{d}_m(t; \theta_k) = \frac{\mathbf{p}(t; \theta) - \mathring{\mathbf{q}}_m(t)}{\|\mathbf{p}(t; \theta) - \mathring{\mathbf{q}}_m(t)\|} \quad (9)$$

The product $\mathbf{d}_m^T \delta_m$ appearing in (7) is, therefore, the projection of δ_m in the direction specified by \mathbf{d}_m . Geometrically, this approximation means that the signal wavefront is essentially planar over distances of the order of $\|\delta_m\|$.

Concatenating the equations in (7) for $m = 1, 2, \dots, M$, we obtain:

$$\boldsymbol{\tau}(t; \theta) \simeq \mathring{\boldsymbol{\tau}}(t; \theta) - \frac{1}{c} \mathbf{D}^T(t; \theta_k) \boldsymbol{\delta}(t) \quad (10)$$

where $\boldsymbol{\tau}(t; \theta)$ is the vector time delays of the signal wavefront from the source to the actual receiver locations:

$$\boldsymbol{\tau}(t; \theta) = \begin{bmatrix} \tau_1(t; \theta) \\ \tau_2(t; \theta) \\ \vdots \\ \tau_M(t; \theta) \end{bmatrix}, \quad (11)$$

$\overset{\circ}{\tau}(t; \theta)$ is the vector delays of the signal from the source to the nominal receiver locations:

$$\overset{\circ}{\tau}(t; \theta) = \begin{bmatrix} \overset{\circ}{\tau}_1(t; \theta) \\ \overset{\circ}{\tau}_2(t; \theta) \\ \vdots \\ \overset{\circ}{\tau}_M(t; \theta) \end{bmatrix}, \quad (12)$$

$\delta(t)$ is the vector array uncertainties:

$$\delta(t) = \begin{bmatrix} \delta_1(t) \\ \delta_2(t) \\ \vdots \\ \delta_M(t) \end{bmatrix}, \quad (13)$$

and $D(t; \theta)$ is the matrix:

$$D(t; \theta) = \begin{bmatrix} \mathbf{d}_1(t; \theta) & & (0) \\ & \mathbf{d}_2(t; \theta) & \\ (0) & & \mathbf{d}_M(t; \theta) \end{bmatrix} \quad (14)$$

where the notation (0) indicates that all the other components of the matrix are zero.

All the information concerning source location and track is contained in the delays, or differential delays, of the signal wavefront to the various receivers. Therefore, source localization can be performed by first measuring the propagation delays, and then translating to the estimates of the source trajectory parameters. Since the array uncertainties enter only in the transition from the delay measurements to the track estimates, we shall concentrate only on that phase of the problem.

Let the vector delay measurements at time $t = t_n$ be modeled as:

$$\begin{aligned} \mathbf{z}_n &= H_n \overset{\circ}{\tau}(t_n; \theta) + \mathbf{v}_n \\ &\simeq H_n \overset{\circ}{\tau}(t_n; \theta) - \frac{1}{c} H_n D^T(t_n; \theta) \delta(t_n) + \mathbf{v}_n \end{aligned} \quad (15)$$

where \mathbf{v}_n represents the measurement errors. In the transition from the first line of (15) into its second line we have invoked Eq (10).

The matrix H_n depends on the measurement scheme employed. If an active system is used (e.g., in radar), the observed data consists of the absolute time-of-arrival (TOA) of the signal wavefront to the various sensors, in which case $H_n = I$, the identity matrix. In the passive case (e.g., in sonar), the observed data consist of time-difference-of-arrival (TDOA) of the source signals to the various sensor pairs, in which case the matrix H_n operates on the vector TOA's to generate the appropriate TDOA's. For example, if one measures the TDOA's $[\tau_i(t) - \tau_M(t)]$ $i = 1, 2, \dots, (M - 1)$ using sensor M as the reference, then H_n is the $(M - 1) \times M$ matrix:

$$H_n^{(k)} = \begin{bmatrix} 1 & & & -1 \\ & \ddots & & \vdots \\ & & (0) & -1 \\ & & & \vdots \\ & (0) & & \vdots \\ & & & 1 & -1 \end{bmatrix}$$

The dependence of H_n on the time index n accounts for situations in which we may want to use different combinations of TOA's or TDOA's at different times (perhaps because of unacceptable SNR conditions at some receiver outputs).

We suppose that the measurement noises \mathbf{v}_n $n = 1, 2, \dots$ are statistically uncorrelated zero-mean and Gaussian with the covariance matrix R_n , depending on the available signal-to-noise ratio (SNR) and on the length of the n^{th} observation interval. We note that R_n is generally a non-diagonal matrix because of the statistical correlation between signal observations at different receiver outputs. A lower bound on R_n that can be used to characterize the minimum attainable mean square error (m.s.e.) can be found in [6].

We must now make additional assumptions concerning the array uncertainties $\delta(t_n)$ $n =$

1, 2, Otherwise, the measurement equation (15) contains too many unknowns, and the problem cannot be solved. We suppose that $\delta(t_n)$ $n = 1, 2, \dots$ satisfy the following stochastic difference equation:

$$\delta(t_n) = \Phi_n \delta(t_{n-1}) + \mathbf{u}_n \quad n = 1, 2, \dots \quad (16)$$

where $\delta(t_0)$ is zero-mean and Gaussian with a pre-specified covariance matrix P_0 , and \mathbf{u}_n $n = 1, 2, \dots$ are uncorrelated zero-mean and Gaussian with the covariances Q_n . This model allows spatial as well as temporal correlation between the array uncertainties. We suppose, without any loss of generality, that $\delta(t_0)$, $\{\mathbf{u}_n\}$, and $\{\mathbf{v}_n\}$ are mutually independent.

The estimation problem may now be stated as follows: Given the observed data $\mathbf{z} = \{\mathbf{z}_n : n = 1, 2, \dots, N\}$, find the best possible estimate of the source trajectory parameters θ . If one interprets "best" in the usual sense of minimizing the mean square error (m.s.e.), the maximum likelihood (ML) estimator has a strong claim to optimality because of its asymptotic efficiency and lack of bias.

Observing the (15) subject to (16) constitutes a stochastic dynamic linear state equation (where $\delta(t_n)$ $n = 0, 1, 2, \dots$ are the state variables), the log-likelihood function, that is the logarithm of the probability density of \mathbf{z} , is readily available in the literature (e.g., [7] [8]):

$$\begin{aligned} L_{\mathbf{z}}(\theta) &= \log f_{\mathbf{z}}(\mathbf{z}; \theta) \\ &= b - \frac{1}{2} \sum_{n=1}^N \left\{ \log \det \Lambda_n(\theta) + \boldsymbol{\mu}_n^T(\theta) \Lambda_n^{-1}(\theta) \boldsymbol{\mu}_n(\theta) \right\} \end{aligned} \quad (17)$$

where b is a constant independent of θ , and

$$\boldsymbol{\mu}_n(\theta) = \mathbf{z}_n - H_n \hat{\mathbf{r}}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \hat{\delta}_{n/n-1} \quad (18)$$

$$\Lambda_n(\theta) = \frac{1}{c^2} H_n D^T(t_n; \theta) P_{n/n-1} D(t_n; \theta) H_n^T + R_n \quad (19)$$

where

$$\hat{\delta}_{n/n-1} = E_{\theta} \{ \delta(t_n) / z_1, z_2, \dots, z_{n-1} \} \quad (20)$$

$$P_{n/n-1} = E_{\theta} \{ (\delta_n - \hat{\delta}_{n/n-1})(\delta_n - \hat{\delta}_{n/n-1})^T / z_1, \dots, z_{n-1} \} \quad (21)$$

where $E_{\theta} \{ \cdot / z_1, z_2, \dots, z_{n-1} \}$ denotes the conditional expectation given z_1, z_2, \dots, z_{n-1} , computed with respect to the parameter value θ . By applying the Kalman filtering equations, $\hat{\delta}_{n/n-1}$ and $P_{n/n-1}$ are computed recursively in n .

Now, the ML estimate $\hat{\theta}_{ML}$ is the θ that maximizes $L_{\mathbf{z}}(\theta)$:

$$\text{Max}_{\theta} L_{\mathbf{z}}(\theta) \longrightarrow \hat{\theta}_{ML} \quad (22)$$

This is a complicated optimization problem that is very difficult to solve. Of course, brute force approach can always be used, evaluating the objective function on a coarse grid to roughly locate the global maximum, and then applying the Gauss method or the Newton-Raphson or some other gradient-search iterative algorithm. However, a grid search involves the computation of the Kalman filtering equation for a dense set of parameter values, and the computation of the gradient requires the differentiation of the Kalman filtering equations. Therefore, these methods tend to be computationally tedious and time consuming.

In this study we develop a computationally efficient iterative algorithm for finding the *ML* estimate of θ , without the need to solve the complicated optimization indicated above. The key idea is the if that array uncertainties $\delta(t_n)$ $n = 0, 1, 2, \dots$ were known to the observer, the actual receiver locations were known, and the estimation problem would significantly simplify. Since the array uncertainties are unknown, they are iteratively estimated within the algorithm, and used to iteratively improve the estimate of the various source trajectory

parameters. Under certain regularity conditions, the proposed algorithm converges monotonically to the desired ML solution.

First, we develop the algorithm for the single source case, and then present the extension to the multiple source case.

III. Development of the Algorithm

In this section we apply the Estimate-Maximize (EM) algorithm [9] to solve the indicated problem. The EM algorithm is basically an iterative method for finding ML parameter estimates. It works with the notion of complete data, and iterates between estimating the log-likelihood of the complete data using the observed (incomplete) data and the current parameter estimate (E-step), and maximizing the estimated log-likelihood function to obtain the new parameter estimate (M-step).

More specifically, let \mathbf{y} denote the complete data, that is related to the observed (incomplete) data \mathbf{z} by some non-invertible (many-to-one) transformation:

$$F(\mathbf{y}) = \mathbf{z} \quad (23)$$

Let $\hat{\theta}^{(l)}$ denote the current estimate of θ after l iterations of the algorithm. Then, the next iteration cycle is specified in two steps as follows:

E-Step: Compute

$$Q(\theta, \hat{\theta}^{(l)}) = E_{\hat{\theta}^{(l)}} \{L_{\mathbf{y}}(\theta)/\mathbf{z}\} \quad (24)$$

M-Step:

$$\underset{\theta}{\text{Max}} \quad Q(\theta, \hat{\theta}^{(l)}) \longrightarrow \hat{\theta}^{(l+1)} \quad (25)$$

where $L_{\mathbf{y}}(\theta) = \log f_{\mathbf{Y}}(\mathbf{y}; \theta)$ is the log-likelihood of the complete data, and $E_{\theta'}\{\cdot/\mathbf{z}\}$ denotes the conditional expectation given \mathbf{z} evaluated with respect to the parameter value θ' .

The EM algorithm thus iteratively uses the latest parameter estimate $\hat{\theta}^{(l)}$ to compute the expected value of $L_{\mathbf{y}}(\theta)$. It then maximizes over θ to get a better parameter estimate. On the next iteration cycle, it uses the improved parameter estimate $\hat{\theta}^{(l+1)}$ to improve the

expectation calculation, and thereby to improve the next parameter estimate. It has been shown [9] [10] that if $Q(\theta, \theta')$ is continuous in θ and θ' , the EM algorithm always converges to a stationary point of the observed log-likelihood $L_z(\theta)$, where each iteration cycle increases the value of $L_z(\theta)$. Of course, as in all "hill climbing" algorithms, the convergence point may not be the global maximum, and thus several starting points or an initial grid search may be needed.

To apply the EM algorithm, we need to specify the complete data \mathbf{y} . Clearly, the choice of complete data is not unique (the transformation $F(\cdot)$ relating \mathbf{y} to \mathbf{z} can be any non-invertible transformation). We want to choose a complete data \mathbf{y} so that the iterative expectation and maximization of the associated log-likelihood function $L_{\mathbf{y}}(\theta)$ is computationally simpler than the direct brute-force maximization of $L_z(\theta)$. As pointed out before, if the array uncertainties were known to the observer, the actual receiver locations were known, and we would have a simpler more convenient likelihood function to compute and to maximize. Therefore, a promising choice of complete data \mathbf{y} are the delay observations $\mathbf{z} = \{z_n : n = 1, 2, \dots, N\}$ jointly with the array uncertainties $\delta = \{\delta(t_n) : n = 0, 1, 2, \dots, N\}$:

$$\mathbf{y} = \begin{bmatrix} \mathbf{z} \\ \delta \end{bmatrix} \quad (26)$$

Applying Baye's rule

$$f_{\mathbf{y}}(\mathbf{y}; \theta) = f_{\mathbf{z}/\Delta}(\mathbf{z}/\delta) \cdot f_{\Delta}(\delta) \quad (27)$$

where $f_{\Delta}(\delta)$ is the probability density of δ , and $f_{\mathbf{z}/\Delta}(\mathbf{z}/\delta)$ is the conditional probability density of \mathbf{z} given δ . Taking the logarithm on both sides of (27) and invoking (15), the log-likelihood of \mathbf{y} is given by:

$$L_{\mathbf{y}}(\theta) = b + \log f_{\mathbf{z}/\Delta}(\mathbf{z}/\delta)$$

$$\begin{aligned}
&= b - \sum_{n=1}^N \| \mathbf{z}_n - H_n \tau(t_n; \theta) \|_{R_n}^2 \\
&\simeq b - \sum_{n=1}^N \left\| \mathbf{z}_n - H_n \hat{\tau}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \delta(t_n) \right\|_{R_n}^2
\end{aligned} \tag{28}$$

where b contains all the terms that are independent of θ , including $\log f_{\Delta}(\delta)$, and where we define:

$$\| \mathbf{r} \|_W^2 = \mathbf{r}^T W^{-1} \mathbf{r} \tag{29}$$

Taking the conditional expectation of $L_{\mathbf{y}}(\theta)$ at a parameter value $\hat{\theta}^{(l)}$ and following straightforward matrix manipulations,

$$\begin{aligned}
Q(\theta, \hat{\theta}^{(l)}) &= E_{\hat{\theta}^{(l)}} \{ L_{\mathbf{y}}(\theta) / \mathbf{z} \} \\
&= b - \frac{1}{2} \sum_{n=1}^N \left\{ \left\| \mathbf{z}_n - H_n^{(k)} \hat{\tau}(t_n; \theta_k) + \frac{1}{c} H_n D^T(t_n; \theta) \hat{\delta}^{(l)}(t_n) \right\|_{R_n}^2 \right. \\
&\quad \left. + \frac{1}{c^2} \text{Tr} \left[R_n^{-1} H_n D^T(t_n; \theta) \mathbf{P}^{(l)}(t_n) D(t_n; \theta) H_n^T \right] \right\}
\end{aligned} \tag{30}$$

where $\text{Tr}[\cdot]$ stands for the trace of the bracketed matrix, and where $\hat{\delta}^{(l)}(t_n)$ and $P^{(l)}(t_n)$ are the conditional mean estimate of $\delta(t_n)$ and its error covariance computed at $\theta = \hat{\theta}^{(l)}$:

$$\hat{\delta}^{(l)}(t_n) = E_{\hat{\theta}^{(l)}} \{ \delta(t_n) / \mathbf{z} \} \tag{31}$$

$$\mathbf{P}^{(l)}(t_n) = E_{\hat{\theta}^{(l)}} \left\{ \left[\delta(t_n) - \hat{\delta}^{(l)}(t_n) \right] \left[\delta(t_n) - \hat{\delta}^{(l)}(t_n) \right]^T / \mathbf{z} \right\} \tag{32}$$

The contribution of the $\text{Tr}[\cdot]$ term appearing in (30) is very small since $P^{(l)}(t_n)$ is of the order of the second moment (covariance) of $\delta(t_n)$. Therefore, to a very good approximation,

$$\begin{aligned}
Q(\theta, \hat{\theta}^{(l)}) &\simeq b - \frac{1}{2} \sum_{n=1}^N \left\| \mathbf{z}_n - H_n \hat{\tau}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \hat{\delta}^{(l)}(t_n) \right\|_{R_n}^2 \\
&\simeq b - \frac{1}{2} \sum_{n=1}^N \left\| \mathbf{z}_n - H_n \hat{\tau}^{(l)}(t_n; \theta) \right\|_{R_n}^2
\end{aligned} \tag{33}$$

where in the transition from the first line of (33) to its second line we have used essentially the same approximation as in (10). The vector

$$\hat{\tau}^{(l)}(t_n; \theta) = \begin{bmatrix} \hat{\tau}_1^{(l)}(t_n; \theta) \\ \hat{\tau}_2^{(l)}(t_n; \theta) \\ \vdots \\ \hat{\tau}_M^{(l)}(t_n; \theta) \end{bmatrix} \quad (34)$$

are the propagation delays computed with respect to the estimated receiver locations:

$$\hat{\tau}_m^{(l)}(t_n; \theta) = \frac{1}{c} \| \mathbf{p}(t_n; \theta) - \hat{\mathbf{q}}_m^{(l)}(t_n) \| \quad (35)$$

where $\hat{\mathbf{q}}_m^{(l)}(t_n)$ is the current estimate of the m^{th} sensor location at $t = t_n$, which is composed of the nominal location compensated by the current estimate of the sensor displacement:

$$\hat{\mathbf{q}}_m^{(l)}(t_n) = \mathring{\mathbf{q}}_m(t_n) + \delta_m^{(l)}(t_n) \quad (36)$$

Recall (33), to compute $Q(\theta, \hat{\theta}^{(l)})$ (E-Step) we only need to compute $\hat{\delta}^{(l)}(t_n)$ $n = 1, 2, \dots, N$.

Therefore, in this setting, the EM algorithm assumes the form:

E-Step: Compute

$$\hat{\delta}^{(l)}(t_n) = E_{\hat{\theta}^{(l)}} \{ \delta(t_n) / \mathbf{z} \} \quad n = 1, 2, \dots, N \quad (37)$$

M-Step:

$$\text{Min}_{\theta} \sum_{n=1}^N \| \mathbf{z}_n - H_n \hat{\tau}^{(l)}(t_n; \theta) \|_{R_n}^2 \longrightarrow \hat{\theta}^{(l+1)} \quad (38)$$

The algorithm is illustrated in Figure 3. We note that the minimization in (38) is similar to the ML problem of estimating source trajectory parameters when the array uncertainties

(sensor locations) are precisely known to the observer. The difference is that in (38) we solve the problem with respect to the estimated sensor locations. Thus, the algorithm iterates back and forth, using the current trajectory estimate $\hat{\theta}^{(l)}$ to improve the estimate of the array uncertainties in the E-Step, and thus to improve the next estimate of the source trajectory parameters. Under the stated regularity condition (that is, the continuity of $Q(\theta, \theta')$ which is clearly satisfied here), the algorithm converges monotonically to the ML estimate of θ or, at least, to a stationary point of the log-likelihood function $L_{\mathbf{z}}(\theta)$, where each iteration cycle increases the likelihood of the estimated parameters.

The conditional expectations required in (37) can be carried out efficiently using the Kalman smoothing equations. Details are presented in Appendix A for the convenience of the reader. We have therefore substituted the direct ML approach that requires the computation of the Kalman filtering equations at a dense set of parameter values, with an iterative procedure that involves the Kalman smoothing equations.

The computation of $\hat{\delta}^{(l)}(t_n)$ simplifies if $\delta(t_n)$ $n = 1, 2, \dots$ are statistically independent. This is a special case of the model in (16) where $\Phi_n = 0$, and it corresponds to a situation in which the time differences $(t_n - t_{n-1})$ exceed the correlation time of the uncertainty process. Recall (15), if $\delta(t_n)$ $n = 1, 2, \dots$ are statistically independent then the pair $(\mathbf{z}_n, \delta(t_n))$ is statistically independent of \mathbf{z}_m $m \neq n$. Therefore, to compute $\hat{\delta}^{(l)}(t_n)$, we only need to compute the conditional expectation of $\delta(t_n)$ with respect to \mathbf{z}_n . Since \mathbf{z}_n and $\delta(t_n)$ are jointly Gaussian, satisfying the linear relation in (15), this conditional expectation is easily evaluated:

$$\begin{aligned} \hat{\delta}^{(l)}(t_n) &= E_{\hat{\theta}^{(l)}} \{ \delta(t_n) / \mathbf{z}_n \} \\ &= -\frac{1}{c} Q_n D(t_n; \hat{\theta}^{(l)}) H_n^T \left[\frac{1}{c^2} H_n D^T(t_n; \hat{\theta}^{(l)}) Q_n D(t_n; \hat{\theta}^{(l)}) H_n^T + R_n \right]^{-1} \end{aligned}$$

$$\left[z_n - H_n \hat{r}(t_n; \hat{\theta}^{(l)}) \right] \quad (39)$$

where Q_n is the a-priori covariance of $\delta(t_n)$. Since $\hat{\delta}^{(l)}(t_n)$ depends only on the measurements z_n at $t = t_n$, it can be computed recursively in n . There is no need to store past measurements for the purpose of estimating the array uncertainties.

IV. Multiple Sources

We now extend the scope by considering an array of M spatially distributed sensors, receiving signals from K spatially distributed sources, as illustrated in Figure 4. Let $\mathbf{p}(t; \theta_k)$ denote the trajectory of the k^{th} signal source. More generally, we could use different functions $\mathbf{p}_k(t; \theta_k)$ to characterize the different tracks.

Let the delay measurements associated with the k^{th} signal source at $t = t_n$ be given by:

$$\begin{aligned} \mathbf{z}_n^{(k)} &= H_n^{(k)} \boldsymbol{\tau}(t_n; \theta_k) + \mathbf{v}_n^{(k)} \\ &\simeq H_n^{(k)} \overset{\circ}{\boldsymbol{\tau}}(t_n; \theta_k) - \frac{1}{c} H_n^{(k)} D^T(t_n; \theta_k) \boldsymbol{\delta}(t_n) + \mathbf{v}_n^{(k)} \end{aligned} \quad (40)$$

where $\boldsymbol{\tau}(t; \theta_k)$ is the vector propagation delays measured from the k^{th} source to the actual receiver locations, $\overset{\circ}{\boldsymbol{\tau}}(t; \theta_k)$ is the vector delays measured relative to the nominal receiver locations, $D(t; \cdot)$ is the matrix defined in (14), and $\boldsymbol{\delta}(t_n)$ is the vector array uncertainties satisfying the stochastic model in (16).

We suppose that the measurement noises $\mathbf{v}_n^{(k)}$ associated with different sources are mutually independent zero-mean and Gaussian with the covariance $R_n^{(k)}$. This assumption means that the signals observed from different sources do not interfere with each other (e.g., by confining them to disjoint frequency bands).

The estimation problem may be stated as follows: Given the observed data $\mathbf{z}_n^{(k)}$ $k = 1, 2, \dots, K$, $n = 1, 2, \dots, N$, find the *ML* estimate of the vector parameters:

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_K \end{bmatrix} \quad (41)$$

Since the various $\mathbf{z}_n^{(k)}$ are statistically correlated due to the array uncertainties, the log-

likelihood function is now even more complicated, and its maximization with respect to all the components of θ is a prohibitive task. Of course, we could decompose this complicated multi-dimensional optimization by processing each delay equation separately to obtain the estimate of the corresponding source location and track. However, by that we completely ignore the geometrical constraint imposed by the various sources to reduce the array uncertainties (the so-called calibration effect [2] [3] [5]) and thus to improve the localization accuracy of each signal source.

We want to apply the EM algorithm to solve the problem. As before, let the complete data \mathbf{y} be composed of the delay observations $\mathbf{z} = \{\mathbf{z}_n^{(k)}\}$ jointly with the array uncertainties $\delta = \{\delta(t_n)\}$. Recall (40), and invoking the statistical independence of the various $\mathbf{v}_n^{(k)}$'s, the log-likelihood of \mathbf{y} is given by:

$$\begin{aligned}
L_{\mathbf{y}}(\theta) &= \log f_{\mathbf{y}}(\mathbf{y}; \theta) \\
&= \log f_{\Delta}(\delta) + \log f_{\mathbf{z}/\Delta}(\mathbf{z}/\delta) \\
&= b - \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{z}_n^{(k)} - H_n^{(k)} \tau(t_n; \theta_k)\|_{R_n^{(k)}}^2 \\
&\simeq b - \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{z}_n^{(k)} - H_n^{(k)} \hat{\tau}(t_n; \theta_k) + \frac{1}{c} H_n^{(k)} D^T(t_n; \theta_k) \delta(t_n)\|_{R_n^{(k)}}^2 \quad (42)
\end{aligned}$$

where b contains all the terms that are independent of θ .

Taking the conditional expectation of $L_{\mathbf{y}}(\theta)$ at a parameter value $\hat{\theta}^{(l)}$, and ignoring terms of the order of $O(\|\delta(t_n)\|^2)$ (see the considerations leading from (28) to (33)),

$$\begin{aligned}
Q(\theta, \hat{\theta}^{(l)}) &= E_{\hat{\theta}^{(l)}} \{L_{\mathbf{y}}(\theta)/\mathbf{z}\} \\
&\simeq b - \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{z}_n^{(k)} - H_n^{(k)} \hat{\tau}(t_n; \theta_k) + \frac{1}{c} H_n^{(k)} D^T(t_n; \theta_k) \hat{\delta}^{(l)}(t_n)\|_{R_n^{(k)}}^2 \\
&\simeq b - \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{z}_n^{(k)} - H_n^{(k)} \hat{\tau}^{(l)}(t_n; \theta_k)\|_{R_n^{(k)}}^2 \quad (43)
\end{aligned}$$

where $\hat{\delta}^{(l)}(t_n)$ is the conditional mean estimate of $\delta(t_n)$ computed at $\theta = \hat{\theta}^{(l)}$, and $\hat{\tau}^{(l)}(t; \cdot)$ is

defined by (34) – (36).

Thus, to compute $Q(\theta, \hat{\theta}^{(l)})$ (E-step) we only need to compute $\hat{\delta}^{(l)}(t_n)$ $n = 1, 2, \dots$, and the maximization of $Q(\theta, \hat{\theta}^{(l)})$ (M-Step) is equivalent to the minimization of each of the terms in the k sum separately. Therefore, the EM algorithm assumes the form:

E-step: Compute

$$\hat{\delta}^{(l)}(t_n) = E_{\hat{\theta}^{(l)}} \{ \delta(t_n) / \mathbf{z} \} \quad n = 1, 2, \dots N \quad (44)$$

M-step: For $k = 1, 2, \dots K$

$$\text{Min}_{\theta_k} \sum_{n=1}^N \| \mathbf{z}_n^{(k)} - H_n^{(k)} \hat{\mathbf{r}}^{(l)}(t_n; \theta_k) \|_{R_n^{(k)}}^2 \longrightarrow \hat{\theta}_k^{(l+1)} \quad (45)$$

The algorithm is illustrated in Figure 5. Once again we note that the minimization in (45) finds the ML location estimate of the k^{th} signal source with respect to the estimated sensor locations. Thus, the algorithm iterates back and forth, using the current estimate $\hat{\theta}^{(l)}$ of the various source location parameters to improve the estimate of the array uncertainties and thus to improve the next source location estimates. Since the algorithm is based on the EM method, it is ensured to converge monotonically to the joint ML estimate of all source location parameters, or at least to a stationary point of the joint log-likelihood function.

All the effect of array callibration enters the E-step, where we use the data available from all signal sources to form the best possible estimate of the array uncertainties. We note that if some of the signal sources have known locations, they are still used in the E-step (with their exact locations). In the M-step, of course, we only re-estimate the unknown location parameters.

Perhaps the most striking feature of the algorithm is that it decomposes the multi-dimensional optimization associated with the direct ML approach into optimizations with

respect to each source location parameters separately leading to a considerable simplification in estimator structure and computations. Because of its parallel structure, the algorithm incorporates any pre-specified number of signal sources with only a linear increase in the computational complexity.

To compute $\hat{\delta}^{(l)}(t_n)$, we need to perform the conditional expectation indicated in (44).

For that purpose, we concatenate the various measurement equations in (40):

$$\begin{aligned} \mathbf{z}_n &= H_n \boldsymbol{\tau}(t_n; \boldsymbol{\theta}) + \mathbf{v}_n \\ &\simeq H_n \overset{\circ}{\boldsymbol{\tau}}(t_n; \boldsymbol{\theta}) - \frac{1}{c} H_n D^T(t_n; \boldsymbol{\theta}) \delta(t_n) + \mathbf{v}_n \end{aligned} \quad (46)$$

where we define:

$$\mathbf{z}_n = \begin{bmatrix} \mathbf{z}_n^{(1)} \\ \mathbf{z}_n^{(2)} \\ \vdots \\ \mathbf{z}_n^{(K)} \end{bmatrix} \quad (47)$$

$$\boldsymbol{\tau}(t; \boldsymbol{\theta}) = \begin{bmatrix} \tau(t; \theta_1) \\ \tau(t; \theta_2) \\ \vdots \\ \tau(t; \theta_K) \end{bmatrix} \quad (48)$$

$$\overset{\circ}{\boldsymbol{\tau}}(t; \boldsymbol{\theta}) = \begin{bmatrix} \overset{\circ}{\tau}(t; \theta_1) \\ \overset{\circ}{\tau}(t; \theta_2) \\ \vdots \\ \overset{\circ}{\tau}(t; \theta_K) \end{bmatrix} \quad (49)$$

$$D(t; \theta) = [D(t; \theta_1) D(t; \theta_2) \dots D(t; \theta_K)] \quad (50)$$

$$H_n = \begin{bmatrix} H_n^{(1)} & & (0) \\ & H_n^{(2)} & \\ (0) & & H_n^{(K)} \end{bmatrix} \quad (51)$$

and

$$\mathbf{v}_n = \begin{bmatrix} \mathbf{v}_n^{(1)} \\ \mathbf{v}_n^{(2)} \\ \vdots \\ \mathbf{v}_n^{(K)} \end{bmatrix} \quad (52)$$

where we note that \mathbf{v}_n $n = 1, 2, \dots$ are statistically independent zero-mean and Gaussian with the covariance matrix:

$$R_n = \begin{bmatrix} R_n^{(1)} & & (0) \\ & R_n^{(2)} & \\ (0) & & R_n^{(K)} \end{bmatrix} \quad (53)$$

Equation (46) is identical in structure to equation (15). Therefore, observing that (46) subject to (16) constitutes a stochastic dynamic linear state equation where $\delta(t_n)$ $n = 1, 2, \dots$ are the state variables, the state estimates $\hat{\delta}^{(l)}(t_n)$ can be calculated efficiently using the Kalman smoothing equations (see Appendix A).

If $\delta(t_n)$ $n = 1, 2, \dots$ are statistically independent, there is no need to use the Kalman smoothing equations and we obtain, in complete analogy with (39),

$$\begin{aligned} \hat{\delta}^{(l)}(t_n) &= E_{\hat{\theta}^{(l)}} \{ \delta(t_n) / \mathbf{z}_n \} \\ &= -\frac{1}{c} Q_n D(t_n; \hat{\theta}^{(l)}) H_n^T \left[\frac{1}{c^2} H_n D^T(t_n; \hat{\theta}^{(l)}) Q_n D(t_n; \hat{\theta}^{(l)}) H_n^T + R_n \right]^{-1} \end{aligned}$$

$$\left[\mathbf{z}_n - H_n \hat{\mathbf{r}}(t_n; \hat{\boldsymbol{\theta}}^{(l)}) \right] \quad (54)$$

where \mathbf{z}_n , $\hat{\mathbf{r}}(t; \boldsymbol{\theta})$, $D(t; \boldsymbol{\theta})$, H_n , and R_n are given by (47), (49), (50), (51), and (53), respectively. Since $\hat{\boldsymbol{\theta}}^{(l)}(t_n)$ depends only on the measurements \mathbf{z}_n at $t = t_n$, they can be computed recursively in n . There is no need to store past measurements for the purpose of estimating the array uncertainties.

V. Gradient-Based Algorithms and Performance Evaluation

As indicated in [9] [11], the rate of convergence of the EM algorithm is exponential (linear). Therefore, near the point of convergence, we may want to use faster converging gradient-based methods such as the Newton-Raphson or the Scoring algorithm. These methods require the computation of the log-likelihood gradient, and the computation of the log-likelihood Hessian or the Fisher's information matrix (FIM), or some approximation of which. The FIM can further be used to assess the m.s.e. of the resulting ML parameter estimates.

The computation of the log-likelihood gradient (score) $\partial L_{\mathbf{z}}(\theta)/\partial\theta$, obtained by differentiating (17), is extremely complicated since it involves the differentiation of the Kalman filtering equations. An alternative approach that yields the same result with much simpler computations is based on Fisher's identity [12]:

$$\frac{\partial}{\partial\theta} L_{\mathbf{z}}(\theta) = E_{\theta} \left\{ \frac{\partial}{\partial\theta} L_{\mathbf{y}}(\theta) / \mathbf{z} \right\} \quad (55)$$

Fisher's identity asserts that the observed data score $\frac{\partial}{\partial\theta} L_{\mathbf{z}}(\theta)$ can be computed by taking the conditional expectation of the complete data score $\frac{\partial}{\partial\theta} L_{\mathbf{y}}(\theta)$. A sketch proof of the identity is presented in Appendix B. A formal proof with the associated regularity conditions can be found in [13].

Consider first the single source case. Differentiating (28)

$$\frac{\partial}{\partial\theta} L_{\mathbf{y}}^T(\theta) = \sum_{n=1}^N \frac{\partial \tau^T(t_n; \theta)}{\partial\theta} H_n^T R_n^{-1} [\mathbf{z}_n - H_n \tau(t_n; \theta)] \quad (56)$$

where the columns $\partial \tau_m^T(t; \theta)/\partial\theta$ of $\partial \tau^T(t; \theta)/\partial\theta$ are closely approximated using Taylor series expansion about the nominal sensor locations:

$$\frac{\partial \tau_m^T(t; \theta)}{\partial\theta} = \frac{\partial \tau_m^T(t; \theta)}{\partial\theta} \Big|_{\hat{\mathbf{q}}_m(t)} + \frac{\partial^2 \tau_m^T(t; \theta)}{\partial \mathbf{q}_m(t) \partial \theta} \Big|_{\hat{\mathbf{q}}_m(t)} \cdot \delta_m(t) + o(\|\delta_m(t)\|^2) \quad (57)$$

Differentiating (2):

$$\frac{\partial \tau_m^T(t; \theta)}{\partial\theta} \Big|_{\hat{\mathbf{q}}_m(t)} = \frac{1}{c} \frac{\partial \mathbf{p}^T(t; \theta)}{\partial\theta} \mathbf{d}_m(t; \theta) \quad (58)$$

Differentiating (4)

$$\frac{\partial^2 \tau_m^T(t; \theta)}{\partial \mathbf{q}_m(t) \partial \theta} \Big|_{\hat{\mathbf{q}}_m(t)} = -\frac{1}{c} \frac{1}{\|\mathbf{p}(t; \theta) - \hat{\mathbf{q}}_m(t)\|} \frac{\partial \mathbf{p}^T(t; \theta)}{\partial \theta} [I - \mathbf{d}_m(t; \theta), \mathbf{d}_m^T(t; \theta)] \quad (59)$$

where $\mathbf{d}_m(t; \theta)$ is the unit vector defined in (9). Substituting (58) and (59) into (57) and ignoring terms of the order of $\|\delta_m(t)\| / \|\mathbf{p}(t; \theta) - \hat{\mathbf{q}}_m(t)\|$ (that is, the ratio of array uncertainties to the source range),

$$\frac{\partial \tau_m^T(t; \theta)}{\partial \theta} \simeq \frac{1}{c} \frac{\partial \mathbf{p}^T(t; \theta)}{\partial \theta} \mathbf{d}_m(t; \theta) = \frac{\partial \hat{\tau}_m^T(t; \theta)}{\partial \theta} \quad (60)$$

Eq. (60) asserts that to a first order approximation, the sensor displacements have negligible effect on the partial derivatives $\partial \tau_m(t; \theta) / \partial \theta$. Therefore, using (60) in (56)

$$\begin{aligned} \frac{\partial}{\partial \theta} L_{\mathbf{y}}^T(\theta) &\simeq \sum_{n=1}^N \frac{\partial \hat{\tau}^T(t_n; \theta)}{\partial \theta} H_n^T R_n^{-1} [\mathbf{z}_n - H_n \tau(t_n; \theta)] \\ &\simeq \sum_{n=1}^N \frac{\partial \hat{\tau}^T(t_n; \theta)}{\partial \theta} H_n^T R_n^{-1} \left[\mathbf{z}_n - H_n \hat{\tau}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \delta(t_n) \right] \end{aligned} \quad (61)$$

Substituting (61) into (55), the observed data score is given by:

$$L_{\mathbf{z}}^T(\theta) = \sum_{n=1}^N \frac{\partial \hat{\tau}^T(t_n; \theta)}{\partial \theta} H_n^T R_n^{-1} \left[\mathbf{z}_n - H_n \hat{\tau}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \hat{\delta}(t_n) \right] \quad (62)$$

where $\hat{\delta}(t_n)$ is the conditional mean estimate of $\delta(t_n)$:

$$\hat{\delta}(t_n) = E_{\theta} \{ \delta(t_n) / \mathbf{z} \} \quad (63)$$

The computation of $\hat{\delta}(t_n)$ can be carried out efficiently using the smoothing equations presented in Appendix A. We have therefore substituted the direct differentiation of (17) that involves the differentiation of the Kalman filtering equations with a much simpler computation that only involves the Kalman smoothing equations.

Taking the second moment (covariance) of the score, the FIM is given by:

$$\begin{aligned}
J(\theta) &= E_{\theta} \left\{ \frac{\partial L_{\mathbf{z}}^T(\theta)}{\partial \theta} \frac{\partial L_{\mathbf{z}}(\theta)}{\partial \theta} \right\} \\
&= \sum_{n=1}^N \sum_{m=1}^M \frac{\partial \hat{\tau}^{\circ T}(t_n; \theta)}{\partial \theta} H_n^T R_n^{-1} E_{\theta} \{ \hat{\mathbf{v}}_n \hat{\mathbf{v}}_m^T \} R_m^{-1} H_m \frac{\partial \hat{\tau}^{\circ}(t_m; \theta)}{\partial \theta}
\end{aligned} \tag{64}$$

where:

$$\hat{\mathbf{v}}_n = \mathbf{z}_n - H_n \hat{\tau}^{\circ}(t_n; \theta) + \frac{1}{c} H_n D^T(t_n; \theta) \hat{\delta}(t_n) \tag{65}$$

Recursive formulae for calculation $E_{\theta} \{ \hat{\mathbf{v}}_n \hat{\mathbf{v}}_k^T \}$ can be found in [14].

To compute the log-likelihood Hessian we must differentiate the expression for the score. In practice, the Hessian may be approximated by replacing derivatives with finite differences. To approximate the partial derivatives of the score, we perturb one coordinate of θ at a time and compute the resulting score according to (62). This approximation requires the computation of the Kalman smoothing equations at closely spaced values of θ , a task that can be simplified by pre-computation of the smoothed error covariance matrices. Another approximation of the Hessian and the FIM, based on score computation, is presented in [14].

In the multiple source case, the score and the FIM are still given by (62) and (64), where \mathbf{z}_n , $\hat{\tau}^{\circ}(t; \theta)$, $D(t; \theta)$, H_n , and R_n are given by (47), (49), (50), (51), and (53), respectively and $\partial \hat{\tau}^{\circ}(t; \theta)/\partial \theta$ is the corresponding block diagonal matrix.

The formule for the score and the FIM are simplified if the uncertainty process $\delta(t_n)$ $n = 1, 2, \dots$ is uncorrelated in time. Recall Eq. (39), we can write in this case

$$\begin{aligned}
\hat{\delta}(t_n) &= E_{\theta} \{ \delta(t_n) / \mathbf{z}_n \} \\
&= -\frac{1}{c} Q_n D(t_n; \theta) H_n^T \left[\frac{1}{c^2} H_n D^T(t_n; \theta) Q_n D(t_n; \theta) H_n^T + R_n \right]^{-1} \\
&\quad \left[\mathbf{z}_n - H_n \hat{\tau}^{\circ}(t_n; \theta) \right]
\end{aligned} \tag{66}$$

Substituting (66) into (62) and following straightforward matrix manipulations, we obtain

the following explicit formula for the score:

$$\frac{\partial}{\partial \theta} L_{\mathbf{z}}^T(\theta) = \sum_{n=1}^N \frac{\partial \overset{\circ}{\mathbf{r}}^T(t_n; \theta)}{\partial \theta} H_n^T \left[\frac{1}{c^2} H_n D^T(t_n; \theta) Q_n D(t_n; \theta) H_n^T + R_n \right]^{-1} [\mathbf{z}_n - H_n \overset{\circ}{\mathbf{r}}(t_n; \theta)] \quad (67)$$

By the virtue of (15),

$$\mathbf{z}_n - H_n \overset{\circ}{\mathbf{r}}(t_n; \theta) \simeq -\frac{1}{c} H_n D^T(t_n; \theta) \delta(t_n) + \mathbf{v}_n \quad (68)$$

Substituting (55) into (54) and taking the second moment, we obtain the following explicit formula for the FIM:

$$J_{\mathbf{z}}(\theta) = \sum_{n=1}^N \frac{\partial \overset{\circ}{\mathbf{r}}^T(t_n; \theta)}{\partial \theta} H_n^T \left[\frac{1}{c^2} H_n D^T(t_n; \theta) Q_n D(t_n; \theta) H_n^T + R_n \right]^{-1} H_n \frac{\partial \overset{\circ}{\mathbf{r}}(t_n; \theta)}{\partial \theta} \quad (69)$$

Once again we note that the expression in (67) and (69) are valid also for the multiple source case. The result in (69) is a generalization of the results developed in [1], [2], [4], [5]. This expression can be used to analyze the attainable m.s.e. for any pre-specified array-source geometry and any given delay measurement scheme, and the realizable improvement in localization accuracy obtainable due to the array calibration effect.

VI. Conclusion

We have developed a computationally efficient iterative algorithm for source localization and tracking using active/passive arrays with uncertainties in sensor locations. We suppose that the available data consist of time delay, or differential time delay, measurements of the signal wavefront across the array. We consider a general scenario in which the array uncertainties may be correlated in time and in space. The proposed algorithm is optimal in the sense that it converges monotonically to the Maximum Likelihood (ML) estimate of the source trajectory parameters. In the case of multiple sources, the algorithm makes an essential use of the information available from all sources to reduce the array uncertainties (the so-called array calibration) and thus to improve the localization accuracy of each signal source. We also derived new expressions for the log-likelihood gradient, the Hessian, and the Fisher's information matrix, that may be used for efficient implementation of gradient based algorithms, and for assessing the mean square error of the resulting ML parameter estimates.

Acknowledgements

This study was supported in part by the Office of Naval Research under contract N00014-85-K-0272. The authors would like to thank Mr. James G. Smith of ONR for the encouragement and support, and Ms. Sue Oliver of the Woods Hole Oceanographic Institution for her excellent secretarial assistance.

References

- [1] P.M. Schultheiss and J.P. Ianniello, "Optimum range and bearing estimation with randomly perturbed array," *J. Acoust. Soc. Am.*, Vol. 68, No. 1, pp. 167-173, July 1980.
- [2] P.M. Schultheiss, E. Ashok and J.P. Ianniello, "Optimum and sub-optimum source localization with sensors subject to random motion," *J. Acoust. Soc. Am.*, Vol. 74, No. 1, pp. 131-142, July 1983.
- [3] N.L. Owsley, "Joint source and sensor location estimation," *Proc. IEEE Int. Conf. on ASSP*, pp. 402, 1984.
- [4] Y. Rockah and P.M. Schultheiss, "Array shape calibration using sources in unknown locations — Part I: Far-field sources," *IEEE Trans. on Acoust., Speech and Signal Processing*, Vol. ASSP-35, pp. 286-299, March 1987.
- [5] Y. Rockah and P.M. Schultheiss, "Array shape calibration using sources in unknown locations — Part II: Near field sources and estimator implementation," *IEEE Trans. on Acoust., Speech and Signal Processing*, Vol. ASSP-35, pp. 724-735, June 1987.
- [6] P.M. Schultheiss and E. Weinstein, "Lower bounds on the localization errors of a moving source observed by a passive array," *IEEE Trans. of Acoust., Speech and Signal Processing*, Vol. ASSP-29, pp. 600-607, 1981.
- [7] F.C. Schweppe, *Uncertain Dynamic Systems*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.
- [8] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: The M.I.T. Press, 1974.

- [9] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum likelihood estimation from incomplete data via the EM-algorithm," Jour. Roy. Stat. Soc., Ser. B., Vol. 39, pp. 1-38, 1977.
- [10] C.F.J. Wu, "On the convergence properties of the EM-algorithm," The Annals of Statistics, No. 11, pp. 95-103, 1983.
- [11] I. Meilijson, "A fast improvement to the EM-algorithm on its own terms," to be published (Jour. Roy. Stat. Soc. Ser. B).
- [12] R.A. Fisher, "Theory of statistical estimation," Proc. Cambridge Phil. Soc., 22, pp. 700-725, 1925.
- [13] M. Segal and E. Weinstein, "A new method for evaluating the log-likelihood gradient (score) of linear dynamic systems," IEEE Trans. on Automatic Control, August 1988.
- [14] M. Segal and E. Weinstein, "A new method for evaluating the log-likelihood gradient, the Hessian, and the Fisher's information matrix of linear dynamic systems," to appear in the IEEE Trans. on Information Theory.

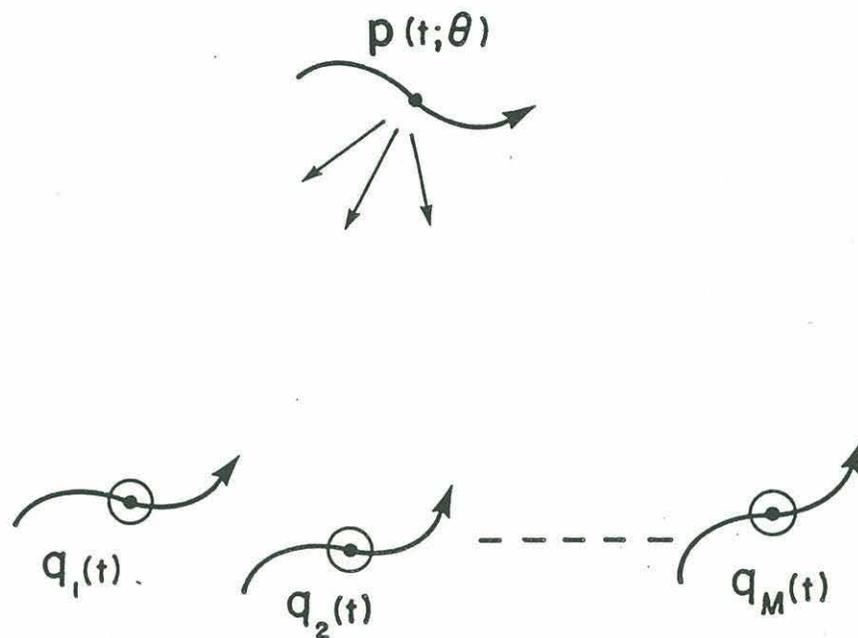


Figure 1

Figure 1: Array-Source Geometry.

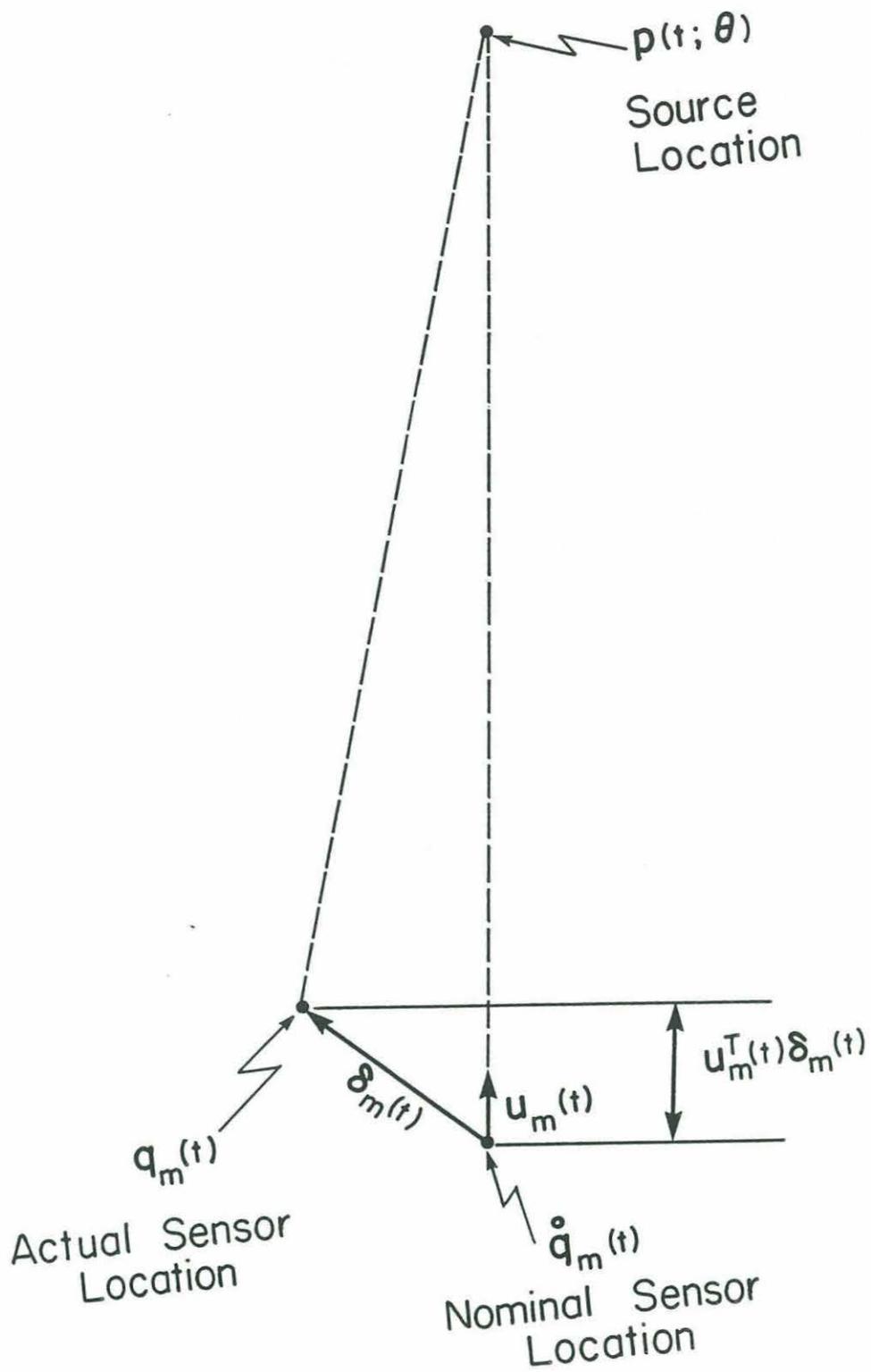


Figure 2

Figure 2: Effect of Sensor Uncertainties on Array-Source Geometry.

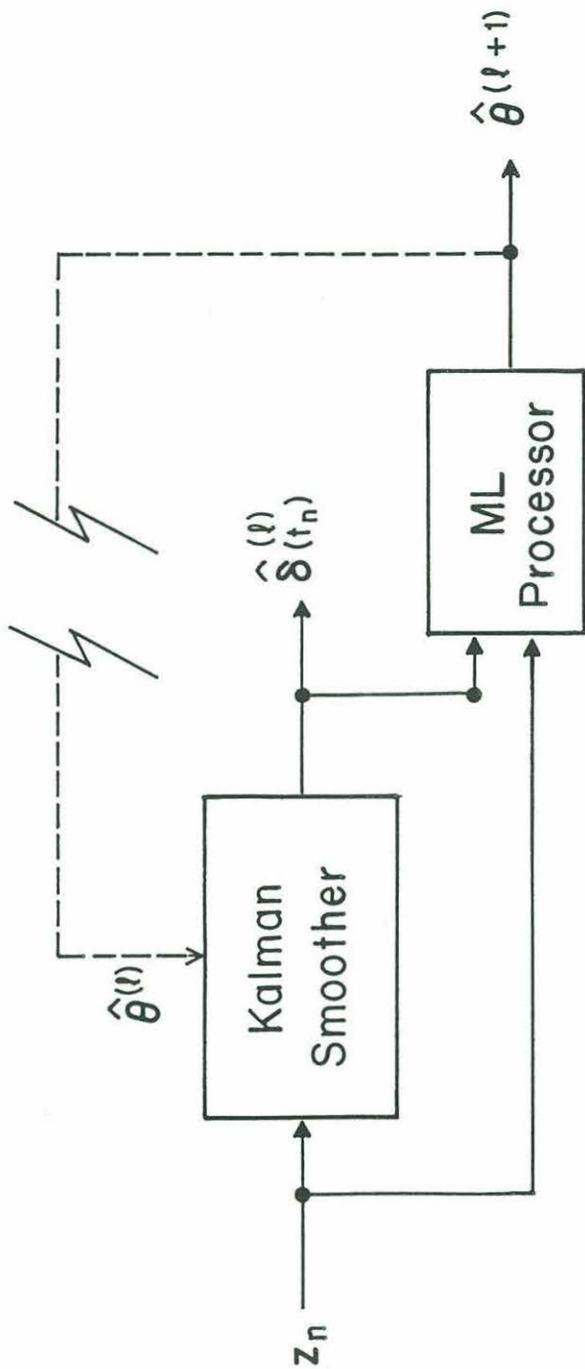


Figure 3: The Proposed Algorithm.

Figure 3

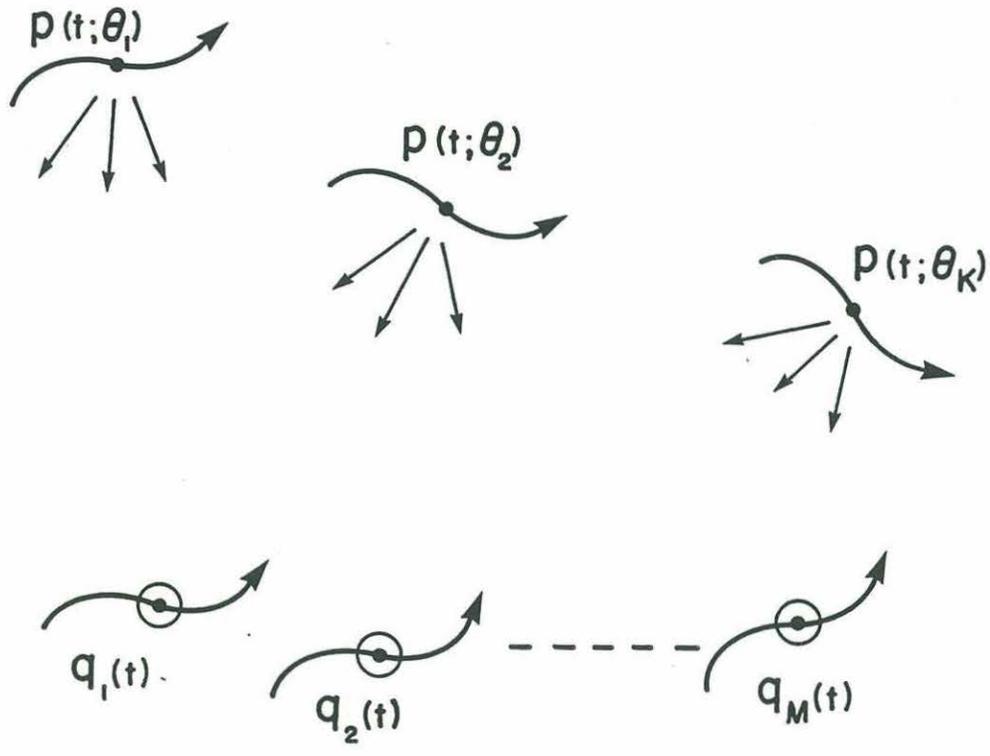


Figure 4

Figure 4: Array-Source Geometry for Multiple Sources.

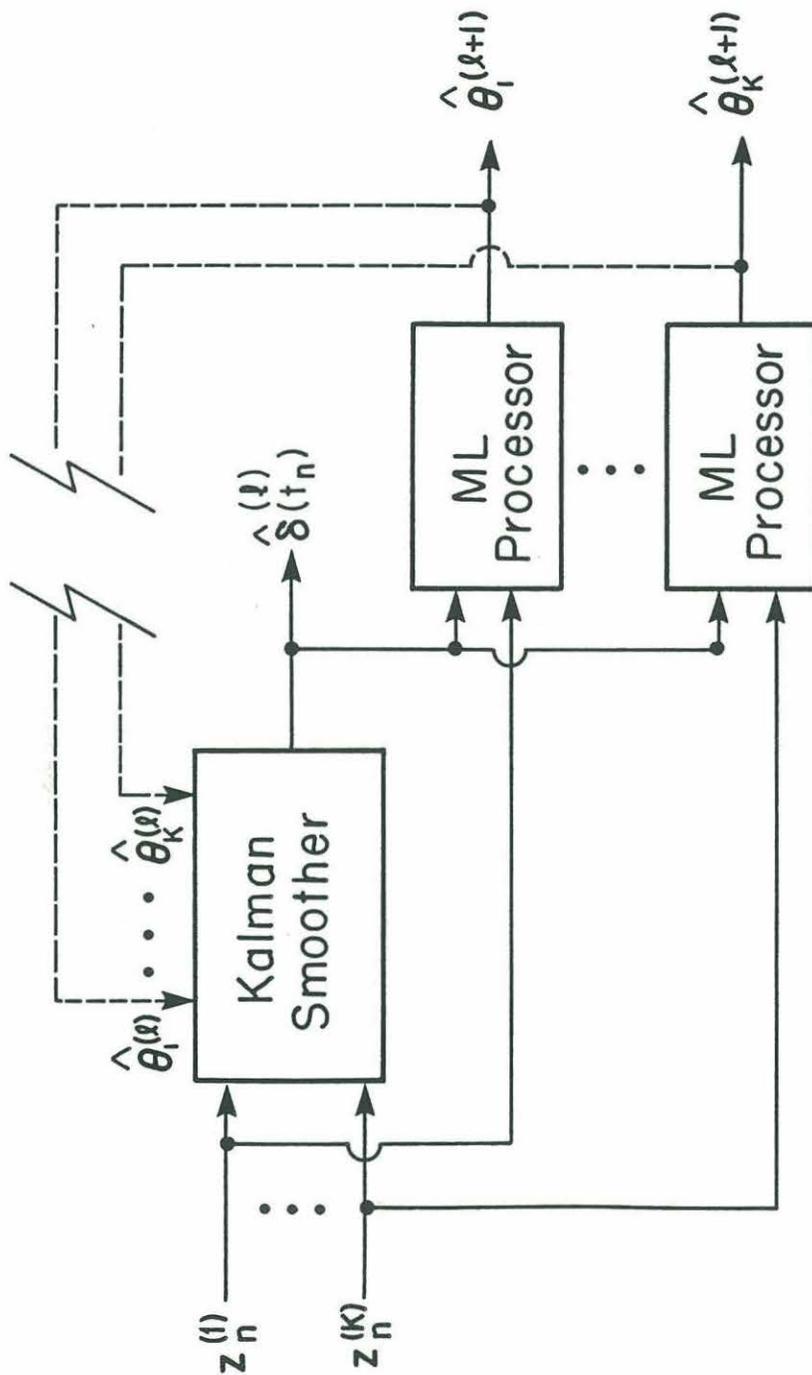


Figure 5: The Proposed Algorithm for Multiple Sources.

Figure 5

Appendix A: Kalman Smoothing Equations for Array Uncertainty Estimation

The array uncertainties $\delta(t_n)$ $n = 1, 2, \dots$ and the observed data z_n $n = 1, 2, \dots$ satisfy the following stochastic dynamic linear state equations:

$$\begin{aligned}\delta(t_n) &= \Phi_n \delta(t_{n-1}) + \mathbf{u}_n \\ z_n &= H_n \delta(t_n) + \mathbf{v}_n \quad n = 1, 2, \dots, N\end{aligned}$$

Define:

$$\begin{aligned}\mu_{n/k} &= E_{\theta} \{ \delta(t_n) / z_1 z_2 \cdots z_k \} \\ P_{n/k} &= E_{\theta} \left\{ \left[\delta(t_n) - \mu_{n/k} \right] \left[\delta(t_n) - \mu_{n/k} \right]^T / z_1, z_2 \cdots z_k \right\}\end{aligned}$$

Propagation Equations:

$$\begin{aligned}\text{For } n &= 1, 2, \dots, N \\ \mu_{n/n-1} &= \Phi_n \mu_{n-1/n-1} & \mu_{0/0} &= 0 \\ P_{n/n-1} &= \Phi_n P_{n-1/n-1} \Phi_n^T + Q_n & P_{0/0} &= P_0\end{aligned}$$

Up-Dating Equations:

$$\begin{aligned}\text{For } n &= 1, 2, \dots, N \\ \mu_{n/n} &= \mu_{n/n-1} + K_n \left[z_n - H_n \mu_{n/n-1} \right] \\ P_{n/n} &= [I - K_n H_n] P_{n/n-1}\end{aligned}$$

where K_n is the Kalman Gain:

$$K_n = P_{n/n-1} H_n^T [H_n P_{n/n-1} H_n^T + R_n]^{-1}$$

Smoothing Equations:

For $n = N, N-1, \dots, 1$

$$\mu_{n-1/n} = \mu_{n-1/n-1} + S_{n-1} [\mu_{n/N} - \Phi_n \mu_{n-1/n-1}]$$

$$P_{n-1/N} = P_{n-1/n-1} + S_{n-1} [P_{n/N} - P_{n/n-1}] S_{n-1}^T$$

where

$$S_{n-1} = P_{n-1/n-1} \Phi_n^T P_{n/n-1}^{-1}$$

The outcome of the smoothing equations are the uncertainty estimates and the associated error covariances:

$$\hat{\delta}(t_n) = \mu_{n/N}$$

$$P(t_n) = P_{n/N}$$

Appendix B: Fisher's Identity

Let \mathbf{z} and $\mathbf{y} = (\mathbf{z}^T; \delta^T)^T$ denote the observed (incomplete) and the complete data vectors, respectively. Using Baye's rule for probability densities,

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{\mathbf{Z}}(\mathbf{z}; \theta) \cdot f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta)$$

Taking the logarithm on both sides of the equation and differentiating with respect to θ ,

$$\frac{\partial}{\partial \theta} \log f_{\mathbf{Z}}(\mathbf{z}; \theta) = \frac{\partial}{\partial \theta} \log f_{\mathbf{Y}}(\mathbf{y}; \theta) - \frac{\partial}{\partial \theta} \log f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta)$$

Taking the conditional expectation given \mathbf{z} ,

$$\frac{\partial}{\partial \theta} \log f_{\mathbf{Z}}(\mathbf{z}; \theta) = E_{\theta} \left\{ \frac{\partial}{\partial \theta} \log f_{\mathbf{Y}}(\mathbf{y}; \theta) / \mathbf{z} \right\} - E_{\theta} \left\{ \frac{\partial}{\partial \theta} \log f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) / \mathbf{z} \right\}$$

Now,

$$\begin{aligned} & E_{\theta} \left\{ \frac{\partial}{\partial \theta} \log f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) / \mathbf{z} \right\} = \\ &= \int \frac{\partial}{\partial \theta} \log f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) \cdot f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) d\delta = \int \frac{\partial}{\partial \theta} f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) d\delta \\ &= \frac{\partial}{\partial \theta} \int f_{\Delta/\mathbf{Z}}(\delta/\mathbf{z}; \theta) d\delta = \frac{\partial}{\partial \theta} (1) = 0 \end{aligned}$$

Therefore,

$$\frac{\partial}{\partial \theta} \log f_{\mathbf{Z}}(\mathbf{z}; \theta) = E_{\theta} \left\{ \frac{\partial}{\partial \theta} \log f_{\mathbf{Y}}(\mathbf{y}; \theta) / \mathbf{z} \right\}$$

or,

$$\frac{\partial}{\partial \theta} L_{\mathbf{z}}(\theta) = E_{\theta} \left\{ \frac{\partial}{\partial \theta} L_{\mathbf{y}}(\theta) / \mathbf{z} \right\}$$

which is Fisher's identity.

DOCUMENT LIBRARY

July 5, 1989

Distribution List for Technical Report Exchange

Attn: Stella Sanchez-Wade
Documents Section
Scripps Institution of Oceanography
Library, Mail Code C-075C
La Jolla, CA 92093

Hancock Library of Biology &
Oceanography
Alan Hancock Laboratory
University of Southern California
University Park
Los Angeles, CA 90089-0371

Gifts & Exchanges
Library
Bedford Institute of Oceanography
P.O. Box 1006
Dartmouth, NS, B2Y 4A2, CANADA

Office of the International
Ice Patrol
c/o Coast Guard R & D Center
Avery Point
Groton, CT 06340

Library
Physical Oceanographic Laboratory
Nova University
8000 N. Ocean Drive
Dania, FL 33304

NOAA/NESDIS Miami Library Center
4301 Rickenbacker Causeway
Miami, FL 33149

Library
Skidaway Institute of Oceanography
P.O. Box 13687
Savannah, GA 31416

Institute of Geophysics
University of Hawaii
Library Room 252
2525 Correa Road
Honolulu, HI 96822

Library
Chesapeake Bay Institute
4800 Atwell Road
Shady Side, MD 20876

MIT Libraries
Serial Journal Room 14E-210
Cambridge, MA 02139

Director, Ralph M. Parsons Laboratory
Room 48-311
MIT
Cambridge, MA 02139

Marine Resources Information Center
Building E38-320
MIT
Cambridge, MA 02139

Library
Lamont-Doherty Geological
Observatory
Columbia University
Palisades, NY 10964

Library
Serials Department
Oregon State University
Corvallis, OR 97331

Pell Marine Science Library
University of Rhode Island
Narragansett Bay Campus
Narragansett, RI 02882

Working Collection
Texas A&M University
Dept. of Oceanography
College Station, TX 77843

Library
Virginia Institute of Marine Science
Gloucester Point, VA 23062

Fisheries-Oceanography Library
151 Oceanography Teaching Bldg.
University of Washington
Seattle, WA 98195

Library
R.S.M.A.S.
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Maury Oceanographic Library
Naval Oceanographic Office
Stennis Space Center
NSTL, MS 39522-5001

Marine Sciences Collection
Mayaguez Campus Library
University of Puerto Rico
Mayaguez, Puerto Rico 00708

REPORT DOCUMENTATION PAGE	1. REPORT NO. WHOI-89-32	2.	3. Recipient's Accession No.
4. Title and Subtitle Optimal Source Localization and Tracking Using Arrays with Uncertainties in Sensor Locations		5. Report Date August 1989	
7. Author(s) Mordechai Segal and Ehud Weinstein		8. Performing Organization Rept. No. WHOI-89-32	
9. Performing Organization Name and Address The Woods Hole Oceanographic Institution Woods Hole, Massachusetts 02543		10. Project/Task/Work Unit No.	
12. Sponsoring Organization Name and Address The Office of Naval Research		11. Contract(C) or Grant(G) No. (C) N00014-85-K-0272 (G)	
15. Supplementary Notes This report should be cited as: Woods Hole Oceanog. Inst. Tech. Rept., WHOI-89-32.		13. Type of Report & Period Covered Technical Report	
16. Abstract (Limit: 200 words) We develop a computationally efficient iterative algorithm for source localization and tracking using active/passive arrays with uncertainties in sensor locations. We suppose that the available data consist of time delay, or differential time delay, measurements of the signal wavefront across the array. We consider a general scenario in which the array uncertainties may be correlated in time and in space. The proposed algorithm is optimal in the sense that it converges monotonically to the Maximum Likelihood (ML) estimate of the source trajectory parameters. In the case of multiple sources, the algorithm makes an essential use of the information available from all sources to reduce the array uncertainties (the so-called array callibration) and thus to improve the localization accuracy of each signal source. We also derive new expressions for the log-likelihood gradient, the Hessian, and the Fisher's information matrix, that may be used for efficient implementation of gradient based algorithms, and for assessing the mean square error of the resulting ML parameter estimates.		14.	
17. Document Analysis a. Descriptors			
1. source localization		4. EM Algorithm	
2. sensor uncertainties		5. Array callibration	
3. Maximum Likelihood			
b. Identifiers/Open-Ended Terms			
c. COSATI Field/Group			
18. Availability Statement Approved for publication; distribution unlimited.		19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 39
		20. Security Class (This Page)	22. Price