

# APPROXIMATE EXPRESSIONS FOR THE MEAN AND THE COVARIANCE OF THE MAXIMUM LIKELIHOOD ESTIMATOR FOR ACOUSTIC SOURCE LOCALIZATION

Vikas C. Raykar and Ramani Duraiswami

Perceptual Interfaces and Reality Lab., Institute for Advanced Computer Studies,  
Department of Computer Science, University of Maryland, CollegePark

## ABSTRACT

Acoustic source localization using multiple microphones can be formulated as a maximum likelihood estimation problem. The estimator is implicitly defined as the minimum of a certain objective function. As a result we cannot get explicit expressions for the mean and the covariance of the estimator. In this paper, we derive approximate expressions for the mean vector and covariance matrix of the estimator using Taylor's series expansion of the implicitly defined estimator. The validity of our expressions is verified by Monte-Carlo simulations. We also study the performance of the estimator for different microphone array configurations.

## 1. INTRODUCTION

The benefits that a microphone array provides over a single microphone are two fold. Using a microphone array we can localize a sound source accurately. Once the source location is known the microphone array can be electronically steered to the source. Broadly three types of methods exist for localizing a sound source [1]: Focalization using a steered beamformer, High resolution spectral estimation methods and Time Difference of Flight (TDOF) based methods. The most commonly used method in practice is the TDOF based method. In this method the signals received by several microphones are processed to estimate the time delays between pairs of microphones. Based on the estimated TDOFs the Maximum Likelihood (ML) estimator for the source location can be derived. Approximate closed form solutions do exist [2]. However the ML estimation is always used as a final step to refine the approximate source locations.

The performance of the ML estimator can be studied in terms of the bias vector and error covariance matrix. The bias and error covariance depends on the noise variance, the number of microphones, the geometry of the array, and the source position. Often for a given microphone array geometry we would like to see how well the estimator performs for different source locations. One way to study this is to do extensive Monte-Carlo Simulations for different source locations. However if we get an analytical expression for the bias and the covariance of the estimator then these studies can be carried out quickly and the estimator can be studied in depth.

The ML estimate for the source location is defined implicitly as the minimum of a certain objective function. Hence it is not possible to get exact analytical expressions for the mean vector and the covariance matrix. However, by using the implicit function theorem and the Taylor's series it is possible to derive approximate expressions for the mean and covariance matrix of implicitly defined estimators [3, 4]. Previous studies [5, 6] have analyzed the ML estimator performance by deriving the Crámer-Rao lower bound (CRLB). We could have derived the CRLB which gives the lower bound on the error covariance matrix of any *unbiased* estimator. However we cannot determine whether our estimator is unbiased.

The following are the novel contributions of this paper: (a) Unlike other approaches which derive the CRLB assuming that

the estimator is unbiased, we derive an approximate expression for both the bias and the error covariance matrix of the estimator. (In [7, 8] we used a similar approach in the context of microphone array position calibration.) (b) We study the performance of the estimator for different array geometries and show that the array performs best within the region enclosed by the microphones. The rest of the paper is organized as follows. In Sec. 2 we precisely state the estimation problem and derive the ML estimator in Sec. 3. In Sec. 4 we derive approximate expressions for the mean and the covariance of the estimator. In Sec. 5 we verify the validity of our expressions by Monte-Carlo simulations and study the estimator for different microphone array geometries.

## 2. PROBLEM STATEMENT

Let us say we have  $M$  microphones and one isotropic source. Let  $\mathbf{m}_i = [mx_i, my_i, mz_i]^T$  represent the three dimensional spatial coordinates of the  $i^{th}$  microphone, where  $mx_i$ ,  $my_i$ , and  $mz_i$  are it's coordinates. Let the source be located at  $\mathbf{s} = [sx, sy, sz]^T$ , where  $sx$ ,  $sy$ , and  $sz$  are the coordinates of the acoustic source. The Time Difference Of Flight (TDOF) for a given pair of microphones and the source is defined as the time difference between the signal received by the two microphones. The TDOF between the  $i^{th}$  and the  $j^{th}$  microphone is given by,

$$TDOF_{ij} = \frac{\|\mathbf{s} - \mathbf{m}_i\| - \|\mathbf{s} - \mathbf{m}_j\|}{c}, \quad (1)$$

where  $c$  the speed of sound and  $\|\mathbf{s} - \mathbf{m}_i\|$  is the Euclidean distance between the source and the  $i^{th}$  microphone. The speed of sound in a given acoustical medium is given by  $c = (331 + 0.6T)m/s$ , where  $T$  is the temperature of the medium in Celsius. In practice the TDOF cannot be estimated accurately. Let  $\widehat{TDOF}_{ij}$  be the estimated TDOF between the  $i^{th}$  and the  $j^{th}$  microphone based on the signal received by each of the microphones. Given  $M$  microphones, the set of  $M(M-1)/2$  TDOF measurements constitute our observations based on which we have to estimate the location of the acoustic source.

## 3. MAXIMUM LIKELIHOOD ESTIMATOR

Assuming an additive Gaussian noise model for the observations we can derive the Maximum Likelihood (ML) estimate for the source location as follows. Let  $\Theta$  be a column vector of length  $P$ , representing all the unknown non-random parameters to be estimated (the source coordinates  $sx$ ,  $sy$ , and  $sz$ ). Let  $\Gamma$  be a column vector of length  $N$ , representing the noisy measurements. Let  $T(\Theta)$  be a column vector of length  $N$ , representing the actual value of the observations. Then our model for the observations is  $\Gamma = T(\Theta) + \eta$ , where  $\eta$  is the zero-mean additive white Gaussian noise vector of length  $N$  where each element has variance  $\sigma_j^2$ . Also let us define  $\Sigma$  to be the  $N \times N$  covariance matrix of the noise vector  $\eta$ . The likelihood function (multivariate Gaussian in our case) of  $\Gamma$  in vector form can be written as:

$$p(\Gamma/\Theta) = (2\pi)^{-\frac{N}{2}} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}[\Gamma - T(\Theta)]^T \Sigma^{-1} [\Gamma - T(\Theta)]}. \quad (2)$$

The ML estimate of  $\Theta$  is the one which maximizes the likelihood ratio (or equivalently the log-likelihood ratio) and is given by

$$\hat{\Theta}_{ML} = \arg_{\Theta} \max F(\Theta, \Gamma)$$

$$F(\Theta, \Gamma) = -\frac{1}{2}[\Gamma - T(\Theta)]^T \Sigma^{-1} [\Gamma - T(\Theta)]. \quad (3)$$

Assuming that each of the measurements are independently corrupted by zero-mean additive white Gaussian noise of variance  $\sigma_j^2$  the ML estimate becomes a weighted nonlinear least squares problem. For independent noise components  $\Sigma$  is a diagonal matrix. Simplifying Eq. 3 we get

$$F(\Theta, \Gamma) = -\frac{1}{2} \sum_{j=1}^N \frac{[\Gamma_j - T_j(\Theta)]^2}{\sigma_j^2}. \quad (4)$$

In our case,  $\Theta$  represents the source coordinates,  $\Gamma$  corresponds to the estimated TDOF measurements and  $T(\Theta)$  corresponds to the actual TDOF. Substituting in Eq. 4 we get the following expression for the source location (removing the constant term and the negative sign and replacing the maximum by the minimum),

$$\hat{s}_{TDOF} = \arg_{\mathbf{s}} \min \sum_{i=1}^{M-1} \sum_{j=i+1}^M \frac{(TDOF_{ij} - TDOF_{ij})^2}{\sigma_{ij}^2}.$$

#### 4. MEAN AND COVARIANCE OF THE ML ESTIMATOR

The ML estimate is defined implicitly as the maximum of a certain error function (Eq. 4). Hence it is not possible to get exact analytical expressions for the mean and the covariance. In this section by using the implicit function theorem and the Taylor's series expansion we to derive approximate expressions for the mean and covariance matrix of the implicitly defined estimator [3, 4].

##### 4.1. Vector Derivatives

In further derivations we need the first, second, and third derivatives of the objective function in Eq. 3 (or Eq. 4) with respect to  $\Theta$  and  $\Gamma$ . The  $P \times 1$  column gradient operator  $\nabla_{\Theta}$  is defined as  $\nabla_{\Theta} F(\Theta, \Gamma) = [\frac{\partial F(\Theta, \Gamma)}{\partial \Theta_1}, \frac{\partial F(\Theta, \Gamma)}{\partial \Theta_2}, \dots, \frac{\partial F(\Theta, \Gamma)}{\partial \Theta_P}]^T$ , where (differentiating Eq. 4),  $\frac{\partial F(\Theta, \Gamma)}{\partial \Theta_k} = \sum_{j=1}^N \frac{\partial T_j(\Theta)}{\partial \Theta_k} \frac{1}{\sigma_j^2} [\Gamma_j - T_j(\Theta)]$ . Similarly the  $N \times 1$  column gradient operator  $\nabla_{\Gamma}$  with respect to  $\Gamma$  is defined as  $\nabla_{\Gamma} F(\Theta, \Gamma) = [\frac{\partial F(\Theta, \Gamma)}{\partial \Gamma_1}, \frac{\partial F(\Theta, \Gamma)}{\partial \Gamma_2}, \dots, \frac{\partial F(\Theta, \Gamma)}{\partial \Gamma_N}]^T$ , where,  $\frac{\partial F(\Theta, \Gamma)}{\partial \Gamma_n} = -\frac{[\Gamma_n - T_n(\Theta)]}{\sigma_n^2}$ . In the vector notation the first derivatives can be conveniently written as,

$$\nabla_{\Theta} F(\Theta, \Gamma) = J^T \Sigma^{-1} [\Gamma - T(\Theta)],$$

$$\nabla_{\Gamma} F(\Theta, \Gamma) = -\Sigma^{-1} [\Gamma - T(\Theta)]. \quad (5)$$

where  $J$  is a  $N \times P$  matrix of partial derivatives of  $T(\Theta)$  called the *Jacobian* of  $T(\Theta)$ .  $[J]_{ij} = \frac{\partial T_i(\Theta)}{\partial \Theta_j}$ . The individual second derivatives are as follows,

$$\frac{\partial^2 F(\Theta, \Gamma)}{\partial \Theta_i \partial \Theta_k} = -\sum_{j=1}^N \frac{1}{\sigma_j^2} \frac{\partial T_j(\Theta)}{\partial \Theta_k} \frac{\partial T_j(\Theta)}{\partial \Theta_i} + \frac{1}{\sigma_j^2} [\Gamma_j - T_j(\Theta)] \frac{\partial^2 T_j(\Theta)}{\partial \Theta_i \partial \Theta_k}.$$

$$\frac{\partial^2 F(\Theta, \Gamma)}{\partial \Gamma_m \partial \Gamma_n} = \begin{cases} 0, & \text{if } m \neq n; \\ -\frac{1}{\sigma_n^2}, & \text{if } m = n. \end{cases}$$

$$\frac{\partial^2 F(\Theta, \Gamma)}{\partial \Gamma_n \partial \Theta_k} = \frac{\partial^2 F(\Theta, \Gamma)}{\partial \Theta_k \partial \Gamma_n} = \frac{1}{\sigma_n^2} \frac{\partial T_n(\Theta)}{\partial \Theta_k}. \quad (6)$$

In our derivations we only need the second derivatives evaluated at  $\Gamma = \bar{\Gamma} = T(\Theta)$  where,  $\bar{\Gamma}$  is the mean of  $\Gamma$ . Evaluating

Eqs. 6 at  $\Gamma_j = T_j(\Theta)$  the second derivatives can be conveniently written in the matrix form as follows,

$$\nabla_{\Theta} \nabla_{\Theta} F(\Theta, \Gamma)|_{\Gamma=\bar{\Gamma}} = -J^T \Sigma^{-1} J.$$

$$\nabla_{\Gamma} \nabla_{\Gamma} F(\Theta, \Gamma)|_{\Gamma=\bar{\Gamma}} = -\Sigma^{-1}.$$

$$\nabla_{\Gamma} \nabla_{\Theta} F(\Theta, \Gamma)|_{\Gamma=\bar{\Gamma}} = \Sigma^{-1} J.$$

$$\nabla_{\Theta} \nabla_{\Gamma} F(\Theta, \Gamma)|_{\Gamma=\bar{\Gamma}} = J^T \Sigma^{-1}. \quad (7)$$

The following are the third derivatives of the objective function which we will be using in our derivation of the mean vector.

$$\frac{\partial^3 F(\Theta, \Gamma)}{\partial \Theta_m \partial \Theta_i \partial \Theta_k} = \sum_{j=1}^N -\frac{1}{\sigma_j^2} \left[ \frac{\partial T_j(\Theta)}{\partial \Theta_k} \frac{\partial^2 T_j(\Theta)}{\partial \Theta_m \partial \Theta_i} + \frac{\partial T_j(\Theta)}{\partial \Theta_i} \frac{\partial^2 T_j(\Theta)}{\partial \Theta_m \partial \Theta_k} + \frac{\partial T_j(\Theta)}{\partial \Theta_m} \frac{\partial^2 T_j(\Theta)}{\partial \Theta_i \partial \Theta_k} \right] + \frac{[\Gamma_j - T_j(\Theta)]}{\sigma_j^2} \frac{\partial^3 T_j(\Theta)}{\partial \Theta_m \partial \Theta_i \partial \Theta_k},$$

$$\frac{\partial^3 F(\Theta, \Gamma)}{\partial \Theta_l \partial \Theta_k \partial \Gamma_n} = \frac{1}{\sigma_n^2} \frac{\partial^2 T_n(\Theta)}{\partial \Theta_l \partial \Theta_k}, \quad \frac{\partial^3 F(\Theta, \Gamma)}{\partial \Theta_k \partial \Gamma_n^2} = 0. \quad (8)$$

Unlike the first and the second derivatives Eqs. 8 cannot be written in a much simpler form without invoking tensor notation. All individual derivatives are listed in Appendix I.

##### 4.2. Estimator Covariance

The ML estimate is the one which maximizes the objective function  $F(\Theta, \Gamma)$  defined in Eq. 3. Assuming that the objective function has a unique global maximum, the maximum can be found by setting the first derivative to zero, i.e.,  $\nabla_{\Theta} F(\Theta, \Gamma)|_{\Theta=\hat{\Theta}_{ML}} = \mathbf{0}$ , where  $\mathbf{0}$  is a zero column vector of length  $P$ . The implicit function theorem guarantees that this implicitly defines a vector valued function  $\hat{\Theta}_{ML} = h(\Gamma) = [h_1(\Gamma), h_2(\Gamma), \dots, h_P(\Gamma)]^T$  that maps the observation vector  $\Gamma$  to the parameter vector  $\hat{\Theta}_{ML}$ . Therefore,

$$\nabla_{\Theta} F(\Theta, \Gamma)|_{\Theta=h(\Gamma)} = \nabla_{\Theta} F(h(\Gamma), \Gamma) = \mathbf{0}. \quad (9)$$

It is not possible to find an analytical expression for  $h(\Gamma)$ , but we can approximate the covariance using the first-order Taylor's series expansion for  $h(\Gamma)$ . Let  $\bar{\Gamma}$  be the mean of  $\Gamma$ . Then the first-order Taylor's series expansion for  $h(\Gamma)$  around  $\bar{\Gamma}$  gives,

$$h(\Gamma) \approx h(\bar{\Gamma}) + [\nabla_{\Gamma} h(\Gamma)]^T |_{\Gamma=\bar{\Gamma}} (\Gamma - \bar{\Gamma}). \quad (10)$$

Taking the covariance on both sides yields

$$Cov[h(\Gamma)] \approx [\nabla_{\Gamma} h(\Gamma)]^T |_{\Gamma=\bar{\Gamma}} Cov(\Gamma) [\nabla_{\Gamma} h(\Gamma)]^T |_{\Gamma=\bar{\Gamma}}. \quad (11)$$

Note we do not know  $h(\Gamma)$ , but the dependence is only through the first-order partial derivatives of  $h(\Gamma)$ . Differentiating Eq. 9 with respect to  $\Gamma$  and evaluating at  $\bar{\Gamma}$  yields

$$\nabla_{\Theta} \nabla_{\Theta} F(h(\bar{\Gamma}), \bar{\Gamma}) [\nabla_{\Gamma} h(\bar{\Gamma})]^T + \nabla_{\Theta} \nabla_{\Gamma} F(h(\bar{\Gamma}), \bar{\Gamma}) = \mathbf{0}. \quad (12)$$

Assuming that  $\nabla_{\Theta} \nabla_{\Theta} F(h(\bar{\Gamma}), \bar{\Gamma})$  is invertible we can write  $[\nabla_{\Gamma} h(\bar{\Gamma})]^T = -[\nabla_{\Theta} \nabla_{\Theta} F(h(\bar{\Gamma}), \bar{\Gamma})]^{-1} \nabla_{\Theta} \nabla_{\Gamma} F(h(\bar{\Gamma}), \bar{\Gamma})$ . Substituting for the derivatives form Eq. 7 we get,  $[\nabla_{\Gamma} h(\bar{\Gamma})]^T = -[-J^T \Sigma^{-1} J]^{-1} J^T \Sigma^{-1}$ . Using this in the covariance expression in Eq. 11, we arrive at the following expression for the covariance of the estimator,

$$Cov \hat{\Theta} = Cov[h(\Gamma)] = [J^T \Sigma^{-1} J]^{-1}. \quad (13)$$

If we assume all the observations have the same variance  $\sigma^2$ , i.e.,  $\Sigma = \sigma^2 I$ , we get  $Cov \hat{\Theta} = \sigma^2 [J^T J]^{-1}$ .

### 4.3. Estimator Mean

Taking the expectation of the first order Taylor series expansion in Eq. 10, we get  $E[h(\Gamma)] \approx h(\bar{\Gamma}) = h(T(\Theta))$ . We made use of the fact that  $\bar{\Gamma} = T(\Theta)$ . We see that the mean is the value given by the estimation procedure when applied to the actual noise free measurements  $T(\Theta)$ . A more accurate expression for the mean can be derived using the second order Taylor's series expansion. The second-order Taylor's series expansion for  $h(\Gamma)$  around  $\bar{\Gamma}$  can be written as,  $h(\Gamma) \approx h(\bar{\Gamma}) + \sum_n \frac{\partial h(\bar{\Gamma})}{\partial \Gamma_n} (\Gamma_n - \bar{\Gamma}_n) + \frac{1}{2} \sum_n \sum_m \frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n \partial \Gamma_m} (\Gamma_n - \bar{\Gamma}_n)(\Gamma_m - \bar{\Gamma}_m)$ . Taking the expectation on both sides of the truncated Taylor's series expansion we get,  $E[h(\Gamma)] \approx h(T(\Theta)) + \frac{1}{2} \sum_n \sum_m \frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n \partial \Gamma_m} Cov(\Gamma_n, \Gamma_m)$ . Since the observations are independent, this simplifies to,

$$E[h(\Gamma)] \approx h(T(\Theta)) + \frac{1}{2} \sum_n \frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n^2} \sigma_n^2 \quad (14)$$

To evaluate,  $\frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n^2}$  we need to differentiate Eq. 12 with respect to  $\Gamma$ . Eq. 12 on a term by term basis can be written as follows,  $\sum_k \frac{\partial^2 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k} \frac{\partial h_k(\bar{\Gamma})}{\partial \Gamma_n} + \frac{\partial^2 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Gamma_n} = 0, j = 1, \dots, P, n = 1, \dots, N$ . Differentiating this with respect to  $\Gamma_n$ ,

$$\sum_k \left[ \sum_l \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Theta_l} \frac{\partial h_l(\bar{\Gamma})}{\partial \Gamma_n} + 2 \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Gamma_n} \right] \frac{\partial h_k(\bar{\Gamma})}{\partial \Gamma_n} + \sum_k \frac{\partial^2 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k} \frac{\partial^2 h_k(\bar{\Gamma})}{\partial \Gamma_n^2} + \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Gamma_n^2} = 0, \quad (15)$$

$j = 1, \dots, P, n = 1, \dots, N$ . From Eq. 8 we have  $\frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Gamma_n^2} = 0$ . Eq. 15 can be written as,

$$\sum_k \frac{\partial^2 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k} \frac{\partial^2 h_k(\bar{\Gamma})}{\partial \Gamma_n^2} = \nabla_{\Theta} \nabla_{\Theta} F(\Theta, \Gamma)|_{\Gamma=\bar{\Gamma}} \frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n^2} - \sum_k \left[ \sum_l \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Theta_l} \frac{\partial h_l(\bar{\Gamma})}{\partial \Gamma_n} + 2 \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Gamma_n} \right] \frac{\partial h_k(\bar{\Gamma})}{\partial \Gamma_n}. \quad (16)$$

Substituting from Eq. 7,  $\frac{\partial^2 h(\bar{\Gamma})}{\partial \Gamma_n^2} = [J^T \Sigma^{-1} J]^{-1} \mathbf{a}_n$  where  $\mathbf{a}_n$  is  $P \times 1$  vector where the  $j^{th}$  element is given by the following,  $[\mathbf{a}_n]_j = \sum_k \left[ \sum_l \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Theta_l} \frac{\partial h_l(\bar{\Gamma})}{\partial \Gamma_n} + 2 \frac{\partial^3 F(h(\bar{\Gamma}), \bar{\Gamma})}{\partial \Theta_j \partial \Theta_k \partial \Gamma_n} \right] \frac{\partial h_k(\bar{\Gamma})}{\partial \Gamma_n}$ . The expression for the mean can now be written as,  $E[h(\Gamma)] \approx h(T(\Theta)) + \frac{1}{2} [J^T \Sigma^{-1} J]^{-1} \sum_n \mathbf{a}_n \sigma_n^2$ . If  $A = [\mathbf{a}_1 \dots \mathbf{a}_N]$  is a  $P \times N$  matrix and  $diag(\Sigma)$  is a  $N \times 1$  vector consisting of the diagonal elements of  $\Sigma$ , then

$$E[h(\Gamma)] \approx h(T(\Theta)) + \frac{1}{2} [J^T \Sigma^{-1} J]^{-1} A \text{diag}(\Sigma). \quad (17)$$

If we assume all the observations have the same variance  $\sigma^2$ , i.e.,  $\Sigma = \sigma^2 I$ , we get  $E[h(\Gamma)] \approx h(T(\Theta)) + 0.5 \sigma^4 [J^T J]^{-1} A \mathbf{1}$  where  $\mathbf{1}$  is a column vector of ones of length  $N$ .

## 5. DISCUSSION

### 5.1. Monte-Carlo Simulations

We perform Monte-Carlo simulations in order to verify the validity of the approximate expressions for the mean vector and the covariance matrix of the estimator. Fig. 1(a) shows a configuration of 4

microphones (solid squares) distributed in a  $4m \times 4m$  square<sup>1</sup>. For the given configuration the actual TDOF was computed and corrupted with zero mean additive gaussian noise of standard deviation  $\sigma = 10^{-4}$  (0.1 milliseconds, around 3.5 cm). The source location was estimated using the Levenberg-Marquardt method for solving non-linear least square problems. This was repeated for 200 trials for different source locations. The estimated source locations are plotted as dots in Fig. 1(a). The 95 % uncertainty ellipses derived from the theoretical expression for the covariance matrix are also plotted around the theoretical mean. The bias vector is indicated as a line from the actual source position to the mean source position. As can be seen the estimated source locations lie within the theoretical 95 % uncertainty ellipses. The bias is negligible to be clearly seen in the figure.

### 5.2. Effect of Microphone Array Geometry

Fig. 1(b)-(d) shows the bias vector and the covariance ellipses for different microphone array configurations. First, we note that as the number of microphones increases the variance and the bias decreases (compare Fig. 1(b) and (c)). This is because as the number of microphones increases ( $O(n)$ ) the number of observations also increases ( $O(n^2)$ ). The estimator performs best within the region bounded by the microphones. The area of the uncertainty ellipses increases as we move further away from the microphone array. The particular orientation is dictated by the geometrical configuration of the microphone array. For a given pair of microphones and a given TDOF, Eq. 1 represents one half of a hyperbola, with  $\frac{\mathbf{m}_i + \mathbf{m}_j}{2}$  as the center with  $\mathbf{m}_i$  and  $\mathbf{m}_j$  being the two focal points and the line joining the two microphones as the axis of symmetry. The actual source location is obtained as the intersection of the hyperbolas in a least square sense. Any error in the estimated TDOF, will cause a error in the intersection whose shape is dictated by the microphone array geometry. Fig. 1(f) show the shape of the error region as an intersection of two hyperbolas (error of 4 cm), corresponding to 2 TDOF pairs for 3 microphones, for two different source positions (one within the region bounded by the 3 microphones and the other outside). For the source position within the region bounded by the microphones the error region is much smaller compared to that of the source far away from the microphone array. In all the plots the bias is indicated as a line from the actual source position to the theoretical mean position. The bias is negligible to be clearly seen in the figure. Fig. 1(e) shows only the bias vectors for a 4 microphone array.

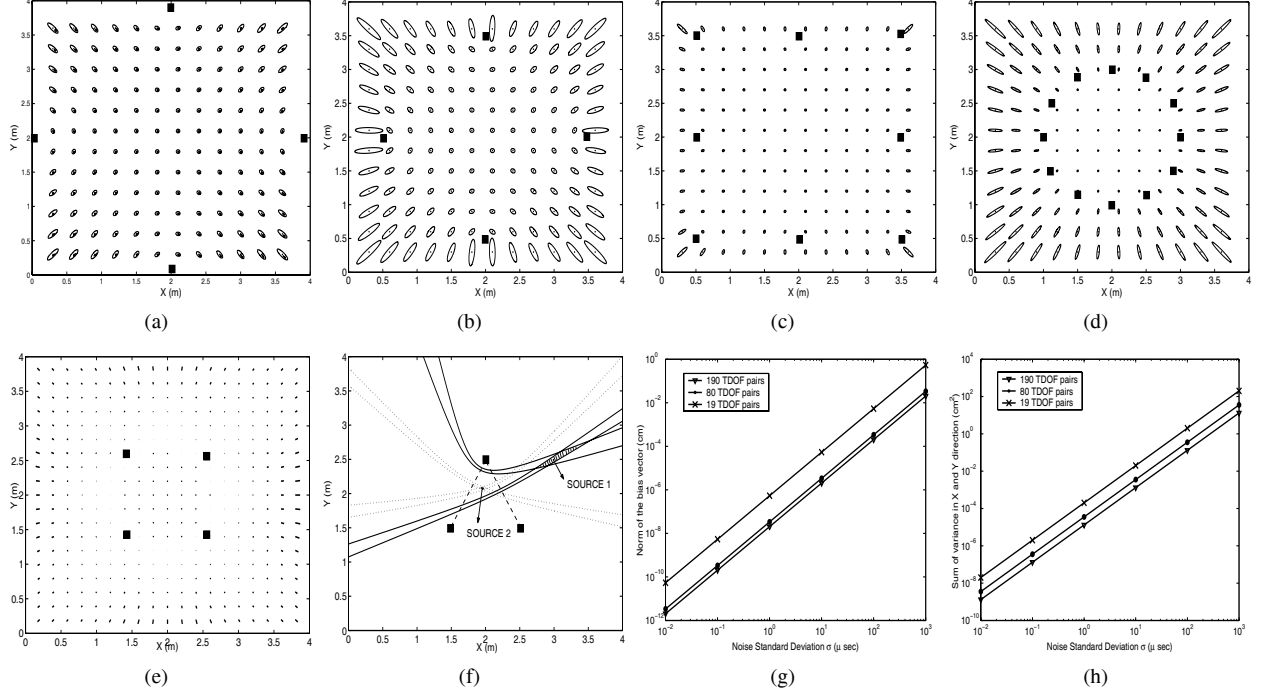
### 5.3. Effect of Number of TDOF pairs

Given  $M$  microphones we can have  $M(M-1)/2$  measurements. Of these only  $M-1$  are linearly independent. We need not use all the TDOF pairs in our estimation procedure. Fig. 1(g) and (h) shows the norm of the bias vector and the sum of the variances in the  $x$  and  $y$  direction as a function of the noise standard deviation  $\sigma$ , for different number of microphone pairs used in our estimation procedure. The graph is for a source location situated at the center of a  $4m \times 4m$  room consisting of 20 microphones distributed randomly. The variance and the bias decrease as the number of microphone pairs are increased.

## 6. CONCLUSION

In this paper we derived an approximate expression for the mean vector and the covariance matrix of the ML estimator for acoustic

<sup>1</sup>For illustration purpose and reasons of space we consider all our microphones to be distributed in two dimensions. Similar results apply in three dimensions.



**Fig. 1.** [Noise standard deviation for all cases is  $\sigma = 10^{-4}$  (0.1 milliseconds, around 3.5 cm). The microphones are represented as filled squares. The actual source positions are marked as dots.] (a) 95% uncertainty covariance ellipses for different source locations for a 4 microphone array. The results of 200 Monte-Carlo trials are shown as dots. (b)-(d) 95% uncertainty covariance ellipses and the bias vector for the following different microphone array configurations: (b) 4 microphones one near each wall, (c) 8 microphones forming a square, and (d) 12 microphones approximately arranged as a circle. The bias is indicated as a line from the actual source position to the theoretical mean position. The bias is negligible to be clearly seen in the figures. (e) shows only the bias vectors for a 4 microphone array. (f) Error region as an intersection of two hyperbolas (error of 4 cm), corresponding to 2 TDOF pairs for 3 microphones, for two different source positions. (g) The norm of the bias vector and (h) the sum of the variances in the  $x$  and  $y$  direction as a function of the noise standard deviation  $\sigma$ , for different number of microphone pairs. The source is situated at the center of a  $4m \times 4m$  room consisting of 20 microphones.

source localization. A microphone array designer can quickly verify how his array performs for different source locations. In this paper we dealt with only one acoustic source. Future work would involve the same analysis with multiple acoustic sources.

## 7. APPENDIX I

$$\frac{\partial TDOF_{ij}}{\partial sx} = \frac{(sx - mx_i)}{c\|s - \mathbf{m}_i\|} - \frac{(sx - mx_j)}{c\|s - \mathbf{m}_j\|}$$

$$\frac{\partial^2 TDOF_{ij}}{\partial sx^2} = -\frac{(sx - mx_i)^2}{c\|s - \mathbf{m}_i\|^3} + \frac{(sx - mx_j)^2}{c\|s - \mathbf{m}_j\|^3} + \frac{1}{c\|s - \mathbf{m}_i\|} - \frac{1}{c\|s - \mathbf{m}_j\|}$$

$$\frac{\partial^2 TDOF_{ij}}{\partial sy \partial sx} = \frac{\partial^2 TDOF_{ij}}{\partial sx \partial sy} = -\frac{(sx - mx_i)(sy - my_i)}{c\|s - \mathbf{m}_i\|^3} + \frac{(sx - mx_j)(sy - my_j)}{c\|s - \mathbf{m}_j\|^3}$$

Similar expressions can be derived with respect to the source location  $sy$  and  $sz$ .

## 8. REFERENCES

- [1] M. Brandstein and D. Ward, *Microphone arrays-Signal Processing Techniques and Applications*, Springer-Verlag, 2001.
- [2] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 12, pp. 1661-1669, 1987.
- [3] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Trans. on Image Processing*, vol. 5, no. 3, pp. 493-506, March 1996.
- [4] A. R. Chowdhury and R. Chellappa, "Stochastic approximation and rate distortion analysis for robust structure and motion estimation," *International Journal of Computer Vision*, vol. 55, no. 1, pp. 27-53, October 2003.
- [5] M.J.D. Rendas and J.M.F. Moura, "Cramer-rao bound for location systems in multipath environments," *IEEE Transactions on Signal Processing*, vol. 39, no. 12, pp. 2593-2610, 1991.
- [6] X. Sheng and Hu Yu-Hen, "Energy based acoustic source localization," in *The 2nd International Workshop on Information Processing in Sensor Networks*, Palo Alto, CA, 2003, pp. 285-300.
- [7] V. C. Raykar, I. Kozintsev, and R. Lienhart, "Position calibration of microphones and loudspeakers in distributed computing platforms," *IEEE Transactions on Speech and Audio Processing (to appear)*.
- [8] V. C. Raykar and R. Duraiswami, "Automatic position calibration of multiple microphones," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, Montreal, Canada, May 2004, vol. IV, pp. 69-72.