CRAMER-RAO BOUND AND OPTIMUM SENSOR ARRAY
FOR SOURCE LOCALIZATION FROM TIME DIFFERENCES OF ARRIVAL

B. Yang and J. Scheuing

Chair of System Theory and Signal Processing
University of Stuttgart, Germany

ABSTRACT

This paper presents a theoretical analysis of the Cramer-Rao lower bound for source localization from time differences of arrival. We derive properties of the Cramer-Rao bound and design optimum sensor arrays which minimize the bound.

1. INTRODUCTION

The technique of source localization by using time differences of arrival (TDOA) plays an important role in many applications like navigation, localization and tracking of acoustic sources, and location services of mobile communication. In the literature, the main research effort was focused on the development of various methods for the source position estimation. A recent review of some methods can be found in [1]. In comparison, relatively less is known about the theoretical performance of these methods. There exist many numerical comparisons between different methods, but the impact of the sensor array geometry on the localization accuracy is still not well understood yet.

This paper aims at a better understanding of the theoretical performance of source localization. We discuss properties of the Cramer-Rao bound for TDOA based localization. We derive necessary and sufficient conditions for the optimum sensor array geometry which minimizes the bound. Examples of optimum arrays for both two-dimensional and three-dimensional localization are given.

The following notations are used in the paper. Matrices and column vectors are represented by boldface and underlined characters. The superscript ^{T} denotes transpose. $\|\cdot\|$ is the Euclidean vector norm. $\mathrm{tr}(\cdot)$ is the matrix trace.

2. PROBLEM FORMULATION

We consider the source localization problem in $\mathbb{R}^D$ ($D = 2, 3$). The sensor array consists of M sensors at the locations $\underline{q}_i \in \mathbb{R}^D$ ($i = 1, \ldots, M$). The source position is $\underline{p} \in \mathbb{R}^D$. The distance between the source and sensor i is $d_i(\underline{p}) = \|\underline{p} - \underline{q}_i\|$ ($i = 1, \ldots, M$). The difference in the distances of sensor i and j from the source is given by $d_{ij}(\underline{p}) = d_i(\underline{p}) - d_j(\underline{p})$. The noisy TDOA measurement between sensor i and j is

thus

$$\tau_{ij} = \frac{1}{v} d_{ij}(\underline{p}) + n_{ij}$$

v is the wave propagation speed and n_{ij} is the TDOA measurement error.

For M sensors, there are a total number of $M(M-1)/2$ possible sensor pairs and TDOA measurements. Let

$$\mathcal{I}_0 = \{(i, j) | 1 \le j < i \le M\} \tag{1}$$

denote the set of all sensor pairs. $\mathcal{I}$ is a subset of $\mathcal{I}_0$ and contains these N ($N \le \frac{M(M-1)}{2}$) sensor pairs whose TDOA measurements are used in source localization. By introducing the $N \times 1$ vectors

$$\underline{\tau} = \left[\begin{array}{c} \tau_{ij} \\ \vdots \end{array} \right]_{(i,j) \in \mathcal{I}}, \underline{d} = \left[\begin{array}{c} d_{ij} \\ \vdots \end{array} \right]_{(i,j) \in \mathcal{I}}, \underline{n} = \left[\begin{array}{c} n_{ij} \\ \vdots \end{array} \right]_{(i,j) \in \mathcal{I}},$$

the signal model for the TDOA measurements becomes

$$\underline{\tau} = \frac{1}{v} \underline{d}(\underline{p}) + \underline{n} \tag{2}$$

The problem of source localization is to estimate the source position vector $\underline{p}$ given $\underline{q}_i, \underline{\tau}$, and v.

3. CRAMER-RAO BOUND

The Cramer-Rao bound (CRB) is a lower bound for the covariance matrix of unbiased estimators. It is often used as a benchmark against which the efficiency of unbiased estimators is tested. It is given by $\mathbf{J}^{-1}$ where

$$\mathbf{J} = \mathrm{E} \left\{ [\underline{\nabla} \ln f(\underline{\tau}; \underline{p})][\underline{\nabla} \ln f(\underline{\tau}; \underline{p})]^T \right\}$$

is the Fisher information matrix. $f(\underline{\tau}; \underline{p})$ is the probability density function (PDF) of $\underline{\tau}$. $\underline{\nabla}$ is the gradient operator with respect to $\underline{p}$. $\mathrm{E}(\cdot)$ denotes the expectation on $\underline{\tau}$.

Assume that the measurement error vector $\underline{n}$ in (2) is Gaussian with zero mean and the full rank covariance matrix $\mathbf{C}$ which is independent of $\underline{p}$. The PDF of $\underline{\tau}$ is

$$f(\underline{\tau}; \underline{p}) = \frac{1}{(2\pi)^{N/2} \sqrt{\det(\mathbf{C})}} \exp \left[-\frac{1}{2} (\underline{\tau} - \frac{1}{v} \underline{d})^T \mathbf{C}^{-1} (\underline{\tau} - \frac{1}{v} \underline{d}) \right] \tag{3}$$

The CRB is well known from [2]

$$\mathbf{J}^{-1} = v^2(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}^T)^{-1} \qquad (4)$$

with

$$
\begin{aligned}
\mathbf{G} &= \nabla \underline{d}^T = [\underline{g}_{ij} \ \ldots]_{(i,j)\in\mathcal{I}}, \\
\underline{g}_{ij} &= \underline{g}_i - \underline{g}_j, \\
\underline{g}_i &= \nabla d_i(\underline{p}) = \frac{\underline{p} - \underline{q}_i}{\|\underline{p} - \underline{q}_i\|} \quad (i = 1, \ldots, M)
\end{aligned}
\qquad (5)
$$

Clearly, $\underline{g}_i$ is a unit-length vector with $\|\underline{g}_i\| = 1$. It points from sensor i to the source. $\mathbf{G}$ is a $D \times N$ matrix. It depends on the source and sensor positions and on the set $\mathcal{I}$ of sensor pairs which are used for source localization.

4. PROPERTIES OF THE CRAMER-RAO BOUND

In this section, we discuss some properties of the CRB.

Prop. 1: One necessary condition for the existence of the CRB is $M \geq D + 1$.
Proof: For the existence of the $D \times D$ inverse Fisher information matrix in (4), the $D \times N$ matrix $\mathbf{G}$ must have rank D. Due to the definition (5), any vectors $\underline{g}_{ij}$ can be written as a linear combination of the $M - 1$ vectors $\underline{g}_{i1}$ $(i = 2, \ldots, M)$: $\underline{g}_{ij} = \underline{g}_{i1} - \underline{g}_{j1}$. Therefore, the maximum number of linearly independent columns in $\mathbf{G}$ is $M - 1$. This means

$$M - 1 \geq \text{rank}(\mathbf{G}) = D$$

Prop. 2: Under the Gaussian error model (3), no unbiased estimators attain the CRB.
Proof: A necessary and efficient condition for an unbiased efficient estimator is the existence of some functions $\mathbf{J}(\cdot)$ and $\underline{h}(\cdot)$ such that [3]

$$\nabla \ln f(\underline{\tau}; \underline{p}) = \mathbf{J}(\underline{p})[\underline{h}(\underline{\tau}) - \underline{p}] \qquad (6)$$

Starting from (3), the gradient vector is calculated to

$$\nabla \ln f(\underline{\tau}; \underline{p}) = \frac{1}{v}\mathbf{G}(\underline{p})\mathbf{C}^{-1}\left[\underline{\tau} - \frac{1}{v}\underline{d}(\underline{p})\right]$$

Since the condition (6) can never be satisfied, there are no efficient estimators for the source position vector $\underline{p}$. Nevertheless, the maximum-likelihood (ML) estimator

$$\hat{\underline{p}}_{ML} = \arg\max_{\underline{p}} f(\underline{\tau}; \underline{p})$$

is able to attain the CRB asymptotically ($N \to \infty$).

One possibility to find the ML estimator is to use a Gauss-Newton iteration [4]. Let

$$
\begin{aligned}
\underline{d}(\underline{p}) &\approx \underline{d}(\underline{p}_k) + \nabla^T \underline{d}(\underline{p})\big|_{\underline{p}=\underline{p}_k} (\underline{p} - \underline{p}_k) \\
&= \underline{d}(\underline{p}_k) + \mathbf{G}^T(\underline{p}_k)(\underline{p} - \underline{p}_k)
\end{aligned}
\qquad (7)
$$

be the truncated Taylor series of $\underline{d}(\underline{p})$ around the solution $\underline{p}_k$ found at the k-th iteration. The quadratic order and all higher-order terms of the Taylor series are ignored. Maximizing the likelihood $f(\underline{\tau}; \underline{p})$ in (3) while using the approximation (7) results in an improved estimate

$$
\begin{aligned}
\underline{p}_{k+1} = \underline{p}_k &+ [\mathbf{G}(\underline{p}_k)\mathbf{C}^{-1}\mathbf{G}^T(\underline{p}_k)]^{-1} \cdot \\
&\mathbf{G}(\underline{p}_k)\mathbf{C}^{-1}[\underline{\tau}v - \underline{d}(\underline{p}_k)]
\end{aligned}
\qquad (8)
$$

The key issue of this method is to find a good initial guess $\underline{p}_0$ to avoid local minima and divergence.

Prop. 3: The more TDOA measurements we use, the smaller the CRB is. Let $\mathcal{I}_i$ $(i = 1, 2)$ be two sets of TDOA measurements and $\mathbf{J}_{\mathcal{I}_i}^{-1}$ the corresponding CRBs. If $\mathcal{I}_1 \subset \mathcal{I}_2$, then $\mathbf{J}_{\mathcal{I}_1}^{-1} \geq \mathbf{J}_{\mathcal{I}_2}^{-1}$, i.e. $\mathbf{J}_{\mathcal{I}_1}^{-1} - \mathbf{J}_{\mathcal{I}_2}^{-1}$ is non-negative definite.
Proof: The assumption $\mathcal{I}_1 \subset \mathcal{I}_2$ implies that all columns of $\mathbf{G}_{\mathcal{I}_1} = [\underline{g}_{ij} \ \ldots]_{(i,j)\in\mathcal{I}_1}$ are also contained in $\mathbf{G}_{\mathcal{I}_2} = [\underline{g}_{ij} \ \ldots]_{(i,j)\in\mathcal{I}_2}$. This means $\mathbf{G}_{\mathcal{I}_1}\mathbf{C}^{-1}\mathbf{G}_{\mathcal{I}_1}^T \leq \mathbf{G}_{\mathcal{I}_2}\mathbf{C}^{-1}\mathbf{G}_{\mathcal{I}_2}^T$ and $\mathbf{J}_{\mathcal{I}_1}^{-1} \geq \mathbf{J}_{\mathcal{I}_2}^{-1}$. This property holds even for correlated TDOA measurements.

In the literature, many source localization methods rely on the spherical model instead of the hyperbolic model. They are spherical intersection [5], spherical interpolation [6] and further improvements of them [2, 1]. These methods are non-iterative and computationally efficient. However, they suffer from the drawback that maximum $M - 1$ sensor pairs $\mathcal{I} = \{(2, 1), \ldots, (M, 1)\}$ (with sensor 1 being the reference sensor) can be used. This limits the accuracy of the source localization. According to Prop. 3, the CRB of these methods is larger than the CRB when we use all $M(M - 1)/2$ sensor pairs.

Therefore, the following two-step procedure is suitable to improve the localization accuracy:

- Use a non-iterative method like any of the above spherical methods to find a reliable initial guess $\underline{p}_0$ for the source position.

- Use one or a few Gauss-Newton-iterations in (8), based on all sensor pairs $\mathcal{I}_0$, to further improve the localization accuracy.

Prop. 4: The CRB does not depend on the range $\|\underline{p} - \underline{q}_i\|$ between the source and sensors. It only depends on the direction of the vectors $\underline{p} - \underline{q}_i$.
Proof: This follows immediately from (5).

Note that this is a property of the CRB. The covariance matrices of source position estimators will depend on both range and direction of $\underline{p} - \underline{q}_i$.

5. OPTIMUM ARRAY GEOMETRY

The geometry of the source localization problem is determined by the sensor and source positions $\underline{q}_i$ and $\underline{p}$. Starting

from the CRB in (4), we now study how to optimize the geometry to minimize the CRB for a given number of sensors. This problem has been addressed in [7, 8] for the narrowband far-field direction-of-arrival (DOA) estimation. To our knowledge, there are no similar works for the TDOA based source localization which is usually a broadband near-field estimation problem.

Since the CRB is a square matrix, we consider the minimization of the trace of the CRB $\text{tr}(\mathbf{J}^{-1})$. It is a lower bound for the sum of variances of unbiased estimators for all elements of the sensor position vector $\underline{p}$. For simplicity, we consider white measurement noise with $\mathbf{C} = \sigma^2 \mathbf{I}$ in this paper. $\mathbf{I}$ is an identity matrix and σ^2 is the variance of the TDOA measurement noise. According to Prop. 3, we have to use the set $\mathcal{I}_0$ of all $M(M-1)/2$ sensor pairs in order to achieve a CRB as small as possible. This means

$$\mathbf{G} = [\underline{g}_{21} \cdots \underline{g}_{M1} \, \underline{g}_{32} \cdots \underline{g}_{M2} \cdots \underline{g}_{M,M-1}] \quad (9)$$

The optimization problem becomes to find M unit-length vectors $\underline{g}_i \in \mathbb{R}^D$ $(i = 1, \ldots, M)$ in such a way that

$$\text{tr}(\mathbf{J}^{-1}) = (v\sigma)^2 \text{tr}[(\mathbf{GG}^T)^{-1}]$$

is minimized.

Theorem 1: If $\mathbf{C} = \sigma^2 \mathbf{I}$ and $\mathcal{I} = \mathcal{I}_0$,

$$\text{tr}(\mathbf{J}^{-1}) \geq (v\sigma)^2 \frac{D^2}{M^2} \quad (10)$$

The equality holds if and only if

C1) $\sum_{i=1}^{M} \underline{g}_i = \underline{0}$ and

C2) the $D \times M$ matrix $\mathbf{g} = [\underline{g}_1 \cdots \underline{g}_M]$ satisfies $\mathbf{gg}^T = (M/D)\mathbf{I}$, i.e. $\mathbf{g}$ has orthogonal row vectors with equal row norm.

Proof: The proof consists of two parts. We first find a lower bound for $\text{tr}[(\mathbf{GG}^T)^{-1}]$ in terms of $\text{tr}(\mathbf{GG}^T)$ because $\text{tr}(\mathbf{GG}^T)$ is simpler to calculate than $\text{tr}[(\mathbf{GG}^T)^{-1}]$. Then we maximize $\text{tr}(\mathbf{GG}^T)$.

Let $\lambda_i > 0$ $(i = 1, \ldots, D)$ be the eigenvalues of the $D \times D$ symmetric and positive definite matrix $\mathbf{GG}^T$. The eigenvalues of $(\mathbf{GG}^T)^{-1}$ are $1/\lambda_i$. It is well known

$$\text{tr}(\mathbf{GG}^T) = \sum_{i=1}^{D} \lambda_i, \quad \text{tr}[(\mathbf{GG}^T)^{-1}] = \sum_{i=1}^{D} 1/\lambda_i$$

According to the Cauchy-Schwarz inequality,

$$D = \begin{bmatrix} \sqrt{\lambda_1} \\ \vdots \\ \sqrt{\lambda_D} \end{bmatrix}^T \begin{bmatrix} 1/\sqrt{\lambda_1} \\ \vdots \\ 1/\sqrt{\lambda_D} \end{bmatrix} \leq \sqrt{\sum_{i=1}^{D} \lambda_i \cdot \sum_{i=1}^{D} 1/\lambda_i}$$
$$= \sqrt{\text{tr}(\mathbf{GG}^T)\text{tr}[(\mathbf{GG}^T)^{-1}]}$$

This means

$$\text{tr}[(\mathbf{GG}^T)^{-1}] \geq \frac{D^2}{\text{tr}(\mathbf{GG}^T)} \quad (11)$$

The equality holds if and only if all eigenvalues of $\mathbf{GG}^T$ are identical $\lambda_i = \lambda$ $(i = 1, \ldots, D)$. This is equivalent to

$$\mathbf{GG}^T = \lambda\mathbf{I} \quad (12)$$

Now we study the term $\text{tr}(\mathbf{GG}^T)$. Due to $\underline{g}_{ij} = \underline{g}_i - \underline{g}_j$, the relationship between the $D \times \frac{M(M-1)}{2}$ matrix $\mathbf{G}$ in (9) and the $D \times M$ matrix $\mathbf{g}$ in C2 is

$$\mathbf{G} = \mathbf{gT}$$

One example for the $M \times \frac{M(M-1)}{2}$ transform matrix $\mathbf{T}$ for $M = 4$ is

$$\mathbf{T} = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

In general, each row of $\mathbf{T}$ has the squared norm $M - 1$ and each pair of rows of $\mathbf{T}$ has the inner product -1. Correspondingly,

$$\mathbf{TT}^T = \begin{bmatrix} M-1 & -1 & \ldots & -1 \\ -1 & M-1 & \ldots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \ldots & M-1 \end{bmatrix} = M\mathbf{I} - \underline{1}\,\underline{1}^T$$

with $\underline{1} = [1, \ldots, 1]^T$. This leads to

$$\mathbf{GG}^T = \mathbf{gTT}^T\mathbf{g}^T = M\mathbf{gg}^T - (\mathbf{g}\underline{1})(\mathbf{g}\underline{1})^T \quad (13)$$

and

$$\begin{aligned} \text{tr}(\mathbf{GG}^T) &= M\text{tr}(\mathbf{gg}^T) - \text{tr}[(\mathbf{g}\underline{1})(\mathbf{g}\underline{1})^T] \\ &= M\sum_{i=1}^{M} \text{tr}(\underline{g}_i\underline{g}_i^T) - \|\mathbf{g}\underline{1}\|^2 \\ &= M^2 - \|\mathbf{g}\underline{1}\|^2 \leq M^2 \quad (14) \end{aligned}$$

The equality holds if and only if the condition C1 is satisfied:

$$\mathbf{g}\underline{1} = \sum_{i=1}^{M} \underline{g}_i = \underline{0} \quad (15)$$

Combining (11) and (14), we obtain

$$\text{tr}[(\mathbf{GG}^T)^{-1}] \geq \frac{D^2}{\text{tr}(\mathbf{GG}^T)} \geq \frac{D^2}{M^2}$$

Combining (12) with (13) and (15), we obtain $\mathbf{gg}^T = c\mathbf{I}$. Since $\text{tr}(\mathbf{gg}^T) = Dc = M$, the condition C2 yields.

5.1. Optimum 2D Array

We have derived the necessary and sufficient conditions C1 and C2 to achieve the minimum CRB. But how does the optimum array look like?

Let $\underline{g}_i = [\cos\alpha_i, \sin\alpha_i]^T$ for $D = 2$. The condition C1 is clearly equivalent to $\sum_{i=1}^{M} e^{j\alpha_i} = 0$. The condition C2 implies $\sum_{i=1}^{M} \cos^2\alpha_i = \sum_{i=1}^{M} \sin^2\alpha_i = \frac{M}{2}$ and $\sum_{i=1}^{M} \sin\alpha_i \cos\alpha_i = 0$. It is easy to verify that a "uniform angular array" (UAA) defined by

$$\alpha_i = \alpha_0 + \frac{2\pi}{M}(i-1) \quad (i = 1, \dots, M) \qquad (16)$$

satisfies both C1 and C2. UAA, however, is sufficient but not necessary as shown by the next theorem.

Theorem 2: A sensor array consists of K subarrays characterized by the $D \times M_k$ matrices $\mathbf{g}_k$ ($k = 1, \dots, K$). If each subarray is optimum in the sense of C1 and C2, then the overall array of $M = \sum_{k=1}^{K} M_k$ sensors with $\mathbf{g} = [\mathbf{g}_1 \ \dots \ \mathbf{g}_K]$ is optimum as well.

Proof: From $\mathbf{g}_k\underline{1} = \underline{0}$ and $\mathbf{g}_k\mathbf{g}_k^T = \text{const} \cdot \mathbf{I}$ for all k, we conclude immediately $\mathbf{g}\underline{1} = \sum_{k=1}^{K} \mathbf{g}_k\underline{1} = \underline{0}$ and $\mathbf{g}\mathbf{g}^T = \sum_{k=1}^{K} \mathbf{g}_k\mathbf{g}_k^T = \text{const} \cdot \mathbf{I}$.

Interestingly, we also know optimum arrays other than superposition of UAAs.

5.2. Optimum 3D Array

In the three-dimensional case ($D = 3$), we again look for "uniform angular arrays" whose vectors $\underline{g}_i$ are "equally" distributed on a unit spherical surface. There are exactly five solutions to this symmetry problem, the so called Platonic solids: tetrahedron, octahedron, cube, icosahedron, and dodecahedron [9]. Table 1 shows these solids and summarizes their number of vertices, edges, and faces. Each of the vectors $\underline{g}_i$ points from the center of solids to one of the vertices. The number of vertices v is identical to the number of sensors M. The vectors $\underline{g}_i$ are known from the literature. It is straightforward to show that all Platonic solids satisfy the optimum array conditions C1 and C2.

Theorem 2 also applies to the three-dimensional case. This means, any superposition of centered Platonic solids is again an optimum sensor array. A superposition of a tetrahedron and an octahedron returns, for example, an optimum array with $M = 10$ sensors.

6. ACKNOWLEDGMENT

The authors would like to thank Mr. Zeile for valuable discussions.

tetrahed.	octahed.	cube	icosahed.	dodecahed.
$M = v = 4$	6	8	12	20
$e = 6$	12	12	30	30
$f = 4$	8	6	20	12

Table 1. Platonic solids and their number of vertices v, edges e, and faces f

7. REFERENCES

[1] Y. Huang, J. Benesty, and G. W. Elko, "Source localization," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds., chapter 9. Kluwer Academic Publishers, 2004.

[2] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Processing*, vol. 42, pp. 1905–1915, 1994.

[3] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*, Prentice-Hall, Englewood Cliffs, 2 edition, 1995.

[4] W. H. Foy, "Position-location solutions by taylor-series estimation," *IEEE Trans. Aerospace and Electron. Systems*, vol. 12, pp. 187–194, 1976.

[5] H. C. Schau and A. Z. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoustic, Speech, and Signal Processing*, vol. 35, pp. 1223–1225, 1987.

[6] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoustic, Speech, and Signal Processing*, vol. 35, pp. 1661–1669, 1987.

[7] C. Chambers, T. C. Tozer, et al., "Temporal and spatial sampling influence on the estimates of superimposed narrowband signals: When less can mean more," *IEEE Trans. Signal Processing*, vol. 44, pp. 3085–3096, 1996.

[8] A. B. Gershman and J. F. Böhme, "A note on most favorable array geometries for DOA estimation and array interpolation," *IEEE Signal Processing Letters*, vol. 4, pp. 232–235, 1997.

[9] H. Cundy and A. Rollett, *Mathematical Models*, Stradbroke, England: Tarquin Pub., 3 edition, 1989.