

Performance of Maximum Likelihood Estimation for Multipath TDOA Passive Ranging

Benjamin Shapo
Ultra Electronics 3Phoenix
Chantilly, VA, USA
ben.shapo@ultra-3pi.com

Christopher Kreucher
Integrity Applications Incorporated
Ann Arbor, MI, USA
ckreuche@umich.edu

Abstract—This paper presents an estimation procedure for exploiting multipath signal propagation in shallow water to perform passive ranging. The scenario of interest is a pair of sensors receiving an underwater acoustic signal that has arrived via a bounce from the ocean floor. This approach also applies in other situations, for example, to multiple airborne sensors passively receiving emissions from an airborne source with signal bounce off the earth's surface. We develop an implicitly-defined maximum likelihood estimator; provide a numerical scheme for calculating the estimate; and derive expressions for the estimator bias and variance. Finally, we show that at low noise, the bias is negligible and the variance is equivalent to the usual Cramer-Rao bound on unbiased estimators.

Keywords—Estimation, Cramer-Rao bound, signal multipath, passive ranging.

I. INTRODUCTION

Passive ranging refers to the process of estimating the range of a non-cooperative emitter by receiving and processing its signal at distant location(s). Passive ranging in acoustic systems (as well as RF systems) is challenging for a number of reasons. A primary difficulty is that the time at which a non-cooperative transmitter emits a signal is unknown, so calculating distance from the receive time is not possible.

Approaches to this problem in acoustic settings often use a coherent array of receiving sensors. In deep water, one can assume that the signal has traversed the direct path between the transmitter and each receive element and undergone either spherical or cylindrical spreading. In this situation, wavefront curvature approaches are often successful [1][2]. These methods work by coherently exploiting measurements of the receive times across the array. The differences in receive times allow for measurement of the curvature of the impinging wavefront, which is related to the radial distance of the emitter. The method degrades as the total array length becomes small relative to the target range.

Other authors have combined bearing measurements of the source over time, possibly from multiple sensors, with models of kinematic behavior [3]-[6]. Often these approaches are cast into a tracking framework, which estimates both the (2D or 3D) position and velocity of the source [6]. It is also possible to gain sensing diversity by controlling receiver motion, and planning the path to optimize localization performance [7]. These approaches address array length insufficiency by moving the sensor or waiting for the source to move.

In this paper, we assume a shallow-water environment that includes bottom-bounce propagation paths rather than direct paths. This approach allows operation even in cluttered environments, where direct propagation may be unavailable. We show that with two sensors at known positions, we can exploit the bounce to develop an estimate of range, and compute the bias and variance of the estimate. We assume standard signal processing techniques can be used to estimate the time difference of arrival (TDOA) for the two paths, with some error. While the details of this processing are not part of this paper, correlation-based techniques comprise one class of algorithms suitable for this task.

The paper proceeds as follows. Section II sets up the multipath ranging scenario and describes our statistical measurement model. Section III derives the maximum likelihood estimator and illustrates its performance in simulation. Finally, in Section IV, we derive the analytical bias and variance of the estimator as a function of range and illustrate the agreement with the observed estimator performance from Section III.

II. PROBLEM SETUP

Figure 1 shows the model scenario for this problem. The water depth is D meters. We assume that this depth is shallow enough that bottom bounce energy is available to the sensor. An acoustic source (right) emits energy that bounces off the ocean bottom and arrives separately at two receivers (left). In our notation, we designate the direct path distance from the source to the closer receiver as R . In addition, the distance between the receivers, projected along the bearing to the target (assumed known) is ΔR , so the direct distance to the second receiver is $R + \Delta R$. These quantities also specify the path lengths (including the bottom-bounce) from source to the two sensors, designated P_1 and P_2 respectively. Our goal is to describe the performance of an algorithm which estimates the range R using these bottom bounces.

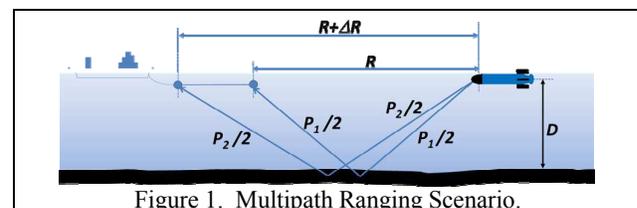


Figure 1. Multipath Ranging Scenario.

We assume a well-behaved ocean floor, where Snell's Law requires that the incidence angle and the reflection angle are equal. This is equivalent to the condition that the bottom-bounce occurs half way between the source and each receiver. The figure also indicates the half path lengths $P_1/2$ and $P_2/2$.

Finally, the approach assumes that we are analyzing short data intervals so that the target is approximately stationary within the duration of the signal processing.

Given these assumptions, we can write expressions involving the path lengths P_1 and P_2 in the following two equations:

$$\left[\frac{P_1(r)}{2}\right]^2 = \left[\frac{R}{2}\right]^2 + D^2 \quad (1)$$

and

$$\left[\frac{P_2(r)}{2}\right]^2 = \left[\frac{(R+\Delta R)}{2}\right]^2 + D^2. \quad (2)$$

The difference in path lengths to the first and second receiving element is then given by

$$P_2 - P_1 = \sqrt{(R+\Delta R)^2 + 4D^2} - \sqrt{R^2 + 4D^2}. \quad (3)$$

Dividing by the speed of sound in the medium, c , gives the measurable quantity called the time difference of arrival (TDOA)

$$\tau(R) \equiv \frac{P_2 - P_1}{c} = \frac{\sqrt{(R+\Delta R)^2 + 4D^2} - \sqrt{R^2 + 4D^2}}{c}. \quad (4)$$

We measure the TDOA $\tau(R)$ in noise. The model we use here is that there are N measurements of the TDOA while the target remains at a constant range corrupted by independent identically distributed Gaussian noise, i.e.,

$$z_i = \tau(R) + w_i, \quad (5)$$

for the N independent measurements $i = 1, \dots, N$.

Let the vector $Z = [z_1, z_2, \dots, z_i, \dots, z_N]$ be defined as the set of all measurements z_i for $i = 1, \dots, N$. The measurement likelihood can then written as

$$f(Z | R) = \frac{1}{\sqrt{2\pi\sigma_w^2}^N} \prod_{i=1}^N \exp\left[-\frac{1}{2\sigma_w^2}(z_i - \tau(R))^2\right], \quad (6)$$

and the log-likelihood is

$$\Phi(Z, R) \equiv \log f(Z | R) = C - \frac{1}{2\sigma_w^2} \sum_{i=1}^N (z_i - \tau(R))^2. \quad (7)$$

where C is a constant that does not involve Z or R .

III. MAXIMUM LIKELIHOOD ESTIMATOR

In this section, we derive an implementation of the maximum likelihood estimator (MLE) [8][9] for R given the measurements Z and illustrate the behavior of the estimator with a set of simulations.

A. Maximum Likelihood Estimator

The MLE selects the parameter(s) that are most likely given the measurements. In this case, the parameter of interest is the range R and the appropriate equation is

$$\hat{R} = \arg \max_R f(Z | R). \quad (8)$$

Maximizing the log-likelihood function $\Phi(Z, R)$ provides an equivalent estimator. Refer back to Equation (7) for the definition of $\Phi(Z, R)$.

In practice, our approach for computing the MLE involves finding the zeros of the derivative of the log-likelihood function, i.e., solving

$$\frac{\partial}{\partial R} \log f(Z | R) = 0, \quad (9)$$

or, in the notation of Equation (7),

$$\frac{\partial}{\partial R} \Phi(R, Z) = 0. \quad (10)$$

We show below that with our bottom-bounce TDOA model, the MLE is defined only implicitly. We address this by solving for the estimated range using a Newton-based iterative procedure [14][15].

With this as background, we compute the MLE as follows. First we note

$$\frac{\partial}{\partial R} \Phi(R, Z) = \frac{1}{\sigma_w^2} \sum_{i=1}^N (z_i - \tau(R)) \frac{\partial \tau(R)}{\partial R}, \quad (11)$$

The MLE is defined by this expression's zeros which are given by

$$(\bar{z} - \tau(R)) \frac{\partial \tau(R)}{\partial R} = 0, \quad (12)$$

where $\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$.

Proceeding, we find

$$\bar{z}^2 = (\tau(R))^2 = \left(\frac{P_2}{c} - \frac{P_1}{c}\right)^2. \quad (13)$$

Multiplying both sides by the constant c^2 and then plugging in expressions for P_1 and P_2 from Equations (1) and (2) yields

$$\begin{aligned} c^2 \bar{z}^2 &= (P_2 - P_1)^2 = P_2^2 + P_1^2 - 2P_1P_2 \\ &= (R + \Delta R)^2 + 4D^2 + R^2 + 4D^2 - \\ &\quad 2\sqrt{(R + \Delta R)^2 + 4D^2}\sqrt{R^2 + 4D^2}. \end{aligned} \quad (14)$$

The expression in Equation (14) defines R implicitly. It lends itself to numerical solution via a standard Newton approach (for example, [10]). If we write

$$\begin{aligned} g(R) &= 2R^2 + 2R\Delta R + \Delta R^2 + 8D^2 \\ &\quad - 2\sqrt{(R + \Delta R)^2 + 4D^2}\sqrt{R^2 + 4D^2} - (c\bar{z})^2, \end{aligned} \quad (15)$$

then we have cast the problem into a Newton-type root finding setting, where $g(R) = 0$. We compute the solution iteratively using a step that is defined by the gradient

$$\begin{aligned} \frac{\partial g}{\partial R} &= 4R + 2\Delta R - 2\sqrt{R^2 + 4D^2} \frac{R + \Delta R}{\sqrt{(R + \Delta R)^2 + 4D^2}} - \\ &\quad 2\sqrt{(R + \Delta R)^2 + 4D^2} \frac{R}{\sqrt{R^2 + 4D^2}}. \end{aligned} \quad (16)$$

The iteration then combines equations (15) and (16) to update from the k^{th} iteration to the $(k+1)^{\text{th}}$, as directed by the standard Newton approach [11].

$$\hat{R}_{k+1} = \hat{R}_k - \frac{g(R_k)}{\left. \frac{\partial g}{\partial R} \right|_{R=R_k}}. \quad (17)$$

B. Empirical Maximum Likelihood Estimator Performance

Figure 2 through Figure 4 illustrate the empirical performance of the Newton iterations described in equations (15) - (17) as a function of the TDOA error σ_w by showing the range estimates for each of 100,000 independent, single-measurement trials. Each trial was seeded with the initial estimate of $\hat{R}_0 = 0$ and terminated after Newton iterations $k+1$ and k agree to within one meter (i.e., the iterations have converged).

The simulations correspond to the scenario shown above in Figure 1, where we have chosen inter-element spacing of $\Delta R = 100\text{m}$, water depth $D = 100\text{m}$ and true range of $R = 5\text{km}$. The TDOA measurements were corrupted with zero-mean Gaussian noise. Figure 2 through Figure 4 show that increasing the TDOA noise results in an increase in MLE bias. Figure 2 plots 100k Monte Carlo trials for the low TDOA value of $\sigma_w = .01\mu\text{sec}$. The true target range is indicated by a green line. The blue dots represent individual trial MLE estimates, and they scatter around the true value. The red overlay shows the mean value of the MLE Monte Carlo trial estimates.

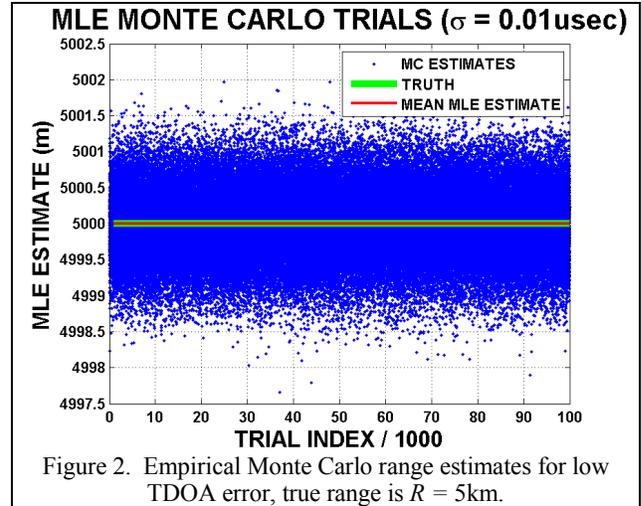


Figure 2. Empirical Monte Carlo range estimates for low TDOA error, true range is $R = 5\text{km}$.

The figure shows clearly that the MLE bias (separation between the red line and the green truth line) is negligible compared to the MLE variance.

Figure 3 shows the same simulation, with higher TDOA error of $\sigma_w = 10\mu\text{sec}$. The significantly larger span of estimates along the y-axis indicates the much higher variance in MLE estimates.

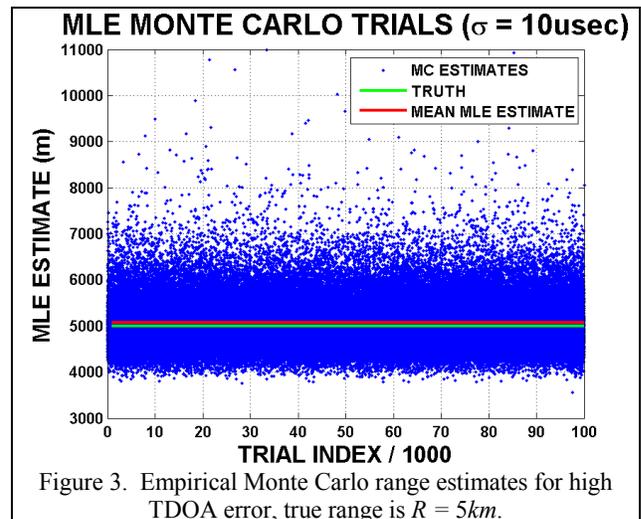


Figure 3. Empirical Monte Carlo range estimates for high TDOA error, true range is $R = 5\text{km}$.

In addition, the bias has grown significantly. While this result may not be readily evident from Figure 3, Figure 4 provides an apples-to-apples comparison (displayed with the same y-axis limits) of the MLEs with the two different TDOA parameters. In the left-hand panel, TDOA error is small ($.01\mu\text{sec}$) and the bias (difference between red and green lines) is not detectable to the eye. However, in the right-hand panel, the TDOA error has increased to $10\mu\text{sec}$ and now the bias (displayed with the same axes as in the left-hand panel) is about 75m . Note also that the MLE variance has increased significantly, demonstrated by the wider extent of the blue dots indicating individual trial estimates.

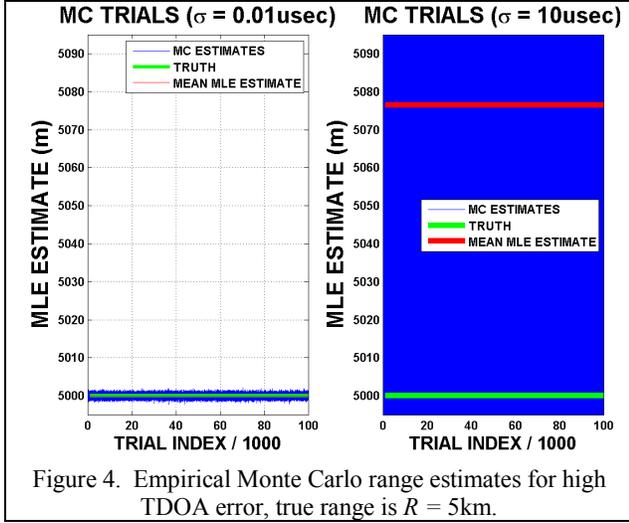


Figure 4. Empirical Monte Carlo range estimates for high TDOA error, true range is $R = 5\text{km}$.

In the following section, we derive expressions that describe the bias and variance of the estimator as a function of R , ΔR , D , and measurement noise σ_w .

IV. ANALYTICAL ESTIMATOR PERFORMANCE

Fessler [12] has studied the bias and variance of implicitly defined estimators such as that given by eq. (14). His approach is to consider notionally a function $\hat{R} = h(Z)$, which provides the estimate of R from the random data Z . Since $h(Z)$ cannot be written explicitly, his approach is to develop the Taylor series expansion of $h(Z)$ and use the terms in the expansion to compute the mean and variance of the estimate \hat{R} . In this section, we use his results to develop second-order expressions for our estimator's bias and variance and compare them to the empirically observed estimator performance based on Monte Carlo simulations.

A. Estimator Bias

In our notation, the second order approximation of the estimator mean is given by

$$E(\hat{R}) \approx h(\bar{Z}) + \frac{1}{2} \sum_{n=1}^N \frac{\partial^2}{\partial z_n^2} h(\bar{Z}) \text{Var}\{z_n\}, \quad (18)$$

where \bar{Z} denotes the mean of Z , and we use the shorthand

$$\frac{\partial^2}{\partial z_n^2} h(\bar{Z}) \equiv \left. \frac{\partial^2}{\partial z_n^2} h(Z) \right|_{Z=\bar{Z}}. \quad (19)$$

In these expressions, $h(\bar{Z})$ is the estimate of R in the absence of noise. Let \bar{R} denote this value, i.e., $\bar{R} = h(\bar{Z})$. The numerical value of \bar{R} is found by running the estimation process described by eq. (17) on noise-free data. We find empirically that $\bar{R} = R_{true}$ for our model.

We require an expression for $\frac{\partial^2}{\partial z_n^2} h(\bar{Z})$ from eq. (18) to

compute the estimator bias. Fessler shows that this expression depends on the partial derivatives of Φ and h [12]. In our setting, these terms are

$$\begin{aligned} \frac{\partial^2}{\partial R^2} \Phi(\bar{R}, \bar{Z}) &= -\frac{N}{\sigma_w^2} \left(\frac{\partial \tau(\bar{R})}{\partial R} \right)^2 \\ \frac{\partial^2}{\partial R \partial z_n} \Phi(\bar{R}, \bar{Z}) &= \frac{N}{\sigma_w^2} \frac{\partial \tau(\bar{R})}{\partial R} \\ \frac{\partial^3}{\partial R^3} \Phi(\bar{R}, \bar{Z}) &= -\frac{3N}{\sigma_w^2} \frac{\partial \tau(\bar{R})}{\partial R} \frac{\partial \tau^2(\bar{R})}{\partial R^2} \\ \frac{\partial}{\partial z_n} h(\bar{Z}) &= \frac{1}{N} \left(\frac{\partial \tau(\bar{R})}{\partial R} \right)^{-1} \\ \frac{\partial^3}{\partial R^2 \partial z_n} \Phi(\bar{R}, \bar{Z}) &= \frac{1}{\sigma_w^2} \frac{\partial \tau^2(\bar{R})}{\partial R^2}, \end{aligned} \quad (20)$$

and

$$\frac{\partial^3}{\partial R \partial z_n^2} \Phi(\bar{R}, \bar{Z}) = 0. \quad (21)$$

This leads directly to the expression

$$\frac{\partial^2}{\partial z_n^2} h(\bar{Z}) = \frac{-\left(\partial^2 \tau(\bar{R}) / \partial R^2 \right)}{N^2 \left(\partial \tau(\bar{R}) / \partial R \right)^3}. \quad (22)$$

Combining this expression with eq. (18) and the observation that $\bar{R} = R_{true}$, we finally have that the estimator bias is given by

$$B(\hat{R}) \approx \frac{\sigma_w^2}{2N} \frac{\partial^2 \tau(R_{true})}{\partial R^2} / \left(\frac{\partial \tau(R_{true})}{\partial R} \right)^3, \quad (23)$$

where

$$\frac{\partial \tau(R)}{\partial R} = \frac{1}{c} \left(\frac{R + \Delta R}{\sqrt{(R + \Delta R)^2 + 4D^2}} - \frac{R}{\sqrt{R^2 + 4D^2}} \right), \quad (24)$$

and

$$\frac{\partial^2 \tau(R)}{\partial R^2} = \frac{4D^2}{c} \left(\frac{1}{\left[(R + \Delta R)^2 + 4D^2 \right]^{3/2}} - \frac{1}{\left[R^2 + 4D^2 \right]^{3/2}} \right). \quad (25)$$

The estimator is asymptotically unbiased, and bias is proportional to the variance of the TDOA noise σ_w^2 .

B. Estimator Variance

The second order approximation of the estimator variance [12] is

$$V(\hat{R}) \approx \frac{\sigma_w^2}{N} \left(\frac{\partial^2}{\partial R \partial z_n} \Phi(\bar{R}, \bar{Z}) \right)^2 / \left(\frac{\partial^2}{\partial R^2} \Phi(\bar{R}, \bar{Z}) \right)^2, \quad (26)$$

which is given explicitly using the results of the calculations given in eq. (20) as

$$V(\hat{R}) \approx \frac{\sigma_w^2}{N} \left(\frac{\partial \tau(R_{true})}{\partial R} \right)^2. \quad (27)$$

The variance is proportional to the TDOA noise variance and, as expected, decreases linearly with the number of measurements N .

This variance can be seen to be equivalent to the usual Cramer-Rao bound (CRB) [13] on unbiased estimators as follows. First, the Fisher information is defined as

$$I(R) = -E \left[\frac{\partial^2}{\partial R^2} \log(f(Z | R)) \right]. \quad (28)$$

From the definitions in eq. (20), we have

$$I(R) = \frac{N}{\sigma_w^2} \left(\frac{\partial \tau(R)}{\partial R} \right)^2 \quad (29)$$

and the CRB is given by

$$\text{CRB} \equiv I^{-1}(R) = \frac{\sigma_w^2}{N} \left(\frac{\partial \tau(R)}{\partial R} \right)^2, \quad (30)$$

which coincides with eq. (27).

We can then write the bias B and standard deviation S at range R (to second order) explicitly by combining the results in equations (23) and (27) with the expressions in equations (24) and (25) as

$$B(R) = \frac{2D^2 \sigma_w^2 c^2}{N} \times \frac{(R^2 + 4D^2)^{3/2} - ((R + \Delta R)^2 + 4D^2)^{3/2}}{\left((R + \Delta R) \sqrt{R^2 + 4D^2} - R \sqrt{(R + \Delta R)^2 + 4D^2} \right)^3} \quad (31)$$

and

$$S(R) = \frac{c \sigma_w}{\sqrt{N}} \times \frac{\sqrt{(R + \Delta R)^2 + 4D^2} \sqrt{R^2 + 4D^2}}{(R + \Delta R) \sqrt{R^2 + 4D^2} - R \sqrt{(R + \Delta R)^2 + 4D^2}}. \quad (32)$$

We see that the bias to standard deviation ratio increases as the standard deviation of the TDOA noise.

C. Empirical Performance

We now compare these analytical expressions for bias and variance with the empirically observed estimator performance from Monte Carlo simulations.

Figure 5 and Figure 6 show the model problem we use to illustrate estimator performance. They highlight two potential spacings for the two sensors and a target at range R . We consider the estimates of range from the both possible pairs of sensing positions illustrated in the figure, indicated as positions (1, 2), and positions (1, 3). Figure 5 shows sensors placed in the closer configuration, and Figure 6 shows the wider sensor spacing.

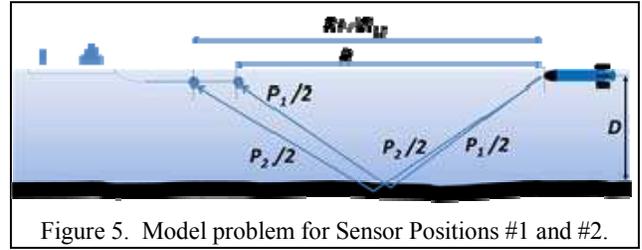


Figure 5. Model problem for Sensor Positions #1 and #2.

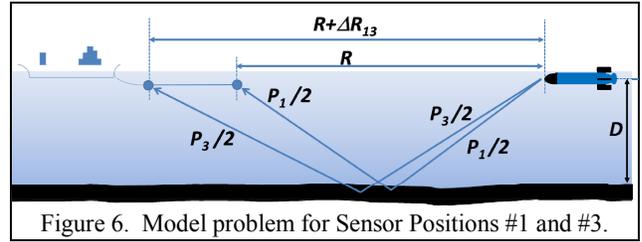


Figure 6. Model problem for Sensor Positions #1 and #3.

For our numerical examples, we used inter-position spacings of $\Delta R_{12}=100m$ (closer spacing at positions (1, 2)) and $\Delta R_{23}=200m$ (wider spacing at positions (1, 3)) and water depth $D=100m$ and considered various discrete horizontal ranges R up to $5km$. Figure 7 shows the single-measurement standard deviation of the estimate with TDOA error standard deviation of $\sigma_w=5\mu sec$. The solid lines in Figure 7 illustrate the analytical estimator standard deviation for both sensor position pairs. The circle symbols show the empirical estimator performance (Eq. (17)) averaged over 100,000 Monte Carlo trials. The realized estimator performance coincides closely with the analytical expressions.

Figure 8 shows the analytically computed estimator bias and the empirically observed bias over the same Monte Carlo trials. Again, the realized performance coincides with the analytical expressions.

As expected, the error indicates larger variance range estimation as range increases. It is also interesting to compare the curves themselves. Ranging using the wider sensor spacing (positions 1 and 3) represents a dramatic performance improvement. This significant difference is due to the larger

spacing (ΔR_{13} vs. ΔR_{12}) between sensor positions 1 and 3, relative to the smaller spacing between positions 1 and 2.

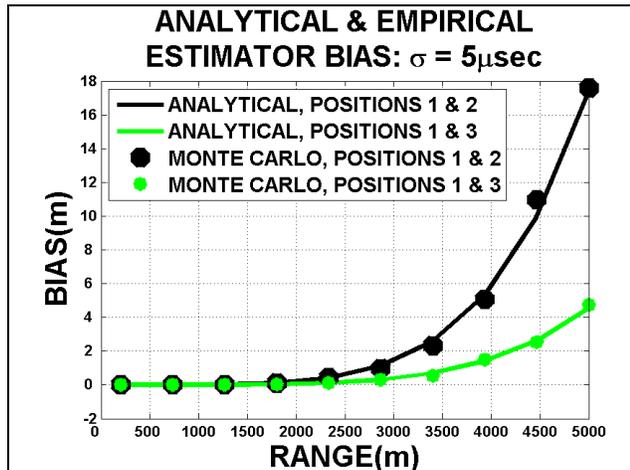


Figure 7. Analytically computed estimator standard deviation (solid lines) and Monte Carlo standard deviation (circles).

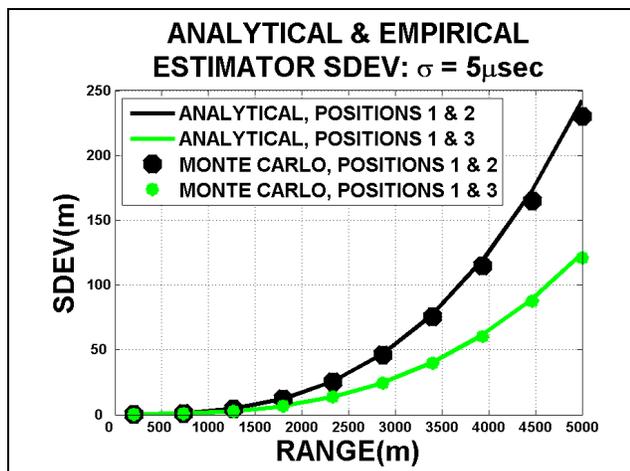


Figure 8. Analytically computed estimator bias (solid lines) and empirically computed bias (circles).

Finally, Figure 9 shows the relationship between estimator bias and variance at a TDOA noise value of $\sigma_w = 5\mu s$, parameterized by the range R to the source. We selected a pair of sensors spaced 100m apart (the smaller sensor spacing ΔR_{12}) and a target at range 5km. The agreement between predicted and empirical bias at the closest range displayed (1267m) appears poor because the bias itself is so low (see Figure 8).

Equations (31) and (32) show that the estimator bias is negligible relative to estimator standard deviation for $\sigma_w \leq 5\mu s$ at the ranges we have considered here. Conversely, bias becomes the dominant source of error at $\sigma_w \geq 1ms$, but the mean squared error at that point is so large that estimator is not viable.

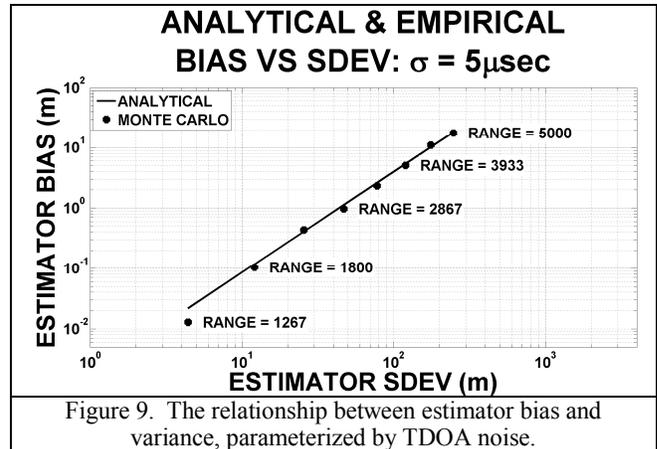


Figure 9. The relationship between estimator bias and variance, parameterized by TDOA noise.

REFERENCES

- [1] H.L. Van Trees, Optimum Array Processing. Wiley, 2002.
- [2] B.G. Ferguson, "Wavefront curvature passive ranging in a temporally varying sound propagation medium," *Proceedings of the Oceans 2001 Conference*, ISBN no. 0-933957-28-9, pp. 2359-2365, November, 2001.
- [3] W.P. Malcolm, A. Doucet, and S. Zollo, "Sequential Monte Carlo tracking schemes for maneuvering targets with passive ranging," *Proceedings of the 5th International Conference on Information Fusion*, vol. 1, pp. 482-488, July, 2002.
- [4] J. M. C. Clark, S. A. Robbiati, and R. B. Vinter, "The shifted Rayleigh mixture filter for bearings-only tracking of maneuvering targets," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, July, 2007.
- [5] C. Kreucher and B. Shapo, "A Fuse-Before-Track Approach to Target State Estimation Using Passive Acoustic Sensors," *Proceedings of the 14th IEEE Conference on Information Fusion*, pages 1889-1896, July 5-8, 2011.
- [6] C. Kreucher and B. Shapo, "Multitarget Detection and Tracking Using Multisensor Passive Acoustic Data," *IEEE Journal of Oceanic Engineering*, 36(2):205-218, April 2011.
- [7] R. Karlsson and F. Gustafsson, "Recursive Bayesian estimation: bearings-only applications," *IEE Proceedings on Radar, Sonar, and Navigation*, vol. 152, no. 5, pp. 305-313, October, 2005.
- [8] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I. Wiley, 1968.
- [9] Bickel, P. J., & Doksum, K. A. *Mathematical Statistics*. Prentice-Hall, 1977.
- [10] WH Press; SA Teukolsky; WT Vetterling; and BP Flannery. *Numerical Recipes: The Art of Scientific Computing* (3rd ed.). Cambridge University Press, 2007.
- [11] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, 2001.
- [12] J. Fessler, "Mean and Variance of Implicitly Defined Biased Estimators (Such as Penalized Maximum Likelihood): Applications to Tomography," *IEEE Transactions on Image Processing*, vol. 5, no. 3, March 1996.
- [13] V. Raykar and R. Duraiswami, "Approximate Expressions for the Mean and Covariance of the Maximum Likelihood Estimator for Acoustic Source Localization," *ICASSP 2005*, pp. 73-76.
- [14] D. Storer, A. Nehorai, "Newton Algorithms for Conditional and Unconditional Maximum Likelihood Estimation of the Parameters of Exponential Signals in Noise," *IEEE Transactions on Signal Processing*, vol. 40, no. 6, June 1992.
- [15] O. Cappe, V. Buchoux, E. Moulines, "Quasi-Newton Method for Maximum Likelihood Estimation of Markov Models," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 2265-2268, May 1998.