Efficient 2D Sensor Location Estimation using Targets of Opportunity

DAVID F. CROUSE
RICHARD W. OSBORNE, III
KRISHNA PATTIPATI
PETER WILLETT
YAAKOV BAR-SHALOM

This paper discusses the Maximum Likelihood (ML) algorithm for the self-localization of passive (angular) or active (angle and range) sensors using targets of opportunity. The approach, which is considered in two dimensions, is appropriate when traditional alternatives, such as the use of known-location targets or satellite navigation systems, are unavailable. It is not assumed that the sensors can "see" each other, though they are assumed to take measurements with respect to a common (biased) axis. Unlike previous ML algorithms, we take into account the circular nature of the angular measurements, allowing for more accurate estimates to be obtained. A simple least-squares method is additionally provided for initialization. Simulations demonstrate that the accuracy of the ML estimator approaches the Cramér-Rao Lower Bound (CRLB), something that similar algorithms have been unable to achieve.

Manuscript received August 18, 2010; revised December 14, 2011; released for publication September 15, 2012.

Refereeing of this contribution was handled by Benjamin Slocumb.

This work was partially supported by the Office of Naval Research under contracts N00014-09-10613 and N00014-10-10412 and by the Army Research Office under contract W911NF-06-1-0467.

Authors' address: The Department of Electrical and Computer Engineering, University of Connecticut, 371 Fairfield Way, U-2157, Storrs, Connecticut 06269, E-mail: ({crouse, rosborne, krishna, willett, ybs}@engr.uconn.edu).

1557-6418/13/$17.00

1. INTRODUCTION

Due to their low cost and ease of deployment, the use of passive acoustic sensors for target tracking has seen increasing popularity. Such systems might consist of individual microphones or hydrophones that self-assemble into arrays [24], or, perhaps, sensors consisting of clusters of microphones or hydrophones, each producing measurements consisting of arrival angles and features/attributes for use in data association [19].[1] The clusters of microphones or hydrophones form individual sensors, which can also be referred to as "nodes" in the system. This paper focusses on the latter scenario, localizing sensors with measurements taken with respect to a common, unknown coordinate axis. Determining which detection on one sensor corresponds to the same target on another sensor (measurement association) might be accomplished, for example, by utilizing acoustic patterns for classification, as has previously been done to aid acoustic tracking [19]. Target tracking is not considered here. The scenarios considered focus on angular noise levels up to 2° (root-mean squared error), which is the accuracy of the sensors in [19], though acoustic sensors can often have significantly worse angular accuracies.

When considering the construction of land-based sensor networks, it cannot be assumed that satellite-based localization systems, such as GPS (USA) or GLONASS (Russia), will be available, and such signals cannot penetrate far underwater. However, many non-satellite-based location estimation algorithms, which have been primarily designed for use in underwater and wireless networks may be used. A number of methods applied to sonar channels are described in [4]. These approaches typically utilize the communication characteristics between sensors and are divided into two categories: range-based and range-free. Range-based methods utilize range (distance) measurements. Range-free schemes do not utilize range information. Both techniques might take advantage of moving anchor nodes that broadcast their position [6, 9, 13, 22].

Our focus is on algorithms for node localization based on the angle-only observations of the nodes, though we do consider the case where range measurements are also available. Estimates based on angle-only measurements are particularly useful when the sensors have a limited broadcast range. Underwater, this might be the case when the sensor network is built using data MULEs (Mobile Ubiquitous LAN[2] Extensions) [21]. A data MULE is a mobile device that approaches the sensors to collect data. In such a network, traditional methods of sensor localization, which rely on communication channels between sensors, are not applicable.

[1] Given multiple targets in a scene, features, such as the range-Doppler profile of different targets, can be used in target tracking algorithms to determine which measurements originated from which targets.

[2] Local Area Network

The node localization algorithm considered in this paper can also be used with networks of wireless land-based acoustic sensors. Though many localization techniques using aspects of the wireless communication channel between sensors exist, estimates of the sensor locations obtained using acoustic data can reasonably be expected to be uncorrelated with those obtained using more traditional means. Multiple uncorrelated estimates can be easily fused, improving the overall accuracy. Under the typical assumption that the noise on the estimates is Gaussian distributed, estimate fusion via the least squares algorithm [1, Ch. 3] requires that the covariance matrix of the individual estimates be known. Thus, in Section 7, the Cramér Rao Lower Bound (CRLB), a lower bound on the error of an unbiased estimator, is derived for the estimation problem at hand. Since the accuracy of the ML estimation method derived in this paper approaches the CRLB, as demonstrated in Section 7, the CRLB should be an accurate approximation for the covariance matrix of the estimate at low noise levels. However, being a lower-bound, in more difficult (nonlinear) estimation problems, or when the signal to noise ratio is low, the CRLB tends to be overly optimistic. The validity of the CRLB for estimate fusion is not considered in this paper.

A maximum likelihood (ML) solution to localizing both passive and active nodes is outlined in [16]. Though sensors often cannot take measurements spanning a full 360°. Additionally, the distribution of the noise corrupting angular measurements often depends upon the geometry of the target with respect to the hydrophone or microphone array taking the measurement. Nonetheless, it is common practice for the noise corrupting angular measurements to be modeled as Gaussian, which is not bounded to a range of 0 to 2π radians. Since the sensors can face any direction, the angular measurements taken by the sensors in a global coordinate system can span the range of 0 to 2π or $-\pi$ to π, depending upon where the boundary is placed. The Gaussian noise approximation is often good *except near the discountinity* (0–2π or $-\pi$–π). Figure 1 illustrates the boundary problem. This paper rederives the ML algorithm accounting for the idiosyncrasies of circular data. Section 2 discusses the signal model and the ML solution is provided in Section 4. Since poor performance is generally observed when using a random initialization, Section 5 discusses the generation of initial estimates without prior information. Though this work focusses on angle-only networks, the case where 2D range measurements are also available is additionally considered. Section 7 demonstrates the performance of the algorithms through simulation and Section 8 summarizes the results.

With the exception of [16], few algorithms have been developed to jointly localize and determine the orientation of angle-only sensors. A significant amount of work has been done regarding localizing users within cellular networks [23], with very little attention paid

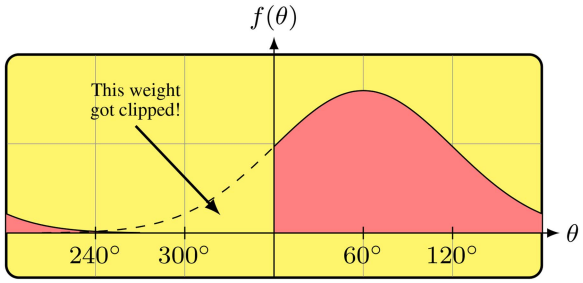


Fig. 1. The traditional linear measurement model as applied to circular data with a mean of 60° does not accurately represent the uncertainty in the likelihood of observations near the 0–2π boundary, as illustrated for the Normal distribution. In this case, a significant portion of the density is clipped. A circular measurement model (illustrated by the dashed line) more accurately reflects the underlying uncertainty.

to the angle-only measurement case. In what has been done, a single user must always be in range of at least two base stations (anchor nodes). Other work has considered similar issues for cellular networks [17], whereby all users are in range of a number of anchor nodes. In our solution, no target ever needs to be simultaneously observed by two nodes of a known location provided that conditions for observability, which are discussed in Section 3, are met.

In [18] an algorithm for localizing sensors that can see their neighbors was developed. The algorithm is deterministic and errors compound with propagation distances. An attempt to mitigate this problem was given in [11], where a linear programming method was used to improve the consistency of the angular measurements between sensors before the deterministic localization algorithm was executed. In both instances, it is generally assumed that if one sensor can see another, then the reverse is true. Thus, these algorithms are not applicable to the case where targets of opportunity are observed. Another method involving semidefinite programming was given in [3], but it requires the use of a heuristic parameter that depends upon the size and geometry of the network.

Many papers dealing with sensor registration only correct for residual bias after an initial estimate has been obtained. Most require full range and angle estimates (see [5] for an extensive bibliography), though some are adaptable to the range-only case [20]. The majority of algorithms only estimate the sensor orientations, though some can also estimate the sensor positions [14]. Most approaches utilize some type of linearization and none of them are applicable to the aforementioned estimation scenarios when no initial estimate is available.

2. DEFINITIONS AND MODELS

We assume that all angular and, if available, range measurements are taken in two dimensions with respect to a common axis, which need not be known. The sensors and the targets are assumed to be individual points in space. The measurements between sensors are as-

sumed to be synchronized. That is, in each scan, measurements from disparate sensors represent the angle from the sensor to the targets at the same time (at the same target position).[3] It should be noted that individual observations may occur simultaneously or at different times if a target is stationary. A measurement of the same target at a different time shall simply be considered as another target in the context of this problem, since tracking is not performed (admittedly such information would help, but here we ignore it). If multiple targets are present at the same time, then classification information may be used to associate measurements between the sensors. We will not address the problem where measurements cannot be associated between sensors, in which case there may be multiple possible solutions for the target location based upon a particular set of observations.

When dealing with angles, it will become necessary to utilize a four-quadrant inverse tangent function with range $(-\pi, \pi]$, which is defined as follows

$$\operatorname{atan}_2[y,x] \triangleq \begin{cases} \arctan\left[\dfrac{y}{x}\right] & x > 0 \\ \arctan\left[\dfrac{y}{x}\right] + \pi & x < 0, \quad y \geq 0 \\ \arctan\left[\dfrac{y}{x}\right] - \pi & y < 0, \quad x < 0 \\ \dfrac{\pi}{2} & x = 0, \quad y > 0 \\ -\dfrac{\pi}{2} & x = 0, \quad y < 0 \\ 0 & x = 0, \quad y = 0 \end{cases} \tag{1}$$

where arctan represents the standard inverse tangent function with range $(-\pi/2, \pi/2)$.

Let $\theta_{s,t}$ and $r_{s,t}$ be the angular and (if available) range measurements from sensor s observing target t. Both shall be assumed corrupted with zero-mean additive noise:

$$\theta_{s,t} = \theta_{s,t}^{\text{true}} + w_{s,t}^{\theta} \tag{2}$$

$$r_{s,t} = r_{s,t}^{\text{true}} + w_{s,t}^{r}. \tag{3}$$

All of the additive noises are assumed independent. The range noise, $w_{s,t}^r$, is assumed to be distributed Gaussian $\mathcal{N}\{0, \sigma_r^2\}$. The Gaussian noise assumption is commonly used despite the fact that one will never measure a negative range.[4] As the targets can be assumed to be far from the sensors compared to the standard deviation of the noise ($> 30\sigma$), this approximation is accurate. However, the use of a Gaussian approximation for noise corrupting angular measurements is more problematic.

When dealing with angular measurements, many traditional statistical concepts need to be redefined due to the "wrapping" of the distribution about the circle and due to problems at the 0–2π boundary. For example, the traditional notions of mean and variance no longer provide useful quantities; a sample mean of $-\pi$ and π would yield zero, which is the worst possible estimate, since $-\pi$ and π represent the same point on the circle. For this reason, a lot of research has been done with regard to statistical methods relating to directional data in multiple dimensions [7, 8, 15]. Our definition of the mean direction and circular standard deviation are based on the following trigonometric moments

$$\alpha \triangleq \mathrm{E}[\cos\theta], \qquad \beta \triangleq \mathrm{E}[\sin\theta]. \tag{4}$$

The mean resultant length, ρ, and mean direction, μ_θ, are defined in terms of these moments through the polar relation

$$\alpha + j\beta = \rho e^{j\mu_\theta}. \tag{5}$$

Thus, the magnitude of $\alpha + j\beta$ is the mean resultant length, and its phase is the circular mean. The circular standard deviation is defined as

$$\sigma_\theta \triangleq \sqrt{-2\ln\rho}. \tag{6}$$

The maximum likelihood sensor localization algorithm derived in [16] using common targets of opportunity assumed that angular measurements of the target locations were corrupted with additive Gaussian noise. That is, the distribution of the angular measurement was

$$p(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\theta-\mu)^2}{2\sigma^2}\right). \tag{7}$$

As will be demonstrated in Section 7, maximization based upon the algorithm in [16] yields results that are often very useful but, depending upon the geometry of the sensors and the observed targets, can also be quite far from the CRLB. In other words, the estimator cannot be statistically efficient without accounting for the circularity of the measurements. This paper accounts for the circularity of the noise.

In the following subsections, we consider three noise distributions appropriate for circular data: the wrapped normal distribution, the clipped mod normal distribution and the von Mises distribution. At low circular standard deviations, less than around 20° or $\pi/9$ radians, all of the distributions are essentially the same, as shown in Fig. 2. However, as the standard deviations increase, the differences become more profound, with the von Mises distribution distinguishing itself the most from the other two.

If an angular measurement is truly corrupted with additive Gaussian noise, which is often a good assumption due to the Central Limit Theorem, then the noise effectively gets wrapped on the region from 0 to 2π or $-\pi$ to π, depending upon where one wishes to make the cut. For this reason, the wrapped normal distribution,

[3] This is equivalent to saying that the sensors are assumed to be synchronized and the propagation delay between a target and a sensor is assumed to be negligible.

[4] The normal distribution is unbounded, implying that there exists a probability, however small, of measuring a negative range.

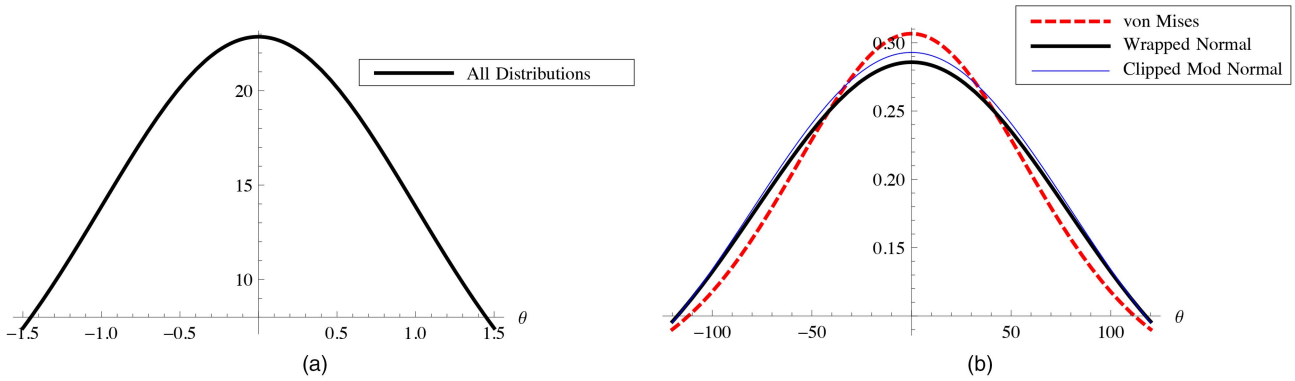


Fig. 2. When the circular standard deviation is low, as in (a), then the distributions shown are essentially the same. Increasing the circular standard deviation, the von Mises differentiates itself first, and only when the circular standard deviation is very high are the two normal distributions significantly different. The plot ranges are $\pm 1.5\sigma_\theta$. (a) $\sigma_\theta = 1°$. (b) $\sigma_\theta = 80°$.

discussed in Section 2.1, is the most natural distribution to use. Moreover, the circular mean and standard deviation in the wrapped normal distribution have an easy-to-understand meaning: they are equal to the mean and standard deviation of the linear normal distribution that got wrapped to the circle [15]. Indeed, if the standard deviation is small and the mean is far from the boundary, then both distributions are nearly the same.

On the other hand, array processing generally does not directly yield angular measurements, but rather unit vectors pointing toward the targets. Processing the measurements in their original (array) coordinates can possibly avoid the boundary issues illustrated in Fig. 1. However, many sensors only provide angular measurements. If one converts such unit vectors having components corrupted with Gaussian noise into angles, then the angular measurements are von Mises distributed [15, pg. 42]. The von Mises distribution is discussed in Section 2.3.

2.1. The Wrapped Normal Distribution

The *wrapped normal distribution* is obtained when Gaussian noise is added to a circular datum. As discussed, for example in [7], [15], if the additive noise is distributed as $\mathcal{N}\{\mu,\sigma^2\}$, then the wrapped normal distribution has the following PDF

$$p(\theta) = \frac{1}{2\pi}\left(1 + 2\sum_{k=1}^{\infty}\rho^{k^2}\cos[k(\theta-\mu)]\right) \tag{8}$$

where

$$\rho = e^{-\sigma^2/2} \tag{9}$$

and $-\pi \le \theta < \pi$. The aforementioned definitions of the mean direction and circular standard deviation are such that $\mu_\theta = \mu$ and $\sigma_\theta = \sigma$.

2.2. The Clipped Mod Normal Distribution

We developed the clipped mod normal distribution as a simple approximation that avoids the infinite sum present in the wrapped normal distribution. It comes from the assumption that almost all of the mass of the Gaussian noise added to the measurement is within $\pm\pi$ of the mean. In this case, shifting the cutting region as far as possible from the mean and discarding the mass that would have been wrapped (in this case, almost nothing) and renormalizing the distribution is a good approximation of the distribution on the circle. Thus, the wrapped normal distribution may be approximated by the following shifted and clipped distribution

$$p(\theta) = \frac{1}{c}\exp\left(-\frac{[m(\theta-\mu)]^2}{2\sigma^2}\right) \tag{10}$$

where

$$c = \Phi\left[\frac{\pi}{\sigma}\right] - \Phi\left[-\frac{\pi}{\sigma}\right] \tag{11}$$

$$m(\theta) = \begin{cases} \theta - 2\pi & \text{if} \quad \theta > \pi \\ \theta + 2\pi & \text{if} \quad \theta < -\pi \\ \theta & \text{otherwise} \end{cases} \tag{12}$$

and $-\pi \le \{\mu,\theta\} < \pi$, Φ is the cumulative distribution of the standard normal distribution, and c is the normalizing constant. Because the function $m(\theta-\mu)$ is squared, other forms for m are also valid. Note, however, that one cannot replace m with the modulo over π or over 2π, because that would assign large penalties to small negative offsets.

2.3. The von Mises Distribution

The von Mises distribution was derived in 1918 by Richard von Mises in an attempt to statistically determine whether the atomic weights of elements were integer multiples of a common base unit of weight, whereby non-integer measurements would be attributed to noise.[5] The von Mises distribution (on the circle) has been widely studied, in part to due to its similarity to the wrapped Cauchy distribution and the wrapped normal

[5]He concluded that the likelihood at the zero point (that the atomic weights are integers) was nine times greater than the average likelihood across the rest of the circle, which did not tell him very much.

distribution. More information on the von Mises distribution and its multidimensional generalizations (von Mises-Fisher distributions) can be found in [7], [8], [15]. The von Mises distribution is given by

$$p(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos[\theta - \mu_\theta]) \tag{13}$$

where μ_θ is the mean direction, $-\pi \leq \{\theta, \mu_\theta\} < \pi$, and $0 \leq \kappa < \infty$. The value κ is a measure of the concentration of the distribution and is inversely proportional to the circular standard deviation. The function I_ν is a modified Bessel function of the first kind. For an integer ν, I_ν is given by

$$I_\nu(\kappa) = \frac{1}{\pi} \int_0^\pi \exp[\kappa \cos[\theta]] \cos[\nu\theta] d\theta. \tag{14}$$

The mean resultant length of the von Mises distribution is

$$\rho = \frac{I_1(\kappa)}{I_0(\kappa)}. \tag{15}$$

The von Mises distribution approximates a wrapped Normal distribution having the same mean and circular standard deviation. Given κ, the circular standard deviation can be calculated from (15) and (6). Finding a value of κ corresponding to a particular circular standard deviation can be performed using a simple numerical search. However, the mean resultant length in (15) can be hard to evaluate for accurate measurements. For example, if $\sigma_\theta = \pi/180$, that is a 1° standard deviation, a reasonable value for an accurate sensor, then $\kappa \approx 3283$. Though ρ is just under one, $I_1(\kappa) \approx 5.81 \times 10^{1423}$, a number that cannot be stored in a computer's double-precision floating point register. However, the ratio in (15) may be expressed as an infinite sum. Methods for computing the ratio are compared in [10] and the most efficient method for this problem is summarized in Appendix B. The calculation of this ratio is important for evaluating the CRLB, as discussed in Section 6.

3. THE OBSERVABILITY OF THE SENSOR LOCATIONS

The requirements for observability in the angle-only case shall be considered. Figure 3 shows a system with three sensors, s_1, s_2, and s_3. Suppose that the location of s_1 is known. In this case, all angles in the system may be preserved by scaling everything around s_1. That is, the locations of s_2 and s_3 are not observable if only s_1 is known. Now suppose that both s_1 and s_3 have known locations. In this case, the locations of the targets t_1 and t_2 can be uniquely determined (we know this from the angle-side-angle theorem of planar geometry).

Given that the locations of s_1 and s_3 are known, if only target t_1 were observed, then, having angle-only measurements, the location of s_2 could not be uniquely determined, because the observed angle at s_2 simply defines a line passing through t_1. If both t_1 and t_2 were observed, then, as shown in Fig. 3(b) angles θ_1, θ_2 and θ_3

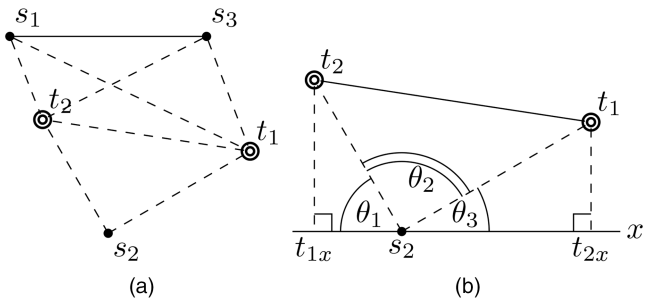


Fig. 3. A system with three sensors and two targets in (a). If the location of two sensors is known, then the locations of the targets may be uniquely determined as well as the third sensor location.

are all known. Similarly, because s_1 and s_3 have known locations, the locations of t_1 and t_2 are known exactly. Thus, the location of s_2 may be solved by considering the intersection of the line passing through t_1 at an angle of θ_3 with respect to the x-axis with the line passing through t_2 at an angle of $\theta_1 + \theta_3$ with respect to the x-axis. Note that this will not work if t_1 and t_2 are collinear with respect to s_2.

All together, in order for the locations of sensors in a network consisting of angle-only observations to be observable, the locations of at least two sensors must be known a priori (anchor nodes). Additionally, at least two targets must be observed by the sensors. However, the anchor nodes need not observe common targets if connected subsets of sensors between them can observe the targets seen by the anchor nodes. Section 5 presents an algorithm for generating initial estimates of the sensor locations by solving a set of linear equations. In that case, this observability criterion manifests itself as a requirement that the matrix in the linear equation be invertible. The fact that the anchor nodes need not be simultaneously seen by all sensors is made clearer through simulation in Section 7 where the anchor nodes never see the same targets simultaneously.

When range measurements are available, the situation becomes simpler, because a single sensor can uniquely identify the location of a target. Thus, the location of any additional sensor seeing the target may also be determined. This means that the location of only one sensor needs to be known. However, if measurements are taken with respect to a common, unknown axis, then two sensor locations must be known in order to resolve the angular ambiguity.

4. MAXIMUM LIKELIHOOD ESTIMATION

This section presents a general formulation of the maximum likelihood estimator using range measurements and angular measurements taken with respect to a common, unknown axis. If range measurements are not available or if the measurement axis is known, then the appropriate terms in the objective function and gradient should be omitted. Defining $w^r_{i,j}$ and $w^\theta_{i,j}$ to be uncorrelated additive noise corrupting the range and angular components of the measurement from sensor i to target

location j, $\flat$ to be the "bias" on the angular measurements (representing the fact that they are taken with respect to an unknown axis), then, using the measurement model given in Section 2, a measurement consisting of a range, $r_{i,j}$ and angle, $\theta_{i,j}$ is given by

$$r_{i,j} = \sqrt{(y_t^j - y_s^i)^2 + (x_t^j - x_s^i)^2} + w_{i,j}^r \tag{16}$$

$$\theta_{i,j} = \text{atan}_2[y_t^j - y_s^i, x_t^j - x_s^i] + \flat + w_{i,j}^\theta \tag{17}$$

where (x_t^j, y_t^j) are the Cartesian coordinates of target j, and (x_s^i, y_s^i) are the coordinates of sensor i. The assumption that $w_{i,j}^r$ and $w_{i,j}^\theta$ be uncorrelated does not always hold. However, there do not appear to be readily available probability distributions that can jointly represent the linear nature of the range measurement and the circular nature of the angular measurement.

Defining the vectors $\mathbf{s}$ and $\mathbf{t}$ to be the sets of all unknown sensor and target locations, the likelihood function is the product of the likelihoods for the ranges and angles

$$\Lambda(\mathbf{s},\mathbf{t},\theta_b) = \prod_{i,j} p_r(r_{i,j}|\mathbf{s},\mathbf{t},\flat) p_\theta(\theta_{i,j}|\mathbf{s},\mathbf{t},\flat) \tag{18}$$

where the product in (18) is over all pairs (i,j) where sensor i observes target j. It is assumed that the range measurements are corrupted with normally distributed noise as

$$p(r_{i,j}|\mathbf{s},\mathbf{t},\flat) \sim \mathcal{N}\left[\sqrt{(y_t^j - y_s^i)^2 + (x_t^j - x_s^i)^2}, (\sigma_{i,j}^r)^2\right]. \tag{19}$$

The distribution of $\theta_{i,j}$ depends upon which model from Section 2 we are using.

Determining the maximum of (18) is equivalent to finding the minimum of the negative log-likelihood of (18), designated as $\lambda(\mathbf{s},\mathbf{t})$, which (discarding constant terms) has the following form:

$$\begin{aligned}\lambda(\mathbf{s},\mathbf{t},\flat) &= \overbrace{\sum_{i,j} \frac{-1}{2(\sigma_{i,j}^r)^2}\left(r_{i,j} - \sqrt{(y_t^j - y_s^i)^2 + (x_t^j - x_s^i)^2}\right)^2}^{\lambda_r} \\ &\quad + \underbrace{\sum_{i,j} K_{s,t} f\left(\theta_{i,j} - \text{atan}_2\left[\frac{y_t^j - y_s^i}{x_t^j - x_s^i}\right] - \flat\right)}_{\lambda_\theta}. \end{aligned} \tag{20}$$

The value $K_{i,j}$ and function $f(\cdot)$ are given depending upon the distribution of $\theta_{i,j}$ according to Table I.

In the simulations, the minimization of (20) was carried out using the Quasi-Newton optimization algorithm [2]. For this the gradient of $\lambda(\mathbf{s},\mathbf{t},\flat)$ is needed. This gradient is the sum of the gradients of λ_r and λ_θ, which are defined in (20). The gradient elements of λ_r are given by

TABLE I
Values and Functions Dependent upon the Distribution of $\theta_{i,j}$ that are used in Expressions for the Likelihood (the function m is defined in (12))

	Clipped Mod Normal	von Mises	Wrapped Normal
$K_{i,j}$	$1/(2(\sigma_{i,j}^\theta)^2)$	$\kappa_{i,j}$	1
$f(\cdot)$	$m(\cdot)^2$	$\cos[\cdot]$	$\log[1 + 2\Sigma_{k=1}^\infty \rho^{k^2} \cos[k(\cdot)]$
$\mathfrak{F}(\cdot)$	$2m(\cdot)$	$\sin[\cdot]$	$(2\Sigma_{k=1}^\infty \rho^{k^2} k \sin[k(\cdot)])/$ $(1 + 2\Sigma_{k=1}^\infty \rho^{k^2} \cos[k(\cdot)])$
$\mathfrak{s}$	-1	1	1

$$\frac{\partial \lambda_r}{\partial a^i} = -\sum_j \frac{r_{i,j} - \sqrt{d_{i,j}}}{(\sigma_{i,j}^r)^2 \sqrt{d_{i,j}}} c_a^r(i,j) \tag{21}$$

$$\frac{\partial \lambda_r}{\partial b^j} = -\sum_i \frac{r_{i,j} - \sqrt{d_{i,j}}}{(\sigma_{i,j}^r)^2 \sqrt{d_{i,j}}} c_b^r(i,j) \tag{22}$$

$$\frac{\partial \lambda_r}{\partial \flat} = 0 \tag{23}$$

with $a \in \{x_s, y_s\}$ and $b \in \{x_t, y_t\}$. The constants are

$$d_{i,j} = (x_t^j - x_s^i)^2 + (y_t^j - y_s^i)^2 \tag{24}$$

$$c_{x_s}^r(i,j) = (x_t^j - x_s^i) \tag{25}$$

$$c_{y_s}^r(i,j) = (y_t^j - y_s^i) \tag{26}$$

$$c_{x_t}^r(i,j) = -(x_t^j - x_s^i) \tag{27}$$

$$c_{y_t}^r(i,j) = -(y_t^j - y_s^i). \tag{28}$$

With the quantities $\mathfrak{F}$ and $\mathfrak{s}$ given by Table I, the gradient elements of $\lambda_\theta(\mathbf{s},\mathbf{t})$ are

$$\frac{\partial \lambda_\theta}{\partial a^i} = \sum_j K_{i,j} \mathfrak{F}(\theta_{i,j} - \text{atan}_2[y_t^j - y_s^i, x_t^j - x_s^i] - \flat) c_a^\theta(i,j) \tag{29}$$

$$\frac{\partial \lambda_\theta}{\partial b^j} = \sum_i K_{i,j} \mathfrak{F}(\theta_{i,j} - \text{atan}_2[y_t^j - y_s^i, x_t^j - x_s^i] - \flat) c_b^\theta(i,j) \tag{30}$$

$$\frac{\partial \lambda_\theta}{\partial \flat} = \mathfrak{s} \sum_{i,j} K_{i,j} \mathfrak{F}(\theta_{i,j} - \text{atan}_2[y_t^j - y_s^i, x_t^j - x_s^i] - \flat) \tag{31}$$

where the c^θ terms are

$$c_{x_s}^\theta(i,j) = (y_t^j - y_s^i)/d_{i,j} \tag{32}$$

$$c_{y_s}^\theta(i,j) = -(x_t^j - x_s^i)/d_{i,j} \tag{33}$$

$$c_{x_t}^\theta(i,j) = -(y_t^j - y_s^i)/d_{i,j} \tag{34}$$

$$c_{y_t}^\theta(i,j) = (x_t^j - x_s^i)/d_{i,j} \tag{35}$$

and $d_{i,j}$ is as defined in (24).

In order to be able to perform likelihood maximization, initial estimates of the quantities being estimated are needed. The following section discusses how these may be obtained.

5. ALGORITHMS FOR GENERATING INITIAL ESTIMATES

5.1. Initial Estimates Without Range Measurements

1) *Joint Estimation of Sensor and Target Locations*: The angular measurement of sensor i observing target j, $\theta_{i,j}$, taken with respect to a known, common axis may be expressed as follows

$$\tan[\theta_{i,j}] = \frac{y_t^j - y_s^i}{x_t^j - x_s^i} \tag{36a}$$

$$\cot[\theta_{i,j}] = \frac{x_t^j - x_s^i}{y_t^j - y_s^i}. \tag{36b}$$

These equations may be rearranged to get

$$x_t^j \tan[\theta_{i,j}] - x_s^i \tan[\theta_{i,j}] - y_t^j + y_s^i = 0 \tag{37a}$$

$$y_t^j \cot[\theta_{i,j}] - y_s^i \cot[\theta_{i,j}] - x_t^j + x_s^i = 0. \tag{37b}$$

Thus, using (37a) and (37b), one can generate a linear system of equations that, in the absence of measurement noise and assuming that a sufficient number of equations are linearly independent, can be solved exactly for the sensor and target locations, given that enough sensors have a priori known locations. Sensors with known locations are necessary for the uniqueness of a nontrivial solution and are simply put on the right-hand side of the equation.

For example, consider the presence of three sensors observing two targets. Assuming that the location of sensors one and three are known a priori, the location of the two targets and the third sensor is given by the linear set of equations in (38), which is derived using (37a).

$$\underbrace{\begin{bmatrix} 0 & 0 & \tan[\theta_{1,1}] & -1 & 0 & 0 \\ -\tan[\theta_{2,1}] & 1 & \tan[\theta_{2,1}] & -1 & 0 & 0 \\ 0 & 0 & \tan[\theta_{3,1}] & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tan[\theta_{1,2}] & -1 \\ -\tan[\theta_{2,2}] & 1 & 0 & 0 & \tan[\theta_{2,2}] & -1 \\ 0 & 0 & 0 & 0 & \tan[\theta_{2,2}] & -1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_s^2 \\ y_s^2 \\ x_t^1 \\ y_t^1 \\ x_t^2 \\ y_t^2 \end{bmatrix}}_{\mathbf{s}}$$

$$= \underbrace{\begin{bmatrix} \tan[\theta_{1,1}]x_s^1 - y_s^1 \\ 0 \\ \tan[\theta_{3,1}]x_s^3 - y_s^3 \\ \tan[\theta_{1,2}]x_s^1 - y_s^1 \\ 0 \\ \tan[\theta_{3,2}]x_s^3 - y_s^3 \end{bmatrix}}_{\mathbf{b}}. \tag{38}$$

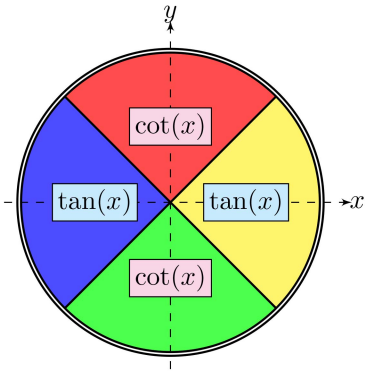


Fig. 4. The quadrants in which one should use the tangent or cotangent so as to minimize the effects of measurement error.

Thus, in this case, the location of the second sensor and the target at both times is the solution to

$$\mathbf{As} = \mathbf{b}. \tag{39}$$

A necessary condition for the observability of the system is that the locations of at least two sensors are known. However, the two sensors with known locations do not necessarily have to observe the same target at the same time for the matrix $\mathbf{A}$ to have full rank.

Due to the use of the tangent in (37a), serious estimation inaccuracies will occur if the targets used for estimation are close to $\pm 90^\circ$ with respect to any of the observing sensors. This is because the measurement error causes the measured angle to be above or below $\pm 90^\circ$, changing a very large positive entry in the $\mathbf{A}$ matrix to a very large negative value or vice versa. This problem can be minimized by using Equation (37b), which uses the cotangent, when the observation is between 45° and 135° or between -45° and -135°, as shown in Fig. 4.

2) *Estimation of the Sensor Locations Alone*: If the target locations are not needed, they can be eliminated from the estimation. We shall once again assume that all angles are taken with respect to a reference direction common for all sensors. In this subsection, we shall also assume that each target is observed simultaneously by at least three sensors with an appropriate (non-collinear) geometry. We define the measurements as being taken with respect to the x-axis in our 2-D coordinate system. For simplicity of notation, let us define the following functions

$$\Delta_{a,b}^T(j) \triangleq \tan[\theta_{a,j}] - \tan[\theta_{b,j}] \tag{40}$$

$$\Delta_{a,b}^C(j) \triangleq \cot[\theta_{a,j}] - \cot[\theta_{b,j}] \tag{41}$$

$$\Psi_{a,b}(j) \triangleq 1 - \cot[\theta_{a,j}]\tan[\theta_{b,j}] \tag{42}$$

where a and b are sensor indices and j is a target index.

As proven in Appendix A, given any three sensors simultaneously observing the target, one can combine (36a) and (40) using (37a) for each sensor to get an expression relating the sensor locations independent

of the Cartesian location of the target. For sensors 1 through 3, this gives us equations (43a)–(43d).

$$0 = y_s^1 \Delta_{2,3}^T(j) + x_s^1 \tan[\theta_{1,j}]\Delta_{3,2}^T(j) + y_s^2 \Delta_{3,1}^T(j) + x_s^2 \tan[\theta_{2,t}]\Delta_{1,3}^T(j) + y_s^3 \Delta_{1,2}^T(j) + x_s^3 \tan[\theta_{3,j}]\Delta_{2,1}^T(j) \quad (43a)$$

$$0 = y_s^1 \cot[\theta_{1,j}]\Delta_{2,3}^T(j) + x_s^1 \Delta_{3,2}^T(j) - y_s^2 \Psi_{1,3}(j) + x_s^2 \tan[\theta_{2,j}]\Psi_{1,3}(j) + y_s^3 \Psi_{1,2}(j) - x_s^3 \tan[\theta_{3,j}]\Psi_{1,2}(j) \quad (43b)$$

$$0 = y_s^1 \cot[\theta_{1,j}]\Psi_{2,3}(j) - x_s^1 \Psi_{2,3}(j) - y_s^2 \cot[\theta_{2,j}]\Psi_{1,3}(j) + x_s^2 \Psi_{1,3}(j) + y_s^3 \Delta_{2,1}^C(j) + x_s^3 \tan[\theta_{3,j}]\Delta_{1,2}^C(j) \quad (43c)$$

$$0 = y_s^1 \cot[\theta_{1,j}]\Delta_{3,2}^C(j) + x_s^1 \Delta_{2,3}^C(j) + y_s^2 \cot[\theta_{2,j}]\Delta_{1,3}^C(j) + x_s^2 \Delta_{3,1}^C(j) + y_s^3 \cot[\theta_{3,j}]\Delta_{2,1}^C(j) + x_s^3 \Delta_{1,2}^C(j). \quad (43d)$$

As was the case in the previous section, the equations derived in this section can be used with multiple observations of the targets over time to reduce the solution of sensor locations to that of solving $\mathbf{As} = \mathbf{b}$, where in this case $\mathbf{s}$ consists of only the sensor locations.

Note that as the number of sensors increases, the number of possible equations that can be written increases rapidly. However, the equations are not all independent. For example, for N sensors observing a common target, there are $\binom{N}{3}$ possible variants of (43a) that can be written depending upon which three targets are put into the equation. However, for $N > 3$ only N of these equations are linearly independent and the rest do not provide any new information, because they are not based on new observations. Linearly dependent equations may be removed by using the Modified Gram-Schmidt Orthonormalization Algorithm [12] or other, similar methods, though, as demonstrated in Section 7, this can hurt the performance of the algorithm.

5.2. Initial Estimates With Range Measurements

1) *Jointly Estimating Sensor and Target Locations*: When range measurements are available, the estimation problem becomes much simpler. Letting the range measurement of sensor i observing target j be $r_{i,j}$, we can write

$$r_{i,j}\cos[\theta_{i,j}] = x_t^j - x_s^i \quad (44)$$

$$r_{i,j}\sin[\theta_{i,j}] = y_t^j - y_s^i. \quad (45)$$

As was true in the angular case we can collect these linear equations and solve them for the sensor and target locations. In this instance, the $\mathbf{A}$ matrix is particularly simple, being composed only of ± 1 and 0 elements.

2) *Estimating the Sensor Locations Alone*: When two sensors simultaneously observe the same target, we can eliminate the target location from the estimation problem by manipulating (45) and (44) to get

$$r_{2,j}\cos[\theta_{2,j}] - r_{1,j}\cos[\theta_{1,j}] = x_s^1 - x_s^2 \quad (46)$$

$$r_{2,j}\sin[\theta_{2,j}] - r_{1,j}\sin[\theta_{1,j}] = y_s^1 - y_s^2. \quad (47)$$

As was the case in Section 5.1.2, we can again use the equations to find initial estimates based on a linear least squares solution.

5.3. Measurements with Respect to an Unknown, Common Axis

We consider the case where all sensors have the same unknown bias in their measurements. This might occur, for example, if all measurements are taken with respect to magnetic north, but the anchor node locations are given in terms of geographic north.

Figure 5 illustrates the scaling uncertainty that arises when only one anchor node is used. The dark circles represent sensors and the open circles anchor nodes. The dotted lines are only present to show that the transforms considered are affine (they do not distort the relative angles). In the noiseless case, if we were to remove the second anchor node and not compensate for the bias, then Figs. 5(b), (c), and (d) are three possible solutions for the system described by $\mathbf{As} = \mathbf{b}$ using the equations for angle-only observations from Section 5.1 (when range measurements are available, then only one solution exists). All of the biased solutions are rotated by the bias angle. As shown in (b) and (c), the figure can be scaled about the single anchor node without changing any of the measured angles (in the case where angles are measured to targets, the apparent locations of the targets are scaled as well). Figure 5(d) comes about due to our use of the tangent and cotangent in Section 5.1, whereby the equations do not change if all angles are flipped 180°.

In the case of only two anchor nodes and angle-only measurements, a method of estimating the sensor locations while correcting for the unknown global rotation is as follows:

1) Find an observation from the first anchor node. Assume that it is at a known, fixed distance from the first anchor node (such as 10 m). Find its location using the bias measurement under this assumption. This shall be a pseudo-anchor node.

2) Perform the sensor location estimation as described in Section 5.1 using the biased measurements, the first anchor node and the previously determined pseudo-anchor node as an anchor node assuming that the location of the second anchor node is unknown.

3) Find the vector between the first anchor node and the true location of the second anchor node (for example, for the scenario in Fig. 5(a), it has been

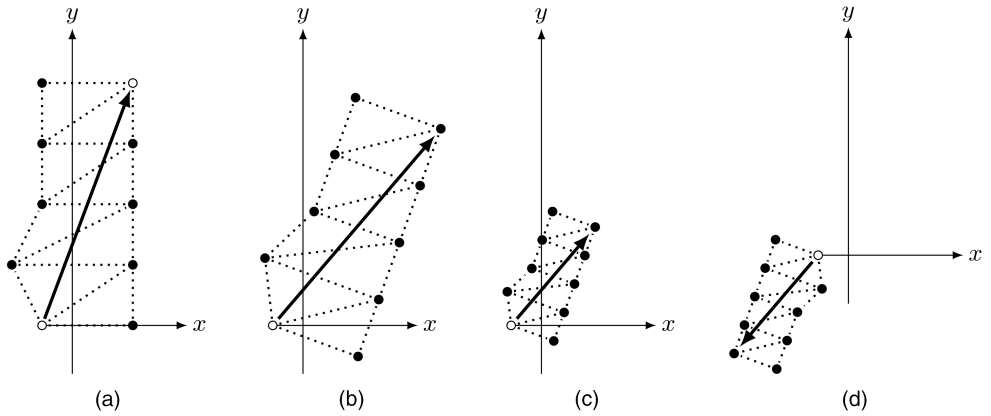


Fig. 5. The open circles represent the anchor sensors, whose locations are known a priori. The dotted lines show that the transformation being considered is affine (does not distort the relative angles between the sensors). The array points from the first (fixed) anchor node to the one that is removed. The array is the vector between the first and last nodes. Subfigure (a) shows the true setup of the problem. Subfigures (b), (c) and (d) show possible solutions when the second anchor node is removed and a common bias is left uncompensated. All scaled solutions are angularly consistent—even, as illustrated in (d), when the scaling factor is negative. (a) True sensor locations. (b) One biased solution. (c) Second biased solution. (d) Third biased solution.

drawn). We shall call this $\mathbf{v}_1$. Also find the vector between the first anchor node and the apparent position of the second anchor node as given by the previous estimation (such as the vectors in Fig. 5(b) or 5(c)). The choice of the pseudo-anchor node rules out the geometry of 5(d)); we shall call this $\mathbf{v}_2$.

4) Evaluate $\theta = \angle\mathbf{v}_2 - \angle\mathbf{v}_1$.

5) Perform the sensor location estimation again using the adjusted angles and both of the true anchor nodes to get a final estimate of the sensor locations.

The algorithm finds a solution for the biased system and then compares how that solution is rotated with respect to the true system. The first step of the aforementioned method creates a pseudo anchor point to set a reference for the scaling of the solution. This is important to make sure that we do not get a solution that is inverted by 180°, as in Fig. 5(d). Moreover, it is necessary for setting the scale of the figure. We would like to find a solution, but one possible solution places the nodes infinitesimally close to the first anchor point. The use of such a solution would be subject to precision problems on any computer.

A similar procedure can be performed if range measurements are available. In this case, there is no need to designate any node as a pseudo anchor node. When more sensors are present, one may break the observations into subsets according to the connectivity between anchor nodes, and calculate separate biases before averaging them. We shall not consider that case here.

6. THE CRAMÉR-RAO LOWER BOUND

In order to evaluate the efficiency of the estimator (how well the estimator is performing compared to a lower bound on the unbiased estimator), the Cramér Rao Lower Bound (CRLB) for the particular scenarios must be calculated [1]. The CRLB provides a lower bound on the covariance matrix of an unbiased estimator as

$$\mathrm{E}\{[\hat{\mathbf{x}} - \mathbf{x}_0][\hat{\mathbf{x}} - \mathbf{x}_0]^T\} \geq J^{-1} \tag{48}$$

where $\mathbf{x}$ is a vector parameter, $\hat{\mathbf{x}}$ is the parameter estimate, $\mathbf{x}_0$ is the true parameter value, and J is the Fisher Information Matrix (FIM).

The FIM is defined as

$$J \triangleq \mathrm{E}\{[\nabla_{\mathbf{x}}\lambda(\mathbf{x})][\nabla_{\mathbf{x}}\lambda(\mathbf{x})]^T\}|_{\mathbf{x}=\mathbf{x}_0}. \tag{49}$$

In the context of the problem at hand, $\mathbf{s}$ and $\mathbf{t}$ correspond to the variable $\mathbf{x}$. The appropriate diagonal entries of J^{-1} provide a lower bound for the mean squared error (MSE) of each estimated parameter, assuming that the estimator is unbiased. The FIM may be estimated by averaging values of $[\nabla_{\mathbf{x}}\lambda(\mathbf{x})][\nabla_{\mathbf{x}}\lambda(\mathbf{x})]^T|_{\mathbf{x}=\mathbf{x}_0}$ across Monte Carlo runs. For the case where the angular measurements have a von Mises distribution, an exact closed-form solution (in terms of modified Bessel functions) for the elements of the FIM will be presented. In the simulations when using the wrapped normal distribution, the CRLB was estimated by averaging the squared gradient across Monte Carlo runs.

Consider the elements of the FIM for normally distributed range measurements and von Mises distributed angular measurements taken with respect to a common, unknown axis (if range or angular measurements are not available or the measurement axis is known, then the appropriate terms may be omitted). Each element of the FIM is the expectation of a product of sums. Note that the expectation of cross terms[6] in the product sums involving angular measurements is zero. This is because the variables in question are independent and, thus the expectation of the product is equal to the product of the

[6]A cross term is a product such that elements involving (i_1, j_1) and (i_2, j_2) are multiplied where $i_1 \neq i_2$ or $j_1 \neq j_2$.

expectations. Based on (29) and (30), the expectations have the following form,

$$\int_{-\pi}^{\pi} \sin[\theta-\mu] e^{\kappa\cos[\theta-\mu]} d\theta = 0. \tag{50}$$

Note that the product of cross terms involving λ_r are zero, because again the variables in question are independent and the expectation can be factorized. Based on (21) and (22), the expectations shall have the following form

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \frac{r-\sqrt{d}}{\sigma^2\sqrt{d}} e^{-(r-\sqrt{d})^2/2\sigma^2} dr = 0. \tag{51}$$

Due to the assumed independence of the noise corrupting the range measurement and that corrupting the angular measurement, all cross terms between derivatives of λ_r and λ_θ are zero. Thus, we can write

$$J = \overbrace{\mathrm{E}\{[\nabla_{\mathbf{x}}\lambda_r(\mathbf{x})][\nabla_{\mathbf{x}}\lambda_r(\mathbf{x})]^T\}}^{J_r} + \overbrace{\mathrm{E}\{[\nabla_{\mathbf{x}}\lambda_\theta(\mathbf{x})][\nabla_{\mathbf{x}}\lambda_\theta(\mathbf{x})]^T\}}^{J_\theta}. \tag{52}$$

Let us compute J_θ. Because all of the cross terms are zero, we only need to concern ourselves with the expectation of the product of the gradient elements with the same (i,j) values. To simplify things, we shall note that

$$I_1(\kappa) = \frac{\kappa}{2\pi} \int_{-\pi}^{\pi} \sin^2[\theta-\mu] e^{\kappa\cos[\theta-\mu]} d\theta \tag{53}$$

which does not depend on μ. Thus, taking the expected value over the elements in the FIM, we get

$$\mathrm{E}\left[\frac{\partial\lambda_\theta}{\partial\alpha_s^{i_1}}\frac{\partial\lambda_\theta}{\partial\beta_s^{i_2}}\right] = \begin{cases} \sum_j \kappa_{i,j}\rho_{i,j} c^\theta_{\alpha_s}(i,j) c^\theta_{\beta_s}(i,j) & \\ \qquad\qquad \text{if} \quad i_1 = i_2 = i & \\ 0 \qquad\quad \text{otherwise} & \end{cases} \tag{54}$$

$$\mathrm{E}\left[\frac{\partial\lambda_\theta}{\partial\alpha_t^{j_1}}\frac{\partial\lambda_\theta}{\partial\beta_t^{j_2}}\right] = \begin{cases} \sum_i \kappa_{i,j}\rho_{i,j} c^\theta_{\alpha_t}(i,j) c^\theta_{\beta_t}(i,j) & \\ \qquad\qquad \text{if} \quad j_1 = j_2 = j & \\ 0 \qquad\quad \text{otherwise} & \end{cases} \tag{55}$$

$$\mathrm{E}\left[\frac{\partial\lambda_\theta}{\partial\alpha_s^{i}}\frac{\partial\lambda_\theta}{\partial\beta_t^{j}}\right] = \kappa_{i,j}\rho_{i,j} c^\theta_{\alpha_s}(i,j) c^\theta_{\beta_t}(i,j) \tag{56}$$

$$\mathrm{E}\left[\left(\frac{\partial\lambda_\theta}{\partial\flat}\right)^2\right] = \sum_{i,j} \kappa_{i,j}\rho_{i,j} \tag{57}$$

$$\mathrm{E}\left[\frac{\partial\lambda_\theta}{\partial\alpha_s^{i}}\frac{\partial\lambda_\theta}{\partial\flat}\right] = \sum_{j} \kappa_{i,j}\rho_{i,j} c^\theta_{\alpha_s}(i,j) \tag{58}$$

$$\mathrm{E}\left[\frac{\partial\lambda_\theta}{\partial\alpha_t^{j}}\frac{\partial\lambda_\theta}{\partial\flat}\right] = \sum_{i} \kappa_{i,j}\rho_{i,j} c^\theta_{\alpha_t}(i,j) \tag{59}$$

where $(\alpha,\beta) \in \{x,y\}$ and the mean resultant lengths are

$$\rho_{i,j} = \frac{I_1(\kappa_{i,j})}{I_0(\kappa_{i,j})}. \tag{60}$$

The calculation of the ratio of modified Bessel functions in the CRLB can be problematic, as mentioned in Section 2.3. An algorithm for calculating the ratio is discussed in Appendix B.

Now let us consider the calculation of J_r.

$$\mathrm{E}\left[\frac{\partial\lambda_r}{\partial\alpha_s^{i_1}}\frac{\partial\lambda_r}{\partial\beta_s^{i_2}}\right] = \begin{cases} \sum_j \frac{1}{d_{i,j}(\sigma^r_{i,j})^2} c^r_{\alpha_s}(i,j) c^r_{\beta_s}(i,j) & \\ \qquad\qquad \text{if} \quad i_1 = i_2 = i & \\ 0 \qquad\quad \text{otherwise} & \end{cases} \tag{61}$$

$$\mathrm{E}\left[\frac{\partial\lambda_r}{\partial\alpha_t^{j_1}}\frac{\partial\lambda_r}{\partial\beta_t^{j_2}}\right] = \begin{cases} \sum_i \frac{1}{d_{i,j}(\sigma^r_{i,j})^2} c^r_{\alpha_t}(i,j) c^r_{\beta_t}(i,j) & \\ \qquad\qquad \text{if} \quad j_1 = j_2 = j & \\ 0 \qquad\quad \text{otherwise} & \end{cases} \tag{62}$$

$$\mathrm{E}\left[\frac{\partial\lambda_r}{\partial\alpha_s^{i}}\frac{\partial\lambda_r}{\partial\beta_t^{j}}\right] = \frac{1}{d_{i,j}(\sigma^r_{i,j})^2} c^r_{\alpha_s}(i,j) c^r_{\beta_t}(i,j) \tag{63}$$

$$\mathrm{E}\left[\left(\frac{\partial\lambda_r}{\partial\flat}\right)^2\right] = \mathrm{E}\left[\frac{\partial\lambda_r}{\partial\alpha_s^{i}}\frac{\partial\lambda_r}{\partial\flat}\right] = \mathrm{E}\left[\frac{\partial\lambda_r}{\partial\alpha_t^{j}}\frac{\partial\lambda_r}{\partial\flat}\right] = 0. \tag{64}$$

7. SIMULATIONS

7.1. The Scenario

We used a scenario involving ten sensors and four targets over 20 time-steps (in the equations, observations of the same target at a different times are treated as separate "targets"). The sensors were placed in x locations in the set of $\{-50, 100\}$ meters and y locations in the set of $\{0, 100, 200, 300, 400\}$ meters, with the exception of the one that would have been at $(-50, 100)$, which was instead set to $(-100, 100)$ in order to break the symmetry of the arrangement so that it would be clear if poor estimates flipped anything. This configuration is shown in Fig. 6. The locations of the sensors at $(-50, 0)$ and $(100, 400)$ were assumed to be known a priori, and they were used as anchor arrays.

The first target was located at an x location of -250 meters and traveled at a constant speed from 20 to 380 meters in y. The second target was placed at an x location of 350 meters and traveled at a constant speed from 0 to 400 meters in y. The third target started at 400 meters in x, traveled at a constant speed to 800 meters by step 10 and came back to 400 meters in the x direction by step 20. In the y direction, it traveled at a constant speed from 20 to 380. The fourth target was placed at a y location of 500 meters and traveled at a constant speed from -600 to 1000 meters in x.

To demonstrate that unlike other algorithms, no target needs to be simultaneously visible to both anchor nodes, and the targets were only visible to a subset of the sensors at each time. From steps 1 to 4, only the sensors at y locations of 0 and 100 meters could see the targets. From steps 5 through 8, the sensors between 0 and 200 meters could see the targets. From steps 9 through 12

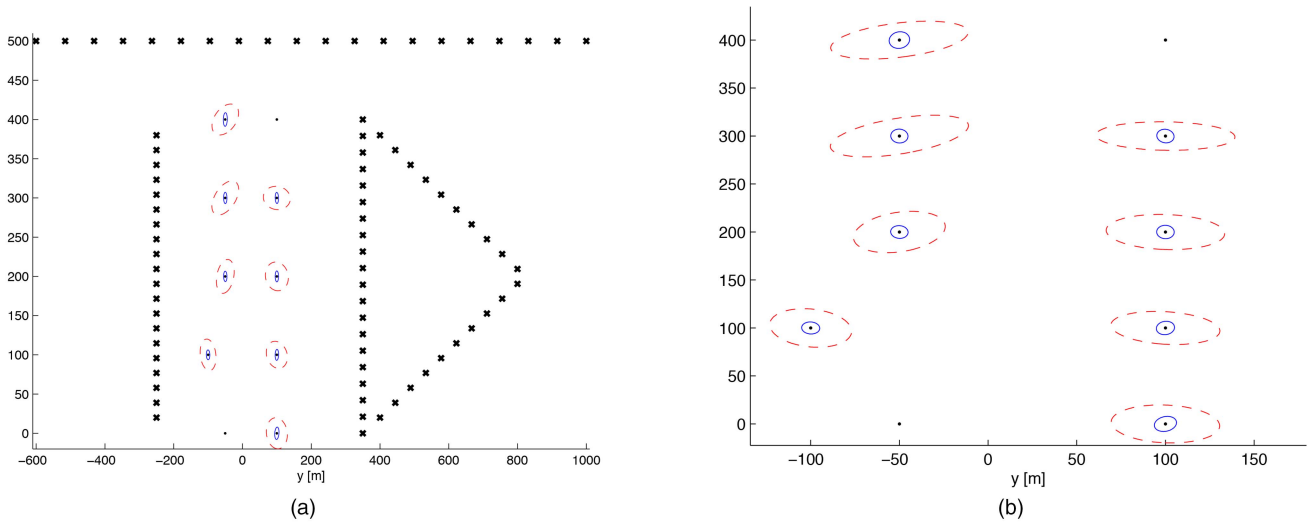


Fig. 6. The scenario showing the true sensor (the dots) and target locations (the ×s) used in the simulations. The anchor nodes are in the upper-right and lower-left corners. The ellipses represent the 99% confidence regions based on the CRLB for the sensor locations when the angular measurement axis is unknown. The dashed (outer) ellipses are using angle-only measurements using a von Mises angular noise distribution with $\sigma_\theta = 2^\circ = \pi/90$; the smaller, solid line ellipses are for the case with both angular and range measurements with $\sigma_r = 7.5$ meters. (a) Overall layout. (b) Magnified ellipses.

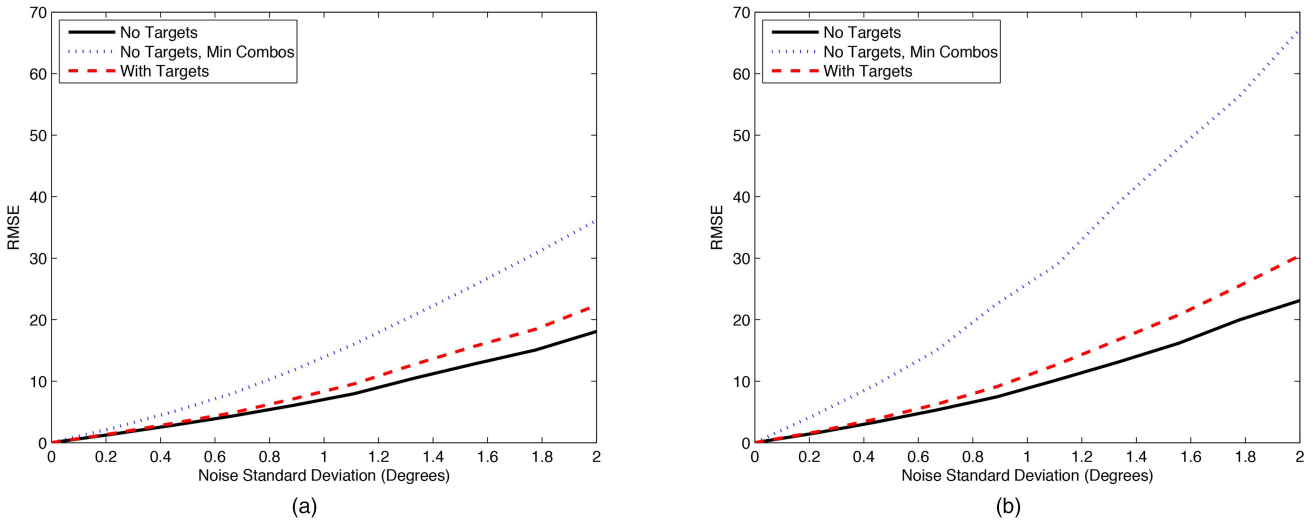


Fig. 7. The RMSE of the estimated sensor locations of the initialization algorithms. In (a) the measurements are taken with respect to a common, known axis. In (b) they are taken with respect to a common, unknown axis. 1000 Monte Carlo runs were performed. (a) Known measurement axis. (b) Unknown measurement axis.

the sensors between 100 and 300 meters could see the targets. From steps 13 through 16 the sensors between 200 and 400 meters could see the targets and from steps 17 through 20 the sensors from 300 to 400 meters in y could see the target. This means that the two sensors with known locations *never both simultaneously saw any target*.

7.2. The Initialization Algorithms

We compared the performance of the angle-only initialization algorithms under both known and unknown measurement axes, as shown in Fig. 7, where the RMSE of the sensor location estimates is shown, averaged over all sensors having unknown locations. The line labeled "No Targets" is the algorithm from Section 5.1.2 where the sensor locations are estimated without explicitly estimating the target locations. The line labeled "No Targets, Min Combos" is the same, except redundant equations of the $\binom{N}{3}$ that could be generated for each set of N sensors observing a common target were eliminated using the Gram-Schmidt algorithm [12]. The line labeled "With Targets" is the RMSE of the sensors when the target locations are jointly estimated, as given in Section 5.1.1. In all cases the solution to $\mathbf{As} = \mathbf{b}$ was found using least squares. The angular measurements were generated using the wrapped normal distribution with independence between sensors. One thousand Monte Carlo runs were performed.

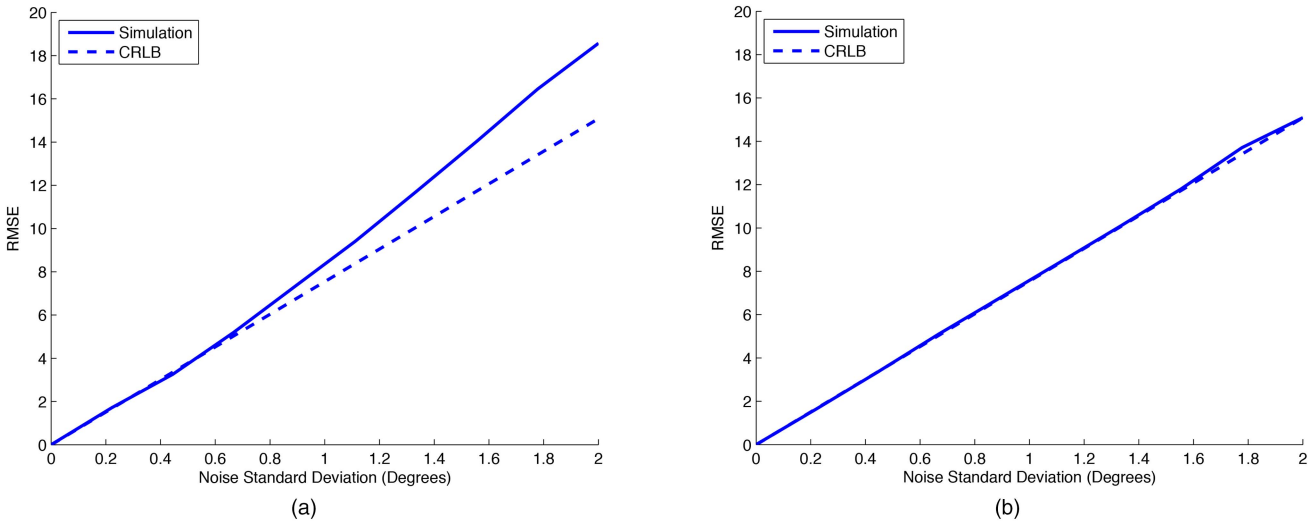


Fig. 8. The RMSE of the estimated sensor locations of the three scenarios compared to the CRLB under wrapped normal noise. In (a) we use the method of [16] utilizing a linear normal PDF that does not account for the circular nature of the measurements. In (b) we use the new ML method utilizing the clipped mod normal distribution assumption. 1,000 Monte Carlo runs were performed. (a) ML not corrected. (b) ML corrected.

The maximum noise standard deviation was set at 2°, corresponding to the accuracy of the acoustic sensors used in [19]. Higher noise standard deviations were found to produce occasionally very bad estimates (outliers). The likelihood of encountering such outliers varies depending upon the geometry. In many practical scenarios, this may not be a problem, since often, coarse estimates of the sensor locations can be obtained when they are placed and the initialization algorithm can be bypassed. The maximum noise standard deviation in the simulations in this paper was chosen sufficiently low such that extremely bad estimates did not occur, explaining the smoothness of the curves in the simulations.

7.3. ML Maximization

We compared the performance of the ML algorithm of [16] that does not take into account the circular nature of the measurements (assuming a linear normal distribution), with our ML algorithm (assuming a clipped normal distribution with the same standard deviation). The least-squares algorithm of Section 5.1.1 estimating both sensor and target locations was used to provide initial estimates. Measurements were generated using a wrapped normal distribution. All measurements were taken with respect to a common, known axis. 1000 Monte Carlo runs were performed. The results are shown in Fig. 8. Since the wrapping of the distributions depends upon where the π, $-\pi$ boundary is placed, rotating the global coordinate system changes the performance. However, if the ensemble of sensors make angular observations over the entire 360° range, then no rotation will exist where the basic linear model is nearly identical to the circular model.

In Table II, we numerically compare the effects of having different amounts of data regarding the scenario, in this case when using the von Mises noise distribution. The results are comparable when using the clipped mod normal distribution. The noise parameters for the sensors were $\sigma_\theta = 1°$ and, when range measurements were available, $\sigma_r = 7.5$ m.

TABLE II
The Average RMSE of the ML Estimates of the Sensor Locations Depending upon the Measurements Available for $\sigma_\theta = 1°$ and $\sigma_r = 7.5$ m as Obtained using the Initialization and Likelihood Maximization Algorithms Compared to the CRLB using von Mises Distributed Noise
(the results when using wrapped normal noise and performing ML maximization assuming the clipped mod normal density are similar; 1,000 Monte Carlo runs were performed)

Measurements		von Mises Simulated	von Mises CRLB
angle	known axis	5.466	5.437
	unknown axis	6.773	6.306
angle+range	known axis	2.037	2.049
	unknown axis	2.171	2.146

8. CONCLUSIONS

The importance of accounting for the circular nature of the data when performing sensor localization was highlighted. If the measurement noise is truly Gaussian, then the resulting noisy measurement will be wrapped on the unit circle, leading to the wrapped normal distribution. We introduced an approximation for the wrapped normal distribution that is very accurate for small to moderate circular standard deviation values. We derived simple linear least squares solutions for the target locations that we then used as initial estimates for performing ML estimation, as well as a method for handling a common, unknown measurement axis. When

using wrapped normally-distributed noise, our estimation method accounting for the circular nature of the data and using the clipped mod normal distribution for estimation proved efficient, whereas a previously introduced (and certainly decent) ML algorithm [16] that uses a linear noise model can diverge from the CRLB, dependent upon the geometry of the sensors and where the $\pi/-\pi$ boundary for the global coordinates is placed. We also quantified the effects of having different information available when estimating the sensor locations, including the availability of range measurements and knowledge of the common measurement axis of the sensors.

APPENDIX A. DERIVATION OF (43A)

Here we derive (43a), which underlies much of the algorithm. The derivations of (43b), (43c) and (43d) are performed similarly. Equation (36a) applied to the first sensor gives us

$$\tan[\theta_{1,j}] = \frac{y_t^j - y_s^1}{x_t^j - x_s^1} \tag{65}$$

$$y_t^j = y_s^1 + (x_t^j - x_s^1)\tan[\theta_{1,j}]. \tag{66}$$

Substituting (66) into (36a) applied to the second and third sensors gives us

$$\tan[\theta_{2,j}] = \frac{y_s^1 - y_s^2 + (x_t^j - x_s^1)\tan[\theta_{1,j}]}{x_t^j - x_s^2} \tag{67}$$

$$\tan[\theta_{3,j}] = \frac{y_s^1 - y_s^3 + (x_t^j - x_s^1)\tan[\theta_{1,j}]}{x_t^j - x_s^3}. \tag{68}$$

Solving (67) for the x location of the targets gives us

$$x_t^j = \frac{y_s^2 - y_s^1 + x_s^1\tan[\theta_{1,j}] - x_s^2\tan[\theta_{2,j}]}{\tan[\theta_{1,j}] - \tan[\theta_{2,j}]}. \tag{69}$$

Substituting (69) back into (68) and simplifying gives us the form of (43a).

APPENDIX B. CALCULATING BESSEL FUNCTION RATIOS

In [10] two methods were considered for converting a continued fraction representation of the ratio of two modified Bessel functions of the first kind into sums, allowing the computation of the ratio of Bessel functions without the overflow problems associated with calculating each function alone. Though the method attributed to Gauss had better asymptotic performance, it was demonstrated that the method attributed to Perron allowed for faster calculation of the ratio of two Bessel functions to an accuracy typically desired on a computer. The method based upon work by Perron is summarized as follows:

$$\frac{I_\nu(x)}{I_{\nu-1}(x)} = \sum_{k=0}^{\infty} c_k \tag{70}$$

where $\{x,\nu\} > 0$ and

$$c_k = \prod_{n=1}^{k} p_k \tag{71}$$

$$c_0 = 1 \tag{72}$$

$$p_1 = \frac{\frac{1}{2}x\left(\nu+\frac{1}{2}\right)}{\left(\nu+\frac{x}{2}\right)\left(\nu+x+\frac{1}{2}\right) - \frac{1}{2}x\left(\nu+\frac{1}{2}\right)} \tag{73}$$

$$p_k = \frac{\frac{1}{2}x\left(\nu+k-\frac{1}{2}\right)(1+p_{k-1})}{\left(\nu+x+\frac{k-1}{2}\right)\left(\nu+x+\frac{k}{2}\right) - \frac{1}{2}x\left(\nu+k-\frac{1}{2}\right)(1+p_{k-1})}. \tag{74}$$

Computation of the ratio of Bessel functions may thus be approximated by summing a suitable number of terms from (70). A suitable termination criterion for $\nu = 1$ is to stop when adding the next increment no longer changes the result. For example using double-precision arithmetic, for $\kappa = 3000$ terms $k \geq 5$ no longer change the result; for $\kappa = 11$ terms $k \geq 44$ no longer change the result.

REFERENCES

[1] Y. Bar-Shalom, X. Li, and T. Kirubarajan
Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software.
J. Wiley and Sons, 2001.

[2] D. P. Bertsekas
Nonlinear Programming.
Athena Scientific, 1999.

[3] P. Biswas and Y. Ye
Semidefinite programming for ad hoc wireless network localization.
In *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, Barkeley, CA, 2004, pp. 46–54.

[4] V. Chandrasekhar, W. Seah, Y. S. Choo, and H. V. Ee
Localization in underwater sensor networks—survey and challenges.
In *Proceedings of the First ACM International Workshop on Underwater Networks*, 2006, pp. 33–40. [Online], available: http://wuwnet.engr.uconn.edu/papers/p033-chandrasekhar.pdf.

[5] D. F. Crouse, Y. Bar-Shalom, and P. Willett
Sensor bias estimation in the presence of data association uncertainty.
In *Proceedings of SPIE, Signal and Data Processing of Small Targets Conference*, vol. 7445-22, Aug. 2009.

[6] M. Erol, L. Vieira, and M. Gerla
AUV-aided localization for underwater sensor networks.
In *Proceedings of the International Conference on Wireless Algorithms, Systems and Applications*, Aug. 2007, pp. 44–54.

[7] N. Fisher
Statistical Analysis of Circular Data.
Cambridge, UK: Cambridge University Press, 1993.

[8] N. Fisher, T. Lewis, and B. Embleton
Statistical Analysis of Spherical Data.
Cambridge: Cambridge University Press, 1987.

[9] A. Galstyan, B. Krishnamachari, K. Lerman, and S. Pattem
Distributed online localization in sensor networks using a moving target.
In *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, 2004, pp. 61–70.

[10] W. Gautschi and J. Slavik
On the computation of modified Bessel function ratios.
Mathematics of Computation, **32**, 143 (July 1978), 865–875.

[11] C. Gentile and J. Shiu
Sensor location through linear programming with arrival angle constraints.
In *Proceedings of the IEEE Information Theory Society Conference on Information Sciences and Systems*, Mar. 2007.

[12] G. H. Golub and C. F. van Loan
Matrix Computations (3rd ed.).
Johns Hopkins University Press, 1996.

[13] W-L. Huang, S-J. Li, and L. Duo
Designing reduced beacon trajectory for sensor localization.
Journal of Zheijang University, **8**, 12 (2007), 1971–1982.

[14] X. Lin, Y. Bar-Shalom, and T. Kirubarajan
Multisensor-multitarget bias estimation for general asynchronous sensors.
IEEE Transactions on Aerospace and Electronic Systems, **41**, 3 (2005), 899–921.

[15] K. V. Mardia and P. E. Jupp
Directional Statistics.
Chichester, England: John Wiley & Sons, 2000.

[16] R. L. Moses, D. Krishnamurthy, and R. Patterson
An auto-calibration method for unattended ground sensors.
In *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, vol. 3, 2002, pp. 2941–2944.

[17] A. Nasipuri and K. Li
A directionality based location discovery scheme for wireless sensor networks.
In *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, 2002, pp. 105–111.

[18] D. Niculescu and B. Nath
Ad-hoc positioning system (APS) using AOA.
In *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies*, San Francisco, CA, Mar. 2003.

[19] V. C. Ravindra, Y. Bar-Shalom, and T. Damarla
Feature-aided localization of ground vehicles using passive acoustic sensor arrays.
In *Proceedings of the 12th International Conference on Information Fusion*, Seattle, July 2009, pp. 70–77.

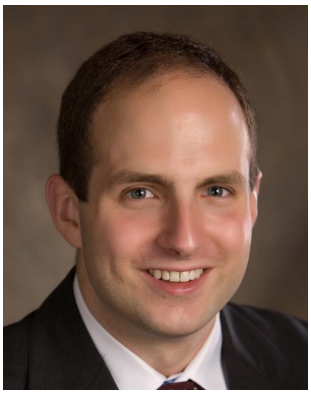
[20] B. Ristic and N. Okello
Sensor registration in ECEF coordinates using the MLR algorithm.
In *Proceedings of the Sixth International Conference of Information Fusion*, Cairns, Queensland, Australia, 2003, pp. 135–142.

[21] R. C. Shah, S. Roy, S. Jain, and W. Brunette
Data MULEs modeling a three-tier architecture for sparse sensor networks.
In *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications*, Seattle, May 2003, pp. 30–41.

[22] K-F. Ssu, C-H. Ou, and H. C. Jiau
Localization with mobile anchor points in wireless sensor networks.
IEEE Transactions on Vehicular Technology, **54**, 3 (May 2005), 1187–1197.

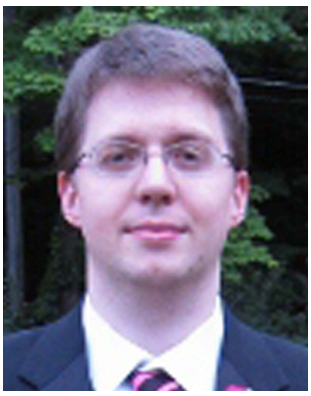
[23] G. Sun, J. Chen, W. Guo, and K. Liu
Signal processing techniques in network-aided positioning: A survey of state-of-the-art positioning designs.
IEEE Signal Processing Magazine, **22**, 4 (July 2005), 12–23.

[24] J. Zhang, M. Walpola, D. Roelant, H. Zhu, and K. Yen
Self-organization of unattended wireless acoustic sensor networks for ground target tracking.
Pervasive and Mobile Computing, **5**, 2 (Apr. 2009), 148–164.



David Frederic Crouse received his B.S., M.S., and Ph.D. degrees in electrical engineering, in 2005, 2008, and 2011, respectively, from the University of Connecticut (UCONN), Storrs, and his B.A. degree in German from the University of Connecticut in 2006.

In 2006 he spent a year with the Ruprecht-Karls-Universität, Heidelberg, Germany. He currently works at the Naval Research Laboratory in Washington, D.C. and is an associate editor for the *IEEE Aerospace and Electronic Systems Magazine*. His research interests lie in the areas of statistical signal processing and tracking.



Richard W. Osborne, III received his B.S., M.S., and Ph.D. degrees from the University of Connecticut in 2004, 2007, and 2012, respectively.

He is currently an assistant research professor in the Electrical and Computer Engineering Department at the University of Connecticut. In addition to his work at the University of Connecticut, he has consulted with a number of outside companies, and helped to develop and teach a short course on tracking and data fusion. His current research interests lie in the fields of estimation theory and information fusion.

Krishna R. Pattipati received his B.Tech. degree in electrical engineering with highest honors from the Indian Institute of Technology, Kharagpur, in 1975, and his M.S. and Ph.D. degrees in systems engineering from UConn, Storrs, in 1977 and 1980, respectively.

He was with ALPHATECH, Inc., Burlington, MA from 1980 to 1986. He has been with the Department of Electrical and Computer Engineering at UConn, where he is currently the UTC Professor of Systems Engineering. His current research activities are in the areas of agile planning, diagnosis and prognosis techniques for cyber-physical systems, multi-object tracking, and combinatorial optimization. A common theme among these applications is that they are characterized by a great deal of uncertainty, complexity, and computational intractability. He is a cofounder of Qualtech Systems, Inc., a firm specializing in advanced integrated diagnostics software tools (TEAMS, TEAMS-RT, TEAMS-RDS, TEAMATE), and serves on the board of Aptima, Inc.

Dr. Pattipati was selected by the IEEE Systems, Man, and Cybernetics (SMC) Society as the Outstanding Young Engineer of 1984, and received the Centennial Key to the Future award. He has served as the Editor-in-Chief of the *IEEE Transactions on Systems, Man, and Cybernetics—Part B* from 1998 to 2001, vice-president for technical activities of the IEEE SMC Society from 1998 to 1999, and as vice-president for conferences and meetings of the IEEE SMC Society from 2000 to 2001. He was corecipient of the Andrew P. Sage Award for the Best SMC Transactions Paper for 1999, the Barry Carlton Award for the Best AES Transactions Paper for 2000, the 2002 and 2008 NASA Space Act Awards for "A Comprehensive Toolset for Model-based Health Monitoring and Diagnosis," and "Real-time Update of Fault-Test Dependencies of Dynamic Systems: A Comprehensive Toolset for Model-Based Health Monitoring and Diagnostics," and the 2003 AAUP Research Excellence Award at UCONN. He also won the best technical paper awards at the 1985, 1990, 1994, 2002, 2004, 2005, and 2011 IEEE AUTOTEST Conferences, and at the 1997, 2004 Command and Control Conference. He is an elected Fellow of IEEE and of the Connecticut Academy of Science and Engineering.

Peter Willett is an IEEE Fellow and has been a faculty member in the Electrical and Computer Engineering Department at the University of Connecticut since 1986. His interests are in statistical signal processing, detection, machine learning, communications, data fusion, and tracking. He was Editor-in-Chief of *IEEE Transactions on AES* from 2006–2011. He has been Associate Editor of the *IEEE AES, SMC-A and SMC-B Transactions*. He remains Associate Editor of the *IEEE AES Magazine*, and ISIF's *Journal of Advances in Information Fusion*, and is also a member of the editorial board of IEEE's *Special Topics in Signal Processing* journal and a Senior Editor of *IEEE Signal Processing Letters*. He is a member of the IEEE AESS Board of Governors and is Vice-Chair of the IEEE Signal Processing Society's Sensor-Array and Multichannel Technical Committee.

Yaakov Bar Shalom was born on May 11, 1941. He received his B.S. and M.S. degrees from the Technion, Israel Institute of Technology, in 1963 and 1967 and the Ph.D. degree from Princeton University in 1970, all in electrical engineering.

From 1970 to 1976 he was with Systems Control, Inc., Palo Alto, CA. Currently he is Board of Trustees Distinguished Professor in the Department of Electrical and Computer Engineering and Marianne E. Klewin Professor in Engineering at the University of Connecticut. He is also Director of the ESP (Estimation and Signal Processing) Lab.

His current research interests are in estimation theory, target tracking and data fusion. He has published over 400 papers and book chapters in these areas and in stochastic adaptive control. He coauthored the monograph *Tracking and Data Association* (Academic Press, 1988), the graduate texts *Estimation and Tracking: Principles, Techniques* and *Software* (Artech House, 1993), *Estimation with Applications to Tracking and Navigation: Algorithms and Software for Information Extraction* (Wiley, 2001), the advanced graduate texts *Multitarget-Multisensor Tracking: Principles and Techniques* (YBS Publishing, 1995), *Tracking and Data Fusion* (YBS Publishing, 2011), and edited the books *Multitarget-Multisensor Tracking: Applications and Advances* (Artech House, Vol. I, 1990; Vol. II, 1992; Vol. III, 2000).

He has been elected Fellow of IEEE for "contributions to the theory of stochastic systems and of multi-target tracking." He has been consulting to numerous companies and government agencies, and originated the series of Multitarget-Multisensor Tracking short courses offered via UCLA Extension, at Government Laboratories, private companies and overseas.

During 1976 and 1977 he served as Associate Editor of the *IEEE Transactions on Automatic Control* and from 1978 to 1981 as Associate Editor of *Automatica*. He was Program Chairman of the 1982 American Control Conference, General Chairman of the 1985 ACC, and Cochairman of the 1989 IEEE International Conference on Control and Applications. During 1983–87 he served as Chairman of the Conference Activities Board of the IEEE Control Systems Society and during 1987–89 was a member of the Board of Governors of the IEEE CSS. He was a member of the Board of Directors of the International Society of Information Fusion (1999–2004) and served as General Chairman of FUSION 2000, President of ISIF in 2000 and 2002 and Vice President of Publications in 2004–11.

In 1987 he received the IEEE CSS Distinguished Member Award. Since 1995 he is a Distinguished Lecturer of the IEEE AESS and has given numerous keynote addresses at major national and international conferences. He is corecipient of the M. Barry Carlton Award for the best paper in the *IEEE Transactions on Aerospace and Electronic Systems* in 1995 and 2000 and recipient of the 1998 University of Connecticut AAUP Excellence Award for Research. In 2002 he received the J. Mignona Data Fusion Award from the DoD JDL Data Fusion Group. He is a member of the Connecticut Academy of Science and Engineering. In 2008 he was awarded the IEEE Dennis J. Picard Medal for Radar Technologies and Applications, and in 2012 the Connecticut Medal of Technology.

He has been listed by academic.research.microsoft (top authors in engineering) as #1 among the researchers in Aerospace Engineering based on the citations of his work.