Concurrent MAP Data Association and Absolute Bias Estimation with an Arbitrary Number of Sensors

Bret D. Kragel, Scott Danford and Aubrey B. Poore

Numerica Corporation, 4850 Hahns Peak Dr., Loveland, Colorado

ABSTRACT

Bias estimation using objects with unknown data association requires concurrent estimation of both biases and optimal data association. This report derives maximum *a posteriori* (MAP) data association likelihood ratios for concurrent bias estimation and data association based on sensor-level track state estimates and their joint error covariance. Our approach is unique for two reasons. First, we include a bias prior that allows estimation of absolute sensor biases, rather than just relative biases. Second, we allow concurrent bias estimation and association for an arbitrary number of sensors. The two-sensor likelihood ratio is derived as a special case of the general M-sensor result.

Keywords: Multiple target tracking, MAP bias estimation, data association, assignment problems, track-to-track association

1. INTRODUCTION

Bias estimation and residual bias mitigation are prerequisites for successful multiple sensor data fusion. Bias estimates can be based on filtering or batch least squares with or without process noise. For training data in the estimation problem, one frequently considers truth data, which may or may not be certain, or targets of opportunity for which the data association is known.[1,2] In addition, some biases are observable and others are not, depending to some extent on the training data. Bias estimates can also be classified as either absolute or relative as discussed by Herman.[1] Once estimates are determined and the sensor tracks and navigation errors are corrected, one must then address the residual bias problem; i.e., the remaining uncertainty in the stochastic bias estimates after measurements and transformations are corrected.[3]

At times, truth objects or targets of opportunity may not be available or may not be sufficient to yield maximum observability of the biases. In this case, one might consider turning to the targets themselves for which data association is unknown. The development of a bias estimation method using objects with unknown data association is a difficult one; however, much progress has been achieved in this direction for what is called the target object map problem as presented in the book by Blackman and Popoli [4, Section 10.5.5] and as further addressed by Levedahl's global nearest pattern algorithm.[5,6] In particular, Levedahl has developed an assignment algorithm that specifically allows optimal assignments in the presence of a persistent bias between two sets of data. The algorithm is based upon a Dijkstra search after organizing the assignment problem hypothesis space into a tree and, as the search is based upon dynamic programming, the solution found is guaranteed to be optimal. Levedahl's algorithm inherently supports finding the k-best rather than just the single best solution, and has proven to be significantly faster than approaches using an auction or JVC algorithm applied to a log-likelihood or χ_2-cost matrix, while also achieving much higher assignment accuracy in the presence of significant bias. Mori and Chong[7] also investigated several algorithms for two-sensor estimation of relative biases.

For the two-sensor joint MAP bias estimation and association problem, Poore et al. pursued a class of algorithms for absolute bias estimation based on the development of a good initial heuristic followed by the application of one of a class of branch and bound algorithms for constructing a guaranteed optimum.[8–10] The purpose of this paper is to formulate a joint multisensor likelihood function including a full bias prior that will allow adaptation of these techniques to the multisensor joint MAP bias estimation and association problem.

The organization of this paper is straightforward. A likelihood ratio appropriate for concurrent bias estimation and data association is formulated in Section 2, and the two-sensor likelihood ratio is derived as a special case. Section 3 provides a note on bias observability. Section 4 summarizes.

Further author information: (Send correspondence to B.D.K.)
B.D.K.: E-mail: bret.kragel@numerica.us, Telephone: 1 970 461 2000
R.S.D.: E-mail: scott.danford@numerica.us, Telephone: 1 970 461 2000
A.B.P.: E-mail: aubrey.poore@numerica.us, Telephone: 1 970 461 2000

Published in Signal and Data Processing of Small Targets 2008, SPIE Proc. 6969, Paper 6969-50.

2. THE MAP ASSIGNMENT FORMULATION

The evaluation of data association hypotheses has been considered by a number of authors, notably Reid,[11] Blackman and Popoli,[4] and Mori and Chong.[12,13] Recently, Kaplan and Blair[14] derived multisensor track-to-track assignment costs under the assumption that the estimation errors for separate sensors reporting on a common object were uncorrelated, while Bar-Shalom and Chen[15] developed likelihood ratios for the multisensor problem that included the cross-correlation between sensors tracking common objects. Our full maximum *a posteriori* formulation for concurrent multisensor track-to-track data association and bias estimation is similar in spirit to that of Mori and Chong,[12] where the case of two sensors without bias was considered.

Consider $n \geq 0$ objects moving independently within a surveillance region, denoted $\mathcal{X}$, common to $M \geq 2$ sensors, where neither the number of objects, nor their true positions x_i, $i = 1, \ldots, n$, is known *a priori*. Each sensor detects and tracks a subset of the objects; that is, we consider the possibility of missed detections. Periodically, the sensors report* the estimated track states (kinematic state and possibly feature and/or attribute data) to a central node for fusion. The reported sensor tracks take the form

$$\begin{aligned}
\hat{x}_{i_1} &= x_{i_1} - b_1 + \nu_{i_1}, \quad i_1 = 1, \ldots, n_1, \\
\hat{x}_{i_2} &= x_{i_2} - b_2 + \nu_{i_2}, \quad i_2 = 1, \ldots, n_2, \\
&\vdots \\
\hat{x}_{i_M} &= x_{i_M} - b_M + \nu_{i_M}, \quad i_M = 1, \ldots, n_M,
\end{aligned}$$

where for Sensor m, $m \in \{1, \ldots, M\}$, n_m denotes the number of tracks reported by the sensor, $\hat{x}_{i_m}$ denotes the estimated state of Track i_m, $i_m = 1, \ldots, n_m$, x_{i_m} is the corresponding true, but unknown, state of the object corresponding to Track i_m, b_m is an unknown translational bias error for Sensor m, and ν_{i_m} denotes the random estimation error.

The problem to be addressed is that of determining which, if any, of the sensor tracks emanate from common objects, while concurrently computing an estimate of the sensor biases, $b = (b_1, b_2, \ldots, b_M)^T$. Generalizing the work by Levedahl,[5,6] where only the *relative* bias between sensors was considered, prior bias estimates $\hat{b}$ can be included by utilizing the covariance $R_b = E[(b - \hat{b})(b - \hat{b})^T]$, which is assumed to be known.

To formalize the subsequent discussion, we follow Poore[16] and let $\mathcal{Z}(m) = \{\hat{x}_{i_m}\}_{i_m=1}^{n_m}$ denote the source tracks from Sensor m, $m = 1, \ldots, M$, while $\mathcal{Z} = \bigcup_{m=1}^{M} \mathcal{Z}(m)$ denotes the set of all source tracks. Similarly, $\mathcal{I} = \bigcup_{m=1}^{M} \mathcal{I}(m)$, with $\mathcal{I}(m) = \{i_m\}_{i_m=1}^{n_m}$, $m = 1, \ldots, M$, denotes the indices assigned to the tracks by the sensors. A partition H of $\mathcal{I}$ and the collection $\mathcal{H}$ of all such partitions are defined by

$$H = \{h_1, \ldots, h_{n(H)} \mid h_k \subset \mathcal{I},\ h_k \neq \emptyset,\ k = 1, \ldots, n(H)\}, \tag{1}$$

$$|h_k \cap \mathcal{I}(m)| \leq 1,\ k = 1, \ldots, n(H),\ m = 1, \ldots, M, \tag{2}$$

$$h_k \cap h_j = \emptyset,\ \text{for } k \neq j, \tag{3}$$

$$\mathcal{I} = \bigcup_{k=1}^{n(H)} h_k, \tag{4}$$

$$\mathcal{H} = \{H \mid H \text{ satisfies (1)} - \text{(4)}\}, \tag{5}$$

where the number $n(H)$ of subsets (or tracks) comprising a particular partition or hypothesis will vary from one hypothesis to another. We shall refer to the partitions $H \in \mathcal{H}$ as *track-to-track data association hypotheses* and the subsets $h_k \subset \mathcal{I}$ as *system tracks*. Conditions (3) and (4) ensure that each sensor track is assigned to one and only one system track, while condition (2) ensures that no system track contains more than one sensor track from any particular sensor; thus, every sensor track is assigned uniquely to a system track and no two tracks from the same sensor are assigned to the same system track.

Note that each system track is composed of a vector $h_k = \{h_{kj}\}_{j=1}^{n(h_k)}$ of sensor indices, where $n(h_k) = |h_k|$ denotes the cardinality of the set h_k; that is, $h_{kj} = i_m$, where i_m implies both a sensor, m, and a sensor track index, $i_m \in$

*The reported track state estimates are presumed to be time-aligned, either because the sensors report all tracks simultaneously, or by propagating the tracks to a common time.

$\{1, \ldots, n_m\}$. Through a slight abuse of notation[†], this allows us to conveniently reference the sensors and sensor tracks comprising each hypothesized track by appending the track notation as a subscript to the component of interest; thus, the notation $m \in h_k$ refers to the set of sensors postulated by H to contribute sensor tracks to System Track k. Similarly, $b_{h_{kj}}$ denotes the bias of the sensor reporting on h_{kj}, while b_{h_k} denotes the vector slice, ordered by sensor index, of the biases of all sensors postulated by H to report on System Track k. For example, if h_k is composed of, say, Track 3 from Sensor 2 and Track 1 from Sensor 5, then $b_{h_k} = (b_2, b_5)^T$.

In a slightly different sense, we use $\hat{x}_{h_k}$ to refer to the vector of sensor tracks, ordered by sensor index, postulated by h_k to emanate from a common object. In this manner, each partition (or hypothesis) H induces a partition of the data $\mathcal{Z}$ via

$$\mathcal{Z}_H = \{\mathcal{Z}_{h_1}, \ldots, \mathcal{Z}_{h_{n(H)}}\}, \quad \mathcal{Z}_{h_k} = \{\hat{x}_{h_{kj}}\}_{j=1}^{n(h_k)}.$$

Clearly, $\mathcal{Z}_{h_k} \cap \mathcal{Z}_{h_j} = \emptyset$ for $k \neq j$ and $\mathcal{Z} = \bigcup_{k=1}^{n(H)} \mathcal{Z}_{h_k}$.

Because estimates of the sensor biases and the assignment of sensor tracks to system tracks affect each other reciprocally, the optimal assignment and bias estimate can be found only by concurrent optimization of both; therefore, we seek the assignment (or hypothesis) $H^* \in \mathcal{H}$, and bias b^* for which

$$\Pr(H^*, b^* \,|\mathcal{Z}) = \max_{H,b} \Pr(H, b \,|\mathcal{Z}); \tag{6}$$

that is, we seek the association hypothesis and bias with the maximum *a posteriori* probability.

2.1 Computation of the Posterior Probabilities

In addition to the track-to-track hypothesis set $\mathcal{H}$, it will prove convenient to consider the set $\mathcal{A}_1$ of all bijective mappings from subsets of object indices $\{1, \ldots, n\}$ onto the set $\{1, \ldots, n_1\}$ of Sensor 1 indices. In this manner, each mapping $\alpha_1 \in \mathcal{A}_1$ hypothesizes the assignment of exactly one distinct truth object to each object reported by Sensor 1; that is, $\text{Dom}(\alpha_1) \subset \{1, \ldots, n\}$, $\text{Rng}(\alpha_1) = \{1, \ldots, n_1\}$ and $|\text{Dom}(\alpha_1)| = n_1$. Likewise, we denote by $\mathcal{A}_m$ the set of all bijective mappings from subsets of $\{1, \ldots, n\}$ onto the set $\{1, \ldots, n_m\}$ of Sensor m indices, $m = 2, \ldots, M$. As was the case for the sensor biases and sensor tracks, we can greatly simplify notation by letting $\alpha_{h_{kj}}^{-1} \triangleq \alpha_{h_{kj}}^{-1}(h_{kj})$ refer to the truth object postulated to generate the sensor track corresponding to h_{kj}[‡].

One possible situation is depicted in Figure 1. In this case, each of two sensors has detected and reported on 3 of the 5 objects that exist in the surveillance region. The Sensor 1 mapping α_1 is indicated by the solid arrows, and the Sensor 2 mapping α_2 by the dashed arrows. For the situation depicted in the figure, there are four system tracks in the correct hypothesis: The first track is composed of Sensor 1, Track 1 and Sensor 2, Track 3, the second track consists of Sensor 1, Track 3 and Sensor 2, Track 2, while the third and fourth tracks are singletons, consisting of Sensor 1, Track 2 and Sensor 2, Track 1, respectively.

Using Bayes' Rule and the total probability theorem, the posterior probability given the reported sensor tracks of any pair (H, b) can be written as

$$\begin{aligned}
\Pr(H, b \,|\mathcal{Z}) &= \frac{1}{p(\mathcal{Z})} p(\mathcal{Z}, b, H) \\
&= \mathcal{C} \sum_{n=0}^{\infty} p(\mathcal{Z}, b, H, n) \\
&= \mathcal{C} \sum_{n=0}^{\infty} \int_{\mathcal{X}^n} p(\mathcal{Z}, b, H, x^n, n)\, dx^n \\
&= \mathcal{C} \sum_{n=0}^{\infty} \sum_{\substack{\alpha_m \in \mathcal{A}_m \\ m=1,\ldots,M}} \int_{\mathcal{X}^n} p\left(\mathcal{Z}, b, H, \{\alpha_m\}_{m=1}^{M}, x^n, n\right) dx^n,
\end{aligned} \tag{7}$$

[†]The abuse of notation arises from allowing h_{kj} to reference a sensor in some contexts and a sensor track in others. Which is meant should always be clear from the context.

[‡]Note that the notation h_{kj} is used here to reference the reporting sensor in the subscript of $\alpha_{h_{kj}}^{-1}(h_{kj})$, but references the sensor track index in the argument. The intent should be clear from the definition of α.

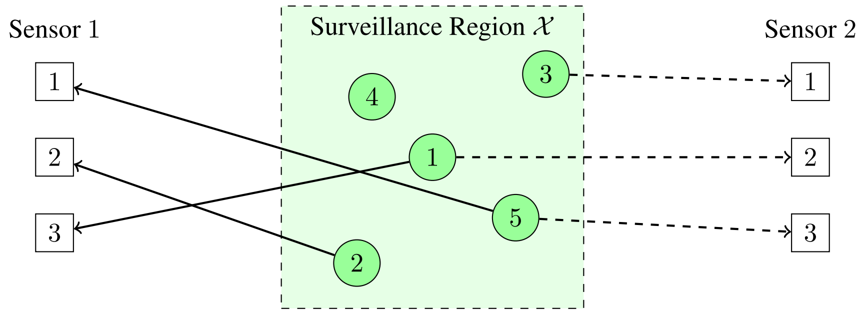

Figure 1. One possible assignment of truth objects to sensor tracks. In this case, there are 5 targets in the surveillance region, though each sensor has detected only 3 of the objects.

where the constant $\mathcal{C} = 1/p(\mathcal{Z})$ is independent of H and b, n is the number of objects in the surveillance region and $x^n = (x_1, \ldots, x_n)$ denotes the joint state of the objects in the surveillance region.

The goal of the expansion in (7) was to obtain an integrand $p\left(\mathcal{Z}, b, H, x^n, \{\alpha_m\}_{m=1}^M, n\right)$ that can be reduced to a form that leads to a computationally tractable assignment problem. We now proceed to derive the desired decomposition. Under the rules of conditional probability, the integrand can be written as

$$\begin{aligned} & p\left(\mathcal{Z}, b, H, \{\alpha_m\}_{m=1}^M, x^n, n\right) \\ & \quad = p\left(\mathcal{Z} \left| H, \{\alpha_m\}_{m=1}^M, b, x^n, n\right.\right) \Pr\left(H \left| \{\alpha_m\}_{m=1}^M, b, x^n, n\right.\right) \Pr\left(\{\alpha_m\}_{m=1}^M \left| b, x^n, n\right.\right) p\left(b, x^n \left| n\right.\right) \Pr(n) && (8) \\ & \quad = p\left(\mathcal{Z} \left| H, \{\alpha_m\}_{m=1}^M, b, x^n\right.\right) \Pr\left(H \left| \{\alpha_m\}_{m=1}^M, n\right.\right) \prod_{m=1}^M \Pr\left(\alpha_m \left| x^n\right.\right) p(b)\, p(x^n) \Pr(n)\,. && (9) \end{aligned}$$

Note that in going from (8) to (9), we have used the fact that the probability of hypothesis H, given the assignments $\alpha_1, \ldots, \alpha_M$, no longer depends on the joint object state x^n or the bias b, but only on whether the hypothesized associations in H are compatible with $\alpha_1, \ldots, \alpha_M$; thus, $\Pr\left(H \left| \{\alpha_m\}_{m=1}^M, b, x^n\right.\right) = \Pr\left(H \left| \{\alpha_m\}_{m=1}^M, n\right.\right)$. The original dependence on x^n is reflected in the term $\Pr\left(\{\alpha_m\}_{m=1}^M \left| x^n\right.\right)$, where the dependence on b was suppressed since we expect that sensor biases will impair only the ability of the sensor to accurately measure the target state, not its ability to detect the target. We have also assumed that the sensors independently detect and track the objects in the surveillance region, so that $\Pr\left(\{\alpha_m\}_{m=1}^M \left| x^n\right.\right) = \prod_{m=1}^M \Pr\left(\alpha_m \left| x^n\right.\right)$. Finally, we note that in terms involving x^n, the dependence on n is superfluous and was dropped in (9) to simplify notation.

We now examine the multiplicative terms of Equation (9) individually using the assumption that the *a priori* object states are independent:

- $p\left(\mathcal{Z} \left| H, \{\alpha_m\}_{m=1}^M, b, x^n\right.\right)$: Assuming the sensors are tracking the objects in their field-of-view independently, the tracks are independent conditioned on H, and

$$\begin{aligned} p\left(\mathcal{Z} \left| H, \{\alpha_m\}_{m=1}^M, b, x^n\right.\right) &= p\left(\{\{\hat{x}_{i_m}\}_{i_m=1}^{n_m}\}_{m=1}^M \left| H, \{\alpha_m\}_{m=1}^M, b, x^n\right.\right) \\ &= \prod_{h_k \in H} p\left(\hat{x}_{h_k} \left| x_{\alpha_{h_k}^{-1}}, b_{h_k}\right.\right). && (10) \end{aligned}$$

  Note that since $\alpha_{h_{k1}}^{-1} = \alpha_{h_{k2}}^{-1} = \cdots = \alpha_{h_{kn(h_k)}}^{-1}$ by hypothesis, it is sufficient to write $\alpha_{h_k}^{-1}$ when conditioning on the truth object in the densities of (10), which will be defined later.

- $\Pr\left(\alpha_m \left| x^n\right.\right)$: We have assigned unique indices to the sensor observations and the objects in $\mathcal{X}$; thus, assuming independence of object states, the probability of a particular assignment of objects to the Sensor m observations, $m = 1, \ldots, M$, is given by

$$\Pr\left(\alpha_m \left| x^n\right.\right) = \frac{1}{n_m!} \prod_{i \in \mathrm{Dom}(\alpha_m)} P_D^m(x_i) \prod_{i \notin \mathrm{Dom}(\alpha_m)} \left(1 - P_D^m(x_i)\right) \qquad (11)$$

System Track	Source Tracks
1	(S_1, T_1) (S_3, T_1)
2	(S_1, T_2)
3	(S_2, T_1)
4	(S_1, T_3) (S_2, T_2) (S_3, T_2)
5	(S_2, T_3) (S_3, T_3)

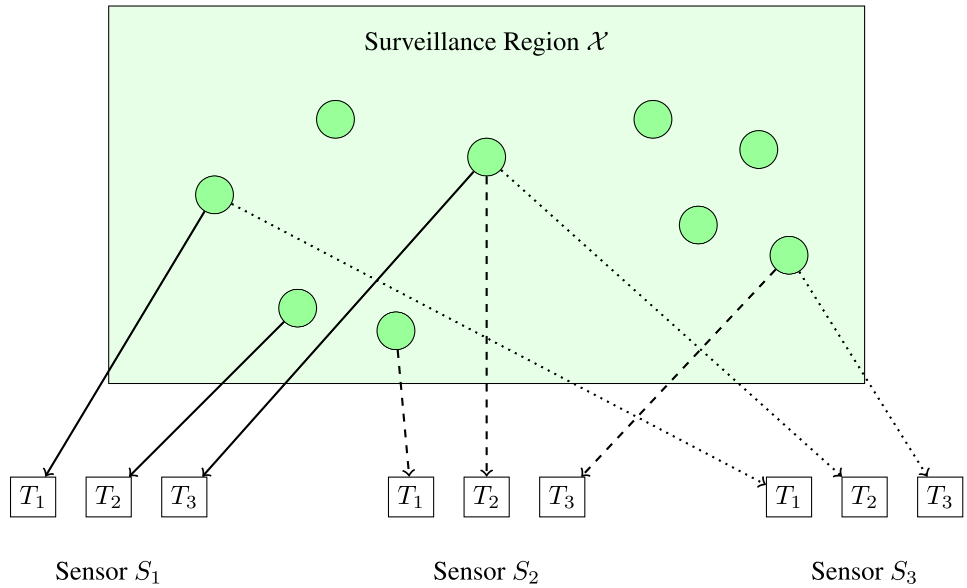

Figure 2. In this illustration, three sensors are each reporting on three objects. The table on the left defines one possible partition of the nine sensor tracks. This hypothesis is plausible only if the mappings from the truth objects to the sensor tracks are compatible with the table as depicted in the figure on the right. Here, the solid lines correspond to α_1, the dashed lines to α_2, and the dotted lines to α_3. If there are n truth objects, then there are n choices for the truth object corresponding to System Track 1, $n-1$ choices for System Track 2, etc., resulting in a total of $n(n-1)\cdots(n-n_D+1) = n!/(n-n_D)!$ compatible mappings, where $n_D = 5$ in this case.

$$= \mathcal{C} \prod_{i \in \text{Dom}(\alpha_m)} P_D^m(x_i) \prod_{i \notin \text{Dom}(\alpha_m)} (1 - P_D^m(x_i)), \tag{12}$$

where $P_D^m(x)$ denotes the probability that Sensor m detects an object at location x. We must divide by $n_m!$ for equality in (11) because we have assigned indices to the observations; that is, $\Pr(\alpha_m \,|x^n)$ is the probability of not only detecting the specified targets, but also of putting the detections into the prescribed order. The transition from (11) to (12) is achieved by noting that as part of the data $\mathcal{Z}$, n_m is independent of the hypothesis H, allowing $1/n_m!$ to be shifted to the constant $\mathcal{C}$.

- $\Pr\left(H \,\middle|\{\alpha_m\}_{m=1}^M, n\right)$: A given hypothesis H postulates the detection of $n_D = n(H)$ distinct objects; thus, given the bijective object-to-track assignments $\{\alpha_m\}_{m=1}^M$, the track-to-track hypothesis H is valid if and only if it is "compatible" with $\{\alpha_m\}_{m=1}^M$; that is,

$$\begin{gathered} n_D = \left| \bigcup_{m=1}^M \text{Dom}(\alpha_m) \right|, \\ \alpha_{h_{k1}}^{-1} = \alpha_{h_{k2}}^{-1} = \cdots = \alpha_{h_{kn(h_k)}}^{-1},\ k = 1, \ldots, n(H), \\ \alpha_{h_{kj}}^{-1} \neq \alpha_{h_{li}}^{-1},\ k \neq l,\ j = 1, \ldots, n(h_k),\ i = 1, \ldots, n(h_l). \end{gathered} \tag{13}$$

We therefore have

$$\Pr\left(H \,\middle|\{\alpha_m\}_{m=1}^M, n\right) = \begin{cases} 1, & \text{if conditions (13) hold,} \\ 0, & \text{otherwise.} \end{cases}$$

It will also prove necessary to know the number of M-tuples $(\alpha_m)_{m=1}^M$ for which the conditions (13) hold. This number can be found simply by counting all possible distinct ordered n_D-tuples from the set $\{1, \ldots, n\}$, which yields

$$n(n-1)\cdots(n-n_D+1) = \frac{n!}{(n-n_D)!} \tag{14}$$

such constructs (see Figure 2).

- $p(x^n)$: Under the assumption that the target states are *a priori* independently distributed throughout the surveillance region, we have

$$p(x^n) = \prod_{i=1}^{n} p(x_i). \tag{15}$$

Setting (10), (12) and (15) into (9) results in

$$\begin{aligned}
&p\left(\mathcal{Z}, b, H, \{\alpha_m\}_{m=1}^M, x^n, n\right) \\
&= \mathcal{C} p(b) \prod_{h_k \in H} p\left(\hat{x}_{h_k} \,\middle|\, x_{\alpha_{h_k}^{-1}}, b_{h_k}\right) \prod_{m=1}^{M} \left[\prod_{i \in \mathrm{Dom}(\alpha_m)} P_D^m(x_i) \prod_{i \notin \mathrm{Dom}(\alpha_m)} \left(1 - P_D^m(x_i)\right) \right] \\
&\qquad \cdot \Pr\left(H \,\middle|\, \{\alpha_m\}_{m=1}^M\right) \prod_{i=1}^{n} p(x_i) \Pr(n) \qquad (16) \\
&= \mathcal{C} p(b) \Pr\left(H \,\middle|\, \{\alpha_m\}_{m=1}^M\right) \Pr(n) \prod_{i \notin \bigcup_{m=1}^M \mathrm{Dom}(\alpha_m)} \prod_{m=1}^{M} \left(1 - P_D^m(x_i)\right) p(x_i) \\
&\qquad \cdot \prod_{h_k \in H} \left[p\left(\hat{x}_{h_k} \,\middle|\, x_{\alpha_{h_k}^{-1}}, b_{h_k}\right) \prod_{m \in h_k} P_D^m\left(x_{\alpha_{h_k}^{-1}}\right) \prod_{m \notin h_k} \left(1 - P_D^m\left(x_{\alpha_{h_k}^{-1}}\right)\right) p\left(x_{\alpha_{h_k}^{-1}}\right) \right], \qquad (17)
\end{aligned}$$

where we have simply rearranged terms in going from (16) to (17).

We can now replace the integrand in (7) with (17). Since we cannot distinguish *a priori* the objects in $\mathcal{X}$, their (not yet specified) *a priori* distribution is invariant with respect to order, allowing us to drop the indices on the object states under integration; thus, (7) takes the simplified form

$$\begin{aligned}
\Pr(H, b \,|\, \mathcal{Z}) &= \mathcal{C} p(b) \sum_{n=0}^{\infty} \Pr(n) \sum_{\substack{\alpha_m \in \mathcal{A}_m \\ m=1,\ldots,M}} \Pr\left(H \,\middle|\, \{\alpha_m\}_{m=1}^M\right) \frac{1}{\hat{\eta}^n} \left(\prod_{i=1}^{n-n_D} \int_{\mathcal{X}} \prod_{m=1}^{M} \left(1 - P_D^m(x)\right) \hat{\eta} p(x)\, dx \right) \\
&\qquad \cdot \prod_{h_k \in H} \int_{\mathcal{X}} p\left(\hat{x}_{h_k} \,|\, x, b_{h_k}\right) \prod_{m \in h_k} P_D^m(x) \prod_{m \notin h_k} \left(1 - P_D^m(x)\right) \hat{\eta} p(x)\, dx \qquad (18) \\
&= \mathcal{C} p(b) \sum_{n=0}^{\infty} \frac{n!}{(n-n_D)!} \frac{k_D^{(n-n_D)}}{\hat{\eta}^n} \Pr(n) \\
&\qquad \cdot \prod_{h_k \in H} \int_{\mathcal{X}} p\left(\hat{x}_{h_k} \,|\, x, b_{h_k}\right) \prod_{m \in h_k} P_D^m(x) \prod_{m \notin h_k} \left(1 - P_D^m(x)\right) \hat{\eta} p(x)\, dx. \qquad (19)
\end{aligned}$$

For later convenience, we have introduced the object intensity $\hat{\eta}$ into (18) by multiplying with $\hat{\eta}^n / \hat{\eta}^n$. In going from (18) to (19), we have simply set $k_D = \prod_{m=1}^M \int_{\mathcal{X}} (1 - P_D^m(x)) \hat{\eta} p(x)\, dx$, since this value is invariant with respect to the hypothesis under our assumptions.

If we now assume, as is common in the literature, that the number of targets in the surveillance region is Poisson distributed with intensity $\hat{\eta}$, then

$$\sum_{n=0}^{\infty} \frac{n!}{(n-n_D)!} \frac{k_D^{(n-n_D)}}{\hat{\eta}^n} \Pr(n) = \sum_{n=n_D}^{\infty} \frac{n!}{(n-n_D)!} \frac{k_D^{(n-n_D)}}{\hat{\eta}^n} \frac{\hat{\eta}^n e^{-\hat{\eta}}}{n!} = e^{-(\hat{\eta} - k_D)},$$

which is also invariant with respect to hypothesis and can be put into the constant $\mathcal{C}$. This leaves us with

$$\Pr(H, b \,|\, \mathcal{Z}) = \mathcal{C} p(b) \prod_{h_k \in H} \int_{\mathcal{X}} p\left(\hat{x}_{h_k} \,|\, x, b_{h_k}\right) \prod_{m \in h_k} P_D^m(x) \prod_{m \notin h_k} \left(1 - P_D^m(x)\right) \hat{\eta} p(x)\, dx.$$

Making the additional assumption (also common in the literature) that the probability of detection is independent of target position, and that the prior density of the true target location is uniform throughout the surveillance region with $p(x) = 1/V$, where V is the volume of the surveillance region $\mathcal{X}$, we can write[17]

$$\Pr(H, b\,|\mathcal{Z}) = \mathcal{C}p(b) \prod_{h_k \in H} \int_{\mathcal{X}} p(\hat{x}_{h_k}\,|x, b_{h_k}) \prod_{m \in h_k} P_D^m \prod_{m \notin h_k} (1 - P_D^m)\, \frac{\hat{\eta}}{V} dx. \tag{20}$$

We now need only compute the integrals in (20) to complete our the derivation of the MAP assignment probabilities (6). When defining a specific form for the density $p(\hat{x}_{h_k}\,|x, b_{h_k})$, we must consider that tracks reported by different sensors on a common target in (20) may be correlated due to common priors and/or common process noise in the sensor estimation filters, and that the sensor tracks emanating from common objects represent dependent rather than independent samples.[18] We allow for this fact by considering joint Gaussian densities of the form

$$p(\hat{x}_{h_k}\,|x, b_{h_k}) = \frac{1}{\left|2\pi P_{h_k}\right|^{1/2}} \exp\left(-\frac{1}{2}\left(\hat{x}_{h_k} + b_{h_k} - B_{h_k}x\right)^T P_{h_k}^{-1}\left(\hat{x}_{h_k} + b_{h_k} - B_{h_k}x\right)\right), \tag{21}$$

where $B_{h_k} \in \mathbb{R}^{p \cdot n(h_k), p}$, $B_{h_k} = [I, \ldots, I]^T$, $I \in \mathbb{R}^{p,p}$ is the identity matrix, p is the dimension of the object state space, and $P_{h_k} \in \mathbb{R}^{p \cdot n(h_k), p \cdot n(h_k)}$ is the joint error covariance matrix for the sensor tracks comprising h_k. If V is large compared to P_{h_k}, the integral (20) can be evaluated with negligible error, leading to the following general result.

THEOREM 2.1. *Assuming a zero-mean Gaussian error for the bias prior*

$$Pr(H, b\,|\mathcal{Z}) = \mathcal{C}\frac{\exp\left(-\frac{1}{2}b^T R_b^{-1} b\right)}{\left|2\pi R_b\right|^{1/2}} \prod_{h_k \in H}\left[\prod_{m \in h_k} P_D^m \prod_{m \notin h_k} (1 - P_D^m)\, \beta_T p(\hat{x}_{h_k}\,|b_{h_k})\right], \tag{22}$$

where $\beta_T \triangleq \hat{\eta}/V$ denotes the target density per unit volume, and

$$p(\hat{x}_{h_k}\,|b_{h_k}) = \frac{1}{\left|2\pi\left(N_{h_k}^T P_{h_k} N_{h_k}\right)\right|^{1/2}} \exp\left(-\frac{1}{2}(\hat{x}_{h_k} + b_{h_k})^T N_{h_k}\left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T(\hat{x}_{h_k} + b_{h_k})\right) \tag{23}$$

for $n(h_k) > 1$, $p(\hat{x}_{h_k}\,|b_{h_k}) = 1$ for $n(h_k) = 1$, and N_{h_k} is any full-rank matrix whose columns span the null space of $B_{h_k}^T$

Proof. To reduce unnecessary notational burden in the following proof, we drop the track subscripts on the matrices and vectors. For symmetric, positive definite $P \in \mathbb{R}^{m,m}$, $z \in \mathbb{R}^n$, $x \in \mathbb{R}^p$ and $B \in \mathbb{R}^{m,p}$ with $p \leq m$ the well-known multivariate Gaussian integration formula yields

$$\int_{\mathbb{R}^p} \exp\left(-\frac{1}{2}(z - Bx)^T P^{-1}(z - Bx)\right) dx$$
$$= \left|2\pi(B^T P^{-1} B)^{-1}\right|^{1/2} \exp\left(-\frac{1}{2} z^T\left(P^{-1} - P^{-1}B\left(B^T P^{-1} B\right)^{-1} B^T P^{-1}\right) z\right). \tag{24}$$

For $p = m$ and full-rank B we have

$$\left(P^{-1} - P^{-1}B\left(B^T P^{-1} B\right)^{-1} B^T P^{-1}\right) = 0. \tag{25}$$

We now show that for $p < m$

$$\left(P^{-1} - P^{-1}B\left(B^T P^{-1} B\right)^{-1} B^T P^{-1}\right) = N\left(N^T P N\right)^{-1} N^T \tag{26}$$

for any full-rank matrices B and N, where the columns of N span the null space of B^T. Note that since P is positive definite, so are P^{-1}, $B^T P^{-1} B$ and $N^T P N$. From $B^T N = 0$, it follows that $N^T B = 0$ and

$$N^T P - N^T B\left(B^T P^{-1} B\right)^{-1} B^T = N^T P$$

$$\Rightarrow N^T P\left[\left(P^{-1} - P^{-1}B\left(B^TP^{-1}B\right)^{-1}B^TP^{-1}\right) - N\left(N^TPN\right)^{-1}N^T\right] = 0;$$

hence

$$P\left[\left(P^{-1} - P^{-1}B\left(B^TP^{-1}B\right)^{-1}B^TP^{-1}\right) - N\left(N^TPN\right)^{-1}N^T\right] \in \text{Null}\left(N^T\right).$$

Since B has full rank and $N^TB = 0$, the columns of B span $\text{Null}\left(N^T\right)$; that is, B is a basis for $\text{Null}\left(N^T\right)$, and,

$$P\left[\left(P^{-1} - P^{-1}B\left(B^TP^{-1}B\right)^{-1}B^TP^{-1}\right) - N\left(N^TPN\right)^{-1}N^T\right] = BA \tag{27}$$

for some matrix $A \in \mathbb{R}^{n,m}$. Left multiplying with B^TP^{-1} leads to

$$B^TP^{-1}BA = \left(B^TP^{-1} - B^TP^{-1}B\left(B^TP^{-1}B\right)^{-1}B^TP^{-1}\right) - \left(B^TN\left(N^TPN\right)^{-1}N^T\right) = 0,$$

and it follows that $A = 0$ since $B^TP^{-1}B$ is positive definite. The assertion (26) now follows immediately from (27).

Using the fact that

$$\int_{\mathcal{X}^{n_k}} p\left(\hat{x}_{h_k}\left|x, b_{h_k}\right.\right) d\hat{x}_{h_k} = 1$$

completes the proof. □

REMARK 2.1. *If the assumption that V is large compared to P_{h_k} fails to hold, likelihoods for tracks containing objects at the boundary of the surveillance region will be inflated. As a result, the optimal hypothesis based on* (22) *will contain more erroneous tracks than would be expected from the problem data.*[19]

REMARK 2.2. *The specific choice of $N_{h_k} \in \mathbb{R}^{p\cdot n(h_k), p\cdot(n(h_k)-1)}$ given by*

$$N_{h_k} = \begin{bmatrix} I & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I \\ -I & \cdots & \cdots & -I \end{bmatrix},$$

or any matrix derived by permutation of the rows of N_{h_k}, where $I \in \mathbb{R}^{p,p}$, yields a particularly useful form for $p\left(\hat{x}_{h_k}\left|b_{h_k}\right.\right)$ in (23).

Equation (22) is the main result of this section and, in principle, $\Pr\left(H, b\left|\mathcal{Z}\right.\right)$ can always be evaluated using (22); however, it is instructive and useful to further examine the two-sensor case. For the special case of two sensors, we can simplify notation by dropping the sensor subscripts on the sensor track indices and simply using i to index the Sensor 1 tracks, while j indexes the Sensor 2 tracks; thus, we consider two sets of sensor tracks

$$\begin{aligned} \hat{x}_i &= x_i - b_1 + \mu_i, \quad i = 1, \ldots, n_1, \\ \hat{x}_j &= x_j - b_2 + \nu_j, \quad j = 1, \ldots, n_2, \end{aligned}$$

where μ_i and ν_j are zero-mean random estimation errors. In this manner, tracks consist either of pairs of sensor tracks (i, j) hypothesized to emanate from a common object, or singleton tracks, denoted $(i, 0)$ and $(0, j)$, consisting of sensor tracks on objects detected only by Sensor 1 or Sensor 2, respectively. Using these conventions, the following corollary follows from Theorem 2.1 and Remark 2.2.

COROLLARY 1. *For $M = 2$ sensors, Equation* (22) *can be specialized to*

$$\begin{aligned} Pr\left(H, b\left|\mathcal{Z}\right.\right) = \mathcal{C}\frac{1}{\left|2\pi R_b\right|^{1/2}}\exp\left(-\frac{1}{2}b^TR_b^{-1}b\right)\prod_{(i,0)\in H}\beta_T P_D^1\left(1 - P_D^2\right)\prod_{(0,j)\in H}\beta_T\left(1 - P_D^1\right)P_D^2 \\ \cdot\prod_{(i,j)\in H}\frac{\beta_T P_D^1 P_D^2}{\left|2\pi\bar{P}_{ij}\right|^{1/2}}\exp\left(-\frac{1}{2}\left(\left(\hat{x}_i + b_1\right) - \left(\hat{x}_j + b_2\right)\right)^T\bar{P}_{ij}^{-1}\left(\left(\hat{x}_i + b_1\right) - \left(\hat{x}_j + b_2\right)\right)\right), \end{aligned}$$

where $b = (b_1, b_2)^T$, $R_b = E[bb^T]$ and $\bar{P}_{ij} = (P_{ii} + P_{jj} - P_{ij} - P_{ji})$, with $P_{ii} = E[\mu_i \mu_i^T]$, $P_{jj} = E[\nu_j \nu_j^T]$, and $P_{ij} = P_{ji}^T = E[\mu_i \nu_j^T]$ denoting the cross-covariance between Track i from Sensor 1 and Track j from Sensor 2, which may arise due to common process noise or common priors in the sensor estimation filters.

Finally, we show that (22) fits easily into the customary formulation of the data association problem as a dimensionless likelihood ratio. Consider the likelihood

$$\Pr\left(H_0, b_0 \,|\mathcal{Z}\,\right) = \mathcal{C}\frac{1}{\left|2\pi R_b\right|^{1/2}} \prod_{m=1}^{M} \left(\beta_T P_D^m \prod_{\substack{\mu=1 \\ \mu\neq m}}^{M} (1 - P_D^{\mu}) \right)^{n_m}$$

where $\mathcal{C}$ is the constant from (22), and $b_0 \triangleq 0$, and H_0 is the null hypothesis postulating that none of the tracks from any of the sensors emanate from a common object. Then, the likelihood ratio

$$\begin{aligned} \Lambda\left(H, b \,|\mathcal{Z}\,\right) &= \frac{\Pr\left(H, b \,|\mathcal{Z}\,\right)}{\Pr\left(H_0, b_0 \,|\mathcal{Z}\,\right)} \\ &= \exp\left(-\frac{1}{2} b^T R_b^{-1} b\right) \prod_{\substack{h_k \in H \\ |h_k|>1}} \frac{\exp\left(-\frac{1}{2}(\hat{x}_{h_k} + b_{h_k}) N_{h_k} (N_{h_k}^T P_{h_k} N_{h_k})^{-1} N_{h_k}^T (\hat{x}_{h_k} + b_{h_k})\right)}{\left|2\pi (N_{h_k}^T P_{h_k} N_{h_k})\right|^{1/2} \beta_T^{|h_k|-1} \prod_{m=1}^{M} (1 - P_D^m)^{|h_k|-1}}, \end{aligned} \tag{28}$$

where β_T was defined in Theorem 2.1, is dimensionless. This follows from the fact that $\left|N_{h_k}^T P_{h_k}^{-1} N_{h_k}\right|^{1/2}$ has a dimension $|h_k| - 1$ times the dimension of the state space, while β_T has a dimension that is the inverse of the state space dimension.

For the two-sensor case (28) reduces to

$$\Lambda\left(H, b \,|\mathcal{Z}\,\right) = \exp\left(-\frac{1}{2} b^T R_b^{-1} b\right) \prod_{(i,j)\in H} \frac{\exp\left(-\frac{1}{2}\left((\hat{x}_i + b_1) - (\hat{x}_j + b_2)\right)^T \bar{P}_{ij}^{-1} \left((\hat{x}_i + b_1) - (\hat{x}_j + b_2)\right)\right)}{\left|2\pi \bar{P}_{ij}\right|^{1/2} \beta_T \left(1 - P_D^1\right)\left(1 - P_D^2\right)}, \tag{29}$$

with b, R_b and $\bar{P}_{ij}$ as defined in Corollary 1. Note that Equation (28) explicitly enumerates only nontrivial tracks, since the likelihood ratios in the product of (28) evaluate to unity for system tracks consisting of only one sensor track.

REMARK 2.3. *Aside from being dimensionless, the likelihood ratio* (28) *has additional desirable qualities. As the probability of detection decreases, the term $(1 - P_D^m)^{|h_k|-1}$ in the denominator of* (28) *becomes larger, favoring system tracks containing fewer source tracks. This is desirable because it prevents spurious associations among densely-spaced objects, where each sensor has detected only a subset of the objects. Likewise, the inclusion of the term $\beta_T^{|h_k|-1}$ in* (28) *prevents spurious associations as the target density in the surveillance region increases.*

3. A NOTE ON BIAS OBSERVABILITY

In this section, we briefly show how the inclusion of a full bias prior allows estimation of absolute sensor biases. The problem (6) is typically solved in log-likelihood space, leading to the following mixed-integer nonlinear programming problem:[9]

$$(H^*, b^*) = \underset{H,b}{\operatorname{argmax}} \left(c_r(b) + \sum_{h_k \in H} c_{h_k}(b) \right), \tag{30}$$

where

$$\begin{aligned} c_r(b) &= \frac{1}{2} b^T R_b^{-1} b, \\ c_{h_k}(b) &= \frac{1}{2}\left(\hat{x}_{h_k} + b_{h_k}\right)^T N_{h_k} \left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T \left(\hat{x}_{h_k} + b_{h_k}\right) + \frac{1}{2}\log\left(\left|2\pi\left(N_{h_k}^T P_{h_k} N_{h_k}\right)\right|\right) + (|h_k| - 1)\kappa, \\ &= \frac{1}{2}\left(F_{h_k}\hat{x} + G_{h_k} b\right)^T N_{h_k} \left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T \left(F_{h_k}\hat{x} + G_{h_k} b\right) + \gamma_{h_k}, \end{aligned}$$

$$\gamma_{h_k} = \frac{1}{2}\log\left(\left|2\pi\left(N_{h_k}^T P_{h_k} N_{h_k}\right)\right|\right) + (|h_k| - 1)\kappa,$$

p is the dimension of the target state space, $\kappa \triangleq \log\left(\beta_T \prod_{m=1}^{M}\left(1 - P_D^m\right)\right)$, M is the number of sensors, and F_{h_k} and G_{h_k} are binary selection matrices defined by $\hat{x}_{h_k} = F_{h_k}\hat{x}$ and $b_{h_k} = G_{h_k} b$, with $\hat{x}$ denoting the vector, ordered by sensor and track, of all sensor tracks.

For any given hypothesis, the optimal bias estimate is given by solution of the linear system

$$\left(R_b^{-1} + \sum_{h_k \in H} G_{h_k}^T N_{h_k}\left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T G_{h_k}\right) b = -\sum_{h_k \in H} F_{h_k}^T N_{h_k}\left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T F_{h_k}\hat{x}. \quad (31)$$

If R_b is of full rank as assumed here, then the system has a unique solution since

$$\sum_{h_k \in H} G_{h_k}^T N_{h_k}\left(N_{h_k}^T P_{h_k} N_{h_k}\right)^{-1} N_{h_k}^T G_{h_k}$$

is positive semidefinite. This is absolute bias estimation. If R_b does not have full rank, then the biases can generally be estimated only as linear combinations of each other; that is, relative bias estimation.

4. CONCLUSION

We derived optimal maximum *a posteriori* (MAP) data association costs for concurrent bias estimation and data association with an arbitrary number of sensors. We also included a bias prior that allows estimation of absolute sensor biases, rather than just relative biases. Finally, we derived the two-sensor cost as a special case of the M-sensor result.

ACKNOWLEDGMENTS

We would like to express our gratitude to Louis Scharf for his insightful comments.

REFERENCES

1. S. M. Herman and A. B. Poore, "Nonlinear least-squares estimation for sensor and navigation biases," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, **6236**, 2006.
2. B. D. Kragel, S. Danford, S. M. Herman, and A. B. Poore, "Bias estimation using targets of opportunity," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, O. E. Drummond, ed., **6699**, (San Diego), August 2007.
3. R. Y. Novoselov, S. M. Herman, S. M. Gadeleta, and A. B. Poore, "Mitigating the effects of residual biases with Schmidt–Kalman filtering," in *Eighth International Conference on Information Fusion*, (Philadelphia, PA), July 2005.
4. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.
5. M. Levedahl, "An explicit pattern matching assignment algorithm," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, **4728**, pp. 461–469, 2002.
6. M. Levedahl, "Method and system for assigning observations." United States Patent US 7,092,924 B1, August 2006.
7. S. Mori and C. Chong, "Comparison of bias removal algorithms in track-to-track association," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, O. E. Drummond, ed., **6699**, SPIE, (San Diego), August 2007.
8. S. Danford, S. Herman, B. Kragel, and A. Poore, "A branch and bound framework for joint map bias estimation and data association," in *Proc. 10th ONR/GTRI Workshop on Target Tracking and Sensor Fusion*, (Monterey, CA), May 2007.
9. S. Danford, B. Kragel, and A. Poore, "Joint MAP bias estimation and data association: Algorithms," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, O. E. Drummond, ed., **6699**, SPIE, (San Diego), August 2007.

10. S. Danford, B. Kragel, and A. Poore, "Joint map bias estimation and data association: Simulations," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, O. E. Drummond, ed., **6699**, SPIE, (San Diego), August 2007.
11. D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control* **24**, pp. 843–854, December 1979.
12. S. Mori and C.-Y. Chong, "Track-to-track association metric i.i.d.-non-Poisson cases," in *Proceedings of the Sixth International Conference on Information Fusion*, pp. 553–559, (Cairns, Queensland, Australia), July 2003.
13. S. Mori, C.-Y. Chong, E. Tse, and R. P. Wishner, "Tracking and classifying multiple targets without a priori identification," *IEEE Transactions on Automatic Control* **AC-31**, pp. 401–409, May 1986.
14. L. M. Kaplan and W. D. Blair, "Assignment costs for multiple sensor track-to-track association," in *Proceedings of International Conference on Information Fusion*, (Stockholm, Sweden), July 2004.
15. Y. Bar-Shalom and H. Chen, "Multisensor track-to-track association for tracks with dependent errors," *Journal of Advances in Information Fusion* **1**, pp. 3–14, July 2006.
16. A. B. Poore, "Multidimensional assignment formulation of data association problems arising from multitarget tracking and multisensor data fusion," *Computational Optimization and Applications* **3**, pp. 27–57, 1994.
17. X. Lin, Y. Bar-Shalom, and T. Kirubarajan, "Multisensor bias estimation with local tracks without a priori association," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, **5204**, (San Diego), August 2003.
18. Y. Bar-Shalom and L. Campo, "The effect of common process noise on the two-sensor fused-track covariance," *IEEE Transactions on Aerospace and Electronic Systems* **AES-22**, pp. 803–805, November 1986.
19. B. D. Kragel, S. M. Herman, and N. J. Roseveare, "Efficiency and sensitivity of methods for assessing ambiguity in data association decisions," in *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, O. E. Drummond, ed., **6969-26**, (Orlando, Florida), March 2008.