Measurement-to-track association and finite-set statistics

Ronald Mahler, Eagan MN, USA, EML: mahlerronald@comcast.net

Abstract
This is a shortened, clarified, and mathematically more rigorous version of the original arXiv version. Its first four main findings remain unchanged from the original: 1) measurement-to-track associations (MTAs) in multitarget tracking (MTT) are heuristic and physically erroneous multitarget state models; 2) MTAs occur in the labeled random finite set (LRFS) approach only as purely mathematical abstractions that do not occur singly; 3) the labeled random finite set (LRFS) approach is not a mathematically obfuscated replication of multi-hypothesis tracking (MHT); and 4) the conventional interpretation of MHT is more consistent with classical than Bayesian statistics. This version goes beyond the original in including the following additional main finding: 5) a generalized, RFS-like interpretation results in a correct Bayesian formulation of MHT, based on MTA likelihood functions and MTA Markov transitions.

1. Introduction
When Reid introduced the multiple-hypothesis tracking (MHT) approach to multiple-target tracking (MTT) in 1979 [8], computer processing was primitive by today’s standards. MHT partially sidestepped this limitation by using measurement-to-track associations (MTAs) to decompose MTT problems into systems of extended Kalman filters. Since then, advances by Reid’s successors have made MHTs the workhorses of MTT. Understandably, however, limited computing power, combined with a lack of suitable mathematical machinery, also impeded proper theoretical rigor. But now, given the existence of greatly improved computational and mathematical tools, it is important to provide MTT with secure theoretical foundations that provide the basis for superior practical MTTs.

Labeled random finite set (LRFS) theory provides both. The RFS IDs/labels model appeared in 1997 in [2, p. 135, 196-197] and subsequently in [7, Sec. 14.5.6]; was systematically expanded into LRFS theory in [13, 14]; which has since been widely adopted. The first general exact closed-form approximation\footnote{In the sense of [6].} of any version of the multitarget Bayes recursive filter (MTBRF)—the GLMB filter—also appeared in [13], a surprising discovery that has been widely adopted or emulated. The first use of Gibbs sampling in LRFS (and probably in MTT) was introduced in [3] followed by [15] and is being increasingly adopted or emulated.

The latest Gibbs-based GLMB filter implementations can near-optimally and simultaneously track over a million 3D targets in real time in significant clutter using off-the-shelf computing equipment [1], another surprising development. Also, GLMB-type filters have: quantifiable approximation errors [14]; linear complexity in the number of measurements [14]; log-linear complexity in the number of hypothesized tracks [10]; and linear complexity in the number of scans in the multi-scan case [12].

This is a shortened, clarified, and mathematically more rigorous version of the original arXiv version. Its first four main findings remain unchanged from the original:

1. MTAs are heuristic and physically erroneous multitarget state models (Section 8).

2. MTAs occur in LRFS, but as purely mathematical abstractions that do not occur singly (Section 9).

3. LRFS is not, as sometimes suggested, a mathematically obfuscated replication of MHT (Section 10).

4. The conventional interpretation of MHT is more consistent with classical than Bayesian statistics (Section 5.1).

This version goes beyond the original in including the following additional main finding:

5. A generalized, RFS-like interpretation results in a correct Bayesian formulation of MHT, based on MTA likelihood functions and MTA Markov transitions (Sections 5.2, 5.3, 6, 7).

What follows are concise summaries of: Bayesian vs. classical statistics (Section 2); LRFS (Section 3); MHT and MTAs (Section 4); MHT conceptual difficulties (Section 5); likelihood functions for MTAs (Section 6); Markov transitions for MTAs (Section 7); the physical reality of MTAs (Section 8); MTAs in MHT vs. MTAs in LRFS (Section 9); and Conclusions (Section 10).
2. Bayesian vs. Classical Statistics

Suppose that physical entities with states $\xi \in \mathcal{X}$ are observed by a sensor that collects measurements $\zeta \in \mathcal{Z}$. The goal is to estimate $\xi$ given $\zeta$. Bayesian and classical statistics both employ a likelihood function (measurement distribution) $L_\zeta(\xi) = f(\zeta | \xi)$ that specifies the probability (or probability density) that measurement $\zeta$ will be collected from $\xi$. Thus $\int f(\zeta | \xi) d\xi = 1$ for all $\xi$.

2.1 Bayesian Statistics. The unknown state $\xi$ is presumed to be a random variable (RV) whose statistics are characterized by a prior distribution, $f_0(\xi)$. An estimate of $\xi$ is determined using some Bayes-optimal state estimator, for example the maximum a posteriori (MAP) estimator:\footnote{This is the only meaning of the term “Bayes-optimal.” An algorithm is not Bayes-optimal simply because it employs Bayes’ rule in some fashion.}

$$\hat{\xi} = \arg \sup_{\xi \in \mathcal{X}} f(\xi | \zeta) \tag{1}$$

where the posterior distribution on $\xi$ is $f(\xi | \zeta) = f(\zeta | \xi) f_0(\xi) / f(\zeta) = \int f(\zeta | \xi) f_0(\xi) d\xi$.

2.2 Classical Statistics. The unknown state $\xi$ is presumed to be a deterministic constant. From a Bayesian perspective, its distribution is $f_0(\xi) = \delta_0(\xi)$ where $\delta_0(\xi)$ is the Dirac delta concentrated at the unknown constant $\xi_0$. An estimate $\hat{\xi}$ of $\xi$ is determined using the maximum likelihood estimator (MLE):

$$\hat{\xi} = \arg \sup_{\xi \in \mathcal{X}} f(\xi). \tag{2}$$

3. LRFS

In LRFS $\mathcal{X} = \mathcal{F}_k(\mathcal{X}_0)$ where $\mathcal{F}_k(\mathcal{X}_0)$ is the set of all labeled finite subsets (LFSs)

$$\mathcal{X} = \{ (x_1, l_1), \ldots, (x_n, l_n) \} \tag{3}$$

of $\mathcal{X} = \mathcal{X}_0 \times L$, where $x_1, \ldots, x_n \in \mathcal{X}_0 = \mathbb{R}^N$ are the targets’ kinematic states and the distinct labels $l_1, \ldots, l_n \in L$ uniquely identify them. Let $0 \leq |S| \leq \infty$ denote the number of elements in a set $S$ and $X_L = \{ l_1, \ldots, l_n \}$ denote the set of labels in $\mathcal{X}$. Then $\mathcal{X}$ is LFS if and only if $| \mathcal{X} | = | X_L |$. A label is a symbol for a discrete state variable: target identity. If there are no labels or, equivalently, if $|L| = 1$ then $\mathcal{F}_k(\mathcal{X}_0) = \mathcal{F}(\mathcal{X}_0)$. A labeled random finite subset (LRFS) $\Xi \subseteq \mathcal{X}$ is an RV on $\mathcal{F}_k(\mathcal{X}_0)$.

Now consider LRFS with the “standard multitarget measurement model” (SMMM), in which a detection process is applied to a sensor signature [7, Sec. 12.3], [4, Sec. 7.2]. Then $\mathcal{Z} = \mathcal{F} (\mathcal{Z}_0)$ is the set of all finite subsets $Z$ of the single-target, single-sensor measurement space $\mathcal{Z}_0$. The likelihood has the form $L_Z (\mathcal{X}) = f (\mathcal{Z} | \mathcal{X})$, where $\int f (\mathcal{Z} | \mathcal{X}) d\mathcal{Z} = 1$ for all LFSs $\mathcal{X}$ and $\int f d\mathcal{Z}$ is an RFS set integral:

$$\int f (\mathcal{Z}) d\mathcal{Z} = \sum_{m \geq 0} \frac{1}{m!} \int f (\{z_1, \ldots, z_m\}) dz_1 \cdots dz_m. \tag{4}$$

The MAP estimator does not exist and must be replaced by (for example) the Bayes-optimal “JoM” estimator [7, Sec. 14.5], [4, Sec. 5.3]

$$\hat{X} = \arg \sup_{\mathcal{X}} \int f (\mathcal{X} | \mathcal{Z}) \frac{c | \mathcal{X} |}{| \mathcal{Z} |!}, \tag{5}$$

where $c > 0$ is a small constant with the same unit of measurement as $\mathcal{X}_0$; and where the posterior distribution is $f (\mathcal{X} | \mathcal{Z}) = f (\mathcal{Z} | \mathcal{X}) f_0 (\mathcal{X}) / f (\mathcal{Z}) = \int f (\mathcal{Z} | \mathcal{X}) f_0 (\mathcal{X}) d\mathcal{X}$.

4. MHT and MTAs

MHT is vector-based rather than set-based, and presumes the SMMM. Its multitarget states are MTAs rather than LFSs.

In this case $\mathcal{X} = \mathcal{X}_0 = \mathbb{R}^{m_0}$ where $\mathcal{X}_0$ is the $n$th Cartesian power of $\mathcal{X}_0$ and “$\otimes$” indicates disjoint union; and where $\mathcal{X}_0 = \{ \phi \}$ where $\phi$ is the null (empty) vector. Likewise, $\mathcal{Z} = \mathcal{Z}_0 = \mathbb{R}^{m_0}$. The sets $\mathcal{X} \in \mathcal{F} (\mathcal{X}_0)$ and $\mathcal{Z} \in \mathcal{F} (\mathcal{Z}_0)$ are replaced by vectors $\mathcal{X} = (x_1, \ldots, x_n) \in \mathcal{X}_0$ and $\mathcal{Z} = (z_1, \ldots, z_m) \in \mathcal{Z}_0$ with presumed orderings $x_1, \ldots, x_n$ and $z_1, \ldots, z_m$.

In what follows, $\{ \mathcal{Z} \} = \{ z_1, \ldots, z_m \}$ will denote the set of entries\footnote{This is an abuse of notation. Strictly speaking, $\{ \mathcal{Z} \}$ denotes a set with a single element $\mathcal{Z}$. The meaning $\{ \mathcal{Z} \} = \{ z_1, \ldots, z_m \}$ will be clear from context.} in $\mathcal{Z}$ and $| \mathcal{Z} |$ the number of those entries. In practice, if $\mathcal{Z} = (z_1, \ldots, z_m)$ then the measurements $z_1, \ldots, z_m$ will always be distinct—i.e., $| \{ \mathcal{Z} \} | = | \mathcal{Z} | = m$.

A real-valued function $f (\mathcal{Z})$ on $\mathcal{Z} \in \mathcal{F} (\mathcal{Z}_0)$ induces a real-valued function $\tilde{f} (\mathcal{Z})$ on $\mathcal{Z} \in \mathcal{Z}_0$ by:

$$\tilde{f} (\mathcal{Z}) = \left\{ \begin{array} { l l } { \frac{1}{| \mathcal{Z} |!} f (| \mathcal{Z} |) } & { \text{if } | \mathcal{Z} | = | \{ \mathcal{Z} \} | = m } \\ { 0 } & { \text{otherwise} } \end{array} \right.. \tag{6}$$

Because $\mathcal{Z}_0 = \mathbb{R}^{m_0}$ is continuously infinite, the RFS set integral $\int f d\mathcal{Z}$ is equivalent to the following integral $\int f d\mathcal{Z}$ over $\mathcal{Z}_0$:

$$\int f (\mathcal{Z}) d\mathcal{Z} = \int \tilde{f} (\mathcal{Z}) d\mathcal{Z} = \sum_{m \geq 0} \int \tilde{f} (\{z_1, \ldots, z_m\}) dz_1 \cdots dz_m, \tag{7}$$

where we have abbreviated $\tilde{f} (\{z_1, \ldots, z_m\}) = \tilde{f} (\{z_1, \ldots, z_m\})$ for all $(z_1, \ldots, z_m) \in \mathcal{Z}_0$.

4.1 MTAs. For conceptual clarity, assume that $1, \ldots, n$ identify all and the only targets present at all times, with spatial distributions of the form $s_1 (x), \ldots, s_n (x)$. Let $\tilde{\mathcal{Z}} = (z_1, \ldots, z_m) \in \mathcal{Z}_0$ be the measurement-vector collected by the sensor. An MTA for $\tilde{\mathcal{Z}}$ is a function

$$\alpha^{\tilde{\mathcal{Z}}} : \{ 1, \ldots, n \} \rightarrow \{ 0, 1, \ldots, m \} \tag{8}$$
such that \( \alpha_i(i) = \alpha_i(i') > 0 \) implies \( i = i' \). The set of all MTAs for \( \mathcal{Z} \) is denoted \( \mathcal{A}_k \) and the number of such MTAs is \( N_m \). Note that because \( \alpha_i \) and the pair \( (\alpha_i, \mathcal{Z}) \) contain exactly the same information, the former is an abbreviation of the latter. Also, whereas \( \mathcal{Z} \) is a constant, \( \alpha_i \in \mathcal{A}_k \) is an RV with \( N_m \) possible realizations.

**Notation 1:** An abbreviated notation will be employed in what follows. If \( \mathcal{Z}_k \) is a measurement-vector collected at time \( t_k \) then \( \alpha_k \) will be abbreviated as \( \alpha_k \in \mathcal{A}_k \), with \( \mathcal{Z}_k \) being implicitly understood.

**Notation 2:** Let \( \mathcal{Z}_{1:k} : Z_1, ..., Z_k \) be a sequence of measurement-vectors at times \( t_1, ..., t_k \) with, by convention, \( Z_{1:0} = \phi \) (null vector). Then \( \alpha_{1:k} : \alpha_1, ..., \alpha_k \) with \( \alpha_i = \alpha_i \) denotes a corresponding sequence of MTAs with, by convention, \( \alpha_{1:0} : \{1, ..., n\} \rightarrow \{0\} \) being the trivial MTA. For a given \( \mathcal{Z}_{1:k} \), the set of such sequences is \( \mathcal{A}_{\mathcal{Z}_{1:k}} \).

**4.2 Physical Interpretation of MTAs.** If \( \alpha_i(0) = 0 \) then the predicted target \( i \) was not detected at time \( t_i \); whereas if \( \alpha_i(0) > 0 \) then the measurement \( z_{\alpha_i} \) in \( \mathcal{Z}_k \) was collected from it. For any \( \tilde{W} \in \mathcal{Z}_0 \) and any \( \alpha \tilde{W} \in \mathcal{A}_{\tilde{W}} \), denote the vector of target-generated measurements in \( \tilde{W} \) by:

\[
\tilde{W}^\alpha = (z_{\alpha(i)} \in \{\tilde{W}\}) : \alpha(i) > 0
\]

where the ordering of the entries in \( \tilde{W}^\alpha \) preserves the ordering in \( \tilde{W} \). Then \( \{\tilde{W}\} - \{\tilde{W}^\alpha\} \) is the set of measurements in \( \tilde{W} \) not generated by any target—i.e., clutter and false detection measurements. In particular, if \( \tilde{W} = \mathcal{Z}_k \) then

\[
\tilde{Z}_k^\alpha = (z_{\alpha_k(i)} \in \{\mathcal{Z}_k\}) : \alpha_k(i) > 0
\]

denotes the vector of target-generated measurements in \( \mathcal{Z}_k \).

**4.3 HOMHT.** There are numerous variants of MHT. For the sake of conceptual clarity, Reid’s “hypothesis-oriented” (HO) version [8] will be presumed—but in the following special case: target number \( n \) is known and constant and the probability of detection \( p_{D,k}(x) \) is state-dependent and usually nonconstant.

Suppose that measurement-vectors \( \mathcal{Z}_{1:k-1} : Z_1, ..., Z_{k-1} \) have previously been collected and that MTAs \( \alpha_{1:k-1} : \alpha_1, ..., \alpha_{k-1} \) have been chosen. Then the spatial distributions at time \( t_k \) are \( \delta_{i,k-1}(x) = \delta_{i,k-1}(x|\alpha_{1:k-1}, \mathcal{Z}_{1:k-1}) \) for \( i = 1, ..., n \). Each of these is predicted to time \( t_k \),

\[
s_{i,k} = \int f_{i,k-1}(x|x_{k-1}) \delta_{i,k-1}(x_{k-1}) dx_{k-1},
\]

where \( f_{i,k-1}(x|x_{k-1}) \) is the single-target Markov density.

Now let \( \tilde{Z}_k = (z_1, ..., z_{m_k}) \) be the new measurement-vector collected at time \( t_k \) and choose an MTA \( \alpha_k : \{1, ..., n\} \rightarrow \{0, 1, ..., m_k\} \). Then the predicted spatial distributions are measurement-updated as follows:

\[
s_{i,k}(x) \propto \left\{ \begin{array}{ll}
f_{i,k}(x) & \text{if } \alpha_k(i) > 0 \\
1 - p_{D,k}(x) & \text{if } \alpha_k(i) = 0 \\
\end{array} \right.
\]

where \( f_k(z|x) \) is the single-target measurement density at time \( t_k \); and where \( s_{i,k}' = s_{i,k}'(x|\alpha_{1:k}, \mathcal{Z}_{1:k}) \).

Given this, the goal of HOMHT (known as “deferred decision logic”) is to, at an appropriate time \( t_j \) (when track uncertainties can most opportunistically be resolved), determine what MTA-sequence \( \alpha_{1:k} : \alpha_1, ..., \alpha_k \) most accurately matches measurements with predicted targets over the time period \( t_1, ..., t_k \) (see Section 4.3).

**4.4 Reid’s Recursion.** Let \( p_{1:k}(\alpha_{1:k}|\mathcal{Z}_{1:k}) \) be the probability of the MTA-sequence \( \alpha_{1:k} \) given the measurement-sequence \( \mathcal{Z}_{1:k} \). Then Reid’s recursion formula for these probabilities is the Bayesian update equation [8, Eq. 8]:

\[
p_{1:k}(\alpha_{1:k}|\mathcal{Z}_{1:k}) = \frac{\tilde{f}_k(\mathcal{Z}_k|\mathcal{Z}_{1:k-1})}{\tilde{f}_k(\mathcal{Z}_k|\mathcal{Z}_{1:k-1})} p_{k-1}(\alpha_{1:k-1}|\mathcal{Z}_{1:k-1})
\]

with Bayesian transition density

\[
\tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1}, \mathcal{Z}_{1:k-1}) = \tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1}, \mathcal{Z}_{1:k-1})
\]

and the number of such MTAs

\[
p_{k}(\alpha_{1:k} | \mathcal{Z}_{1:k-1}) = \frac{p_k(\alpha_{1:k} | \mathcal{Z}_{1:k-1})}{\tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1}, \mathcal{Z}_{1:k-1})}
\]

Now apply the independence assumptions

\[
\tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1}, \mathcal{Z}_{1:k-1}) = \tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1})
\]

where \( p_{k-1}(\alpha_{1:k-1} | \mathcal{Z}_{1:k-1}) = p_{k-1}(\alpha_{1:k-1} | \mathcal{Z}_{1:k-1}) \) since \( \alpha_{k-1} = \alpha_{k-1}^{\mathcal{Z}_{1:k-1}} \) contains the same information as \( \alpha_{k-1}, \mathcal{Z}_{1:k-1} \). Then \( p_{k-1}(\alpha_{1:k-1} | \mathcal{Z}_{1:k-1}) \) is, seemingly, a conventional Bayesian Markov transition [7, Eq. 3.56]; and \( \tilde{f}_k(\mathcal{Z}_k|\alpha_{1:k-1}) \) is, seemingly, a conventional Bayesian measurement density [7, Eq. 3.54].

When \( k = 1 \), Eqs. (13, 14) reduce, also seemingly, to Bayes’ rule:

\[
p_{1}(\alpha_{1}|\mathcal{Z}_1) = \frac{\tilde{f}_1(\mathcal{Z}_1|\alpha_{1})p_{10}(\alpha_{1})}{\tilde{f}_1(\mathcal{Z}_1)}.
\]

Here, \( p_{10}(\alpha_{1}) \) is the Markov transition from the trivial MTA \( \{1, ..., n\} \rightarrow \{0\} \) to \( \alpha_1 \)—but which also serves as the initial distribution.

**4.5 Multitarget State Estimation.** At time \( t_k \), the targets’ spatial distributions will be \( \delta_{i,k}(x|\alpha_{1:k}, \mathcal{Z}_{1:k}) \) for \( i = 1, ..., n \). From a practical point of view, these are meaningless unless the MTAs \( \alpha_j \in \mathcal{A}_j \) for \( j = 1, ..., k \) have been selected so as to best match measurements with targets—i.e., to resemble the “actual MTAs” as much as possible. This selection is the MAP estimate:

\[
\hat{\alpha}_{1:k} = \arg \max_{\alpha_{1:k}} p_k(\alpha_{1:k}|\mathcal{Z}_{1:k}).
\]
where the argmax is taken over the finite number of all \( \alpha_{1:k} \in \mathbb{A}^{\tilde{Z}_{1:k}} \). The multitarget state estimate at time \( t_k \) is then
\[
\hat{x}_k^n = \arg \sup_{x \in \mathbb{X}_0} s_k^n(x | \hat{\alpha}_{1:k}, \tilde{Z}_{1:k}) \quad \text{for} \quad i = 1, \ldots, n.
\]

5. MHT Conceptual Difficulties

The transitions \( s_i^{k-1}(x) \rightarrow s_i^{k-1}(x) \) and \( s_i^{k+1}(x) \rightarrow s_i^k(x) \) for \( i = 1, \ldots, n \) implicitly presume that the \( n \) targets are not state variables. This is equivalent to the following assumption:
\[
\hat{f}_k(x_1, \ldots, x_n | \hat{\alpha}_{1:k}, \tilde{Z}_{1:k}) \cong s_k(x_1 | \hat{\alpha}_{1:k}, \tilde{Z}_{1:k}) \cdots s_n(x_n | \hat{\alpha}_{1:k}, \tilde{Z}_{1:k}).
\]

But this also presumes a questionable interpretation of MTAs, as follows.

5.1 The Conventional MHT Interpretation

Eq. (17) is the following assumption: \( \tilde{Z}_1 \) and \( m_1 \) and the randomness of \( \tilde{Z}_1 \) is known, and the randomness of \( \hat{\alpha}_1 \) as characterized by \( p_{1|0}(\alpha_1) \), involves only the variability of the MTA \( \alpha_1 \in \mathbb{A}^{\tilde{Z}_1} \). Thus the normalization condition for \( p_{1|0}(\alpha_1) \) must be
\[
\sum_{\alpha_1 \in \mathbb{A}^{\tilde{Z}_1}} p_{1|0}(\alpha_1) = 1.
\]

For the same reason, the MAP estimator must be
\[
\hat{\alpha}_1 = \arg \max_{\alpha_1 \in \mathbb{A}^{\tilde{Z}_1}} p_{1|0}(\alpha_1).
\]

But neither can be true. If \( p_{1|0}(\alpha_1) \) is an initial distribution then it encapsulates all information available before the collection \( \tilde{Z}_1 \sim \hat{f}_1(\tilde{Z} | \alpha_1) \) of the first measurement-vector \( Z_1 \). It thus cannot depend on \( Z_1 \) or even \( m_1 \)—which, in turn, chronologically contradicts the fact that \( \alpha_1 = \hat{\alpha}_1 \).

Moreover, this interpretation is peculiar from a Bayesian perspective. If \( \hat{\alpha}_1 \) is a multitarget state with state parameter \( \tilde{Z}_1 \), then it, and \( \tilde{Z}_1 \) in particular, are unknown RVs. But in actuality \( \tilde{Z}_1 \) is known and constant—which means that the distribution of \( \alpha_1 \) must have the form
\[
p_{1|0}(\alpha_1) = \delta_{\tilde{Z}_1}(\tilde{Z})\ p_1(\alpha_1 | \tilde{Z}_1)
\]
where \( \delta_{\tilde{Z}_1}(\tilde{Z}) \) is the Dirac delta concentrated at \( \tilde{Z}_1 \). This denouement is more consistent with classical statistics than Bayesian statistics (see Section 2).

5.2 A General MHT Interpretation

The Bayesian expressions \( \hat{f}_1(\tilde{Z} | \alpha_1) \) and \( p_{1|0}(\alpha_1) \) imply that they are defined for arbitrary finite \( \tilde{Z} \in \mathbb{A}^{\tilde{Z}_1} \) and arbitrary \( \alpha_1 \in \mathbb{A}^{\tilde{Z}_1} \). That is, a probability \( p_{1|0}(\alpha_1) \) must be assigned to each \( \tilde{Z} \) and each \( \alpha_1 \). Thus the true normalization condition must be
\[
\int \sum_{\alpha_1 \in \mathbb{A}^{\tilde{Z}_1}} p_{1|0}(\alpha_1) \ d\tilde{Z} = 1,
\]
where the RFS-like integral \( \int d\tilde{Z} \) is as defined in Eq. (7).

With this interpretation, the chronological contradiction disappears because the RV \( \tilde{Z} \) and thus \( \hat{f}_1(\tilde{Z} | \alpha_1) \) are no longer restricted to the realization \( \tilde{Z} = \tilde{Z}_1 \). This interpretation is what will be adopted hereafter. Its immediate consequence is that we must devise concrete definitions of \( \hat{f}_k(\tilde{Z} | \alpha_1^k) \) and \( p_{1|k-i}(\alpha_1^k | \alpha_{k-i}) \) in Eqs. (15), (16) for all \( \tilde{Z} \in \mathbb{A}^{\tilde{Z}_1} \) and, given \( \tilde{Z} \), for all \( \alpha_1^k \in \mathbb{A}^{\tilde{Z}_1} \). These definitions are the subjects of, respectively, Sections 6 and 7.

5.3 State Estimation in the General Interpretation

The rigorous version of Eq. (18) is the MAP estimate of \( f_k(\alpha_1^{k+1}, \tilde{Z}_1 | \tilde{Z}_k) \) when \( \tilde{Y}_{1:k} = \tilde{Z}_{1:k} \) is held constant:
\[
\hat{\alpha}_1^k = \arg \sup_{\alpha_1 \in \mathbb{A}^{\tilde{Z}_1}} \left[ p_k(\alpha_1^{k+1} | \tilde{Z}_1 | \tilde{Z}_{1:k}) \right]_{\tilde{Y}_{1:k} = \tilde{Z}_{1:k}}.
\]

6. Likelihood Functions for MTAs

6.1 Extended MTAs for Likelihoods

Define the extension of \( \alpha_k = \alpha_k^* : \{1, \ldots, n\} \rightarrow \{0, 1, \ldots, m_k\} \) to \( \alpha^* : \{1, \ldots, n\} \rightarrow \{0, 1, \ldots, m\} \) as follows:
\[
\alpha(i) = \alpha^*(i) = \begin{cases} 
\alpha_k(i) & \text{if } \alpha_k(i) > 0, \\
0 & \text{otherwise}
\end{cases}
\]

Using the notation in Eq. (9), if \( \tilde{Z}_k^* \) is the vector of target-generated measurements in \( \tilde{Z}_k \) then
\[
\{\tilde{Z}^*\} = \{\tilde{Z}_k^*\} \cap \{\tilde{Z}\}
\]
is the set of target-generated measurements in \( Z \); and \( \{\tilde{Z}\} - \{\tilde{Z}^*\} \) is the set of clutter measurements in \( Z \).

6.2 MTA Likelihoods

At time \( t_k \) let \( p_{D,k}(x) \) be the sensor probability of detection; \( f_k(z | x) \) the sensor measurement density; \( \kappa_k(z) \) the intensity function of a Poisson clutter process; \( \lambda_k = \int \kappa_k(z)dz \); \( c_k(z) = \kappa_k(z)/\lambda_k \); and \( k^* = 1 \) if \( Z \) is null and \( k^* = \prod_{i \in \{d\}} k(i) \) otherwise.

Let \( \tilde{Z} = (z_1, \ldots, z_m) \). Then \( \tilde{f}_k(\tilde{Z} | \alpha_k^*) \) is defined to be:
\[
\tilde{f}_k(z_1, \ldots, z_m | \alpha_k) = e^{-\lambda_k} \prod_{i=1}^m \kappa_k(z_i) \prod_{i \in \{d\}} g_k(z_{\alpha(i)}),
\]
where \( \alpha(i) \) was defined in Eq. (23) and where, for \( s_{i}^{k-1}(x) = s_{i}^{k+1}(x) | \alpha_{i-1:k-1}, \tilde{Z}_{i-1:k-1} \) and \( i = 1, \ldots, n \),
\[
g_k(z) = \frac{\int p_{D,k}(x) f_k(z | x) s_{i}^{k-1}(x)dx}{\int p_{D,k}(y) s_{i}^{k-1}(y)dy}
\]
is the probability (density) that \( z \) is generated by the target \( i \) (with target detectability taken into account);\(^{10}\) whereas \( c_k(z) \) is the probability (density) that \( z \) is a clutter measurement. The unitless ratio \( g_k(z) / c_k(z) \) quantifies the degree to which measurement \( z \) is due to target \( i \) rather than clutter. When \( \{\tilde{Z}^*\} \cap \{\tilde{Z}\} = \emptyset \) (i.e., if \( \alpha^* \) has no target-generated measurements) then
\[
\tilde{f}_k(\tilde{Z} | \alpha_k) = \frac{e^{-\lambda_k} \kappa_k^{\tilde{Z}^*}}{|Z|!} = \tilde{r}_k(\tilde{Z})
\]
where $\tilde{n}_k(\tilde{Z})$ denotes the distribution of the Poisson clutter process.

It is verified in Section 11 that $\int f_k(\tilde{Z}|\alpha_k)d\tilde{Z} = 1$ for any $\alpha_k \in \mathcal{A}_k$ and any $\tilde{Z}_k$.

7. Markov Transitions for MTAs

7.1 Predicted MTAs for Markov Transitions. Let $\alpha_{k-1} : \{1, \ldots, n\} \rightarrow \{0,1,\ldots,m_{k-1}\}$ be given, so that the target-generated measurement-vector at time $t_{k-1}$ is $\tilde{Z}^{\alpha_{k-1}} = (z_{\alpha_{k-1}(i)} : \alpha_{k-1}(i) > 0)$. Assume that the single-target measurement density at time $t_k$ has the form

$$f_k(z|x_k) = f_{\nu_k}(z - \eta_k(x_k))$$  \hspace{1cm} (28)

where $\eta_k(x_k) \in \mathbb{R}^n$ is a measurement function and $\mathbf{V}_k \in \mathbb{R}^n$ is zero-mean additive measurement noise.

Let $\hat{x}_{i|k-1} = \arg\sup_{x \in \mathbb{R}^n} N_{k-1}(x)$ for $i = 1, \ldots, n$ be the estimated states of the predicted targets. Then the vector of all predicted target-generated measurements at time $t_k$ is

$$\hat{Z}^k_k = (\eta_k(\hat{x}_{i|k-1}) : i = 1, \ldots, n).$$  \hspace{1cm} (29)

According to $\alpha_{k-1}$, the only targets $i$ that generated measurements at time $t_{k-1}$ were those such that $\alpha_{k-1}(i) > 0$. Assume that these targets are measurement-generating at both times $t_k$ and $t_{k-1}$. Then the actual vector of predicted target-generated measurements at time $t_k$ is

$$\tilde{Z}^k_k = (\eta_k(x_{i|k-1}) : \alpha_{k-1}(i) > 0).$$  \hspace{1cm} (30)

Given this, for any $\tilde{Z} \in \mathcal{Z}_0$, define the predicted MTA $\alpha : \{1, \ldots, n\} \rightarrow \{0,1,\ldots,m\}$ to be

$$\alpha(i) = \alpha^\tau(i) = \begin{cases} \alpha_{k-1}(i) & \text{if} \quad \left\{ \begin{array}{l} \alpha_{k-1}(i) > 0, \\ \eta_k(\hat{x}_{i|k-1}) \in \{\tilde{Z}\} \end{array} \right. \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (31)

The set of target-generated measurements in $\tilde{Z}$ according to $\alpha$ is, therefore, $\{\tilde{Z}^\alpha\} = \{\hat{Z}^k_k\} \cap \{\tilde{Z}\}$; whereas the set of clutter measurements in $\tilde{Z}$ is $\{\tilde{Z}\} - \{\tilde{Z}^\alpha\}$.

7.2 Markov Density for MTAs. Given this, for $\tilde{Z} = (z_1, \ldots, z_m)$ the MTA Markov transition is defined to be:

$$p_{k|k-1}(\alpha^\tau|\alpha_{k-1}) = \frac{e^{-\lambda_k} \kappa_k^m N_{m-1}}{m!} \prod_{i=\alpha(i) > 0} g_k(z_{\alpha(i)}) / c_k(z_{\alpha(i)})$$  \hspace{1cm} (32)

where $\alpha = \alpha^\tau$ was defined in Eq. (31); where $g_k(z)$ is as defined in Eq. (26); and where $\lambda_k$ was defined in Section 4.1. If $\{\hat{Z}^k_k\} \cap \{\tilde{Z}\} = \emptyset$ ($\alpha^\tau$ has no target-generated measurements) then

$$p_{k|k-1}(\alpha^\tau|\alpha_{k-1}) = \tilde{n}_k(\tilde{Z}) \hat{N}_k^{-1}$$  \hspace{1cm} (33)

which does not depend on $\alpha_{k-1}$ at all. Since $\hat{Z}_k$ is yet to be collected, the factor $\hat{N}_k^{-1}$ acknowledges the fact that no $\alpha^\tau \subset \mathcal{A}_k$ can be preferred on an a priori basis, insofar as the prediction $\alpha_{k-1} \rightarrow \alpha^\tau$ is concerned.

It is verified in Section 12 that, for any $\alpha_{k-1} \in \mathcal{A}_k$,\hspace{1cm} (34)

$$\int \sum_{\alpha^\tau \in \mathcal{A}_k} p_{k|k-1}(\alpha^\tau|\alpha_{k-1})d\tilde{Z} = 1.$$  \hspace{1cm} (34)

8. Physical Reality of MTAs

Beyond these purely mathematical issues, a more serious physical one must be addressed: Is an MTA physically real? Stated differently: Is it physically reasonable to claim that specific measurements are generated by specific targets?

This is doubtful. An MTA is a heuristic extrapolation to general multitarget scenarios of the following special case. Suppose that we have a sensor with no missed or false detections. Further suppose that we have $n$ targets with Dirac-delta spatial distributions $s_k(x) = \delta_{x_1}(x), \ldots, s_k(x) = \delta_{x_n}(x)$ that are well-separated with respect to the noise resolution of the sensor, as specified by $L(x) = f_k(z|x)$. In this case it is intuitively clear that there is a permutation $\pi_0$ on $1, \ldots, n$ such that target $i$ generated $z_{\pi_0(i)}$. This is because there is only a small probability that $z_{\pi_0(i)}$ could have originated with any target other than $x_i$. That is, $\pi = \pi_0$ with $f_k(\tilde{Z}|\pi_0) = g_{\pi_0}(z_{\pi_0(1)}) \cdots g_{\pi_0}(z_{\pi_0(n)})$.

Suppose, more generally, that the targets are very close together. Then it is statistically impossible to maintain that any particular measurement is generated by any particular target. Such a claim becomes even more difficult to maintain if the sensor has missed and false detections. But the improbable becomes truly impossible when one considers the following. First, point-detections are discretized mathematical idealizations that have been heuristically extracted from a continuous but noise-contaminated sensor signature. Second, the targets are assumed to be in an idealized “goldilocks zone”: not too near (in which case a single target can generate multiple detections) and not too distant (in which case multiple targets can generate a single detection).

Consequently, to insist that there is a “Bayes-optimal” MTA is to impose a phenomenologically questionable structure upon the modeling of the physical system. Instead, the most that can legitimately be asserted is the following: If targets with state-set $X$ are present, then there is an RFS probability (density) $f_k(Z|X)$ that they will generate a measurement-set $Z$—see Eq. (39). From this it follows that MTAs are not physically valid state representations of multitarget systems.

Is this claim refuted by the fact that the RFS-like analysis in Sections 5-7 demonstrates that MHT has a Bayesian interpretation? No, because Bayesian reasoning can be ap-
plied to even completely imaginary scenarios (e.g., tracking unicorns).

9. MTAs in MHT vs. MTAs in LRFS

MTAs also occur in LRFS likelihoods \( f(Z|X) \). But they do not, as in MTA theory, arise from heuristic intuition. Rather, they arise from a mere change of notation as part of a mathematically rigorous LRFS derivation based on a statistically and phenomenologically rigorous LRFS multitarget measurement model. In particular, the MTAs in \( f(Z|X) \) do not occur as isolated entities but, rather, simultaneously within the terms of a summation (see Eq. (39) and [4, Eq. 7.21]).

Specifically, let \( X = \{x_1, \ldots, x_n\} \) and \( Z = \{z_1, \ldots, z_m\} \) with \(|X| = n\) and \(|Z| = m\). Then the LRFS multitarget measurement model and multitarget calculus is used to derive the following formula:\(^{12}\)

\[
f(Z|X) = e^{-\lambda} \sum_{W_0 \cup W_1 \cup \cdots \cup W_n = Z} \kappa^{Z-W_0} (1 - p_D)^X \cdot \left( \prod_{i=1}^{n} f_i(W_i) \right) \tag{37}
\]

where \( \kappa(Z) = e^{-\lambda |Z|^2} e^Z \); where \( (1 - p_D)^X = \prod_{x \in X} (1 - p_D(x)) \) if \( X \neq \emptyset \) and \( (1 - p_D)^X = 1 \) otherwise; where the summation is taken over all \( W_1, \ldots, W_n \subseteq Z \) such that \(|W_1|, \ldots, |W_n| \leq 1 \);\(^{13}\) and where

\[
f_i(W_i) = \begin{cases} \frac{1}{p_D(x_i)} & \text{if } W_i = \emptyset \\ \frac{1}{1 - p_D(x_i)} & \text{if } W_i = \{z\} \end{cases}
\]

Now, for a given choice of \( W_1, \ldots, W_n \), define \( \alpha : \{1, \ldots, n\} \to \{0, 1, \ldots, m\} \) implicitly by \( \{z_{\alpha(i)}\} = W_i \) if \( W_i \neq \emptyset \) and \( \alpha(i) = 0 \) otherwise. Then \( \alpha \) is an MTA. Conversely, given \( \alpha \), define \( W_i = \{z_{\alpha(i)}\} \) if \( \alpha(i) > 0 \) and \( W_i = \emptyset \) if otherwise. Thus there is a one-to-one correspondence between MTAs \( \alpha \) and lists \( W_1, \ldots, W_n \subseteq Z \) such that \(|W_1|, \ldots, |W_n| \leq 1 \).

After this change from set-theoretic notation to MTA notation, Eq. (37) becomes [4, Eq. 7.21]

\[
f(Z|X) = \kappa(Z) (1 - p_D)^X \cdot \sum_{\alpha} \prod_{\alpha(i) > 0} \frac{p_D(x_i)}{\kappa(z_{\alpha(i)}) \left( 1 - p_D(x_i) \right)} \tag{39}
\]

Note that the right side is invariant with respect to the orderings \( X = \{x_1, \ldots, x_n\} \) and \( Z = \{z_1, \ldots, z_m\} \) chosen for the elements of \( Z \) and \( X \).

10. Conclusion: MHT vs. LRFS

It should be evident at this point that LRFS is not a mathematically obfuscated replication of MHT. LRFS is mathematically, statistically, and physically rigorous, whereas MHT is physically heuristic. LRFS-based MTTs can near-optimally and simultaneously track over a million 3D targets in real time in significant clutter using off-the-shelf computing equipment—a direct consequence of the LRFS reformulation of MTT as a MTBRF on LFSs.

LRFS with SMMM also includes the Bernoulli filter [11], [7, Sec. 14.7], [9] and “dyadic filter” [5], which are non-approximate ECF special cases of the MTBRF when target number cannot exceed 1 or 2, respectively. Neither looks anything like MHT. In particular, the Bernoulli filter is equivalent to a probability hypothesis density (PHD) filter with a nonstandard multitarget motion model.

Moreover, while MHT is limited to SMMM, RFS/LRFS systematically addresses diverse sensor phenomenologies:

1. sensor signature (“raw”) data, such as superpositional or pixelized-imaging sensors [4, Chpts. 19, 20]; and

2. “nonstandard” measurements (features, attributes, natural-language statements, inference rules), as part of a general theory of measurements based on random closed sets (RCSs) [4, Chpt. 22].

11. Appendix: MTA Likelihood Normality

By Eq. (7),

\[
\int f_k(\tilde{Z}|\alpha_k) d\tilde{Z} = \sum_{j \geq 0} \int f_k(z_1, \ldots, z_j|\alpha_k) dz_1 \cdots dz_j.
\]

Abbreviate \( \tilde{W}_j = (z_1, \ldots, z_j) \). Let \( \gamma_j : \{1, \ldots, n\} \to \{0, 1, \ldots, j\} \) be the extension of \( \alpha_k \) to \( \tilde{W}_j \) as defined in Eqn. (23). Then from Eqs. (9, 24, 25) and \( \kappa^{\tilde{W}_j} = \lambda_j^i e_{\tilde{W}_j} \) and \( \{\tilde{Z}^{\gamma_j}\} = \{\tilde{Z}^{\alpha_k}\} \cap \{\tilde{W}_j\} \) the desired result follows:

\[
\int f_k(\tilde{Z}|\alpha_k) d\tilde{Z} = e^{-\lambda_k} \sum_{j \geq 0} \frac{\lambda_j^i}{j!} \int \left( \prod_{l=1}^{i} c_k(z_l) \right) \cdot \left( \prod_{\gamma_k(j) > 0} g_k^b(z_{\gamma_k(j)}) c_k(z_{\gamma_k(j)}) \right) dz_1 \cdots dz_j \tag{41}
\]

\[
= e^{-\lambda_k} \sum_{j \geq 0} \frac{\lambda_j^i}{j!} \int \left( \prod_{l=\gamma_k(j) > 0} c_k(z_l) \right) \cdot \left( \prod_{\gamma_k(j) > 0} g_k^b(\tilde{z}_{\gamma_k(j)}) d\tilde{z}_{\gamma_k(j)} \right) \tag{42}
\]

\[
= e^{-\lambda_k} \sum_{j \geq 0} \frac{\lambda_j^i}{j!} \int \left( \prod_{l=\gamma_k(j) > 0} c_k(z_l) dz_l \right) \cdot \left( \prod_{\gamma_k(j) > 0} g_k^b(z_{\gamma_k(j)}) dz_{\gamma_k(j)} \right) \tag{43}
\]

\[
= e^{-\lambda_k} \sum_{j \geq 0} \frac{\lambda_j^i}{j!} \cdot 1 = 1. \tag{44}
\]

\(^{12}\)For conceptual clarity the unlabeled RFS case is considered here. The LRFS case is similar but slightly more complicated.

\(^{13}\)In which case \( W_0 = Z = (W_1 \cup \ldots \cup W_n) \).
12. Appendix: MTA Markov Transition Normality

By Eq. (7),

\[
\int \sum_{\alpha_\ast \in \mathbb{A}} p_{k|k-1}(\alpha_\ast | \alpha_{k-1}) d \tilde{Z}
\]

\[
= \sum_{j=0} \int \sum_{\alpha_\ast \in \mathbb{A}} p_{k|k-1}(\alpha_{z_1, \ldots, z_j} | \alpha_{k-1}) dz_1 \cdot \ldots \cdot dz_j.
\]

Abbreviate \( \tilde{W}_j = \{z_1, \ldots, z_j\} \). Let \( \gamma_j : \{1, \ldots, n\} \rightarrow \{0, 1, \ldots, j\} \) be the predicted MTA as defined in Eq. (31). Then from Eqs. (30, 31) and \( \tilde{W}_j \), the desired result follows:

\[
= \sum_{j=0} \sum_{\gamma_j \in \tilde{W}_j} e^{-\lambda_k Y_j^{(j)}} \int \left( \prod_{i=1}^{j} c_k(z_i) \right) \cdot \left( \prod_{i=1}^{\gamma_j} g_k(z_{\gamma_j}(i)) \right) dz_1 \cdot \ldots \cdot dz_j
\]

\[
= \sum_{j=0} \sum_{\gamma_j \in \tilde{W}_j} e^{-\lambda_k Y_j^{(j)}} \int \left( \prod_{i=1}^{\gamma_j} c_k(z_i) \right) dz_1 \cdot \ldots \cdot dz_j
\]

\[
= \sum_{j=0} \sum_{\gamma_j \in \tilde{W}_j} e^{-\lambda_k Y_j^{(j)}} \int \left( \prod_{i=1}^{\gamma_j} g_k(z_{\gamma_j}(i)) \right) dz_{\gamma_j}(i)
\]

\[
= \sum_{j=0} \frac{e^{-\lambda_k Y_j^{(j)}} \sum_{\gamma_j \in \tilde{W}_j} \tilde{N}_j^{(j)}}{j!} = \sum_{j=0} \frac{e^{-\lambda_k Y_j^{(j)}}}{j!} = 1. \quad (49)
\]

References
1. M. Beard, B.-T. Vo, and B.-N. Vo, “A solution for large-scale multi-object tracking,” IEEE Trans. Sign. Proc., 68: 2754-2769, 2020.
2. I. Goodman, R. Mahler, and H. Nguyen, Mathematics of Data Fusion, Kluwer Academic Publishers, New York, 1997.
3. Hung Gia Hoang, B.-T. Vo, and B.-N. Vo, “A generalized labeled multi-Bernoulli filter implemented using Gibbs sampling,” 2015, open source: https://www.arxiv.org/abs/1506.00821/.
4. R. Mahler, Advances in Statistical Multisource-Multitarget Information Fusion, Artech House, Norwood, MA, 2014.
5. R. Mahler, “Bayes-optimal tracking of two statistically correlated targets in general clutter,” IET Sign. Proc., 2022, open source: https://doi.org/10.1049/sil2.12150.