Evaluation of Data Association Hypotheses: Non-Poisson I.I.D. Cases

Shozo Mori
ALPHATECH, Inc.
2570 West El Camino Real, Suite 300
Mountain View, CA 94040
U.S.A.
smori@alphatech.com

Chee-Yee Chong
ALPHATECH, Inc.
2570 West El Camino Real, Suite 300
Mountain View, CA 94040
U.S.A.
cchong@alphatech.com

Abstract – *This paper discusses evaluation of data association hypotheses for a general class of multiple target tracking problems. We assume that the number of targets is random and unknown, and given the number of targets, that the joint target state distribution forms a system of independent, identically distributed (i.i.d.) probability distributions. We are particularly interested in the cases in which the probability distribution of the number of targets is not necessarily Poisson. We will show that the Poisson assumption on the number of targets as well as the number of false alarms is not only sufficient but also necessary for the commonly used standard multiplicative hypothesis evaluation formula. Consequently, we claim that the use of the standard multiplicative hypothesis evaluation formula implies, explicitly or implicitly, the Poisson assumption.*

Keywords: Multiple target tracking, data association hypothesis evaluation, i.i.d. cluster, non-Poisson point process.

1 Introduction

Multiple target tracking requires a unique body of theory because (i) the number of objects (and hence the object state space dimension) is random, and (ii) any particular order in which those objects are arranged is either meaningless or otherwise does not contain any useful information. In a general multiple target tracking problem, the system state is a joint target state of an unknown number of targets, while information is given in terms of a collection of measurement sets. A measurement set (called also a scan or a frame) is a unit of information given as a set of measurements in a given measurement space taken by a sensor simultaneously or almost simultaneously. Mathematically, the above two characteristics (i) and (ii) should be identified with those of *random finite sets* ([1, 2]), or equivalently to a certain degree, with those of *finite point processes* ([3]).

Using these mathematical frameworks, any multiple target tracking problem can be stated as a problem for estimating the system state (a random finite set) based on a collection of measurements (random finite sets). In many applications, however, we are required to answer *data association questions* such as, "From which target does this measurement originate?", "Whether or not does this pair of measurements come from the same target?", etc. In fact, a common multiple target tracking practice is a two-step process, in which first (a) a set of data association hypotheses are generated and evaluated, and then (b) under each of selected hypotheses (generally multiple hypotheses), target states are estimated.

English words such as "originate" or "associate" sound very straightforward. Mathematically, however, their meaning is not trivial. Without a precise probabilistic definition we may not be able to apply Bayes rule to evaluate data association hypotheses. Any data association hypothesis is *inherently posterior* since it is formed on a given set of data (observation). In fact, we may view a data association hypothesis as a proposed partition of given data. On the other hand, the state-to-measurement transition can only be defined given a *hidden* random association that we may consider as *prior to* measurement generation. The only clear way to break this "cycle" is to separate measurement "values" from "indices," i.e., to treat the fact that a measurement is the "third" one in the set separately from the fact that its value is "3 meters." This approach was used to derive a solution to a general class of multiple target tracking problems in [4] as an extension of D. B. Reid's multiple hypothesis tracking algorithm [5]. The issues concerning how we should define data association hypotheses were recently re-visited in [6].

In this paper, we will re-visit the general model of multiple target tracking used in [4], and we will provide a solution in which targets are independent and identically distributed (i.i.d.) but the distribution of the number of targets is not necessarily Poisson. We have abundant examples where the probability distribution of the number of targets is not Poisson. Trivial examples include the cases where the number of targets is precisely known or there is a probability-one upper limit on the number of targets. Such information is often provided as collateral or intelligence information.

General non-Poisson problems were first explored in [7] but no significant results had been discovered until we showed, in [8], that the Poisson assumption is not only sufficient but also necessary for the commonly used standard multiplicative hypothesis evaluation formula for track-to-track association problems. This paper will show similar results for a general not-necessarily-Poisson class of multiple target tracking problems, as an extension of the work done in [4], in a batch-processing form. The effect of the Poisson assumption on the number of false alarms in each measurement set will also be examined.

We will start with a single-measurement-set problem (Secs. 2 and 3) and then move to a multiple-measurement-set, static target problem (Sec. 4). We will show the main result (Sec. 5) for a general multiple measurement set, dynamic target problem, followed by the conclusion in Section 6.

2 Target and sensor models

Let the target state space E be a *hybrid space*[1] defined as a direct-product metric space $E = \Re^d \times C$ of a Euclidean space[2] $\Re^d$ and a countable or finite set C with a discrete metric, associated with the *hybrid measure* μ that is the direct-product measure of the Lebesgue measure on $\Re^d$ and the counting measure on C. In this and the next sections, *targets* are defined as an unknown number n of states[3] $(x_i)_{i=1}^n$ in the state space E. We assume the following for targets $(x_i)_{i=1}^n$:

Assumption 1: *The number n of targets is a random nonnegative integer with finite but strictly positive mean $\bar{\nu}$, and given n, the n target states $(x_i)_{i=1}^n$ form a system of i.i.d. random elements in state space E with a common probability density function $\bar{p}$ on E.*

Let E_M be the measurement space that is also a hybrid space with a hybrid measure μ_M. Define a *measurement set* (or *a scan* or *a frame*) as a set of measurements taken by a sensor at the same time or at the approximately same time, and represent it by a sequence $(y_j)_{j=1}^m$ in E_M, where m is the number of measurements in the measurement set (i.e., a random integer) and the j-th measurement value y_j is a random element in E_M.

Before defining our single-measurement-set problem, let us make a general definition of a *random assignment*. Let $\mathcal{N}$ be the set of all the natural numbers.

Definition 1: *A random assignment α is a random finite set in the space $\mathcal{N} \times \mathcal{N}$ of pairs of positive integers, i.e., a random element such that every realization is a set of pairs of positive integers. Denote the domain and the image (range) of α by $\mathrm{Dom}(\alpha) = \bigcup_{j\in\mathcal{N}}\{i \in \mathcal{N} \,|(i,j) \in \alpha\}$ and $\mathrm{Im}(\alpha) = \bigcup_{i\in\mathcal{N}}\{j \in \mathcal{N} \,|(i,j) \in \alpha\}$, respectively. We say α is simple, if it is one-to-one, i.e., if, for every $i \in \mathrm{Dom}(\alpha)$, there exists a unique $j \in \mathcal{N}$ such that $(i,j) \in \alpha$ and $\#(\mathrm{Dom}(\alpha)) = \#(\mathrm{Im}(\alpha))$. When α is simple, we write $j = \alpha(i)$ whenever $(i,j) \in \alpha$.*

In this paper, we consider only a simple random assignment. Fig. 1 illustrates a simple target-to-measurement random assignment.

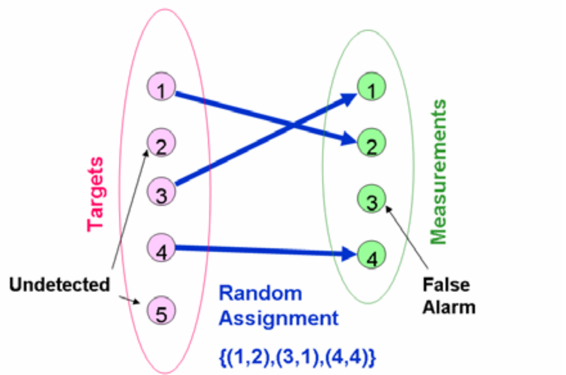

Fig. 1: Target-To-Measurement Random Assignment

Consider a target-to-measurement random assignment a which is a random assignment that, together with the single measurement set $(y_j)_{j=1}^m$, satisfies the following assumption.

Assumption 2: *The target set $(x_i)_{i=1}^n$, the measurement set $(y_j)_{j=1}^m$, and the target-to-measurement random assignment a satisfy the following assumptions:*

2-1: $\mathrm{Dom}(a) \subseteq \{1,\dots,n\}$ *and* $\mathrm{Im}(a) \subseteq \{1,\dots,m\}$ *with probability one. When $i \in \mathrm{Dom}(a)$, we say target i is detected in measurement set $(y_j)_{j=1}^m$. When $j = a(i)$, we say the j-th measurement originates from the i-th target, and when $j \notin \mathrm{Im}(a)$, we say the j-th measurement is a false alarm. Given $\mathrm{Dom}(a)$ and the number m of measurements in the measurement set, the target-to-measurement random assignment a is independent of the target set $(x_i)_{i=1}^n$ and every possible realization of it is equally probable*[4]*, i.e.,*

$P\left(a \,|\mathrm{Dom}(a), m\right) = \left(m - \#(\mathrm{Dom}(a))\right)! / m!,$

[1] The use of hybrid spaces facilitates simultaneous use of continuous and discrete variables, avoiding integration-summation mixtures. Mathematically we can use slightly more general spaces, i.e., locally compact Housdorff spaces satisfying the second axiom of countability without altering any result of this paper.

[2] $\Re = (-\infty, \infty)$ is the set of reals. The counting measure μ_C on any set C is defined as $\mu_C(D) = \#(D)$ for any finite subset D of C and $+\infty$ for any infinite subset D of C. In this paper, $\#(D)$ is the cardinality of (the number of elements in) any set D.

[3] $(x_i)_{i=1}^n$ is a shorthand for a finite sequence $(x_1,\dots,x_n)$.

[4] In this paper, $P(\cdot)$ and $P(\cdot|\cdot)$ are used as generic symbols for unconditional and conditional probability densities or probability mass functions or their mixtures.

2-2: the detection of targets is target-wise independent and determined by a detection probability function that is a measurable function $p_D : E \to [0,1]$,

2-3: the set[5] $\{y_j \mid j \in \{1,\dots,m\} \setminus \mathrm{Im}(a)\}$ *of false alarms in the measurement set* $(y_j)_{j=1}^m$ *is independent from the target states* $(x_i)_{i=1}^n$ *or from any other measurements* y_j. *The number,* $m_{FA} \stackrel{\text{def}}{=} m - \#(\mathrm{Im}(a))$, *of false alarms in the measurement set* $(y_j)_{j=1}^m$, *has a finite mean* ν_{FA}, *and given* m_{FA}, *the values of false alarms form an i.i.d. system of random elements in* E_M *with a common probability density function* p_{FA} *on* E_M, *and*

2-4: the target-to-measurement transition is target-wise independent and defined by a density $p_M(\cdot|\cdot)$ *of a transition probability from the target state space* E *to the measurement space* E_M.

The following lemma given without proof is an immediate consequence of these assumptions.

Lemma 1: *Assumption 2-2 implies*

$$P\left(\mathrm{Dom}(a) \middle| (x_i)_{i=1}^n\right) = \left(\prod_{i \in \mathrm{Dom}(a)} p_D(x_i)\right)\left(\prod_{i \in \{1,\dots,n\} \setminus \mathrm{Dom}(a)} (1 - p_D(x_i))\right) \quad (1)$$

Assumption 2-3 implies

$$\begin{aligned} P\left(m \middle| \mathrm{Dom}(a), (x_i)_{i=1}^n\right) &= P\left(m \middle| \mathrm{Dom}(a)\right) \\ &= P\left(m - \#(\mathrm{Dom}(a))\right) = P(m_{FA}) \end{aligned} \quad (2)$$

and Assumptions 2-3 and 2-4 imply

$$P\left((y_j)_{j=1}^m \middle| a, m, (x_i)_{i=1}^n\right) = \left(\prod_{i \in \mathrm{Dom}(a)} p_M\left(y_{a(i)} \middle| x_i\right)\right)\left(\prod_{j \in \{1,\dots,m\} \setminus \mathrm{Im}(a)} p_{FA}(y_j)\right) \quad (3)$$

3 Single-measurement-set data association hypothesis evaluation

A data association hypothesis on multiple measurement sets is a hypothesis concerning which measurement in one measurement set shares the same origin as which measurement in another measurement set, as we see in the next section. A single-measurement-set counterpart is just a hypothesis that jointly hypothesizes each measurement originating from a target or being a false alarm. In other words, it is nothing but $\mathrm{Im}(a)$ in the following sense.

Definition 2: *A single-measurement-set hypothesis is any subset* J *of* $\{1,\dots,m\}$, *hypothesizing that* y_j *originates from a target if* $j \in J$ *otherwise is a false alarm, or in other words, any realization of a random set* $\lambda \stackrel{\text{def}}{=} \mathrm{Im}(a)$ *in* $\{1,\dots,m\}$.

Then the evaluation of any single-measurement-set hypothesis becomes just a matter of calculating the posterior probability $P\left(\lambda \middle| (y_j)_{j=1}^m\right)$.

Theorem 1: *Under Assumptions 1 and 2, the a posteriori probability of the single-measurement-set hypothesis*[6] λ *conditioned by the single measurement set* $(y_j)_{j=1}^m$ *can be written as*

$$P\left(\lambda \middle| (y_j)_{j=1}^m\right) = P\left((y_j)_{j=1}^m\right)^{-1} \frac{L_{NDT}(\#(\lambda)) L_{NFA}(m - \#(\lambda))}{m!} \left(\prod_{j \in \lambda} \gamma_{NT}(y_j)\right)\left(\prod_{j \in \{1,\dots,m\} \setminus \lambda} \gamma_{FA}(y_j)\right) \quad (4)$$

where the γ_{NT} is the target detection density defined by

$$\gamma_{NT}(y) = \int_E p_M(y|x) p_D(x) \bar{\gamma}(x) \mu(dx) \quad (5)$$

with $\bar{\gamma}(x) = \bar{\nu}\,\bar{p}(x)$, γ_{FA} *is the false alarm density defined by* $\gamma_{FA}(y) = \nu_{FA} p_{FA}(y)$, L_{NDT} *is the likelihood function*[7] *for the number* $n_D = \#(\lambda)$ *of detected targets defined by*

$$L_{NDT}(n_D) = \sum_{n=n_D}^{\infty} P(n) \frac{n!}{\bar{\nu}^n} \frac{\hat{\nu}^{n-n_D}}{(n-n_D)!} \quad (6)$$

with $\hat{\nu} = \int_E (1 - p_D(x)) \bar{\gamma}(x) \mu(dx)$, *and* L_{NFA} *is the likelihood function for the number* $m_{FA} = m - n_D$ *of false alarms defined by* $L_{NFA}(m_{FA}) = P(m_{FA})\left(m_{FA}! / \nu_{FA}^{m_{FA}}\right)$.

Proof: Denote $x = (x_i)_{i=1}^n$ and $y = (y_j)_{j=1}^m$. Then we have $P(\lambda|y) = P(\lambda, y)/P(y)$ and $P(\lambda, y)$ is expanded as

---

[5] "\" is the set-theoretic subtraction operator, i.e., $A \setminus B = \{a \in A \mid a \notin B\}$, for any pair (A, B) of sets.

[6] Intentionally, we will confuse a random element and its realization.

[7] In this paper, the term, "likelihood," as in "track likelihood," used later, is not used in a traditional-statistics sense but rather it is used for any nonnegative function representing some kind of "likeliness."

$$P(\lambda, y) = \sum_{n=0}^{\infty} P(n) \int_{E^n} \sum_a P\left(\lambda \left| a, y, x\right.\right) P\left(y \left| a, m, x\right.\right)$$
$$P\left(a \left| \mathrm{Dom}(a), m, x\right.\right) P\left(m \left| \mathrm{Dom}(a), x\right.\right) \quad (7)$$
$$P\left(\mathrm{Dom}(a) \left| x\right.\right) \left(\prod_{i=1}^{n} \overline{p}(x_i) \mu(dx_i) \right)$$

Given $\lambda = \mathrm{Im}(a)$ and n, there are $n!/n_D!$ realizations of random assignment a that are compatible with λ, i.e., for which $P(\lambda | a, y, x) = P(\lambda | a, m, n) = 1$. $P(y | a, m, x)$ is the probability distribution density of the values of the measurements given by Assumptions 2-3 and 2-4. $P(a | \mathrm{Dom}(a), m, x) = m_{FA}!/m!$ by Assumption 2-1. $P(m | \mathrm{Dom}(a), x) = P(m_{FA})$ by Assumption 2-3. Finally $P\left((x_i)_{i=1}^n\right) = P(n) \prod_{i=1}^n \overline{p}(x_i)$. Hence Eqn. (7) follows from the definitions of L_{NDT}, L_{NFA}, γ_{NT} and γ_{FA}. *Q.E.D.*

Fig. 2 show the contour of $L_{NDT}(n_D) L_{NFA}(m_{FA})$ assuming that both $P(n)$ and $P(m_{FA})$ are uniformly distributed on $\{0,1,...,10\}$, and that $\hat{\nu} = 0.3\overline{\nu}$.

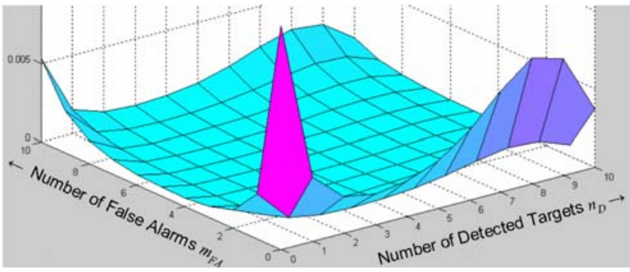

Fig. 2: Joint (n_D, m_{FA}) Likelihood

Using the Poisson assumption, we have the following immediate corollary:

Corollary 1: *Suppose the number n of targets and the number m_{FA} of false alarms in $(y_j)_{j=1}^m$ have Poisson distributions. Then we have*

$$P\left(\lambda \left| (y_j)_{j=1}^m \right.\right) = c\left((y_j)_{j=1}^m\right)^{-1} \left(\prod_{j \in \lambda} \gamma_{NT}(y_j) \right) \left(\prod_{j \in \{1,...,m\} \setminus \lambda} \gamma_{FA}(y_j) \right) \quad (8)$$

where the normalizing constant is given as
$c\left((y_j)_{j=1}^m\right) = P\left((y_j)_{j=1}^m\right) \ m! \ \exp(\tilde{\nu} + \nu_{FA})$.

Proof: The Poisson assumption means that $P(n) = e^{-\overline{\nu}} \overline{\nu}^n / n!$ and $P(m_{FA}) = e^{-\nu_{FA}} \nu_{FA}^{m_{FA}} / m_{FA}!$. Then Eqn. (8) follows from Eqn. (4) immediately. *Q.E.D.*

Definition 3: *The single-measurement-set hypothesis evaluation formula* (8) *is called the single-measurement-set standard multiplicative form*.

Theorem 1 and Corollary 1 show that the difference between Poisson and non-Poisson cases is the existence of an extra factor that generally depends on the number n_D of hypothesized detected targets and the number m of measurements (and hence the number $m_{FA} = m - n_D$ of false alarms). With the Poisson assumption, as shown by (8), the hypothesis evaluation can be done measurement-wise independently using the posterior odds, $\gamma_{NT}(y_j)/\gamma_{FA}(y_j)$. This standard multiplicative hypothesis evaluation form (8) is a commonly used form. The following theorem claims the Poisson assumption is not only sufficient but also necessary for us to use (8).

Theorem 2: *Suppose $P(n) > 0$ for every n (the number of targets), and $P(m_{FA}) > 0$ for every m_{FA} (the number of false alarms), and $\tilde{\nu} \overset{\text{def}}{=} \overline{\nu} - \hat{\nu} > 0$. Then, in order to have the standard multiplicative form* (8) *of the single-measurement-set hypothesis λ, the probability distributions of the number of targets and the false alarms both must be Poisson.*

Proof: Suppose we have the standard multiplicative form (8). Then, for every $m \geq 0$, $L_{NDT}(n_D) L_{NFA}(m - n_D)$ must be constant with respect to $n_D \in \{0,...,m\}$, which implies $L_{NDT}(n_D) = L_{NDT}(0) \beta^{n_D}$ for every $n_D \geq 0$, and $L_{NFA}(m_{FA}) = L_{NFA}(0) \beta^{m_{FA}}$ for every $m_{FA} \geq 0$, for some constant $\beta > 0$. By definition, $\nu_{FA} = \sum_{m_{FA}=1}^{\infty} m_{FA} P(m_{FA})$, and hence, we must have $\beta = 1$. Thus both L_{NDA} and L_{NFA} must be constant, and $P(m_{FA})$ is Poisson.

The necessity for $P(n)$ for the number n of targets follows from the unique existence of a solution to the infinite-dimensional linear equation,

$$S(n_D) = \sum_{n=n_D}^{\infty} \frac{\hat{\nu}^{n-n_D}}{(n - n_D)\ !} R(n) \quad (9)$$

between the two infinite sequences, $(R(n))_{n=0}^{\infty}$ and $(S(n))_{n=0}^{\infty}$. The non-singularity of the linear transformation (9) was proven in [8]. *Q.E.D.*

Remark 1: The definition for an unknown number of targets described in this section is consistent with the definition of a finite point process in [14]. A point process defined as an i.i.d. system given the number of states is called an *i.i.d. cluster* in [14], and with the Poisson assumption, called a *finite Poisson point process*. Generally a point process allows a repeated element, e.g., two targets occupying the same state. A point process without any repeated element is said to be *simple*. As argued in [3], a *simple finite point process* is essentially a *random finite set*. In this paper, this "simple-ness" is guaranteed if the measure of the diagonal sets is zero, i.e., if $d > 0$.

4 Multiple-measurement-set data association hypotheses

Consider K measurement sets $(y_k)_{k=1}^K$, such that $y_k = (y_{kj})_{j=1}^{m_k}$, accompanied by a target-to-measurement random assignment a_k, for each $k \in \{1,\ldots,K\}$. We will impose the following conditional independence assumption:

Assumption 3: *The measurement sets $(y_k)_{k=1}^K$ and the random assignments $(a_k)_{k=1}^K$ are both conditionally independent, i.e.,*

$$P\left((y_k)_{k=1}^K, (a_k)_{k=1}^K \,\middle|\, (x_i)_{i=1}^n\right) = \prod_{k=1}^K P\left(y_k, a_k \,\middle|\, (x_i)_{i=1}^n\right) \tag{10}$$

and Assumptions 2-1 to 2-4 hold for each y_k and a_k, i.e., when a is replaced by a_k, m by m_k, p_D by p_{Dk}, m_{FA} by m_{FAk}, ν_{FA} by ν_{FAk}, p_M by p_{Mk}, and E_M by E_{Mk}, indexed by measurement set indices k.

Then clearly Lemma 1 holds when we replace a by a_k, etc., and the following lemma holds:

Lemma 2: *Under Assumptions 1 and 3, for each k, we have*

$$\begin{aligned} P\left(y_k, a_k \,\middle|\, (x_i)_{i=1}^n\right) = {} & L_{NFAk}\left(m_k - \#\left(\mathrm{Dom}(a_k)\right)\right) \\ & \left(\prod_{j \in \{1,\ldots,m_k\} \setminus \mathrm{Im}(a_k)} \gamma_{FAk}(y_{kj})\right) \\ & \left(\prod_{i \in \mathrm{Dom}(a)} p_{Mk}\left(y_{ka_k(i)} \,\middle|\, x_i\right) p_{Dk}(x_i)\right) \\ & \left(\prod_{i \in \{1,\ldots,n\} \setminus \mathrm{Dom}(a)} \left(1 - p_{Dk}(x_i)\right)\right) \end{aligned} \tag{11}$$

with $L_{NFAk}(m_{FAk}) = P(m_{FAk})\left(m_{FAk}! / \nu_{FAk}^{m_{FAk}}\right)$.

Proof: Eqn. (11) follows from Assumption 3 and a straightforward Bayesian expansion. *Q.E.D.*

The multiple-measurement-set data association hypotheses are then defined as follows:

Definition 4: *A multiple-measurement-set data association λ on the cumulative measurement set $(y_k)_{k=1}^K = \left((y_{kj})_{j=1}^{m_k}\right)_{k=1}^K$ is defined through $(a_k)_{k=1}^K$ as*

$$\lambda = \left\{ \left\{ (k, a_k(i)) \,\middle|\, k \in \{1,\ldots,K\} \right\} \,\middle|\, i \in \bigcup_{k=1}^K \mathrm{Dom}(a_k) \right\} \tag{12}$$

Any realization (sample) of λ is called a data association hypothesis on $(y_k)_{k=1}^K$, and a realization of any member, $\tau = \left\{ (k, a_k(i)) \,\middle|\, k \in \{1,\ldots,K\} \right\}$ for some i, of λ is called track (originating from the i-th target).

Fig. 3 illustrates the definition of a multiple-measurement-set data association hypothesis.

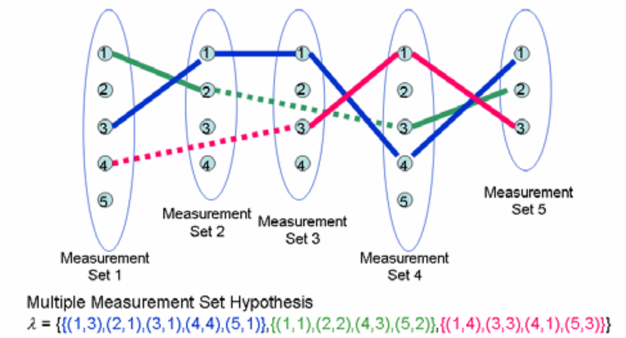

Fig. 3: Data Association Hypothesis

The following theorem is the multiple-measurement-set version of Theorem 1.

Theorem 3: *Under Assumptions 1 and 3, the data association λ is evaluated by*

$$\begin{aligned} P\left(\lambda \,\middle|\, (y_k)_{k=1}^K\right) = {} & P\left((y_k)_{k=1}^K\right)^{-1} \\ & \left(\prod_{k=1}^K \frac{L_{NFAk}\left(m_k - \left(\#\lambda_{|k}\right)\right)}{m_k!} \left(\prod_{j \in \{1,\ldots,m_k\} \setminus \lambda_{|k}} \gamma_{FAk}(y_j)\right)\right) \\ & L_{NDT}\left(\#(\lambda)\right)\left(\prod_{\tau \in \lambda} \ell_K(\tau)\right) \end{aligned} \tag{13}$$

where $\lambda_{|k} = \left\{ j \in \{1,\ldots,m_k\} \,\middle|\, (k,j) \in \cup\lambda \right\}$ for each k, L_{NFAk} is defined in Lemma 2's statement, the number-of-detected-target likelihood $L_{NDT}(n_D)$ is defined by (6) *with the expected number of undetected target $\hat{\nu}$ being replaced by $\hat{\nu}_K$, defined by*

$$\hat{\nu}_K = \int_E \left(\prod_{k=1}^K \left(1 - p_{Dk}(x)\right)\right) \bar{\gamma}(x) \mu(dx) \tag{14}$$

and $\ell_K(\tau)$ is the track likelihood function defined by

$$\ell_K(\tau) = \int_E \left(\prod_{k=1}^K q_k(x;\tau)\right) \bar{\gamma}(x) \mu(dx) \tag{15}$$

with

$$q_k(x;\tau) = \begin{cases} p_{Mk}\left(y_{kj} \,\middle|\, x\right) p_{Dk}(x) & \text{if } (k,j) \in \tau \\ \left(1 - p_{Dk}(x)\right) & \text{otherwise} \end{cases} \tag{16}$$

Proof: As in Theorem 1, we have

$P\left(\lambda \left|(y_k)_{k=1}^K\right.\right) = P\left((y_k)_{k=1}^K\right)^{-1} P\left(\lambda,(y_k)_{k=1}^K\right)$. $P\left(\lambda,(y_k)_{k=1}^K\right)$ is then expanded as

$$P\left(\lambda,(y_k)_{k=1}^K\right) = \sum_{n=\#(\lambda)}^{\infty} P(n) \int_{E^n} \sum_{(a_k)_{k=1}^K} P\left(\lambda \left|(a_k)_{k=1}^K\right.\right) \left(\prod_{k=1}^K P\left(y_k, a_k \left|(x_i)_{i=1}^n\right.\right)\right) \prod_{i=1}^n \overline{p}(x_i)\mu(dx_i) \tag{17}$$

using the conditional independence assumption (Assumption 3). The rest of the expansion to obtain (17) is very similar to that to obtain (4). *Q.E.D.*

Then, with the Poisson assumption, we have

Corollary 2: *Assume that the probability distribution $P(n)$ of the number of targets is Poisson, and for each k, the distribution $P(m_{FAk})$ of the number of the false alarm in measurement set $y_k = (y_{kj})_{j=1}^{m_k}$ is Poisson, in addition to Assumptions 1 and 3. Then we have*

$$P\left(\lambda \left|(y_k)_{k=1}^K\right.\right) = c\left((y_k)_{k=1}^K\right)^{-1} \ell_{FAK}(\lambda)\left(\prod_{\tau\in\lambda} \ell_K(\tau)\right) \tag{18}$$

where

$$\ell_{FAK}(\lambda) = \prod\left\{\gamma_{FAk}(y_{kj}) \left| \begin{array}{l} k\in\{1,\ldots,K\} \text{ and } j\in\{1,\ldots,m_k\} \\ \text{and } (k,j)\notin\bigcup\lambda \end{array}\right.\right\}$$

is the total false alarm likelihood, and

$$c\left((y_k)_{k=1}^K\right) = P\left((y_k)_{k=1}^K\right)\left(\prod_{k=1}^K m_k!\right)\exp\left(\tilde{\nu}_K + \sum_{k=1}^K \nu_{NFAk}\right)$$

is the re-defined normalizing constant with $\tilde{\nu}_K = \overline{\nu} - \hat{\nu}_K$.

Proof: (18) is obtained by inserting Poisson distributions into (13). *Q.E.D.*

Definition 5: *The multiple-measurement-set data association hypothesis evaluation formula* (18) *is called the multiple-measurement-set standard multiplicative form.*

The standard multiplicative form is the batch-processing hypothesis evaluation form. As shown in the next section, its dynamic-target version is commonly used. The Poisson assumptions used to derive this formula are, however, rarely mentioned. The following theorem claims that these Poisson assumptions are indeed essential.

Theorem 4: *Under Assumptions 1 and 3, suppose $P(n) > 0$ for every n, $P(m_{FAk}) > 0$ for every m_{FAk} for each k, and $\tilde{\nu}_K \overset{\text{def}}{=} \overline{\nu} - \hat{\nu}_K > 0$. Then, in order for us to have the multiple-measurement-set standard multiplicative form* (18)*, the Poisson assumptions on the number of targets and the number of false alarms in each measurement set k are both necessary.*

Proof: Suppose the standard multiplicative form (18) holds. Then, for any given $(m_k)_{k=1}^K$, $L_{NDT}(\#(\lambda))$ times $\prod_{k=1}^K L_{NFAk}(m_k - \#(\lambda_k))$ must be constant for all the possible multiple-measurement-set data association hypotheses λ. By letting $m_2 = \cdots = m_K = 0$, varying $m_1 = 0,1,2,\ldots,$ and $\#(\lambda) = \#(\lambda_1) \in \{0,\ldots,m_1\}$, we can prove that both $L_{NDT}(n_D)$ and $L_{NFA1}(m_{FA1})$ must be constant, and $P(m_1 - \#(\lambda))$ is Poisson, as shown in the proof of Theorem 2. By repeating the same argument for $k=2$ to K, we see, for all $k\in\{2,\ldots,K\}$, that $L_{NFAk}(m_{FAk})$ must also be constant, and each $P(m_{FAk})$ is Poisson. We can then prove that $P(n)$ is Poisson, in the exactly same way as we did for Theorem 2, using the non-singular-ness of the linear operator on the infinite dimensional spaces, described by (9). *Q.E.D.*

Remark 2: We should note that, to calculate the posterior probability $P\left(\lambda \left|(y_k)_{k=1}^K\right.\right)$, the application of the usual Bayes rule,

$$P\left(\lambda \left|(y_k)_{k=1}^K\right.\right) = P\left((y_k)_{k=1}^K\right)^{-1} P\left((y_k)_{k=1}^K \left|\lambda\right.\right)P(\lambda),$$

is useless, since the data association hypothesis λ is a hypothesis on data, i.e., a posterior entity, although $P(\lambda)$ can be calculated by taking average over all the possible measurement values in (17). For the same reason, the likelihood approach in the traditional-statistics sense, i.e., the attempt to calculate $P\left((y_k)_{k=1}^K \left|\lambda\right.\right)$ does not make sense, because the definition of hypothesis λ depends on $(y_k)_{k=1}^K$. It is also clear from (13) or (18) that any Poisson distribution $P(n)$ is not a *conjugate prior.*

5 Dynamical target cases

As in the previous section, we consider K measurement sets, $(y_k)_{k=1}^K = \left((y_{kj})_{j=1}^{m_k}\right)_{k=1}^K$ but we now consider dynamical targets. In order to do so, instead of modeling the targets by a set of an unknown number n of elements, $(x_i)_{i=1}^n$, in the target state space E, we model them by a random element, $\left((x_{ik})_{k=0}^K\right)_{i=1}^n$, such that each x_{ik} is a random element in E, i.e., $x_i \overset{\text{def}}{=} (x_{ik})_{k=0}^K$ is a random element in E^{K+1}. In other words, we simply expand the target state space from E to E^{K+1}. We write $x_{ik} = x_i(t_k)$ where t_k is the time when the k-th measurement set $y_k = (y_{kj})_{j=1}^{m_k}$ is taken, for each $k>0$, to make the meaning of the double index for $x_i \overset{\text{def}}{=} (x_{ik})_{k=0}^K$ clear. We should note that we have put an extra time epoch t_0 to put an initial condition for the dynamical targets.

We assume the following:

Assumption 4: *Given the number n of targets, the extended target states $\left(\left(x_i\left(t_k\right)\right)_{k=0}^{K}\right)_{i=1}^{n}$ is an i.i.d. system of random elements in E^{K+1} with a common probability distribution function $\bar{p}\left(\left(x_i\left(t_k\right)\right)_{k=0}^{K}\right)$, with $t_0 \le t_1 \le \cdots \le t_K$, where each t_k is the time when the k-th measurement set $y_k = \left(y_{kj}\right)_{j=1}^{m_k}$ is taken if $k > 0$.*

Assume that the measurement sets $\left(y_k\right)_{k=1}^{K}$ and the random assignment $\left(a_k\right)_{k=1}^{K}$ are conditionally independent in the following sense,

$$P\left(\left(y_k\right)_{k=1}^{K}, \left(a_k\right)_{k=1}^{K} \left| \left(\left(x_i\left(t_k\right)\right)_{k=0}^{K}\right)_{i=1}^{n}\right.\right) = \prod_{k=1}^{K} P\left(y_k, a_k \left| \left(x_i\left(t_k\right)\right)_{i=1}^{n}\right.\right) \tag{19}$$

and Assumptions 2-1 to 2-4 hold for each y_k and a_k, i.e., when a is replaced by a_k, m by m_k, p_D by p_{Dk}, m_{FA} by m_{FAk}, ν_{FA} by ν_{FAk}, p_M by p_{Mk}, and E_M by E_{Mk}.

With this assumption, all the claims made in the previous section remain valid when we replace x_i by $x_i\left(t_k\right)$ or $\left(x_i\left(t_k\right)\right)_{k=1}^{K}$, whichever is appropriate. If we can further impose a commonly used property, the Markovian property (shown below), the evaluation formula becomes one of the most commonly used data association hypothesis evaluation formulae.

Assumption 5: *The common joint probability distribution $\bar{p}\left(\left(x_i\left(t_k\right)\right)_{k=0}^{K}\right)$ is Markovian in the sense,*

$$\bar{p}\left(\left(x_i\left(t_k\right)\right)_{k=0}^{K}\right) = \left(\prod_{k=1}^{K} f_{t_k t_{k-1}}\left(x_i\left(t_k\right) \left| x\left(t_{k-1}\right)\right.\right)\right) \bar{p}_0\left(x_i\left(t_0\right)\right) \tag{20}$$

with an appropriate initial condition probability density $\bar{p}_0$ on E and the density $f_{t_k t_{k-1}}$ of the state transition probability on E for each $k > 0$.

Using Assumptions 4 and 5 in place of Assumption 1, the track likelihood function $\ell_K\left(\tau\right)$ can be calculated from the partial track likelihood function $\ell_k\left(\tau\right)$ defined as

$$\ell_k\left(\tau\right) = \int_{E^{K+1}} \left(\prod_{\kappa=1}^{k} q_\kappa\left(\xi_\kappa ; \tau\right)\right) \bar{\gamma}\left(\left(\xi_k\right)_{k=0}^{K}\right) \prod_{k=0}^{K} \mu\left(d\xi_k\right) \tag{21}$$

with $\bar{\gamma}\left(\left(\xi_k\right)_{k=0}^{K}\right) = \bar{\nu}\, \bar{p}\left(\left(\xi_k\right)_{k=0}^{K}\right)$. With the Markovian assumption (20), the partial track likelihood function $\ell_k\left(\tau\right)$ can be calculated recursively as

$$\ell_k\left(\tau\right) = \ell_{k-1}\left(\tau\right) \tilde{q}_k\left(\tau\right) \tag{22}$$

where

$$\tilde{q}_k\left(\tau\right) = \begin{cases} \int_E p_{Mk}\left(y_{kj} \left| x\right.\right) p_{Dk}\left(x\right) \bar{p}_k\left(x \left| \tau\right.\right) \mu\left(dx\right) & \text{if } \left(k, j\right) \in \tau \\ \int_E \left(1 - p_{Dk}\left(x\right)\right) \bar{p}_k\left(x \left| \tau\right.\right) \mu\left(dx\right) & \text{otherwise} \end{cases} \tag{23}$$

for each $k \in \left\{1, ..., K\right\}$.

Each $\bar{p}_k\left(x \left| \tau\right.\right)$ in (23) is the *a posteriori* probability density of the target state at time t_k, conditioned by $\left(y_{\kappa j}\right)_{\left(\kappa, j\right) \in \tau}$ for $1 \le \kappa < k$, and can be updated to $\hat{p}_k\left(x \left| \tau\right.\right)$, the a posteriori probability distribution density of the target state at time t_k, conditioned by $\left(y_{\kappa j}\right)_{\left(\kappa, j\right) \in \tau}$ for $1 \le \kappa \le k$, by

$$\hat{p}_k\left(x \left| \tau\right.\right) = \begin{cases} \dfrac{p_{Mk}\left(y_{kj} \left| x\right.\right) p_{Dk}\left(x\right) \bar{p}_k\left(x \left| \tau\right.\right)}{\int_E p_{Mk}\left(y_{kj} \left| \xi\right.\right) p_{Dk}\left(\xi\right) \bar{p}_k\left(\xi \left| \tau\right.\right) \mu\left(d\xi\right)} & \text{if } \left(k, j\right) \in \tau \\ \dfrac{\left(1 - p_{Dk}\left(x\right)\right) \bar{p}_k\left(x \left| \tau\right.\right)}{\int_E \left(1 - p_{Dk}\left(\xi\right)\right) \bar{p}_k\left(\xi \left| \tau\right.\right) \mu\left(d\xi\right)} & \text{otherwise} \end{cases} \tag{24}$$

As usual, this *update formula* is accompanied by the *extrapolation formula*

$$\bar{p}_k\left(x \left| \tau\right.\right) = \begin{cases} \int_E f_{t_k t_{k-1}}\left(x \left| \xi\right.\right) \hat{p}_{k-1}\left(\xi \left| \tau\right.\right) \mu\left(d\xi\right) & \text{if } t_k > t_{k-1} \\ \hat{p}_{k-1}\left(x \left| \tau\right.\right) & \text{if } t_k = t_{k-1} \end{cases} \tag{25}$$

and the initial condition, $\hat{p}_0\left(\xi \left| \tau\right.\right) = \bar{p}_0\left(x\right)$.

Then we will state the multiple-measurement-set data association hypothesis evaluation for dynamic cases by the following two theorems and corollary without proof:

Theorem 5: *Under Assumptions 4 and 5, the data association λ is evaluated by Eqn.* (13) *with the track likelihood function $\ell_K\left(\tau\right)$ calculated through Eqns.* (21 – 25).

Corollary 3: *Under Assumptions 4 and 5, the Poisson assumptions on the distribution $P\left(n\right)$ of the number of targets and on the distribution $P\left(m_{FAk}\right)$ of the number of false alarms for each k are a sufficient condition for us to have the multiple-measurement-set standard multiplicative form* (18) *of data association hypothesis evaluation.*

Theorem 6: *Under Assumptions 4 and 5, assume $P\left(n\right) > 0$ for every n, $P\left(m_{FAk}\right) > 0$ for every m_{FAk}, for each k, and $\tilde{\nu}_K = \bar{\nu} - \hat{\nu}_K > 0$. The Poisson assumptions are necessary for the standard multiplicative form* (18).

Remark 3: The standard multiplicative form (18) for multiple-measurement-set data association hypothesis evaluation, modified to dynamical Markovian targets, as shown above, was first introduced in [9]. Subsequently it was used for the *track-oriented multiple hypothesis tracking algorithms* ([10,11]), and for the *multiple-frame assignment* (*multiple-dimension data* association) algorithm ([12,13]). The Poisson assumptions are rarely mentioned, despite the fact, as shown in Theorem 4, that they are necessary to derive this standard multiplicative form. As shown by Theorem 3, whenever the Poisson assumptions are not valid, we must put additional multipliers $L_{NDT}\left(\#(\lambda)\right)$ and $L_{NFAk}\left(m_k - \#(\lambda_k)\right)$ for each measurement set k. Lack of specification of the *a priori* statistics for the number of targets is sometimes explained as the use of the maximum likelihood approach in the traditional statistics sense. However, it is rather clear that, even when we artificially let the a priori probability distribution $P(n)$ converge to a constant function, we will not have the standard multiplicative form (18) for evaluating data association hypotheses, indicating the right hand side of Eqn. (18) does not conform with the likelihood function in the traditional-statistics sense.

6 Conclusion

A general data association hypothesis evaluation formula was derived for a general class of multiple target tracking problems, with i.i.d. targets and false alarms, without the standard Poisson assumptions. The resulting evaluation formula is a slight modification of the formula with the Poisson assumptions, with extra factors that vary explicitly with the assumed numbers of detected targets and false alarms. When these extra factors are constant, we call the evaluation formula "standard multiplicative form." We showed that the Poisson assumptions are not only sufficient but also necessary for us to derive this standard multiplicative form that is commonly used by most of the multiple hypothesis algorithms. This means that the Poisson assumptions are implicitly used by many multiple hypothesis target tracking algorithms, even though they are rarely stated so.

This paper treats a fairly general case of multiple target tracking problems. However, we did not treat the cases where (A) the number of targets changes in time, and (B) there may be merged or split measurement (We have only considered *simple* random assignments). Although it is relatively easy to expand our results to the former cases by including a {unborn,alive,dead} type state component for each target, extension for the latter cases may be difficult. One difficulty is lack of an appropriate mathematical model for merged or split measurements.

We have discussed only batch-processing formulations used for track-oriented multiple hypothesis algorithms or multiple-frame assignment algorithms. We can easily derive a recursive evaluation formula from the batch-processing formula. Therefore, the conclusion is the same for the recursive data association hypothesis evaluation, i.e., the Poisson assumptions are a necessary condition for deriving the commonly used multiplicative formula.

References

[1] Ronald P. S. Mahler. Random sets as a foundation for general data fusion. In *Proceedings of the Sixth Joint Service Data Fusion Symposium*, Vol. I, Part 1, John Hopkins University, Applied Physics Laboratory, Laurel, MD, pp. 357-394, 1993.

[2] Shozo Mori. Random sets in data fusion: multi-object, state-estimation as a foundation of data fusion theory. In *Random Sets – Theory and Applications -,* ed. by John Goutsias, Ronald P.S. Mahler, and Hung T. Nguyen, Springer Verlag, 1996.

[3] Shozo Mori and Chee-Yee Chong. Point process formalism of multitarget tracking. In *Proc. 5th International Conference on Information Fusion*, Annapolis, MD, July 2002.

[4] Shozo Mori, Chee-Yee Chong, Edison Tse, and Richard P. Wishner. Tracking and classifying multiple targets without *a priori* Identification. *IEEE Transaction on Automatic Control*, Vol. AC-31, No. 5, pp. 401-409, 1986.

[5] Donald B. Reid. An algorithm for tracking multiple targets. *IEEE Transaction on Automatic Control*, Vol. AC-24, No. 6, pp. 843-854, 1979.

[6] Shozo Mori and Chee-Yee Chong. What are data association hypotheses?. In *Proc. Workshop on Multiple Hypothesis tracking, A Tribute to Sam Blackman,* San Diego, CA, May 2003.

[7] Shozo Mori and Chee-Yee Chong. A multitarget tracking algorithm - independent but non-Poisson cases. In *Proc. of the 1985 American Control Conference,* Boston, MA, pp. 1054 - 1055, June 1985.

[8] Shozo Mori and Chee-Yee Chong. Track-to-track association metric – i.i.d. but non-Poisson cases -. In *Proc. 6th International Conference on Information Fusion*, Cairns, Australia, July 2003.

[9] Charles L. Morefield. Application of 0-1 integer programming to multi-target tracking problems. *IEEE Transaction on Automatic Control*, Vol. AC-22, No. 3, pp. 302-312, June 1977.

[10] Thomas G. Allen, Thomas Kurien, and Robert B. Washburn. Parallel computer structures for multiobject tracking algorithms on associative processors. In *Proc. American Control Conf.*, Seattle, WA, June, 1986.

[11] George C. Demos, Roberto A. Ribas, Ted J. Broida, and Samuel S. Blackman. Application of MHT to dim moving targets. In *Proc. Symposium on Signal and Data Processing of Small Targets,* Vol.1305 (ed. by Oliver E. Drummond), pp.297-309, April,.1990.

[12] Krishna R. Pattipati, Somnath Deb, and Yaakov Bar-Shalom. Passive multisensor data association using a new relaxation algorithm. In *Multitarget-Multisensor Tracking: Advanced Applications*, ed. by Yaakov Bar-Shalom, Artech House, 1990.

[13] Aubrey B. Poore and Nenad Rijavec. A new class of methods for solving data association problems arising from multitarget tracking. In *Proc. the 1991 American Automatic Control Conference*, Vol. 3, pp. 2302-2304, Boston, MA, 1991.

[14] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes.* Springer-Verlag, 1988.