Three Mathematical Formalisms of Multiple Hypothesis Tracking

SHOZO MORI
CHEE-YEE CHONG
KUO-CHU CHANG

This paper describes three mathematical formalisms, each of which provides a solid foundation for developing multiple hypothesis tracking (MHT) theories and algorithms, as solutions to detection-based multiple target tracking (MTT) problems. The three formalisms are 1) random finite sequence (RFSeq), 2) finite point process (FPP), and 3) random finite set (RFSet) formalisms. We will discuss equivalencies and some subtle differences among them. In addition, we will discuss theoretical consequences of various assumptions on MHT hypothesis evaluation, as well as recent RFSet-based MTT algorithm developments claiming relationship to MHT.

Manuscript received February 1, 2019; revised June 14, 2019, September 22, 2019, November 9, 2019; released for publication 0, 0000.

Refereeing of this contribution was handled by Jason L. Williams.

S. Mori is with Systems & Technology Research, Sacramento, CA, 95823, USA (E-mail: smori@ieee.org).
C.-Y. Chong is an Independent Researcher, Los Altos, CA, 94024, USA (E-mail: cychong@ieee.org).
K.-C. Chang is with the Department of Systems Engineering and Operations Research (SEOR), George Mason University, Fairfax, VA, 22030, USA (E-mail: kchang@gmu.edu).

1557-6418/19/$17.00 © 2019 JAIF

I. INTRODUCTION

This paper is generally concerned with *multiple target tracking* (MTT) problems, as defined in [1]–[3], i.e., problems of tracking a generally unknown number of objects, called *targets*,[1] based on noisy data. Specifically, we are concerned with a particular class of MTT problems, where the information is provided by generally multiple sensors in terms of finite sets of noisy measurements, called *target detections*,[2] without any explicit indication of their origins. This class of problems is sometimes referred to as *point target tracking*, due to the fact that each target is modeled as a point in a target state space, or each target appears and is detected as a point in a sensor measurement space. It may also be referred to as *tracking small targets*,[3] for the same reason. Our focus is on a particular class of solutions based on systematic generation and evaluation of multiple *data association hypotheses*, which hypothesize the number of detected targets and partition the set of all the acquired detections according to their hypothesized common origins, customarily referred to as *multiple hypothesis tracking* (MHT).[4]

To the best of our knowledge, various approaches to MTT problems were first comprehensively described in an MTT survey paper [4]. It referenced the two seminal works: one by C. L. Morefield [5] and the other by D. B. Reid [41] (the work of which was subsequently published as [6][5]). These two works constitute, in our opinion, the first two significant MHT developments. In [5], C. L. Morefield established the MHT foundation by defining *association hypotheses*, each of which is a set of *tracks*, and proposed the best hypothesis selection using a zero–one integer linear programming algorithm. D. B. Reid, in [6], presented a recursive MHT algorithm that propagates and evaluates multiple tracks and hypotheses, both recursively. Subsequently in [7], a generalized recursive MHT algorithm was developed, as a Bayesian optimal solution, showing the optimality as a clear theoretical consequence of mathematical models of targets and sensors, and a set of statistical assumptions, complementing the developments by Morefield [5] and Reid [6].

Generation and evaluation of multiple hypotheses are often considered as an intermediate step toward Bayesian estimation[6] of the states of an unknown number of targets. However, in some applications, the Bayesian estimation of measurement-to-measurement (data-to-data) association is of primary interest and

[1] In MTT, a target is a generic name for any *object* to be tracked.
[2] Also called *contacts, returns*, or simply *measurements* or *observations.*
[3] As opposed to *extended targets* with possibly multiple observations from each single target.
[4] In MHT, by *hypotheses* we always mean data association hypotheses.
[5] The introduction section of [6] contains an excellent summary of the early works on MTT problems.
[6] By *Bayesian estimation*, we mean a process to obtain analytical or numerical expressions of the conditional probability distributions of the states to be estimated, conditioned by available data (information).

an important task by itself. In that regard, the MHT has maintained a unique and important realm within the MTT universe. The evolution of various MHT and MHT-related MTT algorithms over the last 40 years was extensively and comprehensively described in [8], and will not be repeated in this paper. It is not the objective of this paper to describe and compare various MHT algorithms, or to discuss related implementation issues, or to present yet another new algorithm. Instead, our objective is to present three different mathematical formalisms, any one of which provides a solid theoretical foundation to support MHT concept and algorithm developments.

The three formalisms are 1) *random finite sequence* (RFSeq) formalism, 2) *finite point process* (FPP) formalism, and 3) *random finite set* (RFSet) formalism. These three are seemingly quite distinct from each other on the surface, but essentially equivalent to each other in a specific sense, as we discuss in this paper. We will define MHT problem in each formalism, with precise mathematical definitions to commonly used terms that have been often loosely defined, such as "originate from," "associated with," "assigned to," "tracks," "hypotheses," etc. We hope that, showing the uses of these three formalisms, side by side, we will be able to present a clear and precise picture of the past, and the potential future MHT developments. An earlier version of this paper was presented in [9], to which we added some analyses on specific consequences of commonly used assumptions, and our perspectives on relations of a selected set of recently developed RFSet-based MTT algorithms to MHT.

The rest of the paper is organized as follows: Section II presents the three mathematical formalisms for MTT, including all the relevant mathematical concepts in algebra, topology, and probability theories. Section III introduces target and sensor models, in the three formalisms, and defines data association hypotheses, to form a standard MHT problem. Section IV discusses generation and evaluation of association hypotheses, and describes the optimal Bayesian solution to MHT problem, in each of the three formalisms, under a set of commonly used assumptions. It is followed by Section V that discusses relationship of the MHT solution of Section IV with a selected set [36]–[39] of recently developed RFSet-based MTT algorithms. We will state our concluding remarks in Section VI.

II. THREE MATHEMATICAL FORMALISMS

We define MTT as a process of estimating the states of a generally unknown number of objects, called targets, generally changing their states over time with given stochastic dynamics, based on information collected by generally multiple sensors on regular or irregular observation schedules. As we often do, in this paper, any target is identified with its state, i.e., a point in a state space E, which we assume is a *locally compact Hausdorff space satisfying the second axiom of countability* (LCHC2)[7] [10]. Any countable set with discrete topology, as well as any Euclidean space, is LCHC2. Let $\mathcal{B}$ be the collection of Borel sets in E, i.e., the smallest σ-algebra containing all the open sets in E, and we assume that a σ-finite measure μ on the measurable space $(E, \mathcal{B})$ is given. Throughout this paper, we will maintain the measure set $(E, \mathcal{B}, \mu)$ as the target state space. To track n targets, $(x_1, ..., x_n)$, each in E, we use the state space defined as the nth-order direct product $E^n = \overbrace{E \times \cdots \times E}^{n \text{ times}}$ with the direct product topology (inheriting LCHC2), the direct product σ-algebra $\mathcal{B}_n$, and the direct product measure μ^n.

The measure space $(E^n, \mathcal{B}_n, \mu^n)$ with a fixed n provides us with a natural basis for the generalization[8] of the *probabilistic data association* (PDA) ($n = 1$ [12]) and the *joint PDA* (JPDA) ($n \geq 1$ [13]) algorithms to track a fixed number n of targets, each of which has its existence established, as target state x_i within the joint state[9] $(x_i)_{i=1}^n$, and is given a unique distinct *a priori identification* $i \in \{1, ..., n\}$. The main focus of this paper is, however, to present mathematical formalisms to provide a basis for tracking *targets without a priori identification* in the sense that 1) the number n of targets is generally unknown a priori and 2) given n, any particular ordering of the joint states $(x_i)_{i=1}^n$ is arbitrary. These facts necessitate 1) considering all the possible numbers n (any nonnegative integer) of targets and 2) requiring any particular joint target state probability distribution to be *permutable* or *symmetric*, which is an important aspect of this class of MTT that we are exclusively concerned with in this paper.

Remark 1 (Notations: Finite Sequences and Finite Sets): $(x_i)_{i=1}^n \in E^n$ is shorthand[10] of $(x_1, ..., x_n)$, a finite sequence in space E with length n, or an n-tuple of points in E. Sometimes, it will be necessary to use a nested expression to shorten $((y_{11}, ..., y_{1m_1}), ..., (y_{K1}, ..., y_{Km_K}))$ as $((y_{kj})_{j=1}^{m_k})_{k=1}^K$ with double-indexed variables y_{kj}. We also use continuous index, e.g., $(x_t)_{t\in[t_0,\infty)} \in E^{[t_0,\infty)}$ for a function defined on time index set $[t_0, \infty)$. If the index set I is a finite set, by $(x_i)_{i\in I}$, we mean a function x defined on I, but we may also mean a sequence $(x_{i_1}, ..., x_{i_n})$ with an arbitrary enumeration $(i_1, ..., i_n)$ of set I. By $\{x_i\}_{i=1}^n$, we mean $\{x_i\}_{i=1}^n = \bigcup_{i=1}^n \{x_i\}$, which is a set of n elements if x_i's are all distinct, where $\{x\}$ is the singleton with only single element x.

[7] Also known as *locally compact Hausdorff second-countable* topological space. See Remark 2, later in this section, for more explanations on the meaning of this choice of the state space as the basis of our paper.
[8] See Remark 6 in Section III-C, for more comments on PDA and JPDA algorithms, and their generalizations.
[9] See Remark 1.
[10] We consciously avoided the notation such as $(x_i)_{i=1:n}$ or $x_{1:n}$, in favor of $(x_i)_{i=1}^n$, since 1:n or n:m, used as a "colon" MATLAB syntax, is also used for a one-to-many or a many-to-many relationship in database designs, while $(x_i)_{i=1}^n$ is universally used in the mathematical literature (although $\langle x_i \rangle_{i=1}^n$ is used instead in [28]).

Remark 2 (LCHC2): If an LCHC2 space is a vector space, the local compactness implies a finite dimension [11, Th. 1.22, p. 17], so that we are excluding any infinite-dimensional state space in this paper. On the other hand, the countability implies being metrizable and separable (having a countable, dense subset) [10, Ch. 6, p. 241]. Non-Euclidean examples include a hybrid space (direct product of a Euclidean space (e.g., for kinematic sates) and a finite set (e.g., for discrete attribute states)), an ellipsoidal surface (for surface ship tracking), other one- or two-dimensional manifolds (e.g., for targets in road networks), Lie group $SO(3)$ (coupled with Lie algebra $so(3)$), the space of unit quaternions for attitude estimation, etc. Thus, it seems to us that this LCHC2 assumption may specify a necessary mathematical sphere for us to cover for all the application domains that we, engineers, may be interested in, beyond familiar Euclidean spaces. On the other hand, we understand that this LCHC2 assumption allows us to almost "freely" use familiar notions of conditional probability distributions, densities, Bayes rules, stochastic processes, etc., without fear of any mathematical pathology.

A. RFSeq Formalism

Randomness of the number n of targets forces us to consider all the spaces E^n, for $n = 0, 1, 2, ...,$ together, as the direct-sum space[11] $\bigcup_{n=0}^{\infty} E^n$, using the standard convention $E^0 = \{\theta\}$ with the symbol[12] θ for the sequence of the zero length, signifying "nothing," or in our case "no target." Algebraically, $\bigcup_{n=0}^{\infty} E^n$ is the *free monoid* (FM) generated by E (as the set of its alphabets [14][13]), with the concatenation operator $*$ as an associative binary operator, defined by $(x_i)_{i=1}^{n} = (x_i)_{i=1}^{n'} * (x_i)_{i=n'+1}^{n}$ for any $0 \leq n' \leq n$ and any $(x_i)_{i=1}^{n} \in E^n$ with the identity element θ. Topologically, $\bigcup_{n=0}^{\infty} E^n$ is also an LCHC2 with the direct-sum topology, which induces the direct-sum σ-algebra $\bigcup_{n=0}^{\infty} \mathcal{B}_n$, where each $\mathcal{B}_n$ is σ-algebra of Borel sets in E^n, which allows the direct-sum measure $\sum_{n=0}^{\infty} \mu^n$ on it.

Then, we can define an RFSeq as a random element X on the measurable space $(\bigcup_{n=0}^{\infty} E^n, \bigcup_{n=0}^{\infty} \mathcal{B}_n)$. Although, in general, we may not know a priori how many targets exist or we have to track, the number n of all the targets (at least potentially to be detected) is always *finite*, but often with no known a priori upper limit. For each $n = 0, 1, 2, ...,$ let p_n be the probability of the number[14] of targets being n, and given any n, let the joint probability distribution of the n-tuple of target states, $(x_i)_{i=1}^{n} \in E^n$, be $F^{(n)}$, called the *nth-order probability distribution* (n-PDist), so that we can model targets, as a whole, by an RFSeq X, with[15] $F^{(n)}(B) = \text{Prob}\{X \in B | \ell(X) = n\}$ for each n and for each $B \in \mathcal{B}_n$.

Our assumption that the targets are without a priori identification is translated into the assumption that, for each n, $F^{(n)}$ is *permutable*,[16] in the sense $F^{(n)}(B) = F^{(n)}(\pi_a^{(n)}(B))$ for every $B \in \mathcal{B}_n$ and every $a \in A_n$, where A_n is the set of all the permutations on $\{1, ..., n\}$, and $\pi_a^{(n)}((x_i)_{i=1}^{n}) \stackrel{\text{def}}{=} (x_{a(i)})_{i=1}^{n}$ for any $(x_i)_{i=1}^{n} \in E^n$ and any $a \in A_n$. If each permutable probability measure $F^{(n)}$ is absolutely continuous with respect to the product measure μ^n, its Radon–Nikodym derivative $f^{(n)}$, called the *nth-order probability density* (n-PD), is also permutable, in the sense that $f^{(n)}(\pi_a^{(n)}(x)) = f^{(n)}(x)$ for all $x \in E^n$, for any $a \in A_n$.

B. FPP Formalism

In [15, Ch. 5, p. 111], an RFSeq $(x_i)_{i=1}^{n} \in E^n$ with $(p_n, F^{(n)})_{n=0}^{\infty}$ is called an FPP if each n-PDist $F^{(n)}$ is permutable, and is characterized by a sequence $(\mathcal{J}^{(n)})_{n=0}^{\infty}$ of measures, each of which, $\mathcal{J}^{(n)}$, is a finite measure on $(E^n, \mathcal{B}_n)$, defined by $\mathcal{J}^{(n)}(B) = n! p_n F^{(n)}(B)$ for each $B \in \mathcal{B}_n$, called the *nth-order Janossy measure* (n-JM).[17] If n-JM[18] $\mathcal{J}^{(n)}$ is absolutely continuous with respect to the product measure μ^n, its Radon–Nikodym derivative $J^{(n)}$ is called the *nth-order Janossy density* (n-JD), which we can write as $J^{(n)}(x) = n! p_n f^{(n)}(x)$ for every $x \in E^n$, with n-PD $f^{(n)}$ of each n-PDist $F^{(n)}$. Obviously, every n-JM $\mathcal{J}^{(n)}$ is permutable, and so is any n-JD $J^{(n)}$ if it exists.

In this paper, as well as in [9] and [16], however, we present an alternative but equivalent FPP formalism: For each n and each $x \in E^n$, let the equivalence class in E^n, obtained by ignoring the ordering of $x = (x_i)_{i=1}^{n}$, be $[x]$, i.e., $[x] \stackrel{\text{def}}{=} \{\pi_a^{(n)}(x) | a \in A_n\}$ with $\pi_a^{(n)}$ and A_n, as defined earlier. For each n, let us *symbolically* denote $E^n/n! = \{[x] | x \in E^n\}$, using "$n!$" only as a symbol in place of equivalence classes "$[\cdot]$ " or relation "$\sim$."

[11] We assume $E \neq \emptyset$ so that $E^n \neq \emptyset$ for any n, yet we have $E^n \cap E^{n'} = \emptyset$ for any $n \neq n'$.

[12] $\theta \notin E$ ($E^0 = \{\theta\}$) is used as a special symbol (for the empty sequence) throughout this paper. It is also considered as a function whose domain, image (range or codomain), and graph are all the empty set.

[13] A semigroup is a nonempty set with an associative binary operator. A monoid is a semigroup with an identity (unit) element.

[14] We generally assume that the number n of targets is constant, for the reasons explained by Remark 5 in Section III-A.

[15] For any $x \in \bigcup_{n=0}^{\infty} E^n$, by $\ell(x)$ we mean the length of finite sequence x in E, i.e., $\ell(x) = n \Leftrightarrow x \in E^n$, and $\ell(\theta) = 0$.

[16] Synonymous to *symmetric* (*permutation-symmetric*), *exchangeable*, *interchangeable*, etc. See Remark 3.

[17] According to [15], the term *Janossy measure* originated from [17] that references [18]. It is indicated [15, p. 124] that the constant $n!$ in its definition, as it distinctly appears in (1)–(4) also, is included to be *advantageous in simplifying combinatorial formulae*, so that, in a sense, this constant $n!$ uniquely identifies the n-JM, the n-JD, and the JMD (introduced later), distinguishing themselves from other concepts.

[18] In this paper, we use superscripts (n) for n-PDist $F^{(n)}$, n-PD $f^{(n)}$, n-JM $\mathcal{J}^{(n)}$, and n-JD $J^{(n)}$, to signify the fact that they are applied to the nth-order product space E^n, although, customarily, subscripts are used instead as in F_n, f_n, $\mathcal{J}_n$, and J_n.

$E^n/n!$ is a quotient space induced by the quotient map[19] $\varphi_n : x \mapsto [x]$ for each n. We may call $[(x_i)_{i=1}^n]$ an *unordered n-tuple*, while $(x_i)_{i=1}^n$ is an *ordered n-tuple*.

Algebraically, we *may* call $\bigcup_{n=0}^\infty E^n/n!$ the *free commutative monoid* (FCM)[20] generated by E with the commutative operation $*$ defined by $[x] * [x'] = [x * x']$ for every $(x, x') \in E^n \times E^{n'}$, and the identity element $[\theta] = \{\theta\}$. Topologically, each $E^n/n!$ is a quotient topological space. Since each coordinate permutation $\pi_a^{(n)}$ is a homeomorphism (and hence a continuous open map), every open set in $E^n/n!$ can be written as the image $\varphi_n(B)$ of an open set B in E^n, and hence each $E^n/n!$ is LCHC2, and so is their direct sum $\bigcup_{n=0}^\infty E^n/n!$, with the σ-algebra $\bigcup_{n=0}^\infty \mathcal{B}_n/n!$ of Borel sets in it. We may consider $\bigcup_{n=0}^\infty E^n/n!$ as the quotient space through the map $\varphi : \bigcup_{n=0}^\infty E^n \to \bigcup_{n=0}^\infty E^n/n!$ defined by $\varphi(x) = \varphi_n(x)$ for each $x \in E^n$ $(n > 0)$, and $\varphi(\theta) = \varphi_0(\theta) = [\theta] = \{\theta\}$.

Finally, an FPP can be defined as a random element X on a measurable space $(\bigcup_{n=0}^\infty E^n/n!, \bigcup_{n=0}^\infty \mathcal{B}_n/n!)$, with PDist Φ such that

$$\begin{aligned}\Phi\left(\bigcup_{n=0}^{\infty} \varphi(B_n)\right) &= \mathrm{Prob}\left\{X \in \bigcup_{n=0}^{\infty} \varphi(B_n)\right\} \\ &= \sum_{n=0}^{\infty} p_n F^{(n)}\left(\varphi_n^{-1}\left(\varphi_n(B_n)\right)\right) \\ &= \sum_{n=0}^{\infty} \tfrac{1}{n!} \mathcal{J}^{(n)}\left(\varphi_n^{-1}\left(\varphi_n(B_n)\right)\right)\end{aligned} \tag{1}$$

for any $(B_n)_{n=0}^\infty \in \prod_{n=0}^\infty \mathcal{B}_n$. Since each coordinate permutation $\pi_a^{(n)}$ is measurable, every measurable set $\boldsymbol{B} \in \bigcup_{n=0}^\infty \mathcal{B}_n/n!$ in $\bigcup_{n=0}^\infty E^n/n!$ can be expressed as $\bigcup_{n=0}^\infty \varphi(B_n)$ with some $(B_n)_{n=0}^\infty \in \prod_{n=0}^\infty \mathcal{B}_n$. The first equality of (1), therefore, simply states the definition of PDist Φ of a random element X, i.e., $\Phi(\boldsymbol{B}) = \mathrm{Prob}\{X \in \boldsymbol{B}\}$ for any $\boldsymbol{B} \in \bigcup_{n=0}^\infty \mathcal{B}_n/n!$. The second equality of (1) means that, given PDist Φ of FPP X, there exists a series $(p_n, F^{(n)})_{n=0}^\infty$ of probabilities and permutable n-PDists such that[21] $p_n = \Phi(E^n/n!) = \mathrm{Prob}\{\ell(X) = n\}$ for every n, and $p_n F^{(n)}(\varphi_n^{-1}(\boldsymbol{B}_n)) = \Phi(\boldsymbol{B}_n)$ for every $\boldsymbol{B}_n \in \mathcal{B}_n/n!$. It also implies that PDist Φ is uniquely defined by a PDist $(p_n)_{n=0}^\infty$ and any permutable n-PDists $F^{(n)}$ for each n. The third equality is nothing but the definition of each n-JM. For each n and each measurable set $\boldsymbol{B}_n = \varphi_n(B_n)$ in $E^n/n!$, event $\{X \in \boldsymbol{B}_n\}$ can be viewed as the event $\bigcup_{a \in A_n} \{x \in \varphi_n^{-1}(\boldsymbol{B}_n)\}$, where x is an arbitrary enumeration of X. We should note $\varphi_n^{-1}(\varphi_n(B_n)) = \bigcup_{a \in A_n} \pi_a^{(n)}(B_n)$, which is the set of all the enumerations of equivalence classes in $\boldsymbol{B}_n = \varphi_n(B_n)$.

Using the collection $(E^n, \mathcal{B}_n, \mu^n)_{n=0}^\infty$ of the measure spaces, we can define a positive linear functional $\mathcal{L}$, defined on a set of bounded measurable functionals ψ, and a measure $\mathcal{M}$, both on measurable set $(\bigcup_{n=0}^\infty E^n/n!, \bigcup_{n=0}^\infty \mathcal{B}_n/n!)$, such that we have[22]

$$\begin{aligned}\mathcal{L}(\psi) &= \int_{\bigcup_{n=0}^\infty E^n/n!} \psi(X)\, \mathcal{M}(dX) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{E^n} \psi(\varphi(x))\, \mu^n(dx),\end{aligned} \tag{2}$$

where $\mathcal{M}(\boldsymbol{B}_n) = \mu^n(\varphi_n^{-1}(\boldsymbol{B}_n))/n!$ for every $\boldsymbol{B}_n \in \mathcal{B}_n/n!$, for each n. We may call the measure space $(\bigcup_{n=0}^\infty E^n/n!, \bigcup_{n=0}^\infty \mathcal{B}_n/n!, \mathcal{M})$, derived from the state measure space $(E, \mathcal{B}, \mu)$ in this way, the *quotient measure space* (QMS).

It follows from (1) and (2) that, if each n-JM $\mathcal{J}^{(n)}$ of (1) has n-JD $J^{(n)}$, then PDist Φ has the density ϕ, i.e., the Radon–Nikodym derivative of Φ, with respect to the measure $\mathcal{M}$, which we call the *Janossy–Mahler density* (JMD), defined as $\phi(\varphi_n(x)) = J_n(x)$ for every $x \in E^n$, for each n, such that

$$\begin{aligned}\Phi\left(\varphi\left(\bigcup_{n=0}^\infty B_n\right)\right) &= \int_{\varphi(\bigcup_{n=0}^\infty B_n)} \phi(X) \mathcal{M}(dX) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\varphi_n^{-1}(\varphi_n(B_n))} \phi(\varphi_n(x)) \mu^n(dx)\end{aligned} \tag{3}$$

for any $(B_n)_{n=0}^\infty \in \prod_{n=0}^\infty \mathcal{B}_n$.

Remark 3 (FM, FCM, FPP, and JMD): To call a random element on FM[23] an RFSeq, and a random element on FCM an FPP, is our own "invention," which had most probably not seen before our preliminary paper [9] (or its predecessor [16]) was published. The introduction of FM and FCM (or its variation), however, appeared in [15] and [23]. In [15, p. 129], FM $\bigcup_{n=0}^\infty E^n$ is called the *canonical probability space*, and an FPP as a random element in the quotient space $\bigcup_{n=0}^\infty E^n/n!$ is also suggested. In [23], FM $\bigcup_{n=0}^\infty E^n$ is called the *population state space*, and a version of FCM, $\bigcup_{n=0}^\infty E^n/n!$, the *symmetric population state space*. Furthermore, in [23], what we call an FPP was called a *symmetric point process*, while what we call an RFSeq was called simply a *point process*,[24] reflecting the distinction caused by the "commutativity" or "permutability,"[25] or lack of it.

Both in [15, Ch. 5, p. 111] and [23], we may say that an FPP is defined through a series of n-PDist or n-JM,

[19] Also known as identification map, natural map, canonical surjection map, canonical projection map, etc. In topological algebra [19], $E^n/n!$ is called the nth-order symmetric product of E, $\mathrm{SP}^n(E)$.

[20] Cf. [47, Sec. I.6.3, p. 16], for definition of *free commutative semigroup*, which becomes FCM when given an identity element.

[21] For any $[x] \in \bigcup_{n=0}^\infty E^n/n!$, the length $\ell(x)$ of any element in the equivalent class $[x]$ is the same, so that we let $\ell([x]) = \ell(x)$. We have $\Phi(\bigcup_{n=0}^\infty E^n/n!) = \sum_{n=0}^\infty \Phi(E^n/n!) = \sum_{n=0}^\infty p_n = 1$.

[22] We use the convention that $\int_{E^0} q(\xi)\mu^0(\xi) = q(\theta)\mu^0(E^0) = q(\theta)$ for any functional q on $(E^0, \mathcal{B}_0) = (E^0, \{\emptyset, E^0\}) = (\{\theta\}, \{\emptyset, \{\theta\}\})$.

[23] By replacing the target indices by the discrete time indices, an FM can be used as a mathematical model for a discrete-time dynamical process with a variable *end-of-the-process time*, as shown in [21].

[24] According to [23], the term "point process" is attributed to [24].

[25] In [23], the permutability is treated as synonymous to "indistinguishability," which, in our opinion, is misleading to a degree, because, for example, two targets, as realizations of two random points, which do not share the same state, can always be "distinguished," even when the distributions are "identical" and "independent." We would prefer that the distinction is considered as "targets with and without *a priori identifications*," rather than "distinguishability" and "indistinguishability."

rather than a random element itself. A traditional definition of an FPP is, however, as a *random counting measure*[26] [22, Def. 1.1, p. 4], which can represent possibly countably many points. In our MTT applications, however, we do not need to consider any set of countably many points, and therefore, we may say that our definition of FPP, without ever considering a random measure, is justified.[27] A counting measure representation $\mathcal{N}$ of an FPP $X = [(x_i)_{i=1}^n]$, as a random measure on $(E, \mathcal{B})$, can be defined as $\mathcal{N}(B) = \sum_{i=1}^n \mathbb{I}(x_i; B)$ for each $B \in \mathcal{B}$, with an arbitrary enumeration $(x_i)_{i=1}^n$ of FPP X, where $\mathbb{I}$ is the generic indicator function defined as $\mathbb{I}(\xi; A) = 1$ if $\xi \in A$ and zero otherwise for any set A.

We call the probability density ϕ that appears in (3) in our FPP formalism, as well as in the RFSet formalism described later in this section, the JMD, because of 1) its obvious relation to the Janossy densities n-JDs, $(J^{(n)})_{n=0}^\infty$, through $\phi(\varphi(x)) = J^{(n)}(x)$ for any $x \in E^n$, and 2) our understanding that the PD ϕ was first introduced by Dr. R. P. S. Mahler as a single function, as opposed to a series $(J^{(n)})_{n=0}^\infty$ of functions, in his *finite set statistics* (FISST) formalism [25]–[27]. The JMD ϕ is called the *multiobject density function* in [25, Sec. 11.3.3, p. 360] and [26, Sec. 3.2.4, p. 62], and the *global probability density function* in [27, Sec. 4.3.3, p. 162].

C. RFSet Formalism

For each $n > 0$, let $\mathcal{F}_n(E) = \{X \subseteq E | 0 < \#(X) \leq n\}$ and $\tilde{\mathcal{F}}_n(E) = \{X \subseteq E | \#(X) = n\}$. Then, with $\mathcal{F}_0 = \tilde{\mathcal{F}}_0 = \{\emptyset\}$, $\mathcal{F}(E) = \bigcup_{n=0}^\infty \mathcal{F}_n(E) = \bigcup_{n=0}^\infty \tilde{\mathcal{F}}_n(E)$ is the collection of all the finite sets in the state space E. Algebraically, we *may* call $\mathcal{F}(E)$ the *free idempotent commutative monoid* (FICM) with the set-theoretic union[28] as the binary operator on it. For each $n > 0$, redefine the quotient map φ_n as $\varphi_n : E^n \to \mathcal{F}_n(E)$ with $\varphi_n((x_i)_{i=1}^n) = \{x_i\}_{i=1}^n$. It makes $\mathcal{F}_n(E)$ a quotient topological space that is an LCHC2 with its open sets as the collections of the images $\varphi_n(B)$ of all the open sets B in E^n. $\mathcal{F}(E)$ is also LCHC2 as the quotient space induced by the redefined map $\varphi : \bigcup_{n=0}^\infty E^n \to \mathcal{F}(E)$ with $\varphi(x) = \varphi_n(x)$ for all $x \in E^n$ and $\varphi(\theta) = \varphi_0(\theta) = \emptyset$.

An RFSet X can then be defined as a random element on measurable set $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}))$, where $\mathbb{B}(\mathcal{B})$ is the σ-algebra of Borel sets in quotient topological space $\mathcal{F}(E)$. As the PDist Φ of RFSet X, (1) holds for any $(B_n)_{n=0}^\infty \in \prod_{n=0}^\infty \mathcal{B}_n$, with the n-PDist $F^{(n)}$ (and n-JM $\mathcal{J}^{(n)} = n! p_n F^{(n)}$) and the redefined quotient map φ. Exactly in parallel to FPP formalism, through the redefined quotient map φ, we can redefine $\mathcal{L}$ as the positive linear bounded functional on the space of bounded measurable functionals ψ on the measurable space $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}))$, and the measure $\mathcal{M}$ on the measurable space $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}))$, as (2), and the JMD ϕ with respect to the redefined measure space $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}), \mathcal{M})$, as (3).

The *idempotency* of the FICM $\mathcal{F}(E)$, however, poses some peculiar problems: For example, a multidimensional point, $(x_i)_{i=1}^n$ in E^n with $n > 1$, is mapped into a single point in $\mathcal{F}(E)$, when its elements are all identical, i.e., $x_1 = \cdots = x_n$. One way to avoid this peculiarity is to "ignore" such coincidences. Namely, we may assume that the set $D_n \stackrel{\text{def}}{=} \{(x_i)_{i=1}^n \in E^n | x_i = x_{i'} \text{ for some } i \neq i'\}$ of n-tuples with any repeated elements, which we call the *diagonal set in* E^n, has the zero product measure, i.e., $\mu^n(D_n) = 0$. We can then define the JMD ϕ in RFSet formalism by $\phi(\varphi((x_i)_{i=1}^n)) = \phi(\{x_i\}_{i=1}^n) = J^{(n)}((x_i)_{i=1}^n)$ for every $(x_i)_{i=1}^n \in E^n$, for each n.

However, in case where the state space E is countable with discrete topology[29] and counting measure μ, $\mu^2(D_2) = 0$ implies $\mu(E) = 0$, which is obviously not desirable. To remedy the situation, we need to modify (2) and (3) slightly as

$$\begin{cases} \mathcal{L}(\psi) = \int\limits_{\mathcal{F}(E)} \psi(X) \mathcal{M}(dX) \\ \qquad = \sum\limits_{n=0}^{\infty} \frac{1}{n!} \int\limits_{E^n} \psi(\varphi(x)) \tilde{\mu}_n(dx) \\ \Phi\left(\varphi\left(\bigcup_{n=0}^\infty B_n\right)\right) = \int\limits_{\varphi(\bigcup_{n=0}^\infty B_n)} \phi(X) \mathcal{M}(dX) \\ \qquad = \sum\limits_{n=0}^{\infty} \frac{1}{n!} \int\limits_{\varphi^{-1}(\varphi(B_n))} \phi(\varphi(x)) \tilde{\mu}_n(dx) \end{cases} \tag{4}$$

using the modified measure,[30] $\tilde{\mu}_n(B) = \mu^n(B \backslash D_n)$ for every $B \in \mathcal{B}_n$, for each n, to make $\tilde{\mu}_n(D_n) = 0$ without affecting any product measure μ^n, with $\mathcal{M}(\varphi(B)) = \tilde{\mu}_n(\varphi^{-1}(\varphi(B)))/n!$ for any $B \in \mathcal{B}_n$, and with the JMD in RFSet formalism by $\phi(\varphi((x_i)_{i=1}^n)) = \phi(\{x_i\}_{i=1}^n) = J^{(n)}((x_i)_{i=1}^n)$ for every $(x_i)_{i=1}^n \in E^n$. With this modification, each component $\tilde{\mathcal{F}}_n$ of FICM $\mathcal{F}(E)$ as the direct sum $\mathcal{F}(E) = \bigcup_{n=0}^\infty \tilde{\mathcal{F}}_n(E)$ becomes the image of the quotient map φ_n in the $\tilde{\mu}_n$-a.e. sense. In the rest of this paper, whenever the RFSet formalism is used, we assume we are using the modified measures $\tilde{\mu}_n$'s, as in (4).

When E is countable with the discrete topology and the counting measure μ, the measure $\mathcal{M}$ defined in (4), using $\tilde{\mu}_n$'s, becomes a counting measure on $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}))$, where $\mathbb{B}(\mathcal{B})$ becomes the power set of a

[26] A counting measure μ on any measurable space $(E, \mathcal{B})$ is an integer-valued functional defined by $\mu(B) = \#(B)$ for each $B \in \mathcal{B}$. By $\#(A)$, we mean the cardinality of (the number of elements in) any set A, throughout this paper.

[27] In [15, p. 131], it is stated: "The main difficulty with this (Moyal's) approach from our point of view is that it does not extend readily to random measures, which require for their own sake and for applications in later chapter."

[28] The union operator $\cup$ on $\mathcal{F}(E)$ is associative and commutative with the empty set as the unit element. $\mathcal{F}(E)$ is also idempotent, i.e., every $X \in \mathcal{F}(E)$ is an idempotent, because $X * X = X \cup X = X$. Cf. [47, Sec. I.6.3, p. 16], for definition of *free idempotent commutative semigroup*, which becomes FICM when given an identity element.

[29] Namely, every subset B of E is an open set, and hence, σ-algebra of Borel sets $\mathcal{B}$ is the power set of E.

[30] By "\" we mean the set-theoretic subtraction operator, i.e., $A \backslash B = \{a \in A | a \notin B\}$.

countable set[31] $\mathcal{F}(E)$, so that the JMD ϕ, which is the Radon–Nikodym derivative of the probability distribution Φ with respect the counting measure $\mathcal{M}$, becomes the probability mass function (PMF). In such a case, for example, when we define data association, as an RFSet in Section III-C , we will use the generic symbol[32] "P" (instead of ϕ) for such JMD that is nothing but a PMF.

We call the measure space $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}), \mathcal{M})$ derived from the state measure space $(E, \mathcal{B}, \mu)$ for RFSet formalism the QMS. We also call $(\bigcup_{n=0}^{\infty} E^n/n!, \bigcup_{n=0}^{\infty} \mathcal{B}_n/n!, \mathcal{M})$ for FPP formalism QMS. If disambiguation is necessary, we will use FCM-QMS or FICM-QMS.

Remark 4 (QMS): In Section II-B and II-C , we defined the FCM-QMS and FICM-QMS for FPP and RFSet formalisms, respectively, through the quotient map φ, with which we defined the quotient topology and the *quotient measure* $\mathcal{M}$. An alternative, but an equivalent, way to construct these measure spaces may be possible directly from the linear functionals $\mathcal{L}$, defined in (2) or (4), first applied to an appropriate small class of functionals ψ, and then appropriately extended to construct measurable sets and measures, as shown in [28, Ch. 16, p. 419] and [29]. In [25]–[27], what we have defined as RFSet formalism in this paper is called FISST formalism, in which the integral in (4) is called the *set integral*, as its core concept, as we understand. We may interpret the FISST formalism as the one in which the set integral plays this role to construct the appropriate measure space $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}), \mathcal{M})$.

As mentioned in [25, Appendix F, p. 711], FICM $\mathcal{F}(E)$ can be topologized by the relative (subspace) topology as the subset of the space $\mathcal{C}(E)$ of the closed sets in E, with *Fell–Matheron topology*[33] [40, p. 3; 45, p. 398]. Since the quotient map φ is continuous in this topology [20, Prop. 2.4, p. 156], the quotient topology (with which we have introduced RFSet formalism) is stronger[34] than the Fell–Matheron topology. The FISST formalism established in [25]–[27] motivated our definition of an FPP as a random element taking values in FCM $\bigcup_{n=0}^{\infty} E^n/n!$.

In summary, among the three formalisms, the equivalence between RFSeq and FPP is rather obvious. Instead of calling an RFSeq $(x_i)_{i=1}^n$ with permutable n-PDists an FPP, we call a random element on FCM $\bigcup_{n=0}^{\infty} E^n/n!$ an FPP, forcing the permutability on the state space algebraic structure rather than on the n-PDists. By doing so, we put an FPP and an RFSet into almost equivalence, with the same JMD concept. The difference between an FPP and an RFSet is, however, that the former allows *repeated elements*, while the latter does not. An FPP that does not allow any repeated elements is called a *simple* FPP, and it is shown in [15, Prop. 5.4.V, p. 138], the necessary and the sufficient condition for the "simpleness" is $\mathcal{J}^{(n)}(D_n) = 0$; i.e., the n-JM of the diagonal set D_n is zero. In this sense, we may say an RFSet is just a simple FPP.

We should note that the target state space E may be a finite set itself, e.g., when the original state space is approximated by a set of small rectangular cells, as in the target model used in [30]. In that case, it would be *unreasonable* to prohibit any two targets from occupying a single state, so that RFSet formalism becomes inadequate,[35] while FPP formalism may become a perfect alternative. This *idempotency* peculiarity becomes apparent also when we consider the union of two independent RFSets, as we see below.

D. Concatenation, Union, Superposition, and Convolution

As a foundation for MHT, the binary operation on FM, FCM, or FICM, i.e., concatenation or unionization, of random elements, plays crucial roles. Let X_1 and X_2 be two independent RFSets, i.e., two independent random elements in $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}), \mathcal{M})$, with JMDs ϕ_1 and ϕ_2, respectively. Then, as described in [25, Sec. 11.5.3, p. 385], JMD ϕ of the union $X = X_1 \cup X_2$ can be written as

$$\phi(X) = \sum_{X_1 \subseteq X} \phi_1(X_1)\phi_2(X \backslash X_1), \tag{5}$$

which holds true only when each product measure $\tilde{\mu}_{n_1} \times \tilde{\mu}_{n_2}$ of the modified measures satisfies[36] $(\tilde{\mu}_{n_1} \times \tilde{\mu}_{n_2})(D_{n_1+n_2}) = 0$. This condition, guaranteeing that $X_1 \cap X_2 = \emptyset$ with probability 1, is satisfied if $\mu^n(D_n) = 0$, e.g., when target state space E has a *continuous* component such as a Euclidean component.

In FPP formalism, which lacks the idempotency, for any two independent FPPs, $[(x_{1i})_{i=1}^{n_1}]$ and $[(x_{2i})_{i=1}^{n_2}]$, with JMDs ϕ_1 and ϕ_2, respectively, the JMD ϕ of the concatenation $[(x_{1i})_{i=1}^{n_1}] * [(x_{2i})_{i=1}^{n_2}] = [(x_{1i})_{i=1}^{n_1} * (x_{2i})_{i=1}^{n_2}]$ can always be written as

$$\phi([(x_i)_{i=1}^n]) = \sum_{I \subseteq \{1,\dots,n\}} \phi_1([(x_i)_{i \in I}])\phi_2([(x_i)_{i \in \{1,\dots,n\} \backslash I}]), \tag{6}$$

which is translated into the case where the n-JD of the two RFSeqs with permutable n-PDists, with n-JDs,

[31] The countability of E implies the countability of $\mathcal{F}(E)$ under the axiom of countable choice.

[32] As a general "rule," we use the symbols, F and f, for PDist and PD in E ($F^{(n)}$ and $f^{(n)}$ for E^n), P and p for the probability or the PMF for the discrete, or density function of random elements of mixed nature, and ϕ for the JMD for FPP or RFSet formalism.

[33] Also known as hit-or-miss topology. With this topology, $\mathcal{C}(E)$ is a compact Hausdorff space satisfying the second axiom of countability [40, Th. 1-2-1, p. 3], and $\mathcal{F}(E)$ is dense in $\mathcal{C}(E)$ [40, Cor. 2, p. 7].

[34] Hence, our introductions of the linear functional $\mathcal{L}$ and the measure $\mathcal{M}$ are consistent with the Fell–Matheron topology.

[35] In [25, Appendix E, p. 705], it is indicated that the *idempotency* issues should be resolved by the concept of *multisets*.

[36] In Section II-C, we defined $\tilde{\mu}_n(B) = \mu^n(B \backslash D_n)$, which does not necessarily imply $(\tilde{\mu}_{n_1} \times \tilde{\mu}_{n_2})(D_{n_1+n_2}) = 0$, because $\tilde{\mu}_{n_1+n_2} = \tilde{\mu}_{n_1} \times \tilde{\mu}_{n_2}$ does not hold necessarily.

$(J_1^{(n_1)})_{n_1=0}^{\infty}$ and $(J_2^{(n_2)})_{n_2=0}^{\infty}$, is $(J^{(n)})_{n=0}^{\infty}$, as

$$J^{(n)}((x_i)_{i=1}^n) = \sum_{I\subseteq\{1,\ldots,n\}} J_1^{(\#(I))}((x_i)_{i\in I}) J_2^{(n-\#(I))}((x_i)_{i\in\{1,\ldots,n\}\setminus I}). \tag{7}$$

For the rest of this paper, we will denote the right-hand side of (5) for RFSet formalism, or of (6) for FPP formalism, by the *convolution* $\phi_1 \otimes \phi_2$ of two JMDs ϕ_1 and ϕ_2. This convolution can be extended to N independent RFSets or FPPs as $\phi_1 \otimes \cdots \otimes \phi_N$, in an obvious way. When two independent FPPs are represented by random measures $\mathcal{N}_1$ and $\mathcal{N}_2$, the random measure representation $\mathcal{N}$ of the concatenation (the union) is simply the sum $\mathcal{N} = \mathcal{N}_1 + \mathcal{N}_2$, which is called the *superposition* of two FPPs with $\mathcal{N}_1$ and $\mathcal{N}_2$ [42, Ch. 5.1, p. 152].

III. TARGET AND SENSOR MODELS, AND DATA ASSOCIATION HYPOTHESES

Whenever the word MHT is mentioned in the context of MTT, we understand that by a "hypothesis" we mean a *data association hypothesis*, which is the core concept of the MHT. The ultimate goal of any MHT algorithm is to estimate the target states, in either one of the three formalisms described in the previous section, while generation, evaluation, and maintenance of association hypotheses may constitute various intermediate algorithmic steps. In some applications, however, determination of correlation or relations among the data in terms of their origins, retroactively in many cases, is of primary interests and importance. In this section, after describing a general class of target and sensor models for the rest of this paper, association hypotheses will be defined as possible realizations of an RFSet, called *data association*, or simply *association*, in a discrete space.

A. Target Model: Set of Stochastic Processes

Unlike almost all the target models in RFSet (FISST) formalism [25]–[27] where a set of an unknown number of targets is modeled as a stochastic process on FICM, i.e., the collection $\mathcal{F}(E)$ of finite sets in a given target state space E, our target model assumes that

(A1) [Target Model]: The set of targets, as a whole, is modeled as 1) an RFSeq $((x_i(t))_{t\in[t_0,\infty)})_{i=1}^n$, 2) an FPP $[((x_i(t))_{t\in[t_0,\infty)})_{i=1}^n]$, or 3) an RFSet $\{(x_i(t))_{t\in[t_0,\infty)}\}_{i=1}^n$, of stochastic processes on space E, over a continuous time interval $[t_0, \infty)$, with the probability $p_n = P(n)$ of the number of targets being n with a finite mean, so that, for each n, for any N-tuple, $(s_\kappa)_{\kappa=1}^N \in [t_0, \infty)^N$, of distinct times, 1) RFSeq $((x_i(s_\kappa))_{\kappa=1}^N)_{i=1}^n$ has permutable n-PDist $F^{(n)}(\cdot; (s_\kappa)_{\kappa=1}^N)$ and n-JM $\mathcal{J}^{(n)}(\cdot; (s_\kappa)_{\kappa=1}^N)$ on $(E^{Nn}, \mathcal{B}_{Nn}, \mu^{Nn})$ with n-PD $f^{(n)}(\cdot; (s_\kappa)_{\kappa=1}^N)$ and n-JD $J^{(n)}(\cdot; (s_\kappa)_{\kappa=1}^N)$, 2) FPP $[((x_i(s_\kappa))_{\kappa=1}^N)_{i=1}^n]$ has PDist[37] $\Phi(\cdot; (s_\kappa)_{\kappa=1}^N)$ on $(\bigcup_{n=0}^\infty E^{Nn}/n!, \bigcup_{n=0}^\infty \mathcal{B}_{Nn}/n!, \mathcal{M}_N)$ with JMD $\phi(\cdot; (s_\kappa)_{\kappa=1}^N)$, or 3) RFSet $\{(x_i(s_\kappa))_{\kappa=1}^N\}_{i=1}^n$ has PDist $\Phi(\cdot; (s_\kappa)_{\kappa=1}^N)$ on $(\mathcal{F}(E^N), \mathbb{B}(\mathcal{B}_N), \mathcal{M}_N)$ with JMD $\phi(\cdot; (s_\kappa)_{\kappa=1}^N)$, where $\mathcal{M}_N$ is the measure defined by (2) or (4) (for FPP or RFSet formalism) with E replaced by E^N.

Remark 5 (Birth–Death Target Models): There are two significant departures of our target model from the commonly used target models: 1) Targets are modeled as an RFSeq, an FPP, or an RFSet of stochastic processes, each on the target state space $(E, \mathcal{B}, \mu)$, over a continuous time interval $[t_0, \infty)$, rather than a single stochastic process on FM $\bigcup_{n=0}^\infty E^n$ or FCM $\bigcup_{n=0}^\infty E^n/n!$ or FICM $\mathcal{F}(E)$, and 2) the (generally unknown, and hence random) number n of targets is *constant* over the entire time interval $[t_0, \infty)$. We contend that condition 1 is necessary to define data association hypotheses as hypotheses of the *true* association sharing the same origins, to avoid any possibility of target *identities* from *ever* being *switched* by the symmetrization as a consequence of using the FPP or the RFSet formalisms. We believe that condition 2 can be defended from a "first principle" point of view, as we argue in the following.

A *real* birth or death of any target occurs only under very limited circumstances, most probably in battlefield type of environments, where, e.g., missiles are launched or vehicles are destroyed. Even missiles before being launched, however, may exist as ground targets. A destroyed vehicle may be still called a damaged vehicle, or a wreckage, with its existence intact, even if it is dead. In many cases, an emergence of a persistent track is at least partly a result of sensor management, and should not be confused with a target birth,[38] which should be a part of a purely target behavioral model, independent of any sensor.

In most realistic situations, what we call new targets (or newly born targets) are actually those that had remained undetected (and hence existed) but were detected for the first time by a sensor that is capable of detecting them. For the implementation of MHT algorithms, therefore, the issue becomes how to calculate the track initiation likelihood, or the newly detected target likelihood, as discussed in [35] as implementation of track-oriented MHT, or in [36] as a part of overall algorithm complexity as discussed in connection with the RFSet-based algorithms. When hypotheses are evaluated recursively, as discussed in Section III-C , the issue is how to calculate the density of newly detected targets

[37] We are using the same notations for the PDist Φ, and the JMD ϕ for both the FPP and RFSet formalisms. Distinction should be clear from the context. In FPP or RFSet formalism, Assumption A1 implies $\Phi(E^{Nn}/n!; (s_\kappa)_{\kappa=1}^N) = p_n$ or $\Phi(\mathcal{F}_n(E^N); (s_\kappa)_{\kappa=1}^N) = p_n$.

[38] In our opinion, a typical example of this type of confusion can be seen in an assertion made in [2, p. 327]: "A true target is most generally defined to be an object that will persist in the tracking volume for at least several scans." Although "many" true targets may be persistent in any tracking volume (if it is well defined and well designed), we may not know generally, a priori, any "true" target would appear at any time in any portion of any tracking volume, depending on particular *sensor management strategies*.

in the sensor measurement space (as discussed in [41]). In our opinion, the target birth–death model has more often been used for the convenience of the algorithms than for faithfully modeling the targets' behaviors and the sensor detection capabilities.

Moreover, the constant number n can be justified even when the target birth–death is indeed supported by some legitimate reality, by counting all the targets that ever exist in a given time interval, e.g., $[t_0, \infty)$, and by including the augmented discrete states, such as {unborn, alive, dead}, at any given time, with an appropriate target dynamics, within the framework of *multiple models* [46].

B. Sensor Model: Random Assignments and False Alarms

Our sensor model defines available information in the form of a sequence y_k, $k = 1, 2, \ldots$, of *measurement frames*,[39] each of which, $y_k = (y_{kj})_{j=1}^{m_k}$, is an RFSeq, in an appropriate measurement space[40] E_{Mk} with an appropriate measure to let us properly define the m_k-PD, composed of m_k measurements, y_{kj}'s, collected at the same time t_k ($t_0 \leq t_1 \leq t_2 \leq \cdots$) by the same sensor.[41]

The uncertainty of the origin of each measurement y_{kj} is modeled by an unobservable RFSet a_k of pairs of integers in $\{1, \ldots, n\} \times \{1, \ldots, m_k\}$, given the number n of the targets and the number m_k of the measurements in frame k, called the *target-to-measurement assignment* or simply *target assignment* at frame k. Let the domain and the image (range) of a_k be denoted by $\mathrm{Dom}(a_k) = \{i | (i, j) \in a_k \text{ for some } j\}$, and $\mathrm{Im}(a_k) = \{j | (i, j) \in a_k \text{ for some } i\}$. Then, 1) $i \in \mathrm{Dom}(a_k)$ means the ith target is *detected* at frame k, 2) $(i, j) \in a_k$ means[42] the jth measurement of frame k *originates from* the ith target, and 3) $j \notin \mathrm{Im}(a_k)$ means the jth measurement of frame k is a *false alarm* (that does not originate from any target).

Throughout the rest of this paper, we maintain the following two assumptions for each frame k:

(A2) [No Merged or Split Measurement]: There is no *merged measurement*, i.e., $\#(\{i | (i, j) \in a_k\}) = 1$ for any $j \in \mathrm{Im}(a_k)$, and there is no *split measurement*, i.e., $\#(\{j | (i, j) \in a_k\}) = 1$ for any $i \in \mathrm{Dom}(a_k)$.

(A3) [Measurement Ordering]: Given the number m_k of measurements and given the set $\mathrm{Dom}(a_k)$ of indices of detected targets at frame k, the target assignment a_k is independent of the target states at time t_k, and all the $m_k!/(m_k - \#(\mathrm{Dom}(a_k)))!$ possible realizations of a_k, under Assumption A2, are equally probable.

Assumption A2 makes each target assignment a_k a one-to-one function, while Assumption A3 is to best reflect the fact that the actual process of how each sensor orders the measurements $(y_{kj})_{j=1}^{m_k}$ might be very complex and different from sensor to sensor, making any ordering of the measurements not informative.

One of the most basic assumptions for any dynamical state estimation problem to be tractable is *conditional independence* of information. In our MTT cases, that assumption is translated into the conditional independence of the pair (y_k, a_k) of observations y_k and unobservable target assignments a_k, for $k = 1, 2, \ldots$.

(A4) [Conditional Independence]: For any sequence $(y_k, a_k)_{k=1}^K$ of measurement frames and target assignments, we have

$$P\left((y_k, a_k)_{k=1}^K \middle| \varphi\left(((x_i(t))_{t\in[t_0,\infty)})_{i=1}^n\right)\right) = \prod_{k=1}^K P\left(y_k, a_k \middle| \varphi\left((x_i(t_k))_{i=1}^n\right)\right), \tag{8}$$

where $\varphi(x) = x$ for RFSeq formalism, $\varphi(x) = [x]$ for FPP formalism, and $\varphi((x_i)_{i=1}^n) = \{x_i\}_{i=1}^n$ for RFSet formalism, while $\varphi(((x_i(t))_{t\in[t_0,\infty)})_{i=1}^n)$ should be understood as the σ-algebra of events generated by the RFSeq, the FPP, or the RFSet of the entire stochastic processes, modeling targets according to each formalism.

In (8), and in many of subsequent equations, to avoid excessive notational complexities, we will use P or p as the generic symbol for any conditional or unconditional PD whenever its usage will not generate any confusion. However, we should remember that, when the usage of symbol P involves any discrete RFSet such as the target assignment a_k (and also the data association λ_K, defined in Section III-C), its PD is the JMD with respect to the counting measure on the space of subsets of a countable space, and, as discussed in Section II-C , is actually the PMF.

Under Assumption A2, therefore, Assumption A3 can be written as a conditional PMF

$$\begin{aligned} &P(a_k | m_k, \mathrm{Dom}(a_k), X(t_k)) \\ &\quad = P(a_k | m_k, \#(\mathrm{Dom}(a_k))) = \tfrac{(m_k - \#(\mathrm{Dom}(a_k)))!}{m_k!} \end{aligned} \tag{9}$$

for each frame k, where $X(t_k)$ is the target state set in any of the three formalisms. With a straightforward Bayesian expansion, we can write each scan-wise *extended* likelihood function on the right-hand side of (8) as[43]

$$\begin{aligned} &P\left((y_{kj})_{j=1}^{m_k}, a_k \middle| \varphi\left((x_i(t_k))_{i=1}^n\right)\right) = \tfrac{(m_k - \#(\mathrm{Dom}(a_k)))!}{m_k!} \\ &P\left((y_{kj})_{j=1}^{m_k} \middle| a_k, m_k, \varphi\left((x_i(t_k))_{i=1}^n\right)\right) \\ &P\left(m_k \middle| \mathrm{Dom}(a_k), \varphi\left((x_i(t_k))_{i=1}^n\right)\right) \\ &P\left(\mathrm{Dom}(a_k) | \varphi\left((x_i(t_k))_{i=1}^n\right)\right), \end{aligned} \tag{10}$$

[39]Synonymous to scans, measurement sets, data sets, etc.

[40]We generally assume each measurement space E_{Mk} is also LCHC2 so that its conditional PD is well defined as the likelihood function that is a measurable function of the state $X(t_k)$ in $\bigcup_{n=0}^\infty E^n$ or $\bigcup_{n=0}^\infty E^n/n!$ or $\mathcal{F}(E)$. The measurement space E_{Mk} is essentially the field of view of the sensor for frame k, and hence should be compact, or at least bounded.

[41]Generally, one of the multiple sensors.

[42]We also write $a_k(i) = j$ to mean $(i, j) \in a_k$, under Assumption A2.

[43]$P((y_{kj})_{j=1}^{m_k}, a_k | X)$ is the conditional joint PD for RFSeq $(y_{kj})_{j=1}^{m_k}$ in E_{Mk} and RFSet a_k on the space of pairs of integers. Whenever we use any RFSeq such as $(y_{kj})_{j=1}^{m_k}$, we need to remember that the length m_k is a random variable, so that we have $P((y_{kj})_{j=1}^{m_k}) = P((y_{kj})_{j=1}^{m_k} | m_k) P(m_k)$.

the right-hand side of which consists of four factors: 1) the equal probability of each realization of assignment a_k given only set $\mathrm{Dom}(a_k)$ of indices for detected targets, and the number m_k of measurements, as shown by (9), 2) the PD of the values of the measurements $(y_{kj})_{j=1}^{m_k}$ in $E_{\mathrm{M}k}$, given the number m_k of the measurements and their origins specified by a_k, 3) the probability of the number of false alarms of being $m_{\mathrm{FA}k}$ that is equal to $m_k - \#(\mathrm{Dom}(a_k))$ under Assumption A2, and 4) the joint probability of detection/nondetection of the n targets.

We should note that, although each measurement frame $y_k = (y_{kj})_{j=1}^{m_k}$ is modeled as an ordered set (i.e., RFSeq), the assignment is defined on an arbitrarily chosen enumeration of the targets, modeled by RFSeq with permutable n-PDist, or FPP, or RFSet of stochastic processes. Consequently, in (8) and (10), $(x_i(t_k))_{i=1}^{n}$ means target states with an enumeration that is arbitrary but consistent throughout all the measurement frames that we model.

By summing out the assignment a_k in (10), we obtain the measurement frame likelihood function in the *ordinary sense* as

$$
\begin{aligned}
&P\left((y_{kj})_{j=1}^{m_k} \middle| \varphi\left((x_i(t_k))_{i=1}^{n}\right)\right) = \tfrac{1}{m_k!} \sum_{a_k \in \bar{\mathcal{A}}(\{1,\ldots,n\},\{1,\ldots,m_k\})} \\
&\quad P\left((y_{kj})_{j=1}^{m_k} \middle| a_k, m_k, \varphi\left((x_i(t_k))_{i=1}^{n}\right)\right) \\
&\quad (m_{FAk}!)\, P\left(m_{FAk} \middle| \mathrm{Dom}(a_k), \varphi\left((x_i(t_k))_{i=1}^{n}\right)\right) \\
&\quad P\left(\mathrm{Dom}(a_k) | \varphi\left((x_i(t_k))_{i=1}^{n}\right)\right),
\end{aligned}
\tag{11}
$$

where $m_{\mathrm{FA}k} = m_k - \#(\mathrm{Dom}(a_k))$ and $\bar{\mathcal{A}}$ (as well as $\mathcal{A}$ that will be used later) is the symbol for the space of one-to-one functions, which we may call *assignment functions*, defined, for any pair of finite sets I and J, as

$$
\begin{cases}
\mathcal{A}(I,J) \stackrel{\mathrm{def}}{=} \left\{ a : I \to J \middle| \begin{array}{l} I = \mathrm{Dom}(a) \text{ and} \\ \#(\mathrm{Dom}(a)) = \#(\mathrm{Im}(a)) \end{array} \right\}, \\
\bar{\mathcal{A}}(I,J) \stackrel{\mathrm{def}}{=} \left\{ a : D \to J \middle| \begin{array}{l} D = \mathrm{Dom}(a) \subseteq I \text{ and} \\ \#(\mathrm{Dom}(a)) = \#(\mathrm{Im}(a)) \end{array} \right\}.
\end{cases}
\tag{12}
$$

It is significant that we define each measurement frame $(y_{kj})_{j=1}^{m_k}$ as an RFSeq (not as an FPP or an RFSet) so that we can call each measurement as the "jth" measurement at the "kth" frame, to define data association hypotheses in the next section. Apparently, both sides of (11) are permutable with respect to the index $j \in \{1, \ldots, m_k\}$ of measurements $(y_{kj})_{j=1}^{m_k}$, as well as with respect to the index $i \in \{1, \ldots, n\}$ of targets $(x_i(t_k))_{i=1}^{n}$, as the likelihood function and hence $(y_{kj})_{j=1}^{m_k}$ can be considered as an FPP or RFSet with conditional JMD $\phi_{Mk}([(y_{kj})_{j=1}^{m_k}] | [(x_i(t_k))_{i=1}^{n}])$ in FPP formalism, or $\phi_{Mk}(\{y_{kj}\}_{j=1}^{m_k} | \{x_i(t_k)\}_{i=1}^{n})$ in RFSet formalism, dropping $1/m_k!$ from (11), reflecting the fact that the *order* of the measurements does not bear any information.

C. Association and Association Hypotheses

Measurement-to-measurement or *data-to-data* or simply *data association* λ_K over given cumulative frames $(y_k)_{k=1}^{K} = ((y_{kj})_{j=1}^{m_k})_{k=1}^{K}$ is defined, from the multiframe target assignment $(a_k)_{k=1}^{K}$, as

$$
\lambda_K = \left\{ \bigcup_{k=1}^{K} \{(k, j) | (i, j) \in a_k\} \middle| i \in \bigcup_{k=1}^{K} \mathrm{Dom}(a_k) \right\}. \tag{13}
$$

We should note that we define the data association λ_K, not as a partition of the cumulative measurements themselves $(y_k)_{k=1}^{K} = ((y_{kj})_{j=1}^{m_k})_{k=1}^{K}$, but rather as a partition of the cumulative set of measurement indices, $I_K \stackrel{\mathrm{def}}{=} \bigcup_{k=1}^{K} \{k\} \times \{1, \ldots, m_k\}$. Each component of λ_K constitutes the indices of all the measurements originating from the same target, so that $\#(\lambda_K)$ targets are detected in $(y_k)_{k=1}^{K}$, implying $\#(\lambda_K) \le n$, while its complement $I_K \backslash (\bigcup \lambda_K)$ is the set of all the measurement indices for false alarms in $(y_k)_{k=1}^{K}$. We call any realization of association λ_K a *data association hypothesis*[44] or simply an *association hypothesis* or a *hypothesis*. As a consequence of Assumption A2, the set Λ_K of all the association hypotheses on $(y_k)_{k=1}^{K}$ is given by

$$
\Lambda_K = \left\{ \lambda \subseteq \mathcal{T}_K \backslash \{\emptyset\} \middle| \begin{array}{l} \tau \cap \tau' = \emptyset \text{ for any} \\ (\tau, \tau') \in \lambda \times \lambda \text{ such that } \tau \neq \tau' \end{array} \right\}, \tag{14}
$$

where

$$
\mathcal{T}_K \stackrel{\mathrm{def}}{=} \left\{ \tau \subseteq I_K \middle| \begin{array}{l} \#\left(\{ j \in \{1, \ldots, m_k\} | (k, j) \in \tau \}\right) \le 1 \\ \text{for any } k \in \{1, \ldots, K\} \end{array} \right\}, \tag{15}
$$

each member of which is called a *track* on $(y_k)_{k=1}^{K}$; i.e., each hypothesis is a *consistent* (i.e., *nonoverlapping*) set of nonempty tracks.

In Section IV, we will describe issues concerning generation of data association hypotheses, and their evaluation under additional sets of assumptions, completing our definition of MHT in the three mathematical formalisms, which is the main goal of this paper.

Remark 6 (Hypotheses): As we call any possible realization of the data association, i.e., an RFSet, a data association hypothesis, we may call any possible realization of target assignment a_k for each frame k, i.e., any element in $\bar{\mathcal{A}}(\{1, \ldots, n\}, \{1, \ldots, m_k\})$, a *target-to-measurement assignment hypothesis*. The latter type of hypotheses was introduced in the context of the PDA [12] and JPDA [13] algorithms, assuming a fixed number of targets, predating the development of the MHT. In [2, Sec. 7.5.2, p. 431] and [3, Sec. 4.2, p. 113], multiple-scan, non-Gaussian extension of the JPDA algorithms is discussed. In [3, Sec. 4.1.1, p. 109], in order to model an unknown number of

[44] Also known as data-to-data or measurement-to-measurement association hypothesis. We are using two different terms, "association" and "assignment," to make a clear distinction between two random sets λ_K and a_k.

targets, it was proposed to augment target space E to $E \cup \{\theta\}$, where "θ" is the "target does not exist" state, and to use the joint state space $(E \cup \{\theta\})^N$ with a fixed number N (that serves as a priori upper bound N on the number of targets), within the extended JPDA context mentioned earlier.

As seen in (13), each association hypothesis $\lambda \in \Lambda_K$ can be viewed as an equivalence class of multiframe target assignment hypotheses $(a_k)_{k=1}^K \in \prod_{k=1}^K \bar{\mathcal{A}}(\{1, ..., n\}, \{1, ..., m_k\})$, the equivalence defined through the permutation of the target indices. Given the number n of targets and cumulative frames $(y_k)_{k=1}^K$, through (13), each multiframe target assignment $(a_k)_{k=1}^K$ is uniquely determined by a pair (λ, α) of data association $\lambda \in \Lambda_K$ and *track-to-target assignment* (or simply *track assignment*) $\alpha \in \mathcal{A}(\lambda, \{1, ..., n\})$ so that we have $\tau = \bigcup_{k=1}^K \{(k, a_k(\alpha(\tau)))|\alpha(\tau) \in \mathrm{Dom}(a_k)\}$ for any $\tau \in \lambda_K$. The target permutability, assumed by Assumption A1, implies that, given $((y_k)_{k=1}^K, \lambda_K, n)$, every realization of track assignment α in $\mathcal{A}(\lambda, \{1, ..., n\})$ is equally probable. Moreover, in FPP or RFSet formalism, any *arbitrary enumeration* of the targets in (8)–(13) can be viewed as another random assignment from the set of targets, $X = [(x_i)_{i=1}^n]$ or $X = \{x_i\}_{i=1}^n$, to its index set $\{1, ..., n\}$ with $n = \ell(X)$ or $n = \#(X)$, i.e., a random element in $\mathcal{A}(X, \{1, ..., n\})$.

Remark 7 (Merged and Split Measurements): For many sensors, the no-merged-or-split-measurement assumption (A2) is a reasonable assumption. It is very likely that any occasional violation of this assumption may be helped out by an effective *recovery* algorithm. On the other hand, there have been many efforts to generate and probabilistically evaluate, explicitly, merge/split measurement hypotheses, e.g., [34] (merged measurements) and [35] (split measurements).

IV. HYPOTHESIS GENERATION AND EVALUATION

To our best knowledge, the concept of the data association hypothesis, the core of the MHT, as described in the previous section, was first clearly defined in [5], in terms of tracks and hypotheses, together with an algorithm for selecting the single best (*most probable* or *maximum a posteriori probability*) hypothesis in a batch-data-processing mode. Subsequently, an algorithm for simultaneously generating and evaluating tracks and hypotheses, using recursive formulas, was first systematically and comprehensively described in [6]. In this section, we discuss hypothesis generation, and hypothesis evaluation under commonly used assumptions, using the three formalisms described in Section II, and the target/sensor models defined in Section III, which we may view as a form of generalizations of the results described in [5]–[7].

A. Hypothesis Generation and Management

For any pair $((y_k)_{k=1}^{K_1}, (y_k)_{k=1}^{K_2})$ of cumulative frames such that $K_1 < K_2$, we call a track $\tau_1 \in \mathcal{T}_{K_1}$ a *predecessor* of a track $\tau_2 \in \mathcal{T}_{K_2}$ (or τ_2 is a *successor* of τ_1) if $\tau_1 = \{(k, j) \in \tau_2 | k \leq K_1\}$ (including the case $\tau_1 = \oslash$). We call a hypothesis $\lambda_1 \in \Lambda_{K_1}$ a *predecessor* of a hypothesis $\lambda_2 \in \Lambda_{K_2}$ (or λ_2 is a *successor* of λ_1) if, for each track $\tau_2 \in \lambda_2$, there exists a (necessarily unique) predecessor τ_1 in λ_1 or otherwise track τ_2 has an empty predecessor $\tau_1 = \emptyset$ in $\mathcal{T}_{K_1}$. Then, both cumulative collections of tracks and hypotheses, $\bigcup_{k=1}^K \mathcal{T}_k$ and $\bigcup_{k=1}^K \Lambda_k$, respectively, form arborescent (tree) directed graphs through the predecessor–successor relations. For each hypothesis $\lambda \in \Lambda_{K_2}$ and each track $\tau \in \mathcal{T}_{K_2}$, we denote their unique predecessors in Λ_{K_1} and $\mathcal{T}_{K_1}$ by $\lambda_{|K_1}$ and $\tau_{|K_1}$, respectively.

There may be many systematic methods for generating these trees. In [6], D. B. Reid called hypothesis tree generation using each measurement y_{kj} as a level variable[45] the *measurement-oriented approach*, from which the term *measurement-oriented MHT* originated, in contrast to the *target-oriented approach* in which a target-to-measurement assignment tree is generated using each target index as a level variable[46] (e.g., for PDA and JPDA algorithms) with a fixed known number n of targets.

The algorithm described in [5] recursively generates and evaluates tracks (including the track likelihood defined later in this section), in effect, building a track tree. Using a batch-processing form of hypothesis evaluation, it then selects the single best association hypothesis on $(y_k)_{k=1}^K$ based on the a posteriori probability $P(\lambda|(y_k)_{k=1}^K)$ (defined in Section IV-B) for each hypothesis, using a zero–one integer programming technique, where a set of association hypotheses is formed as feasible solutions to a system of binary linear equations. Over the years, it has become customary to call any MHT algorithm using this approach, which originated from [5], a *track-oriented MHT*.

It is well known that the numbers, $\#(\Lambda_K)$ and $\#(\mathcal{T}_K)$, of hypotheses and tracks generally grow very rapidly, at exponential rates in many cases, so that any practical MHT implementation must have reasonable means of controlling the growth. Common methods for controlling the growth of the number of association hypotheses include *gating, pruning, combining*, and *clustering*, as outlined in [6]. The single best hypothesis selection of [5] over sliding windows of consecutive frames has been widely used as means for pruning track trees in a variety of ways for many track-oriented MHT algorithms. Many heuristic methods to control the numbers, $\#(\Lambda_K)$ and $\#(\mathcal{T}_K)$, generally known as *hypothesis management methods*, have been devised in the past 40 years or so, as described in [8].

[45] Assigning each measurement to tracks in hypotheses at each expansion.

[46] Assigning each target to measurements at each expansion.

B. Hypothesis Evaluation: Independence Assumptions

As mentioned in Remark 6 in Section III-C, in any of the three formalisms, an immediate consequence of Assumptions A1–A3 and the definition (13) of data association is as follows: Given the data association $\lambda_K \in \Lambda_K$ on cumulative frame $(y_k)_{k=1}^K$, and given the number n of targets such that $n \geq \#(\lambda_K)$, any one of the equally possible $n!/(n-\#(\lambda_K))!$ track assignments α's in $\mathcal{A}(\lambda_K, \{1, ..., n\})$ will define uniquely a multiframe target assignment $(a_k)_{k=1}^K \in \prod_{k=1}^K \bar{\mathcal{A}}(\{1, ..., n\}, \{1, ..., m_k\})$. Hence, if (13) holds, we have $P(\lambda_K | n, (y_k)_{k=1}^K, (a_k)_{k=1}^K) = 1$ and $P((a_k)_{k=1}^K | n, (y_k)_{k=1}^K, \lambda_K) = (n-\#(\lambda_K))!/n!$. Both are zero otherwise. Hence, we have

$$\begin{aligned} &P(\lambda_K, n | (y_k)_{k=1}^K) \\ &= P((y_k)_{k=1}^K)^{-1} \tfrac{n!}{(n-\#(\lambda_K))!} P((y_k, a_k)_{k=1}^K, n). \end{aligned} \tag{16}$$

On the right-hand side of (16), $(a_k)_{k=1}^K \in \prod_{k=1}^K \bar{\mathcal{A}}(\{1, ..., n\}, \{1, ..., m_k\})$ is any one of the $n!/(n-\#(\lambda_K))!$ multiframe target assignment hypotheses that supports λ_K (through (13)).

Since the sensor model defined in Section III-B allows us to have multiple sensors, the sequence of measurement frame times, $(t_k)_{k=1}^K$, may contain repeated time stamps. We therefore need to consider a subset $[K]$ of $\{1, ..., K\}$ to remove any repeated time, i.e., $[K] \subseteq \{1, ..., K\}$, $\#([K]) = \#(\{t_k\}_{k=1}^K) \leq K$, and $K \in [K]$, for hypothesis evaluation and target state estimation.

Under Assumptions A1–A4, $P((y_k, a_k)_{k=1}^K, n)$ in (16) can be expanded by the target states, in RFSeq formalism, as

$$\begin{aligned} P((y_k, a_k)_{k=1}^K, n) = p_n \int_{E^{\#([K])n}} &\left(\prod_{k=1}^K P\left(y_k, a_k \middle| (x_i(t_k))_{i=1}^n\right)\right) \\ &f^{(n)}\left(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n; (t_\kappa)_{\kappa\in[K]}\right) \\ &\mu^{\#([K])n}\left(((dx_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n\right). \end{aligned} \tag{17}$$

The product, $n! p_n f^{(n)}(\cdot; (t_\kappa)_{\kappa\in(K)})$, which appears when we substitute (17) into (16), is nothing but the n-JD $J^{(n)}(\cdot; (t_\kappa)_{\kappa\in(K)})$, and hence should be replaced by JMD $\phi([((x_i(t_\kappa))_{\kappa\in(K)})_{i=1}^n]; (t_\kappa)_{\kappa\in(K)})$ in FPP formalism, and JMD $\phi(\{(x_i(t_\kappa))_{\kappa\in(K)}\}_{i=1}^n; (t_\kappa)_{\kappa\in(K)})$in RFSet formalism. Each frame-wise extended likelihood function $P(y_k, a_k | \varphi((x_i(t_k))_{i=1}^n))$ (with $\varphi(x) = x$, $\varphi(x) = [x]$, and $\varphi((x_i)_{i=1}^n) = \{x_i\}_{i=1}^n$ for RFSeq, FPP, and RFSet formalisms, respectively) can then be expanded by the sensor model (10).

The a posteriori probabilities of each hypothesis $\lambda_K \in \Lambda_K$ and of the number n of targets are obtained separately through marginalization of (16) with (17). To evaluate them in a practical and hence meaningful way, however, we need to divorce ourselves from target-to-measurement assignments, $(a_k)_{k=1}^K$, which would require a few more assumptions on the target and sensor models, including

(A5) [i.i.d. Targets]: Given the number n of targets, assume the joint probability distribution for the set of targets is i.i.d. with the common single-target joint PD f_{TGT}, in the sense that, for any $(s_\kappa)_{\kappa=1}^N \in [t_0, \infty)^N$ of distinct times, for any $((x_i(s_\kappa))_{\kappa=1}^N)_{i=1}^n \in E^{Nn}$, we have $f^{(n)}(((x_i(s_\kappa))_{\kappa=1}^N)_{i=1}^n; (s_\kappa)_{\kappa=1}^N) = \prod_{i=1}^n f_{\mathrm{TGT}}((x_i(s_\kappa))_{\kappa=1}^N; (s_\kappa)_{\kappa=1}^N)$ in RFSeq formalism, and $\phi_{\mathrm{TGT}}(\varphi(((x_i(s_\kappa))_{\kappa=1}^N)_{i=1}^n); (s_\kappa)_{\kappa=1}^N) = n! p_n \prod_{i=1}^n f_{\mathrm{TGT}}((x_i(s_\kappa))_{\kappa=1}^N; (s_\kappa)_{\kappa=1}^N)$ in FPP ($\varphi(x) = [x]$) or RFSet ($\varphi((x_i)_{i=1}^n) = \{x_i\}_{i=1}^n$) formalism.

Under this i.i.d. assumption, the target model can conveniently be expressed by the *intensity measure density* (IMD),[47] $\gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa=1}^N; (s_\kappa)_{\kappa=1}^N) = \nu f_{\mathrm{TGT}}((\xi_\kappa)_{\kappa=1}^N; (s_\kappa)_{\kappa=1}^N)$, for any N, for any $(\xi_\kappa, s_\kappa)_{\kappa=1}^N \in (E \times [t_0, \infty))^N$, with a priori expected number of targets, $\nu = \sum_{n=1}^\infty n p_n < \infty$.

Another set of independence assumptions is concerned with our sensor model:

(A6) [Independent Detections and i.i.d. False Alarms]: For each measurement frame, $y_k = (y_{kj})_{j=1}^{m_k}$, 1) the target detection is target-wise independent and determined by a common detection probability as a function $p_{\mathrm{D}k}$ of the target state, 2) the target-state-to-measurement transition is also target-wise independent with a common transition probability density[48] $p_{\mathrm{M}k}$, and 3) each false alarm in the frame is independent from the target states and from other false alarms with a common PD, $p_{\mathrm{FA}k}$, while the probability of the number of false alarms in the frame being $m_{\mathrm{FA}k}$ is given as $p_{\mathrm{NFA}k}(m_{\mathrm{FA}k})$ with finite mean $\nu_{\mathrm{FA}k} = \sum_{m_{\mathrm{FA}k}=1}^\infty m_{\mathrm{FA}k} p_{\mathrm{NFA}k}(m_{\mathrm{FA}k}) < \infty$.

By applying Assumption A6 to (10), for each k, we have

$$\begin{aligned} P\left((y_{kj})_{j=1}^{m_k}, a_k \middle| \varphi((x_i(t_k))_{i=1}^n)\right) = \tfrac{L_{\mathrm{FA}k}(\{1,\dots,m_k\}\setminus \mathrm{Im}(a_k))}{m_k!} \\ \left(\prod_{i\in \mathrm{Dom}(a_k)} p_{\mathrm{M}k}(y_{k a_k(i)} | x_i(t_k)) p_{\mathrm{D}k}(x_i(t_k))\right) \\ \left(\prod_{\substack{i=1 \\ i\notin \mathrm{Dom}(a_k)}}^n (1 - p_{\mathrm{D}k}(x_i(t_k)))\right) \end{aligned} \tag{18}$$

[47]For an RFSeq $(x_i)_{i=1}^n$, an FPP $[(x_i)_{i=1}^n]$, or an RFSet $\{x_i\}_{i=1}^n$ in $(E^N, \mathcal{B}_N)$, for any $N = 1, 2, ...$, the intensity measure (IM) Γ is a finite measure on $(E^N, \mathcal{B}_N)$ defined by $\Gamma(B) = \mathbb{E}(\sum_{i=1}^n \mathbb{I}(x_i; B))$ (with the random measure representation $\mathcal{N}$ of FPP formalism, $\Gamma(B) = \mathbb{E}(\mathcal{N}(B))$.), for each $B \in \mathcal{B}_N$, using the generic symbols, $\mathbb{E}$ and $\mathbb{I}$, for mathematical expectation and indicator function. The IMD is its density, i.e., the Radon–Nikodym derivative with respect to the measure μ^N. More commonly used name for IM is the *first-order moment measure* ([15, Sec. 5.4, p. 132], but we prefer IM and IMD because we only use the moment measure of the first order. Another synonym is *expectation measure*. A conditional version of IMD is called *probability hypothesis density* in [25]–[27].

[48]The use of a common $p_{\mathrm{M}k}$, equally for all the m_k measurements, y_{kj}'s, in frame k, may not be justified when each measurement y_{kj} has different measurement error characteristics from others. In that case, we should use the measurement-index-dependent $p_{\mathrm{M}kj}$ in place of $p_{\mathrm{M}k}$ (which we avoid for the sake of simplicity).

with the frame-wise false alarm likelihood, defined for each $I_{\mathrm{FA}k} \subseteq \{1, ..., m_k\}$, as

$$L_{\mathrm{FA}k}(I_{\mathrm{FA}k}) = L_{\mathrm{NFA}k}(\#(I_{\mathrm{FA}k})) \prod_{j \in I_{\mathrm{FA}k}} \gamma_{\mathrm{FA}k}(y_{kj}), \tag{19}$$

where $\gamma_{\mathrm{FA}k}(\eta) = \nu_{\mathrm{FA}k} p_{\mathrm{FA}k}(\eta)$ is the IMD of the false alarms in frame k at each $\eta \in E_{\mathrm{M}k}$, and $L_{\mathrm{NFA}k}(m_{\mathrm{FA}k}) = (m_{\mathrm{FA}k}!/(\nu_{\mathrm{FA}k})^{m_{FA}}) p_{\mathrm{NFA}k}(m_{\mathrm{FA}k})$ is the likelihood on the number $m_{\mathrm{FA}k}$ of false alarms in frame $y_k = (y_{kj})_{j=1}^{m_k}$.

As shown in Appendix A, Assumptions A1–A6 allow us to derive a batch-mode hypothesis evaluation formula, which we call *Morefield form*, in terms of the a posteriori probability of the data association λ_K on cumulative frame $(y_k)_{k=1}^K$, as

$$\begin{aligned} &P(\lambda_K | (y_k)_{k=1}^K) \\ &= C_{\mathrm{M}K}^{-1} L_{\mathrm{NDT}K}(\#(\lambda_K)) \left(\prod_{\tau \in \lambda_K} L_{\mathrm{TRK}K}(\tau) \right) L_{\mathrm{FA}}^{(K)}(\lambda_K) \end{aligned} \tag{20}$$

with

1) the normalizing constant (*Morefield constant*), $C_{\mathrm{M}K} = P((y_k)_{k=1}^K)(\prod_{k=1}^K m_k!)$;

2) track likelihood

$$\begin{aligned} L_{\mathrm{TRK}K}(\tau) = \int_{E^{\#([K])}} &\left(\prod_{k=1}^K q_{\mathrm{MD}k}(\xi_k; \tau) \right) \\ &\gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa \in [K]}; (t_\kappa)_{\kappa \in [K]}) \prod_{\kappa \in [K]} \mu(d\xi_\kappa) \end{aligned} \tag{21}$$

defined for each track $\tau \in \mathcal{T}_K$, derived from the a priori joint IMD γ_{TGT}, and the *extended target-wise state likelihood function* $q_{\mathrm{MD}k}(\cdot; \tau)$, defined by, for each $\xi \in E$,

$$q_{\mathrm{MD}k}(\xi; \tau) = \begin{cases} p_{\mathrm{M}k}(y_j|\xi) p_{\mathrm{D}k}(\xi), \\ \quad \text{if } (k, j) \in \tau \text{ for some } j \in \{1, ..., m_k\}, \\ 1 - p_{\mathrm{D}k}(\xi), \\ \quad \text{if } (k, j) \notin \tau \text{ for any } j \in \{1, ..., m_k\}; \end{cases} \tag{22}$$

3) likelihood $L_{\mathrm{NDT}K}(n_{\mathrm{D}})$ of the cumulative number $n_{\mathrm{D}} = \#(\lambda_K)$ of detected targets defined by

$$L_{\mathrm{NDT}K}(n_{\mathrm{D}}) = \sum\nolimits_{n=n_{\mathrm{D}}}^{\infty} p_n \frac{n!}{\nu^n} \frac{(\hat{\nu}_K)^{n-n_{\mathrm{D}}}}{(n-n_{\mathrm{D}})!} \tag{23}$$

expressed by the a priori mean ν of the number of the targets and the a posteriori expectation $\hat{\nu}_K = L_{\mathrm{TRK}K}(\emptyset)$ of the number of the targets that remain undetected through the cumulative frames $(y_k)_{k=1}^K$;

4) *multiframe false alarm likelihood* $L_{\mathrm{FA}}^{(K)}$, defined by

$$L_{\mathrm{FA}}^{(K)}(\lambda) = \prod_{k=1}^K L_{\mathrm{FA}k}\left(\{j \in \{1, ..., m_k\} | (k, j) \notin \cup\lambda\}\right) \tag{24}$$

for each $\lambda \in \Lambda_K$ through frame-wise false alarm likelihood $L_{\mathrm{FA}k}$ defined by (19).

As shown in [5], Morefield form (20) for evaluating hypotheses can be expressed as a form of zero–one integer programming problem, by enumerating the set $\cup\Lambda_K = \mathcal{T}_K \backslash \{\emptyset\}$ of all the nonempty tracks as $(\tau_i)_{i=1}^{N_{\mathrm{T}}}$, and by mapping the set Λ_K of all the hypotheses into the space $\{0, 1\}^{N_{\mathrm{T}}}$ through $(\xi_i)_{i=1}^{N_{\mathrm{T}}} = (\mathbb{I}(\tau_i; \lambda))_{i=1}^{N_{\mathrm{T}}} \in \{0, 1\}^{N_{\mathrm{T}}}$ for each $\lambda \in \Lambda_K$.

C. More on Hypothesis Evaluation: Markov and Poisson Assumptions

We will now introduce two more commonly used assumptions.

(A7) [Markov Assumption]: The targets are modeled as an RFSeq, an FPP, or an RFSet of independent stochastic processes, with a common a priori joint IMD γ_{TGT}, which is *Markovian*, in the sense that, for any N-tuple $(s_\kappa)_{\kappa=1}^N \in [t_0, \infty)^N$ of distinct times, such that $s_1 < s_2 < \cdots < s_N$, and for any $(\xi_\kappa)_{\kappa=1}^N \in E^N$, we have

$$\begin{aligned} &\gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa=1}^N; (s_\kappa)_{\kappa=1}^N) \\ &= \gamma_{\mathrm{TGT}}(\xi_1; s_1) \prod_{\kappa=2}^N f_{\mathrm{TRN}}(\xi_\kappa | \xi_{\kappa-1}; s_\kappa - s_{\kappa-1}, s_{\kappa-1}) \end{aligned} \tag{25}$$

with a given *state transition probability density* (STPD), $f_{\mathrm{TRN}}(\cdot|\cdot; \Delta s, s)$, on $(E, \mathcal{B}, \mu)$, for each $\Delta s > 0$ and $s \in [t_0, \infty)$.

Markov assumption (A7) enables us to calculate track likelihood $L_{\mathrm{TRK}K}(\tau)$ for each nonempty track $\tau \in \cup\Lambda_K = \mathcal{T}_K \backslash \{\emptyset\}$ defined by (21), recursively as

$$L_{\mathrm{TRK}k}(\tau_{|k}) = \begin{cases} \gamma_{\mathrm{MND}k}(y_{kj}), \\ \quad \text{if } k = k_0(\tau) \text{ with } (k, j) \in \tau, \\ L_{\mathrm{TRK}(k-1)}(\tau_{|(k-1)}) L_{\mathrm{MD}k}(\tau_{|k}), \\ \quad \text{if } k > k_0(\tau) \end{cases} \tag{26}$$

for[49] $k = k_0(\tau), k_0(\tau) + 1, ..., K$, with the measurement IMD $\gamma_{\mathrm{MND}k}$ from newly detected targets, and the measurement-or-no-detection likelihood $L_{\mathrm{MD}k}$, which are defined by

$$\begin{cases} \gamma_{\mathrm{MND}k}(y_{kj}) = \int_E p_{\mathrm{M}k}(y_{kj}|\xi) p_{\mathrm{D}k}(\xi) \bar{\gamma}_k(\xi) \mu(d\xi) \\ L_{\mathrm{MD}k}(\tau_{|k}) = \int_E q_{\mathrm{MD}k}(\xi; \tau_{|k}) \bar{f}_k(\xi | \tau_{|(k-1)}) \mu(d\xi) \end{cases} \tag{27}$$

for each k, any $y_{kj} \in E_{\mathrm{M}k}$, and any track $\tau \in \mathcal{T}_K$ for any $K \geq k$. The recursive calculation (26) of track likelihood can be done in parallel to a recursive process for obtaining the *update*d track target state PD $\hat{f}_k(\cdot|\tau_{|k})$ from the *predicted* $\bar{f}_k(\cdot|\tau_{|(k-1)})$, and generating the next *predicted* $\bar{f}_{k+1}(\cdot|\tau_{|k})$, for every $\xi \in E$, as

$$\begin{cases} \hat{f}_k(\xi|\tau_{|k}) = \begin{cases} \gamma_{\mathrm{MND}k}(y_{kj})^{-1} p_{\mathrm{M}k}(y_{kj}|\xi) p_{\mathrm{D}k}(\xi) \bar{\gamma}_k(\xi), \\ \quad \text{if } k = k_0(\tau) \text{ with } (k, j) \in \tau, \\ L_{\mathrm{MD}k}(\tau_{|k})^{-1} q_{\mathrm{MD}k}(\xi; \tau_{|k}) \bar{f}_k(\xi|\tau_{|(k-1)}), \\ \quad \text{otherwise } (\tau_{|(k-1)} \neq \emptyset), \end{cases} \\ \bar{f}_k(\xi|\tau_{|(k-1)}) = \begin{cases} \int_E f_{\mathrm{TRN}}(\xi|\xi'; t_k - t_{k-1}, t_{k-1}) \\ \quad \hat{f}_{k-1}(\xi'|\tau_{|(k-1)}) \mu(d\xi'), \quad \text{if } t_k > t_{k-1}, \\ \hat{f}_{k-1}(\xi|\tau_{|(k-1)}), \quad \text{if } t_k = t_{k-1}. \end{cases} \end{cases} \tag{28}$$

[49] $k_0(\tau) \stackrel{\text{def}}{=} \min\{k | (k, j) \in \tau \text{ for some } j\}$ is the index of the first frame where track τ obtains a measurement, i.e., track initiation frame.

The *predicted* IMD $\bar{\gamma}_k(\xi)$ of the undetected targets in (27) and (28) is obtained from the similar recursion, along the *updated* IMD $\hat{\gamma}_k(\xi)$, for each $k = 1, 2, ...,$ as

$$\begin{cases} \hat{\gamma}_k(\xi) = (1 - p_{\mathrm{D}k}(\xi))\bar{\gamma}_k(\xi) \\ \bar{\gamma}_k(\xi) = \begin{cases} \int_E f_{\mathrm{TRN}}(\xi|\xi'; t_k - t_{k-1}, t_{k-1})\hat{\gamma}_{k-1}(\xi')\mu(d\xi'), \\ \qquad\qquad \text{if } k > 1 \text{ and } t_k > t_{k-1}, \\ \hat{\gamma}_{k-1}(\xi), \qquad \text{if } k > 1 \text{ and } t_k = t_{k-1}, \\ \gamma_{TGT}(\xi, t_1), \quad \text{if } k = 1 \end{cases} \end{cases} \tag{29}$$

for every $\xi \in E$.

For each $k = 1, ..., K$, let $\bar{\nu}_k = \int_E \bar{\gamma}_k(\xi)\mu(d\xi)$ and $\hat{\nu}_k = \int_E \hat{\gamma}_k(\xi)\mu(d\xi)$. Then, we have $\bar{\nu}_k = \hat{\nu}_{k-1}$ for any $k > 1$ (reflecting our no-birth-no-death target model), $\bar{\nu}_1 = \nu$ is the a priori expectation of the number n of targets, and $\hat{\nu}_K = L_{\mathrm{TRK}K}(\emptyset)$ is the a posteriori expectation of the number of targets that are not detected in any of the K frames, $(y_k)_{k=1}^K$.

Under Markovian assumption (A7), we can *rewrite* Morefield form (20) in a recursive hypothesis evaluation form, as

$$P(\lambda|(y_{k'})_{k'=1}^k) = C_{\mathrm{R}k}^{-1} P(\lambda_{|(k-1)}|(y_{k'})_{k'=1}^{k-1}) \left(\prod_{\substack{\tau\in\lambda \\ \tau_{|(k-1)}\neq\emptyset}} L_{\mathrm{MD}k}(\tau) \right) \cdot \frac{L_{\mathrm{NDT}k}(\#(\lambda))}{L_{\mathrm{NDT}(k-1)}(\#(\lambda_{|(k-1)}))} \left(\prod_{\substack{j=1 \\ \{(k,j)\}\in\lambda}}^{m_k} \gamma_{\mathrm{MND}k}\left(y_{kj}\right) \right) \cdot L_{\mathrm{FA}k} \left(\bigcup_{\substack{j=1 \\ (k,j)\notin\cup\lambda}}^{m_k} \{j\} \right) \tag{30}$$

for each k, for every $\lambda \in \Lambda_k$, where

1) $C_{\mathrm{R}k} = (m_k!)P(y_k|(y_{k'})_{k'=1}^{k-1}) = J_{\mathrm{M}k}^{(m_k)}(y_k|(y_{k'})_{k'=1}^{k-1})$ is the normalizing constant (*Reid constant*);
2) $\lambda_{|(k-1)}$ is the unique predecessor[50] of $\lambda \in \Lambda_k$ in Λ_{k-1};
3) $\tau_{|(k-1)} \in \mathcal{T}_{k-1}$ is the unique predecessor of each track τ in a given hypothesis λ (including the case where $\tau_{|(k-1)} = \emptyset$; in that case, τ is a singleton $\{(k, j)\}$ for some $j \in \{1, ..., m_k\}$, i.e., a new track at frame y_k);
4) $L_{\mathrm{MD}k}(\tau)$ is the track-to-measurement likelihood of *old track* $\tau_{|(k-1)}$ and measurement y_{kj} if $(k, j) \in \tau$, and the missed detection likelihood otherwise (i.e., if $(k, j) \notin \tau$), defined in (27);
5) $L_{\mathrm{NDT}k}$ and $L_{\mathrm{NDT}(k-1)}$ are the likelihoods of the cumulative numbers of detected targets, in $(y_{k'})_{k'=1}^k$ and $(y_{k'})_{k'=1}^{k-1}$, with $\hat{\nu}_k$ and $\hat{\nu}_{k-1}$, as defined by (23);
6) $\gamma_{\mathrm{MND}k}$ is the new detection IMD defined in (27);
7) $L_{\mathrm{FA}k}(I_{\mathrm{FA}k})$ is the false alarm likelihood defined by (19).

We call this recursive hypothesis evaluation formula (30), *Reid form*, which is the non-Poisson extension of the formulas in [6, eq. 16, p. 848] and [7, eq. 19, p. 405], and was presented in [43, Th. 2, p. 231].[51]

Another common assumptions are Poisson assumptions on the a priori PDist $(p_n)_{n=0}^\infty$ of the number n of the targets, and on the PDist $(p_{\mathrm{NFA}k}(m_{\mathrm{FA}k}))_{m_{\mathrm{FA}k}=0}^\infty$ of the number $m_{\mathrm{FA}k}$ of false alarms in each frame $y_k = (y_{kj})_{j=1}^{m_k}$.

(A8) [Poisson Assumptions]: 1) The PDist $(p_n)_{n=0}^\infty$ of the number of targets is Poisson with mean ν, i.e., $p_n = e^{-\nu}\nu^n/n!$, for each $n = 0, 1, 2, ...,$ and 2) for each frame $k = 1, 2, ...,$ the PDist $p_{\mathrm{NFA}k}$ of the number $m_{\mathrm{FA}k}$ of false alarms in frame $y_k = (y_{kj})_{j=1}^{m_k}$ is Poisson with mean $\nu_{\mathrm{FA}k}$, i.e., $p_{\mathrm{NFA}k}(m_{\mathrm{FA}k}) = e^{-\nu_{\mathrm{FA}k}}\nu_{\mathrm{FA}k}^{m_{\mathrm{FA}k}}/m_{\mathrm{FA}k}!$, for each $m_{\mathrm{FA}k} = 0, 1, 2, ...$

With this Poisson assumption (A8), the likelihood functions, $L_{\mathrm{NDT}K}$ for the number of detected targets defined by (23) and likelihood $L_{\mathrm{NFA}k}$ for the number of false alarms at each frame k, both become constants, as $L_{\mathrm{NDT}K} \equiv e^{-(\nu-\hat{\nu}_K)}$ and $L_{\mathrm{NFA}k} \equiv e^{-\nu_{\mathrm{FA}k}}$, respectively. It was proven in [31, Th. 2, p. 1136] that Poisson assumption (A8) is also a necessary condition for those likelihoods to be constants. With this assumption, Morefield form (20) of hypothesis evaluation can be transformed to a linear objective function of a zero–one linear integer programming problem, or equivalently to an objective function for a form of multidimensional assignment algorithm described in [32]. Any hypothesis selection algorithm using Morefield form (20) became the core algorithm for every so-called track-oriented MHT [8].

D. Target State Estimation

Under Assumptions A1–A3, given cumulative frame $(y_k)_{k=1}^K$, for each assumed number n of targets and for each data association $\lambda_K \in \Lambda_K$, there are $n!/(n - \#(\lambda_K))!$ multiframe target assignments $(a_k)_{k=1}^K$'s, each of which *supports* association λ_K (i.e., $(\lambda_K, (a_k)_{k=1}^K)$ satisfies (13)) and is uniquely determined by one of the equally probable $n!/(n - \#(\lambda_K))!$ track assignments $\alpha \in \mathcal{A}(\lambda_K, \{1, ..., n\})$ as mentioned in Remark 6 of Section III. Hence, in RFSeq formalism, we have

$$\begin{aligned} &f^{(n)}\left(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n; (t_\kappa)_{\kappa\in[K]} \left|(y_k)_{k=1}^K\right.\right) P(n|(y_k)_{k=1}^K) \\ &= \sum_{\lambda_K\in\Lambda_K} f^{(n)}\left(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n; (t_\kappa)_{\kappa\in[K]} \left|\lambda_K, (y_k)_{k=1}^K\right.\right) P(\lambda_K, n|(y_k)_{k=1}^K) \\ &= \sum_{\lambda_K\in\Lambda_K} P(\lambda_K, n|(y_k)_{k=1}^K)\,((n - \#(\lambda_K))!/n!) \sum_{\alpha_K\in\mathcal{A}(\lambda_K,\{1,...,n\})} f^{(n)}\left(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n; (t_\kappa)_{\kappa\in[K]} \left|(y_k, a_k)_{k=1}^K\right.\right), \end{aligned} \tag{31}$$

where, within the second summation, $(a_k)_{k=1}^K$ is the multiframe assignment that is uniquely determined by $\lambda_K \in \Lambda_K$ and $\alpha_K \in \mathcal{A}(\lambda_K, \{1, ..., n\})$, such that

[50] For $k = 1$, we use the convention that $\bar{\Lambda}_0 = \{\emptyset\}$ and $P(\bar{\lambda}|(y_{k'})_{k'=1}^0) = P(\bar{\lambda}) = 1$ for $\bar{\lambda} = \emptyset$.

[51] In [43], the statement of Theorem 2 (p. 231) contains a misstatement: $L_{\mathrm{NDT}k}(\#(\lambda))$ in eq. (22) (p. 231) must be replaced by $L_{\mathrm{NDT}k}(\#(\lambda))/L_{\mathrm{NDT}(k-1)}(\#(\bar{\lambda}))$ where $\bar{\lambda} = \lambda_{|(k-1)}$ is the unique predecessor of hypothesis λ, and accordingly, $(\bar{\nu} - \hat{\nu}_k)$ in Corollary 2 of [43, p. 233] should be replaced by $(\hat{\nu}_{k-1} - \hat{\nu}_k)$.

$(k, a_k(\alpha_K(\tau))) \in \tau$ for each $\tau \in \lambda_K$ and $\alpha_K(\tau) \in \mathrm{Dom}(a_k)$. While $P(\lambda_K, n|(y_k)_{k=1}^K)$ can be determined through (16) and (17), $f^{(n)}(\cdot;(t_\kappa)_{\kappa\in[K]}|(y_k,a_k)_{k=1}^K)$ can be expressed by the standard Bayes formula, under conditional independence assumption (A4).

The target permutability of Assumption A1 implies that, once $f^{(n)}(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n;(t_\kappa)_{\kappa\in[K]}|(y_k,a_k)_{k=1}^K)$ is evaluated for any particular $(a_k)_{k=1}^K$ determined by an arbitrary (λ_K, n, α), each term of the second summation of (31) can be obtained by appropriate coordinate permutation defined by each $\alpha \in \mathcal{A}(\lambda_K, \{1, ..., n\})$.

With additional independence assumptions (A5 and A6), in RFSeq formalism, (31) can be rewritten as

$$\hat{f}_K^{(n)}((x_i(t_K))_{i=1}^n) = \sum_{\lambda\in\Lambda_K} P(\lambda|(y_k)_{k=1}^K)P(n|\lambda,(y_k)_{k=1}^K) \frac{(n-\#(\lambda))!}{(\hat{\nu}_K)^{n-\#(\lambda)}} \sum_{\alpha\in\mathcal{A}(\lambda,\{1,...,n\})} \left(\prod_{\tau\in\lambda} \hat{f}_K(x_{\alpha(\tau)}(t_K)|\tau)\right) \left(\prod_{\substack{i=1\\ i\notin\mathrm{Im}(\alpha)}}^n \hat{\gamma}_K(x_i)\right) \tag{32}$$

with

$$p(n|\lambda_K, Y_K) = \begin{cases} (L_{\mathrm{NDT}K}(\#(\lambda_K)))^{-1} \frac{(\hat{\nu}_K)^{n-\#(\lambda_K)}}{(n-\#(\lambda_K))!} \cdot \frac{n!}{\nu^n} p_n, & \text{if } n \geq \#(\lambda_K), \\ 0, & \text{otherwise}, \end{cases} \tag{33}$$

where

1) $\hat{f}_K^{(n)}((x_i(t_K))_{i=1}^n) = J^{(n)}((x_i(t_K))_{i=1}^n;t_K|(y_k)_{k=1}^K)$ is the conditional n-JD of the current state $(x_i(t_K))_{i=1}^n$ conditioned by $(y_k)_{k=1}^K$;

2) $L_{\mathrm{NDT}K}(n_D)$ is the likelihood of the hypothesized number $\#(\lambda)$ of all the detected targets in the K frames being n_D, defined by (23);

3) $\hat{\nu}_K = L_{\mathrm{TRK}K}(\emptyset)$ is the expected number of targets remaining undetected throughout the K frames, defined by (21) through $q_{\mathrm{MD}k}$ as $q_{\mathrm{MD}k}(\cdot;\emptyset) = 1 - p_{\mathrm{D}k}(\cdot)$ for each $k = 1, ..., K$;

4) for nonempty track $\tau \in \cup\Lambda_K = \mathcal{T}_K\backslash\{\emptyset\}$, $\hat{f}_K(\cdot|\tau)$ is the track target (current) state PD defined by

$$\hat{f}_K(\xi_K|\tau) = \left(\int_{E^{\#([K])-1}} \left(\prod_{k=1}^K q_{\mathrm{MD}k}(\xi_k;\tau)\right) \gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa\in[K]};(t_\kappa)_{\kappa\in[K]}) \prod_{\kappa\in[K]\backslash\{K\}} \mu(d\xi_\kappa)\right) \Bigg/ \left(\int_{E^{\#([K])}} \left(\prod_{k=1}^K q_{\mathrm{MD}k}(\xi'_k;\tau)\right) \gamma_{\mathrm{TGT}}((\xi'_\kappa)_{\kappa\in[K]};(t_\kappa)_{\kappa\in[K]}) \prod_{\kappa\in[K]} \mu(d\xi'_\kappa)\right); \tag{34}$$

5) $\hat{\gamma}_K$ is the IMD of the targets remaining undetected after the K frames, defined by

$$\hat{\gamma}_K(\xi_K) = \int_{E^{\#([K])-1}} \left(\prod_{k=1}^K (1-p_{\mathrm{D}k}(\xi_k))\right) \gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa\in[K]};(t_\kappa)_{\kappa\in[K]}) \prod_{\kappa\in[K]\backslash\{K\}} \mu(d\xi_\kappa) \tag{35}$$

for each $\xi_K \in E$.

By dropping the most current target-wise extended track-to-measurement likelihood function $q_{\mathrm{MD}K}(\cdot;\tau)$ from both the denominator and the numerator of (34), we have the prediction PD $\bar{f}_K(\cdot|\tau_{|(K-1)})$, which can then be used for the recursive calculation of the track likelihood by (26), without Poisson (A8) or Markovian (A7) assumptions. Under Markov assumption (A7), with or without Poisson assumption (A8), $\hat{f}_K(\cdot|\tau)$ and $\bar{f}_K(\cdot|\tau_{|(K-1)})$ can be obtained through the familiar recursion of (28). Similarly, by taking out $(1-p_{\mathrm{D}K}(\xi_K))$ from the integrand of (35), we have the predicted IMD $\bar{\gamma}_K$ of the undetected targets, while, with Markovian assumption (A7), $\bar{\gamma}_k$ and $\hat{\gamma}_k$ can be obtained recursively by (29).

With Poisson assumption (A8), with or without Markov assumption (A7), we can rewrite (32) as

$$\hat{f}_K^{(n)}((x_i(t_K))_{i=1}^n) = e^{-\hat{\nu}_K} \sum_{\lambda\in\Lambda_K} P(\lambda|(y_k)_{k=1}^K) \sum_{\alpha\in\mathcal{A}(\lambda,\{1,...,n\})} \left(\prod_{\tau\in\lambda} \hat{f}_K(x_{\alpha(\tau)}(t_K)|\tau)\right) \left(\prod_{\substack{i=1\\ i\notin\mathrm{Im}(\alpha)}}^n \hat{\gamma}_K(x_i)\right) \tag{36}$$

and[52]

$$p(n|\lambda_K, Y_K) = \begin{cases} e^{-\hat{\nu}_K} \frac{(\hat{\nu}_K)^{n-\#(\lambda_K)}}{(n-\#(\lambda_K))!}, & \text{if } n \geq \#(\lambda_K), \\ 0, & \text{otherwise}. \end{cases} \tag{37}$$

In FPP formalism, we have

$$\hat{\phi}_K([(x_i(t_K))_{i=1}^n]) \stackrel{\text{def}}{=} \phi([(x_i(t_K))_{i=1}^n];t_K|(y_k)_{k=1}^K) = \hat{f}_K^{(n)}((x_i(t_K))_{i=1}^n),$$

while

$$\hat{\phi}_K(\{(x_i(t_K))_{i=1}^n\}) \stackrel{\text{def}}{=} \phi(\{(x_i(t_K))_{i=1}^n\},t_K|(y_k)_{k=1}^K) = \hat{f}_K^{(n)}((x_i(t_K))_{i=1}^n)$$

in RFSet formalism, both being expressed by (32) (non-Poisson cases) and (36) (Poisson cases), as the a posteriori JMD.

V. RELATION OF MHT TO RFSET-BASED MTT ALGORITHMS

We understand that the relation between MHT and RFSet-based MTT algorithms has been actively discussed recently, e.g., in [36]–[39]. It was even claimed in

[52] Thus, the a posteriori probability distribution of the number of undetected targets is Poisson.

[38] that MHT can be derived from an RFSet-based algorithm. This development is interesting, as we remember that RFSet-based algorithms started to be developed as *correlation-free*[53] algorithms [44]. As mentioned earlier, it is not our objective to conduct the literature survey. In this section, we will state our perspectives of RFSet-based MTT algorithms, from our MHT viewpoints presented in Sections I–IV. We will use what we consider as typical RFSet-based MTT target and sensor models.

A. RFSet Target and Sensor Models

Throughout this section, we basically maintain all the assumptions made so far, i.e., Assumptions A1–A8. Instead of the RFSet-of-stochastic-processes assumption (A1), however, we assume a discrete-time Markov process $(X_k)_{k=1}^K$ $(K \leq \infty)$ on an FICM-QMS $(\mathcal{F}(E), \mathbb{B}(\mathcal{B}), \mathcal{M})$, with the kth state $X_k \in \mathcal{F}(E)$ at time t_k of the kth measurement frame, such that $t_1 \leq t_2 \leq \cdots$, defined by a transition JMD[54] $\phi_{\mathrm{TRN}k}(X_{k+1}|X_k)$, and a Poisson initial state JMD $\phi(X_1) = e^{-\bar{\nu}_1} \prod_{x \in X_1} \bar{\gamma}_1(x)$ with IMD $\bar{\gamma}_1$ and the expected number of targets, $\bar{\nu}_1 = \int_E \bar{\gamma}_1(x)\mu(dx)$, at time t_1.

As a "typical" RFSet-based model, let us assume that the transition JMD $\phi_{\mathrm{TRN}k}(X_{k+1}|X_k)$ (for $t_{k+1} > t_k$) includes a birth–death term, as[55] $\phi_{\mathrm{TRN}k}(\cdot|X_k) = \phi_{\mathrm{TS}k}(\cdot|X_k) \otimes \phi_{\mathrm{B}k}(\cdot)$, which is the convolution of 1) the *survival-transition* JMD $\phi_{\mathrm{TS}k}(\cdot|X_k)$ defined as a conditional *multiple Bernoulli* (MBe) (or *Poisson binomial*) JMD,

$$\phi_{\mathrm{TS}k}(\cdot|\{x_{ki}\}_{i=1}^n) = \phi_{\mathrm{TSBe}k}(\cdot|x_{k1}) \otimes \cdots \otimes \phi_{\mathrm{TSBe}k}(\cdot|x_{kn}) \tag{38}$$

for each $X_k = \{x_{ki}\}_{i=1}^n$, with each conditional *Bernoulli* (Be) JMD $\phi_{\mathrm{TSBe}k}(\cdot|x_{ki})$, defined by, for any $X \in \mathcal{F}(E)$,

$$\phi_{\mathrm{TSBe}k}(X|x_{ki}) = \begin{cases} 1 - p_{\mathrm{S}k}(x_{ki}), & \text{if } X = \emptyset, \\ f_{\mathrm{T}k}(x_{k+1}|x_{ki})p_{\mathrm{S}k}(x_{ki}), & \text{if } X = \{x_{k+1}\}, \\ 0, & \text{if } \#(X) > 1 \end{cases} \tag{39}$$

assuming target-wise independent, target-state dependent *survival probability* $p_{\mathrm{S}k} : E \to [0,1]$, and a discrete-time STPD[56] $f_{\mathrm{T}k}(x_{k+1}|x_k) = f_{\mathrm{TRN}}(x_{k+1}|x_k; t_{k+1} - t_k, t_k)$ that is target-wise independent, and 2) a *Poisson birth* JMD $\phi_{\mathrm{B}k}(X) = e^{-\nu_{\mathrm{B}k}} \prod_{x \in X} \gamma_{\mathrm{B}k}(x)$, defined through the IMD $\gamma_{\mathrm{B}k}$ with $\nu_{\mathrm{B}k} = \int_E \gamma_{\mathrm{B}k}(x)\mu(dx)$, all generally depending on $k = 1, 2, \ldots$.

In Section III-B, we modeled the kth frame as an RFSeq $(y_{kj})_{j=1}^{m_k}$ with each measurement having a unique label. The likelihood function (11) is, however, permutable with respect to both measurements $(y_{kj})_{j=1}^{m_k}$ and the target states $(x_i(t_k))_{i=1}^n$. Therefore, we can consider each measurement frame as an RFSet $Y_k = \{y_{kj}\}_{j=1}^{m_k}$ in an LCHC2 measure space[57] $(E_{\mathrm{M}k}, \mathcal{B}_{\mathrm{M}k}, \mu_{\mathrm{M}k})$, having a conditional JMD $\phi_{\mathrm{M}k}(\cdot|X_k) = \phi_{\mathrm{MD}k}(\cdot|X_k) \otimes \phi_{\mathrm{FA}k}(\cdot)$ that is the convolution of 1) conditional JMD $\phi_{\mathrm{MD}k}(\cdot|X_k)$ of target detections and 2) Poisson JMD $\phi_{\mathrm{FA}k}(Y_{\mathrm{FA}k}) = e^{-\nu_{\mathrm{FA}k}} \prod_{\eta \in Y_{\mathrm{FA}k}} \gamma_{\mathrm{FA}k}(\eta)$ with $\nu_{\mathrm{FA}k} = \int_{E_{\mathrm{M}k}} \gamma_{\mathrm{FA}k}(\eta)\mu_{\mathrm{M}k}(d\eta)$ for the set of false alarms.

With the independent detection assumption (A6), the target detections are modeled by a conditional MBe JMD,

$$\phi_{\mathrm{MD}k}(\cdot|\{x_{ki}\}_{i=1}^n) = \phi_{\mathrm{MDBe}k}(\cdot|x_{k1}) \otimes \cdots \otimes \phi_{\mathrm{MDBe}k}(\cdot|x_{kn}) \tag{40}$$

for $X_k = \{x_{ki}\}_{i=1}^n$, with each conditional Be JMD, $\phi_{\mathrm{MDBe}k}(\cdot|x_{ki})$, defined as, for any $Y \in \mathcal{F}(E_{\mathrm{M}k})$,

$$\phi_{\mathrm{MDBe}k}(Y|x_{ki}) = \begin{cases} 1 - p_{\mathrm{D}k}(x_{ki}), & \text{if } Y = \emptyset, \\ p_{\mathrm{M}k}(y|x_{ki})p_{\mathrm{D}k}(x_{ki}), & \text{if } Y = \{y\}, \\ 0, & \text{if } \#(Y) > 1. \end{cases} \tag{41}$$

For $k = 1, 2, \ldots$, let $\bar{\phi}_k(X_k)$ and $\hat{\phi}_k(X_k)$ be the predicted and the updated state JMD, i.e.,

$$\begin{cases} \bar{\phi}_k(X_k)\mathcal{M}(dX_k) = \begin{cases} \mathrm{Prob}\{X_k \in dX_k\}, & \text{if } k = 1, \\ \mathrm{Prob}\{X_k \in dX_k|(Y_{k'})_{k'=1}^{k-1}\}, & \text{if } k > 1, \end{cases} \\ \hat{\phi}_k(X_k)\mathcal{M}(dX_k) = \mathrm{Prob}\{X_k \in dX_k|(Y_{k'})_{k'=1}^{k}\}. \end{cases} \tag{42}$$

Then, as shown in Appendix B, with the target and sensor models described earlier, we can prove that predicted JMD $\bar{\phi}_k(X_k)$ and updated JMD $\hat{\phi}_k(X_k)$ can be expressed as convolutions $\bar{\phi}_k = \bar{\phi}_{\mathrm{D}k} \otimes \bar{\phi}_{\mathrm{U}k}$ and $\hat{\phi}_k = \hat{\phi}_{\mathrm{D}k} \otimes \hat{\phi}_{\mathrm{U}k}$, respectively, where 1) conditional JMD $\bar{\phi}_{\mathrm{D}k}$ or $\hat{\phi}_{\mathrm{D}k}$ for the detected targets is written as[58]

$$\begin{cases} \bar{\phi}_{\mathrm{D}k}(X) = \displaystyle\sum_{\substack{\bar{\lambda} \in \bar{\Lambda}_k \\ \#(\bar{\lambda}) = \#(X)}} \bar{p}_k(\bar{\lambda}) \sum_{\bar{\alpha} \in \mathcal{A}(\bar{\lambda}, X)} \left(\prod_{\bar{\tau} \in \bar{\lambda}} \bar{f}_k(\bar{\alpha}(\bar{\tau})|\bar{\tau}) \right) \\ \hat{\phi}_{\mathrm{D}k}(X) = \displaystyle\sum_{\substack{\hat{\lambda} \in \hat{\Lambda}_k \\ \#(\hat{\lambda}) = \#(X)}} \hat{p}_k(\hat{\lambda}) \sum_{\hat{\alpha} \in \mathcal{A}(\hat{\lambda}, X)} \left(\prod_{\hat{\tau} \in \hat{\lambda}} \hat{f}_k(\hat{\alpha}(\hat{\tau})|\hat{\tau}) \right) \end{cases} \tag{43}$$

and 2) conditional JMDs $\bar{\phi}_{\mathrm{U}k}$ and $\hat{\phi}_{\mathrm{U}k}$ for the undetected targets are Poisson JMD as

$$\begin{cases} \bar{\phi}_{\mathrm{U}k}(X) = e^{-\bar{\nu}_k} \displaystyle\prod_{x \in X} \bar{\gamma}_k(x) \text{ with } \bar{\nu}_k = \int_E \bar{\gamma}_k(\xi)\mu(d\xi) \\ \hat{\phi}_{\mathrm{U}k}(X) = e^{-\hat{\nu}_k} \displaystyle\prod_{x \in X} \hat{\gamma}_k(x) \text{ with } \hat{\nu}_k = \int_E \hat{\gamma}_k(\xi)\mu(d\xi) \end{cases} \tag{44}$$

for each $X \in \mathcal{F}(E)$, with undetected target IMD $\bar{\gamma}_k$ and $\hat{\gamma}_k$, where

---

[53] We understand that the "correlation" is a "traditional" U.S. Navy terminology for data association.

[54] We assume that the transition JMD $\phi_{\mathrm{TRN}k}$ is only defined for $t_{k+1} > t_k$, and that $X_{k+1} = X_k$ if $t_{k+1} = t_k$.

[55] We assume, for Section V, that any diagonal set D_n in E^n has zero product measure μ^n, i.e., $\mu^n(D_n) = 0$.

[56] f_{TRN} is the continuous time STPD of Assumption A7, assuming $\Delta s = t_{k+1} - t_k > 0$.

[57] We also assume, for Section V, that any diagonal set $D_{\mathrm{M}km}$ in each product measurement space $E_{\mathrm{M}k}^m$ has zero product measure $\mu_{\mathrm{M}k}^m$, i.e., $\mu_{\mathrm{M}k}^m(D_{\mathrm{M}km}) = 0$ for each m.

[58] We understand that each JMD of (43), a probability-weighted sum of the symmetrized asymmetric PD products, is called generalized multi-Bernoulli (GMBe) in [33].

1) $\bar{\Lambda}_k$ and $\hat{\Lambda}_k$ are sets of association hypotheses, each hypothesis being as a collection of nonempty tracks, each of which is a subset of $\bigcup_{k'=1}^{k-1}\{k'\}\times Y_{k'}$ or $\bigcup_{k'=1}^{k}\{k'\}\times Y_{k'}$;

2) $(\bar{p}_k(\bar{\lambda}))_{\bar{\lambda}\in\bar{\Lambda}_k}$ and $(\hat{p}_k(\hat{\lambda}))_{\hat{\lambda}\in\hat{\Lambda}_k}$ are probabilistic weights;

3) each nonempty track, $\bar{\tau}\in\bigcup\bar{\Lambda}_k$ or $\hat{\tau}\in\cup\hat{\Lambda}_k$, is accompanied by track PD, $\bar{f}_k(\cdot|\bar{\tau})$ or $\hat{f}_k(\cdot|\hat{\tau})$, on the state space E.

Thus, the predicted $\bar{\phi}_k$ is represented by *parameters,* $((\bar{p}_k(\bar{\lambda}))_{\bar{\lambda}\in\bar{\Lambda}_k},(\bar{f}_k(\cdot|\bar{\tau}))_{\bar{\tau}\in\cup\bar{\Lambda}_k},\bar{\gamma}_k)$, the updated JMD $\hat{\phi}_k$ by $((\hat{p}_k(\hat{\lambda}))_{\hat{\lambda}\in\hat{\Lambda}_k},(\hat{f}_k(\cdot|\hat{\tau}))_{\hat{\tau}\in\cup\hat{\Lambda}_k},\hat{\gamma}_k)$, through (43) and (44). Those parameters, which we may call *sufficient statistics*,[59] are recursively calculated as shown in the next two sections.

B. RFSet Filtering Update

The conditional JMD is updated, from $\bar{\phi}_k$ to $\hat{\phi}_k$, by the *Bayes update formula*, as

$$\hat{\phi}_k(X_k)=\frac{\phi_{\mathrm{M}k}(Y_k|X_k)\bar{\phi}_k(X_k)}{\int_{\mathcal{F}(E)}\phi_{\mathrm{M}k}(Y_k|X)\bar{\phi}_k(X)\mathcal{M}(dX)}. \tag{45}$$

As proven in Appendix B, the updated parameter $(\hat{p}(\hat{\lambda}))_{\hat{\lambda}\in\hat{\Lambda}_k}$ of the conditional JMD $\hat{\phi}_k(X_k)$ is obtained from the parameters $((\bar{p}(\bar{\lambda}))_{\bar{\lambda}\in\bar{\Lambda}_k},(\bar{f}_k(\cdot|\bar{\tau}))_{\bar{\tau}\in\cup\bar{\Lambda}_k},\bar{\gamma}_k)$ of the predicted JMD $\bar{\phi}_k(X_k)$ defined in (43) and (44), and from the sensor model defined by (40) and (41) with parameters $(p_{\mathrm{D}k},p_{\mathrm{M}k},\gamma_{\mathrm{FA}k})$, as the Poisson version of Reid form,

$$\hat{p}_k(\hat{\lambda})=C_{\mathrm{R}k}'^{-1}\bar{p}_k(\bar{\lambda})\left(\prod_{\substack{\hat{\tau}\in\hat{\lambda}\\ \hat{\tau}_{|(k-1)}\neq\emptyset}}L_{\mathrm{MD}k}(\hat{\tau})\right)\left(\prod_{\substack{y\in Y_k\\ \{(k,y)\}\in\hat{\lambda}}}\gamma_{\mathrm{MND}k}(y)\right)\left(\prod_{\substack{y\in Y_k\\ (k,y)\notin\cup\hat{\lambda}}}\gamma_{\mathrm{FA}k}(y)\right) \tag{46}$$

for each updated hypothesis $\hat{\lambda}$ in the set $\hat{\Lambda}_k$ that is defined as[60]

$$\hat{\Lambda}_k=\left\{\lambda_{\mathrm{OLD}k}(\bar{\lambda},\bar{a})\cup\lambda_{\mathrm{NEW}k}(Y_{\mathrm{N}k})\;\middle|\;\begin{array}{l}\bar{\lambda}\in\bar{\Lambda}_k,\\ \bar{a}\in\bar{\mathcal{A}}(\bar{\lambda},Y_k)\text{ and}\\ Y_{\mathrm{N}k}\subseteq Y_k\backslash\mathrm{Im}(\bar{a})\end{array}\right\} \tag{47}$$

with

$$\begin{cases}\lambda_{\mathrm{OLD}k}(\bar{\lambda},\bar{a})=\{\bar{\tau}\cup\{(k,\bar{a}(\bar{\tau}))\}|\bar{\tau}\in\mathrm{Dom}(\bar{a})\}\\ \qquad\qquad\qquad\cup(\bar{\lambda}\backslash\mathrm{Dom}(\bar{a})),\\ \lambda_{\mathrm{NEW}k}(Y_{\mathrm{N}k})=\{k\}\times Y_{\mathrm{N}k}.\end{cases} \tag{48}$$

Equation (46) can be obtained by applying Poisson assumption (A8) to Reid form (30). The right-hand side of (46) consists of 1) Poisson version $C'_{\mathrm{R}k}$ of Reid constant, 2) the prior probability $\bar{p}_k(\bar{\lambda})$ of the unique predecessor $\bar{\lambda}=\hat{\lambda}_{|(k-1)}$ of each $\hat{\lambda}$ in $\bar{\Lambda}_k$, 3) the extended track-to-measurement likelihood $L_{\mathrm{MD}k}(\hat{\tau})$ defined in (27) with $\tau_{|k}=\hat{\tau}$, 4) the newly detected target measurement IMD $\gamma_{\mathrm{MND}k}(y)$ defined in (27), and 5) the false alarm IMD $\gamma_{\mathrm{FA}k}(y)$ of the Poisson false alarm JMD $\phi_{\mathrm{FA}k}$. The rest of the parameters for the updated JMD $\hat{\phi}_k(X_k)$ are updated to $(\hat{f}_k(\cdot|\hat{\tau}))_{\hat{\tau}\in\cup\hat{\Lambda}_k}$ and $\hat{\gamma}_k$ in the first equations of (28) and (29) from $(\bar{f}_k(\cdot|\bar{\tau}))_{\bar{\tau}\in\cup\bar{\Lambda}_k}$ and $\bar{\gamma}_k$, respectively.

We should immediately note that (46) is the Reid form for evaluating association hypotheses recursively, shown in [6, eq. 16, p. 848], and that (47) expresses the recursive hypothesis expansion that corresponds almost exactly to the illustration in [6, Fig. 2, p. 846].

We should also note that, as seen in (48), each track (and hence each hypothesis) is defined through the value $y\in Y_k$ of each measurement in each frame Y_k (that is defined as an RFSet), not through the measurement index, as having been done in Section III-C. Since we assume that the diagonal set $D_{\mathrm{M}km}$ in each order-m product measurement space $(E_{\mathrm{M}k})^m$ has zero product measure, we maintain the no-merged-or-split-measurement assumption (A2). Since every measurement frame is *data* or *observation*, we can reorder (or relabel) measurements in the RFSet-modeled frame in any arbitrary way (as we wish), and yet we obtain the same permutable target state likelihoods. For this reason, this difference in the definition of hypothesis is inconsequential, and in that sense, the hypothesis evaluation (46) is exactly the same as Reid form (30), except for the Poisson assumption (A8).

However, there is an important difference that we should note: In Section III-C, we define the hypotheses as possible realizations of an RFSet, which we call "association," so that their evaluation is to calculate the conditional probabilities, while the hypotheses in this section appear only as *parameters* to define *weights* in (43). As shown in Appendix B, the fact that the set of weights, $(\hat{p}_k(\hat{\lambda}))_{\hat{\lambda}\in\hat{\Lambda}_k}$, is indeed in a *unit simplex* is a consequence of the evaluation of the denominator of the right-hand side of the update equation (45), i.e., the normalizing constant, under the induction assumption that $(\bar{p}_k(\bar{\lambda}))_{\bar{\lambda}\in\bar{\Lambda}_k}$ is a set of probabilistic weights.

C. RFSet Filtering Extrapolation

Since the Markov process $(X_1,X_2,...)$ on measure space $(\mathcal{F}(E),\mathbb{B}(B),\mathcal{M})$ is defined in Section V-A with the transition JMD $\phi_{\mathrm{TRN}k}$, for each $t_k<t_{k+1}$, the *predicted* JMD $\bar{\phi}_{k+1}$ of X_{k+1} conditioned on $(Y_{k'})_{k'=1}^{k}$ is obtained by extrapolating the *previously updated* JMD $\hat{\phi}_k$, as

$$\bar{\phi}_{k+1}(X_{k+1})=\int_{\mathcal{F}(E)}\phi_{\mathrm{TRN}k}(X_{k+1}|X_k)\hat{\phi}_k(X_k)\mathcal{M}(dX_k). \tag{49}$$

[59]These *sufficient statistics* are not finite dimensional unless the track PDs and undetected target IMDs have finite-dimensional representations, which, most likely, exist only approximately.

[60]Using convention $\bar{\mathcal{A}}(\emptyset,Y)=\{\theta\}$ with $\mathrm{Dom}(\theta)=\mathrm{Im}(\theta)=\emptyset$.

As shown in Appendix B, the representation (parameters) $((\hat{p}_k(\hat{\lambda}))_{\hat{\lambda}\in\hat{\Lambda}_k}, (\hat{f}_k(\cdot|\hat{\tau}))_{\hat{\tau}\in\cup\hat{\Lambda}_k}, \hat{\gamma}_k)$ of the conditional JMD $\hat{\phi}_k$ is extrapolated to $((\bar{p}_{k+1}(\bar{\lambda}))_{\bar{\lambda}\in\bar{\Lambda}_{k+1}}, (\bar{f}_{k+1}(\cdot|\bar{\tau}))_{\bar{\tau}\in\cup\bar{\Lambda}_{k+1}}, \bar{\gamma}_{k+1})$ for the conditional JMD $\bar{\phi}_{k+1}$ as

1) the extrapolated probabilistic weights

$$\bar{p}_{k+1}(\bar{\lambda}) = \sum_{\substack{\hat{\lambda}\in\hat{\Lambda}_k \\ \hat{\lambda}\supseteq\bar{\lambda}}} \hat{p}_k(\hat{\lambda}) \left(\prod_{\bar{\tau}\in\bar{\lambda}} P_{\mathrm{S}k}(\bar{\tau})\right) \left(\prod_{\hat{\tau}\in\hat{\lambda}\setminus\bar{\lambda}} (1 - P_{\mathrm{S}k}(\hat{\tau}))\right) \tag{50}$$

for each *predicted* hypothesis $\bar{\lambda}$ in

$$\bar{\Lambda}_{k+1} = \bigcup_{\hat{\lambda}\in\hat{\Lambda}_k} \mathcal{F}(\hat{\lambda}) = \{\bar{\lambda} \subseteq \hat{\lambda} | \hat{\lambda} \in \hat{\Lambda}_k\} \tag{51}$$

with the *track survival probability* $P_{\mathrm{S}k}(\hat{\tau})$ defined by

$$P_{\mathrm{S}k}(\hat{\tau}) = \int_E p_{\mathrm{S}k}(x_k)\hat{f}_k(x_k|\hat{\tau})\mu(dx_k) \tag{52}$$

for each $\hat{\tau} \in \cup\hat{\Lambda}_k$;

2) the extrapolated target state PD $\bar{f}_{k+1}(\cdot|\bar{\tau})$ for any surviving track $\bar{\tau} \in \cup\bar{\Lambda}_{k+1} = \cup\hat{\Lambda}_k$ as

$$\bar{f}_{k+1}(x_{k+1}|\bar{\tau}) = P_{\mathrm{S}k}(\bar{\tau})^{-1} \int_E f_{\mathrm{T}k}(x_{k+1}|x_k)p_{\mathrm{S}k}(x_k)\hat{f}_k(x_k|\bar{\tau})\mu(dx_k); \tag{53}$$

3) the predicted IMD of the undetected targets that are either surviving or newly born as

$$\bar{\gamma}_{k+1}(x_{k+1}) = \gamma_{\mathrm{B}k}(x_{k+1}) + \int_E f_{\mathrm{T}k}(x_{k+1}|x_k)p_{\mathrm{S}k}(x_k)\hat{\gamma}_k(x_k)\mu(dx_k). \tag{54}$$

Death of any previously detected target is *hypothesized* by $\hat{\tau} \in \hat{\lambda}\setminus\bar{\lambda}$ through an *updated hypothesis* $\hat{\lambda} \in \hat{\Lambda}_k$ and a *predicted hypothesis* $\bar{\lambda} \in \bar{\Lambda}_{k+1}$ such that $\bar{\lambda} \subseteq \hat{\lambda}$, which may break the tree structure of the hypotheses described in Section IV-A. The extrapolation, as described earlier, should take place only when $t_{k+1} > t_k$. In case $t_{k+1} = t_k$, we should let $\bar{\Lambda}_{k+1} = \hat{\Lambda}_k$, $\bar{p}_{k+1} = \hat{p}_k$, and $\bar{f}_{k+1}(\cdot|\tau) = \hat{f}_k(\cdot|\tau)$ for any $\tau \in \cup\bar{\Lambda}_{k+1} = \cup\hat{\Lambda}_k$, to avoid any unwanted "jumps."

D. Track Continuity

At least to the authors of this paper, it is rather surprising to see that, when we eliminate the birth–death model from the state transition described in Section V-A, i.e., with $\gamma_{\mathrm{B}k} \equiv 0$ and $p_{\mathrm{S}k} \equiv 1$, a purely stochastic-process-on-FICM RFSet model of Section V-A regenerates Reid form (30), exactly by (46), and the state estimation of (36) by (43) and (44), in Section V-B, with Poisson assumption[61] (A8). We should remember that Reid form (30) was derived in Section IV, 1) with target model of an RFSeq (or FPP or RFSet) of stochastic processes and 2) with data association hypotheses defined as the set of all the possible realizations of a random element, called "data association." This validates a claim made in [38]: "The MHT can be derived from an RFSet MTT algorithm." We should note that, however, in the RFSet-based algorithm as described in this section, the hypotheses appear only as "indices," with which probabilistic weights over the GMBer terms of (43) are expressed, not as the probabilistic evaluation of possible realizations of a random element as we defined as *association*.

The target transition model in RFSet formalism, described in Section V-A, does have an appearance that targets may *exchange* their states among them because the state transition is expressed as RFSet-state-to-RFSet transition. In fact, our motivation of using a set of stochastic processes, rather than a single stochastic process on an FICM, is to avoid the *target switches* of this kind, *through the most obvious and explicit way*. To the best of our knowledge, there are at least two known efforts to avoid these target switches. These two are based on two quite different approaches: 1) the introduction of the *labeled* RFSets in [33] and 2) the use of *trajectory states* in [39]. The former adds an extra state element, called a *label* to each single target state, to prevent target exchange during the extrapolation step, according to our interpretation. The latter extends[62] the individual target state to the consecutive series, from the target's birth to the current state, again to prevent the "target exchange," which is characterized as the maintenance of the track continuity in [39].

After having seen the re-creation of Reid form (30) by (46) in this section, sharing the same conclusion by [38], as we understand, we are not quite sure now if all those *precautions* to maintain track continuity are really necessary, or if they are mere *precauzione inutile*. It seems to us, at this point, that the track continuity issues are implicitly taken care of by the use of the concept of the tracks (and hypotheses), which is actually a core concept of the MHT.

VI. CONCLUSIONS

We presented three mathematical formalisms, i.e., RFSeq, FPP, and RFSet formalisms, which provide us with theoretical foundations for MTT problems in general, and the basis for MHT in particular, when generally multiple sensors provide target detections with uncertain origins. MHT, as a concept for providing solutions to MTT problems, has been studied over the last 40 years extensively, as described in [8]. In this paper, using a general class of target and sensor models, we revisited the

[61] The RFSet filtering shown in Sections V-A to V-C should be easily extended to non-Poisson cases.

[62] By this extension, the trajectory-state estimation may be considered as a variable-time-interval smoothing, i.e., estimation of each target trajectory from the moment of the birth to the current state or to the time when the target is killed.

generation and evaluation of data association hypotheses, and provided some new perspectives, by presenting them using, side by side, the three different formalisms. Those three may appear quite differently on the surface but are almost equivalent to each other except for subtle differences, e.g., those caused by *repeated elements* that are allowed in an RFSeq or an FPP but not in an RFSet.

Based on target models that use the concept of a set of stochastic processes, rather than a single stochastic process on FM (RFSeq), FCM (FPP), or FICM (RFSet), and sensor models with RFSeq outputs, we explicitly defined data association as a discrete-valued random element. We called all of its possible realizations data association hypotheses, as defined in Section III-C. The two well-known hypothesis evaluation forms, Morefield form and Reid form, were then derived in Section IV-B and IV-C, with a gradual introduction of commonly used assumptions, (A1–A8). Although hypotheses can be defined without the independence assumptions, the familiar hypothesis–track structure of MHT appeared only after the independence assumptions (A5 and A6) were introduced. The consequences of the other assumptions were rather predictable: the separation of evaluation of the probabilities of the number of newly detected targets and that of data association hypotheses was obtained by the Poisson assumption (A8), and the familiar extrapolation–update recursion structure appeared with the introduction of the Markov assumption (A7).

In Section V, we stated our perspectives on the recently developed RFSet-based MTT algorithms, for which intimate relations to MHT were claimed. We observed that not only the MHT hypothesis–track structure emerged as described in [36], but also the exact Reid form was *surprisingly* re-created from a *pure* RFSet model, which we think is consistent with the claim made by Brekke and Chitre [38]. Our conjecture on the reason for this reappearance of Reid form is the use of the hypothesis/track structure that forces the desired *continuity*, well within the context of MHT.

The MTT algorithm developments based on RFSet formalism, also known as FISST formalism [25]–[27], were relatively new, compared with the long history of FPP formalism, which is claimed to have started with [24]. The authors hope some *old* wisdom may benefit our efforts in advancing MTT technologies further.

APPENDIX A: DERIVATION OF HYPOTHESIS EVALUATION EQUATIONS

Under Assumptions A1–A3, for any cumulative frame $(y_k)_{k=1}^K = ((y_{kj})_{j=1}^{m_k})_{k=1}^K$, it follows from (16) that data association $\lambda_K \in \Lambda_K$ on $(y_k)_{k=1}^K$ can be evaluated as

$$\begin{aligned} P(\lambda_K|(y_k)_{k=1}^K) &= \sum_{n=\#(\lambda_K)}^{\infty} P(\lambda_K, n|(y_k)_{k=1}^K) \\ &= P((y_k)_{k=1}^K)^{-1} \sum_{n=\#(\lambda_K)}^{\infty} \frac{n!}{(n-\#(\lambda))!} P((y_k, a_k)_{k=1}^K, n), \end{aligned} \tag{A.1}$$

where n is the number of targets and $(a_k)_{k=1}^K \in \prod_{k=1}^K \bar{\mathcal{A}}(\{1, \ldots, n\}, \{1, \ldots, m_k\})$ is, for a given (λ_K, n), any one of the $n!/(n - \#(\lambda))!$ multiframe target assignment hypotheses that support λ_K, in the sense that the pair $(\lambda_K, (a_k)_{k=1}^K)$ satisfies (13).

In the RFSeq formalism, with additional assumptions (A4–A6), substitute (18) into (17), and apply $f^{(n)}(((x_i(t_\kappa))_{\kappa\in[K]})_{i=1}^n; (t_\kappa)_{\kappa\in[K]}) = \prod_{i=1}^n f_{\mathrm{TGT}}((x_i(t_\kappa))_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]})$.

Then, we have

$$\begin{aligned} P((y_k, a_k)_{k=1}^K, n) &= P(((y_{kj})_{j=1}^{m_k}, a_k)_{k=1}^K|n)p_n \\ &= p_n \nu^{-n} \left(\prod_{k=1}^K \frac{L_{\mathrm{FA}k}(\{1,\ldots,m_k\}\setminus \mathrm{Im}(a_k))}{m_k!} \right) \\ &\int_{E^{\#([K])n}} \left(\prod_{k=1}^K \left(\prod_{i\in \mathrm{Dom}(a_k)} p_{\mathrm{M}k}(y_{ka_k(i)}|x_i(t_k)) p_{\mathrm{D}k}(x_i(t_k)) \right) \right. \\ &\left. \left(\prod_{\substack{i=1 \\ i\notin \mathrm{Dom}(a_k)}}^n (1 - p_{\mathrm{D}k}(x_i(t_k))) \right) \right) \\ &\prod_{i=1}^n \gamma_{\mathrm{TGT}}((x_i(t_\kappa))_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]}) \mu^{\#([K])}((dx_i(t_\kappa))_{\kappa\in[K]}), \end{aligned} \tag{A.2}$$

where the a priori target state IMD, over $(t_\kappa)_{\kappa\in(K)}$, with a priori mean $\nu = \sum_{n=1}^\infty n p_n < \infty$ of the number of targets, is $\gamma_{\mathrm{TGT}}((x_i(t_\kappa))_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]}) = \nu f_{\mathrm{TGT}}((x_i(t_\kappa))_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]})$, and $L_{\mathrm{FA}k}(I_{\mathrm{FA}k})$ is the false alarm likelihood defined by (19).

Let the integral in (A.2) over the set $E^{\#([K])n}$ be $L_{\mathrm{TGT}K}((y_k, a_k)_{k=1}^K; n)$. Then, when $(a_k)_{k=1}^K$ supports λ_K (i.e., for which (13) holds) with $\#(\lambda) \leq n$, we have

$$\begin{aligned} &L_{\mathrm{TGT}K}((y_k, a_k)_{k=1}^K; n) \\ &= \prod_{i=1}^n \int_{E^{\#([K])}} \left(\prod_{\substack{k=1 \\ i\in \mathrm{Dom}(a_k)}}^K p_{\mathrm{M}k}(y_{ka_k(i)}|x_i(t_k)) p_{\mathrm{D}k}(x_i(t_k)) \right) \\ &\quad \left(\prod_{\substack{k=1 \\ i\notin \mathrm{Dom}(a_k)}}^K (1 - p_{\mathrm{D}k}(x_i(t_k))) \right) \\ &\quad \gamma_{\mathrm{TGT}}((x_i(t_\kappa))_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]}) \\ &\quad \mu^{\#([K])}((dx_i(t_\kappa))_{\kappa\in[K]}) \\ &= \left(\prod_{\tau\in\lambda_K} \int_{E^{\#([K])}} \left(\prod_{k=1}^K q_{\mathrm{MD}k}(\xi_k; \tau) \right) \gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]}) \right. \\ &\quad \left. \mu^{\#([K])}((d\xi_\kappa)_{\kappa\in[K]}) \right) \\ &\quad \left(\int_{E^{\#([K])}} \left(\prod_{k=1}^K (1 - p_{\mathrm{D}k}(\xi_k)) \right) \right. \\ &\quad \left. \gamma_{\mathrm{TGT}}((\xi_\kappa)_{\kappa\in[K]}; (t_\kappa)_{\kappa\in[K]}) \mu^{\#([K])}((d\xi_\kappa))_{\kappa\in[K]} \right)^{n-\#(\lambda_K)} \\ &= \left(\prod_{\tau\in\lambda_K} L_{\mathrm{TRK}K}(\tau) \right) L_{\mathrm{TRK}K}(\emptyset)^{n-\#(\lambda_K)} \\ &= \left(\prod_{\tau\in\lambda_K} L_{\mathrm{TRK}K}(\tau) \right) (\hat{\nu}_K)^{n-\#(\lambda_K)}, \end{aligned} \tag{A.3}$$

where $L_{\mathrm{TRK}K}(\tau)$ is the track likelihood defined by (21) with (22), and $\hat{\nu}_K = L_{\mathrm{TRK}K}(\emptyset)$ is the a posteriori expected number of targets that remain undetected over the K frames. Then, substituting (A.3) into (A.2), and substituting (A.2) into (A.1), we obtain Morefield form (20), which completes the derivation of the hypothesis evaluation in RFSeq formalism.

In FPP or RFSet formalism, the integral in (17), and the n-PD $f^{(n)}$ in its integrand, should be replaced by the set integral defined in (2) or (4), and by the JMD $\phi(\cdot; (t_\kappa)_{\kappa \in (K)})$. The constant $n!$, included in the JMD in either FPP or RFSet formalism, is cancelled out by $1/n!$ included in the definition of the set integral in (2) or (4), resulting in the same expression as the one by (A.3), and hence, we have the same hypothesis evaluation equation, i.e., Morefield form (20).

The recursive form hypothesis evaluation equation, i.e., Reid form (30), can be readily derived from its batch-processing counterpart, Morefield form (20), with or without Markov assumption (A7), and vice versa (i.e., from Reid form to Morefield form).

APPENDIX B : DERIVATION OF SOLUTION TO AN RFSET FILTER

This appendix provides a proof to our assertion that, under the assumptions made in Section V-A, the conditional JMDs, $\bar{\phi}_k$ and $\hat{\phi}_k$, defined by (42), can be written as the convolutions of the GMBer JMDs, $\bar{\phi}_{\mathrm{D}k}$ and $\hat{\phi}_{\mathrm{D}k}$, defined by (43), and the Poisson JMDs, $\bar{\phi}_{\mathrm{U}k}$ and $\hat{\phi}_{\mathrm{U}k}$, of (44), respectively. Our proof is one by mathematical induction, giving a proof to all the update and extrapolation equations (46)–(48) and (50)–(54), together at the same time.

For $k = 1$, we have $\bar{\Lambda}_1 = \{\emptyset\}$, and $\bar{\phi}_{U1} = \bar{\phi}_1$ is the Poisson initial-state JMD. For any $k \geq 1$, let assume, as the induction assumption, that $\bar{\phi}_k$ is the convolution $\bar{\phi}_k = \bar{\phi}_{\mathrm{D}k} \otimes \bar{\phi}_{\mathrm{U}k}$ of the GMBer $\bar{\phi}_{\mathrm{D}k}$ given in (43) and Poisson JMD $\bar{\phi}_{\mathrm{U}k}$ in (44). This convolution can be rewritten as

$$\bar{\phi}_k(\{x_i\}_{i=1}^n) = e^{-\bar{\nu}_k} \sum_{\bar{\lambda} \in \bar{\Lambda}_k} \bar{p}_k(\bar{\lambda}) \sum_{\bar{\alpha} \in \mathcal{A}(\bar{\lambda}, \{1,\ldots,n\})} \left(\prod_{\bar{\tau} \in \bar{\lambda}} \bar{f}_k(x_{\bar{\alpha}(\bar{\tau})} | \bar{\tau}) \right) \left(\prod_{\substack{i=1 \\ i \notin \mathrm{Im}(\bar{\alpha})}}^{n} \bar{\gamma}_k(x_i) \right). \tag{B.1}$$

The JMD likelihood function $\phi_{\mathrm{M}k}$ for frame $Y_k = \{y_{kj}\}_{j=1}^{m_k}$, defined as the convolution of MBe JMD $\phi_{\mathrm{MD}k}$, defined by (40) and (41), and of Poisson JMD $\phi_{\mathrm{FA}k}$, can be written as

$$\phi_{\mathrm{M}k}(\{y_{kj}\}_{j=1}^{m_k} | \{x_i(t_k)\}_{i=1}^n) = e^{-\nu_{\mathrm{FA}k}} \sum_{a \in \bar{\mathcal{A}}(\{1,\ldots,n\},\{1,\ldots,m_k\})} \left(\prod_{i \in \mathrm{Dom}(a)} p_{\mathrm{M}k}(y_{ka(i)} | x_i(t_k)) p_{\mathrm{D}k}(x_i(t_k)) \right) \left(\prod_{\substack{i=1 \\ i \notin \mathrm{Dom}(a)}}^{n} (1 - p_{\mathrm{D}k}(x_i(t_k))) \right) \left(\prod_{\substack{j=1 \\ j \notin \mathrm{Im}(a)}}^{m_k} \gamma_{\mathrm{FA}k}(y_{kj}) \right). \tag{B.2}$$

It follows from (B.1) and (B.2) that

$$\begin{aligned} &\phi_{\mathrm{M}k}(\{y_{kj}\}_{j=1}^{m_k} | \{x_i\}_{i=1}^n) \bar{\phi}_k(\{x_i\}_{i=1}^n) \\ &= e^{-\nu_{\mathrm{FA}k} - \bar{\nu}_k} \sum_{\substack{\bar{\lambda} \in \bar{\Lambda}_k \\ \#(\bar{\lambda}) \leq n}} \bar{p}_k(\bar{\lambda}) \sum_{\bar{\alpha} \in \mathcal{A}(\bar{\lambda},\{1,\ldots,n\})} \sum_{a \in \bar{\mathcal{A}}(\{1,\ldots,n\},\{1,\ldots,m_k\})} \\ &\left(\prod_{i \in \mathrm{Im}(\bar{\alpha}) \cap \mathrm{Dom}(a)} p_{\mathrm{M}k}(y_{ka(i)} | x_i) p_{\mathrm{D}k}(x_i) \bar{f}_k(x_i | \bar{\alpha}^{-1}(i)) \right) \\ &\left(\prod_{i \in \mathrm{Im}(\bar{\alpha}) \setminus \mathrm{Dom}(a)} (1 - p_{\mathrm{D}k}(x_i)) \bar{f}_k(x_i | \bar{\alpha}^{-1}(i)) \right) \\ &\left(\prod_{i \in \mathrm{Dom}(a) \setminus \mathrm{Im}(\bar{\alpha})} p_{\mathrm{M}k}(y_{ka(i)} | x_i) p_{\mathrm{D}k}(x_i) \bar{\gamma}_k(x_i) \right) \\ &\left(\prod_{\substack{i=1 \\ i \notin \mathrm{Dom}(a) \cup \mathrm{Im}(\bar{\alpha})}}^{n} (1 - p_{\mathrm{D}k}(x_i)) \bar{\gamma}_k(x_i) \right) \\ &\left(\prod_{\substack{j=1 \\ j \notin \mathrm{Im}(a)}}^{m_k} \gamma_{FAk}(y_{kj}) \right), \end{aligned} \tag{B.3}$$

where the five product factors, within the second summations over $\bar{\mathcal{A}}(\{1, \ldots, n\}, \{1, \ldots, m_k\})$, correspond to 1) targets detected before detected again, 2) targets detected before but not detected by frame k, 3) targets detected for the first time in frame k, 4) targets not detected before and remaining undetected, and 5) false alarms.

Each of the first three factors in (B.3) in the second summation can be written as the product of a particular assignment likelihood and the updated (or initiated) track PD obtained assuming that assignment. For example, $p_{\mathrm{M}k}(y_{ka(i)}|x_i) p_{\mathrm{D}k}(x_i) \bar{f}_k(x_i|\bar{\tau})$ is the product of the likelihood $\int_E p_{\mathrm{M}k}(y_{ka(i)}|\xi) p_{\mathrm{D}k}(\xi) \bar{f}_k(\xi|\bar{\tau}) \mu(d\xi)$ of track $\bar{\tau}$ being assigned to measurement $y_{ka(i)}$, and the updated track PD $\hat{f}_k(x_i | \bar{\tau} \cup \{(k, y_{ka(i)})\})$.

When we calculate the denominator of (45) by the "set integral" defined in (4), because $\phi_{\mathrm{M}k}(\{y_{kj}\}_{j=1}^{m_k} | \{x_i\}_{i=1}^n) \bar{\phi}_k(\{x_i\}_{i=1}^n)$ is permutable with respect to $(x_i)_{i=1}^n \in E^n$, each term of the second summation over $\bar{\alpha} \in \mathcal{A}(\bar{\lambda}, \{1, \ldots, n\})$ of (B.3) becomes the same values in the integral, i.e., $n!/(n - \#(\bar{\lambda}))!$ times the one term obtained by any arbitrarily chosen $\bar{\alpha} \in \mathcal{A}(\bar{\lambda}, \{1, \ldots, n\})$. Thus, rearranging the summations of (B.3) for the numerator of (45), the updated JMD is calculated as

$$\hat{\phi}_k(\{x_i\}_{i=1}^n) = e^{-\hat{\nu}_k} \sum_{\hat{\lambda} \in \hat{\Lambda}_k} \hat{p}_k(\hat{\lambda}) \sum_{\hat{\alpha} \in \mathcal{A}(\hat{\lambda},\{1,\ldots,n\})} \left(\prod_{\hat{\tau} \in \hat{\lambda}} \hat{f}_k(x_{\hat{\alpha}(\hat{\tau})} | \hat{\tau}) \right) \left(\prod_{\substack{i=1 \\ i \notin \mathrm{Im}(\hat{\alpha})}}^{n} \hat{\gamma}_k(x_i) \right), \tag{B.4}$$

which is nothing but the convolution $\hat{\phi}_k = \hat{\phi}_{\mathrm{D}k} \otimes \hat{\phi}_{\mathrm{U}k}$ of $\hat{\phi}_{\mathrm{D}k}$ defined in (43) and of $\hat{\phi}_{\mathrm{U}k}$ defined in (44), with parameters $((\hat{p}_k(\hat{\lambda}))_{\hat{\lambda} \in \hat{\Lambda}_k}, (\hat{f}_k(\cdot|\hat{\tau}))_{\hat{\tau} \in \cup \hat{\Lambda}_k}, \hat{\gamma}_k)$ defined in (46), (28), and (29).

To derive the extrapolation formulas in Section V-C, we first should observe that $\bar{\phi}_{k+1} = \bar{\phi}_{\mathrm{D}(k+1)} \otimes \bar{\phi}_{\mathrm{U}(k+1)}$

with $\bar{\phi}_{\mathrm{U}(k+1)} = \tilde{\phi}_{\mathrm{U}(k+1)} \otimes \phi_{\mathrm{B}k}$, where

$$\begin{cases} \bar{\phi}_{\mathrm{D}(k+1)}(X_{\mathrm{D}(k+1)}) = \int_{\mathcal{F}(E)} \phi_{\mathrm{TS}k}(X_{\mathrm{D}(k+1)}|X_{\mathrm{D}k}) \\ \qquad\qquad \hat{\phi}_{\mathrm{D}k}(X_{\mathrm{D}k})\mathcal{M}(dX_{\mathrm{D}k}) \\ \tilde{\phi}_{\mathrm{U}(k+1)}(X_{\mathrm{U}(k+1)}) = \int_{\mathcal{F}(E)} \phi_{\mathrm{TS}k}(X_{\mathrm{U}(k+1)}|X_{\mathrm{U}k}) \\ \qquad\qquad \hat{\phi}_{\mathrm{U}k}(X_{\mathrm{U}k})\mathcal{M}(dX_{\mathrm{U}k}) \end{cases} \tag{B.5}$$

implying that $\bar{\phi}_{\mathrm{D}(k+1)}$ and $\tilde{\phi}_{\mathrm{U}(k+1)}$ are independent[63] from each other, since $\hat{\phi}_{\mathrm{D}k}$ and $\hat{\phi}_{\mathrm{U}k}$ are independent from each other. $\phi_{\mathrm{B}k}$ is independent from $\bar{\phi}_{\mathrm{D}(k+1)}$ and from $\tilde{\phi}_{\mathrm{U}(k+1)}$ because $\phi_{\mathrm{TRN}k}(\cdot|X_k) = \phi_{\mathrm{TS}k}(\cdot|X_k) \otimes \phi_{\mathrm{B}k}(\cdot)$.

On the other hand, we can rewrite (38) and (39) as

$$\phi_{\mathrm{TS}k}(\{x_i\}_{i=1}^{n}|\{x'_{i'}\}_{i'=1}^{n'}) = \sum_{\substack{a' \in \bar{\mathcal{A}}(\{1\ldots,n'\},\{1,\ldots,n\}) \\ \#(\mathrm{Dom}(a'))=n}} \left(\prod_{i' \in \mathrm{Dom}(a')} f_{\mathrm{T}k}(x_{a'(i')}|x'_{i'}) p_{\mathrm{S}k}(x'_{i'}) \right) \left(\prod_{\substack{i'=1 \\ i' \notin \mathrm{Dom}(a')}}^{n'} (1 - p_{\mathrm{S}k}(x'_{i'})) \right). \tag{B.6}$$

By substituting the second equation of (43), and (B.6) into the first equation of (B.5), following the definition (4) of the "set integral," we have

$$\begin{aligned} &\bar{\phi}_{\mathrm{D}(k+1)}(\{x_i\}_{i=1}^{n}) \\ &= \int_{\mathcal{F}(E)} \phi_{\mathrm{TS}k}(\{x_i\}_{i=1}^{n}|\{x'_{i'}\}_{i'=1}^{n'}) \\ &\quad \hat{\phi}_{\mathrm{D}k}(\{x'_{i'}\}_{i'=1}^{n'})\mathcal{M}(d\{x'_{i'}\}_{i'=1}^{n'}) \\ &= \sum_{n'=0}^{\infty} \frac{1}{n'!} \sum_{\substack{a' \in \bar{\mathcal{A}}(\{1,\ldots,n'\},\{1,\ldots,n\}) \\ \#(\mathrm{Dom}(a'))=n}} \int_{E^{n'}} \left(\prod_{i' \in \mathrm{Dom}(a')} f_{\mathrm{T}k}(x_{a'(i')}|x'_{i'}) p_{\mathrm{S}k}(x'_{i'}) \right) \\ &\quad \left(\prod_{\substack{i'=1 \\ i' \notin \mathrm{Dom}(a')}}^{n'} (1 - p_{\mathrm{S}k}(x'_{i})) \right) \\ &\quad \left(\sum_{\substack{\hat{\lambda} \in \hat{\Lambda}_k \\ \#(\hat{\lambda})=n'}} \hat{p}_k(\hat{\lambda}) \sum_{\hat{\alpha} \in \mathcal{A}(\hat{\lambda},\{1,\ldots,n'\})} \prod_{\hat{\tau} \in \hat{\lambda}} \hat{f}_k(x'_{\hat{\alpha}(\hat{\tau})}|\hat{\tau}) \right) \\ &\quad \prod_{i'=1}^{n'} \mu(dx'_{i'}) \\ &= \sum_{\hat{\lambda} \in \hat{\Lambda}_k} \hat{p}_k(\hat{\lambda}) \sum_{\substack{a'' \in \bar{\mathcal{A}}(\hat{\lambda},\{1,\ldots,n\}) \\ \#(\mathrm{Dom}(a''))=n}} \frac{1}{\#(\hat{\lambda})!} \sum_{\hat{\alpha} \in \mathcal{A}(\hat{\lambda},\{1,\ldots,\#(\hat{\lambda})\})} \\ &\quad \left(\prod_{\hat{\tau} \in \mathrm{Dom}(a'')} \int_E f_{\mathrm{T}k}(x_{a''(\hat{\tau})}|x'_{\hat{\alpha}(\hat{\tau})}) p_{\mathrm{S}k}(x'_{\alpha(\hat{\tau})}) \hat{f}_k(x'_{\alpha(\hat{\tau})}|\hat{\tau})\mu(dx'_{\alpha(\hat{\tau})}) \right) \\ &\quad \left(\prod_{\hat{\tau} \in \hat{\lambda} \setminus \mathrm{Dom}(a'')} \int_E (1 - p_{\mathrm{S}k}(x'_{\alpha(\hat{\tau})})) \hat{f}_k(x'_{\alpha(\hat{\tau})}|\hat{\tau})\mu(dx'_{\alpha(\hat{\tau})}) \right). \end{aligned} \tag{B.7}$$

[63] More precisely, the RFSets represented by conditional JMT $\bar{\phi}_{\mathrm{D}(k+1)}$ and $\tilde{\phi}_{\mathrm{U}(k+1)}$ are independent.

For given any $\hat{\lambda} \in \hat{\Lambda}_k$, the last summation of (B.7) is over all the enumerations of the tracks in $\hat{\lambda}$. The summation for all the a'''s in $\bar{\mathcal{A}}(\hat{\lambda}, \{1, \ldots, n\})$ such that $\#(\mathrm{Dom}(a'')) = n$ is the summation over all the choices of subsets $\bar{\lambda}$ of $\hat{\lambda}$, such that $\#(\bar{\lambda}) = n \le \#(\hat{\lambda}) = n'$, plus all the possible enumerations of the tracks in the "decimated" hypothesis $\bar{\lambda}$. Hence, we have (51), and we can rewrite (B.7) in the form of first equation of (43) with the index k replaced by $k+1$, with the probabilistic weights $(\bar{p}_{k+1}(\bar{\lambda}))_{\bar{\lambda} \in \bar{\Lambda}_{k+1}}$ defined by (50), and with the track PDs, $(\bar{f}_{k+1}(\cdot|\bar{\tau}))_{\bar{\tau} \in \cup \bar{\Lambda}_{k+1}}$, defined by (53).

Since $\hat{\phi}_{\mathrm{U}k}$ is Poisson and the transition PD $\phi_{\mathrm{TS}k}$ of (38) is target-wise independent, $\tilde{\phi}_{\mathrm{U}(k+1)}$ defined by the second equation of (B.5) is also Poisson with the IMD defined as the second term of the right-hand side of (54). Since $\tilde{\phi}_{\mathrm{U}(k+1)}$ is independent of $\phi_{\mathrm{B}k}$, we have the Poisson JMD $\bar{\phi}_{\mathrm{U}(k+1)} = \tilde{\phi}_{\mathrm{U}(k+1)} \otimes \phi_{\mathrm{B}k}$, which completes the proof by mathematical induction for all the update and prediction formulas in Section V-B and V-C.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan *Estimation with Applications to Tracking and Navigation*. Hoboken, NJ: Wiley, 2001.

[2] S. S. Blackman and R. Popoli *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.

[3] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell *Bayesian Multiple Target Tracking*, 2nd ed. Norwood, MA: Artech House, 2014.

[4] Y. Bar-Shalom "Tracking methods in a multitarget environment," *IEEE Trans. Autom. Control*, vol. 23, no. 4, pp. 618–626, Aug. 1978.

[5] C. L. Morefield "Application of 0–1 integer programming to multi-target tracking problems," *IEEE Trans. Autom. Control*, vol. 23, no. 3, pp. 302–312, Jun. 1977.

[6] D. B. Reid "An algorithm for tracking multiple targets," *IEEE Trans. Autom. Control*, vol. 24, no. 6, pp. 843–854 Dec. 1979.

[7] S. Mori, C.-Y. Chong, E. Tse, and R. P. Wishner "Tracking and classifying multiple targets without a priori identification," *IEEE Trans. Autom. Control*, vol. 31, no. 5, pp. 401–409, May 1986.

[8] C.-Y. Chong, S. Mori, and D. B. Reid "Forty years of multiple hypothesis tracking," *J. Adv. Inf. Fusion*, vol. 14, no. 2, Dec. 2019; an earlier version appeared as C.-Y. Chong, S. Mori, and D. B. Reid "Forty years of multiple hypothesis tracking—A review of key developments," in *Proc. 21st Int. Conf. Inf. Fusion*, Cambridge, UK, Jul. 2018.

[9] S. Mori, C.-Y. Chong, and K. C. Chang "Three formalisms of multiple hypothesis tracking," in *Proc. 19th Int. Conf. Inf. Fusion*, Heidelberg, Germany, Jul. 2016.

[10] J. R. Munkres *Topology*, 2nd ed. London, U.K.: Pearson Education, 2015.

[11] W. Rudin *Functional Analysis*. New York, NY: McGraw-Hill, 1973.

[12] Y. Bar-Shalom and E. Tse
"Tracking in a cluttered environment with probabilistic data association,"
Automatica, vol. 11, pp. 451–460, Sep. 1975.

[13] Y. Bar-Shalom
"Extension of the probabilistic data association filter in multitarget tracking,"
in *Proc. 5th Symp. Nonlinear Estimation*, San Diego, CA, Sep. 1974, pp. 16–21.

[14] M. Lothaire
Combinatorics on Words. Cambridge, U.K.: Cambridge University Press, 1997.

[15] D. J. Daley and D. Vere-Jones
An Introduction to the Theory of Point Processes, vol. I: Elementary Theory and Methods, 2nd ed. Berlin, Germany: Springer, 2003.

[16] S. Mori and C.-Y. Chong
"Point process formalism of multitarget tracking,"
in *Proc. 5th Int. Conf. Inf. Fusion*, Annapolis, MD, Jul. 2002.

[17] S. K. Srinivasan
Stochastic Theory and Cascade Processes. New York, NY: Academic Elsevier, 1969.

[18] L. Janossy
"On the absorption of a nucleon cascade,"
Proc. R. Irish Acad. Sect. A: Math. Phys. Sci., vol. 53, pp. 181–188, 1950.

[19] P. Blagojevic, V. Grujic, and R. Zivaljevic
"Symmetric products of surfaces: a unifying theme for topology and physics,"
in *Proc. Summer School Mod. Math. Phys., SFIN XV (A3)*, Institute of Physics, Belgrade, Serbia, 2002.

[20] E. Michael
"Topologies on spaces of subsets,"
Trans. Amer. Math. Soc., vol. 71, pp. 152–182, 1951.

[21] S. Mori and E. Tse
"A bargaining process modeled as an infinite dynamic non-cooperative game,"
in *Proc. Amer. Control Conf.*, Arlington, VA, Jun. 1982.

[22] A. F. Kaar
Point Processes and Their Statistical Inference. New York, NY: Marcel Dekker, 1986.

[23] J. E. Moyal
"The general theory of stochastic population processes,"
Acta Math., vol. 108, pp. 1–31, 1962.

[24] H. Wald
"Sur les Processus Stationnaires Ponctuels,"
in *Le Calcul des Probabilites et ses Applications*, no. 13. Paris, France: Centre National de la Recherche Scientifique, 1949.

[25] R. P. S. Mahler
Statistical Multisource–Multitarget Information Fusion. Norwood, MA: Artech House, 2007.

[26] R. P. S. Mahler
Advances in Statistical Multisource--Multitarget Information Fusion. Norwood, MA: Artech House, 2014.

[27] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen
Mathematics of Data Fusion. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1997.

[28] H. L. Royden
Real Analysis, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1988.

[29] P. J. Daniell
"A general form of integral,"
Ann. Math., Second Ser., vol. 19, no. 4, pp. 276–294, Jun. 1918.

[30] K. Kastella
"Joint multitarget probabilities for detection and tracking,"
in *Proc. SPIE AEROSENSE '97 Conf. Acquisition, Tracking, Pointing XI*, Orlando, FL, Apr. 1997, vol. 3086, pp. 122–128.

[31] S. Mori and C.-Y. Chong
"Evaluation of data association hypotheses: Non-Poisson i.i.d. cases,"
in *Proc. 7th Int. Conf. Inf. Fusion*, Stockholm, Sweden, Jul. 2004, pp. 1133–1140.

[32] A. B. Poore and N. Rijavec
"A new class of methods for solving data association problems arising from multitarget tracking,"
in *Proc. Amer. Autom. Control Conf.*, Boston, MA, Jun. 1991, vol. 3, pp. 2302–2304.

[33] B.-N. Vo, B.-T. Vo, and D. Phung
"Labeled random finite sets and the Bayes multi-target tracking filter,"
IEEE Trans. Signal Process., vol. 62, no. 24, pp. 6554–6567, Dec. 2014.

[34] S. Mori, K.-C. Chang, and C.-Y. Chong
"Tracking aircraft by acoustic sensors—multiple hypothesis approach applied to possibly unresolved measurements,"
in *Proc. Amer. Control Conf.*, Minneapolis, MN, Jun. 1987, pp. 1099–1103.

[35] S. P. Coraluppi and C. A. Carthel
"Multiple-hypothesis tracking for targets producing multiple measurements,"
IEEE Trans. Aerosp. Electron. Syst., vol. 54, no. 3, pp. 1485–1498, Jun. 2018.

[36] J. Williams
"Marginal multi-Bernoulli filters: RFS derivations of MHT, JIPDA, and association-based MeMBer,"
IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 3, pp. 1664–1687, Jul. 2015.

[37] E. F. Brekke and M. Chitre
"Relationship between finite set statistics and the multiple hypothesis tracker,"
IEEE Trans. Aerosp. Electron. Syst., vol. 54, no. 4, pp. 1902–1917, Jul. 2018.

[38] E. F. Brekke and M. Chitre
"The multiple hypothesis tracker derived from finite set statistics,"
in *Proc. 20th Int. Conf. Inf. Fusion*, Xi'an, China, Jul. 2017.

[39] K. Granström, L. Svensson, Y. Xia, J. Williams, and Á. F. García-Fernández
"Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories,"
in *Proc. Int. Conf. Inf. Fusion*, Cambridge, UK, Jul. 2018.

[40] G. Mathéron
Random Sets and Integral Geometry. Hoboken, NJ: Wiley, 1975.

[41] D. B. Reid
"A multiple hypothesis filter for tracking multiple targets in a cluttered environment," Lockheed Report LMSC-D560254, Lockheed Palo Alto Research Laboratory, Palo Alto, CA, Sep. 1977.

[42] D. Stoyan, W. S. Kendall, and J. Mecke
Stochastic Geometry and Its Applications, 2nd ed. Hoboken, NJ: Wiley, 1995.

[43] S. Mori and C.-Y. Chong
"Data association hypothesis evaluation for i.i.d. but non-Poisson multiple target tracking,"
in *Proc. SPIE Symp. Signal Data Process. Small Targets*, O. Drummond
Ed. Orland, FL, Apr. 2004, vol. 5428, pp. 224–236.

[44] S. Mori
"Random sets in data fusion,"
in *Proc. SPIE Symp. Signal Data Process. Tracking Small Targets*, O. Drummond
Ed. San Diego, CA, Aug. 1997, vol. 3163, pp. 278–289.

[45] I. Molchanov
Theory of Random Sets. Berlin, Germany: Springer, 2005.

[46] H. A. P. Blom and Y. Bar-Shalom
"The interacting multiple model algorithm for systems with Markovian switching coefficients,"
IEEE Trans. Autom. Control, vol. 33, no. 8, pp. 780–783, Aug. 1988.

[47] M. Petrich
Introduction to Semigroups. Durham, NC: Bell & Howell, 1973.

Shozo Mori received the B.S. degree in electric engineering from Waseda University, Tokyo, Japan, in 1970, the M.S. degree in control engineering from Tokyo Institute of Technology, Tokyo, Japan, in 1972, and the Ph.D. degree in engineering economic systems from Stanford University, Stanford, CA, in 1983. He is currently a Consultant with Systems & Technology Research (STR), Sacramento, CA, USA. He was with Advanced Decision & Information Systems (later acquired by Booz-Allen & Hamilton), Mountain View, CA, USA, from 1980 to 1991, Tiburon Systems (later acquired by Texas Instruments, and then by Raytheon Company), San Jose, CA, USA, from 1992 to 1999, Information Extraction & Transport, Arlington, VA, USA, from 2000 to 2003, ALPHATECH (later acquired by BAE Systems), Burlington, MA, USA, from 2003 to 2013, and STR, Woburn, MA, USA, from 2013 to present. His current research interests include multiple target tracking, particularly as applications of theories of random finite sets and finite point processes, distributed data fusion problems in various domains, and applications of theories of information processing and multiple-person decision-making systems to economic and social system analysis. He is an IEEE Life Member. He was a member of Program Committee for MSS National Sensor and Data Fusion Symposium, and was an Area/Assistant Editor for the IEEE Transactions on Systems, Man, and Cybernetics, IEEE Transactions on Aerospace and Electronic Systems, and *Journal of Advances in Information Fusion* published by International Society of Information Fusion. He was the recipient of 2016 Joseph Mignogna Data Fusion Award.

Chee-Yee Chong received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1969, 1970, and 1973, respectively. He was with the Electrical Engineering Faculty, Georgia Tech, Atlanta, GA, USA, until 1980, when he moved to Silicon Valley to join Advanced Decision Systems, which was acquired by Booz Allen Hamilton in 1991. In 2003, he joined ALPHATECH, which was acquired by BAE Systems, and stayed until 2013. He is currently an independent researcher and consultant. He has been involved in all levels of information fusion, including Bayesian multiple hypothesis tracking, optimal track fusion and association, close loop tracking with sensor resource management, analytical model to predict association performance, Bayesian networks for situation assessment, hard and soft data fusion, and graph approaches for data association. He is particularly interested in combining model-based and learning-based approaches for robust data fusion in mission critical systems. He co-founded the International Society of Information Fusion (ISIF), and served as its President in 2004. He was general Co-chair for the 12th International Conference on Information Fusion held in Seattle, WA, USA, in 2009. He was an Associate Editor for the IEEE Transactions on Automatic Control and Information Fusion, and is currently an Associate Editor for *Journal of Advances in Information Fusion* published by ISIF. He was the recipient of the ISIF Yaakov Bar-Shalom Award for a Lifetime of Excellence in Information Fusion in 2016.

Kuo-Chu Chang received the M.S. and Ph.D. degrees in electrical engineering from the University of Connecticut, Storrs, CT, USA, in 1983 and 1986, respectively. From 1983 to 1992, he was a Senior Research Scientist with Advanced Decision Systems (ADS) Division, Booz-Allen & Hamilton, Mountain View, CA, USA. In 1992, he joined the Systems Engineering and Operations Research Department, George Mason University, Fairfax, VA, USA, where he is currently a Professor. His research interests include estimation theory, multisensor data fusion, Bayesian inference, machine learning, and financial engineering. He has more than 35 years of industrial and academic experience and authored or co-authored more than 200 papers in the areas of multitarget tracking, distributed sensor fusion, Bayesian network technologies, and financial engineering. He was an Associate Editor on Tracking/Navigation Systems from 1993 to 1996 and on Large Scale Systems from 1996 to 2006 for the IEEE Transactions on Aerospace and Electronic Systems. He was also an Associate Editor for the IEEE Transactions on Systems, Man, and Cybernetics, from 2002 to 2007. He was the Technical Program Co-chair of the 2009 International Conference on Information Fusion, Seattle, WA, USA. He was elected as an IEEE Fellow for his contribution on "sensor data fusion and Bayesian probabilistic inference" in 2010.