This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

Relationship between Finite Set Statistics and the Multiple Hypothesis Tracker

Edmund Brekke, *Member, IEEE,* and Mandar Chitre, *Senior Member, IEEE*

***Abstract*—The multiple hypothesis tracker (MHT) and finite set statistics (FISST) are two approaches to multitarget tracking which both have been heralded as optimal. In this paper we show that the multitarget Bayes filter with basis in FISST can be expressed in terms the MHT formalism, consisting of association hypotheses with corresponding probabilities and hypothesis-conditional densities of the targets. Furthermore, we show that the resulting MHT-like method under appropriate assumptions (Poisson clutter and birth models, no target-death, linear-Gaussian Markov target kinematics) only differs from Reid's MHT with regard to the birth process.**

***Index Terms*—Multitarget tracking, Data association, MHT, FISST, Random Finite Sets**

I. Introduction

THe standard multitarget tracking problem can be decomposed into the two subproblems of *filtering* and *data association*. Filtering concerns the estimation of a target's kinematic state from a time series of measurements. Data association concerns how one determines which measurements originate from which targets, and which measurements are to be considered useless *clutter*. Many of the tasks carried out in a surveillance system are most appropriately understood in the context of data association. For example, establishing tracks on new targets is fundamentally a data association problem, since a decision to establish a new track amounts to a decision regarding the origin of a sequence of measurements.

A Bayesian formalism is natural in target tracking since prior knowledge about target kinematics is easily quantified by a prior distribution. For filtering without measurement origin uncertainty one can evaluate the posterior probability density function (pdf) of the state given the sequence of measurements so far observed. If the filtering problem is Gaussian as well as linear, then the posterior is a Gaussian, whose sufficient statistics are found by the Kalman filter (KF) [1]. For non-linear and non-Gaussian problems, no closed-form representation of the posterior can be found in general, but the posterior can still be approximated by techniques such as sequential Monte-Carlo methods (SMC) [2].

When data association is involved things get more complicated, and it is not immediately clear that a well-defined posterior exists. Consequently, the concept of optimality becomes problematic for data association problems. Truly Bayes-optimal solutions, i.e., solutions which minimize *a posteriori* expected loss, can only be achieved if the posterior is available. From an intuitive perspective, it seems fairly obvious that Bayes-optimal solutions to the standard multitarget tracking problem must, in one way or another, consider all feasible data association hypotheses.

A milestone was reached in 1979 when Donald Reid proposed the multiple hypothesis tracker (MHT) [3], which evaluates posterior probabilities of all feasible data association hypotheses. Several variants of the MHT exist. In this paper we are solely concerned with Reid's original algorithm. In contrast to many of the later developments, it has two attributes which are very important here. First, Reid's MHT is a recursive method, which updates hypothesis probabilities according to a recursive formula. Second, although one typically would want to find the hypothesis with the highest posterior probability, maximum *a posteriori* (MAP) estimation is never mentioned in [3]. The output of Reid's MHT as described in [3] is simply a collection of possible hypotheses, with corresponding probabilities. In practice, implementations of MHT-methods rely on pruning, simplifications and sliding window techniques to mitigate its exponential complexity. While a practical MHT therefore will be suboptimal, it is commonly believed that the ideal MHT without approximations is optimal in some unspecified sense.

A different approach known as finite set statistics (FISST) was developed by Ronald Mahler [4, 5, 6] in the 2000's. FISST is a reformulation of point process theory tailored to multitarget tracking [7]. In FISST, both targets and measurements are generally treated as random finite sets, i.e., as set-valued random variables. This allows one to express a Bayes-optimal solution to the full tracking problem using a single prediction equation and a single update equation. This optimal recursion, commonly known as the multitarget Bayes filter, is just as intractable as the MHT, but Mahler and coworkers have developed several approximative solutions such as the probability hypothesis density (PHD) filter, the cardinalized probability hypothesis density (CPHD) filter and the multitarget multi-Bernoulli (MeMBer) filter [6, 8].

The relationship between FISST and previously established tracking methods is a controversial topic which has been explored in several papers, although never fully resolved. The integrated probabilistic data association (IPDA) and the joint IPDA (JIPDA) have been linked to the FISST formalism in [9] and [10]. The Set JPDA employs concepts based on FISST to improve the JPDAF [11]. It was shown in [12] that the CPHD filter is equivalent to MHT when maximally one target is present. Several classical tracking methods were discussed from the perspective of point process theory in [13]. The relationship between MHT and FISST has been explored in

E. Brekke is with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, O. S. Bragstads plass 2D, 7034 Trondheim, Norway. e-mail: edmund.brekke@ntnu.no.

M. Chitre is with the Acoustic Research Lab, Tropical Marine Science Institute, National University of Singapore, 14 Kent Ridge Road, Singapore 119222, e-mail: mandar@arl.nus.edu.sg.

Manuscript received October 21, 2016; revised December 15, 2017.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

depth by some authors, most notably Jason Williams and Shozo Mori. In [10], Williams argued that data structures similar to those used in the TO-MHT are implicitly present in the multitarget Bayes filter. Similar ideas were earlier proposed by Mori et al. in [14]. In a more recent paper [15], Mori and coauthors showed that MHT hypothesis probabilities can be derived with basis in random finite sets, although they did not explicitly establish Reid's recursive formulation of MHT. On the other hand, important references in the literature on FISST have criticized the MHT, and argued that association-based approaches such as the MHT may not be consistent with the Bayesian paradigm [6 pp. 340-341, 16 pp. 8-9, 17 p. 14].

A central issue is the treatment of new targets. Reid assumed that a "density of new (i.e., unknown) targets" was given, and could be used to calculate the probabilities of hypotheses assigning measurements to new targets. On the one hand, several papers on MHT (e.g., [18]) refer to Reid's new target density as a birth model, which is a *prior* entity. This interpretation is quite convenient for a couple of reasons, especially since a birth model is part of the standard model of multitarget tracking that has emerged together with FISST. On the other hand, Reid's choice of words indicates that his new target density should be similar to the unknown target density discussed in papers such as [10], which is a *posterior* entity. In this context it is interesting to notice that some recent papers have elaborated the unknown target concept further. For instance, in [19] a distinction is made between unnoticed targets (which eventually are detected at a later time) and ghost targets (which never are detected).

In a recent conference paper [20] we have shown that Reid's MHT indeed can be derived from the multitarget Bayes filter, as parameterized in [10]. This result hinges on the validity of Reid's assumptions (including Gaussian-linear kinematics and no target death), as well as the interpretation of Reid's new target density as the same as the unknown target density in [10].

In this paper we will approach the relationship between FISST and Reid's MHT from a different perspective in two ways. First, we will interpret Reid's new target density as a birth density. We parameterize the resulting multitarget filter entirely as a mixture over association hypotheses. These hypotheses generalize Reid's association hypotheses by also including undiscovered targets. We show that this filter has much of the same structure as Reid's MHT, and involves a hypothesis probability formula which is equal to Reid's formula multiplied by a factor that accounts for undiscovered targets. Second, while the analysis conducted in [20] relied on the transform-domain concept of probability generating functionals (p.g.fl.'s), all the developments of this paper are formulated purely in terms of multiobject densities.

The paper is organized as follows. In Section II we define and explain some key concepts underlying this paper. A very brief introduction to FISST is given in Section III. Assumptions underlying our development of the FISST-MHT are presented in Section IV. The FISST-MHT is then developed in Section V. The special case of Gaussian-linear dynamics is extensively studied in Section VI. Some illustrative examples are provided in Section VII, before a conclusion is given in Section VIII. Mathematical derivations are left for the Appendix. This paper was largely motivated by our work on multihypothesis data association for simultaneous localization and mapping (SLAM), which has previously been reported in [21], [22] and [23].

For the reader's reference, a list of notation used can be found in Table I. Random finite sets are generally denoted by uppercase Greek letters, while their realizations are denoted by uppercase Latin letters. Association hypotheses are denoted by lowercase Greek letters. Sets of association hypotheses are denoted by uppercase calligraphic Latin letters. Some dependencies are omitted for improved readability: In most cases we do not spell out how cardinality numbers (e.g., the number δ_k of detected targets) depend on the association hypotheses. Detailed explanations of notation concerning association hypotheses are given in the next section.

II. Key concepts: Association variables and association hypotheses

Reid's seminal paper [3] developed the concept of association hypotheses through 3 successive refinements of the outcome space, referred to as Number, Configuration and Assignment. The *Number* event concerns how many measurements are associated with existing targets, newborn targets and with clutter. The *Configuration* event concerns how the set of measurements is partitioned into subsets associated with each of these three sources. The *Assignment* event concerns "the specific source of each measurement which has been assigned to be from some previously known target".

In [14], a distinction was made between *data-to-data hypotheses* and *target-to-data hypotheses*. Data-to-data hypotheses concern how measurements at one time step are associated to measurements at other time steps. Target-to-data hypotheses concern how measurements are generated by specific targets. We argue in Remark 4 that Reid's Assignment events are similar to data-to-data hypotheses, while target-to-data hypotheses provide a fourth refinement of the outcome space that plays no role in Reid's MHT.

In this paper we use a novel definition of *association hypotheses*, which is similar, but not identical, to the data-to-data hypotheses of [14], and which extends Reid's Assignment events. The data-to-data hypotheses concern only associations between measurements, while our association hypotheses also concern the numbers of existing targets. The reason why this is needed is that we are going to write the multiobject posterior entirely as a sum over association hypotheses. For the fourth refinement, we use the terminology *association variables* instead of target-to-data hypotheses, in order to emphasize that this refinement is something else than conventional data association hypotheses. In this section we propose precise definitions of both association variables and association hypotheses, before we summarize motivations, interpretations and consequences of these definitions in the following remarks. In the following the symbols 0 and $\emptyset$ should be interpreted as misdetection and non-existing target, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

TABLE I: Notations

Symbol	Description
$a^{\omega_k(i)}$	Score contribution from target i under ω_k
$\mathcal{A}(\theta_{k-1}, n_k)$	Set of predicted association variables
$b(\cdot)$	Kinematic newborn target pdf
$\mathcal{B}(\omega_k)$	Set of currently and previously detected targets
$C, \tilde{C}$	Normalization constants
$c(\cdot)$	Kinematic clutter pdf
$\chi_S(\cdot)$	Indicator function
d_k	Number of targets detected for the first time at k
δ_k	Number of detected targets at time k
$\mathcal{E}(n_k, m_k)$	Set of all feasible ξ's given n_k and m_k
$f(\cdot)$	Generic density (multiobject density or pdf)
F	Kinematic transition matrix
$g^{\pi_k^{1:k-1}(i)}(x_k^i)$	Predicted density of target i
$\mathcal{G}_k^n$	Set of association hypotheses for $n_k = n$
$\mathcal{G}_k$	Set of all association hypotheses at time k
γ	Association variable of the transition density
$h^{\omega_k(i)}(x_k^i)$	Posterior pdf of target i
H	Measurement model matrix
H^*	Complementary measurement model matrix
$\mathcal{I}(\theta_{k-1}, n_k, m_k)$	Set of posterior children hypotheses of θ_{k-1}
$\mathrm{Im}^+(\theta_k)$	Non-zero range of $\omega_k \in \mathrm{Per}(\theta_k)$
k	Current time step
λ	Poisson clutter rate
m_k	Cardinality of Z_k
μ	Poisson rate of newborn targets
μ_v	Expected velocity of newborn targets
n_k	Cardinality of X_k
$\mathbb{N}_0$	Natural numbers including zero
$\mathcal{N}(\cdot\,; \cdot, \cdot)$	Gaussian pdf
ω_k	Posterior association variable at time k
$\mathcal{P}(\theta_{k-1}, n_k, m_k)$	Product set of posterior association variables
P_{D}	Detection probability
P_v	Covariance of newborn target velocity
$\mathrm{Per}(\theta_k)$	Permutations of association hypothesis θ_k
$\mathrm{Pr}(\cdot)$	Generic probability
π_k	Predicted association variable
φ_k	Number of false alarms at time k
P	State covariance
P_{D}	Detection probability
q^{θ_k}	Score of association hypothesis θ_k
Q	Process noise matrix
$r_t(\theta_k)$	Multiplicity of track t in association hypothesis θ_k
R	Measurement noise matrix
S	Surveillance region **or** innovation covariance
Σ_k	Measurement set at time k
t	Track
θ_k	Posterior association hypothesis at time k
$u_k(\theta_k)$	Number of undiscovered targets at time k
$\mathcal{U}_k(n_k)$	Set of posterior association hypotheses
V	Volume of surveillance region
x_k^t	Kinematic state of target t at time k
X_k	Realization of Ξ_k
ξ	Association variable of the measurement model
Ξ_k	Target set at time k
z_k^j	Measurement j at time k
Z_k	Realization of Σ_k
$\#(\pi_k \vert \theta_{k-1})$	Number of feasible π_k's given θ_{k-1}
$\emptyset$	Non-existing target

DEFINITION 1 (association variable)**.** An association variable is a mapping $\omega : \{1, \dots, n_k\} \to (\{\emptyset\} \bigcup \mathbb{N}_0)^k$ such that

- $\omega^l(s) = \omega^l(t) \notin \{\emptyset, 0\} \Rightarrow s = t$ for any $l \in \{1, \dots, k\}$,
- for all $i \in \{1, \dots, n_k\}$ there exists at least one $l \in \{1, \dots, k\}$ for which $\omega^l(i) \neq \emptyset$.
- if $\omega^l(i) \neq \emptyset$ and $l' > l$, then $\omega^{l'}(i) \neq \emptyset$.

The superscript l is the lth coordinate of the mapping ω, or equivalently the lth time step between 0 and the current time step k. Furthermore, the image of ω is denoted $\mathrm{Im}(\omega)$, while $\mathrm{Im}^+(\omega)$ denotes non-zero image of ω, i.e., the collection of all vectors in $\mathrm{Im}(\omega)$ which contain any entries different from 0 or $\emptyset$.

DEFINITION 2 (association hypothesis)**.** Let ω be an association variable. The association hypothesis θ corresponding to ω is then defined as the equivalence class consisting of all association variables $\tilde{\omega}$ for which there exists a permutation mapping $\sigma : \{1, \dots, n_k\} \to \{1, \dots, n_k\}$ such that $\omega^l(\sigma(i)) = \tilde{\omega}^l(i)$ for all $l \in \{1, \dots, k\}$. We signify the relation of permutation-equivalence by the symbol "$\sim$". That is, if such a permutation mapping σ exists for ω and $\tilde{\omega}$, then $\omega \sim \tilde{\omega}$.

REMARK 1. The multitarget Bayes filter is naturally expressed in terms of association variables, and such entities are therefore frequently encountered in [6]. The association variables discussed in [6] are always of the form $\omega : \{1, \dots, n\} \to \{0, \dots, m\}$, where n is the number of targets, and m can be either the number of measurements, or the number of targets at a different time. The association variables γ and ξ defined in (5) and (6) are of this form. Extending the concept to k dimensions instead of only one dimension allows one to assign measurement histories, and not only single measurements, to each target. Such historical association variables are typically of the form $\omega_k : \{1, \dots, n\} \to \{\emptyset, 0, \dots, m_1\} \times \dots \times \{\emptyset, 0, \dots, m_k\}$. Here k is the time index of the last data scan, and m_l is the number of measurements in scan number l. By $\omega_k^l(i) = j$ we understand that target i is believed to have given rise to measurement number j in scan number l, while $\omega_k^l(i) = 0$ indicates that the target was not detected at time l and $\omega_k^l(i) = \emptyset$ indicates that the target did not exist at time l. The notation $\omega_k^{l:l'}(i)$ represents the sequence of measurements associated to target i under ω_k at time steps $l, l+1, \dots, l'$.

REMARK 2. Mappings between sets of integers are not the only possible representation of association variables. Any association variable can also be represented in terms of a $k \times n$-matrix, whose l'th row contains elements from $\{\emptyset\} \bigcup \mathbb{N}_0$ such that no element except possibly 0 or $\emptyset$ is repeated. Any column vector $\omega(i)$ of a historical association variable of the form $\omega : \{1, \dots, n\} \to \{\emptyset, 0, \dots, m_1\} \times \dots \times \{\emptyset, 0, \dots, m_k\}$ constitutes a track.

REMARK 3. It is readily apparent that association hypotheses as defined in this paper correspond to multisets, i.e., as sets with possibly repeated elements. The elements of such a multiset are tracks, i.e., the column vectors of any (matrix-valued) association variable in the equivalence class that constitutes the association hypothesis. Be aware that the target states themselves do not constitute a multiset, since the probability of drawing identical states from a continuous

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

measure is zero. While association hypotheses previously have been defined as sets of tracks [14], the extension to multisets arise when we allow the association hypotheses to also include unknown targets, since there can exist several simultaneously born unobserved targets, which generate identical track-vectors consisting solely of $\emptyset$'s and 0's. Because of this ambiguity (equivalence class versus multiset), the notation $x \in \theta$ may be confusing. We reserve this notation to be used for tracks in the multiset of θ, while membership in the equivalence class of θ will be represented by the notation $\omega \in \mathrm{Per}(\theta)$. That is, $\mathrm{Per}(\theta)$ is the set of all *unique* permutations of the tracks in θ. The cardinality of $\mathrm{Per}(\theta)$ will later be given in (17), followed by an example which illustrates the relationship between association variables and and association hypotheses. The set of all unique $t \in \theta$ is denoted $\mathrm{Im}(\theta)$, and is equal to $\mathrm{Im}(\omega)$ for any $\omega \in \mathrm{Per}(\theta)$. The multiplicity of any $t \in \theta$ is denoted $r_t(\theta)$. For any track on a target which has been detected at some time, we have $r_t(\theta) = 1$.

Remark 4. It is commonly believed that Reid's MHT assigns measurements to targets. This is not the case, because Reid's Assignment event only appends measurements from the last scan to existing measurement histories. Careful inspection of (13) in [3] furthermore reveals that the Assignment events do not distinguish between different permutations of newborn targets. In other words, the newborn targets are treated as a set, and not a list/vector, and their identities are never invoked. Therefore, the Assignment events correspond to the *unordered* association hypotheses, and not to the *ordered* association variables.

III. A very brief review of finite set statistics

FISST can be developed with basis in belief-mass functions. The belief mass function of a random finite set Ξ is the probability

$$\beta_\Xi(S) = \Pr(\Xi \subseteq S). \quad (1)$$

Here S is a subset of the base space, i.e., if $x \in \Xi$ and $x \in \mathbb{R}^n$, then $S \subseteq \mathbb{R}^n$. A multiobject density $f_\Xi(X)$ is a function which produces a non-negative real number from any realization X of Ξ. For a set-function to be a multiobject density it is required that it normalizes to one under the set integral

$$\int f_\Xi(X)\delta X = \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1, \dots, x_n\}) \mathrm{d}x_1 \dots \mathrm{d}x_n = 1.$$

Here n is the cardinality of X, and $f(\{x_1, \dots, x_n\}) = f_\Xi(X)$ under the constraint that $|X| = n$. Furthermore, $f_\Xi(X)$ is related to the belief mass function according to

$$\beta_\Xi(S) = \int_S f_\Xi(X)\delta X.$$

Of this general machinery, we will only be concerned with multiobject densities and the set integral in this paper.

IV. Assumptions

In this section we summarize the assumptions used in the main developments of this paper (Section V), while additional Gaussian-linear assumptions are introduced in Section VI-A. We emphasize that no Gaussian-linear assumptions are relied upon before Section VI.

Many variations of the MHT are reported in the literature. These variations differ both with regard to solution methodology and with regard to underlying assumptions. In this paper, the aim is to re-derive Equation (16) in [3] with basis in FISST. Consequently we employ assumptions similar to those of [3]. In particular, we assume that new targets can be born while we exclude the possibility that already existing targets can die.

During estimation cycle number k we assume that all n_{k-1} targets with states x_{k-1}^i, $i = 1, \dots, n_{k-1}$ remain from the previous cycle. At time k each of these targets have densities

$$f(x_k^i) = \int f(x_k^i | x_{k-1}^i) f(x_{k-1}^i) \mathrm{d}x_{k-1}^i$$

where $f(x_k^i | x_{k-1}^i)$ is the kinematic transition pdf and $f(x_{k-1}^i)$ is the posterior pdf at time $k-1$. In addition to these, β_k new targets appear at time k. The number β_k is Poisson distributed

$$\Pr(\beta_k) = \frac{\mu^{\beta_k} e^{-\mu}}{\beta_k!} \quad (2)$$

while the newborn targets have kinematic pdf $b(x_k^i)$.

At time k we receive m_k measurements $z_k^1, \dots, z_k^{m_k}$. Given that measurement z_k^j is generated by target x_k^i, its kinematic likelihood is given by $f(z_k^j | x_k^i)$. Any target generates a measurement with constant probability P_{D}, otherwise it is unobserved. An unknown number of the m_k measurements are clutter measurements with spatial pdf $c(z_k^j)$. The number φ_k of clutter measurements is also Poisson distributed

$$\Pr(\varphi_k) = \frac{\lambda^{\varphi_k} e^{-\lambda}}{\varphi_k!}$$

Standard independence assumptions (see. e.g. [3], [6] or [24]) apply.

In the framework of FISST, these assumptions can be rephrased in terms of multiobject densities. We leave this for the next section, together with specification of the general form of the multiobject prior used in this paper.

V. The multitarget Bayes filter

The multi-target Bayes recursion from step $k-1$ to time step k involves the random sets Ξ_{k-1} (previous target set, with realization X_{k-1}), Ξ_k (current target set, with realization X_k) and Σ_k (current measurement set, with realization Z_k). Using Bayes' rule, the entire recursion can be written in terms of the prediction equation

$$f_{k|k-1}(X_k | Z_{1:k-1}) = \int f_{\Xi_k|\Xi_{k-1}}(X_k | X_{k-1}) f_{k-1|k-1}(X_{k-1} | Z_{1:k-1}) \delta X_{k-1} \quad (3)$$

and the update equation

$$f_{k|k}(X_k | Z_{1:k}) = \frac{1}{\tilde{C}} f_{\Sigma_k|\Xi_k}(Z_k | X_k) f_{k|k-1}(X_k | Z_{1:k-1}). \quad (4)$$

Here $f_{k|k-1}(X_k | Z_{1:k-1})$ is the multiobject predicted density at time step k and $f_{k|k}(X_k | Z_{1:k})$ is the multiobject posterior density at time step k. The measurement model is encapsulated in the likelihood $f_{\Sigma_k|\Xi_k}(Z_k | X_k)$, while the transition

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

density $f_{\Xi_k|\Xi_{k-1}}(X_k|X_{k-1})$ encapsulates the time evolution (Markov model) of the target set. The constant $\tilde{C}$ is equal to $\int f_{\Sigma_k|\Xi_k}(Z_k|X_k) f_{k|k-1}(X_k|Z_{1:k-1})\delta X_k$.

A. Building blocks: The transition density and the likelihood

Let $X_k = \{x_k^1, \ldots, x_k^{n_k}\}$ and let $Z_k = \{z_k^1, \ldots, z_k^{m_k}\}$. From (13.42) in [6] it is reasonably straightforward to show that the assumptions of Section IV lead to the following multitarget transition density:

$$f_{\Xi_k|\Xi_{k-1}}(X_k|X_{k-1}) = \sum_{\gamma} e^{\mu}\mu^{n_k - n_{k-1}} \prod_{i:\gamma(i)>0} f(x_k^i|x_{k-1}^{\gamma(i)}) \prod_{i:\gamma(i)=0} b(x_k^i). \quad (5)$$

The association variable γ ranges over all functions $\gamma : \{1, \ldots, n_k\} \to \{0, \ldots, n_{k-1}\}$ such that $\gamma(i) = \gamma(j) \neq 0 \Rightarrow i = j$, and such that $\{1, \ldots, n_{k-1}\} \subseteq \mathrm{Im}(\gamma)$. In other words, γ is surjective when restricted to the complement of the pre-image of 0.

Similarly, minor manipulations of (12.139) in [6] lead to the following multiobject likelihood:

$$f_{\Sigma_k|\Xi_k}(Z_k|X_k) = \sum_{\xi} \lambda^{\varphi_k} P_{\mathrm{D}}^{\delta_k}(1 - P_{\mathrm{D}})^{n_k - \delta_k} \prod_{j \notin \mathrm{Im}^+(\xi)} c(z_k^j) \prod_{i:\xi(i)>0} f(z_k^{\xi(i)}|x_k^i). \quad (6)$$

The association variable ξ ranges over all functions $\xi : \{1, \ldots, n_k\} \to \{0, \ldots, m_k\}$ such that $\xi(i) = \xi(j) \neq 0 \Rightarrow i = j$. For future reference we denote this set of ξ's by $\mathcal{E}(n_k, m_k)$. We emphasize that γ and ξ are association variables, and not association hypotheses, according to the definitions in Section II.

B. The prior density

In this section we introduce a general form for the prior multiobject density $f_{k-1|k-1}(X_{k-1}|Z_{1:k-1})$ such that the posterior multiobject density $f_{k|k}(X_k|Z_{1:k})$ remains of the same form after the estimation cycle.

Before introducing this prior, let us discuss the components involved. In order to follow [3] as closely as possible, the prior should be a mixture over all possible association hypotheses θ_{k-1}. For any target set cardinality $n_{k-1} = |X_{k-1}|$, we denote the set of possible association hypotheses by $\mathcal{G}_{k-1}^{n_{k-1}}$. The total set of all association hypotheses is denoted $\mathcal{G}_{k-1} = \bigcup_{n_{k-1}=0}^{\infty} \mathcal{G}_{k-1}^{n_{k-1}}$. The probability of an association hypothesis θ_{k-1} is represented by a non-negative number $q^{\theta_{k-1}}$.

For the proposed prior to be a set-function it must not only be a mixture over all association hypotheses θ_{k-1}, but also over all association variables $\omega_{k-1} \in \mathrm{Per}(\theta_{k-1})$. For each track $\omega_{k-1}(i)$ we assume that an instantaneous track density $f^{\omega_{k-1}(i)}(x_{k-1}^i)$ exists.

Based on all this, we assume the multiobject prior to be of the form

$$f_{k-1|k-1}(X_{k-1}|Z_{1:k-1}) = \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}^{n_{k-1}}} q^{\theta_{k-1}} \sum_{\omega_{k-1} \in \mathrm{Per}(\theta_{k-1})} \prod_{i=1}^{n_{k-1}} f^{\omega_{k-1}(i)}(x_{k-1}^i) \quad (7)$$

with the constraint that

$$\sum_{n_{k-1}=0}^{\infty} \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}^{n_{k-1}}} \frac{q^{\theta_{k-1}}}{\prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!} = 1. \quad (8)$$

In addition to being a set-function, $f_{k-1|k-1}(X_{k-1}|Z_{1:k-1})$ must also normalize to one under the set integral in order to be a multiobject density. One can show that the constraint (8) indeed ensures the correct normalization by counting the number of association variables ω_{k-1} in each $\mathrm{Per}(\theta_{k-1})$. This number is $n_{k-1}!/\prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!$. For brevity, further details of the normalization are omitted.

From this it follows that the prior probability mass contributed by the association hypothesis θ_{k-1} is

$$\Pr(\theta_{k-1}|Z_{1:k-1}) = \frac{q^{\theta_{k-1}}}{\prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!}. \quad (9)$$

In other words, the hypothesis θ_{k-1} has a probability equal to the total probability mass contributed by ω_{k-1}'s within $\mathrm{Per}(\theta_{k-1})$.

The form of (7) is very general and encapsulates several plausible prior densities. First, we note that the empty multiobject density, representing the situation that no targets are present, results when

$$q^{\theta_{k-1}} = \begin{cases} 1 & \text{for } \theta_{k-1} \in \mathcal{G}_{k-1}^0 \\ 0 & \text{for all other } \theta_{k-1}. \end{cases}$$

Second, the Poisson multiobject density

$$f_{\Xi_{k-1}}(X_{k-1}) = e^{-\mu}\mu^{n_{k-1}} \prod_{i=1}^{n_{k-1}} f(x_{k-1}^i)$$

results for example when the following criteria are satisfied:

1) There is a single density $f(\cdot)$ such that $f^{\omega_{k-1}(i)}(x_i) = f(x_{k-1}^i)$ for all association variables $\omega_{k-1}(i)$ and all targets i.
2) One and only one θ_{k-1} is present in each $\mathcal{G}_{k-1}^{n_{k-1}}$, and this association hypothesis constitutes a multiset with a single track repeated n_{k-1} times.
3) $q^{\theta_{k-1}} = e^{-\mu}\mu^{n_{k-1}}$.

C. The predicted density

The predicted multiobject density $f_{k|k-1}(X_k|Z_{1:k-1})$ is found by inserting the transition density (5) and the prior (7) into the prediction equation (3). In a manner similar to the prior density, the predicted density can also be written as a linear mixture over association hypotheses and association variables. Under the no-death assumption of [3], any predicted association hypotheses is entirely given by its parent hypothesis and its current cardinality, i.e., of the form (θ_{k-1}, n_k). Conditional on any such association hypothesis, the predicted density is a linear mixture over association variables in a set $\mathcal{A}(\theta_{k-1}, n_k)$ which consists of all $\pi_k : \{1, \ldots, n_k\} \to (\{\emptyset\} \cup \mathbb{N}_0)^{k-1} \times \{0\}$ such that the following two criteria are satisfied

1) Any vector $\pi_k(i)$ is of one of the two forms
   a) $\pi_k(i) = [\emptyset, \ldots, \emptyset, 0]^\mathsf{T}$
   b) $\pi_k(i) = [t^\mathsf{T}, 0]^\mathsf{T}$ for some track $t \in \theta_{k-1}$.

2) For all unique tracks $t \in \theta_{k-1}$ there are exactly $r_t(\theta_{k-1})$ targets $i \in \{1,\dots,n_k\}$ such that $\pi_k(i) = [t^\mathsf{T}, 0]^\mathsf{T}$.

The k'th row of any π_k is filled with zeros because the measurements in Z_k have not yet been associated to targets.

As shown in Appendix A, we can with this machinery express the predicted density as

$$f_{k|k-1}(X_k|Z_{1:k-1}) = e^{-\mu} \sum_{\theta_{k-1}\in\mathcal{G}_{k-1}} q^{\theta_{k-1}} \mu^{n_k-n_{k-1}} \sum_{\pi_k\in\mathcal{A}(\theta_{k-1},n_k)} \prod_{i=1}^{n_k} g^{\pi_k^{1:k-1}(i)}(x_k^i) \quad (10)$$

where the pdf $g^{\pi_k^{1:k-1}(i)}(x_k^i)$ is given by

$$g^{\pi_k^{1:k-1}(i)}(x_k^i) = \begin{cases} \int f(x_k^i|x) f^{\pi_k^{1:k-1}(i)}(x)\mathrm{d}x & \text{if } \pi_k^{k-1}(i) \neq \emptyset \\ b(x_k^i) & \text{if } \pi_k^{k-1}(i) = \emptyset. \end{cases} \quad (11)$$

Due to the no-death assumption, the sum over $\mathcal{A}(\theta_{k-1},n_k)$ will be empty if $n_k < |\theta_{k-1}|$, and in such a case its value is zero by convention.

We may wish to verify that $f_{k|k-1}(X_k|Z_{1:k-1})$ truly is a multiobject density. It is clearly a set-function, since for any π_k in $\mathcal{A}(\theta_{k-1},n_k)$, all the permutations of π_k are also in $\mathcal{A}(\theta_{k-1},n_k)$. Regarding normalization, let us first notice that the number of association variables in any $\mathcal{A}(\theta_{k-1},n_k)$ is given by the multinomial coefficient

$$|\mathcal{A}(\theta_{k-1},n_k)| = \frac{n_k!}{(n_k-n_{k-1})!\prod_{t\in\mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!}.$$

By recognizing this, and by exploiting the properties of the Poisson distribution, we can reduce the set integral $\int f_{k|k-1}(X_k|Z_{1:k-1})\delta X_k$ to the expression in the prior constraint (8), which by definition equals one. Again, the details are omitted for brevity.

From this it follows that the probability mass contributed by the predicted hypothesis (θ_{k-1},n_k) is

$$\Pr(\theta_{k-1},n_k|Z_{1:k-1}) = \frac{q^{\theta_{k-1}} e^{-\mu}\mu^{n_k-n_{k-1}}}{(n_k-n_{k-1})!\prod_{t\in\mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!}$$

for all (θ_{k-1},n_k) such that $n_{k-1} \leq n_k$.

D. *The posterior density*

According to Bayes' rule for multiobject densities, the posterior density $f_{k|k}(X_k|Z_{1:k})$ is proportional to the product of the likelihood (6) and the predicted density (10). Conditional on any choice of θ_{k-1}, n_k and m_k, the posterior is therefore a linear mixture over elements in the product set $\mathcal{A}(\theta_{k-1},n_k)\times\mathcal{E}(n_k,m_k)$. This product set is isomorphic to a set $\mathcal{P}(\theta_{k-1},n_k,m_k)$ which consists of all association variables $\omega_k : \{1,\dots,n_k\} \to (\{\emptyset\}\cup\mathbb{N}_0)^{k-1}\times\{0,\dots,m_k\}$ which satisfy the two criteria

1) $\omega_k^k(i) = \omega_k^k(j) \neq 0 \Rightarrow i = j$.
2) There exists an association variable $\pi_k \in \mathcal{A}(\theta_{k-1},n_k)$ such that $\omega_k^l(i) = \pi_k^l(i)$ for all previous time steps $l \in \{1,\dots,k-1\}$ and for all targets $i \in \{1,\dots,n_k\}$.

The set $\mathcal{P}(\theta_{k-1},n_k,m_k)$ inherits the permutation-symmetry of $\mathcal{A}(\theta_{k-1},n_k)$ and $\mathcal{E}(n_k,m_k)$. Thus, if ω_k is in $\mathcal{P}(\theta_{k-1},n_k,m_k)$, then all permutations of ω_k are also in $\mathcal{P}(\theta_{k-1},n_k,m_k)$. From this it follows that we can partition $\mathcal{P}(\theta_{k-1},n_k,m_k)$ into permutation-symmetric equivalence classes. Each such equivalence class constitutes a posterior association hypothesis θ_k. Formally, we find the set of posterior association hypotheses θ_k (conditional on θ_{k-1}, n_k and m_k) as the quotient set

$$\mathcal{I}(\theta_{k-1},n_k,m_k) = \mathcal{P}(\theta_{k-1},n_k,m_k))/\sim \quad (12)$$

where "$\sim$" signifies the equivalence relation introduced in Definition 2. Furthermore, we define $\mathcal{G}_k^{n_k}$ as the set of all posterior association hypotheses given that $|X_k| = n_k$. This set is found as the union

$$\mathcal{G}_k^{n_k} = \bigcup_{\theta_{k-1}} \mathcal{I}(\theta_{k-1},n_k,m_k). \quad (13)$$

The sets $\mathcal{I}(\theta_{k-1},n_k,m_k)$ are mutually disjoint under the assumptions employed in this paper.

Having thus established the association hypotheses involved, we can, as shown in Appendix B, write the posterior multiobject density as

$$f_{k|k}(X_k|Z_{1:k}) = \sum_{\theta_k\in\mathcal{G}_k^{n_k}} q^{\theta_k} \sum_{\omega_k\in\mathrm{Per}(\theta_k)} \prod_{i=1}^{n_k} h^{\omega_k(i)}(x_k^i) \quad (14)$$

where the following quantities are involved

$$q^{\theta_k} = \frac{1}{C} q^{\theta_{k-1}} \lambda^{\varphi_k} P_\mathrm{D}^{\delta_k}(1-P_\mathrm{D})^{n_k-\delta_k} \mu^{n_k-n_{k-1}} \prod_{j\notin\mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i\,:\,\omega_k^k(i)>0} a^{\omega_k(i)} \quad (15)$$

$$a^{\omega_k(i)} = \int f(z_k^{\omega_k^k(i)}\,|\,x_k^i)\, g^{\omega_k^{1:k-1}(i)}(x_k^i)\,\mathrm{d}x_k^i \quad (16)$$

$$h^{\omega_k(i)}(x_k^i) = \begin{cases} \frac{f(z_k^{\omega_k^k(i)}\,|\,x_k^i) g^{\omega_k^{1:k-1}(i)}(x_k^i)}{a^{\omega_k(i)}} & \text{if } \omega_k^k(i) > 0 \\ g^{\omega_k^{1:k-1}(i)}(x_k^i) & \text{if } \omega_k^k(i) = 0. \end{cases}$$

The scalar q^{θ_k} represents the score of the hypothesis θ_k, while the scalar $a^{\omega_k(i)}$ represents the contribution from target number i to this score. The pdf $h^{\omega_k(i)}(x_k^i)$ contains the posterior kinematic information of target number i. Furthermore, θ_{k-1} is the parent hypothesis of θ_k, and c is a normalization constant.

REMARK 5 (The hypothesis score is well-defined)**.** The products in (15) are defined in terms of *any* association variable ω_k within the equivalence class $\mathrm{Per}(\theta_k)$. For the first product in (15), the ordering of ω_k is clearly irrelevant since the target index i does not appear. For the second product, we notice that $a^{\omega_k(i)}$ depends on $\omega_k(i)$, but not on i itself, since x_k^i is being integrated out in (16). Therefore, the ordering of ω_k is also irrelevant for the second product, and q^{θ_k} is well-defined.

REMARK 6 (Equivalence of association variables)**.** Any $\omega_k \in \mathrm{Per}(\theta_k)$ gives rise to the same collection of posterior track densities $h^{\omega_k(i)}(\cdot)$. Two association variables $\omega_k^{(1)}$ and $\omega_k^{(2)}$ which both are in $\mathrm{Per}(\theta_k)$ may differ in the sense that $h^{\omega_k^{(1)}(i)}(x_k^i) \neq h^{\omega_k^{(2)}(i)}(x_k^i)$, but one can always find a permutation mapping $\sigma : \{1,\dots,n_k\} \to \{1,\dots,n_k\}$ such that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

$h^{\omega_k^{(1)}(i)}(x_k^i) = h^{\omega_k^{(2)}(\sigma(i))}(x_k^i)$ for all $i \in \{1, \dots, n_k\}$. Insofar as one is interested in identities, one is probably not so interested in *target* identities, but rather in *track* identities, as represented by histories of measurements or histories of state estimates. All $\omega_k \in \mathrm{Per}(\theta_k)$ are equivalent from such a practical perspective, and all the kinematic attributes of the association hypothesis θ_k can be expressed in terms of any association variable $\omega_k \in \mathrm{Per}(\theta_k)$. This means that FISST does not need any track labels beyond what is implicitly coded in the measurement histories to provide tracks.

REMARK 7 (The role of track labels)**.** The idea that track labels are required for continuity of tracks is prevalent. This premise underlies tracking methods based on labeled random finite sets [25, 26]. However, the discussion in Remark 6 shows that labels are not needed for track continuity, insofar as one accepts the definition of track used in this paper. One could nevertheless argue that this reasoning hinges on the availability of metadata which are not present in $f_{k|k}(X_k|Z_{1:k})$ itself, i.e., the hypothesis structures. If this is considered a problem, a solution may be to extend random finite sets to a theory of random finite sets of trajectories [27]. Nevertheless, explicit track labels are not needed in this formulation either. It was shown in [28] that track labels have no impact on the tracking results in the standard multitarget Bayes filter. We refer the reader to Sections III-D and VI-D in [20] for further discussion on this topic.

REMARK 8 (Maintenance of structure)**.** The posterior (14) is of the same form as the prior (7). The maintenance of this form validates that (7) indeed is a reasonable "induction hypothesis". Furthermore, we conclude that association hypotheses are an integral part of the multiobject posterior insofar as the prior can be written on this form for any previous time step.

REMARK 9 (Hypotheses are not state variables)**.** In [6] pp. 340-341, it was claimed that the MHT treats association hypotheses as state variables, and that this leads to some conceptual problems. However, the mixture expression (14), and similar expressions in, e.g., [10, 20, 26], demonstrate that FISST allows association hypotheses to be present in the multiobject posterior without actually entering the state itself (which is the random set Ξ_k with realization X_k). Furthermore, assigning probabilities to the association hypotheses do not change this. Such probabilities merely express a partitioning of the outcome space into a countable set of well-defined events.

In order to calculate the posterior association probabilities inherent in (14), it is necessary to determine the number of association variables ω_k contained in $\mathrm{Per}(\theta_k)$ for any posterior association hypothesis θ_k. This number is given by the multinomial coefficient

$$\#(\omega_k \mid \theta_k) = |\mathrm{Per}(\theta_k)| = \frac{n_k!}{\prod_{t \in \mathrm{Im}(\theta_k)} r_t(\theta_k)!}. \tag{17}$$

EXAMPLE 1 (Enumeration of association variables)**.** A possible association hypothesis could be

$$\theta_k = \left\{ \begin{array}{ccc} 1 & \emptyset & \emptyset \\ 0 & 0 & 0 \end{array} \right\}$$

where curly braces are used to emphasize its multiset nature. According to this hypothesis, the following has happened:

- At time $k = 1$ a single target existed and gave rise to measurement number 1.
- At time $k = 2$ this target was not observed. Furthermore, two new targets were born, but also not observed.

For this association hypothesis, we have three possible association variables, cf. Definition 2 and Remark 3. These are

$$\begin{bmatrix} 1 & \emptyset & \emptyset \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \emptyset & 1 & \emptyset \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \emptyset & \emptyset & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let us then investigate the validity of (17). For the numerator we have that $n_k! = 3! = 6$. For the denominator we have two unique tracks in $\mathrm{Im}(\theta_k)$. For the first of these, $[1, 0]^\mathsf{T}$, we have $r_t(\theta_k)! = 1!$. For the second of these, $[\emptyset, 0]^\mathsf{T}$, we have $r_t(\theta_k)! = 2!$, and thus the demoninator is 2. Consequently, (17) tells us that the number of association variables corresponding to θ_k is $6/2 = 3$.

We can partition the total probability mass into contributions from each parent hypothesis θ_{k-1}, which are further partitioned into contributions from each child hypothesis θ_k. This is done by means of the set integral:

$$\begin{aligned} 1 &= \int f_{k|k}(X_k \mid Z_{1:k}) \delta X_k \\ &= \frac{1}{C} \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}} \sum_{n_k = n_{k-1}}^{\infty} \frac{1}{n_k!} \\ &\quad \sum_{\theta_k \in \mathcal{I}(\theta_{k-1}, n_k, m_k)} q^{\theta_{k-1}} \frac{n_k!}{\prod_{t \in \mathrm{Im}(\theta_k)} r_t(\theta_k)!} P_\mathrm{D}^{\delta_k} (1 - P_\mathrm{D})^{n_k - \delta_k} \\ &\quad \lambda^{\varphi_k} \mu^{n_k - n_{k-1}} \prod_{j \notin \mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i \,:\, \omega_k^k(i) > 0} a^{\omega_k(i)} \\ &= \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}} \sum_{n_k = n_{k-1}}^{\infty} \sum_{\theta_k \in \mathcal{I}(\theta_{k-1}, n_k, m_k)} q^{\theta_k} \frac{1}{\prod_{t \in \mathrm{Im}(\theta_k)} r_t(\theta_k)!}. \end{aligned}$$

Let $u_k(\theta_k)$ denote the number of unknown newborn targets hypothesized by the association hypothesis θ_k. We can then express the the posterior probability of θ_k as

$$\begin{aligned} \Pr(\theta_k \mid Z_{1:k}) &= q^{\theta_k} \frac{1}{\prod_{t \in \mathrm{Im}(\theta_k)} r_t(\theta_k)!} \\ &= \frac{1}{C} \lambda^{\varphi_k} P_\mathrm{D}^{\delta_k} (1 - P_\mathrm{D})^{n_k - \delta_k} \mu^{n_k - n_{k-1}} \frac{\prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!}{\prod_{t \in \mathrm{Im}(\theta_k)} r_t(\theta_k)!} \\ &\quad \left(\prod_{j \notin \mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i \,:\, \omega_k^k(i) > 0} a^{\omega_k(i)} \right) \Pr(\theta_{k-1} | Z_{1:k-1}). \\ &= \frac{1}{C} \lambda^{\varphi_k} P_\mathrm{D}^{\delta_k} (1 - P_\mathrm{D})^{n_k - \delta_k} \mu^{n_k - n_{k-1}} \frac{1}{u_k(\theta_k)!} \\ &\quad \left(\prod_{j \notin \mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i \,:\, \omega_k^k(i) > 0} a^{\omega_k(i)} \right) \Pr(\theta_{k-1} \mid Z_{1:k-1}). \end{aligned} \tag{18}$$

Equations (14) and (18) may be viewed as the key results of this paper.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

VI. Comparison with the standard MHT under Gaussian-linear assumptions

Reid's MHT [3] was derived under standard Gaussian-linear assumptions. In this section we show that the FISST-MHT that was developed in the previous section becomes very similar to Reid's MHT in the Gaussian-linear case.

A. Gaussian-linear assumptions

For the special case of a Gaussian-linear tracking problem we make the following standard assumptions

$$f(x_k|x_{k-1}) = \mathcal{N}(x_k\,;\,Fx_{k-1},Q) \tag{19}$$

$$f(z_k^j|x_k^i) = \mathcal{N}(z_k^j\,;\,Hx_k^i,R) \text{ if } z_k^j \text{ is gen. by } x_k^i \tag{20}$$

$$c(z_k^j) = \frac{1}{V} \text{ if } z_k^j \text{ is clutter} \tag{21}$$

in addition to the assumptions already listed in Section IV. Furthermore, we assume that the state vector and observation matrix are as in Section VII of [3]

$$x_k = [\rho_{x,k},\rho_{y,k},v_{x,k},v_{y,k}]^\mathsf{T}$$
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{22}$$

where ρ denotes position and v denotes velocity.

It makes sense to define the birth density $b(x_k^i)$ as uniform over the surveillance region with constant value $1/V$, as was done in [3]. Notice, though, that this only specifies the birth density over position components of the state vector, and not over velocity components. In order to address this we define another matrix

$$H^* = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{23}$$

such that $H^*x_k = [v_{x,k},v_{y,k}]^\mathsf{T}$. Let us also assume that the velocity of a newborn target is a Gaussian random vector with mean μ_v and covariance P_v. The birth density is then defined as

$$b(x_k^i) = \frac{1}{V}\chi_S(Hx_k^i)\mathcal{N}(H^*x_k^i\,;\,\mu_v,P_v). \tag{24}$$

where S is the surveillance region (whose volume is V), and $\chi_S(\cdot)$ denotes the indicator function of S.

We also make the assumption that R, Q, μ_v and the off-diagonal elements of F are small relative to the size of the surveillance region.

B. Reid's MHT

The key development of [3] was equation (16) in that paper, which specified the posterior probability of an association hypothesis. Using the notation of this paper, we can rewrite this formula as

$$\begin{aligned}\Pr(\theta_k|Z_{1:k}) = &\frac{1}{C}P_\mathrm{D}^{\delta_k}(1-P_\mathrm{D})^{n_k-\delta_k}\left(\frac{\lambda}{V}\right)^{\varphi_k}\left(\frac{\mu}{V}\right)^{n_k-n_{k-1}}\\ &\times\left(\prod_{i\in\mathcal{B}(\omega_k)}\mathcal{N}(z_k^{\omega_k^k(i)}\,;\,z_{k|k-1}^{\omega_k^{1:k-1}(i)},S_{k|k-1}^{\omega_k^{1:k-1}(i)})\right)\\ &\times\Pr(\theta_{k-1}|Z_{1:k-1})\end{aligned} \tag{25}$$

where the set

$$\mathcal{B}(\omega_k) = \big\{i \text{ such that } \omega_k^k(i)>0 \text{ and if } \omega_k^{k-1}(i)=0 \text{ then } \omega_k^l(i)>0 \text{ for some } l>k-1\big\}$$

contains the indices of targets which are currently as well as previously detected, where

$$\begin{aligned} z_{k|k-1}^{\omega_k^{1:k-1}(i)} &= HFx_{k-1|k-1}^{\omega_k^{1:k-1}(i)}\\ S_{k|k-1}^{\omega_k^{1:k-1}(i)} &= H(FP_{k-1|k-1}^{\omega_k^{1:k-1}(i)}F^\mathsf{T}+Q)H^\mathsf{T}+R\end{aligned} \tag{26}$$

and where θ_{k-1} is the parent hypothesis of θ_k. Recalling Remark 5, we emphasize that the product in (25) attains the same value for any $\omega_k\in\mathrm{Per}(\theta_k)$.

Reid assumed Gaussian-linear kinematics in [3], and hence used standard KF formulas to evaluate the posterior kinematic pdf of the target states when conditioned on a particular data association hypothesis. The exact formulas used in [3] are not repeated here. Instead, we show how the FISST-MHT lends itself to Gaussian-linear kinematics in Sections VI-C and VI-D.

C. Kinematics of the Gaussian-linear FISST-MHT

In this and the next subsection we develop a Gaussian-linear FISST-MHT based on the general FISST-MHT of Section V. This subsection is devoted to the state estimation of the Gaussian-linear FISST-MHT, while data association is treated in the next subsection.

State estimation for the Gaussian-linear FISST-MHT entails specification of the posterior target densities $h^{\omega_k(i)}(x_k^i)$ under Gaussian-linear assumptions. Unfortunately, the presence of the non-Gaussian birth density $b(x_k^i)$ means that the posterior target densities will never be truly Gaussian. Nevertheless, under reasonable assumptions, the deviations from Gaussianity will not be large, and KF-based formulas can safely be used. In order to show this we need to treat $h^{\omega_k(i)}(x_k^i)$ differently depending on whether x_k^i is a newborn target, whether it is currently observed, and whether it is previously observed. This leads to 6 different cases as shown in Figure 1.

It should be noted that only cases 2, 5 and 6 were considered in [3]. If we interpret the new target density in [3] as a birth density akin to (2), then the omission of the other cases must be viewed as a flaw in Reid's MHT, whose impact will be studied in the remainder of this paper. If it instead is interpreted as the density of undiscovered targets [20], as Reid arguably intended, then the additional cases are not needed. However, if one wants to be faithful to the standard model, including its target birth model, and at the same time choose the latter interpretation, then a procedure to estimate the unknown target density such as the Poisson-component of [10] must accompany Reid's MHT.

1) Newborn undetected target: Case 1 results when $\omega_k^k(i)=0$ and $\omega_k^{k-1}(i)=\varnothing$. In this case it is readily apparent that $h^{\omega_k(i)}(x_k^i)=b(x_k^i)$.

2) Newborn detected target: Case 2 results when $\omega_k^k(i)>0$ and $\omega_k^{k-1}(i)=\varnothing$. In this case the density $h^{\omega_k(i)}(x_k^i)$ will

Not observed before				Observed before	
Newborn		Previously born			
Not observed now	Observed now	Not observed now	Observed now	Not observed now	Observed now
1	2	3	4	5	6

Fig. 1: Schematic overview of the 6 different cases for the form of the kinematic posterior $h^{\omega_k(i)}(x_k^i)$. Only cases 2, 5 and 6 were considered in Reid's paper [3].

be approximately Gaussian insofar as the measurement noise matrix R is sufficiently small. We find that

$$\begin{aligned} h^{\omega_k(i)}(x_k^i) &= g^{\omega_k^{1:k-1}(i)}(x_k^i)\, b(x_k^i) \\ &= \mathcal{N}(z_k^{\omega_k^k(i)}\,;\, Hx_k^i, R)\, b(x_k^i) \\ &\approx \mathcal{N}(x_k^i\,;\, x_{k|k}^{\omega_k(i)}, P_{k|k}^{\omega_k(i)}) \end{aligned}$$

where

$$x_{k|k}^{\omega_k(i)} = \begin{bmatrix} z_k^{\omega_k^k(i)} \\ \mu_v \end{bmatrix} \text{ and } P_{k|k}^{\omega_k(i)} = \begin{bmatrix} R & 0 \\ 0 & P_v \end{bmatrix}. \tag{27}$$

3) Already existing and so far undetected target: Case 3 results when $\omega_k^l(i) = 0$ for all $l > l'$ and $\omega_k^l(i) = \emptyset$ for all $l \le l'$ for some $l' \in \{1, \dots, k-2\}$. In this case, $h^{\omega_k(i)}(x_k^i)$ is a $k - l'$ times convolution between the transition density $\mathcal{N}(x_k\,;\, Fx_{k-1}, Q)$ and the original birth density. Insofar as the plant noise covariance Q as well as μ_v and the off-diagonal elements of F are small enough, this can be approximated by the birth density itself:

$$\begin{aligned} h^{\omega_k(i)}(x_k^i) = &\int \mathcal{N}(x_k^i\,;\, Fx_{k-1}, Q) \int \mathcal{N}(x_{k-1}\,;\, Fx_{k-2}, Q) \\ &\dots \mathcal{N}(x_{l'+1}\,;\, Fx_{l'}, Q) b(x_{l'}) \mathrm{d}x_{l'} \dots \mathrm{d}x_{k-2} \mathrm{d}x_{k-1} \\ \approx &\, b(x_k^i). \end{aligned}$$

4) Already existing target detected for the first time: Case 4 results when $\omega_k^k(i) > 0$, $\omega_k^{k-1}(i) = 0$ and $\omega_k^l(i) \in \{0, \emptyset\}$ for all $l < k$. For this case, the no-death assumption implies that there exists some $l < k-1$ such that $\omega_k^{l'}(i) = \emptyset$ for all $l' \le l$, and such that that $\omega_k^{l'}(i) = 0$ for all $l' \in \{l, \dots, k-1\}$. Since Case 4 is always preceded by Case 3, we can assume that the predicted target density is approximately equal to the birth density, i.e., that $g^{\omega_k^{1:k-1}(i)}(x_k^i) \approx b(x_k^i)$. It follows from this that Case 4 is equivalent to Case 2, so that

$$h^{\omega_k(i)}(x_k^i) \approx \mathcal{N}(x_k^i\,;\, x_{k|k}^{\omega_k(i)}, P_{k|k}^{\omega_k(i)})$$

where $x_{k|k}^{\omega_k(i)}$ and $P_{k|k}^{\omega_k(i)}$ are as defined in (27).

5) Previously detected target currently undetected: Case 5 results when $\omega_k^k(i) = 0$ and $\omega_k^l(i) > 0$ for some $l < k$. In this case we find that

$$h^{\omega_k(i)}(x_k^i) = \mathcal{N}(x_k^i\,;\, Fx_{k-1|k-1}^{\omega_k^{1:k-1}(i)}, FP_{k-1|k-1}^{\omega_k^{1:k-1}(i)}F^\mathsf{T} + Q).$$

6) Previously detected target currently detected: Case 6 results when $\omega_k^k(i) > 0$ and $\omega_k^l(i) > 0$ for some $l < k$. In this case the standard KF-formulas yield

$$h^{\omega_k(i)}(x_k^i) = \mathcal{N}(x_k^i\,;\, x_{k|k}^{\omega_k(i)}, P_{k|k}^{\omega_k(i)})$$

where

$$\begin{aligned} x_{k|k}^{\omega_k(i)} &= Fx_{k-1|k-1}^{\omega_k^{1:k-1}(i)} + K_k(z_k^{\omega_k^k(i)} - HFx_{k-1|k-1}^{\omega_k^{1:k-1}(i)}) \\ P_{k|k}^{\omega_k(i)} &= (I - K_k H) P_{k|k-1}^{\omega_k^{1:k-1}(i)} \\ P_{k|k-1}^{\omega_k^{1:k-1}(i)} &= FP_{k-1|k-1}^{\omega_k^{1:k-1}(i)} F^\mathsf{T} + Q \\ K_k &= P_{k|k-1}^{\omega_k^{1:k-1}(i)} H^\mathsf{T} (HP_{k|k-1}^{\omega_k^{1:k-1}(i)} H^\mathsf{T} + R)^{-1}. \end{aligned}$$

The formulas of this last case correspond to the KF-based formulas on page 845 of [3]. Other cases were not explicitly treated in [3], but only dealt with through a somewhat vague discussion on page 848 of [3]. Our Cases 2 and 5 can both be viewed as special cases of this general discussion. Cases 1, 3 and 4 were not considered possible at all in [3]. Finally, we emphasize the implications of Remark 6: In order to specify the kinematic attributes of an association hypothesis θ_k, we only need to specify the kinematic attributes of any association variable $\omega_k \in \mathrm{Per}(\theta_k)$, since since all $\omega_k \in \mathrm{Per}(\theta_k)$ are permutations of each other.

D. Probabilities of the Gaussian-linear FISST-MHT

In order to compare our hypothesis probability formula (18) with Reid's hypothesis probability formula (25) we need to substitute expressions for the target contribution scalar $a^{\omega_k(i)}$ under the Gaussian-linear assumptions (19) - (24). With regard to the cases listed in Figure 1 we notice that $a^{\omega_k(i)}$ is only defined for cases 2, 4 and 6, since $a^{\omega_k(i)}$ is only evaluated for $\omega_k^k(i) > 0$.

For cases 2 and 4, i.e., for previously unobserved targets, we find that

$$\begin{aligned} a^{\omega_k(i)} &= \frac{1}{V} \int_S \mathcal{N}(z_k^{\omega_k^k(i)}\,;\, \rho, R) \mathrm{d}\rho \int_{\mathbb{R}^2} \mathcal{N}(v\,;\, \mu_v, P_v) \mathrm{d}v \\ &\approx \frac{1}{V} \end{aligned} \tag{28}$$

where the first equality only holds approximately for Case 4. The approximation on the second line of (28) holds insofar as R is small enough and $z_k^{\omega_k^k(i)}$ is not too close to the edge of the surveillance region.

For Case 6, i.e., for targets which are both previously and currently detected, it is straightforward to show that

$$a^{\omega_k(i)} = \mathcal{N}(z_k^{\omega_k^k(i)}\,;\, z_{k|k-1}^{\omega_k^{1:k-1}(i)}, S_{k|k-1}^{\omega_k^{1:k-1}(i)}) \tag{29}$$

with $z_{k|k-1}^{\omega_k^{1:k-1}(i)}$ and $S_{k|k-1}^{\omega_k^{1:k-1}(i)}$ defined as in (26).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

By inserting (21), (28) and (29) into (18) we obtain the Gaussian-linear hypothesis probability formula

$$\begin{aligned}\Pr(\theta_k|Z_{1:k}) \approx \frac{1}{C}\left(\frac{\lambda}{V}\right)^{\varphi_k} P_{\mathrm{D}}^{\delta_k}(1-P_{\mathrm{D}})^{n_k-\delta_k} \\ \times \frac{1}{u_k(\theta_k)!}\mu^{n_k-n_{k-1}-d_k}\left(\frac{\mu}{V}\right)^{d_k} \\ \times \left(\prod_{i\in\mathcal{B}(\omega_k)} \mathcal{N}(z_k^{\omega_k^k(i)}; z_{k|k-1}^{\omega_k^{1:k-1}(i)}, S_{k|k-1}^{\omega_k^{1:k-1}(i)})\right) \\ \times \Pr(\theta_{k-1}|Z_{1:k-1}). \qquad (30)\end{aligned}$$

By d_k we denote the number of targets which are detected *for the first time* at time step k under the hypothesis θ_k.

E. Relationship between the Gaussian-linear FISST-MHT and Reid's MHT

We make four observations about how the Gaussian-linear FISST-MHT relates to Reid's MHT:

First, we can rest assured that all association hypotheses that are present Reid's MHT are also present in the FISST-MHT. Any such association hypothesis can be interpreted as a set of tracks, i.e., as a set of measurement histories.

Second, any association hypothesis that is present in the FISST-MHT and not in Reid's MHT must include a track on a target that at some point in time existed without having been observed. Such an association hypothesis may possibly, but not necessarily, correspond to a multiset with several repeated tracks. For the Poisson birth model, the FISST-MHT yields an infinitude of such hypotheses while Reid's MHT yields only a finite number of hypotheses.

Third, the expressions for hypothesis probabilities are identical except from two terms which arise from the treatment of unobserved newborn targets. That is, except from the factors $\mu^{n_k-n_{k-1}-d_k}$ and $1/u_k(\pi)!$, the expression in (30) is identical to the expression in (25).

Fourth, the deviation between the two approaches depends on the birth rate μ. As $\mu \to 0$, most hypotheses not present in Reid's MHT will get very small posterior probabilities in the FISST-MHT, and the two approaches become indistinguishable.

VII. Illustrative example

In this section we study a simple example in order to illustrate how the Gaussian-linear FISST-MHT works with hypotheses, and how it differs from Reid's MHT. In this example we are given four consecutive sets of measurements

$$Z_1 = \{z_1^1\},\ \ Z_2 = \{z_2^1, z_2^2\},\ \ Z_3 = \{z_3^1\},\ \ Z_4 = \{z_4^1, z_4^2\}$$

where $z_1^1 = [1,1]^\mathsf{T}$, $z_2^1 = [2,2]^\mathsf{T}$, $z_2^2 = [2,4]^\mathsf{T}$, $z_3^1 = [3,3]^\mathsf{T}$, $z_4^1 = [4,2]^\mathsf{T}$ and $z_4^2 = [4,4]^\mathsf{T}$ as illustrated in Figure 2. The following kinematic matrices are used:

$$F = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ Q = \begin{bmatrix} \frac{0.01}{3} & 0 & \frac{0.01}{2} & 0 \\ 0 & \frac{0.01}{3} & 0 & \frac{0.01}{2} \\ \frac{0.01}{2} & 0 & 0.01 & 0 \\ 0 & \frac{0.01}{2} & 0 & 0.01 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ R = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}.$$

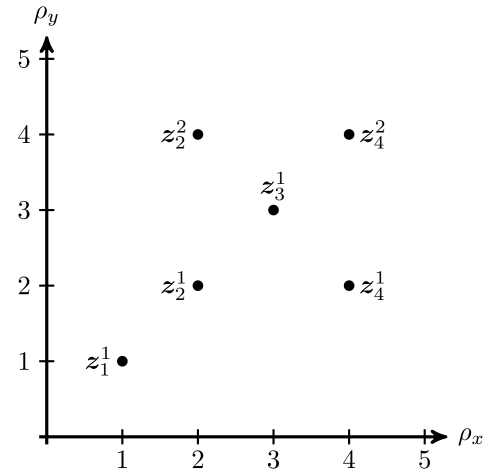

Fig. 2: Simulated scenario.

Furthermore, we use the tuning constants $\lambda = 0.25$, $V = 25$ and $P_{\mathrm{D}} = 0.7$.

It seems fairly likely that at least one target is present, moving from $(1,1)$ at $k = 1$ to $(4,4)$ at $k = 4$. It also seems plausible that a second target is moving from $(2,4)$ at $k = 2$ to $(4,2)$ at $k = 4$. The measurement z_3^1 could have been generated by both these targets. These possibilities can be represented by three association hypotheses:

$$\theta_4^{(1)} = \left\{\begin{matrix} 1 \\ 1 \\ 1 \\ 2 \end{matrix}\right\},\ \theta_4^{(2)} = \left\{\begin{matrix} 1 & \emptyset \\ 1 & 2 \\ 1 & 0 \\ 2 & 1 \end{matrix}\right\},\ \theta_4^{(3)} = \left\{\begin{matrix} 1 & \emptyset \\ 1 & 2 \\ 0 & 1 \\ 2 & 1 \end{matrix}\right\}.$$

We write association hypotheses using braces in order to emphasize their multiset nature, cf. Remark 3. More concretely, the notation of these expressions should be understood as follows:

According to the hypothesis $\theta_4^{(1)}$ there exists one target during time steps 1-4. This target has given rise to the measurements z_1^1, z_2^1, z_3^1 and z_4^2 at time steps 1, 2, 3 and 4, respectively.

According to the hypothesis $\theta_4^{(2)}$ there exist two targets during time steps 2-4, of which one existed also at time step 1. The target which existed already at time step 1 has given rise to the same measurements as the target hypothesized by the hypothesis $\theta_4^{(1)}$. The other target that is hypothesized by the hypothesis $\theta_4^{(2)}$ has give rise to measurement z_2^2 at time step 2 and measurement z_4^1 at time step 4, while being undetected at time step 3.

Let us first investigate the case of a moderate birth rate $\mu = 0.25$ in some detail. Reid's MHT generates 602 hypotheses for the simulated scenario, while the FISST-MHT generates an infinite number of hypotheses in the absence of pruning. By pruning all hypotheses which contain more than 3 simultaneously newborn targets, and all hypotheses whose probabilities fall below 0.001, we reduce the numbers of hypotheses to 27 for Reid's MHT and to 300 for the FISST-MHT.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

The top 7 hypotheses generated by Reid's MHT are

$$\left\{ \begin{array}{c|cc|cc|cc|cc|ccc|c} 1 & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & \varnothing & \varnothing \\ 1 & 1 & \varnothing & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & \varnothing & 2 \\ 1 & 1 & \varnothing & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & \varnothing & 1 \\ 2 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 & 1 \end{array} \right\}$$

with corresponding posterior probabilities

$$\left\{ \begin{array}{ccccccc} 0.499 & 0.349 & 0.044 & 0.044 & 0.031 & 0.022 & 0.002 \end{array} \right\}.$$

The top 7 hypotheses generated by the FISST-MHT are

$$\left\{ \begin{array}{c|cc|cc|cc|cc|cc|cc} 1 & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing & 1 & \varnothing \\ 1 & 1 & \varnothing & 1 & \varnothing & 1 & 2 & 1 & 2 & 1 & \varnothing & 1 & 2 \\ 1 & 1 & \varnothing & 1 & 0 & 1 & 0 & 0 & 1 & 1 & \varnothing & 1 & 0 \\ 2 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 0 & 2 & 0 \end{array} \right\}$$

with corresponding posterior probabilities

$$\left\{ \begin{array}{ccccccc} 0.373 & 0.261 & 0.078 & 0.033 & 0.033 & 0.028 & 0.024 \end{array} \right\}.$$

That is, the top hypothesis of both approaches is $\theta_4^{(1)}$, with posterior probability 0.499 and 0.373 under Reid's MHT and the FISST-MHT, respectively. The hypothesis $\theta_4^{(2)}$ comes as number 3 and 4 under the two MHT's, while the hypothesis $\theta_4^{(3)}$ comes as number 4 and 5. Hypothesis number 6 of the FISST-MHT contains an unobserved target, while hypothesis number 3 of the FISST-MHT contains a target which is born before it is observed. Neither is allowed in Reid's MHT. Except for these two hypotheses, the top 5 hypotheses are identical for both approaches.

The top hypotheses of the FISST-MHT naturally score lower probabilities than the top hypotheses in Reid's MHT, since the FISST-MHT must distribute the probability mass over more hypotheses. Therefore, a direct comparison between the two collections of posterior probabilities is not entirely meaningful. Instead, we should for each Reid-hypothesis θ sum together the probabilities of all FISST-MHT hypotheses which agree with θ insofar as non-Reid elements are ignored. For example, both the first and the sixth FISST-MHT hypotheses above agree with the first Reid-hypothesis in this sense. Merging FISST-MHT hypotheses in this manner yields the following list of *aggregate* posterior probabilities for the FISST-MHT hypotheses corresponding to the top 7 Reid-hypotheses:

$$\left\{ \begin{array}{ccccccc} 0.414 & 0.410 & 0.048 & 0.047 & 0.034 & 0.043 & 0.002 \end{array} \right\}.$$

Finally, for this particular case, we find that only 0.001 of the FISST-MHT's probability mass is assigned to hypotheses not agreeing with any Reid-hypotheses.

Let us then look at how the hypothesis probabilities behave as functions of the birth rate μ. In Figure 3 we have displayed the probability of the generally most plausible hypothesis $\theta_4^{(1)}$ for various μ under both Reid's MHT and the FISST-MHT. Furthermore, the the dashed line also displays the aggregate probability of FISST-MHT hypotheses which agree with θ_4^1. This figure illustrates how the two approaches become identical as $\mu \to 0$. The discrepancy between the aggregate probabilities and the probabilities in Reid's MHT for higher values of μ can be understood from the following perspective: Reid's MHT treats newly discovered targets as newborn, while the FISST-MHT also allows for newly discovered targets to have been born at earlier time steps. Thus, the FISST-MHT gets a higher effective birth rate, which amounts to shifting the red curve rightwards, where its values are lower.

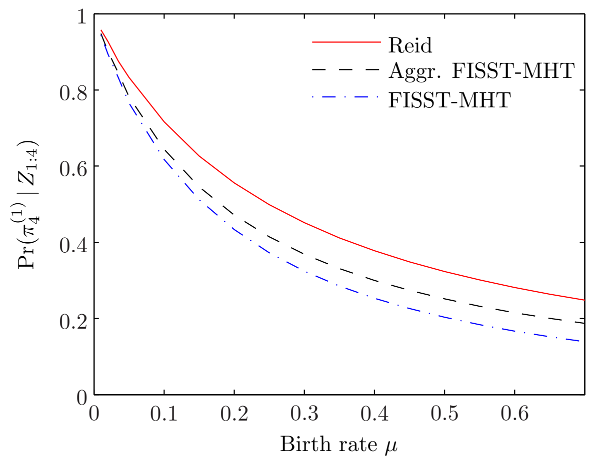

Fig. 3: Posterior probability of the particular hypothesis $\theta_4^{(1)}$ as a function of the target birth rate μ.

A. Details of hypothesis enumeration

As a final example, let us study in greater detail how association hypotheses and association variables are generated from a parent hypothesis. Assume that we somehow know that

$$\theta_2^* = \left\{ \begin{array}{cc} 1 & \varnothing \\ 1 & 2 \end{array} \right\}$$

is the one and only true hypothesis at time step 2. That is, we know that one target originated at $k = 1$ and caused the measurements z_1^1 and z_2^1, and that another target originated at $k = 2$ and caused the measurement z_2^2.

Conditional on this hypothesis, the set $\mathcal{A}(\theta_2^*, n_3)$ of predicted association variables contains 0 elements for $n_3 = 0$ and $n_3 = 1$. It contains 2 elements for $n_3 = 2$ and 6 elements for $n_3 = 3$. It contains 12 elements for $n_3 = 4$, and so on. Looking more closely at the case $n_3 = 4$, we find that $\mathcal{A}(\theta_2^*, n_3)$ contains the following association variables:

$$\left\{ \begin{bmatrix} \varnothing\varnothing 1 \varnothing \\ \varnothing\varnothing 1 2 \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing\varnothing\varnothing 1 \\ \varnothing\varnothing 2 1 \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing 1 \varnothing\varnothing \\ \varnothing 1 \varnothing 2 \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing 1 \varnothing\varnothing \\ \varnothing 1 2 \varnothing \\ 0000 \end{bmatrix}, \right.$$
$$\begin{bmatrix} \varnothing\varnothing\varnothing 1 \\ \varnothing 2 \varnothing 1 \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing\varnothing 1 \varnothing \\ \varnothing 2 1 \varnothing \\ 0000 \end{bmatrix}, \begin{bmatrix} 1\varnothing\varnothing\varnothing \\ 1\varnothing\varnothing 2 \\ 0000 \end{bmatrix}, \begin{bmatrix} 1\varnothing\varnothing\varnothing \\ 1\varnothing 2\varnothing \\ 0000 \end{bmatrix},$$
$$\left. \begin{bmatrix} 1\varnothing\varnothing\varnothing \\ 1 2 \varnothing\varnothing \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing\varnothing\varnothing 1 \\ 2\varnothing\varnothing 1 \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing\varnothing 1 \varnothing \\ 2\varnothing 1 \varnothing \\ 0000 \end{bmatrix}, \begin{bmatrix} \varnothing 1 \varnothing\varnothing \\ 2 1 \varnothing\varnothing \\ 0000 \end{bmatrix} \right\}.$$

The product set $\mathcal{P}(\theta_2^*, n_3, m_3)$ contains 60 posterior association variables for $n_3 = 4$ and $m_3 = 1$ [1]:

$$\left\{ \begin{bmatrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,0\,0 \end{bmatrix}, \begin{bmatrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,0\,1 \end{bmatrix}, \begin{bmatrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,1\,0 \end{bmatrix}, \begin{bmatrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,1\,0\,0 \end{bmatrix}, \ldots \right.$$
$$\left. \ldots, \begin{bmatrix} \emptyset\,1\,\emptyset\,\emptyset \\ 2\,1\,\emptyset\,\emptyset \\ 0\,0\,0\,1 \end{bmatrix}, \begin{bmatrix} \emptyset\,1\,\emptyset\,\emptyset \\ 2\,1\,\emptyset\,\emptyset \\ 0\,0\,1\,0 \end{bmatrix}, \begin{bmatrix} \emptyset\,1\,\emptyset\,\emptyset \\ 2\,1\,\emptyset\,\emptyset \\ 0\,1\,0\,0 \end{bmatrix}, \begin{bmatrix} \emptyset\,1\,\emptyset\,\emptyset \\ 2\,1\,\emptyset\,\emptyset \\ 1\,0\,0\,0 \end{bmatrix} \right\}.$$

On the other hand, the quotient set $\mathcal{I}(\theta_2^*, n_3, m_3)$ contains only 4 different association hypotheses for $n_3 = 4$ and $m_3 = 1$:

$$\left\{ \left\{ \begin{matrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,0\,0 \end{matrix} \right\}, \left\{ \begin{matrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,0\,1 \end{matrix} \right\}, \left\{ \begin{matrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,0\,1\,0 \end{matrix} \right\}, \left\{ \begin{matrix} \emptyset\,\emptyset\,1\,\emptyset \\ \emptyset\,\emptyset\,1\,2 \\ 0\,1\,0\,0 \end{matrix} \right\} \right\}.$$

Notice that only the last hypothesis in $\mathcal{I}(\theta_2^*, n_3, m_3)$ can be interpreted as a set, while the other hypotheses are more appropriately interpreted as multisets due to repeated tracks.

VIII. Conclusion

In [17], Mahler argues against several misconceptions about FISST. Misconception number 10 reads "*The right model of the multitarget state is that used in the multihypothesis tracking (MHT) paradigm, not the RFS paradigm*". In this paper we have, in contrast, shown that there is no conflict between the two paradigms to begin with. Reid's MHT is an approximation of the multitarget Bayes filter, obtained by ignoring target death and ignoring the existence of unobserved targets. In the special case with no target births and no target deaths, and Gaussian-linear kinematics, Reid's MHT is identical to the multitarget Bayes filter if all hypotheses are retained. Also, if all hypotheses involving unknown targets are discarded from the multitarget Bayes filter, then Reid's MHT results. In a related conference paper [20] we have also shown that Reid's MHT follows exactly if it is provided with the unknown target density from the factorization of the multiobject posterior that was proposed in [10]. A key observation underlying these conclusions is that Reid's MHT does not assign measurements to targets, but only to tracks, which here are defined as temporal sequences of measurements. Based on this, a case can be made that practical MHT methods are just as principled approximations of the multitarget Bayes filter as the PHD, CPHD and MeMBer filters. The approximations are of a very different nature, and the suitability of the different approximations may depend on the application.

Both the MHT formulations proposed in [20] and in this paper may be used as a starting point for research on practical MHT methods. Further inspiration from FISST, or generalizations of FISST such as [27], together with techniques for hypothesis aggregation, as studied in, e.g., [19], will be interesting research topics for the near future.

Acknowledgment

The authors wish to thank the reviewers for several suggestions which significantly improved the paper. This work was partially supported by the Research Council of Norway through projects 223254 (Centre for Autonomous Marine Operations and Systems at NTNU) and 244116/O70 (Sensor Fusion and Collision Avoidance for Autonomous Marine Vehicles).

[1]It can be shown that $|\mathcal{E}(n,m)| = |\mathcal{E}(m,n)| = m|\mathcal{E}(n-1,m-1)| + |\mathcal{E}(n-1,m)|$ and that $|\mathcal{E}(n,1)| = n+1$.

Appendix A
Derivation of the predicted density

In this first appendix we show that the predicted density $f_{k|k-1}(X_k \mid Z_{1:k-1})$ is as given by (10) in Section V-C.

To begin with, we evaluate the prediction integral in the Bayes recursion (3) by inserting the multitarget Markov model (5) and the prior (7) into (3). This leads to a quadruple sum over prior target cardinalities n_{k-1}, association variables γ of the multiobject transition density, prior association hypotheses θ_{k-1} and prior association variables ω_{k-1}:

$$\begin{aligned} &f_{k|k-1}(X_k \mid Z_{1:k-1}) \\ &\quad = \int f_{\Xi_k|\Xi_{k-1}}(X_k \mid X_{k-1}) f_{k-1|k-1}(X_{k-1} \mid Z_{k-1}) \delta X_{k-1} \\ &\quad = e^{-\mu} \sum_{n_{k-1}=0}^{\infty} \frac{1}{n_{k-1}!} \sum_{\gamma} \mu^{n_k - n_{k-1}} \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}^{n_{k-1}}} q^{\theta_{k-1}} \\ &\qquad \sum_{\omega_{k-1} \in \mathrm{Per}(\theta_{k-1})} \prod_{i\,:\,\gamma(i)=0} b(x_k^i) \\ &\qquad \int \prod_{i\,:\,\gamma(i)>0} f(x_k^i \mid x_{k-1}^{\gamma(i)}) \\ &\qquad \prod_{s=1}^{n_{k-1}} f^{\omega_{k-1}(s)}(x_{k-1}^s) \mathrm{d}x_{k-1}^1 \cdots \mathrm{d}x_{k-1}^{n_{k-1}}. \end{aligned} \tag{31}$$

It is possible to simplify this expression a great deal. By exploiting the function $g^{\pi_k^{1:k-1}(i)}(\cdot)$ defined in (11) we can write the kinematic integral in (31) more compactly:

$$\begin{aligned} I &= \int \prod_{i\,:\,\gamma(i)>0} f(x_k^i \mid x_{k-1}^{\gamma(i)}) \\ &\qquad \prod_{s=1}^{n_{k-1}} f^{\omega_{k-1}(s)}(x_{k-1}^s)\, \mathrm{d}x_{k-1}^1 \cdots \mathrm{d}x_{k-1}^{n_{k-1}} \\ &= \int \prod_{i\,:\,\gamma(i)>0} f(x_k^i \mid x_{k-1}^{\gamma(i)}) \\ &\qquad \prod_{i\,:\,\gamma(i)>0} f^{\omega_{k-1}(\gamma(i))}(x_{k-1}^{\gamma(i)})\, \mathrm{d}x_{k-1}^1 \cdots \mathrm{d}x_{k-1}^{n_{k-1}} \\ &= \prod_{i\,:\,\gamma(i)>0} \int f(x_k^i \mid x_{k-1}^{\gamma(i)}) f^{\omega_{k-1}(\gamma(i))}(x_{k-1}^{\gamma(i)})\, \mathrm{d}x_{k-1}^i \\ &= \prod_{i\,:\,\gamma(i)>0} g^{\omega_{k-1}(\gamma(i))}(x_k^i). \end{aligned} \tag{32}$$

The second equality of (32) follows because

$$\prod_{s=1}^{n_{k-1}} f^{\omega_{k-1}(s)}(x_{k-1}^s) = \prod_{i\,:\,\gamma(i)>0} f^{\omega_{k-1}(\gamma(i))}(x_{k-1}^{\gamma(i)}). \tag{33}$$

The third equality follows because the multiple integral over all x_{k-1}^i can be factorized into single integrals over each x_{k-1}^i, and the fourth equality is simply due to the definition of $g^{\pi_k^{1:k-1}(i)}(\cdot)$.

In order to bring (31) onto the same form as our proposed expression (10) for the predicted density, we need to reduce the double sum over γ and ω_{k-1} to a single sum over predicted association variables π_k. For this, we provide an alternative definition of π_k, uniquely given by γ and ω_{k-1}, as follows:

$$\pi_k(i) = \begin{cases} [\omega_{k-1}(\gamma(i))^\mathsf{T}, 0]^\mathsf{T} & \text{if } \gamma(i) > 0 \\ [\varnothing, \ldots, \varnothing, 0]^\mathsf{T} & \text{if } \gamma(i) = 0. \end{cases} \tag{34}$$

In order to show that the set of π_k's defined according to (34) is identical to the set $\mathcal{A}(\theta_{k-1}, n_k)$, we show that these π_k's satisfy the requirements in the definition of $\mathcal{A}(\theta_{k-1}, n_k)$, and that all elements in $\mathcal{A}(\theta_{k-1}, n_k)$ must satisfy (34).

Regarding the first requirement in the definition of $\mathcal{A}(\theta_{k-1}, n_k)$, it is clear that for any π_k as defined in (34), all its tracks $\pi_k(i)$ must be on one of the two allowed forms.

Regarding the second requirement, the restricted surjectivity of γ ensures that there is some $i \in \{1, \ldots, n_k\}$ such that $\gamma(i) = s$ for all $s \in \{1, \ldots, n_{k-1}\}$. Consequently, there is one $i \in \{1, \ldots, n_k\}$ such that $\pi_k^{1:k-1}(i) = \omega_{k-1}(s)$ for all tracks $\omega_{k-1}(s)$ in any chosen ordering of θ_{k-1}. This applies also for repeated tracks. Therefore, there are exactly $r_t(\theta_{k-1})$ targets i such that $\pi_k^{1:k-1}(i) = t$ for any distinct track $t \in \theta_{k-1}$ as required. Conversely, for any element $\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)$ we can find a function π_k^* given by (34) such that $\pi_k = \pi_k^*$.

We have thus shown that the sum over γ and ω_{k-1} can be written as a sum over $\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)$. However, there is not a one-to-one correspondence between terms in this double sum and elements in $\mathcal{A}(\theta_{k-1}, n_k)$, since several combinations of γ and ω_{k-1} will lead to the same π_k. We find this number as

$$\#(\omega_{k-1}, \gamma \mid \pi_k, \theta_{k-1}) = \#(\omega_{k-1} | \theta_{k-1}) \#(\gamma | \omega_{k-1}, \pi_k).$$

We recall that

$$\#(\omega_{k-1} | \theta_{k-1}) = \frac{n_{k-1}!}{\prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!}.$$

All of these ω_{k-1}'s are compatible with the given π_k if a suitable γ is chosen. Without the possibility of repeated tracks, γ would be given uniquely by the combination of ω_{k-1} and π_k. However, the presence of repeated tracks means that several γ's may be compatible, and the number of compatible γ's is given by

$$\#(\gamma | \omega_{k-1}, \pi_k) = \prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})! \tag{35}$$

This is illustrated in Example 2. It follows from this that

$$\#(\omega_{k-1}, \gamma \mid \pi_k, \theta_{k-1}) = n_{k-1}!. \tag{36}$$

We can now return to the summations in (31), where (36) enables us to cancel the $n_{k-1}!$ from the denominator of the set integral:

$$\frac{1}{n_{k-1}!} \sum_{\gamma} \sum_{\omega_{k-1} \in \mathrm{Per}(\theta_{k-1})} \mu^{n_k - n_{k-1}} \prod_{i : \gamma(i) = 0} b(x_k^i) \prod_{i : \gamma(i) > 0} g^{\omega_{k-1}(\gamma(i))}(x_k^i)$$
$$= \sum_{\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)} \mu^{n_k - n_{k-1}} \prod_{i=1}^{n_k} g^{\pi_k^{1:k-1}(i)}(x_k^i)$$

Thus we arrive at the desired expression

$$f_{k|k-1}(X_k \mid Z_{1:k-1}) = e^{-\mu} \sum_{\theta_{k-1} \in \mathcal{G}_{k-1}} q^{\theta_{k-1}} \mu^{n_k - n_{k-1}} \sum_{\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)} \prod_{i=1}^{n_k} g^{\pi_k^{1:k-1}(i)}(x_k^i).$$

EXAMPLE 2 (The number of possible permutation mappings)**.** Here we provide an illustratory example of how the formula $\#(\gamma | \omega_{k-1}, \pi_k) = \prod_{t \in \mathrm{Im}(\theta_{k-1})} r_t(\theta_{k-1})!$ works. Let the predicted association variable be

$$\pi_k = \begin{bmatrix} 1 & 0 & 0 & \varnothing & \varnothing \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and let the prior association variable be

$$\omega_{k-1} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

These are compatible if and only if the permutation mapping γ is one of the following:

$$\begin{bmatrix} 3 & 1 & 2 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 3 & 2 & 1 & 0 & 0 \end{bmatrix}.$$

We see that the number of compatible permutation mappings is 2! as claimed by the formula.

APPENDIX B
DERIVATION OF THE POSTERIOR DENSITY

In this second appendix we show that the posterior density is given by (14) in Section V-D. Inserting the likelihood (6) and the predicted density (10) into Bayes rule yields a triple sum over prior association hypotheses θ_{k-1}, predicted association variables π_k and association variables ξ of the likelihood:

$$f_{k|k}(X_k \mid Z_{1:k}) = \frac{1}{C} f_{\Sigma_k|\Xi_k}(Z_k \mid X_k) f_{\Xi_k|\Sigma_{k-1}}(X_k \mid Z_{1:k-1})$$
$$= \frac{1}{C} \Bigg[\sum_{\xi \in \mathcal{E}(n_k, m_k)} \lambda^{\varphi_k} P_\mathrm{D}^{\delta_k} (1 - P_\mathrm{D})^{n_k - \delta_k} \prod_{j \notin \mathrm{Im}^+(\xi)} c(z_k^j) \prod_{i : \xi(i) > 0} f(z_k^{\xi(i)} \mid x_k^i) \Bigg]$$
$$\Bigg[\sum_{\theta \in \mathcal{G}_{k-1}} q^{\theta_{k-1}} \sum_{\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)} \mu^{n_k - n_{k-1}} \prod_{s=1}^{n_k} g^{\pi_k^{1:k-1}(s)}(x_k^s) \Bigg]$$
$$= \frac{1}{C} \sum_{\theta \in \mathcal{G}_{k-1}} q^{\theta_{k-1}} \sum_{\pi_k \in \mathcal{A}(\theta_{k-1}, n_k)} \sum_{\xi \in \mathcal{E}(n_k, m_k)} \lambda^{\varphi_k} \mu^{n_k - n_{k-1}} P_\mathrm{D}^{\delta_k} (1 - P_\mathrm{D})^{n_k - \delta_k} \prod_{j \notin \mathrm{Im}^+(\xi)} c(z_k^j)$$
$$\prod_{i : \xi(i) > 0} f(z_k^{\xi(i)} \mid x_k^i) \prod_{s=1}^{n_k} g^{\pi_k^{1:k-1}(s)}(x_k^s). \tag{37}$$

The last two sums in (37) can be replaced by a single sum over elements in the product set of $\mathcal{A}(\theta_{k-1}, n_k)$ and $\mathcal{E}(n_k, m_k)$. This product set is isomorphic to the set $\mathcal{P}(\theta_{k-1}, n_k, m_k)$ that was introduced in Section V-D. To see this, let us define the association variable ω_k according to

$$\omega_k(i) = [\pi_k^{1:k-1}(i)^\mathsf{T}, \xi(i)]^\mathsf{T} \tag{38}$$

Any such ω_k satisfies the two criteria for membership in $\mathcal{P}(\theta_{k-1}, n_k, m_k)$, and all elements $\mathcal{P}(\theta_{k-1}, n_k, m_k)$ can be constructed in this manner. For any element in $\mathcal{P}(\theta_{k-1}, n_k, m_k)$ the corresponding π_k and ξ are given uniquely. This allows us to write the posterior density as

$$f_{k|k}(X_k \mid Z_{1:k}) = \frac{1}{C} \sum_{\theta_{k-1}\in\mathcal{G}_{k-1}} q^{\theta_{k-1}} \sum_{\omega_k\in\mathcal{P}(\theta_{k-1},n_k,m_k)} \lambda^{\varphi_k} \mu^{n_k-n_{k-1}} P_{\mathrm{D}}^{\delta_k}(1-P_{\mathrm{D}})^{n_k-\delta_k} \prod_{j\notin \mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i\,:\,\omega_k^k(i)>0} f(z_k^{\omega_k^k(i)} \mid x_k^i) \prod_{s=1}^{n_k} g^{\omega_k^{1:k-1}(s)}(x_k^s). \tag{39}$$

According to the definition of the posterior hypothesis collection $\mathcal{I}(\theta_{k-1}, n_k, m_k)$, the second sum in (39) can also be written in terms of a double sum over the sets $\mathcal{I}(\theta_{k-1}, n_k, m_k)$ and $\mathrm{Per}(\theta_k)$. Furthermore, since the sets $\mathcal{I}(\theta_{k-1}, n_k, m_k)$ are mutually disjoint over θ_{k-1}, we can replace the sums over $\mathcal{G}_{k-1}$ and $\mathcal{I}(\theta_{k-1}, n_k, m_k)$ by a single sum over $\mathcal{G}_k^{n_k}$ as defined in (13). From these manipulations, (39) becomes a double sum of the same form as (14).

The last two products in (39) can be rewritten as two products over $i \in \{1, \ldots, n_k\}$ such that $\omega_k^k(i) = 0$ and $\omega_k^k(i) > 0$, respectively. This is done by writing the products in terms of the posterior track density $h^{\omega_k(i)}(x_k^i)$ and the score contribution $a^{\omega_k(i)}$:

$$\prod_{i\,:\,\omega_k^k(i)>0} f(z_k^{\omega_k^k(i)} \mid x_k^i) \prod_{s=1}^{n_k} g^{\omega_k^{1:k-1}(s)}(x_k^s) = \prod_{i\,:\,\omega_k^k(i)=0} h^{\omega_k(i)}(x_k^i) \prod_{i\,:\,\omega_k^k(i)>0} a^{\omega_k(i)} h^{\omega_k(i)}(x_k^i).$$

Thus, the posterior becomes

$$f_{k|k}(X_k|Z_{1:k}) = \frac{1}{C} \sum_{\theta_k\in\mathcal{G}_k^{n_k}} q^{\theta_{k-1}} \sum_{\omega_k\in\mathrm{Per}(\theta_k)} \lambda^{\varphi_k} \mu^{n_k-n_{k-1}} P_{\mathrm{D}}^{\delta_k}(1-P_{\mathrm{D}})^{n_k-\delta_k} \prod_{j\notin \mathrm{Im}^+(\omega_k^k)} c(z_k^j) \prod_{i\,:\,\omega_k^k(i)=0} h^{\omega_k(i)}(x_k^i) \prod_{i\,:\,\omega_k^k(i)>0} a^{\omega_k(i)} h^{\omega_k(i)}(x_k^i). \tag{40}$$

The final expression (14) follows from (40) when all the non-pdf terms in (40) are collected into the scalar q^{θ_k} as defined in (15).

REFERENCES

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

[2] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.

[3] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[4] I. R. Goodman, H. T. Nguyen, and R. Mahler, *Mathematics of Data Fusion*. Springer, 1997.

[5] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, May 2003.

[6] ——, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.

[7] ——, "Multitarget moments and their application to multitarget tracking," in *Proceedings of Workshop on Estimation, Tracking, and Fusion: A Tribute to Y. Bar-Shalom*, Naval Postgraduate School, Monterey, CA, USA, May 2001, pp. 134–166.

[8] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 409–423, Feb. 2009.

[9] S. Challa, M. R. Morelande, D. Musicki, and R. J. Evans, *Fundamentals of object tracking*. Cambridge University Press, 2011.

[10] J. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBer," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1664–1687, July 2015.

[11] L. Svensson, D. Svensson, M. Guerriero, and P. Willett, "Set JPDA filter for multitarget tracking," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4677–4691, Oct. 2011.

[12] M. Ulmke, O. Erdinc, and P. K. Willett, "Gaussian mixture cardinalized PHD filter for ground moving target tracking," in *Proceedings of Fusion 2007*, Quebec, Canada, July 2007.

[13] R. Streit, "How I learned to stop worrying about a thousand and one filters and love analytic combinatorics," in *Proceedings of IEEE Aerospace Conference*, Big Sky, MT, USA, Mar. 2017.

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

[15] S. Mori, C.-Y. Chong, and K. C. Chang, "Three formalisms of multiple hypothesis tracking," in *Proceedings of FUSION*, Heidelberg, Germany, Jul. 2016.

[16] B.-T. Vo, "Random finite sets in multi-object filtering," Ph.D. dissertation, The University of Western Australia, 2008.

[17] R. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2014.

[18] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Applications and Advances*, Y. Bar-Shalom and W. D. Blair, Eds. Artech House, 1990, ch. 3, pp. 219–245.

[19] S. Coraluppi and C. A. Carthel, "If a tree falls in the woods, it does make a sound: multiple-hypothesis tracking with undetected target births," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2379–2388, July 2014.

[20] E. F. Brekke and M. Chitre, "The multiple hypothesis tracker derived from finite set statistics," in *2017 20th International Conference on Information Fusion (Fusion)*, July 2017, pp. 1–8.

[21] E. Brekke and M. Chitre, "Bayesian multi-hypothesis scan matching," in *Proceedings of OCEANS'13*, Bergen, Norway, Jun. 2013.

[22] ——, "A multi-hypothesis solution to data association for the two-frame SLAM problem," *The International Journal of Robotics Research*, vol. 34, no. 1, pp. 43–63, Jan. 2015.

[23] E. Brekke, B. Kalyan, and M. Chitre, "A novel formulation of the Bayes recursion for single-cluster filtering," in *Proceedings of IEEE Aerospace Conference*, Big Sky, MT, USA, Mar. 2014.

[24] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT, USA: YBS Publishing, 2011.

[25] B.-T. Vo and B.-N. Vo, "A random finite set conjugate prior and application to multi-target tracking," in *Proceedings of ISSNIP 2011*, December 2011, pp. 431–436.

[26] ——, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, Jul. 2013.

[27] A. Garcia-Fernandez, L. Svensson, and M. Morelande, "Multiple target tracking based on sets of trajectories," 2016, arXiv e-print.

[28] E. H. Aoki, P. K. Mandal, L. Svensson, Y. Boers, and A. Bagchi, "Labeling uncertainty in multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 3, pp. 1006–1020, June 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2018.2805178, IEEE Transactions on Aerospace and Electronic Systems

Edmund Brekke (S'10-M'11) has an MSc in Industrial Mathematics (2005) and a PhD in Engineering Cybernetics (2010), both awarded by NTNU in Norway. His PhD research covered methods for single-target tracking in clutter. After his PhD studies, Brekke worked as a postdoctoral research fellow at the Acoustic Research Lab at NUS in Singapore. He was appointed an Associate Professor in Sensor Fusion at NTNU in 2014. Brekke's main research interests lie in Bayesian estimation, with applications in target tracking and autonomous vehicles.

Mandar Chitre (S'04-M'05-SM'11) received the 1375 B.Eng. and M.Eng. degrees in electrical engineering from the National University of Singapore (NUS), Singapore, in 1997 and 2000, respectively, the M.Sc. degree in bioinformatics from the Nanyang Technological University (NTU), Singapore, i 2004, and the Ph.D. degree in underwater communications from NUS, in 2006. From 1997 to 1998, he was with the ARL, NUS, Singapore. From 1998 to 2002, he headed the technology division of a regional telecommunications solutions company. In 2003, he rejoined ARL, initially as the Deputy Head (Research) and is currently the Head of the laboratory. He also holds a joint appointment with the Department of Electrical and Computer Engineering, NUS, as a Associate Professor. His current research interests include underwater communications, autonomous underwater vehicles, model-based signal processing, and modeling of complex dynamic systems. Dr. Chitre has served on the technical program committees of the IEEE OCEANS, WUWNet, DTA, and OTC conferences and has served as a reviewer for numerous international journals. He serves as an the Editor-in-Chief for the IEEE Journal of Oceanic Engineering, and has also severed on editorial boards for the IEEE Communications Magazine and the IEEE ICRA conference. He was the Chairman of the Student Poster Committee for IEEE OCEANS'06 in Singapore, and the Chairman for the IEEE Singapore AUV Challenge 2013. He is currently the IEEE Ocean Engineering Society Technology Committee Co-Chair of underwater communication, navigation, and positioning.