A Theoretical Foundation for the Stein-Winter "Probability Hypothesis Density (PHD)" Multitarget Tracking Approach

June 20, 2000

Ronald P.S. Mahler

Lockheed Martin NE&SS, 3333 Pilot Knob Road, Eagan MN 55121

651-456-4819; ronald.p.mahler@lmco.com

ABSTRACT

In several unpublished manuscripts written from 1993 to 1995, Michael Stein, C.L. Winter, and Robert Tenney introduced a multitarget tracking and evidential-accumulation concept called a "Probability Hypothesis Surface" (PHS). A PHS is the graph of a probability distribution—the *Probability Hypothesis Density (PHD)*—that, when integrated over a region in target state space, gives the expected number of targets in that region. The PHD is uniquely defined by this property: Any other density function that satisfies it must be the PHD. In particular, the PHD is the *expected value* of the *point process* of a random track-set—i.e., of the density that, when integrated over a region in state space, gives the exact (random) number of targets in that region. In 1997 in the book *Mathematics of Data Fusion* I sketched the elements of a theoretical foundation for PHS/PHD. The purpose of this paper is to publish a full account of this material for the first time. We show that the PHD is a first-order moment statistic of the random multitarget process and, consequently that from a computational perspective it is a multitarget analog of single-target constant-gain Kalman filters such as the α-β-γ filter.[1]

1.0 INTRODUCTION

In several unpublished manuscripts written during the period from 1993 to 1995, Michael Stein and C.L. Winter (Los Alamos National Laboratory) and Robert Tenney (Alphatech Corp.) introduced a multitarget tracking and evidential-accumulation concept "called a "Probability Hypothesis Surface" or "PHS" **[20,19]**. A PHS is the graph of a certain unnormalized probability distribution—the PHD or Probability Hypothesis Density $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$—that has the following property: Given any region S in target state space, the integral $\int_S \hat{D}_{k|k}(\mathbf{x}|Z^{(k)})d\mathbf{x}$ is the expected number of targets contained in S. This property characterizes the PHD *uniquely.* That is, if $g_k(\mathbf{x})$ is any other density which gives the expected number of targets in S when integrated over S, then it is (no matter how imaginative the name one might assign to it) nothing else but the PHD. For, since $\int_S g_k(\mathbf{x})d\mathbf{x} = \int_S \hat{D}_{k|k}(\mathbf{x}|Z^{(k)})d\mathbf{x}$ for all measurable S then $g_{k|k} = \hat{D}_{k|k}$ almost everywhere. In particular, the PHD is the expected value of the *point process* of a random track-set—i.e., of the density that, when integrated over a region in state space, gives the exact

The work in this paper was supported by the U.S. Army Research Office under contracts DAAH04-94-C-0011 and DAAG55-98-C-0039. The contents do not necessarily reflect the position or policy of the Government. No official endorsement should be inferred.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)
20-06-2000

2. REPORT TYPE
Conference Proceedings

3. DATES COVERED (FROM - TO)
xx-xx-2000 to xx-xx-2000

4. TITLE AND SUBTITLE
A Theoretical Foundation for the Stein- Winter " Probability Hypothesis Density (PHD)" Multitarget Tracking Approach
Unclassified

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Mahler, Ronald P. ;

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Lockheed Martin NE&SS
3333 Pilot Knob Road
Eagan, MN55121

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
Director, CECOM RDEC
Night Vision and Electronic Sensors Directorate, Security Team
10221 Burbeck Rd.
Ft. Belvoir, VA22060-5806

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
See Also ADM201258, 2000 MSS Proceedings on CD-ROM, January 2001

14. ABSTRACT
In several unpublished manuscripts written from 1993 to 1995, Michael Stein, C.L. Winter, and Robert Tenney introduced a multitarget tracking and evidential-accumulation concept called a "Probability Hypothesis Surface" (PHS) .A PHS is the graph of a probability distribution-the Probability Hypothesis Density (PHD)-that, when integrated over a region in target state space, gives the expected number of targets in that region. The PHD is uniquely defined by this property: Any other density function that satisfies it must be the PHD. In particular , the PHD is the expected value of the point process of a random track-set-i.e. , of the density that, when integrated over a region in state space, gives the exact (random) number of targets in that region. In 1997 in the book Mathematics of Data Fusion I sketched the elements of a theoretical foundation for PHS/PHD. The purpose of this paper is to publish a full account of this material for the first time. We show that the PHD is a first-order moment statistic of the random multitarget process and, consequently that from a computational perspective it is a multitarget analog of single-target constant-gain Kalman filters such as the a-fJ-'Y filter.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE
Unclassified	Unclassified	Unclassified

17. LIMITATION OF ABSTRACT
Same as Report (SAR)

18. NUMBER OF PAGES
20

19. NAME OF RESPONSIBLE PERSON
Fenster, Lynn
lfenster@dtic.mil

19b. TELEPHONE NUMBER
International Area Code
Area Code Telephone Number
703767-9007
DSN
427-9007

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

(random) number of targets in that region (see Section 2.4). Stein and Winter devised the PHD concept in part as a structure for a new inference technique called *Weak Evidence Accrual (WEA)* that exploits the additive rather than multiplicative properties of Bayes' rule. We will not discuss this aspect of the PHD approach in great detail here (see, however, Theorem 5). Rather, we will discuss the potential significance of the PHD as a computational strategy for Bayes-optimal multitarget filtering—specifically, as a *a multitarget-tracking analog of constant-gain Kalman filters such as the α-β-γ filter* that results in additive information-update rules of the WEA type.

In 1997 in section 4.3.4, pages 168-170 of the book *Mathematics of Data Fusion* **[7]** I sketched the elements of a theoretical foundation for PHD based on the "finite-set statistics (FISST)" approach described in Chapters 2 and 4 through 8 of that book. Because of page limitations, the full description of this work—specifically, the proofs of the various assertions—had to be cut from the final draft of the book. The purpose of this paper is to publish this material in the open literature for the first time, as well as to show how FISST tools such as the set derivative can be used to develop a Bayes filtering scheme for PHD's. Because of space limitations, it is not possible to provide a summary of FISST and the FISST calculus in this paper. See the monograph *An Introduction to Multisource-Multitarget Statistics and Its Applications* **[11]** and the book chapter *Multisensor-Multitarget Statistics* **[13]** for more details.

1.1 APPROXIMATION WITH STATISTICAL MOMENTS

The theoretical starting-point of *single-target* tracking is the following *Bayesian discrete-time recursive nonlinear filtering equations* (see [8], [17], **[2, pp. 373-377]**, and **[9, p. 174]**:

$$
\begin{aligned}
f_{k+1|k}(\mathbf{x}_{k+1}|Z^k) &= \int f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{x}_k)\, f_{k|k}(\mathbf{x}_k|Z^k) d\mathbf{x}_k \\
f_{k+1|k+1}(\mathbf{x}_{k+1}|Z^{k+1}) &= \frac{f(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})\, f_{k+1|k}(\mathbf{x}_{k+1}|Z^k)}{f(\mathbf{z}_{k+1}|Z^k)} \\
\hat{\mathbf{x}}^{MAP}_{k+1|k+1} &= \arg\sup_{\mathbf{x}} f_{k+1|k+1}(\mathbf{x}|Z^{k+1}), \qquad \hat{\mathbf{x}}^{EAP}_{k+1|k+1} = \int \mathbf{x} \cdot f_{k+1|k+1}(\mathbf{x}|Z^{k+1}) d\mathbf{x}
\end{aligned}
$$

where

(1) $\mathbf{x}_k$ is the target state variable at time-step k and $\mathbf{z}_k$ is the observed measurement at time-step k;

(2) $f_{k|k}(\mathbf{x}_k|Z^k)$ is the Bayes posterior distribution conditioned on the data-stream $Z^k = \{\mathbf{z}_1, ..., \mathbf{z}_k\}$;

(3) $f(\mathbf{z}|\mathbf{x})$ is the sensor likelihood function;

(4) $f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{x}_k)$ is the target Markov transition density that models between-measurements target motion;

(5) $f_{k+1|k}(\mathbf{x}_{k+1}|Z^k)$ is the time-prediction of the posterior $f_{k|k}(\mathbf{x}_k|Z^k)$ to time-step $k+1$;

(6) $(\mathbf{z}_{k+1}|Z^k) = \int f(\mathbf{z}_{k+1}|\mathbf{y}_{k+1})\, f_{k+1|k}(\mathbf{y}_{k+1}|Z^k) d\mathbf{y}_{k+1}$ is the Bayes normalization constant; and

(7) $\hat{\mathbf{x}}^{MAP}_{k+1|k+1}$ and $\hat{\mathbf{x}}^{EAP}_{k+1|k+1}$ are the Bayes-optimal maximum *a posteriori* (MAP) and expected *a posteriori* (EAP) state estimators, respectively.

In all of these formulas, data and state vectors have the form $\mathbf{y} = (y_1, ..., y_n, w_1, ..., w_n)$ where $y_1, ..., y_n$ are continuous variables and $w_1, ..., w_n$ are discrete variables, and we denote the space of all state vectors as $\mathcal{S}$. Integrals of functions of such variables involve both summations and continuous integrals. Since state vectors $\mathbf{x}$ may have discrete components, $f(\mathbf{z}|\mathbf{x})$ can encompass different measurement models for different target types and $f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{x}_k)$ can encompass different motion models for different target types. If the measurement and motion models are linear and Gaussian, the above equations reduce to the Kalman time-update and information-update equations, respectively [8].

All relevant information about the state-vector $\mathbf{x}$ of the target at time-step k is contained in the Bayes posterior density function $f_{k|k}(\mathbf{x}|Z^k)$. Updating it to a new posterior $f_{k+1|k+1}(\mathbf{x}|Z^{k+1})$ using the Bayes filtering equations usually presents a formidable computational challenge. If signal-to-noise ratio (SNR) is high enough that all time-evolving posteriors are not too complex, however, one can *compress* the posterior into a finite number of *summary statistics* and propagate these statistics in time instead of the posterior itself. The two most familiar summary statistics are the first-moment vector and second-moment matrix

$$\hat{\mathbf{x}}_{k|k} = \int \mathbf{x}\; f_{k|k}(\mathbf{x}|Z^k)d\mathbf{x}, \qquad \hat{Q}_{k|k} = \int \mathbf{x}\mathbf{x}^T\; f_{k|k}(\mathbf{x}|Z^k)d\mathbf{x}$$

where "^{T}" denotes matrix tranpose. If SNR is so high that the second-order and higher moments can be neglected then the first moment is approximately a sufficient statistic, $f_{k|k}(\mathbf{x}|Z^k) \cong f_{k|k}(\mathbf{x}|\hat{\mathbf{x}}_{k|k})$. In this case we can propagate $\hat{\mathbf{x}}_{k|k}$ in time instead of the full distribution using a constant-gain Kalman filter such as the α-β-γ filter—i.e., using completely linear equations. Otherwise, if SNR is such that the higher-order moments can be neglected but covariance cannot, then $\hat{\mathbf{x}}_{k|k}$ and $\hat{Q}_{k|k}$ are approximate sufficient statistics, $f_{k|k}(\mathbf{x}|Z^k) \cong f_{k|k}(\mathbf{x}|\hat{\mathbf{x}}_{k|k}, \hat{Q}_{k|k})$, and we can propagate $\hat{\mathbf{x}}_{k|k}$ and $\hat{Q}_{k|k}$ using a Kalman filter.

1.2 STATISTICAL MOMENTS FOR MULTITARGET PROBLEMS

Stein and Winter's PHD approach can be thought of as, in part, an attempt to extend the reasoning just outlined to *multitarget* tracking problems. In such problems the optimal approach would be to write down the following multisensor, multitarget analogs of the Bayes nonlinear filtering equations:

$$\begin{aligned} f_{k+1|k}(X_{k+1}|Z^{(k} & \quad \int f_{k+1|k}(X_{k+1}|X_k)\; f_{k|k}(X_k|Z^{(k)})\delta X_k \\ f_{k+1|k+1}(X_{k+1}|Z^{(k+1)}) & \propto f(Z_{k+1}|X_{k+1})\; f_{k+1|k}(X_{k+1}|Z^{(k)}) \\ \hat{X}^{JoM}_{k+1|k+1} & = \arg\sup_X \frac{c^{|X|}}{|X|!} f_{k+1|k+1}(X|Z^{k+1}) \end{aligned}$$

In this case,

(1′) X_k is the *multitarget state*, i.e. the set of unknown target states (which are also of unknown number) and Z_k is the set of all measurements collected off of all targets at time-step k;

(2′) $f_{k|k}(X_k|Z^{(k)})$ is a *multitarget posterior density* at time-set k conditioned on the time-stream $Z^{(k)} = \{Z_1, ..., Z_k\}$;

(3′) $f(Z|X)$ is the *multisensor, multitarget likelihood function* that describes the likelihood of observing the observation-set Z given that the multitarget system has multitarget state-set X;

(4′) $f_{k+1|k}(X_{k+1}|X_k)$ is the multitarget Markov transition density that describes the likelihood that the targets will have state-set X_{k+1} at time-step $k+1$ given that they had state-set at time-step k;

(5′) $f_{k+1|k}(X_{k+1}|Z^{(k)})$ is the time-prediction of the multitarget posterior $f_{k|k}(X_k|Z^{(k)})$ to time-step $k+1$;

(6′) $f(Z_{k+1}|Z^{(k)}) = \int f(Z_{k+1}|Y)\; f_{k+1|k}(Y|Z^{(k)})\delta Y$ is the Bayes normalization constant; and

(7′) $\hat{X}^{JoM}_{k+1|k+1}$ is a multitarget analog of the MAP estimator (whose direct multitarget extension is undefined).

The multitarget filtering equations cannot be used in the blind fashion just indicated but, rather, require the tools of finite-set statistics (FISST). **[7,11]** A short history of multitarget Bayes filtering can be found in Section 1. 4 below.

In more detail, a multitarget state-set X has the form

$$X = \emptyset, \quad \{\mathbf{x}\}, \quad \{\mathbf{x}_1, \mathbf{x}_2\} \quad \quad \{\mathbf{x}_1 \quad \mathbf{x}_n$$

where $X = \emptyset$ indicates that no target is present, $X = \{\mathbf{x}\}$ indicates that one target with state $\mathbf{x}$ is present, $X = \{\mathbf{x}_1, \mathbf{x}_2\}$ indicates that two targets with states $\mathbf{x}_1, \mathbf{x}_2$ are present, and so on. The Bayes multitarget posterior $f_{k|k}(X|Z^{(k)})$ has the form

$$\begin{aligned} f_{k|k}(\emptyset|Z^{(k)}) &= \text{posterior likelihood that no targets are present} \\ f_{k|k}(\{\mathbf{x}\}|Z^{(k)}) &= \text{post. like. of one target with state } \mathbf{x} \\ \\ f_{k|k}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}|Z^{(k)}) &= \text{post. like. of } n \text{ targets with states } \mathbf{x}_1, \ldots, \mathbf{x}_n \end{aligned}$$

It must sum to one over all multitarget states. That is, let

$$f_{k|k}(0|Z^{(k)}) = f_{k|k}(\emptyset|Z^{(k)}), \qquad f_{k|k}(i|Z^{(k)}) = \frac{1}{i!}\int f_{k|k}(\{\mathbf{x}_1, \ldots, \mathbf{x}_i\}|Z^{(k)}) d\mathbf{x}_1 \quad d\mathbf{x}_i$$

be the marginal posterior probability that there are $i = 0, 1, 2, \ldots$ targets present. Then for $f_{k|k}(X|Z^k)$ to be a multitarget probability density the following quantity, called a *set integral*, must sum to one:

$$1 = \int f_{k|k}(X|Z^{(k)})\delta X = f_{k|k}(0|Z^{(k)}) + f_{k|k}(1|Z^{(k)}) + f_{k|k}(2|Z^{(k)}) + \quad + f_{k|k}(n|Z^{(k)}) +$$

Given the formidable computational complexity of the *single-target* Bayes nonlinear filtering equations, it should be clear that this complexity will be magnified many-fold in multitarget problems. Drastic but intelligent approximation strategies are required. In a recent paper **[14]** I proposed one computational strategy based on a multitarget analog of the familiar Gaussian approximation. In this paper I exploit a different computational analogy: propagating multitarget analogs of first-order (and/or second-order) moments of the time-evolving random track-set. I use the multitarget moments outlined in Section 4.3.4, pages 168-170 of *Mathematics of Data Fusion.* **[7]** Let $\mathbf{x}$ be a fixed target state. Then for any $i \geq 1$ the marginal-density value

$$\frac{1}{i!}\int f_{k|k}(\{\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_i\}|Z^{(k)}) d\mathbf{x}_1 \cdots d\mathbf{x}_i$$

is the total posterior likelihood that the multitarget system has $i+1$ targets *and* that one of these targets has state $\mathbf{x}$. Consequently, for each $\mathbf{x}$ the marginal-density value

$$D_{k|k}(\{\mathbf{x}\}|Z^{(k)}) = \int f_{k|k}(\{\mathbf{x}\} \cup Y|Z^{(k)})\delta Y = \sum_{i=1}^{\infty} \frac{1}{i!}\int f_{k|k}(\{\mathbf{x}, \mathbf{x}_1, \quad \mathbf{x}_i\}|Z^{(k)}) d\mathbf{x}_1 \quad d\mathbf{x}_i$$

is the *total posterior likelihood that the multitarget system contains a target that has state* $\mathbf{x}$. Consequently, $D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$ will tend to have maxima approximately at the locations of the targets.

It can be shown (see Section 2.2) that $\int_S D_{k|k}(\{\mathbf{x}\}|Z^{(k)})d\mathbf{x}$ is the expected number of targets in S—which means that $D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$ is the same thing as the Stein-Winter PHD $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$. Intuitively speaking, just as the value of the probability density function $f_{\mathbf{X}}(\mathbf{x})$ of a continuous random vector $\mathbf{X}$ provides a means of describing the zero-probability event $\Pr(\mathbf{X}=\mathbf{x})$, so the PHD $\hat{D}_{\Gamma}(\mathbf{x})$ of a finite random track-set Γ provides a means of describing the zero-probability event $\Pr(\mathbf{x}\in\Gamma)$ (see Section 2.5). Also, the state vector $\mathbf{x}$ in any PHD should be interpreted as an *accumulated (or compressed) multitarget state* rather than as a conventional single-target state.

Last but not least, from the point of view of *point-process theory*, the PHD is the same thing as the *expectation density* or *first factorial-moment density* of the random track-set Γ at time-step k (see Section 2.4). This means, in particular, that $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$ is a type of least-squares best-fit approximation of the multitarget posterior $f_{k|k}(X|Z^{(k)})$ by a single-target density function (see Section 2.4). In this sense the PHD is a multitarget analog of the single-target first-order moment $\hat{\mathbf{x}}_{k|k}$. If the multitarget sensing situation is benign enough—meaning that signal-to-clutter ratio (SCR) as well as SNR is large, then the PHD will be an approximate sufficient statistic: $f_{k|k}(X|Z^{(k)})\cong f_{k|k}(X|\hat{D}_{k|k})$. In principle, therefore, it should be possible and desirable to propagate $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$ instead of the full multitarget posterior $f_{k|k}(X|Z^{(k)})$, using suitable analogs of the single-target Bayes recursive filtering equations of Section 1.1. Real-time multitarget tracking would then be, from a computational point of view, reduced to the (still very difficult) problem of implementing a real-time single-target nonlinear filter capable of modeling the rather complex time-evolution of the PHD.

That is, what we would like to be able to do is to establish the existence of a diagram of the form

$$\begin{tikzcd}
\cdots \arrow[r] & f_{k|k} \arrow[r, "\text{multitarget prediction}"] & f_{k+1|k} \arrow[d] \arrow[r, "\text{multitarget Bayes' rule}"] & f_{k+1|k+1} \arrow[d, dash] \arrow[r] & \cdots \\
\cdots \arrow[r] & \hat{D}_{k|k} \arrow[r, "\text{PHD prediction??}"] & \hat{D}_{k+1|k} \arrow[r, "\text{PHD Bayes rule??}"] & \hat{D}_{k+1|k+1} \arrow[r] & \cdots
\end{tikzcd}$$

where: (1) the top row portrays the time-sequence of the multitarget Bayes filtering equations; (2) the downward-pointing arrows indicate the replacement of multitarget posteriors by their corresponding PHD's; and (3) the bottom row portrays a recursive time-sequence of filtering operations on PHD's that always yields the result that one would get if one computed multitarget posteriors using only the top row and then transformed them into their PHD approximations.

Our goal, then, is to fill in the "question marks" in the bottom row of the diagram. We will show (Section 3.1) that the PHD $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$ can, under certain assumptions, be time-propagated between measurements to a new PHD $\hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$ using a suitable extension of the first of the two single-target Bayes filtering equations of Section 1.1, one that accounts for multitarget behaviors such as appearance and disappearance of targets. We will also show (Section 3.2) that, given a new multisensor-multitarget observation-set Z_{k+1}, the PHD $\hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$ can be updated to a new PHD $\hat{D}_{k+1|k+1}(\mathbf{x}|Z^{(k+1)})$ using one of two different approximate methods. First, a relatively simple PHD version of Bayes' rule; and second, a more complicated but "additive" update transformation based on the concept of the "approximate multitarget posterior" of a PHD.

1.3 RELATED APPROACHES

The idea of using a single-target density function $g_{k|k}$ (or, more commonly, probability contours of its graph) as a basis for multitarget tracking is a relatively common one. Examples are the Naval Research Laboratory's TABS (Tactical Antisubmarine-warfare Battle-management System) tracker, Metron Corp.'s *Nodestar* tracker, **[22]** and others **[23]**. The work described in this paper differs from earlier work in its systematic and theoretically rigorous treatment of the "correct" $g_{k|k}$ as a PHD: i.e., as a first-order factorial-moment statistic of the multitarget system.

1.4 A SHORT HISTORY OF BAYES MULTITARGET FILTERING

The concept of multitarget Bayesian nonlinear filtering (Section 1.3) is a relatively new one. If one assumes that the number of targets is known beforehand, the earliest exposition appears to be due to Washburn **[24]** in 1987, using a point process formulation (see Section 2.4 for a summary of point process theory).

Date	Author(s)	Theoretical Basis
1991	Miller et. al. [10] "Jump Diffusion"	Stochastic PDEs
1994	Bethel and Paras [3]	Discrete filtering
1994	Mahler [11] "Finite-Set Statistics"	FISST
1996	Stone et. al. [21] "Unified Data Fusion"	Heuristic
1996	Mahler-Kastella [15] "Joint Multitarget Probabilities"	FISST
1997	Portenko et. al. [16]	Point processes

The table summarizes the history of the approach when the number n of targets is *not* known and must be determined along with the individual target states. The earliest work in this case appears to be due to Miller, O'Sullivan, Srivistava, et. al. **[10]**. Their very sophisticated approach requires solution of stochastic diffusion equations on non-Euclidean manifolds. It is also apparently the only approach to deal with continuous evolution of the multitarget state. (All other approaches listed in the table assume discrete state-evolution.) Mahler was apparently the first to systematically deal with the general discrete state-evolution case (Bethel and Paras assume discrete observation and state variables). Kastella's "joint multitarget probabilities (JMP)," introduced at Lockheed Martin in 1996, are a renaming of a number of early core FISST concepts (set integrals, multitarget information metrics, multitarget posteriors, joint multitarget state estimators, etc.) devised two years earlier **[15]**. A "JMP" itself is just a discretization of a FISST (or, for that matter, a Jump-Diffusion) multitarget posterior:

$$n!\, f_{JMP}(\mathbf{x}_1 \quad \mathbf{x}_n|Z) = f_{FISST}(\{\mathbf{x}_1 \quad ,\mathbf{x}_n\}|Z) = n!\, f_{FISST}(\mathbf{x}_1, \ .,\mathbf{x}_n|Z)$$

Stone et. al. have provided a valuable contribution by clarifying the relationship between multitarget Bayes filtering and multi-hypothesis correlation. Nevertheless, their approach is, with regrets, described as "heuristic" in the table for the reasons summarized in **[11, pp. 91-93]**.

1.5 ORGANIZATION OF THE PAPER

The paper is organized as follows. Basic concepts are covered in Section 2: multitarget moment densities (Section 2.1); the properties of multitarget moments and the PHD, including their direct construction using set derivatives (Section 2.2); examples of PHD's (Section 2.3); identification of the PHD as a first-order statistical moment (Section 2.4); and PHD's in the discrete-state case (Section 2.5). Section 3 is devoted to the Bayes filtering equations for the PHD: the time-update equation (Section 3.1), the approximate Bayes-update equation (Section 3.2), and prior PHD's (Section 3.3). Proofs of the theorems are relegated to Section 4. Conclusions may be found in Section 5.

2.0 MULTITARGET MOMENT DENSITIES AND THE PHD

The purpose of this section is to: (1) formally define the concept of a *multitarget moment density function* $D_{k|k}(X|Z^{(k)})$; (2) provide a general procedure for constructing it using the FISST set derivative; and (3) show that the first-order multitarget moment density $D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$ is the same thing as the Stein-Winter Probability Hypothesis Density $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$. I define multitarget moment densities in Section 2.1, describe their major properties in Section 2.2, provide examples of PHD's in Section 2.3, show that the PHD is a first-order statistical moment in Section 2.4 and, in Section 2.5, consider PHD's in the special case when the target state-space is discrete.

2.1 MULTITARGET MOMENT DENSITIES

I begin with a definition:

Definition 1 (Multitarget Moment Densities) [7, p. 169]: The *multitarget moment density* is

$$D_{k|k}(X|Z^{(k)}) = \int f_{k|k}(X \cup W|Z^{(k)})\delta W = \sum_{i=0}^{\infty} \frac{1}{i!} \int f_{k|k}(X \cup \{\mathbf{x}_1, ..., \mathbf{x}_i\})d\mathbf{x}_1 \cdots d\mathbf{x}_i$$

Notice that $D_{k|k}(\emptyset|Z^{(k)}) = 1$. If the number $|X|$ of elements in X is restricted to n then I will call the function $D_{k|k}(\{\mathbf{x}_1, ..., \mathbf{x}_n\}|Z^{(k)})$ the *n'th multitarget moment density.* Also notice that the set integral is well-defined in the sense that $f_{k|k}(X \cup W|Z^{(k)})$ always has the same units as X and so there are no units-mismatch problems of the kind described in **[11, p. 39]**. For any multitarget state $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$, $D_{k|k}(\{\mathbf{x}_1, ..., \mathbf{x}_n\}|Z^{(k)})$ is the marginal-posterior likelihood that, regardless of how many targets there may be in the multitarget system, exactly n of them have states $\mathbf{x}_1, ..., \mathbf{x}_n$.

2.2 COMPUTING PHD'S USING THE SET DERIVATIVE

In this section I show how to construct multitarget moment densities directly from the random multitarget track-set using the FISST set derivative (Theorem 1) and use this fact to show that the PHD and the first-order multitarget moment density are the same thing (Theorem 2). I begin by demonstrating the first result (which will also allow us to compute between-measurement laws of motion for PHD's directly from multitarget motion models, see Section 3.1). Let Γ_k be the random set of current tracks at time-step k, meaning that $f_{k|k}(X|Z^{(k)})$ is the multitarget density corresponding to the belief-mass function $\beta_{k|k}(S|Z^{(k)}) = \Pr(\Gamma_k \subseteq S)$. Then:

Theorem 1 (Computing Multitarget Moments Using the Set Derivative) [7, p. 169]: Let $D_{k|k}(X|Z^{(k)})$ be the multitarget moment density corresponding to the multitarget posterior $f_{k|k}(X|Z^{(k)})$. Then:

$$D_{k|k}(X|Z^{(k)}) = \frac{\delta \beta_{k|k}}{\delta X}(\mathcal{S}|Z^{(k)})$$

for all finite subsets X of (single-target) state space $\mathcal{S}$.

The proof of this assertion can be found in Section 4.1. It should be compared to the similar formula for constructing multitarget posterior densities **[11, pp. 30-31]**:

$$f_{k|k}(X|Z^{(k)}) = \frac{\delta \beta_{k|k}}{\delta X}(\emptyset|Z^{(k)})$$

As already noted, because of Theorem 1 we can show that the first-order moment density and the PHD are the same thing. If $\Gamma_k \cap S$ is the set of tracks contained in S then $|\Gamma_k \cap S|$ is the number of tracks in S and $\mathrm{E}[|\Gamma_k \cap S|]$ is the expected number of tracks in S. Then:

Theorem 2 (The PHD is the First Multitarget Moment Density) [7, p. 169]: For any measurable subset $S \subseteq \mathcal{S}$ of state-vectors,

$$\mathrm{E}[\,|\Gamma_k \cap S|\,] = \int_S D_{k|k}(\{\mathbf{x}\}|Z^{(k)})d\mathbf{x}$$

Consequently, the first multitarget moment density and the Probability Hypothesis Density (PHD) are equal almost everyhere: $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) = D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$.

See Section 4.2 for the proof.

2.3 EXAMPLES OF PHD's

2.3.1 EXAMPLE 1: INFORMATION LOSS IN PHD's. Information is lost when we compress a single-target posterior density $f_{k|k}(\mathbf{x}|Z^k)$ into its first moment $\hat{\mathbf{x}}_{k|k} = \int \mathbf{x}\; f_{k|k}(\mathbf{x}|Z^k)d\mathbf{x}$. Likewise, information is lost when we compress a multitarget posterior $f_{k|k}(X|Z^{(k)})$ into its PHD $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$. For example, suppose that we are trying to determine the locations of two targets on the real-number line based on a single sensor-scan $Z_1 = \{z_1, z_2\}$. Suppose that the multitarget posterior has the form

$$f_{1|1}(\{x_1,x_2\}|Z^{(1)}) = N_{\sigma^2}(x_1 - z_1)N_{\sigma^2}(x_2 - z_2) + N_{\sigma^2}(x_2 - z_1)N_{\sigma^2}(x_1 - z_2)$$

where $N_{\sigma^2}(x) = (\sqrt{2\pi}\sigma)^{-1}\exp(-x^2/2\sigma^2)$ is the normal distribution with variance σ^2. The corresponding PHD and second moment are, respectively,

$$\begin{aligned}
\hat{D}_{1|1}(x|Z^{(1)}) &= \int f_{1|1}(\{x,y\}|Z^{(1)})dy = N_{\sigma^2}(x - z_1) + N_{\sigma^2}(x - z_2) \\
D_{1|1}(\{x_1,x_2\}|Z^{(1)}) &= f_{1|1}(\{x_1,x_2\}|Z^{(1)})
\end{aligned}$$

Note that $\int \hat{D}_{1|1}(x|Z^{(1)})dx = 2$, so that the expected number of targets is two. In general, $\hat{D}_{1|1}$ is bimodal. However, it is easily shown that it is unimodal with maximal value at $x = \frac{1}{2}(z_1 + z_2)$ whenever $|z_1 - z_2| < 2\sigma$. The multitarget posterior $f_{1|1}(\{x_1,x_2\}|Z^{(1)})$, on the other hand, is always unimodal (as a function of a set variable) but fails to distinguish two distinct targets when $|z_1 - z_2| < \sqrt{2}\sigma$. In this case its unique maximal value is located at $x_1 = x_2 = \frac{1}{2}(z_1 + z_2)$. So, for data separations in the range $\sqrt{2}\sigma < |z_1 - z_2| < 2\sigma$ the multitarget posterior is capable of separating two targets whereas the PHD is not. This indicates, not unexpectedly, that a PHD-based multitarget tracker will experience more difficulty with closely-spaced targets than would a tracker based on the full multitarget nonlinear filtering equations of Section 1.2.

2.3.2 EXAMPLE 2: CLUTTER AND PHD's. Extend the previous example by including the effects of clutter. Assume that a single sensor observes two targets with no missed detections but with false alarms governed by the independent clutter process $\kappa(Z)$. If we collect one scan $Z_1 = \{z_1, z_2\}$ consisting of two distinct observations z_1, z_2 and assume a uniform prior then the following multitarget posterior is the result:

$$f_{1|1}(\{x_1,x_2\}|Z^{(1)}) = \frac{2}{m(m-1)} \sum_{\{z_1,z_2\}\subseteq Z} (N_{\sigma^2}(x_1 - z_1)N_{\sigma^2}(x_2 - z_2) + N_{\sigma^2}(x_2 - z_1)N_{\sigma^2}(x_1 - z_2))$$

with $f_{1|1}(X|Z^{(1)}) = 0$ whenever X does not contain exactly two elements. The corresponding PHD is:

$$\hat{D}_{1|1}(x|Z^{(1)}) = \frac{2}{m} \sum_{z \in Z} N_{\sigma^2}(x - z)$$

Notice that $\int \hat{D}_{1|1}(x|Z^{(1)})dx = 2$. As a function of the set $\{x_1, x_2\}$, the multitarget posterior has $C_{m,2} = \frac{1}{2}m(m-1)$ peaks. Each peak corresponds to a different hypothesis regarding which two-element subset of Z are target-generated reports rather than false alarms. Since the PHD condenses multitarget information into a density on single-target state space, it has at most m peaks. If targets are sufficiently separated, each peak corresponds to a different hypothesis about the location of the individual targets.

2.4 THE PHD IS A FIRST-ORDER MOMENT OF A RANDOM SET [7, p. 169]

The purpose of this section is to: (1) show that the multitarget moment density $D_{k|k}(X|Z^{(k)})$ is identical to the "factorial-moment densities" of point process theory; (2) conclude that the PHD is a first moment of the multitarget system; and (3) provide an inversion formula for transforming multitarget moment densities $D_{k|k}(X|Z^{(k)})$ into multitarget posteriors $f_{k|k}(X|Z^{(k)})$. All unreferenced page numbers in this section refer to the textbook by Daly & Vere-Jones [4]. Let Γ_k be a random track-set. Then either the random integer-valued measure

$$N_k(S|Z^{(k)}) = |\Gamma_k \cap S| = \int_S \delta_{\Gamma_k}(\mathbf{x})d\mathbf{x} = \text{ no. of tracks in } \Gamma_k \text{ contained in region } S$$

or its random density function $\delta_{\Gamma_k}(\mathbf{x}) = \sum_{\mathbf{w} \in \Gamma_k} \delta_{\mathbf{w}}(\mathbf{x})$ is called a *multi-dimensional point process.* Point process theory is a special case of random set theory and, in fact, multi-dimensional point processes seem to have been originally defined as random sets rather than as random measures [1]. The statistical behavior of $N_k(S|Z^{(k)})$—or, equivalently, of Γ_k and $\delta_{\Gamma_k}(\mathbf{x})$—is characterized by its family $j_{k,i}(\mathbf{x}_1, ..., \mathbf{x}_i)$ of *Janossy densities* (pp. 122-123). Janossy densities are completely symmetric in all arguments; vanish whenever $\mathbf{x}_i = \mathbf{x}_j$ for some $i \neq j$ (p. 134, Prop. 5.4.IV); and are jointly normalized in the sense that $\sum_{i=0}^{\infty} \frac{1}{i!} j_{k,i}(\mathbf{x}_1, ..., \mathbf{x}_i) = 1$. The multitarget posterior density $f_{k|k}(X|Z^{(k)})$ of Γ_k is the same thing as the family of Janossy densities $j_{k,i}$ of $N_k(S|Z^{(k)})$:

$$j_{k,i}(\mathbf{x}_1, ..., \mathbf{x}_i) = f_{k|k}(\{\mathbf{x}_1, ..., \mathbf{x}_i\}|Z^{(k)})$$

for all distinct $\mathbf{x}_1, ..., \mathbf{x}_i$ (In like manner, the following quantities are just different notations for a multitarget posterior:

$$f_{k|k}(\{\mathbf{x}_1, ..., \mathbf{x}_n\}|Z^{(k)}) = f_{k|k}(\delta_{\mathbf{x}_1} + ... + \delta_{\mathbf{x}_n}|Z^{(k)}) = f_{k|k}(N_{\{\mathbf{x}_1, ..., \mathbf{x}_n\}}|Z^{(k)})$$

where the second and third quantities denote, respectively, posterior probability distributions over all point-process densities δ_X or all point-process measures N_X.)

The expected value

$$M_{k,[1]}(S|Z^{(k)}) = \mathrm{E}[N_{k|k}(S|Z^{(k)})] = \mathrm{E}[\,|\Gamma_k \cap S|\,] = \int_S \mathrm{E}[\delta_{\Gamma_k}(\mathbf{x})]d\mathbf{x}$$

is called the *expectation measure* or *first factorial-moment measure* of $N_k(S|Z^{(k)})$ (p. 130). Its density

$$m_{k,[1]}(\mathbf{x}) = \mathrm{E}[\delta_{\Gamma_k}(\mathbf{x})] = \int \delta_X(\mathbf{x})\ f_{k|k}(X|Z^{(k)})\delta X$$

is called the *expectation density* or *first factorial-moment density.* Higher-order factorial-moment densities $m_{k,[i]}(\mathbf{x}_1,...,\mathbf{x}_i)$ can be defined (pp. 130, 112) and, from Definition 1 (Section 2.1), it follows that (p. 133, equation 5.4.11):

$$D_{k|k}(\{\mathbf{x}_1,...,\mathbf{x}_j\}|Z^{(k)}) = m_{k,[j]}(\mathbf{x}_1,...,\mathbf{x}_j)$$

for distinct $\mathbf{x}_1,...,\mathbf{x}_j \in \mathcal{S}$. That is, the $D_{k|k}(X|Z^{(k)})$ are statistical moments of the random set Γ_k and *the PHD is the first factorial-moment density.* Moreover, the multitarget posterior density $f_{k|k}(X|Z^{(k)})$ can be recovered from the multitarget moment density $D_{k|k}(X|Z^{(k)})$ via the following set integral (p. 133, equation 5.4.12):

Theorem 3: (Inversion formula for multitarget moment densities) [7, p. 169]:

$$f_{k|k}(X|Z^{(k)}) = \int (-1)^{|W|} D_{k|k}(X \cup W|Z^{(k)})\delta W$$

This formula confirms the obvious fact that the multitarget posterior cannot be completely described by any one multitarget moment density $D_{k|k}(X|Z^{(k)})$ and, in particular, by the PHD $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$. Rather, all multitarget moment densities are required to completely recover the information contained in the multitarget posterior.

2.5 THE PHD IN THE DISCRETE-STATE CASE [7, p. 169]

Suppose that (single-target) state space $\mathcal{S}$ is a finite set of target-state cells x. Let Γ be the randomly varying track-set and note that $f_{k|k}(X|Z^{(k)}) = \Pr(\Gamma = X)$. Then:

$$\hat{D}_{k|k}(x|Z^{(k)}) = \sum_{X \ni x} p_\Gamma(X) = \sum_{X \ni x} \Pr(\Gamma = X) = \sum_{X} \Pr(x \in X,\ \Gamma = X) = \Pr(x \in \Gamma$$

This result shows that, in the discrete case, the PHD of a random track-set Γ is the same thing as I.R. Goodman's *one-point covering function* $\mu_\Gamma(x) = \Pr(x \in \Gamma) = \hat{D}(x)$ of the random set Γ. **[5,6]** The existence of this relationship is the reason why, in *Mathematics of Data Fusion*, I used the term "global covering densities" for what in this paper I call "multitarget moment densities." It also shows that in the continuous case, the PHD provides a means of representing the zero-probability event $\Pr(\mathbf{x} \in \Gamma)$ in much the same way that the density $f_{\mathbf{X}}(\mathbf{x})$ of a continuous random vector provides a means of representing the zero-probability event $\Pr(\mathbf{X} = \mathbf{x})$. Furthermore, it is easy to show that

$$\sum_x \hat{D}_{k|k}(x|Z^{(k)}) = \sum_x \sum_X p(x \in X, \Gamma = X) = \sum_X \left(\sum_x p(x \in X) \right) p_\Gamma(X) = \mathrm{E}[\,|\Gamma \cap X|\,]$$

In the fuzzy logic literature, the sum $\sum_x \mu(x)$ is called the "sigma-count" of the fuzzy membership function μ and is interpreted as the "number of elements" in the fuzzy set corresponding to μ.

3.0 RECURSIVE BAYES FILTERING OF THE PHD

In this section we derive recursive filtering equations for the PHD analogous to the single-target Bayes recursive filtering equations of Section 1.1. These equations include between-measurements time-update equations (see Section 3.1) and an *approximate* Bayes information-update equation (see Section 3.2). The construction of prior PHD's is discussed in Section 3.3.

3.1 TIME-UPDATE OF THE PHD

The between-measurements time-evolution of a single-target posterior is described by the first of the single-target Bayes filtering equations of Section 1.1. The purpose of this section is to show how to construct similar laws of motion for PHD's. Let $f_{k+1|k}(Y|X)$ be the multitarget Markov transition density that corresponds to some multitarget motion model **[11, pp. 21-23]** and let $f_{k+1|k}(Y|Z^{(k)})$ be the time-predicted multitarget posterior as computed in the first of the multitarget Bayes filtering equations of Section 1.2. The PHD's corresponding to these two multitarget posteriors are

$$\hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) = \int f_{k|k}(\{\mathbf{x}\} \cup W|Z^{(k)})\delta W, \qquad \hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)}) = \int f_{k+1|k}(\{\mathbf{x}\} \cup W|Z^{(k)})\delta W$$

It might appear that the law of motion relating $\hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$ and $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$ is nothing more complicated than the following simple analog of the first Bayes filtering equation:

$$\hat{D}_{k+1|k}(\mathbf{y}|Z^k) = \int f_{k+1|k}(\mathbf{y}|\mathbf{x})\ \hat{D}_{k|k}(\mathbf{x}|Z^k)d\mathbf{x}$$

In actuality, this evolution law describes only multitarget motion in which *target motions are independent and the number of targets does not change* (Corollary 2 below). We begin, therefore, with a general result that encompasses quite general multitarget motion models that account for "birth" and "death" of targets—i.e., targets that enter and leave the scenario for whatever reason. The result itself is less important than its method of proof (see Section 4.3).

Theorem 4 (Laws of Motion for PHD's): Suppose that between time-step k and time-step $k+1$, the following multitarget motion model is assumed: (1) target motions are statistically independent; (2) the Markov motion model for single targets is $f_{k+1|k}(\mathbf{y}|\mathbf{x})$; (3) the probability that any individual target will not "die" (i.e., not vanish from the scenario) if it has state $\mathbf{x}$ at time-step k is $d_{k+1|k}(\mathbf{x})$; and (4) each target with state $\mathbf{x}$ at time-step k generates, independently of all other targets, new "birth" targets in a fashion governed by a FISST multitarget density $b_{k+1|k}(Y|\mathbf{x})$. Then:

$$\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) = \int \left(d_{k+1|k}(\mathbf{x})f_{k+1|k}(\mathbf{y}|\mathbf{x}) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x})\right)\ \hat{D}_{k|k}(\mathbf{x}|Z^{(k)})d\mathbf{x}$$

where $\hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x})$ denotes the PHD of the multitarget density $b_{k+1|k}(X|\mathbf{x})$.

The proof can be found in Section 4.3. Let $N_{k|k} = \int \hat{D}_{k|k}(\mathbf{y}|Z^{(k)})d\mathbf{x}$ be the expected number of targets in the track-set Γ_k. Then from Theorem 4 it follows that the expected number $N_{k+1|k}$ of time-extrapolated targets is

$$N_{k+1|k} = \int \left(d_{k+1|k}(\mathbf{x})N_{k|k} + N^B_{k+1|k}(\mathbf{x})\right)\ \hat{D}_{k|k}(\mathbf{x}|Z^{(k)})d\mathbf{x}$$

where $N^B_{k+1|k}(\mathbf{x}) = \int \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x})d\mathbf{y}$ is the expected number of birth targets.

The following is a special case of Theorem 4 that employs a simple state-dependent Poisson model to account for the appearance of new targets.

Corollary 1 (Law of Motion for PHD's With Poisson Births): Suppose that between time-step k and time-step $k+1$, the following multitarget motion model is assumed: (1) target motions are statistically independent; (2) the Markov motion model for single targets is $f_{k+1|k}(\mathbf{y}|\mathbf{x})$; (3) the probability that any individual target will not "die" is $d_{k+1|k}$; and (4) each state $\mathbf{x}$ at time-step k

generates, independently of all other targets, new "birth" targets in a Poisson-distributed fashion with Poisson parameter $\lambda_{k+1|k}$ and birth distribution $b_{k+1|k}(\mathbf{y}|\mathbf{x})$. Then:

$$\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) = \int \left(d_{k+1|k} f_{k+1|k}(\mathbf{y}|\mathbf{x}) + \lambda_{k+} \; b_{k+1|k}(\mathbf{y}|\mathbf{x})\right) \hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

The proof of this fact follows immediately from Theorem 4 by noting that

$$b_{k+1|k}(\{\mathbf{y} \quad \mathbf{y}_n\}|\mathbf{x}) = e^{-\lambda_{k} \; |k} \lambda^n_{k+1|k} b_{k+1|k}(\mathbf{y}_1|\mathbf{x}) \quad b_{k+1|k}(\mathbf{y}_n|\mathbf{x}) \qquad b_{k+1|k}(\emptyset|\mathbf{x}) = e$$

and so

$$b_{+1|k}(S|\mathbf{x}_i) = \sum_{i=0}^{\infty} \frac{1}{i!} \int_{S^i} b_{k+1|k}(\{\mathbf{y}_1, \quad \mathbf{y}_n\}|\mathbf{x}) d\mathbf{y}_1 \quad d\mathbf{y}_n = e^{\lambda_{k+1|k}(b_{k+1|k}(S|\mathbf{x}_i)-1)}$$

and therefore

$$\hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}_i) = \left[\frac{\delta b_{k+1|k}}{\delta \mathbf{y}}(S|\mathbf{x}_i) \right. \quad = \left[\lambda_{k+1|k} b_{k+1|k}(\mathbf{y}|\mathbf{x}_i) e^{\lambda_{k+1|k}(b_{k} \; |k(S|\mathbf{x}_i)-1)}\right]_{S=} = \lambda_{k+1|k} b_{k+1|k}(\mathbf{y}|\mathbf{x}_i$$

Note that the time-extrapolated number of targets is $N_{k+1|k} = \left(d_{k+1|k} + \lambda_{k+1|k}\right) N_{k|k}$. We conclude by deriving the law of motion for PHD's whose between-measurements time-evolution is governed by the simplest possible multitarget motion model.

Corollary 2 (Simplest Law of Motion for PHD's): Suppose that between time-step k and time-step $k+1$, the following multitarget motion model is assumed: (1) target number does not change; (2) target motions are statistically independent; and (3) the Markov transition model for the single-target motion model is $f_{k+1|k}(\mathbf{y}|\mathbf{x})$. Then

$$\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) = \int f_{k+1|k}(\mathbf{y}|\mathbf{x}) \; \hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

The proof of this fact results from setting $d_{k+1|k} = 1$ (no targets disappear) and $\lambda_{k+1|k} = 0$ (no targets appear) in Corollary 1. Stated in different words: Given this simple multiarget motion model, the between-measurements time-evolution of the PHD is governed by the same law of motion as that which governs the between-measurements time-evolution of the posterior density of any single target in the multitarget system.

3.2 BAYES INFORMATION-UPDATE OF THE PHD

In the single-target case, when a new measurement $\mathbf{z}_{k+1}$ is collected this information can be incorporated into the time-extrapolated posterior $f_{k+1|k}(\mathbf{y}|Z^k)$ using Bayes' rule (the second of the single-target Bayes filtering equations of Section 1.1). The question that confronts us in this section is as follows. Suppose that we have collected a new multisensor-multitarget observation-set Z_{k+1}. Let $f(Z|X)$ be the multisensor-multitarget likelihood function that corresponds to some multisensor-multitarget sensor model **[11, pp. 17-20]** and let the Bayes-rule update of the time-predicted multitarget posterior be computed as in the second of the multitarget Bayes filtering equations of Section 1.2. Then given the corresponding PHD's

$$\hat{D}_{k+1|k+1}(\mathbf{x}|Z^{(k+1)}) = \int f_{k+1|k+1}(\{\mathbf{x}\} \cup W|Z^{(k)}) \delta W \qquad \hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)}) = \int f_{k+1|k}(\{\mathbf{x}\} \cup W|Z^{(k)}) \delta W$$

what rule will allow us to use Z_{k+1} to update $\hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$ to get $\hat{D}_{k+1|k+1}(\mathbf{x}|Z^{(k+1)})$? As it turns out, it is not possible to construct a simple recursive update for PHD's that faithfully reflects the Bayes update on multitarget posteriors.

3.2.1 INFORMATION-UPDATE USING AN APPROXIMATE BAYES' RULE. We have no choice, then, but to adopt an *approximate* Bayes update step. One possible approximation is suggested by turning to the discrete case described in Section 2.5. *Assume* that current observations depend only on the current condensed state $\mathbf{x}$—i.e, $\Pr(Z_{k+1}|\mathbf{x} \in \Gamma_{k+1}, Z^{(k)}) \cong \Pr(Z_{k+1}|\mathbf{x} \in \Gamma_{k+1})$. Then

$$\hat{D}_{k+1|k+1}(\mathbf{x}|Z^{(k+1)}) = \Pr(x \in \Gamma|Z^{(k+1)}) = \frac{\Pr(Z_{k+1}|x \in \Gamma)\ \Pr(x \in \Gamma|Z^{(k)})}{\Pr(Z_{k+1}|Z^{(k)})} = \frac{\hat{D}(Z_{k+1}|\mathbf{x})\ \hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})}{f_{k+1}(Z_{k+1}|Z^{(k)})}$$

where

$$\hat{D}(Z|\mathbf{x}) = \frac{\int f(Z|\{\mathbf{x}\} \cup W)\ f_0(\{\mathbf{x}\} \cup W)\delta W}{\int f_0(\{\mathbf{x}\} \cup W)\delta W}$$

The likelihood should be "biased" only by the previous PHD $\hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$. Therefore we assume that $f_0(X)$ is a multitarget uniform density $u(X)$ (see Section 3.3 below).

3.2.2 INFORMATION-UPDATE USING THE APPROXIMATE POSTERIOR OF A PHD. The work described in this paper can be viewed from a different perspective that, for lack of space, we can only summarize here. (Details will appear in a subsequent paper.) Let $f_{k|k}(X|Z^{(k)})$ be a multitarget posterior, $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)})$ its associated PHD, and $N_{k|k} = \int \hat{D}_{k|k}(\mathbf{x}|Z^{(k)})d\mathbf{x}$. We want to approximate $f_{k|k}$ by an approximate multitarget posterior $\hat{f}_{k|k} \cong f_{k|k}$ that is covariance-free—i.e., whose multitarget moments are

$$D_{k|k}(\emptyset|Z^{(k)}) = \qquad D_{k|k}(X|Z^{(k)}) = D_{k|k}(\mathbf{x}_1|Z^{(k)}) \quad D_{k|k}(\mathbf{x}_n|Z^{(k)})$$

where $X = \{\mathbf{x}_1 \quad , \mathbf{x}_n\}$ Theorem 3 of Section 2.4 yields $\hat{f}_{k|k}$

$$\begin{aligned}
\hat{f}_{k|k}(X|Z^{(k)}) &= \int (-1)^{|Y|} D_{k|k}(X \cup Y|Z^{(k)})\delta Y \\
& \sum_{i=0}^{\infty} \frac{1}{i!} \int (-1)^i D_{k|k}(\mathbf{x}_1|Z^{(k)}) \cdots D_{k|k}(\mathbf{x}_n|Z^{(k)}) D_{k|k}(\mathbf{y}_1|Z^{(k)}) \quad D_{k|k}(\mathbf{y}_n|Z^{(k)})d\mathbf{y} \quad d\mathbf{y}, \\
& D_{k|k}(\mathbf{x}_1|Z^{(k)}) \quad D_{k|k}(\mathbf{x}_n|Z^{(k)}) \sum_{i=0}^{\infty} \frac{(-N_{k|k})^i}{i!} = e^{-N_{k|k}} D_{k|k}(\mathbf{x}_1|Z^{(k)}) \quad D_{k|k}(\mathbf{x}_n|Z^{(k)})
\end{aligned}$$

where $\hat{f}_{k|k}(\emptyset|Z^{(k)}) = e^{-N_{k|k}}$. Because $\hat{f}_{k|k} \cong f_{k|k}$ we can propagate $\hat{f}_{k|k}$ in place of $f_{k|k}$. Given this, Section 3.1 can be interpreted in a different light. We replace $f_{k|k}$ and $f_{k+1|k}$ by $\hat{f}_{k|k}$ and $\hat{f}_{k+1|k}$ and determine what law of motion $\hat{D}_{k|k} \to \hat{D}_{k+1|k}$ corresponds to the law of motion $\hat{f}_{k|k} \to \hat{f}_{k+1|k}$ specified by the multitarget time-prediction integral of Section 1.2. Theorem 4 emerges as a consequence.

The multitarget Bayes' rule information-update step of Section 1.2) can be interpreted in a similar manner. That is, we replace $f_{k+1|k}$ and $f_{k+1|k+1}$ by their approximations $\hat{f}_{k+1|k}$ and $\hat{f}_{k+1|k+1}$ and then determine what transformation $\hat{D}_{k+1|k} \to \hat{D}_{k+1|k+1}$ of the associated PHD's corresponds to the transformation $\hat{f}_{k+1|k} \rightarrow \hat{f}_{k+1|k+1}$ specified by the multitarget Bayes' rule. Under certain assumptions it is possible to derive formulas for $\hat{D}_{k+1|k} \rightarrow \hat{D}_{k+1|k+1}$.

Theorem 5 For example, suppose that (1) there is a single sensor with (single-target) likelihood function $f(\mathbf{z}|\mathbf{x})$; (2) target observations are independent; (3) the probability of missed detection p_D is state-independent; (4) $N_{k+1|k} < (1-p_D)^{-1}$; and (5) sensor observations are corrupted by independent, state-independent, Poisson false alarms with Poisson parameter λ and distribution $c(\mathbf{z})$. Let $Z_{k+1} = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$ denote the latest scan of data. Then it can be shown that

$$\hat{D}_{k+1|k+1}(\mathbf{x}|Z^{(k+1)}) \cong \sum_{\mathbf{z} \in Z_{k+1}} \frac{\hat{D}_{k+1}(\mathbf{z}|Z^{(k)})}{1+\hat{D}_{k+1}(\mathbf{z}|Z^{(k)})} \hat{D}_{k+1|k+1}(\mathbf{x}|\mathbf{z}, Z^{(k)}) + \frac{(1-p_D)N_{k+1|k}}{(1-(1-p_D)N_{k+1|k})N_{k+1|k}} \hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)})$$

where

$$\hat{D}_{k+1|k+1}(\mathbf{x}|\mathbf{z}, Z^{(k)}) = K^{-1} f(\mathbf{z}|\mathbf{x})\ \hat{D}_{k+1|k}(\mathbf{x}|Z^{(k)}), \qquad \hat{D}_{k+1}(\mathbf{z}|Z^{(k)}) \quad \frac{p_D}{\lambda c(\mathbf{z})} \int f(\mathbf{z}|\mathbf{y}) \hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) dy$$

and where $K = \int f(\mathbf{z}|\mathbf{y}) \hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) dy$. That is, the Bayes-update step for PHD's is *additive*—a property that Stein and Winter call "Weak Evidential Accrual."

3.3 PRIOR PHD's

Let $f_0(X)$ be a prior multitarget density **[11, p. 37]**. Then we can construct the corresponding prior PHD using the definition of a first multitarget moment function (Definition 1): $\hat{D}_0(\mathbf{x}) = \int f_0(\{\mathbf{x}\} \cup W)\delta W$. Alternatively, if we specify a prior random track-set Γ then the prior PHD can be constructed directly from Γ using Theorem 1.

For example, suppose that we have n independent tracks with prior densities $f_1(\mathbf{x}), ..., f_n(\mathbf{x})$ and that the i'th track is believed to exist with probability π_i. Then the prior track-set is $\Gamma = \Gamma_1 \cup ... \cup \Gamma_n$ where $\Gamma_i = \{\mathbf{X}_i\} \cup \emptyset_i^{\pi_i}$. The prior belief-mass function is

$$\beta_0(S) = \Pr(\Gamma \subseteq S) = \Pr(\Gamma_1 \subseteq S) \cdots \Pr(\Gamma_n \subseteq S) = (1 - \pi_1 + \pi_1 p_1(S)) \ \cdots \ (1 - \pi_n + \pi_n p_n(S))$$

where $p_i(S) = \int_S f_i(\mathbf{x}) d\mathbf{x}$. Using Theorem 1 we find that the prior PHD is

$$\begin{aligned} \hat{D}_0(\mathbf{x}) &= \left[\frac{\delta \beta_0}{\delta \mathbf{x}}(S)\right]_{S=\mathcal{S}} = \left[\sum_{i=1} (1-\pi_1+\pi_1 p_1(S)) \cdots \pi_i p_i(S) \cdots (1-\pi_n+\pi_n p_n(S))\right]_{S=\mathcal{S}} \\ &= \pi_1 f_1(\mathbf{x}) + ... + \pi_n f_n(\mathbf{x}) \end{aligned}$$

The prior expected number of target is, therefore, $N_0 = \int \hat{D}_0(\mathbf{x}) d\mathbf{x} = \pi_1 + ... + \pi_n$.

As another example, suppose that $f_0(X)$ is a multitarget uniform distribution **[11, p. 37]**,

$$u(X) = \begin{cases} n! V^{-n}(M+1)^{-1} & \text{if} \quad X \subseteq S \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

Then for $\mathbf{x} \in S$,

$$\begin{aligned} \hat{D}_0(\mathbf{x}) &= \int_S u(\{\mathbf{x}\} \cup W)\delta W = \sum_{i=0}^{M-1} \frac{1}{i!} \int_{S^i} \frac{(i+1)!}{V^{i+1}(M+1)} d\mathbf{x}_1 \cdots d\mathbf{x}_i \\ & \frac{1}{V(M+1)} \sum_{j=0}^{M} j = \frac{}{V(M+1)} \frac{}{2} = \frac{}{2V} \end{aligned}$$

The prior expected number of targets is, therefore, $N_0 = \int_S D_0(\mathbf{x})d\mathbf{x} = \frac{1}{2}M$.

4.0 MATHEMATICAL PROOFS

4.1 PROOF OF THEOREM 1

Let Γ_k be a finite random subset of state space $\mathcal{S}$ and let $D_{k|k}(X|Z^{(k)})$ be its corresponding multitarget moment density. We are to prove that

$$D_{k|k}(X|Z^{(k)}) = \frac{\delta\beta_{k|k}}{\delta X}(\mathcal{S}|Z^{(k)})$$

for all $X \subseteq \mathcal{S}$, where $\beta_{k|k}(S|Z^{(k)}) = \Pr(\Gamma_k \subseteq S)$ is the belief-mass function of Γ_k. Let $\Phi = \frac{\delta\beta_{k|k}}{\delta X}$ and suppose we knew that

$$\Phi(S) = \int_S \frac{\delta\Phi}{\delta W}(\emptyset)\delta W$$

for any measurable $S \subseteq \mathcal{S}$. Then since $W \cap X = \emptyset$ almost everywhere it would follow that

$$D_{k|k}(X|Z^{(k)}) = \int f_{k|k}(X \cup W|Z^{(k)})\delta W = \int \left[\frac{\delta}{\delta W}\left(\frac{\delta\beta_{k|k}}{\delta X}\right)\right](\emptyset)\delta W = \int_{\mathcal{S}} \frac{\delta\Phi}{\delta W}(\emptyset)\delta W = \Phi(\mathcal{S}) = \frac{\delta\beta_{k|k}}{\delta X}(\mathcal{S})$$

and we would be done. So let us prove that $\Phi(S) = \int_S \frac{\delta\Phi}{\delta W}(\emptyset)\delta W$ for $\Phi = \frac{\delta\beta_{|k}}{\delta X}$. First, by Theorem 17 of **[7, p. 155]** we know that

$$\beta_{k|k}(S|Z^{(k)}) = \sum_{i=0}^{\infty} a_i \int_{S^i} f_i(\mathbf{x}_1, \quad \mathbf{x}_i)d\mathbf{x}_1 \quad d\mathbf{x}_i$$

for some real numbers a_k and where f_i is a completely symmetric density in i arguments. Thus if we set

$$\beta_i(S) \triangleq \int_{S^i} f_i(\mathbf{x}_1, \quad ., \mathbf{x}_i)d\mathbf{x}_1 \quad d\mathbf{x}_i$$

it follows that

$$\Phi(S) = \frac{\delta\beta_{k|k}}{\delta X}(S) = \sum_{i=0}^{\infty} a_i \frac{\delta\beta_i}{\delta X}(S) = \sum_{i=0}^{\infty} a_i \Phi_i(S)$$

where $\Phi_i \triangleq \frac{\delta\beta_i}{\delta X}$ for all $i \geq 0$. It is enough to prove that $\Phi_i(S) = \int_S \frac{\delta\Phi_i}{\delta W}(\emptyset)\delta W$ for all $i \geq 0$ since it would then follow that

$$\Phi(S) = \sum_{i=0}^{\infty} a_i \int_S \frac{\delta\Phi_i}{\delta W}(\emptyset)\delta W = \int_S \frac{\delta\left(\sum_{i=0}^{\infty} a_i\Phi_i\right)}{\delta W}(\emptyset)\delta W = \int_S \frac{\delta\Phi}{\delta W}(\emptyset)\delta W$$

as desired. To show that $\Phi_i(S) = \int_S \frac{\delta\Phi_i}{\delta W}(\emptyset)\delta W$, let $Y = \{\mathbf{y}_1, \quad , \mathbf{y}_j\}$ with $|Y| = j$ and recall from Proposition 19 of **[7, p. 159]** that

$$\frac{\delta\beta_i}{\delta Y}(\emptyset) = \delta_{i,|Y|}\ i!\ f_i(\mathbf{y}_1 \quad , \mathbf{y}_i)$$

If $Y = X \cup W$ with $X = \{\mathbf{x}_1 \quad \mathbf{x}_n\}, W = \{\mathbf{w}_1 \quad \mathbf{w}_{j-n}\}$ and $X \cap W = \emptyset$ then

$$\frac{\delta \Phi_i}{\delta W}(\emptyset) = \frac{\delta \beta_i}{\delta (X \cup W)}(\emptyset) = 0$$

if $|W| \neq i - n$ and

$$\frac{\delta \Phi_i}{\delta W}(\emptyset) = i! \, f_i(\mathbf{x}_1 \quad \mathbf{x}_n, \mathbf{w}_1 \quad \mathbf{w}_i$$

otherwise. Accordingly,

$$\begin{aligned}
\int_S \frac{\delta \Phi_i}{\delta W}(\emptyset)\delta W &= \sum_{j=0}^{\infty} \frac{1}{j!} \int_{S^j} \frac{\delta \Phi_i}{\delta \{\mathbf{w}_1, ..., \mathbf{w}_j\}}(\emptyset) d\mathbf{w}_1 \cdots d\mathbf{w}_j \\
& \frac{1}{(i-n)!} \int_{S^{i-n}} \frac{\delta \Phi_i}{\delta \{\mathbf{w}_1, ..., \mathbf{w}_{i-n}\}}(\emptyset) d\mathbf{w}_1 \cdots d\mathbf{w}_{i-n} \\
& \frac{1}{(i-n)!} \int_{S^{i-n}} \frac{\delta \beta_i}{\delta \{\mathbf{w}_1, ..., \mathbf{w}_{i-n}, \mathbf{x}_1, ..., \mathbf{x}_n\}}(\emptyset) d\mathbf{w} \quad d\mathbf{w}_{i-n} \\
& \frac{i!}{(i-n)!} \int_{S^{i-n}} f_i(\mathbf{w}_1, ..., \mathbf{w}_{i-n}, \mathbf{x}_1, ..., \mathbf{x}_n) d\mathbf{w} \quad d\mathbf{w}_{i-n}
\end{aligned}$$

On the other hand

$$\begin{aligned}
\Phi_i(S) &= \frac{\delta \beta_i}{\delta X}(S) = \frac{\delta}{\delta X}\left(\int_{S^i} f_i(\mathbf{w}_1 \quad \mathbf{w}_i) d\mathbf{w}_1 \quad d\mathbf{w}_i\right) \\
& \sum_{1 \le j_1 \neq ... \neq j_i \le i} \int_{S^{i-n}} f_i(\mathbf{w}_1 \quad [\mathbf{w}_1]_{j_1} \quad [\mathbf{w}_n]_j \quad \mathbf{w}_i) \cdot d\mathbf{w}_1 \quad [d\mathbf{w}_{j_1}] \quad [d\mathbf{w}_j \quad d\mathbf{w}
\end{aligned}$$

where the last summation is taken over all distinct $j_1, ..., j_i$ with $1 \le j_1, ..., j_i \le i$, where $[\mathbf{w}]_j$ indicates that the argument $\mathbf{w}_j$ has been replaced by $\mathbf{x}$, and where $[d\mathbf{w}_j]$ indicates that the differential $d\mathbf{w}_j$ is excluded. Since f_i is symmetric and since there are $i! C_{i,n}$ terms in the summation we then get

$$\Phi_i(S) = \frac{i!}{(i-n)!} \int_{S^i} f_i(\zeta \mathbf{w} \quad , \mathbf{w}_{i-n}, \mathbf{x}_1 \quad \mathbf{x}_n) d\mathbf{w}_1 \quad d\mathbf{w}_{i-n}$$

Thus $\Phi_i(S) = \int_S \frac{\delta \Phi_i}{\delta W}(\emptyset)\delta W$ for all $i \ge 0$ and we are done.

4.2 PROOF OF THEOREM 2

We are to prove the following: for any $S \subseteq \mathcal{S}$

$$\mathrm{E}[\,|\Gamma_k \cap S|\,] = \int_S D_{k|k}(\{\mathbf{x}\}|Z^{(k)}) d\mathbf{x}$$

Consequently, the first multitarget moment density and the Probability Hypothesis Density (PHD) are equal almost everyhere: $\hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) = D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$. Let $\delta_W(\mathbf{x}) = \sum_{\mathbf{w} \in W} \delta_{\mathbf{x}}(\mathbf{w})$ and $\delta_\emptyset(\mathbf{x}) = 0$ where $\delta_{\mathbf{x}}(\mathbf{w})$ is the Dirac delta concentrated on $\mathbf{x}$. First note that

$$\int \delta_W(\mathbf{x}) f_{k|k}(W|Z^{(k)}) \delta W = \sum_{i=1}^{\infty} \frac{1}{i!} \int (\delta_{\mathbf{x}}(\mathbf{w} \quad + \quad + \delta_{\mathbf{x}}(\mathbf{w}_i)) \, f_{k|k}(\{\mathbf{w}_1 \quad \mathbf{w}_i\}|Z^{(k)}) d\mathbf{w}_1 \quad d\mathbf{w}$$

$$\sum_{j=0}^{\infty} \frac{1}{j!} \int f_{k|k}(\{\mathbf{x}, \mathbf{y}_1 ..., \mathbf{y}_j\}|Z^{(k)}) d\mathbf{y}_1 \cdots d\mathbf{y}_j$$

$$\int f_{k|k}(\{\mathbf{x}\} \cup Y|Z^{(k)}) \delta Y = D_{k|k}(\{\mathbf{x}\}|Z^{(k)})$$

Now define the indicator function $\mathbf{1}_S(\mathbf{x})$ by $\mathbf{1}_S(\mathbf{x}) = 1$ if $\mathbf{x} \in S$ and $\mathbf{1}_S(\mathbf{x}) = 0$ otherwise. Then

$$\begin{aligned}
\int_S D_{k|k}(\{\mathbf{x}\}|Z^{(k)}) d\mathbf{x} &= \int \mathbf{1}_S(\mathbf{x}) \left(\int \delta_W(\mathbf{x}) f_{k|k}(W|Z^{(k)}) \delta W \right) d\mathbf{x} \\
&\int \left(\int \mathbf{1}_S(\mathbf{x}) \delta_W(\mathbf{x}) d\mathbf{x} \right) f_{k|k}(W|Z^{(k)}) \delta W \\
&\int |S \cap W| \; f_{k|k}(W|Z^{(k)}) \delta W = \mathrm{E}[|\Gamma_k \cap S|]
\end{aligned}$$

4.3 PROOF OF THEOREM 4

We are to show:

$$\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) = \int \left(d_{k+1|k}(\mathbf{x}) f_{k+1|k}(\mathbf{y}|\mathbf{x}) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}) \right) \; \hat{D}_{k|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

First, notice that by Definition 1 (Section 2.1),

$$\begin{aligned}
\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) &= \int f_{k+1|k}(\{\mathbf{y}\} \cup W|Z^{(k)}) \delta W = \int \left(\int f_{k+1|k}(\{\mathbf{y}\} \cup W|X) \; f_{k|k}(X|Z^{(k)}) \delta X \right) \delta W \\
&\int \left(\int f_{k+1|k}(\{\mathbf{y}\} \cup W|X) \delta W \right) \; f_{k|k}(X|Z^{(k)}) \delta X = \int \hat{D}_{k+1|k}(\mathbf{y}|X) \; f_{k|k}(X|Z^{(k)}) \delta X
\end{aligned}$$

where

$$\hat{D}_{k+1|k}(\mathbf{y}|X) = \int f_{k+1|k}(\{\mathbf{y}\} \cup W|X) \delta W$$

Suppose that we know that, given $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_i\}$,

$$\hat{D}_{k+1|k}(\mathbf{y}|X) = \sum_{j=1}^{i} d_{k+1|k}(\mathbf{x}_j) f_{k+1|k}(\mathbf{y}|\mathbf{x}_j) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}_j)$$

then we will have

$$\begin{aligned}
\hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)}) \quad & \int \hat{D}_{k+1|k}(\mathbf{y}|X) \; f_{k|k}(X|Z^{(k)}) \delta X \\
& \sum_{i=1}^{\infty} \frac{1}{i!} \int \sum_{j=1}^{i} \left(d_{k+1|k}(\mathbf{x}_j) f_{k+1|k}(\mathbf{y}|\mathbf{x}_j) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}_j) \right) \; f_{k|k}(\{\mathbf{x}_1, ..., \mathbf{x}_i\}|Z^{(k)}) d\mathbf{x}_1 \quad d\mathbf{x}_i \\
& \int \left(d_{k+1|k}(\mathbf{x}) f_{k+1|k}(\mathbf{y}|\mathbf{x}) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}) \right)
\end{aligned}$$

$$\left(\sum_{j=0}^{\infty}\frac{1}{j!}\int f_{k|k}(\{\mathbf{x},\mathbf{w}_1, \quad \mathbf{w}_j\}|Z^{(k)})d\mathbf{w}_1 \quad d\mathbf{w}_j\right)d\mathbf{x}$$

$$\int\left(d_{k+1|k}(\mathbf{x})f_{k+1|k}(\mathbf{y}|\mathbf{x})+\hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x})\right)\ \hat{D}_{k+1|k}(\mathbf{y}|Z^{(k)})d\mathbf{x}$$

as claimed. We must therefore show that

$$\hat{D}_{k+1|k}(\mathbf{y}|X)=\sum_{i=1}^{n}\left(d_{k+1|k}(\mathbf{x}_i)f_{k+1|k}(\mathbf{y}|\mathbf{x}_i)+\hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}_i)\right)$$

Let $\mathbf{X}^{k+1}=\Phi_k(\mathbf{x},\mathbf{W}^k)$ be the single-target motion model that corresponds to the single-target Markov density $f_{k+1|k}(\mathbf{y}|\mathbf{x})$, let $X=\{\mathbf{x}_1,...,\mathbf{x}_n\}$ be the target states at time-step k, let $\mathbf{W}_1^k,...,\mathbf{W}_n^k$ be i.i.d. copies of the random noise vector $\mathbf{W}^k$, and let $\mathbf{X}_i^{k+1}=\Phi_k(\mathbf{x}_i,\mathbf{W}_i^k)$. Define $\Gamma_i^k=\{\mathbf{X}_i^{k+1}\}\cap\emptyset_i^{d_{k+1|k}(\mathbf{x})}$ where $\emptyset_i^{d_{k+1|k}(\mathbf{x})}$ is a random subset of state space $\mathcal{S}$ such that $\emptyset_i^{d_{k+1|k}(\mathbf{x})}=\emptyset$ with probability $1-d_{k+1|k}(\mathbf{x})$ and $\emptyset_i^{d_{k+1|k}}=\mathcal{S}$ with probability $d_{k+1|k}(\mathbf{x})$. We also assume that $\emptyset_1^{d_{k+1|k}(\mathbf{x})},...,\emptyset_n^{d_{k+1|k}(\mathbf{x})}$ are i.i.d. and that $\emptyset_1^{d_{k+1|k}(\mathbf{x})},...,\emptyset_n^{d_{k+1|k}(\mathbf{x})},\mathbf{W}_1^k,...,\mathbf{W}_n^k$ are statistically independent. Then $\Gamma_i^k=\emptyset$ with probability $1-d_{k+1|k}(\mathbf{x})$ and $\Gamma_i^k=\{\mathbf{X}_i^{k+1}\}$ with probability $d_{k+1|k}(\mathbf{x})$. Furthermore, by assumption each target with state $\mathbf{x}$ at time-step k generates, independently of all other targets, new "birth" targets at time-step $k+1$ in a manner governed by a multitarget density $b_{k+1|k}(X|\mathbf{x})$. Translated into mathematical terms the multitarget motion model is, therefore,

$$\Gamma_{k+1}=\Gamma_1^k\cup \quad \cup\Gamma_n^k\cup B_1^k\cup \quad \cup B_n^k$$

and the corresponding belief-mass function is

$$\begin{aligned}
\beta_{k+1|k}(S|X) &= \Pr(\Gamma_1^k\cup...\cup\Gamma_n^k\cup B_1^k\cup...\cup B_n^k\subseteq S)\\
&= \Pr(\Gamma_1^k\subseteq S)\cdots\Pr(\Gamma_n^k\subseteq S)\ \Pr(B_1^k\subseteq S)\cdots\Pr(B_n^k\subseteq S)\\
&= \left(1-d_{k+1|k}(\mathbf{x}_1)+d_{k+1|k}(\mathbf{x}_1)p_{k+1|k}(S|\mathbf{x}_1)\right)\cdots\left(1-d_{k+1|k}(\mathbf{x}_n)+d_{k+1|k}(\mathbf{x}_n)p_{k+1|k}(S|\mathbf{x}_n)\right)\\
&\quad\cdot b_{k+1|k}(S|\mathbf{x}_1)\cdots b_{k+1|k}(S|\mathbf{x}_n)
\end{aligned}$$

where $b_{k+1|k}(S|\mathbf{x})=\int b_{k+1|k}(X|\mathbf{x})\delta X$ is the belief-mass function corresponding to the multitarget density function $f_{k+1|k}(X|Z^{(k)})$. The first-order set derivative is

$$\begin{aligned}
\frac{\delta\beta_{k+1|k}}{\delta\mathbf{y}}(S|X) &= \sum_{i=1}^{n}\left(1-d_{k+1|k}(\mathbf{x}_1)+d_{k+1|k}(\mathbf{x}_1)p_{k+1|k}(S|\mathbf{x}_1)\right)\\
&\quad \frac{\delta}{\delta\mathbf{y}}\left(1-d_{k+1|k}(\mathbf{x}_i)+d_{k+1|k}(\mathbf{x}_i)p_{k+1|k}(S|\mathbf{x}_i)\right)\\
&\quad \left(1-d_{k+1|k}(\mathbf{x}_n)+d_{k+1|k}(\mathbf{x}_n)p_{k+1|k}(S|\mathbf{x}_n)\right)\\
&\quad \cdot\sum_{i=1}^{n}b_{k+1|k}(S|\mathbf{x}_1)\cdots\frac{\delta}{\delta\mathbf{y}}b_{k+1|k}(S|\mathbf{x}_i)\cdots b_{k+1|k}(S|\mathbf{x}_n)\\
&\quad \sum_{i=1}^{n}\left(1-d_{k+1|k}(\mathbf{x}_1)+d_{k+1|k}(\mathbf{x}_1)p_{k+1|k}(S|\mathbf{x}_1)\right)\quad d_{k+1|k}(\mathbf{x}_i)f_{k+1|k}(\mathbf{y}|\mathbf{x}_i)
\end{aligned}$$

$$(1 - d_{k+1|k}(\mathbf{x}_n) + d_{k+1|k}(\mathbf{x}_n) p_{k+1|k}(S|\mathbf{x}_n))$$
$$\sum_{i=1}^{n} b_{k+1|k}(S|\mathbf{x}_1) \cdots \frac{\delta b_{k+1|k}}{\delta \mathbf{y}}(S|\mathbf{x}_1) \cdots b_{k+1|k}(S|\mathbf{x}_n)$$

Setting $S = \mathcal{S}$ and using the fact that $b_{k+1|k}(\mathcal{S}|\mathbf{x}_i) = 1$ for all i we get

$$\hat{D}_{k+1|k}(\mathbf{y}|X) = \frac{\delta \beta_{k+1|k}}{\delta \mathbf{y}}(\mathcal{S}|X) = \sum_{i=1}^{n} \left(d_{k+1|k}(\mathbf{x}_i) f_{k+1|k}(\mathbf{y}|\mathbf{x}_i) + \hat{b}_{k+1|k}(\mathbf{y}|\mathbf{x}_i) \right)$$

as desired

5.0 CONCLUSIONS

In this paper I have used finite-set statistics (FISST) to provide a systematic and theoretically rigorous theoretical framework for the Stein-Tenney Probability Hypothesis Density (PHD) multitarget tracking approach. After framing the optimal multitarget tracking problem as a multitarget recursive Bayes filtering problem, I showed that the PHD is actually a first-order statistical moment of the time-evolving random track-set Γ_k. Consequently, the PHD approach can be interpreted as a multitarget analog of a constant-gain Kalman filter (e.g., the α-β-γ filter). I showed how PHD's can be computed directly using the "set derivative." I also showed how the conventional single-sensor, single-target recursive Bayes filtering equations can be generalized to the PHD case (though, in the case of the Bayes update step, this generalization can be only approximate).

Given this, from a computational point of view real-time optimal multitarget tracking is reduced to the problem of implementing a real-time single-target nonlinear filter that is capable of modeling the rather complex time-evolution of the PHD. This is in itself a difficult research problem that will not be successfully addressed by thirty-year-old techniques copied from textbooks. Rather, it will require advanced techniques currently under development by a number of researchers. See **[11, pp. 5-6, 15-16]** for a brief discussion of the major computational issues and some of the major computational strategies.)

The approach outlined here can, in principle, be generalized to develop a filtering approach based on the second-order multitarget moment densities $D_{k|k}(\{\mathbf{x}_1, \mathbf{x}_2\}|Z^{(k)})$. Since second moments can be reformulated to have the general form of a covariance, it is aso possible in principle to develop a statistical multitarget analog of the Kalman filter.

BIBLIOGRAPHY

[1] M. Bardin (1984) "Multidimensional Point Processes and Random Closed Sets," *J. Applied Prob.*, vol. 21, pp. 173-178

[2] Y. Bar-Shalom and X.-R. Li (1993) *Estimation and Tracking: Principles, Techniques, and Software*, Artech House

[3] R.E. Bethel and G.J. Paras (1994) "A PDF multitarget-tracker," IEEE Trans AES, Vol. 30 No. 2, pp. 386-403

[4] D.J. Daley and D. Vere-Jones (1988) *An Introduction to the Theory of Point Processes*, Springer-Verlag

[5] I.R. Goodman (1982) "Fuzzy sets as equivalence classes of random sets," in R. Yager, ed., *Fuzzy Sets and Possibility Theory*, Permagon Press, 1982, pp. 327-343

[6] I.R. Goodman and H.T. Nguyen (1985) *Uncertainty Models for Knowledge Based Systems*, North-Holland, Amsterdam

[7] I.R. Goodman, R.P.S. Mahler, and H.T. Nguyen (1997) *Mathematics of Data Fusion*, Kluwer Academic Publishers

[8] Y.C. Ho and R.C.K. Lee (1964) A Bayesian Approach to Problems in Stochastic Estimation and Control, *IEEE Trans. AC*, Vol. AC-9, pp. 333-339

[9] A.H. Jazwinski (1970) *Stochastic Processes and Filtering Theory*, Academic Press

[10] A.D. Lanterman, M.I. Miller, D.L. Snyder, and W.J. Miceli (1994) "Jump-diffusion processes for the automated understanding of FLIR scenes," SPIE Proc., Vol. 2234, pp. 416-427

[11] R. Mahler (2000) *An Introduction to Multisource-Multitarget Statistics and Its Applications*, Lockheed Martin Technical Monograph, March 15 2000, 104 pages

[12] R. Mahler (1999) "Multisource-multitarget detection and acquisition: A unified approach," *SPIE Proc.* Vol. 3809, pp. 218-229

[13] R. Mahler (2001) "Multisensor-Multitarget Statistics," in D. Hall (ed.), *Data Fusion Handbook*, CRC Press, Boca Raton, to appear

[14] R. Mahler (2000) "The Search for Tractable Bayes Multitarget Filters," *SPIE Proc.* Vol. 4048, to appear

[15] S. Musick, K. Kastella, and R. Mahler (1998) "A practical implementation of joint multitarget probabilities," SPIE Proc., Vol. 3374, pp. 26-37

[16] N. Portenko, H. Salehi, and A. Skorokhod (1997) "On optimal filtering of multitarget tracking systems based on point processes observations," Random Operators and Stochastic Equations, Vol. 1, pp. 1-34

[17] H.W. Sorenson (1988) "Recursive Estimation for Nonlinear Dynamic Systems," in J.C. Spall (ed.), *Bayesian Analysis of Statistical Time Series and Dynamic Models*, Marcel Dekker

[18] M.C. Stein (1995) "Recursive Bayesian Fusion for Force Estimation," *Proc. 8th Nat'l Symp. on Sensor Fusion*, Vol. I (Unclassified), Mar. 15-17 1995, Dallas TX, pp. 47-66

[19] M.C. Stein and C.L. Winter (1993), "An Additive Theory of Probabilistic Evidence Accrual," Los Alamos National Laboratories report LA-UR-93-3336

[20] M. Stein and R.R. Tenney, "What's the Difference Between PHS and MHT?", undated Oasis Research and Alphatech working paper.

[21] L.D. Stone, C.A. Barlow, and T.L. Corwin (1999) Bayesian Multiple Target Tracking, Artech House

[22] L.D. Stone, T.L. Corwin, and J.B. Hofmann (1995) "Technical Documentation of Nodestar," Naval Research Laboratory Technical Report NRL/FR/5580-95-9788

[23] K.M. Tao, R. Abileah, and J.D. Lawrence (2000) "Multiple-target tracking in dense noisy environments: a probabilistic mapping perspective," *SPIE Proc.* Vol. 4048, to appear

[24] R.B. Washburn (1987) "A Random Point Process Approach to Multi-Object Tracking," Proc. Amer. Contr. Conf., Vol. 3, June 10-12 1987, Minneapolis, pp. 1846-1852