

Review

# “Statistics 103” for Multitarget Tracking

Ronald Mahler

Random Sets LLC, Eagan, MN 55122, USA; MahlerRonald@comcast.net

Received: 3 December 2018; Accepted: 24 December 2018; Published: 8 January 2019



**Abstract:** The finite-set statistics (FISST) foundational approach to multitarget tracking and information fusion was introduced in the mid-1990s and extended in 2001. FISST was devised to be as “engineering-friendly” as possible by avoiding avoidable mathematical abstraction and complexity—and, especially, by avoiding measure theory and measure-theoretic point process (p.p.) theory. Recently, however, an allegedly more general theoretical foundation for multitarget tracking has been proposed. In it, the constituent components of FISST have been systematically replaced by mathematically more complicated concepts—and, especially, by the very measure theory and measure-theoretic p.p.’s that FISST eschews. It is shown that this proposed alternative is actually a mathematical paraphrase of part of FISST that does not correctly address the technical idiosyncrasies of the multitarget tracking application.

**Keywords:** multitarget tracking; random finite set; point process; finite-set statistics

## 1. Introduction

The finite-set statistics (FISST) foundational approach to multitarget tracking and information fusion—stochastic geometry, random finite sets (RFS’s), belief-mass functions, and set derivatives and integrals—was introduced in the mid-1990s [1]. Its current extended form—probability generating functionals (p.g.fl.’s) and Volterra functional derivatives [2–4]—dates from 2001 [5]. FISST has inspired work by dozens of research groups in at least 20 nations; and FISST publications have been cited tens of thousands of times. A short survey of the FISST state-of-the-art c. 2015 can be found in Ref. [6]. The currently most advanced FISST-based algorithm, the generalized labeled multi-Bernoulli (GLMB) filter [3,7,8], is capable of real-time tracking of over one million 2D targets in clutter using off-the-shelf computing equipment [9].

FISST was devised to be as “engineering-friendly” as possible by *avoiding avoidable mathematical abstraction and complexity* [4]. Few tracking engineers have studied measure theory and far fewer are proficient. Still fewer have studied point process (p.p.) theory (which typically requires proficiency in measure theory), and few are proficient. For this reason, FISST does not employ measure theory or measure-theoretic p.p.’s, because simpler and more practical concepts, such as multitarget density functions, RFS’s, and Volterra functional derivatives, suffice.

Despite its “engineering-friendly” emphasis, FISST has inspired two rather contradictory reactions. Some have insinuated that FISST is probably unnecessary because it will probably turn out to be just a mathematical obfuscation of multi-hypothesis tracker (MHT) theory. Such a stance is quite mistaken and has been addressed in the tutorial [10].

Others, however, have recently intimated that FISST is *insufficiently complex* because it is *insufficiently general*. They have systematically replaced the constituent components of FISST with mathematically more complicated concepts—and, especially, with the very measure theory and measure-theoretic p.p. theory that FISST eschews.

It has been my observation, as well as that of others, that most tracking engineers—even those very familiar with measure and p.p. theories—must invest a great deal of effort to digest such

papers in order to decipher their possible contributions. It is thus entirely possible for mathematical complexity to obscure mathematical or engineering missteps. A central question, therefore, is this: What *engineering* advantages, if any, do sigma-algebras, measures, measure-theoretic p.p.'s, and other mathematically sophisticated concepts offer—especially given that FISST-based algorithms such as the GLMB filter offer unprecedented capability?

This review paper, which is a sequel to Refs. [4,11], is intended to answer this question. It will address the alternative multitarget statistical theory described in Refs. [12–14], with (for the sake of specificity) emphasis on Ref. [13]. This theory will hereafter be referred to as “measure-theoretical point-process multitarget tracking”, or “MPMT” for short. It will be demonstrated that MPMT is a *mathematical paraphrase* of part of FISST—one that, moreover, does not correctly address the technical idiosyncrasies of the multitarget tracking application.

A mathematical paraphrase is a substitution of terminology, notation, concepts, ideas, or processes that are *equivalent in mathematical meaning* to the terminology, notation, concepts, ideas, or processes in the original. In particular, it will be shown that:

1. When applied to practical multitarget tracking, p.p.'s are not “more general” than RFS's.
2. The “regional variance” of Ref. [12] *does* admit a density—thereby refuting the only offered evidence in Refs. [12–14] that MPMT is unavoidable for practical multitarget tracking.
3. When applied to multitarget tracking, the “chain differential” is identical to the Gâteaux and Fréchet derivatives—and thus mathematically equivalent to the FISST functional derivative.

These, and other noteworthy facts that follow, have thus far been overlooked in the tracking literature. It is therefore important that such oversights be carefully addressed.

The paper will address the following replacements of FISST concepts with MPMT concepts: RFS's with p.p.'s (Section 2); FISST densities with “measures” (Section 3); set integrals with measure-theoretic integrals (Section 4); functional derivatives with “chain differentials” (Section 5); the FISST product rule with “Leibniz' Rule” (Section 6); and RFS motion models with p.p. motion models (Section 7). Mathematical derivations can be found in Section 8 and Conclusions in Section 9. The discussions have been made as tutorial as feasible.

## 2. RFS's Replaced by “Point Processes” (p.p.'s)

MPMT replaces RFS's with “... the more general concept of point process” [13] (p. 1324) (This phrase is logically vacuous: “more general” than what? What is implicitly meant is “more general than RFS's”). Specifically, if  $\mathfrak{R}$  denotes the real numbers then:

“... the population of targets is represented by a point process  $\Phi$ , on a single-target state space  $\mathbf{X} \subseteq \mathfrak{R}^d$ , whose elements describe individual target states. A realization of  $\Phi$  is a vector of points  $\phi = (x_1, \dots, x_N)$  depicting a specific multitarget configuration, where  $x_i \in \mathbf{X}$  ... A point process  $\Phi$  is characterized by its probability distribution  $P_\Phi$  on the measurable space  $(X, \mathbf{B}_X)$ , where  $X = \cup_{n \geq 0} \mathbf{X}^n$  is the point process state space, i.e., the space of all the finite vectors of points in  $\mathbf{X}$ , and  $\mathbf{B}_X$  is the Borel  $\sigma$ -algebra on  $X$  ... The probability distribution of a point process is defined as a symmetric function, so that the order of points in a realization is irrelevant for statistical purposes ...”. [13] (p. 1325)

The following subsections address the topics: RFS's are not an “alternative construction” (Section 2.1); RFS's are simpler than simple p.p.'s (Section 2.2); non-RFS p.p.'s are inappropriate for multitarget tracking (Section 2.3); vectors are poor multitarget state representations (Section 2.4); simple p.p.'s produce a flawed mathematical paraphrase of RFS's (Section 2.5); and FISST is actually more general than MPMT (Section 2.6).

### 2.1. RFS's Are Not an "Alternative Construction"

In MPMT, p.p.'s are assumed to be "simple" (i.e., the  $x_1, \dots, x_n$  in  $(x_1, \dots, x_n)$  are distinct), while it is also asserted that an RFS is an "alternative construction" of a simple p.p. that is "also available in the literature" [13] (p. 1325, footnote). *This is misleading.* It is simple p.p.'s that are being proffered as an alternative to RFS's for application to multitarget tracking.

It could be argued to the contrary that, in the pure-mathematics literature of many decades ago, RFS's historically arose as an alternative to the three formulations of p.p.'s originally proposed by Moyal in Ref. [15]. However, any such claim overlooks the following fact: *When signal processing engineers apply concepts drawn from the pure-mathematics literature, they typically create original intellectual property which must be properly acknowledged as such* (Otherwise, why would one need signal processing engineers?) Moyal's paper addressed no practical applications at all, and appeared at the same time as the Kalman filter and nearly 20 years before Reid's seminal MHT paper [16]. Nearly a half-century after [15], FISST was devised as a novel application of stochastic geometry (not p.p. theory) to a specific engineering application: multitarget tracking and information fusion. It is this original application that requires proper attribution.

To state the issue plainly: In the recent engineering tracking literature, the "p.p." model of a random multitarget state in Refs. [12–14] is, historically speaking, being promoted as an *alternative* to the original FISST RFS model of a random multitarget state—not the other way around.

### 2.2. RFS's Are Simpler than Simple p.p.'s

Contrast the definition of a p.p. previously given with the definition of an RFS:

An RFS  $\Xi$  of the single-target state space  $\mathfrak{S}$  is a random variable whose realizations are the finite subsets  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  of  $\mathfrak{S}$  of cardinality  $n \geq 0$ .

This requires only simple concepts easily understood by engineers: random variable, finite set, cardinality. There is no need for measurable spaces, Borel sigma-algebras, or probability measures (symmetric or otherwise) (RFS's do have a measure-theoretic basis, but in practical application it can usually be ignored—see Section 3.1.)

**Remark 1.** *The fact that finite sets are order-free does not mean that we cannot distinguish between targets. In general, a single-target state will have the form  $\mathbf{x} = (\mathbf{u}, \ell)$  where  $\mathbf{u}$  is the kinematic state and  $\ell$  is a uniquely identifying track label [1] (pp. 135, 196–197). This is the basis for the labeled RFS (LRFS) theory of Vo and Vo [7–9]; [3] (Chapter 15). In LRFS theory,  $\mathbf{u}$  and  $\ell$  are random variables and the  $\ell$  are unordered symbols.*

### 2.3. Non-RFS p.p.'s Are Inappropriate for Multitarget Tracking

This is because non-RFS p.p.'s are technically deficient representations of random multitarget states. Every target track must have a unique identifying label—for example, "Bob." Given this,  $\mathbf{x} = (\mathbf{u}, \text{Bob})$  cannot occur more than once in  $(\mathbf{x}_1, \dots, \mathbf{x}_n)$  since, otherwise, "Bob" would be present twice or more simultaneously. Thus all state-p.p.'s must be simple—i.e., they must be RFS's—and so the claim that p.p.'s are "more general [than RFS's]" is false in actual engineering application. And, in any case, immediately after this claim was made all p.p.'s were assumed to be simple.

### 2.4. Vectors Are Poor Multitarget State Representations

There are multiple reasons for this [17] (Section II-A); [18] (Section 4.2.3). (1) The targets in a multitarget population have no natural ordering. Imposing physically extraneous information on it, such as an order, risks the creation of unknown statistical biases. (2) If the system has  $n$  distinct targets  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  then it has  $n!$  vector state representations  $\phi_{X,\pi} = (\mathbf{x}_{\pi 1}, \dots, \mathbf{x}_{\pi n})$  (for permutations  $\pi$  on  $1, \dots, n$ )—whereas ideally there should be a one-to-one correspondence between physical states and their representations. (3) The goal of multitarget algorithms is to produce estimates of the multitarget state that are as close as possible to ground truth. A mathematical distance metric on multitarget states

is required to do so. Assume that there exists a metric  $\rho(\phi_1, \phi_2)$  on *vector states*. It must be constant under permutation of the entries of  $\phi_1$  and  $\phi_2$ —in which case  $\rho(\phi_1, \phi_2) = \rho'(\chi(\phi_1), \chi(\phi_2))$  for some  $\rho'$ , where  $\chi(\phi)$  denotes the set of entries in  $\phi$ . Let  $\phi_\pi = (x_{\pi_1}, \dots, x_{\pi_m})$ . Then  $\rho(\phi_\pi, \phi_{\pi'}) = \rho(\phi_\pi, \phi_\pi) = 0$  for permutations  $\pi \neq \pi'$ , contradicting the definition of a metric. That is: *no metric on vector states exists*. Finite sets have well-known metrics such as “OSPA” [3].

**Remark 2.** *It could be argued that, because the probability distribution of a p.p. is symmetric, “... the order of points in a realization is irrelevant for statistical purposes” [13] (p. 1325). This is immaterial. Distance is an intrinsic, deterministic property of a multitarget state space that is independent of any particular probability distribution on that space.*

**Remark 3.** *It should be pointed out that one of the authors of Refs. [12–14], as a coauthor of Ref. [18] (Section 4.2.3), marshaled similar arguments to similarly criticize vector representation.*

Finally, Ref. [13] appears to employ finite sets while the contrary is claimed. (1) Most obviously: “... the abuse of notation ‘ $x \in \phi$ ’ is used for ‘ $x \in \chi(\phi)$ ’ where  $\chi$  is the function associating a vector of distinct elements to the set composed of the same elements” [13] (footnote 3) (Such “abuse” is required only because vectors have been needlessly substituted in place of finite sets). (2) OSPA is used to measure distance between vectors (This is theoretically problematic since  $\rho(\phi_1, \phi_2) = \rho_{\text{OSPA}}(\chi(\phi_1), \chi(\phi_2))$  cannot be a metric). (3) The set-theoretic notation “ $\Phi_1 \cup \Phi_2$ ” is used to denote the “superposition” (i.e., set-theoretic union) of simple p.p.’s (i.e., RFS’s)  $\Phi_1, \Phi_2$  [13] (Proposition 1).

### 2.5. Simple p.p.’s Produce a Flawed Mathematical Paraphrase of RFS’s

The replacement of every finite subset  $X = \{x_1, \dots, x_n\} \subseteq \mathfrak{S}$  with a vector  $\phi = (x_1, \dots, x_n) \in X$  and every RFS  $\Xi$  with a simple p.p.  $\Phi$  results in a conceptually questionable and unnecessarily complexified mathematical paraphrase of FISST that does not correctly address the technical idiosyncrasies of the multitarget tracking application.

### 2.6. FISST is More General than MPMT

This is because FISST (a) has an integro-differential calculus of possibly nonadditive set functions and their density functions (Section 3.1); and (b) it Bayes-optimally addresses multitarget-multisource information fusion using “hard + soft” data in a unified manner [2] (Chapters 3–7); [3] (Chapter 22). The latter is attributable to the fact that FISST is based on stochastic geometry, which in turn is based on the theory of random closed subsets (RCS’s) [19], which in turn is the basis of FISST’s unification of “hard + soft” information fusion.

## 3. FISST Densities Replaced by “Measures”

MPMT replaces the former with the latter because:

“... a measure-theoretical formulation provides a more general framework that is required to construct certain statistical properties on point processes that can be exploited for practical applications; a recent example is given in [21] for the construction of the regional statistics ...”. [13] (p. 1325)

(The phrase “more general framework” is again logically vacuous: “more general” than what? What is implicitly meant is “more general than FISST”.) Here, “regional statistics” refers to the “regional variance” of Ref. [12]—i.e., the variance of the random integer  $|\Xi \cap S|$ :

$$\text{var}_{\Xi}(S) = E[|\Xi \cap S|^2] - E[|\Xi \cap S|]^2. \quad (1)$$

In Ref. [12] it was claimed that the set function  $\text{var}_{\Xi}(S)$  “... is ... not a measure ... [and so it] does not necessarily admit a density in general ... This fact motivates the measure-theoretical approach ...” This is not the case, because (as we shall see)  $\text{var}_{\Xi}(S)$  *does admit a density*.

First, however, readers should be advised that the meaning of “measure” in Refs. [12–14] is often unclear. For example, since  $\text{var}_{\Xi}(S)$  is *not* a measure, how can it motivate “the measure-theoretical approach”? Sometimes “measure” has its usual meaning: a nonnegative set function  $\mu(S)$  such that  $\mu(\cup_{n \geq 1} S_n) = \sum_{n \geq 1} \mu(S_n)$  for mutually disjoint  $S_n$ . Other times, however, it means *nonadditive* set functions such as  $\text{var}_{\Xi}(S)$ .

The following subsections will address: the elements of FISST (Section 3.1); measure-theoretic p.p. theory (Section 3.2); the FISST density of the regional variance (Section 3.3); why measures are inappropriate for multitarget tracking (Section 3.4); and why the “measure-theoretical formulation” in Refs. [12–14] produces a complexified mathematical paraphrase of FISST (Section 3.5).

### 3.1. Basic Concepts of Finite-Set Statistics

This section is drawn from Ref. [4]. The theoretical basis of *single-target statistics* is the *probability measure*  $p_{\mathbf{X}}(S) = \Pr(\mathbf{X} \in S)$  of a random vector  $\mathbf{X} \in \mathfrak{S}$  (not to be confused with the p.p. single-target state space  $\mathbf{X} = \mathfrak{S}$ ). Single-target tracking requires the *probability density* of  $p_{\mathbf{X}}(S)$ :

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{dp_{\mathbf{X}}}{d\lambda}(\mathbf{x}) \quad (2)$$

where the right side is the Radon-Nikodým derivative of  $p_{\mathbf{X}}(S)$  with respect to Lebesgue measure  $\lambda(S)$  on  $\mathfrak{S} \subseteq \mathbb{R}^N$ . That is:  $f_{\mathbf{X}}(\mathbf{x})$  has the property that  $\int_S f_{\mathbf{X}}(\mathbf{x}) d\lambda(\mathbf{x}) = p_{\mathbf{X}}(S)$  where  $\int_S \cdot d\lambda = \int \cdot \mathbf{1}_S(\mathbf{x}) d\lambda(\mathbf{x})$  and  $\mathbf{1}_S(\mathbf{x})$  is the indicator function of subset  $S \subseteq \mathfrak{S}$ .

The goal of FISST was to *reformulate multitarget tracking as a generalized single-target tracking problem*, with RFS’s  $\Xi$  taking the place of random vectors  $\mathbf{X}$ . The theoretical basis of *multitarget statistics* is the probability measure  $p_{\Xi}(O) = \Pr(\Xi \in O)$  over the Borel-measurable subsets  $O$  of the hyperspace whose elements are the finite subsets of single-target state space  $\mathfrak{S}$ . (A “hyperspace” is a space whose elements are subsets of some other “base space.”) FISST avoids  $p_{\Xi}(O)$  by equivalently replacing it with the stochastic-geometric *belief measure* (a.k.a. belief-mass function)  $\beta_{\Xi}(S) = \Pr(\Xi \subseteq S)$ —a conceptually simple generalization of  $p_{\mathbf{X}}(S) = \Pr(\mathbf{X} \in S)$ .

**Remark 4.** *The belief measure can usually be avoided since it is usually necessary only for motion and measurement modeling—see Section 7.1.*

Multitarget tracking is based on the *multitarget probability density* of  $\beta_{\Xi}(S)$ —i.e., the multitarget analog of Equation (2):

$$f_{\Xi}(X) = \frac{\delta \beta_{\Xi}}{\delta X}(\emptyset) = \left[ \frac{\delta \beta_{\Xi}}{\delta X}(S) \right]_{S=\emptyset} \quad (3)$$

where  $\emptyset$  denotes the empty set and where the right side is a FISST *set derivative* of  $\beta_{\Xi}(S)$  with respect to  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subseteq \mathfrak{S}$ . (The set derivative is a constructivist generalization of the Radon-Nikodým derivative:  $(dp_{\mathbf{X}}/d\lambda)(\mathbf{x}) = (\delta p_{\mathbf{X}}/\delta \{\mathbf{x}\})(\emptyset)$ —see Ref. [4] (Section IV-F).)

A related density is the FISST multitarget factorial moment density:

$$D_{\Xi}(X) = \frac{\delta \beta_{\Xi}}{\delta X}(\mathfrak{S}) = \left[ \frac{\delta \beta_{\Xi}}{\delta X}(S) \right]_{S=\mathfrak{S}}. \quad (4)$$

The special case  $D_{\Xi}(\mathbf{x}) = D_{\Xi}(\{\mathbf{x}\})$  is known as the *probability hypothesis density* (PHD) of  $\Xi$ .

**Remark 5.**  *$f_{\Xi}(X)$  and  $D_{\Xi}(X)$  were defined in 1997 in Ref. [1] using stochastic geometry and set derivatives—not p.p. theory. Likewise for the first derivation [20] of the PHD filter.*

For any real-valued function  $h(\mathbf{x})$  and any finite  $X \subseteq \mathfrak{S}$ , let  $h^X = 1$  if  $X = \emptyset$  and  $h^X = \prod_{\mathbf{x} \in X} h(\mathbf{x})$  otherwise. Then the multitarget analog of  $p_{\mathbf{X}}(S) = \int_S f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$  is

$$\beta_{\Xi}(S) = \int_S f_{\Xi}(X) \delta X = \int \mathbf{1}_S^X \cdot f_{\Xi}(X) \delta X, \quad (5)$$

where the set integral  $\int f(X) \delta X$  of a multitarget density function  $f(X)$  is defined as

$$\int f(X) \delta X = \sum_{n \geq 0} \frac{1}{n!} \int f(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n. \quad (6)$$

The regional set integral  $\int_S f(X) \delta X$  is nonadditive in  $S$  because  $S \mapsto \prod_{\mathbf{x} \in X} \mathbf{1}_S(\mathbf{x})$  is nonadditive (It is not true that integrals must be additive in  $S$ —see, for example, [21]).

The set derivative has the following important property. Let  $\sigma(S)$  be a nonnegative set function defined on the closed subsets  $S \subseteq \mathfrak{S}$ . Then if it exists, its FISST multitarget density is  $\sigma^*(X) = (\delta\sigma/\delta X)(\emptyset)$  since

$$\int_S \frac{\delta\sigma}{\delta X}(\emptyset) \delta X = \sigma(S). \quad (7)$$

### 3.2. Measure-Theoretical p.p. Theory

The “measure-theoretical formulation” of p.p. theory in MPMT is stated as follows:

“The probability distribution  $P_{\Phi}$  [of a simple p.p.  $\Phi$ ] is characterized by its projection measures  $P_{\Phi}^{(n)}$ , for any  $n \geq 0$ . The  $n$ th-order projection measure  $P_{\Phi}^{(n)}$ , for any  $n \geq 1$ , is defined on the Borel  $\sigma$ -algebra of  $\mathbf{X}^n$  and gives the probability for the point process to be composed of  $n$  points, and the probability distribution of these points ... For any  $n \geq 0$ ,  $J_{\Phi}^{(n)}$  denotes the  $n$ th-order Janossy measure ... and is defined as  $J_{\Phi}^{(n)}(B_1 \times \dots \times B_n) = n! P_{\Phi}^{(n)}(B_1 \times \dots \times B_n)$  ... The probability density  $p_{\Phi}$  (respectively (resp.) the  $n$ th-order projection density  $p_{\Phi}^{(n)}$ , the  $n$ th-order Janossy density  $j_{\Phi}^{(n)}$ ) is the Radon-Nikodým derivative of the probability distribution  $P_{\Phi}$  (resp. the  $n$ th-order projection measure  $P_{\Phi}^{(n)}$ , the  $n$ th-order Janossy measure  $J_{\Phi}^{(n)}$ ) with respect to (w.r.t.) some reference measure ... Throughout this article the exploitation of the Janossy measures will be preferred, for they are convenient tools in the context of functional differentiation ... ”. [13] (p. 1325)

The “ $k$ th-order factorial moment measure”  $M_{\Phi}^{(k)}(B_1, \dots, B_k)$  and its density  $m_{\Phi}^{(k)}(x_1, \dots, x_k)$  are also introduced [13] (Equation (20)). MPMT is related to FISST as follows:

$$j_{\Xi}^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{1}{n!} \cdot f_{\Xi}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \quad (8)$$

$$m_{\Xi}^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{1}{n!} \cdot D_{\Xi}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \quad (9)$$

for distinct  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . (If  $\mathbf{x}_1, \dots, \mathbf{x}_n$  are distinct then the factor  $1/n!$  on the right sides of Equations (8) and (9) apportion the probability of  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  equally among the  $n!$  vectors that have the same elements as  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ ).

In summary: the families of multivariate measures  $J_{\Phi}^{(k)}(B_1, \dots, B_k)$  resp.  $M_{\Phi}^{(k)}(B_1, \dots, B_k)$  for  $n \geq 1$ —i.e., the very measures that FISST rejected as unnecessary—have been substituted in place of  $f_{\Xi}(X)$  resp.  $D_{\Xi}(X)$ .

This restoration—and thus MPMT—is allegedly unavoidable because (a) measures are “convenient tools” for “functional differentiation”; and (b) the fact that  $\text{var}_{\Xi}(S)$  does not have a density proves that “... a measure-theoretical formulation provides a more general framework [than FISST] ... for practical applications ... ” Neither assertion is true:  $\text{var}_{\Xi}(S)$  does admit a density (Section 3.3); and measures are unnecessary for functional differentiation (Section 5.4).

Moreover, this restoration strips away a primary FISST insight: that all information about a multitarget system  $\Xi_{k|k}$  at time  $t_k$  can be represented by a single multitarget probability density function  $f_{k|k}(X|Z_{1:k})$ —i.e., the multitarget probability density function of  $\Xi_{k|k}$ . The recent very fast implementations of the GLMB filter have been possible only because advanced stochastic sampling techniques can be applied to  $f_{k|k}(X|Z_{1:k})$ —see Ref. [9] (pp. 1–2).

### 3.3. The FISST Multitarget Density of the Regional Variance

Contrary to the claim in Refs. [12–14],  $\text{var}_{\Xi}(S)$  does admit a density even though it is not an additive measure. Specifically, recall the FISST set derivative (Section 3.1) and define:

$$\text{var}_{\Xi}^*(X) = \frac{\delta \text{var}_{\Xi}}{\delta X}(\emptyset). \quad (10)$$

In Section 8.1 it is shown that this equals 0 unless  $|X| = 2$ , in which case

$$\text{var}_{\Xi}^*(\{\mathbf{x}_1, \mathbf{x}_2\}) = 2D_{\Xi}(\{\mathbf{x}_1, \mathbf{x}_2\}) + 2\delta_{\mathbf{x}_1}(\mathbf{x}_2) \cdot D_{\Xi}(\mathbf{x}_1) - 2D_{\Xi}(\mathbf{x}_1) \cdot D_{\Xi}(\mathbf{x}_2) \quad (11)$$

for distinct  $\mathbf{x}_1, \mathbf{x}_2$ . By Equation (7) it must be the case that

$$\int_S \text{var}_{\Xi}^*(X) \delta X = \text{var}_{\Xi}(S). \quad (12)$$

This fact is verified in Section 8.2 for completeness. Thus  $\text{var}_{\Xi}^*$  is the FISST density of  $\text{var}_{\Xi}$ .

Consequently and contrary to claim, the existence of the regional variance does not prove the unavailability of MPMT. Moreover, the fact that Equation (1) might be easier to use than Equation (11) in some circumstances is meager justification for wholesale adoption of formal measure theory (which in any case is inapplicable to  $\text{var}_{\Xi}(S)$  since it is not an additive measure).

### 3.4. Measures Are Inappropriate for Practical Multitarget Tracking

This is because (a) practical Bayes-optimal multitarget state estimation requires densities; and (b) the multitarget Bayes' rule as specified in Ref. [13] (Equation (22)),

$$P_k(d\phi|Z_{1:k}) = \frac{g_k(Z_k|\phi) \cdot P_{k|k-1}(d\phi|Z_{1:k-1})}{\int g_k(Z_k|\bar{\phi}) \cdot P_{k|k-1}(d\bar{\phi}|Z_{k:k-1})}, \quad (13)$$

requires the density function  $P_{k|k-1}(d\phi|Z_{1:k-1}) = p_{k|k-1}(\phi|Z_{1:k-1})$  in the numerator.

This point is implicitly conceded in [14] (p. 49), where the authors "... assume that ... all [additive] measures studied in this article ... admit densities".

**Remark 6.** Indeed,  $P_{k|k-1}(d\phi|Z_{1:k-1})$  is a mathematically equivalent substitution for the FISST  $f_{k|k-1}(X|Z_{1:k-1})$ ; and  $\int P_{k|k-1}(d\phi|Z_{1:k-1}) = \int p_{k|k-1}(\phi|Z_{1:k-1}) d\lambda_c^{\cup}(\phi)$  is an equivalent substitution for the FISST set integral  $\int f_{k|k-1}(X|Z_{1:k-1}) \delta X$ —see Section 4.

**Remark 7.** It might nevertheless be objected that there is a purely measure-theoretic version of Bayes' rule, the Killianpur-Striebel formula. It is immaterial since it is not employed in Refs. [12–14] despite the "measure-theoretical" emphasis of these papers. And if it had been, it would have only produced another mathematical paraphrase of FISST that begs the question: what significant engineering advances result from using it rather than Bayes' rule?

**Remark 8.** Since Dirac deltas are density functions, even singular measures can have density functions. For example, consider the bivariate measure  $\mu_{\Xi}(S_1, S_2) = E[|\Xi \cap S_1 \cap S_2|]$ . Its density function can be shown to be  $f(x, y) = \delta_y(x) \cdot D_{\Xi}(y)$ . See also Equation (11).

### 3.5. FISST Densities vs. Additive/Nonadditive Measures

For the purposes of multitarget tracking, families of multivariate measures, such as  $J^{(k)}_{\Phi}(B_1, \dots, B_k)$  or  $M^{(k)}_{\Phi}(B_1, \dots, B_k)$  for  $n \geq 1$ , are mathematically equivalent to, but mathematically far more complicated than, the FISST multitarget density functions that they replace, such as  $f_{\Xi}(X)$  and  $D_{\Xi}(X)$ . Consequently, replacing every FISST density with measures (or some other set function) produces a mathematically complexified mathematical paraphrase of FISST that is inappropriate for practical multitarget tracking *since densities are unavoidable*.

## 4. Set Integrals Replaced by Measure-Theoretic Integrals

The set integral  $\int \cdot \delta X$  was described in Section 3.1. MPMT replaces it with an integral  $\int \cdot d\lambda(\phi)$  with respect to an *unspecified* “reference measure”  $\lambda$  [13] (Equation (2)). This is misleading, because  $\lambda$  cannot be arbitrary. If it is to be applicable to multitarget tracking it must be an extension of Lebesgue measure on  $\mathfrak{S} \subseteq \mathbb{R}^N$  to  $\mathfrak{S}^{\infty} = \bigcup_{n \geq 0} \mathfrak{S}^n$ .

The following subsections address: the extension of Lebesgue measure  $\lambda$  on  $\mathfrak{S} \subseteq \mathbb{R}^N$  to a measure  $\lambda_c^{\cup}$  on  $\mathfrak{S}^{\infty}$  (Section 4.1); why the measure-theoretic integral  $\int \cdot d\lambda_c^{\cup}(\phi)$  is problematic from the point of view of practical multitarget tracking (Section 4.2); and why the substitution of  $\int \cdot d\lambda_c^{\cup}(\phi)$  in place of  $\int \cdot \delta X$  in Refs. [12–14] produces a conceptually flawed, complexified mathematical paraphrase of FISST (Section 4.3).

### 4.1. Extending Lebesgue Measure to Multitarget States

The following is drawn from Ref. [2] (Appendices F.3 and F.4). Suppose that  $\mathfrak{S} \subseteq \mathbb{R}^N$  for some  $N$  and let  $\lambda(S)$  be Lebesgue measure on  $\mathfrak{S}$ . How can  $\lambda$  and Equation (2) be extended to  $\mathfrak{S}^{\infty} = \bigcup_{n \geq 0} \mathfrak{S}^n$ ? Begin with  $\lambda$ . Let  $\lambda^n(O')$  be the usual extension of  $\lambda$  to the Cartesian-product space  $\mathfrak{S}^n$  for measurable  $O' \subseteq \mathfrak{S}^n$ . Let  $O \subseteq \mathfrak{S}^{\infty}$  be measurable—i.e.,  $O^{(n)} = O \cap \mathfrak{S}^n$  is measurable in  $\mathfrak{S}^n$  for every  $n \geq 1$ , in which case  $\lambda^n(O^{(n)})$  exists for every  $n \geq 1$ . If the unit of measurement in  $\mathfrak{S}$  is  $\iota$  then the unit of measurement of  $\lambda^n(O^{(n)})$  is  $\iota^n$ . Let  $c > 0$  be a constant whose unit of measurement is  $\iota$ . Define the extension of  $\lambda$  to  $\mathfrak{S}^{\infty}$  as:

$$\lambda_c^{\cup}(O) = \sum_{n \geq 0} \frac{\lambda^n(O^{(n)})}{c^n}. \quad (14)$$

This is well-defined since each term in the sum is unitless.

Next let  $f(\phi)$  be a unitless, nonnegative function of  $\phi \in \mathfrak{S}^{\infty}$  and abbreviate  $f(\mathbf{x}_1, \dots, \mathbf{x}_n) = f((\mathbf{x}_1, \dots, \mathbf{x}_n))$  and  $\int \cdot d\mathbf{x}_1 \cdots d\mathbf{x}_n = \int \cdot d\lambda^n(\mathbf{x}_1, \dots, \mathbf{x}_n)$ . Then it is integrable with respect to  $\lambda_c^{\cup}$  if the following exists:

$$\int_O f(\phi) d\lambda_c^{\cup}(\phi) = \sum_{n \geq 0} \frac{1}{c^n} \int_{O^{(n)}} f(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \cdots d\mathbf{x}_n. \quad (15)$$

Now turn to the generalization of Equation (2). Let  $\mu(O)$  be a probability measure on  $\mathfrak{S}^{\infty}$  and let  $\mu^n$  denote its restriction to  $\mathfrak{S}^n$ . Recall that  $\mu$  is absolutely continuous with respect to (a.c.w.r.t.) another measure  $\mu_0$  if  $\mu(O) = 0$  whenever  $\mu_0(O) = 0$ . If  $\mu$  is a.c.w.r.t.  $\lambda_c^{\cup}$  then  $\mu^n$  is a.c.w.r.t.  $\lambda^n$  for all  $n \geq 1$ . Consequently, by the Radon-Nikodým theorem, for each  $n \geq 1$  there is an almost everywhere unique  $f_n(\phi)$  on  $\phi \in \mathfrak{S}^n$  such that

$$\mu^n(O') = \int_{O'} f_n(\phi) d\lambda^n(\phi) = \int_{O'} f(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \cdots d\mathbf{x}_n \quad (16)$$

for all measurable  $O' \subseteq \mathfrak{S}^n$ . The unit of measurement of  $f_n(\phi)$  is  $\iota^{-n}$ . Define the unitless function  $f_c(\phi) = c^n \cdot f_n(\phi)$  if  $\phi = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ . Then

$$\mu(O) = \mu(\bigcup_{n \geq 0} O^{(n)}) = \sum_{n \geq 0} \mu^n(O^{(n)}) = \sum_{n \geq 0} \int_{O^{(n)}} f_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \cdots d\mathbf{x}_n \quad (17)$$

$$= \sum_{n \geq 0} \frac{1}{c^n} \int_{O^{(n)}} f_c(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \cdots d\mathbf{x}_n = \int_O f_c(\phi) d\lambda_c^\cup(\phi) \quad (18)$$

for all measurable  $O \subseteq \mathfrak{S}^\infty$ . That is:  $f_c(\phi) = (d\mu/d\lambda_c^\cup)(\phi)$  is the Radon-Nikodým density of  $\mu(O)$  w.r.t.  $\lambda_c^\cup$ —i.e., it is the extension of Equation (2) to  $\mathfrak{S}^\infty$  (If  $\mu = P_\Phi$  it is what in Ref. [13] is denoted as  $P_\Phi(d\phi)$  or  $p_\Phi(\phi)$ ).

This is *conceptually troublesome* since  $\mu$  has a different density for each  $c > 0$ . In p.p. theory, the usual resolution of this difficulty is to set  $c = 1 \cdot t$  [22] (pp. 1226–1229). But as we shall now see, this leads to a new conceptual difficulty when applied to multitarget tracking.

#### 4.2. Measure-Theoretic Integrals and Multitarget State Estimation

Define the FISST multitarget density function  $f(X)$  by

$$f(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = n! \cdot f_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (19)$$

for distinct  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . From Equations (6), (18) and (20), the measure-theoretic and set integrals are equivalent:

$$\int f_c(\phi) d\lambda_c^\cup(\phi) = \int f(X) \delta X. \quad (20)$$

Also, the maximum a posteriori estimate

$$\phi_c = \operatorname{argsup}_\phi f_c(\phi) \quad (21)$$

of  $f_c(\phi)$  is equivalent to FISST's JoM (Joint Multitarget) estimate of  $f(X)$  [2] (p. 498):

$$X_c = \operatorname{argsup}_X \frac{c^{|X|}}{|X|!} \cdot f(X). \quad (22)$$

As was explained in Ref. [2] (pp. 499–500), to arrive at an intuitively reasonable  $X_c$  the magnitude of  $c$  should be approximately equal to the accuracy with which any  $\mathbf{x} \in \mathfrak{S}$  can be estimated. Since this argument is fairly lengthy and involved, it cannot be reproduced here.

The *fixed* choice  $c = 1 \cdot t$  will, in general, produce poor JoM estimates of  $f(X)$  (and therefore poor MAP estimates of  $f_c(\phi)$ ). The only reasonable resolution is to attach  $c$  to a particular estimator—JoM—rather than to so fundamental a concept as a multitarget integral.

#### 4.3. Set Integrals vs. Measure-Theoretic Integrals

The measure-theoretic integral  $\int \cdot d\lambda_c^\cup(\phi)$  is mathematically equivalent to but mathematically far more complicated than the set integral  $\int \cdot \delta X$ , which is not measure-theoretic. Also, from the point of view of practical multitarget tracking  $\int \cdot \delta X$  resp.  $f(X)$  resp.  $X_c$  are preferable to  $\int \cdot d\lambda_c^\cup(\phi)$  resp.  $f_c(\phi) = P_\Phi(d\phi)$  resp.  $\phi_c$ . Consequently, replacing every set integral with a measure-theoretic integral, and every multitarget density with a Radon-Nikodým derivative, produces a flawed, complexified mathematical paraphrase of FISST.

### 5. Functional Derivatives Replaced by “Chain Differentials”

In MPMT the former is replaced with the latter “... so that a general chain rule can be determined...” [13] (p. 1326). The *plain meaning* of this phrase is: *the chain differential is necessary for a general chain rule (as applied in Ref. [13] to p.g.fl.'s)*. It is false for two reasons:

1. The FISST functional derivative already has a general chain rule—see [23]; [3] (pp. 78–79).
2. When applied to p.g.fl.'s the chain differential is identical to the Gâteaux and Fréchet derivatives—and thus mathematically equivalent to the FISST functional derivative.

The following subsections address the following topics: probability generating functionals (Section 5.1); differentiation theory (Section 5.2); differentiation of p.g.fl.'s (Section 5.3); equivalence of chain differentials and functional derivatives (Section 5.4); and the chain differential produces a complexified mathematical paraphrase of FISST (Section 5.5).

### 5.1. Probability Generating Functionals

The statistics of an RFS  $\Xi$  are equivalently characterized by  $\beta_{\Xi}(S)$  and  $f_{\Xi}(X)$ . A third fundamental statistical descriptor of  $\Xi$ , the *probability generating functional* (p.g.fl.), is:

$$G_{\Xi}[h] = \int h^X f_{\Xi}(X) \delta X \quad (23)$$

where the notation  $h^X$  was defined in Equation (5). For present purposes the “test function”  $h$  will be assumed to be a nonnegative bounded function, in which case  $0 \leq G_{\Xi}[h] < \infty$ . (FISST follows the practice in Ref. [24] of further assuming that  $0 \leq h(\mathbf{x}) \leq 1$ .) Note that  $G_{\Xi}[\mathbf{1}_S] = \beta_{\Xi}(S)$ .

A great many generating functionals besides the p.g.fl. are used in p.p. theory: characteristic, Laplace, moment, factorial-moment, cumulant, factorial-cumulant, Khinchin, etc., [24]. It was FISST that identified the particular importance of the p.g.fl. for multitarget tracking.

The p.g.fl. finds its greatest use in the derivation of approximate multitarget filters such as the PHD and cardinalized PHD (CPHD) filters. This, in turn, requires a differential calculus of p.g.fl.'s—the subject of the next two subsections.

### 5.2. Differentiation Theory

Let  $A, B$  be (possibly infinite-dimensional) topological linear spaces and let  $\tau: A \rightarrow B$  be a transformation. Then the *Gâteaux differential* is a simple and obvious generalization of the differential quotient of undergraduate calculus:

$$(\delta\tau)(a'; a) = \lim_{\varepsilon \rightarrow 0} \frac{\tau(a' + \varepsilon \cdot a) - \tau(a')}{\varepsilon}. \quad (24)$$

If the function defined by  $a \mapsto (\delta\tau)(a'; a)$  exists and is linear and continuous then  $(\delta\tau)(a'; \cdot)$  is called the *Gâteaux derivative* of  $\tau$  at  $a'$ .

Now recall that a Banach space is a normed topological linear space that is closed with respect to limits. (A norm is a nonnegative function  $\|\mathbf{x}\|$  such that  $\|\mathbf{x}\| = 0$  implies  $\mathbf{x} = 0$  and which satisfies the triangle inequality:  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ .)

Let  $A, B$  be Banach spaces with respective norms  $\|\cdot\|_A$  and  $\|\cdot\|_B$ . If there exists a linear-continuous function  $D_{a'}\tau: A \rightarrow B$  such that

$$\lim_{a \rightarrow 0} \frac{\|\tau(a' + a) - \tau(a') - (D_{a'}\tau)(a)\|_B}{\|a\|_A} = 0 \quad (25)$$

then  $D_{a'}\tau$  is called the *Fréchet derivative* of  $\tau$  at  $a'$ . If the Fréchet derivative exists then so does the Gâteaux derivative, and the two are equal.

The Fréchet derivative admits a chain rule in the following sense. Let  $\psi: B \rightarrow C$  be a second transformation between Banach spaces. If the Fréchet derivatives of  $\tau$  and  $\psi$  exist at  $a'$  resp.  $\tau(a')$  then so does the Fréchet derivative of  $(\tau \circ \psi)(a) = \psi(\tau(a))$  at  $a'$  and it is:

$$(D_{a'}(\tau \circ \psi))(a) = (D_{\tau(a')} \psi)((D_{a'}\tau)(a)). \quad (26)$$

Because the Gâteaux differential does not admit a chain rule in general, Bernard [25] devised a restricted version of it that does: the “chain differential.” It is defined as

$$(\delta^* \tau)(a'; a) = \lim_{n \rightarrow \infty} \frac{\tau(a' + \varepsilon_n \cdot a_n) - \tau(a')}{\varepsilon_n} \quad (27)$$

if the limit exists and is identical for any  $\varepsilon_n \rightarrow 0$  and  $a_n \rightarrow a$ . If the chain differential exists then it is the Gâteaux differential [25] (Proposition 1). If  $a \mapsto (\delta^* \tau)(a'; a)$  exists and is linear and continuous then  $(\delta^* \tau)(a'; \cdot)$  is called the chain derivative of  $\tau$  at  $a'$  [25] (Proposition 1). If the Fréchet derivative exists then it is equal to the chain derivative [25] (Proposition 1).

### 5.3. Differentiation of p.g.fl.'s

The Gâteaux and chain differentials of a p.g.fl.  $G_{\Xi}[h]$  will be notated as, respectively,

$$\frac{\partial G_{\Xi}}{\partial g}[h] = \lim_{\varepsilon \downarrow 0} \frac{G_{\Xi}[h + \varepsilon \cdot g] - G_{\Xi}[h]}{\varepsilon} \quad (28)$$

$$\frac{\partial^* G_{\Xi}}{\partial^* g}[h] = \lim_{n \rightarrow \infty} \frac{G_{\Xi}[h + \varepsilon_n \cdot g_n] - G_{\Xi}[h]}{\varepsilon_n}. \quad (29)$$

Suppose that  $G_{\Xi}[h]$  is Gâteaux differentiable. Since  $g(\mathbf{y}) = \int g(\mathbf{x}) \cdot \delta_{\mathbf{x}}(\mathbf{y}) d\mathbf{x}$  and  $g \mapsto (\partial G_{\Xi} / \partial g)[h]$  is linear and continuous it follows that, intuitively speaking,

$$\frac{\partial G_{\Xi}}{\partial g}[h] = \int g(\mathbf{x}) \frac{\partial G_{\Xi}}{\partial \delta_{\mathbf{x}}}[h] d\mathbf{x} \quad (30)$$

for all  $g$ . If it exists, the quantity

$$\frac{\delta G_{\Xi}}{\delta \mathbf{x}}[h] = \frac{\partial G_{\Xi}}{\partial \delta_{\mathbf{x}}}[h] \quad (31)$$

is Volterra's functional derivative of  $G_{\Xi}[h]$  at  $\mathbf{x}$  [26] (p. 75; p. 24, Equation (3)). Its significance is that it permits the direct derivation of density functions without resort to measures (and for this reason is preferred by the physics community [27,28]). Equation (30) shows that the functional derivative is mathematically equivalent to the Gâteaux derivative.

If  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  with  $|X| = n$  then the iterated functional derivative is

$$\frac{\delta G_{\Xi}}{\delta X}[h] = \frac{\delta^n G_{\Xi}}{\delta \mathbf{x}_1 \dots \delta \mathbf{x}_n}[h]. \quad (32)$$

The set and functional derivatives are related by

$$\frac{\delta \sigma}{\delta X}(S) = \frac{\delta \sigma^+}{\delta X}[1_S] \quad (33)$$

where  $\sigma^+$  is the p.g.fl. of  $\sigma^*(X) = (\delta \sigma / \delta X)(\emptyset)$ :

$$\sigma^+[h] = \int h^X \frac{\delta \sigma}{\delta X}(\emptyset) \delta X. \quad (34)$$

Thus if  $\sigma = \beta_{\Xi}$  then:

$$f_{\Xi}(X) = \frac{\delta G_{\Xi}}{\delta X}[0] = \frac{\delta \beta_{\Xi}}{\delta X}(\emptyset). \quad (35)$$

In MPMT the space of test functions  $h$  is assumed to have the  $L_\infty$  norm  $\|h\|_\infty = \sup_{\mathbf{x} \in \mathfrak{X}} |h(\mathbf{x})|$ —see Ref. [13] (p. 1326, footnote 2). The chain differential is therefore *superfluous* if the Fréchet derivative of  $G_\Xi[h]$  with respect to  $\|\cdot\|_\infty$  exists. If so, it is given by:

$$\lim_{g \downarrow 0} \frac{|G_\Xi[h+g] - G_\Xi[h] - (D_h G_\Xi)[g]|}{\|g\|_\infty} = 0. \quad (36)$$

#### 5.4. Equivalence of Chain Differentials and Functional Derivatives of p.g.fl.'s

Let  $G_\Xi[h]$  be the p.g.fl. of the RFS  $\Xi$ . Since  $(h + \varepsilon g)^X = \sum_{W \subseteq X} h^{X-W} \cdot \varepsilon^{|W|} g^W$  (see Ref. [3] Equation (3.6)), Equation (28) becomes

$$\frac{\partial G_\Xi}{\partial g}[h] = \int \left( \sum_{x \in X} h^{X-\{x\}} g(x) \right) f_\Xi(X) \delta X \quad (37)$$

(see Section 8.3). This is a Gâteaux *derivative* since it is linear and continuous in  $g$ . Equation (37) can be rewritten as

$$\frac{\partial G_\Xi}{\partial g}[h] = \int g(\mathbf{x}) \left( \int h^X f_\Xi(\{\mathbf{x}\} \cup X) \delta X \right) d\mathbf{x} \quad (38)$$

(see Section 8.4). From Equations (30) and (31), the quantity in the parentheses is the functional derivative:

$$\frac{\delta G_\Xi}{\delta \mathbf{x}}[h] = \int h^X f_\Xi(\{\mathbf{x}\} \cup X) \delta X. \quad (39)$$

That is: the Gâteaux and functional derivatives of a p.g.fl. always exist. In Section 8.6 it is additionally shown that the Fréchet derivative of a p.g.fl. always exists and is identical to the Gâteaux derivative:  $(D_h G_\Xi)[g] = (\partial G_\Xi / \partial g)[h]$ .

As for the chain differential of a p.g.fl., it is easily shown (see Section 8.5) that it always exists and is identical to the Gâteaux (and therefore the Fréchet) derivative:

$$\frac{\partial^* G_\Xi}{\partial^* g}[h] = \frac{\partial G_\Xi}{\partial g}[h] = (D_h G_\Xi)[g]. \quad (40)$$

Thus, by Equation (30), the density of the authors' measure  $S \mapsto (\partial G_\Xi / \partial \mathbf{1}_S)[h]$  is the functional derivative. The following two points are therefore established:

1. The general chain rule for p.g.fl.'s is a consequence of the Fréchet derivative, not the superfluous chain differential.
2. The general chain rule for chain derivatives is mathematically equivalent to the general chain rule for functional derivatives and thus produces nothing new.

Specifically, let  $G_\Xi[T[h]] = G_\Psi[h]$  for some RFS  $\Psi$ . Then the chain rule for functional derivatives is Ref. [2] (Equation (11.285)):

$$\frac{\delta}{\delta \mathbf{x}} G_\Xi[T[h]] = \frac{\partial G_\Xi}{\partial \left( \frac{\delta T}{\delta \mathbf{x}}[h] \right)} [T[h]] = \int \frac{\delta T}{\delta \mathbf{x}}[h](\mathbf{w}) \frac{\delta G_\Xi}{\delta \mathbf{w}} [T[h]] d\mathbf{w} \quad (41)$$

and the general chain rule for the functional derivative is Ref. [23]; [3] (Equation (3.91)):

$$\frac{\delta}{\delta X} G_\Xi[T[h]] = \sum_{P \perp X} \frac{\partial G_\Xi}{\partial^{W \in P} \left( \frac{\delta T}{\delta W} [h] \right)} [T[h]] \quad (42)$$

where the summation is taken over all partitions  $P$  of  $X$ .

### 5.5. Functional Derivative vs. Chain Differential

The Gâteaux differential of a p.g.fl., Equation (28), is a simple and obvious generalization of the differential quotient of elementary calculus. The chain differential of a p.g.fl. is more complicated and by no means obvious. It is also identical to the Gâteaux and Frechét derivatives and therefore equivalent to the FISST functional derivative. Consequently: replacing every functional derivative with a chain differential produces a mathematically complexified paraphrase of FISST.

### 6. FISST Product Rule Replaced by “Leibniz’ Rule”

“Leibniz’ Rule” is mathematically equivalent to the FISST product rule since the chain differential is equivalent to the functional derivative. That is, “Leibniz’ rule” [13] (Equation (14))

$$\delta^n (F \cdot G)(h; (\eta)_{i=1}^n) = \sum_{\pi \subseteq \{1, \dots, n\}} \delta^{|\pi|} F(h; (\eta)_{i \in \pi}) \delta^{n-|\pi|} G(h; (\eta)_{i \in \pi^c}) \quad (43)$$

is substituted in place of its equivalent, the FISST product rule [2] (Equation (11.271)); [3] (Equation (3.70)):

$$\frac{\delta}{\delta X} (F[h] \cdot G[h]) = \sum_{W \subseteq X} \frac{\delta F}{\delta(X-W)} [h] \frac{\delta G}{\delta W} [h]. \quad (44)$$

Note the relative conceptual simplicity of the latter compared to the former. There is no mention of the FISST general product rule [2] (Equation (11.274)); [3] (Equation (3.68)):

$$\frac{\delta}{\delta X} (F_1[h] \cdots F_n[h]) = \sum_{W_1 \cup \dots \cup W_n = X} \frac{\delta F_1}{\delta W_1} [h] \cdots \frac{\delta F_n}{\delta W_n} [h]. \quad (45)$$

More generally, there is no mention of the extensive FISST “toolbox” of general multitarget differentiation and integration rules [2] (pp. 383–389); [3] (pp. 69–80). MPMT instead replaces it with a paraphrase consisting of chain-differential rules that, for the purposes of multitarget tracking, are mathematically equivalent to their FISST counterparts [29]. For example, the formula [13] (Equation (16))

$$\delta^k G_{\Phi}(h; \eta_1, \dots, \eta_k) = \sum_{n \geq k} \frac{1}{(n-k)!} \int \prod_{i=1}^k \eta_i(x_i) \prod_{i=k+1}^n h(x_i) J_{\Phi}^{(n)}(dx_1, \dots, x_n) \quad (46)$$

is substituted in place of, and is equivalent to, the decade-old FISST “generalized Radon-Nikodým” formula [2] (Equation (11.251)); [3] (pp. 69–80, 95–97)):

$$\frac{\delta G_{\Xi}}{\delta X} [h] = \int h^W f_{\Xi}(X \cup W) \delta W. \quad (47)$$

Note the conceptual simplicity of the latter compared to the former.

**Remark 9.** Note that Equation (39) is the special case of Equation (47) with  $X = \{x\}$ .

**Remark 10.** The paper [14] does contain three acknowledgements of the FISST “toolbox”: the FISST “generalized product rule for set derivatives” [14] (p. 51), the FISST “multi-target Bayesian recursion” [14] (p. 49), and the FISST “extraction rule . . . for the evaluation of the multitarget density of a RFS” [14] (p. 51). The paper [23] (in which the general chain rule for the FISST functional derivative is derived) is acknowledged [14] (p. 50)—but is cited as a chain differential paper even though it addresses only the functional derivative. These negligible differences aside, the issues raised in this paper apply with full force to [14] (not just [12,13]).

## 7. RFS Motion Models Replaced by MPMT Motion Models

The paper [13] is devoted to a CPHD filter with target spawning, which in turn requires the predicted p.g.fl.  $G_{k|k-1}[h]$ . This formula was derived 14 years earlier in Ref. [30] (p. 1173). An alleged p.p derivation of it is substituted in its place.

The following subsections address: the FISST multitarget motion model (Section 7.1); the FISST predicted p.g.fl. (Section 7.2); the FISST and p.p. multitarget motion models are identical (Section 7.3); and the FISST and p.p. predicted p.g.fl.'s are identical (Section 7.4).

### 7.1. The “Standard” FISST Multitarget Motion Model

This section is drawn from Sections III-C and IV-D of Ref. [4]. Single-target tracking is based on an explicit motion model. It typically has the form  $\mathbf{X}_{k|k-1} = h_{k|k-1}(\mathbf{x}') + \mathbf{W}_{k|k-1}$  where  $\mathbf{X}_{k|k-1}$  is the random predicted target state,  $h_{k|k-1}(\mathbf{x}')$  is the deterministic predicted state given that the target has state  $\mathbf{x}'$  at time  $t_{k-1}$ , and  $\mathbf{W}_{k|k-1}$  is the plant noise. The statistics of this model are characterized by the probability measure  $p_{k|k-1}(S|\mathbf{x}') = \Pr(\mathbf{X}_{k|k-1} \in S | \mathbf{X}_{k-1|k-1} = \mathbf{x}')$ . The Markov density  $f_{k|k-1}(\mathbf{x}|\mathbf{x}')$  is derived from it via calculus.

A fundamental innovation of FISST was to extend this reasoning to multitarget systems. Assume that there is no target spawning—i.e., a target at time  $t_k$  survives or disappears but does not generate new targets. This scenario is described by the RFS “standard” motion model

$$\Xi_{k|k-1} = T_{k|k-1}(\mathbf{x}'_1) \cup \dots \cup T_{k|k-1}(\mathbf{x}'_n) \cup B_{k|k-1}. \quad (48)$$

Here,  $X' = \{\mathbf{x}'_1, \dots, \mathbf{x}'_n\}$  with  $|X'| = n$  is the multitarget state at time  $t_{k-1}$ ; the RFS  $T_{k|k-1}(\mathbf{x}')$  describes the evolution of a target with state  $\mathbf{x}'$ ; and the RFS  $B_{k|k-1}$  describes the newly-appearing targets. Here, either  $T_{k|k-1}(\mathbf{x}') = \emptyset$  (target vanishes with probability  $1 - p_S(\mathbf{x}')$ ) or  $T_{k|k-1}(\mathbf{x}') = \{\mathbf{X}_{k|k-1}\}$  (target survives with probability  $p_S(\mathbf{x}')$ ). Also,  $B_{k|k-1}$  is assumed to be a Poisson RFS.

If targets evolve independently then the belief measure of  $\Xi_{k|k-1}$  factorizes as:

$$\beta_{\Xi_{k|k-1}}(S|X') = \beta_{T_{k|k-1}(\mathbf{x}'_1)}(S) \cdots \beta_{T_{k|k-1}(\mathbf{x}'_n)}(S) \cdot \beta_{B_{k|k-1}}(S). \quad (49)$$

Given this, Equation (35) and the FISST multitarget calculus rules are used to derive an explicit formula for the multitarget Markov density  $f_{k|k-1}(X|X')$  [2] (Equations (13.6)–(13.8)).

If there is target spawning, then  $T_{k|k-1}(\mathbf{x}')$  is replaced by  $T_{k|k-1}(\mathbf{x}') \cup T'_{k|k-1}(\mathbf{x}')$  where the RFS  $T'_{k|k-1}(\mathbf{x}')$  models the targets spawned by a target with state  $\mathbf{x}'$  at time  $t_{k-1}$ .

### 7.2. The FISST Predicted p.g.fl.

Given the formula for  $f_{k|k-1}(X|X')$ , one can derive the formula for the p.g.fl.  $G_{k|k-1}[h]$  (of the predicted multitarget distribution  $f_{k|k-1}(X|Z_{1:k-1})$ ) in terms of the p.g.fl.  $G_{k-1|k-1}[h]$  of the prior distribution  $f_{k-1|k-1}(X|Z_{1:k-1})$ , as follows [2] (Equation (16.399)); [3] (p. 1167):

$$G_{k|k-1}[h] = e_h \cdot G_{k-1|k-1}[(1 - p_S + p_S p_h) \cdot b_h]. \quad (50)$$

Here,  $p_h(\mathbf{x}') = \int h(\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x}|\mathbf{x}') d\mathbf{x}$  describes the surviving targets;  $b_h(\mathbf{x}') = \int h^X \cdot b_{k|k-1}(X|\mathbf{x}') \delta X$  describes the spawned targets; and  $e_h[h] = \int h^X \cdot b_{k|k-1}(X) \delta X$  describes the appearing targets.

### 7.3. The FISST and MPMT Motion Models Are Identical

In Ref. [13] the RFS motion model is implicitly presumed and a “p.p.” paraphrase of it substituted in its place. Specifically,  $T_{k|k-1}(\mathbf{x}')$  is replaced by a surviving “daughter” p.p.  $\Phi_S$ ;  $T'_{k|k-1}(\mathbf{x}')$  is replaced by a “spawning point process”  $\Phi_b$ ;  $B_{k|k-1}$  is replaced by a “spontaneous birth process”  $\Phi_\gamma$ ; and  $\Xi_{k|k-1}$  is replaced by the “predicted multitarget process”  $\Phi_{k|k-1}$ .

#### 7.4. The FISST and MPMT Predicted p.g.fl.'s Are Identical

In Ref. [13] (Equation (62b)) the “Galton-Watson equation” and other formulas are used to derive the predicted p.g.fl.:

$$G_{k|k-1}[h] = G_\gamma[h] \cdot G_{k-1|k-1}[G_s[h|\cdot] \cdot G_b[h|\cdot]] \quad (51)$$

where  $G_\gamma[h]$  is identical to  $e_h$ ;  $G_s[h|\cdot]$  is identical to  $1 - p_S(\cdot) + p_S(\cdot)p_h(\cdot)$ ; and  $G_b[h|\cdot]$  is identical to  $b_h$ . That is: Equation (51) is the result of a “p.p.” derivation that is nearly identical to the FISST derivation, and is exactly the same formula that was derived using FISST 14 years earlier.

### 8. Mathematical Derivations

The theoretical results reported in this section are original. Even so, the results reported in Sections 8.3–8.6—i.e., the existence and equality of the Fréchet, Gâteaux, and chain derivatives of a p.g.fl.—should be regarded, from an intuitive point of view, as nearly obvious. A p.g.fl.  $G[h]$  is a functional analog of a power-series function  $f(x) = \sum_{n \geq 0} a_n x^n$ . (Indeed, it is an instance of what Volterra in Ref. [26] called a “functional power series.”) Since a power-series function is analytic—i.e., its Newtonian derivatives  $(d^n f / dx^n)(x)$  of arbitrary order  $n$  exist, with  $(d^n f / dx^n)(0) = n! \cdot a_n$ —it should not be surprising that p.g.fl.'s are analogously analytic.

#### 8.1. Derivation of the Density Function of the Regional Variance

We are to prove Equation (11). First extend  $\text{var}_\Xi(S)$  to a functional as follows:

$$\text{var}_\Xi^+[h] = \int \left( \sum_{x \in X} h(x) \right)^2 f_\Xi(X) \delta X - \left( \int \left( \sum_{x \in X} h(x) \right) f_\Xi(X) \delta X \right)^2. \quad (52)$$

It is easily seen that  $\text{var}_\Xi(S) = \text{var}_\Xi^+[\mathbf{1}_S]$ . Thus, from Equation (33) we get:

$$\text{var}_\Xi^*(S) = \frac{\delta \text{var}_\Xi}{\delta X}(\emptyset) = \frac{\delta \text{var}_\Xi^+}{\delta X}[0]. \quad (53)$$

By Campbell's theorem [3] (Equation (4.96)), the second term of Equation (52) can be simplified:

$$\text{var}_\Xi^+[h] = \int \left( \sum_{x \in X} h(x) \right)^2 f_\Xi(X) \delta X - \left( \int h(x) D_\Xi(x) dx \right)^2. \quad (54)$$

Taking functional derivatives  $\delta / \delta x_1$  and  $\delta / \delta x_2$  of Equation (56) with  $x_1 \neq x_2$  we get:

$$\frac{\delta \text{var}_\Xi^+}{\delta x_1}[h] = \int 2 \left( \sum_{x \in X} h(x) \right) \left( \sum_{x' \in X} \delta_{x_1}(x') \right) f_\Xi(X) \delta X - 2 \left( \int h(x) D_\Xi(x) dx \right) \left( \int \delta_{x_1}(x) D_\Xi(x) dx \right) \quad (55)$$

$$= \int 2 \left( \sum_{x \in X} h(x) \right) \left( \sum_{x' \in X} \delta_{x_1}(x') \right) f_\Xi(X) \delta X - 2 D_\Xi(x_1) \left( \int h(x) D_\Xi(x) dx \right) \quad (56)$$

and so

$$\frac{\delta^2 \text{var}_\Xi^+}{\delta x_2 \delta x_1}[h] = \int 2 \left( \sum_{x \in X} \delta_{x_2}(x) \right) \left( \sum_{x' \in X} \delta_{x_1}(x') \right) f_\Xi(X) \delta X - 2 D_\Xi(x_1) D_\Xi(x_2). \quad (57)$$

The quadratic version of Campbell's theorem is [3] (Equation (4.102)):

$$\int \left( \sum_{y \in Y} h(y) \right) \left( \sum_{y' \in Y} h'(y') \right) f_\Xi(Y) \delta Y = \int h(y) h'(y) D_\Xi(y) dy + \int h(y_1) h'(y_2) D_\Xi(\{y_1, y_2\}) dy_1 dy_2. \quad (58)$$

Given this and since  $\mathbf{x}_1 \neq \mathbf{x}_2$ ,

$$\int \left( \sum_{\mathbf{y} \in Y} \delta_{\mathbf{x}_1}(\mathbf{y}) \right) \left( \sum_{\mathbf{y}' \in Y} \delta_{\mathbf{x}_2}(\mathbf{y}') \right) f_{\Xi}(Y) \delta Y = \int \delta_{\mathbf{x}_1}(\mathbf{y}) \delta_{\mathbf{x}_2}(\mathbf{y}) D_{\Xi}(\mathbf{y}) d\mathbf{y} + \int \delta_{\mathbf{x}_1}(\mathbf{y}_1) \delta_{\mathbf{x}_2}(\mathbf{y}_2) D_{\Xi}(\{\mathbf{y}_1, \mathbf{y}_2\}) d\mathbf{y}_1 d\mathbf{y}_2 \quad (59)$$

$$= \delta_{\mathbf{x}_1}(\mathbf{x}_2) \cdot D_{\Xi}(\mathbf{x}_1) + D_{\Xi}(\{\mathbf{x}_1, \mathbf{x}_2\}). \quad (60)$$

Thus after setting  $h = 0$ , Equation (57) yields Equation (11).

### 8.2. The Set Integral of the Regional-Variance Density Is the Regional Variance

We are to prove Equation (12). From Equations (6) and (11) the set integral of  $\text{var}^+_{\Xi}(X)$  is

$$\int_S \text{var}^*_{\Xi}(X) \delta X = \int_{S \times S} D_{\Xi}(\{\mathbf{x}_1, \mathbf{x}_2\}) d\mathbf{x}_1 d\mathbf{x}_2 + \int_{S \times S} \delta_{\mathbf{x}_1}(\mathbf{x}_2) D_{\Xi}(\mathbf{x}_1) d\mathbf{x}_1 d\mathbf{x}_2 - \left( \int_S D_{\Xi}(x) dx \right)^2 \quad (61)$$

$$= \int_{S \times S} D_{\Xi}(\{\mathbf{x}_1, \mathbf{x}_2\}) d\mathbf{x}_1 d\mathbf{x}_2 + \int_S D_{\Xi}(x) dx - E[|\Xi \cap S|^2]. \quad (62)$$

Setting  $h = h' = \mathbf{1}_S$  in Equation (58) results in

$$\int_{S \times S} D_{\Xi}(\{\mathbf{y}_1, \mathbf{y}_2\}) d\mathbf{y}_1 d\mathbf{y}_2 = E[|\Xi \cap S|^2] - \int_S D_{\Xi}(\mathbf{y}) d\mathbf{y}. \quad (63)$$

Substituting this into Equation (62) we get, as claimed,  $\text{var}_{\Xi}(S)$ .

### 8.3. The Gâteaux Differential of a p.g.fl.

We are to prove Equation (37). By Equation (28) the Gâteaux differential of  $G_{\Xi}[h]$  is

$$\frac{\partial G_{\Xi}}{\partial g}[h] = \lim_{\varepsilon \downarrow 0} \frac{G_{\Xi}[h + \varepsilon \cdot g] - G_{\Xi}[h]}{\varepsilon}. \quad (64)$$

Since  $(h + \varepsilon g)^X = \sum_{W \subseteq X} h^{X-W} \cdot \varepsilon^{|W|} g^W$  [3] (Equation (3.6)),

$$G_{\Xi}[h + \varepsilon \cdot g] = \int \left( \sum_{W \subseteq X} h^{X-W} \varepsilon^{|W|} g^W \right) f_{\Xi}(X) \delta X \quad (65)$$

$$= G_{\Xi}[h] + \varepsilon \int \left( \sum_{\mathbf{x} \in X} h^{X-\{\mathbf{x}\}} g(\mathbf{x}) \right) f_{\Xi}(X) \delta X + \varepsilon^2 \int \left( \sum_{W \subseteq X, |W| \geq 2} h^{X-W} \varepsilon^{|W|-2} g^W \right) f_{\Xi}(X) \delta X. \quad (66)$$

From this Equation (37) immediately follows.

### 8.4. Formula for the Functional Derivative of a p.g.fl.

We are to prove Equation (38). From the definition of a set integral, Equation (6), we see that

$$\int \left( \sum_{\mathbf{x} \in X} h^{X-\{\mathbf{x}\}} g(\mathbf{x}) \right) f_{\Xi}(X) \delta X = \sum_{n \geq 0} \frac{1}{n!} \int h(\mathbf{x}_2) \cdots h(\mathbf{x}_n) g(\mathbf{x}_1) f_{\Xi}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n \quad (67)$$

$$+ \dots$$

$$\sum_{n \geq 0} \frac{1}{n!} \int h(\mathbf{x}_1) \cdots h(\mathbf{x}_{n-1}) g(\mathbf{x}_n) f_{\Xi}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n$$

$$= \sum_{n \geq 0} \frac{n}{n!} \int g(\mathbf{x}) h(\mathbf{x}_1) \cdots h(\mathbf{x}_{n-1}) f_{\Xi}(\{\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_{n-1}\}) d\mathbf{x} d\mathbf{x}_1 \cdots d\mathbf{x}_{n-1} \quad (68)$$

$$= \sum_{i \geq 0} \frac{1}{i!} \int g(\mathbf{x}) h(\mathbf{x}_1) \cdots h(\mathbf{x}_i) f_{\Xi}(\{\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_i\}) d\mathbf{x} d\mathbf{x}_1 \cdots d\mathbf{x}_i \quad (69)$$

$$= \int g(\mathbf{x}) \left( \int h^X f_{\Xi}(\{\mathbf{x}\} \cup X) \right) \delta X d\mathbf{x}. \tag{70}$$

8.5. The Chain Differential of a p.g.fl.

We are to prove Equation (40). From Equations (65) and (66) and the definition of the chain differential, Equation (29),

$$G_{\Xi}[h + \varepsilon_n g_n] = \int \left( \sum_{W \subseteq X} h^{X-W} \varepsilon_n^{|W|} g_n^W \right) f_{\Xi}(X) \delta X \tag{71}$$

$$= G_{\Xi}[h] + \varepsilon_n \int \left( \sum_{\mathbf{x} \in X} h^{X-\{\mathbf{x}\}} g_n(\mathbf{x}) \right) f_{\Xi}(X) \delta X + \varepsilon_n^2 \int \left( \sum_{W \subseteq X, |W| \geq 2} h^{X-W} \varepsilon_n^{|W|-2} g_n^W \right) f_{\Xi}(X) \delta X. \tag{72}$$

Equation (40) immediately follows from this.

8.6. The Frechét Derivative of a p.g.fl.

We are to show that, with respect to the  $L_{\infty}$  norm  $\|h\|_{\infty} = \sup_{\mathbf{x} \in X} |h(\mathbf{x})|$ , the Frechét derivative of a p.g.fl. exists. We know that if it exists then it must be equal to the Gâteaux derivative, which by Equation (37) is

$$l_h[g] = \int \left( \sum_{\mathbf{x} \in X} h^{X-\{\mathbf{x}\}} g(\mathbf{x}) \right) f_{\Xi}(X) \delta X. \tag{73}$$

Because of Equation (36) we are to show that

$$\lim_{g \downarrow 0} \frac{|G_{\Xi}[h + g] - G_{\Xi}[h] - l_h[g]|}{\|g\|_{\infty}} = 0. \tag{74}$$

However, the left side is easily seen to be

$$\lim_{g \downarrow 0} \frac{\left| \int \sum_{W \subseteq X, |W| \geq 2} h^{X-W} g^W f_{\Xi}(x) \delta X \right|}{\|g\|_{\infty}} = \lim_{g \downarrow 0} \left| \int \sum_{W \subseteq X, |W| \geq 2} h^{X-W} \left( \frac{g^W}{\sup_{\mathbf{x}} g(\mathbf{x})} \right) f_{\Xi}(x) \delta X \right| \tag{75}$$

$$= \left| \int \sum_{W \subseteq X, |W| \geq 2} h^{X-W} \lim_{g \downarrow 0} \left( \frac{g^W}{\sup_{\mathbf{x}} g(\mathbf{x})} \right) f_{\Xi}(x) \delta X \right|. \tag{76}$$

For fixed  $W = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  with  $|W| = n \geq 2$ , the limit in Equation (76) is

$$\lim_{g \downarrow 0} \left( \frac{g(\mathbf{x}_1) \cdots g(\mathbf{x}_n)}{\sup_{\mathbf{x}} g(\mathbf{x})} \right) \leq \lim_{g \downarrow 0} g(\mathbf{x}_1) \cdots g(\mathbf{x}_{n-1}) = 0. \tag{77}$$

9. Conclusions

The finite-set statistics (FISST) approach to multitarget tracking—stochastic geometry, random finite sets (RFS’s), belief-mass functions, and set derivatives—was introduced in the mid-1990s [1]. Its extended form—probability generating functionals (p.g.fl.’s) and Volterra functional derivatives [2–4]—dates from 2001 [5]. An allegedly more general alternative to FISST—herein called “point-process measure-theoretical multitarget tracking” or “MPMT”—has been presented in Refs. [12–14]. Herein it was demonstrated that MPMT is a mathematical paraphrase of part of FISST with the following mathematically equivalent substitutions:

Finite set  $\Rightarrow$  vector; RFS  $\Rightarrow$  simple point process (p.p.); multitarget densities  $\Rightarrow$  “measures” (actually, set functions or families of measures); multitarget probability density  $\Rightarrow$  family

of Janossy measures; multitarget factorial-moment density  $\Rightarrow$  family of factorial-moment measures; multitarget distribution  $f_{k|k}(X|Z_{1:k}) \Rightarrow$  Radon-Nikodým density  $P_{k|k}(d\phi|Z_{1:k})$ ; set integral  $\Rightarrow$  measure-theoretic integral; functional derivative  $\Rightarrow$  chain differential; functional-derivative product rule  $\Rightarrow$  chain-differential “Leibniz’ rule”; functional-derivative chain rules  $\Rightarrow$  chain-derivative chain rules; RFS multitarget motion model  $\Rightarrow$  “p.p.” multitarget motion model; FISST predicted p.g.fl.  $\Rightarrow$  “p.p.” predicted p.g.fl.; and so on.

It was further demonstrated that each of these substitutions is an unnecessary mathematical complexification of the FISST component that it replaces. In particular:

- Vector multitarget-state representation is a mathematically equivalent complexification of finite set representation that is inappropriate for practical multitarget tracking.
- A simple p.p. is a mathematically equivalent complexification of an RFS that is inappropriate for practical multitarget tracking.
- The “regional variance” of Ref. [12] *does* admit a density—thereby refuting the only evidence offered in Refs. [12–14] that MPMT is unavoidable for practical multitarget tracking.
- The measure-theoretic integral is a mathematically equivalent complexification of the FISST set integral that is inappropriate for practical multitarget tracking.
- When applied to practical multitarget tracking, the “chain differential” is a mathematically equivalent complexification of the FISST functional derivative.

Beyond this, FISST is significantly more general than MPMT because it: (a) has an integro-differential calculus of nonadditive set functions and their densities; and (b) provides a provably Bayes-optimal unification of “hard + soft” multitarget information fusion.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Goodman, I.; Mahler, R.; Nguyen, H. *Mathematics of Data Fusion*; Kluwer Academic Publishers: New York, NY, USA, 1997.
2. Mahler, R. *Statistical Multisource-Multitarget Information Fusion*; Artech House Publishers: Norwood, MA, USA, 2007.
3. Mahler, R. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House Publishers: Norwood, MA, USA, 2014.
4. Mahler, R. “Statistics 102” for multisensor-multitarget tracking. *IEEE J. Sel. Top. Sign. Proc.* **2013**, *7*, 376–389. [[CrossRef](#)]
5. Mahler, R. Multitarget moments and their application to multitarget tracking. In Proceedings of the Workshop on Estimation, Tracking, and Fusion, Monterey, CA, USA, 17 May 2001; pp. 134–166. Available online: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a414365.pdf> (accessed on 28 December 2018).
6. Mahler, R. A Brief Survey of Advances in Random-Set Fusion. In Proceedings of the International Conference on Control, Automation and Information Sciences (ICCAIS2015), Changshu, China, 29–31 October 2015.
7. Vo, B.-T.; Vo, B.-N. Labeled random finite sets and multi-object conjugate priors. *IEEE Trans. Signal Proc.* **2013**, *61*, 3460–3475. [[CrossRef](#)]
8. Vo, B.-N.; Vo, B.-T.; Hoang, H.G. An efficient implementation of the generalized labeled multi-Bernoulli filter. *IEEE Trans. Signal Proc.* **2017**, *65*, 1975–1987. [[CrossRef](#)]
9. Beard, M.; Vo, B.-T.; Vo, B.-N. A solution for large-scale multi-object tracking. *arXiv* **2018**, arXiv:1804.06622v1.
10. Mahler, R. Measurement-to-track association and finite-set statistics. *arXiv* **2017**, arXiv:1701.07078.
11. Mahler, R. “Statistics 101” for multisensor, multitarget data fusion. *IEEE Aerosp. Electron. Syst. Mag. Part 2 Tutor.* **2004**, *19*, 53–64. [[CrossRef](#)]
12. Delande, E.; Uney, M.; Houssineau, J.; Clark, D. Regional variance for multi-object filtering. *IEEE Trans. Signal Proc.* **2014**, *62*, 3415–3428. [[CrossRef](#)]

13. Bryant, D.; Delande, E.; Gehly, S.; Houssineau, J.; Clark, D.; Jones, B. The CPHD filter with target spawning. *IEEE Trans. Signal Proc.* **2017**, *65*, 1324–1338. [[CrossRef](#)]
14. Schlangen, I.; Delande, E.; Houssineau, J.; Clark, D. A second-order PHD filter with mean and variance in target number. *IEEE Trans. Signal Proc.* **2018**, *66*, 48–63. [[CrossRef](#)]
15. Moyal, J. The general theory of stochastic population processes. *Acta Math.* **1962**, *108*, 1–31. [[CrossRef](#)]
16. Reid, D. An algorithm for tracking multiple targets. *IEEE Trans. Autom. Control* **1979**, *24*, 843–854. [[CrossRef](#)]
17. Vo, B.-N.; Vo, B.-T.; Pham, N.-T.; Suter, D. Joint detection and estimation of multiple objects from image observations. *IEEE Trans. Signal Proc.* **2010**, *58*, 5129–5241. [[CrossRef](#)]
18. Vo, B.-N.; Vo, B.-T.; Clark, D. Bayesian multiple target filtering using random finite sets. In *Integrated Tracking, Classification, and Sensor Management*; Mallick, M., Krishnamurthy, V., Vo, B.-N., Eds.; Wiley: New York City, NY, USA, 2013; Chapter 3.
19. Stoyan, D.; Kendall, W.; Mecke, J. *Stochastic Geometry and Its Applications*, 2nd ed.; John Wiley & Sons: Chichester, UK, 1995.
20. Mahler, R. A theoretical foundation for the Stein-Winter “Probability Hypothesis Density (PHD)” multitarget tracking approach. In Proceedings of the 2000 MSS National Symposium Sensor & Data Fusion, Kelly AFB, San Antonio, TX, USA, 20–22 June 2000; Volume I, pp. 99–118.
21. Denneberg, D. *Non-Additive Measure and Integral*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1994.
22. Vo, B.-N.; Singh, S.; Doucet, A. Sequential Monte Carlo methods for multi-target filtering with random finite sets. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1224–1245.
23. Clark, D.; Mahler, R. Generalized PHD filters via a general chain rule. In Proceedings of the 15th International Conference on Information Fusion, Singapore, 9–12 July 2012.
24. Daley, D.; Vere-Jones, D. *An Introduction to the Theory of Point Processes, Volume 1: Elementary Theory and Methods*; Springer: New York, NY, USA, 2003.
25. Bernard, P. Chain differentials with an application to the mathematical fear operator. *Nonlinear Anal. Theory Methods Appl.* **2005**, *62*, 1225–1233. [[CrossRef](#)]
26. Volterra, V. *Theory of Functionals and of Integral and Integro-Differential Equations*; Translated by Long, M.; Blackie and Son, Ltd.: London, UK; Glasgow, UK, 1930.
27. Ryder, L. *Quantum Field Theory*, 2nd ed.; Cambridge U. Press: Cambridge, UK, 1996.
28. Engel, E.; Dreizler, R. *Density Functional Theory*; Springer: New York, NY, USA, 2011.
29. Clark, C.; Houssineau, J.; Delande, E. A few calculus rules for chain differentials. *arXiv* **2015**, arXiv:1506.08626v1.
30. Mahler, R. Multitarget filtering via first-order multitarget moments. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1152–1178. [[CrossRef](#)]

