On Point Processes in Multitarget Tracking

Ronald Mahler, Random Sets LLC, Eagan, MN, U.S.A.
RandomSets@comcast.net

***Abstract*— The finite-set statistics (FISST) approach to multitarget tracking—random finite sets (RFS's), belief-mass functions, and set derivatives—was introduced in the mid-1990s. Its current extended form—probability generating functionals (p.g.fl.'s) and functional derivatives—dates from 2001. In 2008, an "elementary" alternative to FISST was proposed, based on "finite point processes" rather than RFS's. This was accompanied by single-sensor and multisensor versions of a claimed generalization of the PHD filter, the "multitarget intensity filter" or "iFilter." Then in 2013 in the *Journal of Advances in Information Fusion* (*JAIF*) and elsewhere, the same author went on to claim that the FISST p.g.fl./functional derivative approach is actually "due to" (a "corollary" of) a 50-year-old pure-mathematics paper by Moyal; and described a "point process" p.g.fl./functional derivative approach to multitarget tracking supposedly based on it. In this paper it is shown that: (1) non-RFS point processes are a phenomenologically erroneous foundation for multitarget tracking; (2) nearly every equation, concept, discussion, derivation, and methodology in the *JAIF* paper originally appeared in FISST publications, without being so attributed; (3) FISST cannot possibly be "due to Moyal"; and (4) the "point process" approach described in *JAIF* differs from FISST only in regard to terminology and notation, and thus in this sense appears to be an obscured, phenomenologically erroneous, and improperly attributed copy of FISST. The paper concludes with the following question: Given the above, do the peer-review standards of the *Journal of Advances in Information Fusion* rise to the level expected of any credible scientific journal? It is also shown that the derivations of the single-sensor and multisensor iFilter appear to have had major errors, as did a subsequent recasting of the multisensor iFilter as a "traffic mapping filter."**

I. Introduction

The finite-set statistics (FISST) approach to multitarget tracking and information fusion—random finite sets (RFS's), belief-mass functions, and set derivatives—was introduced in the mid-1990s [8]. Its current extended form—probability generating functionals (p.g.fl.'s) and functional derivatives—dates from 2001 [23]. It was first systematically described in 2007 in the book [26]. See [28] for a tutorial introduction.

Since 2007, the approach has inspired a considerable amount of research, conducted by many dozens of researchers in at least 19 nations, reported in well over a thousand publications. As a result, progress has been rapid and has proceeded in diverse and sometimes unexpected directions, propelled by many clever new ideas.

These include algorithms for: simultaneous multitarget tracking and sensor registration; multitarget tracking and comprehensive clutter estimation in unknown, dynamic clutter; provably Bayes-optimal and tractable exact closed-form multitarget tracking; unified multitarget track-to-track fusion in ad hoc sensor networks; and unified, provably Bayes-optimal "hard + soft" information fusion.

In addition, FISST-based algorithms have been shown to significantly outperform traditional methods, including: single-target tracking in heavy clutter; track-before-detect (TBD) in pixelized images; simultaneous localization and mapping (SLAM) in robotics; and multitarget tracking using superpositional sensors.

See [15] for a short survey of these advances, or [13] for detailed coverage of most of them.

The purpose of this paper is to address the following issues—and, in the process, correct many scientific misconceptions. In 2008 in [50] and [45], a claimed alternative to FISST was proposed, which was (1) based on "finite point processes" rather than RFS's; and (2) "elementary" because (unlike the FISST p.g.fl./ functional derivative approach) it required only "familiarity with single target Bayesian filtering and with PPP's [Poisson point processes] at an elementary level." In particular, a claim was made for an "elementary" derivation of the probability hypothesis density (PHD) filter. Also proposed were single-sensor [50] and multisensor [45] versions of a claimed generalization of the PHD filter, the "multitarget intensity filter" (later renamed "iFilter").

At the time, the "point process" approach was celebrated by some because it seemed to offer an "elementary" alternative to FISST. Given this, one might have expected to see concrete evidence justifying such expectations—for example, an "elementary" derivation of the cardinalized PHD (CPHD) filter, or a CPHD generalization of the iFilter. However, no such work seems to have appeared.

In 2012 [47] and 2013 [43], [48], [2], the author of [50], [45] appeared to abandon the "elementary" approach in favor of the FISST p.g.fl./ functional derivative approach. He did so, however—and most notably in the *Journal of Advances in Information Fusion* paper [48]—by claiming that FISST is actually "due to" (a "corollary" of) a 50-year-old pure-mathematics paper [33] by Moyal. Then—allegedly directly applying Moyal's paper (substituted in place of its mere "corollary," FISST)—he outlined an alleged "point process" formulation of the p.g.fl./functional derivative approach; and used it to produce allegedly new p.g.fl./functional derivative derivations of the PHD filter and iFilter [48]. In 2013, he and two co-authors [2] employed this formulation, while citing [48] rather than FISST as its source.

This history is recounted in greater detail in Section IV.

Given that the "elementary" approach sufficed ("multitarget intensity filters can be understood in essentially elementary terms" [50, p. 1]), the need for either Moyal's paper or such rederivations is not altogether clear.

Be this as it may, when distilled to their essence the claims in [50], [45], [48], [2] appear to be as follows:

1) *2008*: Mahler's p.g.fl./functional derivative approach

to multitarget tracking is unnecessary because "single target Bayesian filtering and with PPP's at an elementary level" suffice.

2) *2012-2013*: Mahler's p.g.fl./functional derivative approach to multitarget tracking is necessary, but is actually "due to Moyal" rather than Mahler.

Strong scientific claims such as these require strong scientific scrutiny. This paper will investigate their validity in a systematic and detailed manner, hopefully clarifying a number of technical issues about "finite point processes" along the way. The following major points will be demonstrated:

1) *Section III*: A non-RFS point process is—unlike an RFS–a phenomenologically erroneous model of a multitarget system. Moreover, any point process becomes an RFS when it is applied to practical multitarget tracking. Therefore, the "point process" approach in [50], [45] differed from the FISST RFS approach only in terminology and notation and thus, in this sense, appears to have been an obscured and erroneous copy of it.
2) *Sections V, VIII*: The *Journal of Advances in Information Fusion* paper [48] was accepted, and is being cited, as original research. Yet nearly every equation, concept, discussion, derivation, and methodology in it appeared earlier in the FISST publications [3], [22], [23], [24], [25], [26], [27]—but without being so attributed. A few examples (to be discussed in more detail shortly):
    a) *Section V-A*: FISST notation and terminology was systematically changed without attribution. For example, the FISST "cardinality distribution" was changed to "cardinal number density" in [46, p. 45] and then to "canonical number distribution" in [48, p. 128].
    b) *Section V-B*: In [48], the "point process" derivation of a central FISST equation (the p.g.fl. version of Bayes' rule, [26, Eq. (G.427)]) was (except for notation) identical to the 2007 FISST derivation in [26, Eqs. (G.428-G.438)]—but without being so attributed. Yet, in [48] the original FISST equation [26, Eq. (G.427)] was claimed to be a mere special case of this identical "point process" copy.
    c) *Sections V-B through V-F*: In [48], the "point process" approach for applying p.g.fl.'s and functional derivatives to multitarget tracking was (except for terminology and notation) identical to the FISST approach, without being so attributed.
    d) *Section V-F*: In particular, the "point process" derivation of the PHD filter in [48] was (except for terminology and notation) identical to the FISST derivation described in 2004 in [25], without being so attributed. It was also a special case of the FISST derivations of the CPHD filter in [24] and the "general PHD filter" in [3]. The corresponding "point process" rederivation of the iFilter in [48] employed the same basic FISST derivation, without being so acknowledged.
3) *Sections VI, IX*: It is impossible for the FISST approach to multitarget tracking to be "due to," or a mere "corollary" of, a purely measure-theoretic paper that addressed no practical applications at all, and which appeared at the same time as the Kalman filter and nearly 20 years before Reid's seminal MHT paper, [36]. In particular:
    a) *Sections IX-C, IX-D*: The FISST p.g.fl./functional derivative approach is neither "due to Moyal" nor "has exactly the same meaning" as in [33]—because it cannot be found anywhere in [33].
    b) *Section IX-B.5, IX-C*: FISST allows one to *explicitly construct* concrete formulas for the *density functions* of RFS's (as is required for practical application), using Volterra's concept of the *functional derivative* of a p.g.fl.
    c) *Section IX-A.2*: But the paper [33] merely *defined* abstract *multivariate measures* in terms of *Gâteaux derivatives* of p.g.fl.'s, with no means of constructing their density functions. Neither functional derivatives nor the term "functional derivative" appear anywhere in [33].
    d) *Section VI-D*: Reverse-engineering is fundamentally different than engineering. It is easy to know the right things to do—and even easier to claim that these things are actually obvious—if someone else has previously shown you how to do it all in complete detail.
4) *Sections X, XI*: The derivations of both the single- and multisensor iFilters in [50], [45] had major mathematical errors; and the key concept underlying the iFilter appears to be phenomenologically questionable. Moreover, the PHD filter is not a special case of the iFilter; and the multisensor iFilter appears to have demonstrably poorer performance than the RFS alternatives.
5) *Sections XI-E, XI-F, XI-G*: A subsequent recasting [46] of the multisensor iFilter, as a "multisensor traffic mapping filter," also appears to have had major mathematical and conceptual errors.

The paper begins with a short refresher on proper scientific discourse (Section II) and ends with a summary and conclusions (Section VII) in which the following question is posed:

- Do the peer-review standards of the *Journal of Advances in Information Fusion* rise to the level expected of any credible scientific journal?

To achieve a more streamlined exposition, four systematic supporting analyses have been placed as Appendices: a comparison of the contents of the *Journal of Advances in Information Fusion* paper [48] with earlier FISST publications (Section VIII); a comparison of FISST with Moyal's paper [33] (Section IX); the iFilter (Section X); and the multisensor iFilter (Section XI).

II. Preamble: On scientific discourse

It is necessary to begin with the following reminders:

1) Scientific critique is essential to scientific discourse. Technical error can seriously undermine a scientific discipline if it is propagated unnoticed and unremarked.[1]

[1]For an instructive cautionary tale, see K. Kelley and C. Moody, "The booms and busts of molecular electronics," *IEEE Spectrum*, Oct. 2015.

2) A scientifically valid but critical argument is still a scientifically valid argument. In particular, any factually true statement is a scientifically valid statement.
3) To accuse the author of a scientific critique of being "unprofessional," simply because s/he has written a critique, is itself unprofessional. The proper scientific response to open scientific critique is more open scientific critique—not *ad hominem* attacks.

III. RFS'S VS. "POINT PROCESSES"

In this section, the RFS and "point process" formulations of multitarget tracking theory are summarized (Section III-A). The latter is shown to be, for purposes of practical multitarget tracking, mathematically identical to the RFS formulation (Section III-B); (2) phenomenologically erroneous in general (Section III-C); and (3) not theoretically general enough to address "hard + soft" fusion (Section III-D).

A. The "point process" formulation

As previously noted, in 2008 in [50] and [45], an allegedly new approach to multitarget tracking was proposed, based on "finite point processes." In particular, in [50], Kingman's well-known book *Poisson Processes* [11] was cited as an authoritative text on point processes.[2]

However, Kingman employed an *RFS formulation* of point process theory.[3] That is, let $\mathfrak{X}$ be a target state space. Then a point process in $\mathfrak{X}$ is a random variable whose realizations are finite sets—that is, unordered lists $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ such that $\mathbf{x}_1, ..., \mathbf{x}_n \in \mathfrak{X}$ and $n \geq 0$, and $\{\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_j, ..., \mathbf{x}_n\} = \{\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_n\}$ if $\mathbf{x}_i = \mathbf{x}_j$.

Despite Kingman's RFS formulation, in [50] the realizations of a "point process" were defined in non-RFS terms—that is, as vectors $(n, \mathbf{x}_1, ..., \mathbf{x}_n)$ or, alternatively, $(n, \{\mathbf{x}_1, ..., \mathbf{x}_n\})$ for any $n \geq 0$—where (at variance with standard modern notation) $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ here denotes a finite unordered list rather than a finite set.[4]

B. RFS's vs. "point processes" in practical tracking, 1

The $\mathbf{x}_1, ..., \mathbf{x}_n$ in $(n, \mathbf{x}_1, ..., \mathbf{x}_n)$ or $(n, \{\mathbf{x}_1, ..., \mathbf{x}_n\})$ need not be distinct—there can be many copies of any $\mathbf{x}_i$. The number n_i of copies of $\mathbf{x}_i$ is the *multiplicity* of $\mathbf{x}_i$ in $(n, \{\mathbf{x}_1, ..., \mathbf{x}_n\})$. If $n_i = 1$ for any $i = 1, ..., n$, for every realization, then the point process is *simple*. Since an finite unordered list with distinct elements is the same thing as a finite set, *a simple point process is an RFS*.

To be of practical use, the probability density function (p.d.f.) $f(n, \{\mathbf{x}_1, ..., \mathbf{x}_n\})$ of a point process must exist—i.e., be finite-valued. However, a basic result of point process theory is the following. Let $\mathfrak{X} = L \times S$ ($S \subseteq \mathbb{R}^N$, L finite). Then the p.d.f. exists only if the point process is simple—i.e., only if it is an RFS ([4], p. 134, Prop. 5.4.IV). That is:

- *When a point process is applied to practical multitarget tracking, it becomes an RFS.*
- *Consequently, any claimed "point process" alternative to the FISST RFS formulation differs from it only in terminology and notation, and so is in this sense an obscured copy of it.*

C. RFS's vs. "point processes" in practical tracking, 2

In practical multitarget tracking, a target-state must have the form $\mathbf{x} = (\ell, \mathbf{u})$, where $\mathbf{u}$ is the kinematic state and ℓ is a unique track-label—for example, ℓ = 'Bob'. As a consequence, any realization $(n, \mathbf{x}_1, ..., \mathbf{x}_n)$ of a point process that is not a finite set cannot correctly model the state of a multitarget system. This is because

$$(n, \mathbf{x}_1, ..., (\text{Bob}, \mathbf{u}), ..., (\text{Bob}, \mathbf{u}), ..., \mathbf{x}_n)$$

indicates that there are two or more copies of 'Bob' in the scene—a physical impossibility. That is:

- *A non-RFS point process is a phenomenologically invalid model of a random multitarget state.*
- *Consequently, RFS theory (and not general point process theory) is the theoretically correct foundation for practical multitarget tracking.*
- *Therefore, any "point process" alternative to the RFS formulation is, in this sense, an erroneous copy of it.*

Thus when it is stated that "Mahler...refers to finite point processes as random finite sets..." [43, p. 2], it appears that the reverse is true. In [43] and elsewhere, it is RFS's that are being misleadingly referred to as "finite point processes."

D. RFS's vs. "point processes" in information fusion

In addition, FISST is central for the provably Bayes-optimal unification of "hard + soft" information fusion—see Chapter 22 of [13]. Since this unification requires general *random closed sets* and not just random finite sets, "point processes" are not theoretically general enough to address such issues.

IV. FISST IS "DUE TO MOYAL"

The history of "point process" claims about FISST was sketched in the Introduction. The purpose of this section is to recount those claims in greater detail.

In 2012, in a tutorial [47, p. 87] at the International Conference on Information Fusion in Singapore, it was claimed that a central equation of the FISST p.g.fl./functional derivative approach—the p.g.fl. version of Bayes' rule (see Eq. (4) or Eq. (64))—is a mere "corollary" of a 50-year-old pure-mathematics paper [33] by Moyal.[5]

In 2013, this claim was expanded in a paper at the Workshop on Sensor Data Fusion in Bonn, Germany [43, p. 2]:

- "Mahler's tracking contributions are [limited to] the use of the PGFL to derive the Bayes posterior process and the repeated use of the summary statistic called the intensity (or, equivalently, the PHD) to approximate the Bayes

[2] To wit: "For further background on PPP's from a multidimensional perspective, see Kingman [6]"—[50], first paragraph of Section 2.

[3] RFS theory is a widely accepted formulation of point process theory. Besides Kingman, see [1], [37], [42].

[4] [50], discussion following Eq. (1).

[5] A copy of [47, p. 87] can be provided upon request.

posterior point process and thereby close the Bayesian recursion."

Given that (1) Mahler's only "tracking contributions" are the PHD filter and the application of the p.g.fl. Bayes' rule to it, but (2) the p.g.fl. Bayes' rule itself is merely a "corollary" of Moyal's paper [33], it follows that the PHD filter is Mahler's only significant "tracking contribution." In particular, the FISST p.g.fl./multitarget calculus, and the methodologies for applying it to multitarget tracking, are not.

This thesis was elaborated at length in 2013 in a paper [48] published in the *Journal of Advances in Information Fusion* (*JAIF*):[6] There it was claimed that:

- "The finite-set statistics (FISST) concerns functional differentiation of PGFLs, where functional differentiation has exactly the same meaning as in the calculus of variations" [48, p. 121]—where in this respect the paper [33] by Moyal was cited.
- The functional calculus used in FISST is "due to Moyal" [48, p. 121] (and thus, by implication, not to Mahler).

As previously noted, what followed was a p.g.fl./functional calculus approach to multitarget tracking, allegedly based on Moyal's paper [33] (substituted in place of its mere "corollary," FISST); and it was employed to produce allegedly new "point process" p.g.fl./functional derivative derivations of the PHD filter and iFilter. Furthermore, the p.g.fl. Bayes' rule was now claimed to be a significant new result, original with the same author who had previously characterized it [47, p. 87] as a mere "corollary" of Moyal's paper (see Section V-B).

A sequel [2] presented at the 2013 International Conference on Information Fusion in Istanbul, Turkey, cited the yet-to-be-published [48], rather than FISST, as the source of the p.g.fl./functional derivative approach employed in it. In particular, its authors explicitly cited (see [2, Eq. (28)]) the "point process" derivation of the PHD filter in [48, Eqs. (42-52)] rather than any FISST derivation—despite the fact that (see Section V-F) the "point process" derivation was mathematically identical to the FISST derivation described in [25] as well as a special case of the FISST derivations of the CPHD filter in [24] and the "general PHD filter" in [3].

V. A CONTENT ANALYSIS OF [48]

The *Journal of Advances in Information Fusion* paper [48] was accepted, and is being cited, as original research. The purpose of this section is to demonstrate that nearly every equation, concept, discussion, derivation, and methodology in it appeared earlier (in all cases but one, at least five years earlier) in the FISST publications [3], [22], [23], [24], [25], [26], [27]—but without being so attributed. As per the claims made in [47, p. 87] and [43, p. 2], its only explicit reference to FISST was in regard to the PHD filter [48, p. 120] (this being, after all, Mahler's only actual "tracking contribution"). [48, Eqs. (1-17,18-22)] were claimed to be "due to Moyal,"[7] even though [48, Eqs. (6-10,12-17,22)] cannot be found in [33]—but did appear in earlier FISST publications.

An equation-by-equation demonstration of these claims can be found in Section VIII. What follows is an examination of some of the more obvious instances:

1) "Point process" notation/terminology (Section V-A).
2) "Point process" derivation of the p.g.fl. Bayes' rule (Section V-B). *This is of special significance*.
3) "Point process" z-transforms (Section V-C).
4) "Point process" state estimators (Section V-D).
5) "Branching processes" (Section V-E).
6) "Point process" derivation methodology (Section V-F).
7) Other issues (Section V-G).

A. *"Point process" notation and terminology*

In [48], well-known FISST terminology was systematically changed, usually without attribution. A few examples:

1) "random finite set"—changed to "finite point process,"[8] even though (as was shown in Sections III-B and III-C), a "finite point process" is actually an RFS when applied to practical multitarget tracking.
2) set-theoretic union of RFS's—changed to "superposition" of "point processes."
3) "probability hypothesis density (PHD)"—changed to "intensity function"[9] (even though the terminology "PHD" is a historical[10] and widely accepted usage in multitarget tracking).
4) "cardinality distribution"—changed to "cardinal number density"[11] in [46, p. 45] and subsequently to "canonical number distribution" in [48, p. 128].

Also in [48], well-known FISST notation was systematically changed, without attribution. A few typical examples:

1) The FISST $\Xi_{k+1|k}$ (random target state-set) and Σ_{k+1} (random measurement-set)—changed to Ξ, Υ.
2) The FISST $(n!)^{-1} \cdot f_\Xi(\{x_1, ..., x_n\})$—changed to
$$p^{\Xi}_{NX}(n, x_1, ..., x_n) = p^{\Xi}_{N}(n)\, p^{\Xi}_{X|N}(x_1, ..., x_n|n). \quad (1)$$
3) The FISST multitarget likelihood function $f_k(Z|X)$ and multitarget posterior distribution $f_{k|k}(X|Z^{(k)})$ at time t_k—changed to
$$p^{\Upsilon|\Xi}(m, y_1, ..., y_m|n, x_1, ..., x_n)$$
$$p^{\Xi|\Upsilon}(n, x_1, ..., x_n|m, y_1, ..., y_m).$$
4) The FISST notation $f_{k+1}(Z|X) \cdot f_{k+1|k}(X) = f_{k+1}(Z, X)$, for the joint probability distribution of Σ_{k+1} and $\Xi_{k+1|k}$—changed to
$$\begin{aligned} & p^{\Upsilon\Xi}_{MN}(m, n)\, p^{\Xi}_{YX|MN}(y_1, ..., y_m, s_1, ..., s_n|m, n) \\ = \; & p^{\Xi}_{MYNX}(m, y_1, ..., y_m, n, s_1, ..., s_n). \quad (2) \end{aligned}$$

[6]Statements of fact: *JAIF* is the house journal of the International Society of Information Fusion (ISIF)—the organization of which the author of [48] was President when it was submitted for publication.

[7]To wit: "The results presented in this section [Section 3] are due to Moyal" [48, p. 120]

[8]To wit: "Mahler...refers to finite point processes as random finite sets..." [43, p. 2]

[9]To wit: "In tracking applications, [the intensity function] is *sometimes* called the probability hypothesis density (PHD)" (emphasis added) [49, p. 5].

[10]The terminology "probability hypothesis density" was coined by Stein and Winter, not Mahler—see [22].

[11]This terminology is inappropriate because, in mathematics, the term "density" is reserved for continuously infinite spaces.

5) The FISST p.g.fl.'s $F_{k+1}[g,h]$, $G_{k+1|k+1}[h]$, and $G_{k+1}[g|X]$—changed to $G^{\Upsilon\Xi}[g,h]$, $G^{\Xi|\Upsilon}[h]$, and $G^{\Upsilon|\Xi}[g|s_1,...,s_n]$.
6) The FISST notation $\int f_\Xi(X)\delta X$ for a set integral of $f(X)$—changed to

$$\sum_{n=0}^{\infty}\int_{S^n} p^{\Xi}_{NX}(n,s_1,...,s_n)ds_1\cdots ds_n. \tag{3}$$

7) The FISST notation h^X for the power functional—changed to $\prod_{j=1}^n h(s_j)$.

B. "Point process" derivation of p.g.fl. Bayes' rule

Special attention should be paid to this section. The FISST p.g.fl. version of Bayes' rule was derived in Eqs. (G.428-G.438) on p. 757 of [26] and is [26, Eq. (G.427)]:

$$G_{k|k}[h]=\frac{\frac{\delta F}{\delta Z_k}[0,h]}{\frac{\delta F}{\delta Z_k}[0,1]} \tag{4}$$

where [26, Eqs. (G.428,G.429)]

$$F[g,h]=\int g^Z\cdot h^X\cdot f_k(Z|X)\cdot f_{k|k-1}(X)\delta X\delta Y \tag{5}$$

is the joint p.g.fl. of the joint distribution $f_{k|k-1}(Z,X)=f_k(Z|X)\cdot f_{k|k-1}(X)$.

An alleged "point process" derivation of Eq. (4) was presented in [48, Eqs. (23-27)], leading to a "point process" version of Eq. (4) [48, Eq. (28)]:

$$G^{\Upsilon|\Xi}[h|y_1,...,y_m]=\frac{\frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1\cdots\partial y_m}[0,h]}{\frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1\cdots\partial y_m}[0,1]} \tag{6}$$

where [48, Eq. (23)]:

$$\begin{aligned} & G^{\Upsilon\Xi}[g,h] \qquad (7)\\ = & \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}p^{\Upsilon\Xi}_{MN}(m,n)\\ & \times\int_{Y^m}\int_{S^n}\left(\prod_{i=1}^m g(y_i)\right)\left(\prod_{j=1}^n h(s_j)\right)\\ & \times p^{\Upsilon\Xi}_{YX|MN}(y_1,...,y_m,s_1,...,s_n|m,n)\\ & \times dy_1\cdots dy_m ds_1\cdots ds_n. \end{aligned}$$

It was then asserted that:

- Eq. (6) is "valid for general point processes...[whereas a]...specialized version...for multitarget tracking applications is derived in [10,p.757], where it is described as the PGFL 'form of the multitarget corrector'" [48, p. 124] and is "specific to the tracking application" [48, p. 2].

Here, "[10,p.757]" refers to p. 757 of the 2007 FISST textbook [26], where Eq. (4) was derived.

The following should be pointed out:

1) The "point process" derivation [48, Eqs. (23-27)] was (except for notation and terminology) identical to the FISST derivation [26, Eqs. (G.428-G.438)] in "[10,p.757]"—but not so attributed. For example, Eq. (7) is identical to Eq. (5)—but not so attributed.
2) Given this, it is unclear how Eq. (4) could be asserted to be a mere "specialized version" of Eq. (6). Eq. (4) was derived in "[10,p.757]" as a general theorem of probability, and the allegedly fully general "point process" derivation in [48, Eqs. (23-27)] is identical to it. A general theorem of probability does not become less general because it is subsequently applied to multitarget tracking.[12]
3) It is equally unclear how Eq. (6) could be claimed as a significant original contribution in [48] given that, earlier in [47, p. 87], this same equation had been characterized by the same author as a mere "corollary" of Moyal's theoretically general paper [33].

C. "Point process" p.g.fl.'s and z-transforms

In [48, p. 123] the following was stated:

- "In the signal processing literature, $F^{\Xi}(z^{-1})$ is called the z-transform of the sequence of probabilities $p^{\Xi}_N(n)$: $n=0,1,...$"

Notation aside, this is an abbreviated version of the FISST discussions in [26, pp. 343-344, 372]—but not so attributed.

D. "Point process" multitarget state estimators

In [48, Eq. (32)], a Bayes-optimal multitarget state estimator was defined to minimize the posterior multitarget Bayes risk:

$$\hat{\xi}_{\text{Bayes}}=\arg\min_{\zeta\in\mathcal{E}(S)} R(\zeta). \tag{8}$$

Notation and terminology aside, this discussion is identical to the FISST discussions in [8, pp. 189-190], [22, p. 1159, Sec. II-B.7], and [26, p. 63]—but not so attributed. The rigorous definition of Bayes-optimal multitarget state estimators was original with FISST—see [13, pp. 110-111].

In particular, in [48, Eq. (32)] the following is claimed:

- "The MAP estimate is undefined for the posterior pdf $p^{\Xi|\Upsilon}(\xi|\upsilon)$. To see this, it is only necessary to observe that $p^{\Xi|\Upsilon}(\xi_1|\upsilon)$ and $p^{\Xi|\Upsilon}(\xi_2|\upsilon)$ have different units when the realizations ξ_1 and ξ_2 have different numbers of points."

Notation and terminology aside, this is identical to the FISST discussions in [8, pp. 189-190], [26, pp. 494-495], [27, p. 59], and [22, p. 1159, Sec. II-B.7]—but not so attributed.

The failure of the classical state estimators in the multitarget case was an early FISST insight [21, pp. 299-300]. It was by no means obvious at the time—consider, for example, the following doubly erroneous statement by Stone et al. in 1999 [41, pp. 162-163]: "The [multitarget] posterior distribution...represents our knowledge of the number of targets present and their state...From this...we can compute point estimates...*such as maximum a posteriori probability estimates or means*" (emphasis added).

[12]Does Eq. (6) itself become "specific to the tracking application" because it is subsquently so applied?

E. "Branching processes"

In [48, Eqs. (37,38)], Eq. (7) is rewritten in "branching process form," which in turn is described as "central to multitarget tracking applications":

$$G^{\Upsilon\Xi}[g,h] = G^{\Xi}[hT[g|\cdot]] \tag{9}$$

where $T[g|s] = G^{\Upsilon}[g|s]$ is the p.g.fl. of the measurement process for a target with state s. Except for notation, Eq. (9) is identical to the no-clutter special case of the FISST equation [3, Eq. (68,71)],

$$F_{k+1}[g,h] = \kappa_{k+1}[g] \cdot G_{k+1|k}[h \cdot T_{k+1}[g]], \tag{10}$$

but is not so attributed.[13] Moreover, the equations [48, Eqs. (34,35)] used to derive it are identical to to the FISST [24, Eqs. (45,44)]—or, alternatively, to the fifth and sixth equations from top in the second column of the FISST [22, p. 1173].

F. "Point process" derivation methodology

An allegedly original "point process," p.g.fl./functional calculus derivation of the measurement-update for the PHD filter was described in [48, Eqs. (42-52)]. It was then modified (i.e., PHD filter state-transition model replaced by the intermixing model of Section X-C) to produce a "point process," p.g.fl./functional calculus derivation of the iFilter.

However, it is demonstrated in this section that the "point process" PHD filter derivation was, without being so attributed: (1) identical to the one described in 2004 in [25]; (2) the Poisson (PHD filter) special case of the 2007 derivation of the CPHD filter in [24]; and (3) a special case of the 2012 derivation of the "general PHD filter" in [3].

The "point process" derivation in [48, Eqs. (42-52)] begins with the formula for the p.g.fl. for the measurement process of a system of targets with states $s_1, ..., s_n$ [48, Eq. (42)]:

$$G^{\Upsilon|\Xi}[g|s_1, ..., s_n] = G^{\Psi_{\text{clutter}}}[g] \prod_{i=1}^{n} G^{\Upsilon_{\text{target}}}[g|s_i] \tag{11}$$

where the "PGFL of target-originated measurements" [48, Eq. (43)] for a target with state s is:

$$G^{\Upsilon_{\text{target}}}[g|s] = 1 - p^D(s) + p^D(s) \int_Y g(y)p(y|s)dy. \tag{12}$$

Eqs. (11,12) were asserted without proof. However, they are due to the FISST RFS measurement modeling methodology (see Section IX-B.3), without being so attributed. Notation aside, Eq. (11) is identical to the second equation from the bottom in the second column of p. 1173 in the FISST [22]:

$$G_{k+1}[g|X] = G[g|\mathbf{x}_1] \cdots G[g|\mathbf{x}_n] \cdot G_{\Theta}[g] \tag{13}$$

where $G_{\Theta}[g]$ is the p.g.fl. of the clutter process; and also to the FISST [3, Eq. (68)]:

$$G_{k+1}[g|X] = \kappa_{k+1}[g] \cdot T_{k+1}[g]^X. \tag{14}$$

Likewise, Eq. (12) is identical to the third equation from bottom in the second column of p. 1173 in the FISST [22]:

$$G[g|\mathbf{x}] = 1 - p_D(\mathbf{x}) + p_D(\mathbf{x})\, p_g(\mathbf{x}) \tag{15}$$

where $p_g(\mathbf{x}) = \int g(\mathbf{z}) \cdot f_{k+1}(\mathbf{z}|\mathbf{x})d\mathbf{z}$.

From this point on in [48], the FISST derivation methodology (see Section IX-B.6)—and, in particular, the specific FISST derivation decribed in 2004 in [25, Eqs. (20,21)]—is employed without attribution. To wit:

In [48, Eq. (46)], Eqs. (11,12) are used to derive a formula for the joint p.g.fl.:

$$\begin{aligned} & G^{\Upsilon\Xi}[g,h] \\ = \;& \exp\left[\begin{array}{c} -\int_Y \lambda(y)dy + \int_Y g(y)\lambda(y)dy \\ -\int_S f^{\Xi}(s)ds + \int_S h(s)f^{\Xi}(s)ds \\ -\int_S h(s)p^D(s)f^{\Xi}(s)ds \\ +\int_S \int_Y g(y)h(s)p(y|s)P^D(s)f^{\Xi}(s)dyds \end{array}\right]. \end{aligned} \tag{16}$$

Notation aside, this is identical to the seventh equation from the top in the first column of p. 1174 in the FISST [22]:

$$F[g,h] = \exp\left(\lambda c[g] - \lambda + \mu s[hq_D] + \mu s[hp_D p_g] - \mu\right) \tag{17}$$

where $q_D(\mathbf{x}) = 1 - p_D(\mathbf{x})$ and where

$$s[hq_D] = \int h(\mathbf{x})(1 - p_D(\mathbf{x}))s(\mathbf{x})d\mathbf{x} \tag{18}$$

$$s[hq_D] = \int h(\mathbf{x})p_D(\mathbf{x})p_g(\mathbf{x})s(\mathbf{x})d\mathbf{x} \tag{19}$$

$$p_g(\mathbf{x}) = \int g(\mathbf{z})f_{k+1}(\mathbf{z}|\mathbf{x})d\mathbf{z}. \tag{20}$$

Eq. (6)—the "point process" copy of the FISST p.g.fl. Bayes' rule, Eq. (4)—is then applied. First, the functional derivative with respect to the measurements [48, Eq. (48)]:

$$\begin{aligned} & \frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1 \cdots \partial y_m}[g,h] \\ = \;& G^{\Upsilon\Xi}[g,h] \prod_{i=1}^{m} \left(\begin{array}{c} \lambda(y_i) \\ +\int_S h(s)p(y_i|s)P^D(s)f^{\Xi}(s)ds \end{array}\right). \end{aligned} \tag{21}$$

Notation aside, this is identical to to the equation between Eqs. (20,21) of the FISST [25]—but not so attributed:

$$\frac{\delta F_{k+1}}{\delta Z}[g,h] = F_{k+1}[g,h] \prod_{\mathbf{z}\in Z} \left(\lambda c(\mathbf{z}) + D_{k+1|k}[hp_D L_{\mathbf{z}}]\right). \tag{22}$$

It is also the Poisson (PHD filter) special case of the FISST [24, Eq. (124)]. Likewise, [48, Eq. (50)]

$$\begin{aligned} & \frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1 \cdots \partial y_m}[0,1] \\ = \;& G^{\Upsilon\Xi}[0,1] \prod_{i=1}^{m} \left(\begin{array}{c} \lambda(y_i) \\ +\int_S h(s)p(y_i|s)P^D(s)f^{\Xi}(s)ds \end{array}\right) \end{aligned} \tag{23}$$

is the FISST Eq. (22) with $g = 0$ and $h = 1$.

The following is then used to derive the PHD filter measurement-update equation—see Eq. (76)—from Eq. (21) [48, Eq. (29)]:

$$f^{\Upsilon|\Xi}(x) = \frac{\frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1 \cdots \partial y_m \partial x}[0,1]}{\frac{\partial^m G^{\Upsilon\Xi}}{\partial y_1 \cdots \partial y_m}[0,1]}. \tag{24}$$

Notation aside, Eq. (24) is identical to the FISST [24, Eq. (52)] or [26, Eq. (16.412)]—but not so attributed:

$$D_{k+1|k+1}(\mathbf{x}) = \frac{\frac{\delta F}{\delta Z_{k+1}\delta \mathbf{x}}[0,1]}{\frac{\delta F}{\delta Z_{k+1}}[0,1]}. \tag{25}$$

[13] Statement of fact: The author of [48] was aware of the paper [3] and its contents, since he sat in the first row during its presentation.

The numerator of Eq. (24) is computed in [48, Eq. (49)]. As was noted in 2004 in the FISST [25, Eq. (22)], this follows easily from Eq. (25) and Eq. (33) below—without being so attributed. It is also the Poisson (PHD filter) special case of the FISST [24, Eq. (139)]—but not so attributed.

G. Other issues

This section addresses erroneous claims that were made in [48] regarding (1) the origin of p.g.fl.'s (Section V-G.1); (2) formulas for the posterior p.g.f. (Section V-G.2); (3) a "point process" formula for the p.g.f. corresponding to the PHD filter (Section V-G.3); and (4) a statement about "finite point process" state models (Section V-G.4).

1) The origin of p.g.fl.'s : In [48, p. 119] it is stated that:

- "PGFLs for finite point processes were introduced in 1962 by Moyal."

False. In the introduction of his paper, Moyal clearly stated that "...generating functionals were introduced in this connection [point process theory] by Kendall [9] and Bartlett and Kendall [3]..." [33, p. 2]—citing papers published by those authors in the late 1940s and early 1950s. (See also [33], footnote 1 on p. 13.) Daley and Vere-Jones attribute the first use of the p.g.fl. to Bogoliubov in 1946 [5, p. 15].[14] In addition, the p.g.fl. was a point process application of Volterra's "functional power series" concept, which dates from the late 1920's [51, p. 21].

2) The "canonical number distribution" : Following [48, Eq. (21)] it is stated that:

- "The probability $p_N^{\Xi}(n)$ ["canonical number distribution"] is $n!$ times the integral of the ordered pdf $p^{\Xi}(n, x_1, ..., x_n)$ over all $x_1, ..., x_n$." That is:

$$p_N^{\Xi}(n) = n! \int p^{\Xi}(n, x_1, ..., x_n) dx_1 \cdots dx_n. \tag{26}$$

False. For, $p_N^{\Xi}(n)$ is the marginal distribution of $p^{\Xi}(n, x_1, ..., x_n)$ after integrating out $x_1, ..., x_n$:

$$p_N^{\Xi}(n) = \int p^{\Xi}(n, x_1, ..., x_n) dx_1 \cdots dx_n \tag{27}$$

$$= \frac{1}{n!} \int p^{\Xi}(n, \{x_1, ..., x_n\}) dx_1 \cdots dx_n \tag{28}$$

where Eq. (28) follows from [48, Eq. (15)]. Eq. (26) is an erroneous version of the FISST integral formula for the cardinality distribution, [24, Eq. (6)] or [26, Eq. (11.115)]:

$$p_{\Xi}(n) = \frac{1}{n!} \int f_{\Xi}(\{\mathbf{x}_1, ..., \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n. \tag{29}$$

A related error is the following formula for the "posterior pdf of the canonical number"[15] [48, Eq. (31)]:

$$p_N^{\Xi|\Upsilon}(n) = \frac{1}{n!} \frac{d^n}{dx^n} F^{\Xi|\Upsilon}[0]. \tag{30}$$

[14]Statement of fact: The author of [48] was aware of Daley & Vere-Jones' attribution, since I quoted it to him during his tutorial [47] at the 2012 International Conference on Information Fusion.

[15]It is inappropriate to call $p_N^{\Xi|\Upsilon}(n)$ a "pdf" since, in standard mathematical usage, this terminology is reserved for continuous spaces.

Since $F^{\Xi|\Upsilon}[0]$ is a constant, its n'th derivative is 0 and so $p_N^{\Xi|\Upsilon}(n) = 0$ for $n > 0$.[16] Eq. (30) is an erroneous version of the FISST [24, Eq. (168)]:

$$p_{k+1|k+1}(n) = \frac{1}{n!} G_{k+1|k+1}^{(n)}(0). \tag{31}$$

3) A "point process" p.g.f. formula : In [48, Eq. (54)], it is stated that the following formula (for the posterior probability generating function (p.g.f.)), "appears to be new to the PHD literature":

$$\begin{aligned} & F_{\text{PHD}}^{\Xi|\Upsilon}(x) \\ = & \exp\left[(x-1)\int_S (1 - P^D(s) f^{\Xi}(s) ds\right] \\ & \times \prod_{i=1}^m \frac{\lambda(y_i) + x\int_S p(y_i|s) P^D(s) f^{\Xi}(s) ds}{\lambda(y_i) + \int_S p(y_i|s) P^D(s) f^{\Xi}(s) ds}. \end{aligned} \tag{32}$$

False. In 2004 in [25, Eq. (21)], the posterior p.g.fl. for the PHD filter was identified as:

$$\begin{aligned} G_{k+1|k+1}[h] = & e^{D_{k+1|k}[(h-1)(1-p_D)]} \\ & \cdot \prod_{\mathbf{z} \in Z} \frac{\lambda c(\mathbf{z}) + D_{k+1|k}[h p_D L_{\mathbf{z}}]}{\lambda c(\mathbf{z}) + D_{k+1|k}[p_D L_{\mathbf{z}}]}. \end{aligned} \tag{33}$$

Eq. (32) trivially follows from the substitution $h = x$. Eq. (32) is also the Poisson (PHD filter) special case of the well-known FISST formula for the posterior p.g.f. of the CPHD filter [24, Eq. (61)], [26, Eq. (16.326)].

4) "Finite point process" state models : In [48, p. 130] the following is stated:

- "Finite point process models...are only approximate models for multitarget state. Accepting the point process model for the multitarget state as a given, the PHD filter and iFilter are good applications..."

False, in multiple respects. First, an RFS $\Xi \subseteq \mathfrak{X}$ is an *exact* model of the random multitarget state. Second, the statement appears to repeat a common misconception: that the Poisson (PHD filter) *approximation* of Ξ is "the" unique RFS "model" of the random multitarget state. Not only is Ξ the actual RFS *model*, there is a progression of successively more accurate *approximations* of Ξ: the i.i.d.c. approximation (for CPHD filters), the multi-Bernoulli approximation (for multi-Bernoulli filters), and the generalized labeled multi-Bernoulli (GLMB) approximation (for GLMB filters). See [28] for a tutorial summary or [13] for deep detail.

Third, and as was noted in Section III-C, a non-RFS "finite point process" is a phenomenologically erroneous model of the random multitarget state. Thus the very phrase "finite point process models for multitarget state" is misleading at best and an oxymoron at the worst.

VI. Reverse-Engineering vs. Engineering

In Section V (and as documented in equation-by-equation detail in Section VIII), it was shown that the selection of equations, concepts, discussions, derivations, and methodologies chosen in [48] is a subset of FISST's—while not being so attributed. In Section IX, it is demonstrated that:

[16]The same error occurs in [48, Eq. (36)].

1) *Sections IX-C*, *IX-D*: The alleged "point process" p.g.fl./functional calculus approach in [48] is neither "due to Moyal" nor "has exactly the same meaning" as in [33]—because it cannot be found anywhere in [33].
2) *Section IX-C*: It *is*, however, identical to the *heuristic* version of the p.g.fl./functional derivative approach used in FISST—but without being so attributed.

The purpose of this section is to demonstrate that the FISST p.g.fl./functional calculus could not be "due to Moyal" even if it could be found in [33].

Let us begin with an example. In Section V-G.1 it was pointed out that Bartlett's and Kendall's p.g.fl. was an adaptation of Volterra's "functional power series" (f.p.s.) [51, p. 21]. Could we—in imitation of the "point process" insinuations of [47, p. 87], [43, p. 2], [48]—claim that the p.g.fl. is actually "due to Volterra"—a mere "corollary" of [51]? And that we could thereby bypass Bartlett and Kendall by purporting a p.g.fl. theory based only on [51]? Clearly we could not. Bartlett and Kendall devised a novel application of the f.p.s. to point process theory—an application Volterra could not have even known about. Only perfect hindsight would allow us to retroactively declare the f.p.s. to be the first p.g.fl. This observation leads us to the following point:

- The contents of the *JAIF* paper [48] can be portrayed as obvious implications or applications of Moyal's paper [33] only with the benefit of perfect hindsight gained from pre-existing knowledge of FISST—i.e., only with recourse to reverse engineering.

Three specific instances of such hindsight will be examined as case studies: the p.g.fl. (Section VI-A), the functional derivative (Section VI-B), and the FISST derivation methodologies (Section VI-C).

Consider the following analogy: the reverse-engineering of an electronic device. Any particular component of the device might be easily found in one of any number of parts catalogues. *Unacknowledged reverse-engineering* has occurred if one: (1) copies the parts list, (2) copies the schematic for assembling the parts into a functional device, (3) builds the device according that schematic; and (4) portrays the resulting device as either original or obvious.

In like manner, thousands of equations, concepts, discussions, and methodologies (the "parts") can be found in various publications in the point process literature (the "catalogues"). None of these publications addressed multitarget detection and tracking prior to the appearance of FISST.

A. The p.g.fl.

Moyal mentioned two point process "generating functionals": the p.g.fl. and the characteristic functional (c.fl.). However, the p.g.fl. and c.fl. are not the only possible generating functionals. Daley and Vere-Jones [4] describe many others, including the factorial moment-generating functional (f.m.g.fl.), the factorial cumulant generating functional (f.c.g.fl.), and the Laplace functional (L.fl.).

The p.g.fl. is not specially privileged—different purposes require different generating functionals. Snyder and Miller [40] apply point process theory to imaging applications. They employ only the c.fl.—the p.g.fl. is never even mentioned. The well-known textbook [10] by Karr employs only the L.fl.—and, once again, the p.g.fl. is never mentioned.

Out of all of the generating functionals, in [48] it was the p.g.fl. that was selected as the proper choice for multitarget tracking. This insight could not be "due to Moyal" since Moyal's paper did not address even Bayes filtering, let alone multitarget tracking. It is the case, however, that when [48] was submitted for publication the p.g.fl. had been a prominent "part" in the FISST "parts list" for over a decade—see [23], [22], [24], [26], and Section IX-B.4.

B. The functional derivative

Given the p.g.fl., there are many possible functional calculi from which one can choose: Gâteaux differentials, chain differentials, Gâteaux derivatives, Hadamard derivatives, Frechét derivatives, functional derivatives, etc.

Out of all of the possible calculi, in [48] it was the *heuristic* functional derivative—see Section IX-C—that was selected as the proper choice for multitarget tracking. This insight could not be "due to Moyal," since neither the functional derivative (heuristic or otherwise) nor the term "functional derivative" can be found in Moyal's paper—see Sections IX-A.2 and IX-D. Indeed, Moyal employed Gâteaux derivatives, not functional derivatives—see Section IX-A.2. It is the case, however, that when [48] was submitted for publication both the functional derivative and its heuristic version had been prominent "parts" in the FISST "parts list" for over a decade—see [23], [22], [24], [26], and Section IX-B.5.

C. The FISST derivation methodologies

Given the p.g.fl. and functional derivative, one also needs systematic methodologies for successfully applying them to multitarget tracking. Out the thousands of point process concepts and methodologies available in point process "catalogues," it was the concepts of the FISST derivation methodology of Section IX-B.6 that were selected in [48] as the proper choice for multitarget tracking (Section V-F). In addition, in [48] the end-results of the FISST modeling methodology of Section IX-B.3 were assumed without attribution.

None of these could be obvious extrapolations of Moyal's paper to multitarget tracking, given that it appeared at the same time as the Kalman filter and nearly 20 years before Reid's seminal MHT paper, [36]. It is the case, however, that when [48] was submitted for publication these concepts and methodologies had been highly visible components of the FISST "schematic" for over a decade—see [23], [22], [24], [26], and Sections IX-B.3 and IX-B.6.

D. Reverse-engineering

As was noted in the Introduction:

- Reverse-engineering is fundamentally different than engineering. It is easy to know the right things to do (e.g., to know what "parts" to look for in point process "catalogues")—and even easier to claim that these things are actually obvious—if someone else has shown you how to do it all (i.e., has specified the "parts list" and the "schematic") in complete detail.

VII. Conclusions

In this paper the following claims were demonstrated:

1) The "point process" alternative to the FISST RFS formulation of multitarget tracking, originally described in 2008 in [50], [45], appears to have been an obscured and phenomenologically erroneous copy of the RFS approach (Section III). Moreover, [50], [45] appear to have had major mathematical and conceptual errors (Sections X, XI).
2) The 2013 *Journal of Advances in Information Fusion* (*JAIF*) paper [48] was accepted, and is being cited, as original research. Yet nearly every equation, concept, discussion, derivation, and methodology in it first appeared in earlier FISST publications—but without being so attributed (Sections V, VIII).
3) In particular, in [48] it was claimed that a central FISST formula, the p.g.fl. version of Bayes' rule, is a mere special case of an alleged "point process" version of that same formula—even though the derivation of the "point process" formula was mathematically identical to the FISST derivation of the original FISST formula, without being so attributed (Section V-B).
4) The contents of the *JAIF* paper [48] can be portrayed as obvious implications or applications of the paper [33] by Moyal only with the benefit of perfect hindsight gained from pre-existing knowledge of the FISST "parts list" and "schematic"—i.e., only with recourse to reverse engineering (Section VI).
5) Consequently, the "point process" p.g.fl./functional derivative approach to multitarget tracking described in [48] and subsequently in [2] appears to be an obscured, phenomenologically erroneous, and improperly attributed copy of FISST.

Given these facts, the following question seems reasonable:

- Do the peer-review standards of the *Journal of Advances in Information Fusion* rise to the level expected of any scientifically credible journal?

It seems otherwise. It is unclear how, in any competent review, the errors identified in Section V-G could have been overlooked. The one in V-G.4 was obvious; those in Eqs. (26,30) were glaring; and the one in Section V-G.3 should have been blinding to anyone who actually examined Moyal's paper. It is unclear how other facts could have evaded notice: that almost the entirety of [48] had previously appeared in FISST publications; that this went unacknowledged in [48]; and that FISST could not possibly be "due to Moyal."

One explanation: *JAIF* does not require reviewers to be expert in a paper's subject, but merely "respected scientists/mathematicians."[17] Thus: eminences who know little about FISST are, regardless, *JAIF*-qualified to review papers about FISST. This elevation of uninformed eminence over actual expertise suggests a deep misunderstanding at *JAIF* of what proper scientific peer review is all about. The substitution of an appearance of scientific rigor in place of the reality seems almost guaranteed to proliferate scientific misconception and error, or worse—and not just in regard to "finite point processes" vis-a-vis multitarget tracking.

[17] R. Lynch (former ISIF VP of Communications), email, March 3, 2015.

VIII. Appendix: Content Analysis of [48]

Here it is demonstrated that nearly all of the equations, concepts, discussions, derivations, and methodologies in the *Journal of Advances in Information Fusion* paper [48] are, except for notation and terminology, identical to those in the earlier FISST publications [8], [23], [27], [22], [25, Sec. 3.4], [24], [26], and [3].[18] The phrase "is identical to" is shorthand for "is identical to except for notation and terminology."

1) [48, Eq. (1)]: identical to [22, Eq. (44)] or [24, Eq. (15)] or [26, Eq. (11.154)].
2) [48, Eq. (2)]: identical to the result in [22, p. 1165, Prop. 5c] or to the fifth equation from the top in the first column of [22, p. 1174].
3) [48, Eq. (3)]: identical to [22, Eq. (47)] or [3, Eq. (3)].
4) [48, Eq. (5)]: obvious from [48, Eq. (1)].
5) [48, Eq. (6)] (1st equation): identical to [22, Eq. (50)].
6) [48, Eq. (6)] (2nd equation): identical to 1st equation in 1st column of p. 1170 of [22] and to [22, Eq. (108)] with $g = \delta_{\mathbf{x}}$.
7) [48, Eq. (7)]: special case of [26, Eq. (11.251)] with $|Y| = 1$.
8) [48, Eq. (8)]: obvious from [48, Eq. (7)].
9) [48, Eq. (9)]: obvious from [48, Eq. (8)].
10) [48, Eq. (10)]: special case of [26, Eq. (11.251)] with $|Y| = 2$.
11) [48, Eq. (12)](1st equation): identical to [26, Eq. (11.251)].
12) [48, Eq. (12)] (2nd equation): identical to [26, Eq. (11.251)].
13) [48, Eq. (13)]: special case of [24, Eq. (22)] or [26, Eq. (14.278)] with $|X| = 1$.
14) [48, Eq. (14)]: special case of [24, Eq. (22)] or [26, Eq. (14.278)] with $|X| = 2$.
15) [48, Eq. (15)]: identical to [24, Eq. (22)] or to [26, Eq. (14.278)].
16) [48, Eq. (16)]: identical to [3, Eqs. (22,15)] or [26, Eqs. (16.35,16.26)].
17) [48, Eq. (17)]: identical to [22, Eqs. (60,41)] with $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$.
18) [48, Eq. (18)]: properly cited (to Daley & Vere-Jones).
19) [48, Eq. (19)]: properly cited (to Daley & Vere-Jones).
20) [48, Eq. (20)]: identical to [24, Eqs. (10,17)] or [26, Eqs. (11.158-11.163)].
21) [48, Eq. (21)]: identical to [24, Eqs. (10,17)] or [26, Eqs. (11.158-11.163)].
22) The following is then stated: "In the signal processing literature, $F^{\Xi}(z^{-1})$ is called the z-transform of the sequence of probabilities $p_N^{\Xi}(n) : n = 0, 1, ...$" This is identical to [26, pp. 343-344, 372].
23) [48, Eq. (21)]: identical to [24, Eq. (11)] or [26, Eqs. (11.11)].

[18] Statements of fact: The papers [22] and [24] are very well-known. The author of [48] was sitting in the front row when [3] was presented. He once told me that he had read [27]. In [48], he cites [26, pp. 757-756]. In [2] he cites [23].

24) The following is then stated: "The probability $p_N^{\Xi}(n)$ is $n!$ times the integral of the ordered pdf $p^{\Xi}(n, x_1, ..., x_n)$ over all $x_1, ..., x_n$." This is identical to [24, Eq. (6)] or [26, Eq. (11.115)].
25) [48, Eq. (22)]: identical to [24, Eq. (12)] or [26, Eq. (11.164)].
26) [48, Eq. (23)]: identical to 4th equation from top in the 2nd column of p. 1173 in [22]. Also: [48, Eq. (24)] is an obvious consequence of this equation.
27) [48, Eq. (24)]: obvious consequence of [48, Eq. (23)].
28) [48, Eq. (25)]: identical to [26, Eq. (G.434)].[19]
29) [48, Eq. (26)]: identical to [26, Eq. (G.435)]:[20]
30) [48, Eq. (27)]: identical to [26, Eqs. (G.437,G.438)].
31) [48, Eqs. (25,26,27,28)]: identical to [26, Eq. (G.434, G.435, G.436, G. 437, G.438)].
32) [48, Eq. (29)]: identical to [24, Eq. (52)] or [26, Eq. (16.412)].
33) [48, Eq. (30)]: identical to [24, Eq. (46)] or [26, Eq. (16.404)] with $h = x$.
34) [48, Eq. (31)] (1st equation): an erroneous attempt to replicate [24, Eq. (168)].
35) [48, Eq. (31)] (2nd equation): obvious consequence of [48, Eq. (21)].
36) "Posterior pdf of the canonical number" or "canonical number distribution" [48, p. 128] or "cardinal number density" [46, p. 45]: systematic substitutions for the FISST "cardinality distribution."
37) Section 4.4 of [48]: unattributed copy of discussions in [27, p. 59], [8, p. 189-190], and [22, p. 1159, Sec. II-B.7].
38) [48, p. 124]: "The MAP estimate is undefined for the [multitarget] posterior pdf $p^{\Xi|\Upsilon}(\xi|\upsilon)$. To see this, it is only necessary to observe that $p^{\Xi|\Upsilon}(\xi_1|\upsilon)$ and $p^{\Xi|\Upsilon}(\xi_2|\upsilon)$ have different units when the realizations ξ_1 and ξ_2 have different numbers of points" (where $\xi = (n, x_1, ..., x_n)$ and $\upsilon = (m, y_1, ..., y_m)$). An unattributed copy of the discussions in [27, p. 59], [22, p. 1159, Sec. II-B.7], and [26, pp. 494-495].
39) [48, p. 124]: "Pseudo-MAP estimates can be defined using the posterior distribution of the canonical number and intensity functions, or other summary statistics." Unattributed reference to the heuristic multitarget state estimators for the PHD and CPHD filters as described in [26, pp. 595, 640].
40) [48, Eqs. (34,35)]: identical to 5th and 6th equations from top in 2nd column of p. 1173 of [22], or to [24, Eqs. (45,44)].
41) [48, Eq. (36)]: an erroneous attempt to replicate [22, Eq. (26)].
42) [48, Eq. (37)]: identical to special case of [3, Eq. (68)] with $\kappa_{k+1}[g] = 1$ (no clutter), and where [3, Eq. (65)] $T_{k+1}[g](\mathbf{x}) = G_{k+1}[g|\mathbf{x}]$.[21]
43) [48, Eq. (38)]: identical to special case of [3, Eq. (71)] with $\kappa_{k+1}[g] = 1$ (no clutter).
44) [48, Eq. (39)]: identical to special case of [22, Eq. (75)] with $b_{k+1|k}(\mathbf{x}|\mathbf{w}) = 0$ (targets do not spawn other targets).
45) [48, Eq. (42)]: identical to 2nd equation from the bottom in 2nd column of p. 1173 in [22] where $G_{\Theta}[g]$ is the p.g.fl. of the clutter process; and also to [3, Eq. (68)].
46) [48, Eq. (43)]: identical to 3rd equation from bottom in 2nd column of p. 1173 in [22], where $p_g(\mathbf{x}) = \int g(\mathbf{z}) f_{k+1}(\mathbf{z}|\mathbf{x}) d\mathbf{z}$.
47) Also, [48, Eq. (44)] is an obvious consequence of Points 43 and 44.
48) [48, Eq. (45)]: obvious consequence of Points 38 and 43.
49) [48, Eq. (46)]: identical to 7th equation from the top in 1st column of p. 1174 in [22].
50) What follows is a continuation of the unattributed copy of the FISST derivation of the PHD filter, using the FISST methodology described in [25, Sec. 3.4] and [24, pp. 1535-1538]. Eqs. (47-52) are special cases of equations that occur in the derivation of the CPHD filter in [24, pp. 1535-1538]. Some are, except for notation, identical to those in [25, Sec. 3.4]. [48, Eq. (47)] is a special case of the equation between Eqs. (20,21) of [25] with $|Z| = 1$. It is also the PHD filter special case of [24, Eq. (124)] with $m = 1$. [48, Eq. (48)] is identical to the equation between Eqs. (20,21) of [25]. It is also the PHD filter special case of [24, Eq. (124)]. [48, Eq. (49)] is the PHD filter special case of [24, Eq. (139)]. [48, Eq. (50)] is identical to is a special case of the equation between Eqs. (20,21) of [25] with $g = 0$ and $h = 1$. It is also the PHD filter special case of [24, Eq. (126)]. [48, Eq. (51)] is the PHD filter special case of [24, Eq. (141)]. [48, Eq. (52)]:is identical to [24, Eq. (155)] or [22, Eqs. (8-9)] or [26, Eqs. (16.108,16.109)]. [48, Eq. (54)] is identical to [25, Eqs. (21,22)] with $h = x$. It is also a special case of the CPHD filter equation [24, Eq. (156)].
51) Section 5.2: A 2nd unattributed copy of the FISST derivation methodology as described in [25], [24, pp. 1535-1538], [3]—this time applied to the iFilter.
52) [48, Eq. (53)]: identical to [22, Eq. (65)].

IX. Appendix: FISST vs. Moyal

In this section it will be demonstrated that:

1) The alleged "point process" p.g.fl./functional calculus approach in [48] is neither "due to Moyal" nor "has exactly the same meaning" as in [33]—because it cannot be found anywhere in [33].
2) It *is*, however, identical to the *heuristic* version of the p.g.fl./functional derivative approach used in FISST—but without being so attributed.

The section is organized as follows: a summary of Moyal's functional calculus (Section IX-A); a summary of the FISST calculus and methodologies (Section IX-B); heuristic functional derivatives (Section IX-C); and a discussion of "secular functions" (Section IX-D).

[19] *Note*: The "$f_{k+1}(W|X)$" in Eq. (G.434) is a typo—it should be "$f_{k+1}(Z|X)$."

[20] *Note*: The "$f_{k+1}(W|Z^{(k)})$" in Eq. (G.435) is a typo—it should be "$f_{k+1}(Z|Z^{(k)})$."

[21] *Note*: In [3, Eq. (65)], "$f_{k+1}[g|\mathbf{x}]$" is a typo. It should be "$G_{k+1}[g|\mathbf{x}]$".

A. Summary of Moyal's calculus

In [33, p. 2-3], Moyal formulated three versions of point process theory and showed them to be equivalent. For our purposes we need only consider one of them, which was discussed in Section III-A. The statistics of a point process $\mathcal{P}$ are specified by its probability measure $P(O) = \Pr(\mathcal{P} \in O)$ where O is a measurable subset of the space $\mathcal{X}$ of all *finite unordered lists*[22] $\{x_1, ..., x_n\}$ of arbitrary length $n \geq 0$ [33, Eq. (2.5)]. The measure P induces, and is completely determined by, symmetric multivariate measures $P^{(n)}(S_1, ..., S_n)$ for all measurable $S_1, ..., S_n \subseteq \mathfrak{X}$ (known in modern parlance as "Janossy measures" [4]).

1) Measure-theoretic p.g.fl.'s : Let $h(\cdot)$ be a bounded, complex-valued integrable function with norm [33, Eq. (4.1)]

$$\|h\| = \sup_{x \in \mathfrak{X}} |h(x)|. \tag{34}$$

Then the p.g.fl. of $\mathcal{P}$ is ([33, Eq. (4.11)] with $k = 0$):

$$G[h] = \sum_{n=0}^{\infty} \int_{\mathfrak{X}^n} h(x_1) \cdots h(x_n) \cdot P^{(n)}(dx_1 \cdots dx_n) \tag{35}$$

where the integrals are taken with respect to the measures $P^{(n)}(S_1, ..., S_n)$, and where the summation converges for those h such that $G[\|h\|] < \infty$.

2) Gâteaux differentials and derivatives : Moyal employs the "n-th order variation" [33, Eq. (4.11)]

$$\delta^n_{w_1,...,w_n} G[h] = \left[\frac{\partial^n}{\partial \varepsilon_1 \cdots \partial \varepsilon_n} G \left[h + \sum_{i=1}^{n} \varepsilon_i w_i \right] \right]_{\varepsilon_1 = ... = \varepsilon_n = 0} \tag{36}$$

where $w_1(x), ..., w_n(x)$ are bounded, complex-valued functions. The first-order variation is, therefore,

$$\delta_w G[h] = \left[\frac{d}{d\varepsilon} G\left[h + \varepsilon w\right] \right]_{\varepsilon=0} \tag{37}$$

$$= \lim_{\varepsilon \to 0} \frac{G\left[h + \varepsilon w\right] - G\left[h\right]}{\varepsilon}. \tag{38}$$

In the modern mathematical literature, Eq. (38) is known as a *Gâteaux differential*, and is not specific to functionals.

If $\delta_w G[h]$ is linear and continuous in w, then $\delta_w G[h]$ is known as the *Gâteaux derivative* of $G[h]$ at h and is often denoted as $(D_h G)[w]$ [7].

Now suppose that the p.g.fl. $G[h]$ of $\mathcal{P}$ is known. Then the *multivariate measures* of $\mathcal{P}$ can be recovered via iterated Gâteaux differentials [33, Eq. (4.14)]:

$$P^{(n)}(S_1, ..., S_n) = \frac{1}{n!} \cdot \delta^n_{\mathbf{1}_{S_1},...,\mathbf{1}_{S_n}} G[0] \tag{39}$$

where $\mathbf{1}_S(x)$ denotes the indicator function of $S \subseteq \mathfrak{X}$. (Since the left side is presumed to be a measure, Moyal implicitly assumes that the right side is an iterated Gâteaux derivative.) The multivariate factorial moment *measures* of $\mathcal{P}$ can similarly be recovered as [33, Eq. (4.16)]:

$$M_{(n)}(S_1, ..., S_n) = \delta^n_{\mathbf{1}_{S_1},...,\mathbf{1}_{S_n}} G[1]. \tag{40}$$

[22]Moyal used the notation $\{x_1, ..., x_n\}$ to denote a finite unordered list, not (as is now standard practice) a finite set.

In particular, the first factorial moment *measure*—i.e., the measure corresponding to the PHD if the PHD exists—is:

$$M_{(1)}(S) = \delta_{\mathbf{1}_S} G[1]. \tag{41}$$

Moyal does not provide or suggest any methodology for deriving concrete formulas even for his multivariate *measures* in Eqs. (39-41)—let alone for their corresponding density functions (as would be required for application). A less abstract and more practically useful approach is required. FISST was specifically designed to be such an approach.

B. Summary of FISST

This section is organized as follows: belief measures (Section IX-B.1); set derivatives (Section IX-B.2); RFS motion and measurement modeling(Section IX-B.3); the FISST approach to p.g.fl.'s (Section IX-B.4); functional derivatives (Section IX-B.5); the FISST derivation methodology (Section IX-B.6); and heuristic functional derivatives (Section IX-C).

1) Belief measures : As was noted in Sections III-B and III-C, when applied to practical multitarget tracking: (1) a "point process" is the same thing as an RFS and (2) a non-RFS "point process" is phenomenologically erroneous. Moyal's measure-theoretic framework is therefore an unnecessary obfuscation and so FISST employs RFS's from the very outset. If Ξ is an RFS then the Janossy measures $P^{(n)}(S_1, ..., S_n)$ have Janossy densities

$$p_n(x_1, ..., x_n) = \frac{1}{n!} \cdot f_\Xi(\{x_1, ..., x_n\}), \tag{42}$$

where $f_\Xi(X)$ is the FISST notation for a multitarget probability density function.

Moreover, FISST specifically avoids more complicated formulations of RFS theory by adopting the *stochastic geometry* formulation [42]. In this case the statistics of Ξ are characterized by its *belief measure* [26, pp. 711-716]:

$$\beta_\Xi(S) = \Pr(\Xi \subseteq S), \tag{43}$$

which is not only far more physically intuitive than Moyal's measures $\{P^{(n)}(S_1, ..., S_n)\}_{n=0}^{\infty}$, but also allows direct set-theoretic access to the RFS Ξ.

2) Set derivatives : Moreover, the Janossy densities can be constructed from $\beta_\Xi(S)$. Define the first-order *set derivative* of $\beta_\Xi(S)$ at the point $x \in \mathfrak{X}$ to be [8, pp. 144-151], [26, pp. 380-381], [28, Eqs. (54,55)]

$$\frac{\delta \beta_\Xi}{\delta \mathbf{x}}(S) = \lim_{|E'_\mathbf{x}| \searrow 0} \lim_{|E_\mathbf{x}| \searrow 0} \frac{\beta_\Xi(S_{E'_\mathbf{x}} \cup E_\mathbf{x}) - \beta_\Xi(S_{E'_\mathbf{x}})}{|E_\mathbf{x}|} \tag{44}$$

where $E_\mathbf{x}$ and $E'_\mathbf{x}$ are drawn from special families [39, Ch. 10] of small neighborhoods of x with $S_{E'_\mathbf{x}} \stackrel{\text{def.}}{=} S - E'_\mathbf{x}$. If $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ with $|X| = n$, the *general set derivative* is [8, pp. 144-151], [26, pp. 380-381]

$$\frac{\delta \beta_\Xi}{\delta X}(S) = \left\{ \begin{array}{ccc} \beta_\Xi(S) & \text{if} & X = \emptyset \\ \frac{\delta^n \beta}{\delta \mathbf{x}_1 \cdots \delta \mathbf{x}_n}(S) & \text{if} & \text{otherwise} \end{array} \right. . \tag{45}$$

Given this, $f_\Xi(X)$ can be computed as a set derivative:

$$f_\Xi(X) = \frac{\delta \beta_\Xi}{\delta X}(\emptyset). \tag{46}$$

Similarly, the density functions of Moyal's measures $M_{(n)}(S_1, ..., S_n)$ and $M_{(1)}(S)$ can be computed directly as set derivatives of belief measures:

$$M_\Xi(X) = \frac{\delta \beta_\Xi}{\delta X}(\mathfrak{X}), \qquad D_\Xi(\mathbf{x}) = \frac{\delta \beta_\Xi}{\delta \mathbf{x}}(\mathfrak{X}). \tag{47}$$

3) RFS modeling methodology : Belief measures and set derivatives form the basis of a rigorous, systematic methodology for multitarget motion and measurement modeling.

The simplest RFS motion model has the form

$$\Xi_{k|k-1} = T_{k|k-1}(\mathbf{x}'_1) \cup ... \cup T_{k|k-1}(\mathbf{x}'_{n'}) \cup B_{k|k-1} \tag{48}$$

where $X' = \{\mathbf{x}'_1 ..., \mathbf{x}'_{n'}\}$ with $|X'| = n'$ is the multitarget state-set at time t_{k-1}; where $T_{k|k-1}(\mathbf{x}')$ is the RFS model of the survival or disappearance of a target with state $\mathbf{x}'$ at time t_{k-1}; $B_{k|k-1}$ is the RFS model for newly-appearing targets; and $\Xi_{k|k-1}$ is the predicted multitarget state-set.

Likewise, the simplest RFS measurement model is

$$\Sigma_{k|k} = \Upsilon_k(\mathbf{x}_1) \cup ... \cup \Upsilon_k(\mathbf{x}_n) \cup C_k \tag{49}$$

where $X = \{\mathbf{x}_1 ..., \mathbf{x}_n\}$ with $|X| = n$ is the multitarget state-set at time t_k; where $\Upsilon_k(\mathbf{x})$ is the random measurement-set generated by a target with state $\mathbf{x}$ at time t_k; $C_{k|k-1}$ is the RFS modeling the clutter process at time t_k; and $\Sigma_{k|k}$ is the random measurement-set at time t_k.

The statistics of $\Xi_{k|k-1}$ and $\Sigma_{k|k}$ are completely characterized by their belief measures

$$\beta_{k|k-1}(S) = \Pr(\Xi_{k|k-1} \subseteq S) \tag{50}$$

$$\beta_{k|k}(S) = \Pr(\Sigma_{k|k} \subseteq S). \tag{51}$$

Once concrete formulas have been derived for $\beta_{k|k-1}(S)$ and $\beta_{k|k}(S)$, formulas for the multitarget Markov density and multitarget likelihood function can be derived using set derivatives:

$$f_{k|k-1}(X|X') = \frac{\delta \beta_{k|k-1}}{\delta X}(\emptyset|X') = \left[\frac{\delta \beta_{k|k-1}}{\delta X}(S|X')\right]_{S=\emptyset} \tag{52}$$

$$f_{k|k}(Z|X) = \frac{\delta \beta_{k|k}}{\delta Z}(\emptyset|X) = \left[\frac{\delta \beta_{k|k}}{\delta Z}(S|X)\right]_{S=\emptyset}. \tag{53}$$

4) RFS formula for p.g.fl.'s : In 2001 in [23], the belief measure and set derivative were generalized to, respectively, the *probability generating functional* (p.g.fl.) and the *functional derivative*. In FISST the p.g.fl. is defined as:

$$G_\Xi[h] = \int h^X \cdot \frac{\delta \beta_\Xi}{\delta X}(\emptyset) \delta X \tag{54}$$

where the *set integral* is defined by:[23]

$$\int f(X)\delta X = \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{\mathbf{x}_1, ..., \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n \tag{55}$$

and where the *power functional* h^X is defined by $h^X = 1$ if $X = \emptyset$ and $h^X = \prod_{\mathbf{x} \in X} h(\mathbf{x})$ if otherwise. Following modern practice [4], in FISST it is assumed that $0 \le h(\mathbf{x}) \le 1$ identically—in which case $0 \le G_\Xi[h] \le 1$.

[23]The factor $(n!)^{-1}$ is important, because it results in greatly simplified multitarget calculus formulas. It is also often assumed in physics—see [9, pp. 234, 266].

5) Rigorous functional derivatives : The *functional derivative* of a functional $F[h]$ was introduced by Volterra [51, pp. 22-23, 75] and is commonly denoted [7, Eqs. (A.15,A.23,A.24)]:[24]

$$\frac{\delta F}{\delta h(\mathbf{x})}[h] \overset{\text{abbr.}}{=} \frac{\delta F}{\delta \mathbf{x}}[h] \tag{56}$$

where the right side is the abbreviation used in FISST. Note the following:

- The functional derivative is not the Gâteaux differential $\delta_w F[h]$ used by Moyal, Eq. (35).[25]
- Neither the functional derivative nor the terminology "functional derivative" occur anywhere in Moyal's paper.

Rather, if $\delta_w F[h]$ is linear and continuous in $w(\cdot)$ then the following gives the relationship between the functional derivative and the Gâteaux derivative [51, p. 24, Eq. (3)]:

$$\int w(\mathbf{x}) \cdot \frac{\delta F}{\delta \mathbf{x}}[h] d\mathbf{x} = \frac{\partial F}{\partial w}[h] \tag{57}$$

where the right side is the FISST notation[26] for a Gâteaux derivative $(D_h F)[w]$.

The p.g.fl. is a *generalized belief measure* and the functional derivative is a *generalized set derivative* since [26, p. 373]

$$\beta_\Xi(S) = G_\Xi[\mathbf{1}_S] \tag{58}$$

$$\frac{\delta \beta_\Xi}{\delta X}(S) = \frac{\delta G_\Xi}{\delta X}[\mathbf{1}_S]. \tag{59}$$

From Eqs. (46,47), it follows that the density functions of Moyal's measures $P^{(n)}(S_1, ..., S_n)$, $M_{(n)}(S_1, ..., S_n)$, and $M_{(1)}(S)$ can be computed as functional derivatives—which, rigorously speaking, are set derivatives:[27]

$$\frac{\delta G_\Xi}{\delta X}[0] = \frac{\delta \beta_\Xi}{\delta X}(\emptyset) = f_\Xi(X) \tag{60}$$

$$\frac{\delta G_\Xi}{\delta X}[1] = \frac{\delta \beta_\Xi}{\delta X}(\mathfrak{X}) = M_\Xi(X) \tag{61}$$

$$\frac{\delta G_\Xi}{\delta \mathbf{x}}[1] = \frac{\delta \beta_\Xi}{\delta \mathbf{x}}(\mathfrak{X}) = D_\Xi(\mathbf{x}). \tag{62}$$

6) FISST derivation methodology : This methodology is based on the p.g.fl. versions of the multitarget prediction integral and the multitarget Bayes' rule [26]:

$$G_{k|k-1}[h] = \int G_{k|k-1}[h|X'] \cdot f_{k-1|k-1}(X) \delta X \tag{63}$$

$$G_{k|k}[h] = \frac{\frac{\delta F_k}{\delta Z_k}[0,h]}{\frac{\delta F_k}{\delta Z_k}[0,1]} \tag{64}$$

[24]This meaning of the term "functional derivative"—i.e., as Volterra's derivative—is the established usage in the mathematical literature. Besides [7, p. 35], see the entry for "functional derivative" in both Wikipedia and the online *Encyclopedia of Mathematics*.

[25]See [7, Eqs. (A.15,A.23,A.24)] or the entry for "Gâteaux differential" in Wikipedia or the online *Encyclopedia of Mathematics*.

[26]This notation, which is deliberately imitative of undergraduate calculus, is unique to FISST—but was not so attributed in [48].

[27]The theoretically rigorous functional derivative $(\delta G_\Xi / \delta \mathbf{x})[h]$ for a general $h(\cdot)$ is also a set derivative (see [28, Eq. (78)]), but the specifics are unnecessary for current purposes.

where

$$F_k[g,h] = \int h^X \cdot G_k[g|X] \cdot f_{k|k-1}(X)\delta X \quad (65)$$

$$G_{k|k-1}[h|X'] = \int h^X \cdot f_{k|k-1}(X|X')\delta X \quad (66)$$

$$G_k[g|X] = \int g^Z \cdot f_k(Z|X)\delta Z. \quad (67)$$

Given this, approximate RFS filters, such as PHD, CPHD, or multi-Bernoulli filters, can be derived from the formulas for $G_{k|k-1}[h]$ and $G_{k|k}[h]$.

C. Heuristic Functional Derivatives

Suppose, as is commonly done in physics [38, pp. 173-174], that we substitute $w = \delta_{\mathbf{x}}$. Then from Eq. (57) we could compute the functional derivative directly:

$$\frac{\partial F}{\partial \delta_{\mathbf{x}}}[h] = \lim_{\varepsilon \searrow 0} \frac{G_\Xi[h+\varepsilon\delta_{\mathbf{x}}] - G_\Xi[h]}{\varepsilon} = \frac{\delta F}{\delta \mathbf{x}}[h]. \quad (68)$$

However, this substitution is mathematically erroneous in both Moyal's and the FISST frameworks. This is because $\delta_x(\cdot)$ is not only unbounded, it is not even a function since $\delta_x(x) = \infty$. Nevertheless, in FISST it is employed as a purely heuristic adjunct to the rigorous definition, Eq. (44).

In any case, Eq. (68) is not "due to Moyal" or "has exactly the same meaning as in" Moyal's [33] because it cannot be found anywhere in [33]. This, in turn, is because it is mathematically erroneous in Moyal's framework.

D. Functional derivatives and "secular functions"

In the paragraph following [48, Eq. (6)], this difficulty was addressed as follows. There it was stated that:

- "Alternatively, specifying the variation to be a function in a test sequence for the delta function and taking the limit gives the same result."

That is, if $w_{\mathbf{x},n} \to \delta_{\mathbf{x}}$ (in some sense that is not explained) then the Gâteaux derivative in the direction of $\delta_{\mathbf{x}}$ is:

$$\frac{\partial G_\Xi}{\partial \delta_{\mathbf{x}}}[h] \stackrel{\text{def.}}{=} \lim_{n\to\infty} \frac{\partial G_\Xi}{\partial w_{\mathbf{x},n}}[h]. \quad (69)$$

Eq. (69) cannot be "due to Moyal" because it cannot be found anywhere in Moyal's paper. This, in turn, is due to the fact that it is erroneous within the mathematical theory described therein. If $w_n \to \delta_{\mathbf{x}}$ then, for every number $\varepsilon > 0$, there must exist an integer $N > 0$ such that $\|w_n - \delta_{\mathbf{x}}\| < \varepsilon$ for all $n \geq N$, where $\|\cdot\|$ is Moyal's norm, Eq. (34). However, $\|w_n - \delta_{\mathbf{x}}\|$ is mathematically undefined—and would be infinite (and therefore always larger than ε) even if defined.

However, the phrase "a test sequence for the delta function" appears to be an unstated reference to Lighthill's theory of "generalized functions" [12]. For, in a subsequent paper [49], Eq. (69) was recast in terms of the "secular function" defined as follows [49, Eq. (6)]:[28]

$$J_{F,h}(\alpha; \mathbf{x}) = \lim_{n\to\infty} G_\Xi[h + \alpha \cdot w_{\mathbf{x},n}] \quad (70)$$

where (without being so identified in [49]) $w_{\mathbf{x},n}$ are "good functions" [12, p. 16] such that

$$\lim_{n\to\infty} \int w_{\mathbf{x},n}(\mathbf{y}) \cdot w(\mathbf{y}) d\mathbf{y} = w(\mathbf{x}) \quad (71)$$

for any "good function" w. Specifically in [49] the functional derivative is recast as:

$$J^{(1)}_{F,h}(0;\mathbf{x}) = \left[\frac{d}{d\alpha} \lim_{n\to\infty} G_\Xi[h+\alpha\cdot w_{\mathbf{x},n}]\right]_{\alpha=0} \quad (72)$$

$$= \lim_{n\to\infty} \left[\frac{d}{d\alpha} G_\Xi[h+\alpha\cdot w_{\mathbf{x},n}]\right]_{\alpha=0} \quad (73)$$

$$= \lim_{n\to\infty} \frac{\partial G_\Xi}{\partial w_{\mathbf{x},n}} = \frac{\delta G_\Xi}{\delta \mathbf{x}}[h] \quad (74)$$

where the first part of Eq. (74) results from Eq. (37).

However this may be, it is irrelevant. Eq. (69) still cannot be "due to Moyal" because Moyal neither cites [12] nor makes any reference to generalized function theory. Indeed, he may have been unaware of [12] at the time, since it had appeared only four years earlier.

X. Appendix: The "iFilter"

In this section, the iFilter is described (Section XI), along with major mathematical (Section X-B) and phenomenological (Section X-C) issues in its derivation, as well as various claims made for it (Section X-D).

A. iFilter: Concept

Introduced in 2008 in [50], the "multitarget intensity filter"—subsequently renamed the "iFilter"—was promoted as a signal achievement of the "elementary," "point process" approach to multitarget tracking. A two-step process was adopted in [50]. First, an alleged "elementary" re-derivation of the PHD filter was devised. Then, after replacing the PHD filter's clutter and target-appearance models with different ones, this derivation was refashioned into a derivation of the iFilter. Consequently, any error in the derivation of the PHD filter propagates into an error in the derivation of the iFilter.

We begin with summaries of the PHD filter (Section X-A.2) and the iFilter (Section X-A.1).

1) PHD filter models and equations : At any prediction time t_{k-1}, the target-appearance process is assumed to be Poisson with known PHD $N^B_{k|k-1} s^B_{k|k-1}(\mathbf{x})$, where $N^B_{k|k-1}$ is the expected number of appearing targets and $s^B_{k|k-1}(\mathbf{x})$ is the spatial distribution of those appearances. The probability that a target with state $\mathbf{x}'$ at time t_{k-1} will persist into time t_k is $p_S(\mathbf{x}') \stackrel{\text{abbr.}}{=} p_{S,k|k-1}(\mathbf{x}')$. If it persists, the probability

[28] The claimed purpose of "secular functions" is to transform functional derivatives into ordinary Newtonian derivatives, so that conventional computer algebra algorithms can be used to construct explicit formulas for functional derivatives. However, an arsenal of FISST calculus identities (e.g., the general product and chain rules) permit the derivation of extremely complicated functional derivatives (see Chapters 3, 4 of [13]). It was not explained in [49] why conventional computer algebra—even if up to the task, which is questionable—would be more advantageous than the FISST calculus.

(density) that it will transition to state $\mathbf{x}$ is $f_{k|k-1}(\mathbf{x}|\mathbf{x}')$. The PHD filter time-update equation is:

$$D_{k|k-1}(\mathbf{x}) = N^B_{k|k-1} s^B_{k|k-1}(\mathbf{x}) + \int p_S(\mathbf{x}') \cdot f_{k|k-1}(\mathbf{x}|\mathbf{x}') \cdot D_{k-1|k-1}(\mathbf{x}') d\mathbf{x}'. \quad (75)$$

Likewise, at any time t_k with newly-collected measurement-set Z_k, the clutter process is assumed to be Poisson with known PHD (intensity function) $\lambda_k c_k(\mathbf{z})$ where λ_k is the expected number of clutter measurements ("clutter rate") and $c_k(\mathbf{z})$ is the clutter spatial distribution. The probability that a target with state $\mathbf{x}$ generates a measurement is $p_D(\mathbf{x}) \overset{\text{abbr.}}{=} p_{D,k}(\mathbf{x})$, and the probability (density) that this measurement will be $\mathbf{z}$ is $f_k(\mathbf{z}|\mathbf{x})$. The PHD filter measurement-update is:

$$\frac{D_{k|k}(\mathbf{x})}{D_{k|k-1}(\mathbf{x})} = L_{Z_k}(\mathbf{x}) \quad (76)$$

$$= 1 - p_D(\mathbf{x}) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\mathbf{x}) \cdot f_k(\mathbf{z}|\mathbf{x})}{\lambda_k c_k(\mathbf{z}) + \tau_k(\mathbf{z})} \quad (77)$$

where $\tau_k(\mathbf{z}) = \int p_D(\mathbf{x}) \cdot f_k(\mathbf{z}|\mathbf{x}) \cdot D_{k|k-1}(\mathbf{x}) d\mathbf{x}$ and where $L_{Z_k}(\mathbf{x})$ is called the "pseudolikelihood."

2) iFilter models and equations : The following material first appeared in 2012 in [16]. In [50] the target state space $\mathfrak{X}$ is replaced by $\mathfrak{X}^+ = \mathfrak{X} \uplus \mathfrak{X}_\phi$ where $\mathfrak{X}_\phi$ is a space of "clutter targets" from which all clutter measurements are generated. The probability that a "clutter target" will generate a clutter measurement is $p_D(\phi)$, and the probability (density) that it will generate $\mathbf{z}$ is $c_k(\mathbf{z})$.

The appearance of a new target with state $\mathbf{x}$ at time t_k is modeled as the transition of some clutter target to $\mathbf{x}$, as described by

$$\psi_k(\mathbf{x}|\phi) = (1 - \psi_k(\phi|\phi)) \cdot s^B_{k+1|k}(\mathbf{x}), \quad (78)$$

where $\psi_k(\phi|\phi)$ is the probability that a clutter target will transition to another clutter target. Analogously, clutter at time t_k is modeled as the transition of a target at time t_{k-1} with state $\mathbf{x}'$ to to some clutter target, with probability $\psi_k(\phi|\mathbf{x}')$. Thus an actual target with state $\mathbf{x}'$ at time t_{k-1} transitions to an actual target with state $\mathbf{x}$ at time t_k according to

$$\psi_k(\mathbf{x}|\mathbf{x}') = (1 - \psi_k(\phi|\mathbf{x}')) \cdot f_{k+1|k}(\mathbf{x}|\mathbf{x}'). \quad (79)$$

The iFilter propagates not only the PHD $D_{k|k}(\mathbf{x})$ but also, implicitly, the expected number $\dot{N}_{k|k}$ of clutter targets. It was claimed in [50, p. 1] that:

- "[T]he target birth and measurement clutter processes that are assumed specified a priori in [Mahler's PHD papers] are estimated here."

This is untrue. In actuality, $c_k(\mathbf{z})$ and $s^B_{k|k-1}(\mathbf{x})$ are assumed known a priori (in fact, $s^B_{k|k-1}(\mathbf{x})$ is implicitly assumed to be constant[29]) and thus only λ_k and $N^B_{k|k-1}$ are estimated. However, even this turns out to be untrue in general, see Section X-D.

The iFilter time-update equations are

$$D_{k|k-1}(\mathbf{x}) = \psi_{k-1}(\mathbf{x}|\phi) \cdot \dot{N}_{k-1|k-1} + \int \psi_k(\mathbf{x}|\mathbf{x}') \cdot D_{k-1|k-1}(\mathbf{x}') d\mathbf{x}' \quad (80)$$

$$\dot{N}_{k|k-1} = \psi_{k-1}(\phi|\phi) \cdot \dot{N}_{k-1|k-1} + \int \psi_{k-1}(\phi|\mathbf{x}') \cdot D_{k-1|k-1}(\mathbf{x}') d\mathbf{x}' \quad (81)$$

where $\hat{N}^B_{k|k-1} = (1 - \psi_k(\phi|\phi)) \cdot \dot{N}_{k-1|k-1}$ is a claimed estimate of $N^B_{k|k-1}$. The iFilter data-update equations are

$$\frac{D_{k|k}(\mathbf{x})}{D_{k|k-1}(\mathbf{x})} = 1 - p_D(\mathbf{x}) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\mathbf{x}) \cdot f_k(\mathbf{z}|\mathbf{x})}{\hat{\lambda}_k c_k(\mathbf{z}) + \tau_k(\mathbf{z})} \quad (82)$$

$$\frac{\dot{N}_{k|k}}{\dot{N}_{k|k-1}} = 1 - p_D(\phi) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\phi) \cdot c_k(\mathbf{z})}{\hat{\lambda}_k c_k(\mathbf{z}) + \tau_k(\mathbf{z})} \quad (83)$$

where $\hat{\lambda}_k = p_D(\phi) \cdot \dot{N}_{k|k-1}$ is a claimed estimate of λ_k.

B. iFilter: Mathematical issues

In [50, Eq. (27)], the authors introduce $c > 0$ such that

$$f^{Detected}_{k|k}(x) = \frac{c}{m} \sum_{j=1}^{m} \frac{p_{Z_k|X_k}(z_j|x)\; p^D(x)\; f_{k|k-1}(x)}{\lambda^{Target}_{k|k-1}(z_j)}. \quad (84)$$

To reverse-engineer Eq. (76), they must show that $c = m$, where m is the current number of measurements. Without proof, they claimed that (1) c is a random variable whose probability distribution is $L(c) \propto e^{-c} c^m$ [50, Eq. (28)]; and (2) c can be determined by applying the maximum a posteriori (MAP) estimator to $e^{-c} c^m$, resulting in $c = m$. This argument contains three major errors of basic statistical reasoning, first pointed out in 2010 [18]:

1) c is not a random variable. It is actually equal to m, which—since it is a fixed realization of a random integer M—is a constant. Thus c is also a constant.
2) The probability distribution of c is, therefore, not $L(c) \propto e^{-c} c^m$ but rather a Dirac delta $L(c) = \delta_m(c)$ (which, of course, cannot be legitimately assumed since this would require knowing a priori that $c = m$—the very thing that is to be proved).
3) Why a MAP estimator rather than, say, the expected value estimator—in which case we get $c \neq m$?

Thus, unfortunately, the "elementary" derivation of the PHD filter is erroneous. Consequently, so is the "elementary" derivation of the iFilter.[30]

C. iFilter: Phenomenological issues

There is a more fundamental difficulty: the iFilter appears to be questionable from a phenomenological point of view [16]. The key concept underlying it is the notion of modeling clutter (and target disappearances) as transitions of actual targets to "clutter targets," and modeling target appearances as transitions of "clutter targets" to actual targets.

[29] See [50], sentence following Eq. (41)).

[30] The above major errors are not the only ones in [50], just the ones that are most easily explained and understood.

However, an effective multitarget tracker must be able to accurately distinguish targets from background clutter. It is therefore crucial to efficiently exploit differences between (1) clutter versus target measurement-generation statistics, and (2) clutter versus target motion (i.e., state-transition statistics).

But when targets are allowed to become clutter and/or vice-versa, clutter statistics and target statistics become intertwined. This will make it more difficult to distinguish targets from clutter—e.g., when a jet fighter can transition to a radar false alarm or vice-versa. Such difficulties will be even more pronounced when the statistics of the targets depend on target identity—e.g., when a helicopter can transition to a windmill or a tank to a shed, or vice-versa.

In a correct model, actual targets must transition only to actual targets and "clutter targets" only to "clutter targets." This issue is discussed at length in pp. 558-560 of [13].

D. iFilter: Other issues

The following claims were made in [50]:

1) Claim 1: The iFilter can always estimate the target-appearance rate $\hat{N}^B_{k|k-1}$ *: False*, as is shown by the following counterexample [16]. Suppose that there is no clutter, in which case there can be no "clutter targets." In this case $\dot{N}_{k-1|k-1}$—the expected number of "clutter targets"—must be zero, so that $\hat{N}^B_{k|k-1} = 0$ regardless of what $N^B_{k|k-1}$ might be.

2) Claim 2: The iFilter can always estimate the clutter rate λ_k *: False*, as is shown by the following counterexample. Assume that $p_D(\mathbf{x}) = p_D(\phi) = 1$ and that the probability of target disappearance is constant, $\psi_k(\phi|\mathbf{x}) = 1 - p_T$. Let $\int \psi_k(\mathbf{x}|\phi) d\mathbf{x} = 1 - \dot{p}_T$ be the total probability that a single target will appear. Finally, suppose that these happen to be "conjugate" in the sense that $p_T + \dot{p}_T = 1$. Then it is easily shown [16] that $\dot{N}_{k+1|k} = \dot{p}_T \cdot m_k$ and therefore that the estimated clutter rate is $\hat{\lambda}_k = \dot{N}_{k|k-1} = \dot{p}_T \cdot m_k$ for any k. That is, $\hat{\lambda}_k$ is a always a fixed fraction of the current number of measurements, regardless of what λ_k might actually be.

3) Claim 3: The PHD filter is a special case of the iFilter[31] *: False*, because the PHD filter time-update is not a special case of the iFilter time-update. As a counterexample, once again assume that there is no clutter and therefore no clutter generators. Since $\dot{N}_{k-1|k-1} = 0$, the iFilter target-appearance term in Eq. (80) vanishes—whereas the corresponding PHD filter term in Eq. (75) does not.

Furthermore, it is the reverse claim that is true: the iFilter is actually a PHD filter with a questionable state-transition model. Specifically, and has been shown in [16] or Section 631-636 and p. 644 of [13], the iFilter can—using only straightforward algebra—be directly derived as an ordinary PHD filter under the following assumptions: (i) the state space is $\mathfrak{X} \uplus \mathfrak{X}_\phi$; and (ii) the state-transition model on $\mathfrak{X} \uplus \mathfrak{X}_\phi$ is the questionable iFilter intermixing model.

Moreover, and as is shown in [30], [32], [6] and pp. 560-593 of [13], straightforward algebra can be used to derive a "κ-CPHD filter" that can estimate not only λ_k but also $c_k(\mathbf{z})$; as well as a "λ-CPHD filter" special case that estimates only λ_k.

[31]To wit: "Replacing the estimated clutter intensity with the a priori clutter intensity gives the PHD filter..." ([50], p. 6, 1st paragraph).

FISST p.g.fl. methods can additionally be applied to estimate the probability distribution $p^\kappa_k(m)$ on the number m of clutter measurements at time t_k (see [31] or p. 592 of [13]). This approach can be further extended to estimate $N^B_{k|k-1}$ and $s^B_{k|k-1}(\mathbf{x})$ in addition to λ_k, $c_k(\mathbf{z})$, and $p^\kappa_k(m)$ [17].

Finally, suppose that there are few target appearances or disappearances. Then it can be shown that the questionable intermixing state-transition model will be approximately disabled and that, as a consequence, the iFilter will behave like the λ-PHD filter special case of the λ-CPHD filter (see [16]).

4) Claim 4: "Intensity filters" can be understood in essentially elementary terms: False—even if the derivations of the PHD filter and "intensity filter" in [50] were not erroneous. Reverse-engineering of the PHD, CPHD, and other RFS filters has become something of a subspeciality—see pp. 200-201 of [13]. However, the following questions must be posed whenever some claim of engineering superiority vis-a-vis FISST is asserted or implied in any such exercise: Would its authors have been able to independently come up with the correct answer if they had not been able to (so to speak) look it up in the back of the FISST textbook? What does their method accomplish other than re-inventing the wheel?

In the specific case of the "elementary," "point process" approach: it has apparently not been possible to reverse-engineer the CPHD filter using only "single target Bayesian filtering and with PPP's at an elementary level"—or to devise a CPHD filter generalization of the iFilter.

XI. Appendix: The Multisensor iFilter

This section begins with a background summary of multisensor PHD and CPHD filters (Section XI-A). Then the multisensor version of the iFilter is described (Section XI-B), along with major mathematical issues in its derivation (Section XI) and performance issues (Section XI-D). The "multisensor traffic mapping filter" is discussed in Section XI-E), along with major mathematical (Section XI-F) and conceptual (Section XI-G) issues in its derivation.

A. Multisensor PHD and CPHD filters

The PHD and CPHD filters, like the iFilter, presume the existence of a single sensor. Shortly after its introduction, the PHD filter was heuristically extended by several authors to the multisensor case using the "iterated corrector" approach—i.e., by sequentially applying the PHD filter update equation, Eq. (76), once for each sensor. However, the output of this filter depends on sensor order. While this has little effect on performance when sensor probabilities of detection are approximately equal, performance can be significantly affected otherwise (see [34] or pp. 288-289 of [13]).

An exact multisensor PHD filter was introduced in 2009 (see [19] and pp. 283-287 of [13]), but is computationally challenging. Three "parallel combination" (PCAM) PHD/CPHD filters, which are computationally tractable and independent of sensor order, were introduced in 2010 (see [14] and pp. 289-300 of [13]). The simplest and least accurate of these is a "product pseudolikelihood" PHD filter. That is, the pseudolikelihoods of Eq. (77) are constructed for each sensor and then multiplied together to create a joint pseudolikelihood.[32]

B. Multisensor iFilter: Concept

At the same time that the iFilter was introduced in 2008, a "Bayesian" attempt was made to extend both it and the PHD filter to the multisensor case [45]. Specifically, the predicted PHD $D_{k|k-1}(\mathbf{x})$ is updated using each sensor, and the resulting PHDs are averaged. This is equivalent to an "averaged-pseudolikelihood" approach. That is, for the PHD filter, the pseudolikelihoods of Eq. (77) are constructed for each sensor and then averaged (rather than multiplied) to create a joint pseudolikelihood:

$$\frac{D_{k|k}(\mathbf{x})}{D_{k|k-1}(\mathbf{x})} = \overset{+}{L}_{\overset{1}{Z}_k,...,\overset{*}{\overset{n}{Z}}_k}(\mathbf{x}) = \frac{1}{\overset{*}{n}} \sum_{\ell=1}^{\overset{*}{n}} \overset{\ell}{L}_{\overset{\ell}{Z}_k}(\mathbf{x}) \tag{85}$$

where $\overset{*}{n}$ is the number of sensors; where $\overset{\ell}{Z}_k$ denotes the measurement-set collected at time t_k by the ℓ'th sensor; and where $\overset{\ell}{L}_{\overset{\ell}{Z}_k}(\mathbf{x})$ is the PHD filter pseudolikelihood of Eq. (76). An analogous equation holds for the iFilter.

C. Multisensor iFilter: Mathematical issues

As was first pointed out in 2009 [20], the averaged-pseudolikelihood approach appears to be conceptually questionable and, from a Bayesian point of view, mathematically erroneous. Its performance is also demonstrably worse than RFS multisensor PHD/CPHD filters. The discussion first appeared in [28] and is excerpted from pp. 300-309 of [13].

Consider the following special case: a single target is present and the sensors have no clutter and no missed detections. The predicted and updated PHDs are then probability density functions and Eq. (85) becomes

$$\frac{\overset{+}{f}_{k|k}(\mathbf{x})}{f_{k|k-1}(\mathbf{x})} = \frac{1}{\overset{*}{n}} \sum_{\ell=1}^{\overset{*}{n}} \frac{\overset{\ell}{L}_{\overset{\ell}{\mathbf{z}}}(\mathbf{x})}{\overset{\ell}{\tau}_k(\overset{\ell}{\mathbf{z}})} \tag{86}$$

where, given obvious notation,

$$\overset{\ell}{\tau}_k(\overset{\ell}{\mathbf{z}}) = \int \overset{\ell}{p}_D(\mathbf{x}) \cdot \overset{\ell}{L}_{\overset{\ell}{\mathbf{z}}}(\mathbf{x}) \cdot D_{k|k-1}(\mathbf{x}) d\mathbf{x}. \tag{87}$$

From this it is immediately follows that

$$\overset{+}{f}_{k|k}(\mathbf{x}) = \frac{1}{\overset{*}{n}} \sum_{\ell=1}^{\overset{*}{n}} f_{k|k}(\mathbf{x}|\overset{\ell}{\mathbf{z}}) \tag{88}$$

where $f_{k|k}(\mathbf{x}|\overset{\ell}{\mathbf{z}})$ is the posterior distribution updated using the measurement $\overset{\ell}{\mathbf{z}}$ from the ℓ'th sensor.

The following difficulties are apparent:

1) $\overset{+}{f}_{k|k}(\mathbf{x})$ is quite different than the Bayes-optimal solution—i.e., the multisensor, single-target version of Bayes' rule:

$$\overset{\times}{f}_{k|k}(\mathbf{x}) = \frac{\overset{1}{L}_{\overset{1}{\mathbf{z}}}(\mathbf{x}) \cdots \overset{*}{\overset{n}{L}}_{\overset{*}{\overset{n}{\mathbf{z}}}}(\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x})}{\int \overset{1}{L}_{\overset{1}{\mathbf{z}}}(\mathbf{y}) \cdots \overset{*}{\overset{n}{L}}_{\overset{*}{\overset{n}{\mathbf{z}}}}(\mathbf{y}) \cdot f_{k|k-1}(\mathbf{y}) d\mathbf{y}}. \tag{89}$$

2) Since the track distribution $\overset{+}{f}_{k|k}(\mathbf{x})$ is a mixture distribution, it will cause localization accuracy to *decrease* (as compared to the single-sensor track distributions $f_{k|k}(\mathbf{x}|\overset{\ell}{\mathbf{z}})$); and localization accuracy will decrease further with the number of sensors.

The second point is most easily demonstrated using a simple example: two bearing-only sensors in the plane, with respective likelihood functions

$$\overset{1}{L}_z(x,y) = N_{\sigma^2}(z-x) \tag{90}$$

$$\overset{2}{L}_z(x,y) = N_{\sigma^2}(z-y) \tag{91}$$

where $N_{\sigma^2}(z)$ denotes a one-dimensional Gaussian distribution with variance σ^2. That is, the sensors are oriented so as to triangulate the position of a target located at (x,y). For conceptual clarity, let the prior distribution be

$$f_{k|k-1}(x,y) = N_{\sigma_0^2}(x-x_0) \cdot N_{\sigma_0^2}(y-y_0) \tag{92}$$

where σ_0^2 is arbitrarily large, $\sigma_0^2 \to \infty$, in which case $f_{k|k-1}(x,y)$ is uniform. Let z_1, z_2 be the measurements collected by the sensors. Then Bayes' rule yields

$$\overset{\times}{f}_{k|k}(x,y) \cong N_{\sigma^2}(x-z_1) \cdot N_{\sigma^2}(y-z_2) \tag{93}$$

and results in a triangulated localization at (z_1, z_2) with variance $\cong 2\sigma^2$. But with the averaged likelihood,

$$\overset{+}{f}_{k|k}(x,y) \cong \frac{1}{2} \left[\begin{array}{c} N_{\sigma^2}(x-z_1) \cdot N_{\sigma_0^2}(y-y_0) \\ +N_{\sigma_0^2}(x-x_0) \cdot N_{\sigma^2}(y-z_2) \end{array} \right]. \tag{94}$$

This distribution has four "tails" whose lengths increase with the size of its variance, which is $\cong \sigma_0^2 \to \infty$. See pp. 302-307 of [13] for illustrations and details.

Now apply additional bearing-only sensors, all with orientations different from the first two and each other. The variance increases with the number of sensors—whereas it greatly decreases if the multisensor Bayes' rule is used instead.

D. Multisensor iFilter: Performance issues

These analytical assessments have been verified empirically—see [35] and pp. 306-309 of [13]. Specifically, Nagappa and Clark conducted simulations comparing multisensor PHD and CPHD filters such as those in Section XI-A.

In a first set of three-sensor simulations, two sensors had $p_D = 0.95$ and the third $p_D = 0.9$. In decreasing order of performance: PCAM-CPHD, PCAM-PHD, iterated-corrector CPHD, iterated-corrector PHD, averaged-pseudolikelihood PHD. The performance of the averaged-pseudolikelihood PHD filter was particularly poor, with the iterated-corrector PHD filter having intermediate performance. Similar results were observed when the probability of detection of the third sensor was decreased to $p_D = 0.85$ and again to $p_D = 0.7$.

[32]This multisensor PHD filter was originally proposed in 2003 as a heuristic [22, Eq. (106)].

E. The "multisensor traffic mapping filter"

In [45], the "multisensor multitarget intensity filter" was explicitly identified as a multitarget tracking filter. As previously noted, the theoretically problematic nature of this multisensor tracking filter was pointed out in 2009 [20] and its poor tracking performance in 2011 [35].

A year later, the following revelation appeared: the "multisensor multitarget intensity filter" in [45] had not only been "*misidentified* there as a multisensor target filter" (emphasis added), it was actually something quite different: a "multisensor traffic mapping filter" [45, p. 1]. Which is to say: it appears that the "elementary," "point process" derivation of the "multisensor multitarget intensity filter"—i.e., as a multitarget tracking filter—had to have been erroneous at a fundamental conceptual level.

The purpose of the "multisensor traffic mapping filter" is as follows: "estimate, or map, the mean rate at which different regions of a state space generate target detection opportunities in a field of distributed sensors" [45, p. 1]. It was further claimed to be "practical for applications with large numbers of sensors because [its] computational complexity is linear in the numbers of sensors and measurements" [45, p. 1].

The "multisensor traffic mapping filter" can be summarized as follows. Given time t_k, let $\overset{*}{n}_k$ be the number of sensors (abbreviated hereafter as $\overset{*}{n}$), which are assumed independent. Then, for $\ell = l, ..., \overset{*}{n}$, let: (1) $\overset{\ell}{p}_{D,k}(\mathbf{x}) =$ "probability that target [with state] $\mathbf{x}$ is detected by sensor ℓ" [45, p. 46]; (2) $\overset{\ell}{\kappa}_k(\overset{\ell}{\mathbf{z}})$ = clutter intensity function for sensor ℓ; and (3) $\overset{\ell}{f}_k(\overset{\ell}{\mathbf{z}}|\mathbf{x})$ = measurement density for sensor ℓ. Finally, let $\overset{\ell}{\beta}_k(\mathbf{x}) \geq 0$ be such that $\sum_{\ell=1}^{\overset{*}{n}} \overset{\ell}{\beta}_k(\mathbf{x}) = 1$ identically [45, Eq. (17)]. Here, $\overset{\ell}{\beta}_k(\mathbf{x})$ is a "field-level probability" that "represents the fraction of the total number of detection opportunities [at $\mathbf{x}$] for sensor ℓ," and is "very different" than the "sensor-level probability" $\overset{\ell}{p}_{D,k}(\mathbf{x})$ [45, p. 46].

The "multisensor traffic mapping filter" propagates a PHD $D_{k|k}(\mathbf{x})$, which is interpreted as the expected number of "detection opportunities" in an infinitesimal region surrounding $\mathbf{x}$ [45, Eq. (17)]. Suppose that the sensors collect measurement-sets $\overset{1}{Z}_k, ..., \overset{\overset{*}{n}}{Z}_k$ with $|\overset{\ell}{Z}_k| = \overset{\ell}{m}_k$. Then the updated intensity function is given by [45, Eq. (26)]:

$$\frac{D_{k|k}(\mathbf{x})}{D_{k|k-1}(\mathbf{x})} = \sum_{\ell=1}^{\overset{*}{n}} \overset{\ell}{\beta}_k(\mathbf{x}) \left(\begin{array}{c} 1 - \overset{\ell}{p}_{D,k}(\mathbf{x}) \\ + \sum_{\overset{\ell}{\mathbf{z}} \in \overset{\ell}{Z}_k} \frac{\overset{\ell}{p}_{D,k}(\mathbf{x}) \cdot \overset{\ell}{f}_k(\overset{\ell}{\mathbf{z}}|\mathbf{x})}{\overset{\ell}{\kappa}_k(\overset{\ell}{\mathbf{z}}) + \overset{\ell}{\tau}_k(\overset{\ell}{\mathbf{z}})} \end{array} \right) \tag{95}$$

where $\overset{\ell}{\tau}_k(\overset{\ell}{\mathbf{z}}) = \int \overset{\ell}{p}_{D,k}(\mathbf{x}) \cdot \overset{\ell}{f}_k(\overset{\ell}{\mathbf{z}}|\mathbf{x}) \cdot D_{k|k-1}(\mathbf{x}) d\mathbf{x}$. Eq. (95) reduces to Eq. (85) when $\overset{\ell}{\beta}_k(\mathbf{x}) = \overset{*}{n}^{-1}$ for all $\ell = 1, ..., \overset{*}{n}$.

F. "Traffic mapping filter": Mathematical issues

Eq. (95) is erroneous, unfortunately. As a consequence, the claimed computational linearity of the "multisensor traffic mapping filter" is spurious since Eq. (95) is not a theoretically valid measurement-update equation.

The proof of Eq. (95) was based on the following p.g.fl. identity (asserted without proof) [46, Eq. (20)]:

$$G[\overset{1}{g}, ..., \overset{\overset{*}{n}}{g}, h] = \overset{1}{G}[\overset{1}{g}, h] \cdots \overset{\overset{*}{n}}{G}[\overset{\overset{*}{n}}{g}, h] \tag{96}$$

where, for any i,

$$G[\overset{1}{g}, ..., \overset{i}{g}, h] = \int \overset{1}{g}^{\overset{1}{Z}} \cdots \overset{i}{g}^{\overset{i}{Z}} \cdot h^X \cdot f_k(\overset{1}{Z}, ..., \overset{i}{Z}|X) \cdot f_{k|k-1}(X) \delta\overset{1}{Z} \cdots \delta\overset{i}{Z} \delta X \tag{97}$$

is the p.g.fl. of the joint process $(\overset{1}{\Upsilon}, ..., \overset{i}{\Upsilon}, \Sigma)$ [45, p. 46] (where $\overset{j}{Z}$ is a realization of the random measurement-set $\overset{j}{\Upsilon}$ and X a realization of the random state-set Σ). In particular, for $i = 1$, $\overset{\ell}{G}[\overset{\ell}{g}, h]$ is assumed to be "the same as that of a PHD filter" [45, Eq. (19)], using the sensor ℓ. That is,

$$\overset{\ell}{G}[\overset{\ell}{g}, h] = \exp\left(\overset{\ell}{\kappa}_k[\overset{\ell}{g} - 1] + D_{k|k-1}[h(1 + \overset{\ell}{p}_D \overset{\ell}{L}_{\overset{\ell}{g}-1})] \right) \tag{98}$$

where $\overset{\ell}{L}_{\overset{\ell}{g}}(\mathbf{x}) = \int \overset{\ell}{g}(\overset{\ell}{\mathbf{z}}) \cdot \overset{\ell}{f}_k(\overset{\ell}{\mathbf{z}}|\mathbf{x}) d\overset{\ell}{\mathbf{z}}$. Also, given sensor independence,

$$f_k(\overset{1}{Z}, ..., \overset{i}{Z}|X) = \overset{1}{f}_k(\overset{1}{Z}|X) \cdots \overset{i}{f}_k(\overset{i}{Z}|X). \tag{99}$$

However, even given these "traffic process" assumptions, Eq. (96)—and thus Eq. (95)—is true only if $\overset{*}{n} = 1$ (i.e., only for a single sensor).

A detailed counterexample under such assumptions can be found in Appendix K.3 of [13]. However, a simpler counterexample is more informative. Assume that Eq. (96) is true. If so, it is true for the following special case: a single target, and two sensors with no missed detections or false alarms. Under these assumptions, Eq. (96) reduces to

$$\begin{aligned} & \int \overset{1}{g}(\overset{1}{\mathbf{z}}) \cdot \overset{2}{g}(\overset{2}{\mathbf{z}}) \cdot h(\mathbf{x}) \cdot f_k(\overset{1}{\mathbf{z}}|\mathbf{x}) \cdot f_k(\overset{2}{\mathbf{z}}|\mathbf{x}) \\ & \cdot f_{k|k-1}(\mathbf{x}) d\overset{1}{\mathbf{z}} d\overset{2}{\mathbf{z}} d\mathbf{x} \\ = & \left(\int \overset{1}{g}(\overset{1}{\mathbf{z}}) \cdot h(\mathbf{x}) \cdot f_k(\overset{1}{\mathbf{z}}|\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x}) d\overset{1}{\mathbf{z}} d\mathbf{x} \right) \\ & \cdot \left(\int \overset{2}{g}(\overset{2}{\mathbf{z}}) \cdot h(\mathbf{x}) \cdot f_k(\overset{2}{\mathbf{z}}|\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x}) d\overset{2}{\mathbf{z}} d\mathbf{x} \right). \end{aligned} \tag{100}$$

If this is true, it must be true for the following special case: $f_k(\overset{1}{\mathbf{z}}|\mathbf{x}) = f_k(\overset{1}{\mathbf{z}})$ and $f_k(\overset{2}{\mathbf{z}}|\mathbf{x}) = f_k(\overset{2}{\mathbf{z}})$ identically—in which case Eq. (96) reduces to

$$\int h(\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x}) d\mathbf{x} = \left(\int h(\mathbf{x}) \cdot f_{k|k-1}(\mathbf{x}) d\mathbf{x} \right)^2. \tag{101}$$

If this is true, it must be true when $h(\mathbf{x}) = 1/2$ identically—in which case Eq. (96) reduces to $1/2 = 1/4$, a contradiction.

The source of the error appears to have been a mistaken presumption that Eq. (96) is true because Eq. (99) is true.

G. "Traffic mapping filter": Conceptual issues

When closely examined, the "multisensor traffic mapping filter" appears to be a multitarget tracking filter in disguise.

First, it is not clear what real-world problem "traffic mapping" is meant to solve. The $\overset{\ell}{\beta}(\mathbf{x})$ are claimed to model "geometrical considerations concerning the point [$\mathbf{x}$] and the *entire* sensor field" [45, p. 46]. But what are these "geometrical considerations," and what is the precise physical meaning of $\overset{\ell}{\beta}(\mathbf{x})$ in regard to them? How might one construct a concrete formula for $\overset{\ell}{\beta}(\mathbf{x})$?

Second, the term "target detection opportunity" is never defined in a precise manner, mathematically or physically. We are informed only that such "opportunities" are "correlated with regions containing targets" [45, p. 48] and that they include "both detections and missed detections" [45, p. 45].

Given this, it is unclear why "target detection opportunity" is not just a synonym for "target." It is unclear how to count a missed detection as a "target detection opportunity," unless one has previously established the existence of a target, which should have been detected but was not. It is unclear how to count a detection as a "target detection opportunity," unless one has first determined that it was caused by an established (or new) target rather than clutter. That is: not only is a target a "target detection opportunity," one cannot have a "target detection opportunity" in the absence of a target. The two concepts appear to be synonymous.

References

[1] M. Baudin, "Multidimensional point processes and random closed sets," *J. Applied Prob*., 21: 173-178, 1984.

[2] A. Bozdogan, M. Efe, and R. Streit, "Reduced Palm intensity for track extraction," *Proc. 16th Int'l Conf. on Information Fusion*, pp. 1243-1250, Istanbul, Turkey, July 9-12, 2013.

[3] D. Clark and R. Mahler, "Generalized PHD filters via a general chain rule," *Proc. 15th Int'l Conf. on Information Fusion*, Singapore, July 9-12, 2012.

[4] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, First Edition, Springer-Verlag, New York, 1988.

[5] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume 1: Elementary Theory and Methods*, Springer-Verlag, New York, 2003.

[6] A. El-Fallah, A. Zatezalo, R. Mahler, R. Mehra, and W. Pereira, "Background agnostic CPHD tracking of dim targets in heavy clutter," *Proc. SPIE*, Vol. 8745, 2013.

[7] E. Engel and R. Dreizler, *Density Functional Theory*, Springer, 2011.

[8] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen, *Mathematics of Data Fusion*, Kluwer Academic Publishers, New York, 1997.

[9] T. L. Hill, *Statistical Mechanics: Principles and Practical Applications*, Dover Publications, 1987

[10] A. F. Karr, *Point Processes and Their Statistical Inference*, Second Edition, Marcel Dekker, New York, 1991.

[11] J. Kingman, *Poisson Processes*, Oxford University Press, London, 1992.

[12] M. Lighthill, *An Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, New York, 1958.

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

[14] R. Mahler, "Approximate multisensor CPHD and PHD filters," *Proc. 13th Int'l Conf. on Information Fusion*, Edinburgh, Scotland, July 26-29, 2010.

[15] R. Mahler, "A brief survey of advances in random-set fusion," *Proc. 2015 Int'l Conf. on Control, Automation and Information Sciences (ICCAIS2015)*, Changshu, China, Oct. 29-31, 2015.

[16] R. Mahler, "A comparison of 'clutter-agnostic' PHD/CPHD filters," *Proc. SPIE*, Vol. 8392, 2012.

[17] R. Mahler, "CPHD filters for unknown clutter and target-birth processes," *Proc. SPIE*, Vol. 9091, 2014.

[18] R. Mahler, "Linear-complexity CPHD filters," *Proc. 13^{th} Int'l Conf. on Information Fusion*, Edinburgh, Scotland, July 26-29, 2010.

[19] R. Mahler, "The multisensor PHD filter, I: General solution via multi-target calculus," *Proc. SPIE*, Vol. 7336, 2009.

[20] R. Mahler, "The multisensor PHD filter, II: Erroneous solution via 'Poisson magic,"' *Proc. SPIE*, Vol. 7336, 2009.

[21] R. Mahler, "Multisource, multitarget filtering: A unified approach," *Proc. SPIE*, Vol. 3373, pp. 296-307, 1998.

[22] R. Mahler, "Multitarget filtering via first-order multitarget moments," *IEEE Trans. Aerospace and Electronics Systems*, 39(4): 1152-1178, 2003.

[23] R. Mahler, "Multitarget moments and their application to multi-target tracking," *Proc. Workshop on Estimation, Tracking, and Fusion: A Tribute to Y. Bar-Shalom*, pp. 134-166, Naval Postgraduate School, Monterey CA, May 17, 2001. Available at: www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA414365.

[24] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerospace and Electronic Systems*, 43(4): 1523-1543, 2007.

[25] R. Mahler, "Sensor management with non-ideal sensor dynamics," *Proc. 7th Int'l Conf. on Inf. Fusion*, Stockholm, June 28-July 1, 2004.

[26] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, Norwood, MA, 2007.

[27] R. Mahler, "'Statistics 101' for Multisensor, Multitarget Data Fusion," *IEEE Aerospace & Electronics Sys. Mag., Part 2: Tutorials*, 19(1): 53-64, 2004.

[28] R. Mahler, "'Statistics 102' for multisensor-multitarget tracking," *IEEE J. Special Topics in Sign. Proc.*, 7(3): 376-389, 2013.

[29] R. Mahler, "Tracking multiple "bunching" target correlations," *Proc. IEEE Int'l Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI2015)*, San Diego, Sept. 14-16, 2015.

[30] R. Mahler and A. El-Fallah, "CPHD and PHD filters for unknown backgrounds, III: Tractable multitarget filtering in dynamic clutter," *Proc. SPIE*, Vol. 7698, 2010.

[31] R. Mahler and B.-T. Vo, "An improved CPHD filter for unknown clutter backgrounds," *Proc. SPIE*, Vol. 9091, 2014.

[32] R. Mahler, B.-T. Vo, and B.-N. Vo, "CPHD filtering with unknown clutter rate and detection profile," *IEEE Trans. Sign. Proc.*, 59(6): 3497-3513, 2011.

[33] J. E. Moyal, "The general theory of stochastic population processes," *Acta Mathematica*, 108: 1-31, 1962.

[34] S. Nagappa and D. Clark, "On the ordering of the sensors in the iterated-corrector probability hypothesis density (PHD) filter," *Proc. SPIE*, Vol. 8050, 2011.

[35] S. Nagappa, D. Clark, and R. Mahler, "Incorporating track uncertainty into the OSPA metric," *Proc. 14th Int'l Conf. on Information Fusion*, Chicago, July 5-8, 2011.

[36] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Aut. Contr.*, 24(6): 843-854, 1979.

[37] B. Ripley, "Locally finite random sets: foundations for point process theory," *Annals of Prob*., 4(6): 983-994, 1976.

[38] L. H. Ryder, *Quantum Field Theory*, 2nd Edition, Cambridge U. Press, 1996.

[39] G.E. Shilov and B.L. Gurevich, *Integral, Measure and Derivative: A Unified Approach*, Prentice-Hall, 1966.

[40] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, Second Edition, Springer, New York, 1991.

[41] L Stone, C. Barlow, and T. Corwin, *Bayesian Multiple Target Tracking*, First Edition, Artech House, 1999.

[42] D. Stoyan, W. S. Kendall, and J. Meche, *Stochastic Geometry and Its Applications*, Second Edition, John Wiley & Sons, 1995.

[43] R. Streit, "Birth and death in multitarget tracking filters," *Proc. 2013 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF2013)*, Bonn, Germany, Oct. 9-11, 2013.

[44] R. Streit, "Marked multitarget intensity filters," *Proc. 13^{th} Int'l Conf. on Information Fusion*, Edinburgh, Scotland, July 26-29, 2010.

[45] R. Streit, "Multisensor multitarget intensity filter," *Proc. 11th Int'l Conf. on Information Fusion*, pp. 1694-1701, Cologne, Germany, June 30-July 3, 2008.

[46] R. Streit, "Multisensor traffic mapping filters," *Proc. 7th Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, Bonn, Germany, Sept. 4-6, 2012.

[47] R. Streit, "Poisson Point Processes: Tracking, Imaging and Distributed Sensing," Tutorial No. 4, 15th Int'l Conf. on Information Fusion, Singapore, July 9, 2012.

[48] R. Streit, "The probability generating functional for finite point processes, and its application to the comparison of PHD and intensity filters," *Journal of Advances in Information Fusion*, 8(2): 119-132, 2014.

[49] R. Streit, "A technique for deriving multitarget intensity filters using ordinary derivatives," *J. Advances in Information Fusion*, 9(1): 1-10, 2014.

[50] R. Streit and L. Stone, "Bayes derivation of multitarget intensity filters," *Proc. 11th Int'l Conf. on Information Fusion*, pp. 1686-1693, Cologne, Germany, June 30-July 3, 2008.

[51] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*, (trans. M. Long), Blackie and Son, Ltd., London and Glasgow, 1930.