Arithmetic Average Density Fusion—Part II: Unified Derivation for Unlabeled and Labeled RFS Fusion

TIANCHENG LI, Senior Member, IEEE
Northwestern Polytechnical University, Xi’an, China

As a fundamental information fusion approach, the arithmetic average (AA) fusion has recently been investigated for various random finite set (RFS) filter fusions in the context of multisensor, multitarget tracking. It is not a straightforward extension of the ordinary density-AA fusion to the RFS distribution but has to preserve the form of the fusing multitarget density. In this article, we first propose a statistical concept, probability hypothesis density (PHD) consistency, and explain how it can be achieved by the PHD-AA fusion and lead to more accurate and robust detection and localization of the present targets. This forms a both theoretically sound and technically meaningful reason for performing interfilter PHD AA-fusion/consensus while preserving the form of the fusing RFS filter. Then, we derive and analyze the proper AA fusion formulations for most existing unlabeled/labeled RFS filters basing on the (labeled) PHD-AA/consistency. These derivations building on a single coherent framework are theoretically unified, exact, need no approximation, and greatly enable heterogenous unlabeled and labeled RFS density fusion which is separately demonstrated in companion papers.

1. INTRODUCTION

Multitarget tracking (MTT) that involves estimating both the number and states of multiple targets in the presence of clutter and misdetection has been a long standing research topic with a wide range of applications including surveillance, traffic control, and autonomous driving [1]. Of particular significance in real world is the scenario involving a time-varying, random perhaps, number of targets that may appear and disappear from the field of view of the sensors at any time in which the number and even identities of targets need to be estimated jointly with the states, namely joint detection and tracking. One of the most attractive approaches to this problem is the random finite set (RFS) theory [2], which can be dated back to the stochastic flows theory and has led to various probabilistic filters based on different scenario models and needs, such as the probability hypothesis density (PHD) filter [3], [4], the cardinalized PHD (CPHD) filter [5], the multi-Bernoulli (MB) filter [6], the Bernoulli filter [7], [8], the Poisson MB mixture (PMBM) filter [9], the generalized labeled MB (GLMB) filter [10], [11], and their variants.

The growing prevalence of wireless sensor networks [12], [13], [14] has attracted an increasing interest in robust multisensor MTT methods that can take advantage of large quantities of gathered information from multiple sensors for improved estimation accuracy, extended tracking coverage, and enhanced viability/scalability of the local sensors. Existing multisensor MTT approaches can be classified into two main groups differing in the type of the information shared between sensors: data level and estimate level [15]. Briefly speaking, the majority of the former is dedicated to compute the exact multisensor likelihood/posterior. It usually uses a centralized sensor network where the measurement data of all sensors are sent to the fusion center for joint, optimal fusion. This, however, suffers from poor computational scalability (with respect to the number of sensors) due to the intractable multisensor measurement-to-track association problem [16]. By contrast, the latter communicates and fuses the posterior estimates (given in means of various densities) produced by local filters, leading to the suboptimal but robust fusion result. This often realized in the distributed manner offers better computational scalability and system reliability, while the distributed fusion structure is more flexible and less prone to node failures as compared with the centralized fusion.

The so-called arithmetic/geometric average (AA/GA) fusion approach [17] is up to the task.

Based on the fundamental linear opinion pool [18], the AA fusion has an intimate connection with the mixture-type estimator [19], such as the celebrated Gaussian mixture (GM) filter, which is arguably the AA fusion of multiple Gaussian estimators. It is also in tune with the average consensus approach [12], [13], [14] suited for distributed implementation. However, its application for multisensor RFS density fusion, both centralized and distributed, has only appeared for a few years. Prior to this, the dominant
solutions were based on the Bayes or minimum-variance framework. As a non-Bayes method, the AA density fusion was first known due to its conservativeness (namely free of information double-counting regardless of the correlation between sensors) [20], [21], high efficiency for computing, and superiority to deal with misidentification [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39]. Recently, more properties of the AA/GA fusion have been analyzed in [40] and fast convergence in [41]. Nevertheless, the multitarget RFS density that involves the number, states, and even identities of the targets is not a simple extension of the classic single-target density. The RFS filters aim to solve the target detection, state estimation, and even data association jointly, for which the core need of effective fusion is the accuracy and robustness of the detection and localization of each target. Despite the many RFS-AA fusion formulas and flexible implementations [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39] proposed to accommodate diverse RFS densities, a theoretically unified analysis of these formulas is still missing. Existing derivations and reasoning for multisensor RFS density fusion [26], [29], [30], [31], [42] are mainly based on the best fit of the multitarget probability distribution (MPD). They endure the following three theoretical disadvantages and limitations.

1) These derivations rely on the MPD-best-fit property (see Section II-C), which does not directly relate to the real target distribution or the optimal multisensor posterior, and therefore, does not disclose how the linear fusion can deal with false/missing data and gain better accuracy in target detection and localization. Furthermore, it cannot explain why the GA fusion that has the similar MPD-best-fit property performs very differently in comparison with the AA fusion.

2) The MPD-best-fit needs approximation in general because the AA of most forms of RFS MPDs is no longer the same form. Each derivation is, therefore, only applicable to a specific form of MPD. They do not form an exact, unified framework suited for all RFS filters.

3) A complete definition of the divergence between labeled densities that needs to take into account the divergence of both the state distribution and the labels is still missing. Existing definitions, as well as the derivation based on them, take no account of the difference between matched labels—actually a metric for the label difference is still missing—and apply only when the number of labels is the same among the fusing densities.

To overcome the aforementioned limitations and disadvantages, this article contributes in the following aspects.

1) We propose a statistical concept, PHD consistency, which lays a technically unified, intuitively meaningful foundation for (labeled) RFS fusion to explain how the PHD-AA fusion can deal with false/missing data and improve accuracy in target detection and localization.

2) We derive appropriate AA fusion formulas for different (labeled) RFS filters starting from the (labeled) PHD-AA fusion, which all maintain exactly AA calculation in terms of the relevant (labeled) PHDs and ensure PHD consistency. Although the final fusion results turn out to coincide with those separately given in [26], [29], [30], [31], and [42], our derivation is based on an exact, unified framework that is suited for all RFS filters. This unified framework paves a flexible way for either homogeneous or heterogeneous RFS filter fusion; see [43]. It also exposes the essence of existing consensus-driven multisensor RFS filters, which merely seek consensus over their (labeled) PHD rather than over the MPDs, while the MPD-best-fit derivation implies the consensus on the level of MPD overoptimistically.

3) Aiming at labeled PHD consistency, the averaging of labeled PHDs that are functions on single target, single label is appropriate for the general case in which the fusing filters have different numbers of labels, overcoming the deficiency of the labeled MPD-best-fit derivation.

4) Derived from the labeled PHD-AA fusion, the exact AA fusion of the general (δ-)GLMB [10], [11] is given for the first time, which can be easily tailored in more specific forms to accommodate the marginalized GLMB (M-GLMB) [44] and the labeled MB (LMB) [45].

This article is the second part of a series of papers that aim to provide a comprehensive and thorough study of the AA fusion methodology and its application for target tracking. The first is on the fundamental Kalman/particle filter fusion [21], while the third is on advanced heterogeneous unlabeled and labeled RFS filter fusion [43] based on the finding of this work.

The rest of this article is organized as follows. The representative RFS processes and the AA fusion approach are briefly introduced in Section II. The concept of PHD consistency is formally defined in Section III where the suitable AA fusion formulas for PHD and CPHD filters are consequently derived. The unlabeled Bernoulli, MB and MB mixture (MBM), and the labeled GLMB, M-GLMB, and LMB fusion are addressed in Sections IV and V, respectively. Finally, Section VI concludes this article.

II. BACKGROUND AND PRELIMINARY

The essential goal of multisensor cooperative target tracking is to get more accurate and robust detection and localization of the present targets based on the information exchange and fusion among the sensors. In this work, we focus on the estimate-level fusion where what is fused is the first-order moments of the MPDs, namely the PHD, individually calculated by the local off-the-shelf RFS filters. For brevity, the details of local single-sensor filters and of
the intersensor communication are fully omitted. We further omit the time notation and the dependence of the filter estimate on the observation process. Sensors are assumed to be synchronous and perfectly coordinated.

A. RFS Theory

Denote by $\mathcal{X} \subseteq \mathbb{R}^d$ the $d$-dimensional state space and by $\mathcal{X}$ all of the finite subsets of $\mathcal{X}$. The states of a random number of targets are described by an RFS $X = \{x_1, \ldots, x_n\} \subseteq \mathcal{X}$, where $n = |X|$ denotes the random number of targets, namely the cardinality of the set, and $x_i \in \mathcal{X}$ is the state vector of the $i$th target. The random nature of the multitarget RFS $X$ is captured by its MPD, denoted by $f(X)$. For any realization of $X$ with a given cardinality $|X| = n$, denoted by $X_n = \{x_1, \ldots, x_n\}$, [2, eq. 2.36]

$$f(X_n) = n! \rho(n) f(x_1, \ldots, x_n)$$

where the localization densities $f(x_1, \ldots, x_n)$ for $n = 1, 2, \ldots$, are symmetric in their arguments, and the cardinality distribution $\rho(n) \triangleq \Pr(|X| = n) = \int_{|X|=n} f(X) \delta X$ is given by

$$\rho(n) = \frac{1}{n!} \int_{\mathcal{X}^n} f(\{x_1, \ldots, x_n\}) \, dx_1 \ldots dx_n. \quad (2)$$

The set integral in $\mathcal{X}$ is defined as [2, ch. 3.3]

$$\int_{\mathcal{X}} f(X) \delta X = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{X}^n} f(\{x_1, \ldots, x_n\}) \, dx_1 \ldots dx_n = \sum_{n=0}^{\infty} \rho(n)$$

$$= 1 \quad (4)$$

where $f(\emptyset) = \rho(0)$.

The PHD $D(x)$, also known as the first moment density [46, pp. 168 and 169] or the intensity, of the MPD $f(X)$ is a density function on single target $x \in \mathcal{X}$, defined as [2, ch. 4.2.8]

$$D(x) \triangleq \int_{\mathcal{X}} \left( \sum_{y \in \mathcal{X}} \delta_y(x) \right) f(X) \delta X \quad (5)$$

where $\delta_y(x)$ is the Dirac delta function concentrated at $y$. The PHD has clear physical significance as its integral in any region $S \subseteq \mathcal{X}$ gives the expected number $N_S$ of targets in that region, i.e.,

$$N_S = \int_S D(x) \, dx. \quad (6)$$

Therefore, it tells how well the present targets are detected.

B. Some Important RFSs

1) Independent and Identically Distributed Cluster (IIDC): An RFS $X$ is considered as an IIDC RFS if its cardinality is arbitrarily distributed according to $\rho(n)$, and for any finite cardinality $n$, the elements are independently and identically distributed according to the single-target probability density (SPD) $s(x)$. The MPD and PHD associated with the IIDC RFS $X$ are, respectively, given by

$$f^{\text{iidc}}(X_n) = n! \rho(n) \prod_{x \in X_n} s(x) \quad (7)$$

$$D^{\text{iidc}}(x) = \sum_{n=0}^{\infty} n \rho(n) s(x). \quad (8)$$

The IIDC RFS reduces to a Poisson RFS when the cardinality distribution $\rho(n)$ is Poisson. The MPD and PHD of the Poisson RFS $X_n$ with mean $\lambda$ are, respectively, given by

$$f^p(X) = e^{-\lambda} \prod_{x \in X} \lambda s(x) \quad (9)$$

$$D^p(x) = \lambda s(x). \quad (10)$$

2) Bernoulli and MB: The density and PHD of a Bernoulli RFS $X$ with target existence probability $r$ and SPD $s(x)$ are, respectively, given by

$$f^b(X) = \begin{cases} 1-r & X = \emptyset \\ rs(x) & X = \{x\} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

$$D^b(x) = rs(x). \quad (12)$$

An MB RFS $X_n$ is the union of $n$ independent Bernoulli RFSs [6]. Denoting the $l$th Bernoulli component (BC) by $(r_l, s_l(x))$, the MPD and PHD of MB RFS $X_n$ are, respectively, given by

$$f^{mb}(X_n) = \sum_{\omega_{l, l}^{(n)}} \prod_{l=1}^{n} f^b(X^{(l)}) \quad (13)$$

$$D^{mb}(x) = \sum_{l=1}^{n} r_l s_l(x) \quad (14)$$

where $\omega_{l, l}^{(n)}$ denotes the disjoint union. When multiple MBs are used jointly and associated with different measurement-to-track hypotheses or labels, they result in the PMBM [9], the MBM [47], and the GLMB [10, 11], which are all conjugate prior for multitarget filtering.

3) Poisson MB Mixture (PMBM): An PMBM RFS [9] is the convolution of a Poisson RFS and an MBM RFS with MPD

$$f^{mbm}(X) = \sum_{Y \subseteq W = X} f^p(Y) f^{mbm}(W) \quad (15)$$

where the Poisson density $f^p(\cdot)$ represents all undetected targets, and the MBM density $f^{mbm}(\cdot)$ representing potentially detected targets is a weighted sum of MBs and has the form

$$f^{mbm}(X_n) \propto \sum_{j \in \mathcal{J}} \sum_{l=1}^{n} \prod_{l=1}^{n} w_{j,l} f_{j,l}^b(X^{(l)}) \quad (16)$$

where $j$ is an index to the MBs (each corresponding to a global hypothesis), $\mathcal{J}$ is the index set, and $w_{j,l}$ is a coefficient/weight assigned to the $l$th BC in the $j$th MB.
The PHD of the MBM RFS $X_n$ of the MPD (16) is

$$L^{\text{pmbm}}(x) = \sum_{j \in L} \sum_{l=1}^{n} w_{j,l} r_{j,l}(x)$$  (17)

where $(r_{j,l}, s_{j,l}(x))$ denote the parameters of the $l$th BC in the $j$th hypothesis/MB.

As a special PMBM, the Poisson MB (PMB) [48] uses the MB instead of MBM for representing detected targets.

4) Generalized Labeled MB (GLMB): Denote by $\mathcal{L}$ the label space and by $L$ all of the finite subsets of $\mathcal{L}$. A labeled RFS (LRFS) is an RFS whose elements are assigned distinct, discrete labels [10], [11]. A realization of an LRFS with cardinality $n$, multitarget state $X_n$ and label set $L_n = \{l_1, l_2, \ldots, l_n\} \subseteq \mathcal{L}$ is denoted by $\tilde{X}_n = \{(x_1, l_1), (x_2, l_2), \ldots, (x_n, l_n)\} \subseteq \mathbb{X} \times L$. The LRFS is completely characterized by its MPD $\pi(\tilde{X})$. The most-known LRFS density, namely the GLMB, is a mixture of multitarget exponentials with MPD [10], [11]

$$\pi^\delta(\tilde{X}) = \Delta(\tilde{X}) \sum_{c \in C} \omega^c(\mathcal{L}(\tilde{X})) \prod_{(x, l) \in \tilde{X}} s^c(x, l)$$  (18)

where the distinct label indicator $\Delta(\tilde{X}) = \delta_{\tilde{X}}[\mathcal{L}(\tilde{X})]$, $C$ is a discrete index set, the projection function $\mathcal{L}(\mathcal{X} \times L) \rightarrow \mathcal{L}$ is given by $\mathcal{L}((x, l)) = l$, and $\omega^c(L_n)$ is the weight of the hypothesis that $|\tilde{X}| = n$ targets are presented with respective labels $l_1, \ldots, l_n \in L_n$ and respective labeled SPDs $s^c(x, l_1), \ldots, s^c(x, l_n)$, satisfying

$$\sum_{L \subseteq \mathcal{L}} \sum_{c \in C} \omega^c(L) = 1, \int_{\mathcal{X}} s^c(x, l) \, dx = 1.$$  

The labeled PHD (LPHD) of an GLMB is given by [10]

$$\tilde{D}^\delta(\mathbf{x}, l) \equiv \int_{\mathcal{X} \times L} \pi \left( \left. ((x, l) \cup \tilde{X}) \right| \delta \tilde{X} \right)$$

$$= \sum_{c \in C} \sum_{L \subseteq \mathcal{L}} \omega^c(L) \int_{\mathcal{X} \times L} 1_L(l) \omega^c(L)$$  (20)

where $1_x(y) = 1$ if $y \subseteq x$ and $1_x(y) = 0$ otherwise.

A $\delta$-GLMB RFS is a special GLMB RFS with $C = \mathcal{L} \times \Xi$, $\omega^c(L') = \omega^c(\mathcal{L})$, and $s^c(x, l) = s^c(L', l)$, where the pair $(L, \xi)$ represents the hypothesis that the track set $L$ has a history $\xi$ of measurement-track association maps [11]. The MPD of a $\delta$-GLMB RFS $\tilde{X}$ is

$$\pi^\delta(\tilde{X}) = \Delta(\tilde{X}) \sum_{L \subseteq \mathcal{L}} \delta_L(\mathcal{L}(\tilde{X})) \prod_{(x, l) \in \tilde{X}} s^L(x, l).$$  

The LPHD of the $\delta$-GLMB is given as follows [10]:

$$\tilde{D}(\mathbf{x}, l) \equiv \int_{\mathcal{X} \times L} \pi \left( \left. ((x, l) \cup \tilde{X}) \right| \delta \tilde{X} \right)$$

$$= \sum_{L \subseteq \mathcal{L}} 1_L(l) \sum_{\xi \in \Xi} \omega^L(\xi) \prod_{(x, l) \in \tilde{X}} s^L(x, l).$$  (22)

C. Ordinary AA Density Fusion

We briefly review the AA fusion in terms of the ordinary probability density function (PDF). For a set of PDFs $f_i(x)$ produced by estimator $i \in I$ where $I = \{1, 2, \ldots, I\}$ denotes the set of sensors, the AA fusion is simply given as follows:

$$f_{AA}(x) \equiv \sum_{i \in I} w_i f_i(x)$$  (23)

where the fusion weights $w = \{w_1, \ldots, w_I\} \in \mathbb{W}$, and weight space $\mathbb{W} \equiv \{w \in \mathbb{R}^I | w^T 1_I = 1, w_i > 0 \forall i \in I\} \subset \mathbb{R}^I$.

The AA fusion greatly ensures the desired fusion closure for these mixture densities. Another important property of the AA PDF fusion is that the expectation $E_{f_{AA}(x; w)} H(x)$ of an arbitrary function $H(x)$ of $x$ with respect to the AA-fused density $f_{AA}(x; w)$ is given as follows:

$$E_{f_{AA}(x; w)} H(x) = \sum_{i \in I} w_i E_{f_i(x)} H(x).$$  (24)

Arbitrary moments of $x$ can be obtained for appropriate choices of $H(x)$. When $H(x) = x$, (24) reads the unbiasedness of the AA fusion, i.e., the AA of a number of unbiased estimators is guaranteed to be unbiased. It was further pointed out that the AA is a Fréchet mean in the sense of $L_2$ norm error [28], i.e.,

$$f_{AA}(x) = \arg \min_{g \in \mathcal{F}_I} \sum_{i \in I} w_i \int_{\mathcal{X}} (f_i(x) - g(x))^2 \, dx$$  (25)

where $\mathcal{F}_I$ denotes the set of scalar-valued functions in space $\mathcal{X}$: $\mathcal{F}_I = \{f : \mathcal{X} \rightarrow \mathbb{R}\}$.

The AA fusion minimizes the weighted sum of the Kullback–Leibler (KL) divergences of the fusing densities relative to the fused result [21], [49], i.e.,

$$f_{AA}(x) = \arg \min_{g \in \mathcal{F}_I} \sum_{i \in I} w_i D_{KL}(f_i || g)$$  (26)

where the KL divergence $D_{KL}(f || g) = \int_{\mathcal{X}} f(x) \log \frac{f(x)}{g(x)} \, dx$ in the case of continuous functions and $D_{KL}(f || g) = \sum_{n=0}^{\infty} f(n) \log \frac{f(n)}{g(n)}$ in the case of discrete functions.

The aforementioned optimization, also known as minimum information loss (MIL) fusion [29], [31], provides a best fit of the mixture (BfOM) of all fusing distributions with weights, $\{w_i f_i\}_{i \in I}$. Considering that the mixture $\{f_i\}_{i \in I}$ contains the information obtained by all sensors, it is a reasonable substitute of “all information.” This approach best fits the mixture but not the true distribution or the multisensor optimal distribution and so the BfOM/MIL itself does not guarantee any gain in the estimation accuracy or robustness. This is easily evident. By turning the direction of the KL divergence in (26), the minimization will lead to the GA fusion [42], [49], which is significantly different with the AA fusion, i.e.,

$$f_{GA}(x) = \arg \min_{g \in \mathcal{F}_I, g(x) \geq 0} \sum_{i \in I} w_i D_{KL}(f_i || g)$$  (27)

$$= \frac{1}{c} \prod_{i \in I} \tilde{f}_i(x)$$  (28)

denotes the set of sensors, $c = \int_{\mathcal{X}} \prod_{i \in I} f_i(x) \, dx$ is a normalization constant.
However, given the true distribution \( p(x) \), the optimal fusing weights should be given by

\[
\mathbf{w}_{\text{opt}} = \arg \min_{\mathbf{w} \in \mathbb{W}} D_{\text{KL}}(f_{\text{AA}} \| p) = \arg \min_{\mathbf{w} \in \mathbb{W}} \sum_{i \in I} w_i \left( D_{\text{KL}}(f_i \| p) - D_{\text{KL}}(f_i \| f_{\text{AA}}) \right). \tag{29}
\]

This formulation is solidly motivated since the true distribution \( p(x) \) is referred to, which is a fundamentally different to \( (26) \) and \( (27) \). However, the true distribution \( p(x) \) is always unknown. A practically operable alternative is given by ignoring the former part in \( (29) \) and solving the following suboptimal, diversity preference, maximization problem \cite{21}:

\[
\mathbf{w}_{\text{BFoM}} = \arg \max_{\mathbf{w} \in \mathbb{W}} \sum_{i \in I} w_i D_{\text{KL}}(f_i \| f_{\text{AA}}). \tag{30}
\]

Combining \( (26) \) with \( (30) \) yields a maxmin formulation for joint optimization of the fusing form and fusing weights

\[
(\mathbf{w}_{\text{BFoM}}, f_{\text{AA}}) = \arg \max_{\mathbf{w} \in \mathbb{W}} \arg \min_{g \in \mathcal{F}_X} \sum_{i \in I} w_i D_{\text{KL}}(f_i \| g) \tag{31}
\]

which is hereafter referred to as the BFoM optimization.

We note that the aforementioned BFoM formulation is appropriate for probability density but not the labeled densities in general; see Section V-A.

III. PHD-CONSISTENCY AND PHD-AA FUSION

An intuitive idea to resolve RFS density fusion is to extend the PDF-AA fusion to the MPD domain for which, however, the major barrier is from the fact that most RFS MPDs do not admit closed-form AA fusion, e.g., the AA of IIDC/sMBs is no more an IIDC/MB \cite{28,33}. This attempt at existing derivations \cite{26,29,30,31,42} is actually undesirable for the local filters that need to maintain a closed-form filter iteration. To fulfill the need of preserving the form of the fusion MPDs, ad hoc approximation is often needed for fitting the MPD \cite{29,30,31,42}. That being said, only homogeneous MPDs are able to be fused in this way. To overcome this problem, we resort to fusion in terms of the first moment of the MPD rather than the MPD itself. In this section, we first derive the PHD-AA fusion from the MPD-AA fusion. Then, a statistical concept, PHD consistency, is proposed that explains how the PHD-AA fusion can combat false/missing data and improve the accuracy in target detection and localization. Finally, the proper, exact formulation of the AA fusion of the PHD filter and of the CPHD filter is derived.

A. From MPD-AA to Cardinality-AA and PHD-AA

The MPD-AA fusion is given by

\[
f_{\text{AA}}(X) \triangleq \sum_{i \in I} w_i f_i(X) \tag{32}
\]

which corresponds to the following BFoM, c.f., \( (26) \),

\[
f_{\text{AA}}(X) = \arg \min_{g \in \mathcal{F}_X} \sum_{i \in I} w_i D_{\text{KL}}(f_i \| g) \tag{33}
\]

where \( \mathcal{F}_X \triangleq \{ f : \mathbb{X} \to \mathbb{R} \} \) and the set KL divergence is defined as

\[
D_{\text{KL}}(f \| g) \triangleq \int_X f(x) \log \frac{f(x)}{g(x)} \delta x. \tag{34}
\]

Notably, according to the definition of the PHD \( (5) \), the MPD-AA fusion \( (32) \) implies cardinality-AA fusion and PHD-AA fusion, respectively, i.e.,

\[
\rho_{\text{MPD-AA}}(n) = \int_{|X|=n} f_{\text{AA}}(X) \delta X = \sum_{i \in I} w_i \int_{|X|=n} f_i(X) \delta X = \sum_{i \in I} w_i \rho_i(n) \triangleq \rho_{\text{AA}}(n) \tag{35}
\]

\[
D_{\text{MPD-AA}}(x) = \int_x \left( \sum_{w \in \mathbb{W}} \delta_w(x) \right) f_{\text{AA}}(X) \delta X = \sum_{i \in I} w_i D_i(x) \triangleq D_{\text{AA}}(x) \tag{36}
\]

which corresponds to the following BFoM, c.f., \( (33) \),

\[
\rho_{\text{AA}}(n) = \arg \min_{\rho \in \mathbb{R}^N} \sum_{i \in I} w_i D_{\text{KL}}(\rho_i \| \rho) \tag{37}
\]

\[
D_{\text{AA}}(x) = \arg \min_{g \in \mathcal{F}_X} \sum_{i \in I} w_i D_{\text{KL}}(D_i \| g) \tag{38}
\]

s.t. \( \int_X g(x) dx = \sum_{i \in I} w_i \int_X D_i(x) dx \). \tag{39}

While \( (37) \) is directly from \( (26) \), the minimization in \( (38) \) that extends the definition of the KL divergence to the PHD domain is subject to the cardinality average constraint \( (39) \). A mathematical justification of \( (38) \) and \( (39) \) is given in the Appendix.

The cardinality and PHD are of clear physical significance and their average calculation has obvious benefit as explained next. Even more importantly, the PHD is not limited to any specific RFS density forms as the MPD fusion does, and so, the PHD-AA fusion can be straightforwardly applied to any RFS filters. Obviously, the MPD average/consensus implies the PHD average/consensus but not vice versa.

B. Motivation for PHD-AA Fusion

The PHD is a key statistical character of the RFS distribution, and so an accurate PHD estimate is vital in any RFS
filters. On this point, we present the following definition relative to the effectiveness of the multisensor fusion.

**Definition 1 (PHD Consistency)** The accuracy of the multisensor-fused PHD increases with the number of fusing estimators. That is, the PHD consistent fusion should result in statistically more accurate detection of the targets and the fused PHD converges to the first moment of the true MPD as the number of the fusing estimators increases indefinitely.

We note that our aforementioned definition abides with the standard definition of the consistency for parameter estimation, i.e., as the number of data points used increases indefinitely, the resulting sequence of estimates converges in probability to the ground truth. This is different from the concept given in [50] and [51], which means consensus/agreement.

**Lemma 1** Given that the estimates of the number of targets yielded by the fusing estimators are conditionally independent with each other and unbiased everywhere in the state space, the PHD-AA fusion (36) basing on the Bayes inference guarantees PHD consistency.

**Proof** Denoting by $\text{MSE}_S$ the mean-square-error (MSE) of the estimated number $\hat{N}_i^S$ of targets in any region $S \subseteq \mathcal{X}$ obtained by the filter $i \in \mathcal{I}$, the estimated number of targets given by the integral of the AA-fused PHD in $S$ is, c.f. (24)

$$\hat{N}^S_{\text{PHD-AA}} = \int_S D_{AA}(x)dx = \int_S \sum_{i \in I} w_i D_i(x)dx = \sum_{i \in I} w_i \hat{N}_i^S.$$  

The PHD consistency is ensured in the following sense:

$$\text{MSE}_{\text{PHD-AA}}^S = \mathbb{E} \left[ (\hat{N}^S_{\text{PHD-AA}} - N^S)^2 \right] = \mathbb{E} \left[ \left( \sum_{i \in I} w_i \hat{N}_i^S - N^S \right)^2 \right] = \sum_{i \in I} w_i^2 \text{MSE}_i^S$$

$$= \sum_{i \in I} w_i^2 \text{MSE}_i^S$$

where expectation $\mathbb{E}[\cdot]$ is conditional on the observation, $N^S$ denotes the real number of targets in region $S$, and conditional independence and unbiasedness of $\hat{N}_i^S$, $\forall i \in \mathcal{I}$ were used in (40).

By using $w_i = \frac{1}{I} \forall i \in \mathcal{I}$, (41) reduces to

$$\text{MSE}_{AA}^S = \frac{1}{I} \sum_{i \in I} w_i \text{MSE}_i^S$$

$$\leq \frac{1}{I} \max_{i \in I} \{ \text{MSE}_i^S \}$$

where the equation holds iff $\text{MSE}_i^S = \text{MSE}^S_{\text{PHD-AA}} \forall i, j \in \mathcal{I}$ and the consistency is indicated by

$$\lim_{I \to \infty} \Pr \left[ \text{MSE}_{\text{PHD-AA}}^S > \varepsilon \right] = 0$$

for every $\varepsilon > 0$

The PHD-AA fusion can be illustrated in Fig. 1. Beyond the PHD/CPHD filters that directly calculate the PHD in time series, many RFS/LRFS filters are obtained based on PHD matching, e.g., [33], [44], [45], [52], [53], and [54]. The PHD accuracy plays the same important role in these filters. In addition, the cardinality-AA fusion as illustrated in Fig. 2 has the similar property.

**Corollary 1** Assuming that the fusing cardinality estimates $\rho(n)$ are conditionally independent with each other and unbiased $\forall n = 0, 1, 2, \ldots$, the AA-fused cardinality (35) with properly designed fusing weights will gain more accurate estimate in the sense of smaller MSE.

**Proof** The proof is similar to that for Lemma 1. Denote by $\text{MSE}_i^{(n)}$ the MSE of the estimated probability $\rho_i(n)$ of $n$ targets in total obtained by the filter $i \in \mathcal{I}$. Given that $\rho_i(n), i \in \mathcal{I}$ are conditionally statistically independent with each other and unbiased, the cardinality-AA fusion leads to

$$\text{MSE}_{AA}^{(n)} = \mathbb{E} \left[ (\rho_{AA}(n) - \text{Pr}[N = n])^2 \right] = \mathbb{E} \left[ \left( \sum_{i \in I} w_i \rho_i(n) - \text{Pr}[N = n] \right)^2 \right] = \sum_{i \in I} w_i^2 \text{MSE}_i^{(n)}$$

Fig. 1. PHD-AA fusion will lead to statistically more accurate and robust detection of the present targets when the estimate of each filter is statistically unbiased and independent on each other. $\hat{N}_i^S$ denotes the estimated number of targets in the region $S$ yielded by the RFS filter $i = 1, 2, 3$.

Fig. 2. Cardinality-AA fusion will lead to statistically more accurate estimate of the number of targets when the cardinality estimate of each filter is statistically unbiased and independent on each other, $\rho_i$ denotes the estimated cardinality probability mass function yielded by the RFS filter $i = 1, 2, 3$.
where $\Pr[N = n]$ denotes the probability of $n$ targets in total. Then, similar result as (44) can be obtained.

**Remark 1** In addition to the consistency in terms of the estimator accuracy, the AA fusion can reduce the effects of clutter and missed detections [24, 25] since the mixing operation and averaging calculation are insensitive to false and missing data, improving the robustness of the local filters. In the case of low detection profile, the AA fusion has a significant advantage as compared with the Bayes/multiplication-based fusion approaches. This, however, cannot be explained by the BFoM/MIL formulation. Simply, the GA fusion that minimizes similar divergence (27) has, however, been proven vulnerable to local misde-
namination [22, 23, 24, 25, 29, 31, 33, 38] and may even underestimate the number of targets [51, 55, 56].

C. PHD Filter Fusion

By assuming a Poisson RFS formulation of the multitarget states as in (9) where the Poisson parameter $\lambda = \int_X D_{\text{AA}}(x)dx$, the PHD filter [3, 4] recursively propagates the PHD in lieu of the Poisson MPD itself. As such, the PHD-AA fusion is very appropriate for the PHD filter fusion, where the resulting MPD is uniquely determined by the AA-fused PHD

$$f_{\text{PHD-AA}}^p(X) = e^{-\hat{\lambda}_{\text{PHD-AA}}} \prod_{x \in X} D_{\text{AA}}(x).$$

(47)

The PHD-AA fusion (36) for the PHD filter is consistent with the IIDC-AA fusion given in Lemma 2, which also amounts to averaging the Poisson parameters, i.e.,

$$\hat{\lambda}_{\text{PHD-AA}} = \int_X D_{\text{AA}}(x)dx = \sum_{i \in I} w_i \hat{\lambda}_i.$$  

(48)

The PHD filter is actually the first testbed for applying the AA fusion in the context of multisensor RFS density fusion. A pioneering attempt of adopting averaging-for-fusion in the context of RFS filters was proposed with respect to the multitarget likelihood [57], which actually leads to isovariance results as averaging the posteriors if the priors are the same. Based on the average consensus approach and distributed peer-to-peer networks [21, 56], the PHD-AA fusion has been proposed in [23, 24, 25, and 34] based on GM and particles implementation, respectively. Greatly, both can be viewed as a mixture distribution and allows greatly for closed-form GM-PHD-AA and particle-PHD-AA fusion, respectively. For extremely efficient communication and fusion-computation, one may resort to the cardinality-only fusion [58], which shares and fuses only the cardinality (variables in the case of PHD filters and functions in the case of CPHD filters).

Following the same line of thinking, the CPHD filter [5, 59] that recursively propagates the PHD and cardinality distribution can be fused in means of cardinality and PHD fusion, respectively [22, 42]. This coincides with the result of the IIDC-AA fusion given in Lemma 2. Further issues such as limited field of view and asynchronous sensing have been addressed in [34] and [35] and in [36] and [37], respectively.

D. IIDC PHD-AA Fusion

The IIDC (7) point process, which is the backbone of the CPHD filter [5, 59] is completely specified by the cardinality distribution $p(n)$ and the SPD $s(x)$. We have the following important IIDC PHD-AA fusion result.

**Lemma 2** For IIDC MPDs $f_{\text{iidc}}^i(x_i) = n! p_i(n) \prod_{i=1}^n s_i(x_i), \ i \in I$, the cardinality-AA and PHD-AA fusion (36) using fusion weights $w \in W$ result in

$$f_{\text{PHD-AA}}^i(x_i) = n! p_{\text{AA}}(n) \prod_{i=1}^n s_{\text{PHD-AA}}(x_i)$$

(49)

with AA-fused cardinality distribution $p_{\text{AA}}(n)$ as given in (35), and the cardinalized-AA-fused SPD as follows:

$$s_{\text{PHD-AA}}(x) = \frac{1}{\hat{N}_{\text{AA}}} \sum_{i \in I} w_i \hat{N}_i s_i(x)$$

(50)

**Proof** The cardinality-AA results straightforwardly the AA-fused cardinality distribution $p_{\text{AA}}(n)$. The PHD-AA fusion of the IIDCs can be expressed as follows:

$$D_{\text{AA}}^i(x) = \sum_{i \in I} w_i D_i^i(x).$$

(51)

Since the SPD is the normalized PHD, one can easily get (50) from (51), i.e.,

$$s_{\text{PHD-AA}}(x) = \frac{\int_X D_{\text{AA}}^i(x)dx}{\int_X D_{\text{AA}}^i(x)dx} = \frac{1}{\hat{N}_{\text{AA}}} \sum_{i \in I} w_i \int_X D_i^i(x)dx = \frac{1}{\hat{N}_{\text{AA}}} \sum_{i \in I} w_i \hat{N}_i s_i(x)$$

(52)

where as defined in (6), $\hat{N}_i = \int_X D_i(x)dx$.

Note that the AA-fused cardinality and PHD for IIDC fusion were first proposed in [42, Th. 4]. The approximate BFoM/MIL derivation of the fused cardinality (35) and fused SPD (50) can be found in [29, Proposition 1]. Comparably, our derivation based on the PHD consistency imposes no ad hoc approximation and has both a theoretically and intuitively sound interpretation: the PHD-AA fusion is committed to more accurate and robust detection of the present targets.

IV. UNLABELED BERNOULLI/MB/MBM PHD-AA FU-
SION

The Bernoulli distribution is conjugate prior for single-
target filtering that suits for detecting and tracking at most one target in the presence of clutter and misdetection. In fact, the Bernoulli RFS has been an important cornerstone for developing many multitarget filters trackers, which can
be briefly classified into two groups: One is the mixture such as the MBM [47], which admits closed-form AA fusion and the others such as the MB [6] do not. Loosely speaking, when the BCs are labeled, they lead to (δ-) GLMB [10], [11] and its simplified versions, including the M-GLMB [44] and LMB [45], respectively. When an extra Poisson component is used (for representing the locally undetected targets), they lead to the PMBM [9] and PMB [48]. When both label and Poisson components are used, it leads to Poisson LMB [60].

The fusion of these mixtures needs to follow the first principle of target-wise fusion [28], [61], i.e., only the information regarding the same, at least highly potentially the same, target should be fused. It is simply technically nonsensical to fuse the components that correspond to different targets. This target-wise fusion rule plays a key role in producing more accurate localization of the present targets. To comply with this rule, the Poisson component in the PMBM/PMB is not fused [38], [39] since the locally undetected targets are by no means guaranteed to be the same across sensors. For the remaining MBM/MB, BC to BC association, and then, merging are the key in accurately localizing each detected target.

A. Bernoulli PHD-AA Fusion

COROLLARY 2 For a group of Bernoulli densities \( f^b_i(X) \) parameterized by \( (r_i, s_i(x)) \), \( i \in \mathbb{I} \), the PHD-AA fusion (36) using fusion weights \( w_i \) will result in

\[
 f^b_{\text{PHD-AA}}(X) = \begin{cases} 
 1 - r_{\text{PHD-AA}} & X = \emptyset \\
 r_{\text{PHD-AA}} s_{\text{PHD-AA}}(x) & X = \{x\} \\
 0 & \text{otherwise} 
\end{cases}
\]  

(53)

where

\[
 r_{\text{PHD-AA}} = \sum_{i \in \mathbb{I}} w_i r_i 
\]

\[
 s_{\text{PHD-AA}}(x) = \frac{1}{r_{\text{PHD-AA}}} \sum_{i \in \mathbb{I}} w_i r_i s_i(x) 
\]

PROOF Using (12), the Bernoulli PHD-AA is

\[
 D^b_{\text{AA}}(x) = r_{\text{PHD-AA}} s_{\text{PHD-AA}}(x) 
\]

(56)

where \( r_{\text{PHD-AA}} \) and \( s_{\text{PHD-AA}}(x) \) are so defined to ensure PHD consistency by \( D^b_{\text{AA}}(x) = \sum_{i \in \mathbb{I}} w_i D^b_i(x) = \sum_{i \in \mathbb{I}} w_i r_i s_i(x) \). So

\[
 r_{\text{PHD-AA}} = \int_X D^b_{\text{AA}}(x) dx = \int_X \sum_{i \in \mathbb{I}} w_i D^b_i(x) dx = \sum_{i \in \mathbb{I}} w_i r_i 
\]

(57)

\[
 s_{\text{PHD-AA}} = \frac{1}{r_{\text{PHD-AA}}} D^b_{\text{AA}}(x) = \frac{1}{r_{\text{PHD-AA}}} \sum_{i \in \mathbb{I}} w_i r_i s_i(x) 
\]

(58)

In other words, the only BC that ensures PHD-AA fusion is given by the aforementioned Bernoulli-AA fusion. ■

The Bernoulli process with parameters \( (r, s(x)) \) is a special IIDC process with \( \rho(0) = 1 - r \), \( \rho(1) = r \) and \( \rho(n+1) = 0 \). The aforementioned results can be viewed as a special case of Lemma 2 and coincide with the following MPD-AA results given in [27, Th. 1] and [42, Lemma 1].

\[
 f^b_{\text{PHD-AA}}(X) = \sum_{i \in \mathbb{I}} w_i f^b_i(X). 
\]

(59)

B. MB PHD-AA Fusion

Extending the single-target-BC to the multitarget RFS, MB lies in the core of MBM [47], (δ-) GLMB [10], [11], PMB [48], M-GLMB [44], LMB [45], and Poisson LMB [60]. Obviously, the AA fusion of the MBs will lead to a mixture of MBs [28], which should be reduced to a single MB by merging all those BCs corresponding to the same target. That is to carry out the AA fusion in parallel to these associated/matched BCs from different sensors, for which the similarity or divergence between these components plays a key role in evaluating whether the components are close enough and so be merged to one.

Here, we consider the MB-AA fusion from the perspective of PHD-AA/consistency. Consider the MB RFS, which has \( n_i \) BCs in which the \( j \)th BC has existence probability \( r_{i,j} \) and SPD \( s_{i,j}(x) \). A 1-D example consisting of two BCs represented by three SPDs associated with different existing probabilities is given in Fig. 3. The corresponding PHD-AA for the MBs can be expressed as, c.f., (36),

\[
 D^b_{\text{MB}}(x) = \sum_{i \in \mathbb{I}} w_i D^b_i(x) = \sum_{i \in \mathbb{I}} w_i r_i s_i(x) 
\]

(60)

which corresponds to the following fused MB MPD, c.f., (13):

\[
 f^b_{\text{PHD-AA}}(X) = \sum_{i \in \mathbb{I}} r_{i,j} f^b_{\text{PHD-AA}}(X^{(j)}) 
\]

(61)
where each multiple-to-one fused BC \( f_{\text{PHD-AA}}^i(X|\mathcal{U}) \) with parameters \((f_{i,\text{PHD-AA}}, s_{i,\text{PHD-AA}}(x))\) is given in (53) in Corollary 2 for each clustered/associated group of BCs, and the final number \( n_f \) of fused BCs depends on the BC merging/clustering procedure, usually

\[
\max\{n_i\}_{i \in I} \leq n_f \leq \sum_{i \in I} n_i \tag{62}
\]

where the left equation holds iff all BCs from the other sensors are associated with those of a local sensor and the right equation holds iff no BCs are associated with the others.

The readers are referred to [28] for the detail of grouping MBs via clustering or pairwise association [28]. For this purpose, data-driven clustering and Bernoulli merging [62] may be useful. The aforementioned MB-AA fusion has been used for PMB-AA-fusion [38]. In the case of labeled MB, the BC association amounts to the label matching as to be addressed in Section V.

C. MBM PHD-AA Fusion

Since the MBM is a mixture that admits closed-form AA fusion [39], the AA fusion is the perfect match to them. Consider the MBM RFS that has \( I \) MBs, each having \( n_{i,j} \) BCs in which the \( l \)th BC has existence probability \( r_{i,j,l} \) and SPD \( s_{i,j,l}(x) \). The corresponding PHD-AA for the MBMs can be expressed as, c.f., (17),

\[
D_{\text{AA}}^{\text{mbm}}(x) = \sum_{i \in I} w_i D_{l}^{\text{mbm}}(x) = \sum_{i \in I} w_i \sum_{j \in I_j} \sum_{l=1}^{n_{i,j}} f_{i,j,l}s_{i,j,l}(x) = \sum_{(i,j) \in I \times I_j} \sum_{l=1}^{n_{i,j}} f'_{i,j,l}s_{i,j,l}(x) \tag{63}
\]

where the reweighted existence probability of each BC is

\[
f'_{i,j,l} = w_if_{i,j,l}. \tag{64}
\]

It can be seen that the number of MBs in the fused MBM equals the sum of those of the fusing MBMs. To combat with the increase of the MBM size, it is not recommended to communicate and fuse the entire MBM, which is not only communicatively and computationally costly but may result in lower accuracy according to our experience. It is just sufficient to communicate and fuse only a few MBs or even a single MB of the greatest hypothesis weight in each fusing MBM [39]. Furthermore, the BC association and merging should be carried out following the target-wise fusion rule: the BCs regarding the same potential target should be merged, although the data association is usually challenging [61]. Moreover, the BC merging is supposed to be carried out to those from different MBs of different sensors (as addressed in [39, Sec.5.2]) while the BCs in the same MB are supposed to represent different targets and should not be fused.

REMARK 2 BC association and merging are the key operations to execute the target-wise AA fusion of these BC-mixture type filters including MB, PMB, MBM, PMBM (and even the GLMB in spite of the extra labels), and their simplified variants. Based on proper intersensor track/BC association, these filters can all be fused in a similar way that facilitates parallel, closed-form Bernoulli merging. By this, the resulted BC can be more concentrated on its mean.

V. LABELED PHD-AA FUSION

The LRFS filters such as the GLMB filters [10], [11] provide the time-series trajectory information for each target via the labels, which makes them advantageous in comparison with the unlabeled RFS filters by eliminating problems like track fragmentation [16]. However, just because of the use of labels that contain track identification information, the fusion becomes more challenging as it is needed to “coordinate/match” different labels/tracks among the fusing LRFS filters.

DEFINITION 2 (LABEL MATCHING) The labels assigned to the BCs that represent the same target should be matched with each other: Only the matched components be fused and be relabeled the same after the fusion. Two components are deemed irrelevant and should not be involved in the fusion if they are unmatched.

This label-wise fusion can be viewed as a strict implementation of the target-wise fusion rule, which is the prerequisite for the LRFS fusion. Unlike in the unlabeled case where the fusing components that correspond to the same target may not be merged if they are distant, in the case of LRFS fusion, all of the labeled components corresponding to the same target must be matched and relabeled using the same, unique label. This is because a single target should not have multiple labels due to the label uniqueness. The problem of the so-called label inconsistency has been noticed earlier [31], [63], [64] but those given label matching solutions account for merely the state densities at the latest time and ignore the whole-track information of the label. In this work, we assume that labels can be properly matched across the sensors and leave how to do so to the future work.

A. Divergence Between GLMBs/LPHDs

By treating the labels as random, discrete variables, the LRFS integral is conceptually defined as [10], [11], [65]

\[
\int_{\mathcal{X} \times \mathcal{L}} \pi(\bar{X}) \delta\bar{X} = \sum_{n=0}^{\infty} \sum_{L_n \subseteq \mathcal{L}} \frac{1}{n!} \int_{\mathcal{X}^n} \pi(\bar{X}_n) d\mathbf{x}_1, \ldots, d\mathbf{x}_n.
\]

In the realization, however, the labels are deterministic and finite. For a given label set \( L_n = \{l_1, l_2, \ldots, l_n\} \), the KL
Fig. 4. Four LMBs of the same state distribution but different labels indicated by different colors: The so-defined KL divergence (65) applies only between the two LRFS distributions (c) and (d) that have the same number of labels, not the other cases that have different numbers of labels.

divergence of two LRFS densities is reduced to

\[
D_{KL}(\pi_1^{L_1} || \pi_2^{L_2}) = \int_{X} \pi_1(\tilde{X}_n) \log \frac{\pi_1(\tilde{X}_n)}{\pi_2(\tilde{X}_n)} d\tilde{X}_n,
\]  (65)

**Remark 3** The divergence like the aforementioned KL that is defined on functions with the same support is restrictive as it holds only for two labeled MPDs with completely the same labels. This is impractical since simply the labels that unavoidably bear estimate error are usually variant among the filters. Their difference is a part of the divergence that unavoidably bear estimate error are usually variant same labels. This is impractical since simply the labels does not exist with the LPHD.

Moreover, the metric (65) applies only between the LRFS distributions of the same number of labels and different label matching choices may result in different results. This can be illustrated in Fig. 4 where four LMBs have the same unlabeled state distribution but different labels and different potential target-state-estimates. It remains undefined how to calculate the divergence between any other two labeled distributions that have different numbers of labels such as the one given in (a) and (b)/(c)/(d) in Fig. 4. This leads to a theoretical barrier to derive the MPD-AA fusion based on the BFoM (33). This problem, however, does not exist with the LPHD (21) that is rigorously defined on single-target, single-label \(X \times \mathcal{L}\), no matter how many labels each filter has.

We therefore extend the definition of the PHD consistency to the labeled domain as follows. If the PHD consistency to the labeled domain and assume the LPHDs to be fused are properly label-matched (and so the matched components will be relabeled the same after fusion). That is, label matching is the prerequisite of the state density fusion. We reiterate that the LPHD-AA fusion is a necessary but not sufficient condition for the MPD-AA fusion.

\[
\tilde{D}_{MPD-AA}(x, l) = \int_{X \times L} \left( \sum_{w \in \tilde{X}} \delta_w [(x, l)] \right) \pi_{AA}(\tilde{X}) \delta \tilde{X} = \sum_{l \in L} w_l \tilde{D}_l(x, l)
\]

(66)

which has the following BFoM property, c.f., (38):

\[
\tilde{D}_{AA}(x, l) = \arg \min_{\pi \in \mathcal{F}_{xx}} \sum_{l \in L} w_l D_{KL}(\tilde{D}_l || \pi)
\]

s.t. \(\int_X g(x, l) dx = \int_X \tilde{D}_{AA}(x, l) dx\)

(67)

where the scalar-valued function set \(\mathcal{F}_{xx} = \{f : X \times L \rightarrow \mathbb{R}\}\), and the KL divergence with respect to two LPHDs \(\tilde{D}_1(x, l), \tilde{D}_2(x, l)\) of the same label \(l \in L\) is defined as

\[
D_{KL}(\tilde{D}_1 || \tilde{D}_2) = \int_X \tilde{D}_1(x, l) \log \frac{\tilde{D}_1(x, l)}{\tilde{D}_2(x, l)} dx.
\]

(68)

We hereafter start from averaging the state distribution of the LPHDs with matched label as shown in (66) but do not calculate (67) or (68), which applies only to LPHDs of the identical label. However, it is the LPHD consistency not the BFoM that forms the motive of our derivation.

B. Index-Matched GLMB LPHD-AA Fusion

**Lemma 3** For a group of GLMBs \(\pi_{i}^{pl}(\tilde{X}), i \in \mathcal{I}\) with matched index \(c \in C\), defined as follows, c.f., (18):

\[
\pi_{i}^{pl}(\tilde{X}) = \Delta(\tilde{X}) \sum_{c \in C} \omega_{i,c}^{(c)}(L(\tilde{X})) \prod_{(x, l) \in \tilde{X}} s_{i,c}^{(c)}(x, l)
\]

(69)

the corresponding LPHD-AA fusion using fusion weights \(w \in \mathbb{W}\) will result in the following fused GLMB:

\[
\pi_{i}^{LPHD-AA}(\tilde{X}) = \Delta(\tilde{X}) \sum_{c \in C} \omega_{i,c}^{LPHD-AA}(L(\tilde{X})) \prod_{(x, l) \in \til{X}} s_{i,c}^{LPHD-AA}(x, l)
\]

(70)

where

\[
\omega_{i,c}^{LPHD-AA}(L) = \sum_{i \in \mathcal{I}} w_i \omega_{i,c}^{(c)}(L)
\]

(71)

\[
s_{i,c}^{LPHD-AA}(x, l) = \frac{1}{\omega_{i,c}^{LPHD-AA}(L)} \sum_{i \in \mathcal{I}} w_i s_{i,c}^{(c)}(x, l).
\]

(72)
PROOF We first calculate the LPHD \( \tilde{D}_l^\delta(x, l) \) of each GLMB, obtaining \( \forall i \in I \)

\[
\tilde{D}_l^\delta(x, l) = \sum_{c \in \mathbb{C}} s_l^{(c)}(x, l) \sum_{L \subseteq \mathbb{L}} 1_l(l) \omega_l^{(c)}(L).
\]

(73)

Given the matched index \( c \in \mathbb{C} \), their AA using fusion weights \( w \in \mathbb{W} \) is given by

\[
\tilde{D}_\text{AA}^\delta(x, l) = \sum_{i \in I} w_i \tilde{D}_l^\delta(x, l)
\]

(74)

\[
= \sum_{c \in \mathbb{C}} \omega_l^{(c)}(L) \sum_{L \subseteq \mathbb{L}} 1_l(l) \omega_l^{(c)}(L).
\]

(75)

Calculating the state integral of both sides of (74) leads to

\[
\int_{x} \sum_{c \in \mathbb{C}} \omega_l^{(c)}(L) \sum_{L \subseteq \mathbb{L}} 1_l(l) \omega_l^{(c)}(L) dx = \int_{x} \sum_{i \in I} w_i \sum_{c \in \mathbb{C}} s_l^{(c)}(x, l) \sum_{L \subseteq \mathbb{L}} 1_l(l) \omega_l^{(c)}(L) dx.
\]

(76)

which leads to (72) for each matched index \( c \) and label \( l \). \( \blacksquare \)

REMARK 4 The fusion of the GLMBs in Lemma 3 is the LPHD-AA consistency as given in (74), not the LPHD-best-fit like (67). As addressed, the latter holds only in the case when the fusing LPHDs have the same label and the divergence (68) is undefined for the LPHDs of different labels. In contrast, the LPHD-AA (74) can still be calculated for LPHDs realized with actually different labels, \( \tilde{D}_l^\delta(x, l) \), \( i \in I \), i.e.,

\[
\tilde{D}_\text{AA}^\delta(x, l) \triangleq \sum_{i \in I} w_i \tilde{D}_l^\delta(x, l)
\]

(77)

where \( l_i \) denotes the recreated label to relabel all components involved in the right side after fusion.

C. \( \delta \)-GLMB LPHD-AA Fusion With Matched Label Set and Track Association Hypotheses

Given that the labels \( l \in \mathbb{L} \), as well as the relevant history data association hypothesis pair \( (L, \xi) \subseteq \mathbb{L} \times \Xi \), are matched across the sensors, we have the following result.

COROLLARY 3 For a group of \( \delta \)-GLMBs \( \pi_{\text{AA}}^\delta(\tilde{X}), i \in I \) with matched labels \( l \in \mathbb{L} \) and track association hypotheses \( (L, \xi) \subseteq \mathbb{L} \times \Xi \), defined as follows:

\[
\pi_{\text{AA}}^\delta(\tilde{X}) = \Delta(\tilde{X}) \sum_{L \subseteq \mathbb{L}} \delta_L \{L(\tilde{X})\} \sum_{\xi \in \Xi} \omega_{L, \xi}(L) s_{(\xi)}^{(L)}(x, l)
\]

(82)

The corresponding MPD and LPHD are given by

\[
\pi_{\text{AA}}^\delta(\tilde{x}) = \Delta(\tilde{x}) \sum_{l \in \mathbb{L}} \omega_{L, \xi}(L) s_{(\xi)}^{(L)}(x, l).
\]

(83)

D. Simplified GLMB LPHD-AA Fusion

Two computationally cheaper, approximate alternatives to the \( \delta \)-GLMB filter [10], [11] are given by the M-GLMB [44] and the LMB [45]. The former preserves both the LPHD and cardinality distribution, while the latter only matches the unlabeled PHD but not the cardinality distribution [66].

In short, the M-GLMB filter [44] marginalizes the history measurement-track association hypothesis \( \exists \Xi \subseteq \Xi \) of the \( \delta \)-GLMB, namely

\[
\omega_{l, \xi}(l) = \sum_{\xi \in \Xi} \omega_{l, \xi}(l)
\]

(81)

which leads to the following MPD and LPHD:

\[
\pi_{\text{M-GLMB}}^\delta(\tilde{X}) = \Delta(\tilde{X}) \sum_{L \subseteq \mathbb{L}} \delta_L \{L(\tilde{X})\} \sum_{\xi \in \Xi} \omega_{L, \xi}(L) s_{(\xi)}^{(L)}(x, l)
\]

(82)

\[
\pi_{\text{M-GLMB}}^\delta(\tilde{x}) = \sum_{L \subseteq \mathbb{L}} \omega_{L, \xi}(l) s_{(\xi)}^{(L)}(x, l).
\]

(83)
Both M-GLMB-LPHD and LMB-LPHD AA fusion can be derived from Corollary 3 by carrying forward the corresponding simplification that they made on the $\delta$-GLMB.

**Corollary 4** For a group of M-GLMBs $\pi_{m,S}^i(\tilde{x})$, $i \in I$ with matched labels $l \in L$ and label set $L \subseteq \mathbb{L}$,

$$\pi_{m,S}^i(\tilde{x}) = \Delta(\tilde{x}) \sum_{l \in L} \delta_l[\mathcal{L}(\tilde{x})] \omega_i^l \prod_{(x,i) \in \tilde{x}} s_i^l(x, l)$$

the corresponding LPHD-AA fusion using fusion weights $w \in \mathbb{W}$ will result in the fused M-GLMB $\pi_{m,S}^{\delta}(\tilde{x}) = \Delta(\tilde{x}) \sum_{l \in L} \delta_l[\mathcal{L}(\tilde{x})] \omega_i^l \prod_{(x,i) \in \tilde{x}} s_i^l(x, l)$, where

$$\omega_i^l = \sum_{i \in I} w_i \omega_i^l$$

$$s_i^l(x, l) = \frac{1}{\omega_i^l} \prod_{(x,i) \in \tilde{x}} s_i^l(x, l).$$

**Proof** The result is from Corollary 3 by using (81) and replacing $s_{LPHD-AA}(x, l)$ by $s_{LPHD-AA}^l(x, l)$. It has been derived from the constrained MPD-best-fit in [31] based on a compressed expression of the $\delta$-GLMB density [65].

**Corollary 5** For a group of LMBs consisting of labeled BCs with existence probability $\tilde{s}_i(x, l)$ and SPD $s_i(x, l)$, $i \in I$, $l \in L$, where $L$ denotes the common label set of all fusing LMBs, the label-wise LPHD-AA fusion leads to a fused LMB with AA-fused existence probability $\tilde{\omega}_i^l$ and cardinalized-AA-fused SPD $\tilde{s}_{LPHD-AA}(x, l)$ given as follows.\forall l \in L:

$$\tilde{r}_i^l = \frac{\tilde{r}_i^l}{\tilde{r}_{LPHD-AA}}$$

$$\tilde{s}_{LPHD-AA}(x, l) = \sum_{i \in I} s_i^l(x, l).$$

**Proof** The result is from Corollary 3 by using (84) and (85) and also from Lemma 3. It is analogous to Corollary 2, regardless of the use of the label. We omit the proof detail here. Notably, it is not more required that any history information, such as $(L, \tilde{x})$, be matched among sensors, but it is still required that the labels are interferfilter matched.

Corollary 5 has been given based on the constrained MPD-fit [31] and further based on the $L_2$ norm metric (disregarding the labels) instead of the KL divergence in [30, Corollary 3]. In both of them, the divergence between the MPDs does not account for that between labels.

**Remark 5** A key condition for fusing $\delta$-GLMB as shown in Corollary 3 is that the history measurement-track association hypotheses $(L, \tilde{x}) \subseteq \mathbb{L} \times \mathbb{S}$ are well matched among sensors, in addition to the label matching. This involves a process of coordinating and revising perhaps the history estimates, analogous to smoothing, which cannot be satisfied in real time [15, ch.4.4.2]. This has been released to only requiring track set (local hypothesis) matching and label matching in fusing the M-GLMB in Corollary 4 and further to only label matching in the LMB-AA fusion in Corollary 5, respectively.

## VI. Conclusion

In this article, we derive suitable AA fusion formulas of different (labeled) RFS filters based on (labeled) PHD consistency, which theoretically explains how the AA fusion can deal with false/missing data and gain better accuracy in target detection and localization, and provides a theoretically unified and exact framework suited for all forms of either RFS or LRFS filters. Moreover, the limitation of existing MPD-best-fit derivations for labeled density fusion is pointed out and circumvented in our derivation. The exact AA fusion formula for the general GLMB filter is driven for the first time. The key parameters and functions of different (labeled) RFS filters, which are individually averaged for PHD/LPHD-AA fusion, are summarized in Table I. While most results comply with existing formulas in the literature, our derivation is well-motivated, theoretically unified and exact. It exposes the essence of existing consensus-driven multisensor linear fusion RFS filters, which merely seek consensus over the first-order moments of the MPDs, namely the (labeled) PHD, rather than the posteriors. This facilitates heterogeneous RFS filter fusion where different forms of RFS filters seek PHD averaging consensus with each other while preserving their own filter form [43]. There is much to investigate in this direction.

Open-access MATLAB codes for the AA-fusion of some representative RFS filters are available online. Compared with the unlabeled case, the labeled PHD fusion is more challenging due to the need of label matching for which a mathematically solid definition of the divergence between labeled densities or tracks is still missing. Considering the nature of labels, the history information of the track should be taken into account. Scenario-specific label matching strategies need further investigation.

## APPENDIX A

**Proof of (38)–(39)**

Define the following probability densities by normalizing the PHDs taken from (38):

$$d_{AA}(x) = \frac{D_{AA}(x)}{N_{AA}}$$

$$d_i(x) = \frac{D_i(x)}{N_i}$$

---

[1] Available: sites.google.com/site/tianchengli85/matlab-codes/aa-fusion
\begin{equation}
    g'(x) \triangleq \frac{g(x)}{N_g}
\end{equation}

where \( \tilde{N}_A = \int_X D_{AA}(x) \, dx \), \( \tilde{N}_I = \int_X D_I(x) \, dx \), and \( \tilde{N}_g = \int_X g(x) \, dx \).

From (26), it is straightforward to get
\begin{equation}
    d_{AA}(x) = \arg \min_{d \in F_X} \sum_{i \in I} w_i D_{KL}(d_i||d).
\end{equation}

The cost to be minimized in (38) can be expanded as follows:
\begin{equation}
    J(g) \triangleq \sum_{i \in I} w_i D_{KL}(d_i||g) = \sum_{i \in I} w_i \int_{X} D_i(x) \log \frac{D_i(x)}{g(x)} \, dx = \sum_{i \in I} w_i \int_{X} \tilde{N}_I d_i(x) \log \frac{\tilde{N}_I d_i(x)}{\tilde{N}_g g(x)} \, dx
\end{equation}

\begin{equation}
    = \sum_{i \in I} w_i \tilde{N}_I \int_{X} d_i(x) \log \frac{\tilde{N}_I}{\tilde{N}_g} + \sum_{i \in I} w_i D_{KL}(d_i||g')
\end{equation}

where (93) and (94) were used in (97).

It is evident that \( \sum_{i \in I} w_i D_{KL}(d_i||g') \) as well as \( J(g) \) is minimized by setting \( g'(x) = d_{AA}(x) \) according to (95). This proves (38) subject to (39).

REFERENCES

[1] B.-N. Vo et al., “Multitarget Tracking,” in Wiley Encyclopedia of Electrical and Electronics Engineering. Hoboken, NJ, USA: Wiley, 2015.

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

[3] R. P. S. Mahler, “Multitarget Bayes filtering via first-order multitarget moments,” IEEE Trans. Aerosp. Electron. Syst., vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[4] B. N. Vo and W. K. Ma, “The Gaussian mixture probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4901–4914, Nov. 2006.

[5] B. Vo, B. Vo, and A. Cantoni, “Analytic implementations of the cardinalized probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

[6] B.-N. Vo, B.-T. Vo, and A. Cantoni, “The cardinality balanced multi-target multi-Bernoulli filter and its implementations,” IEEE Trans. Signal Process., vol. 57, no. 2, pp. 409–423, Feb. 2009.

[7] B. Vo, D. Clark, B. Vo, and B. Ristic, “Bernoulli forward-backward smoothing for joint target detection and tracking,” IEEE Trans. Signal Process., vol. 59, no. 9, pp. 4473–4477, Sep. 2011.

[8] B. Ristic, B. Vo, B. Vo, and A. Farina, “A tutorial on Bernoulli filters: Theory, implementation and applications,” IEEE Trans. Signal Process., vol. 61, no. 13, pp. 3406–3430, Jul. 2013.

[9] J. L. Williams, “Marginal multi-Bernoulli filters: RFS derivation of, MHT, JPDA, and association-based member,” IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 3, pp. 1664–1687, Jul. 2015.

[10] B.-T. Vo and B.-N. Vo, “Labeled random finite sets and multi-object conjugate priors,” IEEE Trans. Signal Process., vol. 61, no. 13, pp. 3460–3475, Jul. 2013.

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

[12] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” Proc. IEEE, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[13] A. H. Sayed, “Adaptation, learning, and optimization over networks,” Found. Trends Mach. Learn., vol. 7, no. 4/5, pp. 311–801, 2014.

[14] F. Meyer et al., “Message passing algorithms for scalable multitarget tracking,” Proc. IEEE, vol. 106, no. 2, pp. 221–259, Feb. 2018.

[15] K. Da, T. Li, Y. Zhu, H. Fan, and Q. Fu, “Recent advances in multisensor multitarget tracking using random finite set,” Front Inf. Technol. Electron. Eng., vol. 22, no. 1, pp. 5–24, 2021.

[16] B.-N. Vo, B.-T. Vo, and M. Beard, “Multi-sensor multi-object tracking with the generalized labeled multi-Bernoulli filter,” IEEE Trans. Signal Process., vol. 67, no. 23, pp. 5952–5967, Dec. 2019.

[17] T. Li, K. Da, H. Fan, and B. Yu, “Multisensor random finite set information fusion advances, challenges, and opportunities,” in Secure and Digitalized Future Mobility, Y. Cao, O. Kaiwartya, and T. Li, Eds. Boca Raton, FL, USA: CRC Press, 2022, Ch. 3, pp. 32–63.

[18] M. Stone, “The opinion pool,” Ann. Math. Statist., vol. 32, pp. 1339–1342, 1961.

[19] T. Li, H. Liang, B. Xiao, Q. Pan, and Y. He, “Finite mixture modeling in time series: A survey of Bayesian filters and fusion approaches,” Inf. Fusion, vol. 98, 2023, Art. no. 101827.

[20] T. Bailey, S. Julier, and G. Agamennoni, “On conservative fusion of information with unknown non-Gaussian dependence,” in Proc. 15th Int. Conf. Inf. Fusion, 2012, pp. 1876–1883.

[21] T. Li, Y. Song, E. Song, and H. Fan, “Arithmetic average density fusion — Part I: Some statistic and information-theoretic results,” Inf. Fusion, vol. 104, 2024, Art. no. 102199.

[22] J. Y. Yu, M. Coates, and M. Rabbat, “Distributed multi-sensor PHD filter using pairwise gossiping,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2016, pp. 3176–3180.

[23] T. Li, J. Corchado, and S. Sun, “On generalized covariance intersection for distributed PHD filtering and a simple but better alternative,” in Proc. 20th Int. Conf. Inf. Fusion, 2017, pp. 808–815.

[24] T. Li, J. Corchado, and S. Sun, “Partial consensus and conservative fusion of Gaussian mixtures for distributed PHD fusion,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 5, pp. 2150–2163, Oct. 2019.

[25] T. Li and F. Hlawatsch, “A distributed particle- PHD filter using arithmetic-average fusion of Gaussian mixture parameters,” Inf. Fusion, vol. 73, pp. 111–124, 2021.

[26] A. K. Goslar, R. Hoseinezhad, and A. Bab-Hadiashar, “Cauchy-Schwarz divergence-based distributed fusion with Poisson random finite sets,” in Proc. Int. Conf. Control, Automat. Inf. Sci., 2017, pp. 112–116.

[27] T. Li, Z. Liu, and Q. Pan, “Distributed Bernoulli filtering for target detection and tracking based on arithmetic average fusion,” IEEE Signal Process. Lett., vol. 26, no. 12, pp. 1812–1816, Dec. 2019.

[28] T. Li, X. Wang, Y. Liang, and Q. Pan, “On arithmetic average fusion and its application for distributed multi-Bernoulli multitarget tracking,” IEEE Trans. Signal Process., vol. 68, pp. 2883–2896, Apr. 2020.

[29] L. Gao, G. Battistelli, and L. Chisci, “Multiobject fusion with minimum information loss,” IEEE Signal Process. Lett., vol. 27, pp. 201–205, Jan. 2020.

[30] A. K. Goslar et al., “Cooperative sensor fusion in centralized sensor networks using Cauchy-Schwarz divergence,” Signal Process., vol. 167, 2020, Art. no. 107278.

[31] L. Gao, G. Battistelli, and L. Chisci, “Fusion of labeled RFS densities with minimum information loss,” IEEE Trans. Signal Process., vol. 68, pp. 5855–5868, Oct. 2020.

[32] R. K. Ramachandran, N. Fronda, and G. S. Sukhatme, “Resilience in multirobot multitarget tracking with unknown number of targets through reconfiguration,” IEEE Trans. Control Netw. Syst., vol. 8, no. 2, pp. 609–620, Jun. 2021.

[33] K. Da, T. Li, Y. Zhu, and Q. Fu, “Gaussian mixture particle jump-Markov-CPHD fusion for multitarget tracking using sensors with limited views,” IEEE Trans. Signal Inf. Process. Netw., vol. 6, pp. 605–616, Aug. 2020.
[34] T. Li, V. Elvira, H. Fan, and J. M. Corchado, “Local-diffusion-based distributed SMC-PHD filtering using sensors with limited sensing range,” IEEE Sensors J., vol. 19, no. 4, pp. 1580–1589, Feb. 2019.

[35] W. Yi, G. Li, and G. Battistelli, “Distributed multi-sensor fusion of PHD filters with different sensor fields of view,” IEEE Trans. Signal Process., vol. 68, pp. 5204–5218, Sep. 2020.

[36] G. Li, G. Battistelli, L. Chisci, and L. Kong, “Distributed multi-target tracking over an asynchronous multi-sensor network,” in Proc. Radar Conf., 2020, pp. 1–6.

[37] B. Yu, T. Li, S. Ge, and H. Gu, “Robust CPHD fusion for distributed multitarget tracking using asynchronous sensors,” IEEE Sensors J., vol. 22, no. 1, pp. 1030–1040, Jan. 2022.

[38] G. Li, “An IOGM-based Poisson multi-Bernoulli filter and its application to distributed multi-sensor fusion,” IEEE Trans. Aerosp. Electron. Syst., vol. 58, no. 4, pp. 3666–3677, Aug. 2022.

[39] T. Li, Y. Xin, Z. Liu, and K. Da, “Best fit of mixture for computationally efficient Poisson multi-Bernoulli mixture filtering,” Signal Process., vol. 202, 2023, Art. no. 108739.

[40] G. Koliander, Y. El-Laham, P. M. Djurić, and F. Hlawatsch, “Fusion of probability density functions,” Proc. IEEE, vol. 110, no. 4, pp. 404–453, Apr. 2022.

[41] M. Kayaalp, Y. Inan, E. Telatar, and A. H. Sayed, “On the arithmetic and geometric fusion of beliefs for distributed inference,” IEEE Trans. Autom. Control, 2023, to be published, doi: 10.1109/TAC.2023.3330405.

[42] K. Da, T. Li, Y. Zhu, H. Fan, and Q. Fu, “Kullback-Leibler averaging for multitarget density fusion,” in Proc. Int. Symp. Distrib. Comput. Artif. Intell., 2019, pp. 253–261.

[43] T. Li, R. Yan, K. Da, and H. Fan, “Arithmetic average density fusion—Part III: Heterogeneous unlabeled and labeled filter fusion,” IEEE Trans. Aerosp. Electron. Syst., vol. 60, no. 1, pp. 1023–1034, Feb. 2024, doi: 10.1109/TAES.2023.3334223.

[44] C. Fantacci, B.-T. Vo, F. Papini, and B.-N. Vo, “The marginalized tagGLMB filter,” arXiv:1501.00926.

[45] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, “The labeled multi-Bernoulli filter,” IEEE Trans. Signal Process., vol. 62, no. 12, pp. 3246–3260, Jun. 2014.

[46] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen, Mathematics of Data Fusion. Dordrecht, Netherlands: Springer, 1997.

[47] A. F. Garcia-Fernández, J. L. Williams, K. Granström, and L. Svensson, “Poisson multi-Bernoulli mixture filter: Direct derivation and implementation,” IEEE Trans. Aerosp. Electron. Syst., vol. 54, no. 4, pp. 1883–1901, Aug. 2018.

[48] J. L. Williams, “An efficient, variational approximation of the best fitting multi-Bernoulli filter,” IEEE Trans. Signal Process., vol. 63, no. 1, pp. 258–273, Jan. 2015.

[49] R. Kulhavý and F. Kraus, “On duality of regularized exponential and linear forgetting,” Automatica, vol. 32, no. 10, pp. 1403–1415, 1996.

[50] S. J. Julier, T. Bailey, and J. K. Uhlmann, “Using exponential mixture models for suboptimal distributed data fusion,” in Proc. IEEE Nonlinear Stat. Signal Process. Workshop, 2006, pp. 160–163.

[51] M. Üney, J. Housinseine, E. Delande, S. J. Julier, and D. E. Clark, “Fusion of finite set distributions: Pointwise consistency and global cardinality,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 6, pp. 2759–2773, Dec. 2019.

[52] B. Wang, W. Yi, R. Hoseinnezhad, S. Li, L. Kong, and X. Yang, “Distributed fusion with multi-Bernoulli filter based on generalized covariance intersection,” IEEE Trans. Signal Process., vol. 65, no. 1, pp. 242–255, Jan. 2017.

[53] Z. Lu, X. Hu, and T. Kirubarajan, “Labeled random finite sets with moment approximation,” IEEE Trans. Signal Process., vol. 65, no. 13, pp. 3384–3398, Jul. 2017.

[54] A. Saucan, M. J. Coates, and M. Rabbat, “A multisensor multi-Bernoulli filter,” IEEE Trans. Signal Process., vol. 65, no. 20, pp. 5495–5509, Oct. 2017.

[55] M. Gunay, U. Orguner, and M. Demirekler, “Chernoff fusion of Gaussian mixtures based on sigma-point approximation,” IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 6, pp. 2732–2746, Dec. 2016.