Distributed SMC-PHD Fusion for Partial, Arithmetic Average Consensus

Tiancheng Li

***Abstract*—We propose an average consensus approach for distributed SMC-PHD (sequential Monte Carlo-probability hypothesis density) fusion, in which local filters extract Gaussian mixtures (GMs) from their respective particle posteriors, share them (iteratively) with their neighbors and finally use the disseminated GM to update the particle weight. The resulting particle distribution is the arithmetic average of the disseminated GM-posteriors. There are two distinguishable features of our approach compared to exiting approaches. First, a computationally efficient particles-to-GM (P2GM) conversion scheme is developed based on the unique structure of the SMC-PHD updater in which the particle weight can be exactly decomposed with regard to the measurements and misdetection. Only significant components of higher weight are utilized for parameterization and so the disseminated information is only a part of that of local posteriors. The consensus, conditioned on partial information dissemination over the network, is called "partial consensus". Second, importance sampling (IS) is employed to re-weight the local particles for integrating the received GM information, without changing the states of the particles. By this, the local prior PHD and likelihood calculation can be carried out in parallel to the dissemination & fusion procedure.**

To assess the effectiveness of the proposed P2GM parameterization approach and IS approach, two relevant yet new distributed SMC-PHD fusion protocols are introduced for comparison. One uses the same P2GM conversion and GM dissemination schemes as our approach but local particles are regenerated from the disseminated GMs at each filtering iteration - in place of the IS approach. This performs similar to our IS approach (as expected) but prevents any parallelization as addressed above. The other is disseminating the particles between neighbors - in place of the P2GM conversion. This avoids parameterization but is communicatively costly. This protocol, essentially seeking complete (posterior) consensus, however, does not perform better than the GM-dissemination based partial consensus. Different to these arithmetic average consensus approaches, the state-of-the-art exponential mixture density approach that seeks geometric average consensus is also realized for comparison.

***Index Terms*—Distributed tracking, average consensus, PHD filter, particle filter, Gaussian mixture, partial consensus, arithmetic average, geometric average.**

I. Introduction

DISTRIBUTED target tracking (DTT) based on wireless sensor networks (WSNs) has received considerable research interest in the last decade. It basically involves a number of spatially distributed, low-powered, interconnected sensors that are equipped with a signal processing unit, allowing them to carry out sensing, calculation and communication with the neighbors without a fusion center [1], [2]. These sensors cooperatively track the targets based on their local measurement and the information disseminated from the others, which facilitates better estimation accuracy and overcomes many deficiencies that an isolated sensor suffers from such as false and missing measurement and limited fields of view. Typically, it is expected that local nodes reach a single *"consensus"* [3], [4] conditioned on the common information they share after sufficient peer-to-peer (P2P) communication.

To deal with the nonlinearity and non-normal non-Gaussian uncertainty that are involved in the statistical models regarding the targets, the scenario and/or the sensors, the sequential Monte Carlo (SMC) approach provides one of the most vital tools for realizing sequential Bayesian inference (SBI), which is also known as the particle filter (PF); see some recent advances [5]. Realizing the PF on the WSN leads to a quite universal distributed tracking framework, which has invited many specific implementation protocols; see literature reviews offered in [6]–[8] primarily regarding a single target.

For multitarget tracking (MTT) in the presence of false and missing data, measurement-to-target association is typically needed which entails either computationally intensive calculation or ad-hoc strategy (such as gating) design [9], [10]. To overcome this difficulty, random finite set (RFS) has emerged as a powerful and versatile tool. In particular, instead of propagating the full multitarget density which has been considered computationally intractable, the PHD (probability hypothesis density) filter propagates the first order statistical moment of the multitarget RFS [11] and avoids measurement-to-target association. Consequently, many RFS-models based PFs have been developed [12], which have become a new driving force for the flourishing of both the PF algorithm and the RFS filtering family.

Nonetheless, exact implementation of the multisensor PHD filter involves summing over all partitions of the measurements from different sensors which is intractable in computation (if not impossible) and typically, one has to resort to simplifying approximation [13], [14]. Alternatively, immense interest has been seen recently for extending the theory of "average consensus" to DTT, in which the item being estimated may be the *arithmetic average* [2] (AA, akin to the linear opinion pool [3], [4]) or the *geometric average* [15] (GA, akin to the logarithmic opinion pool [16]) of the initial values. AA and GA differ in measuring the distance for calculating the "average". In the former, it is the Euclidean distance while in the latter it is the Kullback-Leibler divergence [17], [18].

Notably, the GA fusion coincides with the covariance intersection approach [17], [19]–[21], a type of Chernoff fusion, which was originally developed for addressing unknown infor-

T. Li is with the School of Sciences, University of Salamanca, 37007 Salamanca, Spain, E-mail: t.c.li@usal.es; tiancheng.li1985@gmail.com.

This work is in part supported by the Marie Skłodowska-Curie Individual Fellowship (H2020-MSCA-IF-2015) under Grant no. 709267

mation correlation among sensors and for avoiding information double-counting. This approach has been widely applied for distributed PHD fusion, based on either the Gaussian mixture (GM) implementation [22]–[25] or the PF [26], [27]. Literally, the GA is also referred to as the Kullback-Leibler average (KLA) [22], [28], exponential/geometric mixture density (EMD/GMD) [17], [27] or generalized covariance intersection (GCI) [29]–[31]).

However, the GA rarely admits closed-form solution for a mixture distribution such as GM, not to say an arbitrary particle distribution. The only existing GA-based SMC-PHD fusion resorts to clustering and converting the particle distribution to continuous distribution approximations [27], which disseminates both particles and the continuous functions, suffering from very intensive communication. Moreover, the GA fusion suffers from several deficiencies such as:

1) Delay in detecting new appearing targets [23];
2) Failure to handle closely distributed targets and/or low SNR background [32];
3) Prone to mis-detection [24], [33] or incorrect data [34].

For GM-PHD average consensus, we have demonstrated that the apparently simple AA offers a promising alternative to the GA, yielding higher filtering accuracy, better reliability in cluttered environment with mis-detection, and lower communication and fusion-computation (C&F) cost [35]. Clearly, the AA is universal and unlimited to the GM filter. As we will show in this paper, it also applies to the RFS-PFs, for which the C&F challenges are:

- *Information dissemination*: it is practically preventable to disseminate the particle set, but instead, efficient parameterization such as particle-to-Gaussian/GM conversion to compromise between approximation accuracy and communication cost is on great demand.
- *Information fusion*: the immediate challenge to the parametric posterior dissemination is how to efficiently integrate/incorporate the parametric information into the particle distribution for improving upon its fitness to the global optimum, which is deemed to be an "average" of the interested initial posteriors.
- *Real time networking*: In the most favorable situation, the C&F should be carried out in parallel with local filter calculations to avoid any time delay to the filter, which is referred to as *real time networking*. However, in almost all existing DTT systems, the filtering calculation (at one stage or another) depends on the shared information gained by the C&F which is exactly how the filters benefit from networking. As a result, local filtering calculation and neighbor-wise communication are performed interactively in time. This may not be allowed in reality.

These challenges motivate our work. On the one hand, for posterior parameterization, we investigate the unique structure of the PHD filter whose posterior can be decomposed with regard to measurements and misdetection. In consistent with the notion of "*partial consensus*" [35], only significant components (in the format of few parameters) are disseminated among sensors while the insignificant components are not involved in C&F. The idea of sharing only a part of the information over the network has appeared earlier in the centralized network with a fusion center [36], [37] and in the diffusion network [38], [39], in which the benefit is primarily limited to saving communication/memory cost. However, we demonstrate (besides [35]) that, the *partial consensus* does not only save communication/memory but also, more importantly, improve the estimation accuracy.

On the other hand, we propose an efficient, fully distributed, means to update weights of local particles (without changing their states) according to the disseminated GMs based on importance sampling (IS), leading to an exact weighted AA of the disseminated GM-posteriors.

The resulting framework enjoys two novel features:

- Local filtering calculation is allowed to be carried out partially in parallel to the C&F, for reducing the network dissemination delay.
- The framework can cooperate seamlessly with the distributed GM-PHD filter [35] for a hybrid sensor network consisting of both GM-PHD and SMC-PHD filters, without any special algorithm design for either.

The remainder of this paper is organized as follows. Primary notations and assumptions are listed next immediately. The basics of the SMC-PHD filter and the particle weight decomposition are given in Section II. The proposed distributed SMC-PHD fusion approach is detailed in Section III. Simulations are given in Section IV. We conclude in Section V.

A. Notation and Assumption

The sensor network is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the set of sensors $\mathcal{V} = \{1, 2, \cdots, N\}$ and the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. In the directed graph, any edge is denoted by an ordered pair of sensors $(a, b) \in \mathcal{E}$, which means node b is directly reachable from node a, where a is called the in-neighbor of b while b is the out-neighbor of a. For any $b \in \mathcal{V}$, denote $\mathcal{N}_b := \{a \in \mathcal{V} | (a, b) \in \mathcal{E}, a \neq b\}$, which is the set of all the in-neighbors of node b excluding node b itself. Undirected graph is a special type of directed graph where for any $(a, b) \in \mathcal{E}$, we must have $(b, a) \in \mathcal{E}$.

The collections of target states and measurements at time k can be represented as finite sets $X_k = \{\mathbf{x}_{k,1}, \cdots, \mathbf{x}_{k,N_k}\}$ and $Z_k = \{\mathbf{z}_{k,1}, \cdots, \mathbf{z}_{k,M_k}\}$, where N_k and M_k are the number of targets and of measurements, respectively. The cardinality (number of elements) of a finite set I is denoted by $|I|$. Therefore, we have $N_k = |X_k|$ and $M_k = |Z_k|$. A Gaussian probability density function (PDF) of a random variable $\mathbf{x}$ with mean $\mathbf{m}$ and covariance $\mathbf{P}$ is denoted by $\mathcal{G}(\mathbf{x}; \mathbf{m}, \mathbf{P})$ and the Kronecker delta function is denoted as $\delta_{\mathbf{y}}(\mathbf{x})$.

At each time k, a random Poisson number of targets appear according to the new-born intensity function $\gamma_k(\mathbf{x})$. We do not particularly consider target spawn. Each target is assumed to evolve and generate measurements independently of others. More specifically, a target with state $\mathbf{x}_{k-1}$ may either disappear with probability $1 - p_{\mathrm{S},k}(\mathbf{x}_{k-1})$, or continue to exist at time k with survival probability $p_{\mathrm{S},k}(\mathbf{x}_{k-1})$ and move to a new state with a transition probability density

$$f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}).$$

A target with state $\mathbf{x}_k \in X_k$ is either miss-detected with probability $1-p_{\mathrm{D,k}}(\mathbf{x}_k)$ or detected with probability $p_{\mathrm{D,k}}(\mathbf{x}_k)$ and generates an measurement $\mathbf{z}_k \in Z_k$ with likelihood

$$g_k(\mathbf{z}_k|\mathbf{x}_k).$$

One target can generate no more than one measurement at each scan. In addition, the number of clutter points at time k is subject to a random Poisson distribution $\kappa_k(\mathbf{z})$ and is independent of the real measurement of targets.

II. Decomposition of Particle PHD

A. RFS and PHD

A RFS variable X is a random variable that takes values as unordered finite sets and is uniquely specified by its cardinality distribution $\rho(n) \triangleq \Pr[|X| = n]$ and a family of symmetric joint distributions $p_n(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ that characterize the distribution of its elements over the state space, conditioned on the set cardinality n. The PDF $f(X)$ of a RFS variable X is given as $f(\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}) = n!\rho(n)p_n(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$.

The PHD $D_{\mathcal{S}}(\mathbf{x})$ of a multitarget RFS variable X with the PDF $f(X)$ in a measurable region $\mathcal{S} \subseteq \mathbb{R}^d$ is given as:

$$D_{\mathcal{S}}(\mathbf{x}) = \int_{\mathcal{S}} \delta_X(\mathbf{x}) f(X) \delta X \,, \tag{1}$$

where $\delta_X(\mathbf{x}) \triangleq \sum_{\mathbf{y}\in X} \delta_{\mathbf{y}}(\mathbf{x})$ and the RFS integral reads: $\int_{\mathcal{S}} f(X)\delta X \triangleq f(\emptyset) + \sum_{n=1}^{\infty} \int_{\mathcal{S}^n} \frac{f(\{\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n\})}{n!} d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n$.

B. PHD Filtering Recursion

Denote by $D_{k|k-1}(\mathbf{x})$ and $D_{k|k}(\mathbf{x})$ the PHD of the prior and posterior point processes $X_k|Z_{1:k-1}$ and $X_k|Z_{1:k}$, respectively. Omitting the conditioning on the measurements for convenience, the PHD prediction-updating recursion can be given as follows [11]:

$$\cdots \to D_{k-1|k-1}(\mathbf{x}) \to D_{k|k-1}(\mathbf{x}) \to D_{k|k}(\mathbf{x}) \to \cdots \tag{2}$$

To be more specific,

1) *Time update step* (to calculate the prior PHD):

$$\begin{aligned} D_{k|k-1}(\mathbf{x}) =& \gamma_k(\mathbf{x}) + \\ & \int p_{\mathrm{S},k}(\mathbf{x}') f_{k|k-1}(\mathbf{x}|\mathbf{x}') D_{k-1|k-1}(\mathbf{x}') d\mathbf{x}' \,. \end{aligned} \tag{3}$$

2) *Measurement update step* (to calculate the posterior PHD):

$$\begin{aligned} D_{k|k}(\mathbf{x}) =& \big(1 - p_{\mathrm{D,k}}(\mathbf{x})\big) D_{k|k-1}(\mathbf{x}) + \\ & \sum_{\mathbf{z}\in Z_k} \frac{p_{\mathrm{D,k}}(\mathbf{x}) g_k(\mathbf{z}|\mathbf{x}) D_{k|k-1}(\mathbf{x})}{\kappa_k(\mathbf{z}) + \int p_{\mathrm{D,k}}(\mathbf{x}) g_k(\mathbf{z}|\mathbf{x}) D_{k|k-1}(\mathbf{x}) d\mathbf{x}} \,. \end{aligned} \tag{4}$$

There have been many implementations of the PHD filter based on different types of PFs since the considered standard implementation [40], including auxiliary PF [41], marginalized PF [42] and box PF [43]. Without loss of generality, we adopt the standard implementation [40] (except the estimate extraction part for which we will employ computationally much faster approaches, to be addressed in Section III-E) without giving its detail here.

C. Particle (posterior) Weight Decomposition

The representation of the posterior PHD $D_{k|k}(\mathbf{x})$ by using J_k particles with state $\mathbf{x}_k^{(j)}$ and nonnegative weight $w_{k|k}^{(j)}$, $j = 1, 2, \cdots, J_k$, can be written as [40]

$$D_{k|k}(\mathbf{x}) \approx \sum_{j=1}^{J_k} w_{k|k}^{(j)} \delta_{\mathbf{x}_k^{(j)}}(\mathbf{x}) \,, \tag{5}$$

where (cf. (4))

$$\begin{aligned} w_{k|k}^{(j)} =& \Big(1 - p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big)\Big) w_{k|k-1}^{(j)} + \\ & \sum_{\mathbf{z}_k\in Z_k} \frac{p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big) g_k\big(\mathbf{z}_k|\mathbf{x}_k^{(j)}\big) w_{k|k-1}^{(j)}}{\kappa_k(\mathbf{z}_k) + \sum_{j=1}^{J_k} p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big) g_k\big(\mathbf{z}_k|\mathbf{x}_k^{(j)}\big) w_{k|k-1}^{(j)}} \,, \end{aligned} \tag{6}$$

and $w_{k|k-1}^{(j)}$ is the prediction weight of particle j (either evolved from time $k-1$ or new born at time k; see [40] for detail) and admits $D_{k|k-1}(\mathbf{x}) \approx \sum_{j=1}^{J_k} w_{k|k-1}^{(j)} \delta_{\mathbf{x}_k^{(j)}}(\mathbf{x})$.

Obviously, (6) can be decomposed with regard to the measurements

$$w_{k|k}^{(j)}(\mathbf{z}_k) \triangleq \begin{cases} \Big(1 - p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big)\Big) w_{k|k-1}^{(j)} & \text{if } \mathbf{z}_k = \mathbf{z}_0 \,, \\ \dfrac{p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big) g_k\big(\mathbf{z}_k|\mathbf{x}_k^{(j)}\big) w_{k|k-1}^{(j)}}{\kappa_k(\mathbf{z}_k) + \sum_{j=1}^{J_k} p_{\mathrm{D,k}}\big(\mathbf{x}_k^{(j)}\big) g_k\big(\mathbf{z}_k|\mathbf{x}_k^{(j)}\big) w_{k|k-1}^{(j)}} & \text{if } \mathbf{z}_k \in Z_k \,. \end{cases} \tag{7}$$

where the pseudo-measurement $\mathbf{z}_0$ is introduced to represent the misdetection.

$w_{k|k}^{(j)}(\mathbf{z}_k)$ implies how much each $\mathbf{z}_k$ contributes to the weight of particle j. Straightforwardly, we have

$$w_{k|k}^{(j)} = \sum_{\mathbf{z}_k\in\{\mathbf{z}_0\}\cup Z_k} w_{k|k}^{(j)}(\mathbf{z}_k) \,. \tag{8}$$

Furthermore, we define the sum of weight components of all particles with regard to measurement $\mathbf{z}_k$ as

$$W_k(\mathbf{z}_k) \triangleq \sum_{j=1}^{J_k} w_{k|k}^{(j)}(\mathbf{z}_k) \,, \tag{9}$$

which indicates the probability that the underlying measurement is from a real target ($\mathbf{z}_k \in Z_k$) or that misdetection occurs ($\mathbf{z}_k = \mathbf{z}_0$). Obviously, $\forall \mathbf{z}_k \in Z_k, W_k(\mathbf{z}_k) \in [0, 1]$ [44] and, the weight sum admits

$$\sum_{\mathbf{z}_k\in\{\mathbf{z}_0\}\cup Z_k} W_k(\mathbf{z}_k) = \sum_{j=1}^{J_k} w_{k|k}^{(j)} \triangleq W_k \,. \tag{10}$$

D. Multitarget RFS Cardinality Estimation

The expectation of the total number of targets N_k conditioned on the PHD is given by its integral in the entire state space, which is approximated by the weight sum W_k of all particles as in (5); see the detailed derivation given in *Appendix A*. That is,

$$\mathrm{E}[N_k|D_{k|k}(\mathbf{x})] \approx W_k \,. \tag{11}$$

For estimate extraction, a common approach to estimate the number of targets is rounding the total weight sum, i.e.,

$$\hat{N}_k = [W_k]\,, \tag{12}$$

where the operator $[\cdot]$ rounds the content to the nearest integer.

III. Our Proposal

We consider now a sensor network where all sensors synchronously observe the scenario, affected with independent noises and clutter. The local sensor performs particle prediction, updating and resampling exactly as in the centralized case, except that an additional C&F scheme is carried out once the posterior PHD is achieved at each filtering iteration, which uses the information disseminated from the other sensors to "re-weight" the underlying particle set for consensus.

In the sequel, we shall only concentrate on the C&F part, which consists of three key components:

- Extract a GM from each local particle set and disseminate them in neighborhood, perhaps in multiple iterations and with mixture reduction applied; see Section III-A.
- Update the weight of local particles to integrate the posterior information carried in the disseminated GM that is comprised of components both received from the other sensors and generalized locally; see Section III-C.
- Seek cardinality AA consensus in neighborhood in parallel to the above schemes of GM dissemination and particle-GM fusion; see Section III-D.

A. Particle to GM Conversion

We propose to extract GMs from the a posteriori particle distribution at each sensor based on the weight decomposition as in (7) as that each measurement corresponds to one Gaussian component (GC). By this, as many as M_k+1 GCs can be obtained which, however, could still be too communicatively costly. Arguably, the GM should contain sufficient information of the potential targets subject to the communication limitation. Following the partial consensus principle [35], only the significant GC, namely corresponding to high $W_k(\mathbf{z})$, should be disseminated for consensus, while the insignificant GC should be less likely involved.

The number of significant measurements can be determined either by the estimated number of targets $\hat{N}_k$ as in (12) or as that of the measurements corresponding to $W_k(\mathbf{z}_k)$ larger than a threshold T_c (usually, $T_\mathrm{c} \in [0.1, 0.5]$). The former is referred to as the *Rank* rule while the latter is the *Threshold* rule (considered as the default in our approach), akin to the notions used in [45] and [35]. Either way, we denote the selected measurements by a subset $Z_{k,\mathrm{T}} \subseteq Z_k$.

For each $\mathbf{z}_k \in Z_{k,\mathrm{T}}$, a GC $\mathcal{G}\big(\mathbf{x};\hat{\mathbf{m}}_k(\mathbf{z}_k),\hat{\mathbf{P}}_k(\mathbf{z}_k)\big)$ weighted by $W_k(\mathbf{z}_k)$ can be extracted from the weight component-based particle set $\big\{(\mathbf{x}_k^{(j)}, w_{k|k}^{(j)}(\mathbf{z}_k))\big\}_{j=1}^{J_k}$, i.e., cf.(5)

$$W_k(\mathbf{z}_k)\mathcal{G}\big(\mathbf{x};\hat{\mathbf{m}}_k(\mathbf{z}_k),\hat{\mathbf{P}}_k(\mathbf{z}_k)\big) \approx \sum_{j=1}^{J_k} w_{k|k}^{(j)}(\mathbf{z}_k)\delta_{\mathbf{x}_k^{(j)}}(\mathbf{x}), \tag{13}$$

where $W_k(\mathbf{z}_k)$ is already given in (9) while the mean and the covariance of the formed GC are given as follows, respectively,

$$\hat{\mathbf{m}}_k(\mathbf{z}_k) = \sum_{j=1}^{J_k} w_{k|k}^{(j)}(\mathbf{z}_k)\mathbf{x}_k^{(j)}\,, \tag{14}$$

$$\hat{\mathbf{P}}_k(\mathbf{z}_k) = \sum_{j=1}^{J_k} w_{k|k}^{(j)}(\mathbf{z}_k)\big(\mathbf{x}_k^{(j)} - \hat{\mathbf{m}}_k(\mathbf{z}_k)\big)\big(\mathbf{x}_k^{(j)} - \hat{\mathbf{m}}_k(\mathbf{z}_k)\big)^{\mathrm{T}}. \tag{15}$$

It is necessary to note that, such a Gaussian approximate can only become accurate when the prior PHD is Gaussian-distributed and the likelihood function is Gaussian. Otherwise, the parameterization is no more than approximation.

Various particles-to-GC/GM converting approaches have been developed for distributed particle filtering [6]–[8], [27], [46]–[48], mostly based on either ad-hoc strategy or sophisticated learning algorithm. Thanks to the unique weight-decomposition property of the SMC-PHD updater, our approach is computationally very efficient and reliable.

We may define the "*partial PHD*" $D_{k,\mathrm{T}}(\mathbf{x})$ as a congregation of all the significant components of the PHD, namely the part to be extracted for parameterization, i.e., cf.(5)

$$D_{k,\mathrm{T}}(\mathbf{x}) \triangleq \sum_{j=1}^{J_k} w_{k,\mathrm{T}}^{(j)}\delta_{\mathbf{x}_k^{(j)}}(\mathbf{x})\,, \tag{16}$$

where $w_{k,\mathrm{T}}^{(j)} = \sum_{\mathbf{z}\in Z_{k,\mathrm{T}}} w_{k|k}^{(j)}(\mathbf{z}_k) \leq w_{k|k}^{(j)}$ denotes all the significant part of the weight of particle j. In contrast, the remaining components $w_{k|k}^{(j)} - w_{k,\mathrm{T}}^{(j)}$ are considered insignificant and will not be involved in the C&F procedure.

Substituting (13) into (16) gives an explicit GM approximation of the partial PHD, i.e.,

$$D_{k,\mathrm{T}}(\mathbf{x}) \approx \sum_{\mathbf{z}\in Z_{k,\mathrm{T}}} W_k(\mathbf{z}_k)\mathcal{G}\big(\mathbf{x};\hat{\mathbf{m}}_k(\mathbf{z}_k),\hat{\mathbf{P}}_k(\mathbf{z}_k)\big)\,. \tag{17}$$

Remark 1 When multiple communication iterations are performed, the local sensor will have an iteration-increasing GM size unless mixture reduction is applied [7]. To save the communication, GM merging [35] may be performed at all or some iterations, e.g., when the size exceeds a predefined upper threshold, which, however, may lead to information double-counting in addition to merging error. For example, if a GC sent from sensor a to sensor b is merged with another GC at sensor b, the resulting fused GC will be sent back to sensor a in the next communication iteration. In this case, appropriate (sensor-oriented) fusion weights shall be designed for fast consensus convergence.

B. Weighted, Arithmetic Average of Partial PHDs

In the following formulation, we will use subscripts a and $b \in \mathcal{N}_a$ to distinguish between two neighboring sensors. In the proposed protocol, the local partial PHD $D_{a,k,\mathrm{T}}(\mathbf{x})$ at sensor a will be linearly averaged with the received GM/partial PHD from the neighbors $D_{b,k,\mathrm{T}}(\mathbf{x}), \forall b \in \mathcal{N}_a$, i.e.,

$$\bar{D}_{a,k,\mathrm{T}}(\mathbf{x}) = \sum_{b\in\{a\}\cup\mathcal{N}_a} \omega_{b\to a} D_{b,k,\mathrm{T}}(\mathbf{x})\,, \tag{18}$$

where the fusion weights $\omega_{b\to a} \geq 0$, $\sum_{b\in\{a\}\cup\mathcal{N}_a}\omega_{b\to a} = 1$ indicating that the fusion result is an "average".

As the key of our approach, we will use the arithmetically averaged partial PHD to replace the local PHD, by means of re-weighting the local particles. To this end, we apply the the Metropolis weights [49], [50] which are given as

$$\omega_{b\to a}=\begin{cases}\frac{1}{1+\max(|\mathcal{N}_a|,|\mathcal{N}_b|)} & \text{if } b\in\mathcal{N}_a, b\neq a\,,\\ 1-\sum_{l\in\mathcal{N}_a}\omega_{l\to a} & \text{if } b=a\,.\end{cases} \tag{19}$$

To note, the purpose of Metropolis weights here is the same to that in the original proposal [49], [50], which is for fast AA convergence and has no explicit connection to issues such as "dividing the common information" or "coping with the correlation" between sensors [51]. Therefore, we will not address on those issues (for which a recent review [52] is available) but instead we concentrate our distributed fusion goal on *"average consensus"*.

Remark 2 The rationale for calculating the AA of PHDs is based on the essential property of the PHD that *the integral of PHD in any region gives the expected number of targets in that region* - cf. (1). This renders the AA calculation as in (18) a meaningful interpretation. Also, there is an important assumption behind the PHD filter: both misdetection and clutter are random and are independent of the real measurement of targets. So, it is unlikely for the same target to be missed in detection, or to say a target does not form a significant GC, in the majority of all sensors, or false alarms coincidentally occur in the same area in the majority of all sensors [35]. Using the principle of "majority rule", the AA can compensate for the false/missing data of a single sensor. Indeed, the AA is provably immune to either false or missing data problem. Exactly because of this, it is reasonable to abandon the insignificant GCs without worrying about misdetection or false alarms, namely partial AA consensus. Doing so does not only save communication but also tend to ameliorate the local signal-to-noise ratio (SNR) [35], [51].

C. Particle Updating w.r.t. GMs based on IS

Denote by $\{(\mathbf{x}_{a,k}^{(j)},\mu_a)\}_{j=1}^{J_{a,k}}$ the uniformly weighted particles yielded by resampling after the PHD updating at sensor a at time k. Given the local weight sum $W_{a,k}$, we have $\mu_a=\frac{W_{a,k}}{J_{a,k}}$.

As long as the resampling scheme adopted is unbiased [53] and $J_{a,k}$ is large enough, the new particle set still admits an appropriate approximation of the posterior PHD, namely

$$D_{a,k}(\mathbf{x})\approx\sum_{j=1}^{J_{a,k}}\mu_a\delta_{\mathbf{x}_{a,k}^{(j)}}(\mathbf{x})\,. \tag{20}$$

Arguably, these particles by assigning appropriate weights can approximate any PHD that has the same support space as $D_{a,k}(\mathbf{x})$. In particular, for $D_{b,k,\mathrm{T}}(\mathbf{x}), b\in\{a\}\cup\mathcal{N}_a$, we have

$$D_{b,k,\mathrm{T}}(\mathbf{x})\approx\sum_{j=1}^{J_{a,k}}w_b\big(\mathbf{x}_{a,k}^{(j)}\big)\delta_{\mathbf{x}_{a,k}^{(j)}}(\mathbf{x})\,. \tag{21}$$

where $w_b\big(\mathbf{x}_{a,k}^{(j)}\big)$ is the new weight assigned to particle $\mathbf{x}_{a,k}^{(j)}$.

Substituting (21) into (18) yields

$$\bar{D}_{a,k,\mathrm{T}}(\mathbf{x})\approx\sum_{j=1}^{J_{a,k}}\bar{w}_{a,k}^{(j)}\delta_{\mathbf{x}_{a,k}^{(j)}}(\mathbf{x}), \tag{22}$$

where

$$\bar{w}_{a,k}^{(j)}=\sum_{b\in\{a\}\cup\mathcal{N}_a}\omega_{b\to a}w_b\big(\mathbf{x}_{a,k}^{(j)}\big). \tag{23}$$

To determine $w_b\big(\mathbf{x}_{a,k}^{(j)}\big)$ to fulfill (21), we employ the classic IS approach. The idea of IS is to choose a proposal distribution $q(\mathbf{x})$ in place of the target probability distribution $p(\mathbf{x})$. The support of $q(\mathbf{x})$ is assumed to cover that of $p(\mathbf{x})$. Rewrite a general integration problem as

$$\int_{\mathcal{S}}f(\mathbf{x})p(\mathbf{x})d\mathbf{x}=\int_{\mathcal{S}}f(\mathbf{x})\frac{p(\mathbf{x})}{q(\mathbf{x})}q(\mathbf{x})d\mathbf{x}, \tag{24}$$

where $f(\mathbf{x})$ is an integrable function in a measurable space $\mathcal{S}$.

The IS [54, Chapter 3.3] is to use a number, to say J, of independent samples drawn from $q(\mathbf{x})$ to obtain a weighted sum to approximate (24):

$$\hat{f}_p=\frac{1}{J}\sum_{i=1}^{J}w(\mathbf{x}^{(i)})f(\mathbf{x}^{(i)}), \tag{25}$$

where the importance weights/ratios are

$$w(\mathbf{x}^{(i)})=\frac{p(\mathbf{x}^{(i)})}{q(\mathbf{x}^{(i)})}. \tag{26}$$

If both $q(\mathbf{x})$ and $p(\mathbf{x})$ are discrete, i.e., the random variable $\mathbf{x}$ can only take on discrete values from a set $\mathcal{X}$, $p(\mathbf{x}^{(i)})$ and $q(\mathbf{x}^{(i)})$ are actually known as the probability mass function (PMF) $\mathrm{pmf}_{\mathbf{x}}(\mathbf{y})$ of the discrete random variable $\mathbf{x}$, which is defined as

$$\mathrm{pmf}_{\mathbf{x}}(\mathbf{y})=\begin{cases}\Pr[\mathbf{x}=\mathbf{y}] & \text{if } \mathbf{y}\in\mathcal{X}\,,\\ \text{undefined} & \text{if } \mathbf{y}\notin\mathcal{X}\,.\end{cases} \tag{27}$$

Here, we extend the PMF definition from the discrete "probabilities" to the "intensity/PHD" so that the value is not limited to be in the scope of $[0,1]$, termed as intensity mass function (IMF), which reads

$$\mathrm{imf}_{\mathbf{x}}(\mathbf{y})=\begin{cases}D[\mathbf{x}=\mathbf{y}] & \text{if } \mathbf{y}\in\mathcal{X}\,,\\ \text{undefined} & \text{if } \mathbf{y}\notin\mathcal{X}\,.\end{cases} \tag{28}$$

where $D[\mathbf{x}=\mathbf{y}]$ is in sharp the weight of the particle $\mathbf{x}=\mathbf{y}$, which can be larger than 1.

Now, the uniformly weighted particles given by resampling as shown in (20) are just samples randomly drawn from the proposal $D_{a,k}(\mathbf{x})$ and the discrete set $\mathcal{X}$ in (28) is just the particle state set $\{\mathbf{x}_{a,k}^{(j)}\}_{j=1}^{J_{a,k}}$. To get the desired PHD distribution $D_{b,k,\mathrm{T}}(\mathbf{x})$ as shown in (21), it is as easy as weighting these particles by

$$w_b\big(\mathbf{x}_{a,k}^{(j)}\big)=\frac{\mathrm{imf}_{b,k,\mathrm{T}}\big(\mathbf{x}_{a,k}^{(j)}\big)}{\mathrm{imf}_{a,k}\big(\mathbf{x}_{a,k}^{(j)}\big)}. \tag{29}$$

where $\mathrm{imf}_{b,k,\mathrm{T}}\big(\mathbf{x}_{a,k}^{(j)}\big)$ is to evaluate the IMF at state $\mathbf{x}_{a,k}^{(j)}$ w.r.t. the GM-PHD disseminated from sensor $b\in\{a\}\cup\mathcal{N}_a$ as in (14), (15) and (9), which is given as

$$\mathrm{imf}_{b,k,\mathrm{T}}\big(\mathbf{x}_{a,k}^{(j)}\big)=\sum_{\mathbf{z}_k\in Z_{b,k,\mathrm{T}}}W_{b,k}(\mathbf{z}_k)\mathcal{G}\big(\mathbf{x}_{a,k}^{(j)};\hat{\mathbf{m}}_b(\mathbf{z}_k),\hat{\mathbf{P}}_b(\mathbf{z}_k)\big), \tag{30}$$

and

$$\mathrm{imf}_{a,k}\big(\mathbf{x}_{a,k}^{(j)}\big)=D_{a,k}[\mathbf{x}=\mathbf{x}_{a,k}^{(j)}], \tag{31}$$

which is nothing else but just the weight of the "mother" particle from which particle $\mathbf{x}_{a,k}^{(j)}$ is resampled, denoted as $w_{a,k|k}^{[j]}$. Therefore, to avoid repeated computing, we need to store the weights of the particles prior to resampling.

Substituting (31) and (30) into (29) yields

$$w_b\big(\mathbf{x}_{a,k}^{(j)}\big) \propto \frac{\sum_{\mathbf{z}_k \in Z_{b,k,\mathrm{T}}} W_{b,k}(\mathbf{z}_k)\mathcal{G}\big(\mathbf{x}_{a,k}^{(j)}; \hat{\mathbf{m}}_b(\mathbf{z}_k), \hat{\mathbf{P}}_b(\mathbf{z}_k)\big)}{w_{a,k|k}^{[j]}}, \tag{32}$$

which is subject to the cardinality consistence (for which, in fact, a separate cardinality AA consensus scheme will be applied in parallel as addressed next), i.e.,

$$\sum_{j=1}^{J_{a,k}} w_b\big(\mathbf{x}_{a,k}^{(j)}\big) = \sum_{\mathbf{z}_k \in Z_{b,k,\mathrm{T}}} W_{b,k}(\mathbf{z}_k). \tag{33}$$

Remark 3 A key for the success of the IS approach is that the support of the proposal $q(\mathbf{x})$ covers that of the target distribution $p(\mathbf{x})$, both distributions better of the similar shape. This is true in our case as that the posteriors obtained at local sensors corresponding to the same multitarget RFS are approximately identical in general.

D. Cardinality AA Consensus

In parallel to the above C&F procedure, the standard AA consensus [49], [50] is also applied to update the local weight sum at each communication iteration, namely cardinality AA consensus or simply cardinality consensus (CC), as follows:

$$\bar{W}_{a,k} = \sum_{b \in \{a\} \cup \mathcal{N}_a} \omega_{b \to a} W_{b,k}\,, \tag{34}$$

where the local weight sum $W_{b,k} = \sum_{j=1}^{J_{b,k}} w_{b,k}^{(j)}$.

The resulting new weight sum will be used for re-scaling the weight $\bar{w}_{a,k}^{(j)}$ of each particle given in (23), i.e., the after-consensus (AC) weight of particle j is

$$w_{a,k}^{(j),\mathrm{AC}} = \frac{\bar{w}_{a,k}^{(j)}}{\sum_{j=1}^{J_{a,k}} \bar{w}_{a,k}^{(j)}} \bar{W}_{a,k}\,. \tag{35}$$

E. Estimate Extraction

Estimates of the targets' states can be extracted in two means. One is carried out prior to the C&F and is purely based on the local particle-posterior, for which the usual estimate extraction procedures proposed for the centralized SMC-PHD filter such as multi-EAP [45] or other computing fast measurement-driven approaches e.g., [55], [56] which actually extract the means of the formed GC as in (14) of selected measurements corresponding to significant weights $W_{a,k}(\mathbf{z}_k)$, which is taken as the default way.

The advantage of this means is that, the estimate extraction does not need to "wait" for the C&F procedure and is therefore able to be carried out timely. The disadvantage is that, the latest information from the other sensors is not used, although the neighbor information has been used in the previous filtering iterations by C&F and is reflected in the prior. We refer to this as "*real-time/before-consensus (BC) estimation*".

The other means to extract estimates is carried out with regard to the disseminated GCs, taking into account the latest information from the other sensors. This is referred to as "*delayed/AC estimation*" as it can only be performed after the C&F. There are two typical ways to do so. One is merging closely-distributed GCs and extracting the mean(s) of the GC with larger weights. The other way is clustering the weighted GCs, and extracting the centroid of each significant cluster, which is taken as the default way in our approach. Either way, the number of estimates can be determined by the consensus on the cardinality which is rounding (34). To note, if the weight sum of one cluster or the weight of a single GC is closer to another integer $n \geq 2$ rather than 1 (indicating multiple targets in that cluster), n estimates should be extracted from that cluster or that GC.

There is still space for optimizing these estimate extraction algorithms on the basis of either the particle set or the disseminated GM. For example, for multisensor data clustering, it is useful to set constraints on the size of cluster [57], to avoid false alarm (e.g., a cluster of too small size) and to deal with overlapped clusters (e.g., a cluster of over large size because of closely-distributed targets). By this, the clustering scheme may automatically determine the number of estimates. Extensions in this regard are however beyond the focus of this work.

F. Parallelization of Local Filtering and C&F

In summary, a complete filtering iteration of the proposed distributed SMC-PHD filter is illustrated in Algorithm 1. We have an important note on the parallel processing of local filtering calculation and the C&F procedure.

Remark 4 The proposed IS which preserves the state of local particles renders filtering-C&F parallelization possible: in parallel to the network C&F at time k, some of local filtering calculations required for the filter iteration $k+1$ can be executed including

1) *Step 1 of Algorithm 1*: calculating the prior PHD as in (3), including generating new-born particles and propagating particles inherent from time k, and
2) *Step 2-1 of Algorithm 1*: calculating the likelihoods $g_{k+1}\big(\mathbf{z}_{a,k+1}|\mathbf{x}_{a,k+1|k}^{(j)}\big)$ and the detection probabilities $p_{\mathrm{D},k+1}\big(\mathbf{x}_{a,k+1|k}^{(j)}\big)$ for all particles $j = 1, 2, \cdots, J_{a,k}$, and the clutter intensities $\kappa_{k+1}(\mathbf{z}_{a,k+1})$ regarding all measurements $\mathbf{z}_k \in Z_{a,k}$.

The parallelization is feasible because all of these calculations do not need the knowledge of the particle weights until *Step 2-2* where the particle weights are needed and so the calculations thereafter can only be performed after (35).

G. Hybrid GM- and SMC-PHD Sensor Network

The proposed distributed SMC-PHD filtering protocol is naturally incorporable to the distributed GM-PHD filter [35]. That is, some sensors operate SMC-PHD filters while the others operate GM-PHD filters, both disseminating and receiving GMs. We demonstrate this hybrid filter network in our simulation in Section IV.

Algorithm 1 Distributed SMC-PHD Iteration at sensor a based on GM for dissemination and IS for fusion

---

Input:

- All statistical models as required, the measurement RFS $Z_{a,k}$ received at time k, and the posterior particle set $\xi_{a,k-1} \triangleq \{(\mathbf{x}_{a,k-1}^{(j)}, w_{a,k-1}^{(j)})\}_{j=1}^{J_{a,k-1}}$ for time $k-1$.

Output:

- State-estimates of the targets and a new particle set $\xi_{a,k} \triangleq \{(\mathbf{x}_{a,k}^{(j)}, w_{a,k}^{(j)})\}_{j=1}^{J_{a,k}}$.

Procedure:

Step 1 Existing particle propagation and new particle generation:

- Update the state of particles $\mathbf{x}_{a,k-1}^{(j)}$ to $\mathbf{x}_{a,k}^{(j)}$, $j = 1, \cdots, J_{a,k-1}$ according to the state transition model $\phi_{a,k|k-1}(\mathbf{x}|\mathbf{x}')$, preserving the weight $w_{a,k|k-1}^{(j)} = w_{a,k-1}^{(j)}$.
- Add new particles $\{(\mathbf{x}_{a,k}^{(j)}, w_{a,k|k-1}^{(j)})\}_{j=J_{a,k-1}+1}^{J'_{a,k}}$ according to the new-born target intensity model $\gamma_{a,k}(\mathbf{x})$.

Step 2 Particle weight updating:

- Re-weight all particles as in (6), yielding a new particle set $\xi'_{a,k} \triangleq \{(\mathbf{x}_{a,k}^{(j)}, w_{a,k}^{(j)})\}_{j=1}^{J'_{a,k}}$, consisting of two steps:
  1) Calculate the likelihood $g_k(\mathbf{z}_{a,k}|\mathbf{x}_{a,k}^{(j)})$ and detection probability $p_{\mathrm{D},k}(\mathbf{x}_{a,k}^{(j)})$ for each particle $j = 1, 2, \cdots, J'_{a,k}$, and the clutter intensity $\kappa_k(\mathbf{z}_{a,k})$ w.r.t. all measurements $\mathbf{z}_{a,k} \in Z_{a,k}$.
  2) Calculate the finally updated weight as in (6), taking into account the prediction weight $w_{a,k|k-1}^{(j)}$.

Step 3 Real-time/BC estimate extraction:

- Extract totally $[W_{a,k}]$ state-estimates as in (14) and (15) for $\mathbf{z}_k \in Z_{a,k}$ of higher weights $W_{a,k}(\mathbf{z}_k)$ as in (9).
- Determine $Z_{a,k,\mathrm{T}} \subset Z_{a,k}$ satisfying that $\forall \mathbf{z}_k \in Z_{a,k,\mathrm{T}}$: $W_{a,k}(\mathbf{z}_k) \geq T_{\mathrm{c}}$ - Threshold rule.

Step 4 Resampling:

- Resample from $\xi'_{a,k}$ to get a uniformly weighted new particle set $\xi_{a,k} \triangleq \{(\mathbf{x}_{a,k}^{(j)}, \mu_a)\}_{j=1}^{J_{a,k}}$ as in (20). Store the initial weight $w_{a,k}^{[j]}$ of the mother particle from which particle j is sampled.

Step 5 Partial consensus via IS:

- Extract a GC for each $\mathbf{z}_k \in Z_{a,k,\mathrm{T}}$ as in (14) and (15).
- Disseminate the GCs and $W_{a,k}$ to the neighbors and receive theirs; this step may be carried out for multiple iterations and if necessary (e.g., when the number of GCs exceeds a specific threshold), mixture reduction may be performed at some iterations.
- Calculate $\{w_b(\mathbf{x}_{a,k}^{(j)})\}_{j=1}^{J_{a,k}}$ w.r.t. all gathered GCs as in (32);
- Calculate $\{\bar{w}_{a,k}^{(j)}\}_{j=1}^{J_{a,k}}$ as in (23), and calculate $\bar{W}_{a,k}$ as in (34);
- Calculate $\{w_{a,k}^{(j),\mathrm{AC}}\}_{j=1}^{J_{a,k}}$ as in (35) as the final AC weight of each particle.

Step 6 Delayed/AC estimate extraction:

- Apply the $k-$means clustering on the gathered GCs and extract the centroid of each cluster as estimates, with the number of estimates given by $[\bar{W}_{a,k}]$.

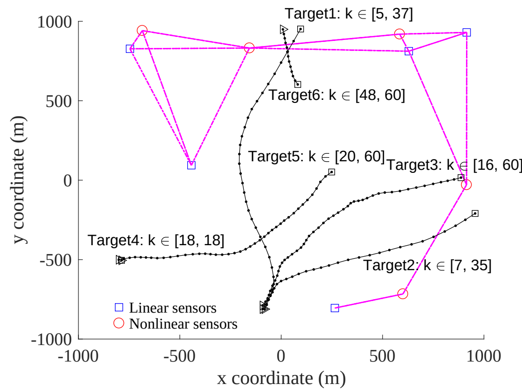


Fig. 1. Tracking scenario: target trajectories (starting at '△' and ending at '□') and a sensor network consisting of both linear and nonlinear sensors.

IV. Simulations

The simulations are set up in a scenario over the planar region $[-1000, 1000]\mathrm{m} \times [-1000, 1000]\mathrm{m}$ which is monitored fully by a connected undirected sensor network. The trajectories of totally 6 targets are given in Fig. 1 with the starting and ending times of each trajectory noted. The target birth process follows a Poisson RFS with intensity function $\gamma_k(\mathbf{x}) = \sum_{i=1}^{3} \lambda_i \mathcal{N}(.; \mathbf{m}_i, \mathbf{Q}_r)$, with Poisson rates $\lambda_1 = \lambda_2 = \lambda_3 = 0.05$ and the Gaussian parameters $\mathbf{m}_1 = [0, 0, 950, -30]^{\mathrm{T}}$, $\mathbf{m}_2 = [-100, 10, -800, 30]^{\mathrm{T}}$, $\mathbf{m}_3 = [-800, 20, -500, 0]^{\mathrm{T}}$, and $\mathbf{Q}_r = \mathrm{diag}([100, 25, 100, 25]^{\mathrm{T}})$, where $\mathrm{diag}(\mathbf{a})$ represents a diagonal matrix with diagonal $\mathbf{a}$.

Each target has a time-constant survival probability $p_{\mathrm{S}}(\mathbf{x}_k) = 0.98$ and the survival target follows a nearly constant velocity motion as given

$$\mathbf{x}_k = \begin{bmatrix} 1 & \Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} \Delta^2/2 & 0 \\ \Delta & 0 \\ 0 & \Delta^2/2 \\ 0 & \Delta \end{bmatrix} \mathbf{u}_k \,, \tag{36}$$

where $\mathbf{x}_k = [p_{x,k}, \dot{p}_{x,k}, p_{y,k}, \dot{p}_{y,k}]^{\mathrm{T}}$ with the position $[p_{x,k}, p_{y,k}]^{\mathrm{T}}$ and the velocity $[\dot{p}_{x,k}, \dot{p}_{y,k}]^{\mathrm{T}}$, the sampling interval $\Delta = 1\mathrm{s}$, and the state transition noise $\mathbf{u}_k \sim \mathcal{G}(\mathbf{0}_2, 25\mathbf{I}_2)$.

We deploy two different types of sensors, with either linear measurement models or nonlinear measurement models, as marked in Fig.1. The linear position measurement model is given as follows

$$\mathbf{z}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} v_{k,1} \\ v_{k,2} \end{bmatrix} \,, \tag{37}$$

with $v_{k,1}$ and $v_{k,2}$ as mutually independent zero-mean Gaussian noise with the same standard deviation of 10.

The FOV (field of view) of each nonlinear sensor is a disc of radius 3000m centralized with the sensor's position $[s_{n,x}, s_{n,y}]^{\mathrm{T}}$, which fully covers the scenario. The range and bearing measurement is given by

$$\mathbf{z}_k = \begin{bmatrix} \sqrt{(p_{x,k} - s_{n,x})^2 + (p_{y,k} - s_{n,y})^2} \\ \arctan\left((p_{y,k} - s_{n,y})/(p_{x,k} - s_{n,x})\right) \end{bmatrix} + \mathbf{v}_k \,, \tag{38}$$

where $\mathbf{v}_k \sim \mathcal{N}(; \mathbf{0}, \mathbf{R}_k)$, with $\mathbf{R}_k = \mathrm{diag}\left([\sigma_r^2, \sigma_\theta^2]^{\mathrm{T}}\right), \sigma_r = 10\mathrm{m}$, $\sigma_\theta = \pi/90$ rad/s.

The linear sensors have the same and constant target detect probability $p_{\mathrm{D}}(\mathbf{x}_k) = 0.95$ while the nonlinear sensors have $p_{\mathrm{D}}(\mathbf{x}_k) = 0.95\mathcal{N}([|p_{x,k} - s_{n,x}|, |p_{y,k} - s_{n,y}|]^{\mathrm{T}}; \mathbf{0}, 6000^2 I_2)/\mathcal{N}(0; 0, 6000^2 I_2)$. Clutter is uniformly distributed over each sensor's FOV with an average rate of 10 points per scan, which indicates clutter intensity $\kappa_k = 10/2000^2$ for the linear sensors and $\kappa_k = 10/3000/2\pi$ for the nonlinear sensors.

Two scenarios have been considered. First, the sensor network consists of only SMC-PHD filters, namely a pure SMC-PHD filter network. Second, each linear sensor operates a GM-PHD filter while each nonlinear sensor operates a SMC-PHD filter, namely a hybrid network consisting of both GM-PHD and SMC-PHD filters. In the next subsection, different C&F schemes are designed for comparison with our approach.

For mixture reduction regarding to the GM: GCs with a weight lower than 10^{-4} will be truncated, any two GCs closer than Mahalanobis-distance $\tau = 4$ will be merged and the maximum number of significant GCs to be transmitted and owned by a sensor is 100. The GC is identified as a significant GC if its weight is larger than $T_{\mathrm{c}} = 0.45$ and will be transmitted among neighbors.

The optimal sub-pattern assignment (OSPA) metric [58] is used to evaluate the estimation accuracy of the filter, with cut-off parameter $c = 1000$ and order parameter $p = 2$. Furthermore, we define

- *Network OSPA*: the average of OSPAs obtained by all sensors in the network at each sampling step;
- *Time-average Network OSPA*: the average of the Network OSPAs over all filtering steps.

To evaluate the communication cost, we record a GC that consists of a weight parameter (1 tuple), a 4-dimensional vector mean (4 tuples), and a 4×4 -dimension symmetric matrix covariance (10 tuples) as 15 tuples and the scale-valued cardinality parameter as 1 tuple. In addition, each weighted particle takes 5 tuples (4 for the state vector and 1 for the weight). For the SMC-PHD filter, $N_{\mathrm{p}} = 200$ particles are assigned for each expected target during the resampling scheme to adjust the number of particles in time series. It is worth noting that strategies such as roughening [59] that is to add a small zero-mean random variable to the state of each resampled particle [5], is useful to increase the diversity of particles after resampling for the PF and is adopted in our implementations.

Each simulation was performed 100 runs with independently generated measurement series, each run consisting of 100 filtering iterations. Different numbers t of neighbor/P2P communication iterations from 0 (without applying any communication between sensors) to 10 or 5 are applied to all consensus schemes.

A. Comparison approachs

1) GM-EMD-IS: The state-of-the-art distributed SMC-PHD fusion given in [27] is based on EMD, which consists of two essential parts: 1) convert the particle distribution to continuous distribution approximations and 2) construct the multitarget EMD. In the former, the work [27] proposed a clustering approach for kernel learning which according to our experience is computationally intensive and unstable. Instead, in our implementation, we use the proposed P2GM strategies (with a much lower threshold $T_{\mathrm{c}} = 0.1$ for selecting a sufficiently large number of GCs for accurate approximation) for generating the continuous kernel distribution (namely GM). In the latter, to construct the multitarget EMD, the IS approach can be applied as is done in our approach to update the fused particles. But there are two key differences:

- In our approach, there is no particle communication between sensors. In the EMD approach [27], the particles are the union of the particle sets that are sampled from neighbor sensors. Because of this, it needs to disseminate both particles and the corresponding kernel/GM function parameters at each communication iteration;
- In our approach the target density is the AA of disseminated GMs while for EMD it is the GA of local PHDs.

More specifically, denoting the union of uniformly weighted particle sets from neighbor sensors as $\left\{\mathbf{x}_{a,k}^{(j)}, \mu_a\right\}_{j=1}^{J_{a,k}}$ (we will present a particle resampling dissemination approach later, which is used here as well), the EMD-IS approach determines the weight of each particle as follows (cf. (23)):

$$\bar{w}_{a,k}^{(j)} \propto \prod_{b\in\{a\}\cup\mathcal{N}_a} \left(w_b\left(\mathbf{x}_{a,k}^{(j)}\right)\right)^{\omega_{b\to a}}, \tag{39}$$

where $w_b\left(\mathbf{x}_{a,k}^{(j)}\right)$ is coherent with (32).

Also, different to our cardinality AA consensus (34), the cardinality consensus is implemented in the GA sense, i.e.

$$\bar{W}_{a,k} = \prod_{b\in\{a\}\cup\mathcal{N}_a} \left(W_{b,k}\right)^{\omega_{b\to a}}. \tag{40}$$

However, we did not employ the sophisticated fusion weight strategy proposed in [27] but still applied the Metropolis weights, which have been widely used for the GCI based GM-PHD fusion since [22]. Therefore, the resulting EMD fusion approach is indeed a modified implementation of that given in [27] and is referred to as "GM-EMD-IS".

Remark 5 The logarithmic fusion is arguably more compelling than the linear fusion for dealing with multi-sensor *likelihood* fusion as it is external Bayesian [18]. However, as addressed in this paper, the items to be fused are the unknown-cross-correlated *PHD*s (which is more meaningful to calculate the AA, cf. Remark 2) and the sum of fusion weights is presumed being unity, Bayesian optimality is impossible. Instead, the problem of false and missing data is more prominent in distributed MTT. This together the consideration on communication and computation form the essential reason we advocate AA rather than GA for *"average consensus"*.

2) Particle Resampling Dissemination (PRD): In contrast to the GA that rarely admits closed-form solution for mixture distributions, an exact solution for AA of PHDs can be easily given by disseminating particles. That is, different to the partial consensus as in (18) and (45), complete consensus is sought by linearly averaging the local PHD $D_{a,k}(\mathbf{x})$ at sensor a with that from the neighbors $D_{b,k}(\mathbf{x}), \forall b \in \mathcal{N}_a$, i.e.,

$$\bar{D}_{a,k}(\mathbf{x}) = \sum_{b\in\{a\}\cup\mathcal{N}_a} \omega_{b\to a} D_{b,k}(\mathbf{x}). \tag{41}$$

However, disseminating all particles requires tremendous communication and will unnecessarily lead to an increasing number of particles at local sensors. To maintain a stable number of particles and to reduce the communication cost, the number $J_{b\to a,k}$ of transmitting particles from sensor $b \in \mathcal{N}_a$ to sensor a can be determined as

$$J_{b\to a,k} = [N_{\mathrm{p}} W_{b\to a,k}], \tag{42}$$

where $N_{\mathrm{p}} = 200$ as specified and $W_{b\to a,k} = \omega_{b\to a} W_{b,k}$, relying on the particle weight sum of the particle-sending sensor b and fusion weights to the particle-receiving sensor a; for $\omega_{b\to a}$, we apply the Metropolis weights again.

These particles are obtained by unbiased (re)sampling retaining the identical representation of the intial particle posterior [60]. Each resampled particle is weighted as

$$w^{(j)}_{b\to a,k} = \frac{W_{b\to a,k}}{J_{b\to a,k}} \approx \frac{1}{N_{\mathrm{p}}}. \tag{43}$$

It is straightforward to have the weight sum of all particles received at sensors (cf. (34))

$$\sum_{b\in\{a\}\cup\mathcal{N}_a} \sum_{j=1}^{J_{b\to a,k}} w^{(j)}_{b\to a,k} = \sum_{b\in\{a\}\cup\mathcal{N}_a} W_{b\to a,k} = \bar{W}_{a,k} \tag{44}$$

We detail this protocol in *Appendix B*, with reference to Algorithm 1. As shown, the resampling scheme plays a key role for "selecting" the particles for dissemination, which strives to gain a trade-off between communication cost and approximation accuracy. We refer to this protocol as PRD. Particularly different to the other protocols, no GM is involved. However, for estimate output, we still apply the similar estimate extraction solution as in (14), (15) and (9) from the particle distribution. Then, the component of weight greater than $T_{\mathrm{e}} = 0.5$ will be identified as a target estimate.

3) GM-Re-Sampling (RS): In place of the IS approach, new particles can be re-generated from the disseminated GMs via a standard sampling algorithm, namely recovering particles from GM after the C&F procedure. We refer to this approach as GM-RS, which is realized for the specific purpose to evaluate the effectiveness of the proposed IS approach.

Whether RS or IS, the GM can be merged to control the mixture size or simply remain unchanged (to avoid any fusion error or information double-counting) during their dissemination at each communication iteration, as addressed in Remark 1. In the latter, it is commonly referred to as *distributed flooding* [7] while in the former we use the *conservative GM merging* (CGMM) scheme [35]. Overall, on disseminating the GMs, there are different combinations of strategies: either RS or IS and either flooding or CGMM.

4) CC-AA: The cardinality consensus based on AA (CC-AA) approach only disseminates the weight sum information between sensors to re-scale the weight of particles.

5) No consensus: We also realize a centralized protocol in which all local filters do not share any information with each other referred to as "no consensus".

As addressed in Section III-E, all these distributed SMC-PHD filtering protocols, except the non-consensus protocol, can extract estimates at two different stages: the BC manner and the AC manner, as addressed in Step 3 and Step 6 in Algorithm 1, respectively. In the former, we simply extract the mean of the extracted GMs as in (14) as the estimates while the number of estimates is determined by (12), which resembles the approaches given in [55], [56]. In the latter, we apply the well-known $k-$means clustering algorithm to the disseminated GCs for estimate extraction, where the number of clusters is specified by rounding (34) and the estimates are given as the center of the formed clusters.

B. WSN of Pure SMC-PHD Filters

In this case, each sensor runs a SMC-PHD filter individually, using either the linear measurement model/data or the nonlinear measurement model/data.

When $t = 5$, the Network OSPA, the online estimated number of targets, and the computing time of different consensus protocols for each filtering step are given in the upper row of Fig. 2, separately. For different numbers of communication iterations from $t = 0$ to $t = 10$, the time-averaged network OSPA and network communication of different consensus protocols are given in the bottom row of Fig. 2, separately. We have the following key findings with regard to the filtering accuracy, communication and computation cost, respectively:

1) On filtering accuracy:

- All C&F protocols improve the filtering accuracy by reducing the filter OSPA as compared to the protocol applying no consensus. The OSPA reduction basically increases with that of the number of the P2P communication iterations, till to a convergent/consensual level.
- Except CC-AA for which both BC and AC outputs are the same, all AC estimates are significantly better than the BC estimates indicating that the C&F is immediately beneficial for improving the local filter accuracy at each filtering step.
- The GM flooding protocols that avoids information double-counting outperform the CGMM protocol.
- As expected, the CC-AA which is the least communicatively and computationally costly, benefits the filter the least among all distributed C&F schemes.
- Roughly speaking, the IS approaches perform similar to the RS approaches in each corresponding category when applying CGMM/Flooding or in BC/AC manners. More precisely, the RS approaches outperform the IS approaches in the BC estimation case (whether Flooding or CGMM is applied), performs very similar in the CGMM-AC case and inferior in the GM flooding-AC case, as compared with the IS approaches.
- Surprisingly, all GM based C&F approaches including CGMM/GM-flooding combined with RS/IS, and EMD, outperform the PRD approach, in both BC and AC manners. The result is intuitively surprising because, despite the communication cost, the particles set contains more complete information of the posterior PHD than the GM extracted from particles. However, *more does not alway mean better*. The partial consensus by which the sensors share only the significant components of the GM rather than all, is supposed to be help reduce the affection of

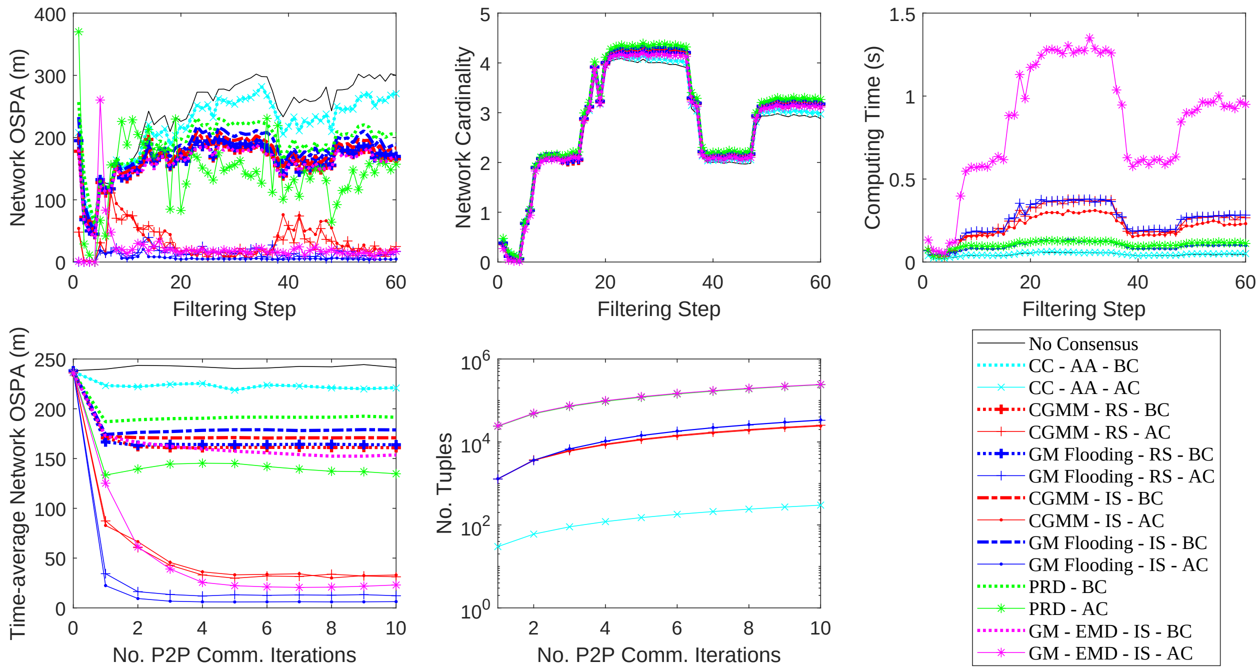


Fig. 2. Average performance of 10 SMC-PHD filters when different number of P2P communication iterations are performed between them. Their connection is shown in Fig.1. The upper row: *Network OSPA* (shown in the left), *online estimated number of targets* (middle) and *computing time* (right) of different consensus protocols for each filtering step, respectively, when 5 P2P communication iterations are applied between neighbors. The bottom row: *Time-averaged network OSPA* (shown in the left) and *network communication cost* (middle) for different numbers of communication iterations. CC-AA: Cardinality consensus based on arithmetic average; CGMM: Conservative Gaussian mixture merging [35], PRD: Particle resampling dissemination; RS: Re-sampling; IS: Importance sampling; BC: Before consensus; AC: After consensus; EMD: Exponential mixture density.

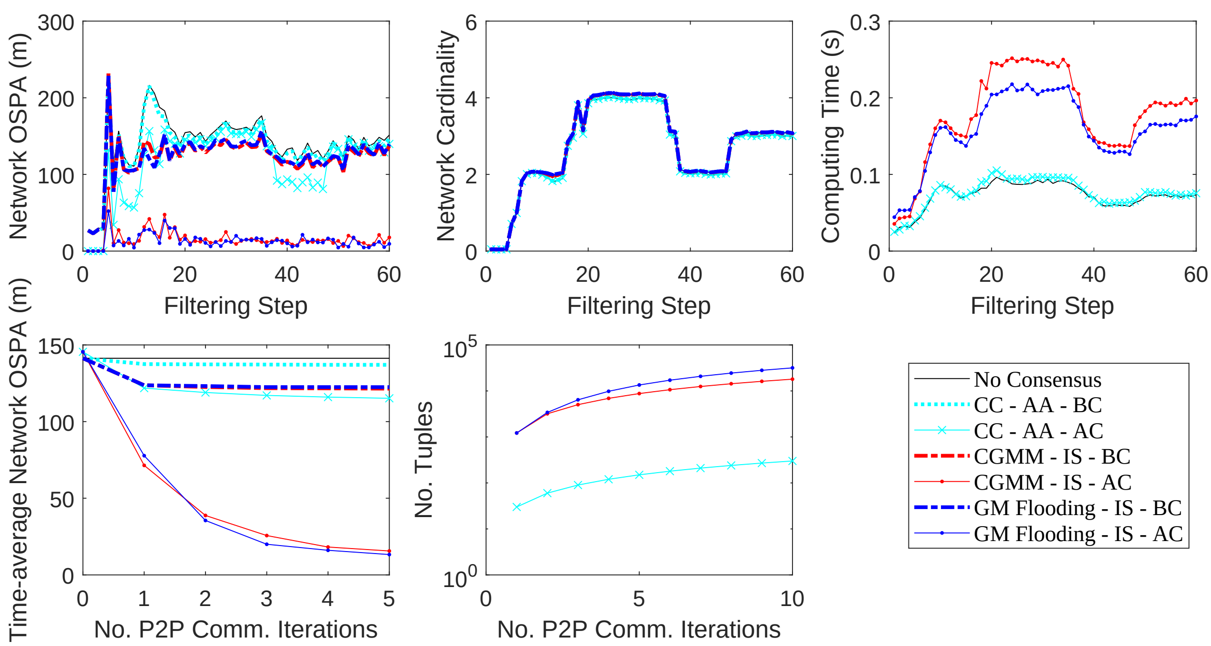


Fig. 3. Average performance of 5 SMC-PHD filters and 5 GM-PHD filters when different number of communication iterations are performed between them. The upper row: *Network OSPA* (shown in the left), *online estimated number of targets* (middle) and *computing time* (right) of different consensus protocols for each filtering step, respectively, when 5 P2P communication iterations are applied. The bottom row: *Time-averaged network OSPA* (shown in the left) and *network communication cost* (middle) for different numbers of communication iterations. CC-AA: Cardinality consensus based on arithmetic average; CGMM: Conservative Gaussian mixture merging [35], IS: Importance sampling; BC: Before consensus; AC: After consensus.

the false alarm, as has been illustrated in [35]. This leads to an unique advantage that the PRD approach or the complete consensus does not have.

- On the AC estimation, the GM-EMD-IS approach performs slightly better than the CGMM-(RS/IS) approaches, inferior to the GM flooding-(RS/IS) approach, while on the BC estimation, the GM-EMD-IS approach performs favorably which is better than all the others (especially for large t).

2) On communication:

- Both the GM-EMD-IS approach and the PRD protocol, which disseminate a large number of weighted particles are unsurprisingly the most communicatively costly.
- Except $t = 1$, the flooding is slightly more communicatively costly than the CGMM which applies mixture reduction during communication.
- The CC-AA approach is computationally ignorable.

3) On computation:

- The EMD approach is the most computationally costly while the CC-AA is the least and is ignorable.
- The CGMM approaches are more computationally costly than the flooding approaches.
- The RS approach is even more computationally costly than the IS approach.
- The PRD approach and the GM flooding-IS approach are similar in computational cost and are very efficient.

With particular regard to the proposed P2GM based IS approach, we have the following conclusions

- The flooding-communication approach achieves the best accuracy benefit among all; in fact, it reaches very close to the maximal accuracy gain that the algorithm converges to by only 1 or 2 communication iterations;
- When CGMM is applied during the C&F, the communication cost will be reduced somewhat (depending on how the merging threshold is set) while the computational cost is increased and the accuracy benefit is reduced.
- The IS approach performs similar to the RS approach in accuracy and communication cost, but computes faster. The RS approach prevents any parallelization of the filtering calculation and the C&F.
- The GA-based EMD approach is significantly more communicatively and computationally intensive than AA based approaches while it only shows insignificant superiority in improving the BC estimation accuracy.

C. Hybrid WSN of SMC-PHD and GM-PHD filters

In this case, we study a hybrid sensor network: each linear sensor operates a GM-PHD filter [35] while each nonlinear sensor operates a SMC-PHD filter. However, the PRD, GM-RS and GM-EMD-IS approaches are based to the PF and do not apply to the GM-PHD filter and therefore, will not be realized here. To integrate the disseminated GM into the local posterior, rather than the IS approach for the SMC-PHD filter (where the local posterior is represented by particles), straightforward GM union is applied for the GM-PHD filter (where the local posterior is a GM). To note, in the distributed GM-PHD filter [35], the fused AA PHD at sensor a is given by linearly averaging the initial complete posterior PHD $D_{a,k}(\mathbf{x})$ (a GM) with the received partial PHD (also a GM, which only represents a part of the corresponding posterior) from the neighbors $D_{b,k,\mathrm{T}}(\mathbf{x}), \forall b \in \mathcal{N}_a$, i.e., (cf. (18))

$$\bar{D}_{a,k}(\mathbf{x}) = \omega_{a\to a} D_{a,k}(\mathbf{x}) + \sum_{b\in\mathcal{N}_a} \omega_{b\to a} D_{b,k,\mathrm{T}}(\mathbf{x}), \quad (45)$$

which is slightly different to the AA implemented for the SMC-PHD filter in (18) as here, the local sensor contributes the whole PHD $D_{a,k}(\mathbf{x})$ rather than only the partial PHD.

To show the simulation result, similar contents (up to $t = 5$) given in Fig. 3 correspond to those in Fig. 2, respectively. The results are highly consistent, confirming the effectiveness of our approach for the hybrid filter network. For example,

- All C&F approachs converge with the increase of the number of P2P communication iterations; the proposed IS-AC approach demonstrates again fast convergence;
- The AC estimation is more accurate than the BC estimation in all approachs;
- The CC-AA is ignorable in either computation or communication;
- The CGMM scheme saves communication but costs more computation as compared to the GM flooding approach.

However, different to what shown in the last simulation, the GM flooding-IS approach performs very similar with the CGMM-IS approach (except $t = 1$) in the sense of OSPA reduction. We conjecture that this is because the CGMM does not cause approximation error for the GM-PHD filter as significantly as it did to the SMC-PHD filter based on P2GM. Therefore, appropriate merging does not have to sacrifice the filter accuracy.

V. Conclusion

We present a "partial, arithmetic average consensus" approach to distributed SMC-PHD fusion. Our approach is composited of two major parts. One part is regarding particles-to-GM conversion, which constructs a GM from the particle set at each local sensor for parameterized information dissemination. The GM represents only the significant part of the particle posterior rather than the complete, for a trade-off between approximation accuracy and communication cost. The disseminated GMs are linearly/arithmetically averaged over the network for consensus. The other part is an importance sampling approach for re-weighting the local particles according to the disseminated GM without changing their states. This allows parallel implementation of the local calculation for the prior PHD and likelihood, and the network communication, combating or even avoiding time delay to the filter. The effectiveness and reliability of our approach have been demonstrated in both regards: the particles-to-GM conversion and the importance sampling.

The proposed distributed SMC-PHD filter can be seamlessly cooperated with the distributed GM-PHD filter, leading to a promising, hybrid sensor framework in which each sensor, using whether linear or nonlinear measurement model, may operate a PHD filter implemented by means of either GM or SMC according to its realistic needs or conditions.

Appendix A
Cardinality Estimation Conditioned on PHD

Denoting the PDF of the multi-target RFS variable $X_k = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ as $f(X_k)$, we have

$$\rho(n) = \int_{|X_k|=n} f(X_k)\delta X_k \tag{46}$$

From (5) and $\int_{\mathbb{R}^d} \delta_{\mathbf{x}_k^{(j)}}(\mathbf{x})d\mathbf{x} = 1$, we have

$$\int_{\mathbb{R}^d} D_k(\mathbf{x})d\mathbf{x} \approx W_k \tag{47}$$

Substituting (1) and (46) to (47) yields

$$\begin{aligned} \int_{\mathbb{R}^d} D_k(\mathbf{x})d\mathbf{x} &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \delta_{X_k}(\mathbf{x})f(X_k)\delta X_k d\mathbf{x} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \delta_{X_k}(\mathbf{x})d\mathbf{x}f(X_k)\delta X_k \\ &= \int_{\mathbb{R}^d} nf(X_k)\delta X_k \\ &= \sum_{n\geq 0} n\rho(n) \end{aligned} \tag{48}$$

Estimation (48) is known as expected a posteriori (EAP) estimate of the number of targets at time k, i.e.,

$$\hat{N}_k^{\text{EAP}} = \int_{\mathbb{R}^d} D_k(\mathbf{x})d\mathbf{x} \approx W_k. \tag{49}$$

Appendix B
PRD based Distributed SMC-PHD Filter

Algorithm 2 Distributed SMC-PHD Iteration at sensor a based on particle resampling and dissemination

Input and **Output** are the same to Algorithm 1.

Procedure:

Step 1-Step 3 are the same as that of Algorithm 1.

Step 4 Resampling:

- Calculate the local particle weight sum $W_{a,k}$.
- Resample from $\xi'_{a,k}$ to get $|\mathcal{N}_a| + 1$ new particle sets $\xi_{a\to b,k} \triangleq \{\mathbf{x}_{a,k}^{(j)}, \mu_a\}_{j=1}^{J_{a\to b,k}}$ for $b \in \{a\} \cup \mathcal{N}_a$ as in (42) and (43).

Step 5 Partial consensus via particle dissemination:

- Disseminate the weighted particles $\xi_{a\to b,k}$ together with parameter $W_{a\to b,k}$, and the local BC estimates yielded in Step 3 to sensor $b \in \mathcal{N}_a$; simultaneously, receive their transmissions.
- Calculate $\bar{W}_{a,k}$ as in (34) and use it to replace $W_{a,k}$.
- The resulting particle set $\xi_{a,k}$ is given by

$$\xi_{a,k} = \cup_{b\in\{a\}\cup\mathcal{N}_a} \xi_{b\to a,k} \tag{50}$$

Step 6: *Iteration, if necessary*:

- Steps 4 and 5 may be carried out for multiple iterations (set $\xi'_{a,k} \leftarrow \xi_{a,k}$ before redo resampling).

Step 7 Delayed/AC estimate extraction:

- Similar to Step 6 of Algorithm 1, but differently the items to be clustered are the BC estimates obtained in Step 3.

Acknowledgment

This draft of this work has been discussed with Prof. Franz Hlawatsch who directed the author's secondment at the Institute of Telecommunications, Vienna University of Technology from July to September 2017. The author would like to acknowledge his comments and suggestion. This work has also been presented in part at the ATR laboratory in NUDT, Changsha China on Dec. 12 2017, which was invited by Dr. Hongqi Fan.

References

[1] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Found. Trends in Machine Learn.*, vol. 7, no. 4-5, pp. 311–801, 2014.

[2] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.

[3] M. H. Degroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[4] C. Genest and J. V. Zidek, "Combining probability distributions: A critique and an annotated bibliography," *Statistical Science*, vol. 1, no. 1, pp. 114–135, 1986.

[5] X. Wang, T. Li, S. Sun, and J. M. Corchado, "A survey of recent advances in particle filters and remaining challenges for multitarget tracking," *Sensors*, vol. 17, no. 12, 2017.

[6] O. Hlinka, F. Hlawatsch, and P. M. Djuric, "Distributed particle filtering in agent networks: A survey, classification, and comparison," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 61–81, Jan 2013.

[7] T. Li, J. Corchado, and J. Prieto, "Convergence of distributed flooding and its application for distributed Bayesian filtering," *IEEE Trans. Signal Inf. Process. Net.*, vol. 3, no. 3, pp. 580–591, 2017.

[8] J. Li and A. Nehorai, "Distributed particle filtering via optimal fusion of Gaussian mixtures," *IEEE Trans. Signal Inf. Process. Net.*, vol. PP, no. 99, pp. 1–1, 2017.

[9] B.-N. Vo, M. Mallick, Y. Bar-shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-T. Vo, *Multitarget Tracking*. John Wiley & Sons, Inc., 2015.

[10] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Trans. Signal Proc.*, vol. 65, no. 13, pp. 3478–3493, July 2017.

[11] 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.

[12] B. Ristic, M. Beard, and C. Fantacci, "An overview of particle methods for random finite set models," *Inf. Fusion*, vol. 31, pp. 110 – 126, 2016.

[13] R. Mahler, "Approximate multisensor CPHD and PHD filters," in *2010 13th Int. Conf. Inf. Fusion*, July 2010, pp. 1–8.

[14] S. Nannuru, S. Blouin, M. Coates, and M. Rabbat, "Multisensor CPHD filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 4, pp. 1834–1854, August 2016.

[15] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, *Belief Consensus and Distributed Hypothesis Testing in Sensor Networks*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 169–182.

[16] M. Bacharach, "Normal bayesian dialogues," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 837–846, 1979.

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

[18] S. Bandyopadhyay and S.-J. Chung, "Distributed Bayesian filtering using logarithmic opinion pool for dynamic sensor networks," *arXiv:1712.04062 [cs.SY]*, 2017.

[19] J. K. Uhlmann, "Dynamic map building and localization: new theoretical foundations," Ph.D. dissertation, University of Oxford, 1995.

[20] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the 1997 American Control Conference*, Jun 1997, pp. 2369–2373.

[21] M. B. Hurley, "An information theoretic justification for covariance intersection and its generalization," in *Proceedings of the Fifth Int. Conf. Inf. Fusion*, July 2002, pp. 505–511.

[22] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE J. Sel. Topics Signal Process*, vol. 7, no. 3, pp. 508–520, June 2013.

[23] 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, December 2016.

[24] W. Yi, M. Jiang, S. Li, and B. Wang, "Distributed sensor fusion for RFS density with consideration of limited sensing ability," in *20th Int. Conf. Inf. Fusion*, July 2017, pp. 1–6.

[25] N. Forti, G. Battistelli, L. Chisci, S. Li, B. Wang, and B. Sinopoli, "Distributed joint attack detection and secure state estimation," *IEEE Trans. Signal Inf. Process. Net.*, vol. PP, no. 99, pp. 1–1, 2017.

[26] M. Uney, D. E. Clark, and S. J. Julier, "Information measures in distributed multitarget tracking," in *14th Int. Conf. Inf. Fusion*, July 2011, pp. 1–8.

[27] ——, "Distributed fusion of PHD filters via exponential mixture densities," *IEEE J. Sel. Topics Signal Process*, vol. 7, no. 3, pp. 521–531, June 2013.

[28] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707 – 718, 2014.

[29] R. P. S. Mahler, "Optimal/robust distributed data fusion: a unified approach," in *Proc. SPIE*, vol. 4052, 2000, pp. 128–138.

[30] D. Clark, S. Julier, R. Mahler, and B. Ristic, "Robust multi-object sensor fusion with unknown correlations," in *Sensor Signal Process. for Defence (SSPD 2010)*, Sept 2010, pp. 1–5.

[31] R. P. S. Mahler, "Toward a theoretical foundation for distributed fusion," in *Distributed Data Fusion for Network-Centric Operations*, D. Hall, C.-Y. Chong, J. Llinas, and M. L. II, Eds. CRC Press, 2012, pp. 199–224.

[32] 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 Proc.*, vol. 65, no. 1, pp. 242–255, Jan 2017.

[33] J. Y. Yu, M. Coates, and M. Rabbat, "Distributed multi-sensor CPHD filter using pairwise gossiping," in *2016 IEEE Int. Conf. Acoustics, Speech and Signal Process*, March 2016, pp. 3176–3180.

[34] I. Hwang, K. Roy, H. Balakrishnan, and C. Tomlin, "A distributed multiple-target identity management algorithm in sensor networks," in *43rd IEEE Conference on Decision and Control*, 2004, pp. 728–734.

[35] T. Li, J. Corchado, and S. Sun, "Partial consensus and conservative fusion of gaussian mixtures for distributed PHD fusion," *arXiv:1711.10783 [cs.SY]*, 2017.

[36] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, April 1996.

[37] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Trans. Signal Proc.*, vol. 56, no. 7, pp. 3229–3235, July 2008.

[38] M. Godavarti and A. O. Hero, "Partial update LMS algorithms," *IEEE Trans. Signal Proc.*, vol. 53, no. 7, pp. 2382–2399, July 2005.

[39] R. Arablouei, S. Werner, Y. F. Huang, and K. Doanay, "Distributed least mean-square estimation with partial diffusion," *IEEE Trans. Signal Proc.*, vol. 62, no. 2, pp. 472–484, Jan 2014.

[40] B. N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, Oct 2005.

[41] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary particle implementation of probability hypothesis density filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1437–1454, July 2010.

[42] Y. Petetin, M. Morelande, and F. Desbouvries, "Marginalized particle PHD filters for multiple object Bayesian filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1182–1196, April 2014.

[43] M. Schikora, A. Gning, L. Mihaylova, D. Cremers, and W. Koch, "Box-particle probability hypothesis density filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 3, pp. 1660–1672, July 2014.

[44] T. Li, S. Sun, M. Boli, and J. M. Corchado, "Algorithm design for parallel implementation of the SMC-PHD filter," *Signal Process.*, vol. 119, pp. 115 – 127, 2016.

[45] T. Li, J. M. Corchado, S. Sun, and H. Fan, "Multi-EAP: Extended EAP for multi-estimate extraction for SMC-PHD filter," *Chin. J. Aeronaut.*, vol. 30, no. 1, pp. 368 – 379, 2017.

[46] M. Coates, "Distributed particle filters for sensor networks," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*. New York, NY, USA: ACM, 2004, pp. 99–107.

[47] X. Sheng, Y. H. Hu, and P. Ramanathan, "Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network," in *Fourth International Symposium on Information Processing in Sensor Networks*, April 2005, pp. 181–188.

[48] D. Gu, "Distributed particle filter for target tracking," in *Proceedings 2007 IEEE Int. Conf. Robot. Autom.*, April 2007, pp. 3856–3861.

[49] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65 – 78, 2004.

[50] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *4th Int. Symp. Inf. Process. Sensor Net.*, April 2005, pp. 63–70.

[51] T. Li, J. Corchado, and S. Sun, "On generalized covariance intersection for distributed phd filtering and a simple but better alternative," pp. 808–815, July 2017.

[52] M. A. Bakr and S. Lee, "Distributed multisensor data fusion under unknown correlation and data inconsistency," *Sensors*, vol. 17, no. 11, 2017.

[53] T. Li, M. Bolić, and P. M. Djurić, "Resampling methods for particle filtering: Classification, implementation, and strategies," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 70–86, May 2015.

[54] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.

[55] L. Zhao, P. Ma, X. Su, and H. Zhang, "A new multi-target state estimation algorithm for PHD particle filter," in *2010 13th Int. Conf. Inf. Fusion*, July 2010, pp. 1–8.

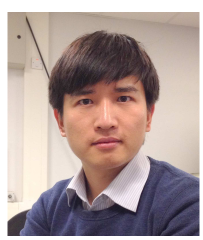
[56] M. Schikora, W. Koch, R. Streit, and D. Cremers, "Sequential monte carlo method for multi-target tracking with the intensity filter," *Advances in Intelligent Signal Process. and Data Mining: Theory and Applications*, vol. 410, pp. 55–87, 2012.

[57] T. Li, F. D. la Prieta Pintado, J. M. Corchado, and J. Bajo, "Multi-source homogeneous data clustering for multi-target detection from cluttered background with misdetection," *Applied Soft Computing*, vol. 60, pp. 436 – 446, 2017.

[58] D. Schuhmacher, B. T. Vo, and B. N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Proc.*, vol. 56, no. 8, pp. 3447–3457, Aug 2008.

[59] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEEE Proc. F Radar Signal Proc.*, vol. 140, no. 2. IET, 1993, pp. 107–113.

[60] T. Li, G. Villarrubia, S. Sun, J. M. Corchado, and J. Bajo, "Resampling methods for particle filtering: identical distribution, a new method, and comparable study," *Front Inform Technol Electron Eng*, vol. 16, no. 11, pp. 969–984, 2015.



Tiancheng Li received his Bachelors degree in Mechanical and Electrical Engineering, with a minor degree in Business Administration, from Harbin Engineering University (China), in 2008, his first PhD in Electrical and Electronic Engineering from London South Bank University (United Kingdom), in 2013 and his second Doctoral degree in Mechatronics Engineering from Northwestern Polytechnical University (China), in 2015. He has been a Research Associate with the BISITE Research Group, University of Salamanca (Spain) since June 2014 and presently, he holds a Marie Skodowska-Curie Individual Fellowship (H2020-MSCA-IF-2015) with the Host University of Salamanca and the Partner Vienna University of Technology (Austria).

He has been on the Editorial Board of two peer-reviewing journals: FITEE (2017-, Frontiers of Information Technology & Electronic Engineering) and ADCAIJ (2016-, Advances in Distributed Computing and Artificial Intelligence Journal) as well as on the organizing and/or technical program committee of several international symposiums and conferences including FUSION 2014-17, ACM-SAC 2015-17 and DCAI 2015-18.

His research interests lie in the general area of statistical signal processing and distributed information fusion, with particular emphasis on novel Markov-free solutions for realistic multisensor multiobject detection, tracking and forecasting by means of sensor data clustering, fitting and mining.