Label Matching: It Is Complicated

Kuangyu Di[1], Tiancheng Li[1], Guchong Li[1], Yan Song[1], Xudong Dang[2]
[1]*Key Lab. Inform. Fusion Tech.*, *Northwestern Polytechnical University*, Xi'an, China
[2]*National Key Laboratory of Radar Signal Processing*, *Xidian University*, Xi'an, China
kydi@mail.nwpu.edu.cn, {t.c.li,guchong.li}@nwpu.edu.cn,syzx@mail.nwpu.edu.cn, dangxudong@xidian.edu.cn

***Abstract*—This paper addresses the intractable track matching problem involved in multi-sensor multi-target tracking using the labeled multi-Bernoulli filters. Unlike the unlabeled density defined in the common state space, the labeled multi-target density is defined in the joint state and label space, where the label contains time-series/history information of the underlying track. To measure the similarity between labeled densities (individual tracks) that is required for inter-sensor track matching and fusion, one has to account for the divergences in both state and label spaces. The challenge, however, arises from the lack of a proper metric to measure the label difference. It requires considering the entire trajectory of the track, encompassing the whole-life information from the birth of the track to the present. In this paper, we provide a solution of comparing and matching labels based on the whole-life time-series state distributions of the labels/tracks, by extending the common divergences like the Cauchy-Schwarz and Kullback-Leibler from distributions at a single time-instant to those over time-series. Representative scenarios are considered for illustration.**

***Index Terms*—Random finite set, multiple target tracking, labeled multi-Bernoulli filter, label matching, track association.**

I. Introduction

MULTI-target tracking (MTT) is a well-established research field with broad applications. The goal of the MTT is to estimate the trajectories of a, perhaps time-varying, number of targets from sensor data. The most known three approaches that have been developed and investigated in this field are 1) the joint probability data association (JPDA) [1]; 2) multiple hypotheses tracking (MHT) [2], [3]; and 3) random finite set (RFS) [4]; see also [5]. It was recently highlighted that Soviet Union Researchers Achkasov [6], [7] in 1970-1972 and Bakut and Ivanchuk [8] in 1976 derived prototypical versions of the JPDA and the probability hypothesis density (PHD) filter, within a single coherent framework originating from statistical mechanics, referred to as stochastic flows (SF) or flow theory; see the the detailed historical review [9], [10]. However, we will stick with the use of the concept of RFS in this work considering its popularity in the present literature. A cutting-edge branch of the RFS/SF approaches is the so-called labeled RFS (LRFS), which aims to estimate the multi-object trajectories (properly linked time-series state estimates for each individual target) in a principled manner [11], [12]. Representative works in this line include the generalized labeled multi-Bernoulli (GLMB) filter [11] and two of its computationally simplified versions known as δ-GLMB filter [12] and the labeled multi-Bernoulli (LMB) [13] filter. It is obviously advantageous to estimate and maintain the trajectory information of each target. The idea of trajectory estimation or labeling has been developed in other forms such as the trajectory RFS [14], [15] and appeared in the non-RFS estimation literature too; see e.g., [16]–[19].

A fundamental issue raised with the filtering and fusion approaches is the metric that calculate the divergence between the filtered densities. For example, the Kullback-Leibler divergence (KLD) is used to derive multi-sensor density fusion rules [20]–[24] and implement filter approximation [25]–[27], while the Cauchy-Schwarz divergence (CSD) is widely adopted in the sensor control scenes [28]–[31]. While many metrics have been used to calculate the variability between multi-target densities in the unlabeled case, less research has been particularly focused on the labeled case where the labeled density is the joint density of the label and state [11]–[13].

This is nontrivial simply because the labels are non-comparable. A heuristics is to match labels first and then compute the divergence between, and fuse, the underlying state densities with matched labels [32]–[37], where the matched labels are considered to be equivalent. However, only the state density at a particular time, namely the marginal distribution of the involving tracks, is taken into account, ignoring the history part of the tracks. This does not meet the nature of the label that indicates the whole-track information of the potential target. Arguably, whenever work with the labels, the whole-track information should be accounted for, as was done in evaluating the performance of track or trajectory estimators [38], [39].

Nevertheless, most of the existing labeled density fusion solutions need to solve label matching firstly [32]–[37] (see also the reason due to the theoretical derivation of the average fusion for RFS filters given in [40]), discard the labels [41], or directly assume that the labels are impractically perfectly coordinated between filters [42]. All of them do not consider the complete track information but only the marginal density of the latest time. This paper is dedicated to fill this gap by proposing a method for measuring the divergence between labeled densities, which takes into account the whole trajectory information in order to carry out label matching. Although some of the discussion and findings we made hold to the other labeled densities, we will in particular pay attention to the LMB densities [13] in the example study.

The rest of the paper is organized as follows. The prelim-

This work was partially supported by National Natural Science Foundation of China (Grant No. 62071389), Natural Science Basic Research Program of Shaanxi Province (Program No. 2023JC-XJ-22) and the Fundamental Research Funds for the Central Universities.

inary of the RFS and LRFS are reviewed in Section II. The difficulty and necessity for comparing the labeled densities are addressed in III. Section IV presents our proposed solutions for calculating track difference and for label matching. Simulation is given in Section V before we conclude in Section VI.

II. Background and Preliminary

A. RFS Theory

We inherit the notations from [13] as follows. The lowercase letters (e.g., x) represent single-target states, while the uppercase letters (e.g., X) represent multi-target states. In addition, we use bold to represent labeled densities or states (e.g., $\mathbf{x}$, $\mathbf{X}$, $\boldsymbol{\pi}$) to distinguish unlabeled ones. Furthermore, the spaces are represented as blackboard bold letters.

Let $\mathcal{X} \subseteq \mathbb{R}^d$ represent the d-dimensional state space, and $\mathbb{X}$ denote all finite subsets of $\mathcal{X} \subseteq \mathbb{R}^d$. Due to the inherent characteristic of RFSs, multi-target states can naturally be represented by an RFS $X \in \mathbb{X}$. The stochastic characteristics of the multi-target RFS X is encapsulated by its multi-target probability distribution, represented as $f(X)$. For a specific realization of X with a given cardinality $|X| = n$, denoted by $X_n = \{x_1, \ldots, x_n\}$ [4, Eq. 2.36] where $x_i \in \mathcal{X}$ represents the state of the i-th target,

$$f(X_n) = n!\rho(n)f(x_1, \ldots, x_n), \tag{1}$$

where $f(x_1, \ldots, x_n)$ for $n = 1, 2, \ldots,$ are the spatial densities, which are symmetric in the arguments, and the cardinality distribution $\rho(n) \triangleq \Pr\{|X| = n\} = \int_{|X|=n} f(X)\delta X$ is given by $\rho(n) = \frac{1}{n!}\int_{\mathcal{X}^n} f(\{x_1, \ldots, x_n\})dx_1, \ldots, x_n$.

The set integral in $\mathbb{X}$ is defined as [4, Eq. Ch. 3.3]

$$\int_{\mathbb{X}} f(X)\delta X = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{X}^n} f(\{x_1, \ldots, x_n\})dx_1, \ldots, x_n \tag{2}$$

$$= \sum_{n=0}^{\infty} \rho(n) \tag{3}$$

$$= 1, \tag{4}$$

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

Denote by $\mathcal{L}$ the label space and by $\mathbb{L}$ all of the finite subsets of $\mathcal{L}$. In the LRFS, targets are distinguished by a unique label and so the labeled density hereafter denoted as $\pi(\mathbf{X})$ is the joint density of the continuous state and the discrete label, where $\mathbf{X}$ is a finite subset set of $\mathbb{X} \times \mathbb{L}$. As shown in Fig. 1, the densities obtained over time, as well as estimate extracted from them, are linked by labels yielding the trajectories of individual targets.

The LRFS integral is conceptually defined as [11], [12]

$$\int_{\mathbb{X}\times\mathbb{L}} \pi(\mathbf{X})\delta\mathbf{X} = \sum_{n=0}^{\infty} \int_{\mathcal{X}^n \times \mathcal{L}^n} \pi(\mathbf{X}_n)\delta\mathbf{X}_n \tag{5}$$

$$= \sum_{n=0}^{\infty} \sum_{L_n \subseteq \mathcal{L}^n} \int_{\mathbb{X}} \pi(\mathbf{X}_n)\delta X_n \tag{6}$$

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

where $\mathbf{X}_n$ is a realization of an LRFS with cardinality n, X_n is the multi-target state and L_n is the label set.

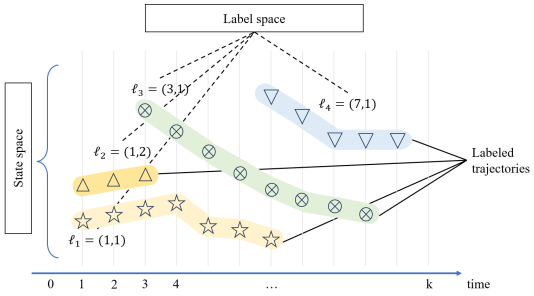


Fig. 1. The track indicated by a time-series of densities with the same label.

B. GLMB RFS

An LRFS is an RFS of elements with discrete and distinct labels. An GLMB is an LRFS with a distribution [11]

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{c\in\mathbb{C}} w^{(c)}(\mathcal{L}(\mathbf{X})) \prod_{(\mathbf{x},\ell)\in\mathbf{X}} p^{(c)}(x, \ell) \tag{7}$$

in the state space $\mathbb{X}$ and label space $\mathbb{X}$, where $\mathbb{C}$ is a discrete index set, $w^{(c)}(L)$ and $p^{(c)}$ satisfy

$$\sum_{L\subseteq\mathbb{L}} \sum_{c\in\mathbb{C}} w^{(c)}(L) = 1, \int p^{(c)}(x, \ell)dx = 1.$$

A δ-GLMB RFS is a special case of GLMB RFS with $\mathbb{C} = \mathcal{F}(\mathbb{L}) \times \Xi, w^{(c)}(L) = w^{(I,\xi)}(L) = w^{(I,\xi)}\delta_I(L), p^{(c)} = p^{(I,\xi)} = p^{(\xi)}$, where (I, ξ) denotes the track set I with an association maps history ξ [12] and Ξ is a discrete space. The distirbution of a δ-GLMB RFS $\mathbf{X}$ is given by

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} w^{(I,\xi)}\delta_I(\mathcal{L}(\mathbf{X})) \prod_{(x,\ell)\in\mathbf{X}} p^{(\xi)}(x, \ell). \tag{8}$$

C. LMB RFS

A LMB RFS can be completely represented by a parameter set $\{(r^{(\zeta)}, p^{(\zeta)}) : \zeta \in \Psi\}$ and the density of LMB RFS is [13]

$$\boldsymbol{\pi}(\{(x_1, \ell_1), \ldots, (x_n, \ell_n)\}) = \delta_n(|\{\ell_1, \ldots, \ell_n\}|) \prod_{\zeta\in\Psi} (1 - r^{(\zeta)}) \times \prod_{j=1}^{n} \frac{1_{\alpha(\Psi)}(\ell_j) r^{(\alpha^{-1}(\ell_j))} p^{(\alpha^{-1}(\ell_j))}(x_j)}{1 - r^{(\alpha^{-1}(l_j))}}, \tag{9}$$

where $\alpha : \Psi \to \mathbb{L}$ is a 1-1 mapping of indexes to labels, $1_Y(x) = 1$ if $x \in Y$ and zero otherwise.

III. Divergence of Labeled Densities

It is nontrivial to define the distance/divergence in the joint state and label space. The definition of the LRFS integral (5) is based on the random state and label in the space $\mathbb{X} \times \mathbb{L}$. In the realization, however, the labels are deterministic and finite. In the first place, the labels, as non-numeric variables, do not have a definition of distance from each other. More specifically, the labels for individual targets are often defined

as the ordered pairs of integers (k, l), where k represents the time of birth, and l represents the sequence number of the target. Obviously, the variable l does not possess an inherent ordering concept within the real number range. This means that, there is no inherent relationship of magnitude between $l = 1$ and $l = 2$. Consequently, the label $\ell = (k, l)$ also does not allow for any distance calculation in the Euclidean space such as $\ell_m - \ell_n$. An extreme case raised in [40] is that two labeled densities of the same state distribution but assigned with different labels, for which it is nontrivial to define an intuitively convincing distance/divergence.

A practically useful yet heuristic way to address the above problem is given by disregarding the labels in comparing two labeled densities. That is, one first matches their labels, assuming that the matched label indicating the same potential target are the same, and then calculate the divergence between the matched components densities in the unlabeled state space (not in the joint state and label space). [1] To this end, inspired by [43], we decompose the labeled density $\boldsymbol{\pi}(\mathbf{X}_n)$ as the joint of the state spatial density $f(X_n)$ conditional on the label set L_n and the label probability mass function $p(L_n)$, i.e.,

$$\boldsymbol{\pi}(\mathbf{X}_n) := p(L_n) f(X_n) \tag{10}$$

$$p(L_n) = \int \boldsymbol{\pi}(\mathbf{X}_n) \delta X_n \tag{11}$$

where $L_n = \{l_1, l_2, \ldots, l_n\}, X_n = \{x_1, x_2, ..., x_n\}, \mathbf{X}_n = \{(x_1, l_1), (x_2, l_2), ..., (x_n, l_n)\}$.

Following this line of think, the divergence between different label-matched densities can be calculated by any suitable metric defined in the state space such as the KLD, CSD and the integral squared distance (ISD). That is, the divergence of two labeled densities $\boldsymbol{\pi}_1(\mathbf{X}_n), \boldsymbol{\pi}_2(\mathbf{X}_n)$ with respective label set L_n and L'_n that are matched can be conceptually defined over the spatial densities as given in Table I.

Remark 1. *The above definitions of the divergence between labeled densities take into account differences in state distribution and ignores differences between labels; matched labels are treated as the same. Specifically, even if the spatial distributions of the densities of the two labels are exactly the same but the labels are different, the divergence should not be zero as the above definitions will indicate.*

Remark 2. *The KLD requires two densities having the same support while CSD and ISD do not. As such the KLD only applies to the case where each non-zero component of one density must be matched with the non-zero component of the other density with the same support. To this end, the number of labels of two labeled densities must be the same and the matching be bijection. This stringent requirement, however, is not required in the case of using the CSD or ISD.*

[1] In fact, these two steps are usually implemented in the reverse order. That is, one first calculate the distance between any two labeled components from two fusing densities. The matching is an optimization issue that is usually solved so that the overall distance between all matched components is the minimal in all possible matching solutions; see Section IV.

IV. Divergence of Tracks and Label Matching

So far, only the state densities at the present time is considered in comparing the labeled densities and in label matching. In this section, we will further address the divergence between labeled tracks that is a series of densities with the same label and the relevant label matching by taking into account the whole-life trajectory information of the tracks. Motivation for doing so will be illustrated firstly. All seek a more reliable and convincing label matching solution.

A. Why Tracks?

Due to the nature of the labels that signify that the current and history estimates with the same label all belong to the same target, incorporating the entire trajectory information for label matching allows for a more comprehensive utilization of the label information. The necessity and advantage to match labels/tracks according to their whole-life information rather than only the present is obvious. In the scenario given in Fig. 2, the tracks exhibit diverse patterns such as:

- Tracks a and a' have similar state estimates at the present moment and their entire trajectories show a consistent trend. However, Track a' starts later than Track a.
- Tracks b and b' have the same track length and display similar state estimations at the current moment, but their trajectories diverge considerably in the first half.
- Tracks c and c' possess similar positions at the current moment, but their overall trends differ significantly.

In all above cases, the tracks are obviously very different from each other although their latest state densities are similar. It will easily lead to wrong matching results if only the latest density is accounted for, which will lead to both problematic multi-sensor track fusion [32], [33], [35] and problematic single-sensor components merging.

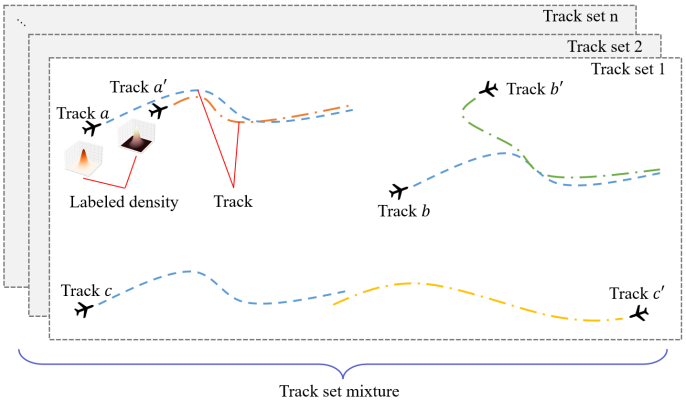


Fig. 2. Very different target trajectories with similar current state estimates.

B. Divergence between Matched Tracks

In the following, we provide a technically sound metric to compare pairwise matched tracks each of which consist of a time-series of densities. To this end, we need to extend the difference between matched labeled-densities to the matched

TABLE I
METRICS OF UNLABELED/SINGLE-LABELED DENSITIES

Metric	Unlabeled expression	Labeled expression
KLD	$D_{\text{KL}}(f\|\|g) = \int_{\mathcal{X}} f(x) \log \frac{f(x)}{g(x)} dx$	$D_{\text{KL}}(\boldsymbol{\pi}_1\|\|\boldsymbol{\pi}_2) = \int_{\mathcal{X}^n} p(L_n) f_1(X_n) \log \frac{p(L_n) f_1(X_n)}{p(L'_n) f_2(X_n)} \delta X_n$
CSD	$D_{\text{CS}}(f\|\|g) = -\log \frac{\int_{\mathcal{X}} f(x) g(x) dx}{\sqrt{\int_{\mathcal{X}} f(x)^2 dx \int_{\mathcal{X}} g(x)^2 dx}}$	$D_{\text{CS}}(\boldsymbol{\pi}_1\|\|\boldsymbol{\pi}_2) = -\log \frac{\int_{\mathcal{X}^n} p(L_n) p(L'n) f_1(X_n) f_2(X_n) \delta X_n}{\sqrt{\int_{\mathcal{X}^n} p(L_n) f_1(X_n)^2 \delta X_n \int_{\mathcal{X}^n} p(L'_n) f_2(X_n)^2 \delta X_n}}$
ISD	$D_{\text{IS}}(f\|\|g) = \int_{\mathcal{X}} (f(x) - g(x))^2 dx$	$D_{\text{IS}}(\boldsymbol{\pi}_1\|\|\boldsymbol{\pi}_2) = \int_{\mathcal{X}^n} (p(L_n) f_1(X_n) - p(L'_n) f_2(X_n))^2 \delta X_n$

tracks, for which the target state spatial distribution information in the entire time series are taken into account.

Denote by $\mathcal{T}_\ell = [\boldsymbol{\pi}_k^\ell]_{k=\alpha_\ell}^{\beta_\ell}$ the track (here represents a density sequence with the same label, different from the standard track definition) corresponding to label ℓ, where $\boldsymbol{\pi}_k^\ell$ is the labeled density at time step k, α_ℓ and β_ℓ denote the start and end times of the track, respectively. The life of the track $\mathcal{T}_\ell$, that is, the time when the track exists is expressed as $\mathcal{I}_l = \{k \in \mathbb{N}; \alpha_\ell \leq k \leq \beta_\ell\}$. For two tracks $\mathcal{T}_{\ell_1}$ and $\mathcal{T}_{\ell_2}$, the track divergence calculation is as follows

$$\begin{aligned} d_{\mathcal{T}}(\mathcal{T}_{\ell_1}, \mathcal{T}_{\ell_2}) = & \sum_{k \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2}} \min(D_{\text{density}}(\boldsymbol{\pi}_k^{\ell_1}, \boldsymbol{\pi}_k^{\ell_2}), \lambda_C) \\ & + \lambda_C \times |\mathcal{I}_{\ell_1} \cup \mathcal{I}_{\ell_2} - \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2}|, \end{aligned} \tag{12}$$

where D_{density} denotes the distance/divergence between the two labeled densities such as the KLD, CSD and ISD defined in Tabel I, λ_C is the cut-off error of unmatched track points, which also serves as a gate controlling the matching and is usually specified by the practitioner. By this gating, we will not consider the possibility of matching two densities that have a divergence larger than λ_C. However, because of the gating, the divergence may not be a distance simply because the triangle inequality does not hold; further analysis and proof will be given in our future work.

As shown, the pairwise matched track divergence (12) consists of two components:

- The first term $\sum_{k \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2}} \min(D_{\text{density}}(\boldsymbol{\pi}_k^{\ell_1}, \boldsymbol{\pi}_k^{\ell_2}), \lambda_C)$ captures the divergence between the overlapping segments of the two tracks.
- The latter term $\lambda_C \times |\mathcal{I}_{\ell_1} \cup \mathcal{I}_{\ell_2} - \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2}|$ represents the non-overlapping segments of the two tracks, where the spatial distribution of one track is estimated while the other track lacks an estimate at that time.

C. Divergence between Matched Track Sets

Denote by $\mathcal{S} = \{\mathcal{T}_\ell\}_{\ell \in L}$ a collection of tracks corresponding to label set L, i.e., a set of labeled densities over time series, where $|\mathcal{S}|$ is the number of the tracks/labels. For two sets consisting of labeled densities $\mathcal{S}^{(1)} = \{\mathcal{T}_{\ell_m^{(1)}}\}_{m=1}^M$ and $\mathcal{S}^{(2)} = \{\mathcal{T}_{\ell_n^{(2)}}\}_{n=1}^N$, where $M = |\mathcal{S}^{(1)}|$ and $N = |\mathcal{S}^{(2)}|$ are the respective number of labels/tracks, we first define a divergence matrix $D = (d_{mn})_{M \times N}$ where $d_{mn} = d_{\mathcal{T}}(\mathcal{T}_{\ell_m^{(1)}}, \mathcal{T}_{\ell_n^{(2)}})$ represents the element of the divergence matrix D corresponding to the m-th row and n-th column which measures the similarity between the densities with label l_m of set 1 and the densities with label l_n of set 2 calculated as given in (12). Based on this, we define two alternative divergences between these two track sets that are matched as supposed.

1) Label matching based divergence: We first define a matching matrix $\phi = (a_{mn})_{M \times N}$ as follows,

$$a_{mn} = \begin{cases} 1, & l_m^{(1)} \text{ matching } l_n^{(2)} \\ 0, & otherwise \end{cases} \tag{13}$$

Without loss of generality, we assume that $M \leq N$, then, the problem of label matching is transformed into the following optimization problem:

$$\begin{aligned} \hat{\phi} &= \arg\min_{\phi} \sum_m \sum_n a_{mn} d_{mn} \\ s.t. &\begin{cases} a_{mn} \in \{0, 1\} \\ \sum_n a_{mn} = 1 \text{ for } m = 1, \ldots, M. \end{cases} \end{aligned} \tag{14}$$

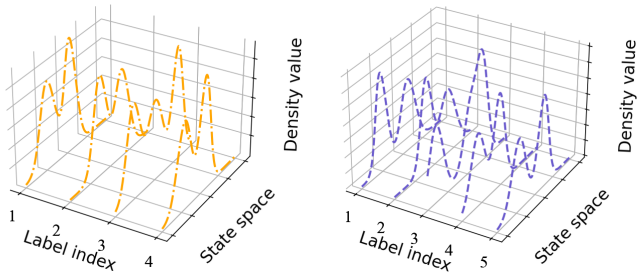
The above model deals with the label matching problem as a standard assignment problem, which can be solved by standard optimizers such as Hungarian algorithm [44]. The divergence between track density sets based on the label matching is defined as

$$d_{\mathcal{S}}(\mathcal{S}^{(1)}, \mathcal{S}^{(2)}) = \sum_n \Big(\sum_m a_{mn} d_{mn} + \lambda_C T_n^{(2)} (1 - \sum_m a_{mn}) \Big), \tag{15}$$

where $T_n^{(2)} = \beta_{\ell_n^{(2)}} - \alpha_{\ell_n^{(2)}}$ is time-length of the track n.

Similar to the track divergence, (15) consists of two parts. The first term $\sum_n \sum_m a_{mn} d_{mn}$ represents the part of the label matching, i.e., the summing of divergences considered to be from the same target. The other term $\sum_n \lambda_C T_n^{(2)} (1 - \sum_m a_{mn})$ is the penalty term for unmatched labels. A common cause for unmatched tracks is that the labeled densities have different numbers of label sets or labels as shown in Fig. 3 in which not all labels can be matched.

2) Matching-insensitive divergence: Instead of defining the divergence like (15) between two track sets based on the label matching solution, one may want the divergence definition unaffected by the subjective matching. To this end, the following



(a) densities of set 1 (b) densities of set 2

Fig. 3. Illustration of two labeled densities with different numbers of labels and so different function supports for which the KLD cannot apply.

divergence is defined,

$$d_{\mathcal{S}}(\mathcal{S}^{(1)},\mathcal{S}^{(2)}) = \sum_m \sum_n w_{mn} d_{mn}, \tag{16}$$

$$w_{mn} \propto \mathcal{N}\left(\frac{d_{mn}}{\min(T_m^{(1)},T_n^{(2)})}; 0, d_{\mathrm{G}}^2\right), \tag{17}$$

$$\text{s.t.} \sum_n w_{mn} = 1, \forall m = 1,\ldots,M, \tag{18}$$

where $\mathcal{N}(x;0,d^2)$ is the Gaussian likelihood function with zero-mean and covariance d^2, d_{G} is a gate specified by the practitioner in order to assign the weight/probability for each matching; this is analogous to, much simpler than, the probabilistic label association solution given by [34]. One may choose according to the previously mentioned matching gate λ_C, such as $d_{\mathrm{G}} = 0.2\lambda_C$. We argue that there is still space for designing the weights w_{mn} in a more principled way. This is an issue similar to calculating the association probability in the JPDA approach [1], [6], [7].

The above definition (16) takes into account all possible matches with weights, making it no more necessary to find the *best matching* as given in (14). It is arguably matching-free, objective in comparison with the subjective matching-based definition (15). The difference between the two divergence definitions can be illustrated in the following scenario.

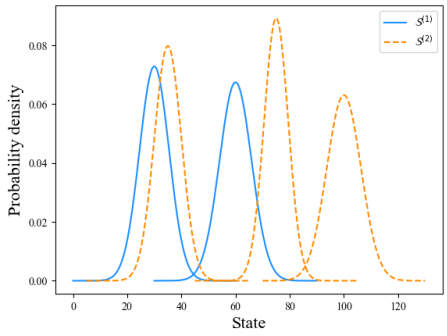


Fig. 4. Example 1. The blue solid line and the yellow dashed line are Gaussian distributions in the track density set $\mathcal{S}^{(1)}$ and $\mathcal{S}^{(2)}$, respectively.

Example 1. *Consider two track density sets $\mathcal{S}^{(1)}$ and $\mathcal{S}^{(2)}$. There are two tracks in set $\mathcal{S}^{(1)}$ and three tracks in set $\mathcal{S}^{(2)}$. For computational simplicity, it is assumed that the tracks have only one time step and have been aligned, and each density is a Gaussian distribution. A concrete implementation is given as $\mathcal{S}^{(1)} = \{[\mathcal{N}_{1,\ell_1^{(1)}}(30;30)],[\mathcal{N}_{1,\ell_1^{(1)}}(60;35)]\}$ and $\mathcal{S}^{(2)} = \{[\mathcal{N}_{1,\ell_1^{(2)}}(35;25)],[\mathcal{N}_{1,\ell_1^{(2)}}(70;20)],[\mathcal{N}_{1,\ell_1^{(2)}}(100;40)]\}$. The distributions of the two sets are shown in Fig. 4.*

Calculate the divergence matrix D using (12) with $\lambda_C = 50$ based on the basic KLD metric and $d_G = 0.2\lambda_C$ in (16), yielding

$$D = \begin{bmatrix} 1.509 & 50 & 50 \\ 13.532 & 6.720 & 21.004 \end{bmatrix}. \tag{19}$$

The optimization problem in (14) is then solved using the Hungarian algorithm

$$\hat{\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{20}$$

Denote by $W = (w_{mn})$ the weight matrix corresponding to (17), then

$$W = \begin{bmatrix} 9.9999\times10^{-1} & 3.7693\times10^{-6} & 3.7693\times10^{-6} \\ 3.0596\times10^{-1} & 6.0984\times10^{-1} & 8.4193\times10^{-2} \end{bmatrix}$$

Then, the label matching based divergence and matching-insensitive divergence are 58.229 and 18.047, respectively. It is clear that there is a significant difference between the two divergences due to the fact that the former performs a globally optimal assignment, adding a direct penalty term for unmatched tracks, while the latter takes into account the likelihood of matching, and under reasonable parameters, tracks with large divergences are given less weights. Therefore the total divergence is relatively smaller.

D. Label Matching for LMB

As addressed so far, the track/label matching can be obtained by solving the optimisation problem formulated in (14) for which the d_{mn} as defined in (12) is the key. More details for calculating (12) are given below for the LMB.

The LMB filter posterior sequence of sensor i upto time t is recorded as $\Pi^{(i)} = [\boldsymbol{\pi}_1^{\ell_l^{(i)}}, \boldsymbol{\pi}_2^{\ell_l^{(i)}}, \ldots, \boldsymbol{\pi}_t^{\ell_l^{(i)}}]_{l=1}^{M_t^{(i)}}$, where $M_t^{(i)}$ is the number of labels/tracks that still survive at time t and $\boldsymbol{\pi}_k^{\ell_l^{(i)}} = \{r^l, p^l\}$ is a Bernoulli component. The estimation is therefore represented by $M_t^{(i)}$ tracks, where the lth track is $\mathcal{T}_{t,l}^{(i)} = [(r_k^l, p_k^l)]_{k=\alpha_l}^{\beta_l}$ with existing period $\mathcal{I}_{t,l}^{(i)} = \{k \in \mathbb{N}; \alpha_l^{(i)} \leq k \leq \beta_l^{(i)}\}$ as defined in section IV-B. Then the divergence between the m-th labeled density of sensor i and the n-th labeled density of sensor j, $d_{mn}^{(ij)}$, can be calculated by (12). Here the track existence probability r^l correspond to the label probability $p(l)$ used therein.

E. Matching between Track Set Mixtures

We further denote a track set mixture by $\mathcal{M} = \{w^I, \{\mathcal{T}_\ell\}_{\ell\in I}\}_{I\in\mathcal{F}(\mathbb{L})}$, where $I \in \mathcal{F}(\mathbb{L})$ is a track set, i.e., a set of labels and w^I is the corresponding weight for track set I, satisfying $\sum w^I = 1$. It amounts to a family of track sets obtained under different hypotheses and associated with different weights (as the case with the δ-GLMB), which can

be further expressed as $\mathcal{M} = \{w^I, \{[\pi_k^\ell]_{k=\alpha_\ell}^{\beta_\ell}\}_{\ell \in I}\}_{I \in \mathcal{F}(\mathbb{L})}$ and illustrated in Fig. 2.

To address the label matching for track set mixtures, one needs to first match track sets, whose divergences from each other have been defined in Sec. IV-C and then match the inner labels between matched track sets. For the (δ-)GLMB, we leave the implementation details to our future work.

Remark 3. *Unlike researches on distance metrics proposed for trajectories sets [14], [15], [38], the elementary distance between densities is still defined in the state space $\mathcal{X} \subseteq \mathbb{R}^d$. The track/trajectory in the study is defined as a series of state densities, more precisely the Bayesian posteriors, instead of a sequence of points [16], [17] or a continuous-time function like [18], [19].*

V. Simulation

In this section, we study the effectiveness of the proposed label matching solution (14) for the GM-LMB filters [13]. Considering the physical meaning and redundancy of labels [45], only the labels that are considered as targets at the current time step are matched.

We considered the region of interest (ROI) given by $[-1000m, 1000m] \times [-1000m, 1000m]$. The target state is denoted as $x_k = [\mathrm{x}_k, \dot{\mathrm{x}}_k, \mathrm{y}_k, \dot{\mathrm{y}}_k, \omega_k]^{\mathrm{T}}$ where $[\mathrm{x}_k, \mathrm{y}_k]^{\mathrm{T}}$ denotes the spatial position, $[\dot{\mathrm{x}}_k, \dot{\mathrm{y}}_k]^{\mathrm{T}}$ represents the velocity and ω_k is the turn rate. The survival single-target movement used by the filters follows a coordinated turn (CT) model with a sampling period of $\Delta = 1s$ and transition density $f_{k|k-1}(x_k|x_{k-1}) = \mathcal{N}(x_k; Fx_{k-1}, Q)$, where

$$F(\omega) = \begin{bmatrix} 1 & \frac{\sin\omega}{\omega} & 0 & -\frac{1-\cos\omega}{\omega} & 0 \\ 0 & \cos\omega & 0 & -\sin\omega & 0 \\ 0 & \frac{1-\cos\omega}{\omega} & 1 & \frac{\sin\omega}{\omega} & 0 \\ 0 & \sin\omega & 0 & \cos\omega & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

and $Q = \mathrm{diag}([I_2 \otimes G, \sigma_u^2])$ with

$$G = \begin{bmatrix} \frac{\sigma_w^2}{4} & \frac{\sigma_w^2}{2} \\ \frac{\sigma_w^2}{2} & \sigma_w^2 \end{bmatrix}, \tag{22}$$

where $\otimes$ is the Kronecher product, $\sigma_w = 5m/s$ and $\sigma_u = (\pi/180)rad/s$. Each of the two sensors $i = 1, 2$ has a target detection probability 0.98 and a linear measurement model as follows

$$z_{i,k} = Hx_k + v_{i,k}, \tag{23}$$

where

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{24}$$

and $v_{i,k} = [v_{i,k}^{(1)}, v_{i,k}^{(2)}]^{\mathrm{T}}$ are independent identical distributed zero-mean Gaussian with standard deviation $5m$.

The simulation scenario is composed of 6 targets of deterministic trajectories which were generated by the mentioned CT model without using process noises and start at different times. The total simulation time is $100s$. The target motion trajectories are shown in Fig. 5.

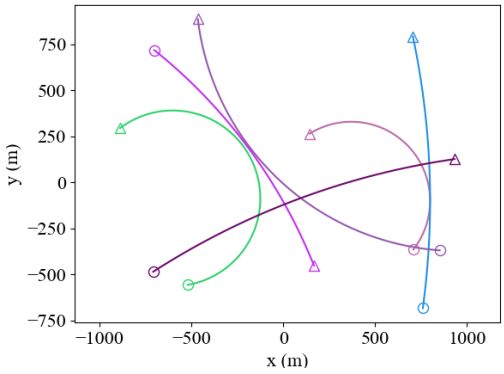


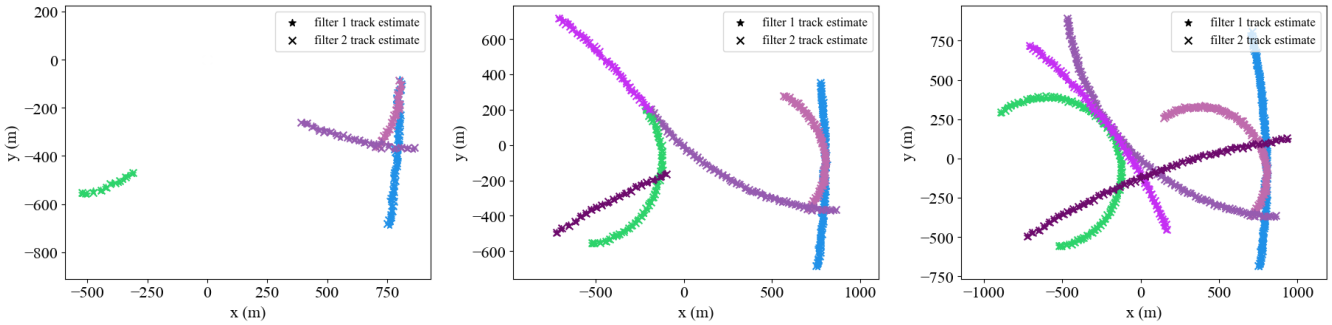
Fig. 5. Ground truth (different colors indicate different tracks, and circles/triangles start/end points).

The scene is set up with two GM-LMB filters. The local filter uses up to 100 BCs and each BC contains up to 10 Gaussion components (GCs). To ensure that the filter operates effectively, BCs with a probability of target existence below 10^{-3} are pruned, so are the GCs with weights below 10^{-5}. The divergences between densities are calculated using the CSD and the coefficient of the cut-off error λ_C is taken as 50.

The label matching results based on (14) for the filtering step $k = 40, 70$ and 100 are shown, respectively. Different markers are used to indicate the track estimates of different filters, namely the peak extraction of the tracks and the same color indicates the matched track between two filters.

As shown in Fig. 6(a), two targets are close at the time, and the other two are separated. The intersecting tracks are quite close in position, velocity magnitude, and velocity direction, and our method yields correct matching. In Fig. 6(b), the occurrence of target crossover on the left leads to a similar estimation of the target position and velocity magnitude by sensor 1 and sensor 2, with two progressively approaching and the third coming from the opposite direction. The proposed method still performs a proper correlation. Fig. 6(c) shows the final label matching result, which is in line with the matching of the tracks that exist at $k = 70$.

Remark 4. *The permutation invariance [15] involved in multi-target tracking may cause certain complications. When the targets are close to each other, the measurement-to-track association is indeed challenging. In this work, however, we do not confront this issue since we only deal with the filter-established tracks and not how they are formed in the filtering process. That being said, our proposed matching and fusion solution based on the whole-life information has advantages in dealing with this problem. As demonstrated in Fig. 7, in the earlier attempts that merely considered the spatial distribution of the targets at a single time-instant, the two intersecting tracks can be easily incorrectly matched. This problem will be highly combated in our proposed method that takes into account the whole-life information. This is just the advantage of estimating the track in comparison with estimating any*



(a) label matching from $k = 1$ to $k = 40$ (b) label matching $k = 1$ to $k = 70$ (c) label matching $k = 1$ to $k = 100$

Fig. 6. The label matching result yielded at different tracking stages of multiple targets, where '$\star$' denotes the estimates of LMB filter 1 and 'x' denotes those of LMB filter 2. Labels are marked by color and the same color given by different filters indicate the matching.

point-state/density.

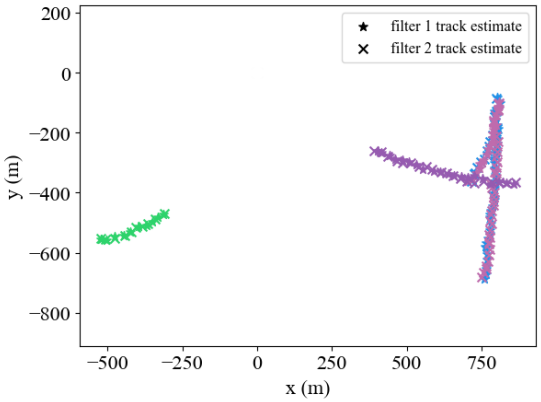


Fig. 7. Label matching from $k = 1$ to $k = 40$ based on current ($k = 40$) state estimation, where '$\star$' denotes the estimates of LMB filter 1 and 'x' denotes those of LMB filter 2. Labels are marked by color and the same color given by different filters indicate the matching.

VI. Conclusion

This paper rethinks the challenge encountered in the computation of the variation or similarity between labeled tracks, namely the time-series densities yielded by the labeled RFS filters. We point out that it is nontrivial to define the divergence between labeled tracks, nor is it easy to match the labels/tracks. Yet, label matching plays a key role in track fusion in the multi-sensor LRFS filtering approaches and has been superficially addressed by merely taking into account of the marginal density of the tracks. Differently, we make an attempt to compare and match tracks by taking into account their whole-life information, resulting in two versions of track divergence, one based on an unique, probably the best label matching solution and the other relies on no single matching choice but consider all matching possibilities. Examples and simulations are considered for illustration and for demonstrating the effectiveness/reasonableness of our approach.

References

[1] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE Journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, 1983.

[2] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[3] R. L. Streit and T. E. Luginbuhl, "Maximum likelihood method for probabilistic multihypothesis tracking," in *Signal and Data Processing of Small Targets 1994*, O. E. Drummond, Ed., vol. 2235, International Society for Optics and Photonics. SPIE, 1994, pp. 394 – 405.

[4] R. P. Mahler, *Advances in statistical multisource-multitarget information fusion*. Artech House, 2014.

[5] B.-n. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley encyclopedia of electrical and electronics engineering*, no. 2015, 2015.

[6] Y. S. Achkasov, "Finding trajectories by a posteriori analysis of flows," *Engineering Cybernetics (Tekhnicheskaya Kibernetika)*, vol. 9, no. 5, pp. 919–926, 1971.

[7] ——, "Measurement of trajectory parameters from tracks in a noise background," *Engineering Cybernetics (Tekhnicheskaya Kibernetika)*, vol. 10, no. 1, pp. 151–157, 1972.

[8] P. A. Bakut and N. A. Ivanchuk, "Calculation of the a posteriori characteristics of the flow of resolved objects," *Engineering Cybernetics (Tekhnicheskaya Kibernetika)*, vol. 14, no. 6, pp. 148–156, 1976.

[9] S. M. O'Rourke, "Closing the ivory valikaition gates: A historical analysis of achkasov's kaita association filters," in *2023 26th International Conference on Information Fusion (FUSION)*, 2023, pp. 1–8.

[10] D. E. Clark, A. Narykov, and R. L. Streit, "Stochastic flows–a primer on early multi-object filtering work with point processes," in *2023 26th International Conference on Information Fusion (FUSION)*. IEEE, 2023, pp. 1–7.

[11] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[12] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the bayes multi-target tracking filter," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6554–6567, 2014.

[13] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labeled multi-bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3246–3260, 2014.

[14] Á. F. García-Fernández, A. S. Rahmathullah, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3917–3928, 2020.

[15] Á. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 1685–1707, 2019.

[16] A. Andriyenko, K. Schindler, and S. Roth, "Discrete-continuous optimization for multi-target tracking," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, June 2012.

[17] E. H. Aoki, P. K. Mandal, L. Svensson, Y. Boers, and A. Bagchi, "Labeling uncertainty in multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 3, p. 1006–1020, 2016.

[18] T. Li, H. Chen, S. Sun, and J. M. Corchado, "Joint smoothing and tracking based on continuous-time target trajectory function fitting," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1476–1483, 2019.

[19] T. Li, Y. Song, and H. Fan, "From target tracking to targeting track: A data-driven yet analytical approach to joint target detection and tracking," *Signal Processing*, vol. 205, p. 108883, 2023.

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

[21] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus cphd filter for distributed multitarget tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508–520, 2013.

[22] M. Üney, J. Houssineau, 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, 2019.

[23] K. Da, T. Li, Y. Zhu, H. Fan, and Q. Fu, "Recent advances in multisensor multitarget tracking using random finite set," *Frontiers of Information Technology & Electronic Engineering*, vol. 22, no. 1, pp. 5–24, 2021.

[24] T. Li, Y. Song, E. Song, and H. Fan, "Arithmetic average density fusion-part I: Some statistic and information-theoretic results," *Information Fusion*, vol. 104, p. 102199, 2024.

[25] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 258–273, 2015.

[26] A. F. García-Fernández and B.-N. Vo, "Derivation of the phd and cphd filters based on direct kullback–leibler divergence minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 21, pp. 5812–5820, 2015.

[27] B. Wang, S. Li, W. Yi, and G. Battistelli, "Performance analysis for parallel grouping-based labeled multi-bernoulli filter," *Signal Processing*, vol. 202, p. 108779, 2023.

[28] M. Jiang, W. Yi, and L. Kong, "Multi-sensor control for multi-target tracking using cauchy-schwarz divergence," in *2016 19th International Conference on Information Fusion (FUSION)*, 2016, pp. 2059–2066.

[29] A. K. Gostar, R. Hoseinnezhad, T. Rathnayake, X. Wang, and A. Bab-Hadiashar, "Constrained sensor control for labeled multi-Bernoulli filter using Cauchy-Schwarz divergence," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1313–1317, 2017.

[30] G. Li, G. Li, and Y. He, "Resolvable group target tracking via multi-Bernoulli filter and its application to sensor control scenario," *IEEE Transactions on Signal Processing*, vol. 70, pp. 6286–6299, 2022.

[31] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam, "Void probabilities and cauchy–schwarz divergence for generalized labeled multi-bernoulli models," *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5047–5061, 2017.

[32] S. Li, G. Battistelli, L. Chisci, W. Yi, B. Wang, and L. Kong, "Computationally efficient multi-agent multi-object tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 260–275, 2019.

[33] L. Gao, G. Battistelli, and L. Chisci, "Fusion of labeled RFS densities with minimum information loss," *IEEE Trnas. Signal Process.*, vol. 68, pp. 5855–5868, Oct. 2020.

[34] T. Kropfreiter and F. Hlawatsch, "A probabilistic label association algorithm for distributed labeled multi-bernoulli filtering," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*. IEEE, 2020, pp. 1–8.

[35] K. Shen, P. Dong, Z. Jing, and H. Leung, "Consensus-based labeled multi-bernoulli filter for multitarget tracking in distributed sensor network," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 722–12 733, 2021.

[36] G. Li, G. Li, and Y. He, "Distributed multiple resolvable group targets tracking based on hypergraph matching," *IEEE Sensors Journal*, vol. 23, no. 9, pp. 9669–9676, 2023.

[37] L. Gao, G. Battistelli, and L. Chisci, "Resilient labeled multi-bernoulli fusion with peer-to-peer sensor network," *Information Fusion*, vol. 100, p. 101965, 2023.

[38] M. Beard, B. T. Vo, and B.-N. Vo, "A solution for large-scale multi-object tracking," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2754–2769, 2020.

[39] Z. Su, H. Ji, C. Tian, and Y. Zhang, "Performance evaluation for multi-target tracking with temporal dimension specifics," *Chinese Journal of Aeronautics*, vol. 37, no. 2, pp. 446–458, 2024.

[40] T. Li, "Arithmetic average density fusion-part II: Unified derivation for unlabeled and labeled RFS fusion," *IEEE Transactions on Aerospace and Electronic Systems*, 2024.

[41] M. Jiang, W. Yi, R. Hoseinnezhad, and L. Kong, "Distributed multi-sensor fusion using generalized multi-bernoulli densities," in *2016 19th International Conference on Information Fusion (FUSION)*, 2016, pp. 1332–1339.

[42] A. K. Gostar, T. Rathnayake, R. Tennakoon, A. Bab-Hadiashar, G. Battistelli, L. Chisci, and R. Hoseinnezhad, "Cooperative sensor fusion in centralized sensor networks using Cauchy-Schwarz divergence," *Signal Processing*, vol. 167, p. 107278, 2020.

[43] F. Papi, B.-N. Vo, B.-T. Vo, C. Fantacci, and M. Beard, "Generalized labeled multi-Bernoulli approximation of multi-object densities," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5487–5497, 2015.

[44] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[45] E. Brekke and M. Chitre, "Relationship between finite set statistics and the multiple hypothesis tracker," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1902–1917, 2018.