

A Bayesian look at the optimal track labelling problem

Edson H. Aoki¹, Yvo Boers², Lennart Svensson³, Pranab Mandal¹, and Arunabha Bagchi¹

¹University of Twente, ²Thales Nederland B.V., ³Chalmers University of Technology

Keywords: Target tracking, Finite Set Statistics, track labelling, particle filter.

Abstract

In multi-target tracking (MTT), the problem of assigning labels to tracks (track labelling) is vastly covered in literature, but its exact mathematical formulation, in terms of Bayesian statistics, has not been yet looked at in detail. Doing so, however, may help us to understand how Bayes-optimal track labelling should be performed or numerically approximated. Moreover, it can help us to better understand and tackle some practical difficulties associated with the MTT problem, in particular the so-called “mixed labelling” phenomenon that has been observed in MTT algorithms.

In this paper, we rigorously formulate the optimal track labelling problem using Finite Set Statistics (FISST), and look in detail at the mixed labeling phenomenon. As practical contributions of the paper, we derive a new track extraction formulation with some nice properties and a statistic associated with track labelling with clear physical meaning. Additionally, we show how to calculate this statistic for two well-known MTT algorithms.

1 Introduction

The track labelling problem is perhaps just as old as the multi-target tracking problem itself. In the display of a radar operator, it is often necessary not only to display the estimated position of the multiple objects (i.e. the tracks), but also attribute a unique label to each track. Ideally, this track label should consistently be associated with the same real-world object, enhancing thus the situational awareness of the operator.

In practice, the feasibility of maintaining this label-to-true target consistency depends on observability conditions. One situation where this consistency is frequently lost is after targets move in close proximity to each other. In this case, the measurements and initial information may not allow us to precisely determine which target is which after the separation. Therefore, if required to make a hard decision to assign labels to tracks, the tracker will frequently make wrong choices. This situation is illustrated in Fig. 1.

This situation where the available information allows more than one labelling possibility is referred as “mixed labelling” by Boers, Sviestins and Driessen [4]. Track extraction methods based on the mean (or, equivalently, on the Minimum Mean Square Error (MMSE) estimate) will result in track coalescence (in exact posterior sense), as observed by Blom et al. [3]. However, even if the chosen track ex-

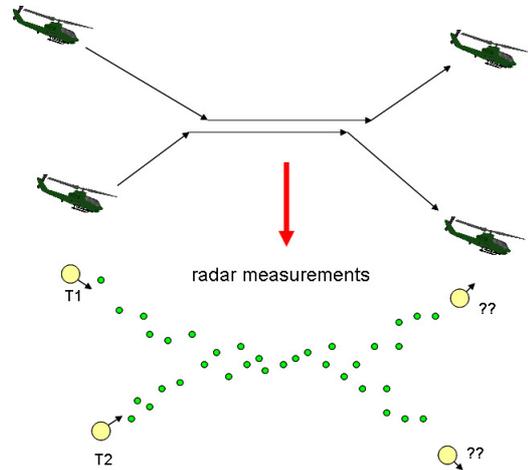


Figure 1: Situation where assignment of labels to tracks is ambiguous

traction method avoids coalescence, two questions – which form the main motivation of this work – remain to be answered:

- **Question 1:** *How does one optimally assign labels T1 and T2 to the two tracks?*
- **Question 2:** *What is the probability that the assignment is incorrect, i.e. that track swap has occurred?* This probability may be useful to the operator; for instance, our decision of shooting down or not an aircraft may be influenced if we know that the aircraft has a considerable probability (say, 40%) of corresponding to someone else!

Some statistics associated with labelling uncertainty are proposed in recent works [2, 5, 7], but the physical interpretation of these quantities is not clear from their description, making it difficult for us to assess whether they are the answer to the proposed questions.

In reality, the questions are also not perfectly clear. What do we exactly mean by probability of incorrect labelling? After all, the tracks will almost never correspond exactly to the true target locations. If the tracks are themselves not “correct”, what shall we understand by “correct labelling”? The difficulty to find both intuitive answers and questions about the track labelling problem urges us to look at it from a more fundamental perspective. This requires a rigorous formulation and analysis of the problem of multi-target tracking and labelling (MTTL) in a Bayesian framework.

This idea of jointly estimating target identities together with states is known for some time in the literature, e.g. in an

early work of Salmond, Fisher and Gordon [12]. However, to rigorously handle general multi-target scenarios with target birth and death, plus unknown number of targets, a more sophisticated mathematical basis is required, such as Finite Set Statistics (FISST) [10]. The idea of using FISST to perform joint multi-object tracking and labelling appears in a number of works, e.g. [9, 14]. In our work, however, we will look at the general track labelling problem rather than a specific algorithm or application.

The organization and contributions of this paper are as follows. In Section 2, we provide a mathematical description of the general MTTL problem using FISST. In Section 3, we provide a mathematical characterization of the “mixed labelling” phenomenon. In Section 4.1, we provide a *statistical description of the labelling error* with clear physical interpretation: the labelling probability (i.e. we give a proper formulation for **Question 2**). In Section 4.2, we propose a conceptual *track extraction scheme* for MTT algorithms which has a number of nice properties, including being applicable to scenarios with target birth and death and giving a proper formulation for **Question 1**. In Section 5, we provide methods to calculate the labelling probability for two well-known MTT algorithms: the Multiple Hypothesis Tracking (MHT) and the Multi-target Sequential Monte Carlo (M-SMC) filter. This means that we also give *answers* to **Questions 1 and 2**. Section 6 draws conclusions.

2 Bayes formulation of the multi-target tracking and labelling (MTTL) problem

Before we describe the formulation, we will present a few notation conventions that will be used throughout this work. An upper-case letter (like X) will denote a vector-valued random variable, and its lower-case counterpart (like x) will, as usual, denote a particular realization. An upper-case bold-faced letter (like \mathbf{X}) will denote a finite set-valued random variable, and its lower-case counterpart will denote the corresponding realization. The probability density of a vector-valued random variable X will be denoted as $p(x)$; the multi-object density of a RFS variable (that we refer to simply as RFS density) will be denoted as $f(\mathbf{x})$.

In the FISST formulation, the multi-target state, rather than being represented by a random vector, is represented by a random finite set (RFS) of form $\mathbf{X}_k = \{X_k^{(1)}, \dots, X_k^{(T_k)}\}$, where k denotes the time index, $X_k^{(i)}$ is a random vector denoting the state of a single target i , and T_k , the number of targets, is also a random variable. A detailed description of FISST and its application to the multi-target tracking problem can be found in [10].

In order to perform labelling jointly with tracking, we need to explicitly add labels to the multi-target state. In other words, the single-target state $X_k^{(i)}$ should have the form

$$X_k^{(i)} = \begin{bmatrix} S_k^{(i)} \\ L_k^{(i)} \end{bmatrix}, \text{ where } L_k^{(i)} \text{ denotes the target's assigned}$$

label, and $S_k^{(i)}$ denotes all other state components (position, velocity, etc.). In FISST, the statistical information about

this RFS state is represented by the RFS density

$$f(\mathbf{x}_k | Z^k) = f \left(\left\{ \left[s_k^{(1)}, l_k^{(1)} \right]', \dots, \left[s_k^{(t_k)}, l_k^{(t_k)} \right]' \right\} \middle| Z^k \right)$$

where Z^k denotes the collection of observations up to and including time k .

With appropriate Markov assumptions, the Bayesian recursion for the RFS density has the form

$$f(\mathbf{x}_k | Z^k) = \frac{f(\mathbf{z}_k | \mathbf{x}_k) f(\mathbf{x}_k | Z^{k-1})}{f(\mathbf{z}_k | Z^{k-1})} \quad (1)$$

where \mathbf{z}_k denotes the most recent set of observations, $f(\mathbf{z}_k | \mathbf{x}_k)$ is the multi-object likelihood function and

$$f(\mathbf{x}_k | Z^{k-1}) = \int f(\mathbf{x}_k | \mathbf{x}_{k-1}) f(\mathbf{x}_{k-1} | Z^{k-1}) \delta \mathbf{x}_{k-1} \quad (2)$$

$$f(\mathbf{z}_k | Z^{k-1}) = \int f(\mathbf{z}_k | \mathbf{x}_k) f(\mathbf{x}_k | Z^{k-1}) \delta \mathbf{x}_k \quad (3)$$

where $\int \dots \delta \mathbf{x}$ denotes a *set integral* (see definition in [10, pp. 361–362]). In order to implement (1), we need to calculate $f(\mathbf{z}_k | \mathbf{x}_k)$ and $f(\mathbf{x}_k | \mathbf{x}_{k-1})$. We will hence have a separate look into these densities.

2.1 The likelihood $f(\mathbf{z}_k | \mathbf{x}_k)$

Let $\mathbf{S}_k = \{S_k^{(1)}, \dots, S_k^{(T_k)}\}$ correspond to the unlabelled multi-target state. We assume that observations are independent of labels, conditioned on the rest of the state, i.e.

$$f(\mathbf{z}_k | \mathbf{x}_k) = f(\mathbf{z}_k | \mathbf{S}_k). \quad (4)$$

We can then construct $f(\mathbf{z}_k | \mathbf{S}_k)$, for various types of observations, using the guidelines in [10, chap. 12]. Note that assumption (4) is not restrictive; we can ensure that it always holds by proper modeling. For instance, if we have observations of “identity-like” information (such as identification friend-or-foe (IFF) messages), this “identity-like” information (in our example, the IFF code) can be explicitly modeled as a state component of $S_k^{(i)}$.

2.2 The state transition density $f(\mathbf{x}_k | \mathbf{x}_{k-1})$

2.2.1 No target births or deaths

Let $p \left(s_k^{(i)} \middle| s_{k-1}^{(j)} \right)$ be the single-target state transition density, i.e. the motion model that describes the transition from the single-target state $s_{k-1}^{(j)}$ to $s_k^{(i)}$. Assuming that single-target dynamics are decoupled, i.e., $f(\mathbf{x}_k | \mathbf{x}_{k-1})$ can be factorized into single-target densities, from [10, chap. 13], we have

$$\begin{aligned} f(\mathbf{x}_k | \mathbf{x}_{k-1}) &= \sum_{\theta \in \Theta_{t_k}} p \left(x_k^{(1)} \middle| x_{k-1}^{(\theta(1))} \right) \dots p \left(x_k^{(t_k)} \middle| x_{k-1}^{(\theta(t_k))} \right) \quad (5) \end{aligned}$$

where Θ_{t_k} is the set of all permutations on $\{1, \dots, t_k\}$. Observe now that

$$p\left(s_k^{(i)}, l_k^{(i)} \mid s_{k-1}^{(j)}, l_{k-1}^{(j)}\right) = \begin{cases} p\left(s_k^{(i)} \mid s_{k-1}^{(j)}\right), & l_k^{(i)} = l_{k-1}^{(j)} \\ 0, & l_k^{(i)} \neq l_{k-1}^{(j)} \end{cases} \quad (6)$$

since a target cannot change its label. Therefore, (5) can be simplified to

$$\begin{aligned} f(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \\ = p\left(s_k^{(1)} \mid s_{k-1}^{(\tilde{\theta}(1))}\right) \dots p\left(s_k^{(t_k)} \mid s_{k-1}^{(\tilde{\theta}(t_k))}\right) \end{aligned} \quad (7)$$

where

$$\tilde{\theta} \in \Theta_{t_k}, \text{ s.t. } l_k^{(1)} = l_{k-1}^{(\tilde{\theta}(1))}, \dots, l_k^{(t_k)} = l_{k-1}^{(\tilde{\theta}(t_k))}.$$

2.2.2 With target births and deaths

Let $p\left(s_k^{(i)} \mid s_{k-1}^{(j)}\right)$ denote again the single-target state transition density and $p_S\left(s_{k-1}^{(i)}\right)$ denote the survival probability, i.e. the probability that target survives from time step $k-1$ to time k , which may depend on $s_{k-1}^{(i)}$ (it is not a density on $s_{k-1}^{(i)}$!). Let us additionally assume that

- the birth process is a multi-object Poisson process (see [10, p. 366]), i.e. the state distributions of appearing targets are mutually independent and the rate that new targets are born is Poisson-distributed with mean μ ;
- the state distributions of appearing targets are independent from the state of existing targets.

Using similar derivations to those made for the scenario without target births/deaths (with details omitted here for the sake of brevity), it is possible to show that $f(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ is given by

$$\begin{aligned} f(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \\ = e^{-\mu} \prod_{m \in \Gamma_b} \mu p_B\left(x_k^{(m)}\right) \prod_{n \in \Phi_{\tilde{\theta}}} \left(1 - p_S\left(s_{k-1}^{(n)}\right)\right) \\ \times \prod_{i \in \Phi_{\tilde{\theta}}} p_S\left(s_{k-1}^{(i)}\right) p\left(s_k^{(\tilde{\theta}(i))} \mid s_{k-1}^{(i)}\right) \end{aligned} \quad (8)$$

where, for $i \in \{1, \dots, t_{k-1}\}$:

$$\begin{aligned} \tilde{\theta}(i) &= \begin{cases} j, & \text{if } l_{k-1}^{(i)} = l_{k-1}^{(j)} \text{ for some } j \in \{1, \dots, t_k\} \\ 0, & \text{otherwise} \end{cases}, \\ \Phi_{\tilde{\theta}} &= \{i \mid i \in \{1, \dots, t_{k-1}\}, \tilde{\theta}(i) > 0\}, \\ \bar{\Phi}_{\tilde{\theta}} &\triangleq \{i \mid i \in \{1, \dots, t_{k-1}\}, i \notin \Phi_{\tilde{\theta}}\}, \\ \Gamma_b &\triangleq \left\{j \mid j \in \{1, \dots, t_k\}, l_k^{(j)} \notin \left\{l_{k-1}^{(1)}, \dots, l_{k-1}^{(t_{k-1})}\right\}\right\} \end{aligned}$$

and $p_B\left(x_k^{(m)}\right)$ is the single-target labelled state density of an appearing target. Its exact form depends on how we decide to assign labels to appearing states.

Attempting to specify $p_B\left(x_k^{(m)}\right)$ leads, however, to a problem. To derive (8), we have assumed that the labelled state

distributions of appearing targets are mutually independent, and that they are also independent from the labelled states of existing targets. Strictly speaking, however, we cannot assume this independence since we must ensure that the labels are at least *mutually different*. One possible “turn-around” to this problem is to draw the label $l_k^{(m)}$ of an appearing target from a continuous distribution (like a simple uniform distribution), which would at least ensure that the label *almost never* corresponds to the label of any other target.

3 The mixed labelling phenomenon

3.1 Mathematical characterization

Mixed labelling corresponds to a situation where there is ambiguity in labelling, i.e. in the assignment of labels ($l_k^{(i)}$) to locations (where a “location” here means simply an unlabelled single-target state $s_k^{(i)}$). We will now describe the phenomenon mathematically¹, using the Bayesian formulation of the MTTL problem from Section 2.

Given a set of locations $\mathbf{s}_k = \left\{s_k^{(1)}, \dots, s_k^{(t)}\right\}$, let

$$\begin{aligned} \Pi_k(\mathbf{s}_k) \\ = \left\{ \mathbf{x}_k \mid \mathbf{x}_k = \left\{ \left[\begin{array}{c} s_k^{(1)} \\ l_k^{(1)} \end{array} \right], \dots, \left[\begin{array}{c} s_k^{(t)} \\ l_k^{(t)} \end{array} \right] \right\}, f(\mathbf{x}_k \mid Z^k) > 0 \right\}. \end{aligned} \quad (9)$$

For a given \mathbf{s}_k , a situation of “no mixed labelling” would be when, for some $\hat{\mathbf{x}}_k \in \Pi_k(\mathbf{s}_k)$, we have

$$f(\hat{\mathbf{x}}_k \mid Z^k) \gg f(\mathbf{x}_k \mid Z^k), \forall \mathbf{x}_k \in \Pi_k(\mathbf{s}_k) \setminus \hat{\mathbf{x}}_k, \quad (10)$$

which means that for a set of unlabelled states $\left\{s_k^{(1)}, \dots, s_k^{(t)}\right\}$, there is only one logical choice of labels to be assigned to these states. Note that two elements $\mathbf{x}_k, \hat{\mathbf{x}}_k$ of $\Pi_k(\mathbf{s}_k)$ have always the same number of dimensions, so their RFS densities are always comparable.

Conversely, a “total mixed labelling” (for a given \mathbf{s}_k) would be when

$$f(\hat{\mathbf{x}}_k \mid Z^k) \approx f(\mathbf{x}_k \mid Z^k), \forall \mathbf{x}_k, \hat{\mathbf{x}}_k \in \Pi_k(\mathbf{s}_k) \quad (11)$$

i.e. all possible labellings are equally probable. In this situation, we can say that there is not a single “correct labelling” for the set of locations $\left\{s_k^{(1)}, \dots, s_k^{(t)}\right\}$.

Naturally, any situation that corresponds to neither (10), nor (11) can be referred to as “partial mixed labelling”.

Remark 3.1 *Mixed labelling, as have we described it, is a characteristic of a set of locations \mathbf{s}_k given the multi-target posterior; i.e. a local property. In practice, for labelling purposes, we are typically only interested in a subset of the elements of the state space of s_k . For instance, we may just be interested in labelling the estimated locations, i.e.*

¹We remark that the provided description is intended to give intuition to problem and not to be strictly rigorous, as we are resorting to operators \gg, \approx .

the tracks $\hat{s}_k = \{\hat{s}_k^{(1)}, \dots, \hat{s}_k^{(t)}\}$ displayed to the operator. Therefore, although it may be possible to describe the phenomenon in a “non-local” manner, we believe that this description suffices for most practical purposes.

3.2 Mixed labelling due to closely spaced targets

The occurrence of mixed labelling when targets separate after moving in close proximity to each other has been empirically observed, as in [4]. When the multi-target Bayes recursion is implemented by a particle filter, mixed labelling manifests itself by particle clouds corresponding to each target intersecting each other, as shown in Fig. 2.

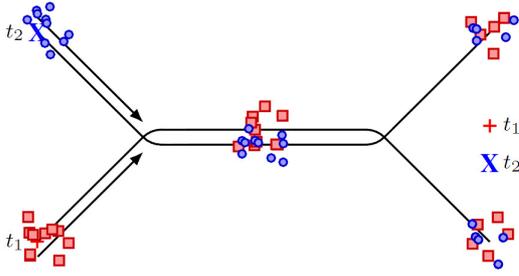


Figure 2: Particle representation of the multi-target distribution in a situation where mixed labelling occurs ([5])

We also verified the occurrence of mixed labelling in such situation (for the two-target case) by performing a theoretical analysis on the exact multi-target Bayes recursion (see details of this analysis in tech. rep. [1, sect. III]).

3.3 “Natural” vs. “artificial” elimination of mixed labelling

Since **Questions 1 and 2** proposed in Section 1 exist because of mixed labelling, one may then ask: instead of bothering ourselves with these questions, why not simply use an algorithm that “eliminates” mixed labelling?

It is very important, however, to remark that mixed labelling, being associated with the *exact multi-target posterior*, is a property of the *physical problem*, not of any particular algorithm!

We have identified some situations where mixed labelling may “naturally” be eliminated, i.e. be eliminated from the exact posterior. These situations are described in detail in tech. rep. [1, sect. III]. An obvious situation of “natural elimination” of mixed labelling is when measurements carry information about the target identities, for instance, the IFF code. Another situation is when one of the state components corresponds to the target classification (e.g. helicopter, fighter aircraft, commercial aircraft), and each target was precisely classified before mixed labelling happened. In this case, mixed labelling may disappear if each target starts exhibiting dynamics unique to their classification.

On the other hand, what may also happen is that mixed labelling still exists in the exact multi-target posterior, but it is not visible in the output of the chosen multi-target tracking algorithm. This “artificial elimination” of mixed la-

labelling, also referred as *self-resolving*, is typical of particle filter and multiple hypotheses implementations of the multi-target Bayes recursion, and has been identified in [4]. **“Self-resolving” should be generally treated as a problem, not as a “solution”**, because it causes a true ambiguity in the posterior to be underestimated by the filter. Some approaches to deal with self-resolving are described in recent works [2, 5, 7].

3.4 Initial mixed labelling

The phenomenon of “initial mixed labelling”, that we describe as mixed labelling affecting tracks originated by appearing targets, to the best of our knowledge, has not been yet discussed in previous literature. We will provide here only a preliminary discussion about the phenomenon.

In Section 2.2.2, we discussed the multi-target state transition density $f(\mathbf{x}_k | \mathbf{x}_{k-1})$ for states containing labels. However, alongside the support of this density, we may have different labels assigned to the same single-states (associated with appearing targets). How that precisely happens depends on the scheme for assigning labels to appearing targets.

What are the practical consequences of initial mixed labelling? The usefulness of labels is to identify, at some time step k , which tracks correspond to which tracks at some previous time step, say j . But if a target has just appeared, it did not originate a track at time j ; hence, which exact label is displayed for this track (and hence any mixed labelling that may be associated with it) is irrelevant. Therefore, it may be reasonable to devise a scheme to perform “artificial” (i.e. at implementation level) elimination of initial mixed labelling.

4 Statistics for optimal track labelling

4.1 The labelling probability

On basis of the Bayesian formulation of the MTTL problem, and the mathematical characterization of the mixed labelling phenomenon, we are ready to propose a *mathematical formulation* for **Question 2** of Section 1, i.e. for the probability of labelling error. We will do that through a sequence of statements.

Definition 4.1 Let M be a random vector. We say that M is a *partial state* of another random vector X , if all entries of M are also entries of X .

Definition 4.2 Let $\{X^{(1)}, \dots, X^{(t)}\}$ be a RFS variable, such that each single-state state is given by $X^{(i)} = [M^{(i)}, N^{(i)}]'$ (or, alternatively, $X^{(i)} = [N^{(i)}, M^{(i)}]'$), i.e. $M^{(i)}$ and $N^{(i)}$ are partial states of $X^{(i)}$. We then define the $M^{(\cdot)} | N^{(\cdot)}$ -*split density* of the RFS variable $\{X^{(1)}, \dots, X^{(t)}\}$ as

$$f_{M^{(\cdot)} | N^{(\cdot)}}(\{x^{(1)}, \dots, x^{(t)}\}) \triangleq \frac{f(\{x^{(1)}, \dots, x^{(t)}\})}{f(\{n^{(1)}, \dots, n^{(t)}\})}. \quad (12)$$

where $x^{(i)} = [m^{(i)}, n^{(i)}]'$ (or $x^{(i)} = [n^{(i)}, m^{(i)}]'$ as appropriate).

Remark 4.3 If $M^{(i)}$ assumes values in a discrete space, it is possible to show (details omitted here; see tech. rep. [1, sect. II.B.1]) that $f_{M^{(\cdot)}|N^{(\cdot)}}(\mathbf{x})$ is the probability mass associated with the finite set $\mathbf{x} = \{x^{(1)}, \dots, x^{(t)}\}$, conditioned on the finite set $\mathbf{n} = \{n^{(1)}, \dots, n^{(t)}\}$, with each element of \mathbf{n} being a partial state of a distinct element of \mathbf{x} . It may sound counterintuitive that \mathbf{x} , not being a discrete variable, is associated with a probability mass. The key point is that the conditioning on \mathbf{n} causes \mathbf{x} to only be able to assume discrete values, specifically, the possible assignments of the partial states $m^{(1)}, \dots, m^{(t)}$ to the partial states $n^{(1)}, \dots, n^{(t)}$.

Definition 4.4 Consider a RFS \mathbf{X}_k described as in Section 2. We define the **labelling probability** as the probability mass of a finite set of labelled target states $\mathbf{x}_k = \left\{ \left[s_k^{(1)}, l_k^{(1)} \right]', \dots, \left[s_k^{(t)}, l_k^{(t)} \right]' \right\}$, conditioned on the finite set of unlabelled states $\mathbf{s}_k = \left\{ s_k^{(1)}, \dots, s_k^{(t_k)} \right\}$ and observations Z^k . The quantity is given by

$$p_l(\mathbf{x}_k | \mathbf{s}_k) \triangleq f_{L(\cdot)|S(\cdot)}(\mathbf{x}_k | Z^k). \quad (13)$$

Remark 4.5 From Remark 4.3, the labelling probability may be interpreted as the posterior probability of an assignment of labels to unlabelled states, in the assumption that these unlabelled states match the true target locations². Definition 4.4 can then be readily used to mathematically formulate **Question 2** of Section 1. For a set of labelled tracks $\hat{\mathbf{x}}_k = \left\{ \hat{x}_k^{(1)}, \dots, \hat{x}_k^{(t)} \right\}$ and the corresponding unlabelled tracks $\hat{\mathbf{s}}_k = \left\{ \hat{s}_k^{(1)}, \dots, \hat{s}_k^{(t)} \right\}$, the probability of error in label-to-track assignment can be described by $1 - p_l(\hat{\mathbf{x}}_k | \hat{\mathbf{s}}_k)$.

4.2 The MMOSPA-MLP estimate

We are now ready to propose a conceptual track extraction scheme specially suited for the optimal tracking problem. Let $\mathbf{S}_k = \left\{ S_k^{(1)}, \dots, S_k^{(t_k)} \right\}$ denote the RFS corresponding to the unlabelled states. In our proposed scheme, the set of labelled tracks $\hat{\mathbf{x}}_k$ is the solution of the problem given by

$$\hat{\mathbf{s}}_k = \arg \inf_{\bar{\mathbf{s}}_k} \int \left(\epsilon_p^{(c)}(\mathbf{s}_k, \bar{\mathbf{s}}_k) \right)^p f(\mathbf{s}_k | Z^k) \delta \mathbf{s}_k \quad (14)$$

$$\hat{\mathbf{x}}_k = \arg \max_{\mathbf{x}_k} p_l(\mathbf{x}_k | \hat{\mathbf{s}}_k) \quad (15)$$

where $\epsilon_p^{(c)}$ is the Optimal Subpattern Assignment (OSPA) metric defined by Schuhmacher, Vo and Vo [13] and c and p are parameters discussed in the same work.

The rationale of the estimate given by (14)–(15) is quite simple. In the first step (14), we obtain the unlabelled tracks, according to the Minimum Mean OSPA (MMOSPA) estimate defined by Guerriero et al. [8]. This corresponds, hence, to an optimal choice (in Mean OSPA sense) of unlabelled tracks, which additionally avoids track coalescence.

²Obviously this assumption is almost never true, but the same often holds for conditional probabilities in general.

In the second step (15), the labelled tracks are obtained by using the previously obtained MMOSPA estimate and choosing the assignment of labels that maximizes the labelling probability according to Definition 4.4. We refer to this two-step scheme as **MMOSPA-MLP** estimate (where MLP stands for Maximum Labelling Probability).

Note that second step (15) also gives, for **Question 1** proposed in Section 1, a proper formulation (in the sense of being mathematically rigorous and having clear physical interpretation).

5 Calculating the labelling probabilities for existing MTT algorithms

We will show how to approximate the labelling probabilities described in Section 4.1 for two existing MTT algorithms. This corresponds to answering **Question 2** proposed in Section 1, and using the MLP step (15), it also corresponds to answering **Question 1**. Calculation of the MMOSPA estimate (14) is not discussed here. Note, however, that the MLP step can be combined with any other method (i.e. other than the MMOSPA step) to obtain a set of unlabelled tracks.

The following relationship, that holds for the labelling probability (derivation omitted here), will be particularly useful:

$$p_l(\mathbf{x}_k | \mathbf{s}_k) = \frac{\int f(\mathbf{x}_k | \mathbf{x}_{k-1}) f(\mathbf{x}_{k-1} | Z^{k-1}) \delta \mathbf{x}_{k-1}}{f(\mathbf{s}_k | Z^{k-1})} \quad (16)$$

where we remind that \mathbf{s}_k also occurs implicitly in \mathbf{x}_k .

We remark that both filtering algorithms suffer from the self-resolving phenomenon described in Section 3.3. This means that the calculated labelling probabilities will gradually lose accuracy after target separation.

5.1 Multi-target Sequential Monte Carlo (M-SMC) filter

The M-SMC filter described in [10, pp. 551–564], with labels treated as state components, as in [9], corresponds to the particle filter implementation of the Bayesian recursion (1). Note that the “Joint Multi-track Particle Filter” described by García-Fernández and Grajal [6] is a similar algorithm, albeit with a different derivation.

The multi-target density $f(\mathbf{x}_k | Z^k)$ is represented by a set of particles $\{\mathbf{x}_k(i), w_k(i)\}_{i=1}^{N_P}$, where $\mathbf{x}_k(i)$ denotes a realization of multi-target state, $w_k(i)$ the particle weight, and N_P the number of particles.

Labelling probabilities can then be calculated by straightforward particle approximation of (16), i.e.

$$p_l(\mathbf{x}_k | \mathbf{s}_k) \propto \sum_{j=1}^{N_P} w_{k-1}(j) f(\mathbf{x}_k | \mathbf{x}_{k-1}(j)). \quad (17)$$

with proportionality turned into an equality by normalization over all $\mathbf{x}_k \in \Pi_k(\mathbf{s}_k)$ (we can do that since $p_l(\mathbf{x}_k | \mathbf{s}_k)$ corresponds to a conditional probability mass; see remark 4.3). Note that the true cardinality of $\Pi_k(\mathbf{s}_k)$ may be very large when target births and deaths are considered, but the

considered values will be restricted by the particle approximation. Even so, additional labelling pruning mechanisms may be necessary.

5.2 Multiple Hypothesis Tracking (MHT)

In the MHT algorithm [11], the multi-target density $f(\mathbf{x}_k|Z^k)$ is represented by a set of hypotheses $\{\mathbf{h}_k(i), w_k(i)\}_{i=1}^{N_H}$ where $\mathbf{h}_k(i)$ denotes an hypothesis on the multi-target state, $w_k(i)$ the hypothesis weight, and N_H the number of hypotheses. Each hypothesis has form $\mathbf{h}_k(i) = \{h_k^{(1)}(i), \dots, h_k^{(t_k(i))}(i)\}$, where the single-target hypothesis $h_k^{(j)}(i)$ is given by a triple:

$$\left(\hat{s}_k^{(j)}(i), l_k^{(j)}(i), P_k^{(j)}(i)\right) \quad (18)$$

where $\hat{s}_k^{(j)}(i)$ and $P_k^{(j)}(i)$ are respectively hypotheses on the mean and the covariance of the unlabelled single-target state $S_k^{(j)}$, and $l_k^{(j)}(i)$ is an hypothesis on the corresponding label.

We can use the following procedure to approximate the labelling probabilities for the MHT. At every time step, we produce a number of samples, say N_P , by sampling from the set of hypotheses $\{\mathbf{h}_k(i), w_k(i)\}_{i=1}^{N_H}$. In other words, for samples $m = 1, \dots, N_P$, we do the following:

1. Choose an hypothesis index i_m using multinomial sampling with probabilities $\{w_k(i)\}_{i=1}^{N_H}$
2. For $j = 1, \dots, t_k(i_m)$, sample

$$s_k^{(j)}(i_m) \sim \mathcal{N}\left(\hat{s}_k^{(j)}(i_m); P_k^{(j)}(i_m)\right) \quad (19)$$

3. Make $\mathbf{x}_k(m) = \left\{ \begin{bmatrix} s_k^{(1)}(i_m) \\ l_k^{(1)}(i_m) \end{bmatrix}, \dots, \begin{bmatrix} s_k^{(t_k(i_m))}(i_m) \\ l_k^{(t_k(i_m))}(i_m) \end{bmatrix} \right\}$

Labelling probabilities are then calculated using (17).

6 Conclusions

In this paper, we produced a detailed mathematical description of the optimal track labelling problem, with practical aspects, such as how to perform labelling and how to characterize the probability of labelling error, also being discussed. A recurring concern in this work was to ensure that the proposed statistics for the problem have clear physical interpretation – such that the user can decide whether they are appropriate or not for his/her application, and interpret their results in case he/she decides to use them.

In a future work, we will describe an algorithm that avoids the self-resolving phenomenon described in Section 3.3, and can also be applied to general scenarios with unknown/time-varying number of targets. We will also plan to look with more depth at the problem of track labelling with target birth taken in consideration, which shall include expanding our analysis on the “initial mixed labelling” phenomenon mentioned in Section 3.4.

Acknowledgements

The research leading to these results has received funding from the EU’s Seventh Framework Programme under grant agreement n° 238710. The research has been carried out in the MC IMPULSE project: <https://mcimpulse.isy.liu.se>.

This research has been also supported by the Netherlands Organisation for Scientific Research (NWO) under the Casimir program, contract 018.003.004. Under this grant Yvo Boers holds a part-time position at the Department of Applied Mathematics at the University of Twente.

We also thank Hans Driessen (Thales Nederland B.V.) and Ronald Mahler (Lockheed Martin) for the contributions.

References

- [1] E. H. Aoki, A. Bagchi, P. Mandal, and Y. Boers. A theoretical analysis of Bayes-optimal multi-target tracking and labelling. Technical Report 1953, University of Twente, Enschede, The Netherlands, 2011.
- [2] H. A. P. Blom and E. A. Bloem. Decomposed particle filtering and track swap estimation in tracking two closely spaced targets. In *Proc. 14th International Conference of Information Fusion*, Chicago, IL, 2011.
- [3] H. A. P. Blom, E. A. Bloem, Y. Boers, and J. N. Driessen. Tracking closely spaced targets: Bayes outperformed by an approximation? In *Proc. 11th International Conference on Information Fusion*, Cologne, Germany, 2008.
- [4] Y. Boers, E. Sviestins, and J. N. Driessen. Mixed labelling in multitarget particle filtering. *IEEE Trans. Aerosp. Electron. Syst.*, 46(2):792–802, 2010.
- [5] D. Crouse, P. Willett, L. Svensson, D. Svensson, and M. Guerriero. The set MHT. In *Proc. 14th International Conference of Information Fusion*, Chicago, IL, 2011.
- [6] Á. García-Fernández and J. Grajal. Multitarget tracking using the Joint Multitrack Probability Density. In *Proc. 12th International Conference on Information Fusion*, Seattle, WA, 2009.
- [7] Á. García-Fernández, M. Morelande, and J. Grajal. Particle filter for extracting target label information when targets move in close proximity. In *Proc. 14th International Conference of Information Fusion*, Chicago, IL, 2011.
- [8] M. Guerriero, L. Svensson, D. Svensson, and P. Willett. Shooting two birds with two bullets: how to find Minimum Mean OSPA estimates. In *Proc. 13th International Conference on Information Fusion*, Edinburgh, UK, 2010.
- [9] W.-K. Ma, B.-N. Vo, S. Singh, and A. Baddeley. Tracking an unknown time-varying number of speakers using TDOA measurements: A random finite set approach. *IEEE Trans. Signal Process.*, 54(9):3291–3304, 2006.
- [10] R. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House, Norwood, MA, 2007.
- [11] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Trans. Autom. Control*, AC-24(6), December 1979.
- [12] D. J. Salmond, D. Fisher, and N. J. Gordon. Tracking and identification for closely spaced objects in clutter. In *Proc. European Control Conf.*, 1997.
- [13] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Process.*, 56(8):3447–3457, 2008.
- [14] A.-T. Vu, B.-N. Vo, and R. Evans. Particle Markov chain Monte Carlo for Bayesian multi-target tracking. In *Proc. 14th International Conference of Information Fusion*, Chicago, IL, 2011.