Bayesian sequential track formation

Ángel F. García-Fernández, Mark R. Morelande, Jesús Grajal

Abstract—This paper presents a theoretical framework for track building in multiple target scenarios from the Bayesian point of view. It is assumed that the number of targets is fixed and known. We propose two optimal methods for building tracks sequentially. The first one uses the labelling of the current multitarget state estimate that minimises the mean square labelled optimal subpattern assignment error. This method requires knowledge of the posterior density of the vector-valued state. The second assigns the labelling that maximises the probability that the current multitarget state estimate is optimally linked with the available tracks at the previous time step. In this case, we only require knowledge of the random finite set posterior density without labels.

Index Terms—Target labelling, multiple target tracking, Bayesian framework, random finite sets

I. INTRODUCTION

Multitarget tracking systems should solve two basic problems. The first one is to estimate the number of targets and their states at the current time. The second one is to connect target state estimates that belong to the same target along time to form tracks. Classic approaches to multitarget tracking that perform these aims are, for example, multiple hypothesis tracking (MHT) [1] and joint probabilistic data association (JPDA) [2]. The random finite set (RFS) framework is another approach to perform these tasks. However, most of the developed algorithms within the RFS framework have focused on the first problem and do not deal with the second problem from a theoretically sound perspective, e.g., probability hypothesis density (PHD) filter, cardinalised PHD (CPHD) filter or the multitarget multi-Bernoulli (MeMBer) filter [3], [4].

In the Bayesian approach, all the information of interest about target states at all time steps is contained in its joint probability density function (PDF) given the measurements [5], which we refer to as the trajectory posterior PDF. This suggests that principled approaches to track building should be obtained from this PDF. Nevertheless, most approaches to multiple target tracking focus on filtering posterior density computation because of the difficulty of approximating the

Ángel F. García-Fernández was with Departamento de Señales, Sistemas y Radiocomunicaciones, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain (email: angel.garciafernandez@curtin.edu.au).

Mark R. Morelande is with the Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia (email: mrmore@unimelb.edu.au).

Jesús Grajal is with Departamento de Señales, Sistemas y Radiocomunicaciones, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain. (email: jesus@gmr.ssr.upm.es).

trajectory posterior PDF. Therefore, the usual approach to track building, which will also be adopted here, is based on filtering posteriors.

Using these filtering posteriors, the conventional approach to track building involves adding a label to each target state. This way we can directly provide labelled multitarget estimates, which form the tracks. Labels were used for track formation in [6] using a vector-based formulation and in [7], [8] using the RFS framework. The approaches of [6] and [7], [8] are equivalent due to a bijection between both representations [9, Appendix B]. For the same reason, for fixed and known number of targets, modelling the multitarget state as a vector is equivalent to a labelled set. However, the tracks produced by these algorithms are sometimes of poor quality. For example, when targets move in close proximity for an extended period of time, tracks based on the maximum a posteriori (MAP) estimator switch uncontrollably between the target states and, tracks based on the minimum mean square error (MMSE) estimator do not provide meaningful target state estimates [10]. An ad-hoc solution to provide jitter-free tracks, based on the minimum mean square optimal subpattern assignment (MSOSPA) estimator [11], [12], was proposed in [13]. This method is not of general applicability as is restricted to MHT.

In this paper, we address this theoretical gap and develop principled approaches to track building in the Bayesian framework for fixed and known number of targets based on the (filtering) posterior PDF. We wish to clarify that it is not the purpose of this paper to consider methods of posterior PDF approximation, e.g., algorithms such as MHT, JPDA or PHD filters. Rather, we assume the availability of the posterior PDF at each time step and develop the tools required for building optimal tracks. As there are in the literature optimal approaches to (unlabelled) target state estimation, e.g., based on minimising the MSOSPA error, we set aside this problem and assume that we are given a collection of single target state estimates at each time step. Track building in this context consists of labelling these target state estimates such that a link between target state estimates along time is established. We develop principled approaches to this problem of general applicability. The results can be used whenever it is desired to build tracks given a sequence of target state estimates. There are many practical applications where this is desired, such as radar, sonar and video surveillance. Our theory of optimal track building only requires a metric that accounts for the labels apart from the posterior PDF. We use the labelled optimal subpattern assignment (LOSPA) metric [14] with some given parameters due to simplified mathematical expressions but other metrics for labelled sets could be used in principle. Algorithms based on these criteria, along with their relationships to existing work, are discussed briefly in

the following paragraphs.

In the first algorithm to optimal track building, we minimise the mean square LOSPA (MSLOSPA) error. Therefore, in this approach, the objective is to choose the labelling of the estimate that is closest to the true multitarget state, according to LOSPA. For low penalty of labelling error in LOSPA, the MSLOSPA error can be decomposed into two terms, the MSOSPA error, which does not take labels into account, plus the mean labelling error cost (MLEC). Considering this decomposition, the labelling of the multitarget estimate that minimises the MSLOSPA error is the one that minimises the MLEC. The MLEC depends on the optimal labelling probabilities. These are simply calculated using probability theory and the definition of optimal labelling of a multitarget state estimate. They also provide useful information, for instance, if it is necessary to distinguish between friendly and enemy vehicles that get in close proximity and then separate [15]. Although many papers use the concepts of labelling and labelling probabilities, these concepts have not been clearly defined in the literature yet. To the authors' knowledge, the fact that the (optimal) labelling probabilities depend on an estimate and a metric was first indicated in [9, Sec. 5.3] for a two-target case. In general, this has been usually overlooked in the literature [15]–[18]. In [19], the labelling probabilities for a two-target case are those previously indicated in [9, Sec. 5.3]. In [20], the authors provide a definition of labelling probabilities. However, a proper definition of optimal target labelling probabilities should first define what optimal target labelling is and the probabilities should just be calculated using probability theory.

In some cases, we are interested in minimising track switching rather than obtaining the labelling of the estimate that is closest to the true multitarget state, e.g, as in [13]. In this paper, we formulate this problem in a principled approach as the selection of the labelling of the estimate that maximises the probability that the current multitarget state is optimally linked with the tracks available at the previous time step. On the whole, this method of track building outperforms the one based on the MSLOSPA error. As opposed to the conventional approach, we show that this track building procedure does not require a labelled state, i.e., we can use an unlabelled set to represent the multitarget state. As we explain in the paper, this result is of great utility in the design of multitarget tracking algorithms.

The remainder of the paper is organised as follows. In Section II, we formulate the problem. The first procedure for track formation based on labelled states is explained in Section III. The second procedure for track building, which minimises track switching and does not require a labelled state, is explained in Section IV. Both approaches to track building are compared in Section V. A numerical example that illustrates the ideas put forward in this paper is given in Section VI. Finally, conclusions are drawn in Section VII.

II. PROBLEM FORMULATION

The multitarget state vector at time k is $\mathbf{X}^k = \left[\left(\mathbf{x}_1^k\right)^T, \left(\mathbf{x}_2^k\right)^T, ..., \left(\mathbf{x}_t^k\right)^T\right]^T \in \mathbb{R}^{tn_x}$ where $\mathbf{x}_j^k \in \mathbb{R}^{n_x}$ is

the state vector at time k for target $j \in \{1, ..., t\}$, t is the known target number and T denotes transpose. The multitarget state can be equivalently represented by a labelled set $\left\{\left[\left(\mathbf{x}_1^k\right)^T, l_1\right]^T, \left[\left(\mathbf{x}_2^k\right)^T, l_2\right]^T, ..., \left[\left(\mathbf{x}_t^k\right)^T, l_t\right]^T\right\}$ where $l_j \in \mathbb{R}$ represents the jth label. Labels are unique, assigned deterministically and do not change with time. This implies that vector and labelled set notations are equivalent for fixed and known number of targets. The labels of the labelled set are implicit in the ordering inherent in the multitarget state vector components.

The multitarget state along time is modelled as a Markov process with a transition density $f\left(\cdot \left|\mathbf{X}^{k-1}\right.\right)$. At time k, the multitarget state vector is observed through noisy measurements, which are usually represented as a vector $\mathbf{z}^k \in \mathbb{R}^{n_{z,1}}$ [6] or as a set $Z^k \subset \mathbb{R}^{n_{z,2}}$ [3]. We use $\mathbf{z}^k$ to denote the measurements in the rest of the paper but our paper is general and we can use Z^k instead. The likelihood of $\mathbf{X}^k$ after observing $\mathbf{z}^k$ is denoted by $\ell\left(\mathbf{z}^k \left|\mathbf{X}^k\right.\right)$.

In Bayesian filtering, all the information of interest about $\mathbf{X}^k$ is included in the posterior PDF $\pi^k\left(\cdot\right)$ of the state given the sequence $\mathbf{z}^{1:k} = \left(\mathbf{z}^1, \mathbf{z}^2, ..., \mathbf{z}^k\right)$ of measurements up to time k. Given a prior $\pi^0\left(\cdot\right)$ at time 0, the transition density and the likelihood at every time step, the posterior PDF can be calculated recursively in two stages: prediction and update [21].

This recursion is usually intractable so it must be approximated. For instance, particle filters (PFs) can be applied for any type of transition density and likelihood function as PFs are general Bayesian filtering algorithms. If there is a point detection measurement model, in which there is a collection of measurements and each measurement can be originated from clutter or a target [3], [5], more specific algorithms such as MHT or JPDA [2] can be used.

However, this paper does not deal with posterior PDF approximation but with how to extract target labelling information and build tracks in a principled way regardless of measurement function or the approximation we perform. We therefore assume that the posterior $\pi^k\left(\cdot\right)$ is known although, in practice, the track building procedures indicated in this paper can be carried out using an arbitrary posterior PDF approximation. In the rest of this section, we formulate the problem of track formation.

Track formation

Multitarget tracking systems should estimate target states and link target state estimates over time to form tracks. The first goal is attained by providing individual target state estimates, which can be represented by sets $\hat{X}^k = \left\{\hat{\mathbf{x}}_1^k, \hat{\mathbf{x}}_2^k, ..., \hat{\mathbf{x}}_t^k\right\}$ $k = 0, 1, 2, ...$, where $\hat{\mathbf{x}}_j^k \in \mathbb{R}^{n_x}$ $j \in \{1, ..., t\}$ is an individual target state estimate at time k. There are in the literature optimal methods to obtain $\hat{X}^k$, e.g., based on minimum mean square OSPA (MMSOSPA) estimator ($p = 2$, $c = \infty$ and

Euclidean distance)[1] [11] or an approximation [15], [22], [23].

In this paper, we set aside this problem, which is already solved theoretically, and develop optimal approaches to link these optimal target state estimates to form tracks. In this context, track formation consists of selecting the labels, i.e., the order in the elements of $\hat{X}^k$, to build the multitarget state estimate $\hat{\mathbf{X}}^k$. The sequence of multitarget state estimates $\hat{\mathbf{X}}^k$ $k = 0, 1, 2, ...$ forms the tracks as it provides the target estimates that belong to the same target at different time steps.

This paper is especially concerned with providing two methods for sequential track building. That is, we have the tracks up to a certain time $k - 1$, which are given by $\hat{\mathbf{X}}^l$ $l = 0, 1, 2, ..., k - 1$. When the measurement at time k is received, we assume we obtain $\hat{X}^k$ at time k. Sequential track building consists of labelling $\hat{X}^k$ without modifying the available tracks. This implies that track building is performed while filtering and we directly provide tracks as output.

III. TRACKS THAT MINIMISE MSLOSPA ERROR

In this section we indicate how to build tracks by selecting the labelling of the multitarget state set estimate $\hat{X}^k$ that minimises the MSLOSPA error under the approximation of low penalty of labelling error. Before explaining this in Section III-D, we first review the LOSPA metric in Section III-A, define the notion of optimal labelling in Section III-B and explain how to calculate the optimal labelling probabilities in Section III-C.

A. Labelled OSPA metric

In order to build tracks, we use the OSPA metric for labelled sets defined in [14]. We represent the permutations of vector $[1, ..., t]^T$ as vectors $\boldsymbol{\phi}_i = [\phi_{i,1}, ..., \phi_{i,t}]^T$ $i \in \{1, ..., t!\}$. Then, the LOSPA distance between multitarget vectors $\mathbf{A}^k = \left[\left(\mathbf{a}_1^k\right)^T, \left(\mathbf{a}_2^k\right)^T, ..., \left(\mathbf{a}_t^k\right)^T\right]^T \in \mathbb{R}^{tn_x}$ and $\mathbf{B}^k = \left[\left(\mathbf{b}_1^k\right)^T, \left(\mathbf{b}_2^k\right)^T, ..., \left(\mathbf{b}_t^k\right)^T\right]^T \in \mathbb{R}^{tn_x}$ is [24]

$$d\left(\mathbf{A}^k, \mathbf{B}^k\right) =$$
$$\left(\frac{1}{t} \min_{i \in \{1,...,t!\}} \left[\sum_{j=1}^{t} b^p\left(\mathbf{a}_j^k, \mathbf{b}_{\phi_{i,j}}^k\right) + \alpha^p \overline{\delta}\left[j - \phi_{i,j}\right]\right]\right)^{1/p} \tag{1}$$

where $\overline{\delta}[\cdot]$ is the complement of the Kronecker delta, i.e., $\overline{\delta}[j] = 0$ if $j = 0$ and $\overline{\delta}[j] = 1$ otherwise, $\alpha > 0$, $1 \leq p < \infty$ and $b(\cdot, \cdot)$ is a metric on the space $\mathbb{R}^{n_x}$. In [14] the authors include another parameter p', we set $p' = p$ for simplicity.

---

[1]The *square* OSPA metric (OSPA to the power of two) with $p = 2$, $c = \infty$ and Euclidean distance is equivalent to the OSPA metric with $p = 1$, $c = \infty$ and square Euclidean distance as base metric. Therefore, $\hat{X}^k$ is also the minimum mean OSPA estimator for the OSPA with $p = 1$, $c = \infty$ and square Euclidean distance. In this paper, we always refer to the case $p = 2$, $c = \infty$ and Euclidean distance. Nevertheless, the whole paper can be rewritten using the OSPA with $p = 1$, $c = \infty$ and square Euclidean distance.

Table I: LOSPA between $\hat{\mathbf{X}}^k$ and $\mathbf{X}^k = [-10, 0, 10]^T$

Estimate $\hat{\mathbf{X}}^k$	LOSPA ($\alpha = 0.1$)	LOSPA ($\alpha = 1$)
$[-10.1, 0.1, 10.1]^T$	0.1	0.1
$[0.1, -10.1, 10.1]^T$	$\sqrt{0.1^2 + 0.02/3}$	$\sqrt{0.1^2 + 2/3}$
$[10.1, -10.1, 0.1]^T$	$\sqrt{0.1^2 + 0.03/3}$	$\sqrt{0.1^2 + 3/3}$

Illustrative example: We illustrate how the LOSPA metric works in a simple example. Let us assume there are three unidimensional targets and the multitarget state is $\mathbf{X}^k = [-10, 0, 10]^T$. That is, target 1 is at -10, target 2 is at 0 and target 3 is at 10. We use the Euclidean metric for $b(\cdot, \cdot)$ with $p = 2$. The LOSPA distance between $\mathbf{X}^k$ and several estimates $\hat{\mathbf{X}}^k$, which only differ in their labelling, are given in Table I. As all the estimates only differ in their labelling, they have the same OSPA distance, which is 0.1. This implies that all the estimates have the same accuracy as regards where the targets are. However, the first estimate is closer in the LOSPA sense than the rest. The higher α is, the more the metric penalises wrong labelling/ordering.

B. Optimal labelling of a multitarget state estimate

In this section, we provide the definition of optimal labelling of a multitarget state estimate. Then, we derive one of its properties and provide an illustrative example. Let $\Gamma_{\phi_i}\left(\mathbf{X}^k\right) = \left[\left(\mathbf{x}_{\phi_{i,1}}^k\right)^T, \left(\mathbf{x}_{\phi_{i,2}}^k\right)^T, ..., \left(\mathbf{x}_{\phi_{i,t}}^k\right)^T\right]^T$ be the permutation of $\mathbf{X}^k$ indicated by ϕ_i over single target states, e.g., for $t = 2$, $\phi_1 = [1, 2]^T$ and $\phi_2 = [2, 1]^T$, $\Gamma_{\phi_1}\left(\mathbf{X}^k\right) = \left[\left(\mathbf{x}_1^k\right)^T, \left(\mathbf{x}_2^k\right)^T\right]^T$ and $\Gamma_{\phi_2}\left(\mathbf{X}^k\right) = \left[\left(\mathbf{x}_2^k\right)^T, \left(\mathbf{x}_1^k\right)^T\right]^T$.

Definition 1. The optimal labelling of a multitarget state estimate $\hat{\mathbf{X}}^k = \left[\left(\hat{\mathbf{x}}_1^k\right)^T, \left(\hat{\mathbf{x}}_2^k\right)^T, ..., \left(\hat{\mathbf{x}}_t^k\right)^T\right]^T$ at time k is the vector $\tilde{\boldsymbol{\phi}} = \left[\tilde{\phi}_1, ..., \tilde{\phi}_t\right]^T$ such that

$$\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}_{i^\star} : i^\star = \arg\min_{i \in \{1,...,t!\}} d\left(\hat{\mathbf{X}}^k, \Gamma_{\phi_i}\left(\mathbf{X}^k\right)\right). \tag{2}$$

As proved in Appendix A, the optimal labelling corresponds to the labels of $\hat{\mathbf{X}}^k$ that minimise the LOSPA between $\mathbf{X}^k$ and $\hat{\mathbf{X}}^k$.

The following property is met for the optimal labelling of multitarget state estimates:

Lemma 2. *The optimal labelling of the multitarget state estimate does not depend on α and can be written as*

$$\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}_{i^\star} : i^\star = \arg\min_{i \in \{1,...,t!\}} \left(\frac{1}{t} \sum_{j=1}^{t} b^p\left(\hat{\mathbf{x}}_j^k, \mathbf{x}_{\phi_{i,j}}^k\right)\right)^{1/p}.$$

Lemma 2 is proved in Appendix B and indicates that the optimal labelling is the order of the true multitarget state that determines the OSPA.

It should be noted that the estimated labelling of $\hat{\mathbf{X}}^k$ is always $[1, ..., t]^T$ (the labelling of a vector is implicit in the order of its components). Therefore, we make the following definition

Table II: Optimal labelling of several multitarget state estimates for $\mathbf{X}^k = [-10, 0, 10]^T$

Estimate $\hat{\mathbf{X}}^k$	Optimal labelling $\tilde{\phi}$	$\tilde{N}$
$[-10.1, 0.1, 10.1]^T$	$[1, 2, 3]^T$	3
$[0.1, -10.1, 10.1]^T$	$[2, 1, 3]^T$	1
$[10.1, -10.1, 0.1]^T$	$[3, 1, 2]^T$	0

Definition 3. The multitarget state estimate $\hat{\mathbf{X}}^k$ has the correct labelling if $\tilde{\phi} = [1, 2, ..., t]^T$ (or equivalently $i^\star = 1$) and the estimate of the jth target in $\hat{\mathbf{X}}^k$ has the correct labelling if $\tilde{\phi}_j = j$.

We also say that $\hat{\mathbf{X}}^k$ and the jth target estimate of $\hat{\mathbf{X}}^k$ have the wrong labelling if they do not have the correct labelling.

Illustrative example: To illustrate the idea of optimal labelling we use the same example as in Section III-A. The optimal labelling of several multitarget state estimates are given in Table II, where $\tilde{N}$ denotes the number of target estimates with the correct labelling. The optimal labelling of $\hat{\mathbf{X}}^k = [-10.1, 0.1, 10.1]^T$ is $\tilde{\phi} = [1, 2, 3]^T$ which means that all targets have the correct labelling. However, if $\hat{\mathbf{X}}^k = [0.1, -10.1, 10.1]^T$, the optimal labelling is $\tilde{\phi} = [2, 1, 3]^T$ as target 1 estimate is closer to target 2 and target 2 estimate is closer to target 1. Only target 3 has the correct labelling.

C. Optimal labelling probabilities

According to Definition 1, the optimal labelling of a multitarget state estimate requires knowledge of the true multitarget state. In practice, the true multitarget state is unknown and what we know is its posterior PDF. This gives rise to the calculation of the optimal labelling probabilities. That is, using the posterior PDF and Definition 1, in this section, we indicate how to calculate the probability $P^k_{\hat{\mathbf{X}}^k}(i)$ that the estimate $\hat{\mathbf{X}}^k$ has optimal labelling ϕ_i for $i \in \{1, ..., t!\}$.

Let $\mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)$ denote the region that includes the values of the multitarget state $\mathbf{X}^k$ such that the optimal labelling of $\hat{\mathbf{X}}^k$ is ϕ_i:

$$\mathcal{S}_i\left(\hat{\mathbf{X}}^k\right) = \left\{ \mathbf{X}^k : i = \underset{l \in \{1,...,t!\}}{\arg\min} \, d\left(\hat{\mathbf{X}}^k, \Gamma_{\phi_l}\left(\mathbf{X}^k\right)\right) \right\} \quad (3)$$

where we have used Definition 1. Because of (3) and Lemma 2, regions $\mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)$ $i \in \{1, ..., t!\}$ constitute the Voronoi diagram given the points $\Gamma_{\phi_i}\left(\hat{\mathbf{X}}^k\right)$ $i \in \{1, ..., t!\}$ [25] and $\bigcup_{i=1}^{t!} \mathcal{S}_i\left(\hat{\mathbf{X}}^k\right) = \mathbb{R}^{tn_x}$.

Equation (3) is quite useful because $P^k_{\hat{\mathbf{X}}^k}(i)$ is just the probability that the target state $\mathbf{X}^k$ belongs to region $\mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)$. As a result, the optimal labelling probabilities become

$$P^k_{\hat{\mathbf{X}}^k}(i) = \int_{\mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)} \pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k \quad i \in \{1, ..., t!\}. \quad (4)$$

Using Definition 3, it is clear that the correct labelling probability (CLP) of $\hat{\mathbf{X}}^k$ is simply $P^k_{\hat{\mathbf{X}}^k}(1)$. For example, for $t = 2$, if $P^k_{\hat{\mathbf{X}}^k}(1) = 1$ and $P^k_{\hat{\mathbf{X}}^k}(2) = 0$, $\hat{\mathbf{X}}^k$ has the correct labelling with probability 1. This means that there is no confusion at all about target labelling given the sequence of measurements $\mathbf{z}^{1:k}$.

D. Labelling with the lowest MSLOSPA error

In this section, we indicate how to select the labelling of a multitarget state set estimate $\hat{X}^k$ that minimises the MSLOSPA under the approximation of low penalty of labelling error (small α). The MSLOSPA of $\hat{\mathbf{X}}^k$ is

$$\text{MSLOSPA}\left(\hat{\mathbf{X}}^k\right) = \text{E}\left[d^2\left(\mathbf{X}^k, \hat{\mathbf{X}}^k\right)\right] \quad (5)$$

where $d^2(\cdot, \cdot)$ is the square LOSPA and this expectation is done w.r.t. $\pi^k(\cdot)$. As shown in Appendix C, using $p = 2$, we get

$$\text{MSLOSPA}\left(\hat{\mathbf{X}}^k\right) \leq \text{MSOSPA}\left(\hat{\mathbf{X}}^k\right) + \text{MLEC}\left(\hat{\mathbf{X}}^k\right) \quad (6)$$

where $\text{MSOSPA}\left(\hat{\mathbf{X}}^k\right)$ and $\text{MLEC}\left(\hat{\mathbf{X}}^k\right)$ denote the MSOSPA error and MLEC of $\hat{\mathbf{X}}^k$, respectively, and the inequality is tight for $\alpha \to 0$. The MSOSPA error of $\hat{\mathbf{X}}^k$ (without cut-off distance and $p = 2$) is

$$\text{MSOSPA}\left(\hat{\mathbf{X}}^k\right) = \text{E}\left[\frac{1}{t} \min_{i \in \{1,...,t!\}} \sum_{j=1}^{t} b^2\left(\mathbf{x}^k_j, \hat{\mathbf{x}}^k_{\phi_{i,j}}\right)\right]. \quad (7)$$

The MLEC of $\hat{\mathbf{X}}^k$ is given by

$$\text{MLEC}\left(\hat{\mathbf{X}}^k\right) = \frac{\alpha^2}{t} \sum_{i=2}^{t!} N_i P^k_{\hat{\mathbf{X}}^k}(i) \quad (8)$$

$$N_i = \sum_{j=1}^{t} \overline{\delta}\left[j - \phi_{i,j}\right] \quad (9)$$

where N_i is the number of targets with wrong labelling in labelling vector ϕ_i.

In most applications it may be expected that localising targets will have higher priority than labelling. In such cases the selected labelling cost α should be small so that localisation errors are penalised more heavily than labelling errors. Then, the MSLOSPA estimator corresponds to the minimum MSOSPA estimator with the labelling that minimises the MLEC. In other words, under the approximation of small α, we can minimise the MSLOSPA estimator based on the minimium MSOSPA estimator with the labelling that minimises the MLEC. Given a multitarget state set estimate $\hat{X}^k$, all its possible labellings $\Gamma_{\phi_i}\left(\hat{\mathbf{X}}^k\right)$ $i = 1, ..., t!$ have the same MSOSPA error, as MSOSPA is not affected by labelling. Therefore, Equation (6) is important because the labelled multitarget state estimate $\Gamma_{\phi_{j^\star}}\left(\hat{\mathbf{X}}^k\right)$ with lowest MSLOSPA error, for small α, is the one with lowest MLEC:

$$j^\star = \underset{j \in \{1,...,t!\}}{\arg\min} \sum_{i=2}^{t!} N_i P^k_{\Gamma_{\phi_j}(\hat{\mathbf{X}}^k)}(i) \quad (10)$$

which is a linear combination of the optimal labelling probabilities. It should be noted that if α is not small, (10) still minimises an upper bound on the MSLOSPA error.
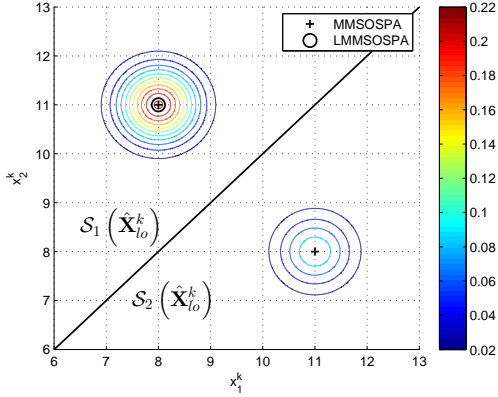
Figure 1: Illustrative example of target labelling and estimation. We show the contour plot of the posterior PDF, the unlabelled MMSOSPA estimates and LMMSOSPA estimate $\hat{\mathbf{X}}_{lo}^k$. We also show regions $\mathcal{S}_1\left(\hat{\mathbf{X}}_{lo}^k\right)$ and $\mathcal{S}_2\left(\hat{\mathbf{X}}_{lo}^k\right)$.

Illustrative example: Let us assume the posterior PDF at time k of two one-dimensional targets is

$$\pi^k\left(\mathbf{X}^k\right) = 0.7\mathcal{N}\left(\mathbf{X}^k; [8,11]^T, 0.25\mathbf{I}_2\right)$$
$$+ 0.3\mathcal{N}\left(\mathbf{X}^k; [11,8]^T, 0.25\mathbf{I}_2\right) \quad (11)$$

where $\mathcal{N}\left(\mathbf{X}^k; \hat{\mathbf{X}}^k, \mathbf{P}^k\right)$ is the Gaussian PDF with mean $\hat{\mathbf{X}}^k$ and covariance matrix $\mathbf{P}^k$ evaluated at $\mathbf{X}^k$. A posterior PDF with this form approximately appears in a tracking scenario with probability of detection one and no clutter after the targets have moved in close proximity and then separated. This situation is chosen for illustrative purposes as the proposed labelling approach is not restricted to such scenarios. The contour plot of the posterior PDF is shown in Figure 1. This figure also includes the MMSOSPA estimate $\{8, 11\}$, labelled MMSOSPA (LMMSOSPA) estimate $\hat{\mathbf{X}}_{lo}^k$, which corresponds to the MMSOSPA estimate with labelling (10), as well as regions $\mathcal{S}_1\left(\hat{\mathbf{X}}_{lo}^k\right)$ and $\mathcal{S}_2\left(\hat{\mathbf{X}}_{lo}^k\right)$.

The MMSOSPA estimate minimises the MSOSPA error. However, this estimate is unlabelled so it must be labelled either with $\hat{\mathbf{X}}_o^k = [8, 11]^T$ or $\Gamma_{\phi_2}\left(\hat{\mathbf{X}}_o^k\right)$ to form tracks. As indicated in Section III-D, the LMMSOSPA provides the best labelling of the MMSOSPA estimate based on minimising the MLEC. For two targets, the labelling that minimises the MLEC comes down to the labelling that has the highest CLP. The CLP of $\hat{\mathbf{X}}_o^k$ is $P_{\hat{\mathbf{X}}_o^k}^k(1)$, which can be calculated integrating (11) in $\mathcal{S}_1\left(\hat{\mathbf{X}}_o^k\right)$, which is the area $x_2^k > x_1^k$, as shown in Figure 1. The CLP for $\Gamma_{\phi_2}\left(\hat{\mathbf{X}}_o^k\right)$ corresponds to the integral of (11) in $\mathcal{S}_1\left(\Gamma_{\phi_2}\left(\hat{\mathbf{X}}_o^k\right)\right) = \mathcal{S}_2\left(\hat{\mathbf{X}}_o^k\right)$, which is the area $x_2^k < x_1^k$. As a result, the CLPs of $\hat{\mathbf{X}}_o^k$ and $\Gamma_{\phi_2}\left(\hat{\mathbf{X}}_o^k\right)$ are 0.7 and 0.3, respectively, so $\hat{\mathbf{X}}_{lo}^k = \hat{\mathbf{X}}_o^k$.

IV. TRACKS THAT MINIMISE SWITCHING

In the previous section, we proposed a way to build tracks based on assigning the best labelling of $\hat{X}^k$ in the MSLOSPA

sense. The aim of this section is to develop a way to form tracks sequentially such that the current multitarget state estimate is optimally linked with the available tracks, i.e., that minimises track switching/jittering. This is done in Section IV-B. In order to get this result, first we indicate how to calculate the probability that a multitrack, i.e., a multitarget track, is optimally linked in Section IV-A.

A. Linked multitrack probability

In this section, we work with the multitrack state estimate $\hat{\mathbf{X}}^{0:k} = \left[\left(\hat{\mathbf{x}}_1^{0:k}\right)^T, \left(\hat{\mathbf{x}}_2^{0:k}\right)^T, ..., \left(\hat{\mathbf{x}}_t^{0:k}\right)^T\right]^T$ where $\hat{\mathbf{x}}_j^{0:k} = \left[\left(\hat{\mathbf{x}}_j^0\right)^T, \left(\hat{\mathbf{x}}_j^1\right)^T, ..., \left(\hat{\mathbf{x}}_j^k\right)^T\right]^T$ is the track estimate for the jth target. We calculate the probability that each track estimate corresponds to the same target, i.e., the probability that the track estimate has the same optimal labelling but it does not matter what labelling. We refer to this probability as linked multitrack probability (LMP) as all the tracks are optimally linked, i.e., they correspond to the same target.

An important difference with the previous section is that optimal labelling probabilities only require knowledge of the (filtering) posterior PDF at the current time while LMPs require the trajectory posterior PDF $\pi^{0:k}(\cdot)$.

Definition 4. Multitrack estimate $\hat{\mathbf{X}}^{0:k}$ has optimal labelling ϕ_i if $\hat{\mathbf{X}}^m$ has optimal labelling ϕ_i for $m \in \{0, ..., k\}$.

In the same way we did in Section III-B, it is useful to define region

$$\mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right) = \left\{\mathbf{X}^{0:k} : d\left(\hat{\mathbf{X}}^m, \Gamma_{\phi_i}\left(\mathbf{X}^m\right)\right)\right.$$
$$\left. < d\left(\hat{\mathbf{X}}^l, \Gamma_{\phi_l}\left(\mathbf{X}^m\right)\right) \ l \neq i, m \in \{0, ..., k\}\right\}$$

that indicates all the true multitrack states $\mathbf{X}^{0:k}$ such that $\hat{\mathbf{X}}^{0:k}$ has optimal labelling ϕ_i. It should be noted that $\mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right) = \mathcal{S}_i\left(\hat{\mathbf{X}}^0\right) \times \mathcal{S}_i\left(\hat{\mathbf{X}}^1\right) \times ... \times \mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)$ where $\times$ denotes the Cartesian product.

Definition 5. Multitrack estimate $\hat{\mathbf{X}}^{0:k}$ is optimally linked if it has an optimal labelling ϕ_i for some $i \in \{1, ..., t!\}$.

If multitrack estimate $\hat{\mathbf{X}}^{0:k}$ has an optimal labelling ϕ_i for some $i \in \{1, ..., t!\}$, all the track estimates $\hat{\mathbf{x}}_1^{0:k}, ..., \hat{\mathbf{x}}_t^{0:k}$ are optimally linked as they have the same optimal label at every time step. Using Definition 5, the probability that $\hat{\mathbf{X}}^{0:k}$ is (optimally) linked, i.e., the LMP, is the probability that $\mathbf{X}^{0:k} \in \bigcup_{i=1}^{t!} \mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right)$. As these regions are disjoint, the LMP becomes

$$P_{\hat{\mathbf{X}}^{0:k}}^{0:k} = \sum_{i=1}^{t!} \int_{\mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right)} \pi^{0:k}\left(\mathbf{X}^{0:k}\right) d\mathbf{X}^{0:k}. \quad (12)$$

In the rest of the paper, unlabelled RFS densities are used. Their argument is a set and we use the symbol ˇ to denote them. It is proved in Appendix D that (12) can be written as

$$P_{\hat{\mathbf{X}}^{0:k}}^{0:k} = \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}^{0:k}\right)} \check{\pi}^{0:k}\left(\left\{\mathbf{x}_1^{0:k}, \mathbf{x}_2^{0:k}, ..., \mathbf{x}_t^{0:k}\right\}\right) d\mathbf{X}^{0:k} \quad (13)$$

where the RFS density over the trajectories is [26]

$$\tilde{\pi}^{0:k}\left(\left\{\mathbf{x}_1^{0:k}, \mathbf{x}_2^{0:k}, ..., \mathbf{x}_t^{0:k}\right\}\right) = \sum_{i=1}^{t!} \pi^{0:k}\left(\Gamma_{\phi_i}\left(\mathbf{X}^{0:k}\right)\right). \quad (14)$$

The most interesting result of LMPs is that they can be calculated using the RFS posterior density over the trajectories. We should bear in mind that computing or approximating the RFS posterior density over the trajectories is quite different to approximating the filtering RFS density, which is the usual approach in the RFS framework. In practice, approximating an RFS density can imply important advantages with respect to approximating a vector valued density. This issue is addressed in Section V-C. It should be noted that we can obtain the linked multitrack probability of a part of the track estimate $\hat{\mathbf{X}}^{m:n}$, where $0 \leq m$ and $m < n \leq k$, by substituting m for 0 and n for k in (12) and (13).

We also want to highlight that Definition 5 and the LMPs of Equation (13) enable us to approach the track building problem in a batch fashion based on the RFS density over the trajectories. Even though this method is not the main purpose of the paper as the paper mainly deals with sequential ways of track building, we think it is insightful and is worth mentioning. As in Section II, let us assume we have the collection of multitarget state estimates $\hat{X}^k = \left\{\hat{\mathbf{x}}_1^k, \hat{\mathbf{x}}_2^k, ..., \hat{\mathbf{x}}_t^k\right\}$ $k = 0, 1, 2, ...$ The estimate at time 0 can be labelled arbitrarily. Then, we can in principle evaluate the LMPs using (13) for all the possible labellings of $\hat{X}^k = \left\{\hat{\mathbf{x}}_1^k, \hat{\mathbf{x}}_2^k, ..., \hat{\mathbf{x}}_t^k\right\}$ $k = 1, 2, ...$ The labelling with highest LMP represents the multitrack with highest probability of being optimally linked.

Illustrative example: Here we extend the example analysed in Section III-D to illustrate the concept of linked multitrack probability. The posterior PDF at time k is given by (11). We assume that the transition density is

$$f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^k\right) = \mathcal{N}\left(x_1^{k+1}; x_1^k, \sigma^2\right)\mathcal{N}\left(x_2^{k+1}; x_2^k, \sigma^2\right). \quad (15)$$

We also assume that at time $k+1$, the measurement is quite uninformative, such that the posterior PDF at time $k+1$ can be approximated as the prior at time $k+1$. Then, the posterior PDF of the trajectory in the interval $k : k+1$ is

$$\pi^{k:k+1}\left(\mathbf{X}^{k:k+1}\right) = \mathcal{N}\left(x_1^{k+1}; x_1^k, \sigma^2\right)\mathcal{N}\left(x_2^{k+1}; x_2^k, \sigma^2\right)\pi^k\left(\mathbf{X}^k\right)$$
$$= 0.7\mathcal{N}\left(\mathbf{x}_1^{k:k+1}; \overline{\mathbf{x}}_1^{k:k+1}, \mathbf{P}^{k+1:k}\right)$$
$$\cdot \mathcal{N}\left(\mathbf{x}_2^{k:k+1}; \overline{\mathbf{x}}_2^{k:k+1}, \mathbf{P}^{k+1:k}\right)$$
$$+ 0.3\mathcal{N}\left(\mathbf{x}_1^{k:k+1}; \overline{\mathbf{x}}_2^{k:k+1}, \mathbf{P}^{k+1:k}\right)$$
$$\cdot \mathcal{N}\left(\mathbf{x}_2^{k:k+1}; \overline{\mathbf{x}}_1^{k:k+1}, \mathbf{P}^{k+1:k}\right)$$

where $\mathbf{x}_i^{k:k+1} = \left[x_i^k, x_i^{k+1}\right]^T$, $\overline{\mathbf{x}}_1^{k:k+1} = [8, 8]^T$, $\overline{\mathbf{x}}_2^{k:k+1} = [11, 11]^T$ and

$$\mathbf{P}^{k+1:k} = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 + \sigma^2 \end{bmatrix}. \quad (16)$$

For simplicity, we analyse the case where $\sigma^2 \to 0$. The multitrack state estimate is $\hat{\mathbf{X}}^{k:k+1} = [8, 8, 11, 11]^T$, i.e., we estimate that target 1 is in position 8 at time k and $k+1$ and target 2 is in position 11 at time k and $k+1$. Note that this is equivalent to using the LMMSOSPA estimator at time k and time $k+1$. Using (12), the LMP is

$$P_{\hat{\mathbf{X}}^{k:k+1}}^{k:k+1} = \int_{\mathcal{L}_1(\hat{\mathbf{X}}^{k:k+1})} \pi^{k:k+1}\left(\mathbf{X}^{k:k+1}\right) d\mathbf{X}^{k:k+1}$$
$$+ \int_{\mathcal{L}_2(\hat{\mathbf{X}}^{k:k+1})} \pi^{k:k+1}\left(\mathbf{X}^{k:k+1}\right) d\mathbf{X}^{k:k+1}$$
$$= \int_{x_1^k < x_2^k, x_1^{k+1} < x_2^{k+1}} \pi^{k:k+1}\left(\mathbf{X}^{k:k+1}\right) d\mathbf{X}^{k:k+1}$$
$$+ \int_{x_1^k > x_2^k, x_1^{k+1} > x_2^{k+1}} \pi^{k:k+1}\left(\mathbf{X}^{k:k+1}\right) d\mathbf{X}^{k:k+1}$$
$$\approx 0.7 + 0.3$$
$$= 1. \quad (17)$$

As indicated in Section III-D, the optimal labelling probabilities of the estimate at time k are 0.7 and 0.3 and they remain unchanged at time $k + 1$. This indicates that we do not know with high certainty which target estimate corresponds to each target at both time steps. Nevertheless, the LMP (17) indicates that the track estimate $\hat{\mathbf{X}}^{k:k+1}$ is optimally linked with a probability that is approximately 1, i.e., the estimate $[8, 8]^T$ belongs to a target and the estimate $[11, 11]^T$ to another target. This information cannot be obtained just from the posterior $\pi^k\left(\cdot\right)$ at time k, we need the trajectory posterior $\pi^{k:k+1}\left(\cdot\right)$ or the RFS trajectory posterior, see (13).

B. Sequential linking of multitarget state estimates

In this section, we study the problem of optimally linking multitarget state estimates sequentially. This means that the current multitarget estimate is optimally linked with the previous tracks so track switching is minimised. We use the following assumption

- A1:

$$\ell\left(\mathbf{z}^k \middle| \mathbf{X}^k\right) = \ell\left(\mathbf{z}^k \middle| \Gamma_{\phi_i}\left(\mathbf{X}^k\right)\right) \quad i \in \{1, ..., t!\} \quad (18)$$

$$f\left(\mathbf{X}^k \middle| \mathbf{X}^{k-1}\right) = \prod_{i=1}^t g\left(\mathbf{x}_i^k \middle| \mathbf{x}_i^{k-1}\right) \quad (19)$$

where $g\left(\cdot \middle| \cdot\right)$ is the transition density for the individual targets.

Assumption A1 implies that measurements do not provide any information about labelling and targets move independently with the same dynamic model. These are usual assumptions in multiple target tracking [3].

We assume that at time k we know the posterior PDF $\pi^k\left(\cdot\right)$ and have an estimate $\hat{\mathbf{X}}^k$ of the state at time k. We also assume we have a set estimate $\Gamma_{\phi_i}\left(\hat{\mathbf{X}}^{k+1}\right)$ $i \in \{1, ..., t!\}$ of the current multitarget state. The optimal sequential linking is performed by selecting the labelling $\check{\phi}$ of $\hat{\mathbf{X}}^k$ that maximises the linked multitrack probability of

$$\hat{\mathbf{X}}_i^{k:k+1} = \left[\left(\hat{\mathbf{X}}^k\right)^T, \left(\Gamma_{\phi_i}\left(\hat{\mathbf{X}}^{k+1}\right)\right)^T\right]^T.$$

which is abbreviated as $P_i^{k:k+1}$. That is,

$$\check{\phi} = \phi_{i^\star} : i^\star = \underset{i \in \{1, ..., t!\}}{\arg\max} P_i^{k:k+1} \quad (20)$$

where

$$P_i^{k:k+1} = \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}_i^{k:k+1}\right)} \check{\pi}^{k:k+1}\left(\left\{\mathbf{x}_1^{k:k+1}, ..., \mathbf{x}_t^{k:k+1}\right\}\right) d\mathbf{X}^{k:k+1} \tag{21}$$

and we have used (13). We refer to $P_i^{k:k+1}$ as sequential linked multitrack probability (SLMP).

It is shown in Appendix E that, under Assumption A1, (21) can be written as

$$\begin{aligned} P_i^{k:k+1} = 1/\rho \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}_i^{k:k+1}\right)} \ell\left(\mathbf{z}^{k+1}\,\middle|\,\mathbf{X}^{k+1}\right) \\ f\left(\mathbf{X}^{k+1}\,\middle|\,\mathbf{X}^k\right) \check{\pi}^k\left(\left\{\mathbf{x}_1^k, ..., \mathbf{x}_t^k\right\}\right) d\mathbf{X}^{k:k+1} \end{aligned} \tag{22}$$

where the normalising constant is given by

$$\begin{aligned} \rho = \frac{1}{t!} \int \ell\left(\mathbf{z}^{k+1}\,\middle|\,\mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1}\,\middle|\,\mathbf{X}^k\right) \\ \check{\pi}^k\left(\left\{\mathbf{x}_1^k, ..., \mathbf{x}_t^k\right\}\right) d\mathbf{X}^{k:k+1}. \end{aligned} \tag{23}$$

The main conclusion of (22) is that, given an estimate at time k and a set estimate at time $k+1$, the labelling of the set estimate at time $k+1$ that maximises the SLMP (minimises track jittering) only depends on the likelihood, the transition density and the RFS posterior density at time k. This implies that the usual RFS formulation, which does not include labels, can also be used to optimally link multitarget state estimates sequentially from the Bayesian point of view. However, this optimal linking has to be done sequentially and we cannot obtain the optimal labelling probabilities. Using RFS densities instead of vector densities implies several benefits that are explained in Section V-C.

LMP approximation: If we are performing Bayesian filtering, we cannot calculate the LMP $P_{\hat{\mathbf{X}}^{k_1:k_2}}^{k_1:k_2}$, which indicates the probability that the partial track estimate $\hat{\mathbf{X}}^{k_1:k_2}$ is optimally linked, for $k_2 > k_1 + 1$ because $P_{\hat{\mathbf{X}}^{k_1:k_2}}^{k_1:k_2}$ requires the posterior PDF of the trajectory. This is not available in Bayesian filtering as Bayesian filtering is only interested in the posterior PDF at the current time step. In fact, if we calculated the posterior PDF of the whole trajectory, it would also make sense to re-estimate the target trajectory $\hat{\mathbf{X}}^{k_1:k_2}$ to take into account the current value of the measurement to update the trajectory. This is Bayesian smoothing rather than filtering.

We suggest using

$$\mathcal{P}_{\hat{\mathbf{X}}^{k_1:k_2}}(k_1, k_2) = \prod_{j=k_1}^{k_2-1} P_{\hat{\mathbf{X}}^{j:j+1}}^{j:j+1} \tag{24}$$

as an approximation to $P_{\hat{\mathbf{X}}^{k_1:k_2}}^{k_1:k_2}$ in a filtering set-up. The benefits of calculating $\mathcal{P}_{\hat{\mathbf{X}}^{k_1:k_2}}(k_1, k_2)$ instead of $P_{\hat{\mathbf{X}}^{k_1:k_2}}^{k_1:k_2}$ are:

- Once we have estimated the multitarget state at a time step between k_1 and k_2, we do not have to re-estimate them once a new measurement is received.
- They are reasonably easy to compute while the computational complexity of calculating $P_{\hat{\mathbf{X}}^{k_1:k_2}}^{k_1:k_2}$ grows exponentially if k_2 increases.
- We can use the RFS framework without labels.

V. COMPARISON BETWEEN BOTH APPROACHES TO TRACK BUILDING

In this section, we summarise some of the results obtained before and indicate the main benefits and drawbacks of calculating the optimal labelling probabilities and SLMPs. In this comparison, we use Assumption A1. We also sketch the benefits of using RFS densities in Section V-C.

A. Tracks that minimise MSLOSPA error

The objective is to label the multitarget state estimate at the current time step to minimise the MSLOSPA error (for small α). Therefore, we optimally link the current multitarget state estimate at the current time step with the true labelled multitarget state. Selecting the labelling of the estimate that minimises the MLEC involves the calculation of $t!$ integrals of a PDF with a state of dimension tn_x.

- The main benefit of computing the optimal labelling probabilities is
  - They indicate the probability that a multitarget labelled state estimate corresponds to a given multitarget labelling. This is important in some scenarios [15].
- The drawbacks are
  - In practice, once the targets have been in close proximity for a sufficiently long time and then separated, the optimal labelling probabilities become even and do not change irrespective of what the targets do. This is due to the fact that the posterior PDF finally becomes permutation invariant [16]. Therefore, we already know that they will not convey useful information in the future. In addition, we cannot build tracks based on them in a rigorous way as any labelling of the target state estimates is optimal.
  - We need to keep a multimodal posterior PDF approximation with up to $t!$ modes due to target state permutations after t targets have been in close proximity and separated. This is computationally expensive, especially, for particle filter implementations [15].

B. Tracks that minimise switching

The objective is to label the multitarget state estimate at the current time step to maximise the SLMP. Therefore, we optimally link the multitarget state estimate at the current time step with the labelled multitarget state estimate at the previous time step. In other words, we minimise track switching. Selecting the labelling of the estimate that minimises the SLMP involves the calculation of $t!$ integrals of a PDF with a state of dimension $2tn_x$.

- The main benefits of computing the SLMPs are
  - They are always useful and carry important information, even if the targets have been in close proximity for a long time and then separated. We can always create tracks based on them and estimate the probability that a part of a track or a whole track is (optimally) linked. This implies that they have the ability to tell us about track confusion from a new

reference time. This is particularly useful if we want to know if there is a new confusion in the tracks after a target crossing, i.e., we do not care about what happened to the tracks in the past and we want to know if track confusion will appear again in the future.

- We only need the RFS posterior density, which implies some advantages as indicated in Section V-C.

- The drawback is
  - They require the calculation of an integral over the posterior of the multitrack state at time steps k and $k+1$. The dimension $2tn_x$ of this integral is twice the dimension tn_x of the integrals usually calculated in filtering, which are done with respect to the state at the current time.

C. Benefits of using RFS densities

We have proved that we can build tracks sequentially using the RFS posterior density and that we can calculate LMPs using the RFS posterior density over the trajectories using (13). These findings can have a major impact on how tracking algorithms are designed. Based on the results of this paper, tracking algorithms can provide tracks based on solid theoretical grounds while exploiting the benefits of approximating RFS densities rather than vector-valued densities (either for the set of target states at time k or the set of target trajectories). The benefits are sketched in the following but we refer the reader to [26].

A vector density $\nu^k\left(\cdot\right)$ belongs to the RFS family of $\check{\pi}^k\left(\{\cdot\}\right)$ if [26]

$$\check{\pi}^k\left(\{\mathbf{x}_1^k, ..., \mathbf{x}_t^k\}\right) = \sum_{i=1}^{t!} \nu^k\left(\Gamma_{\phi_i}\left(\mathbf{X}^k\right)\right)$$

Any vector density that belongs to an RFS family contains the same information as the RFS density. Therefore, the RFS density can be approximated by any $\nu^k\left(\cdot\right)$ that belongs to its family. Under Assumption A1, if we use $\nu^k\left(\cdot\right)$ instead of $\pi^k\left(\cdot\right)$ in the prediction and update steps, the resulting posterior $\nu^{k+1}\left(\cdot\right)$ belongs to the RFS family of $\check{\pi}^{k+1}\left(\{\cdot\}\right)$ [26, Proposition 3]. Importantly, some $\nu^k\left(\cdot\right)$ are more convenient to approximate than others, e.g., they are less multimodal. Therefore, at every time step, we can choose the $\nu^k\left(\cdot\right)$ within the RFS family that suits us. This can have an important effect on performance. An example for Gaussian approximations is provided in [26] but it can be generalised for any type of approximation.

VI. NUMERICAL EXAMPLE

In this section, we illustrate the concepts of CLP and SLMP and the building of tracks based on them in a two-target tracking filtering case. In Section VI-A, we compare different track building procedures when two-targets get in close proximity. In Section VI-B, we analyse the effect of changing the process and measurement noise on the CLP and SLMP.

The state of the jth target at time k is $\mathbf{x}_j^k = \left[x_j^k, \dot{x}_j^k, y_j^k, \dot{y}_j^k\right]^T$ where $\left[x_j^k, y_j^k\right]^T$ is the position vector and $\left[\dot{x}_j^k, \dot{y}_j^k\right]^T$ is the velocity vector. The targets follow a nearly constant velocity model [27]:

$$g\left(\mathbf{x}_j^{k+1} \mid \mathbf{x}_j^k\right) = \mathcal{N}\left(\mathbf{x}_j^{k+1}; \mathbf{F}\mathbf{x}_j^k, \mathbf{Q}\right) \tag{25}$$

$$\mathbf{F} = \mathbf{I}_2 \otimes \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \tag{26}$$

$$\mathbf{Q} = \sigma_u^2 \mathbf{I}_2 \otimes \begin{pmatrix} \tau^3/3 & \tau^2/2 \\ \tau^2/2 & \tau \end{pmatrix} \tag{27}$$

where $j \in \{1, 2\}$, $\mathcal{N}\left(\mathbf{x}; \overline{\mathbf{x}}, \mathbf{Q}\right)$ is the Gaussian PDF evaluated at $\mathbf{x}$ with mean $\overline{\mathbf{x}}$ and covariance matrix $\mathbf{Q}$, τ is the sampling period, σ_u^2 is the continuous-time process noise intensity and $\otimes$ is the Kronecker product.

The PDF of the measurement given the state is

$$\ell\left(\mathbf{z}^k \mid \mathbf{X}^k\right) = \frac{1}{2}\mathcal{N}\left(\mathbf{z}^k; \mathbf{H}_1\mathbf{X}^k, \mathbf{R}\right) + \frac{1}{2}\mathcal{N}\left(\mathbf{z}^k; \mathbf{H}_2\mathbf{X}^k, \mathbf{R}\right) \tag{28}$$

where $\mathbf{R}$ is the covariance matrix of the measurement noise,

$$\mathbf{H}_1 = \mathbf{I}_2 \otimes \mathbf{L}$$

$$\mathbf{H}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \mathbf{L}$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

That is, we have position measurements of the targets without knowing the measurement-to-target association, there are no false alarms and the probability of detection is one. This is equivalent to measuring the set $\left\{\mathbf{z}_1^k, \mathbf{z}_2^k\right\}$ where $\mathbf{z}_j^k$ is the position measurement of one of the targets.

A. Comparison between different track building procedures

We evaluate the CLPs and the SLMPs for three different track building procedures. Estimator 1 (E1) selects the mean of the component with highest weight in the posterior Gaussian mixture. It should be noted that E1 is a widely used track building procedure in MHT, i.e., it is the posterior mean conditional on the most likely hypothesis [1]. Estimator 2 (E2) selects the labelling of E1 that maximises the CLP (minimises MLEC) as indicated in Section III. Estimator 3 (E3) selects the labelling of E1 that maximises the SLMP as indicated in Section IV. E1 does not take into account the optimal track building properties derived in this paper but E2 and E3 do. Therefore, the tracks provided by E2 and E3 are better than the tracks of E1.

The scenario has 79 time steps with parameters $\sigma_u = 1.8\mathrm{m/s}^{3/2}$, $\tau = 0.5$ s, $\mathbf{R} = \sigma_r^2\mathbf{I}_2$, $\sigma_r^2 = 1\,\mathrm{m}^2$, and the target trajectories shown in Figure 2. The trajectory of target one is fixed but the trajectory of target two is moved depending on a parameter called d_x. When $d_x = 0$ m, both target states from time step 30 to 50 are exactly the same but they differ before time step 30 and after time step 50. We analyse the cases $d_x = 0$ m and $d_x = 6$ m. The prior PDF at time 0 is the one in [15]. The posterior PDF is a Gaussian mixture, which can be computed analytically in principle. However, the
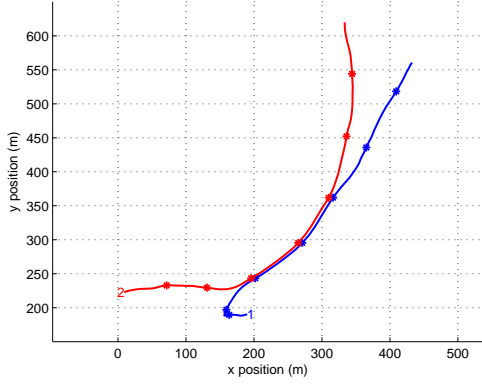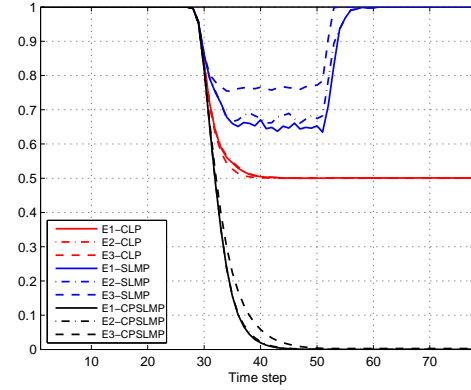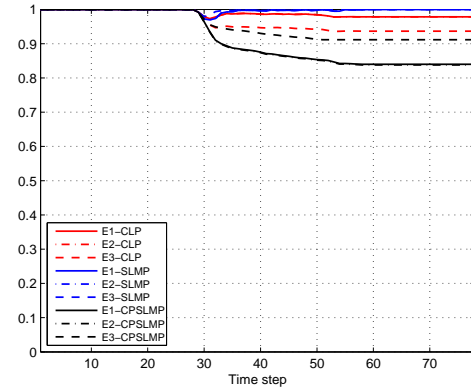
(a)
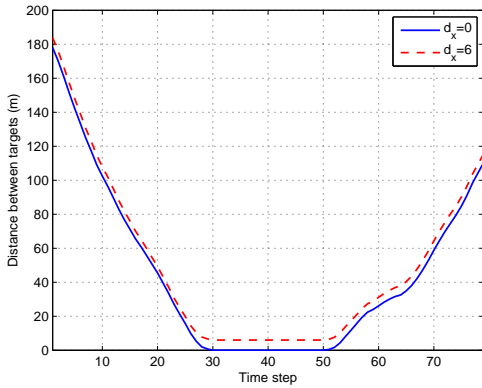
(b)

Figure 2: Scenario of the simulations: (a) Target trajectories for $d_x = 6$ m. Each target is identified at the beginning of its trajectory by a label. The trajectories have an asterisk to indicate the target position every ten time steps. (b) Distance between targets against time for $d_x = 0$ m and $d_x = 6$ m.


(a)

(b)

Figure 3: Correct labelling probabilities and sequential linked multitrack probabilities against time a) $d_x = 0$ m b) $d_x = 6$ m. In the legend, CPSLMP stands for the cumulative product of SLMP.

posterior PDF has an ever-increasing number of components so we use the joining algorithm in [28] to control the number of components of the mixture. With this algorithm, before the targets get close to each other, the average number of posterior mixture components is one for $d_x = 0$ m and $d_x = 6$ m. After they separate, it is two.

The averaged CLP and SLMP against time for $d_x = 0$ m and $d_x = 6$ m using 1000 Monte Carlo runs are shown in Figure 3. The CLP can be obtained in closed-form and the SLMP is approximated using Monte Carlo integration with 20000 samples. For $d_x = 0$ m, the CLP at the end of the simulation is approximately 0.5 for all the estimators in all Monte Carlo runs. This means that we do not know which estimate corresponds to target 1 and target 2. It should be noted that once the CLP is 0.5, it does not convey any new information. However, the SLMP is always meaningful. As of time step 60, the CLP remains unchanged but the SLMP indicates that the multitarget estimates at consecutive times are (optimally) linked with probability 1. The interpretation is that the targets are now separated enough such that the tracks after this time step belong to the same target with probability one although the association with the tracks before

the target crossing is unknown. E3 has a higher probability that two consecutive estimates are linked. This is because it maximises the SLMP. For $d_x = 6$ m, the CLP at the end of the simulation is different for the estimators as the posterior is not permutation invariant. E1 and E2 have the highest probability of correct labelling.

In general, if we are interested in determining where target 1 and target 2 are, we should use E2. However, if we are more interested in building (optimally) connected tracks sequentially with minimum jitter, it is better to use E3. This is illustrated in Figure 4 where the resulting tracks of the estimators for four time steps are plotted for an exemplar run ($d_x = 0$ m). As E3 maximises the SLMP, the resulting four-time-step tracks are more reasonable. E1 and E2 show a jump in the linking of target state estimates. This behaviour happens often as the averaged cumulative product of the SLMPs over the time steps $51, ..., 54$ considered in Figure 4 are, using the results shown in Figure 3 (a): 0.37 for E1, 0.48 for E2 and 0.71 for E3. These results indicate that E3 provides the best tracks, in the sense of minimising switches, followed by E2 and then E1.

B. Process and measurement noise analysis

In this section, we analyse how the CLP and SLMP for the three estimators change depending on the process and
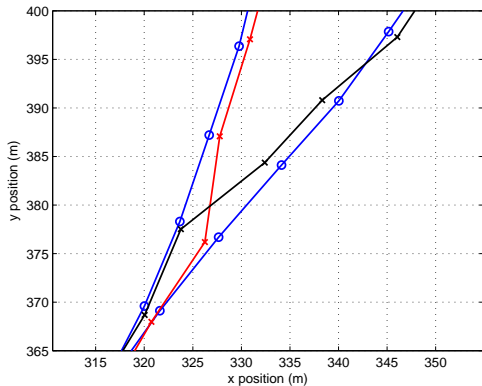
(a)
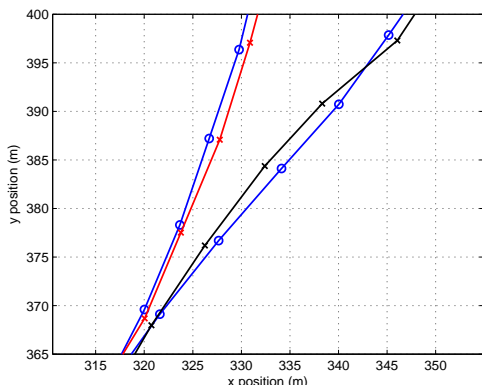
(b)

Figure 4: Track formation for (a) E1 and E2, which have the same output in this exemplar run, (b) E3. The blue circles represent the true target states at time steps $k \in \{51, 52, 53, 54\}$. The true target states that belong to the same target are linked by a blue line. The red crosses indicate the target state estimate for a target. The black crosses indicate the target state estimate for the other target. The red and black lines indicate how the target state estimates are linked to form tracks. E3, which maximises the SLMP, provides more reasonable tracks than E1 and E2.

measurement noise parameters. The prior PDF at time 0 is

$$\pi^0 \left(\mathbf{X}^0 \right) = \mathcal{N} \left(\mathbf{x}_1^0; \overline{\mathbf{x}}_1^0, \mathbf{P}^0 \right) \mathcal{N} \left(\mathbf{x}_2^0; \overline{\mathbf{x}}_2^0, \mathbf{P}^0 \right)$$

where $\overline{\mathbf{x}}_1^0 = [0, 4, 0, -1]^T$, $\overline{\mathbf{x}}_2^0 = [0, 4, -5, 1]^T$ and $\mathbf{P}^0$ is a diagonal matrix whose entries are 0.1 with international system units. The number of time steps in the trajectory is 18 and $\tau = 0.5$ s. In order to provide a more general analysis than in the previous case that only focused on a particular trajectory, we generate a new multitarget trajectory in each Monte Carlo run.

First, we set $\sigma_r = 1$ m and evaluate the CLP and SLMP for $\sigma_u \in \{0.1, 1, 10\} \, \left(\text{m/s}^{3/2} \right)$. To indicate what the trajectories generated from this dynamic model look like, we plot the trajectories used for $\sigma_u = 0.1 \, \text{m/s}^{3/2}$ in Figure 5. The CLP and SLMP against time averaged over all Monte Carlo runs are shown in Figure 6. For $\sigma_u = 0.1 \, \text{m/s}^{3/2}$, CLP and SLMP are nearly one at all time steps. However, as σ_u increases, CLP decreases. This implies that it gets more and more difficult to label the multitarget state estimates properly as
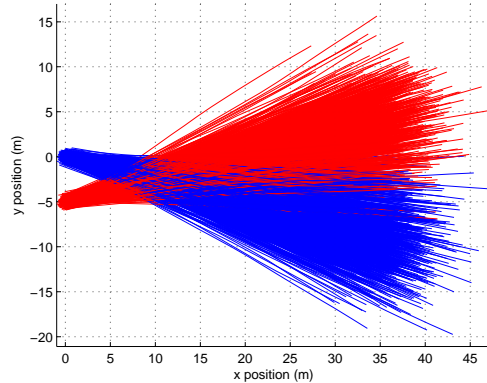

Figure 5: Target trajectories in each Monte Carlo run for $\sigma_u = 0.1 \, \text{m/s}^{3/2}$. Target one trajectories are shown in blue and target two trajectories in red.

the process noise increases. The SLMP gets also lower as σ_u increases and the minimum of the SLMP is attained earlier for $\sigma_u = 10 \, \text{m/s}^{3/2}$ than $\sigma_u = 1 \, \text{m/s}^{3/2}$. This is due to the fact that with higher process noise, the targets are likely to get closer together earlier. It should be noted that SLMP for all σ_u and all estimators is nearly one as of time step 10. This means that the tracks formed as from this time step belong to the same targets with probability close to one. We want to remark that this information is not contained in the posterior density $\pi^k \left(\cdot \right)$ but in $\pi^{k:k+1} \left(\cdot \right)$

Now, we set $\sigma_u = 1 \, \text{m/s}^{3/2}$ and evaluate the CLP and SLMP for $\sigma_r \in \{0.5, 1, 2\}$ (m). The CLP and SLMP against time averaged over all Monte Carlo runs are shown in Figure 7. For $\sigma_r = 0.5$ m, CLP and SLMP are approximately one at all time steps, which indicate that the built tracks belong to the same targets with probability approximately one. As happened with the process noise, performance deteriorates as σ_r increases. This is because the sensor has less capability to separate the states of the two targets so they get mixed up more easily. As expected, the SLMP is always higher for E3 than for E1 and E2 in all simulations.

VII. CONCLUSIONS

In this paper, we have presented an analysis on how we can build optimal tracks from the Bayesian point of view. We have proposed two alternatives. The first method takes the customary approach of labelling the states and finding the best labelling of the multitarget estimate in the MSLOSPA sense (with small α). The second method does not require the labelling of the state and builds tracks by labelling the current multitarget state estimate to maximise the probability that is optimally linked with the multitarget state estimate at the previous time step, i.e., it minimises track jittering. The first approach shows some important drawbacks after the targets have moved in close proximity for a long time and then separated. The second approach does not have this drawback and only requires knowledge of the RFS posterior density.

The results of this paper can have wide applications. First, we can in principle modify commonly used (unlabelled) RFS algorithms, which currently do not build tracks, such that they
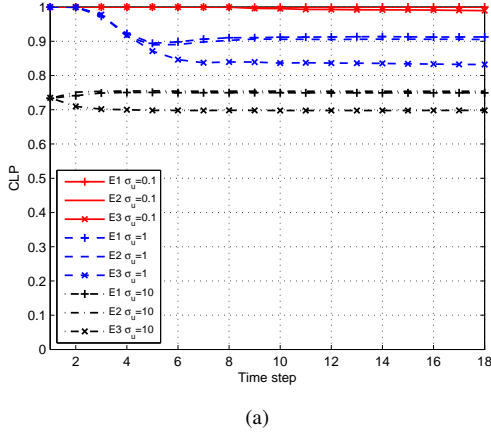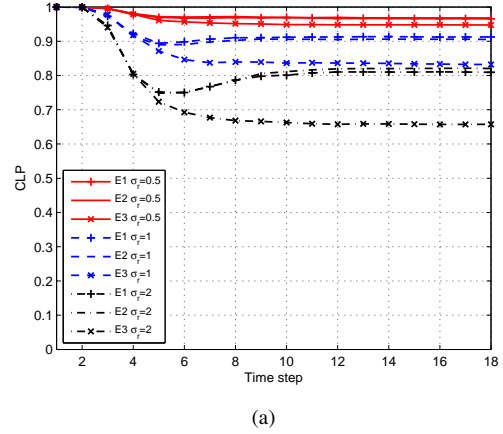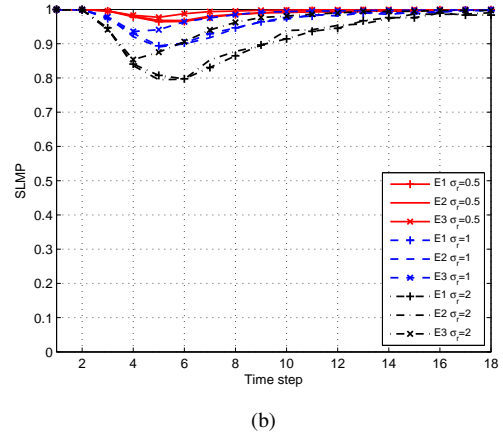
(a)
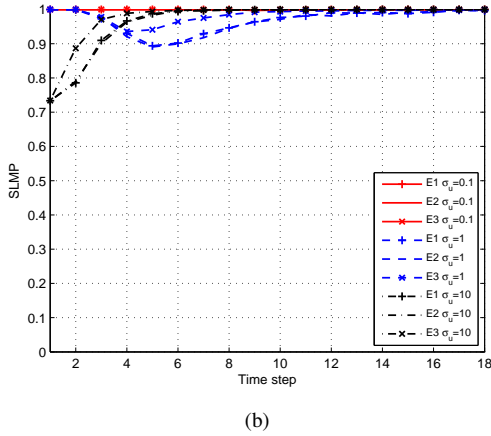
(b)

Figure 6: CLP and SLMP against time using $\sigma_r = 1\,\text{m}$. As σ_u increases, performance deteriorates.

(a)

(b)

Figure 7: CLP and SLMP against time using $\sigma_u = 1\,\text{m/s}^{3/2}$. As σ_r increases, performance deteriorates.

build tracks in a sound, well-defined fashion. Second, we can in principle modify classic algorithms like MHT such that they provide jitter-free tracks based on first principles. In general, as explained in the paper, multiple target tracking algorithms should consider exploiting the advantage of only requiring the RFS posterior density or the RFS trajectory posterior in the smoothing problem to build tracks.

An interesting line of future research is to generalise this analysis to a variable and unknown number of targets.

APPENDIX A

In this appendix, we prove why Definition 1 denotes the optimal labelling of a target. To do so, we use the random set notation as it explicitly includes the labels. The true labelled multitarget state set is

$$X^k = \left\{ \left[\left(\mathbf{x}_1^k\right)^T, 1\right]^T, \left[\left(\mathbf{x}_2^k\right)^T, 2\right]^T, ..., \left[\left(\mathbf{x}_t^k\right)^T, t\right]^T \right\} \tag{29}$$

where we have assumed that the true labels are $[1,...,t]^T$ without loss of generality.

A labelled multitarget state set estimate is

$$\hat{X}^k = \left\{ \left[\left(\hat{\mathbf{x}}_1^k\right)^T, \hat{l}_1\right]^T, \left[\left(\hat{\mathbf{x}}_2^k\right)^T, \hat{l}_2\right]^T, ..., \left[\left(\hat{\mathbf{x}}_t^k\right)^T, \hat{l}_t\right]^T \right\} \tag{30}$$

where the label vector estimate $\left[\hat{l}_1, \hat{l}_2, ...\hat{l}_t\right]^T$ can be any permutation of $[1, ..., t]^T$.

Definition 6. The optimal labelling $\tilde{\boldsymbol{\phi}}' = \left[\tilde{\phi}'_1, ..., \tilde{\phi}'_t\right]^T$ of the labelled multitarget state set estimate is the label vector that minimises the LOSPA distance between (29) and (30):

$$\tilde{\boldsymbol{\phi}}' = \boldsymbol{\phi}_{l^\star} : l^\star$$

$$= \underset{l \in \{1,...,t!\}}{\arg\min}\, d\left(X^k, \left\{ \left[\left(\hat{\mathbf{x}}_1^k\right)^T, \phi_{l,1}\right]^T, ..., \left[\left(\hat{\mathbf{x}}_t^k\right)^T, \phi_{l,t}\right]^T \right\} \right)$$

$$= \underset{l \in \{1,...,t!\}}{\arg\min}\, \underset{i \in \{1,...,t!\}}{\min} \left[\sum_{j=1}^{t} b^p\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{i,j}}^k\right) \right.$$

$$\left. + \alpha^p \bar{\delta}\left[j - \Gamma_{\phi_i}\left(\phi_l\right)_j\right] \right] \tag{31}$$

where $\Gamma_{\phi_i}\left(\phi_l\right)_j$ is the jth component of vector $\Gamma_{\phi_i}\left(\phi_l\right)$.

In the following we prove that Definitions 1 and 6 are equivalent, i.e., $\tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\phi}}'$. We make the change of indices (i is changed by n) such that

$$\phi_n = \Gamma_{\phi_i}\left(\phi_l\right) \tag{32}$$

which implies that

$$\phi_i = \Gamma_{\phi_n}\left(\phi_l^{-1}\right) \tag{33}$$

where ϕ_l^{-1} is the vector that indicates the inverse permutation of ϕ_l. Then, (31) becomes

$$
\begin{aligned}
l^\star &= \arg\min_{l\in\{1,...,t!\}} \min_{n\in\{1,...,t!\}} \left[\sum_{j=1}^t b^p\left(\mathbf{x}_j^k, \Gamma_{\phi_l^{-1}}\left(\hat{\mathbf{X}}^k\right)_{\phi_{n,j}}\right)\right.\\
&\quad \left. + \alpha^p\overline{\delta}\left[j-\phi_{n,j}\right]\right]\\
&= \arg\min_{l\in\{1,...,t!\}} d\left(\mathbf{X}^k, \Gamma_{\phi_l^{-1}}\left(\hat{\mathbf{X}}^k\right)\right)\\
&= \arg\min_{l\in\{1,...,t!\}} d\left(\Gamma_{\phi_l}\left(\mathbf{X}^k\right), \hat{\mathbf{X}}^k\right).
\end{aligned}
$$

We get that $\tilde{\phi} = \tilde{\phi}'$ using (2).

APPENDIX B

In this appendix, we prove Lemma 2. We define vector $\phi^\star = [\phi_1^\star,...,\phi_t^\star]^T$ as

$$
\phi^\star = \phi_{i^\star} : i^\star = \arg\min_{i\in\{1,...,t!\}}\left(\frac{1}{t}\sum_{j=1}^t b^p\left(\hat{\mathbf{x}}_j^k, \mathbf{x}_{\phi_{i,j}}^k\right)\right)^{1/p}. \tag{34}
$$

In the following, we show that $\phi^\star = \tilde{\phi}$ where $\tilde{\phi}$ is the optimal labelling of $\hat{\mathbf{X}}^k$. Using (2), we can write

$$
\begin{aligned}
\tilde{\phi} &= \phi_{l^\star} : l^\star\\
&= \arg\min_{l\in\{1,...,t!\}} d\left(\hat{\mathbf{X}}^k, \Gamma_{\phi_l}\left(\mathbf{X}^k\right)\right)\\
&= \arg\min_{l\in\{1,...,t!\}} \min_{i\in\{1,...,t!\}}\left[\sum_{j=1}^t b^p\left(\hat{\mathbf{x}}_j^k, \Gamma_{\phi_l}\left(\mathbf{X}^k\right)_{\phi_{i,j}}\right)\right.\\
&\quad \left. + \alpha^p\overline{\delta}\left[j-\phi_{i,j}\right]\right] \tag{35}
\end{aligned}
$$

where $\Gamma_{\phi_l}\left(\mathbf{X}^k\right)_{\phi_{i,j}}$ denotes the state vector of target j in the ith permutation of the multitarget state vector $\Gamma_{\phi_l}\left(\mathbf{X}^k\right)$. Let $\phi_n^\star$ $n\in\{1,...,t!\}$ denote the permutations of vector $\phi^\star$, which is given by (34). We can write (35) as

$$
\begin{aligned}
\tilde{\phi} &= \phi_{n^\star}^\star : n^\star\\
&= \arg\min_{n\in\{1,...,t!\}} d\left(\hat{\mathbf{X}}^k, \Gamma_{\phi_n^\star}\left(\mathbf{X}^k\right)\right)\\
&= \arg\min_{n\in\{1,...,t!\}} \min_{i\in\{1,...,t!\}}\left[\sum_{j=1}^t b^p\left(\hat{\mathbf{x}}_j^k, \Gamma_{\phi_n^\star}\left(\mathbf{X}^k\right)_{\phi_{i,j}}\right)\right.\\
&\quad \left. + \alpha^p\overline{\delta}\left[j-\phi_{i,j}\right]\right].
\end{aligned}
$$

The argument of the minimum (in variable i) of the first term isolated is a function of n and is denoted as $i_1^\star(n)$ and the argument of the minimum (in variable i) of the second term isolated is $i_2^\star(n) = 1$ regardless of n. If there exists n^* such that the two arguments of the minima coincide, i.e., $i_1^\star(n^*) = 1$ then, the whole expression is minimised and n^* is obtained. For $n = 1$, the multitarget state vector $\Gamma_{\phi_n^\star}\left(\mathbf{X}^k\right)$ is ordered according to $\phi^\star$ (as $\phi_1^\star = \phi^\star$) and, therefore, the argument of the minimum of the first term is $i^\star(1) = 1$ because of (34). Consequently, $\tilde{\phi} = \phi^\star$.

APPENDIX C

In this appendix, we prove Equation (6). Using (5), we get

$$
\begin{aligned}
&\text{MSLOSPA}\left(\hat{\mathbf{X}}^k\right)\\
&= \mathrm{E}\left[d^2\left(\mathbf{X}^k, \hat{\mathbf{X}}^k\right)\right]\\
&= \int \frac{1}{t}\min_{l\in\{1,...,t!\}}\left[\sum_{j=1}^t b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{l,j}}^k\right) + \alpha^2\overline{\delta}\left[j-\phi_{l,j}\right]\right]\\
&\quad \pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k\\
&= \int_{\mathcal{S}_1(\hat{\mathbf{X}}^k)} \frac{1}{t}\sum_{j=1}^t b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{1,j}}^k\right)\pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k\\
&\quad + \frac{1}{t}\sum_{i=2}^{t!}\int_{\mathcal{S}_i(\hat{\mathbf{X}}^k)}\min_{l\in\{1,...,t!\}}\sum_{j=1}^t\\
&\quad \left[b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{l,j}}^k\right) + \alpha^2\overline{\delta}\left[j-\phi_{l,j}\right]\right]\pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k
\end{aligned}
$$

It is met that

$$
\begin{aligned}
&\min_{l\in\{1,...,t!\}}\sum_{j=1}^t\left[b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{l,j}}^k\right) + \alpha^2\overline{\delta}\left[j-\phi_{l,j}\right]\right]\\
&\leq \sum_{j=1}^t\left[b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{i,j}}^k\right) + \alpha^2\overline{\delta}\left[j-\phi_{i,j}\right]\right]
\end{aligned}
$$

for $\mathbf{X}^k \in \mathcal{S}_i\left(\hat{\mathbf{X}}^k\right)$ and the inequality is tight for $\alpha \to 0$ because of (3). Then,

$$
\begin{aligned}
&\text{MSLOSPA}\left(\hat{\mathbf{X}}^k\right)\\
&\leq \int_{\mathcal{S}_1(\hat{\mathbf{X}}^k)}\frac{1}{t}\sum_{j=1}^t b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{1,j}}^k\right)\pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k\\
&\quad + \frac{1}{t}\sum_{i=2}^{t!}\int_{\mathcal{S}_i(\hat{\mathbf{X}}^k)}\sum_{j=1}^t\left[b^2\left(\mathbf{x}_j^k, \hat{\mathbf{x}}_{\phi_{i,j}}^k\right) + \alpha^2\overline{\delta}\left[j-\phi_{i,j}\right]\right]\\
&\quad \pi^k\left(\mathbf{X}^k\right) d\mathbf{X}^k\\
&= \text{MSOSPA}\left(\hat{\mathbf{X}}^k\right) + \text{MLEC}\left(\hat{\mathbf{X}}^k\right)
\end{aligned}
$$

where $\text{MLEC}\left(\hat{\mathbf{X}}^k\right)$ is the mean labelling error cost (MLEC) of the estimate $\hat{\mathbf{X}}^k$, see (8), and $\text{MSOSPA}\left(\hat{\mathbf{X}}^k\right)$ is the mean square OSPA distance of the estimate $\hat{\mathbf{X}}^k$, see (7).

APPENDIX D

In this appendix, we prove (13). We make the change of variables $\mathbf{Y}^m = \Gamma_{\phi_i}\left(\mathbf{X}^m\right)$ $\left(\mathbf{X}^m = \Gamma_{\phi_i}^{-1}\left(\mathbf{Y}^m\right)\right)$ in each integral in (12) noting that

$$
\mathbf{X}^{0:k} \in \mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right) \longleftrightarrow \mathbf{Y}^{0:k} \in \mathcal{L}_1\left(\hat{\mathbf{X}}^{0:k}\right) \tag{36}
$$

because

$$
\begin{aligned}
&d\left(\hat{\mathbf{X}}^m, \Gamma_{\phi_i}\left(\mathbf{X}^m\right)\right) < d\left(\hat{\mathbf{X}}^l, \Gamma_{\phi_l}\left(\mathbf{X}^m\right)\right) l \neq i, m \in \{0,...,k\}\\
&\longleftrightarrow d\left(\hat{\mathbf{X}}^m, \mathbf{Y}^m\right) < d\left(\hat{\mathbf{X}}^l, \Gamma_{\phi_l}\left(\mathbf{Y}^m\right)\right) l \neq 1, m \in \{0,...,k\}.
\end{aligned}
$$

Then, (12) becomes

$$P_{\hat{\mathbf{X}}^{0:k}}^{0:k} = \sum_{i=1}^{t!} \int_{\mathcal{L}_i\left(\hat{\mathbf{X}}^{0:k}\right)} \pi^{0:k}\left(\mathbf{X}^{0:k}\right) d\mathbf{X}^{0:k}$$

$$= \sum_{i=1}^{t!} \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}^{0:k}\right)} \pi^{0:k}\left(\Gamma_{\phi_i}^{-1}\left(\mathbf{X}^{0:k}\right)\right) d\mathbf{X}^{0:k}$$

$$= \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}^{0:k}\right)} \breve{\pi}^{0:k}\left(\left\{\mathbf{x}_1^{0:k}, \mathbf{x}_2^{0:k}, ..., \mathbf{x}_t^{0:k}\right\}\right) d\mathbf{X}^{0:k} \quad (37)$$

where the RFS density over the trajectories $\breve{\pi}^{0:k}\left(\left\{\mathbf{x}_1^{0:k}, ..., \mathbf{x}_t^{0:k}\right\}\right)$ is given by (14).

APPENDIX E

In this appendix we prove (22). The RFS trajectory posterior density from k to $k+1$ can be written as

$$\breve{\pi}^{k:k+1}\left(\left\{\mathbf{x}_1^{k:k+1}, ..., \mathbf{x}_t^{k:k+1}\right\}\right)$$

$$= \sum_{i=1}^{t!} \pi^{k:k+1}\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k:k+1}\right)\right)$$

$$= \sum_{i=1}^{t!} \frac{1}{\rho_i} \ell\left(\mathbf{z}^{k+1} \middle| \Gamma_{\phi_i}\left(\mathbf{X}^{k+1}\right)\right)$$

$$f\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k+1}\right) \middle| \Gamma_{\phi_i}\left(\mathbf{X}^{k}\right)\right) \pi^{k}\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k}\right)\right) \quad (38)$$

where

$$\rho_i = \int \ell\left(\mathbf{z}^{k+1} \middle| \Gamma_{\phi_i}\left(\mathbf{X}^{k+1}\right)\right) f\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k+1}\right) \middle| \Gamma_{\phi_i}\left(\mathbf{X}^{k}\right)\right)$$

$$\pi^{k}\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k}\right)\right) d\mathbf{X}^{k:k+1}. \quad (39)$$

We apply a change of variables in (39) using the function $\Gamma_{\phi_i}^{-1}\left(\cdot\right)$ as in Appendix D. Under Assumption A1, we can use (18) and (19), and we get:

$$\breve{\pi}^{k:k+1}\left(\left\{\mathbf{x}_1^{k:k+1}, ..., \mathbf{x}_t^{k:k+1}\right\}\right)$$

$$= \frac{\sum_{i=1}^{t!} \ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \pi^{k}\left(\Gamma_{\phi_i}\left(\mathbf{X}^{k}\right)\right)}{\int \ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \pi^{k}\left(\mathbf{X}^{k}\right) d\mathbf{X}^{k:k+1}}$$

$$= \frac{\ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \breve{\pi}^{k}\left(\left\{\mathbf{x}_1^{k}, ..., \mathbf{x}_t^{k}\right\}\right)}{\int \ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \pi^{k}\left(\mathbf{X}^{k}\right) d\mathbf{X}^{k:k+1}}$$

$$= \frac{\ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right) f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \breve{\pi}^{k}\left(\left\{\mathbf{x}_1^{k}, ..., \mathbf{x}_t^{k}\right\}\right)}{\rho}$$

$$(40)$$

where the normalising constant ρ is given by (23). Substituting (40) into (21), we get

$$P_i^{k:k+1} = 1/\rho \int_{\mathcal{L}_1\left(\hat{\mathbf{X}}_i^{k:k+1}\right)} \ell\left(\mathbf{z}^{k+1} \middle| \mathbf{X}^{k+1}\right)$$

$$f\left(\mathbf{X}^{k+1} \middle| \mathbf{X}^{k}\right) \breve{\pi}^{k}\left(\left\{\mathbf{x}_1^{k}, ..., \mathbf{x}_t^{k}\right\}\right) d\mathbf{X}^{k:k+1}.$$

REFERENCES

[1] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, Jan. 2004.

[2] 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, Jul. 1983.

[3] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[4] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic efficiency of the PHD in multitarget/multisensor estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 553–564, 2013.

[5] W. Koch and F. Govaers, "On accumulated state densities with applications to out-of-sequence measurement processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2766–2778, 2011.

[6] A. F. García-Fernández, J. Grajal, and M. R. Morelande, "Two-layer particle filter for multiple target detection and tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1569–1588, July 2013.

[7] 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 Transactions on Signal Processing*, vol. 54, no. 9, pp. 3291–3304, Sept. 2006.

[8] 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, July 2013.

[9] A. F. García-Fernández, "Detection and tracking of multiple targets using wireless sensor networks," Ph.D. dissertation, Universidad Politécnica de Madrid, 2011. [Online]. Available: http://oa.upm.es/9823/

[10] H. Blom, E. Bloem, Y. Boers, and H. Driessen, "Tracking closely spaced targets: Bayes outperformed by an approximation?" in *11th International Conference on Information Fusion*, July 2008, pp. 1–8.

[11] M. Guerriero, L. Svensson, D. Svensson, and P. Willett, "Shooting two birds with two bullets: How to find minimum mean OSPA estimates," in *13th Conference on Information Fusion*, July 2010, pp. 1–8.

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

[13] D. F. Crouse, P. Willett, and Y. Bar-Shalom, "Developing a real-time track display that operators do not hate," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3441–3447, July 2011.

[14] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo, "A metric for performance evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3452–3457, July 2011.

[15] A. F. García-Fernández, M. R. Morelande, and J. Grajal, "Particle filter for extracting target label information when targets move in close proximity," in *14th International Conference on Information Fusion*, 2011, pp. 795–802.

[16] H. Blom and E. Bloem, "Permutation invariance in Bayesian estimation of two targets that maneuver in and out formation flight," in *12th International Conference on Information Fusion*, July 2009, pp. 1296–1303.

[17] ——, "Decomposed particle filtering and track swap estimation in tracking two closely spaced targets," in *14th International Conference on Information Fusion*, 2011, pp. 787–794.

[18] D. F. Crouse, P. Willett, L. Svensson, D. Svensson, and M. Guerriero, "The set MHT," in *14th International Conference on Information Fusion*, 2011, pp. 2134–2141.

[19] R. Georgescu, P. Willett, L. Svensson, and M. Morelande, "Two linear complexity particle filters capable of maintaining target label probabilities for targets in close proximity," in *15th International Conference on Information Fusion*, 2012, pp. 2370–2377.

[20] E. H. Aoki, Y. Boers, L. Svensson, P. Mandal, and A. Bagchi, "A Bayesian look at the optimal track labelling problem," in *9th IET Data Fusion Target Tracking Conference*, 2012, pp. 1–6.

[21] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[22] D. F. Crouse, P. Willett, M. Guerriero, and L. Svensson, "An approximate minimum MOSPA estimator," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 3644–3647.

[23] M. Baum, P. Willett, and U. Hanebeck, "Calculating some exact MMOSPA estimates for particle distributions," in *15th International Conference on Information Fusion*, 2012, pp. 847–853.

[24] Ángel F. García-Fernández, M. R. Morelande, and J. Grajal, "Labelled OSPA metric for fixed and known number of targets." [Online]. Available: http://arxiv.org/abs/1404.3041

[25] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, 2000.

[26] L. Svensson, D. Svensson, M. Guerriero, and P. Willett, "Set JPDA filter for multitarget tracking," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4677–4691, Oct. 2011.

[27] Y. Bar-Shalom, T. Kirubarajan, and X. R. Li, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[28] D. Salmond, "Mixture reduction algorithms for point and extended object tracking in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 667–686, April 2009.
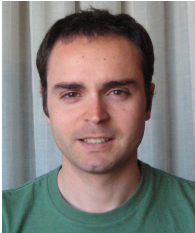
Ángel F. García-Fernández received the telecommunication engineering degree (with honours) and the Ph.D. degree from Universidad Politécnica de Madrid, Madrid, Spain, in 2007 and 2011, respectively.

He is currently a Research Associate in the Department of Electrical and Computer Engineering at Curtin University, Perth, Australia. His main research activities and interests are in the area of Bayesian nonlinear inference and radar imaging.
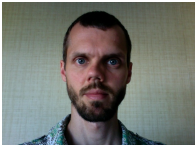
Mark R. Morelande received the B.Eng. degree in aerospace avionics from Queensland University of Technology, Brisbane, Australia in 1997 and the Ph.D. in electrical engineering from Curtin University of Technology, Perth, Australia in 2001.

In 2001 he was a Postdoctoral Fellow at the Centre for Eye Research, Queensland University of Technology. From 2002-2005 he was a Research Fellow at the Cooperative Research Centre for Sensor, Signal and Information Processing, University of Melbourne. He is now a Senior Research Fellow in the Melbourne Systems Laboratory, also at The University of Melbourne. His research interests include non-stationary signal analysis and target tracking with particular emphasis on multiple target tracking and sequential Monte Carlo methods.

Jesús Grajal was born in Toral de los Guzmanes (León), Spain, in 1967. He received the Ingeniero de Telecomunicación and the Ph.D. degrees from the Technical University of Madrid, Madrid, Spain in 1992 and 1998, respectively.

Since 2001 he has been an Associate Professor at the Signals, Systems, and Radiocommunications Department of the Technical School of Telecommunication Engineering of the same University. His research activities are in the area of hardware-design for radar systems, radar signal processing and broadband digital receivers for radar and spectrum surveillance applications.