

# Sensor Management for Search and Track Using the Poisson Multi-Bernoulli Mixture Filter

PER BOSTRÖM-ROST 

DANIEL AXEHILL , Senior Member, IEEE

GUSTAF HENDEBY , Senior Member, IEEE  
Linköping University, Linköping, Sweden

**A sensor management method for joint multitarget search and track problems is proposed, where a single user-defined parameter allows for a tradeoff between the two objectives. The multitarget density is propagated using the Poisson multi-Bernoulli mixture filter, which eliminates the need for a separate handling of undiscovered targets and provides the theoretical foundation for a unified search and track method. Monte Carlo simulations of two scenarios are used to evaluate the performance of the proposed method.**

Manuscript received June 16, 2020; revised November 5, 2020; released for publication February 8, 2021. Date of publication March 3, 2021; date of current version October 11, 2021.

DOI. No. 10.1109/TAES.2021.3061802

Refereeing of this contribution was handled by F. Govaers.

This work was supported in part by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, in part by the Industry Excellence Center LINKSIC funded by the Swedish Governmental Agency for Innovation Systems (VINNOVA), and in part by Saab AB.

Authors' address: The authors are with the Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden, E-mail: (per.bostrom-rost@liu.se; daniel.axehill@liu.se; gustaf.hendeby@liu.se).  
(Corresponding author: Per Boström-Rost.)

0018-9251 © 2021

## I. INTRODUCTION

This article considers sensor management for multi-target tracking. The problem is to control a sensor such that it maintains accurate estimates of the states of the discovered targets, and at the same time, searches for targets that are still undiscovered. Complicating factors for the problem are, apart from the sensor's limited field of view (FoV) and sensing range, that the number of targets is both unknown and timevarying, the measurements are corrupted by noise, there are misdetections, false alarms, and unknown measurement origins.

Classical approaches to multitarget tracking include multiple hypothesis tracking (MHT) [1] and the joint probabilistic data association filter [2]. In recent years, much attention has been focused on methods based on the theory of random finite sets (RFSs) [3], [4], which enables a Bayesian approach to the problem of multitarget tracking. In general, the exact multitarget posterior is intractable to compute, but several tractable algorithms have been developed based on different approximations. Examples of these include the probability hypothesis density (PHD) filter [5], the generalized labeled multi-Bernoulli (GLMB) filter [6], [7], and the Poisson multi-Bernoulli mixture (PMBM) filter [8]. In contrast to the GLMB filter, the PMBM filter operates on unlabeled sets of target states and does not formally provide track continuity. An alternative method is the PMBM tracker [9], which estimates the posterior density over the set of target trajectories and thereby, provides track continuity without introducing labels.

Sensor management, *i.e.*, optimization of control inputs for sensors, becomes interesting in situations where the state of the sensor affects the quality of the measurements. The RFS framework has been used in several sensor management applications with focus on improving the estimation performance of discovered targets [10]–[13]. The PHD representation has also been used to plan sensor trajectories with the aim of discovering the previously undetected targets [14].

Both planning for improving the tracking performance and planning the search for new targets are arguably difficult problems. Even more challenging is the problem of planning for jointly maintaining track of discovered targets and searching for undiscovered targets, as the two objectives are often competing for the same sensor resources. In [15], the PHD filter is used in combination with a Voronoi-based control method to enable a team of mobile sensors to concurrently search for and track an unknown number of targets. The same problem is studied in [16] and [17], where information-theoretic objective functions are used to prioritize between searching and tracking. Whereas [15] considers only a single time step, the methods in [16] and [17] are nonmyopic and optimize sensor control inputs over a horizon of multiple time steps. Both [16] and [17] assume known measurement origins, track each detected target separately, and make use of an occupancy grid filter to represent the presence of undetected targets.

In this article, a new nonmyopic sensor management approach based on the PMBM filter<sup>1</sup> is proposed for multitarget tracking. The PMBM density, which has been shown to be a conjugate prior for the multitarget Bayes filter [8], allows for an elegant separation of the set of targets into two disjoint subsets: targets that have been detected, and targets that are still undiscovered. This provides a solid theoretical foundation for a unified planning algorithm for search and track. The explicit modeling of undetected targets in the PMBM filter is exploited to tradeoff between searching for new targets and tracking detected targets. Furthermore, the underlying PMBM filter eliminates the need of a separate handling of undiscovered targets and known measurement origins as in [16] and [17].

#### A. Notation

Vectors and scalars are described by lower case letters  $x$ , matrices and sets by upper case letters  $X$ , and index sets by blackboard bold letters  $\mathbb{X}$ . The cardinality of a set  $X$  is denoted  $|X|$ . The operator  $\uplus$  denotes the disjoint set union, *i.e.*,  $X^1 \uplus X^2 = X$  means  $X^1 \cup X^2 = X$  and  $X^1 \cap X^2 = \emptyset$ . The inner product of  $a(x)$  and  $b(x)$  is  $\langle a, b \rangle = \int a(x)b(x) dx$ .

#### B. Outline

In Section II, the general sensor management problem for multitarget tracking is presented. Section III briefly reviews the RFS-based target tracking and gives an introduction to the PMBM filter. The details of the proposed PMBM-based sensor management method are presented in Section IV, followed by an approach to solve the planning problem using Monte Carlo tree search (MCTS) in Section V. Results from simulation studies are presented in Section VI, and finally, Section VII concludes this article.

### II. PROBLEM FORMULATION

The problem of finding the optimal sequence of control inputs for a sensor can be formulated as a stochastic optimal control problem [18]. A problem of this type is often referred to as a partially observable Markov decision process (POMDP) [19], [20], which provides a framework for planning under uncertainty. This section outlines general models of the system dynamics and measurements. These models are used to formulate the problem of sensor management for multitarget tracking as a stochastic optimal control problem.

#### A. Dynamics and Measurement Models

Consider a controllable sensor of which the dynamics are governed by the model

$$s_{k+1} = g(s_k, u_k) \quad (1)$$

where  $s_k$  is the state of the sensor and  $u_k \in \mathcal{U}$  is the control input applied at time  $k$ , and  $\mathcal{U}$  is the set of admissible control

inputs. The state of the sensor is assumed to be perfectly known at all times.

The models of the multitarget dynamics and measurements presented here are inspired by [10]. Let  $X_k$  denote the set of targets at time  $k$ . The dynamics of the multitarget state  $X_k$  is defined as follows. As time progresses from  $k$  to  $k+1$ , each individual target  $x_k \in X_k$  generates a new set  $S_{k+1|k}(x_k)$ , which is empty if the target terminates and contains a single element  $x_{k+1}$  if the target survives. By representing the set of newborn targets with  $B_{k+1}$ , the multitarget state transition equation is given by

$$X_{k+1} = f(X_k) = \bigcup_{x_k \in X_k} S_{k+1|k}(x_k) \cup B_{k+1}. \quad (2)$$

Note that the multitarget set  $X_k$  includes *both* the detected targets and targets that have not yet been detected by the sensor.

The multitarget measurement model is defined as follows. At time  $k$ , each target  $x_k \in X_k$  generates a measurement set  $D_k(x_k, s_k)$ . This set contains a single element if the target is detected by the sensor with state  $s_k$ , and is empty if the target is undetected. By denoting the set of false alarms by  $F_k(s_k)$ , the full set of measurements  $Z_k$  generated by  $X_k$  is given by the multitarget measurement equation

$$Z_k = h(X_k, s_k) = \bigcup_{x_k \in X_k} D_k(x_k, s_k) \cup F_k(s_k). \quad (3)$$

#### B. Stochastic Optimal Control Problem

The state of the multitarget set  $X_k$  is estimated using the measurements obtained by the sensor system. Let  $\pi_{k|k} = \pi_{k|k}(X_k | Z_{1:k}, s_{1:k})$  denote the posterior multitarget state density at time  $k$  after observing the measurement sets  $Z_{1:k}$  obtained at sensor states  $s_{1:k}$ . Here,  $Z_{1:k}$  is the sequence of measurement sets up until and including time  $k$  and  $s_{1:k}$  is defined analogously for the sensor state. Furthermore, let  $f_{k|k-1}(\cdot)$  and  $f_{k|k}(\cdot)$  denote the prediction and update steps of the multitarget Bayes filter [3], which will be described in Section III-B.

The considered sensor management problem can then be formulated as follows. Given the current sensor state  $s_k$ , the multitarget density  $\pi_{k|k}$ , and a planning horizon  $N$ , an optimal sequence of sensor control inputs  $u_{k:k+N-1}$  is obtained by solving the following stochastic optimal control problem:

$$\begin{aligned} & \underset{u_{k:k+N-1}}{\text{minimize}} && \mathbf{E}_{Z_{k+1:k+N}} \left[ \sum_{t=k+1}^{k+N} \ell(\pi_{t|t}) \right] \\ & \text{subject to} && u_t \in \mathcal{U}, \\ & && s_{t+1} = g(s_t, u_t) \\ & && \pi_{t+1|t+1} = f_{k|k}(f_{k|k-1}(\pi_{t|t}), Z_{t+1}, s_{t+1}) \\ & && X_{t+1} = f(X_t) \\ & && Z_t = h(X_t, s_t) \end{aligned} \quad (4)$$

where  $\ell(\cdot)$  is the stage cost function, which is used to tradeoff between tracking discovered targets and searching for new targets. The argument of the stage cost function is

<sup>1</sup>The proposed method relies on the PMBM filter, but could with minor adaptations instead be based on the PMBM tracker. However, the benefit of this from a practical planning point of view is limited.

the resulting multitarget density at the corresponding time step, *i.e.*,  $\pi_{t|t}$ . Since this density depends on the measurement sets  $Z_{k+1:t}$ , which at the time of planning are not yet obtained, the expectation of the objective function is taken with respect to the future measurement sets  $Z_{k+1:k+N}$ . This approach is commonly used in sensor management applications [10] [12], [21].

Since the targets maneuver and the scenario changes over time, it is necessary to replan the sequence of sensor control inputs online as new measurements are obtained. This is done in a standard receding horizon fashion [22].

- 1) At time  $k$ , update the multitarget state density to  $\pi_{k|k}$  using the set of measurements  $Z_k$  of the set of targets  $X_k$ , obtained by the sensor system with state  $s_k$ .
- 2) Compute a sequence of control inputs  $u_{k:k+N-1}$  by solving (4).
- 3) Apply the first element of the computed sequence of control inputs,  $u_k$ , to the sensor system, and discard the remaining elements. At time  $k+1$ , the process is repeated from step 1.

### III. BACKGROUND ON MULTITARGET FILTERING

The problem formulated in the previous section requires the estimation of a multitarget state density, for which methods based on the RFS theory can be applied. This section provides an overview of RFS models, multitarget filtering using the RFS, and the PMBM filter.

#### A. RFS Models

An *RFS* is a set with a random number of elements that are themselves random [3]. This means that the cardinality of an RFS is a discrete random variable, and its elements are random variables. This makes the RFS framework convenient for representing sets of multiple targets and multiple sensor measurements. In this article, the following RFSs will be of interest.

1) *Poisson Point Process (PPP)*: A PPP is an RFS of which the cardinality is Poisson distributed with rate  $\mu$  and the elements, given the cardinality, are independent and identically distributed according to  $p(x)$ . The rate  $\mu$  and distribution  $p(x)$  form the intensity  $\lambda(x)$  of the PPP as  $\lambda(x) = \mu p(x)$ . The density of a PPP  $X$  is given by [23]

$$\pi(X) = e^{-(\lambda,1)} \prod_{x \in X} \lambda(x) \quad (5)$$

where the notation  $\langle a, b \rangle = \int a(x)b(x)dx$  is used for the inner product of  $a(x)$  and  $b(x)$ . In this work, PPPs are used to model clutter measurements, target birth, and undetected targets.

2) *Bernoulli RFS*: The cardinality of a Bernoulli RFS is Bernoulli distributed with parameter  $r \in [0, 1]$ . It is either empty, with probability  $1 - r$ , or, with probability  $r$ , contains a single element with density  $p(x)$ . Thus, the density

of a Bernoulli RFS  $X$  is [23]

$$\pi(X) = \begin{cases} 1 - r, & X = \emptyset \\ rp(x), & X = \{x\} \\ 0, & |X| \geq 2. \end{cases} \quad (6)$$

As the Bernoulli RFS captures both the uncertainty regarding a target's existence (through the parameter  $r$ ) and state (through the density  $p(x)$ ), the Bernoulli RFS provides a convenient representation for single targets.

3) *Multi-Bernoulli (MB) RFS*: The disjoint union of a fixed number of Bernoulli RFSs is an MB RFS. Its density is defined by the parameters  $\{r^i, p^i\}_{i \in \mathbb{I}}$ , where  $\mathbb{I}$  is an index set

$$\pi(X) = \begin{cases} \sum_{\cup_{i \in \mathbb{I}} X^i = X} \prod_{i \in \mathbb{I}} \pi^i(X^i), & |X| \leq |\mathbb{I}| \\ 0, & |X| > |\mathbb{I}|. \end{cases} \quad (7)$$

The notation  $X^1 \uplus X^2 = X$  denotes disjoint union, *i.e.*,  $X^2 \cup X^2 = X$  and  $X^2 \cap X^2 = \emptyset$ .

4) *Multi-Bernoulli Mixture (MBM) RFS*: A normalized, weighted sum of MB RFSs is referred to as an MBM RFS. Its density can be stated as [24]

$$\pi(X) = \sum_{j \in \mathbb{J}} w^j \sum_{\cup_{i \in \mathbb{I}} X^i = X} \prod_{i \in \mathbb{I}} \pi^{j,i}(X^i) \quad (8)$$

and is defined by the set of parameters  $\{w^j, \{r^{j,i}, p^{j,i}\}_{i \in \mathbb{I}}\}_{j \in \mathbb{J}}$ , where  $\mathbb{J}$  is an index set for the MB components of the MBM,  $\mathbb{I}^j$  is an index set for the Bernoullis of the  $j$ th MB RFS, and  $w^j$  is the weight of the  $j$ th MB.

#### B. Multitarget Filtering Using the RFS Theory

In the RFS-based filtering approach, the multitarget state and set of measurements at the time  $k$  are modeled as two RFSs denoted  $X_k$  and  $Z_k$ , respectively. The aim is to estimate the posterior multitarget state density  $\pi_{k|k}(X_k | Z_{1:k}, s_{1:k})$ .

Similarly as in the standard single-target case, the multitarget posterior density can be computed recursively via prediction and measurement update steps. With the Bayes multitarget filter [3], the posterior multitarget density at time  $k-1$  is propagated in time using the prediction step

$$\pi_{k|k-1}(X_k | Z_{1:k-1}, s_{1:k-1}) = \int \phi_X(X_k | X_{k-1}) \pi_{k-1|k-1}(X_{k-1} | Z_{1:k-1}, s_{1:k-1}) \delta X_{k-1} \quad (9)$$

where  $\phi_X(X_k | X_{k-1})$  is the multitarget transition density. The standard multitarget transition model [3] is used, which means that  $\phi_X(X_k | X_{k-1})$  models a Markovian process for individual targets with transition density  $p(x_k | x_{k-1})$  and state-dependent probability of survival  $p_s(x)$ , combined with a PPP birth process with intensity  $\lambda^b(x)$ .

Given a measurement set  $Z_k$  obtained using a sensor with the state  $s_k$ , the predicted density is updated using the update step

$$\pi_{k|k}(X_k | Z_{1:k}, s_{1:k}) = \frac{\phi_Z(Z_k | X_k, s_k) \pi_{k|k-1}(X_k | Z_{1:k-1}, s_{1:k-1})}{\int \phi_Z(Z_k | X_k, s_k) \pi_{k|k-1}(X_k | Z_{1:k-1}, s_{1:k-1}) \delta X_k} \quad (10)$$

where  $\phi_Z(Z_k | X_k, s_k)$  is the sensor-dependent multitarget measurement set likelihood function. The standard multitarget measurement model [3] is used, which means that  $\phi_Z(Z_k | X_k, s_k)$  models noisy measurements of individual targets with state-dependent probability of detection  $p_D(x | s)$ , combined with PPP clutter with intensity  $\lambda^{fa}(z | s)$ . At most one measurement is generated for each target in each time step, and each measurement is the result of at most one target. Each measurement is independent of all other targets and other measurements conditioned on the same target; the single target measurement likelihood is  $p(z | x, s)$ . The integrals in (9) and (10) are set integrals, defined in [3].

### C. PMBM Filter

The RFS-based filter used in this work is the PMBM filter, which is derived in [8] and [25]. This section gives an overview of the key concepts in the PMBM filter based on [24] and [25].

1) *Interpretation:* In the PMBM filter, the set of targets  $X$  is divided into two disjoint subsets  $X^u \uplus X^d = X$ , corresponding to the undetected targets  $X^u$  and the targets that have been detected at least once  $X^d$ . The distribution of the targets is modeled by the union of a PPP component and an MBM component.

The MBM component models the potentially detected targets and considers all possible data association hypotheses. Each measurement at each time step may be either the first detection of a new target, a new measurement of a target that has previously been detected, or a false alarm. Each measurement thus results in a new potentially detected target whose distribution is Bernoulli, as it might not exist. For each potentially detected target, there are single target hypotheses that represent sequences of target-to-measurement associations. A global hypothesis, represented by an MB component in the MBM, consists of a number of single target hypotheses from different potentially detected targets. The unknown data associations lead to an intractably large number of components in the MBM density, and approximations are necessary for tractability, *e.g.*, using a track-oriented MHT formulation [25].

An interesting unique feature of the PMBM filter is the explicit modeling of targets that are not yet detected. The distribution of these targets is represented by the PPP component. This enables the user to incorporate prior knowledge about where new targets are likely to appear and behave before they are detected by the sensor. An example of how the intensity of undetected targets might be distributed is given in Fig. 1. This explicit modeling of undetected targets can also be used in sensor management applications in order to guide the sensor toward areas where new targets are likely to be discovered.

2) *Mathematical Representation:* The PMBM density is a combination of a PPP density  $\pi^P$  on the set of undetected targets  $X^u$  and an MBM density  $\pi^{MBM}$  on the set of detected

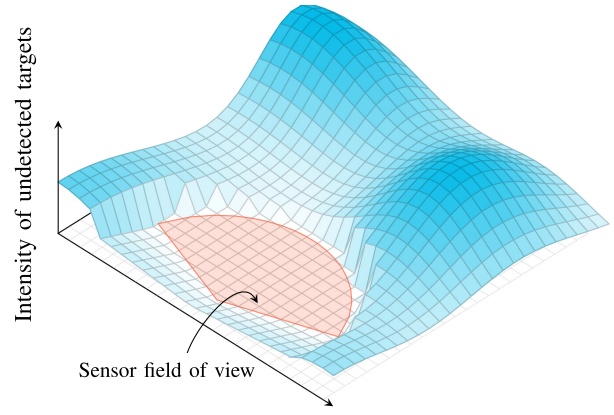


Fig. 1. The intensity of the Poisson component in the PMBM density indicates parts of the state space where undetected targets are likely to be found. This can be exploited when planning the search for new targets.

targets  $X^d$ . The full multitarget density is given by

$$\pi^{PMBM}(X) = \sum_{X^u \uplus X^d = X} \pi^P(X^u) \pi^{MBM}(X^d) \quad (11)$$

where  $\pi^P(X^u)$  is given by (5) and  $\pi^{MBM}(X^d)$  is given by (8). The full PMBM density is completely defined by the following set of parameters:

$$\lambda^u, \{w^j, \{r^{j,i}, p^{j,i}\}_{i \in \mathbb{I}^j}\}_{j \in \mathbb{J}} \quad (12)$$

where  $\lambda^u$  is the intensity of undetected targets, and  $\mathbb{J}$  is an index set for the global hypotheses, *i.e.*, the components of the MBM. The  $j$ th MB component has weight  $w^j$  and density  $\pi^j$  and consists of  $|\mathbb{I}^j|$  Bernoulli components. The Bernoulli density of the  $i$ th potentially detected target under the  $j$ th global hypothesis is defined by the probability of the existence  $r^{j,i}$  and spatial density  $p^{j,i}$ .

The PMBM density has been shown to be a multitarget tracking conjugate prior [8], which means that the prediction (9) and measurement update (10) steps correspond to computing new values for the set of parameters in (12). The equations for the update steps are outlined in Appendix A. For derivations and more details, the reader is referred to [8] and [25].

There are several ways to extract a set of estimated target states from the PMBM density [25]. The method used in this article is based on the global hypothesis with the highest weight. Given the set of PMBM parameters (12), the index of the global hypothesis with the highest weight is obtained as

$$j^* = \arg \max_{j \in \mathbb{J}} w^j. \quad (13)$$

From the corresponding MB component with parameters  $\{r^{j^*,i}, p^{j^*,i}\}_{i \in \mathbb{I}^{j^*}}$ , all Bernoulli components with probability of existence larger than a threshold  $\tau$  are selected, and the expected values of the target states are included in the set of target estimates  $\hat{X}$ , *i.e.*,

$$\hat{X} = \{\hat{x}^{j^*,i}\}_{i \in \mathbb{I}^{j^*}: r^{j^*,i} > \tau} \quad (14a)$$

$$\hat{x}^{j^*,i} = \mathbf{E}[p^{j^*,i}]. \quad (14b)$$



#### IV. PMBM-BASED SENSOR MANAGEMENT

This section formulates a tractable approximation of the original problem (4) based on the PMBM filter that allows for a tradeoff between tracking and searching. The planning for tracking of discovered targets is based on the best hypothesis in the MBM component of the PMBM density, whereas the intensity of the Poisson component is used to guide the sensor toward regions where the probability of discovering undetected targets is high.

##### A. Assumptions

We begin by introducing a number of simplifying assumptions. First of all, it is assumed that the multitarget posterior  $\pi_{k|k}$  is described by a PMBM density defined by the parameters in (12). The detection and survival probabilities are approximated as

$$p_D(x|s) \approx p_D(\mathbf{E}[x]|s) \quad (15a)$$

$$p_S(x) \approx p_S(\mathbf{E}[x]) \quad (15b)$$

and the birth intensity  $\lambda^b$  is modeled as a Gaussian mixture

$$\lambda^b(x) = \sum_{i=1}^{N^b} w^{b,i} \mathcal{N}(x; x^{b,i}, P^{b,i}). \quad (16)$$

Then, by using the prediction and measurement update steps of the extended Kalman filter (EKF) [26] to compute the predictions and measurement updates of the PMBM filter, the intensity of the undetected targets  $\lambda^u$  is a Gaussian mixture and the Bernoulli densities  $p^{j,i}$  are Gaussian

$$\lambda^u(x) = \sum_{i=1}^{N^u} w^{u,i} \mathcal{N}(x; x^{u,i}, P^{u,i}) \quad (17a)$$

$$p^{j,i}(x) = \mathcal{N}(x; \hat{x}^{j,i}, P^{j,i}). \quad (17b)$$

Furthermore, it is assumed that the set of admissible control inputs  $\mathcal{U}$  is a finite set. With these assumptions in place, we are ready to discuss planning for tracking and searching.

##### B. Tracking Discovered Targets

From the tracking perspective, the objective of the planning problem is to reduce the uncertainty in the state estimates of the discovered targets. Each admissible sequence of control inputs results in a unique sensor state trajectory, along which corresponding measurement sets can be predicted. Using these measurement sets together with the prediction and update steps of the PMBM filter, it is possible to predict the evolution of the state estimates. These predictions can be used to compare future uncertainties in the state estimates after applying different sequences of control inputs.

1) *Approximations to Reduce the Computational Cost:* The predicted ideal measurement set approach [27] is employed to avoid the computationally expensive expectation over the future measurement sets in (4). This means that the planning problem is solved under the assumption that the future measurement sets  $Z_t$ , for  $t = k + 1, \dots, k + N$ , are

generated under ideal conditions, *i.e.*, every target within FoV generates a noise-free measurement and there are no false alarms. Thus, given sensor state  $s_t$  and a set of predicted target states  $\hat{X}_t^{\tau_p}$ , the predicted ideal measurement set  $\hat{Z}_t$  is computed as

$$\hat{Z}_t = \bigcup_{\hat{x} \in \hat{X}_t^{\tau_p} : \hat{x} \in \text{FoV}(s_t)} \{ \arg \max_z p(z | \hat{x}, s_t) \}. \quad (18)$$

The set of predicted target states  $\hat{X}_t^{\tau_p}$  is extracted from the best hypothesis in the MBM component of the PMBM density. The index of the MB corresponding to this hypothesis is denoted  $j^*$  and is found using (13). The extraction is similar to the estimation step of the PMBM filter (14), but with a different threshold on the probability of existence  $\tau_p < \tau$  as

$$\hat{X}_t^{\tau_p} = \{ \hat{x}_{t|k}^{j^*,i} \}_{i \in \mathbb{I}^{j^*} : r_{k|k}^{j^*,i} > \tau_p}. \quad (19)$$

Newly detected targets typically have a low probability of existence and the lower threshold  $\tau_p$  allows for including measurements of these targets in the predicted ideal measurement set. This is a way to encourage the sensor to confirm or reject new possible tracks.

Note that the predicted ideal measurement sets are generated with known measurement origins. This can be exploited to decrease the computational cost related to solving the planning problem. As the predicted measurements match hypothesis  $j^*$  exactly except for targets with probability of the existence lower than  $\tau_p$ , this hypothesis will with high probability remain as the most likely hypothesis after updating the PMBM density using these measurements. Thus, with little risk of introducing errors, no other hypothesis than the one with index  $j^*$  need to be updated and no new association hypotheses need to be introduced.

For notational compactness, we define the function  $\hat{h}(\cdot)$  and write the predicted ideal measurement set as

$$\hat{Z}_t = \hat{h}(s_t) \quad (20)$$

where the dependence on  $\hat{X}_t^{\tau_p}$  is implicit.

2) *Stage Cost:* As estimates of individual target states are Gaussian distributed, a measure on the corresponding covariance matrices gives an indication of the tracking performance. Several scalar measures of covariance matrices exist, see, *e.g.*, [28]. Here, the state estimation accuracy is measured by summing the trace of the covariance matrix of each track

$$\ell^{\text{track}}(\pi_{t|t}) = \sum_{i \in \mathbb{I}^{j^*}} \text{tr}(W P_{t|t}^{j^*,i} W^T) \quad (21)$$

where  $W$  is a matrix of appropriate dimensions used to extract and weigh the relevant components of the covariance matrix. The variable  $P_{t|t}^{j^*,i}$  is the covariance matrix of track  $i$  in hypothesis  $j^*$  at time  $t$  given the actual measurements up until time  $k$  and predicted ideal measurements from time  $k + 1$  to  $t$ .

### C. Searching for New Targets

The intensity of the Poisson component in the PMBM density provides a convenient representation of where undetected targets are likely to be found. This is illustrated in Fig. 1. Areas with high birth intensity and parts of the surveillance volume that have not been visited by the sensor in a long time will have a higher intensity of undetected targets.

The intensity  $\lambda^u$  is a deterministic function of the sensor state trajectory that can be computed using the prediction and update steps of the PMBM filter. Thus, the resulting intensity of undetected targets after applying different sequences of control inputs can be predicted and evaluated based on a suitable performance measure.

1) *Stage Cost*: As the objective of the planning problem from the search perspective is to discover new targets, we propose to use the expected number of undetected targets as stage cost. This is given by the integral of the intensity of undetected targets, which is equal to the sum of the weights in the Gaussian mixture

$$\ell^{\text{search}}(\pi_{t|t}) = \int \lambda_{t|t}^u(x) dx = \sum_{i=1}^{N_t^u} w_{t|t}^{u,i}. \quad (22)$$

Here,  $\lambda_{t|t}^u$  is the resulting intensity of undetected targets at time  $t$  given the actual sensor trajectory up until time  $k$  and planned trajectory from time  $k+1$  to  $t$ , and  $N_t^u$  is the number of components in the Gaussian mixture at time  $t$ .

### D. Resulting Optimal Control Problem

For the full planning problem, the proposed stage cost function is a weighted sum of  $\ell^{\text{track}}(\cdot)$  and  $\ell^{\text{search}}(\cdot)$ , i.e.,

$$\ell(\pi_{t|t}) = \ell^{\text{track}}(\pi_{t|t}) + \eta \ell^{\text{search}}(\pi_{t|t}) \quad (23)$$

where  $\eta$  is a user-defined weight used to tradeoff between tracking and searching. The parameter  $\eta$  can be considered as the cost of not detecting a target, as  $\ell^{\text{search}}(\cdot)$  corresponds to the expected number of undetected targets.

The resulting deterministic approximation of the original problem (4), which is solved when planning the sensor control input at time  $k$ , is formulated as follows:

$$\begin{aligned} & \underset{u_{k:k+N-1}}{\text{minimize}} && \sum_{t=k+1}^{k+N} \ell(\pi_{t|t}) \\ & \text{subject to} && u_t \in \mathcal{U} \\ & && s_{t+1} = g(s_t, u_t) \\ & && \pi_{t+1|t+1} = \rho_{k|k}(\rho_{k|k-1}(\pi_{t|t}), \hat{Z}_{t+1}, s_{t+1}) \\ & && \hat{Z}_t = \hat{h}(s_t) \end{aligned} \quad (24)$$

where  $\rho_{k+1|k}(\cdot)$  and  $\rho_{k|k}(\cdot)$  represent the prediction and measurement update steps of the PMBM filter, respectively.

If  $\mathcal{U}$  is a finite set of admissible controls, the solution to (24) can in the theory be obtained by explicitly enumerating all possible control sequences  $u_{k:k+N-1}$  and selecting the one that results in the smallest objective function value. However, the computational complexity of this method is prohibitive for many problems. For larger problems, we

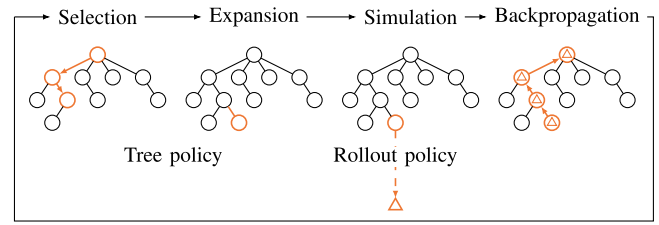


Fig. 2. Overview of the MCTS method. Figure adapted from [29].

propose to use a MCTS method [29] that incrementally builds a search tree and allocates its search time to the most promising parts of the tree. Such a method is described next.

### V. MONTE CARLO TREE SEARCH (MCTS)

The standard MCTS [29] is a stochastic method for finding the optimal actions in search problems. This is done by iteratively building a search tree until some predefined computational budget is reached, after which the best control input at the root node is returned. Each iteration consists of four phases: selection, expansion, simulation, and backpropagation [29]. These phases are often grouped into a *tree policy* (selection and expansion) and a *rollout policy* (simulation). The backpropagation step is used to update node statistics that affect the tree policy decisions in future iterations. Fig. 2 illustrates one iteration of the MCTS method and shows how the search tree is iteratively constructed.

The MCTS method tailored for the PMBM-based sensor management is outlined in Algorithm 1. Each node  $n$  in the tree holds information about how many times the node has been visited so far during the search  $c_v(n)$ , the corresponding sensor state and PMBM density, and the average total cost  $\hat{J}(n)$  of control input sequences that pass through the node. An edge in the search tree represents a control input applied to the sensor.

In the selection phase (Algorithm 1, line 3), a node  $n$  that still has unvisited children is selected for expansion. This is done by starting at the root node and recursively selecting child nodes until a node with at least one unvisited child is reached. Child nodes are selected using the upper confidence bound for trees (UCT) policy [30], i.e., a child node  $n'$  of the node  $n$  is selected as

$$\arg \max_{n' \in \text{children of } n} \left\{ -\hat{J}(n') + \varepsilon \sqrt{\frac{\log c_v(n)}{c_v(n')}} \right\} \quad (25)$$

where the second term encourages selection of control inputs that have rarely been tried, and  $\varepsilon$  is an exploration parameter to tradeoff between exploration and exploitation.

In the expansion phase (Algorithm 1, line 4), a new child node  $n'$  is added to the node selected for expansion  $n$ . This corresponds to selecting a previously untried control input, updating the state of the sensor, generating the predicted ideal measurement set, and predicting and updating the PMBM density.

A simulation (Algorithm 1, line 5) is then run from the new node  $n'$  until the planning horizon, computing the stage costs along the way. This produces an estimate  $\Delta$  of

---

**Algorithm 1:** Overview of MCTS For PMBM-Based Sensor Management.

---

```
1: Initialize tree  $\mathcal{T}$  with root node  $n_0$  with current sensor state and PMBM density
2: for  $N_{\text{iter}}$  iterations do
3:    $n \leftarrow \text{SELECTNODE}(\mathcal{T})$  ▷ Select node to expand using UCT policy
4:    $n' \leftarrow \text{EXPANDTREE}(n)$  ▷ Add new child to node  $n$ ; simulate measurements and update PMBM density
5:    $\Delta \leftarrow \text{SIMULATE}(n')$  ▷ Perform rollout until planning horizon; compute total cost
6:    $\mathcal{T} \leftarrow \text{BACKPROPAGATE}(\mathcal{T}, n, \Delta)$  ▷ Update node statistics
7: end for
8: return  $\arg \min_{n \in \text{children of } n_0} \hat{J}(n)$ 
```

---

the total cost of a control input sequence that starts with the control inputs leading up to  $n'$ . During the simulation phase, the control inputs are selected according to a rollout policy, which could correspond to selecting control inputs at random or based on heuristic for the specific problem [31]. In the backpropagation phase (Algorithm 1, line 6), the result  $\Delta$  is used to update the statistics of the nodes selected for this iteration. Each node's visit count is incremented and its average total cost updated according to  $\Delta$ .

## VI. SIMULATION STUDY

In this section, the proposed PMBM-based sensor management method is evaluated using a series of simulated scenarios. The achieved performance is compared for the following three different stage cost functions:

- 1) the proposed stage cost function (23) for search and track (ST);
- 2) the stage cost function (21) for track only (TO); and
- 3) the stage cost function (22) for search only (SO).

The PMBM-based method is also compared to the grid-based method proposed in [17]. The method from [17] assumes known measurement origins, which simplifies the problem as the true measurement-to-target associations are given. A Gaussian mixture implementation of the Bernoulli filter [32] is used as an underlying tracker for each individual target, and an occupancy grid filter is used to represent undetected targets.

### A. Performance Evaluation

The generalized optimal subpattern assignment (GOSPA) metric [33] is a unified performance metric for multitarget tracking. It penalizes both localization errors for detected targets and errors due to missed and false targets. According to [33], with  $\alpha = 2$  and given  $c > 0$ ,  $1 \leq p < \infty$ , a metric  $d(\cdot)$  in the single target space, the ground truth set  $X = \{x^1, \dots, x^{|X|}\}$  and its estimate  $\hat{X} = \{\hat{x}^1, \dots, \hat{x}^{|\hat{X}|}\}$ , the GOSPA metric can be written as

$$\begin{aligned} & d_p^{(c,2)}(X, \hat{X}) \\ &= \min_{\gamma \in \Gamma} \left( \sum_{(i,j) \in \gamma} d(x^i, \hat{x}^j)^p + \frac{c^p}{2} (|X| + |\hat{X}| - 2|\gamma|) \right)^{\frac{1}{p}} \end{aligned} \quad (26)$$

where  $\Gamma$  is the set of all possible assignment sets  $\gamma$  between  $\{1, \dots, |X|\}$  and  $\{1, \dots, |\hat{X}|\}$ . Such an assignment set satisfies  $\gamma \subseteq \{1, \dots, |X|\} \times \{1, \dots, |\hat{X}|\}$ ,  $(i, j), (i, j') \in \gamma \Rightarrow j = j'$  and  $(i, j), (i', j) \in \gamma \Rightarrow i = i'$ , where the last two properties ensure that every  $i$  and  $j$  gets at most one assignment [33]. Note that  $|\gamma|$  is the number of properly detected targets and that the definition implies that  $|\gamma| \leq \min\{|X|, |\hat{X}|\}$ .

Given the optimal assignment set  $\gamma^*$ , the GOSPA metric can be decomposed as [33]

$$\begin{aligned} & d_p^{(c,2)}(X_k, \hat{X}_k) \\ &= \left( d^{\text{loc}}(X_k, \hat{X}_k, \gamma^*) + d^{\text{miss}}(X_k, \gamma^*) + d^{\text{false}}(\hat{X}_k, \gamma^*) \right)^{\frac{1}{p}} \end{aligned} \quad (27)$$

where  $d^{\text{loc}}(\cdot)$  is the cost related to the localization of properly assigned targets, and  $d^{\text{miss}}(\cdot)$  and  $d^{\text{false}}(\cdot)$  are the costs related to missed and false targets. The three terms in the decomposed metric are computed as

$$d^{\text{loc}}(X, \hat{X}, \gamma^*) = \sum_{(i,j) \in \gamma^*} d(x^i, \hat{x}^j)^p \quad (28a)$$

$$d^{\text{miss}}(X, \gamma^*) = \frac{c^p}{2} (|X| - |\gamma^*|) \quad (28b)$$

$$d^{\text{false}}(\hat{X}, \gamma^*) = \frac{c^p}{2} (|\hat{X}| - |\gamma^*|) \quad (28c)$$

where  $|X| - |\gamma^*|$  and  $|\hat{X}| - |\gamma^*|$  are the number of missed and false targets, respectively. Note that the choice of the maximum allowable localization cost  $c$  defines the cost of not detecting a target as  $c^p/2$ , which can be used to find an appropriate value for the weight parameter  $\eta$  in the objective function (23).

To evaluate the overall performance in the simulation study, the GOSPA is used as the main metric. For the parameters, we use  $c = 50$ ,  $p = 2$ , and define the base distance  $d(x, y)$  as the 2-norm of the position components of  $x - y$ . We also report the following:

- 1) the normalized localization error

$$d^{\text{nle}}(X, \hat{X}, \gamma^*) = \left( \frac{1}{|\gamma^*|} d^{\text{loc}}(X, \hat{X}, \gamma^*) \right)^{\frac{1}{p}} \quad (29)$$

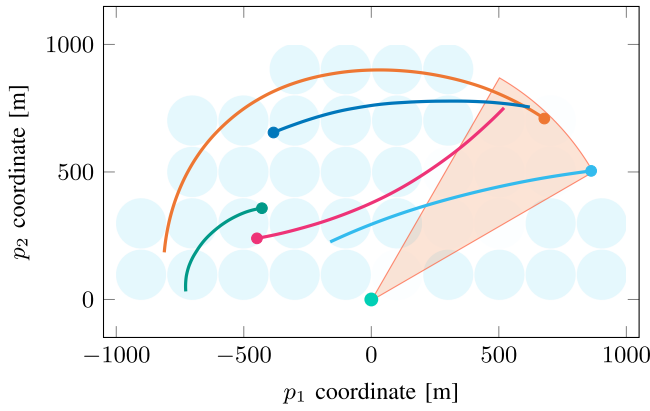


Fig. 3. Example of simulated scenario. There are a total of five targets appearing every 80 s. Two targets disappear at time steps 280 and 360, respectively. The targets' initial positions are marked with filled circles. The FoV is orange and the light blue circles illustrate the Gaussian mixture intensity of the undetected targets.

*i.e.*, the average localization error of properly assigned targets as a localization performance measure;

- 2) the number of missed targets as a target discovery performance measure;
- 3) the number of false targets; and
- 4) the number of lost targets.

#### B. Beam Pointing for a Stationary Sensor

In this scenario, the pointing direction of a stationary range-bearing sensor is controlled to track a varying number of moving targets. The sensor is located at the origin and the surveillance area is defined by the half disk with radius 1 km. A new target with initial position uniformly sampled in the surveillance area appears every 80 s. One target disappears at time 280, and another one at time 360. The full scenario runs for 400 s. One realization is illustrated in Fig. 3.

The state of an individual target  $x = [\tilde{x}^T, \omega]^T$  consists of its two-dimensional (2-D) position and velocity  $\tilde{x} = [p_1, v_1, p_2, v_2]^T$  and turn rate  $\omega$ . Each target follows a coordinated turn model [34],  $p(x_k | x_{k-1}) = \mathcal{N}(x_k; F_{CT}(\omega_{k-1})x_{k-1}, Q)$ , where the state transition matrix is given by

$$F_{CT}(\omega) = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & -\frac{1-\cos(\omega T)}{\omega} & 0 \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) & 0 \\ 0 & \frac{1-\cos(\omega T)}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} & 0 \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (30)$$

and the covariance matrix is given by

$$\begin{bmatrix} \sigma_w^2 G G^T & 0 \\ 0 & \sigma_u^2 \end{bmatrix}, \quad G = I_2 \otimes \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix} \quad (31)$$

where  $I_2$  is the  $2 \times 2$  identity matrix and  $\otimes$  is the Kronecker product. The sampling period is  $T = 1$  s,  $\sigma_w = 0.01$  m/s<sup>2</sup> is the standard deviation of the acceleration noise, and  $\sigma_u = 0.01\pi/180$  rad/s<sup>2</sup> is the standard deviation of the turn rate

noise. The survival probability is assumed constant  $p_s = 0.99$  for each target.

The sensor's FoV is limited, both in angle and range. The beamwidth is  $\theta_{BW} = \pi/6$  rad centered around  $s_k$  and the maximum sensing range is  $r_{\max} = 1$  km. The center of the beam defines the state of the sensor, which is controlled as

$$s_{k+1} = g(s_k, u_k) = u_k \quad (32)$$

*i.e.*, the beam can be moved instantaneously from any pointing direction to any other. The set of admissible control inputs, or pointing directions, is defined as

$$\mathcal{U} = \left\{ \frac{n\pi}{180} \text{ rad}, n \in \{0, 1, \dots, 180\} \right\}. \quad (33)$$

The probability of detection is 0.9 for targets within the sensor's FoV and zero elsewhere. A detected target produces a noisy range-bearing measurement  $z$  with likelihood  $p(z | x) = \mathcal{N}(z; h(x), R)$ , where

$$h(x) = \begin{bmatrix} \sqrt{p_1^2 + p_2^2} \\ \arctan\left(\frac{p_2}{p_1}\right) \end{bmatrix} \text{ and } R = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix} \quad (34)$$

with  $\sigma_r = 5$  m and  $\sigma_\theta = \pi/180$  rad. The Poisson clutter is uniformly distributed in the FoV with 0.5 expected false alarms per time step.

The Poisson birth intensity is a Gaussian mixture tuned to approximate a uniform intensity in the surveillance region. It is given by  $\lambda^b(x) = \sum_{i=1}^{N_b} w^{b,i} \mathcal{N}(x; x^{b,i}, P^{b,i})$ , with  $N_b = 40$ ,  $w^{b,i} = 3.125 \times 10^{-4}$  (corresponding to a new target appearing every 80 s), and  $P^{b,i} = \text{diag}([100, 2, 100, 2, \pi/180])^2$ . The mean of the Gaussian components are  $x^{b,i} = [p_1^{b,i}, 0, p_2^{b,i}, 0, 0]^T$ , where the position components are selected as the grid points of  $p \in \{-900, -700, \dots, 900\} \times \{100, 300, \dots, 900\}$  that satisfy  $\|p\|_2 \leq r_{\max}$ . This gives an unnormalized Gaussian mixture intensity of undetected targets, as illustrated in Fig. 3.

Since the beam can be shifted instantaneously between any pointing directions, the planning algorithm only decides one control input at a time, *i.e.*, the planning horizon is  $N = 1$ . This means that the planning problem (24) can be solved in a reasonable time using an exhaustive search in each iteration of the receding horizon control loop. For the GOSPA calculations, we use the maximum allowable localization error  $c = 50$ , which gives the weight parameter in the ST case  $\eta = 50^2/2 = 1250$ . The weight matrix  $W$  in  $\ell_{\text{track}}(\cdot)$  extracts the position components of the covariance matrices.

Fig. 4 shows the performance measures at each time step averaged over 100 Monte Carlo runs. The results confirm that the proposed stage cost function ST achieves better overall performance than SO and TO in terms of the GOSPA metric. The TO case gives accurate state estimates of the detected targets, as indicated by the normalized localization error, but at the expense of a higher number of missed targets. The ST case performs better than the SO case in terms of missed targets, which might be surprising. The reason is that in the SO case, the sensor is pointed toward



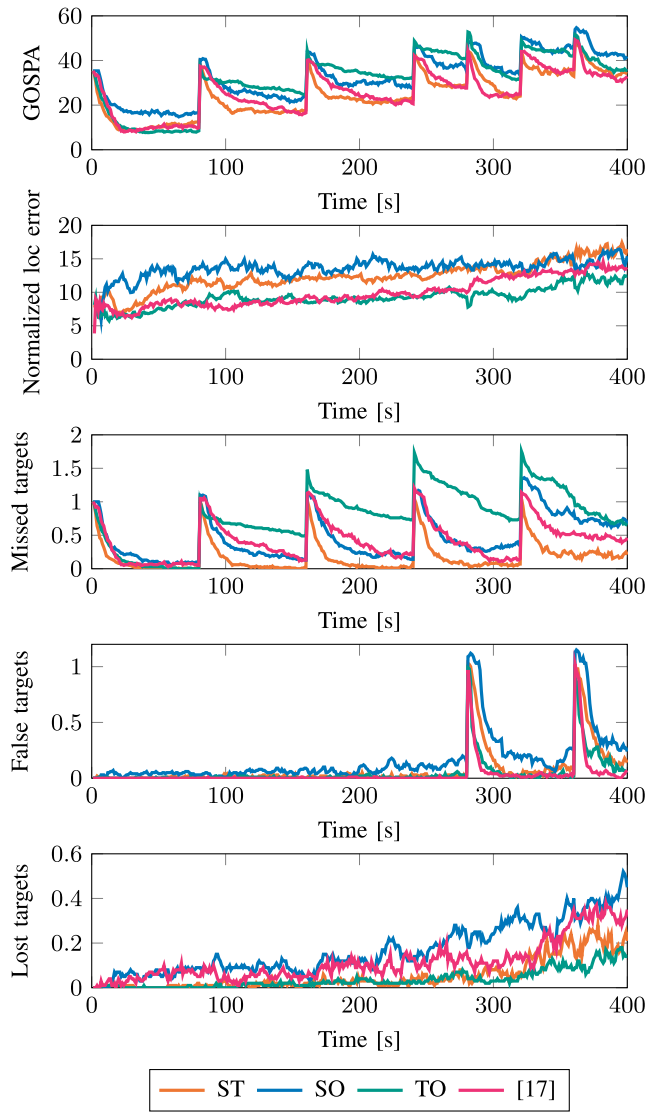


Fig. 4. Comparison of performance measures in the beam pointing scenario for the cases of planning based on search and track (ST), search only (SO), track only (TO), and the method proposed in [17].

areas with a high intensity of undetected targets, which makes it more prone to lose track of targets that it has already discovered. There is no significant difference between the three cases in terms of the number of false targets, but TO is marginally faster than the other cases in terms of removing terminated targets. The results also indicate that ST and TO perform equally well in avoiding losing track of targets. The simulations indicate that the proposed method detects new targets faster than [17], while [17] is faster in terms of removing terminated targets. This indicates that [17] focuses more on tracking existing targets than searching for new ones, compared to ST. Both methods yield similar performance in terms of GOSPA metric.

### C. Trajectory Planning for a Mobile Sensor

The second scenario runs for 3000 s and includes two targets in a surveillance area defined by a rectangle of 1500 m  $\times$  750 m. The trajectory of a mobile sensor with

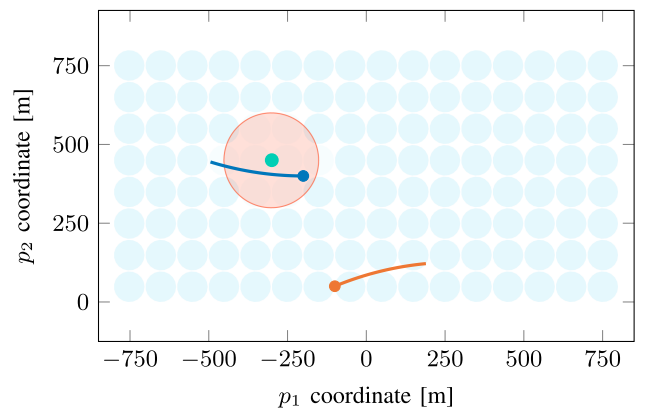


Fig. 5. True target trajectories and initial sensor position with the FoV shown in orange. The targets' initial positions are marked with filled circles. The light blue circles illustrate the Gaussian mixture intensity of the undetected targets.

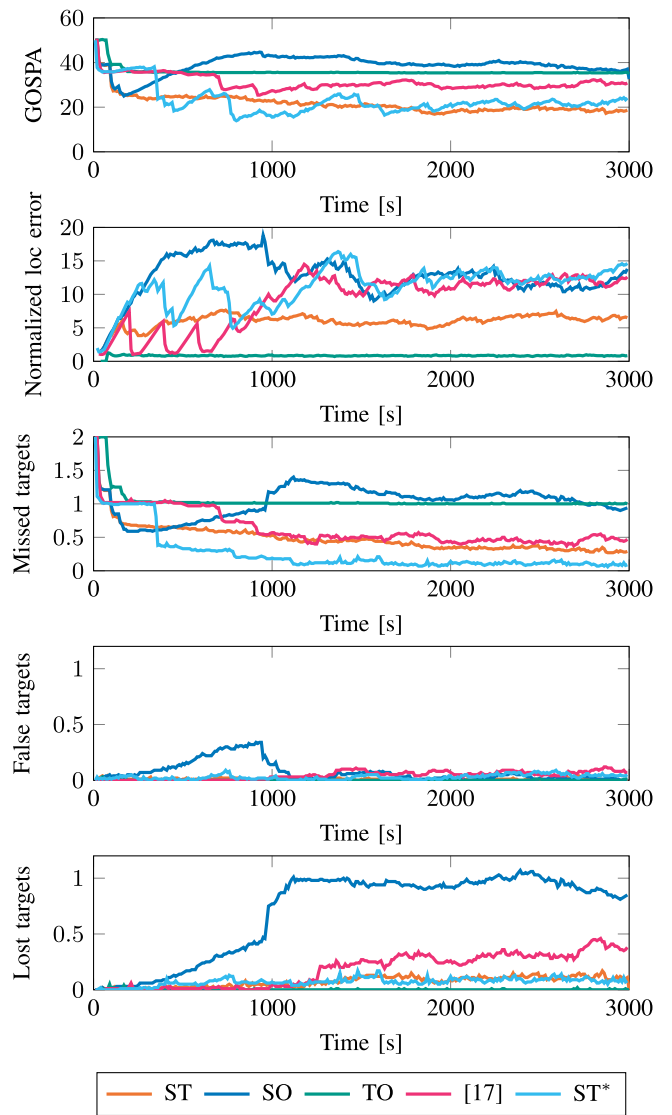


Fig. 6. Comparison of performance measures in the trajectory planning scenario for the cases of planning based on search and track (ST), search only (SO), track only (TO), the method proposed in [17], and search and track with a smaller set of admissible control sequences (ST\*).

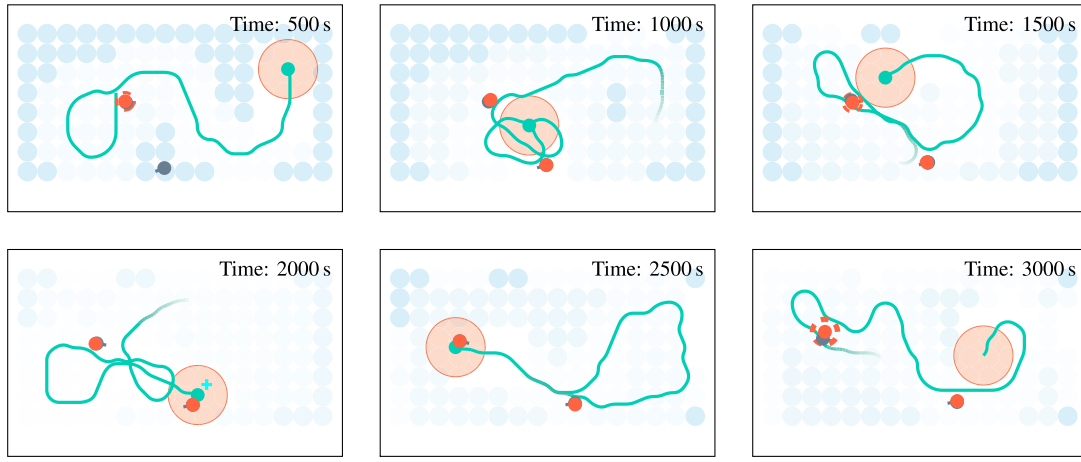


Fig. 7. Snapshots of a simulated scenario where the sensor trajectory (green path) is planned using the proposed method for joint search and track. The shaded orange area is the sensor's FoV, filled gray circles are true target positions, filled orange circles are estimated target positions, dashed orange lines indicate covariances, and light blue circles represent the intensity of undetected targets.

limited sensing range is controlled to discover and track the targets. As illustrated in Fig. 5, one of the targets is outside the sensor's sensing range at the start of the scenario and the sensor has to move away from the first target in order to discover the second one.

The state of an individual target  $x = [p_1, v_1, p_2, v_2]^T$  consists of its 2-D position and velocity and its dynamics are modeled by a nearly constant velocity model with discrete white noise acceleration. The transition density is given by  $p(x_k | x_{k-1}) = \mathcal{N}(x_k; F_{cv}x_{k-1}, Q)$ , where

$$F_{cv} = I_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad Q = \sigma_w^2 G G^T, \quad G = I_2 \otimes \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix} \quad (35)$$

where  $T = 10$  s and  $\sigma_w = 0.05$  m/s<sup>2</sup>. The survival probability is assumed constant  $p_s = 0.99$  for each target.

The Poisson birth intensity is given by  $\lambda^b(x) = \sum_{i=1}^{N_b} w^{b,i} \mathcal{N}(x; x^{b,i}, P^{b,i})$ , with  $N_b = 128$ ,  $w^{b,i} = 5.2 \times 10^{-5}$  (corresponding to a new target appearing every 1500 s), and  $P^{b,i} = \text{diag}([50, 2, 50, 2])^2$ . The mean of the Gaussian components are  $x^{b,i} = [p_1^{b,i}, 0, p_2^{b,i}, 0]^T$ , where the position components defines the grid  $p \in \{-750, -650, \dots, 750\} \times \{50, 150, \dots, 750\}$ . This gives a Gaussian mixture intensity of undetected targets, as illustrated in Fig. 5.

The sensor has a limited sensing range of  $r_{\max} = 150$  m. For targets within sensing range, the probability of detection is 0.9. Each detection results in a linear measurement of the position of the corresponding target. The measurement likelihood has Gaussian density  $p(z|x) = \mathcal{N}(z; Hx, R)$  with parameters

$$H = I_2 \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad R = I_2. \quad (36)$$

The Poisson clutter is uniformly distributed in the FoV with five expected false alarms per time step.

The mobile sensor moves according to differential drive dynamics with constant speed of 5 m/s and the control input, which is updated every 10 s, determines the heading rate.

For the planning problem, the set of admissible control inputs is given by  $\mathcal{U} = \{\omega \text{ rad/s}, \omega \in \{0, \pm\frac{\pi}{8}, \pm\frac{\pi}{4}\}\}$  and the planning horizon is  $N = 10$  steps. This gives a total of  $|\mathcal{U}|^N \approx 10^7$  possible sequences of control inputs to evaluate. Hence, the problem is too large to be solved using an exhaustive search and instead the MCTS method described in Section V is applied. The implemented MCTS search alternates between two different rollout policies. The first one selects control inputs  $u \in \mathcal{U}$  at random. The second one is a heuristic policy that recursively selects the control input  $u \in \mathcal{U}$  that minimizes the distance between the mobile sensor and the target with the most uncertain state estimate. A maximum allowable localization error of  $c = 50$  m is used for the GOSPA calculations, which gives the weight parameter in the ST case  $\eta = 50^2/2 = 1250$ . The weight matrix  $W$  in  $\ell^{\text{track}}(\cdot)$  extracts the position components of the covariance matrices. Based on trial and error, the MCTS exploration parameter is  $\varepsilon = 2500$  and the number of iterations is  $N_{\text{iter}} = 100$ .

The objective function used in [17] does not involve any tuning parameters but requires an exhaustive search to be performed. For this reason, a smaller set of admissible control sequences or maneuvers  $\mathcal{U}^*$  is constructed. Each maneuver  $u \in \mathcal{U}^*$  lasts for 120 s and consists of a course change selected from the set  $\{n\pi/9 \text{ rad}, n \in \{-9, -8, \dots, 9\}\}$  performed with turn rate  $\pi/40$  rad/s followed by a straight path. For a fair comparison, the proposed method for search and track is evaluated also using the smaller set of maneuvers  $\mathcal{U}^*$ . This combination of planning method and set of maneuvers is denoted ST\*.

Fig. 6 shows the results averaged over 100 Monte Carlo runs. It is clear that ST performs better than both SO and TO

in terms of the GOSPA. With TO, the sensor remains close to the blue target in Fig. 5 and never discovers the orange target. Thus, the normalized localization error is low for TO and the target is not lost, but one target is never detected. SO results in more false targets than the other cases. As this case focuses on discovering new targets, the time between measurement updates for the existing tracks can be quite large. When the estimation error for a certain target becomes greater than the GOSPA cutoff distance  $c = 50$  m, it is counted both as a false target and a lost target, which is seen between time 400 and 1000 s. At around time 1000 s, the existence probability of the first target decreases to below the estimation threshold  $\tau$ , after which it is removed and no longer counted as a false target. The simulation results indicate that in this scenario, the proposed method ST\* performs better than [17] in terms of the GOSPA and that both methods result in similar normalized localization errors.

Fig. 7 shows snapshots from one realization where the ST case is used to plan the sensor trajectory. The figures indicate that the proposed objective function yields the desired behavior: the sensor searches for new targets in areas with high intensity of undetected targets and maintains accurate state estimates of the discovered targets.

## VII. CONCLUSION

This article has considered the problem of jointly searching and tracking an unknown and varying number of targets. The problem was formulated as a multiobjective stochastic optimal control problem, where the conflicting objectives of search and track are competing for the same sensor resources. A method was designed to minimize both the uncertainty in the state estimates of discovered targets and the number of undiscovered targets, with a user-defined parameter determining the relative importance of the two objectives.

The PMBM filter was proposed as underlying framework when solving the optimization problem, as it provides a theoretical formulation of the multitarget tracking problem that comprises a notion of both known and undiscovered targets. As such, it not only supports the proposed method, but also provides a solid theoretical foundation for the previously proposed similar approaches that treated the two separately.

In two simulation studies, the proposed method has been illustrated and shown to provide reasonable results. The proposed method is promising for future full-scale experiments.

## APPENDIX A PMBM FILTER RECURSION

The PMBM filter recursion are here given in the style of [24]. The derivation of the filter equations and more details can be found in [8] and [25].

### A. Prediction

Given a posterior PMBM density with parameters

$$\lambda^u, \{w^j, \{r^{j,i}, p^{j,i}\}_{i \in \mathbb{I}^j}\}_{j \in \mathbb{J}} \quad (37)$$

and the standard dynamic model (see Section III-B) with single-target transition density  $p_{k+1,k}$ , the predicted density is on the same form with parameters

$$\lambda_+^u, \{w_+^j, \{r_+^{j,i}, p_+^{j,i}\}_{i \in \mathbb{I}_+^j}\}_{j \in \mathbb{J}_+} \quad (38)$$

where

$$\lambda_+^u(x) = \lambda^b(x) + \langle \lambda^u, p_s p_{k+1,k} \rangle \quad (39a)$$

$$w_+^j = w^j \quad (39b)$$

$$r_+^{j,i} = \langle p^{j,i}, p_s \rangle r^{j,i} \quad (39c)$$

$$p_+^{j,i}(x) = \frac{\langle p^{j,i}, p_s p_{k+1,k} \rangle}{\langle p^{j,i}, p_s \rangle} \quad (39d)$$

and  $w_+^j = w^j$ ,  $\mathbb{I}_+^j = \mathbb{I}^j$ , and  $\mathbb{J}_+ = \mathbb{J}$ .

### B. Measurement Update

1) *Data Association*: As the true origins of measurements in are unknown, association hypotheses are required. Let  $\mathbb{M}$  be an index set for the elements of the measurement set  $Z$ , i.e.,

$$Z = \{z^m\}_{m \in \mathbb{M}} \quad (40)$$

and let  $\mathcal{A}^j$  be a collection of all possible association hypotheses  $A$  for the  $j$ th global hypothesis, i.e., the  $j$ th MB, of which the targets are indexed by  $\mathbb{I}^j$ . Then, an association hypothesis  $A \in \mathcal{A}^j$  is a partition of  $\mathbb{M} \cup \mathbb{I}^j$  into nonempty disjoint subsets  $C \in A$ , called index cells [24].

The standard assumptions in multitarget tracking that the targets are independent of each other implies that an index cell contains at most one target index and at most one measurement index, i.e.,  $|C \cap \mathbb{I}^j| \leq 1$  and  $|C \cap \mathbb{M}| \leq 1$  for all  $C \in A \in \mathcal{A}^j$ . In the following, let  $i_C$  and  $m_C$  denote the target and measurement indices corresponding to index cell  $C$ .

2) *Update Equations*: Given a prior PMBM density with parameters according to (38), a set of measurements  $Z$ , and the standard measurement model (see Section III-B) with measurement likelihood and probability of missed detection given by

$$l_z(z|x) = p_D(x)p(z|x) \quad (41a)$$

$$q_D(x) = 1 - p_D(x) \quad (41b)$$

the updated density is a PMBM density given by

$$\pi^{\text{PMBM}}(X|Z) = \sum_{X^u \cup X^d = X} \pi^P(X^u) \pi^{\text{MBM}}(X^d) \quad (42a)$$

$$\pi^P(X^u) = e^{-(q_D \lambda_+^u, 1)} \prod_{x \in X^u} q_D(x) \lambda_+^u(x) \quad (42b)$$

$$\pi^{\text{MBM}}(X^d) = \sum_{j \in \mathbb{J}_+} \sum_{A \in \mathcal{A}^j} w_A^j \pi_A^j(X^d) \quad (42c)$$

$$\pi_A^j(X^d) = \sum_{\cup_{C \in A} X^{i_C} = X^d} \prod_{C \in A} \pi_C^j(X^{i_C}) \quad (42d)$$

where the weights are given by

$$w_A^j = \frac{w_+^j \prod_{C \in A} \mathcal{L}_C}{\sum_{j \in \mathbb{J}} \sum_{A \in \mathcal{A}^j} w_+^j \prod_{C \in A} \mathcal{L}_C} \quad (43a)$$

$$\mathcal{L}_C = \begin{cases} \lambda_+^u(z^{mC} | \cdot) + \langle \lambda_+^u, l_z(z^{mC} | \cdot) \rangle, & \text{if } C \cap \mathbb{I}^j = \emptyset, C \cap \mathbb{M} \neq \emptyset \\ r_+^{j,iC} \langle p_+^{j,iC}, l_z(z^{mC} | \cdot) \rangle, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} \neq \emptyset \\ 1 - r_+^{j,iC} + r_+^{j,iC} \langle p_+^{j,iC}, q_D \rangle, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} = \emptyset \end{cases} \quad (43b)$$

and the densities  $\pi_C^j(X)$  are Bernoulli densities with parameters

$$r_C^j = \begin{cases} \frac{\langle \lambda_+^u, l_z(z^{mC} | \cdot) \rangle}{\lambda_+^u(z^{mC}) + \langle \lambda_+^u, l_z(z^{mC} | \cdot) \rangle}, & \text{if } C \cap \mathbb{I}^j = \emptyset, C \cap \mathbb{M} \neq \emptyset \\ 1, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} \neq \emptyset \\ \frac{r_+^{j,iC} \langle p_+^{j,iC}, q_D \rangle}{1 - r_+^{j,iC} + r_+^{j,iC} \langle p_+^{j,iC}, q_D \rangle}, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} = \emptyset \end{cases} \quad (44a)$$

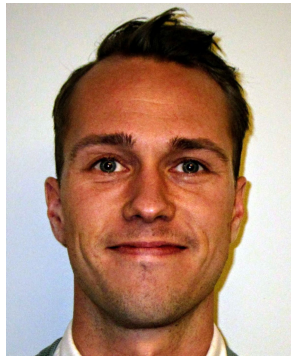
$$p_C^j(x) = \begin{cases} \frac{l_z(z^{mC} | x) \lambda_+^u(x)}{\langle \lambda_+^u, l_z(z^{mC} | \cdot) \rangle}, & \text{if } C \cap \mathbb{I}^j = \emptyset, C \cap \mathbb{M} \neq \emptyset \\ \frac{l_z(z^{mC} | x) p_+^{j,iC}(x)}{\langle p_+^{j,iC}, l_z(z^{mC} | \cdot) \rangle}, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} \neq \emptyset \\ \frac{q_D(x) p_+^{j,iC}(x)}{\langle p_+^{j,iC}, q_D \rangle}, & \text{if } C \cap \mathbb{I}^j \neq \emptyset, C \cap \mathbb{M} = \emptyset. \end{cases} \quad (44b)$$

## REFERENCES

- [1] S. Blackman and R. Popoli  
*Design and Analysis of Modern Tracking Systems*. Norwood, MA, USA: Artech House, 1999.
- [2] Y. Bar-Shalom, P. K. Willett, and X. Tian  
*Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT, USA: YBS publishing, 2011.
- [3] R. Mahler  
*Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.
- [4] R. Mahler  
*Advances in statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2014.
- [5] R. Mahler  
Multitarget Bayes filtering via first-order multitarget moments  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.
- [6] B.-T. Vo and B.-N. Vo  
Labeled random finite sets and multi-object conjugate priors  
*IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3460–3475, Jul. 2013.
- [7] B.-N. Vo, B.-T. Vo, and D. Phung  
Labeled random finite sets and the Bayes multi-target tracking filter  
*IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6554–6567, Dec. 2014.
- [8] J. L. Williams  
Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBer  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1664–1687, Jul. 2015.
- [9] K. Granström, L. Svensson, Y. Xia, J. Williams, and Á. F. García-Fernández  
Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories  
*In Proc. 21st Int. Conf. Inf. Fusion*, Cambridge, U.K., 2018, pp. 1–5.
- [10] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam  
Void probabilities and Cauchy-Schwarz divergence for generalized labeled multi-Bernoulli models  
*IEEE Trans. Signal Process.*, vol. 65, no. 19, pp. 5047–5061, Oct. 2017.
- [11] H. G. Hoang and B. T. Vo  
Sensor management for multi-target tracking via multi-Bernoulli filtering  
*Automatica*, vol. 50, no. 4, pp. 1135–1142, 2014.
- [12] B. Ristic and B.-N. Vo  
Sensor control for multi-object state-space estimation using random finite sets  
*Automatica*, vol. 46, no. 11, pp. 1812–1818, 2010.
- [13] H. Van Nguyen, H. Rezaatofghi, B.-N. Vo, and D. C. Ranasinghe  
Online UAV path planning for joint detection and tracking of multiple radio-tagged objects  
*IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5365–5379, Oct. 2019.
- [14] J. Olofsson, G. Hendeby, T. R. Lauknes, and T. A. Johansen  
Multi-agent informed path planning using the probability hypothesis density  
*Auton. Robots*, vol. 44, pp. 913–925, 2020.
- [15] P. M. Dames  
Distributed multi-target search and tracking using the PHD filter  
*Auton. Robots*, vol. 44, no. 3, pp. 673–689, 2020.
- [16] B. Charrow, N. Michael, and V. Kumar  
Active control strategies for discovering and localizing devices with range-only sensors  
*In Algorithmic Foundations of Robotics XI*, H. L. Akin, N. M. Amato, V. Isler, and A. F. van der Stappen, Eds. Cham, Switzerland: Springer, 2015, pp. 55–71.
- [17] H. Van Nguyen, H. Rezaatofghi, B.-N. Vo, and D. C. Ranasinghe  
Multi-objective multi-agent planning for jointly discovering and tracking mobile objects  
*In Proc. 34th AAAI Conf. Artif. Intell.*, New York, NY, USA, 2020, pp. 7227–7235.
- [18] D. P. Bertsekas  
*Dynamic Programming and Optimal Control*, vol. 1. Belmont, MA, USA: Athena Scientific 2005, vol. 1.
- [19] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra  
Planning and acting in partially observable stochastic domains  
*Artif. Intell.*, vol. 101, no. 1/2, pp. 99–134, 1998.
- [20] D. A. Castañón and L. Carin  
Stochastic control theory for sensor management  
*In Foundations and Applications of Sensor Management*, A. O. Hero, D. A. Castañón, D. Cochran, and K. Kastella, Eds. New York, NY, USA: Springer, 2008, pp. 7–32.
- [21] P. Boström-Rost, D. Axehill, and G. Hendeby  
Informative path planning for active tracking of agile targets  
*In Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, 2019, pp. 1–11.
- [22] J. M. Maciejowski  
*Predictive Control: With Constraints*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.
- [23] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina  
A tutorial on Bernoulli filters: Theory, implementation and applications  
*IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3406–3430, Jul. 2013.
- [24] K. Granström, M. Fatemi, and L. Svensson  
Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 208–225, Feb. 2020.
- [25] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson  
Poisson multi-Bernoulli mixture filter: Direct derivation and implementation  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.
- [26] A. H. Jazwinski  
*Stochastic Processes and Filtering Theory*. New York, NY, USA: Academic Press, 1970.



- [27] R. Mahler  
Multitarget sensor management of dispersed mobile sensors  
In *Theory and Algorithms for Cooperative Systems*, D. Grundel, R. Murphey, and P. Pardalos, Eds. Singapore: World Scientific Publishing Co, 2004, ch. 12, pp. 239–310.
- [28] C. Yang, L. Kaplan, and E. Blasch  
Performance measures of covariance and information matrices in resource management for target state estimation  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2594–2613, Jul. 2012.
- [29] C. B. Browne  
*et al.*, A survey of Monte Carlo tree search methods  
*IEEE Trans. Comput. Intell. AI Games*, vol. 4, no. 1, pp. 1–43, Mar. 2012.
- [30] L. Kocsis and C. Szepesvári  
Bandit based Monte-Carlo planning  
In *Proc. Eur. Conf. Mach. Learn.*, Berlin, Germany, 2006, pp. 282–293.
- [31] S. James, G. Konidaris, and B. Rosman  
An analysis of Monte Carlo tree search  
In *Proc. 31st AAAI Conf. Artif. Intell.*, San Francisco, CA, USA, 2017, pp. 3576–3582.
- [32] B.-N. Vo, C. M. See, N. Ma, and W. T. Ng  
Multi-sensor joint detection and tracking with the Bernoulli filter  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1385–1402, Apr. 2012.
- [33] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson  
Generalized optimal sub-pattern assignment metric  
In *Proc. 20th Int. Conf. Inf. Fusion*, Xi'an, China, 2017, pp. 1–8.
- [34] X. R. Li and V. P. Jilkov  
Survey of maneuvering target tracking. Part I: Dynamic models  
*IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.



**Per Boström-Rost** received the M.Sc. degree in applied physics and electrical engineering, in 2013 and the Lic.Eng. degree in electrical engineering, in 2019, both from Linköping University, Linköping, Sweden, where he is currently working toward the Ph.D. degree with the Division of Automatic Control, Department of Electrical Engineering.

His main research interests include sensor fusion and sensor management with applications to target tracking.



**Daniel Axehill** (Senior Member, IEEE) received the M.Sc. degree in applied physics and electrical engineering, the Lic.Eng. degree in automatic control and the Ph.D. degree in automatic control, all from Linköping University, Linköping, Sweden, in 2003, 2005, and 2008, respectively.

In year 2006, he spent three months with the University of California, Los Angeles, Los Angeles, CA, USA. From January 2009 and to November 2010, he was a Postdoctoral Researcher with the Automatic Control Laboratory, ETH Zurich, Zurich, Switzerland. He is currently an Associate Professor with the Division of Automatic Control, Linköping University. His research interests include optimization, optimal control, motion planning, hybrid systems, and applications of control.



**Gustaf Hendeby** (Senior Member, IEEE) received the M.Sc. degree in applied physics and electrical engineering and the Ph.D. degree in automatic control, both from Linköping University, Linköping, Sweden, in 2002 and 2008, respectively.

He is currently an Associate Professor with the Division of Automatic Control, Department of Electrical Engineering, Linköping University. He worked as a Senior Researcher with the German Research Center for Artificial Intelligence in 2009–2011, and a Senior Scientist with Swedish Defense Research Agency and an Adjunct Associate Professor with Linköping University, in 2011–2015. His main research interests include stochastic signal processing and sensor fusion with applications to nonlinear problems, target tracking, and simultaneous localization and mapping. He has experience of both theoretical analysis as well as implementation aspects.