

# Information-Seeking Sensor Selection for Ocean-of-Things

Augustin A. Saucan<sup>1</sup>, *Member, IEEE*, and Moe Z. Win<sup>2</sup>, *Fellow, IEEE*

**Abstract**—We propose a general sensor selection (SS) methodology for ocean-of-things (OoT) where a sensing network performs multiobject tracking (MOT) under resource constraints. SS methods address the combinatorial problem of determining the best subset of sensors that maximizes a suitable reward function for a fixed cardinality. The novelty of this article is twofold. First, we propose a tractable information-theoretic reward function for MOT-OoT with an unknown and time-varying number of objects such as ocean vessels. A tractable reward function is essential in order to rapidly evaluate a sensor subset, which is crucial in the high-dimensional problems encountered in OoT. Second, we propose a general cross-entropy SS (CE-SS) methodology that efficiently estimates the probabilities of sensor activations and determines the optimal sensor subset according to the proposed reward function and under the imposed cardinality constraint. The CE-SS algorithm avoids exhaustive searching over the space of all sensor subsets, which is intractable for most OoT applications. The CE-SS methodology, coupled with the proposed reward function, is capable of selecting sensors that lead to more accurate estimates than random selection for both the number of vessels and their trajectories. We demonstrate the effectiveness of our method via numerical simulation in several scenarios, including multivessel tracking for OoT with an emulated network of acoustic sensors deployed off the coast of Italy.

**Index Terms**—Bayesian inference, Cauchy–Schwarz divergence, Internet-of-Things, multiobject tracking, multi-Bernoulli filter, Ocean-of-Things, sensor selection.

## I. INTRODUCTION

OCEAN-OF-THINGS (OoT) is an emerging vision for an extensive network of small, low-cost floating devices (floats) that aim to provide persistent maritime monitoring and situational awareness for oceans [1]. OoT aims to integrate thousands of heterogeneous floats equipped with diverse sensing capabilities into a sensing network, while providing access to a selected subset of the collected data in order to perform monitoring tasks. Thus, OoT represents an extension of the Internet-of-Things concept [2] to maritime applications.

Manuscript received December 3, 2019; revised February 29, 2020; accepted April 12, 2020. Date of publication May 6, 2020; date of current version October 9, 2020. The fundamental research described in this paper was supported in part by the Defense Advanced Research Projects Agency under Contract N66001-19-C-4004 and in part by the Army Research Office through the MIT Institute for Soldier Nanotechnologies under Contract W911NF-13-D-0001. (*Corresponding author: Moe Z. Win.*)

Augustin A. Saucan is with the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: asaucan@mit.edu).

Moe Z. Win is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

Digital Object Identifier 10.1109/JIOT.2020.2992509

One such important application is multiobject tracking (MOT), where the objective is to sequentially estimate the number and the states of several objects of interest.

Point processes [3]–[11], and labeled random finite sets (LRFSSs) more specifically, have emerged as a unified framework for Bayesian MOT that incorporate uncertainties in both object states and their number [12]. The resulting MOT algorithms are capable of detecting objects and estimating their entire trajectories throughout a region of interest from noisy observations. At each sensor, a nonideal detection process is considered; that is, an object could be misdetections and additional false-alarm measurements (referred to as clutter) are generated. Upon detection, a single measurement per object is obtained. Sensor networks [13]–[23], and contextual or other soft information [24], can be employed to improve tracking. An additional challenge to misdetections and clutter is measurement origin uncertainty (MOU), i.e., the associations between objects and measurements at each sensor are unknown [25]. Various solutions for tracking with MOU have been proposed in the past, from which the labeled multi-Bernoulli (LMB) model [22], [26] has emerged as a highly efficient algorithm.

A challenge specific to OoT is the selection of a subset of  $M$  sensing devices out of a total of  $N$  devices in the network, referred to as the sensor selection (SS) problem [27]. Resource constraints, such as the limited battery power of the floats, and the constrained communication channel between the floats and the centralized fusion node drive the need for SS with  $M < N$ . Various SS solutions were proposed in the past. The authors of [27] and [28] leveraged the linear and Gaussian state-space statistics of a single object of interest to develop highly efficient algorithms based on convex programming and tree pruning techniques. However, in our MOT for OoT application, multiple objects need to be tracked and the observation equations of the float sensors are nonlinear. For example, range values are measured by floats equipped with single hydrophones, while angles of arrival are observed in the case of floats equipped with hydrophone arrays. Based on minimizing the posterior covariance with an added sparsity-promoting term for the set of active sensors, a sensor-scheduling algorithm for tracking a single object was proposed in [29]. A stochastic sensor-scheduling algorithm for linear dynamic systems was proposed in [30] by finding the optimal sensor activation probabilities via quasiconvex programming. Of related interest are sensor control strategies for target tracking. Control strategies for unmanned-aerial-vehicles performing object tracking were reported in [31], by

relying on a discretized action space and a covariance measure for the cost function. In [32], a team of robots is maneuvered to track a single object under a nonlinear observation model via nonconvex programming.

Sensor selection or, more generally, resource allocation and sensor control with LRFSSs for MOT, has been successfully employed in [33]–[38]. Related topics on selecting multi-paths, antennas, measurements, and relays can be found in [39]–[43]. In [33], the Rényi divergence was employed to measure information gain for sensor control. The Cauchy–Schwarz (CS) divergence between two Poisson point processes was shown to have a tractable expression in [36]. In contrast, for generalized LMB models, the CS divergence is given in [37] and is shown to involve a weighted sum over all subsets of the label space, which is intractable and, in practice requires truncations. In [38], the LMB model was employed for the trajectory planning of an unmanned aerial vehicle that performs MOT. In [44], a second-order uncertainty measure was constructed for multi-Bernoulli posterior densities and is employed for sensor control and management. Most previous works resort to a discretized control space and direct maximization via exhaustive enumeration of all control actions or subset selection combinations. This is intractable in OoT applications where the exhaustive enumeration of all sensor subsets is unattainable. Relay selection and power allocation for underwater communication networks are studied in [45], whereas relay placement and flow allocation are addressed in [46]. Node clustering for energy management is addressed for the Internet of Underwater Things in [47].

In this article, we propose an efficient SS method for MOT in OoT. The novelty of this article is twofold. First, the unknown and time-varying number of objects are modeled as an LMB model and we derive a tractable expression for the CS divergence between the predicted LMB and posterior LMB models in order to quantify the information gain for a specific sensor subset. An advantage of our derived expression for the CS divergence is that it does not involve summations over subsets of labels as in the case of [37], and it only involves integration over the single target space. Second, we propose a cross-entropy (CE) algorithm that searches for the optimum sensor subset, i.e., the subset that maximizes the CS divergence for a fixed resource budget  $M < N$ . As opposed to solutions that explicitly model energy consumption, e.g., [45]–[47], here the cardinality of the sensor subset is directly imposed since the energy consumption in OoT is dominated by the satellite-uplink procedure and is nearly identical for all floats. The CE algorithm is an efficient stochastic search method with provable convergence that avoids the intractability of the brute-force methods used in the past. This is especially relevant for OoT since the search space for the SS combinatorial problem is extremely high, e.g., for  $N = 50$  and  $M = 10$  there are more than  $10^{10}$  possible combinations. The resulting CE method for SS is general and can be employed in conjunction with any MOT filter that provides a list of predicted object tracks composed of a probability density function and a probability of existence for each track. This covers not only the LMB but also a large spectrum of MOT filters, ranging from the probabilistic data association filter [25] and integrated probability data

TABLE I  
MATHEMATICAL NOTATIONS OF INTEREST

Symbol	Definition
$\mathbb{N}, \mathbb{R}$	set of natural numbers, space of reals
$\mathbb{X}$	object state space (in general $\mathbb{X} \triangleq \mathbb{R}^d$ )
$\mathbb{L}$	space of object labels, i.e., identifiers
$\mathbb{Z}_i$	measurement space of the $i$ -th sensor
$x, x$	scalar random variable and its realization
$\mathbf{x}, \mathbf{x}$	random vector and its realization
$\mathbb{X}, \mathcal{X}$	random set and its realization
$ \cdot $	set cardinality (number of elements)
$\{i : j\}$	set of integers $\{i, i + 1, \dots, j - 1, j\}$
$\mathbf{u}_{i:j}$	the subvector $[u_i, \dots, u_j]^T$
$\delta(\cdot)$	Kronecker delta operator
$\mathbb{1}_{\mathcal{A}}(\cdot)$	indicator function of set $\mathcal{A}$
$\langle \cdot, \cdot \rangle$	inner product of functions on $\mathbb{X}$
$\ \cdot\ _p$	$L^p$ vector norm on $\mathbb{X}$
$\mathcal{F}(\mathbb{X})$	set of finite subsets of $\mathbb{X}$
$\mathbb{D}_{\text{KL}}\{\cdot, \cdot\}$	Kullback–Leibler divergence
$\mathbb{D}_{\text{CS}}\{\cdot, \cdot\}$	Cauchy–Schwarz divergence
$\mathcal{E}_n(\cdot)$	elementary symmetric function of order $n$

association filter [48] to the graphical model-based approaches of [22]. In this article, we exemplify our CE-SS algorithm in conjunction with a multisensor LMB-MOT filter. In addition, we provide the parallel multisensor LMB update equations, since to the best of our knowledge only the sequentially updated multisensor LMB was provided in [22, Sec. XIII.B]. We also provide a sequential Monte Carlo (SMC) implementation of the CS divergence expression in order to account for the nonlinear and/or non-Gaussian state-space equations. The advantages of our proposed method are showcased via numerical experiments for SS in realistic OoT scenarios.

The remainder of this article is organized as follows. General notation and background on LRFSS are given in Section II, while the system model is described in Section III. The multisensor LMB filter is derived in Section IV. The CS divergence for LMB models is derived in Section V along with an SMC implementation of the CE-SS reward function. The CE algorithm for SS is presented in Section VI. The results of our numerical simulations are presented in Section VII.

## II. BACKGROUND

The general notation used throughout this article is defined in Table I. Several specific definitions are given next. The Kronecker delta is extended to sets as  $\delta_{\mathcal{X}}(\mathcal{Y}) = 1$  if  $\mathcal{X} = \mathcal{Y}$  and  $\delta_{\mathcal{X}}(\mathcal{Y}) = 0$  otherwise. A set indicator function is defined as  $\mathbb{1}_{\mathcal{X}}(\mathcal{Y}) = 1$  if  $\mathcal{Y} \subseteq \mathcal{X}$  and  $\mathbb{1}_{\mathcal{X}}(\mathcal{Y}) = 0$  otherwise. For compactness,  $\mathbb{1}_{\mathcal{X}}(\mathbf{x})$  is employed in lieu of  $\mathbb{1}_{\mathcal{X}}(\{\mathbf{x}\})$ . On the vector space  $\mathbb{X}$ , the inner product of real-valued functions  $f$  and  $g$  is defined as  $\langle f, g \rangle \triangleq \int_{\mathbb{X}} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}$ . The space  $L^2(\mathbb{X})$  represents the space of real-valued functions  $f$  defined on  $\mathbb{X}$  such that  $\langle f, f \rangle$  is finite. The  $L^2$  norm for a function  $f$  is defined as  $\|f\|_2 = \sqrt{\langle f, f \rangle}$  while the  $L^0$  norm of a vector  $\mathbf{u}$  is

$\|\mathbf{u}\|_0 = |\{u_i : u_i \neq 0\}|$ . For discrete probability distributions  $p$  and  $q$  defined on the same space  $\mathcal{X}$ , the Kullback–Leibler divergence is  $\mathbb{D}_{\text{KL}}(p, q) = \sum_{x \in \mathcal{X}} p(x) \ln(p(x)/q(x))$ . The notation  $\mathcal{N}(\mathbf{x}; \mathbf{m}, \mathbf{M})$  represents the standard normal probability density function of variable  $\mathbf{x}$ , mean vector  $\mathbf{m}$ , and covariance matrix  $\mathbf{M}$ .

### A. LRFS Background

The label space  $\mathbb{L}$  is a countable set whose elements represent object identifiers. An unordered collection of objects is modeled by an LRFS  $\mathbb{X} \triangleq \{(\mathbf{x}^{(1)}, l^{(1)}), \dots, (\mathbf{x}^{(n)}, l^{(n)})\}^1$  [49], where the number of objects, i.e., the set cardinality  $|\mathbb{X}| = n$ , is a random variable on  $\mathbb{N}$ . Furthermore, the object states  $\{\mathbf{x}^{(i)}\}$  are random vectors on  $\mathbb{X}$  and the object labels  $\{l^{(i)}\}$  are random variables that take values in  $\mathbb{L}$ . In other words, an LRFS is a random variable taking values in  $\mathcal{F}(\mathbb{X} \times \mathbb{L})$ . For a realization  $\mathcal{X} = \{(\mathbf{x}^{(1)}, l^{(1)}), \dots, (\mathbf{x}^{(n)}, l^{(n)})\} \in \mathcal{F}(\mathbb{X} \times \mathbb{L})$  of the LRFS  $\mathbb{X}$ , the projection  $\mathcal{L}_{\mathcal{X}} = \{l^{(1)}, \dots, l^{(n)}\}$  and the distinct label indicator  $\Delta_{\mathcal{X}}$  are introduced, such that  $\Delta_{\mathcal{X}} = 1$  if  $|\mathcal{X}| = |\mathcal{L}_{\mathcal{X}}|$  and  $\Delta_{\mathcal{X}} = 0$  otherwise. Given a real-valued function  $h$  and any set  $\mathcal{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_{|\mathcal{Y}|}\}$ , the set exponential is defined as  $[h(\cdot)]^{\mathcal{Y}} = \prod_{\mathbf{y} \in \mathcal{Y}} h(\mathbf{y})$ .

Several constructs exist that characterize the probability distribution of random finite sets (RFSs), e.g., belief-mass functions and multiobject probability densities (for brevity, hereafter referred to as belief functions and belief densities, respectively) [50]–[52], void probabilities [3], and probability measures and probability densities [53]. Relevant to this article is the set integral of a function  $f: \mathcal{F}(\mathbb{X} \times \mathbb{L}) \rightarrow \mathbb{R}$ , defined as [51]

$$\int f(\mathcal{X}) \delta \mathcal{X} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(l_1, \dots, l_n) \in \mathbb{L}^n} \int_{\mathbb{X}^n} f(\{(\mathbf{x}_1, l_1), \dots, (\mathbf{x}_n, l_n)\}) d(\mathbf{x}_1, \dots, \mathbf{x}_n). \quad (1)$$

Subsequently, for any closed set  $\mathcal{A} \subset \mathbb{X} \times \mathbb{L}$  the belief function is introduced as  $\beta(\mathcal{A}) \triangleq \mathbb{P}\{\mathbb{X} \subset \mathcal{A}\}$  and the belief density is defined as any set function  $\pi: \mathcal{F}(\mathbb{X} \times \mathbb{L}) \rightarrow \mathbb{R}_+$  such that  $\beta(\mathcal{A}) = \int \pi(\mathcal{X}) \mathbb{1}_{\mathcal{A}}(\mathcal{X}) \delta \mathcal{X}$ .

### B. Labeled Bernoulli RFS

For a fixed label  $l \in \mathbb{L}$ , a labeled Bernoulli RFS  $\mathbb{X}$  is parametrized by the probability of existence  $r(l)$  and state probability density  $p(\cdot, l)$ . Furthermore,  $\mathbb{X}$  is a singleton, i.e.,  $\mathbb{X} = \{(\mathbf{x}, l)\}$ , where  $\mathbf{x}$  has the probability density function  $p(\cdot, l)$ , with probability  $r(l)$  and  $\mathbb{X} = \emptyset$  with probability  $1 - r(l)$ . Thus, the belief density can be written as

$$\pi_{\mathbb{X}}(\mathcal{X}) = \begin{cases} 1 - r(l), & \text{if } \mathcal{X} = \emptyset \\ r(l) p(\mathbf{x}, l), & \text{if } \mathcal{X} = \{(\mathbf{x}, l)\} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The corresponding cardinality distribution  $p(n) = \mathbb{P}\{|\mathbb{X}| = n\}$  follows directly from (2) and is given by the Bernoulli law with parameter  $r(l)$ . The Bernoulli LRFS of (2) can model up to a single object of interest with uncertainties in both object presence (or existence) and object position.

<sup>1</sup>The set  $\{a^{(1)}, \dots, a^{(n)}\}$ , where  $a^{(i)} \in \mathcal{A} \forall i = 1, 2, \dots, n$  and  $|\mathcal{A}| = n$ , represents an arbitrary enumeration of the elements in the set  $\mathcal{A}$ .

### C. Labeled Multi-Bernoulli RFS

An LMB is obtained as the union of independent labeled Bernoulli RFSs and is parameterized by  $\{r(l), p(\cdot, l)\}_{l \in \mathbb{L}}$ . The multitarget density of an LMB  $\mathbb{X}$  is [49]

$$\pi_{\mathbb{X}}(\mathcal{X}) = \Delta_{\mathcal{X}} w(\mathcal{L}_{\mathcal{X}}) [p(\cdot)]^{\mathcal{X}} \quad (3)$$

where  $w(\mathcal{L})$  is the probability of observing the subset of tracks  $\mathcal{L} \subseteq \mathbb{L}$  and is given by

$$w(\mathcal{L}) = \left[ \prod_{l \in \mathbb{L} \setminus \mathcal{L}} (1 - r(l)) \right] \left[ \prod_{l \in \mathcal{L}} \mathbb{1}_{\mathbb{L}}(l) r(l) \right]. \quad (4)$$

The cardinality distribution  $p(n) = \mathbb{P}\{|\mathbb{X}| = n\}$  of the LMB is given by the multi-Bernoulli law [51, p. 369]. Furthermore, if we denote  $\mathbb{L} = \{l_1, \dots, l_L\}$  and for  $n \in \{0 : L\}$ , the LMB cardinality distribution has the expression

$$p(n) = \left[ \prod_{i=1}^L (1 - r(l_i)) \right] \mathfrak{E}_n \left( \frac{r(l_1)}{1 - r(l_1)}, \dots, \frac{r(l_L)}{1 - r(l_L)} \right) \quad (5)$$

and  $p(n) = 0$  otherwise. The elementary symmetric function  $\mathfrak{E}_n(\cdot)$  of order  $n$  is defined as [51, p. 369]

$$\mathfrak{E}_n(y_1, \dots, y_L) \triangleq \sum_{\mathcal{L} \subset \{1:L\}, |\mathcal{L}|=n} \left[ \prod_{l \in \mathcal{L}} y_l \right] \quad (6)$$

and has the property

$$\sum_{n=0}^L \mathfrak{E}_n(y_1, \dots, y_L) = (1 + y_1) \cdots (1 + y_L). \quad (7)$$

For compactness,  $\mathfrak{E}_n(\mathbf{y}_{1:L})$  will also be employed in lieu of  $\mathfrak{E}_n(y_1, \dots, y_L)$ . Three remarks are in order regarding equation (3). First, the LMB of (3) models up to  $|\mathbb{L}|$  objects where the exact number of objects and their states are random. Second, the LMB belief density expression of (3) has a simpler expression than the corresponding belief density of the unlabeled multi-Bernoulli RFS of [51, p. 368, eq. (11.134)]. This simplification is generated by the usage of labels. More precisely, the correspondence between the elements of a realization  $\mathcal{X}$  of  $\mathbb{X}$  and the Bernoulli components in  $\mathbb{L}$  is immediate via the distinct labels  $\mathcal{L}_{\mathcal{X}}$ . Finally, marking Bernoulli components with distinct labels allows for object tracking across different time steps and for the inference of entire trajectories of objects in MOT filters as exemplified in the filter of Section IV.

## III. OOT SYSTEM OVERVIEW

MOT-OoT applications aim to sequentially estimate the number and the states of multiple objects of interest across time from measurements provided by an extensive heterogeneous network of sensing devices, such as sensor floats and potentially other submerged sensing devices. In the rest of this article, we will refer to a sensing device as a sensor. This is without loss of generality since multiple sensors mounted on the same device can be viewed as a single metasensor. In the sequel, we present the MOT-OoT model assumptions for the dynamical system of objects and sensors.

- A1) The number of objects is unknown and time varying and the multiobject collection is modeled as an LMB  $X_k$  at time step  $k$  and defined on the label space  $\mathbb{L}_k$ .
- A2) The objects evolve independently from time  $k$  to time  $k + 1$ . Furthermore, an object with label  $l$  and state  $\mathbf{x}_k$  at time  $k$  survives with probability  $P_{k+1}^S(\mathbf{x}_k, l)$  to time  $k + 1$ . Upon survival, the target state vector evolves according to the Markovian kernel  $f_{k+1|k}(\cdot|\mathbf{x}_k, l)$ . Independently of surviving objects, birthed objects at time  $k+1$  are modeled via an LMB with parameters  $\{(r_{B,k+1}(l), p_{B,k+1}(\cdot, l))\}_{l \in \mathbb{B}_{k+1}}$ .
- A3) The sensing network is composed of  $N$  sensors. The network can be multimodal, i.e., the sensors can have different sensing modalities, e.g., single hydrophones as range-only sensors and hydrophone arrays as range and bearing sensors.
- A4) An object with label  $l$  and state vector  $\mathbf{x}$  is detected by sensor  $i$  with probability  $P_{i,k}^D(\mathbf{x}, l)$ . Upon detection, a measurement vector  $\mathbf{z}_{i,k} \in \mathbb{Z}_i$  is generated according to a device-specific measurement model with probability density function  $g_{i,k}(\cdot|\mathbf{x}, l)$ .
- A5) Independent from the object-originated measurements, false alarm measurements are generated according to a Poisson point process model with an intensity function  $f_{i,k}^{\text{FA}}: \mathbb{Z}_i \rightarrow \mathbb{R}_+$  for each sensor  $i \in \{1 : N\}$ .
- A6) The measurement set  $\mathcal{Z}_{i,k} \triangleq \{\mathbf{z}_{i,k}^1, \dots, \mathbf{z}_{i,k}^{m_{i,k}}\}$  of sensor  $i$  with  $|\mathcal{Z}_{i,k}| = m_{i,k}$  and  $\mathbf{z}_{i,k} \in \mathbb{Z}_i$  contains both object-originated and false alarm measurements and is affected by MOU; that is, the associations between the objects of  $\mathcal{X}_k$  and the measurements of  $\mathcal{Z}_{i,k}$  are not known. A measurement is either clutter or generated by an object, while an object can generate at most one measurement per sensor.
- A7) The sensor measurement sets are conditionally independent given the object set, i.e.,  $\pi(\mathcal{Z}_{1,k}, \dots, \mathcal{Z}_{N,k} | X_k = \mathcal{X}_k) = \pi(\mathcal{Z}_{1,k} | \mathcal{X}_k) \cdots \pi(\mathcal{Z}_{N,k} | \mathcal{X}_k)$ .
- A8) The sensors communicate with a centralized fusion node that performs MOT. However, due to communication and other resource constraints, only a set of  $M (< N)$  sensors are allowed to uplink their measurements to the data cloud and subsequently to the centralized fusion node at any given time instant. The set of all valid sensor subsets for SS is defined as  $\mathbb{S}_M = \{\mathcal{S} \subset \{1 : N\} : |\mathcal{S}| = M\}$ .

In order to cope with MOU and similar to [54], [55], and [22, Sec. IV-B], the association maps  $a_{i,k}$  are introduced as

$$a_{i,k}(l) = \begin{cases} m \in \{1 : |\mathcal{Z}_{i,k}|\}, & \text{if object } l \text{ generated } \mathbf{z}_{i,k}^{(m)} \text{ at} \\ & \text{sensor } i \text{ and time } k \\ 0, & \text{if object } l \text{ was misdetected} \\ & \text{by sensor } i \text{ and time } k. \end{cases} \quad (8)$$

In light of Assumption A6), let  $\mathcal{A}_{i,k}$  denote the set of all valid association maps for sensor  $i$ , that is,  $a_{i,k} \in \mathcal{A}_{i,k}$  if  $a_{i,k}(l) > 0$  implies  $a_{i,k}(l) \neq a_{i,k}(l')$  for  $l \neq l'$ . In other

words, a valid association map  $a_{i,k} \in \mathcal{A}_{i,k}$  can be viewed as a function that either assigns a distinct measurement index from the set  $\{1 : |\mathcal{Z}_{i,k}|\}$  or a misdetection index 0 to each Bernoulli label. Combining Assumptions A4)–A7) and the association maps (8), the multiobject likelihood for sensor  $i$  can be written as [49, Proposition 7]

$$g_{i,k}(\mathcal{Z}_{i,k} | \mathcal{X}) = e^{-(f_{i,k}^{\text{FA}}, 1)} \left[ f_{i,k}^{\text{FA}}(\cdot) \right]^{\mathcal{Z}_{i,k}} \times \sum_{a_{i,k} \in \mathcal{A}_{i,k}} \delta_{a_{i,k}^{-1}(\{0 : |\mathcal{Z}_{i,k}|\})}(\mathcal{L}\mathcal{X}) \times \left[ \prod_{(\mathbf{x}, l) \in \mathcal{X}} \psi_{i,k}(\mathbf{x}, l; \mathcal{Z}_{i,k}, a_{i,k}(l)) \right] \quad (9)$$

where  $a_{i,k}^{-1}(\mathcal{B}) \triangleq \{l \in \mathbb{L}_k : a_{i,k}(l) \in \mathcal{B}\}$  is the preimage of the set  $\mathcal{B}$  under the map  $a_{i,k}$  and

$$\psi_{i,k}(\mathbf{x}, l; \mathcal{Z}_{i,k}, m) = \begin{cases} \frac{P_{i,k}^D(\mathbf{x}, l) g_{i,k}(\mathbf{z}_{i,k}^m | \mathbf{x}, l)}{f_{i,k}^{\text{FA}}(\mathbf{z}_{i,k}^m)}, & \text{if } m > 0 \\ 1 - P_{i,k}^D(\mathbf{x}, l), & \text{if } m = 0. \end{cases} \quad (10)$$

Due to the MOU of A6), the likelihood function in (9) involves a summation over all valid associations between the elements of  $\mathcal{L}\mathcal{X}$  and the measurements from sensor  $i$ . Additionally for any valid subset of sensors  $\mathcal{S} \in \mathbb{S}_M$  and by employing A7), the multisensor-likelihood function is defined as

$$g_k^{\mathcal{S}}(\mathcal{Z}_k^{\mathcal{S}} | \mathcal{X}) \triangleq \prod_{s \in \mathcal{S}} g_{s,k}(\mathcal{Z}_{s,k} | \mathcal{X}) \quad (11)$$

where the compact notation  $\mathcal{Z}_k^{\mathcal{S}} \triangleq (\mathcal{Z}_{s_1,k}, \dots, \mathcal{Z}_{s_M,k})$  for  $\mathcal{S} \triangleq \{s_1, \dots, s_M\}$  was employed. As per Assumption A8), the number of sensors employed in the multisensor-likelihood function (11) is limited to  $M$ .

Consider at time  $k$  a sequence of sensor subsets  $\mathcal{S}_{1:k} = (\mathcal{S}_1, \dots, \mathcal{S}_k)$ , where  $\mathcal{S}_i \in \mathbb{S}_M \forall i \in \{1 : k\}$ , the corresponding random measurement sets  $\mathcal{Z}_{1:k}^{\mathcal{S}_{1:k}} \triangleq (\mathcal{Z}_1^{\mathcal{S}_1}, \dots, \mathcal{Z}_k^{\mathcal{S}_k})$  and their realizations  $\mathcal{Z}_{1:k}^{\mathcal{S}_{1:k}} \triangleq (\mathcal{Z}_1^{\mathcal{S}_1}, \dots, \mathcal{Z}_k^{\mathcal{S}_k})$ . According to the above assumptions, an LMB is employed to model the collection of objects at each time step  $k$  and we aim to sequentially estimate the posterior belief density  $\pi_k^{\mathcal{S}_{1:k}}(\mathcal{X}) \triangleq \pi_{X_k}(\mathcal{X} | \mathcal{Z}_{1:k}^{\mathcal{S}_{1:k}} = \mathcal{Z}_{1:k}^{\mathcal{S}_{1:k}})$  of the LMB  $X_k$  given all previous and current measurement sets from the selected sensors. A block diagram of the proposed MOT-OoT with SS is presented in Fig. 1. The LMB filter propagates the LRFS posterior belief density through predict and update steps as described in Section IV. The predicted belief density also serves to select the optimal sensor subset  $\mathcal{S}_k$  in the sense of maximizing the CS information gain, as presented in Section V.

#### IV. MULTISENSOR LMB FILTER FOR OoT-MOT

In [26] and [22, Sec. XIII.A], the authors presented single-sensor and sequential multisensor LMB filters, where the latter is obtained by sequentially applying the single-sensor update equation for each sensor. In this section, the filtering equations for the joint multisensor LMB filter are presented. These equations predict and jointly update the LMB with the information from a subset of sensors that is selected by the CE-SS method,

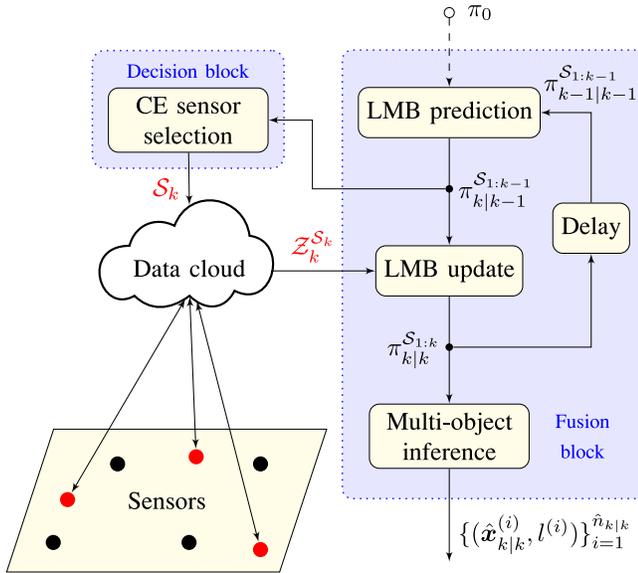


Fig. 1. Block diagram showcasing the proposed LMB filter with sensor selection. The CE-SS method serves as a decision block for selecting the sensors to be activated at the current time step  $k$  (depicted in red). The filter serves as a fusion block which employs the data from the selected sensors to update the LMB parameters. Note that the belief density  $\pi_0$  is only used to initialize the filter at time  $k = 0$ .

as shown in Fig. 1. The predict and update recursion starts from an initial LMB density  $\pi_0(\mathcal{X})$ . Furthermore, at each time  $k$ , based on the predicted density  $\pi_{k-1|k-1}^{S_{1:k-1}}(\mathcal{X})$ , the CE-SS method (presented in Section V) selects the best subset of sensors  $\mathcal{S}_k$ . Subsequently, the selected sensors send their measurement sets  $\mathcal{Z}_{i,k} \forall i \in \mathcal{S}_k$  to the data cloud and serve to perform the LMB update.

### A. Filter Prediction

The filtering recursion starts at time  $k - 1$  from an LMB with belief density  $\pi_{k-1|k-1}^{S_{1:k-1}}(\mathcal{X})$  and parameters  $\{(r_{k-1|k-1}^{S_{1:k-1}}(l), p_{k-1|k-1}^{S_{1:k-1}}(\cdot, l))\}_{l \in \mathbb{L}_{k-1}}$ . The prediction step incorporates object disappearance, the object kinematic model, and object births. Note that the prior LMB parameters also depend on the sensor subsets  $\mathcal{S}_{1:k-1}$ . Following Assumption A2), the surviving objects are modeled as an LMB with parameters  $\{(r_{S,k}^{S_{1:k-1}}(l), p_{S,k}^{S_{1:k-1}}(\mathbf{x}, l))\}_{l \in \mathbb{L}_{k-1}}$  [26] where

$$\eta_S(l) = \langle P_k^S(\cdot, l), p_{k-1|k-1}^{S_{1:k-1}}(\cdot, l) \rangle \quad (12a)$$

$$r_{S,k}^{S_{1:k-1}}(l) = r_{k-1|k-1}^{S_{1:k-1}}(l) \eta_S(l) \quad (12b)$$

$$p_{S,k}^{S_{1:k-1}}(\mathbf{x}, l) = \frac{\langle P_k^S(\cdot, l) f_{k|k-1}(\mathbf{x}|\cdot, l), p_{k-1|k-1}^{S_{1:k-1}}(\cdot, l) \rangle}{\eta_S(l)} \quad (12c)$$

As stated in Assumption A2), object births are modeled via a birth LMB, with parameters  $\{(r_{B,k}(l), p_{B,k}(\cdot, l))\}_{l \in \mathbb{B}_k}$ , that is independent of the surviving objects. The predicted LRFS is given by the union of the surviving LMB and birth LMB, and the resulting LRFS is an LMB with the following parameters  $\{(r_{k|k-1}^{S_{1:k-1}}(l), p_{k|k-1}^{S_{1:k-1}}(\cdot, l))\}_{l \in \mathbb{L}_k} = \{(r_{S,k}^{S_{1:k-1}}(l), p_{S,k}^{S_{1:k-1}}(\cdot, l))\}_{l \in \mathbb{L}_{k-1}} \cup \{(r_{B,k}(l), p_{B,k}(\cdot, l))\}_{l \in \mathbb{B}_k}$ , where

$\mathbb{L}_k = \mathbb{L}_{k-1} \cup \mathbb{B}_k$  and  $\mathbb{L}_{k-1} \cap \mathbb{B}_k = \emptyset$ . Note that label  $l \in \mathbb{B}_k$  is typically constructed as the tuple  $(k, i)$ , where  $k$  is the time sample and  $i$  is an index used to distinguish between the objects born at time  $k$ .

### B. Filter Update

The update step is achieved via Bayes' theorem for RFSs [49]

$$\pi_{k|k}^{S_{1:k}}(\mathcal{X}) = \frac{g_k^{S_k}(\mathcal{Z}_k^{S_k}|\mathcal{X}) \pi_{k-1|k-1}^{S_{1:k-1}}(\mathcal{X})}{\int g_k^{S_k}(\mathcal{Z}_k^{S_k}|\mathcal{Y}) \pi_{k-1|k-1}^{S_{1:k-1}}(\mathcal{Y}) \delta\mathcal{Y}} \quad (13)$$

However, updating the predicted Bernoulli components using (13) does not lead to an LMB density [22], [26], [55]. The augmented association map  $\bar{a}_{i,k}: \mathbb{L}_k \rightarrow \{-1 : |\mathcal{Z}_{i,k}|\}$  extends  $a_{i,k}$  to the case  $\bar{a}_{i,k}(l) = -1$  of Bernoulli component  $l$  not present (i.e., object disappearance). Let  $-1 = (-1, \dots, -1)$  be the  $M$ -tuple with all  $-1$  entries. In addition, for a sensor subset  $\mathcal{S} = \{s_1, \dots, s_M\} \in \mathbb{S}_M$  let  $\bar{a}_k^{\mathcal{S}} \triangleq (\bar{a}_{s_1,k}, \dots, \bar{a}_{s_M,k})$  be the multisensor association map. The set of extended multisensor measurement indices is denoted by  $\mathbb{M}_k^{\mathcal{S}} \triangleq \{-1\} \uplus \times_{s \in \mathcal{S}} \{0 : |\mathcal{Z}_{s,k}|\}$ , while the set of all possible multisensor association maps (valid and nonvalid) is denoted by  $\mathcal{M}_k^{\mathcal{S}} \triangleq \{\bar{a}_k^{\mathcal{S}} : \bar{a}_k^{\mathcal{S}}(l) \in \mathbb{M}_k^{\mathcal{S}} \forall l \in \mathbb{L}_k\}$ .<sup>2</sup> The set of valid multisensor association maps  $\bar{\mathcal{A}}_k^{\mathcal{S}}$  contains maps  $\bar{a}_k^{\mathcal{S}} \in \mathcal{M}_k^{\mathcal{S}}$  such that  $\bar{a}_{s,k}(l) = \bar{a}_{s,k}(l') > 0$  implies  $l = l'$  for any  $s \in \mathcal{S}$ . Note that if object  $l$  is not present, a valid augmented multisensor association map has to simultaneously map  $\bar{a}_{s,k}(l) = -1 \forall s \in \mathcal{S}$ . Note that  $\bar{\mathcal{A}}_k^{\mathcal{S}} \subseteq \mathcal{M}_k^{\mathcal{S}}$ , with equality holding only if  $|\mathbb{L}_k| = 1$ . Subsequently, for compactness the superscript  $\mathcal{S}$  for  $\bar{a}_k^{\mathcal{S}}$  is dropped since this will be evident from the definition  $\bar{a}_k \in \bar{\mathcal{A}}_k^{\mathcal{S}}$ . Recalling that  $\mathcal{Z}_k^{\mathcal{S}} \triangleq (\mathcal{Z}_{s_1,k}, \dots, \mathcal{Z}_{s_M,k})$  with  $\mathcal{S} = \{s_1, \dots, s_M\}$ , for  $\bar{a}_k(l) \neq -1$  the sensor subset pseudolikelihood is defined as

$$\bar{\psi}_k^{\mathcal{S}}(\mathbf{x}, l; \mathcal{Z}_k^{\mathcal{S}}, \bar{a}_k(l)) = \prod_{s \in \mathcal{S}} \psi_{s,k}(\mathbf{x}, l; \mathcal{Z}_{s,k}, \bar{a}_{s,k}(l)) \quad (14)$$

For each multisensor assignment index  $\mathbf{m} \in \mathbb{M}_k^{\mathcal{S}}$ , the updated Bernoulli parameters are

$$\eta_{k|k}^{\mathbf{m}}(l; \mathcal{S}_{1:k}) = \begin{cases} r_{k|k-1}^{S_{1:k-1}}(l) \langle \bar{\psi}_k^{\mathcal{S}}(\cdot, l; \mathcal{Z}_k^{\mathcal{S}}, \mathbf{m}), p_{k-1|k-1}^{S_{1:k-1}}(\cdot, l) \rangle, & \text{if } \mathbf{m} \neq -1 \\ 1 - r_{k|k-1}^{S_{1:k-1}}(l), & \text{if } \mathbf{m} = -1 \end{cases} \quad (15)$$

and, for  $\bar{a}_k(l) \neq -1$ , we define

$$p_{k|k}^{\bar{a}_k(l)}(\mathbf{x}, l; \mathcal{S}_{1:k}) = \bar{\psi}_k^{\mathcal{S}}(\mathbf{x}, l; \mathcal{Z}_k^{\mathcal{S}}, \bar{a}_k(l)) \times \frac{r_{k|k-1}^{S_{1:k-1}}(l) p_{k-1|k-1}^{S_{1:k-1}}(\mathbf{x}, l)}{\eta_{k|k}^{\bar{a}_k(l)}(l; \mathcal{S}_{1:k})} \quad (16)$$

<sup>2</sup>The notation  $\uplus$  represents the disjoint union operator.

The resulting multisensor LRFS posterior belief density of (13) is a mixture of LMB densities with belief density

$$\begin{aligned} \pi_{k|k}^{\mathcal{S}_{1:k}}(\mathcal{X}) &= \Delta_{\mathcal{X}} \sum_{\bar{\mathbf{a}}_k \in \mathcal{M}_k^{\mathcal{S}_k}} \phi_{\bar{\mathbf{a}}_k}^{\mathcal{S}_k}(\mathcal{L}_{\mathcal{X}}) \\ &\times P(\bar{\mathbf{a}}_k; \mathcal{S}_{1:k}) \left[ \prod_{(\mathbf{x}, l) \in \mathcal{X}} p_{k|k}^{\bar{\mathbf{a}}_k(l)}(\mathbf{x}, l; \mathcal{S}_{1:k}) \right] \end{aligned} \quad (17)$$

where we introduced the indicator  $\phi_{\bar{\mathbf{a}}_k}^{\mathcal{S}_k}(\mathcal{L}_{\mathcal{X}}) = \prod_{s \in \mathcal{S}_k} \delta_{\bar{\mathbf{a}}_{s,k}^{-1}(\{0; \mathcal{Z}_{s,k}\})}(\mathcal{L}_{\mathcal{X}})$  and the multisensor association map probability

$$P(\bar{\mathbf{a}}_k; \mathcal{S}_{1:k}) \propto \mathbb{1}_{\bar{\mathcal{A}}_k^{\mathcal{S}_k}}(\bar{\mathbf{a}}_k) \prod_{l \in \mathbb{L}_k} \eta_{k|k}^{\bar{\mathbf{a}}_k(l)}(l; \mathcal{S}_{1:k}). \quad (18)$$

Note that (17) is indeed a mixture of LMB belief densities. Similar to [26] and [22, Sec. XIII.A], we employ the approximation

$$P(\bar{\mathbf{a}}_k; \mathcal{S}_{1:k}) \approx \prod_{l \in \mathbb{L}_k} \hat{P}_l(\bar{\mathbf{a}}_k(l); \mathcal{S}_{1:k}) \quad (19)$$

where  $\hat{P}_l(\bar{\mathbf{a}}_k(l); \mathcal{S}_{1:k})$  is the marginal probability of associating  $\bar{\mathbf{a}}_k(l)$  to label  $l$  [22, eq. (43)]. Computationally efficient algorithms for evaluating marginal probabilities for data association problems are proposed in [22] and rely on graphical models coupled with message passing. By employing (19) in (17), the interlabel dependence-forcing term  $\mathbb{1}_{\bar{\mathcal{A}}_k^{\mathcal{S}_k}}(\bar{\mathbf{a}}_k)$  disappears, leading to a separable sum over the individual-label assignments  $\bar{\mathbf{a}}_k(l)$ , i.e.,  $\sum_{i,j} f(i) g(j) = [\sum_i f(i)] [\sum_j g(j)]$ . Finally, note the following alternative expression for  $\phi_{\bar{\mathbf{a}}_k}^{\mathcal{S}_k}$  as

$$\phi_{\bar{\mathbf{a}}_k}^{\mathcal{S}_k}(\mathcal{L}_{\mathcal{X}}) = \left[ \prod_{l \in \mathbb{L}_k \setminus \mathcal{L}_{\mathcal{X}}} \delta_{-1}(\bar{\mathbf{a}}_k(l)) \right] \left[ \prod_{l \in \mathcal{L}_{\mathcal{X}}} \mathbb{1}_{\mathbb{M}_k^{\mathcal{S}_k} \setminus \{-1\}}(\bar{\mathbf{a}}_k(l)) \right]$$

that is, the labels not contained in  $\mathcal{L}_{\mathcal{X}}$  have to be mapped to  $-1$  while the rest can be mapped to any sensor measurement or misdetection. Thus, the following LMB approximation to the full posterior (17) is obtained as

$$\begin{aligned} \hat{\pi}_{k|k}^{\mathcal{S}_{1:k}}(\mathcal{X}) &= \Delta_{\mathcal{X}} \left[ \prod_{l \in \mathbb{L}_k \setminus \mathcal{L}_{\mathcal{X}}} (1 - \hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l)) \right] \\ &\times \prod_{(\mathbf{x}, l) \in \mathcal{X}} \left[ \mathbb{1}_{\mathbb{L}_k}(l) \hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l) \hat{p}_{k|k}^{\mathcal{S}_{1:k}}(\mathbf{x}, l) \right]. \end{aligned} \quad (20)$$

The parameters  $\{\hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l), \hat{p}_{k|k}^{\mathcal{S}_{1:k}}(\mathbf{x}, l)\}_{l \in \mathbb{L}_k}$  of the approximating Bernoulli components are identified as

$$\hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l) = \sum_{m \in \mathbb{M}_k^{\mathcal{S}_k} \setminus \{-1\}} \hat{P}_l(m; \mathcal{S}_{1:k}) \quad (21a)$$

$$\hat{p}_{k|k}^{\mathcal{S}_{1:k}}(\mathbf{x}, l) = \sum_{m \in \mathbb{M}_k^{\mathcal{S}_k} \setminus \{-1\}} \frac{\hat{P}_l(m; \mathcal{S}_{1:k})}{\hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l)} p_{k|k}^m(\mathbf{x}, l; \mathcal{S}_{1:k}). \quad (21b)$$

The approximating LMB is employed to perform object inference and as the prior belief density for the subsequent time samples.

### C. Multiobject Inference

Multiobject inference from the posterior LMB is achieved in one of two ways.

1) *Threshold Method*: The first method involves comparing the Bernoulli probabilities of existence with a specified threshold  $\xi_{\text{th}} \in (0, 1)$  (typically  $\xi_{\text{th}} = 0.5$ ) and if  $\hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l) > \xi_{\text{th}}$ , the object state  $(\hat{\mathbf{x}}_{k|k}, l)$  is inferred as

$$\hat{\mathbf{x}}_{k|k} = \int_{\mathbb{X}} \mathbf{x} \hat{p}_{k|k}^{\mathcal{S}_{1:k}}(\mathbf{x}, l) d\mathbf{x}. \quad (22)$$

The estimated number of objects is given as

$$\hat{n}_{k|k} = \left| \left\{ l \in \mathbb{L}_k : \hat{r}_{k|k}^{\mathcal{S}_{1:k}}(l) > \xi_{\text{th}} \right\} \right|. \quad (23)$$

2) *MAP Cardinality Estimate*: The second method calls for an initial evaluation of the multi-Bernoulli cardinality distribution  $\hat{p}_{k|k}(n)$  of (5) and then performing the MAP cardinality estimate  $\hat{n}_{k|k} = \text{argmax}_n \hat{p}_{k|k}(n)$ . Subsequently, the state estimates (22) are inferred for the  $\hat{n}_{k|k}$  Bernoulli components with highest probabilities of existence.

## V. INFORMATION SEEKING OOT SENSOR SELECTION

In this section, the proposed information-theoretic sensor selection methodology for LMB models is described. As seen in the block diagram of Fig. 1, the predicted LMB density  $\pi_{k|k-1}$  is used to select the sensor subset that leads to the highest CS information gain for the update procedure. The CS divergence is derived for LMB LRFS models in Section V-A and a reward function for SS in OoT is proposed in Section V-B. In Section V-C, an SMC evaluation is given for the single-object space integrals involved in the reward function.

### A. Cauchy–Schwarz Divergence for LMB Densities

The CS divergence between two LRFS models with belief densities  $\pi_a(\cdot)$  and  $\pi_b(\cdot)$ , defined on the set  $\mathcal{F}(\mathbb{X} \times \mathbb{L})$ , is given by [36]

$$\mathbb{D}_{\text{CS}}\{\pi_a, \pi_b\} = -\ln \left( \frac{\int \kappa^{|\mathcal{X}|} \pi_a(\mathcal{X}) \pi_b(\mathcal{X}) \delta \mathcal{X}}{\sqrt{\int \kappa^{|\mathcal{X}|} \pi_a^2(\mathcal{X}) \delta \mathcal{X} \int \kappa^{|\mathcal{X}|} \pi_b^2(\mathcal{X}) \delta \mathcal{X}}} \right) \quad (24)$$

where for consistency with [36], the measurement unit  $\kappa$  on the space  $\mathbb{X}$  has been explicitly introduced. The unit  $\kappa$  can be dropped in the case of a unitless space  $\mathbb{X} \subseteq \mathbb{R}^d$ . The CS divergence is based on the inequality of the same name and was introduced for probability density functions  $f, g \in L^2(\mathbb{X})$  as  $\mathbb{D}_{\text{CS}}\{f, g\} = -\ln(\langle f, g \rangle / (\|f\|_2 \|g\|_2))$  in [56]. The CS divergence is a distance measure between  $f$  and  $g$ , that is,  $\mathbb{D}_{\text{CS}}\{f, g\} = \mathbb{D}_{\text{CS}}\{g, f\}$ ,  $\mathbb{D}_{\text{CS}}\{f, g\} \geq 0$ , and  $\mathbb{D}_{\text{CS}}\{f, g\} = 0$  whenever  $f = g$  almost everywhere. From a geometrical perspective, the CS divergence can be seen as the ‘‘information difference’’ between two stochastic models from the angle subtended by their density functions [36]. Additionally, in [56], the CS divergence is interpreted as an approximation to the Kullback–Leibler divergence.

*Lemma 1:* Considering two LMB-LRFS models  $\pi_a(\cdot)$  and  $\pi_b(\cdot)$  with parameters  $\{(r_a(l), p_a(\cdot, l))\}_{l \in \mathbb{L}}$  and  $\{(r_b(l), p_b(\cdot, l))\}_{l \in \mathbb{L}}$ , where the probability density functions  $p_a(\cdot, l) \in L^2(\mathbb{X})$  and  $p_b(\cdot, l) \in L^2(\mathbb{X}) \forall l$ , the CS divergence is given by

$$\mathbb{D}_{\text{CS}}\{\pi_a, \pi_b\} = - \sum_{l \in \mathbb{L}} \ln \left( \frac{K_{ab}(l)}{\sqrt{K_{aa}(l) K_{bb}(l)}} \right) \quad (25)$$

where for compactness we defined

$$K_{ab}(l) = (1 - r_a(l))(1 - r_b(l)) + r_a(l) r_b(l) \kappa \langle p_a(\cdot, l), p_b(\cdot, l) \rangle \quad (26a)$$

$$K_{aa}(l) = (1 - r_a(l))^2 + r_a^2(l) \kappa \langle p_a(\cdot, l), p_a(\cdot, l) \rangle \quad (26b)$$

$$K_{bb}(l) = (1 - r_b(l))^2 + r_b^2(l) \kappa \langle p_b(\cdot, l), p_b(\cdot, l) \rangle. \quad (26c)$$

The proof of Lemma 1 is given in Appendix A. Note that the inner products of (26a)–(26c) can be obtained in a closed form when the probability densities  $p_a(\cdot, l)$  and  $p_b(\cdot, l)$  are Gaussian; whereas they can also be easily evaluated via Monte Carlo methods when particle representations are employed for  $p_a(\cdot, l)$  and  $p_b(\cdot, l)$ .

In [35] and [57], the CS divergence between two LMB models was employed for sensor control. However, no exact form such as the one in Lemma 1 is reported. Instead, the LMB densities are approximated with Poisson RFSs via first-order moment preservation. Subsequently, the CS divergence between the approximating Poisson RFSs is easily evaluated due to [36]. The resulting divergence is used as an approximation to the CS divergence for LMB RFSs. In this article, we employ the exact form for the CS divergence as given in Lemma 1 as a measure of SS information gain for MOT in OoT, as described next.

### B. Information Gain for OoT-SS

Information-theoretic SS involves defining a suitable reward function over the set  $\mathbb{S}_M$  of all sensor subsets of cardinality  $M$ . Since the updated LMB parameters (21b) depend on the sensor subset  $\mathcal{S}_k$ , a divergence measure between the predicted and the posterior LMB densities quantifies the information gained by a specific sensor subset  $\mathcal{S}_k$  with respect to the initial density. In this article, the CS divergence between the predicted and the posterior LMB densities will be used as a reward function to optimize over the set  $\mathbb{S}_M$ . More precisely, the sensor reward function is defined as

$$R(\mathcal{S}_k, \mathcal{Z}_k^{\mathcal{S}_k}) = \mathbb{D}_{\text{CS}} \left\{ \pi_{k|k-1}^{\mathcal{S}_k}(\cdot), \hat{\pi}_{k|k}^{\mathcal{S}_k}(\cdot) \right\} \quad (27)$$

where the LMB posterior  $\hat{\pi}_{k|k}^{\mathcal{S}_k}(\cdot)$  is given in (20). Note however that the reward in (27) depends on the measurement set  $\mathcal{Z}_k^{\mathcal{S}_k}$  from the candidate subset  $\mathcal{S}_k$ . Thus, the use of (27) as a reward function in searching for the optimal subset would violate Assumption A8) of Section III. Ideally, an optimal subset can be selected by maximizing a new reward function

$$R(\mathcal{S}_k) = \int R(\mathcal{S}_k, \mathcal{Z}) \pi_{\mathcal{Z}_k^{\mathcal{S}_k}}(\mathcal{Z}) \delta \mathcal{Z} \quad (28)$$

that marginalizes (27) over all realizations of the random measurement set  $\mathcal{Z}_k^{\mathcal{S}_k}$ . The following two challenges prohibit

an analytical and a computationally tractable solution to the aforementioned expectation  $R(\mathcal{S}_k)$ .

- 1) The space of measurement sets  $\mathcal{Z}_k^{\mathcal{S}_k}$  is given by  $\mathcal{F}(\mathbb{Z}_{s_1}) \times \cdots \times \mathcal{F}(\mathbb{Z}_{s_M})$ , i.e., the Cartesian product of all finite subsets of the measurement spaces  $\mathbb{Z}_s$  of each sensor  $s \in \mathcal{S}_k$ .
- 2) For each  $\mathcal{Z}_k^{\mathcal{S}_k}$ , the marginal association probabilities  $\hat{P}_l(\bar{a}_k(l); \mathcal{S}_{1:k})$  need to be evaluated for all  $l \in \mathbb{L}_k$ .

In order to address these two challenges, we propose an approximate LMB  $\tilde{\pi}_{k|k}^{\mathcal{S}_k}(\cdot)$  to the posterior LMB of (20). Two simplifying assumptions are undertaken in the construction of  $\tilde{\pi}_{k|k}^{\mathcal{S}_k}(\cdot)$ .

- (i) For a sensor  $s \in \{1 : N\}$  and for each label  $l \in \mathbb{L}_k$ , the object-expected measurement vector is defined as

$$\tilde{z}_{s,k}^{(l)} = \int_{\mathbb{Z}_s} \mathbf{z} \int_{\mathbb{X}} g_{s,k}(\mathbf{z}|\mathbf{x}, l) p_{k|k-1}^{\mathcal{S}_k}(\mathbf{x}, l) d\mathbf{x} d\mathbf{z} \quad (29)$$

while for a sensor subset  $\mathcal{S}_k = \{s_1, \dots, s_M\}$ , the object-expected multisensor measurement is defined as  $\tilde{\mathcal{Z}}_k^{l, \mathcal{S}_k} \triangleq (\{\tilde{z}_{s_1,k}^{(l)}\}, \dots, \{\tilde{z}_{s_M,k}^{(l)}\})$ .

- (ii) Each Bernoulli component is updated independently of the others with its corresponding object-expected multisensor measurement, that is, component  $l \in \mathbb{L}_k$  is updated with  $\tilde{\mathcal{Z}}_k^{l, \mathcal{S}_k}$ , in order to simplify the evaluation of the marginal association probabilities (19).

These simplifying assumptions are similar to those employed in [33] and lead to the LMB update expressions of Lemma 2 presented next.

*Lemma 2:* The two aforementioned simplifications counter the intractable complexity of (29) and lead to the approximating LMB  $\tilde{\pi}_{k|k}^{\mathcal{S}_k}(\cdot)$  defined by the parameters  $\{(\tilde{r}_{k|k}^{\mathcal{S}_k}(l), \tilde{p}_{k|k}^{\mathcal{S}_k}(\mathbf{x}, l))\}_{l \in \mathbb{L}_k}$  which are given by

$$\tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{l, \mathcal{S}_k} | \mathbf{x}, l) = \prod_{s \in \mathcal{S}_k} \left[ 1 - p_{s,k}^{\text{D}}(\mathbf{x}, l) + \frac{p_{s,k}^{\text{D}}(\mathbf{x}, l) g_{s,k}(\tilde{z}_{s,k}^{(l)} | \mathbf{x}, l)}{f_{s,k}^{\text{FA}}(\tilde{z}_{s,k}^{(l)})} \right] \quad (30a)$$

$$\tilde{\eta}_k^{\mathcal{S}_k}(l) = \int \tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{l, \mathcal{S}_k} | \mathbf{x}, l) p_{k|k-1}^{\mathcal{S}_k}(\mathbf{x}, l) d\mathbf{x} \quad (30b)$$

$$\tilde{r}_{k|k}^{\mathcal{S}_k}(l) = \frac{r_{k|k-1}^{\mathcal{S}_k}(l) \tilde{\eta}_k^{\mathcal{S}_k}(l)}{1 - r_{k|k-1}^{\mathcal{S}_k}(l) + r_{k|k-1}^{\mathcal{S}_k}(l) \tilde{\eta}_k^{\mathcal{S}_k}(l)} \quad (30c)$$

$$\tilde{p}_{k|k}^{\mathcal{S}_k}(\mathbf{x}, l) = \frac{\tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{l, \mathcal{S}_k} | \mathbf{x}, l) p_{k|k-1}^{\mathcal{S}_k}(\mathbf{x}, l)}{\tilde{\eta}_k^{\mathcal{S}_k}(l)}. \quad (30d)$$

The proof of Lemma 2 is given in Appendix B. The approximating LMB  $\tilde{\pi}_{k|k}^{\mathcal{S}_k}(\cdot)$  of (30c) and (30d) is employed to evaluate the CS divergence

$$\tilde{R}(\mathcal{S}_k) = \mathbb{D}_{\text{CS}} \left\{ \pi_{k|k-1}^{\mathcal{S}_k}(\cdot), \tilde{\pi}_{k|k}^{\mathcal{S}_k}(\cdot) \right\}. \quad (31)$$

The reward function  $\tilde{R}(\mathcal{S}_k)$  no longer depends on the random set  $\mathcal{Z}_k^{\mathcal{S}_k}$  and is employed in our proposed SS selection method for MOT in OoT. In the following, an SMC evaluation of  $\tilde{R}(\mathcal{S}_k)$  is presented.

### C. SMC Implementation for OoT-SS

In nonlinear and/or non-Gaussian state-space systems, the object densities  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l)$  are represented via weighted particle sets  $\{(w_{k|k-1}^{(l,j)}, \mathbf{x}_{k|k-1}^{(l,j)})\}_{j=1}^{J(l)}$  that are sequentially propagated in time through the predict (Section IV-A) and update (Section IV-B) steps. The particle sets are considered normalized, i.e.,  $\sum_{j=1}^{J(l)} w_{k|k-1}^{(l,j)} = 1 \forall l \in \mathbb{L}_k$ . The explicit dependence of the particle weights and positions on the sets  $\mathcal{S}_{1:k-1}$  was dropped for compactness. The reader is directed to [26], [53], and [58]–[60] for an in-depth presentation of particle filters and SMC implementations of RFS filters (including the LMB). In this section, the SMC evaluation of the reward function of (31) is given.

The inner product  $\langle \tilde{p}_{k|k}^{\mathcal{S}_{1:k}}(\cdot, l), p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\cdot, l) \rangle$  required in the CS divergence of Lemma 1 is evaluated as

$$\begin{aligned} & \langle \tilde{p}_{k|k}^{\mathcal{S}_{1:k}}(\cdot, l), p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\cdot, l) \rangle \\ &= \int \frac{\tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}, l) p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l)}{\tilde{\eta}_k^{\mathcal{S}_{1:k}}(l)} p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l) d\mathbf{x} \quad (32a) \\ &\approx \sum_{j=1}^{J(l)} \frac{w_{k|k-1}^{(l,j)} \tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}_{k|k-1}^{(l,j)}, l) \hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}_{k|k-1}^{(l,j)}, l)}{\tilde{\eta}_k^{\mathcal{S}_{1:k}}(l)} \end{aligned} \quad (32b)$$

where  $\hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{k|k-1}^{(l)}, \mathbf{C}_{k|k-1}^{(l)})$  is the Gaussian density that provides a moment-matched approximation to the particle set of  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l)$ . This leads to the parameters

$$\begin{aligned} \boldsymbol{\mu}_{k|k-1}^{(l)} &= \sum_{j=1}^{J(l)} w_{k|k-1}^{(l,j)} \mathbf{x}_{k|k-1}^{(l,j)} \\ \mathbf{C}_{k|k-1}^{(l)} &= \sum_{j=1}^{J(l)} w_{k|k-1}^{(l,j)} [\mathbf{x}_{k|k-1}^{(l,j)} - \boldsymbol{\mu}_{k|k-1}^{(l)}][\mathbf{x}_{k|k-1}^{(l,j)} - \boldsymbol{\mu}_{k|k-1}^{(l)}]^T. \end{aligned}$$

Note that the square of  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\cdot, l)$  appears in (32a), which cannot be approximated directly via the substitution  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l) \approx \sum_{j=1}^{J(l)} w_{k|k-1}^{(l,j)} \delta_{\mathbf{x}_{k|k-1}^{(l,j)}}(\mathbf{x})$ . Instead, the Monte Carlo evaluation in (32b) is obtained by first approximating  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\cdot, l)$  via a smooth kernel density estimate (e.g., a Gaussian function)  $\hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l)$ , and subsequently by evaluating the product  $\tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}, l) \hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}, l)$  over the particle set  $\{(w_{k|k-1}^{(l,j)}, \mathbf{x}_{k|k-1}^{(l,j)})\}_{j=1}^{J(l)}$ . For improved accuracy, more complex models such as Gaussian mixtures can be employed as a smooth-function approximation for  $p_{k|k-1}^{\mathcal{S}_{1:k-1}}(\cdot, l)$ . Similarly

$$\begin{aligned} & \langle \tilde{p}_{k|k}^{\mathcal{S}_{1:k}}(\cdot, l), \tilde{p}_{k|k}^{\mathcal{S}_{1:k}}(\cdot, l) \rangle \\ &\approx \sum_{j=1}^{J(l)} \frac{w_{k|k-1}^{(l,j)} \left[ \tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}_{k|k-1}^{(l,j)}, l) \right]^2 \hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}_{k|k-1}^{(l,j)}, l)}{\left( \tilde{\eta}_k^{\mathcal{S}_{1:k}}(l) \right)^2}. \end{aligned}$$

Employing these SMC approximations and ignoring the terms that do not depend on  $\mathcal{S}_k$  in (31), the following SMC-SS

reward function is obtained

$$\hat{R}(\mathcal{S}_k) = - \sum_{l \in \mathbb{L}_k} \ln \left( \frac{1 + q^2(l) \sum_{j=1}^{J(l)} \alpha^{(l,j)}}{\sqrt{1 + q^2(l) \sum_{j=1}^{J(l)} \beta^{(l,j)}}} \right) \quad (33)$$

where

$$\begin{aligned} q(l) &= r_{k|k-1}^{\mathcal{S}_{1:k-1}}(l) [1 - r_{k|k-1}^{\mathcal{S}_{1:k-1}}(l)]^{-1} \\ \alpha^{(l,j)} &= w_{k|k-1}^{(l,j)} \tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}_{k|k-1}^{(l,j)}, l) \hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}_{k|k-1}^{(l,j)}, l) \\ \beta^{(l,j)} &= w_{k|k-1}^{(l,j)} \left[ \tilde{g}_k^{\mathcal{S}_k}(\tilde{\mathcal{Z}}_k^{\mathcal{S}_k} | \mathbf{x}_{k|k-1}^{(l,j)}, l) \right]^2 \hat{p}_{k|k-1}^{\mathcal{S}_{1:k-1}}(\mathbf{x}_{k|k-1}^{(l,j)}, l). \end{aligned}$$

In what follows, we present an algorithm that searches for the subset of sensors  $\mathcal{S}_k$  with maximum reward  $\hat{R}(\mathcal{S}_k)$  of (33) for OoT-MOT applications.

## VI. CROSS-ENTROPY FOR SENSOR SELECTION

The CE method was originally designed as an adaptive importance sampling scheme for the estimation of rare-event probabilities [61]. Subsequently in [62], CE was extended for solving both continuous and combinatorial optimization problems. The main challenge of SS problems is the evaluation of all possible combinations of sensor subsets and their associated rewards. For example, selecting  $M = 10$  sensors from  $N = 50$  entails  $\binom{N}{M} > 10^{10}$  valid sensor combinations and it becomes impractical to evaluate all combinations in most applications. In contrast, the CE method, as will be shown further on, only requires the evaluation of the reward function (33) for the sensor subsets that were sampled by the CE search procedure. Since the number of distinct subsets sampled by CE is in general much smaller than the total number of possible sensor subsets [i.e.,  $\binom{N}{M}$ ], the CE method can incur a much smaller computational overhead than brute-force search while still maintaining theoretical convergence guarantees [63], [64].

### A. Cross-Entropy Algorithm

CE methods rely on the construction of a random sequence of solutions, which converge to the optimal solution with probability arbitrarily close to one [63]. This procedure involves iterating the following two steps. First, given a sampling mechanism, samples are generated that represent potential solutions to the optimization problem. Second, the drawn samples are used to update the sampling mechanism (typically, parameters of the sampling distribution) in order to obtain better solutions at the next iteration with respect to the optimization functional or some suitably chosen reward function. The updated parameters are obtained by minimizing the Kullback–Leibler distance (or equivalently the cross-entropy) between the optimal importance distribution and the parametric family of distributions used for sampling. A comprehensive review of CE methods is given in [64]. Based on [64, Ch. 2.4], in [65], a CE method was proposed for the multidimensional assignment problem, which was applied for multisensor multiobject tracking both in a centralized algorithm [66] and in a decentralized algorithm in [67]. Furthermore, CE was applied to other difficult combinatorial problems, such as the traveling

salesman, the bipartition, and the maximum cut problems in [62] and [68].

In order to achieve a simple mechanism for sampling sensor subsets, marginal activation probabilities are constructed for each sensor. Let  $\mathbf{u} = [u_1, \dots, u_N]^T \in \mathcal{U} \triangleq \{0, 1\}^N$  be the random activation vector for sensors, i.e., for the  $i$ th sensor  $u_i = 1$  denotes ‘‘active’’ while  $u_i = 0$  denotes ‘‘inactive.’’ Furthermore, let  $\mathcal{U}_M = \{\mathbf{u} \in \mathcal{U}, \|\mathbf{u}\|_0 = M\}$  be the space of valid activation vectors where only  $M (< N)$  sensors are active. The marginal activation probability for the  $i$ th sensor is denoted with  $\mathbb{P}\{u_i = 1\} = v_i$ , while the vector of activation probabilities is denoted with  $\mathbf{v} = [v_1, \dots, v_N]^T$ . Sampling activations independently for each sensor according to  $\mathbf{v}$  leads to the probability of drawing a specific sample  $\mathbf{u} \in \mathcal{U}$  as

$$p(\mathbf{u}; \mathbf{v}) \triangleq \mathbb{P}\{\mathbf{u} = [u_1, \dots, u_N]^T\} = \prod_{i=1}^N v_i^{u_i} (1 - v_i)^{1 - u_i}. \quad (34)$$

The resulting activation vectors  $\mathbf{u} \in \mathcal{U}$  are not guaranteed to be valid, i.e.,  $\|\mathbf{u}\|_0 = M$  is not guaranteed. Thus, in order to obtain highly rewarding sensor subsets and valid activation vectors, the modified reward function is introduced

$$R_{\text{CE}}(\mathbf{u}) = \begin{cases} \hat{R}(\{i : i \in \{1 : N\}, u_i = 1\}), & \text{if } \mathbf{u} \in \mathcal{U}_M \\ -\infty, & \text{otherwise.} \end{cases} \quad (35)$$

In effect,  $R_{\text{CE}}(\mathbf{u})$  assigns the reward (33) to valid activation vectors  $\mathbf{u} \in \mathcal{U}_M$  and a value of  $-\infty$  to nonvalid activation vectors  $\mathbf{u} \notin \mathcal{U}_M$ .

The CE method for the deterministic problem of finding the SS that maximizes (33) starts by constructing an associated stochastic problem. More specifically, for a fixed threshold  $\varrho$ , the CE-associated stochastic problem [65] aims to estimate the tail probability

$$\mathbb{P}\{R_{\text{CE}}(\mathbf{u}) \geq \varrho\} \triangleq \sum_{\mathbf{u} \in \mathcal{U}} \mathbb{1}_{\{R_{\text{CE}}(\mathbf{u}) \geq \varrho\}}(\mathbf{u}) p(\mathbf{u}; \mathbf{v}). \quad (36)$$

For sufficiently large  $\varrho$ , (36) is the probability of observing sensor-activation vectors  $\mathbf{u}$  that lead to high rewards  $R_{\text{CE}}(\mathbf{u})$ . CE iteratively constructs a sequence of thresholds  $\varrho^{(k)}$  and parameter vectors  $\mathbf{v}^{(k)}$ , for  $k = 1, 2, \dots, K$ , which correspond to distributions (34) that assign a high-probability mass to high-scoring sensor-activation vectors. Intuitively, a distribution  $p(\mathbf{u}; \mathbf{v})$  which maximizes the tail probability (36), for a given threshold  $\varrho$ , is desirable since it leads to sampling sensor subsets with high reward. The CE iteration starts with an initial probability distribution  $\mathbf{v}^{(0)}$  and subsequently, at each iteration  $k$ , performs the following steps.

- 1) *Sample* a number  $T$  of sensor activation vectors  $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(T)}$  according to  $p(\mathbf{u}; \mathbf{v}^{(k)})$  given in (34).
- 2) *Adaptively estimate the threshold*  $\varrho^{(k+1)}$  as the  $(1 - \tau)$  quantile of the set  $R_{\text{CE}}(\mathbf{u}^{(1)}), \dots, R_{\text{CE}}(\mathbf{u}^{(T)})$ , that is

$$\frac{1}{T} \sum_{t=1}^T \mathbb{1}_{[\hat{\varrho}^{(k+1)}, +\infty)}(R_{\text{CE}}(\mathbf{u}^{(t)})) = \tau. \quad (37)$$

In practice, the sample activation vectors are sorted  $R_{\text{CE}}(\mathbf{u}^{(\sigma_1)}) \geq R_{\text{CE}}(\mathbf{u}^{(\sigma_2)}) \geq \dots \geq R_{\text{CE}}(\mathbf{u}^{(\sigma_T)})$  (ties can be broken arbitrarily), with  $\sigma$  a suitable permutation

of  $\{1 : T\}$ . Subsequently, only the  $H = \lceil \tau T \rceil$  highest scoring vectors are retained.

- 3) *Update marginal activation probabilities* via the Kullback–Leibler divergence (or equivalently cross-entropy) minimization. For fixed  $\hat{\varrho}^{(k+1)}$  and  $\mathbf{v}^{(k)}$ , the following program is constructed

$$\begin{aligned} \min_{\mathbf{v}} \quad & \mathbb{D}_{\text{KL}}\left\{\mathbb{1}_{[\hat{\varrho}^{(k+1)}, +\infty)}(R_{\text{CE}}(\mathbf{u})) \frac{p(\mathbf{u}; \mathbf{v}^{(k)})}{C}, p(\mathbf{u}; \mathbf{v})\right\} \\ \text{s.t.} \quad & v_i \in [0, 1] \quad \forall i \in \{1 : N\} \end{aligned}$$

where  $C \triangleq \sum_{\mathbf{u} \in \mathcal{U}} \mathbb{1}_{[\hat{\varrho}^{(k+1)}, +\infty)}(R_{\text{CE}}(\mathbf{u})) p(\mathbf{u}; \mathbf{v}^{(k)})$  is a normalization constant.<sup>3</sup> The optimization program is approximated using the current sample set to yield

$$\begin{aligned} \max_{\mathbf{v}} \quad & \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{[\hat{\varrho}^{(k+1)}, +\infty)}(R_{\text{CE}}(\mathbf{u}^{(t)})) \ln(p(\mathbf{u}^{(t)}; \mathbf{v})) \\ \text{s.t.} \quad & v_i \in [0, 1] \quad \forall i \in \{1 : N\}. \end{aligned} \quad (38)$$

Further employing the probability expression of (34) and by denoting  $n_i^1 \triangleq |\{t : t \in \{1 : H\}, u_i^{(\sigma_t)} = 1\}|$ , from (38), the convex program is obtained

$$\begin{aligned} \max_{\mathbf{v}} \quad & \sum_{i=1}^N \left[ n_i^1 \ln(v_i) + (H - n_i^1) \ln(1 - v_i) \right] \\ \text{s.t.} \quad & v_i \in [0, 1] \quad \forall i \in \{1 : N\} \end{aligned} \quad (39)$$

which leads to the solution  $v_i^* = n_i^1 / H \quad \forall i \in \{1 : N\}$ .

- 4) *Smoothing* of the updated probability density parameters via

$$\mathbf{v}^{(k+1)} = \gamma^{(k+1)} \mathbf{v}^* + (1 - \gamma^{(k+1)}) \mathbf{v}^{(k)} \quad (40)$$

where  $\{\gamma^{(k)} : k \in \{1 : K\}, \gamma^{(k)} \in (0, 1]\}$  is a sequence of smoothing parameters.

As suggested in [65], the sampling weights are initialized with the uniform distribution, i.e.,  $v_i^{(0)} = 1/N \quad \forall i$ . In Algorithm 1, the CE-SS algorithm for MOT in OoT is presented. The algorithm iterates the main steps of sampling and updating the vector of activation probabilities  $\mathbf{v}$  until convergence. Convergence implies a small change in the updated marginals or reaching a maximal number of iterations  $K$ . For a small  $\nu \in \mathbb{R}_+$ , the stopping criterion  $\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|_2 \leq \nu$  is adopted for the CE iterations. Step 2 of Algorithm 1 involves sampling of activation vectors  $\{\mathbf{u}^{(t)}\}_{t=1}^T$ . Sampling directly from (34) involves  $N$  independent Bernoulli trials with probabilities  $v_i$ . However, this is inefficient since many sampled activation vectors may not be valid, i.e.,  $\mathbf{u} \notin \mathcal{U}_M$ , and consequently do not contribute to the update of the probability vector  $\mathbf{v}$  due to the reward function (35). To avoid generating irrelevant samples, in Section VI-B, we propose an efficient approach for sampling sensor activation vectors conditioned on a fixed cardinality.

<sup>3</sup>Note that  $C > 0$  since, by construction of  $\hat{\varrho}^{(k+1)}$ , there are at least  $H$  outcomes  $\mathbf{u} \in \mathcal{U}$  with  $R_{\text{CE}}(\mathbf{u}) \geq \hat{\varrho}^{(k+1)}$  and positive probability mass  $p(\mathbf{u}; \mathbf{v}^{(k)}) > 0$ .

**Algorithm 1** CE-SS Algorithm for OoT

- 1: Initialize  $k \leftarrow 0$  and set  $v_i^{(0)} \leftarrow 1/N$ .
- 2: Draw  $T$  independent samples  $\{\mathbf{u}^{(t)}\}_{t=1}^T$  according to  $p(\mathbf{u}; \mathbf{v}^{(k)})$  (detailed in Sec. VI-B).
- 3: Compute  $R_{\text{CE}}(\mathbf{u}^{(t)})$  for  $t = 1, \dots, T$ .
- 4: Let  $H \leftarrow \lfloor \tau T \rfloor$  and sort the sequence  $\{R_{\text{CE}}(\mathbf{u}^{(t)})\}_{t=1}^T$  as  $R_{\text{CE}}(\mathbf{u}^{(\sigma_1)}) \geq R_{\text{CE}}(\mathbf{u}^{(\sigma_2)}) \geq \dots \geq R_{\text{CE}}(\mathbf{u}^{(\sigma_T)})$ .
- 5: Update  $v_i^* \leftarrow |\{t : t \in \{1 : H\}, u_i^{(\sigma_t)} = 1\}|/H \forall i$ .
- 6: Smooth probability distribution parameters as in (40).
- 7: Increment  $k$  and iterate steps 2-6 until convergence or  $k = K$ .
- 8: Return highest rewarding activation vector  $\mathbf{u}^{(*)}$  from all samples and all iterations.

**B. Conditional Sampling of Sensor-Activation Vectors**

A direct algorithm for sampling valid sensor-activation vectors involves sampling from (34) via  $N$  independent Bernoulli trials with probabilities  $v_i$  and then accepting the sample if  $\|\mathbf{u}\|_0 = M$ . The procedure is repeated until  $T$  valid samples are obtained. An equivalent algorithm samples from the conditional densities

$$p(\mathbf{u} \mid \|\mathbf{u}\|_0 = M) = p(u_1 \mid \|\mathbf{u}\|_0 = M) \times \prod_{i=2}^N p(u_i \mid \mathbf{u}_{1:i-1}, \|\mathbf{u}\|_0 = M). \quad (41)$$

Furthermore by denoting  $q_i \triangleq [v_i/(1-v_i)]$  and constructing the vector  $\mathbf{q} = [q_1, \dots, q_N]^T$ , the conditional densities can be expressed as

$$p(u_1 \mid \|\mathbf{u}\|_0 = M) = \frac{\sum_{\mathbf{u}_{2:N}} p(\mathbf{u}, \|\mathbf{u}\|_0 = M)}{\sum_{\mathbf{u}} p(\mathbf{u}, \|\mathbf{u}\|_0 = M)} = \frac{q_1^{u_1} \sum_{\mathbf{u}_{2:N}} q_2^{u_2} \dots q_N^{u_N} \delta_M(\|\mathbf{u}\|_0)}{\sum_{\mathbf{u}} q_1^{u_1} q_2^{u_2} \dots q_N^{u_N} \delta_M(\|\mathbf{u}\|_0)} \quad (42)$$

where the Kronecker operator  $\delta_M(\|\mathbf{u}\|_0) = 1$  if  $\|\mathbf{u}\|_0 = M$  and 0 otherwise. Noting that  $\mathfrak{E}_M(\mathbf{q}) = \sum_{\mathbf{u}} q_1^{u_1} q_2^{u_2} \dots q_N^{u_N} \delta_M(\|\mathbf{u}\|_0)$  is the elementary symmetric function of (6), and for the two choices  $u_1 \in \{0, 1\}$ , (42) becomes

$$p(u_1 \mid \|\mathbf{u}\|_0 = M) = \begin{cases} q_1 \frac{\mathfrak{E}_{M-1}(\mathbf{q}_{2:N})}{\mathfrak{E}_M(\mathbf{q})}, & \text{if } u_1 = 1 \\ \frac{\mathfrak{E}_M(\mathbf{q}_{2:N})}{\mathfrak{E}_M(\mathbf{q})}, & \text{if } u_1 = 0. \end{cases} \quad (43)$$

Conditionally on  $\mathbf{u}_{1:i-1}$  and by denoting  $m_{i-1} = \sum_{j=1}^{i-1} u_j$ , the  $i$ th conditional density becomes

$$p(u_i \mid \mathbf{u}_{1:i-1}, \|\mathbf{u}\|_0 = M) = q_i^{u_i} \frac{\mathfrak{E}_{M-m_{i-1}-u_i}(\mathbf{q}_{i+1:N})}{\mathfrak{E}_{M-m_{i-1}}(\mathbf{q}_{i:N})} \quad (44)$$

for  $i < N$ . While  $\mathbb{P}\{u_N = 1 \mid \mathbf{u}_{1:N-1}, \|\mathbf{u}\|_0 = M\} = 1$  if  $m_{N-1} = M - 1$  and 0 otherwise. The conditional sampling algorithm initially samples  $u_1$  via (42) followed by  $u_i$  via (44) for  $i = 2, \dots, N - 1$  or until  $\|\mathbf{u}_{1:i}\|_0 = M$ . Note that  $\mathfrak{E}_0(\mathbf{q}_{i:N}) = 1 \forall i$ ,  $\mathfrak{E}_j(\mathbf{q}_{N-j+1:N}) = \prod_{t=N-j+1}^N q_t$  for  $j \in \{1 : M\}$ , and by convention  $\mathfrak{E}_i(\mathbf{q}_{N-j+1:N}) = 0$  for

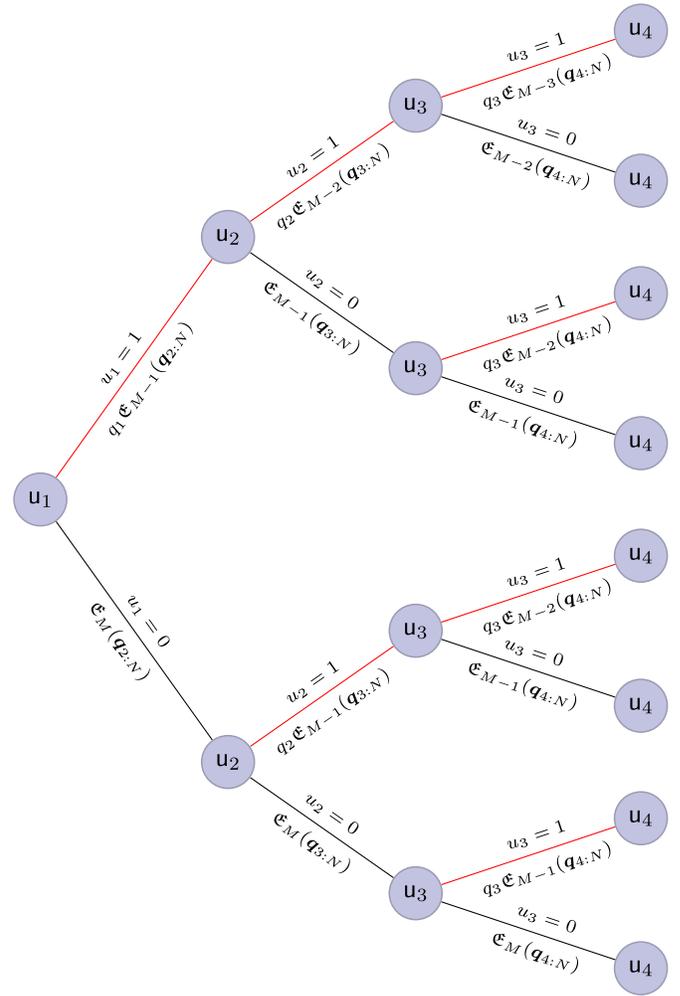


Fig. 2. Binary probability tree showcasing the conditional sampling procedure for the first four sensor activation indicators  $\mathbf{u}_{1:4}$ . At each level,  $i \in \{1:4\}$ , a binary decision, i.e.,  $u_i = 1$  (red edge) or  $u_i = 0$  (black edge), is taken with probabilities proportional to the values inscribed under the corresponding edges.

$i > j$  and  $j \in \{1 : M\}$ . This ensures that a sample with  $\|\mathbf{u}\|_0 = M$  is always obtained. This sampling process is captured by the binary tree depicted in Fig. 2. From (44), the conditional activation  $u_i = 1$  of the  $i$ th sensor can be achieved via Bernoulli sampling with success probability  $b_{i,M-m_{i-1}}$ , where  $b_{i,j} \in [0, 1]$  is given by

$$b_{i,j} = \frac{q_i \mathfrak{E}_{j-1}(\mathbf{q}_{i+1:N})}{q_i \mathfrak{E}_{j-1}(\mathbf{q}_{i+1:N}) + \mathfrak{E}_j(\mathbf{q}_{i+1:N})} \quad (45)$$

for  $i < N$  and  $b_{N,1} = 1$ .

The algorithm for the conditional sampling of valid sensor activation vectors  $\mathbf{u} \in \mathcal{U}_M$  is presented in Algorithm 2. The function call  $\text{Ber}(p)$  denotes a Bernoulli draw with success probability  $p$ . From (6), the values of the elementary symmetric function  $\mathfrak{E}_j(\mathbf{q}_{i:N})$  are easily obtained via the recursion at line 6 of Algorithm 2. Subsequently, a table containing the probabilities of success  $b_{i,j}$  is formed at line 7. A valid activation  $u_i$  for the  $i$ th sensor is sampled at line 13 as a Bernoulli draw. The worst case computational complexity of Algorithm 2 is  $\mathcal{O}(TN)$  for sampling a number of  $T$  vectors. The computational complexity

---

**Algorithm 2** Conditional Sampling of Valid Sensor Activations
 

---

```

1: Input vector of weights  $\mathbf{q} \leftarrow [q_1, \dots, q_N]^T$ ,  $T$ ,  $N$ , and  $M$ .
2: Set  $\mathfrak{E}_0(\mathbf{q}_{i:N}) \leftarrow 1$ ,  $b_{i,j} \leftarrow 1$  for  $i \in \{1 : N\}$ ,  $j \in \{1 : M\}$ .
3: Set  $\mathfrak{E}_j(\mathbf{q}_{N-j+1:N}) \leftarrow \prod_{t=N-j+1}^N q_t$  for  $j \in \{1 : M\}$ .
4: for  $i = (N - 1)$  down to 1 do
5:   for  $j = 1$  to  $\min(M, N - i)$  do
6:      $\mathfrak{E}_j(\mathbf{q}_{i:N}) \leftarrow q_i \mathfrak{E}_{j-1}(\mathbf{q}_{i+1:N}) + \mathfrak{E}_j(\mathbf{q}_{i+1:N})$ 
7:     Compute  $b_{i,j}$  via (45).
8:   end for
9: end for
10: for  $t = 1$  to  $T$  do
11:   Set  $i \leftarrow 1$ ,  $j \leftarrow M$ , and  $\mathbf{u}^{(t)} \leftarrow [0, \dots, 0]^T$ .
12:   while  $(i \leq N) \wedge (j > 0)$  do
13:     Sample  $u_i^{(t)} \leftarrow \text{Ber}(b_{i,j})$ .
14:     Set  $j \leftarrow j - u_i^{(t)}$  and  $i \leftarrow i + 1$ .
15:   end while
16: end for
17: Return the set of valid activation vectors  $\{\mathbf{u}^{(t)}\}_{t=1}^T$ .

```

---

of evaluating the reward function  $R_{\text{CE}}(\mathbf{u})$  for an activation vector is  $\mathcal{O}(\sum_{l \in \mathbb{L}_k} J(l))$  since the terms in (33) involve computing the parameters  $\alpha^{(l,j)}$  and  $\beta^{(l,j)}$  for all labels and particles. Hence, step 3 of Algorithm 1 has a complexity of  $\mathcal{O}(T \sum_{l \in \mathbb{L}_k} J(l))$ . For  $K$  iterations, the computational complexity of Algorithm 2 is  $\mathcal{O}(KT(N + \sum_{l \in \mathbb{L}_k} J(l)))$ . Note that the complexity of the algorithm is linear with both the number of sensors  $N$  and the number of objects  $|\mathbb{L}_k|$ .

The convergence of CE algorithms was studied for generic combinatorial optimization problems in [63], where both necessary and sufficient conditions are given. More specifically, Algorithm 1 eventually generates an optimal solution with probability 1 if: 1) the initial probabilities verify  $v_i^{(0)} \in (0, 1)$  for  $i \in \{1 : N\}$  and 2) the smoothing sequence verifies  $\gamma^{(k)} \in (0, 1] \forall k$  and  $\sum_{k=1}^{\infty} \gamma^{(k)} < \infty$ . For the more common-use case of a constant smoothing parameter  $\gamma^{(k)} = \gamma \in (0, 1]$  and initial probabilities  $v_i^{(0)} \in (0, 1)$  for  $i \in \{1 : N\}$ , Algorithm 1 converges almost surely to a unit mass located in  $\mathcal{U}_M$ , while the probability of generating an optimal activation vector can be made arbitrarily close to 1. Note that in both cases, the convergence properties hold as the number  $K$  of iterations tends to infinity. To the best of our knowledge, a convergence-rate analysis of CE algorithms seems to be lacking in the literature.

## VII. NUMERICAL EXPERIMENTS

Three simulation scenarios are presented to assess the performance of our proposed algorithm. First, a static scenario is considered where the objects and sensors are stationary. The objective of the static scenario is to assess the convergence and computational speedup of the CE algorithm with respect to a brute-force search algorithm. Furthermore, a Monte Carlo comparison is conducted to assess the relative improvement

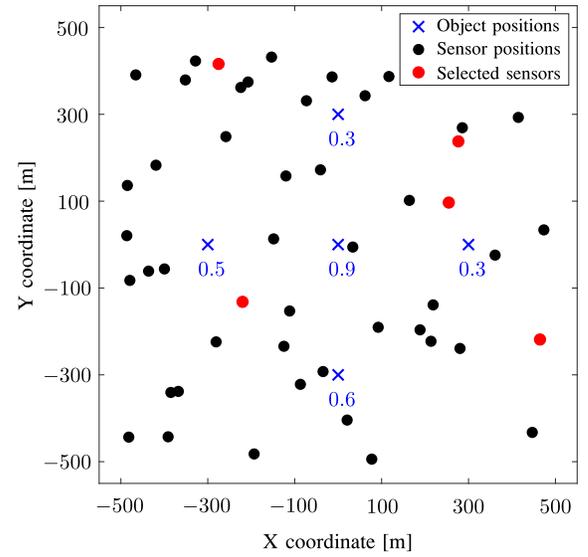


Fig. 3. Static SS scenario with five objects,  $N = 50$  sensors and  $M = 5$  selected sensors according to the highest CE reward function.

of our proposed method with respect to a uniform-SS method. The second scenario emulates an MOT application where a group of objects is being tracked in time over a grid of sensors. Here, the tracking performance achieved by the multi-sensor LMB filter of Section IV coupled with our proposed SS algorithm is compared to the performance of the same filter but coupled with a uniform SS algorithm. Finally, we showcase the OoT tracking of three vessels in an emulated scenario off the coast of Italy with the multisensor LMB filter in conjunction with our proposed cross-entropy SS (CE-SS) algorithm. In all scenarios, the sensors provide range-only measurements which correspond to single-hydrophone floats. For a sensor  $s$  placed at coordinates  $(x_s, y_s)$ , the range-only measurement equation for an object with coordinates  $(x, y)$  is

$$\mathbf{z}_s = \sqrt{(x - x_s)^2 + (y - y_s)^2} + \mathbf{b}_s \quad (46)$$

where the additive noise  $\mathbf{b}_s$  is a zero-mean Gaussian random variable with variance  $\sigma_s^2$  and is independent of the measurement noise  $\mathbf{b}_i$  of other sensors  $i \neq s$  (in order for Assumption A7) of Section III to hold).

### A. Static SS Scenario

An initial numerical simulation is carried out for a group of five static objects with the goal of selecting  $M = 5$  out of a total of  $N = 50$  sensors that lead to the highest information gain according to the CS divergence of (33). The object set is described as an LMB where the Bernoulli components are parametrized via Gaussian densities with corresponding means showcased in Fig. 3 and standard deviation of 10 m in both  $X$  and  $Y$  directions. The Bernoulli probabilities of existence are showcased adjacent to their mean positions. The sensors provide range-only measurements according to (46) and are placed uniformly within the Region of Interest (RoI), and their parameters are also drawn uniformly as follows. Their probabilities of detection are uniformly drawn within  $[0.5, 1]$ , clutter measurement rates (i.e., average number of

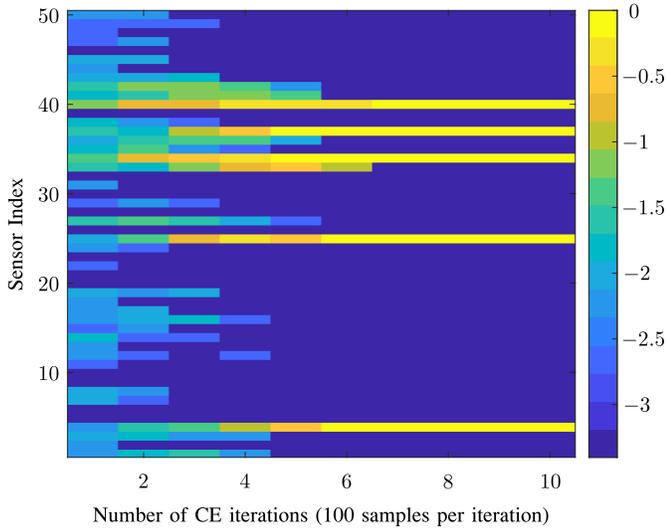


Fig. 4. Logarithm of marginal sensor selection probabilities for the CE-SS algorithm.

TABLE II  
RELATIVE IMPROVEMENT  $\Upsilon_{M,N}$  OF CE-SS OVER UNIFORM SAMPLING

		Total number of sensors $N$			
		100	500	1000	5000
$M$	10	21.5%	38.2%	45.9%	56.5%
	20	24.5%	39.2%	46.9%	57.9%
	30	25.7%	39.6%	46.2%	58.4%

clutter measurements per time step) are uniformly drawn within  $[5, 15]$ , and the measurement noise variance  $\sigma_s^2$  is uniformly drawn within  $[5, 15] \text{ m}^2$ . The goal of this static scenario is to showcase a high-dimensional SS problem typical of OoT and where brute-force enumeration procedures are intractable in practice. A subsequent goal is to evaluate the speedup and convergence of the proposed CE-SS algorithm. For the SS problem of Fig. 3, the total number of valid combinations is 2 118 760. CE-SS was run for ten iterations with  $T = 100$  sensor subsets being sampled at each iteration, thus leading to a total of 1000 samples. Furthermore, only 330 samples were distinct in one simulation with the best scoring sensor subset being highlighted in red in Fig. 3. Additional computational savings are made by employing the memoization of function calls for the CE evaluation (33) of sensor subsets. The total runtime for the brute-force search procedure is 6.8 min while for CE-SS it is 0.18 s, leading to a speedup of more than 220 times. Simulations were performed in MATLAB on a laptop equipped with an Intel i7 processor and 12 GB of RAM. Both the brute-force search and CE-SS algorithms reach the same optimal sensor subset and the marginal sensor selection probabilities for CE are showcased in Fig. 4 as a function of the iteration number. Note the fast convergence of CE, i.e., less than ten iterations are required to converge to the same solution as the brute-force search.

A Monte Carlo analysis of the static simulation scenario is conducted in the following. The proposed CE-SS algorithm is compared with a reference algorithm that selects a subset of

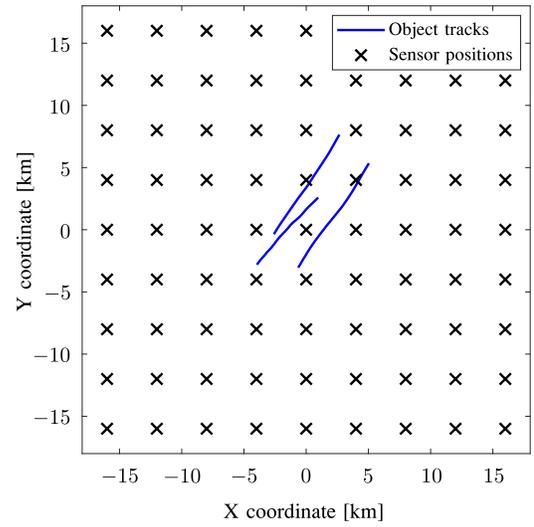


Fig. 5. Sensor grid and an instance of tracks for three closely spaced objects. The time evolution of the number of objects is also shown in the lower panel.

sensors uniformly at random. The parameters of the CE-SS algorithm are: the number of samples per iteration  $T = 300$ , the maximum number of iterations  $K = 20$ , the smoothing parameter  $\gamma = 0.999$ , and the threshold  $\tau = 0.3$ . In order to ensure a fair comparison, the uniform-SS algorithm performs  $T \times K = 6000$  samples where each sample consists of  $M$  sensors uniformly selected from the available  $N$  sensors. For each Monte Carlo run  $i$ , the uniform-SS algorithm returns the sample with the highest reward  $\overline{R}_U^{(i)}$  out of the total  $T \times K$  samples; while the CE-SS algorithm returns the sample with the highest reward  $\overline{R}_{CE}^{(i)}$  and the sample with the lowest reward  $\underline{R}_{CE}^{(i)}$ . For each pair  $(M, N)$ , a total of  $I = 100$  independent Monte Carlo runs are performed and a relative-improvement score of CE-SS over the uniform SS algorithm is defined as

$$\Upsilon_{M,N} = \frac{\sum_{i=1}^I [\overline{R}_{CE}^{(i)} - \overline{R}_U^{(i)}]}{\sum_{i=1}^I [\overline{R}_{CE}^{(i)} - \underline{R}_{CE}^{(i)}]} 100 \text{ [\%]}. \quad (47)$$

The relative-improvement score  $\Upsilon_{M,N}$  provides a measure of reward gain (in terms of the CS reward of Section V) for the CE-SS over the uniform algorithm when the reward of the optimal subset is unknown or hard to compute, e.g., for high  $N$ . In Table II, the relative-improvement scores are provided for various pairs  $(M, N)$ . The exploration of the sensor-subset space is significantly hindered at high  $N$ . However, the increase of  $\Upsilon_{M,N}$  with  $N$  highlights that the adaptive sampling scheme of CE-SS is more efficient at exploring such high-dimensional spaces than uniform sampling.

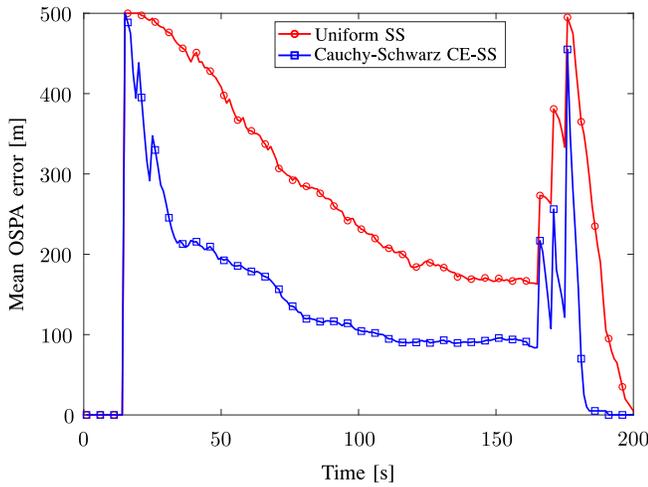


Fig. 6. Mean OSPA for the proposed and reference SS algorithms when the sensor characteristics are uniformly sampled.

### B. Dynamic SS Scenario

In this section, filtering results are presented in a dynamic MOT setting for the multisensor LMB filter of Section IV with our proposed information theoretic CE-SS algorithm and a reference method that uniformly selects a subset of sensors at each time step. A 2-D scenario is simulated involving three objects with trajectories depicted in Fig. 5 over a simulation length of 200 time samples with sampling period  $T_s = 1$  s. Object state vectors are constructed as  $\mathbf{x} = [x, y, \dot{x}, \dot{y}]^T$ , where  $x$  and  $y$  represent the 2-D position coordinates of the object and  $\dot{x}$  and  $\dot{y}$  are its velocities along the two axes. The dynamic model for an object  $l$  is given by the white-noise acceleration model [69, Ch. 6.3.2] and has the transition kernel  $f_{k+1|k}(\mathbf{x}|\mathbf{y}; l)$  is  $\mathcal{N}(\mathbf{x}; \mathbf{F}_{k+1}\mathbf{y}, \mathbf{Q}_k)$  with the state transition matrix

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{I}_2 & T_s \mathbf{I}_2 \\ \mathbf{0}_2 & \mathbf{I}_2 \end{bmatrix} \quad (48)$$

where  $\mathbf{0}_n$  and  $\mathbf{I}_n$  are the zero and identity matrices of size  $n$ . The process noise is Gaussian with the covariance matrix

$$\mathbf{Q}_k = \sigma_o^2 \begin{bmatrix} \frac{T_s^4}{4} \mathbf{I}_2 & \frac{T_s^3}{2} \mathbf{I}_2 \\ \frac{T_s^3}{2} \mathbf{I}_2 & T_s^2 \mathbf{I}_2 \end{bmatrix} \quad (49)$$

where the acceleration noise is taken to be  $\sigma_o^2 = 0.01 \text{ m}^2/\text{s}^4$ . The optimum subpattern assignment (OSPA) error metric [70] was introduced as a distance measure between two sets and is commonly employed in MOT to quantify the estimation error between an estimated set and the ground-truth set of objects. The OSPA metric employs two parameters, the cutoff  $c$  and order  $p$ . Note that OSPA incorporates both errors in cardinality and state estimates. The cutoff controls the impact of cardinality errors in the overall OSPA while a truncated  $L^p$  distance is employed between the estimated and ground-truth state vectors.

A fixed grid of  $N = 81$  range-only sensors, as depicted in Fig. 5, is considered where the sensor characteristics are chosen randomly at the beginning of each simulation. More specifically, the measurement noise standard deviation  $\sigma_s$  is

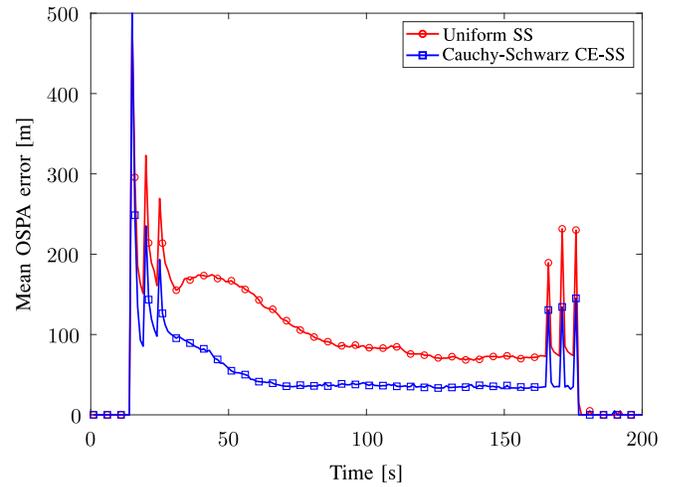


Fig. 7. Mean OSPA for the proposed and reference SS algorithms when the sensor characteristics are range dependent.

chosen uniformly in  $[50, 500]$  m and the probability of detection is uniform over the RoI and is set to either 0.1 or 0.9 with equal probabilities. The goal of the SS algorithms is to select the  $M = 3$  best sensors at each time step. A Monte Carlo simulation is conducted where the object tracks, sensor characteristics, and sensor measurements are sampled anew in each simulation for a total of 100 independent runs. The mean OSPA error metric with cutoff  $c = 500$  m and order  $p = 1$  is presented as a function of time in Fig. 6. Note that only position coordinates are considered in the evaluation of the OSPA error. Observe the error spikes that correspond with the times of object births and deaths. Furthermore, the proposed CE-SS algorithm achieves a lower mean OSPA error than uniform SS. The time-averaged mean OSPA for the CE-SS algorithm is 133 m while 262 m is obtained for the reference method, leading to an average performance improvement of 49%.

In a second simulation, a range-dependent measurement noise is employed where the standard deviation  $\sigma_s(r)$  at distance  $r$  (both expressed in meters) from sensor  $s$  is given by  $\sigma_s(r) = 50 \exp(r/8000)$  while the probability of detection is kept constant at 0.9 for all sensors. In this case, the sensors are accurate at close ranges but exhibit a sharp decline in precision at farther distances from the sensor's position. The time-dependent mean OSPA curves are presented in Fig. 7 and which show the improved performance of the proposed CE-SS algorithm with respect to the reference method. The time-averaged mean OSPA is now 45 m for the proposed CE-SS while a value of 94 m is obtained for the reference method, thus the proposed method achieves a performance improvement of 52%.

### C. SS for Vessel Tracking

In this section, an example is provided of multivesel tracking where our proposed CE-SS algorithm is employed for selecting a subset from the available sensing devices, i.e., floats, at any given time instant. The ground-truth trajectories of the vessels are obtained from their global-positioning-system coordinates, whereas the trajectories of  $N = 590$  floats, subject to ocean currents, are obtained via numerical simulation. The vessels and floats are deployed off

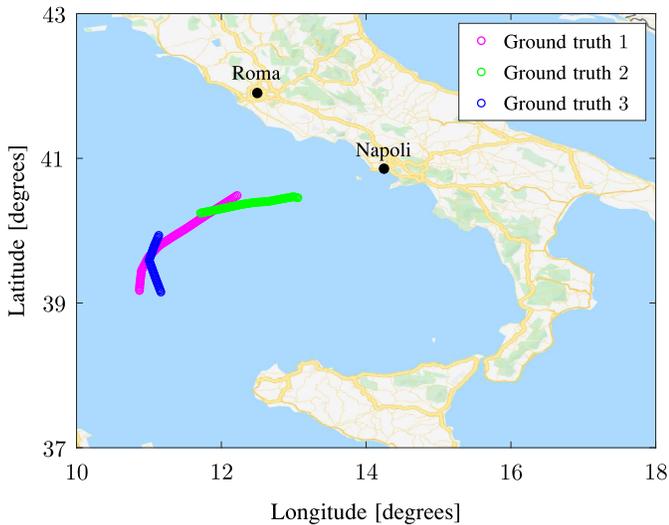


Fig. 8. Ground-truth tracks of three vessels off the coast of Italy over a 9 hour period (background image courtesy of Google Maps).

the coast of Italy and their trajectories are depicted in Fig. 8 over a total of 90 time steps with a  $T_s = 0.1$  hour sampling interval. Fig. 9 presents a detailed image of the vessel trajectories and the float tracks that result from ocean drift.

The floats are equipped with sensors capable of providing range measurements with a maximum measurement range of 10 km. Other sensor characteristics are: Gaussian measurement noise with the standard deviation of 100 m, probability of detection of 0.9, and uniformly distributed clutter with a rate of 1 measurement per time step and per sensor. A discretized nearly constant velocity model [69, Ch. 6.2.2] is assumed for the vessel kinematics with a process noise intensity of  $0.001 \text{ m}^2/\text{s}^3$ . The SMC-LMB filter of Section IV is employed for vessel tracking in conjunction with our proposed CE-SS algorithm. A number of  $5 \times 10^4$  particles are employed to model each of the object posterior densities while the CE-SS parameters are set to:  $T = 300$ ,  $K = 20$ ,  $\gamma = 0.999$ , and  $\tau = 0.3$ . A number of  $M = 30$  floats are selected from the available floats via our proposed CE-SS algorithm. In addition to the  $M = 30$  CE-SS floats, floats which are at the edge of the RoI are uniformly activated at random in order to detect entering vessels, i.e., birthed vessels. In Fig. 10, the resulting estimated vessel tracks can be observed overlaid onto the ground-truth tracks and the CE-SS active floats at every 1 hour interval. There is a good agreement between the ground-truth and estimated tracks for all three vessels. Furthermore, on average 12% of floats (cumulated CE-SS and birth) are active at any time instant. This results in a large reduction of the communication overhead, which is crucial for efficient OoT systems subject to battery-life and/or communication constraints.

### VIII. CONCLUSION

In this article, we provided a general information-seeking sensor selection methodology for Bayesian multiobject tracking in OoT, where the objective is to sequentially estimate the number and the states of a collection of objects observed

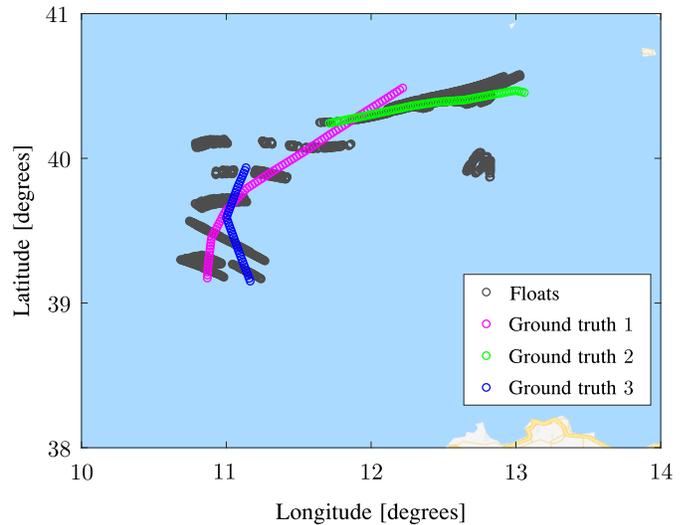


Fig. 9. Detail of Fig. 8 showcasing the vessel ground truth and the trajectories of all 590 floats over the same 9 hour period (background image courtesy of Google Maps).

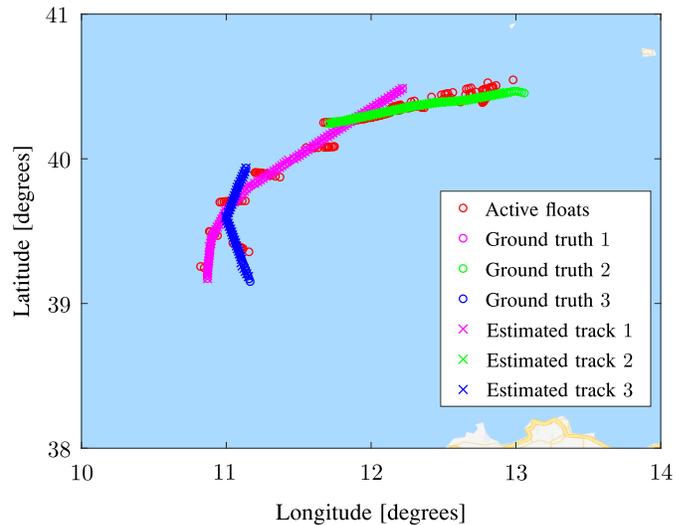


Fig. 10. Estimated and ground-truth tracks of the three vessels and CE-SS-selected floats. The positions of the selected floats are shown at each 1 hour interval (background image courtesy of Google Maps).

by only a subset of sensors. We derive a tractable information divergence measure between the predicted and posterior multiobject models in order to quantify the information gain of a specific sensor subset. Subsequently, we propose a cross-entropy stochastic search algorithm to find the optimum sensor subset according to the proposed divergence measure and under the specified resource constraint. The cross-entropy algorithm alleviates the need for brute-force search algorithms which would otherwise be intractable in most OoT applications. We demonstrate the performance gain of our proposed methodology via numerical experiments.

### APPENDIX A PROOF OF LEMMA 1

By employing the set-integral definition (1), together with (3) and (4) for  $\pi_a$  and  $\pi_b$ , the logarithm numerator in (24)

becomes

$$\begin{aligned}
& \int \kappa^{|\mathcal{X}|} \pi_a(\mathcal{X}) \pi_b(\mathcal{X}) \delta \mathcal{X} \\
&= \int \Delta_{\mathcal{X}} w_a(\mathcal{L}_{\mathcal{X}}) w_b(\mathcal{L}_{\mathcal{X}}) \kappa^{|\mathcal{X}|} [p_a(\cdot) p_b(\cdot)]^{\mathcal{X}} \delta \mathcal{X} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(l_1, \dots, l_n) \in \mathbb{L}^n} \delta_n(|\{l_1, \dots, l_n\}|) w_a(\{l_1, \dots, l_n\}) \\
&\quad \times w_b(\{l_1, \dots, l_n\}) \\
&\quad \times \int_{\mathbb{X}^n} \prod_{i=1}^n \kappa p_a(\mathbf{x}_i, l_i) p_b(\mathbf{x}_i, l_i) d(\mathbf{x}_1, \dots, \mathbf{x}_n) \\
&\stackrel{(i)}{=} \sum_{n=0}^{|\mathbb{L}|} \sum_{\mathcal{L} \subset \mathbb{L}, |\mathcal{L}|=n} w_a(\mathcal{L}) w_b(\mathcal{L}) \prod_{l \in \mathcal{L}} [\kappa \langle p_a(\cdot, l), p_b(\cdot, l) \rangle] \\
&= \left[ \prod_{l \in \mathbb{L}} (1 - r_a(l))(1 - r_b(l)) \right] \sum_{n=0}^{|\mathbb{L}|} \sum_{\mathcal{L} \subset \mathbb{L}, |\mathcal{L}|=n} \\
&\quad \times \prod_{l \in \mathcal{L}} \left[ \kappa \frac{r_a(l)}{1 - r_a(l)} \frac{r_b(l)}{1 - r_b(l)} \langle p_a(\cdot, l), p_b(\cdot, l) \rangle \right] \\
&\stackrel{(ii)}{=} \prod_{l \in \mathbb{L}} K_{ab}(l).
\end{aligned}$$

At line (i), all summation terms over tuples without  $n$  distinct labels are dropped due to  $\delta_n(|\{l_1, \dots, l_n\}|)$  as well as terms with more than  $|\mathbb{L}|$  labels. Subsequently, at line (ii), the property of the elementary symmetric function given in (7) was employed. Similar derivations ensue for the two denominator terms in (24) and the form (25) immediately follows.

#### APPENDIX B PROOF OF LEMMA 2

Under the simplifying assumptions of Section V-B, for a single Bernoulli  $l$  and the object-expected multisensor measurement  $\tilde{z}_k^{l, S_k}$ , the set of extended multisensor measurement indices becomes  $\mathbb{M}_k^{S_k} = \{-1\} \uplus \{0, 1\}^M$  and the set of valid multisensor maps becomes  $\bar{\mathcal{A}}_k^{S_k} = \mathcal{M}_k^{S_k}$  since there is no contention for measurements among multiple objects. Letting  $S_k = \{s_1, \dots, s_M\}$ , note that

$$\begin{aligned}
& \sum_{m \in \mathbb{M}_k^{S_k} \setminus \{-1\}} \bar{\psi}_k^{S_k}(\mathbf{x}, l; \tilde{z}_k^{l, S_k}, \mathbf{m}) \\
&= \sum_{m_1 \in \{0, 1\}} \dots \sum_{m_M \in \{0, 1\}} \prod_{i=1}^M \left\{ \left( 1 - P_{s_i, k}^D(\mathbf{x}, l) \right)^{1 - m_i} \right. \\
&\quad \left. \times \left[ \frac{P_{s_i, k}^D(\mathbf{x}, l) g_{s_i, k}(\tilde{z}_{s_i, k}^{(l)} | \mathbf{x}, l)} \right]^{m_i} \right\} \\
&= \prod_{i=1}^M \left[ 1 - P_{s_i, k}^D(\mathbf{x}, l) + \frac{P_{s_i, k}^D(\mathbf{x}, l) g_{s_i, k}(\tilde{z}_{s_i, k}^{(l)} | \mathbf{x}, l)}{f_{s_i, k}^{\text{FA}}(\tilde{z}_{s_i, k}^{(l)})} \right]
\end{aligned} \tag{50}$$

where a generalized form of the binomial theorem was employed and which leads to (30a). Combining (50) with the

linearity of the inner product and since each Bernoulli component  $l \in \mathbb{L}_k$  is updated independently (i.e., as if  $|\mathbb{L}_k| = 1$ ), the normalization constant of (18) now becomes

$$\begin{aligned}
C(l) &\triangleq \sum_{m \in \mathbb{M}_k^{S_k}} \eta_{k|k}^m(l; \mathcal{S}_{1:k}) \\
&= 1 - r_{k|k-1}^{S_{1:k-1}}(l) + r_{k|k-1}^{S_{1:k-1}}(l) \langle \bar{g}_k^{S_k}(\tilde{z}_k^{l, S_k} | \cdot, l), p_{k|k-1}^{S_{1:k-1}}(\cdot, l) \rangle.
\end{aligned}$$

Subsequently, from the definitions of (18)–(21b), the result of (50) and, since there is no contention, i.e.,  $P(\bar{\mathbf{a}}_k; \mathcal{S}_{1:k}) = \hat{P}_l(\bar{\mathbf{a}}_k(l); \mathcal{S}_{1:k})$ , the probability of existence becomes

$$\begin{aligned}
& \bar{r}_{k|k}^{S_{1:k}}(l) \\
&= \sum_{m \in \mathbb{M}_k^{S_k} \setminus \{-1\}} \frac{r_{k|k-1}^{S_{1:k-1}}(l) \langle \bar{\psi}_k^{S_k}(\cdot, l; \tilde{z}_k^{l, S_k}, \mathbf{m}), p_{k|k-1}^{S_{1:k-1}}(\cdot, l) \rangle}{C(l)} \\
&= \frac{r_{k|k-1}^{S_{1:k-1}}(l) \langle \bar{g}_k^{S_k}(\tilde{z}_k^{l, S_k} | \cdot, l), p_{k|k-1}^{S_{1:k-1}}(\cdot, l) \rangle}{1 - r_{k|k-1}^{S_{1:k-1}}(l) + r_{k|k-1}^{S_{1:k-1}}(l) \langle \bar{g}_k^{S_k}(\tilde{z}_k^{l, S_k} | \cdot, l), p_{k|k-1}^{S_{1:k-1}}(\cdot, l) \rangle}
\end{aligned}$$

while from (30c) is obtained by using the notation in (30b). Similarly, from definition (21b) and by using (30a), the probability density of (30d) follows. Additionally, note that the resulting equations (30c) and (30d) are similar to the multisensor single-Bernoulli filter update equations of [71].

#### ACKNOWLEDGMENT

The authors wish to thank A. Davis and S. Van Liew of Draper Laboratory for pointing to the current fundamental research challenges in maritime localization and tracking for Smart Oceans. The initial discussion with Draper Laboratory inspired the emulation of the scenarios presented in Section VII-C. The authors also would like to thank R. Cohen and D. Gaglione for their helpful suggestions.

#### REFERENCES

- [1] J. Waterston. (2017). *Ocean of Things*. Accessed: Apr. 25, 2020. [Online]. Available: <https://www.darpa.mil/program/ocean-of-things>
- [2] A. Zanello, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [3] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*. New York, NY, USA: Springer, 1988.
- [4] M. Z. Win, P. C. Pinto, A. Giorgetti, M. Chiani, and L. A. Shepp, "Error performance of ultrawideband systems in a Poisson field of narrowband interferers," in *Proc. IEEE Int. Symp. Spread Spectr. Technol. Appl.*, Manaus, Brazil, Aug. 2006, pp. 410–416.
- [5] M. Z. Win, "A mathematical model for network interference," in *Proc. IEEE Commun. Theory Workshop*, Sedona, AZ, USA, May 2007, pp. 99–110.
- [6] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proc. IEEE*, vol. 97, no. 2, pp. 205–230, Feb. 2009.
- [7] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless network," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.
- [8] D. Dardari, A. Conti, C. Buratti, and R. Verdone, "Mathematical evaluation of environmental monitoring estimation error through energy-efficient wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 7, pp. 790–802, Jul. 2007.
- [9] A. Rabbachin, A. Conti, and M. Z. Win, "Wireless network intrinsic secrecy," *IEEE/ACM Trans. Netw.*, vol. 23, no. 1, pp. 56–69, Feb. 2015.
- [10] F. Zabini and A. Conti, "Inhomogeneous Poisson sampling of finite-energy signals with uncertainties in  $\mathbb{R}^d$ ," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4679–4694, Sep. 2016.

- [11] H. ElSawy, A. Sultan-Salem, M.-S. Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 167–203, 1st Quart., 2017.
- [12] A.-A. Saucan, T. Chonavel, C. Sintes, and J.-M. L. Caillec, "CPHD-DOA tracking of multiple extended sonar targets in impulsive environments," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1147–1160, Mar. 2016.
- [13] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.
- [14] M. Z. Win *et al.*, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.
- [15] A. A. Saucan, M. J. Coates, and M. Rabbat, "A multisensor multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5495–5509, Oct. 2017.
- [16] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.
- [17] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018.
- [18] U. A. Khan, S. Kar, and J. M. F. Moura, "Distributed sensor localization in random environments using minimal number of anchor nodes," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 2000–2016, May 2009.
- [19] S. Safavi, U. A. Khan, S. Kar, and J. M. F. Moura, "Distributed localization: A linear theory," *Proc. IEEE*, vol. 106, no. 7, pp. 1204–1223, Jul. 2018.
- [20] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, "Efficient multi-sensor localization for the Internet-of-Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 153–167, Sep. 2018.
- [21] R. Niu, A. Vempaty, and P. K. Varshney, "Received-signal-strength-based localization in wireless sensor networks," *Proc. IEEE*, vol. 106, no. 7, pp. 1166–1182, Jul. 2018.
- [22] F. Meyer *et al.*, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.
- [23] P. Sharma, A.-A. Saucan, D. J. Bucci, and P. K. Varshney, "Decentralized Gaussian filters for cooperative self-localization and multi-target tracking," *IEEE Trans. Signal Process.*, vol. 67, no. 22, pp. 5896–5911, Nov. 2019.
- [24] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.
- [25] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Syst. Mag.*, vol. 29, no. 6, pp. 82–100, Dec. 2009.
- [26] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3246–3260, Jun. 2014.
- [27] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.
- [28] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," *Automatica*, vol. 48, no. 10, pp. 2482–2493, 2012.
- [29] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Optimal periodic sensor scheduling in networks of dynamical systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3055–3068, Jun. 2014.
- [30] C. Li and N. Elia, "Stochastic sensor scheduling via distributed convex optimization," *Automatica*, vol. 58, pp. 173–182, Aug. 2015.
- [31] J. Haugen and L. Imsland, "Monitoring moving objects using aerial mobile sensors," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 2, pp. 475–486, Mar. 2016.
- [32] K. Zhou and S. I. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 678–695, Aug. 2011.
- [33] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1521–1529, Apr. 2011.
- [34] W. Burgard, D. Fox, and S. Thrun, "Active mobile robot localization by entropy minimization," in *Proc. EUROBOT*, 1997, pp. 155–162.
- [35] A. K. Gostar, R. Hoseinnezhad, T. Rathnayake, X. Wang, and A. Bab-Hadiashar, "Constrained sensor control for labeled multi-Bernoulli filter using Cauchy–Schwarz divergence," *IEEE Signal Process. Lett.*, vol. 24, no. 9, pp. 1313–1317, Sep. 2017.
- [36] H. G. Hoang, B.-N. Vo, B.-T. Vo, and R. Mahler, "The Cauchy–Schwarz divergence for Poisson point processes," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4475–4485, Aug. 2015.
- [37] 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.
- [38] H. V. 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.
- [39] M. Z. Win, G. Chrisikos, and N. R. Sollenberger, "Performance of Rake reception in dense multipath channels: Implications of spreading bandwidth and selection diversity order," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1516–1525, Aug. 2000.
- [40] M. Z. Win, N. C. Beaulieu, L. A. Shepp, B. F. Logan, and J. H. Winters, "On the SNR penalty of MPSK with hybrid selection/maximal ratio combining over IID Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 1012–1023, Jun. 2003.
- [41] A. Conti, W. M. Gifford, M. Z. Win, and M. Chiani, "Optimized simple bounds for diversity systems," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2674–2685, Sep. 2009.
- [42] S. Bartoletti, A. Giorgetti, M. Z. Win, and A. Conti, "Blind selection of representative observations for sensor radar networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1388–1400, Apr. 2015.
- [43] A. Bletsas, H. Shin, and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, Sep. 2007.
- [44] A. K. Gostar, R. Hoseinnezhad, A. Bab-Hadiashar, and W. Liu, "Sensor-management for multitarget filters via minimization of posterior dispersion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 6, pp. 2877–2884, Dec. 2017.
- [45] F. Xing, H. Yin, X. Ji, and V. C. M. Leung, "Joint relay selection and power allocation for underwater cooperative optical wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 251–264, Jan. 2020.
- [46] L. Liu, M. Ma, C. Liu, and Y. Shu, "Optimal relay node placement and flow allocation in underwater acoustic sensor networks," *IEEE Trans. Wireless Commun.*, vol. 65, no. 5, pp. 2141–2152, May 2017.
- [47] M. T. R. Khan, S. H. Ahmed, and D. Kim, "AUV-aided energy-efficient clustering in the Internet of underwater things," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 4, pp. 1132–1141, Dec. 2019.
- [48] D. Mušicki and R. Evans, "Joint integrated probabilistic data association: JIPDA," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 40, no. 3, pp. 1093–1099, Jul. 2004.
- [49] 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.
- [50] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. Chichester, U.K.: Wiley, 1996.
- [51] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.
- [52] I. Molchanov, *Theory of Random Sets*. London, U.K.: Springer, 2005.
- [53] B.-N. Vo, S. R. Sigh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.
- [54] J. L. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2942–2959, Oct. 2014.
- [55] 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.
- [56] K. Kampa, E. Hasanbelliu, and J. C. Principe, "Closed-form Cauchy–Schwarz PDF divergence for mixture of Gaussians," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2011, pp. 2578–2585.
- [57] A. K. Gostar, R. Hoseinnezhad, and A. Bab-Hadiashar, "Multi-Bernoulli sensor control using Cauchy–Schwarz divergence," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2016, pp. 651–657.
- [58] P. M. Djuric *et al.*, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.
- [59] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2602–2612, Oct. 2003.
- [60] B.-N. Vo and B.-T. Vo, "Multi-sensor multi-object tracking with the generalized labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 5952–5967, Dec. 2019.
- [61] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," *Eur. J. Oper. Res.*, vol. 99, no. 1, pp. 89–112, 1996.

- [62] R. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodol. Comput. Appl. Probab.*, vol. 1, no. 2, pp. 127–190, Sep. 1999.
- [63] A. Costa, O. D. Jones, and D. Kroese, "Convergence properties of the cross-entropy method for discrete optimization," *Oper. Res. Lett.*, vol. 35, no. 5, pp. 573–580, Sep. 2007.
- [64] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. New York, NY, USA: Springer-Verlag, 2004.
- [65] D. M. Nguyen, H. A. L. Thi, and T. P. Dinh, "Solving the multidimensional assignment problem by a cross-entropy method," *J. Comb. Optim.*, vol. 27, no. 4, pp. 808–823, 2014.
- [66] J. Y. Yu, A. A. Saucan, M. Coates, and M. Rabbat, "Algorithms for the multi-sensor assignment problem in the  $\delta$ -generalized labeled multi-Bernoulli filter," in *Proc. Int. Workshop Comput. Adv. Multi Sensor Adapt. Process. (CAMSAP)*, Dec. 2017, pp. 1–5.
- [67] A. A. Saucan and P. K. Varshney, "Distributed cross-entropy  $\delta$ -GLMB filter for multi-sensor multi-target tracking," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, 2018, pp. 1559–1566.
- [68] R. Y. Rubinstein, "Cross-entropy and rare events for maximal cut and partition problems," *ACM Trans. Model. Comput. Simulat.*, vol. 12, no. 1, pp. 27–53, Jan. 2002.
- [69] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Algorithms and Software for Information Extraction. Hlerampao*. Hoboken, NJ, USA: Wiley, Jul. 2001.
- [70] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.
- [71] B. T. 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.



**Augustin A. Saucan** (Member, IEEE) received the Ph.D. and M.Sc. in electrical engineering from Institut Mines-Télécom Atlantique (formerly Télécom Bretagne), Brest, France, in 2016 and 2012, respectively.

He is a postdoctoral associate at the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology (MIT). Prior to joining MIT, he was a postdoctoral researcher at Syracuse University in 2017–2019 and at the McGill University in 2016–2017. His research interests

includes wireless networks and IoT with emphasis on statistical signal processing, localization and tracking, reinforcement learning, and finite set statistics.

Dr. Saucan is an active member of the IEEE Communications Society and served as a reviewer for several journals, including IEEE Transaction on Signal Processing, IEEE Transactions on Aerospace and Electronic Systems, and IEEE Transactions on Signal and Information Processing over Networks.



**Moe Z. Win** (Fellow, IEEE) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyoo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Università degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.