IEEE Open Journal of Signal Processing

Received 24 November 2021; accepted 15 February 2022. Date of publication 4 March 2022;
date of current version 18 April 2022. The review of this article was arranged by Associate Editor Eric L. Miller.

Digital Object Identifier 10.1109/OJSP.2022.3154684

Cooperative Localization and Multitarget Tracking in Agent Networks with the Sum-Product Algorithm

MATTIA BRAMBILLA [1] (Member, IEEE), DOMENICO GAGLIONE [3], GIOVANNI SOLDI [3], RICO MENDRZIK [4], GABRIELE FERRI [3] (Senior Member, IEEE), KEVIN D. LEPAGE [3], MONICA NICOLI [2] (Senior Member, IEEE), PETER WILLETT [5] (Fellow, IEEE), PAOLO BRACA [3] (Senior Member, IEEE), AND MOE Z. WIN [6] (Fellow, IEEE)

[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, 20133 Milan, Italy
[2]Dipartimento di Ingegneria Gestionale (DIG), Politecnico di Milano, 20156 Milan, Italy
[3]NATO STO Centre for Maritime Research and Experimentation (CMRE), 19126 La Spezia, Italy
[4]Ibeo Automotive Systems GmbH, 22143 Hamburg, Germany
[5]University of Connecticut, Storrs, CT 06269 USA
[6]Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA

CORRESPONDING AUTHOR: MATTIA BRAMBILLA (e-mail: mattia.brambilla@polimi.it)

This work was supported in part by the NATO Allied Command Transformation (ACT) under the DKOE Project and in part by the U.S. Army Research Office through the MIT Institute for Soldier Nanotechnologies under Contract W911NF-13-D-0001. Parts of this paper were previously presented at ICASSP 2020, Barcelona, Spain, May 2020.

ABSTRACT This paper addresses the problem of multitarget tracking using a network of sensing agents with unknown positions. Agents have to both localize themselves in the sensor network and, at the same time, perform multitarget tracking in the presence of clutter and miss detection. These two problems are jointly resolved using a holistic and centralized approach where graph theory is used to describe the statistical relationships among agent states, target states, and observations. A scalable message passing scheme, based on the sum-product algorithm, enables to efficiently approximate the marginal posterior distributions of both agent and target states. The proposed method is general enough to accommodate a full multistatic network configuration, with multiple transmitters and receivers. Numerical simulations show superior performance of the proposed joint approach with respect to the case in which cooperative self-localization and multitarget tracking are performed separately, as the former manages to extract valuable information from targets. Lastly, data acquired in 2018 by the NATO Science and Technology Organization (STO) Centre for Maritime Research and Experimentation (CMRE) through a network of autonomous underwater vehicles demonstrates the effectiveness of the approach in a practical application.

INDEX TERMS Belief propagation, factor graph, maritime surveillance, message passing, probabilistic data association.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

Detecting unknown *targets*, understanding their intentions, and taking reactive countermeasures are common tasks in situational awareness (SA) applications [1]–[7]. Depending on the specific use case, different types of sensors (acoustic, radio frequency, optical, etc. [8]) may be used to sense the environment and provide the desired information. Most of SA applications use multiple cooperative sensors, rather than a single one, to infer the presence and kinematics of targets. Indeed, cooperation dramatically increases the perception capabilities of an SA system, as it relies on a larger dataset of observations (or *measurements*) of the targets [9]–[13]. Examples can be found in several domains such as underwater surveillance networks [14]–[17], connected vehicles [18]–[20], and Internet of Things (IoT) [21]–[24]. Mobility of sensors can further improve the performance of target detection and localization by fusing spatial sensing

under different geometries, also enabling the design of optimized sensor trajectories [25]. However, this requires the sensors to localize themselves continuously. Cooperative self-localization techniques based on belief propagation, also known as the sum-product algorithm (SPA) [26], [27], have been recently proposed, with computational complexity that linearly scales with the number of cooperative sensors [28]–[34]. Additional advantages of SPA methods include the ability to address non-linear and non-Gaussian models and to cope with unknown and time-varying hyperparameters [35].

The advanced capability of a surveillance system to firstly detect and localize, and then track over time a number of hypothesized targets which behave as non-cooperative entities (i.e., that do not deliberately share information with the surveillance system) is referred to as multitarget tracking. Usually, the presence of these targets represents a dangerous situation or a potential threat; e.g., targets can be vulnerable road users in a vehicular scene, intruding ships in the maritime domain, intruding aircraft in the aerospace domain, or thieves in an IoT surveillance system. It follows that the development of robust, reliable, scalable and efficient multitarget tracking algorithms becomes of paramount importance, as safety issues are involved. Abundant literature on multitarget tracking is available, starting from the pioneering works in [36]–[39] through very recent studies such as [40]–[48]. Approaches based on SPA have been proposed as well, both with stationary sensors — whose location is either known [35], [49] or unknown [50] — and with mobile ones [18], [51]–[53]. However, not all of them handle typical multitarget tracking challenges like the presence of clutter-generated measurements (i.e., *false alarms*), missed detections, and measurement origin uncertainty [54], i.e., the problem of unknown association between targets and measurements. Focusing on multitarget tracking algorithms with mobile sensors, the cited works are affected by the following limitations: in [51] sensors do not localize themselves cooperatively; in [52] the maximum number of targets that can be tracked simultaneously is limited and needs to be set a priori; in [18] and [53] the number of targets is time-invariant and known, and, in addition, in [18] neither false alarms nor missed detections are considered, and in [53] the association between targets and measurements is assumed known. Random finite sets (RFSs) constitute an alternative framework for the development of multitarget tracking methods[1] both with stationary [40], [41], [48] and mobile sensors [45]. In particular, in [45] the authors develop a Poisson multi-Bernoulli multitarget tracking filter that jointly estimates the uncertain mobile sensor states and target states using two types of measurements: sensor state measurements, e.g., global navigation satellite system (GNSS) measurements, and target measurements. However, this algorithm, as well as those cited above, are not suitable for a full multistatic network configuration, and do not consider the case of signal reflections from mobile sensors, thus easing the data association problem.

The above issues have been partially addressed in [55], that indeed represents the preliminary study at the basis of this research; it is purpose of this paper to further extend that work as detailed in the next subsection.

B. CONTRIBUTIONS AND PAPER ORGANIZATION

The method we present here is based on a general framework where the concept of *agents*, rather than sensors, is introduced to address the SA task. An agent is a device with sensing and communication functionalities (i.e., transmitting and/or receiving acoustic, radio, or optical signals), along with motion and navigation capabilities. The connectivity is used to set up a cooperative and centralized processing platform. We propose a SPA-based technique that extends the state-of-the-art methods by combining cooperative self-localization and multitarget tracking in a unified centralized framework. In particular, moving agents, whose states are unknown, are capable of jointly localizing themselves by continuously estimating their states and, at the same time, detecting and tracking an unknown, arbitrary, and time-varying number of targets by exploiting multiple types of measurements and in presence of clutter, miss detection and association uncertainty. A fully distributed approach based on consensus strategies [52], [56]–[60] can be adopted and customized for the proposed scheme; this study is not included here and left to future work. We focus the attention on a holistic and centralized approach for cooperative self-localization and multitarget tracking, validated by real world experiments. The main contributions of this paper, which advances the work in [55] where an ad-hoc scenario with defined roles for transmitter and receiver agents is considered, are the following:

- a more general formulation is provided in which moving agents can both sense the environment, thus producing multiple types of measurements, and communicate with each other;
- all *objects*, i.e., both agents and targets, whose states are unknown and need to be estimated, can reflect signals transmitted by a certain agent and thus produce measurements. Therefore, the data association problem is not limited to targets only as in [18], [51]–[53], [55], but it involves agents as well;
- the factor graph underlying the stochastic problem formulation is carefully derived and all the SPA messages are detailed;
- the proposed algorithm is validated in a real underwater scenario using data acquired by a network of autonomous vehicles.

The proposed SPA-based cooperative self-localization and multitarget tracking algorithm inherits the low computational complexity of the SPA-based multitarget tracking algorithm developed in [49], that scales linearly in the number of sensors and quadratically in the number of targets; the difference with [49] is that the number of sensors is here the number of transmitter-receiver pairs.

[1]For the interested reader, similarities and differences between the SPA-based and the RFS-based derivation of multitarget tracking algorithms are described in [49].

The remainder of the paper is organized as follows. Section II describes the scenario we are considering, the related mathematical representation, and the connection with practical use cases. The joint cooperative self-localization and multitarget tracking problem is formulated in Section III. Section IV details the proposed SPA-based algorithm, which is assessed using simulated and real data in Section V. Finally, concluding remarks are drawn in Section VI.

C. NOTATION

Throughout this paper, column vectors are denoted by boldface lower-case letters (e.g., $\boldsymbol{a}$) and matrices by boldface upper-case letters (e.g., $\boldsymbol{A}$). $\mathbf{I}$ and $\mathbf{0}$ denote the identity matrix and the vector of all zeros, respectively, with the size determined by the subscript or from the context. We write $\mathrm{diag}(a_1, \ldots, a_N)$ for an $N \times N$ diagonal matrix with diagonal entries $a_1, \ldots, a_N$. The transpose of a matrix $\boldsymbol{A}$ is written as $\boldsymbol{A}^{\mathrm{T}}$. The Euclidean norm of vector $\boldsymbol{a}$ is denoted by $\|\boldsymbol{a}\|$. For a two-dimensional (2D) vector $\boldsymbol{a}$, $\angle\boldsymbol{a}$ is the angle defined clockwise and such that $\angle\boldsymbol{a}=0$ for $\boldsymbol{a} = [0\ 1]^{\mathrm{T}}$. Sets are denoted by calligraphic letters (e.g., $\mathcal{A}$) and $|\mathcal{A}|$ indicates the cardinality of the set. The symbol $\propto$ denotes equality up to a constant factor. The Dirac delta function is denoted with $\delta(\cdot)$; the Kronecker delta is denoted with $\delta_{a,b}$, and is equal to 1 if $a = b$, and 0 otherwise.

II. PROBLEM DESCRIPTION AND SYSTEM MODEL

Hereafter, we provide a high-level description of the scenario under consideration. Let us suppose to have a set of agents and that each agent is equipped with an on-board device that provides noisy (and possibly incomplete) observations of the agent's own state, referred to as *navigation data*, and with a transmitter and/or a receiver. For convenience, we will refer to an agent equipped with a transmitter as a *Tx-agent*, and to an agent equipped with a receiver as an *Rx-agent*; if an agent is equipped with both a transmitter and a receiver (i.e., a transceiver), we will use Tx-agent or Rx-agent depending on its role in each specific context. The transmitter is used to broadcast a signal as, for example, an acoustic signal used in sonar, or an electromagnetic signal used in radar; we assume that each agent is aware of the signal transmitted by any other agent and that all these signals are orthogonal in some domain (time, frequency, or code) so that interferences can be neglected.

The signal broadcast by a Tx-agent and received by an Rx-agent can be used by the latter to extract a noisy and incomplete observation of the state of the Tx-agent; this type of measurement is referred to as *inter-agent* measurement. It is out of the scope of this paper to describe how this measurement is obtained, however we here provide few examples: the information about the Tx-agent state can be directly encoded into the transmitted signal and retrieved by the Rx-agent; or the Rx-agent can compute the time-of-arrival (ToA) or the angle-of-arrival (AoA) of the received signal, thus obtaining an observation of the Tx-agent state.

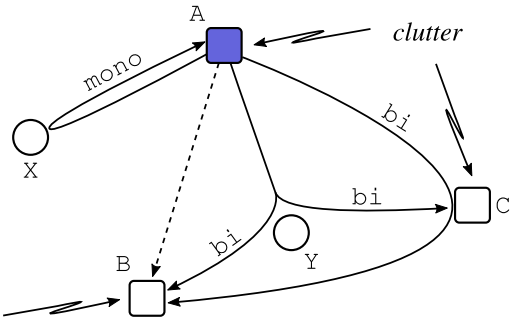


FIGURE 1. Exemplary illustration of the considered scenario. Agents are depicted as squares, targets as circles. Agent A (purple square) is equipped with a transceiver; agents B and C (white squares) are equipped with receivers only. Dashed lines represent inter-agent measurements; solid lines represent MOT measurements, either monostatic or bistatic; solid zig-zag lines represent clutter-generated MOT measurements.

The signal broadcast by a Tx-agent may also reach an Rx-agent after being reflected by a target or another agent present in the scene; in this case, the received signal can be used by the Rx-agent to extract a noisy and incomplete observation of the object (i.e., either target or agent) that caused the reflection. This type of measurement is referred to as *multi-object tracking* (MOT) measurement; as before, this can be obtained by computing, for example, the ToA or the AoA of the received signal. If Tx-agent and Rx-agent coincide, then this MOT measurement is acquired in a *monostatic* configuration; otherwise, if Tx-agent and Rx-agent are different, the MOT measurement is acquired in a *bistatic* configuration. Note that an MOT measurement can be clutter-generated if not caused by the reflection from an object; finally, we here assume that an Rx-agent is able to distinguish between inter-agent and MOT measurements, however it is not able to distinguish if an MOT measurement is originated by a target, an agent, or clutter.

Fig. 1 illustrates an exemplary scenario with three agents A, B, and C, depicted as squares, and two targets X and Y, depicted as circles. Agent A is equipped with a transceiver, and agents B and C are equipped with a receiver only. Tx-agent A informs Rx-agent B of its own location by an inter-agent measurement, represented by means of a dashed line; as mentioned before, the information on A's position might be encoded within the signal, or extracted from the signal by B. Moreover, the signal transmitted by Tx-agent A bounces off target Y and agent C, and is received by Rx-agent B. Therefore, B has available an inter-agent and two bistatic MOT measurements — each sketched as a solid line labeled as "bi" — due to the signal transmitted by Tx-agent A. However, agent B is unable to a priori associate the two MOT measurements to Y and C, respectively; this measurement origin uncertainty needs to be handled. Finally, we observe that the signal reflected by target Y also generates a bistatic MOT measurement at Rx-agent C, and that the signal that bounces off target X is reflected back to Rx-agent A, thus generating a monostatic MOT measurement, pictured as a solid line labeled as "mono". Note that Fig. 1 depicts one possible instance of

inter-agent and MOT measurements produced in this configuration of agents and targets, and that many others are possible.

A first objective of this paper is to show that the challenging bistatic geometry can be fruitfully exploited to enhance the localization of agents along with the detection and tracking of targets. This requires a mathematical representation of the agents, the targets, and the available observations, as detailed in the following subsections.

A. AGENT STATE VECTOR, AGENT PAIRS, AND POTENTIAL TARGETS

Let $\mathcal{A} \triangleq \{1, \ldots, A\}$ be the set of agents, whose cardinality A is known and time-invariant.[2] The state (e.g., position, velocity, heading) of agent $a \in \mathcal{A}$ at time step $t = 1, 2, \ldots$ is represented by the vector $\boldsymbol{s}_{a,t}$ whose evolution in time is given by the following kinematic model

$$\boldsymbol{s}_{a,t} = \boldsymbol{\varepsilon}_a\left(\boldsymbol{s}_{a,t-1}, \boldsymbol{u}_{a,t}\right), \tag{1}$$

where $\boldsymbol{u}_{a,t}$ is a process noise independent across a and t that accounts for the motion uncertainty of the agent [54, Sec. 1.5]. The function $\boldsymbol{\varepsilon}_a(\cdot)$ and the statistics of $\boldsymbol{u}_{a,t}$ define the agent state transition pdf $\tau_a(\boldsymbol{s}_{a,t}|\boldsymbol{s}_{a,t-1})$. We denote the joint agent state vector at time t as $\boldsymbol{s}_t \triangleq [\boldsymbol{s}_{1,t}^{\mathrm{T}}, \ldots, \boldsymbol{s}_{A,t}^{\mathrm{T}}]^{\mathrm{T}}$, and the joint agent state vector at all times as $\boldsymbol{s}_{1:t} \triangleq [\boldsymbol{s}_1^{\mathrm{T}}, \ldots, \boldsymbol{s}_t^{\mathrm{T}}]^{\mathrm{T}}$.

We indicate with $\mathcal{T} \subseteq \mathcal{A}$ the set of Tx-agents, and with $\mathcal{R} \subseteq \mathcal{A}$ the set of Rx-agents. Note that $\mathcal{T} \cup \mathcal{R} = \mathcal{A}$, and that $\mathcal{T} \cap \mathcal{R}$ represents the set of agents equipped with both a transmitter and a receiver. When $|\mathcal{T}| = 1$ and $|\mathcal{R}| \geq 1$ the *network* configuration is referred to as bistatic, otherwise when $|\mathcal{T}| > 1$ and $|\mathcal{R}| \geq 1$ the *network* is referred to as multistatic. We formally consider the Cartesian product set $\mathcal{R} \times \mathcal{T}$ of all the possible pairs (j_1, j_2) such that $j_1 \in \mathcal{R}$ and $j_2 \in \mathcal{T}$.

We observe that agents j_1 and j_2 might also coincide. To establish an (arbitrary) order among the agent pairs, we introduce the index set $\mathcal{J} \triangleq \{1, \ldots, J\}$, with $J = |\mathcal{R}||\mathcal{T}|$, and define an indexing function $\phi : \mathcal{J} \to \mathcal{R} \times \mathcal{T}$, such that $\phi(j)$ represents the j-th agent pair (j_1, j_2). This order, though arbitrary, is used to sequentially process the MOT measurements collected by the agent pairs as described later.

Furthermore, as done in [49], we account for a time-varying unknown number of targets by introducing the concept of *potential target* (PT). The set of PTs at time t is $\mathcal{K}_t \triangleq \{1, \ldots, K_t\}$; the existence of PT $k \in \mathcal{K}_t$ at time t is indicated by the binary variable $r_{k,t} \in \{0, 1\}$, i.e., $r_{k,t} = 1$ if the PT exists and $r_{k,t} = 0$ otherwise, and the state (e.g., position and velocity) of PT $k \in \mathcal{K}_t$ is denoted as $\boldsymbol{x}_{k,t}$, and is formally considered also if $r_{k,t} = 0$. We combine the state and existence variables of PT k into the *augmented* state vector $\boldsymbol{y}_{k,t} \triangleq [\boldsymbol{x}_{k,t}^{\mathrm{T}}, r_{k,t}]^{\mathrm{T}}$, and define the joint vector of all the PTs at time t as $\boldsymbol{y}_t \triangleq [\boldsymbol{y}_{1,t}^{\mathrm{T}}, \ldots, \boldsymbol{y}_{K_t,t}^{\mathrm{T}}]^{\mathrm{T}}$. We observe that the states $\boldsymbol{x}_{k,t}$ of nonexisting PTs (i.e., for which $r_{k,t} = 0$) are obviously irrelevant; thus, all the pdfs defined for the PT augmented states, i.e., $f(\boldsymbol{y}_{k,t}) = f(\boldsymbol{x}_{k,t}, r_{k,t})$, are such that

$$f(\boldsymbol{x}_{k,t}, r_{k,t} = 0) = f_{k,t}\, f_{\mathrm{D}}(\boldsymbol{x}_{k,t}),$$

where $f_{k,t} \in [0, 1]$ is a constant, and $f_{\mathrm{D}}(\boldsymbol{x}_{k,t})$ is an arbitrary "dummy pdf".

B. OBSERVATIONS

As stated above, at any time t an agent $a \in \mathcal{A}$ might collect navigation data from an on-board device, and produce two kinds of measurements: inter-agent and MOT measurements.

1) NAVIGATION DATA

The navigation data $\boldsymbol{g}_{a,t}$ collected by agent $a \in \mathcal{A}$ at time t is an observation made by a of its own state, e.g., acquired with an on-board system, such as GNSS or inertial navigation system (INS). It is modeled as

$$\boldsymbol{g}_{a,t} = \boldsymbol{\theta}_a\left(\boldsymbol{s}_{a,t}, \boldsymbol{n}_{a,t}\right), \tag{2}$$

where $\boldsymbol{n}_{a,t}$ is a noise term, independent across a and t, modeling the finite accuracy of the on-board system. Note that also the measurement model $\boldsymbol{\theta}_a(\cdot)$ depends on the agent a, since agents may be equipped with different types of on-board systems (e.g., GNSS or INS). The function $\boldsymbol{\theta}_a(\cdot)$ and the statistics of $\boldsymbol{n}_{a,t}$ define the likelihood function $\mathfrak{g}_a(\boldsymbol{g}_{a,t}|\boldsymbol{s}_{a,t})$. We indicate with $\mathcal{A}_t^{\boldsymbol{g}} \subseteq \mathcal{A}$ the set of agents that have navigation data available at time t, and we define the stacked vector of all navigation data from all agents at time t as $\boldsymbol{g}_t$, and the stacked vector of all navigation data from all agents at all times as $\boldsymbol{g}_{1:t} \triangleq [\boldsymbol{g}_1^{\mathrm{T}}, \ldots, \boldsymbol{g}_t^{\mathrm{T}}]^{\mathrm{T}}$.

2) INTER-AGENT MEASUREMENTS

The inter-agent measurement produced at time t by Rx-agent $a \in \mathcal{R}$ using the signal transmitted by Tx-agent $a' \in \mathcal{T} \setminus \{a\}$ is modeled as

$$\boldsymbol{\rho}_t^{(a,a')} = \boldsymbol{\vartheta}\left(\boldsymbol{s}_{a,t}, \boldsymbol{s}_{a',t}, \boldsymbol{w}_t^{(a,a')}\right), \tag{3}$$

where $\boldsymbol{w}_t^{(a,a')}$ is an inter-agent measurement noise term independent across a, a', and t. The function $\boldsymbol{\vartheta}(\cdot)$ and the statistics of $\boldsymbol{w}_t^{(a,a')}$ define the likelihood function $\mathfrak{d}(\boldsymbol{\rho}_t^{(a,a')}|\boldsymbol{s}_{a,t}, \boldsymbol{s}_{a',t})$. We indicate with $\mathcal{R}_t^{(a)} \subseteq \mathcal{R} \setminus \{a\}$ the set of Rx-agents that produce an inter-agent measurement using the signal transmitted by Tx-agent a at time t, and with $\mathcal{T}_t^{(a)} \subseteq \mathcal{T} \setminus \{a\}$ the set of Tx-agents that provide an inter-agent measurement to Rx-agent a at time t. Moreover, we define the stacked vector of all inter-agent measurements acquired by Rx-agent a at time t as $\boldsymbol{\rho}_t^{(a)}$, the stacked vector of all inter-agent measurements acquired by all Rx-agents at time t as $\boldsymbol{\rho}_t$, and the vector of all inter-agent measurements acquired by all Rx-agents at all times as $\boldsymbol{\rho}_{1:t} \triangleq [\boldsymbol{\rho}_1^{\mathrm{T}}, \ldots, \boldsymbol{\rho}_t^{\mathrm{T}}]^{\mathrm{T}}$. We remark that, since the Rx-agent a is aware of the signal transmitted by the Tx-agent a', there is no uncertainty on the origin of the inter-agent measurement $\boldsymbol{\rho}_t^{(a,a')}$, thus no data association is required.

[2] Note that the number A of agents may also be modeled as being time-variant — yet, known —, however this would not change the overall theoretical framework.

3) MOT MEASUREMENTS

Let us consider the j-th agent pair (j_1, j_2), with $j \in \mathcal{J}$, $j_1 \in \mathcal{R}$, and $j_2 \in \mathcal{T}$, such that[3] $\phi(j) = (j_1, j_2)$, and assume that Rx-agent j_1 produces $M_t^{(j)}$ MOT measurements from the signal broadcast by the Tx-agent j_2; we indicate with $\boldsymbol{z}_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)} \triangleq \{1, \ldots, M_t^{(j)}\}$, the m-th MOT measurement produced by the j-th agent pair at time t. We recall that an MOT measurement can be generated by the signal transmitted by Tx-agent j_2 reflecting off either an existing PT (i.e., for which $r_{k,t} = 1$) or another agent (i.e., other than j_1 and j_2), or it can be generated by clutter; and that it can be either bistatic, if $j_1 \neq j_2$, or monostatic, if $j_1 = j_2$. An MOT measurement $\boldsymbol{z}_{m,t}^{(j)}$ generated by PT k is modeled as

$$\boldsymbol{z}_{m,t}^{(j)} = \gamma_{\texttt{mono}}\left(\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{v}_{m,t}^{(j)}\right) \tag{4}$$

if acquired in a monostatic configuration; and modeled as

$$\boldsymbol{z}_{m,t}^{(j)} = \gamma_{\texttt{bi}}\left(\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}, \boldsymbol{v}_{m,t}^{(j)}\right) \tag{5}$$

if acquired in a bistatic configuration, where $\boldsymbol{v}_{m,t}^{(j)}$ is a noise term independent across j, m and t. The functions $\gamma_{\texttt{mono}}(\cdot)$ and $\gamma_{\texttt{bi}}(\cdot)$, and the statistics of $\boldsymbol{v}_{m,t}^{(j)}$ define the likelihood functions $f_{\texttt{mono}}(\boldsymbol{z}_{m,t}^{(j)}|\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t})$ and $f_{\texttt{bi}}(\boldsymbol{z}_{m,t}^{(j)}|\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t})$, respectively. In case the MOT measurement $\boldsymbol{z}_{m,t}^{(j)}$ is generated by an agent $j' \neq j_1, j_2$, the models in (4)–(5) still apply as long as $\boldsymbol{x}_{k,t}$ is replaced by $\boldsymbol{s}_{j',t}$.

For convenience, we stack all MOT measurements produced at agent pair j at time t into the vector $\boldsymbol{z}_t^{(j)} \triangleq [\boldsymbol{z}_{1,t}^{(j)\mathrm{T}}, \ldots, \boldsymbol{z}_{M_t^{(j)},t}^{(j)\mathrm{T}}]^{\mathrm{T}}$, all MOT measurements produced at all agent pairs at time t into the vector $\boldsymbol{z}_t \triangleq [\boldsymbol{z}_t^{(1)\mathrm{T}}, \ldots, \boldsymbol{z}_t^{(J)\mathrm{T}}]^{\mathrm{T}}$, and all MOT measurements produced at all agent pairs at all times into the vector $\boldsymbol{z}_{1:t} \triangleq [\boldsymbol{z}_1^{\mathrm{T}}, \ldots, \boldsymbol{z}_t^{\mathrm{T}}]^{\mathrm{T}}$. Furthermore, we define the vector of numbers of MOT measurements produced at all agent pairs at time t as $\boldsymbol{m}_t = [M_t^{(1)}, \ldots, M_t^{(J)}]^{\mathrm{T}}$, and the vector of numbers of MOT measurements produced at all agent pairs at all times as $\boldsymbol{m}_{1:t} \triangleq [\boldsymbol{m}_1^{\mathrm{T}}, \ldots, \boldsymbol{m}_t^{\mathrm{T}}]^{\mathrm{T}}$. Finally, we recall that the MOT measurements, unlike inter-agent measurements, have *unknown* origins, i.e., it is unknown if a given MOT measurement is generated by an object — either target or agent — and by which object.

C. LEGACY PTS AND NEW PTS

Following [49, Sec. VIII-B], each PT at time t and agent pair j is either a "legacy" PT or a "new" PT. A legacy PT is a PT that has already been introduced in the past, either at current time t at any previous agent pair $j' < j$, or at any previous time $t' < t$. We denote with $\mathcal{L}_t^{(j)} \triangleq \{1, \ldots, L_t^{(j)}\}$ the set of $L_t^{(j)}$ legacy PTs at time t at agent pair j, and indicate with $\underline{\boldsymbol{y}}_{\ell,t}^{(j)} \triangleq [\underline{\boldsymbol{x}}_{\ell,t}^{(j)\mathrm{T}}, \underline{r}_{\ell,t}^{(j)}]^{\mathrm{T}}$ the augmented state of legacy PT $\ell \in \mathcal{L}_t^{(j)}$, and with $\underline{\boldsymbol{y}}_t^{(j)} \triangleq [\underline{\boldsymbol{y}}_{1,t}^{(j)\mathrm{T}}, \ldots, \underline{\boldsymbol{y}}_{L_t^{(j)},t}^{(j)\mathrm{T}}]^{\mathrm{T}}$ the joint legacy PT augmented state vector.

New PTs model those targets that are detected for the first time by agent pair j at time t. Each new PT corresponds to an MOT measurement $\boldsymbol{z}_{m,t}^{(j)}$; therefore, the number of new PTs at time t at agent pair j is $M_t^{(j)}$. The augmented state of a new PT is denoted by $\overline{\boldsymbol{y}}_{m,t}^{(j)} \triangleq [\overline{\boldsymbol{x}}_{m,t}^{(j)\mathrm{T}}, \overline{r}_{m,t}^{(j)}]^{\mathrm{T}}$, $m \in \mathcal{M}_t^{(j)}$, and $\overline{r}_{m,t}^{(j)} = 1$ thus means that MOT measurement m was generated by a target that was never detected before, namely, a *newly detected* target. We define the joint augmented state vector of all new PTs at time t at agent pair j as $\overline{\boldsymbol{y}}_t^{(j)} \triangleq [\overline{\boldsymbol{y}}_{1,t}^{(j)\mathrm{T}}, \ldots, \overline{\boldsymbol{y}}_{M_t^{(j)},t}^{(j)\mathrm{T}}]^{\mathrm{T}}$, and the joint augmented state vector of all new PTs introduced at time t as $\overline{\boldsymbol{y}}_t \triangleq [\overline{\boldsymbol{y}}_t^{(1)\mathrm{T}}, \ldots, \overline{\boldsymbol{y}}_t^{(J)\mathrm{T}}]^{\mathrm{T}}$.

Legacy PTs and new PTs at time t at agent pair j, become legacy PTs at the next agent pair $j+1$, if $j < J$, or at the first agent pair at the next time step $t+1$, if $j = J$; in the latter case, this operation also implies performing the prediction from $t-1$ to t of the PT states. It then follows that the number of legacy PTs grows as $L_t^{(j)} = L_t^{(j-1)} + M_t^{(j-1)}$, where $L_t^{(1)} = K_{t-1}$, i.e., the number of legacy PTs at time t at the first agent pair $j = 1$ is equal to the number of PTs at time $t-1$. Analogously, we can reinterpret the vector $\underline{\boldsymbol{y}}_t^{(j)}$ of all the legacy PT augmented states at time t at agent pair j, as the vector stacking all the legacy PT augmented states at time t at the previous agent pair $j-1$, and the new PT augmented states introduced at time t at the previous agent pair $j-1$, that is, $\underline{\boldsymbol{y}}_t^{(j)} = [\underline{\boldsymbol{y}}_t^{(j-1)\mathrm{T}}, \overline{\boldsymbol{y}}_t^{(j-1)\mathrm{T}}]^{\mathrm{T}}$. This correspondence between legacy and new PTs at agent pair $j-1$, and legacy PTs at agent pair j, will hereafter be referred to as "PT mapping". Given the sequential construction of the joint legacy PT augmented state vector $\underline{\boldsymbol{y}}_t^{(j)}$ shown above, the vector $\boldsymbol{y}_t$ of all the PT augmented states at time t introduced in Section II-A can now be formally defined as $\boldsymbol{y}_t \triangleq [\underline{\boldsymbol{y}}_t^{(1)\mathrm{T}}, \overline{\boldsymbol{y}}_t^{(1)\mathrm{T}}, \overline{\boldsymbol{y}}_t^{(2)\mathrm{T}}, \ldots, \overline{\boldsymbol{y}}_t^{(J)\mathrm{T}}]^{\mathrm{T}} = [\underline{\boldsymbol{y}}_t^{(2)\mathrm{T}}, \overline{\boldsymbol{y}}_t^{(2)\mathrm{T}}, \ldots, \overline{\boldsymbol{y}}_t^{(J)\mathrm{T}}]^{\mathrm{T}} = \cdots = [\underline{\boldsymbol{y}}_t^{(J)\mathrm{T}}, \overline{\boldsymbol{y}}_t^{(J)\mathrm{T}}]^{\mathrm{T}}$. The number of PTs at time t, after all the MOT measurements are incorporated, is therefore $K_t = L_t^{(1)} + \sum_{j=1}^{J} M_t^{(j)} = L_t^{(J)} + M_t^{(J)}$. Note that the set of PTs at time $t = 0$ is assumed to be empty, i.e., $K_0 = 0$, and so is the set of legacy PTs at time $t = 1$ at the first agent pair $j = 1$, i.e., $L_1^{(1)} = K_0 = 0$.

We now provide a brief example that, with the support of Fig. 2, shows how new PTs are introduced and how the PT mapping works. Let us consider one Tx-agent a', i.e., $\mathcal{T} = \{a'\}$, and two Rx-agents, i.e., $\mathcal{R} = \{1, 2\}$; thus we have $\mathcal{J} = \{1, 2\}$, and we assume that the first agent pair $j = 1$ is $(1, a')$, and the second agent pair $j = 2$ is $(2, a')$. Note that this choice is arbitrary. Furthermore, we assume that at time $t-1$ the number of PTs is $K_{t-1} = 3$, i.e., $\boldsymbol{y}_{t-1} = [\boldsymbol{y}_{1,t-1}^{\mathrm{T}}, \boldsymbol{y}_{2,t-1}^{\mathrm{T}}, \boldsymbol{y}_{3,t-1}^{\mathrm{T}}]^{\mathrm{T}}$, represented as green circles in the first row of Fig. 2, and that at time t the number of MOT measurements at the agent pair $j = 1$ is $M_t^{(1)} = 4$, and the number of MOT measurements at agent pair $j = 2$ is $M_t^{(2)} = 2$. The PTs at

[3] Note that, in the remainder of the paper, we will be referring to a specific agent pair with "j" and "(j_1, j_2)" interchangeably, without using the indexing function $\phi(\cdot)$.

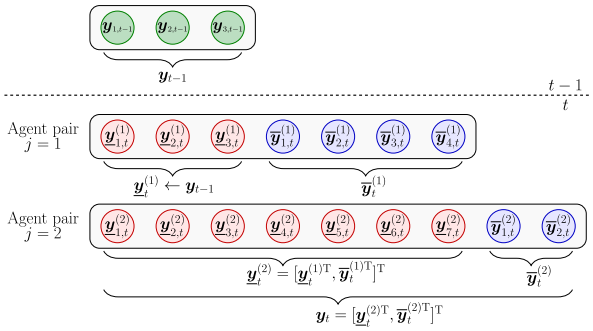


FIGURE 2. Illustration that describes the introduction of new PTs and the PT mapping between times $t-1$ and t, and between agent pairs $j-1$ and j. Green circles represent PTs at previous time $t-1$; red circles represent legacy PTs at time t at agent pair j; blue circles represent new PTs introduced at time t at agent pair j.

previous time $t-1$ become — once PTs' states prediction is performed — legacy PTs at time t at agent pair $j=1$, and are represented as red circles in the second row of Fig. 2; therefore, $L_t^{(1)} = K_{t-1} = 3$ and we formally have that the first PT at time $t-1$ becomes the first legacy PT at time t at agent pair $j=1$, i.e., $\underline{\boldsymbol{y}}_{1,t}^{(1)} \leftarrow \boldsymbol{y}_{1,t-1}$, and so forth for the other PTs, i.e., $\underline{\boldsymbol{y}}_{2,t}^{(1)} \leftarrow \boldsymbol{y}_{2,t-1}$ and $\underline{\boldsymbol{y}}_{3,t}^{(1)} \leftarrow \boldsymbol{y}_{3,t-1}$. Moreover, since the number of MOT measurements at agent pair $j=1$ is $M_t^{(1)} = 4$, four new PTs are introduced at this stage, i.e., $\overline{\boldsymbol{y}}_{1,t}^{(1)}$, $\overline{\boldsymbol{y}}_{2,t}^{(1)}$, $\overline{\boldsymbol{y}}_{3,t}^{(1)}$, and $\overline{\boldsymbol{y}}_{4,t}^{(1)}$, represented as blue circles in the second row of Fig. 2. Then, legacy PTs and new PTs at agent pair $j=1$ become legacy PTs at agent pair $j=2$, represented as red circles in the third row of Fig. 2. Therefore, $L_t^{(2)} = L_t^{(1)} + M_t^{(1)} = 3+4 = 7$, and we formally have that the first legacy PT at agent pair $j=1$ becomes the first legacy PT at agent pair $j=2$, i.e., $\underline{\boldsymbol{y}}_{1,t}^{(2)} \leftarrow \underline{\boldsymbol{y}}_{1,t}^{(1)}$, and so forth for the other legacy PTs at agent pair $j=1$, and that the first new PT at agent pair $j=1$ becomes the fourth legacy PT at agent pair $j=2$, i.e., $\underline{\boldsymbol{y}}_{4,t}^{(2)} \leftarrow \overline{\boldsymbol{y}}_{1,t}^{(1)}$, and so forth for the other new PTs at agent pair $j=1$. The vector of all the legacy PT augmented states at time t and agent pair $j=2$ is thus $\underline{\boldsymbol{y}}_t^{(2)} = [\underline{\boldsymbol{y}}_{1,t}^{(1)\mathrm{T}}, \ldots, \underline{\boldsymbol{y}}_{3,t}^{(1)\mathrm{T}}, \overline{\boldsymbol{y}}_{1,t}^{(1)\mathrm{T}}, \ldots, \overline{\boldsymbol{y}}_{4,t}^{(1)\mathrm{T}}]^{\mathrm{T}}$. This is an example of PT mapping from agent pair $j-1$ to agent pair j. Moreover, since the number of MOT measurements at agent pair $j=2$ is $M_t^{(2)} = 2$, two new PTs are introduced at this stage, i.e., $\overline{\boldsymbol{y}}_{1,t}^{(2)}$ and $\overline{\boldsymbol{y}}_{2,t}^{(2)}$, represented as blue circles in the third row of Fig. 2. The number of PTs at time t is thus $K_t = L_t^{(2)} + M_t^{(2)} = 7+2 = 9$; these nine PTs will then become legacy PTs at time $t+1$ at agent pair $j=1$.

Note that using this mechanism the number of PTs grows indefinitely over time. To keep a tractable number of PTs, a sub-optimal pruning step is performed once all the MOT measurements at time t are processed; details are provided in Section IV-B6. Finally, for the reader's convenience, the main sets introduced in this section are summarized in Table 1.

D. PRACTICAL USE CASES

This section links the concepts of agents, targets, and observations introduced so far to practical use cases. The intention is to ground the reader in the contextualization and identification of potential applications. In this regard, we provide three examples in popular areas of interest.

- *Maritime situational awareness:* in an acoustic underwater wireless sensor network, agents can be nearly stationary communication gateways (either anchored on the seabed or floating), and underwater or surface mobile vehicles (either manned or unmanned) with communication capabilities that form a multistatic active sonar system: one of these vehicles is equipped with a sonar source, while the others act as receivers and use arrays of hydrophones [14]. Agent state $\boldsymbol{s}_{a,t}$ includes position and possibly other kinematic parameters, such as velocity and acceleration; in case of 2D position $\check{\boldsymbol{s}}_{a,t}$, velocity $\dot{\check{\boldsymbol{s}}}_{a,t}$, and acceleration $\ddot{\check{\boldsymbol{s}}}_{a,t}$, the agent state becomes $\boldsymbol{s}_{a,t} = [\check{\boldsymbol{s}}_{a,t}^{\mathrm{T}}, \dot{\check{\boldsymbol{s}}}_{a,t}^{\mathrm{T}}, \ddot{\check{\boldsymbol{s}}}_{a,t}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$. For surface agents, navigation data $\boldsymbol{g}_{a,t}$ can be obtained through GNSS, in which case the noise-free navigation data can be $\boldsymbol{g}_{a,t} = \check{\boldsymbol{s}}_{a,t}$; whereas for underwater agents, the noise-free navigation data can be provided in terms of acceleration by the INS, in which case $\boldsymbol{g}_{a,t} = \ddot{\check{\boldsymbol{s}}}_{a,t}$. Mobile vehicles produce inter-agent measurements through different localization techniques [61], with which the range between two agents can be determined; furthermore, using the array of hydrophones the bearing can also be obtained. In such a case, the noise-free inter-agent measurement produced by agent pair (a, a') is $\boldsymbol{\rho}_t^{(a,a')} = [\|\check{\boldsymbol{s}}_{a,t} - \check{\boldsymbol{s}}_{a',t}\|, \angle(\check{\boldsymbol{s}}_{a',t} - \check{\boldsymbol{s}}_{a,t})]^{\mathrm{T}}$. Similarly, the MOT measurement produced by agent a acting as receiver due to the signal transmitted by the sonar source a' that reflects off a target (e.g., submarine) whose 2D position is $\check{\boldsymbol{x}}_{k,t}$, generally consists of range and bearing information, that is, $\boldsymbol{z}_{m,t}^{(a,a')} = [\|\check{\boldsymbol{s}}_{a,t} - \check{\boldsymbol{x}}_{k,t}\| + \|\check{\boldsymbol{s}}_{a',t} - \check{\boldsymbol{x}}_{k,t}\|, \angle(\check{\boldsymbol{x}}_{k,t} - \check{\boldsymbol{s}}_{a,t})]^{\mathrm{T}}$.
- *Cooperative intelligent transportation system:* in this scenario, agents are mobile land vehicles and road infrastructure nodes equipped with sensing and communication systems. Common devices used for sensing include radar, lidar, camera and ultrasound technologies, while the communication among the agents is achieved through ITS-G5/DSRC[4] or C-V2X[5] [62]. GNSS is the most widely adopted technology to obtain navigation

TABLE I Summary of the Sets Introduced in Section II

DESCRIPTION	SYMBOL
Agents	$\mathcal{A}$
Tx-agents	$\mathcal{T} \subseteq \mathcal{A}$
Rx-agents	$\mathcal{R} \subseteq \mathcal{A}$
Agent pairs	$\mathcal{J}$
PTs at time t	$\mathcal{K}_t$
Legacy PTs at time t at agent pair j	$\mathcal{L}_t^{(j)}$
MOT measurements/new PTs at time t at agent pair j	$\mathcal{M}_t^{(j)}$

[4]Dedicated Short-Range Communications
[5]Cellular-Vehicle-to-Everything

data, often combined with INS for the identification of abrupt kinematic events. Inter-agent measurements can rely on V2X communications to collect range and bearing (if antenna arrays are used) information, and multipath signals can be used to detect the presence of scattering targets (e.g., pedestrians, cyclists, other vehicles) in the environment.

- *Drone swarm:* agents are flying drones equipped with a GNSS receiver, INS, ultra wideband (UWB) technology, and a camera. Navigation data from GNSS and INS is paired with UWB ranging for a cooperative localization of the agents, while targets are detected by cameras with overlapping fields of view. Targets can vary from one application to another, as drone swarms are used in a variety of areas, e.g., smart cities, agriculture, environmental monitoring or mapping, military operations, search and rescue.

These use cases represent only a subset of applications in which the proposed cooperative self-localization and multi-target tracking algorithm can be employed.

III. STOCHASTIC PROBLEM FORMULATION

This section describes the detection and state estimation of a PT at time t, the state estimation of an agent, the MOT measurement model and data association problem (caused by the MOT measurement-origin uncertainty), and summarizes the assumptions used in the proposed formulation. Then, a factorization of the joint posterior pdf of the PT augmented states, agent states, and data association variables (introduced in the next section) is provided; this factorization is eventually used to compute the marginal posterior pdfs $f(\boldsymbol{s}_{a,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ and $f(\boldsymbol{x}_{k,t}, r_{k,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$.

A. AGENT SELF-LOCALIZATION AND TARGET TRACKING

The objective of this work is the cooperative self-localization of agents, jointly with the detection and localization of PTs. This task is performed with a Bayesian approach based on navigation data $\boldsymbol{g}_{1:t}$, inter-agent measurements $\boldsymbol{\rho}_{1:t}$, and MOT measurements $\boldsymbol{z}_{1:t}$, that boils down to the computation of the marginal posterior pdfs $f(\boldsymbol{s}_{a,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$, $\forall a \in \mathcal{A}$, and $f(\boldsymbol{x}_{k,t}, r_{k,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$, $\forall k \in \mathcal{K}_t$.

The detection and state estimation of a PT at time t is performed once all the MOT measurements are processed. The detection of a PT $k \in \mathcal{K}_t$ amounts to calculating the marginal posterior existence probability $f(r_{k,t}=1|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}) = \int f(\boldsymbol{x}_{k,t}, r_{k,t}=1|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})\mathrm{d}\boldsymbol{x}_{k,t}$, and comparing it with a suitably chosen threshold P_{ex}; that is, if $f(r_{k,t}=1|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}) > P_{\mathrm{ex}}$, the existence of PT k is confirmed. Then, for each detected PT k, an estimate of its state $\boldsymbol{x}_{k,t}$ is provided by the minimum mean square error (MMSE) estimator

$$\hat{\boldsymbol{x}}_{k,t}^{\mathrm{MMSE}} \triangleq \int \boldsymbol{x}_{k,t} f(\boldsymbol{x}_{k,t}|r_{k,t}=1, \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})\mathrm{d}\boldsymbol{x}_{k,t},$$

where $f(\boldsymbol{x}_{k,t}|r_{k,t}=1, \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}) = f(\boldsymbol{x}_{k,t}, r_{k,t}=1|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}) / f(r_{k,t}=1|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$. Likewise, an estimate of the agent state $\boldsymbol{s}_{a,t}$, $a \in \mathcal{A}$, is provided by the MMSE estimator

$$\hat{\boldsymbol{s}}_{a,t}^{\mathrm{MMSE}} \triangleq \int \boldsymbol{s}_{a,t} f(\boldsymbol{s}_{a,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})\mathrm{d}\boldsymbol{s}_{a,t}.$$

B. MOT MEASUREMENT MODEL AND DATA ASSOCIATION

As mentioned above (cf. Section II-B3), the MOT measurements $\boldsymbol{z}_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)}$, produced at time t at agent pair j, have unknown origins. Specifically, we make the assumption — known as *point-target assumption* — that each MOT measurement $\boldsymbol{z}_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)}$, at time t at agent pair j, originates either from a legacy PT or agent, hereafter aggregated under the term *legacy object*, or from a new PT (i.e., a PT never detected before), or from clutter (i.e., a false alarm), and it cannot originate from more than one source (legacy object, new PT, or clutter) simultaneously. Conversely, each legacy object or new PT can generate at most one MOT measurement at time t at agent pair j [54]. To handle this uncertainty, firstly we introduce the joint legacy object state vector $\boldsymbol{o}_t^{(j)} \triangleq [\boldsymbol{o}_{1,t}^{(j)\mathrm{T}}, \ldots, \boldsymbol{o}_{O_t^{(j)},t}^{(j)\mathrm{T}}]^{\mathrm{T}}$, with $O_t^{(j)} \triangleq L_t^{(j)} + A$, as the vector stacking at time t the legacy PT states at agent pair j, and the agent states. That is, $\boldsymbol{o}_{i,t}^{(j)} = \underline{\boldsymbol{x}}_{\ell,t}^{(j)}$ if $i = \ell$ and $\ell \in \mathcal{L}_t^{(j)}$, and $\boldsymbol{o}_{i,t}^{(j)} = \boldsymbol{s}_{a,t}$ if $i = L_t^{(j)} + a$ and $a \in \mathcal{A}$. We observe that the vector $\boldsymbol{o}_t^{(j)}$ includes the state vectors of the Rx-agent and the Tx-agent at $i = L_t^{(j)} + j_1$ and $i = L_t^{(j)} + j_2$, respectively, which cannot generate any MOT measurement at agent pair j. Therefore, we assume that a legacy object $i \in \mathcal{O}_t^{(j)} \triangleq \{1, \ldots, O_t^{(j)}\}$ is "detected" by the agent pair j — i.e., it generates a measurement $\boldsymbol{z}_{m,t}^{(j)}$ at the agent pair j — with probability $P_{\mathrm{d}}^{(j)}(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t})$, defined for $i \neq L_t^{(j)} + j_1$ and $i \neq L_t^{(j)} + j_2$ as

$$P_{\mathrm{d}}^{(j)}\left(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}\right) \triangleq \begin{cases} P_{\mathrm{d,mono}}^{(j)}\left(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}\right) & j_1 = j_2, \\ P_{\mathrm{d,bi}}^{(j)}\left(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}\right) & j_1 \neq j_2, \end{cases} \tag{6}$$

and for $i = L_t^{(j)} + j_1$ or $i = L_t^{(j)} + j_2$ as

$$P_{\mathrm{d}}^{(j)}\left(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}\right) = 0, \tag{7}$$

where $P_{\mathrm{d,mono}}^{(j)}(\cdot)$ is the monostatic detection probability of agent $j_1 = j_2$, and $P_{\mathrm{d,bi}}^{(j)}(\cdot)$ is the bistatic detection probability of the agent pair j, with $j_1 \neq j_2$. Secondly, following [49], we introduce: *(i)* the set $\mathcal{N}_t^{(j)}$ of MOT measurements generated by new PTs at time t at agent pair j, that is, $\mathcal{N}_t^{(j)} \triangleq \{m \in \mathcal{M}_t^{(j)} : \overline{r}_{m,t}^{(j)} = 1\}$; *(ii)* the legacy object-oriented association vector $\boldsymbol{\alpha}_t^{(j)} \triangleq [\alpha_{1,t}^{(j)}, \ldots, \alpha_{O_t^{(j)},t}^{(j)}]^{\mathrm{T}}$; and *(iii)* the MOT measurement-oriented association vector $\boldsymbol{\beta}_t^{(j)} \triangleq [\beta_{1,t}^{(j)}, \ldots, \beta_{M_t^{(j)},t}^{(j)}]^{\mathrm{T}}$. Here, $\alpha_{i,t}^{(j)}$, $i \in \mathcal{O}_t^{(j)}$, is defined as $m \in \mathcal{M}_t^{(j)}$ if legacy object i generates MOT measurement m, and 0 if legacy object i does not generate any MOT measurement; and $\beta_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)}$, is defined as i if MOT measurement m

originates from legacy object i, and 0 if MOT measurement m does not originate from any legacy object. The point-target assumption can therefore be expressed by the indicator function $\Phi(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)})$, defined as

$$\Phi\left(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}\right) \triangleq \Psi\left(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}\right) \prod_{m \in \mathcal{N}_t^{(j)}} \Gamma\left(\beta_{m,t}^{(j)}\right), \tag{8}$$

where

$$\Gamma\left(\beta_{m,t}^{(j)}\right) \triangleq \begin{cases} 0 & \beta_{m,t}^{(j)} \in \mathcal{O}_t^{(j)}, \\ 1 & \beta_{m,t}^{(j)} = 0, \end{cases} \tag{9}$$

and

$$\Psi\left(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}\right) \triangleq \prod_{i \in \mathcal{O}_t^{(j)}} \prod_{m \in \mathcal{M}_t^{(j)}} \psi\left(\alpha_{i,t}^{(j)}, \beta_{m,t}^{(j)}\right), \tag{10}$$

with

$$\psi(\alpha_{i,t}^{(j)}, \beta_{m,t}^{(j)}) \triangleq \begin{cases} 0 & \alpha_{i,t}^{(j)} = m \text{ and } \beta_{m,t}^{(j)} \neq i\,, \\ & \quad \text{or } \alpha_{i,t}^{(j)} \neq m \text{ and } \beta_{m,t}^{(j)} = i\,, \\ 1 & \text{otherwise}\,. \end{cases}$$

(We observe that, since the product in (8) is over the set $\mathcal{N}_t^{(j)}$, the indicator function $\Phi(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)})$ formally depends also on the existence variables $\overline{r}_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)}$). Stated differently, valid associations described by $\boldsymbol{\alpha}_t^{(j)}$ and $\boldsymbol{\beta}_t^{(j)}$ are those for which $\Phi(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}) = 1$; and we note that $\Psi(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)})$ is 0 if an MOT measurement is associated with two or more legacy objects (and, vice versa, if a legacy object is associated with two or more MOT measurements), and 1 otherwise; and that the product over $m \in \mathcal{N}_t^{(j)}$ of $\Gamma(\beta_{m,t}^{(j)})$ is 0 if any MOT measurement generated by a new PT is also associated with a legacy object, and 1 otherwise. For convenience, we also define the vectors $\boldsymbol{\alpha}_t \triangleq [\boldsymbol{\alpha}_t^{(1)\mathrm{T}}, \ldots, \boldsymbol{\alpha}_t^{(J)\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\beta}_t \triangleq [\boldsymbol{\beta}_t^{(1)\mathrm{T}}, \ldots, \boldsymbol{\beta}_t^{(J)\mathrm{T}}]^{\mathrm{T}}$, as well as $\boldsymbol{\alpha}_{1:t} \triangleq [\boldsymbol{\alpha}_1^{\mathrm{T}}, \ldots, \boldsymbol{\alpha}_t^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\beta}_{1:t} \triangleq [\boldsymbol{\beta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\beta}_t^{\mathrm{T}}]^{\mathrm{T}}$.

Finally, if the MOT measurement m is generated by legacy object i ($i \neq L_t^{(j)} + j_1$ and $i \neq L_t^{(j)} + j_2$), the statistical dependence of $\boldsymbol{z}_{m,t}^{(j)}$ on the legacy object state $\boldsymbol{o}_{i,t}^{(j)}$, the Rx-agent state $\boldsymbol{s}_{j_1,t}$, and the Tx-agent state $\boldsymbol{s}_{j_2,t}$, is given by

$$\mathrm{f}\left(\boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}\right) \triangleq \begin{cases} f_{\mathtt{mono}}\left(\boldsymbol{z}_{m,t}^{(j)} | \boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}\right) & j_1 = j_2, \\ f_{\mathtt{bi}}\left(\boldsymbol{z}_{m,t}^{(j)} | \boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}\right) & j_1 \neq j_2, \end{cases} \tag{11}$$

where the likelihoods $f_{\mathtt{mono}}(\boldsymbol{z}_{m,t}^{(j)} | \boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t})$ and $f_{\mathtt{bi}}(\boldsymbol{z}_{m,t}^{(j)} | \boldsymbol{o}_{i,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t})$ were introduced in Section II-B3. If the MOT measurement m is generated by a new PT, the statistical dependence of $\boldsymbol{z}_{m,t}^{(j)}$ on the new PT state $\overline{\boldsymbol{x}}_{m,t}^{(j)}$, Rx-agent state $\boldsymbol{s}_{j_1,t}$, and Tx-agent state $\boldsymbol{s}_{j_2,t}$, is still described by the likelihoods in (11) in which the legacy object state $\boldsymbol{o}_{i,t}^{(j)}$ is replaced by the new PT state $\overline{\boldsymbol{x}}_{m,t}^{(j)}$.

C. ASSUMPTIONS

The assumptions underlying the proposed stochastic formulation — besides the point-target assumption — are here summarized: some are basic assumptions commonly used in multisensor multitarget tracking [54], while others belong to the specific formulation borrowed from [49], [53].

A1) The joint agent state vector $\boldsymbol{s}_t$ evolves over time according to a first-order Markov model, and each agent state vector $\boldsymbol{s}_{a,t}$, evolves independently [54]; therefore, the joint agent state transition pdf $f(\boldsymbol{s}_t|\boldsymbol{s}_{t-1})$ factorizes as

$$f(\boldsymbol{s}_t|\boldsymbol{s}_{t-1}) = \prod_{a \in \mathcal{A}} \tau_a(\boldsymbol{s}_{a,t}|\boldsymbol{s}_{a,t-1})\,, \tag{12}$$

where $\tau_a(\boldsymbol{s}_{a,t}|\boldsymbol{s}_{a,t-1})$ is a known state-transition pdf (cf. Section II-A).

A2) The joint PT augmented state vector $\boldsymbol{y}_t$ evolves over time according to a first-order Markov model, and each PT augmented state vector $\boldsymbol{y}_{k,t}$ evolves independently [54]. Recalling that for each PT augmented state $\boldsymbol{y}_{k,t-1}$, $k \in \mathcal{K}_{t-1}$, at time $t-1$, there is one legacy PT augmented state $\underline{\boldsymbol{y}}_{\ell,t}^{(1)}$, $\ell \in \mathcal{L}_t^{(1)}$, at the first agent pair at time t (in other words, $\mathcal{L}_t^{(1)} = \mathcal{K}_{t-1}$), the joint PT augmented state transition pdf is given by

$$f\left(\underline{\boldsymbol{y}}_t^{(1)}|\boldsymbol{y}_{t-1}\right) = \prod_{k \in \mathcal{K}_{t-1}} f\left(\underline{\boldsymbol{x}}_{k,t}^{(1)}, \underline{r}_{k,t}^{(1)}|\boldsymbol{x}_{k,t-1}, r_{k,t-1}\right). \tag{13}$$

Let us recall from Section II-C that the set of PTs at time $t = 0$ is empty, i.e., $\mathcal{K}_0 = \varnothing$, and so is the set of legacy PTs at time t=1 at the first agent pair $j = 1$, i.e., $\mathcal{L}_1^{(1)} = \varnothing$. Therefore, the state transition pdf in (13) cannot be performed at time $t = 1$, and we formally introduce $f(\underline{\boldsymbol{y}}_1^{(1)}|\boldsymbol{y}_0) = 1$ for future use. An expression of $f(\underline{\boldsymbol{x}}_{k,t}^{(1)}, \underline{r}_{k,t}^{(1)}|\boldsymbol{x}_{k,t-1}, r_{k,t-1})$ is provided in [49, Sec. VIII-C], and is here reported for completeness. If PT k does not exist at time $t-1$, i.e., if $r_{k,t-1} = 0$, it cannot exist as legacy PT at time t, i.e, $\underline{r}_{k,t}^{(1)} = 0$, and thus its state pdf is $f_{\mathrm{D}}(\underline{\boldsymbol{x}}_{k,t}^{(1)})$. That is,

$$f\left(\underline{\boldsymbol{x}}_{k,t}^{(1)}, \underline{r}_{k,t}^{(1)}|\boldsymbol{x}_{k,t-1}, r_{k,t-1} = 0\right) = \begin{cases} f_{\mathrm{D}}\left(\underline{\boldsymbol{x}}_{k,t}^{(1)}\right) & \underline{r}_{k,t}^{(1)} = 0\,, \\ 0 & \underline{r}_{k,t}^{(1)} = 1\,. \end{cases}$$

Conversely, if PT k exists at time $t-1$, i.e., if $r_{k,t-1} = 1$, it survives as legacy PT with probability $p_{\mathrm{s}}(\boldsymbol{x}_{k,t-1})$, and its state $\underline{\boldsymbol{x}}_{k,t}^{(1)}$ is distributed according to the state transition pdf $f(\underline{\boldsymbol{x}}_{k,t}^{(1)}|\boldsymbol{x}_{k,t-1})$. Thus,

$$f\left(\underline{\boldsymbol{x}}_{k,t}^{(1)}, \underline{r}_{k,t}^{(1)}|\boldsymbol{x}_{k,t-1}, r_{k,t-1} = 1\right)$$

$$= \begin{cases} \left(1 - p_{\mathrm{s}}(\boldsymbol{x}_{k,t-1})\right) f_{\mathrm{D}}\left(\underline{\boldsymbol{x}}_{k,t}^{(1)}\right) & \underline{r}_{k,t}^{(1)} = 0\,, \\ p_{\mathrm{s}}(\boldsymbol{x}_{k,t-1}) f\left(\underline{\boldsymbol{x}}_{k,t}^{(1)} | \boldsymbol{x}_{k,t-1}\right) & \underline{r}_{k,t}^{(1)} = 1\,. \end{cases}$$

The state transition pdf is defined by the PT kinematic model, that is,

$$\underline{\boldsymbol{x}}_{k,t}^{(1)} = \boldsymbol{\varsigma}\left(\boldsymbol{x}_{k,t-1}, \boldsymbol{e}_{k,t}\right), \tag{14}$$

and by the statistics of the process noise $\boldsymbol{e}_{k,t}$.

A3) The joint agent state vector $\boldsymbol{s}_t$ and the joint PT augmented state vector $\boldsymbol{y}_t$ evolve independently over time [53], [54].

A4) The states of legacy PTs and new PTs at time t at agent pair j are independent [49].

A5) The number of new PTs at time t at agent pair j is a priori (i.e., before the MOT measurements are observed) Poisson distributed with mean $\mu_{\mathrm{n}}^{(j)}$. The states of new PTs are independent and identically distributed according to the prior pdf $f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{m,t}^{(j)})$ [49].

A6) Given the agent states and PT augmented states at time $t-1$, the observations (navigation data, inter-agent and MOT measurements), association variables, agent states, and PT augmented states at time t, are conditionally independent of all the past ($t' < t$) variables [49], [53].

A7) Given the agent states and legacy PT augmented states at time t, the new PT augmented states, observations, and association variables at time t, are conditionally independent of all the past ($t' < t$) agent states and PT augmented states [49], [53].

A8) Similarly, given the agent states at time t, and the legacy PT augmented states at time t at agent pair j, the new PT augmented states, observations, and association variables at time t at current and future agent pairs, are conditionally independent of all the past ($j' < j$) variables [49], [53].

A9) The navigation data and the inter-agent measurements are conditionally independent of each other, and of all the other variables, given the joint agent state vector $\boldsymbol{s}_t$ [54]. Then, recalling their measurement models (cf. (2) and (3)), and in particular that the noise terms are independent across agent states and agent pairs, respectively, it follows that the joint navigation data likelihood $f(\boldsymbol{g}_t|\boldsymbol{s}_t)$ factorizes as

$$f(\boldsymbol{g}_t|\boldsymbol{s}_t) = \prod_{a \in \mathcal{A}_t^{\boldsymbol{g}}} \mathfrak{g}_a(\boldsymbol{g}_{a,t}|\boldsymbol{s}_{a,t}), \tag{15}$$

and that the joint likelihood $f(\boldsymbol{\rho}_t|\boldsymbol{s}_t)$ factorizes as

$$f(\boldsymbol{\rho}_t|\boldsymbol{s}_t) = \prod_{a \in \mathcal{R}} \prod_{a' \in \mathcal{T}_t^{(a)}} \mathfrak{d}\left(\boldsymbol{\rho}_t^{(a,a')}|\boldsymbol{s}_{a,t}, \boldsymbol{s}_{a',t}\right). \tag{16}$$

A10) The number of false alarm MOT measurements at time t at agent pair j is Poisson distributed with mean $\mu_{\mathrm{c}}^{(j)}$. False alarm MOT measurements are independent and identically distributed according to the pdf $f_{\mathrm{c}}^{(j)}(\boldsymbol{z}_{m,t}^{(j)})$ [49], [54].

A11) The agent-originated and PT-originated MOT measurements at time t at agent pair j, are conditionally independent of each other, and conditionally independent of the false alarm MOT measurements, given the agent states $\boldsymbol{s}_{j_1,t}$ and $\boldsymbol{s}_{j_2,t}$, PT augmented states, and association variables [49], [53].

D. JOINT POSTERIOR PDF

The posterior pdfs $f(\boldsymbol{y}_{k,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$, $k \in \mathcal{K}_t$, and $f(\boldsymbol{s}_{a,t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$, $a \in \mathcal{A}$, introduced in Section III-A, are marginal densities of the joint posterior pdf $f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$. By using assumptions (A3), (A6), (A7), (A8), and (A9), this joint posterior pdf can be factorized as (details are provided in the Appendix)

$$\begin{aligned} f\left(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t}|\boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}\right) &\propto f\left(\boldsymbol{s}_0\right) \\ &\times \prod_{t'=1}^{t} f\left(\boldsymbol{s}_{t'}|\boldsymbol{s}_{t'-1}\right) f\left(\underline{\boldsymbol{y}}_{t'}^{(1)}|\boldsymbol{y}_{t'-1}\right) f\left(\boldsymbol{g}_{t'}|\boldsymbol{s}_{t'}\right) f\left(\boldsymbol{\rho}_{t'}|\boldsymbol{s}_{t'}\right) \\ &\times \prod_{j=1}^{J} f\left(\overline{\boldsymbol{y}}_{t'}^{(j)}, \boldsymbol{\alpha}_{t'}^{(j)}, \boldsymbol{\beta}_{t'}^{(j)}, \boldsymbol{z}_{t'}^{(j)}, M_{t'}^{(j)}|\underline{\boldsymbol{y}}_{t'}^{(j)}, \boldsymbol{s}_{t'}\right), \end{aligned} \tag{17}$$

for some prior pdf $f(\boldsymbol{s}_0)$, where we recall that $f(\underline{\boldsymbol{y}}_1^{(1)}|\boldsymbol{y}_0) = 1$ (see assumption (A2)). Then, observing that the description of the data association given by $\boldsymbol{\alpha}_t^{(j)}$ and $\boldsymbol{\beta}_t^{(j)}$ is redundant once $M_t^{(j)}$ is observed — indeed, $\boldsymbol{\alpha}_t^{(j)}$ can be derived from $\boldsymbol{\beta}_t^{(j)}$, and vice versa, when $M_t^{(j)}$ is known [49] —, each factor $f(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, \boldsymbol{z}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t)$ can be further expressed as

$$\begin{aligned} &f\left(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, \boldsymbol{z}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) \\ &\quad = f\left(\boldsymbol{z}_t^{(j)}|\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}, \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) \\ &\qquad \times f\left(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) \\ &\quad = f\left(\boldsymbol{z}_t^{(j)}|\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, M_t^{(j)}, \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) \\ &\qquad \times f\left(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right). \end{aligned} \tag{18}$$

Hereafter, following the derivations in [49], we provide expressions for the prior data association pdf $f(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t)$, and the joint MOT measurements likelihood $f(\boldsymbol{z}_t^{(j)}|\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, M_t^{(j)}, \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t)$.

1) PRIOR DATA ASSOCIATION PDF

By using assumptions (A4), (A5), (A10), and the point-target assumption, the pdf $f(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t)$ can be expressed as

$$f\left(\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, M_t^{(j)}|\underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) = C\left(M_t^{(j)}\right) \Psi\left(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}\right)$$

$$\times \prod_{\ell \in \mathcal{L}_t^{(j)}} q_1\left(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}\right)$$
$$\times \prod_{a \in \mathcal{A}} h_1\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}\right)$$
$$\times \prod_{m \in \mathcal{M}_t^{(j)}} \upsilon_1\left(\overline{\boldsymbol{y}}_{m,t}^{(j)}, \beta_{m,t}^{(j)}\right). \tag{19}$$

Here, $C(M_t^{(j)})$ is a normalization factor that depends only on the number of MOT measurements $M_t^{(j)}$ (see [49], [63] for details), $\Psi(\boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)})$ is defined in (10), and the functions $q_1(\cdot)$, $h_1(\cdot)$, and $\upsilon_1(\cdot)$ represent the contributions to the prior data association pdf of the legacy PTs, the agents, and the new PTs, respectively. The derivation of the pdf in (19) closely follows the derivation of the pdf in [49, Eq. (60)] and is thus omitted. The main difference is given by the product over the agents $a \in \mathcal{A}$ of the function $h_1(\cdot)$, that is a direct consequence of the involvement of the agents in the data association procedure. Next we provide detailed definitions of the functions $q_1(\cdot)$, $h_1(\cdot)$, and $\upsilon_1(\cdot)$. The function $q_1(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}) = q_1(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)})$ is defined for $\underline{r}_{\ell,t}^{(j)} = 1$ as

$$q_1\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)} = 1, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}\right) \triangleq \begin{cases} \dfrac{P_{\mathrm{d}}^{(j)}(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t})}{\mu_{\mathrm{c}}^{(j)}} & \alpha_{\ell,t}^{(j)} \in \mathcal{M}_t^{(j)}, \\ 1 - P_{\mathrm{d}}^{(j)}(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}) & \alpha_{\ell,t}^{(j)} = 0, \end{cases} \tag{20}$$

and for $\underline{r}_{\ell,t}^{(j)} = 0$ as

$$q_1\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)} = 0, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}\right) \triangleq \delta_{\alpha_{\ell,t}^{(j)},0}. \tag{21}$$

The function $h_1(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)})$ is similarly defined as

$$h_1\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}\right) \triangleq \begin{cases} \dfrac{P_{\mathrm{d}}^{(j)}(\boldsymbol{s}_{a,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t})}{\mu_{\mathrm{c}}^{(j)}} & \alpha_{L+a,t}^{(j)} \in \mathcal{M}_t^{(j)}, \\ 1 - P_{\mathrm{d}}^{(j)}(\boldsymbol{s}_{a,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}) & \alpha_{L+a,t}^{(j)} = 0, \end{cases} \tag{22}$$

where, with an abuse of notation, the number of legacy PTs at time t at agent pair j, i.e., $L_t^{(j)}$, is simply referred to as L. We observe that if agent $a \in \mathcal{A}$ is either the Rx-agent j_1, or the Tx-agent j_2, that is, $a = j_1$ or $a = j_2$, according to (7) it follows that $h_1(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; M_t^{(j)}) = \delta_{\alpha_{L+a,t}^{(j)},0}$, which intuitively means that no MOT measurements can be associated to the Rx-agent and the Tx-agent. Finally, $\upsilon_1(\overline{\boldsymbol{y}}_{m,t}^{(j)}, \beta_{m,t}^{(j)}) = \upsilon_1(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)}, \beta_{m,t}^{(j)})$ is defined for $\overline{r}_{m,t}^{(j)} = 1$ as

$$\upsilon_1\left(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)} = 1, \beta_{m,t}^{(j)}\right) \triangleq \Gamma\left(\beta_{m,t}^{(j)}\right) \frac{\mu_{\mathrm{n}}^{(j)}}{\mu_{\mathrm{c}}^{(j)}} f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{m,t}^{(j)}) = \begin{cases} 0 & \beta_{m,t}^{(j)} \in \mathcal{O}_t^{(j)}, \\ \dfrac{\mu_{\mathrm{n}}^{(j)}}{\mu_{\mathrm{c}}^{(j)}} f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{m,t}^{(j)}) & \beta_{m,t}^{(j)} = 0, \end{cases}$$

and for $\overline{r}_{m,t}^{(j)} = 0$ as

$$\upsilon_1^{(j)}\left(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)} = 0, \beta_{m,t}^{(j)}\right) \triangleq f_{\mathrm{D}}(\overline{\boldsymbol{x}}_{m,t}^{(j)}).$$

Note that the function $\upsilon_1(\cdot)$ incorporates the indicator function $\Gamma(\cdot)$ defined in (9); and that the combined use in (19) of the functions $\Psi(\cdot)$ and $\upsilon_1(\cdot)$ describes the point-target assumption as done by the indicator function $\Phi(\cdot)$ defined in (8).

2) JOINT MOT MEASUREMENTS LIKELIHOOD

By using assumptions (A10), (A11), and the point-target assumption, the joint MOT measurements likelihood $f(\boldsymbol{z}_t^{(j)}|\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, M_t^{(j)}, \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t)$ can be expressed as

$$f\left(\boldsymbol{z}_t^{(j)}|\overline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, M_t^{(j)}, \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right) = C\left(\boldsymbol{z}_t^{(j)}\right)$$
$$\times \prod_{\ell \in \mathcal{L}_t^{(j)}} q_2\left(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right)$$
$$\times \prod_{a \in \mathcal{A}} h_2\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right)$$
$$\times \prod_{m \in \mathcal{M}_t^{(j)}} \upsilon_2\left(\overline{\boldsymbol{y}}_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}\right). \tag{23}$$

Here, $C(\boldsymbol{z}_t^{(j)})$ is a normalization factor that depends only on the MOT measurements vector $\boldsymbol{z}_t^{(j)}$ (see [49], [63] for details), and the functions $q_2(\cdot)$, $h_2(\cdot)$, and $\upsilon_2(\cdot)$ embed the MOT measurement likelihoods related to the legacy PTs, the agents, and the new PTs, respectively. The derivation of the likelihood in (23) closely follows the derivation of the likelihood in [49, Eq. (64)] and is thus omitted. The main difference is given by the product over the agents $a \in \mathcal{A}$ of the function $h_2(\cdot)$, that represents the likelihoods of the MOT measurements when these are generated by the agents. Next we provide detailed definitions of the functions $q_2(\cdot)$, $h_2(\cdot)$, and $\upsilon_2(\cdot)$. The function $q_2(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}) = q_2(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)})$ is defined for $\underline{r}_{\ell,t}^{(j)} = 1$ as

$$q_2\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)} = 1, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \triangleq \begin{cases} \dfrac{\mathfrak{f}(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)})}{f_{\mathrm{c}}^{(j)}\left(\boldsymbol{z}_{m,t}^{(j)}\right)} & \alpha_{\ell,t}^{(j)} \in \mathcal{M}_t^{(j)}, \\ 1 & \alpha_{\ell,t}^{(j)} = 0, \end{cases} \tag{24}$$

and for $\underline{r}_{\ell,t}^{(j)} = 0$ as

$$q_2\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)} = 0, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \triangleq 1\,. \tag{25}$$

The function $h_2(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)})$ is similarly defined for $a \neq j_1$ and $a \neq j_2$ as

$$h_2\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \triangleq \begin{cases} \dfrac{\mathfrak{f}(\boldsymbol{s}_{a,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)})}{f_{\mathrm{c}}^{(j)}\left(\boldsymbol{z}_{m,t}^{(j)}\right)} & \alpha_{L+a,t}^{(j)} \in \mathcal{M}_t^{(j)}, \\ 1 & \alpha_{L+a,t}^{(j)} = 0, \end{cases} \tag{26}$$

and for $a = j_1$ or $a = j_2$ as

$$h_2\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \triangleq 1. \tag{27}$$

Finally, $\upsilon_2(\overline{\boldsymbol{y}}_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}) = \upsilon_2(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)})$ is defined as

$$\upsilon_2\left(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}\right) \triangleq \begin{cases} \dfrac{\mathfrak{f}\left(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}\right)}{f_{\mathrm{c}}^{(j)}\left(\boldsymbol{z}_{m,t}^{(j)}\right)} & \overline{r}_{m,t}^{(j)} = 1\,, \\ 1 & \overline{r}_{m,t}^{(j)} = 0\,. \end{cases} \tag{28}$$

The final factorization of the joint posterior pdf $f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ is obtained by inserting (19) and (23) into (18), and (12)–(13), (15)–(16), (18), into (17). Its expression is reported in (29), shown at the bottom of this page, where the equality $\mathcal{L}_t^{(1)} = \mathcal{K}_{t-1}$ has been used (see assumption (A2)), and where $q(\cdot) \triangleq q_1(\cdot)q_2(\cdot)$, $h(\cdot) \triangleq h_1(\cdot)h_2(\cdot)$, and $\upsilon(\cdot) \triangleq \upsilon_1(\cdot)\upsilon_2(\cdot)$. For the reader's convenience, we summarize in Table 2 all the variables involved in this factorization.

TABLE II Summary of the Variables Used in (29)

Description	Symbol
Agent state	$\boldsymbol{s}_{a,t}$
Legacy PT augmented state	$\underline{\boldsymbol{y}}_{\ell,t}^{(j)}$
New PT augmented state	$\overline{\boldsymbol{y}}_{m,t}^{(j)}$
Navigation data	$\boldsymbol{g}_{a,t}$
Inter-agent measurement	$\boldsymbol{\rho}_t^{(a,a')}$
MOT measurement	$\boldsymbol{z}_{m,t}^{(j)}$
Legacy object-oriented DA variable	$\alpha_{\ell,t}^{(j)}$
MOT measurement-oriented DA variable	$\beta_{m,t}^{(j)}$

IV. THE PROPOSED ALGORITHM

Direct marginalization of the joint posterior pdf $f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ for the computation of the marginal posterior pdfs $f(\boldsymbol{y}_{k,t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ and $f(\boldsymbol{s}_{a,t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ is generally infeasible, as it requires high-dimensional integration and summation. Approximations of these marginal posterior pdfs, called *beliefs*, can be efficiently obtained by applying the SPA on a factor graph [26], [27] carefully devised from the factorization in (29).

A. SUM-PRODUCT ALGORITHM: OVERVIEW AND NOTATION

Here, we briefly review factor graphs and the SPA. Let us consider the generic problem of estimating D parameter vectors $\boldsymbol{\lambda}_d$, $d \in \{1, \ldots, D\}$, from a vector of observations $\boldsymbol{\pi}$. In the Bayesian setting, these vectors are random, and the estimation of $\boldsymbol{\lambda}_d$ is based on the posterior pdf $f(\boldsymbol{\lambda}_d | \boldsymbol{\pi})$. This pdf is a marginal pdf of the joint posterior pdf $f(\boldsymbol{\lambda} | \boldsymbol{\pi})$, where $\boldsymbol{\lambda} \triangleq [\boldsymbol{\lambda}_1^{\mathrm{T}}, \ldots, \boldsymbol{\lambda}_D^{\mathrm{T}}]^{\mathrm{T}}$. The joint posterior pdf is assumed to be the product of certain lower-dimensional factors, i.e.,

$$f(\boldsymbol{\lambda} | \boldsymbol{\pi}) \propto \prod_{\ell} \kappa_\ell\left(\boldsymbol{\lambda}^{(\ell)}; \boldsymbol{\pi}\right), \tag{30}$$

$$\begin{aligned} & f\left(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} \middle| \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}\right) \\ & \propto f(\boldsymbol{s}_0) \prod_{t'=1}^{t} \underbrace{\left(\prod_{a \in \mathcal{A}} \tau_a\left(\boldsymbol{s}_{a,t'} \middle| \boldsymbol{s}_{a,t'-1}\right)\right)}_{\text{Agents' States Prediction}} \underbrace{\left(\prod_{a \in \mathcal{A}_{t'}^{\boldsymbol{g}}} \mathfrak{g}_a\left(\boldsymbol{g}_{a,t'} \middle| \boldsymbol{s}_{a,t'}\right)\right)\left(\prod_{a \in \mathcal{R}} \prod_{a' \in \mathcal{T}_{t'}^{(a)}} \mathfrak{d}\left(\boldsymbol{\rho}_{t'}^{(a,a')} \middle| \boldsymbol{s}_{a,t'}, \boldsymbol{s}_{a',t'}\right)\right)}_{\text{Agents' Cooperative Self-Localization}} \\ & \times \underbrace{\left(\prod_{\ell \in \mathcal{L}_{t'}^{(1)}} f\left(\underline{\boldsymbol{y}}_{\ell,t'}^{(1)} \middle| \underline{\boldsymbol{y}}_{\ell,t'-1}\right)\right)}_{\text{PTs' Augmented States Prediction}} \prod_{j=1}^{J} \underbrace{\left(\prod_{\ell \in \mathcal{L}_{t'}^{(j)}} q\left(\underline{\boldsymbol{y}}_{\ell,t'}^{(j)}, \alpha_{\ell,t'}^{(j)}, \boldsymbol{s}_{j_1,t'}, \boldsymbol{s}_{j_2,t'}; \boldsymbol{z}_{t'}^{(j)}\right) \prod_{m \in \mathcal{M}_{t'}^{(j)}} \psi\left(\alpha_{\ell,t'}^{(j)}, \beta_{m,t'}^{(j)}\right)\right)}_{\text{MOT Measurements Evaluation and Data Association}} \\ & \times \underbrace{\left(\prod_{a \in \mathcal{A}} h\left(\boldsymbol{s}_{a,t'}, \alpha_{L+a,t'}^{(j)}, \boldsymbol{s}_{j_1,t'}, \boldsymbol{s}_{j_2,t'}; \boldsymbol{z}_{t'}^{(j)}\right) \prod_{m \in \mathcal{M}_{t'}^{(j)}} \psi\left(\alpha_{L+a,t'}^{(j)}, \beta_{m,t'}^{(j)}\right)\right) \prod_{m \in \mathcal{M}_{t'}^{(j)}} \upsilon\left(\overline{\boldsymbol{y}}_{m,t'}^{(j)}, \beta_{m,t'}^{(j)}, \boldsymbol{s}_{j_1,t'}, \boldsymbol{s}_{j_2,t'}; \boldsymbol{z}_{m,t'}^{(j)}\right)}_{\text{MOT Measurements Evaluation and Data Association (continued)}} \end{aligned} \tag{29}$$

where each argument $\boldsymbol{\lambda}^{(\ell)}$ comprises certain parameter vectors $\boldsymbol{\lambda}_d$, and each $\boldsymbol{\lambda}_d$ can appear in several $\boldsymbol{\lambda}^{(\ell)}$. The factorization (30) can be represented by a factor graph, which is constructed as follows: each parameter variable $\boldsymbol{\lambda}_d$ is represented by a variable node; each factor $\kappa_\ell(\cdot)$ is represented by a factor node; and variable node "$\boldsymbol{\lambda}_d$" and factor node "κ_ℓ" are adjacent, i.e., connected by an edge, if $\boldsymbol{\lambda}_d$ is an argument of $\kappa_\ell(\cdot)$.

The SPA algorithm aims at computing the marginal posterior pdfs $f(\boldsymbol{\lambda}_d|\boldsymbol{\pi})$ in an efficient way, and is based on the factor graph representing the factorization of $f(\boldsymbol{\lambda}|\boldsymbol{\pi})$ in (30). For each node in the factor graph, certain messages are calculated, each of which is then passed to one of the adjacent nodes. Let $\mathcal{V}_\ell$ denote the set of indices d of all those variable nodes "$\boldsymbol{\lambda}_d$" that are adjacent to factor node "κ_ℓ". Then, factor node "κ_ℓ" passes the following message to variable node "$\boldsymbol{\lambda}_d$" with $d \in \mathcal{V}_\ell$:

$$\zeta_{\kappa_\ell \to \boldsymbol{\lambda}_d}(\boldsymbol{\lambda}_d) = \int \kappa_\ell\left(\boldsymbol{\lambda}^{(\ell)}; \boldsymbol{\pi}\right) \prod_{\substack{d' \in \mathcal{V}_\ell \\ d' \neq d}} \eta_{\boldsymbol{\lambda}_{d'} \to \kappa_\ell}(\boldsymbol{\lambda}_{d'})\, \mathrm{d}\boldsymbol{\lambda}_{-d}\,. \tag{31}$$

Here, $\int \ldots \mathrm{d}\boldsymbol{\lambda}_{-d}$ denotes integration with respect to all $\boldsymbol{\lambda}_{d'}$, $d' \in \mathcal{V}_\ell$, except $\boldsymbol{\lambda}_d$, and the messages $\eta_{\boldsymbol{\lambda}_d \to \kappa_\ell}(\boldsymbol{\lambda}_d)$ are calculated as described later. If the factorization (30) involves (also) discrete variables, then the respective integrations in (31) have to be replaced with summations. Furthermore, let $\mathcal{F}_d$ be the set of the indices ℓ of all those factors nodes "κ_ℓ" that are adjacent to variable node "$\boldsymbol{\lambda}_d$". Then, variable node "$\boldsymbol{\lambda}_d$" passes the following message to factor node "κ_ℓ" with $\ell \in \mathcal{F}_d$:

$$\eta_{\boldsymbol{\lambda}_d \to \kappa_\ell}(\boldsymbol{\lambda}_d) = \prod_{\substack{\ell' \in \mathcal{F}_d \\ \ell' \neq \ell}} \zeta_{\kappa_{\ell'} \to \boldsymbol{\lambda}_d}(\boldsymbol{\lambda}_d).$$

For a factor graph with loops, the calculation of the messages is usually repeated in an iterative manner. There is no unique order — or *schedule* — of message calculation, and different orders may lead to different results. Finally, for each variable node "$\boldsymbol{\lambda}_d$," a belief $\tilde{f}(\boldsymbol{\lambda}_d)$ is calculated by multiplying all the incoming messages (passed from all the adjacent factor nodes) and normalizing the resulting product function such that $\int \tilde{f}(\boldsymbol{\lambda}_d)\mathrm{d}\boldsymbol{\lambda}_d = 1$. The belief $\tilde{f}(\boldsymbol{\lambda}_d)$ provides the desired approximation of the marginal posterior pdf $f(\boldsymbol{\lambda}_d|\boldsymbol{\pi})$.

B. SPA-BASED JOINT LOCALIZATION AND TRACKING

The factor graph derived from the factorization (29) contains loops; therefore, a message calculation schedule needs to be selected. The proposed algorithm is based of the following rules: *(i)* messages are not sent backward in time; *(ii)* MOT measurements produced by the agent pairs are processed sequentially according to the arbitrary order established by the indexing function $\phi(\cdot)$, from agent pair $j = 1$ through agent pair $j = J$; and *(iii)* iterative message passing is only performed for agents' cooperative self-localization, and MOT measurements data association within each agent pair j. Note that, similarly to the order of the SPA messages, different orders of the agent pairs may lead to different outcomes of the SPA-based algorithm. Within the proposed framework, the order selection is further complicated by the bistatic geometry of some agent pairs; indeed, a bistatic agent pair might have a favorable geometry to observe a particular target, but not necessarily *all* the targets. The selection of the optimal agent pair order is still an open problem and needs to be tailored according to the specific application and, in particular, to the system's architecture and specifications. These aspects are not addressed in this paper. Finally, we observe that a parallel implementation, suitable for a distributed approach, is however possible as similarly done in [49, Sec. IX-B].

Fig. 3 shows a block diagram that provides an intuitive representation of the proposed method. First, the predictions of the joint agent state vector $\boldsymbol{s}_{t-1}$ and the joint PT augmented state vector $\boldsymbol{y}_{t-1}$ from $t-1$ to t are performed; these operations are described by the red boxes "agents' states prediction" and "PTs' augmented states prediction," respectively. The predicted agent states are then updated by using the navigation data $\boldsymbol{g}_t$ and the inter-agent measurements $\boldsymbol{\rho}_t$ through the "agents' cooperative self-localization". Next, the MOT measurements $\boldsymbol{z}_t^{(j)}$ produced at agent pair j are processed through the "MOT measurements evaluation and DA" box using the updated agent states $\boldsymbol{s}_t$ and the legacy PT augmented states $\underline{\boldsymbol{y}}_t^{(j)}$; this is performed sequentially, from agent pair $j = 1$ to agent pair $j = J$. Once the MOT measurements at agent pair j are processed, the beliefs of the PT states (both legacy and new) and the agent states are updated before the MOT measurements of the next agent pair $j+1$ are processed. All these operations, represented by the red boxes in Fig. 3, correspond to macro-factors in the factorization of the joint posterior pdf in (29). The green boxes refer to the PT mapping operation (cf. Section II-C) that is carried out between any two consecutive agent pairs, i.e., $j-1$ and j, at current time t.

Combining the rules for the message schedule stated above, and the generic SPA rules for calculating messages and beliefs described in Section IV-A, we provide the expressions of the SPA messages for each of these operations in what follows; for clarity, the titles of the next four subsections recall the operations described by the red boxes in Fig. 3. The SPA messages are exchanged on the factor graphs in Fig. 4, Fig. 5, and Fig. 6; the first one relates to the agents' states prediction and cooperative self-localization, the second one to the PTs' augmented states prediction, and the last one to the MOT measurements evaluation and data association. We observe that the agent state variable nodes at time t in Fig. 4, i.e., variable nodes "$\boldsymbol{s}_a$," coincide with the agent state variable nodes in Fig. 6; analogously, the legacy PT augmented state variable nodes at time t in Fig. 5, i.e., variable nodes "$\underline{\boldsymbol{y}}_\ell$," coincide with the legacy PT augmented state variable nodes in Fig. 6 when $j = 1$. The structure of the factor graph in Fig. 4 changes according to the availability, at time t, of the navigation data and the inter-agent measurements; a general case is there illustrated, in which all the agents have navigation data, and inter-agent measurements between any two agents are available. Similarly, the structure of the factor graph in

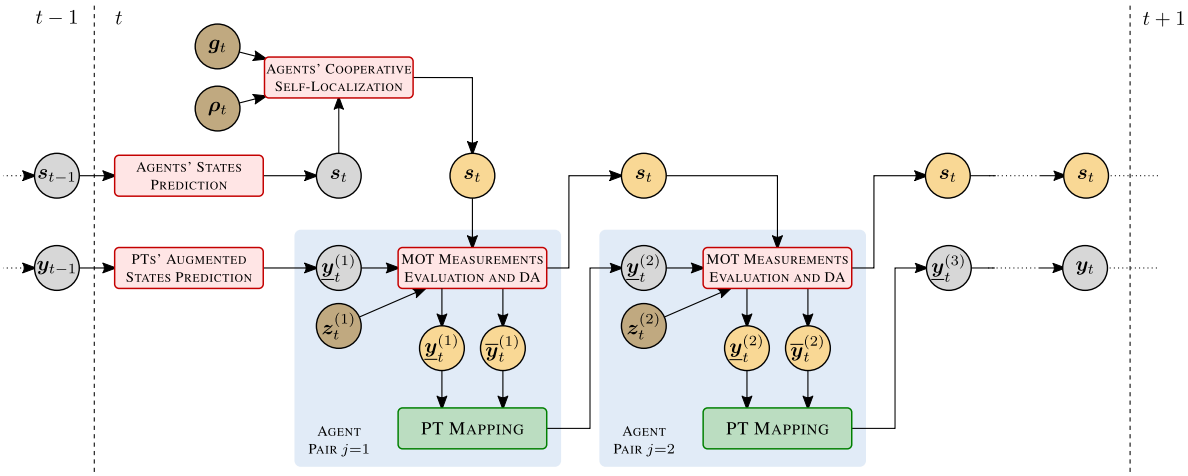


FIGURE 3. Block diagram providing the sequence of operations performed by the proposed method at time t. The red boxes represent the operations performed by the SPA-based algorithm as stated by the factorization of the joint posterior pdf in (29); the green boxes represent the PT mapping operations carried out between two consecutive agent pairs. The nodes represent random vectors: in particular, brown nodes are observations (i.e., navigation data, inter-agent and MOT measurements), and yellow nodes indicate random vectors whose *beliefs* are updated through observations, following either the "agents' cooperative self-localization" or the "MOT measurements evaluation and DA" operations. The arrows link the random vectors to the operations that involve them.

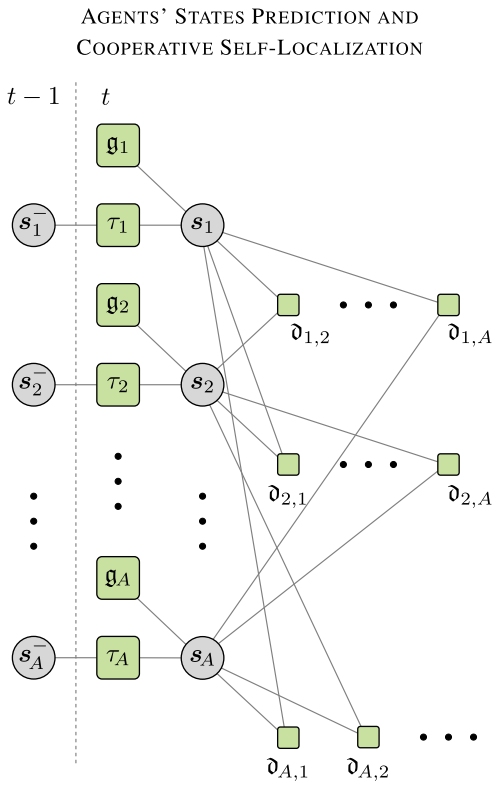


FIGURE 4. Factor graph representing the agents' states prediction and agents' cooperative self-localization portions of the factorization in (29) for one time step t. Grey circles are variable nodes, and green squares are factor nodes. The following short notations are used: $s_a^- \triangleq s_{a,t-1}$; $s_a \triangleq s_{a,t}$; $\mathfrak{g}_a \triangleq \mathfrak{g}_a(\boldsymbol{g}_{a,t}|\boldsymbol{s}_{a,t})$; $\tau_a \triangleq \tau_a(\boldsymbol{s}_{a,t}|\boldsymbol{s}_{a,t-1})$; and $\mathfrak{d}_{a,a'} \triangleq \mathfrak{d}(\boldsymbol{\rho}_t^{(a,a')}|\boldsymbol{s}_{a,t},\boldsymbol{s}_{a',t})$.

Fig. 6 changes, at each time t and agent pair j, according to the number of legacy PTs and number of MOT measurements, as well as to the kind of MOT measurements, that is, monostatic, if $j_1 = j_2$, or bistatic, if $j_1 \neq j_2$; the latter is there illustrated, with two separate variable nodes for the Rx-agent, i.e., "$\boldsymbol{s}_{j_1}$," and the Tx-agent, i.e., "$\boldsymbol{s}_{j_2}$".

1) AGENTS' STATES PREDICTION

The prediction of the state of agent $a \in \mathcal{A}$ is performed by computing the message $\zeta_{\tau_a \to \boldsymbol{s}_a}(\boldsymbol{s}_{a,t})$ from factor node "τ_a" to variable node "$\boldsymbol{s}_a$" in Fig. 4. The expression of this message is as follows

$$\zeta_{\tau_a \to \boldsymbol{s}_a}(\boldsymbol{s}_{a,t}) = \int \tau_a(\boldsymbol{s}_{a,t}|\boldsymbol{s}_{a,t-1}) \tilde{f}_J(\boldsymbol{s}_{a,t-1})\, \mathrm{d}\boldsymbol{s}_{a,t-1},$$

where $\tilde{f}_J(\boldsymbol{s}_{a,t-1})$ is the belief of the agent state at previous time $t-1$, computed at the last agent pair J; its expression is provided later in Section IV-B5.

2) AGENTS' COOPERATIVE SELF-LOCALIZATION

The variable nodes "$\boldsymbol{s}_a$," and factor nodes "$\mathfrak{d}_{a,a'}$" and "$\mathfrak{d}_{a',a}$" in Fig. 4 define a loopy graph. Therefore, the messages related to the agents' cooperative self-localization are iteratively computed as follows. At each iteration $n = 1, \ldots, N_{\mathrm{SL}}$ of the agents' cooperative self-localization loop, the messages $\eta^{(n)}_{\boldsymbol{s}_a \to \mathfrak{d}_{a,a'}}(\boldsymbol{s}_{a,t})$ and $\eta^{(n)}_{\boldsymbol{s}_a \to \mathfrak{d}_{a',a}}(\boldsymbol{s}_{a,t})$ are calculated as

$$\begin{aligned}\eta^{(n)}_{\boldsymbol{s}_a \to \mathfrak{d}_{a,a'}}\left(\boldsymbol{s}_{a,t}\right) &= \zeta_{\tau_a \to \boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \zeta_{\mathfrak{g}_a \to \boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \\ &\quad \times \left(\prod_{\substack{a'' \in \mathcal{T}_t^{(a)} \\ a'' \neq a'}} \zeta^{(n-1)}_{\mathfrak{d}_{a,a''} \to \boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \right) \\ &\quad \times \left(\prod_{a'' \in \mathcal{R}_t^{(a)}} \zeta^{(n-1)}_{\mathfrak{d}_{a'',a} \to \boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \right), \end{aligned} \tag{32}$$

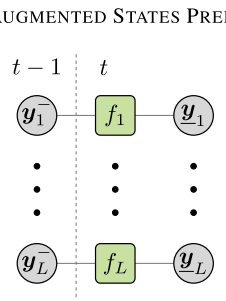


FIGURE 5. Factor graph representing the PTs' augmented states prediction portion of the factorization in (29) for one time step t. Grey circles are variable nodes, and green squares are factor nodes. The following short notations are used: $L \triangleq L_t^{(1)}$; $\underline{y}_\ell^- \triangleq \boldsymbol{y}_{\ell,t-1}$; $\underline{\boldsymbol{y}}_\ell \triangleq \underline{\boldsymbol{y}}_{\ell,t}^{(1)}$; and $f_\ell \triangleq f(\underline{\boldsymbol{y}}_{\ell,t}^{(1)}|\boldsymbol{y}_{\ell,t-1})$.

and

$$\eta_{\boldsymbol{s}_a\to\mathfrak{d}_{a',a}}^{(n)}\left(\boldsymbol{s}_{a,t}\right) = \zeta_{\tau_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right)\zeta_{\mathfrak{g}_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \times\left(\prod_{a''\in\mathcal{T}_t^{(a)}}\zeta_{\mathfrak{d}_{a,a''}\to\boldsymbol{s}_a}^{(n-1)}\left(\boldsymbol{s}_{a,t}\right)\right) \times\left(\prod_{\substack{a''\in\mathcal{R}_t^{(a)}\\ a''\neq a'}}\zeta_{\mathfrak{d}_{a'',a}\to\boldsymbol{s}_a}^{(n-1)}\left(\boldsymbol{s}_{a,t}\right)\right). \tag{33}$$

We observe that the factor node $\mathfrak{d}_{a,a'}$ refers to the likelihood of the inter-agent measurement $\boldsymbol{\rho}_t^{(a,a')}$ produced by Rx-agent a using the signal transmitted by Tx-agent a'; vice versa, the factor node $\mathfrak{d}_{a',a}$ refers to the likelihood of the inter-agent measurement $\boldsymbol{\rho}_t^{(a',a)}$ produced by Rx-agent a' using the signal transmitted by Tx-agent a. Therefore, the agent state $\boldsymbol{s}_{a,t}$ acts as Rx-agent for the message $\eta_{\boldsymbol{s}_a\to\mathfrak{d}_{a,a'}}^{(n)}$, and as Tx-agent for the message $\eta_{\boldsymbol{s}_a\to\mathfrak{d}_{a',a}}^{(n)}$. In both (32) and (33), the message $\zeta_{\mathfrak{g}_a\to\boldsymbol{s}_a}(\boldsymbol{s}_{a,t})$, that is passed from factor node "$\mathfrak{g}_a$" to variable node "$\boldsymbol{s}_a$," is computed as

$$\zeta_{\mathfrak{g}_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) = \begin{cases}\mathfrak{g}_a\left(\boldsymbol{g}_{a,t}\middle|\boldsymbol{s}_{a,t}\right) & a\in\mathcal{A}_t^{\boldsymbol{g}},\\ 1 & a\notin\mathcal{A}_t^{\boldsymbol{g}},\end{cases}$$

and messages $\zeta_{\mathfrak{d}_{a,a'}\to\boldsymbol{s}_a}^{(n)}(\boldsymbol{s}_{a,t})$ and $\zeta_{\mathfrak{d}_{a',a}\to\boldsymbol{s}_a}^{(n)}(\boldsymbol{s}_{a,t})$ are calculated as

$$\zeta_{\mathfrak{d}_{a,a'}\to\boldsymbol{s}_a}^{(n)}\left(\boldsymbol{s}_{a,t}\right) = \int\mathfrak{d}\left(\boldsymbol{\rho}_t^{(a,a')}\middle|\boldsymbol{s}_{a,t},\boldsymbol{s}_{a',t}\right)\eta_{\boldsymbol{s}_{a'}\to\mathfrak{d}_{a,a'}}^{(n)}\left(\boldsymbol{s}_{a',t}\right)\mathrm{d}\boldsymbol{s}_{a',t}, \tag{34}$$

and

$$\zeta_{\mathfrak{d}_{a',a}\to\boldsymbol{s}_a}^{(n)}\left(\boldsymbol{s}_{a,t}\right) = \int\mathfrak{d}\left(\boldsymbol{\rho}_t^{(a',a)}\middle|\boldsymbol{s}_{a,t},\boldsymbol{s}_{a',t}\right)\eta_{\boldsymbol{s}_{a'}\to\mathfrak{d}_{a',a}}^{(n)}\left(\boldsymbol{s}_{a',t}\right)\mathrm{d}\boldsymbol{s}_{a',t}. \tag{35}$$

The iteration constituted by (32)–(35) is initialized by $\zeta_{\mathfrak{d}_{a,a'}\to\boldsymbol{s}_a}^{(0)}(\boldsymbol{s}_{a,t}) = 1$ and $\zeta_{\mathfrak{d}_{a',a}\to\boldsymbol{s}_a}^{(0)}(\boldsymbol{s}_{a,t}) = 1$. Eventually, at iteration $n = N_{\mathrm{SL}}$, the belief of each agent state $a\in\mathcal{A}$ after cooperative self-localization is calculated as:

$$\tilde{f}_0(\boldsymbol{s}_{a,t}) = \frac{1}{C_{a,t}^{(0)}}\zeta_{\tau_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right)\zeta_{\mathfrak{g}_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right) \times\left(\prod_{a'\in\mathcal{R}_t^{(a)}}\zeta_{\mathfrak{d}_{a',a}\to\boldsymbol{s}_a}^{(N_{\mathrm{SL}})}\left(\boldsymbol{s}_{a,t}\right)\right)\left(\prod_{a'\in\mathcal{T}_t^{(a)}}\zeta_{\mathfrak{d}_{a,a'}\to\boldsymbol{s}_a}^{(N_{\mathrm{SL}})}\left(\boldsymbol{s}_{a,t}\right)\right), \tag{36}$$

where $C_{a,t}^{(0)}$ is a normalization constant defined such that $\int\tilde{f}_0(\boldsymbol{s}_{a,t})\,\mathrm{d}\boldsymbol{s}_{a,t} = 1$.

3) PTS' AUGMENTED STATES PREDICTION

Let us recall that for each PT augmented state $\boldsymbol{y}_{\ell,t-1}$, $\ell\in\mathcal{K}_{t-1}$, at time $t-1$, there is one legacy PT augmented state $\underline{\boldsymbol{y}}_{\ell,t}^{(1)}$, $\ell\in\mathcal{L}_t^{(1)}$, at the first agent pair at time t. The prediction of PT ℓ is performed by computing the message $\zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}(\underline{\boldsymbol{y}}_{\ell,t}^{(1)})$ from factor node "f_ℓ" to variable node "$\underline{\boldsymbol{y}}_\ell$" in Fig. 5 as

$$\begin{aligned}\zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}\left(\underline{\boldsymbol{y}}_{\ell,t}^{(1)}\right) &= \zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)}\right)\\ &= \sum_{r_{\ell,t-1}\in\{0,1\}}\int f\left(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)}\middle|\boldsymbol{x}_{\ell,t-1},r_{\ell,t-1}\right)\\ &\qquad\times\tilde{f}_J\left(\boldsymbol{x}_{\ell,t-1},r_{\ell,t-1}\right)\mathrm{d}\boldsymbol{x}_{\ell,t-1},\end{aligned} \tag{37}$$

where $\tilde{f}_J(\boldsymbol{x}_{\ell,t-1},r_{\ell,t-1})$ is the belief of the PT augmented state at previous time $t-1$ computed at the last agent pair J; the computation of this belief is detailed in the Section IV-B5. Note that, from (37) and the fact that $\tilde{f}_J(\boldsymbol{x}_{\ell,t-1},r_{\ell,t-1})$ is normalized, it follows that $\zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)})$ is normalized too, i.e., $\sum_{\underline{r}_{\ell,t}^{(1)}\in\{0,1\}}\int\zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)})\mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(1)} = 1$.

4) MOT MEASUREMENTS EVALUATION AND DATA ASSOCIATION

The MOT measurements evaluation and data association steps are performed at each agent pair j, sequentially from $j=1$ to $j=J$. With reference to Fig. 6, the MOT measurements evaluation step consists in computing the following messages: from factor nodes "q_ℓ" to variable nodes "α_ℓ," i.e., $\zeta_{q_\ell\to\alpha_\ell}(\alpha_{\ell,t}^{(j)})$; from factor nodes "$h_a$" to variable nodes "$\alpha_a^\star$," i.e., $\zeta_{h_a\to\alpha_a^\star}(\alpha_{L+a}^{(j)})$; and from factor nodes "υ_m" to variable nodes "β_m," i.e., $\zeta_{\upsilon_m\to\beta_m}(\beta_{m,t}^{(j)})$. We recall that the MOT measurements can be either monostatic, if $j_1 = j_2$, or bistatic, if $j_1\neq j_2$; next, we provide expressions of these messages for both cases by using the Dirac delta and the Kronecker delta.

The message $\zeta_{q_\ell\to\alpha_\ell}(\alpha_{\ell,t}^{(j)})$ is computed as

$$\begin{aligned}\zeta_{q_\ell\to\alpha_\ell}\left(\alpha_{\ell,t}^{(j)}\right) &= \sum_{\underline{r}_{\ell,t}^{(j)}\in\{0,1\}}\iiint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)},\alpha_{\ell,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_t^{(j)}\right)\\ &\quad\times\tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)}\right)\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\ &\quad\times\left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}}\mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\,\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}.\end{aligned} \tag{38}$$

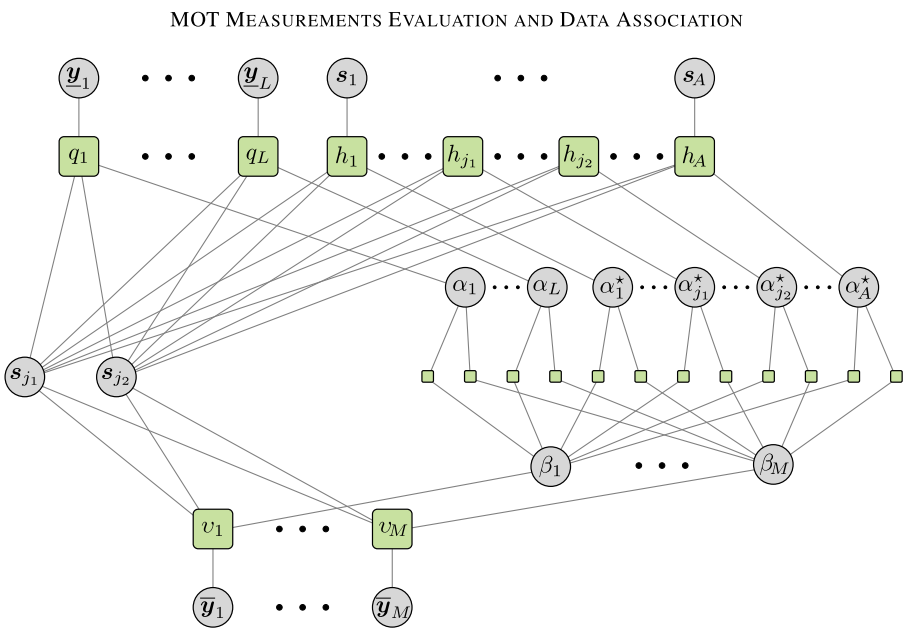


FIGURE 6. **Factor graph representing the MOT measurements evaluation and data association portion of the factorization in (29) for one time step t and agent pair j. Grey circles are variable nodes, and green squares are factor nodes. The following short notations are used:** $L \triangleq L_t^{(j)}$; $\underline{\boldsymbol{y}}_\ell \triangleq \underline{\boldsymbol{y}}_{\ell,t}^{(j)}$; $\boldsymbol{s}_a \triangleq \boldsymbol{s}_{a,t}$; $q_\ell \triangleq q(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)})$; $h_a \triangleq h(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)})$; $\alpha_\ell \triangleq \alpha_{\ell,t}^{(j)}$; $\alpha_a^\star \triangleq \alpha_{L+a,t}^{(j)}$; $M \triangleq M_t^{(j)}$; $\beta_m \triangleq \beta_{m,t}^{(j)}$; $\upsilon_m \triangleq \upsilon(\bar{\boldsymbol{y}}_{m,t}^{(j)}, \beta_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)})$; **and** $\bar{\boldsymbol{y}}_m \triangleq \bar{\boldsymbol{y}}_{m,t}^{(j)}$.

For monostatic MOT measurements, i.e., for $j_1 = j_2$, the Kronecker delta δ_{j_1,j_2} is 1; moreover, according to (20)–(21) and (6)–(7), and to (24)–(25) and (11), the function $q(\cdot) = q_1(\cdot)\, q_2(\cdot)$ does not depend on the vector $\boldsymbol{s}_{j_2,t}$ (note that this also applies to function $h(\cdot) = h_1(\cdot)\, h_2(\cdot)$ according to (22) and (26)–(27), and function $\upsilon(\cdot)$ according to (28) and (11)). Therefore, the message in (38) in the case of monostatic MOT measurements particularizes as

$$\begin{aligned}
\zeta_{q_\ell \to \alpha_\ell}\left(\alpha_{\ell,t}^{(j)}\right) &= \sum_{\underline{r}_{\ell,t}^{(j)} \in \{0,1\}} \iiint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}; \boldsymbol{z}_t^{(j)}\right) \\
&\quad \times \tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right) \\
&\quad \times \delta\left(\boldsymbol{s}_{j_2,t}\right) \mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_1,t} \mathrm{d}\boldsymbol{s}_{j_2,t} \\
&= \int \delta\left(\boldsymbol{s}_{j_2,t}\right) \mathrm{d}\boldsymbol{s}_{j_2,t} \sum_{\underline{r}_{\ell,t}^{(j)} \in \{0,1\}} \iint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}; \boldsymbol{z}_t^{(j)}\right) \\
&\quad \times \tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right) \mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_1,t} \\
&= \sum_{\underline{r}_{\ell,t}^{(j)} \in \{0,1\}} \iint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}; \boldsymbol{z}_t^{(j)}\right) \\
&\quad \times \tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right) \mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_1,t},
\end{aligned}$$

where we used the fact that $\int \delta(\boldsymbol{s}_{a,t})\, \mathrm{d}\boldsymbol{s}_{a,t} = 1$. For bistatic MOT measurements, i.e., for $j_1 \neq j_2$, instead, the Kronecker delta δ_{j_1,j_2} is 0, and the message in (38) becomes

$$\begin{aligned}
&\zeta_{q_\ell \to \alpha_\ell}\left(\alpha_{\ell,t}^{(j)}\right) \\
&\quad = \sum_{\underline{r}_{\ell,t}^{(j)} \in \{0,1\}} \iiint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \\
&\qquad \times \tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right) \\
&\qquad \times \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right) \mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_1,t}\, \mathrm{d}\boldsymbol{s}_{j_2,t}.
\end{aligned}$$

We note that the message in (38) depends on the beliefs at the previous agent pair $j-1$ of the Rx-agent and Tx-agent states, i.e., $\tilde{f}_{j-1}(\boldsymbol{s}_{j_1,t})$ and $\tilde{f}_{j-1}(\boldsymbol{s}_{j_2,t})$, respectively, and of the legacy PTs augmented states, i.e., $\tilde{f}_{j-1}(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)})$. These beliefs are initialized at $j = 1$ with the prediction messages of the agent states and PT augmented states described in Section IV-B2 and Section IV-B3, respectively. Specifically, the agent states beliefs are initialized with $\tilde{f}_0(\boldsymbol{s}_{j_1,t})$ and $\tilde{f}_0(\boldsymbol{s}_{j_2,t})$ as computed in (36), and the PT augmented states beliefs are initialized with $\tilde{f}_0(\underline{\boldsymbol{x}}_{\ell,t}^{(1)}, \underline{r}_{\ell,t}^{(1)}) = \zeta_{f_\ell \to \underline{\boldsymbol{y}}_\ell}(\underline{\boldsymbol{x}}_{\ell,t}^{(1)}, \underline{r}_{\ell,t}^{(1)})$ as computed in (37). Finally, we recall that before processing the MOT measurements at any agent pair $j > 1$, legacy and new PTs at agent pair $j-1$ are mapped into legacy PTs at agent pair j. The PT mapping is formally described by the following expression

$$\begin{aligned}
\tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) &= \sum_{r_{k,t}^{(j-1)} \in \{0,1\}} \delta_{r_{k,t}^{(j-1)}, \underline{r}_{\ell,t}^{(j)}} \\
&\quad \times \int \tilde{f}_{j-1}\left(\boldsymbol{x}_{k,t}^{(j-1)}, r_{k,t}^{(j-1)}\right) \delta\left(\boldsymbol{x}_{k,t}^{(j-1)} - \underline{\boldsymbol{x}}_{\ell,t}^{(j)}\right) \mathrm{d}\boldsymbol{x}_{k,t}^{(j-1)},
\end{aligned}$$

where, with an abuse of notation, $\boldsymbol{x}_{k,t}^{(j-1)}$ and $r_{k,t}^{(j-1)}$ represent the state and the existence variable of either a legacy or a new PT at agent pair $j-1$.

The message $\zeta_{h_a\to\alpha_a^\star}(\alpha_{L+a}^{(j)})$ is expressed as

$$\begin{aligned}\zeta_{h_a\to\alpha_a^\star}&\left(\alpha_{L+a,t}^{(j)}\right)\\&=\iiint h\left(\boldsymbol{s}_{a,t},\alpha_{L+a,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_t^{(j)}\right)\\&\quad\times\tilde{f}_{j-1}\left(\boldsymbol{s}_{a,t}\right)\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\&\quad\times\left(\delta(\boldsymbol{s}_{j_2,t})\right)^{\delta_{j_1,j_2}}\,\mathrm{d}\boldsymbol{s}_{a,t}\,\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}.\end{aligned}$$

We observe that, according to (22), (26), and (27), if agent a is either the Rx-agent j_1, or the Tx-agent j_2, the function $h(\cdot)$ is non-zero — specifically, equal to 1 — if and only if $\alpha_{L+a,t}^{(j)}=0$. Thus, if $a=j_1$ or $a=j_2$, then $\zeta_{h_a\to\alpha_a^\star}(\alpha_{L+a,t}^{(j)})=\delta_{\alpha_{L+a,t}^{(j)},0}$; this follows from the fact that Rx-agent j_1 and Tx-agent j_2 cannot produce MOT measurements at agent pair j.

Finally, the message $\zeta_{\upsilon_m\to\beta_m}(\beta_{m,t}^{(j)})$ is calculated as

$$\begin{aligned}\zeta_{\upsilon_m\to\beta_m}&\left(\beta_{m,t}^{(j)}\right)\\&=\sum_{\bar{r}_{m,t}^{(j)}\in\{0,1\}}\iiint\upsilon\left(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)},\beta_{m,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_{m,t}^{(j)}\right)\\&\quad\times\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\&\quad\times\left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}}\mathrm{d}\overline{\boldsymbol{x}}_{m,t}^{(j)}\,\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}.\end{aligned}$$

The data association step is an iterative procedure that converts the messages $\zeta_{q_\ell\to\alpha_\ell}(\alpha_{\ell,t}^{(j)})$, $\zeta_{h_a\to\alpha_a^\star}(\alpha_{L+a,t}^{(j)})$, and $\zeta_{\upsilon_m\to\beta_m}(\beta_{m,t}^{(j)})$, into the following messages: $\eta_{\alpha_\ell\to q_\ell}(\alpha_{\ell,t}^{(j)})$, from variable nodes "$\alpha_\ell$" to factor nodes "$q_\ell$"; $\eta_{\alpha_a^\star\to h_a}(\alpha_{L+a,t}^{(j)})$, from variable nodes "$\alpha_a^\star$" to factor nodes "$h_a$"; and $\eta_{\beta_m\to\upsilon_m}(\beta_{m,t}^{(j)})$, from variable nodes "$\beta_m$" to factor nodes "$\upsilon_m$". This iterative procedure is described in [49, Sec. IX-A3], and expressions of the messages are provided therein.

5) BELIEFS CALCULATION

Once the MOT measurements produced at agent pair j are incorporated, the information they provide is used to eventually update the agent states and PT augmented states beliefs. For the legacy PT augmented states, the following messages are computed from factor nodes "q_ℓ" to variable nodes "$\underline{\boldsymbol{y}}_\ell$," $\ell\in\mathcal{L}_t^{(j)}$, that is,

$$\begin{aligned}\zeta_{q_\ell\to\underline{\boldsymbol{y}}_\ell}\left(\underline{\boldsymbol{y}}_{\ell,t}^{(j)}\right)&=\zeta_{q_\ell\to\underline{\boldsymbol{y}}_\ell}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)}\right)=\sum_{\alpha_{\ell,t}^{(j)}=0}^{M_t^{(j)}}\eta_{\alpha_\ell\to q_\ell}\left(\alpha_{\ell,t}^{(j)}\right)\\&\quad\times\iint q(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)},\alpha_{\ell,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_t^{(j)})\\&\quad\times\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\&\quad\times\left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}}\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}.\end{aligned}$$

Then, the updated beliefs are obtained as

$$\tilde{f}_j\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)}\right)=\frac{1}{\underline{C}_{\ell,t}^{(j)}}\tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)}\right)\,\zeta_{q_\ell\to\underline{\boldsymbol{y}}_\ell}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)}\right),$$

where $\underline{C}_{\ell,t}^{(j)}$ is a normalization constant defined such that $\sum_{\underline{r}_{\ell,t}^{(j)}\in\{0,1\}}\int\tilde{f}_j(\underline{\boldsymbol{x}}_{\ell,t}^{(j)},\underline{r}_{\ell,t}^{(j)})\,\mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}=1$. (For clarity, we recall that $\tilde{f}_0(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)})=\zeta_{f_\ell\to\underline{\boldsymbol{y}}_\ell}(\underline{\boldsymbol{x}}_{\ell,t}^{(1)},\underline{r}_{\ell,t}^{(1)})$.) Similarly, for the new PT augmented states, the messages from factor nodes "υ_m" to variable nodes "$\overline{\boldsymbol{y}}_m$" are computed as

$$\begin{aligned}\zeta_{\upsilon_m\to\bar{\boldsymbol{y}}_m}&\left(\bar{\boldsymbol{y}}_{m,t}^{(j)}\right)=\zeta_{\upsilon_m\to\bar{\boldsymbol{y}}_m}\left(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)}\right)\\&=\sum_{\beta_{m,t}^{(j)}=0}^{L_t^{(j)}+A}\eta_{\beta_m\to\upsilon_m}\left(\beta_{m,t}^{(j)}\right)\\&\quad\times\iint\upsilon\left(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)},\beta_{m,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_{m,t}^{(j)}\right)\\&\quad\times\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\&\quad\times\left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}}\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}\end{aligned}$$

and the updated beliefs are obtained as

$$\tilde{f}_j\left(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)}\right)=\frac{1}{\overline{C}_{m,t}^{(j)}}\zeta_{\upsilon_m\to\bar{\boldsymbol{y}}_m}\left(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)}\right),$$

where the normalization constant $\overline{C}_{m,t}^{(j)}$ is defined such that $\sum_{\bar{r}_{m,t}^{(j)}\in\{0,1\}}\int\tilde{f}_j(\overline{\boldsymbol{x}}_{m,t}^{(j)},\bar{r}_{m,t}^{(j)})\,\mathrm{d}\overline{\boldsymbol{x}}_{m,t}^{(j)}=1$.

For the agent states $a\in\mathcal{A}$, the messages $\zeta_{h_a\to\boldsymbol{s}_a}(\boldsymbol{s}_{a,t})$ passed from the factor nodes "h_a" to the variable nodes "$\boldsymbol{s}_a$" are calculated according to

$$\begin{aligned}\zeta_{h_a\to\boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right)&=\sum_{\alpha_{L+a,t}^{(j)}=0}^{M_t^{(j)}}\eta_{\alpha_a^\star\to h_a}\left(\alpha_{L+a,t}^{(j)}\right)\\&\quad\times\iint h\left(\boldsymbol{s}_{a,t},\alpha_{L+a,t}^{(j)},\boldsymbol{s}_{j_1,t},\boldsymbol{s}_{j_2,t};\boldsymbol{z}_t^{(j)}\right)\\&\quad\times\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right)\left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}}\\&\quad\times\left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}}\,\mathrm{d}\boldsymbol{s}_{j_1,t}\,\mathrm{d}\boldsymbol{s}_{j_2,t}.\end{aligned}\tag{39}$$

As before, we observe that if agent a is either the Rx-agent j_1, or the Tx-agent j_2, i.e., $a=j_1$ or $a=j_2$, the function $h(\cdot)$

is non-zero if and only if $\alpha_{L+a,t}^{(j)} = 0$, from which it follows that $\zeta_{h_a \to \boldsymbol{s}_a}(\boldsymbol{s}_{a,t}) = \eta_{\alpha_a^\star \to h_a}(\alpha_{L+a,t}^{(j)} = 0)$. Additionally, for Rx-agent j_1 and Tx-agent j_2 further computations are needed, as the following messages have to be calculated: $\zeta_{q_\ell \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$ and $\zeta_{q_\ell \to \boldsymbol{s}_{j_2}}(\boldsymbol{s}_{j_2,t})$, from factor nodes "$q_\ell$" to variable nodes "$\boldsymbol{s}_{j_1}$" and "$\boldsymbol{s}_{j_2}$," respectively; $\zeta_{\upsilon_m \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$ and $\zeta_{\upsilon_m \to \boldsymbol{s}_{j_2}}(\boldsymbol{s}_{j_2,t})$, from factor nodes "$\upsilon_m$" to variable nodes "$\boldsymbol{s}_{j_1}$" and "$\boldsymbol{s}_{j_2}$," respectively; and $\zeta_{h_a \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$ and $\zeta_{h_a \to \boldsymbol{s}_{j_2}}(\boldsymbol{s}_{j_2,t})$, from factor nodes "$h_a$" to variable nodes "$\boldsymbol{s}_{j_1}$" and "$\boldsymbol{s}_{j_2}$," respectively. Hereafter, we provide the expressions of these messages for the Rx-agent j_1; the messages related to the Tx-agent j_2 can be derived similarly by substituting j_1 with j_2 and vice versa. The messages $\zeta_{q_\ell \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$, $\ell \in \mathcal{L}_t^{(j)}$, are defined as

$$\begin{aligned}\zeta_{q_\ell \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t}) = &\sum_{\alpha_{\ell,t}^{(j)}=0}^{M_t^{(j)}} \sum_{\underline{r}_{\ell,t}^{(j)} \in \{0,1\}} \eta_{\alpha_\ell \to q_\ell}\left(\alpha_{\ell,t}^{(j)}\right) \\ &\times \iint q\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}, \alpha_{\ell,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \\ &\times \tilde{f}_{j-1}\left(\underline{\boldsymbol{x}}_{\ell,t}^{(j)}, \underline{r}_{\ell,t}^{(j)}\right) \left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}} \\ &\times \left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}} \mathrm{d}\underline{\boldsymbol{x}}_{\ell,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_2,t};\end{aligned}$$

the messages $\zeta_{\upsilon_m \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$, $m \in \mathcal{M}_t^{(j)}$, are defined as

$$\begin{aligned}\zeta_{\upsilon_m \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t}) = &\sum_{\beta_{m,t}^{(j)}=0}^{L_t^{(j)}+A} \sum_{\overline{r}_{m,t}^{(j)} \in \{0,1\}} \eta_{\beta_m \to \upsilon_m}\left(\beta_{m,t}^{(j)}\right) \\ &\times \iint \upsilon\left(\overline{\boldsymbol{x}}_{m,t}^{(j)}, \overline{r}_{m,t}^{(j)}, \beta_{m,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_{m,t}^{(j)}\right) \\ &\times \left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}} \\ &\times \left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}} \mathrm{d}\overline{\boldsymbol{x}}_{m,t}^{(j)}\, \mathrm{d}\boldsymbol{s}_{j_2,t};\end{aligned}$$

finally, the messages $\zeta_{h_a \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t})$, $a \in \mathcal{A}$, are defined as

$$\begin{aligned}\zeta_{h_a \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t}) = &\sum_{\alpha_{L+a,t}^{(j)}=0}^{M_t^{(j)}} \eta_{\alpha_a^\star \to h_a}\left(\alpha_{L+a,t}^{(j)}\right) \\ &\times \iint h\left(\boldsymbol{s}_{a,t}, \alpha_{L+a,t}^{(j)}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}; \boldsymbol{z}_t^{(j)}\right) \\ &\times \tilde{f}_{j-1}\left(\boldsymbol{s}_{a,t}\right) \left(\tilde{f}_{j-1}\left(\boldsymbol{s}_{j_2,t}\right)\right)^{1-\delta_{j_1,j_2}} \\ &\times \left(\delta\left(\boldsymbol{s}_{j_2,t}\right)\right)^{\delta_{j_1,j_2}} \mathrm{d}\boldsymbol{s}_{a,t}\, \mathrm{d}\boldsymbol{s}_{j_2,t}. \qquad (40)\end{aligned}$$

We observe that the expression of the message in (40) is consistent with the expression in (39) in that, if $a = j_1$, then $\zeta_{h_a \to \boldsymbol{s}_{j_1}}(\boldsymbol{s}_{j_1,t}) = \eta_{\alpha_a^\star \to h_a}(\alpha_{L+a,t}^{(j)} = 0)$. Eventually, the updated belief of the Rx-agent state is computed as

$$\begin{aligned}\tilde{f}_j\left(\boldsymbol{s}_{j_1,t}\right) = &\frac{1}{C_{j_1,t}^{(j)}} \tilde{f}_{j-1}\left(\boldsymbol{s}_{j_1,t}\right) \prod_{\ell \in \mathcal{L}_t^{(j)}} \zeta_{q_\ell \to \boldsymbol{s}_{j_1}}\left(\boldsymbol{s}_{j_1,t}\right) \\ &\times \prod_{m \in \mathcal{M}_t^{(j)}} \zeta_{\upsilon_m \to \boldsymbol{s}_{j_1}}\left(\boldsymbol{s}_{j_1,t}\right) \prod_{a \in \mathcal{A}} \zeta_{h_a \to \boldsymbol{s}_{j_1}}\left(\boldsymbol{s}_{j_1,t}\right),\end{aligned}$$

while the belief for any other agent state, i.e., $a \in \mathcal{A} \setminus \{j_1, j_2\}$, is computed as

$$\tilde{f}_j\left(\boldsymbol{s}_{a,t}\right) = \frac{1}{C_{a,t}^{(j)}} \tilde{f}_{j-1}\left(\boldsymbol{s}_{a,t}\right)\, \zeta_{h_a \to \boldsymbol{s}_a}\left(\boldsymbol{s}_{a,t}\right),$$

where the normalization constant $C_{a,t}^{(j)}$ is defined such that $\int \tilde{f}_j(\boldsymbol{s}_{a,t})\, \mathrm{d}\boldsymbol{s}_{a,t} = 1$; we recall that $\tilde{f}_0(\boldsymbol{s}_{a,t})$ is given in (36).

6) IMPLEMENTATION DETAILS

The SPA-based joint localization and tracking algorithm detailed in this section is implemented following a particle-based approach, where each pdf is described by a set of N_P particles. This choice is particularly appropriate for non-Gaussian settings and pdf, which intrinsically appear in data association problems [35], [51], [55]. Although both agent states and PT states are here unknown — unlike other SPA-based approaches that assume perfect knowledge of the sensors' position [35], [49] — the proposed algorithm scales quadratically with N_P by virtue of a proper stacking of the particles. Furthermore, it scales linearly with the number of MOT measurements, number of agent pairs, and number of iterations of the agents' cooperative self-localization loop and the data association loop, and quadratically with the number of legacy PTs.

Moreover, as mentioned in Section II-C, in order to keep a tractable number of PTs over time, a pruning step is performed. Specifically, once all the MOT measurements at time t are processed, any PT $k \in \mathcal{K}_t$ with existence probability $f(r_{k,t} = 1 | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$ smaller than a threshold P_pr, is removed from the set $\mathcal{K}_t$ and it is not carried over to the next time $t+1$ as legacy PT. Besides this pruning step, the number K_t of PTs at time t is unbounded.

V. EXPERIMENTAL RESULTS

In this section, performance results of the proposed joint cooperative self-localization and multitarget tracking approach are provided. The maritime domain is considered as application scenario, thus the kinematics of agents and targets, the accuracy of measurements, and other parameters are chosen accordingly. First, in Section V-A, a simulated scenario is used to show how to take advantage of target information for the localization of agents. Then, in Section V-B, an application to real maritime data is presented.

A. SIMULATED SCENARIO

1) SET-UP

The simulated scenario is shown in Fig. 7. An area of 10×10 km is surveyed by $A = 4$ agents over 50 time steps; the time step duration is $T_\mathrm{s} = 30$ s. Agents $a = 1$, 2, and 3 move counterclockwise along a circle of radius 3.5 km and center (0,0) at a constant radial velocity of 0.69 m/s, while agent $a = 4$ is anchored at the center. The stationary agent is the only Tx-agent of the considered scenario, while the moving ones are

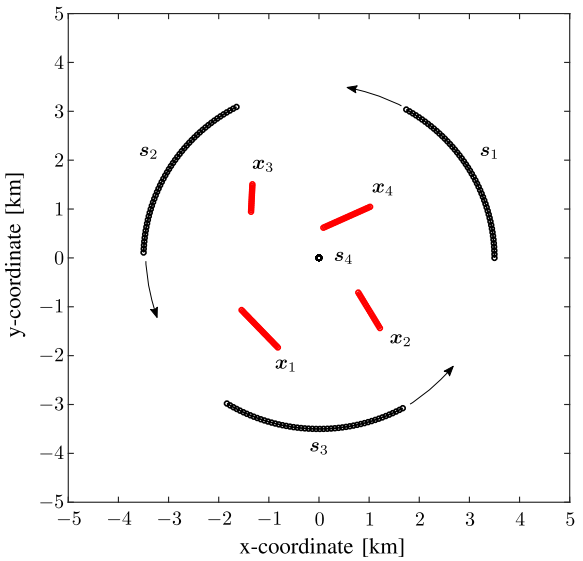


FIGURE 7. Illustration of the simulated scenario (the time index t is omitted). The black circles are the agents' positions over time: the arrows represent the counterclockwise directions of agents $a = 1, 2$, and 3, while agent $a = 4$ is anchored. The red circles are the targets's positions over time.

all Rx-agents, leading to the following sets: $\mathcal{R} = \{1, 2, 3\}$ and $\mathcal{T} = \{4\}$. Agents can communicate and sense over the whole area, i.e., there is no limitation on the sensing/communication range. The agent state $\boldsymbol{s}_{a,t} = [\check{\boldsymbol{s}}_{a,t}^{\mathrm{T}}, \dot{\check{\boldsymbol{s}}}_{a,t}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^4$ comprises both position, i.e., $\check{\boldsymbol{s}}_{a,t}$, and velocity, i.e., $\dot{\check{\boldsymbol{s}}}_{a,t}$, over a 2D space, and its dynamics is modeled with a nearly constant velocity (NCV) model, that is (cf. (1)), $\boldsymbol{s}_{a,t} = \boldsymbol{\varepsilon}_a(\boldsymbol{s}_{a,t-1}, \boldsymbol{u}_{a,t}) = \boldsymbol{A}\boldsymbol{s}_{a,t-1} + \boldsymbol{W}\boldsymbol{u}_{a,t}$, where $\boldsymbol{A} \in \mathbb{R}^{4\times4}$ and $\boldsymbol{W} \in \mathbb{R}^{4\times2}$ are as in [64, Sec. 6.3.2], and the process noise term $\boldsymbol{u}_{a,t}$ is Gaussian distributed with mean $\boldsymbol{0}$ and time-invariant covariance matrix $\omega_{\mathrm{A}}^2\mathbf{I}_2$, with per-component standard deviation (std) $\omega_{\mathrm{A}} = 0.1$ m/s^2.

The scenario also includes four mobile targets, each moving at a constant speed randomly drawn from $[-1.54, 1.54]$ m/s. They appear and disappear at different times, thus are detectable respectively in the following time intervals: $t \in [5, 35]$, $t \in [10, 40]$, $t \in [20, 40]$, and $t \in [30, 45]$. The PT state $\boldsymbol{x}_{k,t} = [\check{\boldsymbol{x}}_{k,t}^{\mathrm{T}}, \dot{\check{\boldsymbol{x}}}_{k,t}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^4$ comprises both position, i.e., $\check{\boldsymbol{x}}_{k,t}$, and velocity, i.e., $\dot{\check{\boldsymbol{x}}}_{k,t}$, over a 2D space, and its dynamics (cf. (14)) follows an NCV model similar to the one adopted for the agent states, with a per-component process noise standard deviation of $\omega_{\mathrm{T}} = 0.1$ m/s^2.

The prior pdf $f(\boldsymbol{s}_0)$, assuming that the states of the agents at time $t = 0$ are independent, can be written as

$$f(\boldsymbol{s}_0) = \prod_{a\in\mathcal{A}} f(\boldsymbol{s}_{a,0}) = \prod_{a\in\mathcal{A}} f(\check{\boldsymbol{s}}_{a,0})\, f(\dot{\check{\boldsymbol{s}}}_{a,0}),$$

where the pdf $f(\check{\boldsymbol{s}}_{a,0})$ is uniform over a circle of radius 150 m around the true position, and $f(\dot{\check{\boldsymbol{s}}}_{a,0})$ is uniform on the 2D interval $[-2.57, 2.57] \times [-2.57, 2.57]$ m/s.

Agents cooperatively localize themselves by combining navigation data $\boldsymbol{g}_{a,t}$ and inter-agent measurements $\boldsymbol{\rho}_t^{(a,a')}$. The former are available at agents $a = 3$ and 4 at all times, i.e., $\mathcal{A}_t^{\boldsymbol{g}} = \mathcal{A}^{\boldsymbol{g}} = \{3, 4\}$, and provide only position information, that is (cf. (2)), $\boldsymbol{g}_{a,t} = \boldsymbol{\theta}_a(\boldsymbol{s}_{a,t}, \boldsymbol{n}_{a,t}) = \check{\boldsymbol{s}}_{a,t} + \boldsymbol{n}_{a,t}$, $a \in \mathcal{A}^{\boldsymbol{g}}$. The noise term $\boldsymbol{n}_{a,t}$ has a Gaussian distribution with mean $\boldsymbol{0}$ and time-invariant covariance matrix $\sigma_a^2\mathbf{I}_2$, with

$$\sigma_a = \begin{cases} 20\,\mathrm{m} & a = 3\,, \\ 5\,\mathrm{m} & a = 4\,. \end{cases}$$

The latter are range-bearing measurements, available at all times and among all agents, defined as (cf. (3)):

$$\begin{aligned}\boldsymbol{\rho}_t^{(a,a')} &= \boldsymbol{\vartheta}\left(\boldsymbol{s}_{a,t}, \boldsymbol{s}_{a',t}, \boldsymbol{w}_t^{(a,a')}\right) \\ &= \begin{bmatrix} 2\|\check{\boldsymbol{s}}_{a,t} - \check{\boldsymbol{s}}_{a',t}\| \\ \angle(\check{\boldsymbol{s}}_{a',t} - \check{\boldsymbol{s}}_{a,t}) \end{bmatrix} + \boldsymbol{w}_t^{(a,a')},\end{aligned}$$

where the noise term $\boldsymbol{w}_t^{(a,a')}$ is Gaussian distributed with mean $\boldsymbol{0}$ and time-invariant covariance matrix $\mathrm{diag}(\varsigma_{\rho,\mathrm{r}}^2, \varsigma_{\rho,\mathrm{b}}^2)$, equal for all pairs (a, a'); the standard deviations are set to $\varsigma_{\rho,\mathrm{r}} = 20$ m and $\varsigma_{\rho,\mathrm{b}} = 1$ deg. The number of iterations of the agents' cooperative self-localization loop is set to $N_{\mathrm{SL}} = 5$.

At time t and agent pair j, agents and PTs give rise to MOT measurements $\boldsymbol{z}_{m,t}^{(j)}$, $m \in \mathcal{M}_t^{(j)}$, as described in Section II-B3. We recall that an MOT measurement can be monostatic ($j_1 = j_2$) or bistatic ($j_1 \neq j_2$), and can derive — unless it is a false alarm — from a reflection from a PT or from an agent (except that from the Rx-agent j_1 and the Tx-agent j_2). We thus have four possible MOT measurement models. A monostatic MOT measurement $\boldsymbol{z}_{m,t}^{(j)}$ comprises range and bearing information and, assuming that it rises from PT k, is modeled as (cf. (4))

$$\begin{aligned}\boldsymbol{z}_{m,t}^{(j)} &= \gamma_{\mathtt{mono}}\left(\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{v}_{m,t}^{(j)}\right) \\ &= \begin{bmatrix} 2\|\check{\boldsymbol{s}}_{j_1,t} - \check{\boldsymbol{x}}_{k,t}\| \\ \angle(\check{\boldsymbol{x}}_{k,t} - \check{\boldsymbol{s}}_{j_1,t}) \end{bmatrix} + \boldsymbol{v}_{m,t}^{(j)}.\end{aligned} \tag{41}$$

Similarly, a bistatic MOT measurement has a bistatic range and bearing information, with the latter representing the AoA of the reflected signal at the Rx-agent; therefore, assuming that it is generated by PT k, it is modeled as (cf. (5))

$$\begin{aligned}\boldsymbol{z}_{m,t}^{(j)} &= \gamma_{\mathtt{bi}}\left(\boldsymbol{x}_{k,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}, \boldsymbol{v}_{m,t}^{(j)}\right) \\ &= \begin{bmatrix} \|\check{\boldsymbol{s}}_{j_1,t} - \check{\boldsymbol{x}}_{k,t}\| + \|\check{\boldsymbol{s}}_{j_2,t} - \check{\boldsymbol{x}}_{k,t}\| \\ \angle(\check{\boldsymbol{x}}_{k,t} - \check{\boldsymbol{s}}_{j_1,t}) \end{bmatrix} + \boldsymbol{v}_{m,t}^{(j)}.\end{aligned} \tag{42}$$

The noise term $\boldsymbol{v}_{m,t}^{(j)}$ in both (41) and (42) is Gaussian distributed with mean $\boldsymbol{0}$ and time-invariant covariance matrix $\mathrm{diag}(\varsigma_{\mathrm{z,r}}^2, \varsigma_{\mathrm{z,b}}^2)$, equal for all agent pairs j; the standard deviations are set to $\varsigma_{\mathrm{z,r}} = 20$ m and $\varsigma_{\mathrm{z,b}} = 1$ deg. The analogous cases for monostatic/bistatic MOT measurements originating from agent's reflection can be easily derived, thus omitted. Note that all the observations, sent to a fusion centre, refer to a common spatial reference system. The detection probability is assumed constant among each agent pair, regardless of the type of MOT measurement (monostatic or bistatic); moreover, it is independent of

TABLE III Simulation Parameters

PARAMETER	SYMBOL	VALUE
Number of agents	A	4
Set of Rx-agents	$\mathcal{R}$	$\{1,2,3\}$
Set of Tx-agents	$\mathcal{T}$	$\{4\}$
Set of agents with navigation data	$\mathcal{A}^g$	$\{3,4\}$
Process noise std	ω_{A}, ω_{T}	0.1 m/s^2
Navigation data std, agent $a=3$	σ_3	20 m
Navigation data std, agent $a=4$	σ_4	5 m
Range information std	$\varsigma_{\rho,\mathrm{r}}$, $\varsigma_{\mathrm{z,r}}$	20 m
Bearing information std	$\varsigma_{\rho,\mathrm{b}}$, $\varsigma_{\mathrm{z,b}}$	1 deg
Detection probability	P_{d}	0.7
Mean number of false alarms	μ_{c}	3
New PT, prior position std	ς_{n}	500 m
Mean number of new PTs	μ_{n}	0.1
Survival probability	p_{s}	0.99
Pruning threshold	P_{pr}	0.01
PT existence threshold	P_{ex}	0.75
Number of particles	N_{P}	1000

the legacy object state, Rx-agent state, and Tx-agent state, that is, (cf. (6)) $P^{(j)}_{\mathrm{d,mono}}(\boldsymbol{o}^{(j)}_{i,t}, \boldsymbol{s}_{j_1,t}) = P^{(j)}_{\mathrm{d,bi}}(\boldsymbol{o}^{(j)}_{i,t}, \boldsymbol{s}_{j_1,t}, \boldsymbol{s}_{j_2,t}) = P_{\mathrm{d}} = 0.7$. Furthermore, the mean number of false alarms is $\mu^{(j)}_{\mathrm{c}} = \mu_{\mathrm{c}} = 3$, and the false alarm distribution $f^{(j)}_{\mathrm{c}}(\boldsymbol{z}^{(j)}_{m,t}) = f_{\mathrm{c}}(\boldsymbol{z}^{(j)}_{m,t})$ is uniform over the entire surveillance area. Finally, the new PT state $\bar{\boldsymbol{x}}^{(j)}_{m,t}$ related to MOT measurement $\boldsymbol{z}^{(j)}_{m,t}$ as described in Section II-C, is distributed according to $f_{\mathrm{n}}(\bar{\boldsymbol{x}}^{(j)}_{m,t}) = f(\breve{\boldsymbol{x}}^{(j)}_{m,t}) f(\dot{\breve{\boldsymbol{x}}}^{(j)}_{m,t})$, with Gaussian pdf $f(\breve{\boldsymbol{x}}^{(j)}_{m,t})$ centered around the Cartesian MOT measurement (converted from the range-bearing space) and with covariance matrix $\varsigma_{\mathrm{n}}^2\mathbf{I}$, with per-component standard deviation $\varsigma_{\mathrm{n}} = 500$ m, and uniform pdf $f(\dot{\breve{\boldsymbol{x}}}^{(j)}_{m,t})$ on the 2D interval $[-1.54\,,\,1.54] \times [-1.54\,,\,1.54]$ m/s; the mean number of new PTs is $\mu^{(j)}_{\mathrm{n}} = \mu_{\mathrm{n}} = 0.1$.

All agent and PT state pdfs are described by sets of $N_{\mathrm{P}} = 1000$ particles. Furthermore, the pruning threshold is set to $P_{\mathrm{pr}} = 0.01$. Table 3 summarizes the main parameters used for the performance evaluation in the simulated scenario.

2) DISCUSSION

In this settings, we compare the performance of the proposed algorithm that jointly performs cooperative self-localization and multitarget tracking, with respect to the case in which the two tasks are performed independently. To ease the notation and for the reader's convenience, we refer to the proposed method as joint localization and tracking (JLT), and to the alternative approach as separate localization and tracking (SLT). The difference between SLT and JLT is that the former estimates the agent states at time t only once by running the cooperative self-localization; then, the multitarget tracking task is performed considering the estimated agent states as *true* states. This is equivalent to running the classical SPA-based multitarget tracking algorithm described in [49] in which the

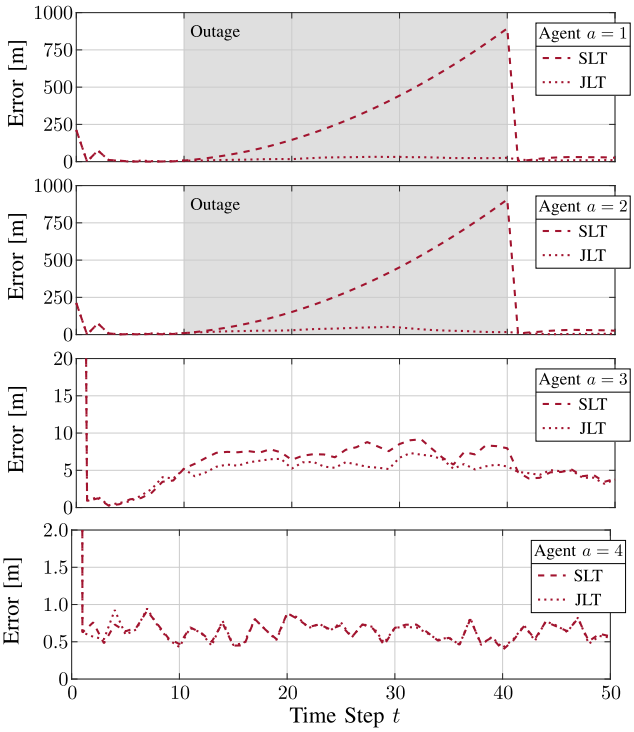


FIGURE 8. Performance comparison between JLT and SLT in terms of agents localization error over time.

states (i.e., positions) of the sensors are assumed known. The proposed JLT algorithm, instead, repeatedly estimates the agent states after the MOT measurements from each agent pair j are processed, as described in Section IV-B and shown in Fig. 3. On the figures, JLT-related quantities are reported with dotted lines, while SLT-related quantities with dashed lines.

It is important to emphasize that in the JLT approach targets are not only unknown objects to be localized, but they represent a valuable information correspondingly used by the agents to refine their own localization through the proposed cooperative mechanism. To highlight this, in the simulation we induce an outage condition to agents $a = 1$, and 2 in the time interval $t \in [10, 40]$, meaning that they do not have availability of inter-agent measurements from the Tx-agent $a = 4$. In practical use cases, outage can be related to perturbations on the communication channel due to hardware malfunctioning, interference, or hacking. Results are averaged over 250 Monte Carlo iterations.

In Fig. 8 we provide the position error over time for each agent individually, for both JLT and SLT. The most interesting result is related to the agents in outage condition, i.e., $a = 1$, and 2: we observe that the JLT is still able to localize them by exploiting the target information. The SLT, instead, can only rely on motion prediction, that leads to high position errors. Considering practical implementation, the use of JLT might allow, for instance, the recovery of the agent, which would be hardly achieved in case of SLT. As a second comment, we note that JLT outperforms SLT in the localization of agent $a = 3$, while there are no benefits for the localization of agent $a = 4$, since it is anchored and navigation data are extremely precise

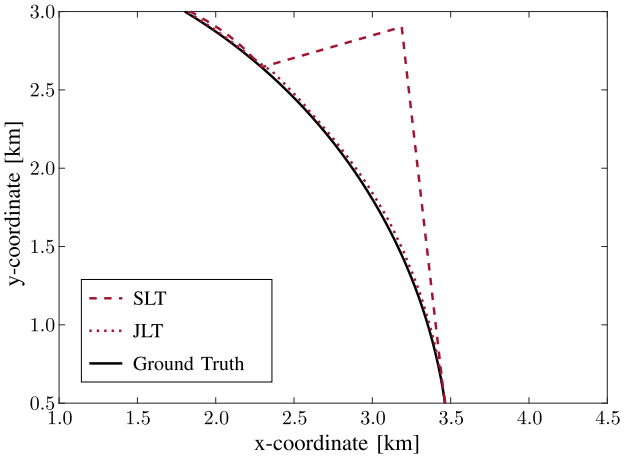


FIGURE 9. **Zoomed reconstructed trajectory of agent $a = 1$. The black solid line is the ground truth trajectory, the red dashed line indicates the SLT estimated trajectory, and the red dotted line refers to the JLT estimated trajectory.**

by default. To show the effect of relying on motion prediction only (SLT) with respect to a profitable use of target implicit information (JLT), in Fig. 9 we zoom in on the reconstructed trajectory of agent $a = 1$, highlighting the huge difference between the two methodologies. Note that the effect that the target location information supports the localization of the agents has been previously demonstrated in [53]. However, as mentioned in Section I-A, the algorithm proposed in [53] is limited by the fact that the number of targets that can be tracked is time-invariant and has to be known in advance, and it assumes a perfect knowledge of the association between targets and measurements.

After the analysis on agent localization, we now focus on the capability of the proposed technique to perform multitarget tracking. Results are given in terms of mean optimal subpattern assignment (MOSPA) error [65] of order 1 and cut off parameter of 5000 m in Fig. 10(a), and in terms of estimated mean number of detected targets in Fig. 10(b); the MOSPA error accounts for localization errors for correctly detected targets, and errors for missed targets and false targets. Results in Fig. 10(a) indicate superior tracking capabilities of the JLT over the SLT, in particular after the outage that starts at $t = 10$ (300 s). Indeed, the outage does not allow the SLT to accurately estimate the positions of agents $a = 1$, and 2; this directly affects the multitarget tracking task, since the MOT measurements they produce cannot be effectively used. After $t = 30$ (900 s), the difference in MOSPA error between JLT and SLT increases, because of the appearance of the last target that is not promptly detected by the SLT. The tardiness in target detection is also observable in Fig. 10(b), where we report the mean number of detected PTs over time.

Results demonstrate how target information is of high importance for practical application, e.g., in maritime surveillance. We proved the advantages of considering a joint framework for cooperative self-localization and multitarget tracking rather than perform the two tasks independently.

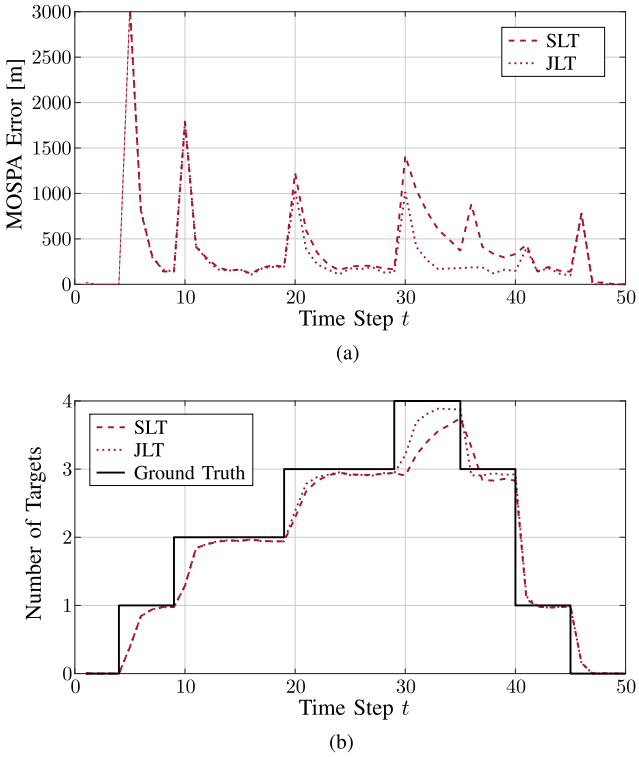


FIGURE 10. **Performance comparison between the multitarget tracking capabilities of JLT and SLT in the simulated scenario in terms of (a) MOSPA error and (b) estimated mean number of detected targets over time.**

B. APPLICATION TO REAL MARITIME DATA

This section presents the performance of JLT and SLT assessed in a real maritime application. We consider a hybrid, autonomous, robotic network developed by NATO Centre for Maritime Research and Experimentation (CMRE) for surveillance applications. The network consists of mobile and fixed gateways that form the communication infrastructure, and of autonomous underwater vehicles (AUVs) capable of detecting and tracking possible threats, and communicating the acquired data to the command and control center [14], [66]. The data we use was gathered during the littoral continuous active sonar trial conducted off the coast of Piombino, near Livorno, Italy, in November 2018 (LCAS18) [67]. During this trial the network consisted of $A = 6$ agents, classified as follows:

- surface agents: a towed sonar source ($a = 1$), two stationary and co-located acoustic modems ($a = 2$, and 3), and a nearly-stationary waveglider ($a = 4$);
- underwater agents: two AUVs ($a = 5$, and 6), named as Groucho and Harpo, each towing a uniform linear array of microphones and equipped with an acoustic modem for communications.

The classification of Rx-agents and Tx-agents is chosen as follows: the set of Tx-agents is constituted by all surface and underwater agents, i.e., $\mathcal{T} = \{1, 2, \ldots, 6\}$, while the set of Rx-agents comprises all agents but the towed sonar source, i.e., $\mathcal{R} = \{2, 3, \ldots, 6\}$. Practically, inter-agent measurements are available between all the Tx-agents and Rx-agents, while MOT measurements are only produced by the AUVs, i.e.,

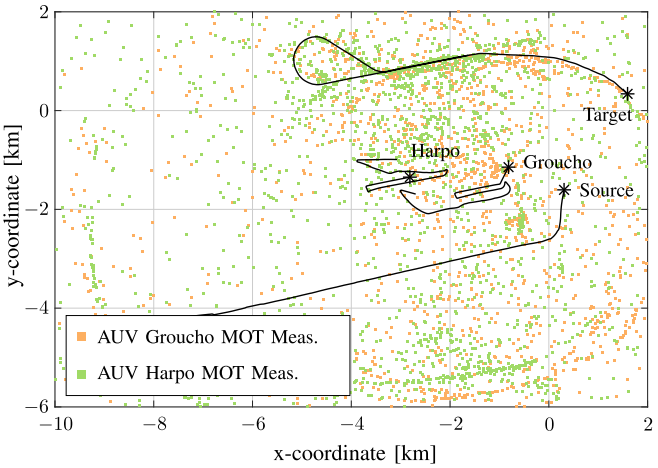


FIGURE 11. MOT measurements (converted in x-y coordinates) produced by AUVs Groucho and Harpo, respectively in orange and green, during the LCAS18 campaign. The black solid lines are the ground truth trajectories of the sonar source, the AUVs Groucho and Harpo, and the target; the asterisks mark the initial positions of these trajectories.

$a = 5$, and 6, using the signal transmitted by the towed sonar source, i.e., $a = 1$; this means that the agent pairs involved in the sequential processing of the MOT measurements are $(j_1, j_2) = (5, 1)$ and $(j_1, j_2) = (6, 1)$, from which it then follows that $J = \{1, 2\}$. Fig. 11 shows all the MOT measurements produced by Groucho and Harpo during the trial, respectively in orange and green, and the ground-truth trajectories of the sonar source, the NATO research vessel Leonardo acting as target, and the AUVs. The ground-truth position of the former two is provided by their on-board GNSS receivers, whereas the ground-truth of Groucho and Harpo is provided by their INS. We note that in this real application, the observations (navigation data, inter-agent and MOT measurements) are not guaranteed to be acquired at the same time, nor they are available at all times because of the challenging propagation conditions posed by the underwater environment. Nevertheless, the flexibility of the proposed algorithm allows to perform each single task (e.g., the agents' cooperative self-localization or the processing of the MOT measurements at a specific agent pair) only when the relevant observations are available.

Navigation data are available for all the surface agents by means of on-board GNSS receivers; the standard deviation of the position information is set to 5 m, and the standard deviation of the velocity information (used by the towed sonar source only) is set to 0.1 m/s. Range and bearing characterize both inter-agent and MOT measurements. Standard deviations of range and bearing information are set to 70 m and 7 deg, respectively. Additionally, inter-agent and MOT measurements are duplicated because of the port-starboard ambiguity, i.e., the inability of the AUVs to discriminate if a signal comes from the port side or from the starboard side due to the intrinsic cylindrical symmetry of the towed array [15]. The other parameters are as reported in Table 3, with the exception of the mean number of false alarms set to $\mu_c = 9$, and the number of particles set to $N_P = 250$.

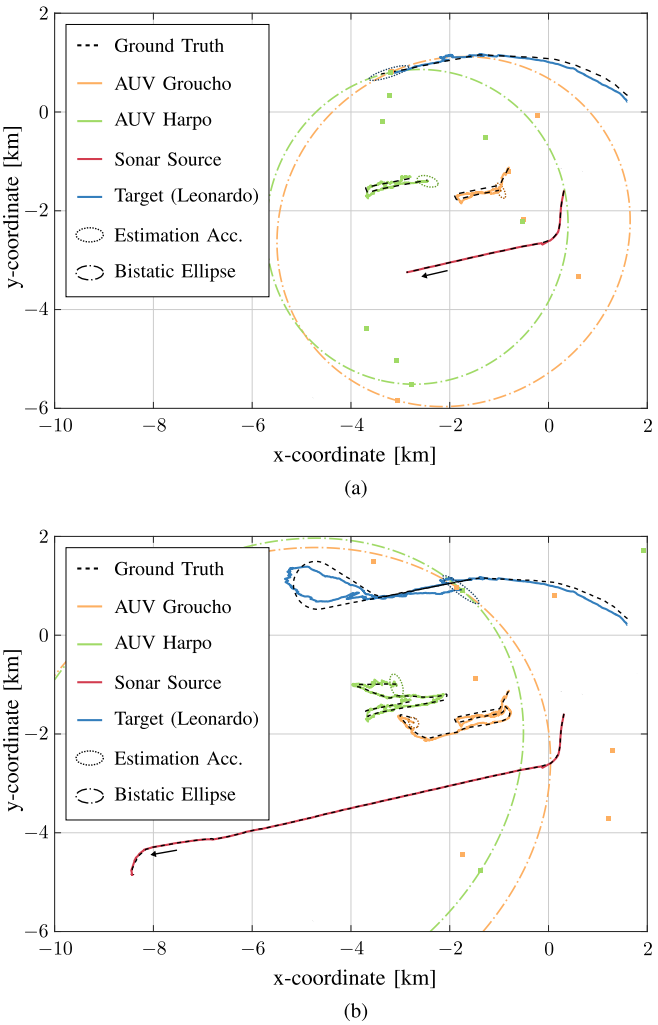


FIGURE 12. Behavior of the JLT using the LCAS18 data set. The figures are snapshots of the output produced by JLT after (a) 37 minutes and (b) 87 minutes since the beginning of the trial, and include MOT measurements, estimated trajectories, and estimated current positions of agents and target. Solid lines represent the estimated trajectory up to current time. Dotted ellipses indicate the estimate and accuracy of agents and target current location (note that the ellipse associated to the sonar source is hardly visible, as its estimated position is very accurate due to the availability of navigation data; the black arrow indicates the direction of movement). Dotted-dashed ellipses are examples of bistatic ellipses associated to Groucho's and Harpo's MOT measurements closest to the target.

Figs. 12(a) and 12(b) are snapshots of the output produced by JLT after 37 and 87 minutes, respectively, since the beginning of the trial. The images show the estimated trajectories of the non-stationary agents and the target, and estimate and accuracy (i.e., 95% confidence interval) of their current positions. Examples of bistatic ellipses associated to Groucho's and Harpo's MOT measurements closest to the target are also provided. A bistatic ellipse is the locus of points in which the sum of the distances from the sonar source and the AUV (i.e., the foci) is constant and equal to the bistatic range component of the AUV's MOT measurement (cf. (42)) [68]. Because of the port-starboard ambiguity two actual contacts are visible on each bistatic ellipse, one close to the target and the other along the specular direction. These snapshots show that it is possible

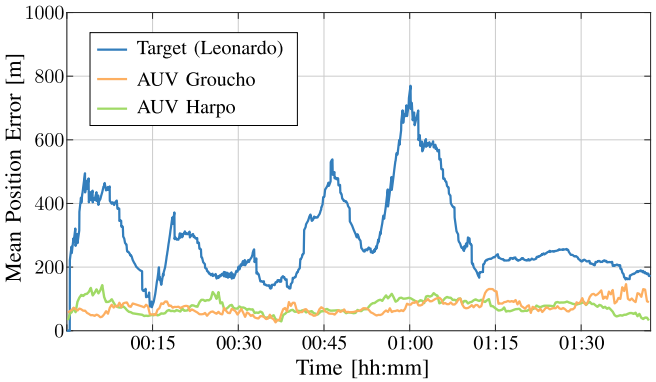


FIGURE 13. Real data experiment. Mean error on the JLT estimated positions of the target (in blue), AUV Groucho (in orange) and AUV Harpo (in green).

TABLE IV Real Data Experiment. Performance Comparison Between JLT and SLT

MEAN POSITION ERROR		JLT	SLT
Target (Leonardo)	Time-averaged	285.4 m	267.7 m
	Max	770.0 m	842.0 m
AUV Groucho	Time-averaged	75.9 m	117.2 m
	Max	143.4 m	189.8 m
AUV Harpo	Time-averaged	74.2 m	79.6 m
	Max	146.6 m	165.9 m
TIME-AVERAGED MOSPA ERROR		348.2 m	405.1 m

to joint localize agents (towed sonar source and AUVs) and target through a fusion of navigation data, inter-agent and MOT measurements. Fig. 13 shows the position errors over time averaged over 20 Monte Carlo iterations obtained with JLT for the target,[6] Groucho, and Harpo (the Monte Carlo iterations only differ for the drawing of the particles). The higher mean position error for the target is clearly due to the fact that only MOT measurements are used to estimate its state; moreover, the peaks occur when either the target itself is maneuvering and/or one or both the AUVs are maneuvering. Table 4 reports a comparison between JLT and SLT in terms of time-averaged and maximum mean position errors, and time-averaged MOSPA error with order 1 and cutoff parameter of 1000 m. JLT outperforms SLT in terms of time-averaged and maximum mean position errors for both the AUVs. These results confirm the benefit of exploiting the target location information for estimating the agents' states. Regarding the tracking of the target, JLT provides a slightly higher time-averaged mean position error, but a lower maximum mean position error and a lower time-averaged MOSPA error, that accounts for both missed targets and false targets. Lastly, we compare in Fig. 14 the JLT estimated velocities of the AUVs — along the two Cartesian coordinates — and the

[6]For each Monte Carlo iteration, the position error for the target is computed as the distance between its ground-truth trajectory and the PT that is consistently closest to it over time.

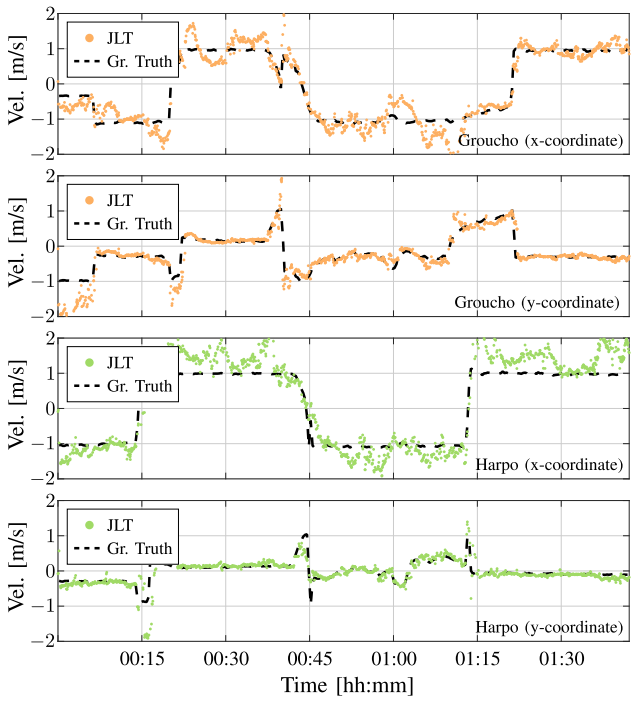


FIGURE 14. Real data experiment. Comparison between the Cartesian components of the JLT estimated velocities of Groucho (top figures) and Harpo (bottom figures), and their ground truth velocities provided by the on-board INS.

velocities provided by the INS: results show a good flexibility to abrupt heading variations and quite accurate velocity estimation overall.

The assessment on real data proves the ability of the proposed SPA-based algorithm to jointly perform cooperative self-localization and multitarget tracking. The results provides evidence of how to profitably take advantage of intrinsic target information to perform agents localization.

VI. CONCLUSION

In this work, we developed a joint technique for cooperative self-localization and multitarget tracking in agent networks. The proposed algorithm is general enough to be tailored to any multi-agent system, where agents are equipped with diverse perception sensors and communication devices. The proposed method performs the mandatory tasks of self-determining the network topology (i.e., the agent network localization) and detecting and tracking an unknown and arbitrary number of targets, where existence probabilities are used to declare their actual presence or to opt for their removal (pruning). The developed technique takes advantage of target information to update and refine the agent positions, assigning an opportunistic role to targets. This latter benefit might not be so relevant in case of a large availability of navigation data and/or inter-agent measurements, but it has been proven through simulations to be of utmost importance in case of malfunctioning or outage conditions. An important aspect of the proposed algorithm is the flexibility: the algorithm intrinsically handles time-variant properties of agents and network topology (such

as a connection of a new agent, or the disappearance of an extant one) and it also admits the coexistence of different types of observations (navigation data, inter-agent and MOT measurements). Lastly, the extension of the data association problem to include agents as well (and not only targets) allows us to consider realistic conditions of signal propagation, where reflections from both agents and targets are unavoidably present and affect the signal processing chain.

The joint cooperative self-localization and multitarget tracking method proposed in this article assumes a centralized architecture of information exchange in a complex network made by several agents. Future work will include the study of distributed/decentralized architectures where the exchange of local target/agent states among the agents, rather than of observations with a centralized node, is more convenient, e.g., in terms of robustness. Promising paradigms for distributed architectures are the consensus networks [52], [56], [60], [69] and the adaptive networks [70], [71].

APPENDIX

Here we derive the factorization in (17) of the joint posterior pdf $f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t})$. Since the MOT measurements $\boldsymbol{z}_{1:t}$ are observed, hence known, the joint vector of numbers of MOT measurements $\boldsymbol{m}_{1:t}$ is also known, that is $f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}) = f(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} | \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}, \boldsymbol{m}_{1:t})$. Then, we obtain the factorization in (46), shown at the bottom of this page, by using assumption (A6) in the third step. Recalling from Section II-C that $\boldsymbol{y}_t$ is the vector stacking the legacy PT augmented states at the first agent pair and all the new PT augmented states introduced at time t, that is, $\boldsymbol{y}_t = [\underline{\boldsymbol{y}}_t^{(1)\mathrm{T}}, \bar{\boldsymbol{y}}_t^{\mathrm{T}}]^{\mathrm{T}}$, each factor $f(\boldsymbol{y}_t, \boldsymbol{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t | \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1})$ of the product in (46) can be further expressed as

$$
\begin{aligned}
& f\left(\boldsymbol{y}_t, \boldsymbol{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1}\right) \\
&\quad = f\left(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t, \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1}\right) \\
&\qquad \times f\left(\underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t \middle| \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1}\right) \\
&\quad = f\left(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t\right) \\
&\qquad \times f\left(\underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t \middle| \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1}\right) \qquad (43) \\
&\quad = f(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t | \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t) \\
&\qquad \times f(\underline{\boldsymbol{y}}_t^{(1)} | \boldsymbol{y}_{t-1}) f(\boldsymbol{s}_t | \boldsymbol{s}_{t-1}) \qquad (44) \\
&\quad = f(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{z}_t, \boldsymbol{m}_t | \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t) f(\boldsymbol{g}_t | \boldsymbol{s}_t) f(\boldsymbol{\rho}_t | \boldsymbol{s}_t) \\
&\qquad \times f(\underline{\boldsymbol{y}}_t^{(1)} | \boldsymbol{y}_{t-1}) f(\boldsymbol{s}_t | \boldsymbol{s}_{t-1}), \qquad (45)
\end{aligned}
$$

$$
\begin{aligned}
& f\left(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t} \middle| \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}, \boldsymbol{m}_{1:t}\right) \\
&\quad \propto f\left(\boldsymbol{y}_{1:t}, \boldsymbol{s}_{0:t}, \boldsymbol{\alpha}_{1:t}, \boldsymbol{\beta}_{1:t}, \boldsymbol{g}_{1:t}, \boldsymbol{\rho}_{1:t}, \boldsymbol{z}_{1:t}, \boldsymbol{m}_{1:t}\right) \\
&\quad = f\left(\boldsymbol{y}_t, \boldsymbol{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \boldsymbol{y}_{1:t-1}, \boldsymbol{s}_{0:t-1}, \boldsymbol{\alpha}_{1:t-1}, \boldsymbol{\beta}_{1:t-1}, \boldsymbol{g}_{1:t-1}, \boldsymbol{\rho}_{1:t-1}, \boldsymbol{z}_{1:t-1}, \boldsymbol{m}_{1:t-1}\right) \\
&\qquad \times f\left(\boldsymbol{y}_{1:t-1}, \boldsymbol{s}_{0:t-1}, \boldsymbol{\alpha}_{1:t-1}, \boldsymbol{\beta}_{1:t-1}, \boldsymbol{g}_{1:t-1}, \boldsymbol{\rho}_{1:t-1}, \boldsymbol{z}_{1:t-1}, \boldsymbol{m}_{1:t-1}\right) \\
&\quad = f\left(\boldsymbol{y}_t, \boldsymbol{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{g}_t, \boldsymbol{\rho}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \boldsymbol{y}_{t-1}, \boldsymbol{s}_{t-1}\right) f\left(\boldsymbol{y}_{1:t-1}, \boldsymbol{s}_{0:t-1}, \boldsymbol{\alpha}_{1:t-1}, \boldsymbol{\beta}_{1:t-1}, \boldsymbol{g}_{1:t-1}, \boldsymbol{\rho}_{1:t-1}, \boldsymbol{z}_{1:t-1}, \boldsymbol{m}_{1:t-1}\right) \\
&\quad = f\left(\boldsymbol{s}_0\right) \prod_{t'=1}^{t} f\left(\boldsymbol{y}_{t'}, \boldsymbol{s}_{t'}, \boldsymbol{\alpha}_{t'}, \boldsymbol{\beta}_{t'}, \boldsymbol{g}_{t'}, \boldsymbol{\rho}_{t'}, \boldsymbol{z}_{t'}, \boldsymbol{m}_{t'} \middle| \boldsymbol{y}_{t'-1}, \boldsymbol{s}_{t'-1}\right) \qquad (46)
\end{aligned}
$$

$$
\begin{aligned}
& f\left(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{z}_t, \boldsymbol{m}_t \middle| \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t\right) \\
&= f\left(\bar{\boldsymbol{y}}_t^{(J)}, \boldsymbol{\alpha}_t^{(J)}, \boldsymbol{\beta}_t^{(J)}, \boldsymbol{z}_t^{(J)}, \boldsymbol{M}_t^{(J)} \middle| \underline{\boldsymbol{y}}_t^{(1)}, \bar{\boldsymbol{y}}_t^{(1)}, \ldots, \bar{\boldsymbol{y}}_t^{(J-1)}, \boldsymbol{\alpha}_t^{(1)}, \ldots, \boldsymbol{\alpha}_t^{(J-1)}, \boldsymbol{\beta}_t^{(1)}, \ldots, \boldsymbol{\beta}_t^{(J-1)}, \boldsymbol{z}_t^{(1)}, \ldots, \boldsymbol{z}_t^{(J-1)}, \boldsymbol{M}_t^{(1)}, \ldots, \boldsymbol{M}_t^{(J-1)}, \boldsymbol{s}_t\right) \\
&\quad \times f\left(\bar{\boldsymbol{y}}_t^{(1)}, \ldots, \bar{\boldsymbol{y}}_t^{(J-1)}, \boldsymbol{\alpha}_t^{(1)}, \ldots, \boldsymbol{\alpha}_t^{(J-1)}, \boldsymbol{\beta}_t^{(1)}, \ldots, \boldsymbol{\beta}_t^{(J-1)}, \boldsymbol{z}_t^{(1)}, \ldots, \boldsymbol{z}_t^{(J-1)}, \boldsymbol{M}_t^{(1)}, \ldots, \boldsymbol{M}_t^{(J-1)}\right) \\
&= f\left(\bar{\boldsymbol{y}}_t^{(J)}, \boldsymbol{\alpha}_t^{(J)}, \boldsymbol{\beta}_t^{(J)}, \boldsymbol{z}_t^{(J)}, \boldsymbol{M}_t^{(J)} \middle| \underline{\boldsymbol{y}}_t^{(J)}, \boldsymbol{\alpha}_t^{(1)}, \ldots, \boldsymbol{\alpha}_t^{(J-1)}, \boldsymbol{\beta}_t^{(1)}, \ldots, \boldsymbol{\beta}_t^{(J-1)}, \boldsymbol{z}_t^{(1)}, \ldots, \boldsymbol{z}_t^{(J-1)}, \boldsymbol{M}_t^{(1)}, \ldots, \boldsymbol{M}_t^{(J-1)}, \boldsymbol{s}_t\right) \\
&\quad \times f\left(\bar{\boldsymbol{y}}_t^{(1)}, \ldots, \bar{\boldsymbol{y}}_t^{(J-1)}, \boldsymbol{\alpha}_t^{(1)}, \ldots, \boldsymbol{\alpha}_t^{(J-1)}, \boldsymbol{\beta}_t^{(1)}, \ldots, \boldsymbol{\beta}_t^{(J-1)}, \boldsymbol{z}_t^{(1)}, \ldots, \boldsymbol{z}_t^{(J-1)}, \boldsymbol{M}_t^{(1)}, \ldots, \boldsymbol{M}_t^{(J-1)}\right) \\
&= f\left(\bar{\boldsymbol{y}}_t^{(J)}, \boldsymbol{\alpha}_t^{(J)}, \boldsymbol{\beta}_t^{(J)}, \boldsymbol{z}_t^{(J)}, \boldsymbol{M}_t^{(J)} \middle| \underline{\boldsymbol{y}}_t^{(J)}, \boldsymbol{s}_t\right) \\
&\quad \times f\left(\bar{\boldsymbol{y}}_t^{(1)}, \ldots, \bar{\boldsymbol{y}}_t^{(J-1)}, \boldsymbol{\alpha}_t^{(1)}, \ldots, \boldsymbol{\alpha}_t^{(J-1)}, \boldsymbol{\beta}_t^{(1)}, \ldots, \boldsymbol{\beta}_t^{(J-1)}, \boldsymbol{z}_t^{(1)}, \ldots, \boldsymbol{z}_t^{(J-1)}, \boldsymbol{M}_t^{(1)}, \ldots, \boldsymbol{M}_t^{(J-1)}\right) \\
&= \prod_{j=1}^{J} f\left(\bar{\boldsymbol{y}}_t^{(j)}, \boldsymbol{\alpha}_t^{(j)}, \boldsymbol{\beta}_t^{(j)}, \boldsymbol{z}_t^{(j)}, \boldsymbol{M}_t^{(j)} \middle| \underline{\boldsymbol{y}}_t^{(j)}, \boldsymbol{s}_t\right). \qquad (47)
\end{aligned}
$$

where assumption (A7) is used in (43), (A3) in (44), and (A9) in (45). Finally, by using assumption (A8) and the fact that new PTs at agent pairs $1, \ldots, j-1$ become legacy PTs at agent pairs $j, \ldots, J$, the joint pdf $f(\bar{\boldsymbol{y}}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, z_t, \boldsymbol{m}_t | \underline{\boldsymbol{y}}_t^{(1)}, \boldsymbol{s}_t)$ in (45) can be factorized as in (47). Eventually, by inserting (47), shown at the bottom of previous page, into (45), and (45) into (46), we obtain the factorization in (17).

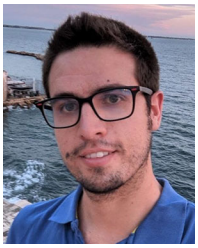
ACKNOWLEDGMENT

The collection of LCAS18 sea trial data used in this paper was made possible by the LCAS Multi National - Joint Research Project, with participants CMRE (NATO), DSTG (AUS), DRDC (CAN), DSTL (GBR), CSSN (ITA), FFI (NOR), DTA (NZL), and ONR (USA).

REFERENCES

[1] M. Z. Win *et al.*, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.
[2] D. Gaglione *et al.*, "Bayesian information fusion and multitarget tracking for maritime situational awareness," *IET Radar, Sonar Navigation*, vol. 14, no. 12, pp. 1845–1857, Dec. 2020.
[3] P. Braca, R. Goldhahn, G. Ferri, and K. D. LePage, "Distributed information fusion in multistatic sensor networks for underwater surveillance," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4003–4014, Jun. 2016.
[4] 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.
[5] 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.
[6] G. Soldi *et al.*, "Space-based global maritime surveillance. Part I: Satellite technologies," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 9, pp. 8–28, Sep. 2021.
[7] G. Soldi *et al.*, "Space-based global maritime surveillance. Part II: Artificial intelligence and data fusion techniques," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 9, pp. 30–42, Sep. 2021.
[8] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *Commun. Surv. Tut.*, vol. 21, no. 3, pp. 2568–2599, Apr. 2019.
[9] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
[10] S. Bartoletti, A. Conti, A. Giorgetti, and M. Z. Win, "Sensor radar networks for indoor tracking," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 157–160, Apr. 2014.
[11] Y. Zheng, N. Cao, T. Wimalajeewa, and P. K. Varshney, "Compressive sensing based probabilistic sensor management for target tracking in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6049–6060, Nov. 2015.
[12] N. Cao, S. Choi, E. Masazade, and P. K. Varshney, "Sensor selection for target tracking in wireless sensor networks with uncertainty," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5191–5204, Oct. 2016.
[13] G. Soatti, M. Nicoli, N. Garcia, B. Denis, R. Raulefs, and H. Wymeersch, "Implicit cooperative positioning in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 12, pp. 3964–3980, Dec. 2018.
[14] G. Ferri *et al.*, "Cooperative robotic networks for underwater surveillance: An overview," *IET Radar, Sonar Navigation*, vol. 11, no. 12, pp. 1740–1761, Dec. 2017.
[15] P. Braca, P. Willett, K. LePage, S. Marano, and V. Matta, "Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1864–1878, Apr. 2014.
[16] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Netw.*, vol. 3, no. 3, pp. 257–279, Mar. 2005.
[17] Y. Li, Y. Wang, W. Yu, and X. Guan, "Multiple autonomous underwater vehicle cooperative localization in anchor-free environments," *IEEE J. Ocean. Eng.*, vol. 44, no. 4, pp. 895–911, Oct. 2019.
[18] M. Brambilla, M. Nicoli, G. Soatti, and F. Deflorio, "Augmenting vehicle localization by cooperative sensing of the driving environment: Insight on data association in urban traffic scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1646–1663, Apr. 2020.
[19] S. Zhang *et al.*, "Distributed direct localization suitable for dense networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1209–1227, Apr. 2020.
[20] S. Sridhar and A. Eskandarian, "Cooperative perception in autonomous ground vehicles using a mobile-robot testbed," *IET Intell. Transp. Syst.*, vol. 13, no. 10, pp. 1545–1556, Sep. 2019.
[21] Y. Ma, C. Tian, and Y. Jiang, "A multitag cooperative localization algorithm based on weighted multidimensional scaling for passive UHF RFID," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6548–6555, Mar. 2019.
[22] 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.
[23] 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.
[24] A. A. Saucan and M. Z. Win, "Information-seeking sensor selection for ocean-of-things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10072–10088, Oct. 2020.
[25] P. Braca, R. Goldhahn, K. D. LePage, S. Marano, V. Matta, and P. Willett, "Cognitive multistatic AUV networks," in *Proc. 17th Int. Conf. Inf. Fusion*, Salamanca, Spain, 2014, pp. 1–7.
[26] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
[27] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proc. IEEE*, vol. 95, no. 6, pp. 1295–1322, Jun. 2007.
[28] A. T. Ihler, J. W. Fisher, R. L. Moses, and A. S. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 809–819, Apr. 2005.
[29] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.
[30] A. Conti, M. Guerra, D. Dardari, N. Decarli, and M. Z. Win, "Network experimentation for cooperative localization," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 467–475, Feb. 2012.
[31] S. Li, M. Hedley, and I. B. Collings, "New efficient indoor cooperative localization algorithm with empirical ranging error model," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1407–1417, Jul. 2015.
[32] 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.
[33] R. Mendrzik and G. Bauch, "Position-constrained stochastic inference for cooperative indoor localization," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 454–468, Sep. 2019.
[34] B. Teague, Z. Liu, F. Meyer, A. Conti, and M. Win, "Network localization and navigation with scalable inference and efficient operation," *IEEE Trans. Mobile Comput.*, early access, Nov. 30, 2020, doi: 10.1109/TMC.2020.3035511.
[35] G. Soldi, F. Meyer, P. Braca, and F. Hlawatsch, "Self-tuning algorithms for multisensor-multitarget tracking using belief propagation," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3922–3937, Aug. 2019.
[36] R. W. Sittler, "An optimal data association problem in surveillance theory," *IEEE Trans. Mil. Electron.*, vol. ME-8, no. 2, pp. 125–139, Apr. 1964.
[37] R. Singer, R. Sea, and K. Housewright, "Derivation and evaluation of improved tracking filter for use in dense multitarget environments," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 423–432, Jul. 1974.
[38] Y. Bar-Shalom, "Tracking methods in a multitarget environment," *IEEE Trans. Autom. Control*, vol. AC-23, no. 4, pp. 618–626, Aug. 1978.
[39] 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.
[40] M. Jiang, W. Yi, R. Hoseinnezhad, and L. Kong, "Distributed multi-sensor fusion using generalized multi-Bernoulli densities," in *Proc. 19th Int. Conf. Inf. Fusion*, Heidelberg, Germany, 2016, pp. 1332–1339.
[41] 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.

[42] L. Gao, G. Battistelli, and L. Chisci, "Event-triggered distributed multitarget tracking," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 570–584, Sep. 2019.
[43] Y. Yu and Y. Liang, "Distributed multitarget tracking based on diffusion strategies over sensor networks," *IEEE Access*, vol. 7, pp. 129802–129814, 2019.
[44] X. Li, C. Zhao, X. Lu, and W. Wei, "Underwater bearings-only multitarget tracking based on modified PMHT in dense-cluttered environment," *IEEE Access*, vol. 7, pp. 93678–93689, 2019.
[45] M. Fröhle, C. Lindberg, K. Granström, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint target-sensor state tracking," *IEEE Trans. Intell. Veh.*, vol. 4, no. 4, pp. 609–621, Aug. 2019.
[46] M. Baradaran Khalkhali, A. Vahedian, and H. Sadoghi Yazdi, "Multitarget state estimation using interactive Kalman filter for multivehicle tracking," *IEEE Intell. Transp. Syst. Mag.*, vol. 21, no. 3, pp. 1131–1144, Mar. 2020.
[47] S. He, H. Shin, and A. Tsourdos, "Distributed joint probabilistic data association filter with hybrid fusion strategy," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 1, pp. 286–300, Jan. 2020.
[48] A. K. Gostar *et al.*, "Centralized cooperative sensor fusion for dynamic sensor network with limited field-of-view via labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 69, pp. 878–891, Dec. 2020.
[49] F. Meyer *et al.*, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.
[50] V. Savic, H. Wymeersch, and E. G. Larsson, "Target tracking in confined environments with uncertain sensor positions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 870–882, Feb. 2016.
[51] F. Meyer and M. Z. Win, "Joint navigation and multitarget tracking in networks," in *Proc. IEEE Int. Conf. Commun. Workshops*, Kansas City, MO, USA, 2018, pp. 1–6.
[52] P. Sharma, 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.
[53] F. Meyer, O. Hlinka, H. Wymeersch, E. Riegler, and F. Hlawatsch, "Distributed localization and tracking of mobile networks including noncooperative objects," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 1, pp. 57–71, Mar. 2016.
[54] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT, USA: YBS Publishing, 2011.
[55] R. Mendrzik *et al.*, "Joint multitarget tracking and dynamic network localization in the underwater domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Barcelona, Spain, 2020, pp. 4890–4894.
[56] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Mar. 2007.
[57] S. Das and J. M. F. Moura, "Distributed Kalman filtering with dynamic observations consensus," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4458–4473, Sep. 2015.
[58] S. Das and J. M. F. Moura, "Consensus innovations distributed Kalman filter with optimized gains," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 467–481, Jan. 2017.
[59] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 508–520, Mar. 2013.
[60] G. Papa, R. Repp, F. Meyer, P. Braca, and F. Hlawatsch, "Distributed Bernoulli filtering using likelihood consensus," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 2, pp. 218–233, Jun. 2019.
[61] G. Mao, B. Fidan, and B. D. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, Jul. 2007.
[62] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X technologies toward the internet of vehicles: Challenges and opportunities," *Proc. IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020.
[63] F. Meyer *et al.*, "Message passing algorithms for scalable multitarget tracking - Supplementary material," 2019. [Online]. Available: https://winslab.lids.mit.edu/procieee_mtt_suppl_mat-pdf/
[64] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. New York, NY, USA: Wiley, 2001.
[65] D. Schuhmacher, B. Vo, and B. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.
[66] G. Ferri, A. Munafó, and K. D. LePage, "An autonomous underwater vehicle data-driven control strategy for target tracking," *IEEE J. Ocean. Eng.*, vol. 43, no. 2, pp. 323–343, Feb. 2018.
[67] G. Ferri *et al.*, "Cooperative autonomy in the CMRE ASW multistatic robotic network: Results from LCAS18 trial," in *Proc. OCEANS*, Marseille, France, 2019, pp. 1–10.
[68] N. Willis, *Bistatic Radar*, Raleigh, NC, USA: SciTech Publishing, 2005.
[69] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic optimality of running consensus in testing binary hypotheses," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 814–825, Feb. 2010.
[70] V. Matta, P. Braca, S. Marano, and A. H. Sayed, "Diffusion-based adaptive distributed detection: Steady-state performance in the slow adaptation regime," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4710–4732, Aug. 2016.
[71] A. H. Sayed, "Adaptive networks," *Proc. IEEE*, vol. 102, no. 4, pp. 460–497, Apr. 2014.



MATTIA BRAMBILLA (Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunication engineering and the Ph.D. degree (*cum laude*) in information technology from the Politecnico di Milano, Milan, Italy, in 2015, 2017, and 2021, respectively. He was a Visiting Researcher with the NATO Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy, in 2019. In 2021, he joined the Faculty of Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, as a Research Fellow. His research interests include signal processing, statistical learning, and data fusion for cooperative localization and communication. He was the recipient of the Best Student Paper Award at the 2018 IEEE Statistical Signal Processing Workshop.

DOMENICO GAGLIONE received the B.Sc. and M.Sc. degrees (*s*umma cum laude) in telecommunications engineering from the Università degli Studi di Napoli Federico II, Naples, Italy, in 2011 and 2013, respectively, and the Ph.D. degree from the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K., in 2017. He is currently a Research Scientist with NATO STO Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy. His research interests include statistical signal processing with emphasis on state estimation, data fusion, and multisensor multitarget tracking. He was the recipient of the Best Student Paper Award at the 2015 IEEE International Radar Conference (RadarCon), Arlington, VA, USA, and of the NATO STO Scientific Achievement Award in 2020, as a Member of the Data Knowledge and Operational Effectiveness (DKOE) Research Group at CMRE for Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness.

GIOVANNI SOLDI received the master's degree in applied mathematics from the University of Milan, Milan, Italy, in 2011, and the Ph.D. degree in signal processing from Télécom ParisTech, Paris, France, in 2016. Since 2016, he has been a Scientist with the Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy. His research interests include statistical signal processing with a focus on state estimation, data fusion, and multisensor-multitarget tracking techniques. He was the recipient of the NATO STO Scientific Achievement Award, in 2020, as a Member of the Data Knowledge and Operational Effectiveness (DKOE) Research Group, CMRE for Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness.

RICO MENDRZIK received the M.Sc. and Ph.D. degrees in electrical engineering from the Hamburg University of Technology, Hamburg, Germany, in 2015 and 2019, respectively. He is currently with Ibeo Automotive Systems GmbH where he is managing the Ground Truth Environment Perception and Software Interfaces Group. He visited the Cognitive Reconfigurable Embedded Systems (CORES) Lab, University of California, Los Angeles, CA, USA, as a Visiting Researcher in 2017. From 2018 to 2019, he was appointed a Visiting Researcher with Wireless Information and Network Sciences Laboratory (WINS Lab), Massachusetts Institute of Technology, Cambridge, MA, USA. He also visited the NATO Science and Technology Organization - Centre for Maritime Research and Experimentation (STO CMRE) in 2019. His main research interests include positioning, environment perception, estimation theory, and statistical signal processing.

GABRIELE FERRI (Senior Member, IEEE) received the Laurea degree (M.S.) in computer engineering (with Hons.) from the University of Pisa, Pisa, Italy, in 2003, and the Ph.D. degree in robotics jointly from the Scuola Superiore Sant'Anna, Pisa, Italy, and IMT Advanced Studies Lucca, Lucca, Italy, in 2008. From 2003 to 2005, he was a Software/System Engineer with WASS Company (a Finmeccanica Group Company), Livorno, Italy. In 2007, he was a Visiting Researcher with Woods Hole Oceanographic Institution, Woods Hole, MA, USA, working on robot autonomy for deep-water hydrothermal vents prospecting. After a period as a Postdoctoral Investigator with Scuola Superiore Sant'Anna, where he was the Project Leader of the HydroNet project, in 2012, he became a Research Scientist with CMRE, La Spezia, Italy, for investigating and testing at sea novel cooperative autonomy strategies for hybrid, heterogeneous, robotics networks in underwater surveillance applications. He was a Scientist in Charge in many experimental trials involving the deployment of complex multiple-assets systems in the field. He has been the Technical Director of the Student Autonomous Underwater Vehicle Challenge-Europe robotics competition since 2014, and was the Director of the European multidomain robotics competitions such as euRathlon 2015 Grand Challenge and the European Robotics League Emergency Robots 2017. He has been the euRobotics Marine Robotics Topic Group Coordinator since August 2020. His research interests include robotic systems for surveillance and exploration, cooperative and environmental-aware autonomy, and robot navigation.

KEVIN D. LEPAGE received the S.M. and Ph.D. degrees in ocean engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1988 and 1992, respectively. He is currently a Principal Scientist with the NATO Science and Technology Organization Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy. His work is directed towards capturing requirements for unmanned solutions for maritime capability and fielding and testing technical demonstrations of demonstrators at sea. While with CMRE, he has been a Scientist in Charge of ten scientific sea trails and NATO naval ASW exercises where unmanned systems have been deployed and tested. In 2021, Dr LePage was the recipient of the NATO STO Scientific Excellence Award for the Littoral Continuous Active Sonar Multi-National-Joint Research Project and in 2017 the STO Scientific Achievement Award for Development and Demonstration of Networked Autonomous ASW, both as Team Leader. He is a Fellow and the Past Chair of the Underwater Acoustics Technical Committee of the Acoustical Society of America.

MONICA NICOLI (Senior Member, IEEE) received the M.Sc. (Hons.) and Ph.D. degrees in communication engineering from the Politecnico di Milano, Milan, Italy, in 1998 and 2002, respectively. She was a Visiting Researcher with ENI Agip, from 1998 to 1999, and Uppsala University, Uppsala, Sweden, in 2001. In 2002, she joined the Politecnico di Milano as a Faculty Member. She is currently an Associate Professor in telecommunications with the Department of Management, Economics and Industrial Engineering. Her research interests include signal processing, machine learning, and wireless communications, with emphasis on smart mobility and Internet of Things (IoT) applications. She was the recipient of the Marisa Bellisario Award, in 1999, and a co-recipient of the best paper awards of the IEEE Symposium on Joint Communications and Sensing, in 2021, the IEEE Statistical Signal Processing Workshop, in 2018, and the *IET Intelligent Transport Systems Journal*, in 2014. She is an Associate Editor for IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. She was an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*, from 2010 to 2017, and a Lead Guest Editor for the Special Issue on Localization in Mobile Wireless and Sensor Networks, in 2011.

PETER WILLETT (Fellow, IEEE) received the B.A.Sc. degree in engineering science from the University of Toronto, Toronto, ON, Canada, in 1982, and the Ph.D. degree from Princeton University, Princeton, NJ, USA, in 1986. He has been a Faculty Member with the Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT, USA, since 1986. Since 1998, he has been a Professor, and since 2003 an IEEE Fellow. His primary research interests include statistical signal processing, detection, machine learning, communications, data fusion, and tracking.

PAOLO BRACA (Senior Member, IEEE) received the Laurea degree (*summa cum laude*) in electronic engineering and the Ph.D. (Hons.) degree in information engineering from the University of Salerno, Fisciano, Italy, in 2006 and 2010, respectively. In 2009, he was a Visiting Scholar with the ECE Department, University of Connecticut, Storrs, CT, USA. From 2010 to 2011, he was a Postdoctoral Associate with the University of Salerno. In 2011, he joined the NATO STO CMRE, where he is currently a Senior Scientist and a Project Manager. Furthermore, he led a number of research projects funded by the European Commission, by the U.S. Office of Naval Research (ONR), and other institutions. He has coauthored more than 200 publications in international scientific journals, conference proceedings, and NATO technical reports. His research interests include statistical signal processing with emphasis on detection and estimation theory, wireless sensor network, multi-agent algorithms, target tracking and data fusion, adaptation and learning over graphs, radar (sonar) signal processing, and machine learning. He was awarded with the National Scientific Qualification to function as Associate and a Full Professor with Italian Universities, in 2017 and 2018, respectively. He is in the technical committee of the major international conferences in the field of signal processing and data fusion. He was the recipient of the Best Student Paper Award (first runner-up) at FUSION conference in 2009, and the NATO Science and Technology Organization (STO) Scientific Achievement Award (SAA) 2017 for its contribution to the Development and Demonstration of Networked Autonomous ASW. He coauthored the paper received the Best Paper Award (first runner-up) at the SSPD conference in 2019. He was also the recipient of the NATO STO SAA, in 2020, as a Team Leader for the Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness. He coauthored the article receiving the Young Scientist Contest Award at the Signal Processing Symposium 2021. He is an Associate Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, *ISIF Journal of Advances in Information Fusion*, and *IET Radar, Sonar and Navigation*. In 2017, he was a Lead Guest Editor of the Special Issue Sonar Multi-Sensor Applications and Techniques in IET Radar, Sonar and Navigation. He swas an Associate Editor for the *IEEE Signal Processing Magazine*, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and *EURASIP Journal Advances on Signal Processing*.

MOE Z. WIN (Fellow, IEEE) is currently a Professor with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, 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 research interests include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He was 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 the recipient of two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, coauthored 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.