

Robust Multitarget Tracking in Interference Environments: A Message-Passing Approach

Xianglong Bai, Hua Lan, Zengfu Wang*, Quan Pan, Yuhang Hao, Can Li

Abstract—Multitarget tracking in the interference environments suffers from the nonuniform, unknown and time-varying clutter, resulting in dramatic performance deterioration. We address this challenge by proposing a robust multitarget tracking algorithm, which estimates the states of clutter and targets simultaneously by the message-passing (MP) approach. We define the non-homogeneous clutter with a finite mixture model containing a uniform component and multiple nonuniform components. The measured signal strength is utilized to estimate the mean signal-to-noise ratio (SNR) of targets and the mean clutter-to-noise ratio (CNR) of clutter, which are then used as additional feature information of targets and clutter to improve the performance of discrimination of targets from clutter. We also present a hybrid data association which can reason over correspondence between targets, clutter, and measurements. Then, a unified MP algorithm is used to infer the marginal posterior probability distributions of targets, clutter, and data association by splitting the joint probability distribution into a mean-field approximate part and a belief propagation part. As a result, a closed-loop iterative optimization of the posterior probability distribution can be obtained, which can effectively deal with the coupling between target tracking, clutter estimation and data association. Simulation results demonstrate the performance superiority and robustness of the proposed multitarget tracking algorithm compared with the probability hypothesis density (PHD) filter and the cardinalized PHD (CPHD) filter.

Index Terms—Robust multitarget tracking, message passing, mean-field approximation, belief propagation, radar interference.

I. INTRODUCTION

In radar target tracking, interference is often present in the received signals. Interference may arise in different forms, such as objects that are not of interest (e.g., precipitation, vegetation, soil), and electronic countermeasures (e.g., suppression jamming, chaff jamming), etc [1]. In many scenarios, interference suppression and clutter elimination techniques may not be as effective as expected and result in some nonuniform, unknown and time-varying clutter. The state-of-the-art algorithms assume that the distributions of both the spatial position and the number of clutter are known and fixed. Specifically, clutter is uniformly distributed in the space of radar measurements, the number of which follows a Poisson distribution with a fixed mean [2]. In specific applications, if the tracking algorithm uses a clutter distribution that does not match the real clutter distribution, it may lead to missed

target tracking, increased false tracks and computational complexity. To this end, robust multitarget tracking (RMTT) can be achieved by joint clutter estimation and target tracking (JCETT), which estimates the time-varying states of moving targets and unknown clutter simultaneously from measurements, leading to a significant performance improvement of multitarget tracking (MTT). Furthermore, to obtain a better performance in dense clutter and low SNR environments, it is desired to incorporate signal strength information into MTT.

However, RMTT in interference environments is complicated by the following factors. (a) *Clutter modelling and estimation*: Modern radars suffer from strong unwanted interference from natural environments and countermeasures. For example, atmosphere, e.g., water, fog, snow, smoke, is usually volumetric scattering and can result in illuminating clutter. The received signals of atmosphere may vary by several reasons, such as frequency, suspended particle sizes, and concentrations of atmospheric particles. In addition, the chaff interference is ejected from an aircraft or ship and blooms into a large reflector hovering in the cloud, which is affected by a couple of factors, such as radar cross section (RCS), shape of chaff clouds and atmospheric phenomena, etc [1]. Consequently, there are different types of clutter distribution in the background of target tracking: one uniform distribution generated by noise; several nonuniform distributions generated by interference, each of which we refer to as a clutter component. The non-homogeneous distributions make clutter modelling and estimation very challenging. (b) *Data association*: MTT in interference environments is difficult due to the unknown association between measurements and targets as well as the unknown association between measurements and clutter components. It is often assumed that at each time, a target can generate at most one measurement, a clutter component can potentially generate multiple measurements, and a measurement can be generated by either one target or one clutter component. Unfortunately, the exponential complexity of data association makes it incredibly challenging. In addition, the data association and the state estimation of target and clutter are highly coupled, i.e., erroneous data association deteriorates the state estimation of target and clutter, and the inaccurate state estimation of target and clutter leads to the data association risk.

The majority of the RMTT literature considering clutter estimation focused on finite set statistics. Mahler *et al.* [3] learned the clutter intensity while target tracking by proposing an adaptive PHD/CPHD filter. Assuming unknown and time-varying clutter, Beard *et al.* [4] proposed a bootstrap CPHD filter, which performed comparatively as well as the matched CPHD filter. Kim and Song [5] proposed a PHD filter with

The authors are with the School of Automation, Northwestern Polytechnical University, Xi'an 710129, China. This work was supported in part by the National Natural Science Foundation of China under Grant 61873211, Grant 61790552, Grant U21B2008, and in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2021JM-06.

* Corresponding author: Zengfu Wang.

clutter intensity estimation in dynamic cluttered environments. In environments with unknown and non-homogeneous clutter, Vo *et al.* [6] proposed a robust multi-Bernoulli (MB) filter to adaptively learn clutter parameters by modelling clutter based on multiple independent generators. In the vein of [6], Gostar *et al.* [7] proposed an MB filter with clutter estimation for sensor selection. In [8]–[10], the clutter intensity was modelled by a finite mixture model, and estimated in the random finite set scheme by the expectation maximization method and the Markov chain Monte Carlo method.

There is a large volume of published studies devoted to MTT using signal strength-related information to achieve RMTT. To improve the discrimination ability for closely spaced targets, assuming that the target SNR was known and fixed during tracking, the problem of data association between measurements and targets based on measured spatial and amplitude information was addressed in [11], [12]. Then, the amplitude likelihood was marginalized over a range of possible target SNR values and the corresponding amplitude information was integrated into the general PHD and CPHD filters [13], and cardinality balanced multitarget multi-Bernoulli (CBMeMBer) filter [14]. The amplitude information was used in multi-object filtering to estimate the target kinematic state and RCS in [15] for robust ground target tracking. Later, the amplitude information was used in MTT for a cluttered environment, and SNR estimation algorithms were proposed based on the maximum a posteriori method [16] and the sequential Monte Carlo method [17]. To improve the discrimination between targets and clutter, the joint amplitude likelihood of the sea clutter neighbouring cells is calculated, which is then integrated into a labelled MB filter [18]. Yang *et al.* [19] incorporated the amplitude information into the simultaneously clutter estimation and target tracking. Ristic *et al.* [20] proposed an MB filter for maritime target tracking using amplitude information.

However, most of the existing RMTT methods for JCETT only estimate the spatial distribution or the amplitude distribution of clutter, but not both. As far as we know, the previously proposed JCETT method estimating both the spatial state and return power of clutter is limited to the method proposed in [19]. The method in [19] adopted the random finite set filters for clutter estimation and target tracking, where the spatial and SNR of target and clutter were modelled as an inverse Gamma Gaussian (IGG) distribution. Unfortunately, the method [19] did not provide details on the spatial state estimation for nonuniform clutter and also resulted in an unappealing separate and sequential estimation framework for clutter and target. As we mentioned before, the target and clutter state estimation and data association are highly coupled and affect each other. It is highly demanded to develop a closed-loop iterative optimization framework for JCETT.

The Bayesian inference algorithms for probabilistic graphical models, such as (loopy) belief propagation (BP) [21] and variational inference [22], have been used for MTT. Chen *et al.* [23] considered the data association problem as a maximum a posteriori configuration problem, and solved it by the max-product BP. In contrast to [23], Williams *et al.* [24] considered the data association problem as a posteriori probability esti-

mation problem, and solved it by the sum-product BP. Later, the data association method proposed in [24] was extended to the multi-scan version [25], the multi-path version [26], and labeled MB filtering [27]. Meyer *et al.* [28], [29] proposed a scalable multi-sensor MTT algorithm by using BP, which was then extended to the scheme of self-tuning the unknown model parameters [30]. In addition, the BP method was also used to extended target tracking [31], [32], cooperative self-localization and MTT [33], joint registration and fusion of heterogeneous sensors [34], fusion of sensor measurements and target-provided information in MTT [35].

In our previous work, the unified MP method [36], which combines the virtues of (loopy) belief propagation (BP) [21] and Mean-field (MF) approximation [22] while circumventing their drawbacks, has been used for multi-path environment [37], maneuvering target tracking [38] and over-the-horizon radar network fusion [39]. These algorithms were derived by representing the joint probability density functions (PDFs) associated with the system models by a factor graph. We decomposed the factor graph into an MF part and a BP part, and the posterior PDFs of the corresponding hidden variables were approximated by BP and MF. As a result, a closed-loop iterative optimization of the posterior probability distribution were obtained, which can effectively deal with the coupling between latent variables.

In this paper, we propose an MP-based framework and algorithm for RMTT in interference environments using strength information of measurements. We formulate RMTT as an JCETT problem including all target states, clutter states and data association. In particular, we use the Swerling-I and Swerling-III models to represent the RCS fluctuations of target and clutter, and adopt the Rayleigh likelihood for the strength information. Furthermore, the clutter intensity is modelled as a finite mixture model. An enabling technique for our methods is the combined formulation of data association which can reason over correspondence between targets, clutter, and measurements. By this new formulation of the RMTT problem, the statistical structure of RMTT is represented by a factor graph. Finally, we use the unified MP algorithm to solve the problem, in which the MF approximation and BP are used in the MF part and the BP part of the factor graph, respectively. Different from our previous work in [38], [39], the factor graph constructed in this paper has several new subgraphs, including the target SNR subgraph, the clutter spatial state and mean CNR subgraph, the clutter mixing weight subgraph, and the combined data association subgraph. The modelling and message passing of these subgraphs makes the problem more challenging. The summary contributions of this paper are as follows:

- We formulate a Bayesian statistical framework for the RMTT problem involving all the hidden variables of targets, clutter and data association. In particular, the target and clutter state are modelled in both the spatial and the power domains, which can improve the performance of RMTT. We also develop a combined data association which can reason over correspondence between targets, clutter, and measurements.
- We use the MP approach combined with the MF approx-

imation and BP to solve the RMTT problem, referred to MP-RMTT. In MP-RMTT, the estimations of all hidden variables are optimized by a closed-loop iterative architecture, i.e., the target and clutter state estimations in the previous iteration are utilized to improve the data association, and then the new data association is used to optimize the target and clutter state estimations, which is capable for handling the coupling issue between the hidden variables.

The rest of the paper is organized as follows. The problem formulation of RMTT is described in Section II. In Section III, the proposed MP-RMTT algorithm is derived. Section IV evaluates the performance of MP-RMTT via simulations. At last, Section V concludes this paper.

II. PROBLEM FORMULATION

In this section, we present the system models firstly. After that, we state the RMTT problem to be solved.

A. Target State Modelling

At time k , let $\mathbf{X}_k^t = \{\mathbf{X}_{i,k}^t\}_{i=1}^{N_T}$ be the target *joint augmented state*, where N_T is the number of targets and is determined in initialization stage, and $\mathbf{X}_{i,k}^t = [(\mathbf{x}_{i,k}^t)^\top \sigma_{i,k}^t]^\top$ is the *augmented state* of target i , consisting of the target kinematic state $\mathbf{x}_{i,k}^t$ and mean SNR $\sigma_{i,k}^t$. The kinematic state $\mathbf{x}_{i,k}^t$ contains the position and velocity of target i . In addition, let target mean SNR $\sigma_{i,k}^t = S_{i,k}^t/N_0$, where $S_{i,k}^t$ is the expected target signal power and N_0 is the expected noise power. The SNR represented in log scale is $\text{SNR}(\text{dB}) = 10 \log_{10}(\sigma_{i,k}^t)$ [40]. For each target i , we define the target kinematic state sequence and the target mean SNR sequence over time 1 to K as $\mathbf{x}_{i,1:K}^t = \{\mathbf{x}_{i,k}^t\}_{k=1}^K$ and $\sigma_{i,1:K}^t = \{\sigma_{i,k}^t\}_{k=1}^K$, respectively. We define the target joint augmented state sequence over time 1 to time K as $\mathbf{X}_{1:K}^t = \{\mathbf{X}_k^t\}_{k=1}^K$.

For a same target, assuming that the PDF for the kinematic state, $p(\mathbf{x}_{i,k}^t)$, and the PDF for the mean SNR, $p(\sigma_{i,k}^t)$, are independent, the PDF of target augmented state is factorized as $p(\mathbf{X}_{i,k}^t) = p(\mathbf{x}_{i,k}^t)p(\sigma_{i,k}^t)$. $p(\mathbf{x}_{i,k}^t)$ is chosen to be a Gaussian distribution which is conjugate prior of the spatial measurement likelihood (which will be detailed in Section II-C). $p(\sigma_{i,k}^t)$ is chosen to be the IG distribution, which is conjugate prior of the strength measurement likelihood (which will be detailed in Section II-C) [15]. We refer the reader to [15], [19] for details of the IG distribution.

By assuming that each target augmented state evolves independently with a first-order Markov dynamic model, the PDF of $\mathbf{X}_{1:K}^t$ is

$$p(\mathbf{X}_{1:K}^t) = \prod_{i=1}^{N_T} p(\mathbf{X}_{i,1}^t) \prod_{k=2}^K p(\mathbf{X}_{i,k}^t | \mathbf{X}_{i,k-1}^t), \quad (1)$$

where $p(\mathbf{X}_{i,1}^t)$ is a prior PDF at time 1 and $p(\mathbf{X}_{i,k}^t | \mathbf{X}_{i,k-1}^t)$ is the transition PDF of target augmented state. Assuming that the target kinematic state transition PDFs $p(\mathbf{x}_{i,k}^t | \mathbf{x}_{i,k-1}^t)$ and the target mean SNR transition PDFs $p(\sigma_{i,k}^t | \sigma_{i,k-1}^t)$ are independent, one has $p(\mathbf{X}_{i,k}^t | \mathbf{X}_{i,k-1}^t) = p(\mathbf{x}_{i,k}^t | \mathbf{x}_{i,k-1}^t)p(\sigma_{i,k}^t | \sigma_{i,k-1}^t)$. The state transition PDF of

target kinematic state $p(\mathbf{x}_{i,k}^t | \mathbf{x}_{i,k-1}^t)$ can be determined by the dynamic model of each target [38]. Assume that the target mean SNR varies slowly. Given the PDF of target mean SNR at time $k-1$, we define $p(\sigma_{i,k}^t | \sigma_{i,k-1}^t) = \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|k-1}^t, \beta_{i,k|k-1}^t)$, where $\alpha_{i,k|k-1}^t = (\alpha_{i,k-1}^t + u^t - 1)/u^t$, $\beta_{i,k|k-1}^t = \beta_{i,k-1}^t/u^t$ with $\alpha_{i,k-1}^t$ and $\beta_{i,k-1}^t$ being parameters of the PDF $p(\sigma_{i,k-1}^t)$, and u^t being a forgetting factor [19]. We assume that the target mean SNR can be predicted in reverse by the same transition PDFs as the forward prediction.

At time k , let $\mathbf{S}_k = \{s_{i,k}\}_{i=1}^{N_T}$ be the target joint visibility state, where $s_{i,k} \in \{0, 1\}$ is a binary random variable and indicates the presence of target i if $s_{i,k} = 1$ or the absence of target i if $s_{i,k} = 0$. We define the target visibility state sequence over time 1 to time K of target i as $s_{i,1:K} = \{s_{i,k}\}_{k=1}^K$. We also define the sequence of target joint visibility state over time 1 to time K as $\mathbf{S}_{1:K} = \{\mathbf{S}_k\}_{k=1}^K$. By assuming that the appearance or the disappearance of each targets are independent and the visibility state of each target transits based on the Markov process, the PDF of $\mathbf{S}_{1:K}$ can be written as

$$p(\mathbf{S}_{1:K}) = \prod_{i=1}^{N_T} p(s_{i,1}) \prod_{k=2}^K p(s_{i,k} | s_{i,k-1}), \quad (2)$$

where $p(s_{i,1})$ is the prior PDF as a Bernoulli distribution, and the transition PDF $p(s_{i,k} | s_{i,k-1})$ is represented by a matrix

$$\mathbf{T}_k = \begin{bmatrix} p(s_{i,k} = 1 | s_{i,k} = 1) & p(s_{i,k} = 1 | s_{i,k} = 0) \\ p(s_{i,k} = 0 | s_{i,k} = 1) & p(s_{i,k} = 0 | s_{i,k} = 0) \end{bmatrix} \\ = \begin{bmatrix} p_s & p_b \\ 1 - p_s & 1 - p_b \end{bmatrix}, \quad (3)$$

where p_s is the target survival probability and p_b is the target birth probability.

B. Clutter Modelling

The clutter is modelled in the space of measurements. We assume that there are one uniform clutter component and multiple nonuniform clutter components distributed over the entire surveillance region. Let $\tau = 0$ and $\tau = 1, \dots, N_C$ be the indices of the uniform and the nonuniform clutter component respectively, and N_C is the maximum possible number of nonuniform clutter components and is determined in the initialization stage.

We define the clutter joint augmented state as $\mathbf{X}_k^c = \{\mathbf{X}_{\tau,k}^c\}_{\tau=0}^{N_C}$, where $\mathbf{X}_{0,k}^c = \sigma_{0,k}^c$ is the state of the uniform clutter component and $\mathbf{X}_{\tau,k}^c = \{\tilde{\mathbf{x}}_{\tau,k}^c \sigma_{\tau,k}^c\}$ is the augmented state of nonuniform clutter component τ . Here, $\sigma_{\tau,k}^c = S_{\tau,k}^c/N_0$ denotes the mean clutter-to-noise ration (CNR) of the clutter component τ , $\tau = 0, \dots, N_C$, where $S_{\tau,k}^c$ is the expected clutter power. The CNR represented in log scale is $\text{CNR}(\text{dB}) = 10 \log_{10}(\sigma_{\tau,k}^c)$ [40]. Note that we assume the expected power of the uniform clutter is equal to the power of background noise, so the CNR of the uniform clutter is 1. In fact, since the background noise is unknown, we estimate the CNR of the uniform clutter as well. Then, $\tilde{\mathbf{x}}_{\tau,k}^c = \{\mathbf{x}_{\tau,k}^c \mathbf{D}_{\tau,k}^c\}$ is the spatial state of the nonuniform clutter component τ , $\tau = 1, \dots, N_C$, where $\mathbf{x}_{\tau,k}^c$ and $\mathbf{D}_{\tau,k}^c$ are

the corresponding position and shape parameters, respectively. At time k , we define $\mathbf{\Pi}_k = \{\pi_{\tau,k}\}_{\tau=0}^{N_C}$ as the clutter joint mixing weights, where $\pi_{\tau,k}$ is the mixing weight of clutter component τ and satisfies $0 \leq \pi_{\tau,k} \leq 1$ and $\sum_{\tau=0}^{N_C} \pi_{\tau,k} = 1$. We define the clutter joint spatial state as $\tilde{\mathbf{X}}_k^c = \{\tilde{\mathbf{x}}_{\tau,k}^c\}_{\tau=1}^{N_C}$ and the clutter joint CNR as $\sigma_k^c = \{\sigma_{\tau,k}^c\}_{\tau=1}^{N_C}$. We define the spatial state sequence, the mean CNR sequence, the mixing weight sequence of clutter component τ over time 1 to time K as $\tilde{\mathbf{X}}_{\tau,1:K}^c = \{\tilde{\mathbf{x}}_{\tau,k}^c\}_{k=1}^K$, $\sigma_{\tau,1:K}^c = \{\sigma_{\tau,k}^c\}_{k=1}^K$, $\pi_{\tau,1:K}^c = \{\pi_{\tau,k}^c\}_{k=1}^K$, respectively. We define the clutter joint augmented state sequence and the clutter joint mixing weights sequence over time 1 to time K as $\mathbf{X}_{1:K}^c = \{\mathbf{X}_k^c\}_{k=1}^K$ and $\mathbf{\Pi}_{1:K}^c = \{\mathbf{\Pi}_k^c\}_{k=1}^K$, respectively.

We assume the PDF of clutter spatial state $p(\tilde{\mathbf{x}}_{i,k}^c)$ and the PDF of clutter mean CNR $p(\sigma_{i,k}^c)$ are independent, and the PDF of the clutter augmented state can be factorized as $p(\mathbf{X}_{i,k}^c) = p(\tilde{\mathbf{x}}_{i,k}^c)p(\sigma_{i,k}^c)$. Since the position measurement likelihood given the clutter spatial state (will be detailed in Section II-C) is a Gaussian distribution, we define the PDF $p(\tilde{\mathbf{x}}_{i,k}^c)$ as a Gaussian-Wishart (GW) distribution which is the conjugate prior of the mean and covariance for a Gaussian distribution. We refer the reader to [41] for details of the Wishart distribution. Therefore, we introduce the GW prior governing the spatial parameters of each nonuniform clutter component, given by

$$p(\tilde{\mathbf{x}}_{\tau,k}^c) = p(\mathbf{x}_{\tau,k}^c, \mathbf{D}_{\tau,k}^c) = p(\mathbf{x}_{\tau,k}^c | \mathbf{D}_{\tau,k}^c) p(\mathbf{D}_{\tau,k}^c) = \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k}^c, (\beta_{\tau,k}^c \mathbf{D}_{\tau,k}^c)^{-1}) \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k}^c, \nu_{\tau,k}^c). \quad (4)$$

Assuming that the clutter and targets have the same likelihood function on signal strength, the PDF of clutter mean CNR $p(\sigma_{i,k}^c)$ is also chosen to be the IG distribution. In addition, since the likelihood of the data association event given the clutter mixing weights (will be detailed in Section II-C) is a multinomial distribution, we choose a Dirichlet distribution on the clutter mixing weights $\text{Dir}(\mathbf{\Pi}_k | \boldsymbol{\alpha}_k) = C(\boldsymbol{\alpha}_k) \prod_{\tau=0}^{N_C} \pi_{\tau,k}^{\alpha_{\tau,k}}$, where $C(\boldsymbol{\alpha}_k)$ is the normalization constant, and $\boldsymbol{\alpha}_k = \{\alpha_{\tau,k}\}_{\tau=0}^{N_C}$ with $\alpha_{\tau,k}$ as the prior number of points associated with clutter component τ .

By assuming that each clutter augmented state transits independently based on the Markov process, the PDF of $\mathbf{X}_{1:K}^c$ can be written as

$$p(\mathbf{X}_{1:K}^c) = \prod_{\tau=1}^{N_T} p(\mathbf{X}_{\tau,1}^c) \prod_{k=2}^K p(\mathbf{X}_{\tau,k}^c | \mathbf{X}_{\tau,k-1}^c), \quad (5)$$

where $p(\mathbf{X}_{\tau,1}^c)$ is the prior PDF at time 1 and $p(\mathbf{X}_{\tau,k}^c | \mathbf{X}_{\tau,k-1}^c)$ is the transition PDF of clutter augmented state. Assume that the clutter spatial transition PDFs $p(\tilde{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c)$ and the clutter mean CNR transition PDFs $p(\sigma_{\tau,k}^c | \sigma_{\tau,k-1}^c)$ are independent, one has $p(\mathbf{X}_{\tau,k}^c | \mathbf{X}_{\tau,k-1}^c) = p(\tilde{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c) p(\sigma_{\tau,k}^c | \sigma_{\tau,k-1}^c)$. In some real applications, compared with the high speed of targets, the dynamics of the interference mainly changes on shape rather than on spatial movement. For example, the atmosphere moves slowly while its shape varies as the meteorological conditions; the chaff cloud hovers in the air while its scattering volume spreads [1]. Thus, we assume that the position of the nonuniform clutter changes within a relatively

small region and the shape of the nonuniform clutter varies over time. Accordingly, we define the transition PDF of the clutter spatial state as the GW distribution as well, that is,

$$p(\tilde{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c) = \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c, (\beta_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c \mathbf{D}_{\tau,k}^c)^{-1}) \times \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c, \nu_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c), \quad (6)$$

where $\hat{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c = \hat{\mathbf{x}}_{\tau,k-1}^c$, $\beta_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c = \beta_{\tau,k-1}^c$, $\mathbf{W}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c = \xi \mathbf{W}_{\tau,k-1}^c$, $\nu_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c = \xi(\nu_{\tau,k-1}^c - m - 1) + m + 1$, with $\hat{\mathbf{x}}_{\tau,k-1}^c$, $\beta_{\tau,k-1}^c$, $\mathbf{W}_{\tau,k-1}^c$, $\nu_{\tau,k-1}^c$ being the parameters of $p(\tilde{\mathbf{x}}_{\tau,k-1}^c)$, $m = 2$ being the dimension of the position measurement, and ξ being a forgetting factor [42]. Again, we assume that the clutter spatial state can be predicted in reverse by the same transition PDFs as the forward prediction. The transition PDF $p(\sigma_{\tau,k}^c | \sigma_{\tau,k-1}^c)$ and the inverse transition PDF $p(\sigma_{\tau,k-1}^c | \sigma_{\tau,k}^c)$ of clutter mean CNR are the same as those of target mean SNR with corresponding forgetting factor u^c . By assuming that each clutter component appears and disappears independently and transits based on the Markov process, the PDF of $\mathbf{\Pi}_{1:K}^c$ can be represented as

$$p(\mathbf{\Pi}_{1:K}^c) = \prod_{\tau=1}^{N_T} p(\pi_{\tau,1}) \prod_{k=2}^K p(\pi_{\tau,k} | \pi_{\tau,k-1}), \quad (7)$$

where $p(\pi_{\tau,1})$ is the prior PDF at time 1 and $p(\pi_{\tau,k} | \pi_{\tau,k-1})$ is transition PDF of clutter mixing weight, which is also defined as the Dirichlet distribution, that is $p(\pi_{\tau,k} | \pi_{\tau,k-1}) = \text{Dir}(\pi_{\tau,k} | \alpha_{\tau,k} | \pi_{\tau,k-1})$. Assuming that the mixing weights change slowly and $\alpha_{\tau,k-1} / \sum_{\tau'=0}^{N_C} \alpha_{\tau',k-1} = \alpha_{\tau,k} / \sum_{\tau'=0}^{N_C} \alpha_{\tau',k}$. We define $\alpha_{\tau,k} | \pi_{\tau,k-1} = \kappa M_{k-1} \alpha_{\tau,k-1} / \sum_{\tau'=0}^{N_C} \alpha_{\tau',k-1}$ by a heuristic approach, where κ is a balance parameter tuning the effects of the prior knowledge and M_{k-1} is the total number of clutter at time $k-1$. We assume that the clutter mixing weights can be predicted in reverse by the same transition PDFs as the forward prediction.

C. Measurement Modelling

At time k , let $\mathbf{Y}_k = \{\mathbf{Y}_{j,k}\}_{j=1}^{N_{M,k}}$ be the measurements, which are generated from radar range-azimuth-strength map using the constant false alarm detector, followed by a peak extraction scheme and detection with a threshold $d > 0$, and $N_{M,k}$ is the number of measurements. Each measurement $\mathbf{Y}_{j,k} = [\mathbf{y}_{j,k}^T, m_{j,k}]^T$ consists of two elements: (i) the spatial information in the polar coordinates $\mathbf{y}_{j,k} = [r_{j,k}, \xi_{j,k}]^T$; (ii) the measured signal strength $m_{j,k} = S_{j,k} / N_0 > d$, where $S_{j,k}$ is received signal power. We define the measurement sequence over time 1 to time K as $\mathbf{Y}_{1:K} = \{\mathbf{Y}_k\}_{k=1}^K$.

For the target-originated measurement, the spatial measurement likelihood of $\mathbf{y}_{j,k}$ given $\mathbf{x}_{i,k}$ is denoted as $p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k})$, which can be determined by the spatial measurement equation [2]. The RCS fluctuations of target can be captured by the Swerling-I and Swerling-III models [1], and the corresponding PDF of strength $m_{j,k}$ in noise background given the target mean SNR $\sigma_{i,k}^t$ can be represented by the general Rayleigh distribution [40]

$$\mathcal{R}(m_{j,k}; \sigma_{i,k}^t, n) = \frac{2m_{j,k}^{2n-1}}{(\sigma_{i,k}^t + 1)^n} \exp\left(-n \frac{m_{j,k}^2}{\sigma_{i,k}^t}\right), \quad (8)$$

where $n = 1$ and $n = 2$ denotes the Swerling-I and Swerling-III models, respectively. Using the approximate expression, the corresponding detection probability is [19]

$$p_D^d = \exp\left(-n \frac{d^2}{\sigma_{i,k}^t + 1}\right). \quad (9)$$

The general Rayleigh PDF after thresholding becomes

$$\mathcal{R}^d(m_{j,k}; \sigma_{i,k}^t, n) = \frac{2m_{j,k}^{2n-1}}{(\sigma_{i,k}^t + 1)^n} \exp\left(-n \frac{m_{j,k}^2 - d^2}{\sigma_{i,k}^t + 1}\right). \quad (10)$$

Likewise in [15], [19], we assume that the measured position and the measured strength information are independent of each other; the corresponding likelihood function can be decomposed as

$$p(\mathbf{Y}_{j,k} | \mathbf{X}_{i,k}^t) = p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k}^t) \mathcal{R}^d(m_{j,k}; \sigma_{i,k}^t, n). \quad (11)$$

We use the finite mixture model to model the spatial distribution of clutter, given by

$$p(\mathbf{y}_{j,k} | \tilde{\mathbf{X}}_k^c, \mathbf{\Pi}_k) = \pi_{0,k} U(V_G) + \sum_{\tau=1}^{N_C} \pi_{\tau,k} p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c), \quad (12)$$

where $U(V_G) = 1/V_G$ is the uniform distribution representing the spatial distribution of uniform clutter component $\tau = 0$; $p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c) = \mathcal{N}(\mathbf{y}_{j,k}; \mathbf{x}_{\tau,k}^c, (\mathbf{D}_{\tau,k}^c)^{-1})$ is the Gaussian distribution with mean $\mathbf{x}_{\tau,k}^c$ and precision matrix $\mathbf{D}_{\tau,k}^c$ representing the spatial distribution of nonuniform clutter component τ . Note that we use precision matrix rather than covariance matrix as this somewhat simplifies the mathematics. By assuming that the nonuniform clutter component is caused by the volumetric scattering interference and the clutter point is dominated by the underlying scattering body, we interpret $\mathbf{x}_{\tau,k}^c$ and $(\mathbf{D}_{\tau,k}^c)^{-1}$ as the centroid and shape of each corresponding nonuniform clutter region, respectively.

We assume that the clutter and targets have the same power fluctuates and the PDF of strength likelihood function is also a Rayleigh distribution after thresholding,

$$\mathcal{R}^d(m_{j,k}; \sigma_{i,k}^c, n) = \frac{2m_{j,k}^{2n-1}}{(\sigma_{i,k}^c + 1)^n} \exp\left(-n \frac{m_{j,k}^2 - d^2}{\sigma_{i,k}^c + 1}\right). \quad (13)$$

As noted in Eq. (12), the measurement likelihood is dependent on both clutter augmented states and clutter mixing weights; this coupling issue will make MTT difficult to solve. In the following, the measurement likelihood is decoupled by introducing a data association event.

D. Data Association

Let $\mathbf{A}_k = \mathbf{A}_k^t \cup \mathbf{A}_k^c$ be the joint data association events at time k , where $\mathbf{A}_k^t = \{a_{i,j,k}^t\}_{i=1}^{N_T} \{a_{\tau,j,k}^c\}_{\tau=0}^{N_C}$ is the joint data association events between measurements and targets, $\mathbf{A}_k^c = \{a_{\tau,j,k}^c\}_{\tau=0}^{N_C} \{a_{i,j,k}^t\}_{i=1}^{N_T}$ is the joint data association events between measurements and clutter. The binary association variable $a_{i,j,k}^t$ denotes that the measurement $\mathbf{Y}_{j,k}$ is generated by target i if $a_{i,j,k}^t = 1$; likewise, the binary association variable

$a_{\tau,j,k}^c$ denotes that the measurement $\mathbf{Y}_{j,k}$ belongs to clutter component τ if $a_{\tau,j,k}^c = 1$, given as

$$a_{i,j,k}^t = \begin{cases} 1, & \text{if } \mathbf{Y}_{j,k} \text{ is generated by target } i, \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

$$a_{\tau,j,k}^c = \begin{cases} 1, & \text{if } \mathbf{Y}_{j,k} \text{ belongs to clutter component } \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

In particular, $a_{i,0,k}^t$ denotes that the target i is missed and $a_{\tau,0,k}^c$ represents that the clutter component τ disappears. For the convenience of description, the following data association event sets are defined, $\mathbf{A}_{j,k} = \{a_{i,j,k}^t\}_{i=1}^{N_T} \cup \{a_{\tau,j,k}^c\}_{\tau=0}^{N_C}$, $\mathbf{A}_{i,k}^t = \{a_{i,j,k}^t\}_{j=0}^{N_{M,k}}$, and $\mathbf{A}_{\tau,k}^c = \{a_{\tau,j,k}^c\}_{j=0}^{N_{M,k}}$. We define the joint data association sequence over time 1 to time K as $\mathbf{A}_{1:K} = \{\mathbf{A}_k\}_{k=1}^K$.

A valid joint data association event \mathbf{A}_k satisfies the following three constraints: (a) Each measurement is originated from at most one target or belongs to at most one clutter component, denoted as $I_{j,k}(\mathbf{A}_{j,k})$; (b) Each target can generate at most one measurement, denoted as $E_{i,k}^t(\mathbf{A}_{i,k}^t)$; (c) Each clutter component either generates measurements or not, denoted as $E_{\tau,k}^c(\mathbf{A}_{\tau,k}^c)$. According to the above three constraints, the following constraint equations are obtained.

$$I_{j,k}(\mathbf{A}_{j,k}) = \begin{cases} 1, & \text{if } \sum_{a_{i,j,k} \in \mathbf{A}_{j,k}} a_{i,j,k} = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

$$E_{i,k}^t(\mathbf{A}_{i,k}^t) = \begin{cases} 1, & \text{if } \sum_{a_{i,j,k}^t \in \mathbf{A}_{i,k}^t} a_{i,j,k}^t = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

$$E_{\tau,k}^c(\mathbf{A}_{\tau,k}^c) = \begin{cases} 0, & \text{if } \exists j > 0 \text{ such that} \\ & a_{\tau,0,k}^c = 1 \text{ and } a_{\tau,j,k}^c > 0, \\ 1, & \text{otherwise.} \end{cases} \quad (18)$$

Define the following set of constraints.

$$\mathbf{I}(\mathbf{A}_{1:K}) = \prod_{k=1}^K \mathbf{I}_k(\mathbf{A}_k) = \prod_{k=1}^K \prod_{j=1}^{N_{M,k}} I_{j,k}(\mathbf{A}_{j,k}), \quad (19)$$

$$\begin{aligned} \mathbf{E}(\mathbf{A}_{1:K}) &= \prod_{k=1}^K \mathbf{E}_k(\mathbf{A}_k) = \prod_{k=1}^K \mathbf{E}_k^t(\mathbf{A}_k^t) \mathbf{E}_k^c(\mathbf{A}_k^c) \\ &= \prod_{k=1}^K \prod_{i=1}^{N_T} E_{i,k}^t(\mathbf{A}_{i,k}^t) \prod_{\tau=0}^{N_C} E_{\tau,k}^c(\mathbf{A}_{\tau,k}^c). \end{aligned} \quad (20)$$

The joint prior probability of data association sequence $\mathbf{A}_{1:k}$ given the joint target visibility state sequence $\mathbf{S}_{1:k}$ is

$$\begin{aligned} p(\mathbf{A}_{1:K} | \mathbf{S}_{1:K}) &= \prod_{k=1}^K p(\mathbf{A}_k^t | \mathbf{S}_k) \\ &= \prod_{k=1}^K \prod_{i=1}^{N_T} P_d(s_{i,k})^{1-a_{i,0,k}^t} (1 - P_d(s_{i,k}))^{a_{i,0,k}^t}, \end{aligned} \quad (21)$$

where $P_d(s_{i,k})$ represents the detection probability of target i given $s_{i,k}$, i.e., $P_d(s_{i,k} = 1) = \hat{p}_{D,i,k}^d$ and $P_d(s_{i,k} = 0) = \varepsilon$ ($0 < \varepsilon \ll 1$) [39], where $\hat{p}_{D,i,k}^d$ is estimated detection

probability of target i calculated by Eq. (9) with the estimated target mean SNR $\mathbb{E}(\sigma_{i,k}^t)$ and detection threshold d .

The probability that the measurement $\mathbf{Y}_{j,k}$ belongs to the clutter component τ can be represented by the clutter mixing weights $\pi_{\tau,k}$, i.e., $p(a_{\tau,j,k}^c = 1) = \pi_{\tau,k}$. The conditional distribution of the data association event given the clutter mixing weights can be written as a multinomial distribution,

$$p(\mathbf{A}_{1:K}|\mathbf{\Pi}_{1:K}) = \prod_{k=1}^K p(\mathbf{A}_k|\mathbf{\Pi}_k) = \prod_{k=1}^K \prod_{\tau=0}^{N_C} \prod_{j=1}^{N_M} \pi_{\tau,k}^{a_{\tau,j,k}^c}. \quad (22)$$

Similarly, the conditional PDF of the spatial measurement $\mathbf{y}_{j,k}$ given the clutter joint spatial state $\tilde{\mathbf{X}}_k^c$ and the joint data association event $\mathbf{A}_{\tau,k}^c$ can be represented by

$$p(\mathbf{y}_{j,k}|\tilde{\mathbf{X}}_k^c, \mathbf{A}_{\tau,k}^c) = \prod_{\tau=0}^{N_C} p(\mathbf{y}_{j,k}|\tilde{\mathbf{x}}_{\tau,k}^c)^{a_{\tau,j,k}^c}. \quad (23)$$

We have therefore found an equivalent formulation of the finite mixture model (as Eq. (12)) involving a data association event, leading to significant simplifications of our MP method that will be presented in Section III. We can represent the conditional distribution of the measurement sequence $\mathbf{Y}_{1:k}$ given the target joint augmented state sequence $\mathbf{X}_{1:k}^t$, clutter joint augmented state sequence $\mathbf{X}_{1:k}^c$, and data association sequence $\mathbf{A}_{1:k}$ as

$$\begin{aligned} p(\mathbf{Y}_{1:K}|\mathbf{X}_{1:K}^t, \mathbf{X}_{1:K}^c, \mathbf{A}_{1:K}) &= \prod_{k=1}^K p(\mathbf{Y}_k|\mathbf{X}_k^t, \mathbf{X}_k^c, \mathbf{A}_k) \\ &= \prod_{k=1}^K \prod_{j=0}^{N_{M,k}} \prod_{\tau=0}^{N_C} p(\mathbf{y}_{j,k}|\tilde{\mathbf{x}}_{\tau,k}^c)^{a_{\tau,j,k}^c} \prod_{i=1}^{N_T} p(\mathbf{y}_{j,k}|\mathbf{x}_{i,k}^t)^{a_{i,j,k}^t}. \end{aligned} \quad (24)$$

Note that, by introducing the data association \mathbf{A}_k , \mathbf{X}_k^t is conditionally independent of \mathbf{E}_k and \mathbf{X}_k^c is conditionally independent of $\mathbf{\Pi}_k$, and we obtain a new measurement likelihood function as in Eq. (24). Furthermore, the new measurement likelihood function can simplify the derivation of message passing rules (The details will be given in Section III).

E. The Joint PDF and Problem Statement

Let $\Theta_{1:K} = \{\mathbf{X}_{1:K}^t, \mathbf{X}_{1:K}^c, \mathbf{S}_{1:K}, \mathbf{\Pi}_{1:K}, \mathbf{A}_{1:K}\}$ denote the collection of all the latent variables. The joint posterior PDF $\mathcal{L}(\Theta_{1:K})$ can be factorized as

$$\begin{aligned} \mathcal{L}(\Theta_{1:K}) &= \frac{p(\mathbf{X}_{1:K}^t, \mathbf{X}_{1:K}^c, \mathbf{S}_{1:K}, \mathbf{\Pi}_{1:K}, \mathbf{A}_{1:K}, \mathbf{Y}_{1:K})}{p(\mathbf{Y}_{1:K})} \\ &\propto p(\mathbf{Y}_{1:K}|\mathbf{X}_{1:K}^t, \mathbf{X}_{1:K}^c, \mathbf{A}_{1:K}) \\ &\quad \times p(\mathbf{X}_{1:K}^t)p(\mathbf{S}_{1:K})p(\mathbf{X}_{1:K}^c)p(\mathbf{\Pi}_{1:K}) \\ &\quad \times p(\mathbf{A}_{1:K}|\mathbf{S}_{1:K})p(\mathbf{A}_{1:K}|\mathbf{\Pi}_{1:K})\mathbf{E}(\mathbf{A}_{1:K})\mathbf{I}(\mathbf{A}_{1:K}). \end{aligned} \quad (25)$$

Insert Eq. (1) for $p(\mathbf{X}_{1:K}^t)$, Eq. (2) for $p(\mathbf{S}_{1:K})$, Eq. (5) for $p(\mathbf{X}_{1:K}^c)$, Eq. (7) for $p(\mathbf{\Pi}_{1:K})$, Eq. (19) for $\mathbf{I}(\mathbf{A}_{1:K})$, Eq. (20) for $\mathbf{E}(\mathbf{A}_{1:K})$, Eq. (21) for $p(\mathbf{A}_{1:K}|\mathbf{S}_{1:K})$, Eq. (22) for $p(\mathbf{A}_{1:K}|\mathbf{\Pi}_{1:K})$, and Eq. (24) for $p(\mathbf{Y}_{1:K}|\mathbf{X}_{1:K}^t, \mathbf{X}_{1:K}^c, \mathbf{A}_{1:K})$, yielding the factorization of Eq. (25) as Eq. (26).

The aim of RMTT is to simultaneously estimate $\mathbf{X}_{1:K}^t$ (target augmented state estimation), $\mathbf{S}_{1:K}$ (target detection), $\mathbf{X}_{1:K}^c$ (clutter augmented state estimation) and $\mathbf{\Pi}_{1:K}$ (clutter

$$\begin{aligned} &\mathcal{L}(\Theta_{1:K}) \\ &\propto \prod_{k=1}^K \prod_{j=0}^{N_{M,k}} \prod_{\tau=0}^{N_C} p(\mathbf{y}_{j,k}|\tilde{\mathbf{x}}_{\tau,k}^c)^{a_{\tau,j,k}^c} \prod_{i=1}^{N_T} p(\mathbf{y}_{j,k}|\mathbf{x}_{i,k}^t)^{a_{i,j,k}^t} \\ &\quad \times \prod_{i=1}^{N_T} p(\mathbf{x}_{i,1}^t)p(\mathbf{s}_{i,1}) \prod_{k=2}^K p(\mathbf{x}_{i,k}^t|\mathbf{x}_{i,k-1}^t)p(\mathbf{s}_{i,k}|\mathbf{s}_{i,k-1}) \\ &\quad \times \prod_{\tau=0}^{N_C} p(\mathbf{x}_{\tau,1}^c)p(\pi_{\tau,1}) \prod_{k=2}^K p(\mathbf{x}_{\tau,k}^c|\mathbf{x}_{\tau,k-1}^c)p(\pi_{\tau,k}|\pi_{\tau,k-1}) \\ &\quad \times \prod_{k=1}^K \mathbf{E}_k(\mathbf{A}_k)\mathbf{I}_k(\mathbf{A}_k) \prod_{k=1}^K p(\mathbf{A}_k|\mathbf{S}_k)p(\mathbf{A}_k|\mathbf{\Pi}_k). \end{aligned} \quad (26)$$

mixing weights estimation), given measurements $\mathbf{Y}_{1:K}$ with unknown $\mathbf{A}_{1:K}$. The posterior PDFs of $\mathbf{X}_{1:K}^t$, $\mathbf{S}_{1:K}$, $\mathbf{X}_{1:K}^c$ and $\mathbf{\Pi}_{1:K}$ can be obtained by marginalizing $\mathcal{L}(\Theta_{1:K})$ in the Bayesian framework. Unfortunately, the marginalizing of data association is exponentially complex, making exact solution computationally prohibitive. Therefore, the combined BP-MF approximation is adopted in the next section.

III. MP-BASED APPROACH FOR RMTT

In this section, we first present the framework for solving the RMTT problem using the combined BP-MF MP algorithm. Then, the message update rules and the approximate beliefs of all hidden variables, including target joint augmented states, clutter joint augmented states, target visibility states, clutter mixing weights and data association, are derived. Finally, we present the initialization, implementation and computational complexity for the proposed RMTT algorithm.

A. Combined BP-MF MP Approach for RMTT

The combined BP-MF MP approach performs inference on a probabilistic graphical model (typically, a factor graph) by exchanging messages. The first step to apply the combined BP-MF approach is to design the factor graph corresponding to the joint PDF of the problem to be solved, for instance, Eq. (26). This is accomplished by associating each random variable or vector in the joint PDF to a variable node in the factor graph and associating each function defined on random variables (including PDFs, conditional PDFs, constrains etc.) to a factor node in the factor graph, and connect a variable node with a factor node if the latter is a function of the former. The factor graph is then split into an MF region and a BP region to maximize their advantages and circumvent their disadvantages. Specifically, BP can deal effectively with hard constraints and has good approximations to marginal PDFs, but is not suitable for situations containing both continuous and discrete hidden variables. MF approximation can guarantee convergence and is straightforward to derive for conjugate-exponential models, but is incompatible with hard constraint. The message update rules and the beliefs of each hidden variables are eventually derived by solving the constrained minimum region-based free energy problem on the factor graph via the Lagrangian

relaxation method. For the detailed derivation of the combined BP-MF MP approach, the reader is referred to [36]. Following the ideas of the combined MP approach, next we present our proposed MP-based method for solving the RMTT problem.

By observing Eq. (26), we define the factor nodes $f_{Y_k} \triangleq p(\mathbf{Y}_k | \mathbf{X}_k^t, \mathbf{X}_k^c, \mathbf{A}_k)$, $f_{X_k^t} \triangleq \prod_{i=1}^{N_T} p(\mathbf{X}_{i,k}^t | \mathbf{X}_{i,k-1}^t)$, $f_{X_k^c} \triangleq \prod_{\tau=1}^{N_C} p(\mathbf{X}_{\tau,k}^c | \mathbf{X}_{\tau,k-1}^c)$, $f_{S_k} \triangleq \prod_{i=1}^{N_T} p(s_{i,k} | s_{i,k-1})$, $f_{\Pi_k} \triangleq \prod_{\tau=1}^{N_C} p(\pi_{\tau,k} | \pi_{\tau,k-1})$, $f_{A_k^t} \triangleq p(\mathbf{A}_k^t | \mathbf{S}_k)$, $f_{A_k^c} \triangleq p(\mathbf{A}_k^c | \Pi_k)$, $f_{E_k} \triangleq \mathbf{E}_k(\mathbf{A}_k)$, $f_{I_k} \triangleq \mathbf{I}_k(\mathbf{A}_k)$, the set of variable nodes $\mathcal{I} \triangleq \{\mathbf{X}_k^t, \mathbf{X}_k^c, \mathbf{S}_k, \Pi_k, \mathbf{A}_k\}_{k=1}^K$, and the set of factor nodes $\mathcal{F} \triangleq \{f_{Y_k}, f_{X_k^t}, f_{X_k^c}, f_{S_k}, f_{\Pi_k}, f_{A_k^t}, f_{A_k^c}, f_{E_k}, f_{I_k}\}_{k=1}^K$. The corresponding factor graph is illustrated in Fig. 1.

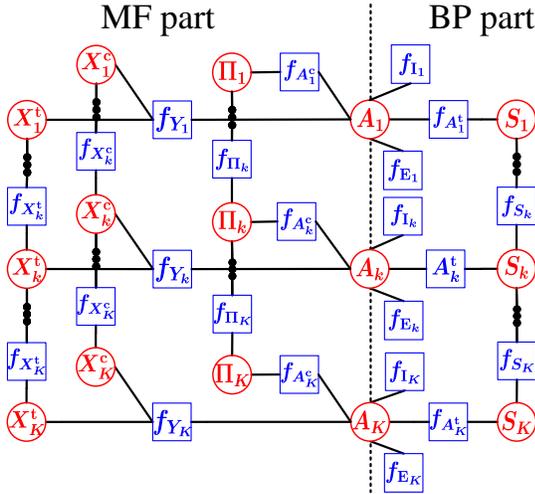


Fig. 1: Factor graph represented by Eq. (26).

We next describe how the factor graph in Fig. 1 is divided into a BP part and an MF part. BP is used to estimate the target visibility state in order to obtain a good approximation for the corresponding posterior PDFs. BP is also used for data association, which contains hard constraints as in Eq. (16)-Eq. (18). The MF approximation is used for the estimation of target joint augmented state, clutter joint augmented state and clutter mixture weight, as these hidden variables are conjugate-exponential models and therefore simple MP rules can be obtained. Accordingly, we split \mathcal{F} into \mathcal{F}_{BP} and \mathcal{F}_{MF} with

$$\begin{aligned} \mathcal{F}_{BP} &= \{f_{S_k}, f_{A_k^t}, f_{E_k}, f_{I_k}\}_{k=1}^K, \\ \mathcal{F}_{MF} &= \{f_{Y_k}, f_{X_k^t}, f_{X_k^c}, f_{\Pi_k}, f_{A_k^c}\}_{k=1}^K. \end{aligned} \quad (27)$$

We have

$$\mathcal{I}_{BP} = \{\mathbf{S}_k, \mathbf{A}_k\}_{k=1}^K, \quad \mathcal{I}_{MF} = \{\mathbf{X}_k^t, \mathbf{X}_k^c, \Pi_k, \mathbf{A}_k\}_{k=1}^K. \quad (28)$$

Then, the joint posterior PDF $\mathcal{L}(\Theta_{1:K})$ can be written as

$$\mathcal{L}(\Theta_{1:K}) = \underbrace{\prod_{k=1}^K f_{Y_k} f_{X_k^t} f_{X_k^c} f_{\Pi_k} f_{A_k^c}}_{\text{MF region}} \times \underbrace{\prod_{k=1}^K f_{S_k} f_{A_k^t} f_{E_k} f_{I_k}}_{\text{BP region}}. \quad (29)$$

The beliefs of hidden variables are approximated by [36]

$$b_{\mathbf{X}}(\mathbf{X}_{1:K}^t) \propto \prod_{\alpha \in \mathcal{S}_{MF}(\mathbf{X}_{1:K}^t)} m_{\alpha \rightarrow \mathbf{X}_{1:K}^t}^{\text{MF}}(\mathbf{X}_{1:K}^t), \quad (30)$$

$$b_{\mathbf{X}}(\mathbf{X}_{1:K}^c) \propto \prod_{\alpha \in \mathcal{S}_{MF}(\mathbf{X}_{1:K}^c)} m_{\alpha \rightarrow \mathbf{X}_{1:K}^c}^{\text{MF}}(\mathbf{X}_{1:K}^c), \quad (31)$$

$$b_{\mathbf{S}}(\mathbf{S}_{1:K}) \propto \prod_{\alpha \in \mathcal{S}_{BP}(\mathbf{S}_{1:K})} m_{\alpha \rightarrow \mathbf{S}_{1:K}}^{\text{BP}}(\mathbf{S}_{1:K}), \quad (32)$$

$$b_{\Pi}(\Pi_{1:K}) \propto \prod_{\alpha \in \mathcal{S}_{MF}(\Pi_{1:K})} m_{\alpha \rightarrow \Pi_{1:K}}^{\text{MF}}(\Pi_{1:K}), \quad (33)$$

$$b_{\mathbf{A}}(\mathbf{A}_{1:K}) \propto \prod_{\alpha \in \mathcal{S}_{BP}(\mathbf{A}_{1:K})} m_{\alpha \rightarrow \mathbf{A}_{1:K}}^{\text{BP}}(\mathbf{A}_{1:K}) \prod_{\alpha \in \mathcal{S}_{MF}(\mathbf{A}_{1:K})} m_{\alpha \rightarrow \mathbf{A}_{1:K}}^{\text{MF}}(\mathbf{A}_{1:K}). \quad (34)$$

Next, we present the detailed derivations of the beliefs for each hidden variables and the messages in Eqs. (30)-Eqs. (34). Before we present the details, we introduce the following proposition [36], [39], which is required in the derivations of the beliefs.

Proposition 1. For all messages passed from variable nodes i to factor nodes α in the MF region, $n_{i \rightarrow \alpha} = b_i(\mathbf{x}_i)$.

B. Derivations of Belief $b_{\mathbf{X}}(\mathbf{X}_{1:K}^t)$

Fig. 2 shows the target tracking and mean SNR estimation subgraph corresponding to the belief $b_{\mathbf{x}}(\mathbf{x}_{1:K}^t)$ and $b_{\sigma}(\sigma_{i,1:K}^t)$, $i = 1, \dots, N_T$, which contains the variable nodes $\mathbf{x}_{i,k}^t$ and $\sigma_{i,k}^t$, $k = 1, \dots, K$.

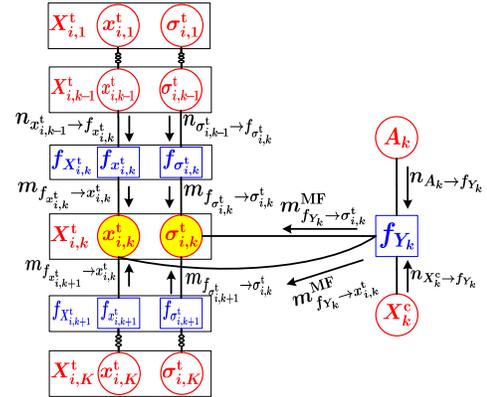


Fig. 2: The subgraph for target state estimation.

1) *The belief of target kinematic state:* The belief $b_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ can be derived as

$$b_{\mathbf{x}}(\mathbf{x}_{i,k}^t) \propto \underbrace{m_{f_{\mathbf{x}_{i,k}^t}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} m_{f_{Y_k} \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}}}}_{\vec{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)} \times \underbrace{m_{f_{\mathbf{x}_{i,k+1}^t}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}}}}_{\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)}. \quad (35)$$

Initialize $\vec{b}_{\mathbf{x}}(\mathbf{x}_{i,1}^t) = \mathcal{N}(\hat{\mathbf{x}}_{i,1}^t; \hat{\Sigma}_{i,1|1}^t, \mathbf{P}_{i,1|1}^t)$ as a Gaussian distribution at time 1, the forward messages $\vec{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ then can be calculated from time 2 to time K . The factor-to-variable messages corresponding to $\vec{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ include

$$\begin{aligned} & m_{f_{\mathbf{x}_{i,k}^t}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} \\ & = \exp \left(\int n_{\mathbf{x}_{i,k-1}^t \rightarrow \mathbf{x}_{i,k}^t} \ln p(\mathbf{x}_{i,k}^t | \mathbf{x}_{i,k-1}^t) d\mathbf{x}_{i,k-1}^t \right), \end{aligned} \quad (36)$$

$$m_{f_{Y_k} \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} = \exp \left(\sum_{j=1}^{N_{M,k}} n_{\mathbf{A}_k \rightarrow \mathbf{Y}_k} \ln p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k}^t, \mathbf{A}_k) \right). \quad (37)$$

According to Eq. (35), we have $n_{\mathbf{x}_{i,k-1}^t \rightarrow f_{\mathbf{x}_{i,k}^t}} = \overrightarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k-1}^t)$ since the backward message $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k-1}^t)$ is not available yet, and $n_{\mathbf{A}_k \rightarrow f_{\mathbf{Y}_k}} = b_{\mathbf{A}}(a_{i,j,k}^t)$. Note that the belief $\overrightarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k-1}^t) = \mathcal{N}(\mathbf{x}_{i,k-1}^t; \hat{\mathbf{x}}_{i,k-1|k-1}^t, \mathbf{P}_{i,k-1|k-1}^t)$ is also a Gaussian distribution. Then, Eq. (36) and Eq. (37) can be derived as

$$m_{f_{\mathbf{x}_{i,k}^t}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} \propto \mathcal{N}(\mathbf{x}_{i,k}^t; \hat{\mathbf{x}}_{i,k|k-1}^t, \mathbf{P}_{i,k|k-1}^t), \quad (38)$$

$$m_{f_{\mathbf{Y}_k}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} = \prod_{j=1}^{N_{M,k}} \mathcal{N}(y_{j,k}; h(\mathbf{x}_{i,k}^t), \mathbf{R}_k)^{\hat{a}_{i,j,k}^t}, \quad (39)$$

where $\hat{\mathbf{x}}_{i,k|k-1}^t$ and $\mathbf{P}_{i,k|k-1}^t$ are calculated according to the transition PDF $p(\mathbf{x}_{i,k}^t | \mathbf{x}_{i,k-1}^t)$; $\hat{a}_{i,j,k}^t$ is the expectation of $a_{i,j,k}^t$ taken over $b_a(a_{i,j,k}^t)$. Eq. (39) can be derived as $m_{f_{\mathbf{Y}_k}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} = \mathcal{N}(\bar{\mathbf{y}}_{i,k}; h(\mathbf{x}_{i,k}^t), \bar{\mathbf{R}}_{i,k})$ with [39]

$$\bar{\mathbf{y}}_{i,k} = \frac{\sum_{j=1}^{N_{M,k}} \hat{a}_{i,j,k}^t \mathbf{y}_{j,k}}{1 - \hat{a}_{i,0,k}^t}, \quad \bar{\mathbf{R}}_{i,k} = \frac{\mathbf{R}_k}{1 - \hat{a}_{i,0,k}^t}. \quad (40)$$

Substitute Eq. (38) and Eq. (39) into Eq. (35), yielding

$$\overrightarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t) = \mathcal{N}(\mathbf{x}_{i,k}^t; \hat{\mathbf{x}}_{i,k|k}^t, \mathbf{P}_{i,k|k}^t), \quad (41)$$

where $\hat{\mathbf{x}}_{i,k|k}^t$ and $\mathbf{P}_{i,k|k}^t$ are calculated by the Unscented Kalman Filter (UKF) [43].

Next, the backward message $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ is calculated from time K to time 1. Again, we need initial message $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,K}^t)$. We take $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,K}^t) = 1$ for all $\mathbf{x}_{i,K}^t$. Then, $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ from time $K-1$ to time 1 can be derived as

$$\begin{aligned} \overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t) &= m_{f_{\mathbf{x}_{i,k+1}^t}^t \rightarrow \mathbf{x}_{i,k}^t}^{\text{MF}} \\ &= \exp\left(\int n_{\mathbf{x}_{i,k+1}^t \rightarrow f_{\mathbf{x}_{i,k}^t}^t} \ln p(\mathbf{x}_{i,k+1}^t | \mathbf{x}_{i,k}^t) d\mathbf{x}_{i,k+1}^t\right). \end{aligned} \quad (42)$$

We multiply $\overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ by $\overrightarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t)$ and use a nonlinear fixed-interval smoother, deriving that $b_{\mathbf{x}}(\mathbf{x}_{i,k}^t) \propto \overrightarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t) \overleftarrow{b}_{\mathbf{x}}(\mathbf{x}_{i,k}^t) \propto \mathcal{N}(\mathbf{x}_{i,k}^t; \hat{\mathbf{x}}_{i,k|K}^t, \mathbf{P}_{i,k|K}^t)$ with $\hat{\mathbf{x}}_{i,k|K}^t$ and $\mathbf{P}_{i,k|K}^t$ being calculated by the Unscented Rauch-Tung-Striebel Smoother (URTSS) [43].

2) *The belief of target mean SNR*: The belief $b_{\sigma}(\sigma_{i,k}^t)$ can be calculated as

$$b_{\sigma}(\sigma_{i,k}^t) \propto \underbrace{m_{f_{\sigma_{i,k}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}}}_{\overrightarrow{b}_{\sigma}(\sigma_{i,k}^t)} \underbrace{m_{f_{\mathbf{Y}_k}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}}}_{\overleftarrow{b}_{\sigma}(\sigma_{i,k}^t)} \times m_{f_{\sigma_{i,k+1}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}}. \quad (43)$$

Initializing $\overrightarrow{b}_{\sigma}(\sigma_{i,1}^t) = \mathcal{I}(\sigma_{i,1}^t; \alpha_{i,1|1}^t, \beta_{i,1|1}^t)$ as an IG distribution at time 1, the forward messages $\overrightarrow{b}_{\sigma}(\sigma_{i,k}^t)$ then can be calculated from time 2 to time K . The factor-to-variable messages corresponding to $\overrightarrow{b}_{\sigma}(\sigma_{i,k}^t)$ include

$$\begin{aligned} m_{f_{\sigma_{i,k}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} \\ = \exp\left(\int n_{\sigma_{i,k-1}^t \rightarrow f_{\sigma_{i,k}^t}^t} \ln p(\sigma_{i,k}^t | \sigma_{i,k-1}^t) d\sigma_{i,k-1}^t\right), \end{aligned} \quad (44)$$

$$m_{f_{\mathbf{Y}_k}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} = \exp\left(\sum_{j=1}^{N_{M,k}} n_{\mathbf{A}_k \rightarrow f_{\mathbf{Y}_k}^t} \ln p(m_{j,k} | \sigma_{i,k}^t, \mathbf{A}_k)\right). \quad (45)$$

We have $n_{\sigma_{i,k-1}^t \rightarrow f_{\sigma_{i,k}^t}^t} = \overrightarrow{b}_{\sigma}(\sigma_{i,k-1}^t)$ according to Proposition 1, $\overrightarrow{b}_{\sigma}(\sigma_{i,k-1}^t) = \mathcal{I}(\sigma_{i,k-1}^t; \alpha_{i,k-1|k-1}^t, \beta_{i,k-1|k-1}^t)$ is an IG distribution, and $n_{\mathbf{A}_k \rightarrow f_{\mathbf{Y}_k}^t} = b_{\mathbf{A}}(a_{i,j,k}^t)$. According to $p(\sigma_{i,k}^t | \sigma_{i,k-1}^t)$, we get $m_{f_{\sigma_{i,k}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} \propto \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|k-1}^t, \beta_{i,k|k-1}^t)$, where $\beta_{i,k|k-1}^t = \beta_{i,k-1|k-1}^t / u^t$ and $\alpha_{i,k}^t = (\alpha_{i,k-1|k-1}^t + u^t - 1) / u^t$. Eq. (45) is derived as

$$m_{f_{\mathbf{Y}_k}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} = \prod_{j=1}^{N_{M,k}} \mathcal{R}^d(m_{j,k}; \sigma_{i,k}^t, n)^{\hat{a}_{i,j,k}^t}, \quad (46)$$

with $\hat{a}_{i,j,k}^t$ being the expectation of $a_{i,j,k}^t$ taken over belief $b_{\mathbf{A}}(a_{i,j,k}^t)$. Then the belief $\overrightarrow{b}_{\sigma}(\sigma_{i,k}^t)$ can be derived as

$$\begin{aligned} \overrightarrow{b}_{\sigma}(\sigma_{i,k}^t) &\propto \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|k-1}^t, \beta_{i,k|k-1}^t) \\ &\times \prod_{j=1}^{N_{M,k}} \mathcal{R}^d(m_{j,k}; \sigma_{i,k}^t, n)^{\hat{a}_{i,j,k}^t} \\ &= \prod_{j=1}^{N_{M,k}} \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,j,k|k}^t, \beta_{i,j,k|k}^t)^{\hat{a}_{i,j,k}^t}, \end{aligned} \quad (47)$$

where $\beta_{i,j,k|k}^t = \beta_{i,k|k-1}^t + nm_{j,k}^2 - nd^2$ and $\alpha_{i,j,k|k}^t = \alpha_{i,k|k-1}^t + n$. By applying the product of the IG distributions, Eq. (47) is derived as $\overrightarrow{b}_{\sigma}(\sigma_{i,k}^t) = \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|k}^t, \beta_{i,k|k}^t)$ with

$$\beta_{i,k|k}^t = \frac{\sum_{j=1}^{N_{M,k}} \hat{a}_{i,j,k}^t \beta_{i,j,k|k}^t}{1 - \hat{a}_{i,0,k}^t}, \quad \alpha_{i,k|k}^t = \frac{\sum_{j=1}^{N_{M,k}} \hat{a}_{i,j,k}^t \alpha_{i,j,k|k}^t}{1 - \hat{a}_{i,0,k}^t}. \quad (48)$$

Next, the backward message $\overleftarrow{b}_{\sigma}(\sigma_{i,k}^t)$ is calculated from time K to time 1, starting with $\overleftarrow{b}_{\sigma}(\sigma_{i,K}^t) = 1$ for $\sigma_{i,K}^t$, which can be calculated as

$$\begin{aligned} \overleftarrow{b}_{\sigma}(\sigma_{i,k}^t) &= m_{f_{\sigma_{i,k+1}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} \\ &= \exp\left(\int m_{f_{\sigma_{i,k+1}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} \ln p(\sigma_{i,k+1}^t | \sigma_{i,k}^t) d\sigma_{i,k+1}^t\right). \end{aligned} \quad (49)$$

Note that $m_{f_{\sigma_{i,k+1}^t}^t \rightarrow \sigma_{i,k}^t}^{\text{MF}} = b_{\sigma}(\sigma_{i,k+1}^t)$, which is calculated by Eq. (43) and must be an IG distribution, i.e., $b_{\sigma}(\sigma_{i,k+1}^t) = \mathcal{I}(\sigma_{i,k+1}^t; \alpha_{i,k+1|K}^t, \beta_{i,k+1|K}^t)$. By the definition of $p(\sigma_{i,k+1}^t | \sigma_{i,k}^t)$, we have $\overleftarrow{b}_{\sigma}(\sigma_{i,k}^t) \propto \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|k+1}^t, \beta_{i,k|k+1}^t)$, which is also an IG distribution, where $\beta_{i,k|k+1}^t = \beta_{i,k+1|K}^t / u^t$ and $\alpha_{i,k|k+1}^t = (\alpha_{i,k+1|K}^t + u^t - 1) / u^t$.

Finally, the belief $b_{\sigma}(\sigma_{i,k}^t)$ is calculated by an IG smoother as $b_{\sigma}(\sigma_{i,k}^t) = \overrightarrow{b}_{\sigma}(\sigma_{i,k}^t) \overleftarrow{b}_{\sigma}(\sigma_{i,k}^t)$. By applying the product of IG, we have $b_{\sigma}(\sigma_{i,k}^t) = \mathcal{I}(\sigma_{i,k}^t; \alpha_{i,k|K}^t, \beta_{i,k|K}^t)$, with $\alpha_{i,k|K}^t = \alpha_{i,k|k}^t + \alpha_{i,k|k+1}^t + 1$ and $\beta_{i,k|K}^t = \beta_{i,k|k}^t + \beta_{i,k|k+1}^t$.

C. Derivations of Belief $b_{\mathbf{X}}(\mathbf{X}_{1:K}^c)$

Fig. 3 shows the clutter spatial state and mean CNR estimation subgraph corresponding to the belief $b_{\mathbf{X}}(\tilde{\mathbf{X}}_{\tau,1:K}^c)$ and $b_{\sigma}(\sigma_{\tau,1:K}^c)$, $\tau = 0, \dots, N_C$, which contains the variable nodes $\tilde{\mathbf{x}}_{\tau,k}^c$ and $\sigma_{\tau,k}^c$, $k = 1, \dots, K$.

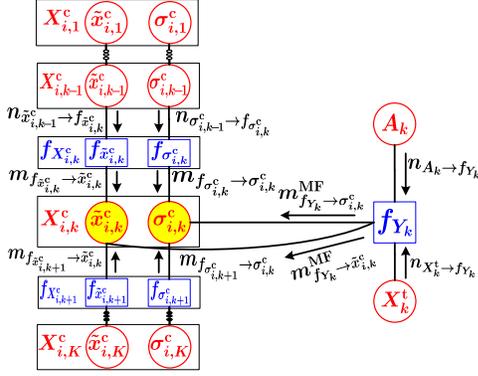


Fig. 3: The clutter estimation subgraph of $b_{\mathbf{x}^c}(\tilde{\mathbf{x}}_{\tau,k}^c)$.

1) *The belief of clutter spatial state:* The belief $b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ can be calculated as

$$b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) \propto \underbrace{m_{f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}}}_{\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)} m_{f_{Y_k} \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} \times \underbrace{m_{f_{\tilde{\mathbf{x}}_{\tau,k+1}^c}^c \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}}}_{\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)}. \quad (50)$$

We initialize $\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,1}^c)$ as a GW distribution at time 1, i.e., $\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,1}^c) \propto \mathcal{W}(\mathbf{D}_{\tau,1}^c; \mathbf{W}_{\tau,1|1}^c, v_{\tau,1|1}^c) \times \mathcal{N}(\mathbf{x}_{\tau,1}^c; \hat{\mathbf{x}}_{\tau,1|1}^c, (\beta_{\tau,1|1}^c \mathbf{D}_{\tau,1}^c)^{-1})$. Then the forward messages $\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ can be calculated from time 2 to time K , which consist of factor-to-variable messages as follows.

$$m_{f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} = \exp \left(\int n_{\tilde{\mathbf{x}}_{\tau,k-1}^c \rightarrow f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c} \ln p(\tilde{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c) d\tilde{\mathbf{x}}_{\tau,k-1}^c \right), \quad (51)$$

$$m_{f_{Y_k} \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} = \exp \left(\sum_{j=1}^{N_{M,k}} n_{A_k \rightarrow f_{Y_k}} \ln p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c, A_k) \right). \quad (52)$$

According to Proposition 1, we have $n_{\tilde{\mathbf{x}}_{\tau,k-1}^c \rightarrow f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c} = \vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k-1}^c)$ and $n_{A_k \rightarrow f_{Y_k}} = b_A(a_{\tau,j,k}^c)$. Similarly, the belief $\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k-1}^c)$ is also a GW distribution, given by,

$$\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k-1}^c) \propto \mathcal{W}(\mathbf{D}_{\tau,k-1}^c; \mathbf{W}_{\tau,k-1|k-1}^c, v_{\tau,k-1|k-1}^c) \times \mathcal{N}(\mathbf{x}_{\tau,k-1}^c; \hat{\mathbf{x}}_{\tau,k-1|k-1}^c, (\beta_{\tau,k-1|k-1}^c \mathbf{D}_{\tau,k-1}^c)^{-1}). \quad (53)$$

By the definition of $p(\tilde{\mathbf{x}}_{\tau,k}^c | \tilde{\mathbf{x}}_{\tau,k-1}^c)$, we get

$$m_{f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} \propto \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k|k-1}^c, v_{\tau,k|k-1}^c) \times \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k|k-1}^c, (\beta_{\tau,k|k-1}^c \mathbf{D}_{\tau,k}^c)^{-1}), \quad (54)$$

where $\hat{\mathbf{x}}_{\tau,k|k-1}^c = \hat{\mathbf{x}}_{\tau,k-1|k-1}^c$, $\beta_{\tau,k|k-1}^c = \beta_{\tau,k-1|k-1}^c$, $\mathbf{W}_{\tau,k|k-1}^c = \xi \mathbf{W}_{\tau,k-1|k-1}^c$, $v_{\tau,k|k-1}^c = \xi(v_{\tau,k-1|k-1}^c - m - 1) + m + 1$. Eq. (52) is calculated as

$$m_{f_{Y_k} \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} = \prod_{j=1}^{N_{M,k}} \mathcal{N}(\mathbf{y}_{j,k}; \mathbf{x}_{\tau,k}^c, (\mathbf{D}_{\tau,k}^c)^{-1})^{\hat{a}_{\tau,j,k}^c}, \quad (55)$$

where $\hat{a}_{\tau,j,k}^c$ is the expectation of $a_{\tau,j,k}^c$ taken over belief $b_A(a_{\tau,j,k}^c)$. Substituting Eq. (54) and Eq. (55) into Eq. (50) and according to [41], $\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ is calculated as

$$\vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) \propto \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k|k}^c, (\beta_{\tau,k|k}^c \mathbf{D}_{\tau,k}^c)^{-1}) \times \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k|k}^c, v_{\tau,k|k}^c), \quad (56)$$

where we define

$$\begin{aligned} \beta_{\tau,k|k}^c &= \beta_{\tau,k|k-1}^c + N_{\tau,k}^c, & v_{\tau,k|k}^c &= v_{\tau,k|k-1}^c + N_{\tau,k}^c, \\ \hat{\mathbf{x}}_{\tau,k|k}^c &= \frac{1}{\beta_{\tau,k}^c} (\beta_{\tau,k|k-1}^c \hat{\mathbf{x}}_{\tau,k|k-1}^c + N_{\tau,k}^c \bar{\mathbf{x}}_{\tau,k}^c), \\ (\mathbf{W}_{\tau,k|k}^c)^{-1} &= (\mathbf{W}_{\tau,k|k-1}^c)^{-1} + N_{\tau,k}^c \Xi_{\tau,k}^c + \\ &\frac{\beta_{\tau,k|k-1}^c N_{\tau,k}^c}{\beta_{\tau,k|k-1}^c + N_{\tau,k}^c} (\bar{\mathbf{x}}_{\tau,k}^c - \hat{\mathbf{x}}_{\tau,k|k-1}^c) (\bar{\mathbf{x}}_{\tau,k}^c - \hat{\mathbf{x}}_{\tau,k|k-1}^c)^T. \end{aligned} \quad (57)$$

In Eq. (57), we use three statistics of the measurement data set, given by

$$\begin{aligned} N_{\tau,k}^c &= \sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c, & \bar{\mathbf{x}}_{\tau,k}^c &= \frac{1}{N_{\tau,k}^c} \sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c \mathbf{y}_{j,k}, \\ \Xi_{\tau,k}^c &= \frac{1}{N_{\tau,k}^c} \sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c (\mathbf{y}_{j,k} - \bar{\mathbf{x}}_{\tau,k}^c) (\mathbf{y}_{j,k} - \bar{\mathbf{x}}_{\tau,k}^c)^T. \end{aligned} \quad (58)$$

Then, the backward message $\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ is calculated from time K to time 1, starting with $\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,K}^c) = 1$ for all $\tilde{\mathbf{x}}_{\tau,K}^c$ at time K , which can be calculated as

$$\begin{aligned} \overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) &= m_{f_{\tilde{\mathbf{x}}_{\tau,k+1}^c}^c \rightarrow \tilde{\mathbf{x}}_{\tau,k}^c}^{\text{MF}} \\ &= \exp \left(\int n_{\tilde{\mathbf{x}}_{\tau,k+1}^c \rightarrow f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c} \ln p(\tilde{\mathbf{x}}_{\tau,k+1}^c | \tilde{\mathbf{x}}_{\tau,k}^c) d\tilde{\mathbf{x}}_{\tau,k+1}^c \right), \end{aligned} \quad (59)$$

where $n_{\tilde{\mathbf{x}}_{\tau,k+1}^c \rightarrow f_{\tilde{\mathbf{x}}_{\tau,k}^c}^c} = b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k+1}^c)$ is a GW distribution, i.e., $\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k+1}^c) = \mathcal{N}(\mathbf{x}_{\tau,k+1}^c; \hat{\mathbf{x}}_{\tau,k+1|K}^c, (\beta_{\tau,k+1|K}^c \mathbf{D}_{\tau,k+1}^c)^{-1}) \times \mathcal{W}(\mathbf{D}_{\tau,k+1}^c; \mathbf{W}_{\tau,k+1|K}^c, v_{\tau,k+1|K}^c)$. Via the definition of $p(\tilde{\mathbf{x}}_{\tau,k+1}^c | \tilde{\mathbf{x}}_{\tau,k}^c)$, $\overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ is also a GW distribution

$$\begin{aligned} \overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) &= \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k|k+1}^c, (\beta_{\tau,k|k+1}^c \mathbf{D}_{\tau,k}^c)^{-1}) \\ &\quad \times \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k|k+1}^c, v_{\tau,k|k+1}^c), \end{aligned} \quad (60)$$

where $\hat{\mathbf{x}}_{\tau,k|k+1}^c = \hat{\mathbf{x}}_{\tau,k+1|K}^c$, $\beta_{\tau,k|k+1}^c = \beta_{\tau,k+1|K}^c$, $\mathbf{W}_{\tau,k|k+1}^c = \xi \mathbf{W}_{\tau,k+1|K}^c$, $v_{\tau,k|k+1}^c = \xi(v_{\tau,k+1|K}^c - m - 1) + m + 1$.

Finally, the belief $b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$ is calculated by a GW smoother as $b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) \propto \vec{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) \overleftarrow{b}_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c)$. By applying the product of GW distribution, we have

$$\begin{aligned} b_{\mathbf{x}}(\tilde{\mathbf{x}}_{\tau,k}^c) &= \mathcal{N}(\mathbf{x}_{\tau,k}^c; \hat{\mathbf{x}}_{\tau,k|K}^c, (\beta_{\tau,k|K}^c \mathbf{D}_{\tau,k}^c)^{-1}) \\ &\quad \times \mathcal{W}(\mathbf{D}_{\tau,k}^c; \mathbf{W}_{\tau,k|K}^c, v_{\tau,k|K}^c), \end{aligned} \quad (61)$$

where $\beta_{\tau,k|K}^c = \beta_{\tau,k|k}^c + \beta_{\tau,k|k+1}^c$, $\hat{\mathbf{x}}_{\tau,k|K}^c = (\beta_{\tau,k|k}^c \hat{\mathbf{x}}_{\tau,k|k}^c + \beta_{\tau,k|k+1}^c \hat{\mathbf{x}}_{\tau,k|k+1}^c) / \beta_{\tau,k|K}^c$, $(\mathbf{W}_{\tau,k|K}^c)^{-1} = (\mathbf{W}_{\tau,k|k}^c)^{-1} + (\mathbf{W}_{\tau,k|k+1}^c)^{-1}$ and $v_{\tau,k|K}^c = v_{\tau,k|k}^c + v_{\tau,k|k+1}^c - m - 1$.

2) *The belief of clutter mean CNR*: The belief $b_\sigma(\sigma_{\tau,k}^c)$ is derived by

$$b_\sigma(\sigma_{\tau,k}^c) \propto \underbrace{m_{f_{\sigma_{\tau,k}^c}^c \rightarrow \sigma_{\tau,k}^c}^{\text{MF}}}_{\vec{b}_\sigma(\sigma_{\tau,k}^c)} m_{f_{\mathbf{Y}_k} \rightarrow \sigma_{\tau,k}^c}^{\text{MF}} \times \underbrace{m_{f_{\sigma_{\tau,k+1}^c}^c \rightarrow \sigma_{\tau,k}^c}^{\text{MF}}}_{\overleftarrow{b}_\sigma(\sigma_{\tau,k}^c)}. \quad (62)$$

Using the same initialisation and MP relus as target mean SNR estimation, the forward message $\vec{b}_\sigma(\sigma_{\tau,k}^c)$ can be calculated as $\vec{b}_\sigma(\sigma_{\tau,k}^c) \propto \mathcal{I}(\sigma_{\tau,k}^c; \alpha_{\tau,k|k}^c, \beta_{\tau,k|k}^c)$, where

$$\beta_{\tau,k|k}^c = \frac{\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c \beta_{\tau,j,k|k}^c}{1 - \hat{a}_{\tau,0,k}^c}, \alpha_{\tau,k|k}^c = \frac{\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c \alpha_{\tau,j,k|k}^c}{1 - \hat{a}_{\tau,0,k}^c}, \quad (63)$$

where $\beta_{\tau,j,k|k}^c = \beta_{\tau,k|k-1}^c + nm_{j,k}^2 - nd^2$, $\alpha_{\tau,j,k|k}^c = \alpha_{\tau,k|k-1}^c + n$.

Again, referring to target mean SNR estimation, the backward message $\overleftarrow{b}_\sigma(\sigma_{\tau,k}^c)$ can be calculated as $\overleftarrow{b}_\sigma(\sigma_{\tau,k}^c) \propto \mathcal{I}(\sigma_{\tau,k}^c; \alpha_{\tau,k|k+1}^c, \beta_{\tau,k|k+1}^c)$, where $\beta_{\tau,k|k+1}^c = \beta_{\tau,k+1|K}^c / u^c$ and $\alpha_{\tau,k|k+1}^c = (\alpha_{\tau,k+1|K}^c + u^c - 1) / u^c$.

Finally, we have $b_\sigma(\sigma_{\tau,k}^c) = \mathcal{I}(\sigma_{\tau,k}^c; \alpha_{\tau,k|K}^c, \beta_{\tau,k|K}^c)$ by an IG smoother, where $\alpha_{\tau,k|K}^c = \alpha_{\tau,k|k}^c + \alpha_{\tau,k|k+1}^c + 1$ and $\beta_{\tau,k|K}^c = \beta_{\tau,k|k}^c + \beta_{\tau,k|k+1}^c$.

D. Derivations of Belief $b_S(\mathbf{S}_{1:K})$

Fig. 4 shows the subgraph for the estimation of target visibility state $s_{i,k}$, $k = 1, \dots, K$. The belief $b_s(s_{i,k})$ can

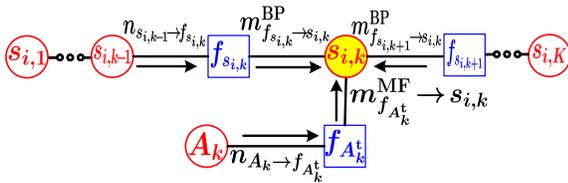


Fig. 4: The subgraph for target visibility state estimation.

be calculated as

$$b_s(s_{i,k}) \propto \underbrace{m_{f_{s_{i,k}} \rightarrow s_{i,k}}^{\text{BP}}}_{\vec{b}_s(s_{i,k})} m_{f_{A_k} \rightarrow s_{i,k}}^{\text{BP}} \times \underbrace{m_{f_{s_{i,k+1}} \rightarrow s_{i,k}}^{\text{BP}}}_{\overleftarrow{b}_s(s_{i,k})}. \quad (64)$$

Initializing $\vec{b}_s(s_{i,1})$ as a Bernoulli distribution at time 1, the forward messages $\vec{b}_s(s_{i,k})$ then can be calculated from time 2 to time K , which consist of factor-to-variable messages

$$m_{f_{s_{i,k}} \rightarrow s_{i,k}}^{\text{BP}} = \sum_{s_{i,k-1}=0}^1 p(s_{i,k} | s_{i,k-1}) n_{s_{i,k-1} \rightarrow f_{s_{i,k}}} \quad (65)$$

$$= \mathbf{T}_k \vec{b}_s(s_{i,k-1}),$$

$$m_{f_{A_k} \rightarrow s_{i,k}}^{\text{BP}} = \sum_{a_{i,0,k}^t=0}^1 p(a_{i,0,k}^t | s_{i,k}) n_{a_{i,0,k}^t \rightarrow f_{A_k}} = \boldsymbol{\xi}_k(s_{i,k}), \quad (66)$$

where $n_{a_{i,0,k}^t \rightarrow f_{A_k}} = b_a(a_{i,0,k})$. Substituting Eq. (65) and Eq. (66) into Eq. (64), the belief $\vec{b}_s(s_{i,k})$ is written as

$\vec{b}_s(s_{i,k}) = \mathbf{T}_k b_s(s_{i,k-1}) \boldsymbol{\xi}_k(s_{i,k})$. The belief $\vec{b}_s(s_{i,1:K})$ from time 1 to time K can be derived as

$$\vec{b}_s(s_{i,1:K}) = \prod_{k=1}^K \vec{b}_s(s_{i,k}) = \vec{b}_s(s_{i,1}) \boldsymbol{\xi}_1(s_{i,1}) \prod_{k=2}^K \mathbf{T}_k \boldsymbol{\xi}_k(s_{i,k}). \quad (67)$$

Next, the backward message $\overleftarrow{b}_s(s_{i,k})$ is calculated from time K to time 1, starting with $\overleftarrow{b}_s(s_{i,K}) = 1$ for all $s_{i,K}$, which is given by

$$\overleftarrow{b}_s(s_{i,k}) = m_{f_{s_{i,k+1}} \rightarrow s_{i,k}}^{\text{BP}} = \sum_{s_{i,k+1}=0}^1 m_{f_{s_{i,k+2}} \rightarrow s_{i,k+1}}^{\text{BP}} m_{f_{A_{k+1}} \rightarrow s_{i,k+1}}^{\text{BP}} p(s_{i,k+1} | s_{i,k}). \quad (68)$$

Thus, the belief $b_s(s_{i,1:K})$ can be recognized as a hidden Markov model (HMM) with observation sequence $\{\boldsymbol{\xi}_1(s_{i,1}), \dots, \boldsymbol{\xi}_1(s_{i,K})\}$, transition matrix \mathbf{T}_k and initial probability $\vec{b}_s(s_{i,1})$. We adopt the forward-backward algorithm [44] to estimate the belief $b_s(s_{i,1:K})$, resulting an HMM smoother. The track management can be achieved by comparing the belief $b_s(s_{i,k})$ with the track confirmation and deletion thresholds.

E. Derivations of Belief $b_\Pi(\Pi_{1:K})$

Fig. 5 shows the clutter mixing weights estimation subgraph corresponding to the belief $b_\pi(\pi_{\tau,1:K})$, $\tau = 0, \dots, N_T$, which contains the variable nodes $\pi_{\tau,k}$, $k = 1, \dots, K$. The belief

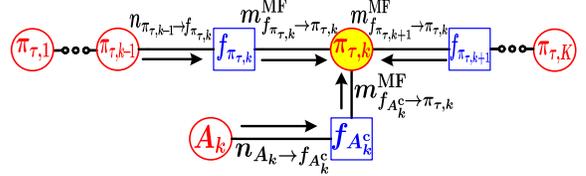


Fig. 5: The mixing weights estimation subgraph of $b_\Pi(\Pi_{1:K})$.

$b_\pi(\pi_{\tau,k})$ is given by

$$b_\pi(\pi_{\tau,k}) \propto \underbrace{m_{f_{\pi_{\tau,k}} \rightarrow \pi_{\tau,k}}^{\text{MF}}}_{\vec{b}_\pi(\pi_{\tau,k})} m_{f_{A_k} \rightarrow \pi_{\tau,k}}^{\text{MF}} \times \underbrace{m_{f_{\pi_{\tau,k+1}} \rightarrow \pi_{\tau,k}}^{\text{MF}}}_{\overleftarrow{b}_\pi(\pi_{\tau,k})}. \quad (69)$$

Initializing $\vec{b}_\pi(\pi_{\tau,1}) \propto \text{Dir}(\pi_{\tau,1}; \alpha_{\tau,1|1})$ as a Dirichlet distribution at time 1, the forward messages $\vec{b}_\pi(\pi_{\tau,k})$ then can be calculated from time 2 to time K , which consist of factor-to-variable messages

$$m_{f_{\pi_{\tau,k}} \rightarrow \pi_{\tau,k}}^{\text{MF}} = \int n_{\pi_{\tau,k-1} \rightarrow f_{\pi_{\tau,k}}} \ln p(\pi_{\tau,k} | \pi_{\tau,k-1}) d\pi_{\tau,k}, \quad (70)$$

$$m_{f_{A_k} \rightarrow \pi_{\tau,k}}^{\text{MF}} = \exp \left(\sum_{j=1}^{N_{M,k}} n_{A_k \rightarrow f_{A_k^c}} \ln p(a_{\tau,j,k}^c | \pi_{\tau,k}) \right). \quad (71)$$

Note that $n_{\pi_{\tau,k-1} \rightarrow f_{\pi_{\tau,k}}} = \vec{b}_\pi(\pi_{\tau,k-1})$ is a Dirichlet distribution, i.e., $\vec{b}_\pi(\pi_{\tau,k-1}) = \text{Dir}(\pi_{\tau,k-1}; \alpha_{\tau,k-1|k-1})$. By the definition of $p(\pi_{\tau,k} | \pi_{\tau,k-1})$, the predicted belief also follows a Dirichlet distribution, i.e., $m_{f_{\pi_{\tau,k}} \rightarrow \pi_{\tau,k}}^{\text{MF}} \propto \text{Dir}(\pi_{\tau,k}; \alpha_{\tau,k|k-1})$, where $\alpha_{\tau,k|k-1} = \kappa M_{k-1} \alpha_{\tau,k-1|k-1} / \sum_{\tau'=0}^{N_C} \alpha_{\tau',k-1|k-1}$

and M_{k-1} is the estimated total number of clutter at time $k-1$. Then, the message $m_{f_{A_k^c} \rightarrow \pi_{\tau,k}}^{\text{MF}}$ is calculated as

$$\begin{aligned} m_{f_{A_k^c} \rightarrow \pi_{\tau,k}}^{\text{MF}} &= \exp \left(\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c \ln \pi_{\tau,k} \right) \\ &= (\pi_{\tau,k})^{\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}^c}. \end{aligned} \quad (72)$$

Thus, the forward message $\vec{b}_{\pi}(\pi_{\tau,k})$ is calculated as

$$\begin{aligned} \vec{b}_{\pi}(\pi_{\tau,k}) &\propto \text{Dir}(\pi_{\tau,k-1}; \alpha_{\tau,k|k-1}) (\pi_{\tau,k})^{\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}} \\ &\propto (\pi_{\tau,k})^{\sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k} + \alpha_{\tau,k|k-1}}. \end{aligned} \quad (73)$$

We recognize $\vec{b}_{\pi}(\pi_{\tau,k}) = \text{Dir}(\pi_{\tau,k} | \alpha_{\tau,k|k})$ as a Dirichlet distribution with $\alpha_{\tau,k|k} = \sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k} + \alpha_{\tau,k|k-1}$.

Next, the backward message $\overleftarrow{b}_{\pi}(\pi_{\tau,k})$ is calculated from time K to time 1, starting with $\overleftarrow{b}_{\pi}(\pi_{\tau,K}) = 1$ for all $\pi_{\tau,K}$, which is given by

$$\begin{aligned} \overleftarrow{b}_{\pi}(\pi_{\tau,k}) &= m_{f_{\pi_{\tau,k+1}} \rightarrow \pi_{\tau,k}}^{\text{MF}} \\ &= \exp \left(\int m_{f_{\pi_{\tau,k+1}} \rightarrow \pi_{\tau,k}}^{\text{MF}} \ln p(\pi_{\tau,k+1} | \pi_{\tau,k}) d\pi_{\tau,k+1} \right). \end{aligned} \quad (74)$$

Note that $m_{f_{\pi_{\tau,k+1}} \rightarrow \pi_{\tau,k}}^{\text{MF}} = b_{\pi}(\pi_{\tau,k+1})$ is a Dirichlet distribution, given by $b_{\pi}(\pi_{\tau,k+1}) = \text{Dir}(\pi_{\tau,k+1}; \alpha_{\tau,k+1|K})$. By the definition of $p(\pi_{\tau,k+1} | \pi_{\tau,k})$, we have $\overleftarrow{b}_{\pi}(\pi_{\tau,k}) \propto \text{Dir}(\pi_{\tau,k} | \alpha_{\tau,k|k+1})$, which is also a Dirichlet distribution, where $\alpha_{\tau,k|k+1} = \kappa M_{k+1} \alpha_{\tau,k+1|K} / \sum_{\tau'=0}^{N_C} \alpha_{\tau',k+1|K}$, where M_{k+1} is the estimated total number of clutter at time $k+1$.

Finally, the belief $b_{\pi}(\pi_{\tau,k})$ is calculated by a Dirichlet smoother as $b_{\pi}(\pi_{\tau,k}) = \overleftarrow{b}_{\pi}(\pi_{\tau,k}) \overrightarrow{b}_{\pi}(\pi_{\tau,k})$. By applying the product of Dirichlet distribution, we have $b_{\pi}(\pi_{\tau,k}) = \text{Dir}(\pi_{\tau,k} | \alpha_{\tau,k|K})$, where $\alpha_{\tau,k|K} = \alpha_{\tau,k|k} + \alpha_{\tau,k|k+1}$. The estimated number of clutter component τ is $M_{\tau,k} = M_k \alpha_{\tau,k|K}$ with $M_k = \sum_{\tau=0}^{N_C} \sum_{j=1}^{N_{M,k}} \hat{a}_{\tau,j,k}$ being the estimated total number of clutter.

F. Derivations of Beliefs $b_{\mathbf{A}}(\mathbf{A}_{1:K})$

Fig. 6 and Fig. 7 show the data association global subgraph and the corresponding local MP subgraph for the belief of \mathbf{A}_k , respectively, which contains the variables nodes $a_{i,j,k}^t$ and $a_{\tau,j,k}^c$, $i = 1, \dots, N_T$, $\tau = 0, \dots, N_C$, $j = 0, \dots, N_{M,k}$, $k = 1, \dots, K$. The belief $b_{\mathbf{A}}(\mathbf{A}_k)$ can be computed as

$$\begin{aligned} b_{\mathbf{A}}(\mathbf{A}_k) &\propto m_{f_{\mathbf{Y}_k} \rightarrow \mathbf{A}_k}^{\text{MF}} \times m_{f_{\mathbf{A}_k^t} \rightarrow \mathbf{A}_k}^{\text{BP}} \times m_{f_{\mathbf{A}_k^c} \rightarrow \mathbf{A}_k}^{\text{BP}} \\ &\quad \times m_{f_{\mathbf{I}_k} \rightarrow \mathbf{A}_k}^{\text{BP}} \times m_{f_{\mathbf{E}_k} \rightarrow \mathbf{A}_k}^{\text{BP}}. \end{aligned} \quad (75)$$

The belief $m_{f_{\mathbf{Y}_k} \rightarrow \mathbf{A}_k}^{\text{MF}}$ is computed as

$$m_{f_{\mathbf{Y}_k} \rightarrow \mathbf{A}_k}^{\text{BP}} = \prod_{j=0}^{N_{M,k}} \prod_{i=1}^{N_T} m_{f_{\mathbf{Y}_k} \rightarrow a_{i,j,k}^t}^{\text{MF}} \prod_{\tau=0}^{N_C} m_{f_{\mathbf{Y}_k} \rightarrow a_{\tau,j,k}^c}^{\text{MF}}. \quad (76)$$

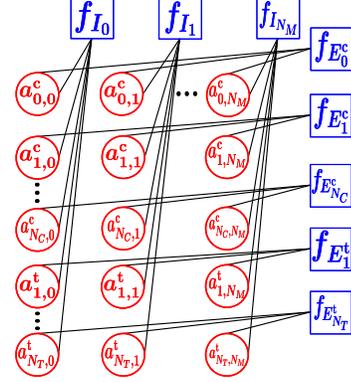


Fig. 6: The global data association subgraph of $b_{\mathbf{A}}(\mathbf{A}_k)$. The time index k is omitted for the simplicity. In addition, each variable node $a_{i,j,k}^t$ ($a_{\tau,j,k}^c$) also connects with the factor nodes $f_{\mathbf{y}_{j,k}}$ and $f_{\mathbf{A}_k^t}$ ($f_{\mathbf{A}_k^c}$) as in Fig. 7, which are not shown here for the simplicity.

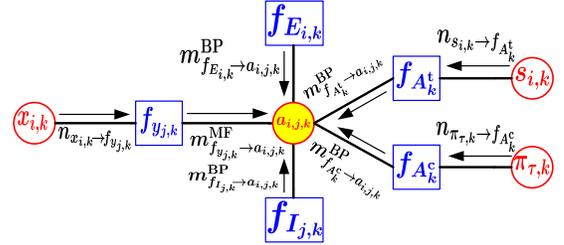


Fig. 7: The local subgraph of data association. The $\mathbf{x}_{i,k}$, $\mathbf{E}_{i,k}$ and $a_{i,j,k}$ represent $\mathbf{x}_{i,k}^t$ ($\mathbf{x}_{i,k}^t$), $\mathbf{E}_{i,k}^t$ ($\mathbf{E}_{i,k}^t$) and $a_{i,j,k}^t$ ($a_{\tau,j,k}^c$) when calculate the association between measurements and targets (clutter).

The message $m_{f_{\mathbf{Y}_k} \rightarrow a_{i,j,k}^t}^{\text{MF}}$ is calculated as

$$\begin{aligned} m_{f_{\mathbf{Y}_k} \rightarrow a_{i,j,k}^t}^{\text{MF}} &= \exp \int n_{\mathbf{X}_{i,k}^t \rightarrow f_{\mathbf{Y}_k}} \ln p(\mathbf{Y}_{j,k} | \mathbf{X}_{i,k}^t) a_{i,j,k}^t d\mathbf{X}_{i,k}^t \\ &= \exp \left(\int n_{\mathbf{x}_{i,k}^t \rightarrow f_{\mathbf{Y}_k}} \ln p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k}^t) a_{i,j,k}^t d\mathbf{x}_{i,k}^t + \int n_{\sigma_{i,k}^t \rightarrow f_{\mathbf{Y}_k}} \ln p(m_{j,k}; \sigma_{i,k}^t) a_{i,j,k}^t d\sigma_{i,k}^t \right). \end{aligned} \quad (77)$$

Since $n_{\mathbf{x}_{i,k}^t \rightarrow f_{\mathbf{Y}_k}} = b(\mathbf{x}_{i,k}^t)$ and $n_{\sigma_{i,k}^t \rightarrow f_{\mathbf{Y}_k}} = b(\sigma_{i,k}^t)$, Eq. (77) can be calculated as

$$\begin{aligned} m_{f_{\mathbf{Y}_k} \rightarrow a_{i,j,k}^t}^{\text{MF}} &= \exp [a_{i,j,k}^t (\mathbb{E}[\ln p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k}^t)] + \mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^t)])], \end{aligned} \quad (78)$$

where,

$$\begin{aligned} \mathbb{E}[\ln p(\mathbf{y}_{j,k} | \mathbf{x}_{i,k}^t)] &\stackrel{\text{c}}{=} \mathbb{E}[(\mathbf{y}_{j,k} - \mathbf{H}\mathbf{x}_{i,k}^t)^T \mathbf{R}_k^{-1} (\mathbf{y}_{j,k} - \mathbf{H}\mathbf{x}_{i,k}^t)] \\ &\stackrel{\text{c}}{=} \text{Tr} \left(\mathbf{R}_k^{-1} (\mathbf{H}\mathbf{P}_{i,k}^t \mathbf{H}^T + (\mathbf{H}\hat{\mathbf{x}}_{i,k}^t - \mathbf{y}_{j,k})(\mathbf{H}\hat{\mathbf{x}}_{i,k}^t - \mathbf{y}_{j,k})^T) \right), \end{aligned} \quad (79)$$

and $\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^t)]$ can be calculated as (The details are

shown in Appendix V-A),

$$\begin{aligned} & \mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^t)] \\ &= (2n-1) \ln m_{j,k} - nm_{j,k}^2 \frac{\alpha_{i,k}^t}{\beta_{i,k}^t} + \frac{n}{2\alpha_{i,k}^t - 4}. \end{aligned} \quad (80)$$

The message $m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}}$ is calculated as,

$$\begin{aligned} & m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}} \\ &= \exp \int n_{\mathbf{X}_{\tau,k}^c \rightarrow f_{\mathbf{Y}_k}} \ln p(\mathbf{Y}_{j,k} | \mathbf{X}_{\tau,k}^c)^{a_{\tau,j,k}^c} d\mathbf{X}_{\tau,k}^c \\ &= \exp \left(\int n_{\tilde{\mathbf{x}}_{\tau,k}^c \rightarrow f_{\mathbf{Y}_k}} \ln p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c)^{a_{\tau,j,k}^c} d\tilde{\mathbf{x}}_{\tau,k}^c + \int n_{\sigma_{\tau,k}^c \rightarrow f_{\mathbf{Y}_k}} \ln p(m_{j,k}; \sigma_{\tau,k}^c)^{a_{\tau,j,k}^c} d\sigma_{\tau,k}^c \right). \end{aligned} \quad (81)$$

With $n_{\tilde{\mathbf{x}}_{\tau,k}^c \rightarrow f_{\mathbf{Y}_k}} = b(\tilde{\mathbf{x}}_{\tau,k}^c)$ and $n_{\sigma_{\tau,k}^c \rightarrow f_{\mathbf{Y}_k}} = b(\sigma_{\tau,k}^c)$, Eq. (81) can be derived as

$$\begin{aligned} & m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}} \\ &= \exp \left[a_{\tau,j,k}^c \left(\mathbb{E}[\ln p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c)] + \mathbb{E}[\ln p(m_{j,k}; \sigma_{\tau,k}^c)] \right) \right], \end{aligned} \quad (82)$$

where $\mathbb{E}[\ln p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c)] = \ln(1/V_G)$ if $\tau = 0$; otherwise,

$$\begin{aligned} & \mathbb{E}[\ln p(\mathbf{y}_{j,k} | \tilde{\mathbf{x}}_{\tau,k}^c)] = -\frac{D}{2} \ln 2\pi + \frac{1}{2} \mathbb{E}[\ln |\mathbf{D}_{\tau,k}^c|] \\ & \quad - \frac{1}{2} \mathbb{E}[(\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c)^T \mathbf{D}_{\tau,k}^c (\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c)], \end{aligned} \quad (83)$$

where,

$$\mathbb{E}[\ln |\mathbf{D}_{\tau,k}^c|] = \sum_{j=1}^m \psi \left(\frac{v_{\tau,k}^c + 1 - j}{2} \right) + m \ln 2 + \ln |\mathbf{W}_{\tau,k}^c|, \quad (84)$$

$$\begin{aligned} & \mathbb{E}[(\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c)^T \mathbf{D}_{\tau,k}^c (\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c)] \\ &= m \beta_{\tau,k}^{c-1} + v_{\tau,k}^c (\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c)^T \mathbf{W}_{\tau,k}^c (\mathbf{y}_{j,k} - \tilde{\mathbf{x}}_{\tau,k}^c), \end{aligned} \quad (85)$$

where $\psi(a) = d \ln \Gamma(a) / da$ is the digamma function [41]. Then, $\mathbb{E}[\ln p(m_{j,k}; \sigma_{\tau,k}^c)]$ is derived as (The details are given in Appendix V-A),

$$\begin{aligned} & \mathbb{E}[\ln p(m_{j,k}; \sigma_{\tau,k}^c)] \\ &= (2n-1) \ln m_{j,k} - nm_{j,k}^2 \frac{\alpha_{\tau,k}^c}{\beta_{\tau,k}^c} + \frac{n}{2\alpha_{\tau,k}^c - 4}. \end{aligned} \quad (86)$$

The belief $m_{f_{A_k^t} \rightarrow A_k}^{\text{BP}}$ is calculated as

$$m_{f_{A_k^t} \rightarrow A_k}^{\text{BP}} = \prod_{i=1}^{N_T} m_{f_{A_k^t} \rightarrow a_{i,0,k}^t}^{\text{BP}}, \quad (87)$$

where,

$$\begin{aligned} & m_{f_{A_k^t} \rightarrow a_{i,0,k}^t}^{\text{BP}} = \sum_{s_{i,k}=0}^1 p(a_{i,0,k}^t | s_{i,k}) n_{s_{i,k} \rightarrow f_{A_k^t}} \\ &= \begin{cases} \sum_{s_{i,k}=0}^1 (1 - P_d(s_{i,k})) m_{f_{s_{i,k}} \rightarrow s_{i,k}}^{\text{BP}}, & \text{if } a_{i,0,k}^t = 1, \\ \sum_{s_{i,k}=0}^1 P_d(s_{i,k}) m_{f_{s_{i,k}} \rightarrow s_{i,k}}^{\text{BP}}, & \text{otherwise.} \end{cases} \end{aligned} \quad (88)$$

The belief $m_{f_{A_k^c} \rightarrow A_k}^{\text{MF}}$ is calculated as

$$m_{f_{A_k^c} \rightarrow A_k}^{\text{MF}} = \prod_{\tau=1}^{N_C} m_{f_{A_k^c} \rightarrow a_{\tau,0,k}^c}^{\text{MF}}, \quad (89)$$

where,

$$m_{f_{A_k^c} \rightarrow a_{\tau,j,k}^c}^{\text{MF}} = \exp \int n_{\pi_{\tau,k} \rightarrow f_{A_k^c}} \ln p(a_{\tau,j,k}^c | \pi_{\tau,k}) d\pi_{\tau,k}. \quad (90)$$

Note that $n_{\pi_{\tau,k} \rightarrow f_{A_k^c}} = b_{\pi}(\pi_{\tau,k})$. We have

$$\begin{aligned} & m_{f_{A_k^c} \rightarrow a_{\tau,j,k}^c}^{\text{MF}} = \exp \left(a_{\tau,j,k}^c \mathbb{E}[\ln \pi_{\tau,k}] \right) \\ &= \exp \left(a_{\tau,j,k}^c \Gamma(\alpha_{\tau,k}^c) - \Gamma \left(\sum_{\tau'=1}^{N_C} \alpha_{\tau',k}^c \right) \right). \end{aligned} \quad (91)$$

The belief $m_{f_{I_k}^{\text{BP}} \rightarrow A_k}$ is calculated as

$$m_{f_{I_k}^{\text{BP}} \rightarrow A_k} = \prod_{j=1}^{N_{M,k}} \prod_{i=1}^{N_T} m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t} \prod_{\tau=0}^{N_C} m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c}. \quad (92)$$

The belief $m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t}$ and $m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c}$ are derived as

$$\begin{aligned} & m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t} = \sum_{\mathbf{A}_{j,k} \setminus \{a_{i,j,k}^t\}} f_{I_{j,k}}(\mathbf{A}_{j,k}) \\ & \quad \times \prod_{i'=1 \setminus i}^{N_T} n_{a_{i',j,k}^t \rightarrow f_{I_{j,k}}} \prod_{\tau=0}^{N_C} n_{a_{\tau,j,k}^c \rightarrow f_{I_{j,k}}}, \end{aligned} \quad (93)$$

$$\begin{aligned} & m_{f_{I_{j,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c} = \sum_{\mathbf{A}_{j,k} \setminus \{a_{\tau,j,k}^c\}} f_{I_{j,k}}(\mathbf{A}_{j,k}) \\ & \quad \times \prod_{i=1}^{N_T} n_{a_{i,j,k}^t \rightarrow f_{I_{j,k}}} \prod_{\tau'=0 \setminus \tau}^{N_C} n_{a_{\tau',j,k}^c \rightarrow f_{I_{j,k}}}, \end{aligned} \quad (94)$$

where $n_{a_{i,j,k}^t \rightarrow f_{I_{j,k}}}$ and $n_{a_{\tau,j,k}^c \rightarrow f_{I_{j,k}}}$ are

$$\begin{aligned} & n_{a_{i,j,k}^t \rightarrow f_{I_{j,k}}} = m_{f_{\mathbf{Y}_k}^c \rightarrow a_{i,j,k}^t}^{\text{MF}} m_{f_{A_k^t} \rightarrow a_{i,j,k}^t}^{\text{BP}} m_{f_{E_k^t} \rightarrow a_{i,j,k}^t}^{\text{BP}}, \\ & n_{a_{\tau,j,k}^c \rightarrow f_{I_{j,k}}} = m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}} m_{f_{A_k^c} \rightarrow a_{\tau,j,k}^c}^{\text{MF}} m_{f_{E_k^c} \rightarrow a_{\tau,j,k}^c}^{\text{BP}}. \end{aligned} \quad (95)$$

The belief $m_{f_{E_k}^{\text{BP}} \rightarrow A_k}$ is calculated as

$$m_{f_{E_k}^{\text{BP}} \rightarrow A_k} = \prod_{j=0}^{N_{M,k}} \prod_{i=1}^{N_T} m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t} \prod_{\tau=0}^{N_C} m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c}. \quad (96)$$

The belief $m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t}$ and $m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c}$ are calculated as

$$\begin{aligned} & m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{i,j,k}^t} = \sum_{\mathbf{A}_{i,k}^t \setminus \{a_{i,j,k}^t\}} f_{E_{i,k}}^t(\mathbf{A}_{i,k}^t) \prod_{j'=1 \setminus j}^{N_{M,k}} n_{a_{i,j',k}^t \rightarrow f_{E_{i,k}}^t}, \\ & m_{f_{E_{i,k}}^{\text{BP}} \rightarrow a_{\tau,j,k}^c} = \sum_{\mathbf{A}_{\tau,k}^c \setminus \{a_{\tau,j,k}^c\}} f_{E_{i,k}}^c(\mathbf{A}_{\tau,k}^c) \prod_{j'=1 \setminus j}^{N_{M,k}} n_{a_{\tau,j',k}^c \rightarrow f_{E_{i,k}}^c}, \end{aligned} \quad (97)$$

where $n_{a_{i,j,k}^t \rightarrow f_{E_{i,k}}^t}$ and $n_{a_{\tau,j,k}^c \rightarrow f_{E_{i,k}}^c}$ are

$$\begin{aligned} & n_{a_{i,j,k}^t \rightarrow f_{E_{i,k}}^t} = m_{f_{\mathbf{Y}_k}^c \rightarrow a_{i,j,k}^t}^{\text{MF}} m_{f_{A_k^t} \rightarrow a_{i,j,k}^t}^{\text{BP}} m_{f_{I_k}^{\text{BP}} \rightarrow a_{i,j,k}^t}, \\ & n_{a_{\tau,j,k}^c \rightarrow f_{E_{i,k}}^c} = m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}} m_{f_{A_k^c} \rightarrow a_{\tau,j,k}^c}^{\text{MF}} m_{f_{I_k}^{\text{BP}} \rightarrow a_{\tau,j,k}^c}. \end{aligned} \quad (98)$$

By applying some mathematical tricks in Appendix V-B, we obtain the simplified messages as follows,

$$\beta_{i,j,k}^t = \frac{\theta_{i,j,k}^t}{\theta_{0,j,k}^c + \sum_{i'=1 \setminus i}^{N_T} \theta_{i',j,k}^t \eta_{i',j,k}^t + \sum_{\tau=1}^{N_C} \theta_{\tau,j,k}^c \eta_{\tau,j,k}^c}, \quad (99)$$

$$\beta_{\tau,j,k}^c = \frac{\theta_{\tau,j,k}^c}{\theta_{0,j,k}^c + \sum_{i=1}^{N_T} \theta_{i,j,k}^t \eta_{i,j,k}^t + \sum_{\tau'=1 \setminus \tau}^{N_C} \theta_{\tau',j,k}^c \eta_{\tau',j,k}^c}, \quad (100)$$

$$\eta_{i,j,k}^t = \frac{1}{\sum_{j'=0 \setminus j}^{N_{M,k}} \beta_{i,j',k}^t}, \quad \eta_{\tau,0,k}^c = \frac{1}{\prod_{j'=1}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}, \quad (101)$$

$$\eta_{\tau,j,k}^c = \frac{\prod_{j'=1 \setminus j}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}{\beta_{\tau,0,k}^c + \prod_{j'=1 \setminus j}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}, \quad (102)$$

where $\beta_{i,j,k}^t$, $\beta_{\tau,j,k}^c$, $\theta_{i,j,k}^t$, $\theta_{\tau,j,k}^c$, $\eta_{i,j,k}^t$, $\eta_{\tau,j,k}^c$ are defined in Eq. (110) and Eq. (112). The values of all messages $\beta_{i,j,k}^t$, $\beta_{\tau,j,k}^c$, $\eta_{i,j,k}^t$, and $\eta_{\tau,j,k}^c$ are initialized as one and updated via Eq. (99)-Eq. (102). The message-updating procedure is terminated after the difference between two successive messages is less than δ_T^{BP} or a maximum number of iterations $r_{\text{max}}^{\text{BP}}$ is reached. The data association may have multiple local optimal solutions corresponding to the multiple fixed points of the message update rules, which may prevent convergence. To this end, we introduce a damping to the message update, i.e., $\mu = \gamma \mu^{\text{old}} + (1 - \gamma) \mu^{\text{new}}$, where μ denote a message and γ is a damping factor. A higher value of γ leads to a slower convergence rate but often with more stable convergence.

The proposed hybrid data association algorithm can be viewed as a generalization of existing data association algorithms solved with BP, which contain the one-to-one constraints that model one target generates at most one measurement and one-to-multi constraints that model one clutter component potentially generates a large number of measurements at one scan. Specifically, the hybrid data associations degenerates into one-to-one data associations [24], [28]–[30] when only targets exist and degenerates into one-to-multi data associations [31], [32] when only clutter exist. Different from the algorithms in [24], [28]–[32] that use the two multimodal random variables to represent a redundant formulation of data association, instead as in [37]–[39], the data association event is modelled as a binary random variable and the corresponding constraints are established, then the marginal association PDFs are inferred by BP.

G. Initialization, Implementation, Computational Complexity

We introduce a scheme for initializing the beliefs of target joint augmented state $\mathbf{X}_{1:K}^t$, target visibility state $\mathbf{S}_{1:K}$, clutter joint augmented state $\mathbf{X}_{1:K}^c$, and clutter mixing weights $\mathbf{\Pi}_{1:K}$. The target track is initialized by a two-point method. We assume that the maximum speed of any target in Cartesian coordinates is less than v_{max} , and the number of consecutive missing measurements of any tracks is less than L_{max} . Suppose that there is a measurement generated from a target at scan k , then the measurement can be used to initialize a potential track with another measurement at scan $k + \delta$, $0 < \delta \leq L_{\text{max}}$ within distance δv_{max} . We initialize the IG distributions of target mean SNR and clutter mean CNR using the maximum likelihood estimation, and let $\alpha = 3$ and $\beta = 2 \sum_{n=1}^N m_n^2 / N$, where N is the total number of initial measurements generated from the corresponding target or clutter and m_n is the strength of measurement n . The belief of target visibility state is initialized as $b_s(s_{i,1} = 1) = f_s$ with

f_s as the initial target visibility probability. The parameters of the spatial distribution of the nonuniform clutter component, i.e., GW, are initialized by clustering technologies, such as variational mixture of Gaussians clustering [41]. Given that the number of measurements generated by each clutter component varies slowly, we initialize the clutter mixing weight parameter $\alpha_{\tau,1}$ with the initial mixing weights of clutter component τ .

Given the initial beliefs on $\mathbf{X}_{1:K}^t$, $\mathbf{S}_{1:K}$, $\mathbf{X}_{1:K}^c$, and $\mathbf{\Pi}_{1:K}$, we can iteratively calculate $b_{\mathbf{X}}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{S}}(\mathbf{S}_{1:K})$, $b_{\mathbf{X}}(\mathbf{X}_{1:K}^c)$, $b_{\mathbf{\Pi}}(\mathbf{\Pi}_{1:K})$, and $b_{\mathbf{A}}(\mathbf{A}_{1:K})$ in principle by running MP described in this section. The proposed MP algorithm is illustrated in Fig. 8 and summarised as **Algorithm 1**, which are explained as follows. Firstly, initialize the belief $b_{\mathbf{X}}^{(0)}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{S}}^{(0)}(\mathbf{S}_{1:K})$, $b_{\mathbf{X}}^{(0)}(\mathbf{X}_{1:K}^c)$, and $b_{\mathbf{\Pi}}^{(0)}(\mathbf{\Pi}_{1:K})$. At the l th iteration, the belief $b_{\mathbf{A}}^{(l)}(\mathbf{A}_{1:K}^c)$ is inferred using the incoming messages $b_{\mathbf{X}}^{(l-1)}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{S}}^{(l-1)}(\mathbf{S}_{1:K})$, $b_{\mathbf{X}}^{(l-1)}(\mathbf{X}_{1:K}^c)$, and $b_{\mathbf{\Pi}}^{(l-1)}(\mathbf{\Pi}_{1:K})$ calculated in the $(l-1)$ th iteration, as the red lines in Fig. 8. Then $b_{\mathbf{A}}^{(l)}(\mathbf{A}_{1:K}^c)$ is used to infer $b_{\mathbf{X}}^{(l)}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{S}}^{(l)}(\mathbf{S}_{1:K})$, $b_{\mathbf{X}}^{(l)}(\mathbf{X}_{1:K}^c)$, and $b_{\mathbf{\Pi}}^{(l)}(\mathbf{\Pi}_{1:K})$, as the black lines in Fig. 8. The algorithm terminates when the MP converges or when the maximum number of iterations is reached.

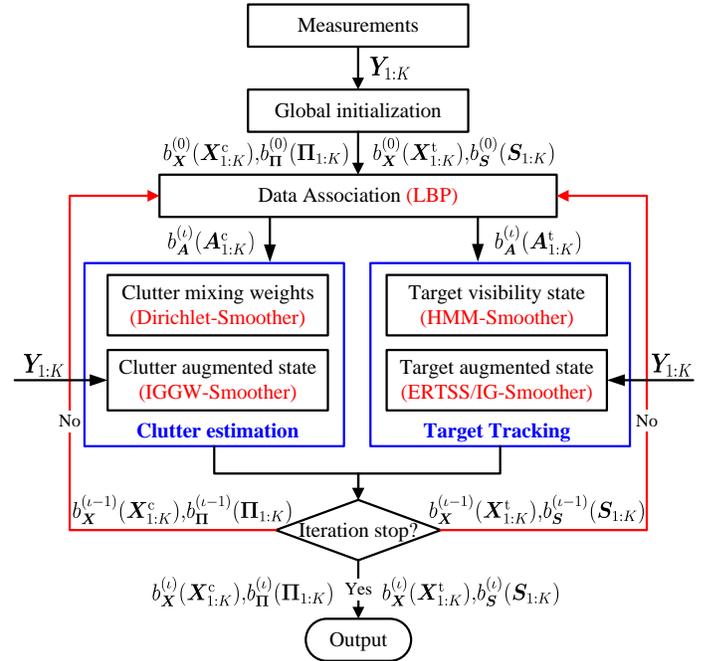


Fig. 8: The flowchart of MP-RMTT.

The complexity associated to the proposed MP-RMTT algorithm is discussed next. Note that the proposed MP-RMTT is an iterative and batch process algorithm with the batch window length K and the number of iterations N_r among five subgraphs, and the corresponding computational cost is provided in Table. I and described below. The estimation of target joint augmented state is solved by an ERTSS and IG smoother with a computational cost $c_x^t = \mathcal{O}(2KN_T)$. The estimation of clutter joint augmented state is carried out by an IG smoother and GW smoother (referred to as IGGW smoother) with a computational cost $c_x^c = \mathcal{O}(2KN_C)$. The forward and backward algorithm is used to estimate the target

Algorithm 1 MP-RMTT algorithm

Input $Y_{1:K}$, the maximum number of iterations r_{\max}^{MP} ;
Output $b_{\mathbf{X}}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{X}}(\mathbf{X}_{1:K}^c)$, $b_{\mathbf{S}}(\mathbf{S}_{1:K})$, $b_{\Pi}(\Pi_{1:K})$, and $b_{\mathbf{A}}(\mathbf{A}_{1:K})$;
 1: **Initialization:** initialize $b_{\mathbf{X}}^{(0)}(\mathbf{X}_{1:K}^t)$, $b_{\mathbf{X}}^{(0)}(\mathbf{X}_{1:K}^c)$, $b_{\mathbf{S}}^{(0)}(\mathbf{S}_{1:K})$, $b_{\Pi}^{(0)}(\Pi_{1:K})$, the tracks number N_T and clutter components number N_C .
 2: **for** $l = 1 : r_{\max}^{\text{MP}}$ **do**
 3: **Data association:** estimate $b_a^{(l)}(a_{i,j,k}^t)$ and $b_a^{(l)}(a_{i,j,k}^c)$, $i = 1, \dots, N_T$, $j = 1, \dots, N_M$, $\tau = 0, \dots, N_C$, $k = 1, \dots, K$ via Eq. (99)-Eq. (102);
 4: **Clutter mixing weights estimation:** calculate $\bar{b}_{\pi}(\pi_{\tau,k})$, $k = 1, \dots, K$, $\tau = 0, \dots, N_C$ via Eq. (69);
 5: **Target visibility estimation:** calculate $b_s(s_{i,k})$, $k = 1, \dots, K$, $i = 1, \dots, N_T$, via Eq. (64);
 6: **Clutter augmented state estimation:** calculate $\bar{b}_x(\tilde{x}_{\tau,k}^c)$ and $b_{\sigma}(\sigma_{\tau,k}^c)$, $k = 1, \dots, K$, $\tau = 0, \dots, N_C$, via Eq. (50) and Eq. (62), respectively;
 7: **Target augmented state estimation:** calculate $\bar{b}_x(x_{i,k}^t)$ and $b_{\sigma}(\sigma_{i,k}^t)$, $k = 1, \dots, K$, $i = 1, \dots, N_T$, via Eq. (35) and Eq. (43), respectively.
 8: **end for**

visibility state and the computational cost is $c_s = \mathcal{O}(KN_T)$. The estimation of clutter mixing weights is solved by a Dirichlet Smoother with a computational cost $c_{\pi} = \mathcal{O}(KN_C)$. The data association is solved by the LBP with a computational cost $c_a = \mathcal{O}(KN_a(N_T + N_C)N_M)$, where N_a is the maximum number of BP iterations. Overall, the computational complexity is $c_{\text{total}} = N_r(c_x^t + c_x^c + c_s + c_{\pi} + c_a)$.

TABLE I: Computational complexity.

Hidden variables	Equation	Computational complexity
$b_{\mathbf{X}}(\mathbf{X}_{1:K}^t)$	(35), (43)	$\mathcal{O}(2KN_rN_T)$
$b_{\mathbf{X}}(\mathbf{X}_{1:K}^c)$	(50), (62)	$\mathcal{O}(2KN_rN_C)$
$b_{\mathbf{S}}(\mathbf{S}_{1:K})$	(64)	$\mathcal{O}(KN_rN_T)$
$b_{\Pi}(\Pi_{1:K})$	(69)	$\mathcal{O}(KN_rN_C)$
$b_{\mathbf{A}}(\mathbf{A}_{1:K})$	(99)-(102)	$\mathcal{O}(KN_rN_a(N_T + N_C)N_M)$

IV. SIMULATION AND ANALYSIS

A. Scenario Configuration

1) *Scenario parameters:* We consider a simulation scenario with five targets together with interference from the environment and countermeasures within a two-dimensional surveillance region. As shown in Fig. 9, all the five targets appear and disappear at $k = 1$ and $k = 340$, respectively. The trajectories of Target 1 and Target 2 can be split into two intervals. The two targets are separated spatially during the first interval. In the second interval, both targets move in parallel at a distance of 20 m. The other three targets (Target 3-Target 5) cross at $k = 175$. Furthermore, there are three elliptical non-uniform clutter components generated by the interference in the simulation scenario, as shown in Fig. 9. Clutter Component 1 is generated by the chaff cloud jamming affected by Target 1 and Target 2, and Clutter

Component 3 is generated by the chaff cloud jamming affected by Target 3, Target 4 and Target 5. Clutter Component 1 appears and disappears at $k = 130$ and $k = 170$, respectively. Clutter Component 3 appears and disappears at $k = 172$ and $k = 225$, respectively. Clutter Component 2 is generated by the atmospheric interference, e.g., thick fog, from $k = 227$ to $k = 280$. By the assumptions of high target speeds and stable atmospheric environment, the position of nonuniform clutter moves within a relatively small area and the shape of the clutter may change in different interference stages.

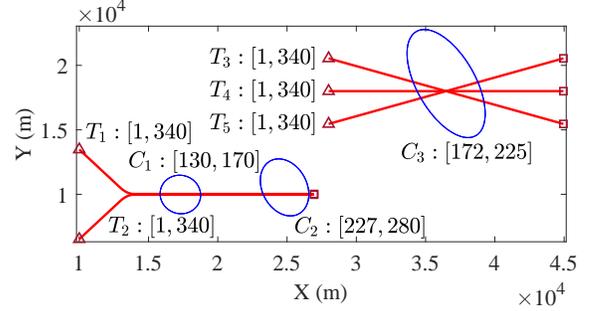


Fig. 9: True trajectories of targets represented by red lines and shapes of nonuniform clutter components represented by blue ellipses. $T_i (C_{\tau}) : [k_1, k_2]$ illustrates that Target i (Clutter Component τ) appears and disappears at k_1 and k_2 . Start/Stop positions of each target are indicated by Δ/\square .

In the first interval of Target 1 and Target 2, both of the two targets follow the constant velocity motion moving close to each other, then follow the constant acceleration motion until the velocity in the Y-axis is zero. In the second interval of Target 1 and Target 2, both of the two targets follow the constant velocity motion. Since the performance evaluation focus on the second interval of Target 1 and Target 2, it is assumed that Target 1 and Target 2 follow the constant velocity motion over all of their interval during the tracking processing for the sake of convenience. Target 3, Target 4 and Target 5 follow constant velocity motion over all of their interval. The parameters of the constant velocity motion are

$$\mathbf{F} = I_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \mathbf{Q} = \sigma_v^2 \times I_2 \otimes \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix}, \quad (103)$$

where $T = 1.25$ s is the sampling period, and $\sigma_v^2 = 0.01$ m/s² is the variance of the driving processes. The parameters of targets are given in Table II unless noted otherwise.

TABLE II: Configurations of targets.

ID	Initial state	Lifetime	SNR
1	$[10000, 40, 13465, -40]^T$	[1, 340]	10/3
2	$[10000, 40, 6570, 40]^T$	[1, 340]	50/3
3	$[28000, 40, 20543, -12]^T$	[1, 340]	10/3
4	$[28000, 40, 18000, 0]^T$	[1, 340]	50/3
5	$[28000, 40, 15458, 12]^T$	[1, 340]	10/3

SNR = $S_{i,k}^t/N_0$ as defined instead of being expressed in log scale.

Each of the nonuniform clutter component is elliptical and spatially follows a two-dimensional Gaussian distribution. The

centroid and axes of the ellipse are used to determine the mean and covariance of the corresponding Gaussian distribution. The configurations of clutter are shown in Table III.

TABLE III: Parameters of clutter.

ID	Number	Mean	Covariance	Lifetime	CNR
0	30	/	/	[1, 340]	1
1	20	$\begin{bmatrix} 20 \text{ km} \\ 30^\circ \end{bmatrix}$	$\begin{bmatrix} 1 \text{ km} & 0 \\ 0 & 3^\circ \end{bmatrix}^2$	[130, 170]	20/3
2	20	$\begin{bmatrix} 27 \text{ km} \\ 23^\circ \end{bmatrix}$	$\begin{bmatrix} 1 \text{ km} & 0 \\ 0 & 3^\circ \end{bmatrix}^2$	[227, 280]	10
3	30	$\begin{bmatrix} 41 \text{ km} \\ 27^\circ \end{bmatrix}$	$\begin{bmatrix} 1 \text{ km} & 0 \\ 0 & 3^\circ \end{bmatrix}^2$	[172, 225]	20/3

Covariance denotes the initial covariance $\mathbf{D}_{\tau,0}^c$ of clutter component τ . Number denotes the initial numbers $\lambda_{\tau,0}$ of clutter component τ . CNR = $S_{\tau,k}^c/N_0$ as defined instead of being expressed in log scale.

The sensor measurements consist of position information and strength information with detection threshold $d = 0.715$. The spatial measurements have zero-mean white Gaussian noise with covariance $\mathbf{R} = \text{diag}(20 \text{ m}, 0.6^\circ)^2$. The measured signal strength is sampled by the inversion method and the acceptance-rejection method for Swerling-I model and Swerling-III model [41], respectively.

2) *Algorithm Parameters*: The forgetting factors of target mean SNR transition PDF and clutter mean CNR transition PDF are $\mu^t = \mu^c = 1.05$. The forgetting factor of clutter spatial transition PDF is $\xi = 0.99$. The balance parameter of clutter mixing weight transition PDF is $\kappa = 5$. For parameters of MP, window length $K = 7$ and sliding step $s = 3$, $\delta_T^{\text{MP}} = 10^{-3}$, the maximum number of iterations $r_{\text{max}}^{\text{MP}} = 3$. For BP in the data association, the iterative convergence threshold $\delta_T^{\text{BP}} = 10^{-6}$, the maximum number of iterations $r_{\text{max}}^{\text{BP}} = 1000$ and the damping factor $\gamma = 0.9$. The detection probability $P_d(s_{i,k} = 0) = 0.01$ if the target is not visible. The target birth probability is set to $p_b = 0.15$ and the target survival probability is set to $p_s = 0.8$. For track initialization, the maximum speed of any target in Cartesian coordinate is $v_{\text{max}} = [120 \text{ m/s } 120 \text{ m/s}]^T$, and the number of consecutive missing measurements of any tracks is less than $L_{\text{max}} = 3$. The initial target visibility state is $f_s = 0.5$. The thresholds of track confirmation and deletion are set to 0.75 and 0.5.

3) *Performance Evaluation*: The performance metrics are given by

- Correct associations rate of targets and measurements (CAR);
- Number of false tracks (NFT) [45];
- Mean optimal subpattern assignment [46] for target position estimation (MOSPA) with order and cutoff parameters as $p = 2$ and $c = 631 \text{ m}$ respectively;
- Relative SNR error (RSE), given by $\Delta\sigma_{i,k}/\sigma_{i,k}$;
- Total number of nonuniform clutter point (TNNC);
- Root Mean Squares Error for nonuniform clutter position estimation (RMSE);
- Wasserstein distance for nonuniform clutter shape estimation (WD), given by $\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^2 + \text{tr}\left\{\mathbf{D}_k + \hat{\mathbf{D}}_k - 2\sqrt{\sqrt{\mathbf{D}_k}\hat{\mathbf{D}}_k\sqrt{\mathbf{D}_k}}\right\}$, where the square root of a matrix \mathbf{D} is defined as the matrix \mathbf{Y} for $\mathbf{D} = \mathbf{Y}^T\mathbf{Y}$.

The values of metrics are average of 100 Monte Carlo runs.

4) *Simulation Scenarios*: We consider three scenarios to demonstrate the performance of MP-RMTT. We outline the three scenarios in Table IV. We compare MP-RMTT against the PHD, CPHD, and CBMeMBer filter with integrated clutter estimation using measured strength information, which use the methods as in [19] with the nonuniform clutter spatial estimation method in [10], referred to MP-RMTT, PHD, CPHD, and CBMeMBer, respectively. We also compare the first iteration output of MP-RMTT, referred to MP-RMTTst. In all of the algorithms, the same scenario input is used.

TABLE IV: Simulation scenarios.

Scenario	Clutter Spatial distribution	Nonuniform Clutter
1	Uniform	\
2	Uniform and Nonuniform	Unvarying
3	Uniform and Nonuniform	Varying

B. Scenario 1: Data Association Aided by AI

The first scenario focus on the capability of target mean SNR estimation and the discrimination ability for closely spaced targets aided by the obtained target mean SNR estimation. There are Target 1 and Target 2, and only uniform clutter. The history of measurements generated by targets and clutter, as well as the target trajectory estimation obtained by MP-RMTT are shown in Fig. 10. During the first interval, MP-RMTT can obtain nearly stable SNR estimation for each target. Since the trajectories of the two targets are very close in the second interval, it is obvious that the two targets cannot be distinguished by spatial measurement alone. The tracking result in Fig. 10 shows that the two targets can be distinguished and tracked by MP-RMTT with strength information.

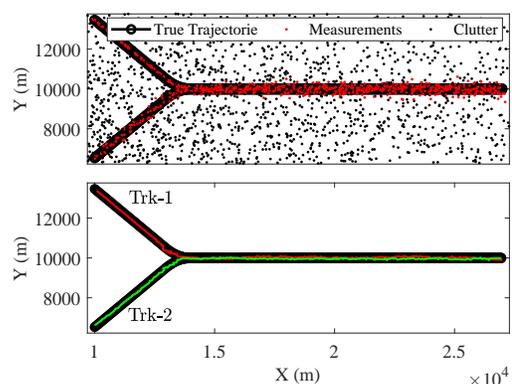


Fig. 10: The history of measurements generated by targets and clutter, and trajectories estimation of targets (bottom) obtained by MP-RMTT. The black and colored lines represent the target true trajectories and valid tracks, respectively. This representation is also used in the figures below.

To analyze the performance of target mean SNR estimation and target tracking with strength information, the evolution of the MOSPA distance and RSE are evaluated, as shown in Fig. 11. It can be seen that stable SNR estimations are obtained for all algorithms as the RSE converges to a small value. The

presented results also demonstrate that MP-RMTT performs better than the PHD, CPHD, and CBMeMber filters.

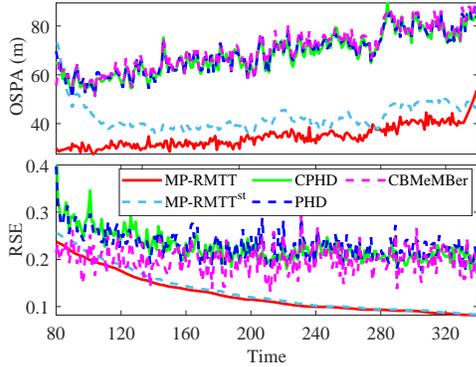


Fig. 11: Monte Carlo average of MOSPA distance (top) and RSE (bottom) of targets.

The obtained MOSPA distance and CAR results w.r.t. different target SNRs are listed in Table. V. In the case of the MP-RMTT that solely exploits kinematic information (referred to MP-RMTT-KI) will inevitably lose the CAR during the second interval. As expected, the obtained CAR take values around 50%. The tracking algorithm utilizing the strength information can obtain an CAR of 80% when the SNR of Target 1 is three times larger than that of Target 2. As the SNR of the second target increases, the ability to discriminate also improves. The MP-RMTT with the Swerling-III model achieves almost optimal CAR with 97% when the SNR of Target 1 is nine times larger than that of Target 2. Note that, MP-RMTT is superior to PHD, CPHD and CBMeMber in terms of the MOSPA metric.

C. Scenario 2: RMTT for Unvarying Clutter

The second scenario focus on the RMTT capability in unvarying clutter background. There are all five targets, one uniform clutter component, and three nonuniform clutter components. The history of measurements generated by targets and clutter are shown in Fig. 12a. The tracking and nonuniform clutter estimation results of MP-RMTT are shown in Fig. 12b. The black and colored ellipsoids represent the covariance matrices of each nonuniform clutter component. This representation is also used in the figures below. The presented results demonstrate that MP-RMTT tracks all five targets correctly, and no false track is generated. Meanwhile, MP-RMTT successfully estimates the spatial and shape of the nonuniform clutter.

To reveal the performance improvements brought by the MP-RMTT, three simulation tests are performed. In the first simulation, the spatial distribution of clutter is set to a uniform distribution where the expected number of clutter per scan is known, referred to MP-RMTT-NCE. In the second simulation, MP-RMTT, PHD and CPHD are used for clutter estimation and target tracking, and the first iteration output of MP-RMTT is also considered. The algorithm knows the clutter intensity exactly in the third simulation, referred to MP-RMTT-KCE.

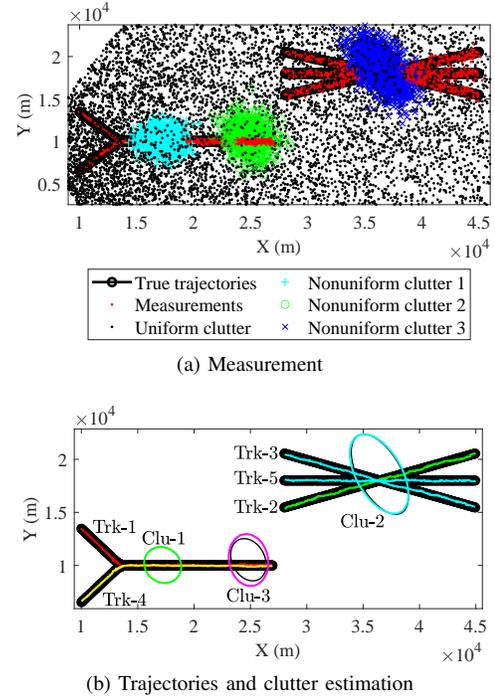


Fig. 12: Measurements, tracks and clutter estimation obtained by MP-RMTT.

The MOSPA of target position and NFT are given in Fig. 13. It is clearly seen that MP-RMTT has comparable performance to algorithms with known clutter intensity perfectly. On the contrary, the performance without clutter estimation is severely degraded. The TNNC, RMSE and WD metrics for clutter estimation are given in Fig. 14, which shows that MP-RMTT has the best clutter estimation performance. The results indicate that the proposed MP-RMTT has comparable performance to an algorithm with a perfectly known clutter distribution, demonstrating the effectiveness and robustness of the algorithm.

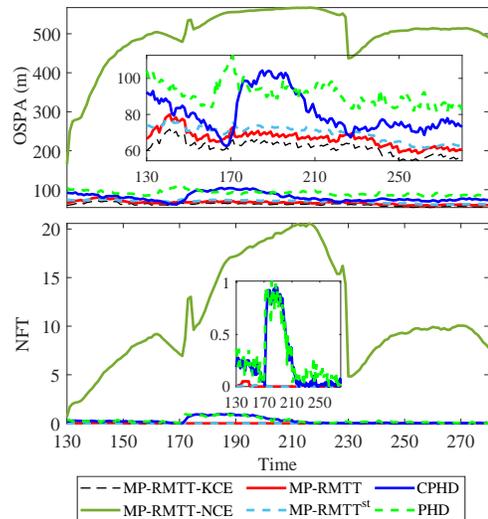


Fig. 13: Monte Carlo MOSPA (top) and NFT (bottom) for target tracking.

TABLE V: Performance comparison in different SNRs

Algorithms	SNR = [10/3 10] dB				SNR = [10/3 20] dB				SNR = [10/3 30] dB			
	Swerling-I		Swerling-II		Swerling-I		Swerling-II		Swerling-I		Swerling-II	
	OSAP	CAR	OSAP	CAR	OSAP	CAR	OSAP	CAR	OSAP	CAR	OSAP	CAR
MP-RMTT-KI	49.2	0.51	49.7	0.52	49.5	0.48	48.7	0.51	51.8	0.46	48.2	0.48
MP-RMTT	42.0	0.80	39.0	0.90	36.8	0.92	35.0	0.97	37.2	0.94	34.7	0.98
PHD	71.3	\	71.6	\	70.9	\	70.3	\	69.7	\	70.8	\
CPHD	158.5	\	153.2	\	159.1	\	154.0	\	157.2	\	153.4	\
CBMeMBer	80.2	\	79.5	\	79.3	\	78.2	\	80.0	\	78.7	\

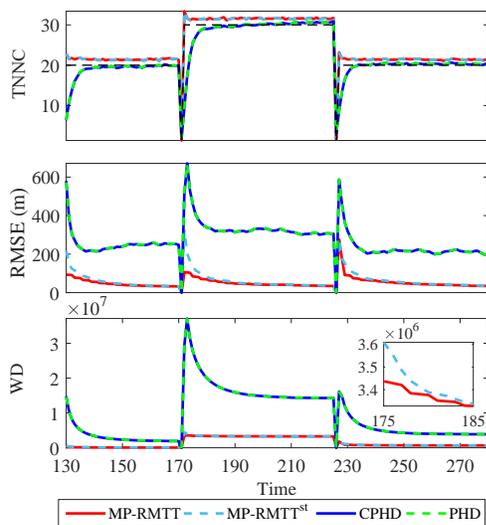


Fig. 14: Monte Carlo TNNC (top), RMSE (middle) and WD (bottom) for clutter estimation.

Fig. 15 depicts the performance comparison of target tracking w.r.t. different number of clutter. As shown in the top of Fig. 15, MP-RMTT performs better on MOSPA compared to other algorithms. The MP-RMTTst and MP-RMTT have comparable performance on the false track acceptance rate (as shown in the middle of Fig. 15), which are better than PHD and CPHD. In terms of RSE (as shown in the bottom of Fig. 15), MP-RMTT is slightly better than MP-RMTTst, besides, PHD and CPHD are worst. Meanwhile, by comparing the MOSPA and RSE of MP-RMTTst and MP-RMTT, it is clearly seen that the performance is improved by the closed-loop process, illustrating the effectiveness of the closed-loop iterative framework. Overall, MP-RMTT outperforms the other algorithms. It is due to the fact that the closed-loop iterative and batch processing of MP-RMTT can improve the performance and robustness of target tracking.

Fig. 16 shows the performance comparison of clutter estimation w.r.t. different number of clutter. The performance of MP-RMTTst is comparable to the performance of MP-RMTT w.r.t. the number of clutter (as shown in the top of Fig. 16), which is better than PHD and CPHD. In terms of clutter RMSE (as shown in the middle of Fig. 16), which is decreased slightly as the increase of the number of clutter, MP-RMTT outperforms the MP-RMTTst slightly; PHD and CPHD are the worst. The WD in the bottom of Fig. 16 shows that,

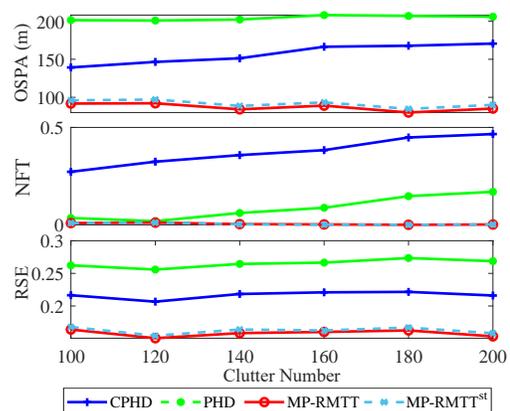


Fig. 15: Performance comparison of target tracking w.r.t. different number of clutter.

MP-RMTT is superior to the other two algorithms. Because PHD and CPHD use the same clutter estimation algorithm, they have comparable performance. MP-RMTT outperforms both PHD and CPHD.

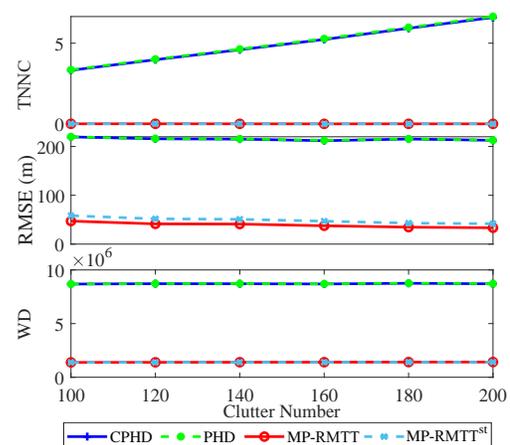


Fig. 16: Performance comparison of clutter estimation w.r.t. different number of clutter.

D. Scenario 3: RMTT for Varying Clutter

The third scenario focuses on the RMTT in varying clutter background. The scenario contains all five targets, a uniform clutter component and three nonuniform clutter components.

Assume that the number and the shape of nonuniform clutter components are time-varying and given by

$$\lambda_{\tau,k} = \left(1 + \frac{1}{2} \sin\left(\frac{k - k_{\text{start}}^{\tau}}{k_{\text{end}}^{\tau} - k_{\text{start}}^{\tau}} \pi\right)\right) \lambda_{\tau,0}, \quad (104)$$

$$\mathbf{D}_{\tau,k}^c = \left(\frac{1}{3} + \frac{k - k_{\text{start}}^{\tau}}{k_{\text{end}}^{\tau} - k_{\text{start}}^{\tau}}\right) \mathbf{D}_{\tau,0}^c,$$

where the initial number $\lambda_{\tau,0}$, initial covariance $\mathbf{D}_{\tau,0}^c$, start time k_{start}^{τ} and end time k_{end}^{τ} of the nonuniform clutter component τ are represented in Table. III. The results reveal that the number of each nonuniform clutter component starts at $\lambda_{\tau,0}$ and then varies as a sinusoidal function, and that the shape starts at $\mathbf{D}_{\tau,0}^c/3$ and then increases linearly. The history of detections and the clutter estimation as well as target tracking results are shown in Fig. 17. This demonstrates that MP-RMTT can obtain target trajectory and nonuniform clutter estimates in cross and close target scenarios.

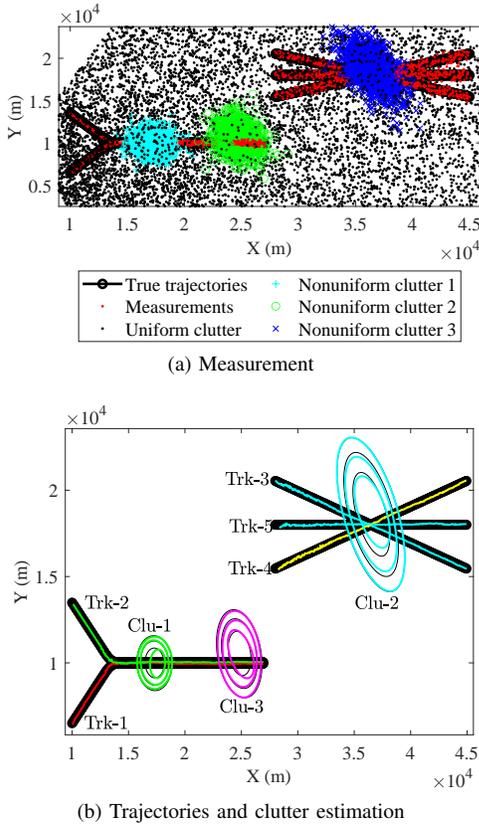


Fig. 17: Measurements, the trajectories and nonuniform clutter estimation obtained by MP-RMTT. The three ellipses of each nonuniform clutter component from small to large represent the covariance matrix of the clutter region at $k = k_{\text{start}}^{\tau}$, $k = k_{\text{start}}^{\tau} + 20$ and $k = k_{\text{start}}^{\tau} + 40$ respectively, where k_{start}^{τ} is start time of the clutter component τ .

The MOSPA of target position and NFT are given in Fig. 18. As the same as in the second scenario, excellent performance is achieved when the clutter distribution is perfectly known, and the performance without clutter estimation is severely degraded. It is worth noting that as the clutter varies, the NFT of PHD and CPHD increases, even up to two per scan,

resulting in a dramatic increase in MOSPA. The performance of MP-RMTT does not degrade as the clutter varies and is comparable to the optimum. The TNNC, RMSE and WD metrics for clutter estimation are given in Fig. 19. The TNNC metric of PHD and CPHD deteriorates as clutter varies. It can be seen that MP-RMTT has the best clutter estimation performance. These results demonstrate that the proposed MP-RMTT algorithm does not degrade as the clutter varies.

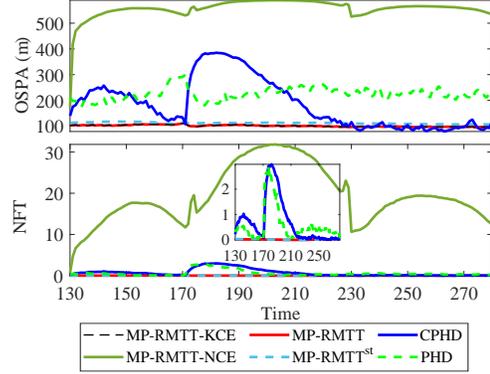


Fig. 18: Monte Carlo MOSPA (top) and NFT (bottom) for target tracking.

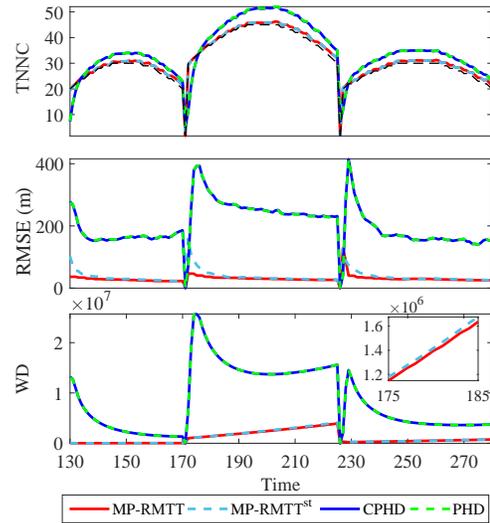


Fig. 19: Monte Carlo TNNC (top), RMSE (middle) and WD (bottom) for clutter estimation.

Fig. 20 shows the target tracking performance w.r.t. different target SNR. Fig. 21 illustrates the clutter estimation performance w.r.t. different numbers of clutter. Likewise in Scenario 2, the results show that MP-RMTT outperforms PHD and CPHD, and improves with the increasing number of iterations, illustrating the robustness of MP-RMTT.

V. CONCLUSIONS

We proposed and demonstrated the application of the MP method to the problem of RMTT by JCETT using measurements with strength information. The proposed MP-based method exhibits outstanding and robust tracking and clutter

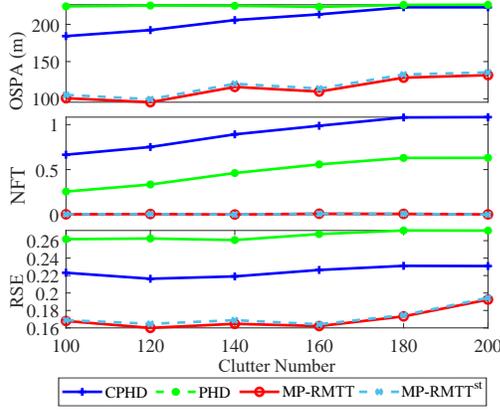


Fig. 20: Performance comparison of target tracking w.r.t. different number of clutter.

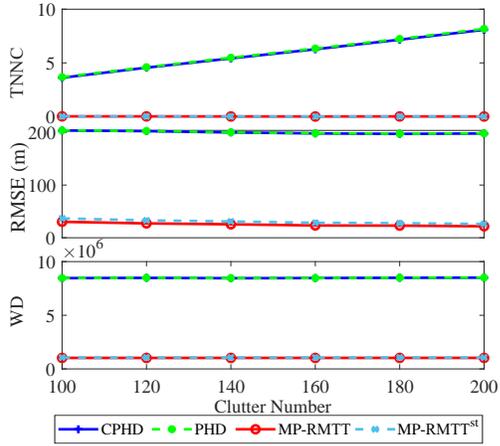


Fig. 21: Clutter estimation performance comparison w.r.t. different number of clutter.

estimation performance. This was achieved through the establishment of a closed-loop iterative framework for target tracking and clutter estimation. Simulation results in three scenarios with different target and clutter distributions showed that the proposed method is superior to methods without clutter estimation and previously proposed methods. In particular, the proposed MP algorithm has a significant improvement in clutter estimation performance relative to PHD and CPHD filters. Promising future research directions are the multi-sensor extension of the proposed method in multi-sensor tracking scenarios as well as using other feature information to discriminate between targets and clutter, such as micro-motion and polarization information.

APPENDIX

A. Derivation of $\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^t)]$ and $\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^c)]$

Because $\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^t)]$ and $\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k}^c)]$ are both expectations of Rayleigh distributions, for the sake of brevity, we denote $\sigma_{i,k}^t$ and $\sigma_{i,k}^c$ here by $\sigma_{i,k}$. The expectation

$\mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k})]$ is derived as follows.

$$\begin{aligned} \mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k})] &= \mathbb{E}[\ln \mathcal{R}^d(m_{j,k}; \sigma_{i,k})] \\ &\stackrel{c}{=} \mathbb{E}[(2n-1) \ln m_{j,k} - n \ln \sigma_{i,k} - n \frac{m_{j,k}^2}{\sigma_{i,k}}] \\ &= (2n-1) \ln m_{j,k} - n \mathbb{E}[\ln \sigma_{i,k}] - nm_{j,k}^2 \mathbb{E}\left[\frac{1}{\sigma_{i,k}}\right]. \end{aligned} \quad (105)$$

The closed-form solution of the expectation $\mathbb{E}[\ln \sigma_{i,k}]$ is hard to obtain. To this end, we use a second-order Taylor expansion of $\ln(\sigma_{i,k})$, given by

$$\ln(\sigma_{i,k}) \approx \ln(\mathbb{E}[\sigma_{i,k}]) + \frac{2\sigma_{i,k}}{\mathbb{E}[\sigma_{i,k}]} - \frac{\sigma_{i,k}^2}{2\mathbb{E}[\sigma_{i,k}]^2} - \frac{3}{2}. \quad (106)$$

Substituting Eq. (105) into Eq. (106), yields

$$\begin{aligned} \mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k})] &= (2n-1) \ln m_{j,k} - \\ &n \ln(\mathbb{E}[\sigma_{i,k}]) + \frac{n\mathbb{E}[\sigma_{i,k}^2]}{2\mathbb{E}[\sigma_{i,k}]^2} - \mathbb{E}\left[\frac{nm_{j,k}^2}{\sigma_{i,k}}\right] - \frac{n}{2}, \end{aligned} \quad (107)$$

where,

$$\begin{aligned} \mathbb{E}[\sigma_{i,k}^{-1}] &= \frac{\alpha_{i,k}}{\beta_{i,k}}, \quad \mathbb{E}[\sigma_{i,k}] = \frac{\beta_{i,k}}{\alpha_{i,k} - 1}, \\ \mathbb{E}[\sigma_{i,k}^2] &= \frac{\beta_{i,k}^2}{(\alpha_{i,k} - 1)(\alpha_{i,k} - 2)}. \end{aligned} \quad (108)$$

Substituting Eq. (108) into Eq. (107), yields

$$\begin{aligned} \mathbb{E}[\ln p(m_{j,k}; \sigma_{i,k})] &= (2n-1) \ln m_{j,k} - nm_{j,k}^2 \frac{\alpha_{i,k}}{\beta_{i,k}} + \frac{n}{2(\alpha_{i,k} - 2)}. \end{aligned} \quad (109)$$

B. Derivation of Data Association

We find that it is convenient to define the following messages

$$\begin{aligned} \theta_{i,j,k}^t &= \frac{m_{f_{\mathbf{Y}_k}^t \rightarrow a_{i,j,k}^t}^{\text{MF}}(1) m_{f_{A_k^t}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(1)}{m_{f_{\mathbf{Y}_k}^t \rightarrow a_{i,j,k}^t}^{\text{MF}}(0) m_{f_{A_k^t}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(0)}, \\ \theta_{\tau,j,k}^c &= \frac{m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}}(1) m_{f_{A_k^c}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}}(1)}{m_{f_{\mathbf{Y}_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}}(0) m_{f_{A_k^c}^c \rightarrow a_{\tau,j,k}^c}^{\text{MF}}(0)}, \end{aligned} \quad (110)$$

$$\begin{aligned} \alpha_{i,j,k}^t &= \frac{m_{f_{I_k}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(1)}{m_{f_{I_k}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(0)}, \quad \alpha_{\tau,j,k}^c = \frac{m_{f_{I_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{BP}}(1)}{m_{f_{I_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{BP}}(0)}, \\ \rho_{i,j,k}^t &= \frac{n_{a_{i,j,k}^t \rightarrow f_{I_{i,k}^t}}(1)}{n_{a_{i,j,k}^t \rightarrow f_{I_{i,k}^t}}(0)}, \quad \rho_{\tau,j,k}^c = \frac{n_{a_{\tau,j,k}^c \rightarrow f_{I_{\tau,k}^c}}(1)}{n_{a_{\tau,j,k}^c \rightarrow f_{I_{\tau,k}^c}}(0)}, \end{aligned} \quad (111)$$

$$\begin{aligned} \eta_{i,j,k}^t &= \frac{m_{f_{E_k}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(1)}{m_{f_{E_k}^t \rightarrow a_{i,j,k}^t}^{\text{BP}}(0)}, \quad \eta_{\tau,j,k}^c = \frac{m_{f_{E_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{BP}}(1)}{m_{f_{E_k}^c \rightarrow a_{\tau,j,k}^c}^{\text{BP}}(0)}, \\ \beta_{i,j,k}^t &= \frac{n_{a_{i,j,k}^t \rightarrow f_{E_{i,k}^t}}(1)}{n_{a_{i,j,k}^t \rightarrow f_{E_{i,k}^t}}(0)}, \quad \beta_{\tau,j,k}^c = \frac{n_{a_{\tau,j,k}^c \rightarrow f_{E_{\tau,k}^c}}(1)}{n_{a_{\tau,j,k}^c \rightarrow f_{E_{\tau,k}^c}}(0)}. \end{aligned} \quad (112)$$

Using the definition in Eq. (111), $\alpha_{i,j,k}^t$ and $\alpha_{i,j,k}^c$ can be calculated as follows.

$$\alpha_{i,j,k}^t = \frac{\Gamma_{i,j,k}^t}{\Upsilon_{i,j,k}^t + \Phi_{i,j,k}^t} = \frac{1}{\sum_{i'=1 \setminus i}^{N_T} \rho_{i',j,k}^t + \sum_{\tau=0}^{N_C} \rho_{\tau,j,k}^c}, \quad (113)$$

$$\alpha_{\tau,j,k}^c = \frac{\Gamma_{\tau,j,k}^c}{\Upsilon_{\tau,j,k}^c + \Phi_{\tau,j,k}^c} = \frac{1}{\sum_{i=1}^{N_T} \rho_{i,j,k}^t + \sum_{\tau'=0 \setminus \tau}^{N_C} \rho_{\tau',j,k}^c},$$

where

$$\begin{aligned} \Gamma_{i,j,k}^t &= \prod_{i'=1 \setminus i}^{N_T} \rho_{i',j,k}^t(0) \prod_{\tau=0}^{N_C} \rho_{\tau,j,k}^c(0), \\ \Upsilon_{i,j,k}^t &= \sum_{i'=1 \setminus i}^{N_T} \rho_{i',j,k}^t(1) \prod_{i''=1 \setminus i, i''}^{N_T} \rho_{i'',j,k}^t(0) \prod_{\tau=0}^{N_C} \rho_{\tau,j,k}^c(0), \\ \Phi_{i,j,k}^t &= \sum_{\tau=0}^{N_C} \rho_{\tau,j,k}^c(1) \prod_{i'=1 \setminus i}^{N_T} \rho_{i',j,k}^t(0) \prod_{\tau'=0 \setminus \tau}^{N_C} \rho_{\tau',j,k}^c(0), \\ \Gamma_{\tau,j,k}^c &= \prod_{i=1}^{N_T} \rho_{i,j,k}^t(0) \prod_{\tau'=0 \setminus \tau}^{N_C} \rho_{\tau',j,k}^c(0), \\ \Upsilon_{\tau,j,k}^c &= \sum_{i=1}^{N_T} \rho_{i,j,k}^t(1) \prod_{i'=1 \setminus i}^{N_T} \rho_{i',j,k}^t(0) \prod_{\tau'=1 \setminus \tau}^{N_C} \rho_{\tau',j,k}^c(0), \\ \Phi_{\tau,j,k}^c &= \sum_{\tau'=0 \setminus \tau}^{N_C} \rho_{\tau',j,k}^c(1) \prod_{i=1}^{N_T} \rho_{i,j,k}^t(0) \prod_{\tau''=0 \setminus \tau, \tau''}^{N_C} \rho_{\tau'',j,k}^c(0). \end{aligned} \quad (114)$$

By Eq. (95), Eq. (111) and Eq. (112), we have

$$\begin{aligned} \beta_{i,j,k}^t &= \theta_{i,j,k}^t \alpha_{i,j,k}^t, \quad \beta_{\tau,j,k}^c = \theta_{\tau,j,k}^c \alpha_{\tau,j,k}^c, \\ \rho_{i,j,k}^t &= \theta_{i,j,k}^t \eta_{i,j,k}^t, \quad \rho_{\tau,j,k}^c = \theta_{\tau,j,k}^c \eta_{\tau,j,k}^c. \end{aligned} \quad (115)$$

Substituting Eq. (115) into Eq. (113), we obtain

$$\begin{aligned} \beta_{i,j,k}^t &= \frac{\theta_{i,j,k}^t}{\theta_{0,j,k}^c + \sum_{i'=1 \setminus i}^{N_T} \theta_{i',j,k}^t \eta_{i',j,k}^t + \sum_{\tau=1}^{N_C} \theta_{\tau,j,k}^c \eta_{\tau,j,k}^c}, \\ \beta_{\tau,j,k}^c &= \frac{\theta_{\tau,j,k}^c}{\theta_{0,j,k}^c + \sum_{i=1}^{N_T} \theta_{i,j,k}^t \eta_{i,j,k}^t + \sum_{\tau'=1 \setminus \tau}^{N_C} \theta_{\tau',j,k}^c \eta_{\tau',j,k}^c}. \end{aligned} \quad (116)$$

Using the definition in Eq. (112), $\eta_{i,j,k}^t$ and $\eta_{i,j,k}^c$ can be calculated as follows

$$\begin{aligned} \eta_{i,j,k}^t &= \frac{\prod_{j'=0 \setminus j}^{N_{M,k}} \beta_{i,j',k}^t(0)}{\sum_{j'=0 \setminus j}^{N_{M,k}} \beta_{i,j',k}^t(1) \prod_{j''=0 \setminus j, j''}^{N_{M,k}} \beta_{i,j'',k}^t(0)} \\ &= \frac{1}{\sum_{j'=0 \setminus j}^{N_{M,k}} \beta_{i,j',k}^t}, \end{aligned} \quad (117)$$

$$\begin{aligned} \eta_{\tau,j,k}^c &= \frac{\Psi_{\tau,j,k}^c}{\beta_{\tau,0,k}^c(1) \prod_{j'=1 \setminus j}^{N_{M,k}} \beta_{\tau,j',k}^c(0) + \Psi_{\tau,j,k}^c} \\ &= \frac{\prod_{j'=1 \setminus j}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}{\beta_{\tau,0,k}^c + \prod_{j'=1 \setminus j}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}, \end{aligned} \quad (118)$$

$$\eta_{\tau,0,k}^c = \frac{\prod_{j'=1}^{N_{M,k}} \beta_{\tau,j',k}^c(0)}{\sum_{\mathbf{A}_i^c \setminus \mathbf{a}_{\tau,0,k}^c} \prod_{j'=1}^{N_{M,k}} \beta_{\tau,j',k}^c} = \frac{1}{\prod_{j'=1}^{N_{M,k}} (1 + \beta_{\tau,j',k}^c)}, \quad (119)$$

where

$$\Psi_{\tau,j,k}^c = \beta_{\tau,0,k}^c(0) \sum_{\mathbf{A}_i^c \setminus \mathbf{a}_{\tau,j,k}^c, \mathbf{a}_{\tau,0,k}^c} \prod_{j'=1 \setminus j}^{N_{M,k}} \beta_{\tau,j',k}^c. \quad (120)$$

REFERENCES

- [1] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of modern radar: basic principles*. NC Raleigh: SciTech Publishing, 2010.
- [2] Y. Bar-Shalom and X. Li, *Multitarget-multisensor tracking: principles and techniques*. YBS publishing, Storrs, 1995.
- [3] R. P. Mahler, B. T. Vo, and B. N. Vo, "CPHD filtering with unknown clutter rate and detection profile," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3497–3513, 2011.
- [4] M. Beard, B. T. Vo, and B. N. Vo, "Multitarget filtering with unknown clutter density using a bootstrap GMCPHD filter," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 323–326, 2013.
- [5] W. C. Kim and T. L. Song, "Interactive clutter measurement density estimator for multitarget data association," *IET Radar, Sonar and Navigation*, vol. 11, no. 1, pp. 125–132, 2017.
- [6] B. T. Vo, B. N. Vo, R. Hoseinnezhad, and R. P. S. Mahler, "Robust multi-Bernoulli filtering," *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 399–409, 2013.
- [7] A. K. Gostar, R. Hoseinnezhad, and A. Bab-Hadiashar, "Multi-Bernoulli sensor-selection for multi-target tracking with unknown clutter and detection profiles," *Signal Processing*, vol. 119, pp. 28–42, 2015.
- [8] F. Lian, C. Han, and W. Liu, "Estimating unknown clutter intensity for PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 2066–2078, 2010.
- [9] X. Chen, R. Tharmarasa, M. Pelletier, and T. Kirubarajan, "Integrated clutter estimation and target tracking using Poisson point processes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1210–1235, 2012.
- [10] W. Liu, Y. Chen, H. Cui, and C. Wen, "A nonuniform clutter intensity estimation algorithm for random finite set filters," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 6, pp. 2911–2925, 2018.
- [11] D. Lerro and Y. Bar-Shalom, "Interacting multiple model tracking with target amplitude feature," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 2, pp. 494–509, 1993.
- [12] L. M. Ehrman and P. R. Mahapatra, "Impact of noncoherent pulse integration on RCS-assisted tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1573–1579, 2009.
- [13] D. Clark, B. Ristić, B. N. Vo, and B. T. Vo, "Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 26–37, 2010.
- [14] F. Yang, W. Zhang, Y. Liang, Y. Su, and X. Yao, "Cardinality balanced multi-target multi-Bernoulli filter for target tracking with amplitude information," in *Proceedings of 19th International Conference on Information Fusion*. ISIF, 2016, pp. 958–964.
- [15] M. Mertens, M. Ulmke, and W. Koch, "Ground target tracking with RCS estimation based on signal strength measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 205–220, 2016.
- [16] S. H. Bae, J. Park, and K. J. Yoon, "Joint estimation of multi-target signal-to-noise ratio and dynamic states in cluttered environment," *IET Radar, Sonar and Navigation*, vol. 11, no. 3, pp. 539–549, 2017.
- [17] S. H. Bae, "Survey of amplitude-aided multi-target tracking methods," *IET Radar, Sonar and Navigation*, vol. 13, no. 2, pp. 243–253, 2019.
- [18] J. Sun, C. Liu, Q. Li, and X. Chen, "Labelled multi-Bernoulli filter with amplitude information for tracking marine weak targets," *IET Radar, Sonar and Navigation*, vol. 13, no. 6, pp. 983–991, 2019.
- [19] B. Yang, J. Wang, C. Yuan, J. Thiyyagalingam, and T. Kirubarajan, "Multi-object Bayesian filters with amplitude information in clutter background," *Signal Processing*, vol. 152, pp. 22–34, 2018.
- [20] B. Ristic, L. Rosenberg, D. Y. Kim, and R. Guan, "Bernoulli filter for tracking maritime targets using point measurements with amplitude," *Signal Processing*, vol. 181, p. 107919, 2021.
- [21] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.
- [22] C. Zhang, J. Butepage, H. Kjellstrom, and S. Mandt, "Advances in variational inference," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 2008–2026, 2019.

- [23] L. Chen, M. J. Wainwright, M. Cetin, and A. S. Willsky, "Data association based on optimization in graphical models with application to sensor networks," *Mathematical and computer modelling*, vol. 43, no. 9-10, pp. 1114–1135, 2006.
- [24] J. L. Williams and R. A. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2942–2959, 2014.
- [25] —, "Multiple scan data association by convex variational inference," *IEEE Transactions on Signal Processing*, vol. 66, no. 8, pp. 2112–2127, 2018.
- [26] S. Sun, H. Lan, Z. Wang, Q. Pan, and H. Zhang, "The application of sum-product algorithm for data association," in *Proceedings of 19th International Conference on Information Fusion*. ISIF, 2016, pp. 416–423.
- [27] T. Kropfleiter, F. Meyer, and F. Hlawatsch, "A fast labeled multi-Bernoulli filter using belief propagation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 2478–2488, 2019.
- [28] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3478–3493, 2017.
- [29] F. Meyer, T. Kropfleiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 121–259, 2018.
- [30] G. Soldi, F. Meyer, P. Braca, and F. Hlawatsch, "Self-tuning algorithms for multisensor-multitarget tracking using belief propagation," *IEEE Transactions on Signal Processing*, vol. 67, no. 15, pp. 3922–3937, 2019.
- [31] F. Meyer and M. Z. Win, "Scalable data association for extended object tracking," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 491–507, 2020.
- [32] F. Meyer and J. L. Williams, "Scalable detection and tracking of geometric extended objects," *IEEE Transactions on Signal Processing*, vol. 69, pp. 6283–6298, 2021.
- [33] P. Sharma, A. A. Saucan, D. J. Bucci, and P. K. Varshney, "Decentralized Gaussian filters for cooperative self-localization and multi-target tracking," *IEEE Transactions on Signal Processing*, vol. 67, no. 22, pp. 5896–5911, 2019.
- [34] D. Cormack, I. Schlangen, J. R. Hoggood, and D. E. Clark, "Joint registration and fusion of an infra-red camera and scanning radar in a maritime context," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1357–1369, 2019.
- [35] D. Gaglione, P. Braca, G. Soldi, F. Meyer, F. Hlawatsch, and M. Z. Win, "Fusion of sensor measurements and target-provided information in multitarget tracking," *IEEE Transactions on Signal Processing*, vol. 70, pp. 322–336, 2022.
- [36] E. Riegler, G. E. Kirkelund, C. N. Manchón, M. A. Badiu, and B. H. Fleury, "Merging belief propagation and the mean field approximation: A free energy approach," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 588–602, 2013.
- [37] H. Lan, S. Sun, Z. Wang, Q. Pan, and Z. Zhang, "Joint target detection and tracking in multipath environment: a variational Bayesian approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 2136–2156, 2019.
- [38] H. Lan, J. Ma, Z. Wang, Q. Pan, and X. Xu, "A message passing approach for multiple maneuvering target tracking," *Signal Processing*, vol. 174, p. 107621, 2020.
- [39] H. Lan, Z. Wang, X. Bai, Q. Pan, and K. Lu, "Measurement-level target tracking fusion for over-the-horizon radar network using message passing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 3, pp. 1600–1623, 2021.
- [40] M. I. Skolnik, *Introduction to radar systems, third edition*. New York: McGraw-Hill, 2002.
- [41] C. Bishop, *Pattern recognition and machine learning*. Stat Sci, 2006.
- [42] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, "A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 594–601, 2018.
- [43] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge university press, 2013.
- [44] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [45] A. A. Gorji, R. Tharmarasa, and T. Kirubarajan, "Performance measures for multiple target tracking problems," in *Proceedings of the 14th International Conference on Information Fusion*. ISIF, 2011, pp. 1–8.
- [46] D. Schuhmacher, B. T. Vo, and B. N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.