Proceedings Paper:
Liu, Y., Li, X., Yang, L. et al. (2 more authors) (2022) On the fixed-interval smoothing for jump Markov nonlinear systems. In: Proceedings of the 2022 25th International Conference on Information Fusion (FUSION). 2022 25th International Conference on Information Fusion (FUSION), 04-07 Jul 2022, Linköping, Sweden. Institute of Electrical and Electronics Engineers . ISBN 9781665489416

https://doi.org/10.23919/FUSION49751.2022.9841334

On the Fixed-Interval Smoothing for Jump Markov Nonlinear Systems

Yi Liu[1◇], Xi Li[2◇], Le Yang[1*], Lyudmila Mihaylova[3], Yanbo Xue[4]
1. Department of Electrical and Computer Engineering, University of Canterbury, Christchurch NZ
2. State Key Lab of Complex Electromagnetic Environment Effects on Electronics and Information Systems, National University of Defense Technology, Changsha China
3. Department of Automatic Control and System Engineering, University of Sheffield, Sheffield UK
4. Career Science Lab, Beijing China
◇: Equal contributors, *: corresponding author, le.yang@canterbury.ac.nz

***Abstract*—This paper considers the fixed-interval smoothing for jump Markov systems. An optimal backward-time recursive equation for computing the joint posterior of the state vector and model index is established first. A suboptimal algorithm is then developed to approximate the new Bayesian smoother under nonlinear state-space models with additive Gaussian noise. The proposed method utilizes the well-known assumed density filtering with Gaussian assumption and the expression for the quotient of two Gaussian densities to compute the smoothing posterior. It eliminates the need for finding the inverse of the state dynamics and can handle singular process noise covariance, compared with several existing multiple model smoothers. Promising results are obtained in simulations using a maneuvering target tracking task.**

I. Introduction

Fixed-interval smoothing estimates the system states in a time interval using all the measurements collected within that interval [1], [2]. It provides better state estimates at the cost of some delay over filtering that utilizes only the measurements received before and up to the sampling time of the state [3]. For linear Gaussian state-space models, the Rauch-Tung-Striebel (RTS) smoother [4] gives the optimal solution. For nonlinear state-space models, Gaussian smoothing [5] or optimization-based methods [6] may be used.

The multiple model approach can be integrated into fixed-interval smoothing to improve performance, when abrupt changes are present in the system state due to load disturbances, additive faults [7], [8] and target maneuvering [9]. In this way, a bank of state-space models is used and the system switches among them according to a Markov chain at each sampling time. The considered system now becomes a jump Markov system. Fixed-interval smoothing for jump Markov systems has been applied in track extraction for multi-target tracking [10], maneuvering target tracking [11]–[14], sensor registration in radar networks [15], articulatory inversion [16] and air traffic control trajectory reconstruction [17].

A few algorithms are available for the fixed-interval smoothing for jump Markov systems. The two-filter approach [18], which is based on the optimal Frazer-Potter linear smoother [19], [20], combines the output of a forward-time interacting multiple model (IMM) filter and a backward-time IMM filter to find the smoothed states. Besides requiring the calculation of the inverse of the state dynamics, this technique assumes that the measurement likelihood can be well approximated using the state posterior computed by the backward-time IMM filter, which may not be satisfied in practice [2], [21].

In [10], a generalized pseudo-Bayesian of order 2 (GPB2)-type smoother was developed. It computes the joint posterior of the state vector and model index through assuming that the smoothed state is accurate enough so that its posterior can be approximated using the Dirac delta function. This degrades the performance when the noise becomes large. In [22], a IMM-type smoother was developed based on several strong assumptions on the state-space models [23]. The method from [23]–[25] combines the output of the forward-time IMM filter and RTS smoothers to produce the smoothed states. Evaluating the measurement likelihood is still needed when calculating the posterior probability of the model index. This is done by the backward-time IMM filter, similar to the two-filter approach. In [26], the algorithm in [23] was generalized to the case where the process and measurement noises are both non-Gaussian. In [27], we developed a closed-form algorithm for fixed-interval smoothing under linear Gaussian state-space models. It approximates directly the backward-time Bayesian recursive equation for calculating the joint posterior of the state vector and model index for enhanced performance. This method requires linear state dynamics and the process noise covariance being invertible, which limits its applications.

This paper significantly extends the algorithm in [27] so that we can handle nonlinear state dynamics and singular process noise covariance matrix. For this purpose, we first derive a new optimal backward-time recursive equation for calculating the joint posterior of the state vector and model index. Next, a suboptimal algorithm that approximates, using the assumed density filtering (ADF) with Gaussian assumption [5], the new Bayesian smoother is developed under nonlinear state-space models with additive Gaussian noise. The expression for the quotient of two Gaussian densities [28], [29] is utilized to obtain a closed-form solution. In the theoretical development, comparisons between the proposed method and the smoothers developed in [10], [23] and [27] is performed. Simulation experiments using a maneuvering target tracking task demonstrate the good performance of the proposed algorithm.

The rest of this paper is organized as follows. Section II presents the optimal backward-time recursive equation for computing the joint posterior of the state vector and model index. Section III gives the proposed smoothing algorithm under nonlinear state-space models. Section IV shows the simulation results. Section V concludes the paper.

II. Optimal Backward-time Recursive Smoothing

A. Problem Formulation

Let M_t^j denote that the jth state-space model of the jump Markov system in consideration is in effect during the sampling period $(t-1, t]$. Under model M_t^j, the system state $\mathbf{x}_t$ and measurements $\mathbf{z}_t$ at time instant t are sampled from the following model-matched state prediction probability density function (PDF) and measurement likelihood:

$$\mathbf{x}_t \sim p(\mathbf{x}_t|\mathbf{x}_{t-1}, M_t^j) \text{ and } \mathbf{z}_t \sim p(\mathbf{z}_t|\mathbf{x}_t, M_t^j). \tag{1}$$

Here, $j = 1, 2, ..., r$ and r is the number of models that the jump Markov system admits. M_t^j evolves according to a homogeneous Markov chain with transition probabilities

$$p(M_t^j|M_{t-1}^i) = p_{ij}, \tag{2}$$

where $i, j = 1, 2, ..., r$ and $\sum_{j=1}^{M} p_{ij} = 1$. The model switching is assumed to be independent of the state dynamics.

We need to obtain the state posterior $p(\mathbf{x}_t|\mathbf{z}_{1:k})$ to achieve fixed-interval smoothing, where $t = 1, 2, ..., k-1$, k is the interval length and $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_k\}$ represents the set of measurements collected within the time interval. For this purpose, we first find the joint posterior of $\mathbf{x}_t$ and M_t^j, $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$, and then marginalize out M_t^j using

$$p(\mathbf{x}_t|\mathbf{z}_{1:k}) = \sum_{j=1}^{r} p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}). \tag{3}$$

B. Optimal Recursive Equation for Computing $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$

We shall derive an optimal backward-time recursive equation for computing the joint posterior $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$, $j = 1, 2, ..., r$, to enable the evaluation of (3). To begin, we assume that the joint posterior of the state $\mathbf{x}_{t+1}$ and model index M_{t+1}^i at time instant $t+1$, which is $p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})$, $i = 1, 2, ..., r$, is already available. As a result, the joint posterior $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$ can be found via [10], [27]

$$p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}) = \sum_{i=1}^{r} \int p(\mathbf{x}_t, M_t^j, \mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})\mathrm{d}\mathbf{x}_{t+1} =$$
$$\sum_{i=1}^{r} \int p(\mathbf{x}_t, M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})\mathrm{d}\mathbf{x}_{t+1}. \tag{4}$$

To obtain the second equality in (4), we utilized the fact that due to the Markov property, $[\mathbf{x}_t^T, M_t^j]^T$ is independent of the measurements in $\{\mathbf{z}_{t+1}, \mathbf{z}_{t+2}, ..., \mathbf{z}_k\}$ given $[\mathbf{x}_{t+1}^T, M_{t+1}^i]^T$.

To evaluate the summands in (4), we note that

$$\begin{aligned} &p(\mathbf{x}_t, M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) \\ &= p(M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{x}_t, \mathbf{z}_{1:t}) p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) \\ &= p(M_t^j|M_{t+1}^i, \mathbf{x}_t, \mathbf{z}_{1:t}) p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}). \end{aligned} \tag{5}$$

Here, the second equality again comes from the Markov property that the model index in effect during $(t-1, t]$, M_j^t, becomes independent of the system state at time $t+1$, $\mathbf{x}_{t+1}$, when M_{t+1}^i and $\mathbf{x}_t$ are both given.

Putting (5) into (4) reveals that to find the summands in (4), we actually need to compute the integral

$$\int p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})\mathrm{d}\mathbf{x}_{t+1}. \tag{6}$$

Following the way leading to the second equality in (4) yields

$$p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) = p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:k}). \tag{7}$$

After applying (7) and substituting $p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k}) = p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:k}) p(M_{t+1}^i|\mathbf{z}_{1:k})$, we can re-write (6) as

$$\begin{aligned} &\int p(\mathbf{x}_t|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})\mathrm{d}\mathbf{x}_{t+1} \\ &\qquad = p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k}) p(M_{t+1}^i|\mathbf{z}_{1:k}). \end{aligned} \tag{8}$$

The expression for $p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k})$ is given in (9) shown at the bottom of this page, where the Bayes theorem is used to obtain the equality. Specifically, $p(\mathbf{x}_{t+1}|\mathbf{x}_t, M_{t+1}^i)$ is the prediction PDF of the state vector $\mathbf{x}_{t+1}$ under model M_{t+1}^i. Besides, $p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})$ is the model-matched filtering posterior of $\mathbf{x}_t$ *after* the model mixing [3], which is indeed computed during the forward-time multiple model filtering process. The denominator $p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t}) = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t, M_{t+1}^i) p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})\mathrm{d}\mathbf{x}_t$ is the normalization factor. Therefore, (9) has the same functional form as the single model optimal Bayesian smoothing equation [2], [30]. This enables using e.g., the computationally efficient nonlinear RTS smoother [5] to compute $p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k})$, which will be utilized in this work for smoothing algorithm development and has been used in the smoother in [23].

We proceed to write $p(M_t^j|M_{t+1}^i, \mathbf{x}_t, \mathbf{z}_{1:t})$ in (5) as

$$p(M_t^j|M_{t+1}^i, \mathbf{x}_t, \mathbf{z}_{1:t}) = \frac{p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t}) p(M_t^j|M_{t+1}^i, \mathbf{z}_{1:t})}{p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})}, \tag{10}$$

where $p(\mathbf{x}_t|M_{t+1}^i, M_t^j, \mathbf{z}_{1:t}) = p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})$ has been applied. $p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})$ is the model-matched filtering posterior of the state $\mathbf{x}_t$, and $p(M_t^j|M_{t+1}^i, \mathbf{z}_{1:t})$ is the mixing probability [3]. These two terms, as well as $p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})$ in the denominator, are all computed during the forward-time multiple

$$p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k}) = \int \frac{p(\mathbf{x}_{t+1}|\mathbf{x}_t, M_{t+1}^i) p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})}{p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t})} p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:k})\mathrm{d}\mathbf{x}_{t+1}. \tag{9}$$

model filtering process and can be considered known when the fixed-interval smoothing is performed.

Using the results in (5)-(10), we can express (4) as

$$p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}) = \sum_{i=1}^{r} \frac{p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k})}{p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t})} \cdot h_{t|k}^{ji}, \tag{11}$$

where $j = 1, 2, ..., r$, and

$$h_{t|k}^{ji} = p(M_t^j|M_{t+1}^i, \mathbf{z}_{1:t}) \cdot p(M_{t+1}^i|\mathbf{z}_{1:k}). \tag{12}$$

Note that the above theoretical development that leads to (11) is *exact*. This completes the derivation of the new optimal backward-time recursive equation for computing the joint posterior $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$.

C. Comparison with the Optimal Smoother in [10] and [27]

The optimal fixed-interval smoother, based on which the smoothing algorithms in [10] and [27] are developed, stems from (4) as well. It has a functional form different from that in (11) because when evaluating $p(\mathbf{x}_t, M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t})$ in (5), it applies the Bayes theorem directly to arrive at

$$\begin{aligned} &p(\mathbf{x}_t, M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) \\ &= \frac{p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{x}_t, M_t^j)p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:t})}{p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:t})} \\ &\propto \frac{p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{x}_t)p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})}{p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t})}. \end{aligned} \tag{13}$$

The term in the third row of the above equation, however, is not a posterior of the state $\mathbf{x}_t$. This can be verified by noting that the denominator can be expressed as

$$\begin{aligned} p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t}) &= \sum_{j=1}^{r} p(\mathbf{x}_{t+1}, M_t^j|M_{t+1}^i, \mathbf{z}_{1:t}) \\ &= \sum_{j=1}^{r} p(\mathbf{x}_{t+1}|M_{t+1}^i, M_t^j, \mathbf{z}_{1:t})p(M_t^j|M_{t+1}^i, \mathbf{z}_{1:t}), \end{aligned} \tag{14}$$

which is not equal to $\int p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{x}_t)p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})\mathrm{d}\mathbf{x}_t = p(\mathbf{x}_{t+1}|M_{t+1}^i, M_t^j, \mathbf{z}_{1:t})$. As a result, the integral in (4) cannot be evaluated using e.g, the nonlinear RTS smoother from [5].

To bypass the aforementioned difficulty, [10] introduces

$$\begin{aligned} &p(\mathbf{x}_t, M_t^j|\mathbf{x}_{t+1}, M_{t+1}^i, \mathbf{z}_{1:t}) \\ &\propto \frac{p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{x}_t)p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t})}{p(\mathbf{x}_{t+1}|M_{t+1}^i, M_t^j, \mathbf{z}_{1:t})} \frac{p(\mathbf{x}_{t+1}|M_{t+1}^i, M_t^j, \mathbf{z}_{1:t})}{p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t})}. \end{aligned} \tag{15}$$

The first term on the right hand side of (15) is now a posterior of $\mathbf{x}_t$. [10] further simplifies (15) through replacing $\mathbf{x}_{t+1}$ in the second term on its right hand side with the *mode* of the smoothing posterior $p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:k})$. As expected, the associated approximation error would increase when the noise level becomes larger. On the other hand, in [27], we put (13) into (4) and evaluated the result using the expression for the quotient of two Gaussian densities for linear Gaussian state-space models. But the application of the developed algorithm is limited to the scenario with linear state transition models and non-singular process noise covariance.

III. Proposed Smoothing Algorithm

A. Algorithm Development

This subsection gives a suboptimal fixed-interval smoothing algorithm that approximates the optimal Bayesian smoother in (11) when the state-space models the jump Markov system admits are nonlinear with additive Gaussian noise. In this case, the model-matched state prediction PDF and measurement likelihood in (1) would become

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}, M_t^j) = \mathcal{N}(\mathbf{x}_t; \mathbf{f}_j(\mathbf{x}_{t-1}), \mathbf{Q}_{t-1}^j), \tag{16a}$$

$$p(\mathbf{z}_t|\mathbf{x}_t, M_t^j) = \mathcal{N}(\mathbf{z}_t; \mathbf{h}_j(\mathbf{x}_t), \mathbf{R}_t^j), \tag{16b}$$

where $\mathbf{f}_j(\cdot)$ and $\mathbf{h}_j(\cdot)$ are the nonlinear state transition function and measurement function under model M_t^j. $\mathcal{N}(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate Gaussian PDF with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ on the random vector $\mathbf{y}$.

Assume that the forward-time filtering is carried out first using the IMM filter [3]. It employs the ADF with Gaussian assumption (i.e., Gaussian filtering [1], [2], [5], [31]) in place of the linear Kalman filter (KF) to handle the nonlinearity in the state-space models. Let the model-matched filtering posterior of the state and posterior probability of the model index obtained in the *update* stage of the IMM filter be

$$p(\mathbf{x}_t|M_t^j, \mathbf{z}_{1:t}) = \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|t}^j, \mathbf{P}_{t|t}^j), \tag{17a}$$

$$p(M_t^j|\mathbf{z}_{1:t}) = w_{t|t}^j, \tag{17b}$$

where $t = 1, 2, ..., k$ and $j = 1, 2, ..., r$. In the *prediction* stage of the IMM filter, the filtering posterior of $\mathbf{x}_t$ after model mixing, the mixing probability, and the model-matched prediction PDF of $\mathbf{x}_{t+1}$ are calculated. We denote them as

$$p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:t}) = \mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|t}^i, \bar{\mathbf{P}}_{t|t}^i), \tag{18a}$$

$$p(M_t^j|M_{t+1}^i, \mathbf{z}_{1:t}) = \bar{w}_{t|t}^{ji}, \tag{18b}$$

$$p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:t}) = \mathcal{N}(\mathbf{x}_{t+1}; \boldsymbol{\mu}_{t+1|t}^i, \mathbf{P}_{t+1|t}^i). \tag{18c}$$

The forward-time IMM filter stops at time instant k. The proposed algorithm then starts the backward-time smoothing with evaluating (11) using $t = k-1$ and $p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k}) = p(\mathbf{x}_k|M_k^i, \mathbf{z}_{1:k})p(M_k^i|\mathbf{z}_{1:k}) = \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{k|k}^i, \mathbf{P}_{k|k}^i) \cdot w_{k|k}^i$ (see (17)). The obtained result is substituted back into (11) to find $p(\mathbf{x}_{k-2}, M_{k-2}^j|\mathbf{z}_{1:k})$, $j = 1, 2, ..., r$. This process continues until $p(\mathbf{x}_1, M_1^j|\mathbf{z}_{1:k})$ is found and the desired fixed-interval multiple model smoothing is accomplished.

We are ready to present the steps for evaluating (11). As the joint posterior $p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k})$ is a scaled Gaussian PDF at $t = k-1$, we express it using the following general form:

$$\begin{aligned} p(\mathbf{x}_{t+1}, M_{t+1}^i|\mathbf{z}_{1:k}) &= p(\mathbf{x}_{t+1}|M_{t+1}^i, \mathbf{z}_{1:k})p(M_{t+1}^i|\mathbf{z}_{1:k}) \\ &= \mathcal{N}(\mathbf{x}_{t+1}; \boldsymbol{\mu}_{t+1|k}^i, \mathbf{P}_{t+1|k}^i) \cdot w_{t+1|k}^i, \end{aligned} \tag{19}$$

where $t = 1, 2, ..., k-1$.

Step-1: We evaluate $p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k})$ in (11) first. As pointed out below (8), its definition (9) has the same functional form as the single model optimal Bayesian smoother. From

(17), (18a), (18c) and (19), all the terms in the integrand in (9) are Gaussian densities. We can follow [5] to arrive at [23]

$$p(\mathbf{x}_t|M_{t+1}^i, \mathbf{z}_{1:k}) = \mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|k}^i, \bar{\mathbf{P}}_{t|k}^i), \tag{20}$$

where

$$\bar{\boldsymbol{\mu}}_{t|k}^i = \bar{\boldsymbol{\mu}}_{t|t}^i + \mathbf{G}_t^i(\boldsymbol{\mu}_{t+1|k}^i - \boldsymbol{\mu}_{t+1|t}^i), \tag{21a}$$

$$\bar{\mathbf{P}}_{t|k}^i = \bar{\mathbf{P}}_{t|t}^i + \mathbf{G}_t^i(\mathbf{P}_{t+1|k}^i - \mathbf{P}_{t+1|t}^i)(\mathbf{G}_t^i)^T. \tag{21b}$$

$\bar{\boldsymbol{\mu}}_{t|t}^i$ and $\bar{\mathbf{P}}_{t|t}^i$ are the mean and covariance of the filtering posterior of the state $\mathbf{x}_t$ after model mixing (see (18a)), while $\boldsymbol{\mu}_{t+1|t}^i$ and $\mathbf{P}_{t+1|t}^i$ are the mean and covariance of the prediction PDF of $\mathbf{x}_{t+1}$ given in (18c). The gain matrix $\mathbf{G}_t^i$ is

$$\mathbf{G}_t^i = \mathbf{C}_t^i(\mathbf{P}_{t+1|t}^i)^{-1}, \tag{22}$$

where $\mathbf{C}_t^i$ is the cross covariance equal to

$$\mathbf{C}_t^i = \int (\mathbf{x}_t - \bar{\boldsymbol{\mu}}_{t|t}^i)(\mathbf{f}_i(\mathbf{x}_t) - \boldsymbol{\mu}_{t+1|t}^i)^T \mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|t}^i, \bar{\mathbf{P}}_{t|t}^i)\mathrm{d}\mathbf{x}_t. \tag{23}$$

Step-2: Substituting (17), (18a) and (20) into (11) yields

$$p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}) = \sum_{i=1}^r \frac{\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|t}^j, \mathbf{P}_{t|t}^j)\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|k}^i, \bar{\mathbf{P}}_{t|k}^i)}{\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|t}^i, \bar{\mathbf{P}}_{t|t}^i)} \cdot h_{t|k}^{ji}. \tag{24}$$

We proceed to evaluate the summands in (24). Specifically, by the product rule for two Gaussian densities, we have that

$$\begin{aligned}&\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|t}^j, \mathbf{P}_{t|t}^j)\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|k}^i, \bar{\mathbf{P}}_{t|k}^i) \\ &\quad = \mathcal{N}(\boldsymbol{\mu}_{t|t}^j; \bar{\boldsymbol{\mu}}_{t|k}^i, \mathbf{P}_{t|t}^j + \bar{\mathbf{P}}_{t|k}^i)\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|k}^{ji}, \bar{\mathbf{P}}_{t|k}^{ji}),\end{aligned} \tag{25}$$

where

$$\bar{\boldsymbol{\mu}}_{t|k}^{ji} = \bar{\boldsymbol{\mu}}_{t|k}^i + \bar{\mathbf{P}}_{t|k}^i(\mathbf{P}_{t|t}^j + \bar{\mathbf{P}}_{t|k}^i)^{-1}(\boldsymbol{\mu}_{t|t}^j - \bar{\boldsymbol{\mu}}_{t|k}^i), \tag{26a}$$

$$\bar{\mathbf{P}}_{t|k}^{ji} = \left((\mathbf{P}_{t|t}^j)^{-1} + (\bar{\mathbf{P}}_{t|k}^i)^{-1}\right)^{-1}. \tag{26b}$$

Next, applying the expression for the quotient of two Gaussian densities [27]–[29]

$$\frac{\mathcal{N}(\mathbf{y}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)}{\mathcal{N}(\mathbf{y}; \boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)} = \frac{|\boldsymbol{\Sigma}_a|}{|\boldsymbol{\Sigma}_a - \boldsymbol{\Sigma}_c|} \cdot \frac{\mathcal{N}(\mathbf{y}; \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_b)}{\mathcal{N}(\boldsymbol{\mu}_a; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_a - \boldsymbol{\Sigma}_c)},$$

where $\boldsymbol{\Sigma}_b = (\boldsymbol{\Sigma}_c^{-1} - \boldsymbol{\Sigma}_a^{-1})^{-1}$ and $\boldsymbol{\mu}_b = \boldsymbol{\Sigma}_b(\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c - \boldsymbol{\Sigma}_a^{-1}\boldsymbol{\mu}_a)$, we arrive at

$$\frac{\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|k}^{ji}, \bar{\mathbf{P}}_{t|k}^{ji})}{\mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_{t|t}^i, \bar{\mathbf{P}}_{t|t}^i)} = \frac{|\bar{\mathbf{P}}_{t|t}^i|}{|\bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji}|} \frac{\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|k}^{ji}, \mathbf{P}_{t|k}^{ji})}{\mathcal{N}(\bar{\boldsymbol{\mu}}_{t|t}^i; \bar{\boldsymbol{\mu}}_{t|k}^{ji}, \bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji})}. \tag{27}$$

Here, $\boldsymbol{\mu}_{t|k}^{ji}$ and $\mathbf{P}_{t|k}^{ji}$ are equal to

$$\mathbf{P}_{t|k}^{ji} = \left((\bar{\mathbf{P}}_{t|k}^{ji})^{-1} - (\bar{\mathbf{P}}_{t|t}^i)^{-1}\right)^{-1}, \tag{28a}$$

$$\boldsymbol{\mu}_{t|k}^{ji} = \mathbf{P}_{t|k}^{ji}\left((\bar{\mathbf{P}}_{t|k}^{ji})^{-1}\bar{\boldsymbol{\mu}}_{t|k}^{ji} - (\bar{\mathbf{P}}_{t|t}^i)^{-1}\bar{\boldsymbol{\mu}}_{t|t}^i\right). \tag{28b}$$

Step-3: Putting (25) and (27) into (24) yields

$$p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}) = \sum_{i=1}^r d_{t|k}^{ji} \cdot \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|k}^{ji}, \mathbf{P}_{t|k}^{ji}), \tag{29}$$

where the weights $d_{t|k}^{ji}$ are defined as

$$d_{t|k}^{ji} = h_{t|k}^{ji} \cdot \frac{|\bar{\mathbf{P}}_{t|t}^i|}{|\bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji}|} \cdot \frac{\mathcal{N}(\boldsymbol{\mu}_{t|t}^j; \bar{\boldsymbol{\mu}}_{t|k}^i, \mathbf{P}_{t|t}^j + \bar{\mathbf{P}}_{t|k}^i)}{\mathcal{N}(\bar{\boldsymbol{\mu}}_{t|t}^i; \bar{\boldsymbol{\mu}}_{t|k}^{ji}, \bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji})}, \tag{30}$$

and $h_{t|k}^{ji} = \bar{w}_{t|t}^{ji} \cdot w_{t+1|k}^i$ according to (12), (18b) and (19).

It is clear from (29) that the joint smoothing posterior at time t would be a Gaussian mixture with r components, although the joint smoothing posterior at time $t+1$ is a scaled Gaussian density (see (19)). To maintain the computational tractability, we approximate $p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k})$ in (29) using

$$p(\mathbf{x}_t, M_t^j|\mathbf{z}_{1:k}) \approx \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|k}^j, \mathbf{P}_{t|k}^j) \cdot w_{t|k}^j. \tag{31}$$

such that it has the same functional form as (19). Invoking the method of moment matching [32], $\boldsymbol{\mu}_{t|k}^j$ and $\mathbf{P}_{t|k}^j$ are found via

$$\boldsymbol{\mu}_{t|k}^j = \sum_{i=1}^r \frac{d_{t|k}^{ji}}{\sum_{l=1}^r d_{t|k}^{jl}} \cdot \boldsymbol{\mu}_{t|k}^{ji}, \tag{32a}$$

$$\mathbf{P}_{t|k}^j = \sum_{i=1}^r \frac{d_{t|k}^{ji}}{\sum_{l=1}^r d_{t|k}^{jl}} \left(\mathbf{P}_{t|k}^{ji} + (\boldsymbol{\mu}_{t|k}^{ji} - \boldsymbol{\mu}_{t|k}^j)(\boldsymbol{\mu}_{t|k}^{ji} - \boldsymbol{\mu}_{t|k}^j)^T\right). \tag{32b}$$

The smoothing posterior probability $w_{t|k}^j$ is equal to

$$w_{t|k}^j = \frac{\sum_{i=1}^r d_{t|k}^{ji}}{\sum_{l=1}^r \sum_{i=1}^r d_{t|k}^{li}}. \tag{33}$$

This completes the development of the suboptimal fixed-interval algorithm under nonlinear state-space models with Gaussian noise. We can see that different from the smoother in [27], the proposed method does not require the inverse of the process noise covariance $\mathbf{Q}_{t-1}^i$ and the state transition can be nonlinear.

B. Algorithm Implementation

The validity of (29) requires that $\mathbf{P}_{t|k}^{ji}$ is positive definite. We can show that this is true when $t = k-1$. Specifically, from the definition of $\mathbf{P}_{t|k}^{ji}$ in (28a), we need to prove that $\bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji}$ is positive definite. Putting (26b) and (21b) yields

$$\begin{aligned}\bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^{ji} &= \bar{\mathbf{P}}_{t|t}^i - \left((\mathbf{P}_{t|t}^j)^{-1} + (\bar{\mathbf{P}}_{t|k}^i)^{-1}\right)^{-1} \\ &= \bar{\mathbf{P}}_{t|t}^i - \bar{\mathbf{P}}_{t|k}^i + \bar{\mathbf{P}}_{t|k}^i(\mathbf{P}_{t|t}^j + \bar{\mathbf{P}}_{t|k}^i)^{-1}\bar{\mathbf{P}}_{t|k}^i \\ &= \mathbf{G}_t^i(\mathbf{P}_{t+1|t}^i - \mathbf{P}_{t+1|k}^i)(\mathbf{G}_t^i)^T + \bar{\mathbf{P}}_{t|k}^i(\mathbf{P}_{t|t}^j + \bar{\mathbf{P}}_{t|k}^i)^{-1}\bar{\mathbf{P}}_{t|k}^i.\end{aligned} \tag{34}$$

The second equality in (34) is obtained via invoking the matrix inversion lemma. We can see that in the third equality of (34), the second term is positive definite. The first term, on the other hand, is at least positive semidefinite when $t = k-1$. This is because in this case, $\mathbf{P}_{t+1|t}^i - \mathbf{P}_{t+1|k}^i = \mathbf{P}_{k|k-1}^i - \mathbf{P}_{k|k}^i$ is the difference between the covariances of the prediction PDF and filtering posterior of the state $\mathbf{x}_k$. It is at least positive semidefinite under Gaussian filtering, which completes the

proof that $\bar{\mathbf{P}}^i_{t|t} - \bar{\mathbf{P}}^{ji}_{t|k}$ in (34) and as a result, $\mathbf{P}^{ji}_{t|k}$ in (28a) are both positive definite when $t = k-1$.

Nevertheless, the above observation may not hold for $t < k-1$, as observed in the simulation experiments. To address this aspect, the uncertainty-injection (UI) technique [33], [34] is employed to scale up the covariance $\bar{\mathbf{P}}^i_{t|t}$ using

$$\bar{\mathbf{P}}^i_{t|t} \leftarrow \lambda \cdot \bar{\mathbf{P}}^i_{t|t}, \tag{35}$$

where $\lambda > 1$ is determined by increasing its value until $\mathbf{P}^{ji}_{t|k}$ in (28a) becomes positive definite. In this case, we set $d^{ji}_{t|k}$ in (30) to be $d^{ji}_{t|k} = h^{ji}_{t|k} = \bar{w}^{ji}_{t|t} \cdot w^i_{t+1|k}$ (see the definition of $h^{ji}_{t|k}$ in (12) and the discussion under (30)).

The computation steps that the proposed fixed-interval smoother performs at each time instant $t < k$ are summarized in **Algorithm 1** shown below. It can be seen that to find the joint posterior for each model index M^j_t, we need to evaluate r smoothers in closed form. Thus, the proposed fixed-interval smoother indeed has a computational complexity close to that of the GPB2-type algorithms from [10], [27] and the method given in Algorithm 1 of [23].

Algorithm 1: One Step of the Proposed Smoother

Input: $\{\boldsymbol{\mu}^j_{t|t}, \mathbf{P}^j_{t|t}, w^j_{t|t}\}_{j=1,2,...,r}$, $\{\bar{w}^{ji}_{t|t}\}_{i=1,2,...,r,\ j=1,2,...,r}$
$\{\boldsymbol{\mu}^i_{t+1|t}, \mathbf{P}^i_{t+1|t}\}_{i=1,2,...,r}$, $\{\bar{\boldsymbol{\mu}}^i_{t|t}, \bar{\mathbf{P}}^i_{t|t}\}_{i=1,2,...,r}$
$\{\boldsymbol{\mu}^i_{t+1|k}, \mathbf{P}^i_{t+1|k}, w^i_{t+1|k}\}_{i=1,2,...,r}$

Output: $\{\boldsymbol{\mu}^j_{t|k}, \mathbf{P}^j_{t|k}, w^j_{t|k}\}_{j=1,2,...,r}$

for $i = 1, 2, ..., r$ **do**
  Calculate $\mathbf{C}^i_t$ and $\mathbf{G}^i_t$ using (23) and (22);
  Compute $\bar{\boldsymbol{\mu}}^i_{t|k}$ and $\bar{\mathbf{P}}^i_{t|k}$ using (21);
end
for $j = 1, 2, ..., r$ **do**
  for $i = 1, 2, ..., r$ **do**
    Compute $\bar{\boldsymbol{\mu}}^{ji}_{t|k}$ and $\bar{\mathbf{P}}^{ji}_{t|k}$ using (26);
    $\lambda \leftarrow 1$;
    Increase λ until $\lambda \cdot \bar{\mathbf{P}}^i_{t|t} - \bar{\mathbf{P}}^{ji}_{t|k}$ is positive definite;
    Evaluate, with $\bar{\mathbf{P}}^i_{t|t}$ replaced by $\lambda \cdot \bar{\mathbf{P}}^i_{t|t}$, (28) and (30) to find $\boldsymbol{\mu}^{ji}_{t|k}$, $\mathbf{P}^{ji}_{t|k}$ and $d^{ji}_{t|k}$;
    if $\lambda > 1$ **then** $d^{ji}_{t|k} = \bar{w}^{ji}_{t|t} \cdot w^i_{t+1|k}$;
  end
  Calculate $\boldsymbol{\mu}^j_{t|k}$ and $\mathbf{P}^j_{t|k}$ using (32);
  Find $w^j_{t|k}$ using (33);
end

C. Comparison with the Smoother from Algorithm 1 of [23]

The smoother presented in Algorithm 1 of [23] is closely related to the technique developed in Section III.A in the sense that it also uses $\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{ji}_{t|k}, \mathbf{P}^{ji}_{t|k})$ given in (27) to evaluate the model-matched state smoothing posterior $p(\mathbf{x}_t|M^j_t, \mathbf{z}_{1:k})$. But it computes $p(\mathbf{x}_t|M^j_t, \mathbf{z}_{1:k})$ and the posterior probability of the model index, $p(M^j_t|\mathbf{z}_{1:k})$, *sequentially*, instead of finding them jointly as in (29). Sepcifically, the Algorithm 1 from [23] finds $p(\mathbf{x}_t|M^j_t, \mathbf{z}_{1:k})$ via

$$p(\mathbf{x}_t|M^j_t, \mathbf{z}_{1:k}) = \sum_{i=1}^{r} \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{ji}_{t|k}, \mathbf{P}^{ji}_{t|k}) p(M^i_{t+1}|M^j_t, \mathbf{z}_{1:k}), \tag{36}$$

where $p(M^i_{t+1}|M^j_t, \mathbf{z}_{1:k})$ is approximated using

$$p(M^i_{t+1}|M^j_t, \mathbf{z}_{1:k}) \propto p_{ji} \cdot \mathcal{N}(\boldsymbol{\mu}^j_{t|t}; \boldsymbol{\mu}^{b,i}_{t|k}, \mathbf{P}^j_{t|t} + \mathbf{P}^{b,i}_{t|k}). \tag{37}$$

Here, p_{ji} is the model transition probability defined in (2), and $\boldsymbol{\mu}^j_{t|t}$ and $\mathbf{P}^j_{t|t}$ are the mean and covariance of the model-matched filtering posterior of $\mathbf{x}_t$ (see (17)). $\boldsymbol{\mu}^{b,i}_{t|k}$ and $\mathbf{P}^{b,i}_{t|k}$ are the mean and covariance of the maximum likelihood estimate $\text{argmax}_{\mathbf{x}_t} p(\mathbf{z}_{t+1:k}|\mathbf{x}_t, M^i_{t+1})$. They are equal to [18], [23]

$$\mathbf{P}^{b,i}_{t|k} = \left((\bar{\mathbf{P}}^i_{t|k})^{-1} - (\bar{\mathbf{P}}^i_{t|t})^{-1}\right)^{-1}, \tag{38a}$$

$$\boldsymbol{\mu}^{b,i}_{t|k} = \mathbf{P}^{b,i}_{t|k}\left((\bar{\mathbf{P}}^i_{t|k})^{-1}\bar{\boldsymbol{\mu}}^i_{t|k} - (\bar{\mathbf{P}}^i_{t|t})^{-1}\bar{\boldsymbol{\mu}}^i_{t|t}\right), \tag{38b}$$

where $\bar{\boldsymbol{\mu}}^i_{t|t}$ and $\bar{\mathbf{P}}^i_{t|t}$ are defined in (18a), while $\bar{\boldsymbol{\mu}}^i_{t|k}$ and $\bar{\mathbf{P}}^i_{t|k}$ are given in (21).

$\mathbf{P}^{b,i}_{t|k}$ in (38a) may not always be positive definite[1]. In fact, we can show by following the same approach leading to (34) that substituting (21b) yields

$$\bar{\mathbf{P}}^i_{t|t} - \bar{\mathbf{P}}^i_{t|k} = \mathbf{G}^i_t(\mathbf{P}^i_{t+1|t} - \mathbf{P}^i_{t+1|k})(\mathbf{G}^i_t)^T. \tag{39}$$

In other words, the inverse of $\mathbf{P}^{b,i}_{t|k}$ would only be positive semidefinite when $t = k-1$, if the number of measurements in $\mathbf{z}_k$ is smaller than the dimensionality of the state $\mathbf{x}_k$. In this case, [23] considers the Gaussian density $\mathcal{N}(\boldsymbol{\mu}^j_{t|t}; \boldsymbol{\mu}^{b,i}_{t|k}, \mathbf{P}^j_{t|t} + \mathbf{P}^{b,i}_{t|k})$ in (37) to be a *flat* prior such that $p(M^i_{t+1}|M^j_t, \mathbf{z}_{1:k}) \approx p_{ji}$. At the same time, the posterior probability for the model index is set to be $p(M^j_t|\mathbf{z}_{1:k}) = w^j_{t|t}$, $j = 1, 2, ...r$, which are just the model posterior probabilities computed by the forward-time IMM filter. This approximation may lead to the performance degradation if $\mathbf{P}^{b,i}_{t|k}$ being singular occurs frequently (see the Simulation Results section).

IV. Simulation Results

A. Simulation Setup

We adopt a simulation scenario very similar to the one employed in [27], [35]. A stationary sensor at the origin measures the range and bearing of a target to estimate its trajectory. At sampling instant t, the measurements are related to the target state vector $\mathbf{x}_t = [x_t, y_t, \dot{x}_t, \dot{y}_t]^T$ through

$$\mathbf{y}_t = \begin{bmatrix} \sqrt{x_t^2 + y_t^2} \\ \tan^{-1}(x_t/y_t) \end{bmatrix} + \mathbf{w}_t, \tag{40}$$

where $[x_t, y_t]^T$ is the target position, $[\dot{x}_t, \dot{y}_t]^T$ is the target velocity, and $\tan^{-1}(\cdot)$ denotes the four-quadrant inverse tangent function. $\mathbf{w}_t$ is the zero-mean Gaussian measurement noise vector with covariance $\mathbf{R} = \text{diag}(\sigma_r^2, \sigma_\theta^2)$, where $\sigma_r = 100$m and $\sigma_\theta = 0.5^o$.

[1]Compared with $\mathbf{P}^{b,i}_{t|k}$, $\mathbf{P}^{ji}_{t|k}$ is less likely to be singular, due to the presence of $\mathbf{P}^j_{t|t}$ (see (34) and the discussion below).

To realize the fixed-interval smoothing of the target state $\mathbf{x}_t$, we collect $k = 200$ measurements with a sampling period of $T = 3$s over an interval of 600s. At the beginning of this time interval, the target is located at $[234.92\text{km}, 85.50\text{km}]^T$ and moving with an initial velocity $[-141.4\text{m/s}, -141.4\text{m/s}]^T$. Within the interval, the target motion follows the constant velocity (CV) model most of the time. Mathematically, in this case, the target state $\mathbf{x}_t$ evolves according to [9]

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{v}_{t-1}, \tag{41}$$

where the state transition matrix $\mathbf{F}$ is

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_2 & T \cdot \mathbf{I}_2 \\ \mathbf{O} & \mathbf{I}_2 \end{bmatrix}. \tag{42}$$

The target makes two turns within the time interval, according to the following constant turn (CT) model [9]

$$\mathbf{x}_t = \mathbf{F}(\omega) \cdot \mathbf{x}_{t-1} + \mathbf{v}_{t-1}, \tag{43}$$

where the state transition matrix $\mathbf{F}(w)$ is

$$\mathbf{F}(\omega) = \begin{bmatrix} 1 & 0 & \sin(\omega T)/\omega & -(1-\cos(\omega T))/\omega \\ 0 & 1 & (1-\cos(\omega T))/\omega & \sin(\omega T)/\omega \\ 0 & 0 & \cos(\omega T) & -\sin(\omega T) \\ 0 & 0 & \sin(\omega T) & \cos(\omega T) \end{bmatrix}. \tag{44}$$

The first turn lasts for 18 seconds from 200s to 218s, and it has an acceleration of 1g with a turn rate of $\omega = -0.05$rad/s. The second turn lasts for two minutes from 480s to 600s, and it has an acceleration of 0.5g with a turn rate of $\omega = 0.022$rad/s. Fig. 1 shows a sample target motion trajectory.

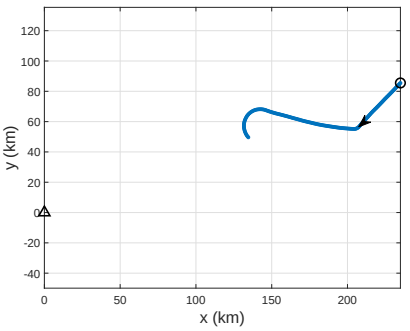

Fig. 1. A sample target trajectory. The triangle denotes the sensor. The circle gives the initial position of the target. The blue curve is the target trajectory.

In (41) and (43), $\mathbf{v}_{t-1}$ is the Gaussian process noise with zero mean and covariance $\mathbf{Q} = \sigma_v^2 \mathbf{G}\mathbf{G}^T$, where $\mathbf{G} = \left[\frac{1}{2}T^2 \cdot \mathbf{I}_2, T \cdot \mathbf{I}_2\right]^T$ and $\sigma_v = 1$m/s.

We compare the estimation accuracy of the fixed-interval smoother proposed in Section III.A with that of five benchmark methods. They include the forward-time IMM filter [3], two-filter approach [18], GPB2-type technique from [10], the smoother from Algorithm 1 of [23] and the method from [27][2].

[2]The process noise covariance $\mathbf{Q} = \sigma_v^2 \mathbf{G}\mathbf{G}^T$ is *singular*, which renders the smoother proposed in [27] inapplicable here. To allow the performance comparison, diagonal loading is applied such that the smoother from [27] in fact uses $\mathbf{Q} + \text{blkdiag}(T^2 \cdot \mathbf{I}_2, \mathbf{O}_2)$ as the process noise covariance.

All the algorithms employ the same bank of $r = 7$ state-space models. These models share the measurement equation in (40), but their process equations are different. Specifically, one state-space model adopts the CV model in (41) while the other six state-space models use the CT models in (43) with turn rates set to ± 0.02rad/s, ± 0.033rad/s and ± 0.1rad/s. The model switching is assumed to follow a Markov chain with a known transition probability matrix $\mathbf{P}$ whose diagonal elements are 0.8 and off-diagonal elements are 0.0333.

The output of the forward-time IMM filter is utilized by the five simulated smoothers to achieve fixed-interval state smoothing. We initialize the IMM filter using the target range and bearing measurements obtained at the first sampling time. The initialization process is the same as the one used in [27], [35] and thus, it is omitted here for brevity.

To handle the nonlinearity in the measurement equation in (40), the Gaussian filtering framework using the cubature rules-based numerical integration [36] is applied. Note that in the current simulation setup, the state dynamics are still linear. The purpose of this setting is to enable the performance comparison between the proposed technique and the smoother from [27], as well as the two-filter approach that requires finding the inverse of the state dynamics. When the state dynamics become nonlinear, to implement the proposed smoother in Section III.A, the cross covariance $\mathbf{C}_t^i$ defined in (23) may need to be evaluated numerically as well. In this case, again, the cubature rules could be used to compute $\mathbf{C}_t^i$.

B. Results and Analysis

We quantify the performance of the six algorithms in consideration in terms of their target position estimation root mean square errors (RMSEs) and target velocity estimation RMSEs. The results are shown in Figs. 2 and 3, and they are obtained through averaging over 2000 Monte Carlo ensemble runs. In the figures, three vertical lines are added to indicate the starting time (200s) and ending time (218s) of the first turn, as well as the starting time (480s) of the second turn.

According to Figs. 2 and 3, the five smoothers all provide evident performance gain over the forward-time IMM filter ('IMM Filter'), which is expected because more measurements are explored in smoothing. Besides, the smoothers offer similar target velocity RMSEs but the GPB2-type smoother from [10] ('GPB2 Smoother') yields the largest target position RMSEs. This may come from the use of the approximation that the smoothing state posterior is concentrated around its mode (see the discussion below (15)), which may become invalid under the relatively large noise level used in the simulation. The performance of the fixed-interval smoother proposed in this paper ('Proposed') is close to that of the two-filter approach [18] ('Two-Filter') and the smoother from [27] ('Method from [27]'). But the new algorithm eliminates the need to find the inverse of the state dynamics and no longer requires the process noise covariance being invertible. Besides, they are superior to the smoother from Algorithm 1 of [23] ('Method from [23]') in terms of greatly reduced target position RMSEs.

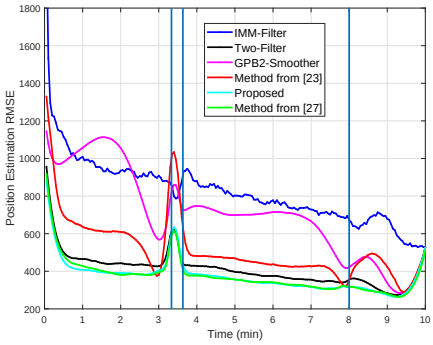

Fig. 2. Comparison of target position RMSEs.

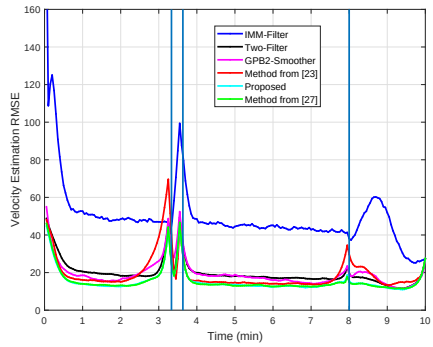

Fig. 3. Comparison of target velocity RMSEs.

The observed performance improvement over the smoother from [23] may be explained by examining the *average* posterior probabilities of the model indexes. For this purpose, we plot in Figs. 4, 5 and 6 the posterior probabilities for the CV model, CT model with turn rate $w = -0.033$rad/s and CT model with $w = 0.02$rad/s. The two CT models are selected because they have turn rates close to those of the first turn ($w = -0.05$rad/s) and the second turn ($w = 0.022$rad/s).

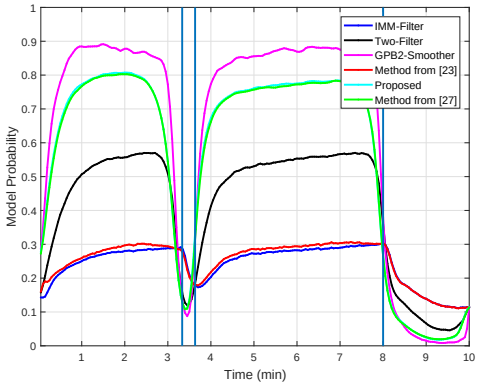

Fig. 4. Comparison of the CV model probabilities.

We can see from Figs. 4-6 that the posterior model probability found by the smoother from [23] is very close to that of

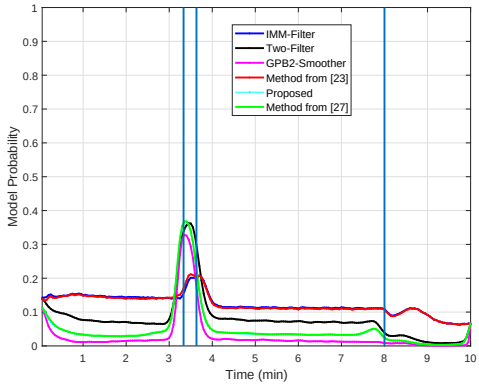

Fig. 5. Comparison of the CT model probabilities ($\omega = -0.033$rad/s).

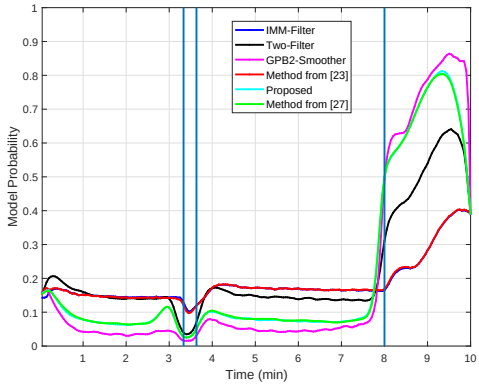

Fig. 6. Comparison of the CT model probabilities ($\omega = 0.02$rad/s).

the forward-time IMM filter. This indicates that the covariance matrices of the maximum likelihood estimates, which are $\mathbf{P}^{b,i}_{t|k}$ given in (38a), become non-invertible frequently (see Section III.C). The underlying reason is possibly that we use a relatively large bank of $r = 7$ state-space models, which increases the chance that there is at least one singular covariance matrix $\mathbf{P}^{b,i}_{t|k}$.

On the other hand, the proposed smoother, as well as the one from [27] and the two-filter approach, better identifies the target motion mode. Specifically, from 1s to 200s and from 218s to 480s when the target motion follows the CV model, the posterior CV model probabilities calculated by the proposed smoother are significantly bigger than that of the method from [23] (see Fig. 4). The observations from Figs. 5 and 6 are similar. They demonstrate, respectively, that the proposed smoother performs better in finding the correct target motion mode during the first turn from 200s to 218s and the second turn from 480s to 600s. The improved performance may be owing to two factors. First, the covariance of the model-matched smoothing posterior computed by the proposed smoother, which is $\mathbf{P}^{ji}_{t|k}$ in (28a), is less likely to be singular (see the analysis in Section III.B). Second, even when $\mathbf{P}^{ji}_{t|k}$ is singular, thanks to the optimal smoother (11) it is based

on, the proposed smoother uses $d^{ji}_{t|k} = \bar{w}^{ji}_{t|t} \cdot w^{i}_{t+1|k}$ to calculate the model probability (see (33) and the discussion under (35)). This still enables the utilization of the information from the forward-time IMM filter contained in the mixing probability $\bar{w}^{ji}_{t|t}$ and the information from smoothing included in $w^{i}_{t+1|k}$. In contrast, the method of [23] explores the model probability obtained during the forward-time IMM filtering only when one covariance $\mathbf{P}^{b,i}_{t|k}$ is not invertible. This may lead to the degraded estimation performance observed in Fig. 2.

V. Conclusions

For the problem of fixed-interval state smoothing for jump Markov systems, this paper first derived a new optimal backward-time recursive equation that can calculate the posterior of the state vector and model index jointly. We then developed a closed-form algorithm that uses Gaussian filtering techniques to approximate the newly obtained Bayesian smoother under nonlinear state-space models with additive Gaussian noise. The proposed optimal smoother and the suboptimal smoothing algorithm were contrasted against several existing approaches to highlight their underlying differences. The good performance of the proposed smoothing algorithm in a maneuvering target tracking task in terms of improved estimation accuracy over multiple benchmark methods was illustrated using simulations.

References

[1] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[2] S. Särkkä, *Bayesian Filtering and Smoothing*. New York: Cambridge University Press, 2013.

[3] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: Theory, algorithms and software*. New York: Wiley, 2001.

[4] H. Rauch, F. Tung, and C. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3, no. 8, pp. 1445–1450, Aug. 1965.

[5] S. Särkkä and J. Hartikainen, "On Gaussian optimal smoothing of nonlinear state space models," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1938–1941, Aug. 2010.

[6] A. Aravkin, J. Burke, and G. Pillonetto, "Optimization viewpoint on Kalman smoothing with applications to robust and sparse estimation," in *Compressed Sensing & Sparse Filtering*, A. Carmi, L. S. Mihaylova, and S. Godsill, Eds. Springer, 2014, pp. 237–280.

[7] H. Ohlsson, F. Gustafsson, L. Ljung, and S. Boyd, "State smoothing by sum-of-norms regularization," in *Proc. IEEE Conf. Decision and Control (CDC)*, Atlanta, GA, USA, Dec. 2010, pp. 2880–2885.

[8] ——, "Smoothed state estimates under abrupt changes using sum-of-norms regularization," *Automatica*, vol. 48, pp. 595–605, Aug. 2012.

[9] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.

[10] W. Koch, "Fixed interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, pp. 2–14, Jan. 2000.

[11] V. P. Jilkov, X. R. Li, and L. Lu, "Performance enhancement of the IMM estimation by smoothing," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Annapolis, MD, USA, Jul. 2002, pp. 713–720.

[12] M. Morelande and B. Ristic, "Smoothed state estimation for nonlinear Markovian switching systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, pp. 1309–1325, Oct. 2008.

[13] M. Malleswaran, V. Vaidehi, S. Irwin, and B. Robin, "IMM-UKF-TFS model-based approach for intelligent navigation," *The Journal of Navigation*, vol. 66, pp. 859–877, Nov. 2013.

[14] W. Ali, Y. Li, M. Raja, and N. Ahmed, "Generalized pseudo Bayesian algorithms for tracking of multiple model underwater maneuvering target," *Applied Acoustics*, vol. 166, pp. 1–12, Sept. 2020.

[15] Z. Li and H. Leung, "An expectation maximization-based simultaneous registration and fusion algorithm for radar networks," in *Proc. Canadian Conf. on Electrical and Computer Engineering (CCECE)*, Ottawa, ON, Canada, May 2006, pp. 31–35.

[16] İ. Özbek, M. Hasegawa-Johson, and M. Demirekler, "On improving dynamic state space approaches to articulatory inversion with MAP based parameter estimation," *IEEE Trans. Audio, Speech and Language Process.*, vol. 20, pp. 67–81, Jan. 2012.

[17] J. García, J. A. Besada, J. M. Molina, and G. de Miguel, "Model-based trajectory reconstruction with IMM smoothing and segmentation," *Information Fusion*, vol. 22, pp. 127–140, Mar. 2015.

[18] R. Helmick, W. Blair, and S. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Trans. Inf. Theory*, vol. 41, pp. 1845–1855, Jun. 1995.

[19] D. C. Fraser and J. E. Potter, "The optimal linear smoother as a combination of two optimum linear filters," *IEEE Trans. Automatic Control*, vol. 14, pp. 387–390, Aug. 1969.

[20] D. Q. Mayne, "A solution of the smoothing problem for linear dynamic systems," *Automatica*, vol. 4, pp. 73–92, Dec. 1966.

[21] M. Briers, A. Doucet, and S. Maskell, "Smoothing algorithms for state-space models," *Annals of the Institute of Statistcal Mathematics*, vol. 62, pp. 61–89, 2010.

[22] N. Nadarajah, R. Tharmarasa, M. McDonald, and T. Kirubarajan, "IMM forward filtering and backward smoothing for manuevering target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, pp. 2673–2678, Jul. 2012.

[23] R. Lopez and P. Danès, "Low-complexity IMM smoothing for jump Markov nonlinear systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, pp. 1261–1272, Jun. 2017.

[24] ——, "Exploiting Rauch-Tung-Striebel formulae for IMM-based smoothing of Markovian switching systems," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 3953–3956.

[25] ——, "A fixed-interval smoother with reduced complexity for jump Markov nonlinear systems," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Salamanca, Spain, Jul. 2014, pp. 1–8.

[26] Y. Yang, Y. Qin, Y. Yang, Q. Pan, and Z. Li, "A Gaussian mixture smoother for Markovian jump linear systems with non-Gaussian noises," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Cambridge, UK, Jul. 2018, pp. 2564–2571.

[27] X. Li, Y. Liu, L. Yang, L. S. Mihaylova, and B. Deng, "Enhanced fixed-interval smoothing for Markovian switching systems," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Sun City, South Africa, Nov. 2021.

[28] J. M. Hernández-Lobato, "Balancing flexibility and robustness in machine learning: Semiparametric methods and sparse linear models," Ph.D. dissertation, Universidad Autónoma de Madrid, 2010.

[29] D. Acar and U. Orguner, "Information decorrelation for an interacting multiple model filter," in *Proc. Intl. Conf. Information Fusion (FUSION)*, Cambridge, UK, Jul. 2018, pp. 1527–1534.

[30] G. Kitagawa, "Non-Gassian state-space modeling of nonstationary time series," *J. of the American Statistical Association*, vol. 82, pp. 1032–1041, Dec. 1987.

[31] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 45, pp. 910–927, May 2000.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.

[33] S. V. Vaerenbergh, M. Lázaro-Gredilla, and I. Santamaría, "Kernel recursive least-squares tracker for time-varying regression," *IEEE Trans. Neural Networks and Learning Systems*, vol. 23, pp. 1313–1326, Aug. 2012.

[34] L. Yang, K. Wang, and L. S. Mihaylova, "Online sparse multi-output Gaussian process regression and learning," *IEEE Trans. Signal and Info. Process. over Networks*, vol. 5, pp. 258–272, Jun. 2019.

[35] X. Li, Y. Liu, L. S. Mihaylova, L. Yang, S. Weddell, and F. Guo, "Enhanced multiple model GPB2 filtering using variational inference," in *Proc. Intl. Conf. Information Fusion (FUSION)*, Ottawa, ON, Canada, Jul. 2019, pp. 2572–2579.

[36] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.