Low-Complexity IMM Smoothing for Jump Markov Nonlinear Systems

Rémy Lopez, Patrick Danès

Low-Complexity IMM Smoothing for Jump Markov Nonlinear Systems

Rémy Lopez⋆ and Patrick Danès†
⋆ Collecte Localisation Satellites; 11, rue Hermès, 31520 Ramonville Saint-Agne, France
† CNRS, LAAS; Univ. Toulouse; Toulouse, France
{rlopez@cls.fr, patrick.danes@laas.fr}

***Abstract*—A suboptimal algorithm to fixed-interval and fixed-lag smoothing for Markovian switching systems is proposed. It infers a Gaussian mixture approximation of the smoothing pdf by combining the statistics produced by an IMM filter into an original backward recursive process. The number of filters and smoothers is equal to the constant number of hypotheses in the posterior mixture. A comparison, conducted on simulated case studies, shows that the investigated method performs significantly better than equivalent algorithms.**

***Index Terms*—Nonlinear Markovian switching systems; Interacting Multiple Model (IMM) filtering and smoothing; Rauch-Tung-Striebel formulae; Target tracking.**

I. Introduction

Many estimation or change detection problems are stated in the context of discrete-time jump Markov systems. Such systems are described by a bank of state space models, sharing the same state vector and corresponding to admissible modes of operation, together with a finite-state Markov chain featuring the transitions between modes. At each time k, the exact posterior probability density function (pdf) of the state vector conditioned on the measurements up to time k' comes as a mixture of the set of all posterior pdfs conditioned on the observations up to k' and on the possible mode sequences up to k, weighted by the posterior probability of these mode sequences. The computational complexity thus grows exponentially with k, so that approximations are needed to make the problem tractable [1][2].

In the filtering context, *i.e.* when $k = k'$, the number of hypotheses composing the above mixture can be reduced by merging those ones which are conditioned on similar mode sub-sequences up to time $k-n$. Generalized Pseudo-Bayesian filters of order n (GPB$_n$) fall into this paradigm. For a bank of M models, they involve M^n filters. However, the most standard approach is undoubtedly the Interacting Multiple Model (IMM) filter [3], which propagates over time a M-hypotheses Gaussian mixture approximation to the posterior pdf at the complexity of GPB$_1$, but with a performance similar to GPB$_2$. Though initially designed for linear jump Markov systems, GPB$_n$ and IMM are widely used in the nonlinear case [4][5]. They can rely on extended [6] or unscented [7] (mode-conditioned) Kalman filters, or can be applied to non-Gaussian state-space models with particle filters [8]. IMM filtering is still an active research area, see for instance its recent independent extensions to heterogeneous-order models, *i.e.* to models which share only parts of their state vectors [9][10].

Fixed-interval smoothing refers to estimating the posterior pdf of the state vector at each time k when measurements are assimilated over a fixed interval of length $T \geq k$. In contrast, fixed-lag smoothing considers the estimation of the state posterior pdf at each time k from measurements up to $k+n$, where n denotes the fixed lag length. Smoothing constitutes a fundamental problem as it helps to improve the estimation performance in comparison to filtering, though at the cost of some delay. In the field of target tracking for instance, delivering a location estimate by assimilating subsequent observations drastically reduces the associated error [11]. In the single-model non-Gaussian case, many schemes were considered, either based on particle filters [12], [13], [14], [15] or within the Random Finite Set paradigm [16]. Under Gaussian or Gaussian sum approximations with jump Markov systems, closed-form solutions to fixed-interval smoothing were proposed in [17], [18], [19].

The aim of this work is to show how the quantities produced by a forward-time IMM filter up to time k' enable a closed-form approximation of the smoothing posterior density at times $k < k'$ by requiring only M filters/smoothers for a bank of M models. The proposed method extends the papers [20], [21] in that it enriches the original algorithm. In comparison, [18], [17] run M^2 smoothers for a bank of M models. Ref. [19] runs M smoothers but displays significantly lower performances than the investigated method.

The paper is organized as follows. Section II states a fixed-interval multiple model smoothing problem. Then, Section III reviews the theoretical foundations of the proposed strategy and positions it with respect to the literature. The main result, *i.e.* a constructive IMM-based fixed-interval smoothing algorithm, constitutes Section IV. Fixed-lag smoothing is considered in Section V. After simulation examples in Section VI comparing the proposed method to the equivalent existing algorithms [17], [18], [19], the paper ends with a conclusion and prospects.

II. Problem Statement

Notations are standard. $(.)^T$ denotes the transpose operator. $P(.)$, $p(.)$ and $E[.]$ respectively term a probability, a probability density function (pdf), and an expectation. $\mathcal{N}(\bar{x}, X)$ stands for the (real) Gaussian distribution with mean $\bar{x}$ and covariance X and $\mathcal{N}(x; \bar{x}, X)$ is the associated pdf on x. The weighted squared norm $\|a\|_R^2 = a^T R a$, with R a symmetric positive definite matrix, is also referred to throughout the text.

The considered nonlinear jump Markov system admits M modes, which constitute the set $\mathscr{M}$. At each time k, $m_k = j$ or m_k^j denotes the event that mode $j \in \mathscr{M}$ is in effect during the sampling period $(t_{k-1}, t_k]$. The sequence of modes follows an homogeneous finite-state Markov chain. Under the event m_k^j, the dynamics of the base (continuous) state x_k and its relationship with the measurement z_k are described by the stochastic nonlinear state space model

$$x_k = f_{k-1}^j(x_{k-1}) + q_{k-1}^j, \qquad z_k = h_k^j(x_k) + r_k^j, \tag{1}$$

where $f_{k-1}^j(.)$, $h_k^j(.)$ are given and q_{k-1}^j, r_k^j account for dynamics and measurement noises. The global (hybrid) state vector at time k will henceforth be termed $\xi_k = (x_k, m_k)$. ξ_k^j will stand for (x_k, m_k^j).

The (given) initial and transition probabilities of modes are

$$P(m_0^j) = \mu_0^j; \; P(m_{k+1}^i | m_k^j) = \pi_{ji}. \tag{2}$$

Similarly, conditioned on mode j, the base state vector at initial time $k = 0$ and the noises are assumed jointly Gaussian and of (given) statistics, with $\delta_{k,k'}$ the Kronecker symbol,

$$\forall k, k', \; E \begin{bmatrix} x_0 | m_0^j \\ q_k^j \\ r_{k'}^j \end{bmatrix} = \begin{pmatrix} \hat{x}_{0|0}^j \\ 0 \\ 0 \end{pmatrix}; \tag{3}$$

$$E \left[\begin{pmatrix} x_0 | m_0^j - \hat{x}_{0|0}^j \\ q_k^j \\ r_k^j \end{pmatrix} \begin{pmatrix} x_0 | m_0^j - \hat{x}_{0|0}^j \\ q_{k'}^j \\ r_{k'}^j \end{pmatrix}^T \right] = \begin{pmatrix} P_{0|0}^j & 0 & 0 \\ 0 & Q_k^j \delta k,k' & 0 \\ 0 & 0 & R_k^j \delta k,k' \end{pmatrix}. \tag{4}$$

As a result, the pdf of the base state x_0 at initial time is a Gaussian mixture. The transition and observation densities associated to (1) and conditioned on the active mode m_k^j active in the sampling interval $(t_{k-1}, t_k]$ write as

$$p(x_k | x_{k-1}, m_k^j) = \mathscr{N}(x_k; f_{k-1}^j(x_{k-1}), Q_{k-1}^j), \tag{5}$$

$$p(z_k | x_k, m_k^j) = \mathscr{N}(z_k; h_k^j(x_k), R_k^j). \tag{6}$$

As aforementioned, a mixture with an exponentially increasing number of hypotheses (densities) would be required in the filtering pdf at further time k, in that $p(x_k|z_{1:k}) = \sum_{j_{0:k} \in \mathscr{M}^{k+1}} p(x_k | m_{0:k} = j_{0:k}, z_{1:k}) P(m_{0:k} = j_{0:k} | z_{1:k})$, where $v_{a:b}$ is a shortcut for the sequence $v_a, \ldots, v_b$ between times a and b. A similar exponential complexity in the number of modes occurs in the exact form of the smoothing density, be it fixed-interval (*i.e.* $p(x_k|z_{1:T})$, with $T \geq k \geq 0$ the fixed interval length), fixed lag (*i.e.* $p(x_k|z_{1:k+n})$, with $n \geq 1$ the fixed lag length) or fixed-point (*i.e.* $p(x_j|z_{1:k})$, with j fixed and $k \geq j$).

As in the single-model case [2], two views can be adopted for fixed-interval smoothing. Ref. [22] consists in fusing the estimates and covariances produced by a forward conventional IMM filter and a modified backward IMM filter. Some difficulties lie in the need to set an inverse dynamics model, especially if (1) is nonlinear, and initialize the backward filter with a flat prior so as to prevent the assimilation of common data into both filters. More recently [18] proposed a second smoothing scheme based on a GPB$_2$ running M^2 forward filters whose estimates are recombined through a Rauch-Tung-Striebel (RTS) backward-time recursion [23] with M^2 smoothers. In comparison to the above two-filter strategy, this approach allows the use of non-invertible dynamics models. Moreover the backward-time pass is simply initialized with the filtered estimate at the end of the fixed interval. This paper rather follows this alternative viewpoint of IMM-based smoothing through RTS backward-time recursions.

III. Theoretical Foundations

This section thoroughly reviews the theoretical foundations of the Interacting Multiple Model filtering and of the possible fixed-interval smoothing backward-time recursions.

A. The Interacting Multiple Model filter

The recursion cycle of the celebrated IMM filter was first outlined in [3]:

1) $\forall i \in \mathscr{M}, \{P(m_k^j|z_{1:k})\}_{j \in \mathscr{M}} \xrightarrow{\text{Prediction}} P(m_{k+1}^i|z_{1:k})$
2) $\forall i \in \mathscr{M}, \{p(x_k|m_k^j, z_{1:k})\}_{j \in \mathscr{M}} \xrightarrow{\text{Interaction}} p(x_k|m_{k+1}^i, z_{1:k})$
3) $\forall i \in \mathscr{M}, p(x_k|m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Prediction}} p(x_{k+1}|m_{k+1}^i, z_{1:k})$
4) $\forall i \in \mathscr{M}, p(x_{k+1}|m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Update}} p(x_{k+1}|m_{k+1}^i, z_{1:k+1})$
5) $\forall i \in \mathscr{M}, P(m_{k+1}^i|z_{1:k}) \xrightarrow{\text{Update}} P(m_{k+1}^i|z_{1:k+1})$

The first step of the cycle should be read as "Compute the predicted mode probability $P(m_{k+1}^i|z_{1:k})$ $\forall i \in \mathscr{M}$ at time $k+1$ from the set of posterior mode probabilities $\{P(m_k^j|z_{1:k})\}_{j \in \mathscr{M}}$ at time k", and so forth. The IMM filter enables the propagation over time of approximations to the mode probabilities $\{\mu_{k|k}^j \approx P(m_k^j|z_{1:k})\}_{j \in \mathscr{M}}$ and of Gaussian approximations to the mode-conditioned filtering pdfs $\{p(x_k|m_k^j, z_{1:k}) \approx \mathscr{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)\}_{j \in \mathscr{M}}$, so that $p(x_k|z_{1:k}) \approx \sum_{j \in \mathscr{M}} \mu_{k|k}^j \mathscr{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)$. Its reasonable complexity comes from its internal computation of the mixing probabilities $\{\mu_{k|k}^{j|i} \approx P(m_k^j|m_{k+1}^i, z_{1:k})\}_{(i,j) \in \mathscr{M} \times \mathscr{M}}$, from which Gaussian approximations to the mode-conditioned prior pdfs $\{p(x_k|m_{k+1}^i, z_{1:k}) \approx \mathscr{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i)\}_{i \in \mathscr{M}}$ are deduced. Starting from these last pdfs, only M independent filters (matched to the modes $\{m_{k+1} = i\}_{i \in \mathscr{M}}$) need to be run between times k and $k+1$ in order to get $\{p(x_{k+1}|m_{k+1}^i, z_{1:k}) \approx \mathscr{N}(x_{k+1}; \hat{x}_{k+1|k}^i, P_{k+1|k}^i)\}_{i \in \mathscr{M}}$, $\{p(x_{k+1}|m_{k+1}^i, z_{1:k+1}) \approx \mathscr{N}(x_{k+1}; \hat{x}_{k+1|k+1}^i, P_{k+1|k+1}^i)\}_{i \in \mathscr{M}}$ and update $\{\mu_{k+1|k+1}^i = P(m_{k+1}^i|z_{1:k+1})\}_{i \in \mathscr{M}}$, leading to $p(x_{k+1}|z_{1:k+1})$.

B. Smoothing using backward-time recursions

The posterior state densities $\{p(x_T|m_T^i, z_{1:T})\}_{i \in \mathscr{M}}$ and mode probabilities $\{P(m_T^i|z_{1:T})\}_{i \in \mathscr{M}}$ at time T are the starting point. Given $\{p(x_{k+1}|m_{k+1}^i, z_{1:T})\}_{i \in \mathscr{M}}$ and $\{P(m_{k+1}^i|z_{1:T})\}_{i \in \mathscr{M}}$, the smoothing steps of the backward recursion can be conducted in three ways. The first two are drawn from the existing literature while the last one is the new approach investigated in this paper.

Backward smoothing recursion - SR1:

1) $\forall (i,j) \in \mathscr{M}^2$,
$p(x_{k+1}|m^i_{k+1}, z_{1:T}) \xrightarrow{\text{Smoothing}} p(x_k|m^i_{k+1}, m^j_k, z_{1:T})$
2) $\forall j \in \mathscr{M}$,
$\{p(x_k|m^i_{k+1}, m^j_k, z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Interaction}} p(x_k|m^j_k, z_{1:T})$
3) $\forall j \in \mathscr{M}, \{P(m^i_{k+1}|z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Smoothing}} P(m^j_k|z_{1:T})$

This recursion cycle was proposed in [17] and [18]. It uses a total of M^2 smoothers for M admissible modes. More specifically, step 1 writes as:

$$\forall (i,j) \in \mathscr{M}^2,\ p(x_k|m^i_{k+1}, m^j_k, z_{1:T}) =$$
$$p(x_k|m^j_k, z_{1:k}) \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m^i_{k+1}) p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:T})}{p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:k})} \mathrm{d}x_{k+1}$$

where $p(x_{k+1}|x_k, m^i_{k+1})$ is the transition density between k and $k+1$ associated to (1) and conditioned on the active mode m^i_{k+1} at time $k+1$. The densities $p(x_k|m^j_k, z_{1:k})$ and $p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:k})$ are computed by a forward GPB$_2$ filter. Finally, the smoothed density $p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:T})$ is approximated by $p(x_{k+1}|m^i_{k+1}, z_{1:T})$ so as to start the recursion cycle.

Backward smoothing recursion - SR2:

1) $\forall j \in \mathscr{M}$,
$\{p(x_{k+1}|m^i_{k+1}, z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Interaction}} p(x_{k+1}|m^j_k, z_{1:T})$
2) $\forall j \in \mathscr{M}, p(x_{k+1}|m^j_k, z_{1:T}) \xrightarrow{\text{Smoothing}} p(x_k|m^j_k, z_{1:T})$
3) $\forall j \in \mathscr{M}, \{P(m^i_{k+1}|z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Smoothing}} P(m^j_k|z_{1:T})$

The major advantage over the scheme SR1 lies in a lower computational load as M smoothers are used for M modes in the smoothing step 2. This reduced complexity is exploited in [19] but at the cost of strong approximations in the development of the algorithm. Step 1 is indeed rewritten as

$$p(x_{k+1}|m^j_k, z_{1:T}) = \sum_{i\in\mathscr{M}} p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:T}) P(m^i_{k+1}|m^j_k, z_{1:T}).$$

The authors claim that in the first term of the sum "*the condition on m^j_k [...] can be ignored due to Markov property*" so that $p(x_{k+1}|m^i_{k+1}, m^j_k, z_{1:T}) = p(x_{k+1}|m^i_{k+1}, z_{1:T})$ [19, Eq. 12]. Incidentally, this equality precludes the exponentially growing complexity of the problem. It should be rather considered as an approximation like in [17] and [18]. The development of step 2 is then conducted using the equality [19, Eq. 11]:

$$\forall j \in\mathscr{M},\ p(x_k|m^j_k, z_{1:T}) =$$
$$p(x_k|m^j_k, z_{1:k}) \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m^j_k) p(x_{k+1}|m^j_k, z_{1:T})}{p(x_{k+1}|m^j_k, z_{1:k})} \mathrm{d}x_{k+1}.$$

The authors further claim that "*the term $p(x_{k+1}|x_k, m^j_k)$ [...] corresponds to the state transition density of model m^j_k*". However, this contradicts the hypothesis of [19, Eq. 1] in that m^j_k terms the active mode that governs the state transition between $k-1$ and k. This last hypothesis is used in the present paper (see (1) and (5)) and in the cited references too.

Backward smoothing recursion - SR3: The present paper investigates an alternative method with a linear number of smoothers.

1) $\forall i \in \mathscr{M}, p(x_{k+1}|m^i_{k+1}, z_{1:T}) \xrightarrow{\text{Smoothing}} p(x_k|m^i_{k+1}, z_{1:T})$
2) $\forall j \in \mathscr{M}, \{p(x_k|m^i_{k+1}, z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Interaction}} p(x_k|m^j_k, z_{1:T})$
3) $\forall j \in \mathscr{M}, \{P(m^i_{k+1}|z_{1:T})\}_{i\in\mathscr{M}} \xrightarrow{\text{Smoothing}} P(m^j_k|z_{1:T})$

The smoothing equation of the first step is now given by

$$\forall i \in \mathscr{M}, p(x_k|m^i_{k+1}, z_{1:T}) =$$
$$p(x_k|m^i_{k+1}, z_{1:k}) \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m^i_{k+1}) p(x_{k+1}|m^i_{k+1}, z_{1:T})}{p(x_{k+1}|m^i_{k+1}, z_{1:k})} \mathrm{d}x_{k+1}$$

where $p(x_k|m^i_{k+1}, z_{1:k})$ and $p(x_{k+1}|m^i_{k+1}, z_{1:k})$ are computed by an IMM filter. The pdf $p(x_{k+1}|m^i_{k+1}, z_{1:T})$ is known from the previous recursion and $p(x_{k+1}|x_k, m^i_{k+1})$ is the genuine transition density between k and $k+1$ as m^i_{k+1} is active over the sampling interval $(t_k, t_{k+1}]$. The other equations of this algorithm are detailed in the following section.

IV. Fixed-interval smoother for jump Markov systems

As aforementioned, the aim is to approximate the smoothing pdf of the jump Markov system (1)–(2)–(3)–(4) as a M-hypotheses Gaussian mixture according to

$$\begin{aligned} p(x_k|z_{1:T}) &= \sum_{j\in\mathscr{M}} P(m^j_k|z_{1:T}) p(x_k|m^j_k, z_{1:T}) \\ &\approx \sum_{j\in\mathscr{M}} P(m^j_k|z_{1:T}) \mathscr{N}(x_k; \hat{x}^j_{k|T}, P^j_{k|T}). \end{aligned} \tag{7}$$

For jump Markov systems, the global (hybrid) state vector ξ_k is independent of $z_{k+1:T}$ when conditioned on ξ_{k+1} so that $p(\xi_k|\xi_{k+1}, z_{1:T}) = p(x_k, m_k|x_{k+1}, m_{k+1}, z_{1:T})$ is equal to $p(\xi_k|\xi_{k+1}, z_{1:k}) = p(x_k, m_k|x_{k+1}, m_{k+1}, z_{1:k})$. By marginalizing over m_k, one gets the equality

$$p(x_k|x_{k+1}, m_{k+1}, z_{1:T}) = p(x_k|x_{k+1}, m_{k+1}, z_{1:k}) \tag{8}$$

which is conditioned only on the active mode over the sampling period ending at t_{k+1}.

All distributions are henceforth approximated by Gaussians. From the statistics $\{\hat{x}^j_{k|k}, P^j_{k|k}, \mu^j_{k|k}\}_{j\in\mathscr{M}}$ and $\{\bar{x}^i_{k|k}, \bar{P}^i_{k|k}\}_{i\in\mathscr{M}}$ produced by an IMM filter at times $k = 0, \ldots, T$, together with $\{\hat{x}^i_{k+1|k}, P^i_{k+1|k}\}_{i\in\mathscr{M}}$ produced at times $k+1 = 1, \ldots, T$, the proposed algorithm recursively determines the smoothing mode-conditioned densities $\{p(x_k|m^j_k, z_{1:T}) \approx \mathscr{N}(x_k; \hat{x}^j_{k|T}, P^j_{k|T})\}_{j\in\mathscr{M}}$ and the smoothed mode probabilities $\{\mu^j_{k|T} = P(m^j_k|z_{1:T})\}_{j\in\mathscr{M}}$ for $k = T-1, \ldots, 0$.

A. Step 1 of SR3: mode-matched smoothing

Theorem 1. *From the knowledge of $p(x_{k+1}|m^i_{k+1}, z_{1:T}) \approx \mathscr{N}(x_{k+1}; \hat{x}^i_{k+1|T}, P^i_{k+1|T})$ and $\mu^i_{k+1|T} = P(m^i_{k+1}|z_{1:T})$ at time $k+1$, the mean and covariance of the smoothed mixing density $p(x_k|m^i_{k+1}, z_{1:T}) \approx \mathscr{N}(x_k; \bar{x}^i_{k|T}, \bar{P}^i_{k|T})$ are first determined with the Rauch-Tung-Striebel formulae*

$$G^i_k = C^i_{k,k+1} \left(P^i_{k+1|k}\right)^{-1} \tag{9}$$
$$\bar{x}^i_{k|T} = \bar{x}^i_{k|k} + G^i_k \left(\hat{x}^i_{k+1|T} - \hat{x}^i_{k+1|k}\right) \tag{10}$$
$$\bar{P}^i_{k|T} = \bar{P}^i_{k|k} + G^i_k \left(P^i_{k+1|T} - P^i_{k+1|k}\right) \left(G^i_k\right)^T \tag{11}$$

where

$$C^i_{k,k+1} = \int \left(x_k - \bar{x}^i_{k|k}\right)\left(f^i_k(x_k) - \hat{x}^i_{k+1|k}\right)^T \mathcal{N}(x_k; \bar{x}^i_{k|k}, \bar{P}^i_{k|k}) dx_k. \quad (12)$$

Proof. The equations (9)-(12) can be demonstrated by following exactly the proof of [24, Sec. II.A] with all densities conditioned on m^i_{k+1}, and by using the property (8). An approximation of the integral (12) can be easily evaluated by means of the unscented transform, as suggested in [24]. □

In contrast to the single-model smoother, equations (9), (10), (11) do not end the recursion cycle because the smoothing density of x_k is conditioned on m^i_{k+1} instead of m^j_k. The following interaction stage bridges the gap between the Gaussian approximations to the mode-conditioned smoothing densities $\mathcal{N}(x_k; \bar{x}^i_{k|T}, \bar{P}^i_{k|T}) \approx p(x_k|m^i_{k+1}, z_{1:T})$ and $\mathcal{N}(x_k; \hat{x}^j_{k|T}, P^j_{k|T}) \approx p(x_k|m^j_k, z_{1:T})$. Two options are hereafter investigated.

B. Step 2 of SR3: a mode interaction with M^2 combinations

Using the total probability theorem, the targeted mode-conditioned smoothing density $p(x_k|m^j_k, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}^j_{k|T}, P^j_{k|T})$ can be expressed as a mixture of densities conditioned on the sequence of modes over two consecutive sampling periods, namely

$$p(x_k|m^j_k, z_{1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m^j_k, m^i_{k+1}, z_{1:T}) P(m^i_{k+1}|m^j_k, z_{1:T}). \quad (13)$$

The two forthcoming theorems enable its computation.

Theorem 2. *The first two moments of $p(x_k|m^j_k, m^i_{k+1}, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}^{ji}_{k|T}, P^{ji}_{k|T})$ are obtained by forward-time IMM filtering and backward-time Rauch-Tung-Striebel recursions as follows:*

$$\hat{x}^{ji}_{k|T} = P^{ji}_{k|T}\left[\left(\bar{P}^i_{k|T}\right)^{-1}\bar{x}^i_{k|T} - \left(\bar{P}^i_{k|k}\right)^{-1}\bar{x}^i_{k|k} + \left(P^j_{k|k}\right)^{-1}\hat{x}^j_{k|k}\right], \quad (14)$$

with

$$P^{ji}_{k|T} = \left[\left(\bar{P}^i_{k|T}\right)^{-1} - \left(\bar{P}^i_{k|k}\right)^{-1} + \left(P^j_{k|k}\right)^{-1}\right]^{-1}. \quad (15)$$

Proof. Following [22], the Bayes formula and the Markov property of the mode sequence lead to

$$p(x_k|m^j_k, m^i_{k+1}, z_{1:T}) \propto p(z_{k+1:T}|x_k, m^i_{k+1}) p(x_k|m^j_k, z_{1:k}). \quad (16)$$

Similarly, the following holds

$$\begin{aligned} & p(x_k|m^i_{k+1}, z_{1:T}) \propto p(z_{k+1:T}|x_k, m^i_{k+1}) p(x_k|m^i_{k+1}, z_{1:k}) \\ & \Leftrightarrow p(z_{k+1:T}|x_k, m^i_{k+1}) \propto \frac{p(x_k|m^i_{k+1}, z_{1:T})}{p(x_k|m^i_{k+1}, z_{1:k})}, \end{aligned} \quad (17)$$

which yields the final equality

$$p(x_k|m^j_k, m^i_{k+1}, z_{1:T}) \propto \frac{p(x_k|m^i_{k+1}, z_{1:T})}{p(x_k|m^i_{k+1}, z_{1:k})} p(x_k|m^j_k, z_{1:k}). \quad (18)$$

All the involved densities are approximated by Gaussians. So, the logarithm of $p(x_k|m^j_k, m^i_{k+1}, z_{1:T})$ writes as $C - \frac{1}{2}J(x_k)$, with C a constant and

$$J(x_k) = \|x_k - \bar{x}^i_{k|T}\|^2_{(\bar{P}^i_{k|T})^{-1}} - \|x_k - \bar{x}^i_{k|k}\|^2_{(\bar{P}^i_{k|k})^{-1}} + \|x_k - \hat{x}^j_{k|k}\|^2_{(P^j_{k|k})^{-1}}. \quad (19)$$

The mean $\hat{x}^{ji}_{k|T}$ of $p(x_k|m^j_k, m^i_{k+1}, z_{1:T})$ is also its mode and comes as the minimum of $J(x_k)$, which leads to (14)–(15). □

Interestingly, Eq. (13) of the interaction step is common with [22, Eq. 73] albeit [22] evaluates the smoothed estimate $\hat{x}^{ji}_{k|T}$ by combining the estimates produced by a conventional IMM filter and a backward-time IMM filter restricted to linear systems with invertible state transition matrix. Moreover, this backward-time IMM filter requires to be initialized at final time T with no prior information. Note that the maximum likelihood estimate $\arg\max_{x_k} p(z_{k+1:T}|x_k, m^i_{k+1})$ writes as

$$\hat{x}^{b,i}_{k|k+1} = P^{b,i}_{k|k+1}\left[\left(\bar{P}^i_{k|T}\right)^{-1}\bar{x}^i_{k|T} - \left(\bar{P}^i_{k|k}\right)^{-1}\bar{x}^i_{k|k}\right] \quad (20)$$

with

$$P^{b,i}_{k|k+1} = \left[\left(\bar{P}^i_{k|T}\right)^{-1} - \left(\bar{P}^i_{k|k}\right)^{-1}\right]^{-1}, \quad (21)$$

and is nothing else but the "one-step backward-time predicted estimate and error covariance" computed by the backward-time IMM filter of [22].

Theorem 3. *The smoothed mixing probabilities $\{\bar{\mu}^{i|j}_{k+1|T} = P(m^i_{k+1}|m^j_k, z_{1:T})\}_{(i,j)\in\mathcal{M}^2}$ involved in (13) are expressed as*

$$\bar{\mu}^{i|j}_{k+1|T} = \frac{P(m^i_{k+1}|m^j_k) p(z_{k+1:T}|m^j_k, m^i_{k+1}, z_{1:k})}{p(z_{k+1:T}|m^j_k, z_{1:k})} = \frac{\pi_{ji}\Lambda_{ji}}{d_j}, \quad (22)$$

where the likelihood

$$\Lambda_{ji} = p(z_{k+1:T}|m^j_k, m^i_{k+1}, z_{1:k}) \quad (23)$$

can be approximated by

$$\begin{gathered} \Lambda_{ji} \approx \mathcal{N}(\Delta^{ji}_k; 0, D^{ji}_k), \\ \Delta^{ji}_k = \hat{x}^{b,i}_{k|k+1} - \hat{x}^j_{k|k},\ D^{ji}_k = P^{b,i}_{k|k+1} + P^j_{k|k}, \end{gathered} \quad (24)$$

and

$$d_j = p(z_{k+1:T}|m^j_k, z_{1:k}) = \sum_{i\in\mathcal{M}} \pi_{ji}\Lambda_{ji} \quad (25)$$

stands for the normalizing constant.

Proof. Eq. (22) is straightforward. The approximation (24) has been proposed in [22]. □

The posterior mean $\hat{x}^j_{k|T}$ and covariance $P^j_{k|T}$ of the mode-conditioned smoothing density (13) are eventually computed via their moment-matched approximations

$$\hat{x}^j_{k|T} = \sum_{i\in\mathcal{M}} \bar{\mu}^{i|j}_{k+1|T} \hat{x}^{ji}_{k|T} \quad (26)$$

$$P^j_{k|T} = \sum_{i\in\mathcal{M}} \bar{\mu}^{i|j}_{k+1|T}\left[P^{ji}_{k|T} + (\hat{x}^{ji}_{k|T} - \hat{x}^j_{k|T})(\hat{x}^{ji}_{k|T} - \hat{x}^j_{k|T})^T\right]. \quad (27)$$

These last equations end the smoothing recursion.

C. Step 2 of SR3: an alternative mode interaction with M combinations

Instead of combining the M^2 filtering densities $p(x_k|m^j_k, z_{1:k})$ and $p(x_k|m^i_{k+1}, z_{k+1:T})$, another option to build $p(x_k|m^j_k, z_{1:T})$ is to directly fuse the M filtering densities $p(x_k|m^j_k, z_{1:k})$ and $p(x_k|m^j_k, z_{k+1:T})$. This alternative option was pointed out in the conclusion of [22], and can be solved by the following theorem.

Theorem 4. *The backward filtering density* $p(x_k|m_k^j, z_{k+1:T}) \approx \mathcal{N}(x_k; \bar{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$ *is computed by the backward-time mixing step*

$$p(x_k|m_k^j, z_{k+1:T}) = \sum_{i\in\mathcal{M}} p(x_k|m_{k+1}^i, z_{k+1:T}) P(m_{k+1}^i|m_k^j, z_{k+1:T}) \quad (28)$$

where $p(x_k|m_{k+1}^i, z_{k+1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i})$ *and the backward mixing probabilities* $P(m_{k+1}^i|m_k^j, z_{k+1:T}) = \mu_{k+1|T}^{b,i|j}$ *are equal to* $P(m_{k+1}^i|m_k^j, z_{1:T}) = \bar{\mu}_{k+1|T}^{i|j}$*. The corresponding first two moments of* $p(x_k|m_k^j, z_{k+1:T})$ *come as*

$$\bar{x}_{k|k+1}^{b,j} = \sum_{i\in\mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} \hat{x}_{k|k+1}^{b,i} \quad (29)$$

$$\bar{P}_{k|k+1}^{b,j} = \sum_{i\in\mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} \left[P_{k|k+1}^{b,i} + (\hat{x}_{k|k+1}^{b,i} - \bar{x}_{k|k+1}^{b,j})(\hat{x}_{k|k+1}^{b,i} - \bar{x}_{k|k+1}^{b,j})^T\right]. \quad (30)$$

Proof. Eq. (28) and its moment matched approximations (29)-(30) were introduced in [22] as the mixing step of a backward IMM filter. The equality $P(m_{k+1}^i|m_k^j, z_{1:T}) = P(m_{k+1}^i|m_k^j, z_{k+1:T})$ comes from the Markov properties of the mode sequence. □

Note that the mean $\bar{x}_{k|k+1}^{b,j}$ and the covariance $\bar{P}_{k|k+1}^{b,j}$ can be obtained using (20)-(21). Eventually, the recursion is closed with the combination of $p(x_k|m_k^j, z_{1:k})$ and $p(x_k|m_k^j, z_{k+1:T})$, as follows.

Theorem 5. *The mean and the covariance of the smoothed density* $p(x_k|m_k^j, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j)$ *are*

$$\hat{x}_{k|T}^j = P_{k|T}^j \left[\left(P_{k|k}^j\right)^{-1} x_{k|k}^j + \left(\bar{P}_{k|k+1}^{b,j}\right)^{-1} \bar{x}_{k|k+1}^{b,j}\right] \quad (31)$$

$$P_{k|T}^j = \left[\left(P_{k|k}^j\right)^{-1} + \left(\bar{P}_{k|k+1}^{b,j}\right)^{-1}\right]^{-1}. \quad (32)$$

Proof. By Bayes formula, the density $p(x_k|m_k^j, z_{1:T})$ can be rewritten as

$$p(x_k|m_k^j, z_{1:T}) \propto p(x_k|m_k^j, z_{1:k}) p(z_{k+1:T}|x_k, m_k^j).$$

Following a reasoning similar to proof of Theorem 2 and using the fact that the model-conditioned likelihood $p(z_{k+1:T}|x_k, m_k^j)$ is equal to $\mathcal{N}(x_k; \bar{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$, (31) and (32) hold. □

D. Step 3 of SR3, smoother output and algorithm implementation

The posterior smoothed mode probability $\mu_{k|T}^j = P(m_k^j|z_{1:T})$ of mode j at time k is given by

$$\mu_{k|T}^j = \frac{p(z_{k+1:T}|m_k^j, z_{1:k}) P(m_k^j|z_{1:k})}{\sum_{j\in\mathcal{M}} p(z_{k+1:T}|m_k^j, z_{1:k}) P(m_k^j|z_{1:k})} = \frac{d_j \mu_{k|k}^j}{\sum_{j\in\mathcal{M}} d_j \mu_{k|k}^j}. \quad (33)$$

For output purposes, the overall smoothing density $p(x_k|z_{1:T})$ in (7) can then be approximated by its moment-matched Gaussian pdf $\mathcal{N}(x_k; \hat{x}_{k|T}, P_{k|T})$, where

$$\hat{x}_{k|T} = \sum_{j\in\mathcal{M}} \mu_{k|T}^j \hat{x}_{k|T}^j \quad (34)$$

$$P_{k|T} = \sum_{j\in\mathcal{M}} \mu_{k|T}^j \left[P_{k|T}^j + (\hat{x}_{k|T}^j - \hat{x}_{k|T})(\hat{x}_{k|T}^j - \hat{x}_{k|T})^T\right]. \quad (35)$$

For detection issues, the MAP mode estimate $\hat{j}_k$ at time t_k writes as

$$\hat{j}_k = \underset{j=1,\dots,M}{\arg\max}\ \mu_{k|T}^j. \quad (36)$$

Importantly, the covariance $P_{k|k+1}^{b,i}$ may not be defined during the first steps of the backward recursion because $\left(P_{k|k+1}^{b,i}\right)^{-1} = \left(\bar{P}_{k|T}^i\right)^{-1} - \left(\bar{P}_{k|k}^i\right)^{-1}$ may not be invertible. This situation occurs when the size of the measurement vector z_k is smaller than that of the state vector x_k and/or when the assimilated measurements do not carry enough information to provide an estimate $\hat{x}_{k|k+1}^{b,i}$ and to endow $P_{k|k+1}^{b,i}$ with finite eigenvalues. In this case, the density $\mathcal{N}(.; 0, D_k^{ji})$ can be viewed as a flat prior. Therefore, the smoothed mixing probabilities and the posterior mode probabilities are approximated by $\bar{\mu}_{k+1|T}^{i|j} \approx P(m_{k+1}^i|m_k^j, z_{1:k}) = \pi_{ji}$ for all indexes i, j, and to $\mu_{k|T}^j \approx \mu_{k|k}^j$ for all j until $\left(P_{k|k+1}^{b,i}\right)^{-1}$ becomes nonsingular.

The complete mode-conditioned smoothing and interaction steps expressed in Theorems 1 to 5 are summarized in Algorithm 1. For clarity, the details of the interaction step are presented separately in Algorithm 2 for M^2 combinations (Interaction 1) and in Algorithm 3 for M combinations (Interaction 2). Interestingly, Interaction 2 is computationally cheaper. While the first option uses $O(M^2)$ matrix inversions, Interaction 2 requires only $O(M)$ of them to build $p(x_k|m_k^j, z_{1:T})$. Nevertheless, Interaction 2 absolutely requires that $\hat{x}_{k|k+1}^{b,i}$ and $P_{k|k+1}^{b,i}$ of (20)-(21) are explicitly defined. Thus, whatever the selected interaction type, the first recursion steps of the complete algorithm have always to be performed with Interaction 1 until $\left(P_{k|k+1}^{b,i}\right)^{-1}$ becomes invertible.

V. Extension to fixed-lag smoothing

Even when real-time processing is required, a smoothed estimation of the system state vector can be provided with a small delay by means of a fixed-lag smoother. For the single model case, the fixed-lag smoothing pdf $p(x_k|z_{1:k+n})$ at time k, with $n \geq 1$ the lag length, can be obtained via two approaches. The first method boils down to building an augmented state vector X_{k+n} as the concatenation of the state vectors between k and $k+n$, *i.e.* $X_{k+n} = (x_{k+n}^T, \dots, x_k^T)^T$ [2]. The smoothed state estimate is then computed by running a filter on the augmented state, prior to selecting the part which corresponds to x_k. The dynamics and observation models need to be extended so as to comply with the augmented state vector, and the Kalman filter needs to be modified as proposed in [2, Eq. 8.66 to 8.70] to keep a computation load linearly growing with n. The second solution [24] is to use directly the Rauch-Tung-Striebel smoother equations over the interval beginning at k and ending at $k+n$. Once the filtered state estimate at time $k+n$ is known, a Rauch-Tung-Striebel recursion is launched for $k+n-1, \dots, k$ as a simple post-processing. For this solution too, the computation load grows linearly with the lag length.

In the framework of jump Markov systems, the most investigated solution is state augmentation. This approach is used in [25] with a basic linear IMM filter applied to a state augmented jump Markov system. Interestingly, the density of

Algorithm 1 ONE STEP OF THE FIXED-INTERVAL SMOOTHER FOR JUMP MARKOV SYSTEMS

$\left[\{\hat{x}^j_{k|T}, P^j_{k|T}, \mu^j_{k|T}\}_{j\in\mathscr{M}}\right] =$ **FIXED-INTERVAL-IMM-SMOOTHER**$\Big(\{\hat{x}^i_{k+1|T}, P^i_{k+1|T}, \mu^i_{k+1|T}\}_{i\in\mathscr{M}}, \{\hat{x}^j_{k|k}, P^j_{k|k}, \mu^j_{k|k}\}_{j\in\mathscr{M}}, \{\bar{x}^i_{k|k}, \bar{P}^i_{k|k}\}_{i\in\mathscr{M}}, \{\hat{x}^i_{k+1|k}, P^i_{k+1|k}\}_{i\in\mathscr{M}}\Big)$

FOR $i \in \mathscr{M}$ **DO** {1. Mode-matched smoothing}
  Smoother gain: $G^i_k = C^i_{k,k+1}\left(P^i_{k+1|k}\right)^{-1}$ where $C^i_{k,k+1} = \int \left(x_k - \bar{x}^i_{k|k}\right)\left(f^i_k(x_k) - \hat{x}^i_{k+1|k}\right)^T \mathscr{N}(x_k; \bar{x}^i_{k|k}, \bar{P}^i_{k|k}) dx_k$.
  Smoothed mixing mean: $\bar{x}^i_{k|T} = \bar{x}^i_{k|k} + G^i_k\left(\hat{x}^i_{k+1|T} - \hat{x}^i_{k+1|k}\right)$.
  Smoothed mixing covariance: $\bar{P}^i_{k|T} = \bar{P}^i_{k|k} + G^i_k\left(P^i_{k+1|T} - P^i_{k+1|k}\right)\left(G^i_k\right)^T$.
END FOR
FOR $i \in \mathscr{M}$ **DO** {2. Mode interaction}
  One-step backward predicted information matrix: $\left(P^{b,i}_{k|k+1}\right)^{-1} = \left(\bar{P}^i_{k|T}\right)^{-1} - \left(\bar{P}^i_{k|k}\right)^{-1}$.
  One-step backward predicted information vector: $\left(P^{b,i}_{k|k+1}\right)^{-1}\hat{x}^{b,i}_{k|k+1} = \left(\bar{P}^i_{k|T}\right)^{-1}\bar{x}^i_{k|T} - \left(\bar{P}^i_{k|k}\right)^{-1}\bar{x}^i_{k|k}$.
  FOR $j \in \mathscr{M}$ **DO**
    Two-mode conditioned likelihood: $\Lambda_{ji} \approx \mathscr{N}(\Delta^{ji}_k; 0, D^{ji}_k)$ with $\Delta^{ji}_k = \hat{x}^{b,i}_{k|k+1} - \hat{x}^j_{k|k}$ and $D^{ji}_k = P^{b,i}_{k|k+1} + P^j_{k|k}$.
    Smoothed mixing probability: $\bar{\mu}^{i|j}_{k+1|T} = \frac{\pi_{ji}\Lambda_{ji}}{\sum_i \pi_{ji}\Lambda_{ji}}$ $\Big(\approx \pi_{ji}$ if $\left(P^{b,i}_{k|k+1}\right)^{-1}$ is not invertible$\Big)$.
  END FOR
END FOR
IF Interaction 1 is selected OR $\left(P^{b,i}_{k|k+1}\right)^{-1}$ is not invertible THEN
  $\left[\{\hat{x}^j_{k|T}, P^j_{k|T}\}_{j\in\mathscr{M}}\right] =$ **MODE-INTERACTION-1**$\Big(\{\left(P^j_{k|k}\right)^{-1}\hat{x}^j_{k|k}, \left(P^j_{k|k}\right)^{-1}\}_{j\in\mathscr{M}}, \{\left(P^{b,i}_{k|k+1}\right)^{-1}\hat{x}^{b,i}_{k|k+1}, \left(P^{b,i}_{k|k+1}\right)^{-1}\}_{i\in\mathscr{M}}, \{\bar{\mu}^{i|j}_{k+1|T}\}_{(i,j)\in\mathscr{M}^2}\Big)$
ELSE
  $\left[\{\hat{x}^j_{k|T}, P^j_{k|T}\}_{j\in\mathscr{M}}\right] =$ **MODE-INTERACTION-2**$\Big(\{\hat{x}^j_{k|k}, P^j_{k|k}\}_{j\in\mathscr{M}}, \{\hat{x}^{b,i}_{k|k+1}, P^{b,i}_{k|k+1}\}_{i\in\mathscr{M}}, \{\bar{\mu}^{i|j}_{k+1|T}\}_{(i,j)\in\mathscr{M}^2}\Big)$
FOR $j \in \mathscr{M}$ **DO** {3. Smoother output}
  Smoothed mode probability: $\mu^j_{k|T} = \frac{d_j \mu^j_{k|k}}{\sum\limits_{j\in\mathscr{M}} d_j \mu^j_{k|k}}$, with $d_j = \sum_i \pi_{ji}\Lambda_{ji}$ $\Big(\mu^j_{k|T} \approx \mu^j_{k|k}$ if $\left(P^{b,i}_{k|k+1}\right)^{-1}$ is not invertible$\Big)$
END FOR
Overall smoothed mean: $\hat{x}_{k|T} = \sum_{j\in\mathscr{M}} \mu^j_{k|T}\hat{x}^j_{k|T}$.
Overall smoothed covariance: $P_{k|T} = \sum_{j\in\mathscr{M}} \mu^j_{k|T}\left[P^j_{k|T} + (\hat{x}^j_{k|T} - \hat{x}_{k|T})(\hat{x}^j_{k|T} - \hat{x}_{k|T})^T\right]$.

Algorithm 2 MODE INTERACTION WITH M^2 COMBINATIONS

$\left[\{\hat{x}^j_{k|T}, P^j_{k|T}\}_{j\in\mathscr{M}}\right] =$ **MODE-INTERACTION-1**$\Big(\{\hat{y}^j_{k|k}, Y^j_{k|k}\}_{j\in\mathscr{M}}, \{\hat{y}^{b,i}_{k|k+1}, Y^{b,i}_{k|k+1}\}_{i\in\mathscr{M}}, \{\bar{\mu}^{i|j}_{k+1|T}\}_{(i,j)\in\mathscr{M}^2}\Big)$

FOR $i \in \mathscr{M}$ **DO** {1. Combination}
  FOR $j \in \mathscr{M}$ **DO**
    Two-mode conditioned smoothed covariance: $P^{ji}_{k|T} = \left[Y^{b,i}_{k|k+1} + Y^j_{k|k}\right]^{-1}$.
    Two-mode conditioned smoothed mean: $\hat{x}^{ji}_{k|T} = P^{ji}_{k|T}\left[\hat{y}^{b,i}_{k|k+1} + \hat{y}^j_{k|k}\right]$.
  END FOR
END FOR
FOR $j \in \mathscr{M}$ **DO** {2. Mixing}
  Mode-conditioned smoothed mean: $\hat{x}^j_{k|T} = \sum_{i\in\mathscr{M}} \bar{\mu}^{i|j}_{k+1|T}\hat{x}^{ji}_{k|T}$.
  Mode-conditioned smoothed covariance: $P^j_{k|T} = \sum_{i\in\mathscr{M}} \bar{\mu}^{i|j}_{k+1|T}\left[P^{ji}_{k|T} + (\hat{x}^{ji}_{k|T} - \hat{x}^j_{k|T})(\hat{x}^{ji}_{k|T} - \hat{x}^j_{k|T})^T\right]$.
END FOR

the augmented state $p(X_{k+n}|m^r_{k+n}, z_{1:k+n})$ is conditioned on the model m^r_{k+n} running at $k+n$ so that the smoothed estimate of x_k is also conditioned on m^r_{k+n}. A nonlinear solution is proposed in [11], where an IMM filter is based on a bank of Unscented Kalman filters (UKFs). The authors show how, thanks to the fact that the corresponding augmented state models are partially linear, the unscented transform underlying the UKF can be performed with the same reduced number of sigma points whatever the lag length n. A fixed-lag smoother for linear jump Markov systems using an IMM filter without state augmentation was proposed by [26] for $n = 1$ only, and by [27] for arbitrary lag lengths but under the assumption that there are no model jumps over the interval $(k, k+n]$.

In light of Section IV, the fixed-lag smoothed estimate at time k can be obtained by simply applying the two-step recursion—Rauch-Tung-Striebel smoothing and mode interaction SR3—for $k+n-1, \ldots, k$ after the IMM filtering has been performed until time $k+n$. As already stated, this approach can be considered as an independent additional processing on the real time filtered estimates and does not require any change to the existing models. The smoothed estimate at k is conditioned on m^j_k instead of m^r_{k+n} as for the augmented state approach, and model jumps are taken into account over the interval $(k, k+n]$ through the mode interaction stage. Notice that in Algorithm 1, the smoother gain G^i_k, the inverses $\left(\bar{P}^i_{k|k}\right)^{-1}$ and $\left(P^j_{k|k}\right)^{-1}$, the products $\left(\bar{P}^i_{k|k}\right)^{-1}\bar{x}^i_{k|k}$ and $\left(P^j_{k|k}\right)^{-1}\hat{x}^j_{k|k}$ may be precomputed and stored while performing the IMM filtering in order to get a faster implementation of the multiple-model fixed-lag smoothing recursion.

VI. SIMULATION EXAMPLES

A. Fixed-interval smoothing

A simulated 2D target tracking example is presented to examine the estimation errors and the posterior mode probabilities produced by the proposed Interactive Multiple Model

Algorithm 3 MODE INTERACTION WITH M COMBINATIONS

$\left[\{\hat{x}^j_{k|T}, P^j_{k|T}\}_{j\in\mathscr{M}}\right]$ = **MODE-INTERACTION-2** $\left(\{\hat{x}^j_{k|k}, P^j_{k|k}\}_{j\in\mathscr{M}}, \{\hat{x}^{b,i}_{k|k+1}, P^{b,i}_{k|k+1}\}_{i\in\mathscr{M}}, \{\bar{\mu}^{i|j}_{k+1|T}\}_{(i,j)\in\mathscr{M}^2}\right)$

FOR $j \in \mathscr{M}$ **DO** {1. Backward IMM mixing}
    Mixing backward predicted mean: $\bar{x}^{b,j}_{k|k+1} = \sum_{i\in\mathscr{M}} \bar{\mu}^{i|j}_{k+1|T}\hat{x}^{b,i}_{k|k+1}$.
    Mixing backward predicted covariance: $\bar{P}^{b,j}_{k|k+1} = \sum_{i\in\mathscr{M}} \bar{\mu}^{i|j}_{k+1|T}\left[P^{b,i}_{k|k+1} + (\hat{x}^{b,i}_{k|k+1} - \bar{x}^{b,j}_{k|k+1})(\hat{x}^{b,i}_{k|k+1} - \bar{x}^{b,j}_{k|k+1})^T\right]$.
END FOR
FOR $j \in \mathscr{M}$ **DO** {2. Combination}
    Mode-conditioned smoothed mean: $\hat{x}^j_{k|T} = \left(\bar{P}^{b,j}_{k|k+1}\right)^{-1}\bar{x}^{b,j}_{k|k+1} + \left(P^j_{k|k}\right)^{-1}\hat{x}^j_{k|k}$.
    Mode-conditioned smoothed covariance: $P^j_{k|T} = \left[\left(\bar{P}^{b,j}_{k|k+1}\right)^{-1} + \left(P^j_{k|k}\right)^{-1}\right]^{-1}$.
END FOR

Rauch-Tung-Striebel (IMM-RTS) smoother. In order to compare it with [22], an invertible state dynamics is considered.

The system state is defined as $x = \left[\mathsf{x},\ \mathsf{y},\ \dot{\mathsf{x}},\ \dot{\mathsf{y}}\right]^T$ where (x, y) term the Cartesian coordinates of the target and $(\dot{\mathsf{x}} = \frac{d\mathsf{x}}{dt}, \dot{\mathsf{y}} = \frac{d\mathsf{y}}{dt})$ stand for its velocities. The mode set contains two discrete-time correlated random walks: a first one with a high diffusion parameter $D_1 = 5^2\,\mathrm{m}^2\mathrm{s}^{-3}$ (maneuvering mode 1) and a second one with a lower diffusion parameter $D_2 = 0.5^2\,\mathrm{m}^2\mathrm{s}^{-3}$ (nearly Constant Velocity or CV mode 2). The state space equations write as

$$x_k = \begin{pmatrix} 1 & 0 & \Delta t_k & 0 \\ 0 & 1 & 0 & \Delta t_k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q^j_{k-1} \tag{37}$$

$$\text{and } Q^j_{k-1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2D_j\Delta t_k & 0 \\ 0 & 0 & 0 & 2D_j\Delta t_k \end{pmatrix} \text{ for } j = 1, 2 \tag{38}$$

with $\Delta t_k = t_k - t_{k-1}$. The vector z_k gathers the noisy measured range and bearing of the target at time t_k and is sampled for $k = 1, \ldots, T$ with period $\Delta t_k = 5$s. Thus, the output equation common to all modes is

$$z_k = \begin{bmatrix} \sqrt{\mathsf{x}_k^2 + \mathsf{y}_k^2} \\ \arctan(\frac{\mathsf{y}_k}{\mathsf{x}_k}) \end{bmatrix} + r_k, \text{ with } R_k = \mathrm{diag}([250^2\,\mathrm{m}^2, 0.02\,\mathrm{rad}^2]). \tag{39}$$

The probability transition matrix is set to

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \begin{pmatrix} 0.97 & 0.03 \\ 0.03 & 0.97 \end{pmatrix}. \tag{40}$$

The target is tracked for 90 steps (or 450 s) on a randomly generated trajectory. It evolves first according to the maneuvering mode 1, then the nearly CV mode 2 and finally the maneuvering mode 1 again. The switching times between modes occur at the deterministic values of $k = 30$ and $k = 60$. At initial time $k = 0$, the prior mode probabilities are assumed equal to each other and the initial position and velocity estimates of the base state x_0 are arbitrarily set to $[2000, 2000, 0, 0]^T$ with covariance $P_{0|0} = \mathrm{diag}([1, 1, 100, 100])$ for all modes. The algorithm was evaluated over 50 Monte Carlo runs. An example of trajectory is displayed in Fig. 1. Our IMM smoother is compared to the IMM/GPB$_2$ filtering solutions [3], [28], the GPB$_2$-RTS smoothing solution [29], the IMM-RTS smoothing solution [19] and the IMM two-filter smoothing solution [22]. The latter requires a backward-time IMM filter initialized with no prior information. As proposed by [22], the backward initialization at final time T is performed by setting for all modes the position estimate $\left[\hat{\mathsf{x}}^b_{T|T}, \hat{\mathsf{y}}^b_{T|T}\right]^T$ and associated

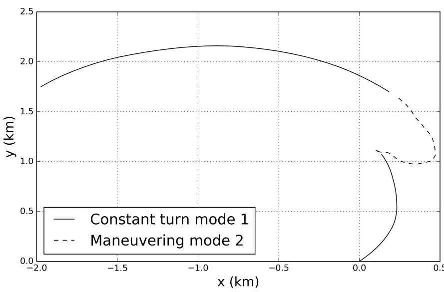

Fig. 1. Example of randomly generated trajectory

covariance to the final measurement z_T and its covariance; the final velocity estimate $\left[\hat{\dot{\mathsf{x}}}^b_{T|T}, \hat{\dot{\mathsf{y}}}^b_{T|T}\right]^T$ is set to 0 with the arbitrary large associated covariance matrix $10^6\mathbb{I}_2\,\mathrm{m}^2\mathrm{s}^{-2}$; the modes are assumed equiprobable at the terminal time.

In Table I, the time-averaged empirical root-mean-square errors (RMSE) for the position and the velocity are shown, as well as the observed time-averaged wrong detection probability (*i.e.* the average probability of selecting the wrong mode with the MAP of (36)). These quantities are also

TABLE I
TIME-AVERAGED VALUES OF: RMSE FOR POSITION (M) AND VELOCITY (M.S^{-1}); WRONG DETECTION PROBABILITY.

Method	Pos.	Vel.	Wrong detect.
IMM filter	221.1	26.2	0.22
GPB$_2$ filter	221.1	26.2	0.22
IMM Two-filter smoother [22]	136.5	12.8	0.11
GPB$_2$-RTS smoother [17], [18]	145.5	13.3	0.18
IMM-RTS smoother [19]	145.7	14.8	0.40
Our IMM-RTS smoother (Interact. 1)	135.3	12.8	0.12
Our IMM-RTS smoother (Interact. 2)	136.1	12.8	0.12

displayed for each time step of the simulation in Fig. 2(a), Fig. 2(b) and Fig. 2(c) (only for the IMM-RTS smoother with Interaction 1, the GPB$_2$-RTS smoother and the IMM two-filter smoother). The results show that our IMM-RTS smoother is well-behaved, independently of the interaction type, with a significant reduction of the RMSE errors in comparison to the filtering solutions. The detection of the active model is also more efficient. The accuracy of the smoother is similar to the two-filter smoothing solution. The GPB$_2$-RTS smoother displays a lower performance both in terms of RMSE and wrong detection probabilities. The first reason is that the algorithm makes the crude approximation $p(x_{k+1}|m^i_{k+1}, z_{1:T}) \approx p(x_{k+1}|m^j_k, m^i_{k+1}, z_{1:T})$

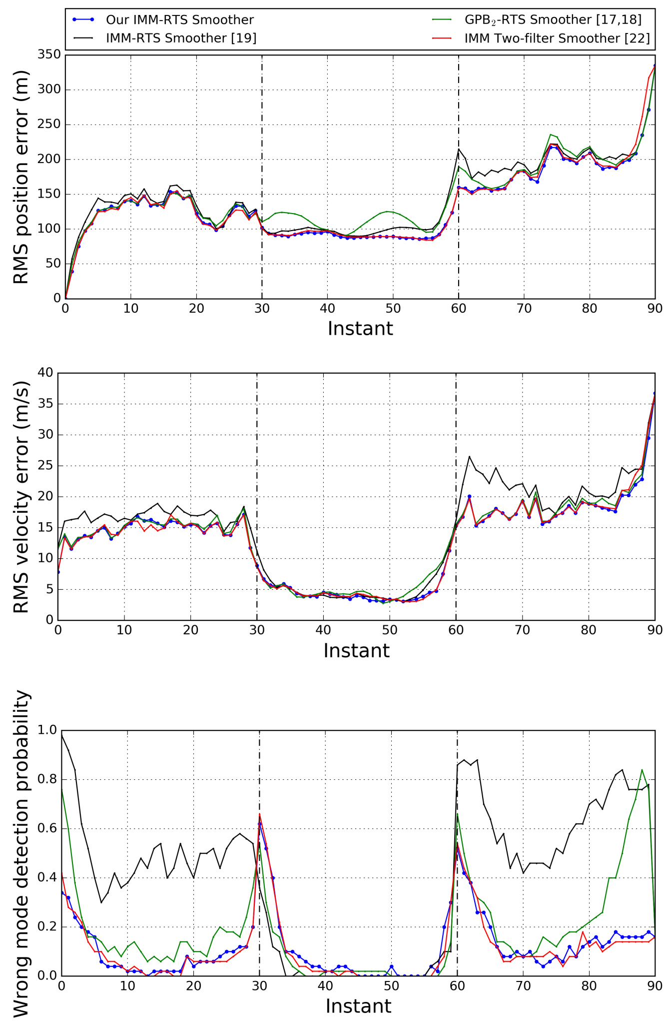

Fig. 2. Comparison of the IMM filter and the smoothers: (from top to bottom) (a) RMSE in Position, (b) RMSE in Velocity and (c) Mode error probability.

to start each step of the backward-time recursion. In light of section IV-B, it might be more efficient to write $p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:T}) \propto \frac{p(x_{k+1}|m_{k+1}^i, z_{1:T})}{p(x_{k+1}|m_{k+1}^i, z_{1:k})} p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:k})$ and build a Gaussian approximation similarly to Theorem 2. Secondly, the smoothed mode probabilities are computed by using $p(m_k^j|z_{1:T}) = \sum_{i\in\mathscr{M}} p(m_k^j|m_{k+1}^i, z_{1:T}) p(m_{k+1}^i|z_{1:T})$ with $p(m_k^j|m_{k+1}^i, z_{1:T}) = \int p(m_k^j|x_{k+1}, m_{k+1}^i, z_{1:k}) p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1}$. The last integral is proportional to $\int p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:k}) p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1}$ and is approximated by evaluating $p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:k})$—whose Gaussian estimate is provided by the GPB$_2$ filter—at the mean of $p(x_{k+1}|m_{k+1}^i, z_{1:T})$. In other words, $p(x_{k+1}|m_{k+1}^i, z_{1:T})$ is assimilated to a Dirac delta function. The IMM-RTS smoother of [19] shows RMSE and wrong detection probabilities larger than the investigated solution too. As presented in section III, that method is indeed based on overly strong approximations which influence the associated algorithm. They occur in the interaction and smoothing steps: the equality $p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) = p(x_{k+1}|m_{k+1}^i, z_{1:T})$ is used and $p(x_{k+1}|x_k, m_k)$ is assumed to be the transition density between k and $k+1$ instead of $p(x_{k+1}|x_k, m_{k+1})$.

Concerning the respective computation times, the GPB$_2$-RTS smoother [17], [18] entails approximations which require a lower number of matrix inversions, and is faster in this example. The same holds for the IMM-RTS smoother [19], though at the expense of a lack of theoretical soundness. Nevertheless, the approach proposed in this paper involves a lower number of filters and smoothers in addition to providing improved accuracy. It is interesting when time and measurement updates are computationally expensive.

B. Fixed-lag smoothing

The evaluation of the proposed multiple-model smoother is completed by an example with nonlinear dynamics in the context of fixed-lag smoothing. The algorithm is compared in this section to the augmented state IMM smoother of [11] ("Morelande et al."). Two discrete-time coordinated turn (CT) models with unknown turn rates ω_j, with $j = 1, 2$, are considered. The target evolves in 2D and the state vector is now $x = \left[\mathsf{x},\ \dot{\mathsf{x}},\ \mathsf{y},\ \dot{\mathsf{y}}, \omega\right]^T$. The discrete-time dynamics of the CT models [30] is given by

$$x_k = \begin{pmatrix} 1 & \frac{\sin\omega_{k-1}\Delta t_k}{\omega_{k-1}} & 0 & -\frac{1-\cos\omega_{k-1}\Delta t_k}{\omega_{k-1}} & 0 \\ 0 & \cos\omega_{k-1}\Delta t_k & 0 & -\sin\omega_{k-1}\Delta t_k & 0 \\ 0 & \frac{1-\cos\omega_{k-1}\Delta t_k}{\omega_{k-1}} & 1 & \frac{\sin\omega_{k-1}\Delta t_k}{\omega_{k-1}} & 0 \\ 0 & \sin\omega_{k-1}\Delta t_k & 0 & \cos\omega_{k-1}\Delta t_k & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q_{k-1}^j \qquad (41)$$

with $Q_{k-1}^j = \mathrm{diag}([\sigma_{\mathsf{x},j}^2 M, \sigma_{\mathsf{y},j}^2 M, \sigma_{\omega,j}^2])$ and $M = \begin{pmatrix} \Delta t_k^3/3 & \Delta t_k^2/2 \\ \Delta t_k^2/2 & \Delta t_k \end{pmatrix}$. For the first mode, we set $\sigma_{\mathsf{x},1}^2 = 1$, $\sigma_{\mathsf{y},1}^2 = 1$, $\sigma_{\omega,1}^2 = 0$ (constant turn rate and nearly constant velocity mode 1) and, for the second mode, $\sigma_{\mathsf{x},2}^2 = 100$, $\sigma_{\mathsf{y},2}^2 = 100$, $\sigma_{\omega,2}^2 = 1.75 \times 10^{-3}$ (maneuvering mode 2). The output equation is given by

$$z_k = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{smallmatrix}\right) x_k + r_k \text{ with } R_k = 10^2 \mathbb{I}_2 \text{ m}^2. \qquad (42)$$

and the probability transition matrix is identical to (40).

As previously, the trajectory and measurements are randomly generated. The target is observed for 90 steps with a sampling period of $\Delta t_k = 1$s for all k. The initial state value is $\hat{x}_{0|0} = \left[0, 30, 0, 30, 0.05\right]^T$ with $P_{0|0} = \mathrm{diag}([10, 1, 10, 1, 0.001])$ for all modes. The switching times are set to $k = 30$ and $k = 60$. The maneuvering model is running between $k = 30$ and $k = 60$. The constant turn and velocity model is active before $k = 30$ and after $k = 60$. Fig. 3 gives a trajectory sample. Comparisons between the methods are performed over

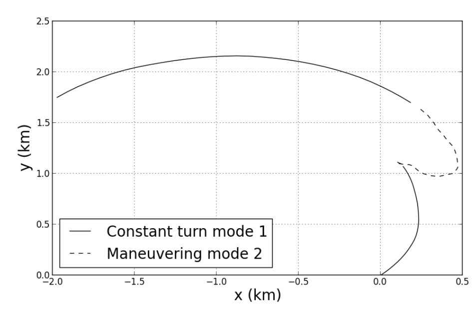

Fig. 3. Example of randomly generated trajectory for coordinated turn model

50 Monte Carlo runs. The time averaged empirical root-mean-squared errors are plotted as a function of the time lag for the position, the velocity and the turn rate in Fig. 4(a), Fig. 4(b) and Fig. 4(c) respectively. The observed time averaged wrong mode detection probability is displayed in Fig. 4(d). Regarding the RMSE, the IMM-RTS and Morelande's smoothers display a similar accuracy whatever the time lag. Besides, the wrong mode detection probabilities are higher with the IMM-RTS smoother for time lags 1 and 2. This pitfall arises from the impossibility to compute the smoothed mode probabilities until $\left(P_{k|k+1}^{b,i}\right)^{-1}$ becomes nonsingular in the backward-time recursion. As the state and measurement vector sizes are 5 and 2, respectively, a time lag of at least 3 must be reached before $\left(P_{k|k+1}^{b,i}\right)^{-1}$ is invertible. After that, the smoothed probabilities can be conveniently evaluated, and the IMM-RTS smoother reaches a slightly better detection rate of the active mode than Morelande's algorithm. The IMM-RTS smoother is faster in this example than the augmented-state implementation.

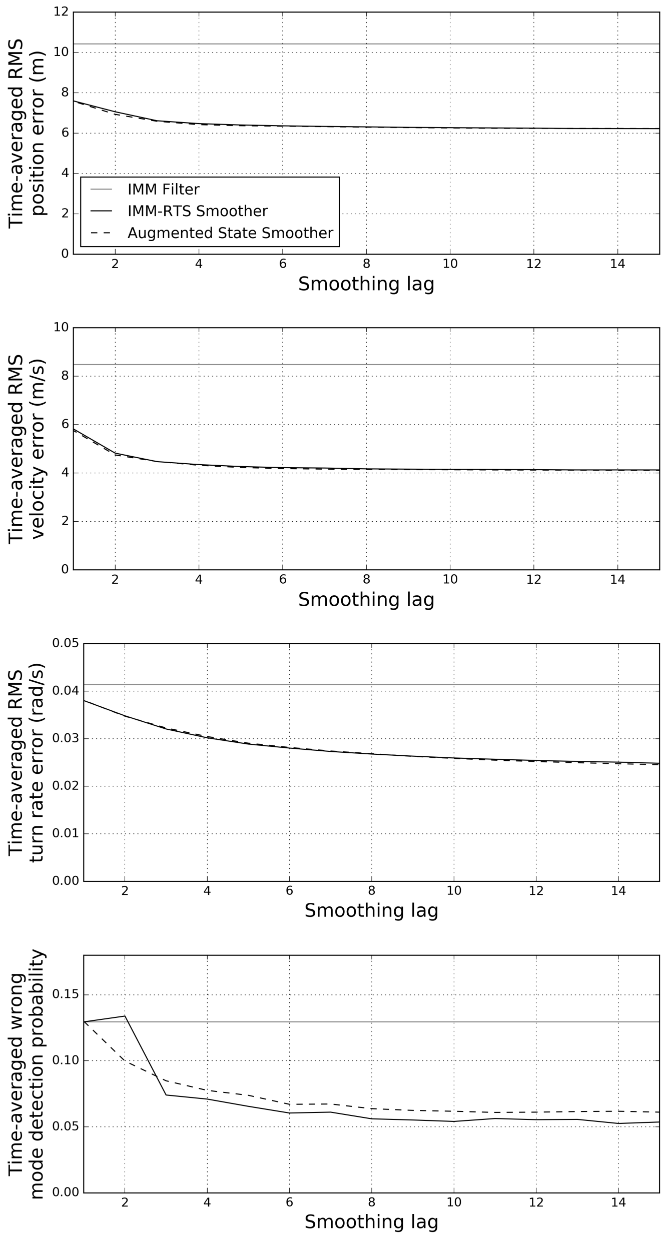

Fig. 4. Comparison of the IMM filter and the fixed-lag smoothers: (from top to bottom) (a) time averaged position RMSE, (b) time averaged velocity RMSE, (c) time averaged turn rate RMSE and (d) time averaged wrong mode detection probability.

VII. Conclusion and Prospects

This paper investigated a suboptimal fixed-interval smoothing algorithm based on a forward-time IMM filtering and a backward-time recursive process. Each recursion consists of a smoothing step and involves Rauch-Tung-Striebel equations adapted to jump Markov systems together with a specific interaction step to allow mode cooperation. The first smoothing stage runs only M Rauch-Tung-Striebel smoothers in parallel, each one being conditioned on one of the M possibly active modes within the sampling period $(t_k, t_{k+1}]$. Its results are then combined with interactions related to the M^2 admissible pairs of models over the successive sampling periods $(t_{k-1}, t_k]$ and $(t_k, t_{k+1}]$. Two complementary combination types are investigated, the second one being computationally cheaper. The recursions SR3 introduced above are not approximations: Step 1 is derived without approximations (end of Section III-B), the exact equation for Step 2 is (14) and Step 3 comes from (33). Approximations are only done when deriving a tractable solution by assuming that the pdf is Gaussian. In contrast, SR1 and SR2 consider approximations both in the recursion cycle and for the shape of the pdf. An example of tracking of a maneuvering target shows that the proposed smoother performs significantly better than the IMM filter [3], the GPB_2-RTS smoother [29], the IMM-RTS smoother [19] and equally well as the two-filter based scheme [22]. Unlike the latter, the proposed algorithm is suited to nonlinear dynamics and measurement equations. In the context of fixed-lag smoothing, comparisons with the solution of [11] displayed a similar accuracy and a better detection of the active mode for a sufficiently large time lag.

Future work will concentrate on adapting the proposed approach to a bank of heterogeneous-order models, *i.e.* to models which share only parts of their respective state vectors [9][10].

References

[1] Y. Bar-Shalom and X. Li, *Estimation and Tracking : Principles, Techniques and Software*. YBS Eds., Norwood, 1993.

[2] F. Gustafsson, *Adaptive Filtering and Change Detection*. John Wiley & Sons, 2000.

[3] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[4] L. Cork and R. Walker, "Sensor fault detection for UAVs using a nonlinear dynamic model and the IMM-UKF algorithm," in *IEEE Conf. on Information, Decision and Control (IDC'07)*, Adelaide, Australia, 2007, pp. 230–235.

[5] R. Toledo-Moreo, M. Zamora-Izquierdo, and A. Gomez-Skarmeta, "IMM-EKF based road vehicle navigation with low cost GPS/INS," in *IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI'2006)*, Heidelberg, Germany, 2006, pp. 433–438.

[6] H. Sorenson, Ed., *Kalman Filtering: Theory and Applications*. IEEE Press, 1985.

[7] S. Julier and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Int. Symp. on Aerospace/Defense Sensing, Simul. and Controls*, vol. 3, 1997.

TABLE II
NOTATION

Symbol	Meaning
$(.)^T$; $\|a\|_R^2$	transpose operator; weighted square norm $a^T Ra$ (with $R = R^T \geq 0$)
$P(.)$; $p(.)$; $E[.]$	probability; probability density function (pdf); expectation
$\mathscr{N}(\bar{x},X)$; $\mathscr{N}(x;\bar{x},X)$	(real) Gaussian distribution with mean $\bar{x}$ and covariance X; associated pdf on x
x_k; z_k;	base (continuous) state vector at time k; measurement vector at time k
$v_{a:b}$	shortcut for the sequence $v_a,\ldots,v_b$ between times a and b
m_k; $m_k = j = m_k^j$; M; $\mathscr{M}$;	mode (discrete) state vector; event "mode j is in effect during the sampling period $(t_{k-1},t_k]$"; number of modes; set of modes
q_{k-1}^j; r_k^j	dynamics and measurement noises in the state and output equations under the event m_k^j (those noises are Gaussian, zero-mean, white, mutually independent, and independent of the Gaussian base vector at $k=0$)
$\xi_k = (x_k,m_k)$; $\xi_k^j = (x_k,m_k^j)$	global (hybrid) state vector at time k; global state vector at time k under the event $m_k = j$
$P(m_0^j) = \mu_0^j$; $P(m_{k+1}^i\|m_k^j) = \pi_{ji}$	initial probability of mode j at $k=0$; transition probability between modes j and i (independent of the time index k)
$p(x_k\|z_{1:k'})$ with $k' > k$	posterior smoothing pdf at time k, conditioned on measurements up to time k'
$P(m_k^j\|z_{1:k})$; $\mu_{k\|k}^j$	posterior filtering probability of mode j; approximation computed by the IMM filter
$p(x_k\|m_k^j,z_{1:k})$; $\mathscr{N}(x_k;\hat{x}_{k\|k}^j,P_{k\|k}^j)$	posterior filtering pdf of the base state vector at time k conditioned on the event m_k^j (or "mode-conditioned" posterior pdf); Gaussian approximation computed by the IMM filter, with filtered moments $\hat{x}_{k\|k}^j,P_{k\|k}^j$ ($\hat{x}_{0\|0}^j,P_{0\|0}^j$ term the prior mean and covariance of the base vector at initial time $k=0$ under the event m_0^j)
$p(x_{k+1}\|m_{k+1}^i,z_{1:k})$; $\mathscr{N}(x_{k+1};\hat{x}_{k+1\|k}^i,P_{k+1\|k}^i)$	posterior prediction pdf of the base state vector at time $k+1$ conditioned on the event m_{k+1}^i; Gaussian approximation computed by the IMM filter, with moments $\hat{x}_{k+1\|k}^i,P_{k+1\|k}^i$
$P(m_k^j\|m_{k+1}^i,z_{1:k})$; $\mu_{k\|k}^{j\|i}$	mixing probability; approximation computed by the IMM filter
$p(x_k\|m_{k+1}^i,z_{1:k})$; $\mathscr{N}(x_k;\bar{x}_{k\|k}^i,\bar{P}_{k\|k}^i)$	mode-conditioned prior (*i.e.*, conditioned on the next event m_{k+1}^i) pdf; Gaussian approximation computed by the IMM filter, with moments $\bar{x}_{k\|k}^i,\bar{P}_{k\|k}^i$
$P(m_k^j\|z_{1:T})$; $\mu_{k\|T}^j$	smoothed probability probability of mode j; approximation computed by our IMM-RTS smoother
$p(x_k\|m_k^j,z_{1:T})$; $\mathscr{N}(x_k;\hat{x}_{k\|T}^j,P_{k\|T}^j)$	mode-conditioned smoothing pdf of the base state vector at time k conditioned on the event m_k^j; Gaussian approximation computed by our IMM-RTS smoother, with moments $\hat{x}_{k\|T}^j,P_{k\|T}^j$
$p(x_k\|m_{k+1}^i,z_{1:T})$; $\mathscr{N}(x_k;\bar{x}_{k\|T}^i,\bar{P}_{k\|T}^i)$	smoothed mixing pdf of the base state vector at time k; Gaussian approximation computed by our IMM-RTS smoother
$p(x_k\|m_k^j,m_{k+1}^i,z_{1:T})$; $\mathscr{N}(x_k;\hat{x}_{k\|T}^{ji},P_{k\|T}^{ji})$	smoothed mixing pdf conditioned on the sequence of modes over two consecutive sampling periods (or "two-mode conditioned smoothed pdf"), involved in the mode interaction stage (when based on M^2 combinations); Gaussian approximation in our IMM-RTS smoother
$P(m_{k+1}^i\|m_k^j,z_{1:T})$; $\bar{\mu}_{k+1\|T}^{i\|j}$	smoothed mixing probability involved in the mode interaction stage (when based on M^2 combinations); approximation in our IMM-RTS smoother
$p(x_k\|m_k^j,z_{k+1:T})$; $\mathscr{N}(x_k;\bar{x}_{k\|k+1}^{b,j},\bar{P}_{k\|k+1}^{b,j})$	mode-conditioned backward filtering density; Gaussian approximation, with backward filtered moments $\bar{x}_{k\|k+1}^{b,j},\bar{P}_{k\|k+1}^{b,j}$
$p(x_k\|m_{k+1}^i,z_{k+1:T})$; $\mathscr{N}(x_k;\hat{x}_{k\|k+1}^{b,i},P_{k\|k+1}^{b,i})$	mixing backward prediction pdf (or "mode-conditioned one-step backward-time prediction pdf" in [22]) involved in the mode interaction stage (when based on M combinations); Gaussian approximation in our IMM-RTS smoother
$P(m_{k+1}^i\|m_k^j,z_{k+1:T})$; $\mu_{k+1\|T}^{b,i\|j}$	backward mixing probability involved in the mode interaction stage (when based on M combinations); approximation in our IMM-RTS smoother; note that $P(m_{k+1}^i\|m_k^j,z_{k+1:T}) = P(m_{k+1}^i\|m_k^j,z_{1:T})$ so that $\mu_{k+1\|T}^{b,i\|j} = \bar{\mu}_{k+1\|T}^{i\|j}$

[8] Y. Boers and J. Driessen, "Interacting multiple model particle filter," *IEE Proc. on Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 344–349, Oct. 2003.

[9] T. Yuan, Y. Bar-Shalom, P. Willet, E. Mozeson, S. Pollak, and D. Hardiman, "A multiple IMM estimation approach with unbiased mixing for thrusting projectiles," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3250–3267, 2012.

[10] R. Lopez, P. Danès, and F. Royer, "Extending the IMM filter to heterogeneous-order state space models," in *IEEE Conf. on Decision and Control (CDC'2010)*, Atlanta, GA, 2010, pp. 7369–7374.

[11] M. Morelande and B. Ristic, "Smoothed state estimation for nonlinear Markovian switching systems," *IEEE Trans. Aerospace and Electronic Systems*, vol. 44, no. 4, pp. 1309–1325, 2008.

[12] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, p. 1, Mar. 1996.

[13] W. Fong, S. Godsill, A. Doucet, and M. West, "Monte Carlo smoothing with application to audio signal enhancement," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 438–449, Feb. 2002.

[14] S. Godsill, A. Doucet, and M. West, "Monte Carlo smoothing for nonlinear time series," *Journal of the American Statistical Association*, vol. 99, no. 465, pp. 156–168, Mar. 2004.

[15] P. Fearnhead, D. Wyncoll, and J. Tawn, "A sequential smoothing algorithm with linear computational cost," Tech. Rep., 2008.

[16] B.-N. Vo, B.-T. Vo, and R. Mahler, "Closed-form solutions to forward-backward smoothing," *IEEE Trans. on Signal Processing*, vol. 60, no. 1, pp. 2 –17, Jan. 2012.

[17] W. Koch, "Fixed-interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 36, no. 1, pp. 2—14, 2000.

[18] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012.

[19] N. Nadarajah, R. Tharmarasa, M. McDonald, and T. Kirubarajan, "IMM forward filtering and backward smoothing for maneuvering target tracking," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2673–2678, 2012.

[20] R. Lopez and P. Danès, "Exploiting Rauch-Tung-Striebel Formulae for IMM-based Smoothing of Markovian Switching Systems," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'2012)*. IEEE, 2012, pp. 3953–3956.

[21] ——, "A fixed-interval smoother with reduced complexity for jump Markov nonlinear systems," in *IEEE Int. Conf. on Information Fusion (FUSION'2014)*, 2014, pp. 1–8.

[22] R. Helmick, W. Blair, and S. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Trans. on Information Theory*, vol. 41, no. 6, pp. 1845–1855, 1995.

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

[24] S. Särkkä and J. Hartikainen, "On Gaussian optimal smoothing of non-linear state space models," *IEEE Trans. on Automatic Control*, vol. 55, no. 8, pp. 1938–1941, 2010.

[25] B. Chen and J. Tugnait, "Interacting multiple model fixed-lag smoothing algorithm for Markovian switching systems," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 36, no. 1, pp. 243–250, 2000.

[26] R. Helmick, W. Blair, and S. Hoffman, "One-step fixed-lag smoothers for Markovian switching systems," *IEEE Trans. on Automatic Control*, vol. 41, no. 7, pp. 1051–1056, 1996.

[27] V. Jilkov, X. Li, and L. Lu, "Performance enhancement of the IMM estimation by smoothing," in *IEEE Int. Conf. on Information Fusion (FUSION'2002)*, vol. 1, Annapolis, MD, 2002, pp. 713–720.

[28] X. Li and V. Jilkov, "Survey of maneuvering target tracking. Part V: Multiple-model methods," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.

[29] B. Mesot and D. Barber, "A simple alternative derivation of the expectation correction algorithm," *IEEE Signal Processing Letters*, vol. 16, no. 2, pp. 121–124, Feb. 2009.

[30] X. Li and V. Jilkov, "Survey of maneuvering target tracking. Part I: Dynamic models," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

Rémy Lopez received a MSc in mathematical and modeling engineering (2009) and the PhD in automatic control (2013) from the National Institute for Applied Science (Toulouse, France). Since 2009, he has been with Collecte Localisation Satellites (Ramonville-Saint-Agne, France), His research interests include target tracking, statistical signal processing and machine learning.

Patrick Danès is a full professor at Univ Paul Sabatier and LAAS-CNRS, Toulouse, France. He got his PhD (control) in 1995, and his accreditation to supervise research (robotics, signal processing and control) in 2010. His areas of research are robot audition and estimation/control problems in or outside robotics.