Received: 6 July 2021 | Revised: 7 September 2021 | Accepted: 25 September 2021

IET Radar, Sonar & Navigation

DOI: 10.1049/rsn2.12179

ORIGINAL RESEARCH PAPER

A marginalised particle filter with variational inference for non-linear state-space models with Gaussian mixture noise

Cheng Cheng[1] | Jean-Yves Tourneret[2] | Xiaodong Lu[1]

[1]School of Astronautics, Northwestern Polytechnical University, Xi'an, China

[2]University of Toulouse, ENSEEIHT-IRIT-TéSA, Toulouse Cedex 7, France

Correspondence
Cheng Cheng, School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China.
Email: cheng.cheng@nwpu.edu.cn.

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61901380

Abstract

This work proposes a marginalised particle filter with variational inference for non-linear state-space models (SSMs) with Gaussian mixture noise. A latent variable indicating the component of the Gaussian mixture considered at each time instant is introduced to specify the measurement mode of the SSM. The resulting joint posterior distribution of the state vector, the mode variable and the parameters of the Gaussian mixture noise is marginalised with respect to the noise variables. The marginalised posterior distribution of the state and mode is then approximated by using an appropriate marginalised particle filter. The noise parameters conditionally on each particle system of the state and mode variable are finally updated by using variational Bayesian inference. A simulation study is conducted to compare the proposed method with state-of-the-art approaches in the context of positioning in urban canyons using global navigation satellite systems.

KEYWORDS

Gaussian mixture noise, marginalised particle filter, non-Gaussian noise, non-linear state-space models, variational inference

1 | INTRODUCTION

Non-linear state-space models (SSMs), composed of a non-linear system and measurement equations, are applied to a wide variety of practical applications including global navigation satellite system (GNSS) positioning [1], radar target tracking [2] and communication systems [3]. A central problem when using these models is recursively inferring the state vector based on a sequence of measurements. In many practical applications, the measurement noise has non-Gaussian statistics resulting from measurement outliers, for instance when the GNSS receiver is affected by multipath (MP) signals in urban canyons [4] or when there are irregular electromagnetic wave reflections from the target surface in radar target tracking [5].

The probability density function (pdf) of a non-Gaussian noise can be well-approximated by using a finite sum of Gaussian pdfs according to the density approximation theorem [6]. As a consequence, different kinds of filters with noises based on Gaussian mixture models (GMMs) have been investigated in the literature for solving the state estimation problem in the presence of non-Gaussian measurement noise. One attempt to solve this problem is based on the multiple model approach, such as the Gaussian sum filter [7, 8] where each component of the GMM corresponds to a possible noise distribution and the posterior estimates of the state vector can be obtained by using a bank of Kalman filters. Mixture reduction strategies such as the interactive multiple models (IMM) [9–11] have also been investigated to try to prevent the number of components of the joint posterior from growing exponentially over time. When the parameters of the GMM are unknown a priori, the expectation-maximisation (EM) algorithm can be embedded into the IMM for simultaneously estimating the unknown state and noise parameters [12]. In order to adaptively determine the number of GMM components, variational Bayesian (VB) EM approaches have been studied in Ref [13]. These approaches estimate the state vector in the VB-expectation step and then update the posterior distributions of the GMM in the VB-maximisation step. However, to our knowledge, they have only been considered in a batch manner. A flexible Bayesian non-parametric model embedded into the Rao–Blackwellized particle filter has also been considered in Ref [14], modelling the measurement noise as an infinite mixture of Gaussian vectors. However, the price to pay with this kind of approach is its high-computational complexity, which may limit its use in some practical applications.

This work studies a new marginalised particle filter (MPF) for non-linear SSMs with observation noise distributed according to a mixture of Gaussian vectors. The MPF studied in Ref [15] for mixed linear/non-linear systems is introduced for state and measurement equations depending linearly on a subset of the state vector and non-linearly on the other state variables. The idea of this technique is to marginalise the distribution of interest with respect to the state variables appearing linearly in the state and measurement equations, allowing the linear components to be processed using analytical methods (such as the Kalman filter) and to handle the non-linear components by sequential Monte Carlo (SMC) techniques [16]. The MPF has been used successfully in many applications including navigation using an inertial navigation system and other aided positioning and multiple target tracking [17, 18]. For instance, an MPF has been proposed in Ref [19, 20] to estimate jointly the number of targets and their states using a sequential algorithm. An overview of the sequential Monte Carlo methods for tracking different objects was investigated in Ref [21]. In addition, the MPF has been applied to track-before-detect for a target tracking problem in Ref [22, 23]. The MPF was also combined with a variational Bayes approximation in Ref [24–26] to estimate the state and unknown measurement noise parameters in non-linear SSMs. For example, the means and covariances of the measurement noise were assumed to be slowly time-varying in Ref [24]. This assumption allowed the derivation of a marginalised adaptive particle filter to infer the unknown noise parameters jointly with the state vector by using maximum entropy with a Kullback–Leibler (KL) divergence constraint. Finally, it is interesting to mention the robust particle filter proposed in Ref [25, 26] to address cases where the additive measurement noise has a Student-*t* distribution.

This work introduces a new strategy for state estimation in non-linear SSMs inspired by the MPF. The main innovations with respect to the existing approaches are summarised below:

- We introduce a latent variable indicating the component of the GMM used for the measurement noise at each time instant and specify the measurement mode of the system.
- The joint posterior distribution of the state vector, the latent variable and the GMM parameters is derived and its parameters are estimated using an MPF with variational inference.

More precisely, the proposed approach is decomposed into two steps: (a) The joint posterior distribution of the state vector, the latent variable and the GMM parameters is marginalised with respect to the noise variables. The resulting marginalised posterior distribution is approximated by using an appropriate MPF; (b) the noise parameters conditionally on each particle generated in (a) are updated by using VB inference.

The work is organised as follows: The problem of state estimation for a non-linear SSM with a GMM measurement noise is presented in Section 2. Section 3 studies a new MPF with variational inference for jointly estimating the state vector, the mode variable and the GMM parameters of this kind of non-linear SSM. The performance of the proposed approach is evaluated in Section 4 and compared with the state-of-the-art approaches in the context of positioning in urban canyons using GNSS. Conclusions are finally drawn in Section 5.

2 | PROBLEM FORMULATION

In this work, we consider the following non-linear discrete-time SSM related to a hidden state vector $\boldsymbol{x}_t \in \mathbb{R}^{n_x}$ and the measurement vector $\boldsymbol{y}_t \in \mathbb{R}^{n_y}$

$$\boldsymbol{x}_t = f(\boldsymbol{x}_{t-1}) + \boldsymbol{\omega}_t, \tag{1}$$

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t) + \boldsymbol{v}_t, \tag{2}$$

where $t = 1, \ldots, T$ denotes the tth time instant, $h(\cdot)$ is a known non-linear measurement function, $f(\cdot)$ is a known linear or non-linear transition function and $\boldsymbol{\omega}_t$ is the process noise distributed according to a zero-mean Gaussian distribution of covariance matrix $\boldsymbol{Q}$. The non-Gaussian measurement noise is approximated using a finite GMM with density

$$p(\boldsymbol{v}_t|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1}\right), \tag{3}$$

where K denotes the number of Gaussian components in the mixture, π_k is the mixing coefficient for the kth component with $\sum_{k=1}^{K} \pi_k = 1$ and $\pi_k > 0$, $\mathcal{N}\left(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1}\right)$ denotes a Gaussian distribution with mean vector $\boldsymbol{\mu}_k$ and precision (inverse covariance) matrix $\boldsymbol{\Lambda}_k$ for the kth component, $\boldsymbol{\mu} = \{\boldsymbol{\mu}_k\}_{k=1}^{K}$ and $\boldsymbol{\Lambda} = \{\boldsymbol{\Lambda}_k\}_{k=1}^{K}$. Accordingly, the state and measurement equation in (1) and (2) can be defined by the following conditional pdfs:

$$\boldsymbol{x}_t \sim p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(f(\boldsymbol{x}_{t-1}), \boldsymbol{Q}), \tag{4}$$

$$\boldsymbol{y}_t \sim \sum_{k=1}^{K} \pi_k \mathcal{N}\left(h(\boldsymbol{x}_t) + \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1}\right). \tag{5}$$

Since each component of $\boldsymbol{y}_t$ can be considered as a measurement mode with a specified distribution, the measurement switches among different modes corresponding to the different GMM components. When working with mixture distributions, it is common to introduce a latent variable r_k indicating the mode of $\boldsymbol{y}_t$ at the tth time instant with a given probability $u_{k,t}$ defined as [27]

$$u_{k,t} \triangleq \Pr(r_t = k), \tag{6}$$

where $k = 1, \ldots, K$, $\sum_{k=1}^{K} u_{k,t} = 1$ and $u_{k,t} > 0$. On using these notations, the vector $\boldsymbol{u}_t = (u_{1,t}, \ldots, u_{K,t})^{\mathrm{T}}$ contains the probabilities that a measurement vector at the tth time instant belongs to the different GMM components, which are time-varying quantities. Moreover, the vector of mixing coefficients $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)^{\mathrm{T}}$ contains the probabilities that the measurements in $\boldsymbol{y}_{1:t} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$ are associated with the different GMM components, which are assumed to be constant quantities in this work.

When considering Bayesian estimators, assigning conjugate priors to $\boldsymbol{\pi}, \boldsymbol{\mu}$ and $\boldsymbol{\Lambda}$ simplifies the analysis. In this work, we have assigned a conjugate Dirichlet distribution $\mathrm{Dir}(\cdot)$ to $\boldsymbol{\pi}$ and a normal Wishart distribution $\mathcal{NW}(\cdot)$ to $\{\boldsymbol{\mu}, \boldsymbol{\Lambda}\}$ [28] leading to

$$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}) = C(\boldsymbol{\alpha}) \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}, \tag{7}$$

$$\begin{aligned} p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= \prod_{k=1}^{K} \mathcal{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \boldsymbol{m}_k, \beta_k, \boldsymbol{W}_k, \nu_k), \\ &= \prod_{k=1}^{K} \mathcal{N}\left(\boldsymbol{\mu}_k | \boldsymbol{m}_k, (\beta_k \boldsymbol{\Lambda}_k)^{-1}\right) \mathcal{W}(\boldsymbol{\Lambda}_k | \boldsymbol{W}_k, \nu_k), \end{aligned} \tag{8}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^{\mathrm{T}}$ ($\alpha_k > 1$ and $k = 1, \ldots, K$) is a concentration parameter vector, $C(\boldsymbol{\alpha})$ is the normalisation constant for the Dirichlet distribution, $\mathcal{W}(\boldsymbol{\Lambda}_k | \boldsymbol{W}_k, \nu_k)$ is the Wishart distribution with precision matrix $\boldsymbol{W}_k$ and ν_k degrees of freedom (with $\nu_k > n_y + 1$, where n_y is the dimension of the measurement vector). Accordingly, the joint prior distribution for the parameters of the noise GMM is defined as follows:

$$p(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}) \prod_{k=1}^{K} \mathcal{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \boldsymbol{m}_k, \beta_k, \boldsymbol{W}_k, \nu_k). \tag{9}$$

When the measurement vector $\boldsymbol{y}_t$ switches between different modes, the latent variable r_t is time-varying. In addition, the parameters of the GMM used for the measurement noise are generally difficult to specify a priori. Thus, the aim of this work is to infer the states $\boldsymbol{x}_t$ jointly with the unknown measurement mode variables and the GMM noise parameters, leading to the joint posterior distribution of all unknown variables given the measurement sequence $\boldsymbol{y}_{1:t}$, denoted as $p(\boldsymbol{x}_t, r_t, \boldsymbol{\theta} | \boldsymbol{y}_{1:t})$ with $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Lambda}\}$. This posterior distribution is studied in the next section.

3 | A MARGINALISED PARTICLE FILTER FOR NON-LINEAR SSMs WITH A GMM FOR THE MEASUREMENT NOISE

According to the MPF concept [15], the joint posterior distribution of the state and unknown parameters $p(\boldsymbol{x}_t, r_t, \boldsymbol{\theta} | \boldsymbol{y}_{1:t})$ can be factorised according to the following chain rule of probability:

$$\begin{aligned} p(\boldsymbol{x}_t, r_t, \boldsymbol{\theta} | \boldsymbol{y}_{1:t}) &= p(\boldsymbol{\theta} | \boldsymbol{x}_t, r_t, \boldsymbol{y}_{1:t}) p(\boldsymbol{x}_t, r_t | \boldsymbol{y}_{1:t}), \\ &= p(\boldsymbol{\theta} | \boldsymbol{x}_t, r_t, \boldsymbol{y}_{1:t}) p(r_t | \boldsymbol{x}_t, \boldsymbol{y}_{1:t}) p(\boldsymbol{x}_t | \boldsymbol{y}_{1:t}), \end{aligned} \tag{10}$$

where the three pdfs on the right-hand side of Equation (10) can be calculated recursively. More precisely, we propose to approximate the posterior pdf $p(\boldsymbol{x}_t | \boldsymbol{y}_{1:t})$ by using an empirical density following the principle of particle filters. The pdf of the mode variables $p(r_t | \boldsymbol{x}_t, \boldsymbol{y}_{1:t})$ can then be computed conditionally on the state $\boldsymbol{x}_t$. Finally, the pdf of the noise parameters is shown to be a Gaussian mixture $p(\boldsymbol{\mu}, \ \boldsymbol{\Lambda} | \boldsymbol{x}_t, r_t, \boldsymbol{y}_{1:t})$ conditionally on the state $\boldsymbol{x}_t$ and the mode variable r_t, which can be obtained using VB inference. The different steps required to compute these quantities are described in the next sections.

3.1 | Updating $(\boldsymbol{x}_t, r_t)$ samples based on the marginalised particle filter

As explained before, the joint posterior pdf of $(\boldsymbol{x}_t, r_t)$ can be factorised as follows:

$$p(\boldsymbol{x}_t, r_t | \boldsymbol{y}_{1:t}) = p(r_t | \boldsymbol{x}_t, \boldsymbol{y}_{1:t}) p(\boldsymbol{x}_t | \boldsymbol{y}_{1:t}), \tag{11}$$

where the mode variable r_t has been marginalised out in the second term of the right hand side. This work proposes to approximate $p(\boldsymbol{x}_t | \boldsymbol{y}_{1:t})$ by using a particle filter, that is,

$$p(\boldsymbol{x}_t | \boldsymbol{y}_{1:t}) \approx \sum_{i=1}^{N} \omega_t^i \delta\left(\boldsymbol{x}_t - \boldsymbol{x}_t^i\right), \tag{12}$$

where N is the number of particles, $\delta(\cdot)$ is the Dirac delta function, $\boldsymbol{x}_t^i$ is the ith particle and ω_t^i is the corresponding weight at the tth time instant. According to Equation (6), $p(r_t | \boldsymbol{x}_t, \boldsymbol{y}_{1:t})$ can be obtained by computing the mode probabilities conditionally on the state samples $\{\boldsymbol{x}_t^i\}_{i=1}^{N}$:

$$u_{k,t}^i = \Pr\left(r_t = k | \boldsymbol{x}_t^i, \boldsymbol{y}_t\right), \tag{13}$$

where $k = 1, \ldots, K$ and $i = 1, \ldots, N$. As a consequence, the joint pdf $p(\boldsymbol{x}_t, r_t | \boldsymbol{y}_{1:t})$ is approximated by using a set of weighted particles, leading to $\{\omega_t^i, \boldsymbol{x}_t^i, \boldsymbol{u}_t^i\}_{i=1}^{N}$, whose update is described in the rest of this section.

According to the theory of particle filters [29], the weight ω_t^i can be updated as follows:

$$\omega_t^i \propto \frac{p\left(\boldsymbol{y}_t | \boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right) p\left(\boldsymbol{x}_t^i | \boldsymbol{y}_{1:t-1}\right)}{q\left(\boldsymbol{x}_t^i | \boldsymbol{x}_{t-1}^i, \boldsymbol{y}_{1:t}\right)} \omega_{t-1}^i, \tag{14}$$

where $\boldsymbol{x}_t^i \sim q\left(\boldsymbol{x}_t | \boldsymbol{x}_{1:t-1}^i, \boldsymbol{y}_{1:t}\right)$, $q\left(\boldsymbol{x}_t | \boldsymbol{x}_{1:t-1}^i, \boldsymbol{y}_{1:t}\right)$ is an importance distribution, which is chosen as the bootstrap proposal [16] in this work, that is,

$$\boldsymbol{x}_t^i \sim p\left(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}^i\right), \quad i=1,\dots,N. \tag{15}$$

Accordingly, Equation (14) can be rewritten as follows:

$$\omega_t^i \propto p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right)\omega_{t-1}^i, \tag{16}$$

where $p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right)$ can be computed after the following marginalizations:

$$\begin{aligned} &p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right) \\ &= \iiint \prod_{k=1}^{K} \left\{p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i,\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i\right)p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t-1}\right)\right\}^{\mathbb{I}\left(r_t^i=k\right)} \\ &\quad p\left(r_t^i|\boldsymbol{y}_{1:t-1}\right)\mathrm{d}r_t\mathrm{d}\boldsymbol{\mu}_k\mathrm{d}\boldsymbol{\Lambda}_k, \end{aligned} \tag{17}$$

where $\mathbb{I}(r_t=k)$ is an indicator for the kth component of the noise GMM (equal to 1 when $r_t=k$ and 0 otherwise) and $r_t \in \{1,\dots,K\}$ is the discrete variable indicating the mixture component. Note that $p\left(r_t^i|\boldsymbol{y}_{1:t-1}\right)$ and $p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t-1}\right)$ are the predictive distributions for the indicator variable and the mean and precision matrix of the kth Gaussian component for the ith particle. The predictive distribution $p\left(r_t^i|\boldsymbol{y}_{1:t-1}\right)$ can be defined according to Equation (6), that is,

$$\widehat{u}_{k,t|t-1}^i = \Pr\left(r_t^i=k|\boldsymbol{y}_{1:t-1}\right), \tag{18}$$

where $\widehat{u}_{k,t|t-1}^i$ denotes the predicted probability of the kth measurement mode for the ith particle at the tth time instant. It can be assumed that the mode sequence has the Markov property. Thus, the joint pdf of the mode sequence can be defined as follows [12, 30]:

$$\Pr(r_t=k, r_{t-1}=l|\boldsymbol{y}_{1:t-1}) = \Pi(k|l)\Pr(r_{t-1}=l|\boldsymbol{y}_{1:t-1}), \tag{19}$$

with

$$\Pi_{lk} = \Pi(k|l) \triangleq \Pr(r_t=k|r_{t-1}=l), \tag{20}$$

where Π_{lk} is the transition probability from the lth to the kth measurement mode. By marginalising the above expression over r_{t-1}, $\widehat{u}_{k,t|t-1}^i$ can be expressed as follows:

$$\widehat{u}_{k,t|t-1}^i = \sum_{l=1}^{K} \Pi_{lk}\widehat{u}_{l,t-1}^i, \tag{21}$$

where $i=1,\dots,N$ and $k=1,\dots,K$, $\widehat{u}_{l,t-1}^i$ denotes the estimated probability of the lth measurement mode for the ith particle at the $(t-1)$th time instant. After replacing Equation (18) into Equation (17), the following result is obtained:

$$\begin{aligned} p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right) &= \sum\nolimits_{k=1}^{K} \widehat{u}_{k,t|t-1}^i \\ &\iint p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i,\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i, r_t=k\right)p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t-1}\right)\mathrm{d}\boldsymbol{\mu}_k\mathrm{d}\boldsymbol{\Lambda}_k. \end{aligned} \tag{22}$$

The conditional pdf of the measurement associated with the kth GMM component can be determined from Equation (5), leading to

$$p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i,\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i, r_t=k\right) = \mathcal{N}\left(h\left(\boldsymbol{x}_t^i\right)+\boldsymbol{\mu}_k^i, \left(\boldsymbol{\Lambda}_k^i\right)^{-1}\right). \tag{23}$$

Since the prior distribution for $\{\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k\}$ follows a normal Wishart distribution as defined in Equation (8), the predictive distribution $p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t-1}\right)$ can be defined as follows:

$$\begin{aligned} &p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t-1}\right) \\ &\quad = \mathcal{NW}\left(\boldsymbol{\mu}_{k,t}^i,\boldsymbol{\Lambda}_{k,t}^i|\widehat{\boldsymbol{m}}_{k,t|t-1}^i, \widehat{\beta}_{k,t|t-1}^i, \widehat{\boldsymbol{W}}_{k,t|t-1}^i, \widehat{\nu}_{k,t|t-1}^i\right), \end{aligned} \tag{24}$$

where $\widehat{\boldsymbol{m}}_{k,t|t-1}^i$, $\widehat{\beta}_{k,t|t-1}^i$, $\widehat{\boldsymbol{W}}_{k,t|t-1}^i$ and $\widehat{\nu}_{k,t|t-1}^i$ are the predicted hyperparameters of the normal and Wishart distributions for the ith particle at the tth time instant. In order to maintain the conjugacy for the distribution of the noise parameters, $\widehat{\boldsymbol{m}}_{k,t|t-1}$, $\widehat{\beta}_{k,t|t-1}$, $\widehat{\boldsymbol{W}}_{k,t|t-1}$ and $\widehat{\nu}_{k,t|t-1}$ can be propagated as follows:

$$\begin{aligned} &\widehat{\boldsymbol{m}}_{k,t|t-1}^i = \widehat{\boldsymbol{m}}_{k,t-1}^i, \ \widehat{\beta}_{k,t|t-1}^i = \widehat{\beta}_{k,t-1}^i, \\ &\widehat{\boldsymbol{W}}_{k,t|t-1}^i = \widehat{\boldsymbol{W}}_{k,t-1}^i, \ \widehat{\nu}_{k,t|t-1}^i = \widehat{\nu}_{k,t-1}^i, \end{aligned} \tag{25}$$

where $i=1,\dots,N$ and $k=1,\dots,K$, $\widehat{\boldsymbol{m}}_{k,t-1}^i$, $\widehat{\beta}_{k,t-1}^i$, $\widehat{\boldsymbol{W}}_{k,t-1}^i$ and $\widehat{\nu}_{k,t-1}^i$ are the estimated hyperparameters of the normal and Wishart distributions for the ith particle at the $(t-1)$th time instant. The a priori quantities $\widehat{\boldsymbol{m}}_{k,t|t-1}$, $\widehat{\beta}_{k,t|t-1}$, $\widehat{\boldsymbol{W}}_{k,t|t-1}$ and $\widehat{\nu}_{k,t|t-1}$ are then used as the initial values for VB inference (detailed in the Section 3.2), leading to the a posteriori quantities $\widehat{\boldsymbol{m}}_{k,t}^i$, $\widehat{\beta}_{k,t}^i$, $\widehat{\boldsymbol{W}}_{k,t}^i$ and $\widehat{\nu}_{k,t}^i$. As a consequence, $p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t}\right)$ is also a normal Wishart distribution denoted as

$$\begin{aligned} &p\left(\boldsymbol{\mu}_k^i,\boldsymbol{\Lambda}_k^i|\boldsymbol{y}_{1:t}\right) \\ &\quad = \mathcal{NW}\left(\boldsymbol{\mu}_{k,t}^i,\boldsymbol{\Lambda}_{k,t}^i|\widehat{\boldsymbol{m}}_{k,t}^i, \widehat{\beta}_{k,t}^i, \widehat{\boldsymbol{W}}_{k,t}^i, \widehat{\nu}_{k,t}^i\right). \end{aligned} \tag{26}$$

Since integration over the product of a normal Wishart prior and a Gaussian distribution leads to a Student-t distribution $St(\cdot)$ [31], inserting Equations (23) and (24) into Equation (22) leads to

$$\begin{aligned} &p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^i, \boldsymbol{y}_{1:t-1}\right) \\ &\quad = \sum\nolimits_{k=1}^{K} u_{k,t|t-1}^i St\left(\boldsymbol{y}_t|\widehat{\boldsymbol{m}}_{k,t|t-1}^i, \widehat{\boldsymbol{L}}_{k,t|t-1}^i, \widehat{\nu}_{k,t|t-1}^i - n_y + 1\right), \end{aligned} \tag{27}$$

where $\widehat{\boldsymbol{L}}_{k,t|t-1}^i = \frac{1+\widehat{\beta}_{k,t|t-1}^i}{\left(\widehat{\nu}_{k,t|t-1}^i - n_y+1\right)\widehat{\beta}_{k,t|t-1}^i}\widehat{\boldsymbol{W}}_{k,t|t-1}^i$ and $St\left(\boldsymbol{y}_t|\widehat{\boldsymbol{m}}_{k,t|t-1}^i, \widehat{\boldsymbol{L}}_{k,t|t-1}^i, \widehat{\nu}_{k,t|t-1}^i - n_y + 1\right)$ is a Student-t distribution defined as follows:

$$St\left(\boldsymbol{y}_t|\widehat{\boldsymbol{m}}^i_{k,t|t-1},\widehat{\boldsymbol{L}}^i_{k,t|t-1},\widehat{\nu}^i_{k,t|t-1}-n_y+1\right)\propto\left|1+\left(\boldsymbol{\xi}^i_t-\widehat{\boldsymbol{m}}^i_{k,t|t-1}\right)^{\mathrm{T}}\frac{\widehat{\beta}^i_{k,t|t-1}\left(\widehat{\boldsymbol{W}}^i_{k,t|t-1}\right)^{-1}}{1+\widehat{\beta}^i_{k,t|t-1}}\left(\boldsymbol{\xi}^i_t-\widehat{\boldsymbol{m}}^i_{k,t|t-1}\right)\right|^{-\frac{\widehat{\nu}^i_{k,t|t-1}+1}{2}} \tag{28}$$

with $\boldsymbol{\xi}^i_t=\boldsymbol{y}_t-h(\boldsymbol{x}^i_t)$. Since $p(r_t=k|\boldsymbol{x}_t,\boldsymbol{y}_{1:t})\propto p(\boldsymbol{y}_t|\boldsymbol{x}_t,\boldsymbol{y}_{1:t-1},r_t=k)p(r_t=k|\boldsymbol{y}_{1:t-1})$, the probability of the kth measurement mode for the ith particle at the tth time instant can be obtained as follows:

$$\widehat{u}^i_{k,t}=\frac{\widehat{\gamma}^i_{k,t}}{\sum_{k=1}^K\widehat{\gamma}^i_{k,t}},\quad k=1,\ldots,K \tag{29}$$

where

$$\widehat{\gamma}^i_{k,t}=\widehat{u}^i_{k,t|t-1}St\left(\boldsymbol{y}_t|\widehat{\boldsymbol{m}}^i_{k,t|t-1},\widehat{\boldsymbol{L}}^i_{k,t|t-1},\widehat{\nu}^i_{k,t|t-1}-n_y+1\right). \tag{30}$$

As a consequence, the maximum a posteriori (MAP) estimator of the indicator r_t for the ith particle at the tth time is

$$\widehat{r}^i_t=\underset{r_t=\{1,\ldots,K\}}{\operatorname{argmax}}\ \widehat{u}^i_{k,t},\quad i=1,\ldots,N. \tag{31}$$

3.2 | Calculating $p\left(\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t}\right)$ using VB inference

According to the mean-field theory in VB inference [32], the joint pdf of the noise parameters $q(\boldsymbol{\theta})$ can be factorised into single-variable factors, that is, $q(\boldsymbol{\theta})=q(\boldsymbol{\pi})q(\boldsymbol{\mu},\boldsymbol{\Lambda})$. According to the factorised approximation, the logarithm of the marginal likelihood $\ln p\left(\boldsymbol{y}_t|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{\theta},\boldsymbol{y}_{1:t-1}\right)$ can be expressed as follows [33]:

$$\ln p\left(\boldsymbol{y}_t|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{\theta},\boldsymbol{y}_{1:t-1}\right)=\mathcal{L}+\mathrm{KL}(q\|p), \tag{32}$$

with

$$\mathcal{L}=\int q(\boldsymbol{\theta})\ln\frac{p\left(\boldsymbol{y}_t,\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t-1}\right)}{q(\boldsymbol{\theta})}\mathrm{d}\boldsymbol{\theta}, \tag{33}$$

and

$$\mathrm{KL}(q\|p)=\int q(\boldsymbol{\theta})\ln\frac{q(\boldsymbol{\theta})}{p\left(\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t}\right)}\mathrm{d}\boldsymbol{\theta}, \tag{34}$$

where $\mathcal{L}$ is a variational objective function used in VB inference, $\mathrm{KL}(q\|p)$ is the Kullback–Leibler (KL) divergence between the true posterior and its approximation. Accordingly, the joint pdf $p\left(\boldsymbol{y}_t,\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t-1}\right)$ can be expressed, according to Equations (5), (7) and (8), as follows:

$$\begin{aligned}p\left(\boldsymbol{y}_t,\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t-1}\right)&=C(\boldsymbol{\alpha})\\&\times\prod_{k=1}^K\Big\{\pi_k^{\alpha_k-1}\mathcal{N}\left(\boldsymbol{\mu}_k|\boldsymbol{m}_k,(\beta_k\boldsymbol{\Lambda}_k)^{-1}\right)\mathcal{W}(\boldsymbol{\Lambda}_k|\boldsymbol{W}_k,\nu_k)\\&\times\left\{\pi_k p\left(\boldsymbol{y}_t|\boldsymbol{x}^i_t,\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k,r^i_t=k\right)\right\}^{\mathbb{I}(r_t=k)}\Big\}.\end{aligned} \tag{35}$$

Considering that the KL divergence is non-negative, minimisation the KL divergence can be achieved by maximising the lower bound $\mathcal{L}$, which results in the computation of expectations with respect to $q(\boldsymbol{\pi})$, $q(\boldsymbol{\mu},\boldsymbol{\Lambda})$ in turn, that is,

$$\ln q(\boldsymbol{\pi})=\mathrm{E}_{\boldsymbol{\theta}\backslash\boldsymbol{\pi}}\left[\ln p\left(\boldsymbol{y}_t,\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t-1}\right)\right], \tag{36}$$

$$\ln q(\boldsymbol{\mu},\boldsymbol{\Lambda})=\mathrm{E}_{\boldsymbol{\theta}\backslash\boldsymbol{\mu},\boldsymbol{\Lambda}}\left[\ln p\left(\boldsymbol{y}_t,\boldsymbol{\theta}|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{y}_{1:t-1}\right)\right], \tag{37}$$

where $\mathrm{E}_{\boldsymbol{\theta}_t\backslash\mathcal{X}}$ denotes the expectation with respect to the variational distributions of all variables in $\boldsymbol{\theta}$, except those contained in $\mathcal{X}$. Finally, the variational distributions can be approximated as follows:

$$\ln q(\boldsymbol{\pi})\approx\sum_{k=1}^K\left\{\mathbb{I}\left(r^i_t=k\right)+\alpha_k-1\right\}\ln\pi_k+C_1, \tag{38}$$

$$\begin{aligned}\ln q(\boldsymbol{u},\boldsymbol{\Lambda})&\approx\ln p\left(\boldsymbol{y}_t|\boldsymbol{x}^i_t,r^i_t,\boldsymbol{\mu},\boldsymbol{\Lambda}\right)+\ln p(\boldsymbol{\mu},\boldsymbol{\Lambda})+C_2,\\&\propto\sum_{k=1}^K\mathbb{I}\left(r^i_t=k\right)\left\{\ln\left(\mathcal{N}\left(h\left(\boldsymbol{x}^i_t\right)+\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k^{-1}\right)\right)\right.\\&\left.+\ln(\mathcal{NW}(\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k))\right\},\end{aligned} \tag{39}$$

where C_1 is a constant summarising the terms independent of $\boldsymbol{\pi}$ and C_2 contains the terms independent of $\boldsymbol{\mu}$ and $\boldsymbol{\Lambda}$.

The hyperparameters α^i_k, $\boldsymbol{m}^i_k$, β^i_k, $\boldsymbol{W}^i_k$, ν^i_k for the variational distributions of the GMM noise parameters conditionally on the ith weighted particle $\left\{\omega^i_t,\boldsymbol{x}^i_t,\boldsymbol{u}^i_t\right\}$ can be finally updated at the tth time instant as follows:

$$\widehat{\alpha}^i_{k,t}=\widehat{\alpha}^i_{k,t|t-1}+\mathbb{I}\left(\widehat{r}^i_t=k\right), \tag{40}$$

$$\widehat{\beta}^{i}_{k,t} = \widehat{\beta}^{i}_{k,t|t-1} + \mathbb{I}\left(\widehat{r}^{i}_{t} = k\right), \tag{41}$$

$$\widehat{\boldsymbol{m}}^{i}_{k,t} = \frac{1}{\widehat{\beta}^{i}_{k,t|t-1} + \mathbb{I}\left(\widehat{r}^{i}_{t} = k\right)} \left(\widehat{\beta}^{i}_{k,t|t-1}\widehat{\boldsymbol{m}}^{i}_{k,t|t-1} + \mathbb{I}\left(\widehat{r}^{i}_{t} = k\right)\left(\boldsymbol{\xi}^{i}_{t} - \widehat{\boldsymbol{m}}^{i}_{k,t|t-1}\right)\right), \tag{42}$$

$$\widehat{\nu}^{i}_{k,t} = \widehat{\nu}^{i}_{k,t|t-1} + \mathbb{I}\left(\widehat{r}^{i}_{t} = k\right), \tag{43}$$

$$\left(\widehat{\boldsymbol{W}}^{i}_{k,t}\right)^{-1} = \left(\widehat{\boldsymbol{W}}^{i}_{k,t|t-1}\right)^{-1} + \frac{\mathbb{I}\left(\widehat{r}^{i}_{t} = k\right)\widehat{\beta}^{i}_{k,t|t-1}}{\widehat{\beta}^{i}_{k,t|t-1} + \mathbb{I}\left(\widehat{r}^{i}_{t} = k\right)} \left(\boldsymbol{\xi}^{i}_{t} - \widehat{\boldsymbol{m}}^{i}_{k,t|t-1}\right)\left(\boldsymbol{\xi}^{i}_{t} - \widehat{\boldsymbol{m}}^{i}_{k,t|t-1}\right)^{\mathrm{T}}, \tag{44}$$

where $k = 1, \ldots, K$ and $i = 1, \ldots, N$. According to the estimate of the indicating variable $\widehat{r}^{i}_{t}$ in Equation (31), the indicator $\mathbb{I}\left(\widehat{r}^{i}_{t} = k\right)$ appearing in the above equations is defined as follows:

$$\mathbb{I}\left(\widehat{r}^{i}_{t} = k\right) = \begin{cases} 1 & \widehat{r}^{i}_{t} = k, \\ 0 & \widehat{r}^{i}_{t} \neq k. \end{cases} \tag{45}$$

According to Equations (40–44), the estimation accuracy for the hyperparameters associated with the GMM noise parameters depends on the estimated indicator r_t (i.e. the identification accuracy of the measurement mode), while an accurate estimation for the noise parameters facilitates the identification of the measurement mode according to Equations (30) and (31).

Remark 1: As for particle filters, resampling is conducted to rejuvenate the particles and reduce the effects of degeneracy. Resampling does not have to be performed at every run of the proposed approach. In this work, the resampling procedure is implemented only when the estimated effective sample size $\widehat{N}_{\mathrm{eff}} = \frac{1}{\sum_{i=1}^{N}\left(\omega^{i}_{t}\right)}$ is below the user-defined threshold N_{thres} (for more details, see Ref [16, 34, 44]).

Algorithm 1 Proposed marginalised particle filter for non-linear SSMs with Gaussian mixture noise

Inputs: $\left\{\omega^{i}_{t-1}, \boldsymbol{x}^{i}_{t-1}, \left\{\widehat{\alpha}^{i}_{k,t-1}, \widehat{\beta}^{i}_{k,t-1}, \widehat{\boldsymbol{m}}^{i}_{k,t-1}, \widehat{\nu}^{i}_{k,t-1}, \widehat{\boldsymbol{W}}^{i}_{k,t-1}\right\}_{k=1}^{K}\right\}_{i=1}^{N}$

Outputs:
$\left\{\omega^{i}_{t}, \boldsymbol{x}^{i}_{t}, \widehat{r}^{i}_{t}, \left\{\widehat{\alpha}^{i}_{k,t}, \widehat{\beta}^{i}_{k,t}, \widehat{\boldsymbol{m}}^{i}_{k,t}, \widehat{\nu}^{i}_{k,t}, \widehat{\boldsymbol{W}}^{i}_{k,t}\right\}_{k=1}^{K}\right\}_{i=1}^{N}$

for $i = 1, \ldots, N$ **do**
  Generate $\boldsymbol{x}^{i}_{t}$ and compute $\widehat{\boldsymbol{u}}^{i}_{t|t-1}$ according to Equations (15) and (21), respectively.
  Compute $\widehat{\boldsymbol{m}}^{i}_{k,t|t-1}$, $\widehat{\beta}^{i}_{k,t|t-1}$, $\widehat{\boldsymbol{W}}^{i}_{k,t|t-1}$ and $\widehat{\nu}^{i}_{k,t|t-1}$ according to Equation (25).
  Compute the weight ω^{i}_{t} for the ith particle $\boldsymbol{x}^{i}_{t}$ according to Equations (16) and (27).
  Compute $\widehat{\boldsymbol{u}}^{i}_{t}$ by using Equation (29), and then determine $\widehat{r}^{i}_{t}$ according to Equation (31).
  Compute $\left\{\widehat{\alpha}^{i}_{k,t}, \widehat{\beta}^{i}_{k,t}, \widehat{\boldsymbol{m}}^{i}_{k,t}, \widehat{\nu}^{i}_{k,t}, \widehat{\boldsymbol{W}}^{i}_{k,t}\right\}_{k=1}^{K}$ conditionally on the ith particle according to Equations (40–44).
end for
Normalise $\widetilde{\omega}^{i}_{t} = \omega^{i}_{t}/\sum_{i=1}^{N}\left(\omega^{i}_{t}\right)$ and evaluate $\widehat{N}_{\mathrm{eff}} = 1/\sum_{i=1}^{N}\left(\widetilde{\omega}^{i}_{t}\right)$.
if $\widehat{N}_{\mathrm{eff}} < N_{\mathrm{thres}}$ **then**
  Perform particle resampling.
end if
Recursion: $t = t + 1$.

Finally, the GMM noise parameters at the tth time instant can be obtained as follows:

$$\widehat{\pi}_{k,t} = \sum_{i=1}^{N} \omega^{i}_{t} \frac{\widehat{\alpha}^{i}_{k,t}}{\sum_{k=1}^{K}\widehat{\alpha}^{i}_{k,t}}, \tag{46}$$

$$\widehat{\boldsymbol{\mu}}_{k,t} = \sum_{i=1}^{N} \omega^{i}_{t}\widehat{\boldsymbol{m}}^{i}_{k,t}, \tag{47}$$

$$\widehat{\boldsymbol{\Lambda}}_{k,t} = \sum_{i=1}^{N} \omega^{i}_{t}\left(\left(\widehat{\nu}^{i}_{k,t}\widehat{\boldsymbol{W}}^{i}_{k,t}\right)^{-1} + \left(\widehat{\boldsymbol{m}}^{i}_{k,t} - \widehat{\boldsymbol{\mu}}_{k,t}\right)\left(\widehat{\boldsymbol{m}}^{i}_{k,t} - \widehat{\boldsymbol{\mu}}_{k,t}\right)^{\mathrm{T}}\right)^{-1}, \tag{48}$$

where $k = 1, \ldots, K$. The proposed MPF for non-linear SSMs with a GMM measurement noise is summarised in Algorithm 1.

4 | EXPERIMENTS

This section validates the proposed MPF for non-linear SSMs with a GMM measurement noise in the context of GNSS-based positioning in urban canyons. The pseudo-range (PR) measurement noise in urban canyons is considered as a non-Gaussian stochastic process due to the fact that the tracking loop within the GNSS receiver is impacted by interferences, for example, due to multipath (MP) signals that severely impair the GNSS-based positioning accuracy [36]. One important remaining challenge for the application of GNSS in urban environments is reduction of the impact of MP signals on positioning methods. Generally, reflected MP signals affecting the received direct GNSS signal can be divided into two kinds of signals: (a) MP interferences defined as the sum of the direct signal and the delayed reflections handled by the GNSS receiver; (b) non-line-of-sight (NLOS) signals resulting from a unique reflected signal received and tracked by the GNSS receiver [37]. In addition, an important

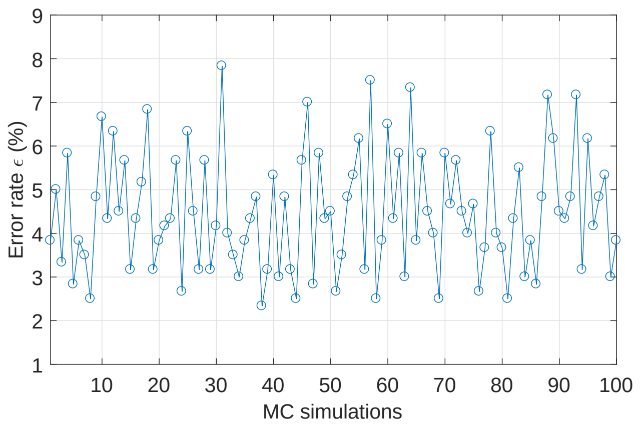

FIGURE 1 Error rate for the maximum a posteriori estimates of r_t over 100 MC simulations (proposed approach)

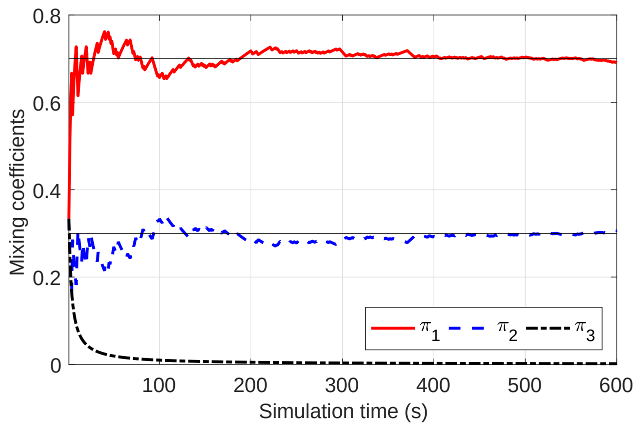

FIGURE 2 Means of the mixing coefficients estimates over 100 MC simulations (proposed approach)

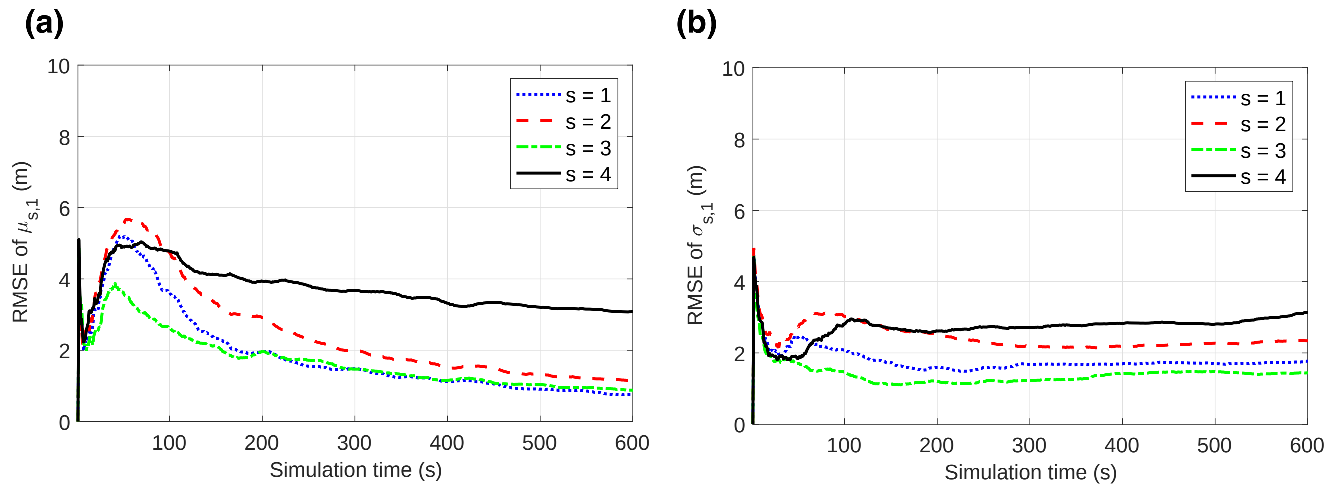

FIGURE 3 Root mean square errors of $\mu_{s,1}$ and $\sigma_{s,1}$ for $s \in \{1, 2, 3, 4\}$

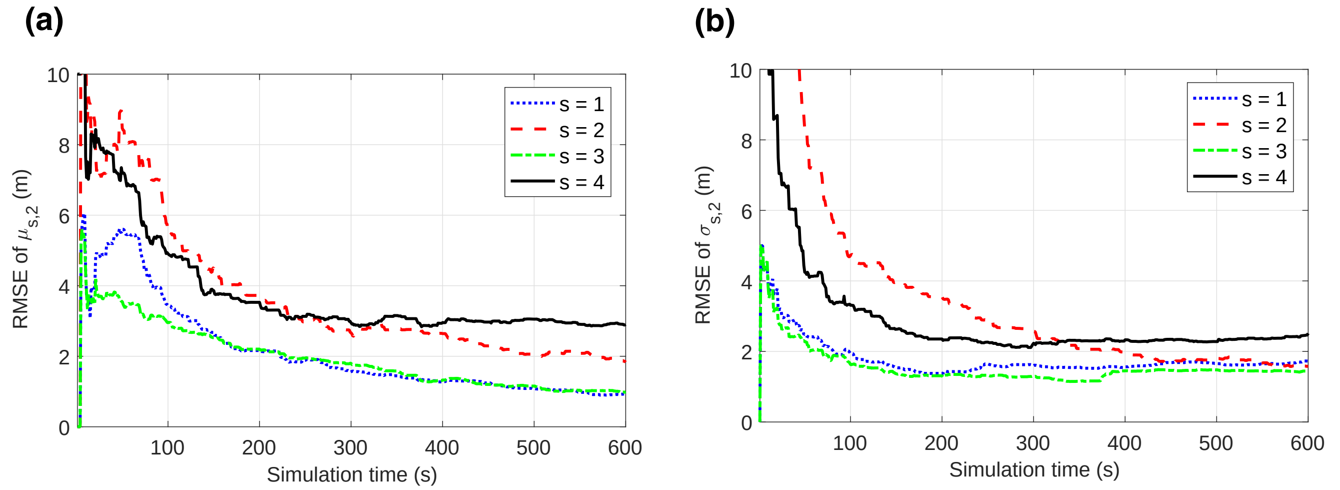

FIGURE 4 Root mean square errors of $\mu_{s,2}$ and $\sigma_{s,2}$ for $s \in \{1, 2, 3, 4\}$

property of MP signals is that they not only depend on the relative position of the receiver (which is generally moving) and GNSS satellites but also on the environment where the receiver is located, especially in urban canyons [38]. Accordingly, these two reception situations can occur both separately or jointly inside the tracking loop of the GNSS

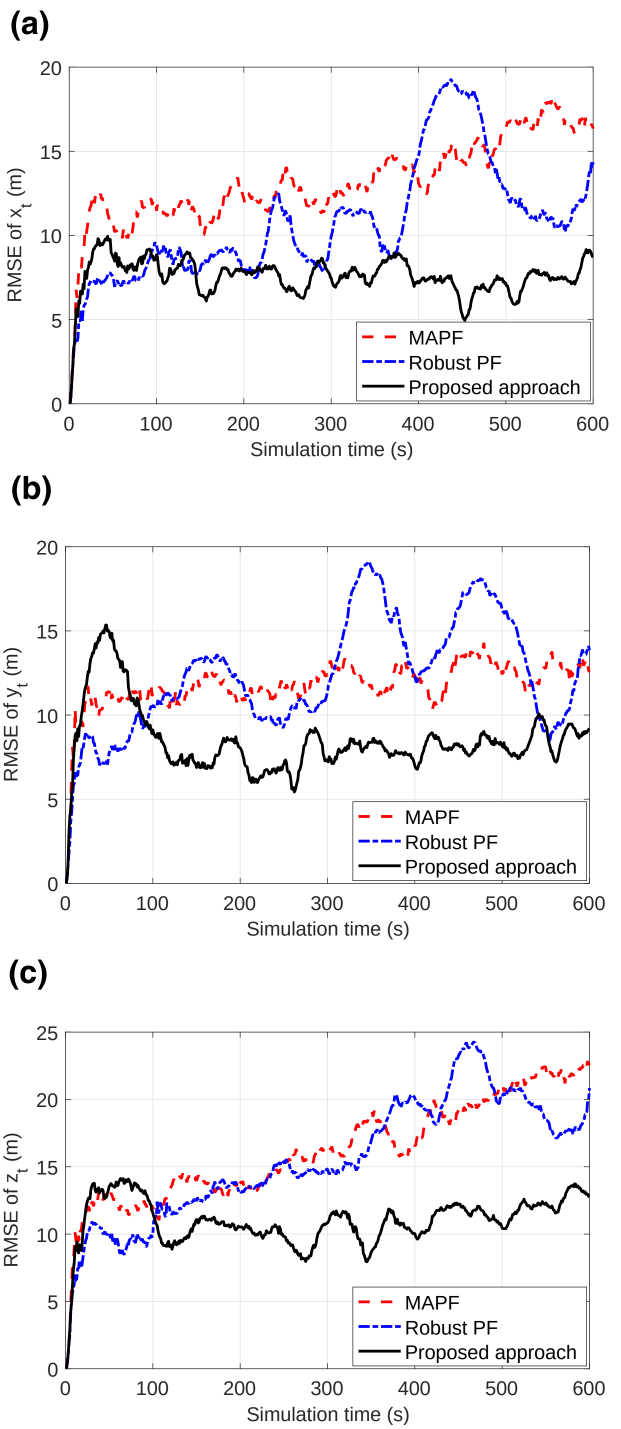

FIGURE 5 Root mean square errors of position estimates with the different approaches

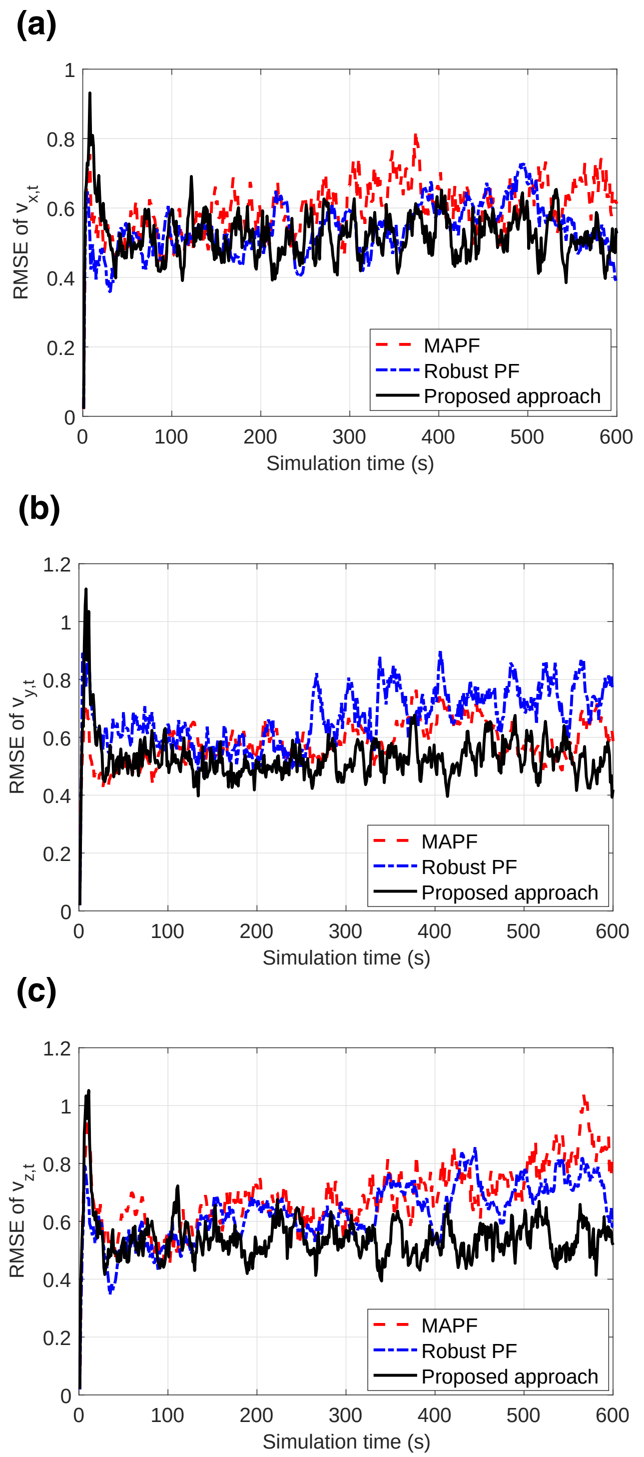

FIGURE 6 Root mean square errors of velocity estimates with the different approaches

receiver, leading to a PR measurement model that switches between different modes [39].

4.1 | Simulation scenario

Since the GNSS positioning solution depends on the dynamic level of the vehicle, a second-order model (i.e. a constant velocity model) is used to describe the dynamics of the vehicle in the earth-centred earth-fixed (ECEF) frame. Moreover, the GNSS receiver clock offset and its drift are taken into account. Therefore, the state vector considered in this simulation is defined as follows [40]:

$$\boldsymbol{x}_t = \left(x_t, v_{x,t}, y_t, v_{y,t}, z_t, v_{z,t}, b_t, d_t\right)^{\mathrm{T}}, \tag{49}$$

where $t = 1, \ldots, \infty$ is the tth sampling time instant, $\boldsymbol{p}_t = (x_t, y_t, z_t)^{\mathrm{T}}$ and $\boldsymbol{v}_t = (v_{x,t}, v_{y,t}, v_{z,t})^{\mathrm{T}}$ are the vehicle position and velocity in the ECEF frame (Cartesian coordinates), and b_t and d_t are the GNSS receiver clock offset and drift. The velocity can be reasonably modelled as a random walk, for example, $\ddot{x} = e_x$, where e_x is a zero-mean Gaussian noise of variance σ_a^2. For short-term applications in which the periodical clock resets of the GNSS receiver are not taken into account, the GNSS receiver clock offset b_t and its drift d_t can also be modelled as random walks, that is, $\dot{b}_t = d_t + e_b$ and $\dot{d}_t = e_d$, where e_b and e_d are zero-mean Gaussian white noises of variances σ_b^2 and σ_d^2. Based on the above assumptions, the discrete-time state model that describes the propagation of the vehicle state $\boldsymbol{x}_t$ can be formulated as follows:

$$\boldsymbol{x}_t = \boldsymbol{F}_{t|t-1}\boldsymbol{x}_{t-1} + \boldsymbol{e}_t, \tag{50}$$

where $\boldsymbol{e}_t = \left(e_x,\, e_y,\, e_z,\, e_b,\, e_d\right)^{\mathrm{T}}$ is the zero-mean Gaussian white noise vector of covariance matrix $\boldsymbol{Q}$, that is, $\boldsymbol{e}_t \sim \mathcal{N}(0, \boldsymbol{Q})$. Expressions of the matrices $\boldsymbol{F}_{t|t-1}$ and $\boldsymbol{Q}$ can be found in Ref [41].

Considering that the atmospheric propagation errors can be compensated within the GNSS receivers, the PR measurement model of the sth in-view satellite at the tth time is defined as

$$y_{s,t} = \|\boldsymbol{p}_{s,t} - \boldsymbol{p}_t\| + b_t + v_{s,t}, \tag{51}$$

where $y_{s,t}$ ($s = 1, \ldots, N_{\mathrm{s}}$, N_{s} being the number of in-view satellites) is the PR measurement associated with the sth in-view satellite, $\|\cdot\|$ is the Euclidean norm, $\boldsymbol{p}_{s,t} = \left(x_{s,t}, y_{s,t}, z_{s,t}\right)^{\mathrm{T}}$ is the sth satellite position in the ECEF frame, and v_s denotes the measurement noise of the sth in-view satellite. In the absence of MP signals, the measurement noise has a Gaussian distribution with zero mean and variance $\sigma_{s,1}^2$ (referred to as a nominal value). For different reception situations, in the presence of MP signals, two possible models are considered according to the availability of the direct signal, that is, variance jumps affecting the zero-mean Gaussian white noise in the context of MP interferences and mean value jumps for NLOS signals [42, 43]. Thus, the measurement noise in the presence of MP signals has a Gaussian distribution with non-zero mean $\mu_{s,2}$ and variance $\sigma_{s,2}^2$. Accordingly, the PR measurement noise $v_{s,t}$ can be modelled as a GMM with two components, that is, $v_{s,t} \sim \pi_1\mathcal{N}\left(0, \sigma_{s,1}^2\right) + \pi_2\mathcal{N}\left(\mu_{s,2}, \sigma_{s,2}^2\right)$.

The state space model defined in Equation (50) has been simulated with the following parameters: process noise standard deviation $\sigma_a = 0.1\ m/s^2$, clock offset deviation $\sigma_b = 3c \times 10^{-10}\ m$ and drift standard deviation $\sigma_d = 2\pi c \times 10^{-10}\ m/s$, where $c = 3 \times 10^8\ m/s$ denotes the velocity of light. The simulation is conducted using $N_{\mathrm{s}} = 4$ in-view satellites. The noise-free GNSS PR measurements have been computed based on an almanac file including all useful satellite orbit data in the simulations. The standard deviation of the measurement noise in the absence of MP signals is set to $\sigma_{s,1} = 10\ m$, where $s \in \{1, 2, 3, 4\}$. In this simulation, the second satellite PR measurement ($s = 2$) is affected by MP interferences, resulting in $\mu_{2,2} = 0$ and $\sigma_{2,2} = 30\ m$, while the fourth satellite PR measurement ($s = 4$) is simultaneously affected by the mean value and variance changes, resulting in $\mu_{4,2} = 10\ m$ and $\sigma_{4,2} = 20\ m$. Accordingly, the noise GMM associated with all in-view satellite measurements is defined as $\pi_1\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + \pi_2\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ where $\boldsymbol{\mu}_1 = \mathrm{diag}[0, 0, 0, 0]$, $\boldsymbol{\Sigma}_1 = \mathrm{diag}\left[10^2, 10^2, 10^2, 10^2\right]$ and $\boldsymbol{\mu}_2 = \mathrm{diag}[0, 0, 0, 10]$, $\boldsymbol{\Sigma}_2 = \mathrm{diag}\left[10^2,\ 30^2, 10^2, 20^2\right]$, where $\mathrm{diag}\left[\rho_1, \ldots, \rho_{N_{\mathrm{s}}}\right]$ denotes a diagonal matrix with diagonal elements $\rho_1, \ldots, \rho_{N_{\mathrm{s}}}$. In addition, the mixture coefficients are set to $\pi_1 = 0.7$ and $\pi_2 = 0.3$, respectively.

4.2 | Validation of the proposed algorithm

In order to validate the capability of the proposed approach for adaptively determining the number of components, the number of GMM components is set to $K = 3$. The transition probability matrix Π used for computing the predicted mode probabilities is defined as [39]

$$\Pi = \begin{pmatrix} 0.98 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}.$$

Accordingly, the initial values for each Gaussian component are set to $\pi_{k,0} = 1/K$, $\alpha_{k,0} = 1$, $\boldsymbol{m}_{k,0} = (0, 0, 0)^{\mathrm{T}}$, $\beta_{k,0} = 1$, $\boldsymbol{W}_{k,0} = \mathrm{diag}\left[0.1^2, 0.1^2, 0.1^2\right]$ and $\nu_{k,0} = N_{\mathrm{s}}$ with $k \in \{1, 2, 3\}$. The simulation time and the number of particles are set to $T = 600s$ and $N = 2000$.[1] The different algorithms are run for $N_{\mathrm{m}} = 100$ Monte Carlo (MC) runs. The MAP estimates of r_t and the root mean square errors (RMSEs) of the states for the different MC runs are defined as follows:

$$\widehat{r}_t = \underset{r_t=\{1,\ldots,K\}}{\mathrm{argmax}} \frac{1}{N_{\mathrm{m}}} \sum_{m=1}^{N_{\mathrm{m}}} \widehat{u}_{k,t}(m), \tag{52}$$

and

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_{\mathrm{m}}} \sum_{m=1}^{N_{\mathrm{m}}} \left(\widehat{\boldsymbol{x}}_t(m) - \boldsymbol{x}_t\right)^2}, \tag{53}$$

where $\widehat{u}_{k,t}(m)$ and $\widehat{\boldsymbol{x}}_t(m)$ are the mth estimates ($m = 1, \ldots, N_{\mathrm{m}}$) of the measurement mode probability and the state at the tth time instant. All algorithms have been coded

[1] The number of particles has been chosen to meet a compromise between the estimation accuracy of the proposed approach and its computational load. A larger number of particles might be considered for other applications.

TABLE 1 Execution times for the three approaches with different numbers of particles

	Execution times (s)		
N	MAPF	Robust PF	Proposed method
2000	89.42	169.50	128.24
4000	189.13	274.63	223.42
6000	273.17	545.31	407.21

using MATLAB and run on a laptop with Intel i-5 and 8 GB RAM.

Figure 1 displays the error rates ϵ^2 for the MAP estimates of r_t for the 100 MC simulations. The corresponding averaged value of the error rates is 4.47%, demonstrating that the measurement mode can be well-identified by using the proposed MPF. In addition, the means of the mixing coefficients computed using 100 MC simulations are depicted in Figure 2. It is clear that the means of $\widehat{\pi}_{1,t}(m)$ and $\widehat{\pi}_{2,t}(m)$ gradually converge to the true values (i.e. 0.7 and 0.3), whereas the mean of $\widehat{\pi}_{3,t}(m)$ approaches zero when time increases ($m = 1, \ldots, N_{\mathrm{m}}$). These results show that the number of mixture components and the values of the mixing coefficients can be adaptively determined by the proposed approach.

The RMSEs of the mean and standard deviation of each Gaussian component (i.e. $\widehat{\mu}_{s,1}$, $\widehat{\sigma}_{s,1}$, and $\widehat{\mu}_{s,2}$, $\widehat{\sigma}_{s,2}$ for $s = \{1, 2, 3, 4\}$) obtained with the proposed approach are depicted in Figures 3 and 4. The RMSEs associated with the first ($s = 1$) and third ($s = 3$) PR measurements are obviously smaller than those associated with the second ($s = 2$) and fourth ($s = 4$) PR measurements. This is due to the fact that the first and third PR measurements were not affected by MP signals, contrary to the other two PR measurements. Moreover, the estimation results obtained for the second and fourth PR measurements show that the false identification of the measurement mode impairs the estimation accuracy of the GMM parameters, especially for the fourth PR measurement, where the noise mean and variance change simultaneously, due to the presence of MP.

4.3 | Comparison with the state-of-the art approaches

The position and velocity estimation accuracies obtained with the proposed method are compared to those obtained using the marginalised adaptive particle filter (MAPF) of Ref [24] and the robust particle filter (PF) of Ref [25]. Figures 5 and 6 display the RMSEs of position and velocity estimations obtained with the different methods. As shown in Figure 5, since the estimates for the parameters of the noise GMM are inaccurate at the beginning of the simulation, the corresponding position estimation accuracy is reduced for the proposed approach when compared to those obtained with the MAPF and the robust PF. However, it is clear that the proposed approach provides a better positioning estimation accuracy when time increases. Figure 6 shows that the velocity estimation obtained with the proposed method is very competitive with respect to the state-of-the-art approaches. Overall, these results show the interest of using the proposed approach for mitigating the impact of MP signals on GNSS-based positioning in urban canyons.

The computational complexities of the MAPF, the robust PR and the proposed approach are $\mathcal{O}(NT)$, $\mathcal{O}(\kappa_{\max}NT)$,[3] and $\mathcal{O}(KNT)$. Table 1 shows the execution time of a single MC run for the three approaches with different numbers of particles. One can conclude that the increased estimation accuracy is obtained at the price of a moderate computational cost, which will be acceptable in many practical applications.

5 | CONCLUSION

This work studied a new marginalised particle filter with variational inference for non-linear SSMs with a Gaussian mixture model (GMM) for the observation noise. A latent variable was introduced for indicating the measurement mode of the SSM, corresponding to a specific component of the GMM. The joint posterior distribution of the state, the latent variable and the parameters of the GMM was then derived and sampled using a new marginalised particle filter combined with variational Bayesian inference. The samples generated sequentially by this particle filter were finally used to estimate the unknown parameters of the SSM. A simulation study was conducted for a realistic GNSS-based positioning problem in urban canyons. The proposed approach was compared to the state-of-the-art approaches, more precisely to the marginalised adaptive particle filter and the robust particle filter, providing accurate estimates of the state and noise parameters at the price of a moderate computational cost. In particular, the number of mixture components and the GMM parameters (including the mixing coefficients and the parameters of each Gaussian distribution) were adaptively estimated by the proposed approach in a recursive way. Our future work will be devoted to implementing the proposed approach for SSMs with other mixture noise model, such as the Gaussian-uniform mixture noise [35]. In addition, testing of the proposed algorithm in the context of radar target tracking is also an interesting prospect.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grants 61901380.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[2] $\epsilon = \frac{N_{\mathrm{F}}}{N_{\mathrm{T}}} \times 100\%$, where N_F and N_{T} denote the number of false MAP estimates and the total number of samples for r_t in the time interval T.

[3] $\kappa_{\max}$ denotes the maximum number of iterations for the VB inference step of the robust PF.

ORCID

Cheng Cheng https://orcid.org/0000-0002-3726-657X

REFERENCES

1. Misra, P., Enge, P.: Global Positioning System: Signals, Measurements and Performance, ch. PVT Estimation, pp. 200–231. Ganga-Jamuna Press, Lincoln, Massachusetts (2001)
2. Bar-Shalom, Y., Li, X.R., Kirubarajan, T.: Estimation with Appplication to Tracking and Navigation: Theory, Algorithm and Software, ch. State Estimation for Nonlinear Dynamic System, pp. 371–394. John Weily & Sons, New York (2004)
3. Wan, E.A., van de Merwe, R.: Unscented Kalman filter. In: Haykin, S. (ed.) Kalman filter and neural networks, ch. 7, pp. 221–268. John Weily & Sons, New York (2001)
4. Steingass, A., Krach, B., Crisci, M.: Robustness versus accuracy: multipath effects on land mobile satellite navigation. IET Radar Sonar Navig. 11(3), 520–527 (2017)
5. Li, H.W., Wang, J.: Particle filter for manoeuvring target tracking via passive radar measurements with glint noise. IET Radar Sonar Navig. 6(3), 180–189 (2012)
6. Plataniotis, K.N., Hatzinakos, D.: Gaussian mixtures and their applications to signal processing. In: Stergiopoulos, S. (ed.) Advanced Signal Processing Handbook, ch. 3. CRC Press, Boca Raton, Florida (2001)
7. Bilik, I., Tabrikian, L.: Maneuvering target tracking in the presence of glint using the nonlinear Gaussian mixture kalman filter. IEEE Trans. Aero. Electron. Syst. 46(1), 246–262 (2010)
8. Gokce, M., Kuzuoglu, M.: Unscented Kalman filter-aided Gaussian sum filter. IET Radar Sonar Navig. 9(5), 589–599 (2015)
9. Mazor, E., et al.: Interacting multiple model methods in target tracking: A survey. IEEE Trans. Aero. Electron. Syst. 34(1), 103–123 (1998)
10. Pishdad, L., Labeau, F.: Approximate MMSE estimator for linear dynamic systems with Gaussian mixture noise. IEEE Trans. Automat. Contr. 62(5), 2457–2463 (2017)
11. Liu, K.L., et al.: Distributed track-to-track fusion for non-linear systems with Gaussian mixture noise. IET Radar Sonar Navig. 13(5), 740–749 (2019)
12. Huang, D.L., Leung, H.: Maximum likelihood state estimation of semi-Markovian switch system in non-Gaussian measurement noise. IEEE Trans. Aero. Electron. Syst. 46(1), 133–146 (2010)
13. Zhu, H., Leung, H., He, Z.S.: State estimation in unknown non-Gaussian measurement noise using variational Bayesian technique. IEEE Trans. Aero. Electron. Syst. 49(4), 2601–2614 (2013)
14. Rabaoui, A., et al.: Dirichlet process mixture for density estimation in dynamic nonlinear modeling: Application to GPS positioning in urban canyons. IEEE Trans. Signal Process. 60(4), 1638–1655 (2012)
15. Schön, T., Gustafsson, F., Nordlund, P.-J.: Marginalized particle filters for mixed linear/nonlinear state-space models. IEEE Trans. Signal Process. 53(7), 2279–2289 (2005)
16. Cappé, O., Godsill, S.J., Moulines, E.: An overview of existing methods and recent advances in sequential Monte Carlo. Proc. of the IEEE. 95(5), 899–924 (2007)
17. Gustafsson, F., et al.: Particle filters for positioning, navigation, and tracking. IEEE Trans. Signal Process. 50(2), 425–437 (2002)
18. Zhou, H.Y., et al.: Rao-Blackwellised particle filtering for low-cost encoder/INS/GNSS integrated vehicle navigation with wheel slipping. IET Radar Sonar Navig. 13(11), 1890–1898 (2019)
19. Särkkä, S., Vehtari, A., Lampinen, J.: Rao-blackwellized particle filter for multiple target tracking. Inform. Fusion. 8(1), 2–15 (2007)
20. Kokkala, J., Särkkä, S.: Combining particle MCMC with Rao-Blackwellized Monte Carlo data association for parameter estimation in multiple target tracking. Digit. Signal Process. 47, 86–95 (2015)
21. Mihaylova, L., et al.: Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking. Digit. Signal Process. 25, 1–16 (2014)
22. Su, H.T., et al.: Rao-Blackwellised particle filter based track-before-detect algorithm. IET Radar Sonar Navig. 2(2), 169–176 (2008)
23. Rollason, M., Salmond, D.: Particle filter for track-before-detect of a target with unknown amplitude viewed against a structured scene. IET Radar Sonar Navig. 12(6), 1588–1596 (2018)
24. Özkan, E., et al.: Marginalized adaptive particle filtering for nonlinear models with unknown time-varying noise parameters. Automatica. 49(6), 1566–1575 (2013)
25. Xu, D.J., Shen, C., Shen, F.: A robust particle filtering algorithm with non-Gaussian measurement noise using Student-t distribution. IEEE Signal Process Lett. 21(1), 30–34 (2014)
26. Cheng, C., Tourneret, J.-Y., Lu, X.D.: A Rao-blackwellized particle filter with variational inference for state estimation with measurement model uncertainties. IEEE Access. 8, 55665–55675 (2020)
27. Bishop, C.M.: Pattern Recognition and Machine Learning, ch. Approximation Inference, pp. 474–485. Springer Publishing, New York (2006)
28. Murphy, K.P.: Machine Learning: A Probabilistic Perspective, ch. Mixture models and the EM algorithm, pp. 337–344. The MIT Press, Cambrigde (2012)
29. Arulampalam, M.S., et al.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process. 50(2), 174–188 (2002)
30. Özkan, E., et al.: Recursive maximum likelihood identification of jump Markov nonlinear systems. IEEE Trans. Signal. Process. 63(3), 754–765 (2015)
31. Murphy, K.P.: Conjugate Bayesian Analysis of the Gaussian Distribution. Univ. British ColumbiaTech. Rep., Vancouver, BC, USA (2007)
32. Tzikas, D.G., Likas, A.C., Galatsanos, N.P.: The variational approximation for Bayesian inference. IEEE Signal Process. Mag. 25(6), 131–146 (2008)
33. Fox, C.W., Roberts, S.J.: A tutorial on variational Bayesian inference. Artif. Intell. Rev. 38, 85–95 (2012)
34. Doucet, A., Godsill, S., Andrieu, C.: On sequential Monte Carlo sampling methods for Bayesian filtering. Stat. Comput. 10, 197–208 (2000)
35. Li, T.C., Bolic, M., Djuric, P.M.: Resampling Methods for Particle Filtering: Classification, implementation, and strategies. IEEE Signal Process. Mag. 32(3), 70–86 (2015)
36. Ward, P.W., Betz, J.W., Hegarty, J.: Interference, multipath, and scintillation. In: Kaplan, E., Hegarty, C. (eds.) Understanding GPS: Principles and Application, ch. 6, pp. 279–295. Artech House, Norwood, Massachusetts (2006)
37. Groves, P.D.: Principles of GNSS, Inertial and Multisensor Integrated Navigation Systems, 2nd ed, ch. GNSS: Advanced Techniques, pp. 458–465. Artech House, London (2013)
38. Cheng, C., Tourneret, J.-Y.: An EM-based multipath interference mitigation in GNSS receivers. Signal Process. 162, 142–152 (2019)
39. Marais, J., et al.: GNSS accuracy enhancement based on pseudo range error estimation in an urban propagation environment. Expert Syst. Appl. 40, 5956–5964 (2013)
40. Giremus, A., Tourneret, J., Doucet, A.: A fixed-lag particle filter for the joint detection/compensation of interference effects in GPS navigation. IEEE Trans. Signal. Process. 58(12), 6066–6079 (2010)
41. Cheng, C., et al.: Detecting, estimating and correcting multipath biases affecting GNSS signals using a marginalized likelihood ratio-based method. Signal Process. 118, 221–234 (2016)
42. Spangenberg, M., et al.: Detection of variance changes and mean value jumps in measurement noise for multipath mitigation in urban navigation. J Navig. 57(1), 35–52 (2010)
43. Qin, H.L., Xue, X., Yang, Q.: GNSS multipath estimation and mitigation based on particle filter. IET Radar Sonar Navig. 13(9), 1588–1596 (2019)
44. Brunot, M.: A Gaussian uniform mixture model for robust Kalman filtering. IEEE Trans. Aero. Electron. Syst. 56(4), 2656–2665 (2020)

How to cite this article: Cheng, C., Tourneret, J.-Y., Lu, X.: A marginalised particle filter with variational inference for non-linear state-space models with Gaussian mixture noise. IET Radar Sonar Navig. 1–11 (2021). https://doi.org/10.1049/rsn2.12179