IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 3, 2020, accepted March 16, 2020, date of publication March 19, 2020, date of current version March 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981948

A Rao-Blackwellized Particle Filter With Variational Inference for State Estimation With Measurement Model Uncertainties

CHENG CHENG[1], JEAN-YVES TOURNERET[2], (Fellow, IEEE), AND XIAODONG LU[1]
[1]School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China
[2]ENSEEIHT-IRIT-TéSA, University of Toulouse, 31071 Toulouse, France

Corresponding author: Cheng Cheng (cheng.cheng@nwpu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61901380 and Grant 61761136001.

ABSTRACT This paper develops a Rao-Blackwellized particle filter with variational inference for jointly estimating state and time-varying parameters in non-linear state-space models (SSM) with non-Gaussian measurement noise. Depending on the availability of the conjugate prior for the unknown parameters, the joint posterior distribution of the state and unknown parameters is approximated by using an auxiliary particle filter with a probabilistic changepoint model. The distribution of the SSM parameters conditionally on each particle is then updated by using variational Bayesian inference. Experiments are first conducted on a modified nonlinear benchmark model to compare the performance of the proposed approach with other state-of-the-art approaches. Finally, in the context of GNSS multipath mitigation, the proposed approach is evaluated based on data obtained from a measurement campaign conducted in a street urban canyon.

INDEX TERMS Joint state and parameter estimation, Rao-blackwellized particle filter, state-space models, variational inference.

I. INTRODUCTION

State-space models (SSMs), composed of dynamic and measurement equations, are applied to a wide variety of signal processing problems, especially in positioning, tracking and navigation [1], [2]. A central problem when using these models is to recursively infer the state based on a sequence of measurements. In general, sensors delivering measurements are assumed to be in their nominal state of work, i.e., the parameters in the measurement equation have to be exactly specified a priori. However, in realistic contexts, these parameters can be time-varying due to abruptly changing environment, leading to the problem of state estimation in the presence of model uncertainty.

Early attempts to solve this problem are based on Gaussian mixture approximations, such as the interacting multiple model (IMM) algorithm [3]. Since the reliability of the IMM is dependent on the number and choice of models, some approaches for jointly estimating state and parameter for SSMs have been proposed. Another solution for joint state and parameter estimation is to assign a prior distribution to the model parameters and augment the state vector by including these unknown parameters. When it cannot be computed in closed form, the joint posterior distribution of all variables can be approximated by using sequential Monte Carlo (SMC) techniques [4]–[6]. Another possibility is to exploit the expectation-maximization (EM) algorithm, i.e., the posterior distribution of the state is obtained in the expectation step and then the maximum likelihood estimator of the unknown parameters is updated in the maximization step [7]–[10]. Recently, variational inference-based approaches in SSMs have been extensively studied for state estimation in the presence of model uncertainty [11]. In the case of linear SSMs, approximate separable distributions for state and noise parameters can be updated by iteratively solving the coupled equations by using variational Bayesian (VB) inference [12]–[14]. A VB-based Kalman filter was also proposed for linear SSMs with non-stationary heavy-tailed noises [15], [16]. In the case of nonlinear SSMs, the noise parameters are usually marginalized out and the posterior distribution of the state is approximated by using an appropriate particle filter (PF). Then the distribution of the noise

The associate editor coordinating the review of this manuscript and approving it for publication was Yingsong Li.

parameters conditionnally on each state particle can be calculated by using VB inference [17], [18].

Considering that state and measurement equations depend linearly on a subset of the state vector, a Rao-Blackwellized particle filter (RBPF) (also known as marginalized particle filter) was proposed in [19], [20]. The filter was 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 the Rao-Blackwellization is to marginalize the distribution of interested with respect to the state variables appearing linearly in the state and measurement equations, allowing these linear components to be processed using analytical methods (such as the Kalman filter) and the non-linear components by SMC techniques [21]. Rao-Blackwelization has been used successfully in many applications including navigation using an inertial navigation system and a terrain-aided positioning [22], multiple target tracking [23] and simultaneous localization and mapping (SLAM) [24]. Some theoretical works derived the asymptotic variance of the Rao-Blackwell particle filter in order to determine in which cases it is interesting to use an RBPF in place of a standard PF with an increased number of particles [25]. When there is no linear sub-structure in the state vector, which is the case in this work, the standard Rao-Blackwellization cannot be applied directly.

Several approaches can be found in the literature for jointly estimating state and static parameter in the frame of the RBPF. The common way is to replace analytical methods with different types of SMC samplers for implementing the parameter estimation, such as the particle Markov chain Monte Carlo [26], the twisted particle filter [27] and the nested particle filter [28], [29]. However, these approaches cannot be easily applied to cases where the parameters appearing in the measurement equation are time-varying. To address this problem, an RBPF based on a Dirichlet process mixture (DPM) was studied in [30]. In addition, the adaptive parameter filter was proposed in [31] to implement a probabilistic changepoint model in the frame of the RBPF for estimating jointly the state vector and the time-varying unknown parameters.

This paper studies an RBPF with variational inference for jointly estimating state and time-varying parameters in nonlinear state-space models with measurement noise distributed according to a Student-t distribution. The interest of using this Student-t distribution is to obtain a robust estimator of the state vector. The idea of the resulting RBPF filter is to marginalize the joint distribution of interest with respect to the parameters of the Student-t distribution, to sample the remaining state and unknown parameters by using an auxiliary particle filter (APF) with a probabilistic changepoint model, and to sample the Student-t parameters by using VB inference. The proposed approach combines the advantages of the adaptive parameter estimation (APE) filter introduced in [31] and the robust PF studied in [18] to build a joint state and time-varying parameter estimator in the presence of non-Gaussian measurement noise.

The paper is organized as follows: The problem of jointly estimating state and time-varying parameters in non-linear SSMs with non-Gaussian measurement noise is presented in Section II. Section III studies the proposed new RBPF based on VB inference. The performance of this filter is evaluated in Section IV, first using simulated data based on a modified nonlinear benchmark model, and then using experimental data in the context of GNSS multipath mitigation in urban canyons. Conclusions are finally reported in Section V.

II. PROBLEM FORMULATION

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

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{\omega}_k \tag{1}$$

$$\boldsymbol{y}_k = h_{\boldsymbol{\theta}_k}(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{2}$$

where $k = 1, \ldots, K$ denotes the kth time instant, $f(\cdot)$ is the state transition function, $\boldsymbol{\omega}_k$ denotes the process noise with a zero mean Gaussian distribution of covariance matrix $\boldsymbol{Q}_k$, $h_{\boldsymbol{\theta}_k}(\cdot)$ is the measurement function depending on a vector of parameters $\boldsymbol{\theta}_k \in \mathbb{R}^{n_\theta}$ and $\boldsymbol{v}_k$ denotes the measurement noise. In practice, the unknown parameter vector $\boldsymbol{\theta}_k$ can be time-varying due to abruptly changing environments leading to different sensor functioning conditions. In addition, the measurement noise $\boldsymbol{v}_k$ can be non-Gaussian due to abruptly changing environments leading to outliers corrupting the measurements. Since the Student-t distribution is heavy-tailed compared to the Gaussian distribution and is robust to outliers, this work assumes that the measurement noise has a Student-t distribution. This distribution is obtained by integrating an infinite mixture of Gaussians with the same mean and scaled precision matrices [32], i.e.,

$$\mathrm{St}(\boldsymbol{v}_k|\boldsymbol{\Lambda}_k, \nu_k) = \int_0^\infty \mathcal{N}\left(\boldsymbol{v}_k|0, (\kappa_k \boldsymbol{\Lambda}_k)^{-1}\right) \times \gamma(\kappa_k|\nu_k/2, \nu_k/2)\, \mathrm{d}\kappa_k \tag{3}$$

where $\mathrm{St}(\cdot)$ denotes the student-t distribution and $\mathcal{N}(\cdot)$ is the multivariate Gaussian probability density function (pdf) (here with zero mean and precision matrix $\kappa_k \boldsymbol{\Lambda}_k$), κ_k is an auxiliary precision scalar, $\gamma(\cdot)$ is the gamma distribution with shape and inverse scale parameters both equal to $\frac{\nu_k}{2}$ and ν_k adjusts the thickness of the distribution tail. Independent conjugate gamma distributions are assigned a priori to $\boldsymbol{\Lambda}_k$, κ_k and ν_k, i.e., the joint prior of $(\boldsymbol{\Lambda}_k, \kappa_k, \nu_k)$ is defined as

$$q(\boldsymbol{\Lambda}_k, \kappa_k, \nu_k) = \prod_{s=1}^{N_\mathrm{d}} \gamma\left(\lambda_{s,k}|\alpha_{s,k}, \beta_{s,k}\right) \times \gamma\left(\kappa_k|\nu_k^1, \nu_k^2\right)\gamma(\nu_k|a_k, b_k) \tag{4}$$

where $\boldsymbol{\Lambda}_k = \mathrm{diag}\left[\lambda_{1,k}, \cdots, \lambda_{N_\mathrm{d},k}\right]$ (a diagonal matrix with diagonal elements $\lambda_{1,k}, \cdots, \lambda_{N_\mathrm{d},k}$), N_d is the dimension of $\boldsymbol{\Lambda}_k$, $\alpha_k, \beta_k, \nu_k^1, \nu_k^2, a_k, b_k$ are the hyperparameters of the prior distributions at the kth time instant.

Considering that both the parameter vector $\boldsymbol{\theta}_k$ and the noise statistical parameters $\boldsymbol{\xi}_k = \{\boldsymbol{\Lambda}_k, \kappa_k, \nu_k\}$ are time-varying in abruptly changing environments, the aim of this paper is to infer the states $\boldsymbol{x}_k$ and the time-varying parameters $\{\boldsymbol{\theta}_k, \boldsymbol{\xi}_k\}$ given measurements $\boldsymbol{y}_{1:k} = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_k\}$.

III. A RAO-BLACKWELLIZED PARTICLE FILTER WITH VARIATIONAL INFERENCE

Since the joint posterior distribution of all unknown variables $p(\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{\xi}_k|\boldsymbol{y}_{1:k})$ has a complex expression and cannot be calculated in closed form, we propose to study a PF approximating this posterior by using sequential importance sampling. The joint posterior distribution of the state and unknown parameters can be factorized as follows

$$p(\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{\xi}_k|\boldsymbol{y}_{1:k}) = p(\boldsymbol{\xi}_k|\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{y}_{1:k})\, p(\boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k}) \quad (5)$$

where the noise statistical parameters $\boldsymbol{\xi}_k$ have been marginalized out in the second term of the right hand side. We propose to approximate $p(\boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k})$ by using an empirical density following the principle of PFs

$$p(\boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i \delta\left((\boldsymbol{x}_k, \boldsymbol{\theta}_k) - \left(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i\right)\right) \quad (6)$$

where N is the number of particles, $\delta(\cdot)$ is the Dirac delta function, $(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i)$ is the ith particle and ω_k^i is the corresponding weight at the kth time instant, which can be updated as follows [33]

$$\omega_k^i \propto p(\boldsymbol{y}_k, \boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1})\, \omega_{k-1}^i. \quad (7)$$

The conditional density $p(\boldsymbol{y}_k, \boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1})$ can be obtained by the following marginalization

$$p(\boldsymbol{y}_k, \boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1}) = \int p(\boldsymbol{y}_k, \boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{\xi}_k, \boldsymbol{y}_{1:k-1}) \times p(\boldsymbol{\xi}_k|\boldsymbol{y}_{1:k-1})\, \mathrm{d}\boldsymbol{\xi}_k \quad (8)$$

where $p(\boldsymbol{\xi}_k|\boldsymbol{y}_{1:k-1})$ is the predictive distribution of the noise statistical parameters. Replacing (6) in (5) leads to

$$p(\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{\xi}_k|\boldsymbol{y}_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i p\left(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k}\right) \times \delta\left((\boldsymbol{x}_k, \boldsymbol{\theta}_k) - \left(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i\right)\right) \quad (9)$$

where the conditional distribution $p(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k})$ is the distribution of the noise parameter vector conditionally on the ith particle $(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i)$. In this paper, we propose to approximate the distribution $p(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k})$ using VB inference. As a consequence, samples $(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i)$ are generated using an APF, and then the posterior distribution of $\boldsymbol{\xi}_k$ conditionally on $(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i)$ is calculated according to VB inference and replaced in (9), allowing the joint distribution of the state, parameters and hyperparameters to be determined.

A. UPDATING (x_k, θ_k) SAMPLES BASED ON THE AUXILIARY PARTICLE FILTER

Assuming that the propagation models for $\boldsymbol{x}_k$ and $\boldsymbol{\theta}_k$ are independent, the posterior distribution $p(\boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k})$ can be recursively updated according to Bayes' rule

$$p(\boldsymbol{x}_k, \boldsymbol{\theta}_k|\boldsymbol{y}_{1:k}) \propto p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{\theta}_k)\, p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})\, p(\boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1}) \quad (10)$$

where the predictive distribution of the state $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})$ can be obtained by using the Chapman-Kolmogorov equation. This paper assumes that the predictive distribution $p(\boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1})$ can depend on $\boldsymbol{\theta}_{k-1}$ or on a non-informative prior distribution (i.e., can change to a value independent of $\boldsymbol{\theta}_{k-1}$). More precisely, the probabilistic model for the changepoint locations proposed in [34] is used for adaptively estimating the time-varying parameter vector $\boldsymbol{\theta}_k$ when a changepoint has occurred. We also assume that there is a probability η (fixed to an a priori value for a given application) of a changepoint at each time instant. As a consequence, the predictive distribution of $\boldsymbol{\theta}_k$ can be defined as follows

$$p(\boldsymbol{\theta}_k|\boldsymbol{y}_{1:k-1}) = \begin{cases} \phi_{\boldsymbol{\theta}_{k-1}}(\boldsymbol{\theta}_k) & \text{with probability } 1-\eta \\ \psi(\boldsymbol{\theta}_k) & \text{with probability } \eta \end{cases} \quad (11)$$

where $\phi_{\boldsymbol{\theta}_{k-1}}(\cdot)$ denotes a prior distribution for $\boldsymbol{\theta}_k$ given $\boldsymbol{\theta}_{k-1}$ (corresponding to the absence of changepoint) and $\psi(\cdot)$ denotes a non-informative prior distribution (corresponding to the presence of changepoint). Note that, as explained in [31], a large value for change probability η will introduce excess parameters from the diffuse prior $\psi(\boldsymbol{\theta})$ when no changepoint has occurred, whereas too small value for η will lead to difficulties in handling time-varying parameters. Based on the previous assumptions, we propose to sample $(\boldsymbol{x}_k, \boldsymbol{\theta}_k)$ based on the APF by using the following steps:

1) Sample $\boldsymbol{x}_k^i$ according to a proposal distribution $q(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^i, \boldsymbol{y}_{1:k})$ for $i = 1, \ldots, N$ where $q(\cdot)$ is the optimal importance distribution introduced in [20].
2) Propagate the previous particles $\boldsymbol{\xi}_{k-1}^i$ by using the evolution equation of the noise parameters to create new particles $\boldsymbol{\xi}_k^{i,-}$, where the superscript "-" means that the quantity is computed a priori.
3) Generate two sets of weights for presence and absence of changepoint at time k, i.e.,

$$\omega_{k,1}^i \propto p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_{k,1}^i, \boldsymbol{\xi}_k^{i,-}\right) \text{ with } \boldsymbol{\theta}_{k,1}^i \sim \phi_{\boldsymbol{\theta}_{k-1}}(\boldsymbol{\theta}) \quad (12)$$

$$\omega_{k,2}^i \propto p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_{k,2}^i, \boldsymbol{\xi}_k^{i,-}\right) \text{ with } \boldsymbol{\theta}_{k,2}^i \sim \psi(\boldsymbol{\theta}) \quad (13)$$

where $i = 1, \ldots, N$. Thus N equally-weighted particles $\{\boldsymbol{x}_{k-1}^i, \boldsymbol{\theta}_{k-1}^i\}_{i=1}^N$ at the $(k-1)$th time instant lead to $2N$ samples at the kth time instant.
4) In order to estimate $\boldsymbol{\theta}_k$, N particles are selected from the $2N$ samples with probabilities proportional to $\left\{(1-\eta)\,\omega_{k,1}^1, \cdots, (1-\eta)\,\omega_{k,1}^N\right\}$ and $\left\{\eta\omega_{k,2}^1, \cdots, \eta\omega_{k,2}^N\right\}$.

5) The weights for the selected N particles are updated according to the APF, i.e.,

$$\omega_k^i = \begin{cases} \dfrac{p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^{j^i}, \boldsymbol{\theta}_{k,1}^{j^i}, \boldsymbol{\xi}_k^{j^i,-}\right)}{\omega_{k,1}^{j^i}} \\ \quad \text{when } j^i \in \{1, \cdots, N\} \\ \dfrac{p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^{j^i}, \boldsymbol{\theta}_{k,2}^{j^i}, \boldsymbol{\xi}_k^{j^i,-}\right)}{\omega_{k,2}^{j^i}} \\ \quad \text{when } j^i \in \{N+1, \cdots, 2N\} \end{cases} \tag{14}$$

where $i = 1, \cdots, N$ and j^i denotes the index of the ith selected particle.

B. CALCULATING $p\left(\xi_k|x_k^i, \theta_k^i, y_{1:k}\right)$ BASED ON VARIATIONAL BAYESIAN INFERENCE

The posterior distribution $p\left(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k}\right)$ conditionally on the ith particle $\left(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i\right)$ cannot be calculated in closed form. According to the VB inference, this posterior can be approximated by another distribution $q\left(\boldsymbol{\xi}_k\right)$, which is factorized into single-variable factors based on the mean-field theory [35], i.e., $q\left(\boldsymbol{\xi}_k\right) = q\left(\boldsymbol{\Lambda}_k\right) q\left(\kappa_k\right) q\left(\nu_k\right)$. According to the VB approximation, the log marginal likelihood $\ln p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{\xi}_k\right)$ can be defined by using the following identity

$$\ln p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{\xi}_k\right) = \mathcal{L} + \mathrm{KL}\left(q||p\right) \tag{15}$$

with

$$\mathcal{L} = \int q\left(\boldsymbol{\xi}_k\right) \ln \frac{p\left(\boldsymbol{y}_k, \boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k-1}\right)}{q\left(\boldsymbol{\xi}_k\right)} \mathrm{d}\boldsymbol{\xi}_k \tag{16}$$

and

$$\mathrm{KL}\left(q||p\right) = \int q\left(\boldsymbol{\xi}_k\right) \ln \frac{q\left(\boldsymbol{\xi}_k\right)}{p\left(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k}\right)} \mathrm{d}\boldsymbol{\xi}_k \tag{17}$$

where $\mathcal{L}$ is a variational lower bound for the log-marginal likelihood, $\mathrm{KL}\left(q||p\right)$ is the Kullback-Leibler (KL) divergence between the true posterior and its approximation. Considering that the KL divergence is non-negative, minimizing the KL divergence can be achieved by maximizing the lower bound $\mathcal{L}$, which results in computing expectations with respect to $q\left(\boldsymbol{\Lambda}_k\right)$, $q\left(\kappa_k\right)$ and $q\left(\nu_k\right)$ in turn. Straightforward computations lead to

$$\begin{aligned} \ln q\left(\boldsymbol{\Lambda}_k\right) &\propto \mathrm{E}_{\kappa_k, \nu_k}\left[\ln p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{\Lambda}_k, \kappa_k\right) + \ln p\left(\boldsymbol{\Lambda}_k\right)\right] \\ &= \sum_{s=1}^{N_\mathrm{s}} \left(\alpha_{s,k} - \frac{1}{2}\right) \ln \lambda_{s,k} \\ &\quad - \left(\beta_{s,k} + \frac{1}{2}\mathrm{E}\left[\kappa_k\right]\left(y_{s,k} - h_{s,\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)^2\right) \lambda_{s,k} \end{aligned} \tag{18}$$

$$\begin{aligned} \ln q\left(\kappa_k\right) &\propto \mathrm{E}_{\boldsymbol{\Lambda}_k, \nu_k}\left[\ln p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{\Lambda}_k, \kappa_k\right) + \ln p\left(\kappa_k|\nu_k\right)\right] \\ &= \frac{\mathrm{E}\left[\nu_k\right] - 1}{2} \ln \kappa_k - \frac{\mathrm{E}\left[\nu_k^i\right]}{2} \kappa_k \\ &\quad - \frac{\left(\boldsymbol{y}_k - h_{\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)^\mathrm{T} \mathrm{E}\left[\boldsymbol{\Lambda}_k\right]\left(\boldsymbol{y}_k - h_{\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)}{2} \kappa_k \end{aligned} \tag{19}$$

$$\begin{aligned} \ln q\left(\nu_k\right) &\propto \mathrm{E}_{\kappa_k}\left[\ln p\left(\kappa_k|\nu_k\right) + \ln p\left(\nu_k\right)\right] \\ &= \left(a_k - \frac{1}{2}\right) \ln \nu_k \\ &\quad - \left(b_k + \frac{\mathrm{E}\left[\kappa_k\right]}{2} - \frac{\mathrm{E}\left[\ln \kappa_k\right]}{2} - \frac{1}{2}\right) \nu_k \end{aligned} \tag{20}$$

where $i = 1, \cdots, N_\mathrm{s}$, $y_{s,k}$ and $h_{s,\boldsymbol{\theta}_k^i}(\cdot)$ denote the sth measurement and the corresponding nonlinear function. The expectations in the above equations can be expressed as follows

$$\begin{aligned} &\mathrm{E}\left[\lambda_{s,k}\right] = \frac{\alpha_{s,k}}{\beta_{s,k}} \quad \mathrm{E}\left[\kappa_k\right] = \frac{\nu_k^1}{\nu_k^2} \\ &\mathrm{E}\left[\ln \kappa_k\right] = \Psi\left(\nu_k^1\right) - \ln \nu_k^2 \quad \mathrm{E}\left[\nu_k\right] = \frac{a_k}{b_k} \end{aligned} \tag{21}$$

where $s = 1, \ldots, N_\mathrm{s}$ and $\Psi(\cdot)$ denotes the digamma function. According to (18)-(20), the hyperparameters of the posterior distribution conditionally on the ith particle $\left(\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i\right)$ can be updated as follows [13]

$$\alpha_{s,k}^{i,+} = \alpha_{s,k}^{i,-} + \frac{1}{2} \tag{22}$$

$$\beta_{s,k}^{i,+} = \beta_{s,k}^{i,-} + \frac{1}{2}\mathrm{E}\left[\kappa_k^i\right]\left(y_{s,k} - h_{s,\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)^2 \tag{23}$$

$$\nu_k^{1,i,+} = \frac{\mathrm{E}\left[\nu_k^i\right] + 1}{2} \tag{24}$$

$$\begin{aligned} \nu_k^{2,i,+} &= \frac{\mathrm{E}\left[\nu_k^i\right]}{2} \\ &\quad + \frac{\left(\boldsymbol{y}_k - h_{\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)^\mathrm{T} \mathrm{E}\left[\boldsymbol{\Lambda}_k^i\right]\left(\boldsymbol{y}_k - h_{\boldsymbol{\theta}_k^i}\left(\boldsymbol{x}_k^i\right)\right)}{2} \end{aligned} \tag{25}$$

$$a_k^{i,+} = a_k^{i,-} + \frac{1}{2} \tag{26}$$

$$b_k^{i,+} = b_k^{i,-} + \frac{\mathrm{E}\left[\kappa_k^i\right]}{2} - \frac{\mathrm{E}\left[\ln \kappa_k^i\right]}{2} - \frac{1}{2} \tag{27}$$

where $s = 1, \cdots, N_\mathrm{d}$, $i = 1, \cdots, N$. Note that the superscript "+" in the above equation means that the quantity is computed a posteriori. In order to maintain the conjugacy for the distribution of the noise statistical parameters, $\alpha_{s,k}$, $\beta_{s,k}$, a_k, b_k are propagated as follows [12]

$$\begin{aligned} \alpha_{s,k}^{i,-} &= \rho\alpha_{s,k-1}^{i,+} \quad \beta_{s,k}^{i,-} = \rho\beta_{s,k-1}^{i,+} \\ a_k^{i,-} &= \rho a_{k-1}^{i,+} \quad b_k^{i,-} = \rho b_{k-1}^{i,+} \end{aligned} \tag{28}$$

where ρ is a forgetting factor, which can be considered as a weight balancing the historical estimates and the current information. Accordingly, a small forgetting factor can track faster changes in the noise statistical parameters whereas a larger value of this factor induces a slower response [17].

Thus $p\left(\boldsymbol{\xi}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \boldsymbol{y}_{1:k}\right)$ can be approximated by iteratively calculating (22)-(27) until an iteration stopping rule is satisfied. Note that the updates for ν_k^1 and ν_k^2 in (24) and (25) are functions of α_k, β_k, a_k and b_k, whereas the updates for β_k

and b_k in (23) and (27) require to know the values of ν_k^1 and ν_k^2. Therefore, the estimates of ν_k^1 and ν_k^2 in (24) and (25) at the $(r+1)$th iteration are calculated by using the estimates of α_k, β_k, a_k and b_k at the rth iteration. Finally, note that the estimates of β_k and b_k at the $(r+1)$th iteration can be implemented based on the last estimates of ν_k^1 and ν_k^2. The proposed RBPF with VB inference for joint state and time-varying parameter estimation is summarized in Alg. 1.

IV. EXPERIMENTAL RESULTS

A. ILLUSTRATIVE EXAMPLE

In order to evaluate the proposed RBPF with variational inference, we use the following modified benchmark model [6] for the illustrations

$$x_k = 0.5\, x_{k-1} + 25\, x_{k-1}/\left(1 + x_{k-1}^2\right) + 8\cos(1.2\, k) + \omega_k$$
$$y_k = x_k^2/a + v_k$$

where $\omega_k \sim \mathcal{N}(0, 0.5)$, a is a time-varying parameter taking values $\{5, -6, 7, 14, 5\}$ with changes occurring at time instants $\{0, 40, 80, 120, 160\}$. The noise v_k is the non-Gaussian measurement noise whose distribution is $(1-\varepsilon)\mathcal{N}(0,1) + \varepsilon\mathcal{D}$ where ε denotes the outlier probability, $\mathcal{D}$ denotes the outlier distribution, which was set to $\mathrm{U}(-20, 20)$ where $\mathrm{U}(\cdot)$ is the uniform distribution [13], [36]. We compare the proposed approach with the interacting multiple model (IMM) algorithm [1] and the APE filter [31]. More details about these algorithms are provided below

- IMM Algorithm [1]: The models in the IMM differ by the choice of the parameter a where 10 equally spaced values of a are sampled in the range $[-18, 18]$ and the measurement noise associated with each model is assumed to be distributed according to a $\mathcal{N}(0,1)$ distribution. Moreover, the unscented Kalman filter is embedded in the IMM in order to estimate the state for the nonlinear SSM.
- APE filter with known variance [31]: The APE filter is combined to the APF with a changepoint model and the measurement noise model is assumed to be distributed according to a $\mathcal{N}(0,1)$ distribution.
- APE filter with unknown variance [31]: In the presence of measurement outliers, the estimation accuracy of the APE filter can be improved by considering that the measurement noise is a zero-mean Gaussian distribution with unknown variance, which is estimated with other parameters, i.e., the APE filter with unknown variance is combined to the APF with a changepoint model and the conjugate prior for the variance of Gaussian distribution is updated by using the state particles and measurements.

In the changepoint model applied to the APE filter and proposed strategy, the predictive distribution of a conditionally upon a_{k-1} is assumed to be a mixture of multivariate Gaussian distributions based on the kernel smoothing method [4], i.e., $\phi_{a_{k-1}}(a) \approx \sum_{i=1}^N \omega_{k-1}^i \mathcal{N}\left(a|\zeta_{k-1}^i, h^2\, V_{k-1}\right)$ where

---

Algorithm 1 Proposed RBPF With Variational Inference

Inputs: $\left\{\boldsymbol{x}_{k-1}^i, \boldsymbol{\theta}_{k-1}^i, \left\{\alpha_{s,k-1}^i, \beta_{s,k-1}^i\right\}_{s=1}^{N_\mathrm{d}}, a_{k-1}^i, b_{k-1}^i\right\}_{i=1}^N$

Outputs: $\left\{\boldsymbol{x}_k^i, \boldsymbol{\theta}_k^i, \left\{\alpha_{s,k}^i, \beta_{s,k}^i\right\}_{s=1}^{N_\mathrm{d}}, \nu_k^{1,i}, \nu_k^{2,i}, a_k^i, b_k^i\right\}_{i=1}^N$

% **Step 1**: Generate samples of $(\boldsymbol{x}_k, \boldsymbol{\theta}_k)$ by using the APF.

Generate $\boldsymbol{x}_k^i \sim p\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^i\right)$ by using (1), and then compute $\left\{\alpha_{s,k}^{i,-}, \beta_{s,k}^{i,-}\right\}_{s=1}^{N_\mathrm{d}}, a_k^{i,-}, b_k^{i,-}$ according to (28) and generate $\nu_k^{i,-} \sim \gamma\left(\nu_k|a_k^{i,-}, b_k^{i,-}\right)$ where $i = 1, \ldots, N$.

Generate $\left\{\lambda_{s,k}^i \sim \gamma\left(\lambda_k|\alpha_{s,k}^{i,-}, \beta_{s,k}^{i,-}\right)\right\}_{s=1}^{N_\mathrm{d}}$ and $\kappa_k^i \sim \gamma\left(\kappa_k|\frac{\nu_k^{i,-}}{2}, \frac{\nu_k^{i,-}}{2}\right)$ where $i = 1, \ldots, N$.

Generate $\boldsymbol{\theta}_{k,1}^i \sim \phi_{\boldsymbol{\theta}_{k-1}}(\boldsymbol{\theta})$ and $\boldsymbol{\theta}_{k,2}^i \sim \psi(\boldsymbol{\theta})$, and then compute $\omega_{k,1}^i \propto p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_{k,1}^i, (\kappa_k^i\boldsymbol{\Lambda}_k^i)^{-1}\right)$ and $\omega_{k,2}^i \propto p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^i, \boldsymbol{\theta}_{k,2}^i, (\kappa_k^i\boldsymbol{\Lambda}_k^i)^{-1}\right)$ where $\boldsymbol{\Lambda}_k^i = \mathrm{diag}\left[\lambda_{1,k}^i, \cdots, \lambda_{N_\mathrm{d},k}^i\right]$ and $i = 1, \cdots, N$.

for $i = 1, \ldots, N$ **do**
  Sample index j^i in $\{1, \cdots, 2N\}$ with probabilities $\left\{(1-\eta)\,\omega_{k,1}^i\right\}_{i=1}^N$ and $\left\{\eta\omega_{k,2}^i\right\}_{i=N+1}^{2N}$;
  Propagate $\boldsymbol{x}_k^i \sim p\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{j^i}\right)$;
  If $j^i \in \{1, \cdots, N\}$, update $\boldsymbol{\theta}_k^i \sim \phi_{\boldsymbol{\theta}_{k,1}^{j^i}}(\boldsymbol{\theta})$ and compute $\tilde{\omega}_k^i \propto \frac{p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^{j^i}, \boldsymbol{\theta}_{k,1}^{j^i}, \left(\kappa_k^{j^i}\boldsymbol{\Lambda}_k^{j^i}\right)^{-1}\right)}{\omega_{k,1}^{j^i}}$;
  If $j^i \in \{N+1, \cdots, 2N\}$, set $\boldsymbol{\theta}_k^i = \boldsymbol{\theta}_{k,2}^{j^i}$ and compute $\tilde{\omega}_k^i \propto \frac{p\left(\boldsymbol{y}_k|\boldsymbol{x}_k^{j^i}, \boldsymbol{\theta}_{k,2}^{j^i}, \left(\kappa_k^{j^i}\boldsymbol{\Lambda}_k^{j^i}\right)^{-1}\right)}{\omega_{k,2}^{j^i}}$;
end for

% **Step 2**: Calculating $q(\boldsymbol{\lambda}_k, \kappa_k, \nu_k)$ by using VB inference.

Initialize $\alpha_{s,k}^{i,+}(0) = \alpha_{s,k}^{j^i,-}$, $\beta_{s,k}^{i,+}(0) = \beta_{s,k}^{j^i,-}$, $a_k^{i,+}(0) = a_k^{j^i,-}$ and $b_k^{i,+}(0) = b_k^{j^i,-}$ where $s = 1, \cdots, N_\mathrm{d}$ and $i = 1, \ldots, N$.

for $i = 1, \ldots, N$ **do**
  for $r = 1, \ldots, r_{\max}$ **do**
    Compute $\nu_k^{1,i,+}(r)$, $\nu_k^{2,i,+}(r)$ according to (24) and (25);
    Compute $\left\{\alpha_{s,k}^{i,+}(r), \beta_{s,k}^{i,+}(r)\right\}_{s=1}^{N_\mathrm{d}}$, $a_k^{i+}(r)$ and $b_k^{i+}(r)$ according to (22), (23), (26) and (27);
    if the parameters change by less than 0.001 **then**
      stop the iteration;
    else
      set $r = r + 1$;
    end if
  end for
end for

Normalize $\omega_k^i = \tilde{\omega}_k^i / \sum_{i=1}^N \left(\tilde{\omega}_k^i\right)$ and perform particle resampling.

% **Step 3**: (Recursion) $k = k + 1$.

Go to Step 1.

$\zeta_{k-1}^i = \tau a_{k-1}^i + (1-\tau)\,\bar{a}$, $V_{k-1} = \sum_{i=1}^N \omega_{k-1}^i\left(a_{k-1}^i - \bar{a}\right)^2$, $\bar{a} = \sum_{i=1}^N \omega_{k-1}^i a_{k-1}^i$, h^2 is the kernel smoothing parameter, $\tau = \sqrt{1-h^2}$ is the shrinkage parameter of the kernel mean,

TABLE 1. Parameters used for the simulations on synthetic data.

Changepoint occurrence probability η	0.05
Non-informative prior pdf $\psi(a)$	$U(-20, 20)$
Kernel smoothing parameter h^2	0.01
Number of particles N	200
Forgetting factor ρ	$1 - \exp(-4)$
Maximum number of VB iteration $r_{\max}$	10
VB hyperparameters $\alpha_0, \beta_0, a_0, b_0$	1

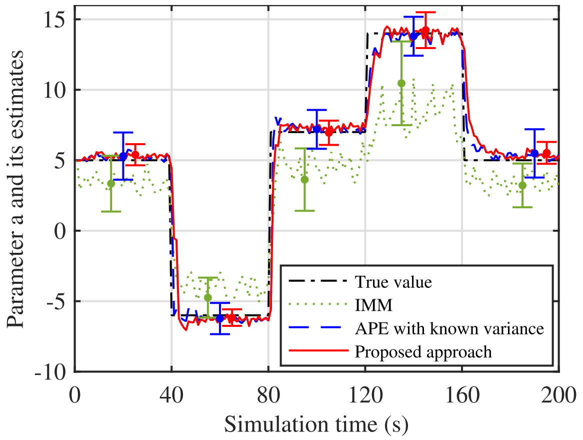

FIGURE 1. Mean of the estimates $\widehat{a}$ for 100 MC runs in the absence of outliers.

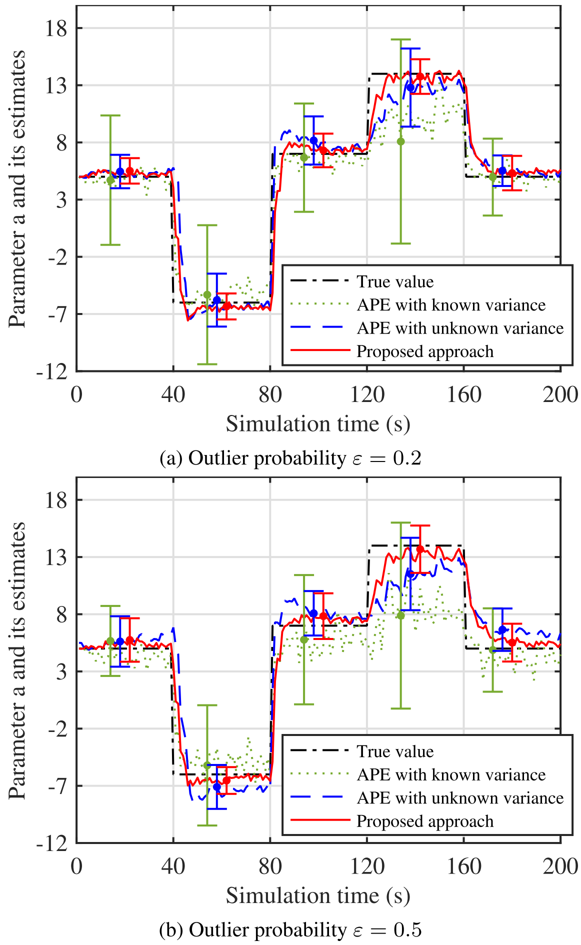

FIGURE 2. Mean of the estimates $\widehat{a}$ for two different outlier probabilities (100 MC runs).

while the non-informative prior pdf $\psi(a)$ was assumed to be a uniform distribution. The related parameters used in all test scenarios are provided in Table 1. $N_m = 100$ Monte Carlo simulations have been run for any approach to compute the average root mean square errors (ARMSEs) of the estimates defined by $\sqrt{(KN_m)^{-1}\sum_{k=1}^{K}\sum_{m=1}^{N_m}(\widehat{x}_k(m) - x_k)^2}$, where $\widehat{x}_k(m)$ is the mth state estimate and K is the number of time instants. All algorithms have been coded using MATLAB and run on a laptop with Intel i-5 and 8 GB RAM.

Fig. 1 displays the means and standard deviations of $\widehat{a}$ [1] computed using 100 MC simulations for the IMM, the APE filter with known variance and the proposed approach in the absence of outliers (i.e., $\varepsilon = 0$). Considering that the changepoint model for parameter a is taken into account both in the proposed approach and in the APE filter, the estimation accuracies for parameter a obtained with these two approaches are obviously better than the one obtained with the IMM. Since the Student-t distribution provides almost the same solution as the Gaussian distribution in the absence of outliers, similar mean and standard deviation for the estimator $\widehat{a}$ should be obtained with the proposed approach and the APE filter.

The means and standard deviations of $\widehat{a}$ computed using 100 MC simulations for different outlier probabilities are depicted in Fig. 2. It is clear that the estimation performance of the APE filter with known variance for parameter a severely degrades in the presence of measurement outliers, whereas the estimation performance of the APE filter is robust to outliers when the variance of measurement noise is assumed to be time-varying and jointly estimated with the state. However, the performance of the proposed approach is less affected by the presence of outliers due to the Student-t distribution, which is more robust to outliers than the Gaussian distribution. The ARMSEs of the state estimates obtained with the three approaches for different outlier probabilities are reported in Table 2. In presence or absence of outliers, the ARMSEs obtained with the proposed approach and the APE filter are obviously smaller than the one obtained with the IMM. Moreover, the state estimation accuracy of the proposed approach is better than that of the APE filter in the absence of outliers. Finally, thanks to the good robustness of the Student-t distribution to outliers, the proposed approach provides more accurate state estimates for any outlier probability.

Table 3 shows the execution times for 100 MC runs by using different numbers of particles for the proposed approach and the APE filter with unknown variance. The main difference between the proposed approach and the APE filter with unknown variance is that the computation of the

[1] The estimator of the parameter a is denoted as $\widehat{a}$

TABLE 2. ARMSEs of state estimates for different outlier probabilities.

Outlier probability	Methods		
	Proposed approach	APE filter	IMM
$\varepsilon = 0$	2.85	3.36 [a]	11.86
$\varepsilon = 0.2$	3.52	3.95 [b]	11.96
$\varepsilon = 0.5$	3.89	4.47 [b]	12.01

[a]Using the APE filter with known variance
[b]Using the APE filter with unknown variance

TABLE 3. Execution times for different numbers of particles.

N	Execution times for 100 MC runs (s)		
	Proposed method with $r_{\max} = 10$	Proposed method with stopping rule	APE filter [a]
200	255.68	207.37	142.06
400	520.34	453.13	288.47
800	1023.34	928.34	612.82

[a]Using the APE filter with unknown variance

VB iterations is embedded in the APF before performing the update of the unknown parameter distribution, which leads to a higher computation load for APF, as reported in Table 3. Note that the computational cost of the proposed approach was evaluated in the following two situations: (a) the algorithm is stopped when the stopping rule in Line 15 of Alg. 1 is satisfied, (b) the number of variational iterations is set to its maximum value $r_{\max}$. It is clear that the corresponding computational cost can be efficiently reduced by introducing the stopping rule of Line 15 of Alg. 1 in the proposed approach.

B. GNSS MULTIPATH MITIGATION IN URBAN CANYONS

This section validates the proposed RBPF with VB inference in the context of multipath (MP) interference mitigation for GNSS receivers in urban canyons. MP signals are mainly due to the fact that a signal transmitted by a navigation satellite is very likely to be reflected or diffracted and can follow different paths before arriving at the GNSS receiver [37]. Depending on the availability of the direct signal, reflected MP signals affecting the received direct GNSS signal can be divided into two situations: (a) MP interfereces resulting from the sum of the direct signal and of 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 [38], [39]. Accordingly, the pseudo-range measurement noise error in the MP interference situation is considered as a non-Gaussian stochastic process, while a mean value jump is present on the pseudo-range measurement in the NLOS signal situation [40]–[42]. Since our aim is to estimate the pseudo-range measurement error resulting from MP rather than to identify the type of MP, the impact of MP signals on the pseudo-range measurement in this work is formulated as an unknown mean value jump and a non-Gaussian measurement noise distributed according to the Student-t distribution in this work. Assuming the atmospheric propagation errors can be compensated within the GNSS receiver, the pseudo-range measurement model collected from N_{d} in-view satellites in the presence of MP can be defined using the following compact form

$$\boldsymbol{y}_k = h_{\boldsymbol{\theta}_k}(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{29}$$

with

$$h_{\boldsymbol{\theta}_k}(\boldsymbol{x}_k) = \left(h_{1,\theta_{1,k}}(\boldsymbol{x}_k), \ldots, h_{N_{\mathrm{d}},\theta_{N_{\mathrm{d}},k}}(\boldsymbol{x}_k)\right)^{\mathrm{T}} \tag{30}$$

and

$$h_{s,\theta_{s,k}}(\boldsymbol{x}_k) = \| \boldsymbol{p}_{s,k} - \boldsymbol{p}_k \| + b_k + \theta_{s,k} \tag{31}$$

where $s = 1, \ldots, N_{\mathrm{d}}$, $\boldsymbol{x}_k$ is state vector for describing the dynamic of the vehicle in the earth-centered earth-fixed (ECEF) frame at the kth time instant, $\boldsymbol{y}_k = \left(y_{1,k}, \ldots, y_{N_{\mathrm{d}},k}\right)^{\mathrm{T}}$ is the pseudo-range measurement vector composed of all in-view satellites, $\boldsymbol{p}_{s,k} = \left(x_{s,k}, y_{s,k}, z_{s,k}\right)^{\mathrm{T}}$ and $\boldsymbol{p}_k = (x_k, y_k, z_k)^{\mathrm{T}}$ are the sth satellite and vehicle positions in the ECEF frame, b_k is the GNSS receiver clock offset, $\| \cdot \|$ and $(\cdot)^{\mathrm{T}}$ denote the Euclidean norm and the transpose of a vector, respectively. In the presence of MP, $\boldsymbol{\theta}_k = \left(\theta_{1,k}, \cdots, \theta_{N_{\mathrm{d}},k}\right)^{\mathrm{T}}$ denotes the mean value jumps appearing on the pseudo-range measurement of all in-view satellites ($\theta_{s,k} = 0$ in the absence of MP) and $\boldsymbol{v}_k$ is the measurement noise distributed according to the Student-t distribution (as in (3)). In practice, an important property of MP signals is that they are not only depend on the relative position of the receiver and GNSS satellites, but also on the environment where the receiver is located, especially in urban canyons. Thus the mean value jump $\boldsymbol{\theta}_k$ and the measurement noise parameters $\boldsymbol{\xi}_k = \{\boldsymbol{\Lambda}_k, \kappa_k, \nu_k\}$ can be considered as unknown time-varying model parameters, which need to be estimated jointly with the vehicle state $\boldsymbol{x}_k$. In addition, the state vector $\boldsymbol{x}_k$ considered in this paper is defined as follows [42], [43]

$$\boldsymbol{x}_k = (x_k, \dot{x}_k, y_k, \dot{y}_k, z_k, \dot{z}_k, b_k, d_k)^{\mathrm{T}} \tag{32}$$

where $k = 1, \ldots, K$ denotes the kth sampling time instant, $\dot{\boldsymbol{p}}_k = (\dot{x}_k, \dot{y}_k, \dot{z}_k)^{\mathrm{T}}$ is the vehicle velocity in the ECEF frame, d_k is the GNSS receiver clock drift. The velocity can be reasonably modelled as a random walk, e.g., $\ddot{x} = e_x$ 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_k and its drift d_k can also be modelled as random walks, i.e., $\dot{b}_k = d_k + e_b$ and $\dot{d}_k = e_d$ where e_b and e_d are zero-mean Gaussian white noises of variance σ_b^2 and σ_d^2. Based on the above assumptions, the discrete-time state model which describes the propagation of the vehicle state $\boldsymbol{x}_k$ can be formulated as

$$\boldsymbol{x}_k = \boldsymbol{F}_{k|k-1}\boldsymbol{x}_{k-1} + \boldsymbol{e}_k \tag{33}$$

where $k = 1, \ldots, K$ denotes the kth sampling time instant, $\boldsymbol{e}_k = \left(e_x,\ e_y,\ e_z,\ e_b,\ e_d\right)^{\mathrm{T}}$ is a zero mean Gaussian noise vector of covariance matrix $\boldsymbol{Q}_k$. Details on the matrices $\boldsymbol{F}_{k|k-1}$ and $\boldsymbol{Q}_k$ can be found in [44].

The experimental data was collected during a measurement campaign carried out in Toulouse center (France) and was

FIGURE 3. Urban canyon trajectory used in the proposed experiments (obtained with Google Earth).

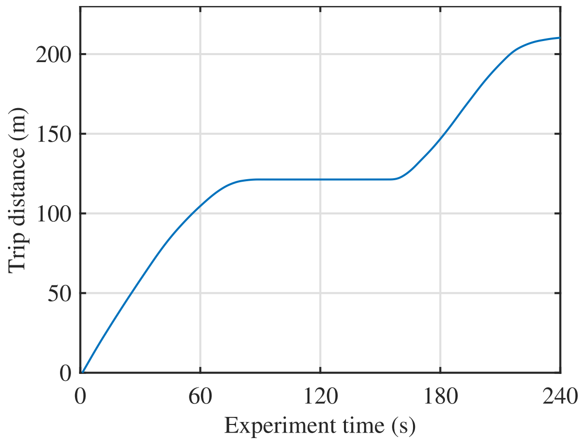

FIGURE 4. Distance to the origin versus time.

used to evaluate the method proposed in [42]. A synchronized integrated navigation system composed of a Novatel receiver coupled to a tactical grade IMAR IMU was used to provide a reference trajectory. Taking advantage of a ground reference station, differential corrections were performed to obtain position accuracy close to 1 m for the reference trajectory, which is considered as the ground truth. For assessing the algorithm performance, the vehicle was equipped with a UBLOX 6T receiver. This receiver delivered not only the position, velocity and time solution, but also, for each satellite, the raw pseudo-range and Doppler frequency measurements, as well as the navigation message. It allowed us to compute satellite locations, and to perform timing and propagation correction on the measured pseudo-range. Data were collected in street urban canyons during which the receiver was strongly affected by MP signals, and post-processed using Matlab.

Fig. 3 shows the trajectory considered in our measurement campaign (lasting 240 s). Fig. 4 displays the evolution of the distance to the starting point of the trajectory (considered in our experiment) versus time, where the original point is defined as the initial position on the trajectory and the trip distance represents the horizontal distance travelled from the initial position. It is clear that the distance to the origin does

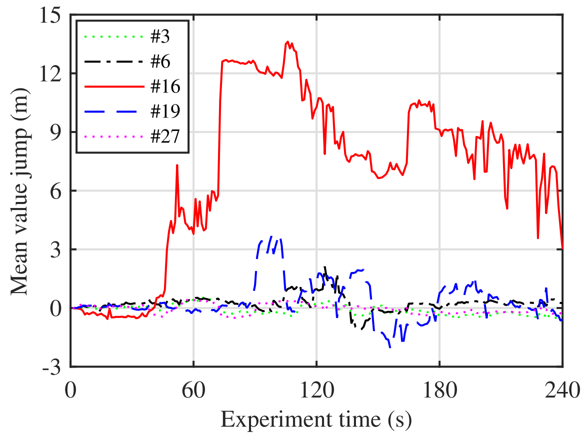

FIGURE 5. Typical example of estimated mean value jumps.

not change during the time interval (78 s, 159 s), as the vehicle is stopped in the middle of two buildings. As it appears at the LOS frequency (the Doppler frequency related to the vehicle velocity is zero), pseudo-range measurements are severely impacted by mean value jump during this period. The in-view satellites observed during the experience are satellites #3, #6, #19, #26, and #27 (i.e., $N_\mathrm{d} = 5$). The non-informative prior used for the mean value jump is a uniform distribution defined by $\psi\left(\theta_{s,k}\right) = \mathrm{U}\left(-20\,\mathrm{m}, 20\,\mathrm{m}\right)$,[2] where $s = 1, \cdots, N_\mathrm{d}$. In order to evaluate the proposed algorithm, we propose to compare its positioning estimation accuracy with that obtained using the standard APF and the Dirichlet process mixture (DPM)-based RBPF studied in [30]. In [30], the pseudo-range measurement noise in the absence of MP is assumed to have a zero mean Gaussian distribution with a nominal standard deviation, whereas the mean and standard deviation of the measurement noise can change abruptly or evolve slowly during a long period of time in the presence of MP, leading to time-varying measurement noise parameters. The DPM-approach of [30] was considered in this work since it provided interesting results for state-space models with time-varying parameters, in particular for MP mitigation in urban canyons. The standard deviations of the process noise, the clock offset and drift noises were set to $\sigma_a = 1\,\mathrm{m/s}^2$, $\sigma_b = 3\mathrm{c} \times 10^{-10}\,\mathrm{m}$ and $\sigma_d = 2\pi\mathrm{c} \times 10^{-10}\,\mathrm{m/s}$, respectively, where $\mathrm{c} = 3 \times 10^8\,\mathrm{m/s}$ denotes the velocity of light. The nominal standard deviation of the pseudo-range measurement noise used for the standard APF was adjusted by cross-validation and was set to 4 m, and the number of particles for all three approaches is 5000. The other parameters used in the proposed approach are reported in Table 1.

Fig. 5 displays typical estimates for the mean value jumps $\widehat{\boldsymbol{\theta}}_k$ impacting the pseudo-range measurements of all in-view satellites during the measurement campaign. According to

[2]The sign of the mean value jump depends on the value of the MP signal carrier phase relative to the direct signal. When the relative carrier phase belongs to $\left(-90°, 90°\right)$, the sign of the mean value jump is positive. When the relative carrier phase belongs to $\left(-180°, -90°\right)$ or $\left(90°, 180°\right)$, the sign of the mean value jump is negative.

TABLE 4. Elevation angles for in-view satellites.

In-view satellite	Satellite elevation
satellite #3	81.9° - 82.4°
satellite #6	77.9° - 77.5°
satellite #16	56.1° - 55.5°
satellite #19	60.8° - 60.3°
satellite #27	82.9° - 85.9°

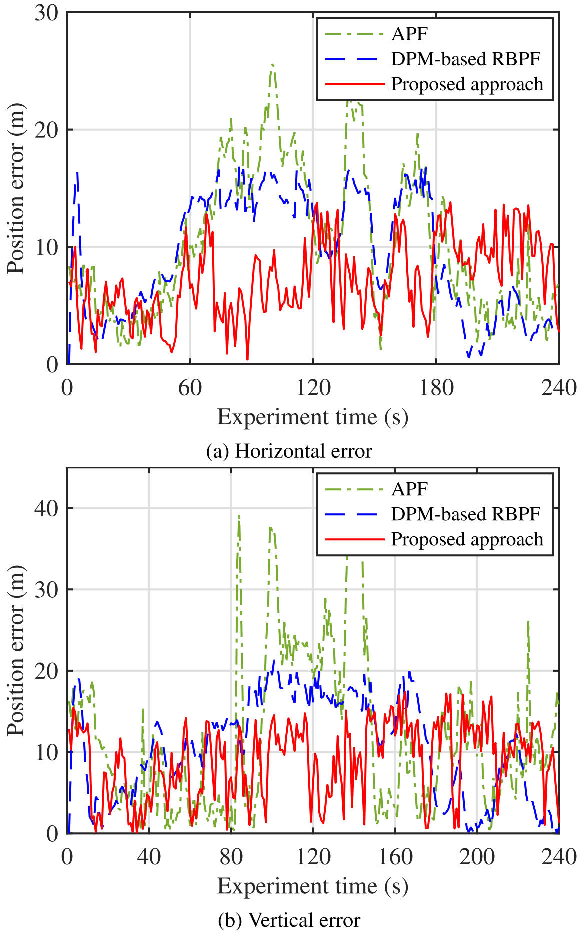

FIGURE 6. Positioning errors versus time.

the results obtained with the proposed approach, the pseudo-range measurement of satellites #3, #6 and #27 are less affected by MP than the other received signals. Conversely, the mean value jumps of satellites #16 and #19 are more important, confirming the presence of MP. The elevation angles of all in-view satellites during the measurement campaign are reported in Table 4. Note that the elevation angles for satellites #3, #6 and #27 are larger than 75°, whereas the elevation angles for satellites #16 and #19 are less than or equal to 60°. In general, the signals acquired with low elevation angles (for satellites #16 and #19) are more likely to be affected by MP, which seems to be the case here.

Fig. 6 displays the positioning errors (horizontal and vertical errors versus time) for the different approaches. It is clear that MP signals severely impair the positioning solution if these interferences are not processed within the receiver. Conversely, the positioning errors obtained with the DPM-based RBPF and the proposed approach remains lower than 20 m, confirming that the MP signals affecting the pseudo-range measurements have been mitigated. We think that the slightly better positioning accuracy obtained with the proposed approach with respect to the DPM-based RBPF is due to a good estimation of the locations of the mean value jumps, which is possible thanks to the use of the change-point model defined in (11). Note that the MP amplitude, delay and phase are constant with respect to those of the direct signal when the vehicle stops in the middle of two buildings (leading to a constant jump). As a consequence, the impact resulting from the MP signal is the maximum during this period (i.e., 78 s-159 s).

V. CONCLUSION

This paper proposed a Rao-Blackwellized particle filter with variational inference for jointly estimating state and time-varying parameters in non-linear state-space models with non-Gaussian measurement noise. The proposed approach embedded the variational Bayesian inference in an auxiliary particle filter with a probabilistic changepoint model to build a joint state and time-varying parameter estimator in the presence of non-Gaussian measurement noise. A simulation study was conducted by using a modified nonlinear benchmark model in order to compare the performance of the proposed approach with the interacting multiple model and the adaptive parameter estimation filter. Since the Student-t distribution considered in the proposed approach is more robust to outliers, the proposed approach provided better estimate accuracy for the state and unknown parameters when compared to the other strategies. Finally, the proposed approach was validated by processing GNSS receiver data collected from a measurement campaign carried out in an urban environment and proved its efficiency for MP mitigation, resulting in improved positioning accuracy.

One of the drawbacks of the proposed approach is the requirement that a conjugate prior of the unknown parameters needs to be available for optimizing the objective function in the proposed variational Bayesian inference. Recent work on stochastic variational inference might solve this problem by applying stochastic optimization to the objective function [45]. Applying stochastic variational inference for estimating the state and parameters of the changepoint model investigated in this work is currently under investigation. Validating the proposed method using data from other measurement campaigns and high-dimensional state-space models is also an interesting prospect.

ACKNOWLEDGMENT

The authors would like to thank Audrey Giremus from the University of Bordeaux for helpful discussions on Dirichlet process mixtures.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Appplication to Tracking Navigation: Theory, Algorithm Software*. New York, NY, USA: Weily, 2004.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

[3] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 1, pp. 103–123, Jan. 1998.

[4] M. West, "Approximating posterior distributions by mixtures," *J. Roy. Stat. Soc. B, Methodol.*, vol. 55, pp. 409–422, Jan. 1993.

[5] G. Storvik, "Particle filters for state-space models with the presence of unknown static parameters," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 281–289, Feb. 2002.

[6] F. Caron, M. Davy, E. Duflos, and P. Vanheeghe, "Particle filtering for multisensor data fusion with switching observation models: Application to land vehicle positioning," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2703–2719, Jun. 2007.

[7] A. Zia, T. Kirubarajan, J. P. Reilly, D. Yee, K. Punithakumar, and S. Shirani, "An EM algorithm for nonlinear state estimation with model uncertainties," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 921–936, Mar. 2008.

[8] T. B. Schön, A. Wills, and B. Ninness, "System identification of nonlinear state-space models," *Automatica*, vol. 47, no. 1, pp. 39–49, Jan. 2011.

[9] N. Kantas, A. Doucet, S. S. Singh, J. Maciejowski, and N. Chopin, "On particle methods for parameter estimation in state-space models," *Stat. Sci.*, vol. 30, no. 3, pp. 328–351, Aug. 2015.

[10] E. Ozkan, F. Lindsten, C. Fritsche, and F. Gustafsson, "Recursive maximum likelihood identification of jump Markov nonlinear systems," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 754–765, Feb. 2015.

[11] V. Šmídl and A. Quinn, "On-line inference of time-variant parameters," in *The Variational Bayes Method in Signal Processing*. New York, NY, USA: Springer, 2006, pp. 145–176.

[12] S. Sarkka and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational Bayesian approximations," *IEEE Trans. Autom. Control*, vol. 54, no. 3, pp. 596–600, Mar. 2009.

[13] H. Zhu, H. Leung, and Z. He, "A variational Bayesian approach to robust sensor fusion based on Student-t distribution," *Inf. Sci.*, vol. 221, pp. 201–214, Feb. 2013.

[14] Z. Li, J. Zhang, J. Wang, and Q. Zhou, "Recursive noise adaptive extended object tracking by variational Bayesian approximation," *IEEE Access*, vol. 7, pp. 151168–151179, 2019.

[15] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, "A novel robust Student's t-based Kalman filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1545–1554, Jun. 2017.

[16] Y. Huang, Y. Zhang, Y. Zhao, and J. A. Chambers, "A novel robust Gaussian–Student's t mixture distribution based Kalman filter," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3606–3620, Jul. 2019.

[17] E. Özkan, V. Šmídl, S. Saha, C. Lundquist, and F. Gustafsson, "Marginalized adaptive particle filtering for nonlinear models with unknown time-varying noise parameters," *Automatica*, vol. 49, no. 6, pp. 1566–1575, Jun. 2013.

[18] D. Xu, C. Shen, and F. Shen, "A robust particle filtering algorithm with non-Gaussian measurement noise using Student-t distribution," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 30–34, Jan. 2014.

[19] G. Casella and C. P. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[20] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[21] O. Cappe, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, vol. 95, no. 5, pp. 899–924, May 2007.

[22] T. Schon, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, Jul. 2005.

[23] S. Särkkä, A. Vehtari, and J. Lampinen, "Rao-Blackwellized particle filter for multiple target tracking," *Inf. Fusion*, vol. 8, no. 1, pp. 2–15, Jan. 2007.

[24] S. Thrun, "Particle filters in robotics," in *Proc. Uncertain. Artif. Intell.*, Alberta, WC, Canada, 2002, pp. 511–518.

[25] A. M. Johansen, N. Whiteley, and A. Doucet, "Exact approximation of Rao-Blackwellised particle filters," in *Proc. 6th IFAC Symp. Syst. Ident.*, Brussels, Belgium, 2012, pp. 488–493.

[26] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 72, no. 3, pp. 269–342, Jun. 2010.

[27] J. Ala-Luhtala, N. Whiteley, K. Heine, and R. Piche, "An introduction to twisted particle filters and parameter estimation in non-linear state-space models," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4875–4890, Sep. 2016.

[28] D. Crisan and J. Míguez, "Nested particle filters for online parameter estimation in discrete-time state-space Markov models," *Bernoulli*, vol. 24, no. 4A, pp. 3039–3086, Nov. 2018.

[29] C. A. Naesseth, F. Lindsten, and T. B. Schon, "High-dimensional filtering using nested sequential Monte Carlo," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4177–4188, Aug. 2019.

[30] A. Rabaoui, N. Viandier, E. Duflos, J. Marais, and P. Vanheeghe, "Dirichlet process mixtures for density estimation in dynamic nonlinear modeling: Application to GPS positioning in urban canyons," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1638–1655, Apr. 2012.

[31] C. Nemeth, P. Fearnhead, and L. Mihaylova, "Sequential Monte Carlo methods for state and parameter estimation in abruptly changing environments," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1245–1255, Mar. 2014.

[32] C. M. Bishop, "Probability distributions," in *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006, pp. 102–105.

[33] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[34] A. Kong, J. S. Liu, and W. H. Wong, "Sequential imputations and Bayesian missing data problems," *J. Amer. Stat. Assoc.*, vol. 89, no. 425, pp. 278–288, Mar. 1994.

[35] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.

[36] Q. Wu, Y. Li, Y. V. Zakharov, W. Xue, and W. Shi, "A kernel affine projection-like algorithm in reproducing kernel Hilbert space," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, early access, Oct. 15, 2019, doi: 10.1109/TCSII.2019.2947317.

[37] E. Kaplan and C. Hegarty, "Interference, multipath, and scintillation," in *Understanding GPS: Principles and Applications*, 2nd ed. Norwood, MA, USA: Artech House, 2006, pp. 279–285.

[38] P. D. Groves, "Mitigation of multipath interference and nonline-of-sight reception," in *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, 2nd ed. London, U.K.: Artech House, 2013, pp. 458–461.

[39] Y. Li, Z. Jiang, O. M. O. Osman, X. Han, and J. Yin, "Mixed norm constrained sparse APA algorithm for satellite and network echo channel estimation," *IEEE Access*, vol. 6, pp. 65901–65908, 2018.

[40] L. Cheng, K. Wang, M. F. Ren, and X. Y. Xu, "Comprehensive analysis of multipath estimation algorithms in the framework of information theoretic learning," *IEEE Access*, vol. 6, pp. 5521–5530, 2018.

[41] M. Spangenberg, V. Calmettes, O. Julien, J.-Y. Tourneret, and G. Duchateau, "Detection of variance changes and mean value jumps in measurement noise for multipath mitigation in urban navigation," *Navigation*, vol. 57, no. 1, pp. 35–52, Mar. 2010.

[42] C. Cheng, J.-Y. Tourneret, Q. Pan, and V. Calmettes, "Detecting, estimating and correcting multipath biases affecting GNSS signals using a marginalized likelihood ratio-based method," *Signal Process.*, vol. 118, pp. 221–234, Jan. 2016.

[43] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*. Lincoln, MA, USA: Ganga-Jamuna Press, 2001.

[44] A. Giremus, J.-Y. Tourneret, and V. Calmettes, "A particle filtering approach for joint Detection/Estimation of multipath effects on GPS measurements," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1275–1285, Apr. 2007.

[45] C. Zhang, J. Bütepage, H. Kjellström, and S. Mandt, "Advances in variational inference," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 2008–2026, Aug. 2019.

CHENG CHENG was born in Xi'an, China, in 1986. He received the B.S. and M.S. degrees from Northwestern Polytechnical University, Xi'an, in 2008 and 2011, respectively, and the Ph.D. degree from the Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France, in 2015. He is currently an Assistant Professor with the School of Astronautics, Northwestern Polytechnical University. His research interests include Bayesian filtering for various navigation systems and aerospace engineering.

JEAN-YVES TOURNERET (Fellow, IEEE) received the Ingénieur degree in electrical engineering from the Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIHT) de Toulouse, in 1989, and the Ph.D. degree from the National Polytechnic Institute, Toulouse, in 1992. He is currently a Professor with the University of Toulouse (ENSEEIHT) and a member of the IRIT Laboratory (UMR 5505 of the CNRS). His research activities are centered around statistical signal and image processing with a particular interest to Bayesian and Markov chain Monte Carlo (MCMC) methods. He has been involved in the organization of several conferences, including the European Conference on Signal Processing EUSIPCO'02 (a Program Chair), the International Conference ICASSP'06 (plenaries), the Statistical Signal Processing Workshop SSP'12 (international liaisons), the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2013 (local arrangements), the Statistical Signal Processing Workshop SSP'2014 (special sessions), the Workshop On Machine Learning For Signal Processing MLSP'2014 (special sessions). He was the General Chair of the CIMI workshop on optimization and statistics in image processing hole in Toulouse, in 2013 (with F. Malgouyres and D. Kouamé) and of the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP, in 2015 (with P. Djuric) and in 2019 (with D. Brie). He was a member of different technical committees, including the Signal Processing Theory and Methods (SPTM) Committee of the IEEE Signal Processing Society, from 2001 to 2007 and 2010 to 2015, and the EURASIP SAT committee on Theoretical and Methodological Trends in Signal Processing, from 2015 to 2019. In 2019, he joined the board of directors of EURASIP. He has been serving as an Associate Editor for the IEEE Transactions on Signal Processing, from 2008 to 2011 and 2015 to 2019, and the *EURASIP Journal on Signal Processing*, from 2013 to 2019.

XIAODONG LU was born in Xi'an, China, in 1979. He received the B.S., M.S., and Ph.D. degrees from Northwestern Polytechnical University, Xi'an, in 2000, 2003, and 2006, respectively. He is currently an Associate Professor with the School of Astronautics, Northwestern Polytechnical University. His research interests include sensor data fusion, integrated navigation, and multitarget cooperative detection systems.