



## Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is a publisher's version published in: <https://oatao.univ-toulouse.fr/27093>

**Official URL** : <https://doi.org/10.1109/TSP.2012.2218811>

### To cite this version :

Closas, Pau and Fernández-Prades, Carles and Vilà-Valls, Jordi Multiple Quadrature Kalman Filtering. (2012) IEEE Transactions on Signal Processing, 60 (12). 6125-6137. ISSN 1053-587X

Any correspondence concerning this service should be sent to the repository administrator:

[tech-oatao@listes-diff.inp-toulouse.fr](mailto:tech-oatao@listes-diff.inp-toulouse.fr)

RESEARCH

Open Access



# NLOS mitigation in indoor localization by marginalized Monte Carlo Gaussian smoothing

Jordi Vilà-Valls<sup>1\*</sup> and Pau Closas<sup>2</sup>

## Abstract

One of the main challenges in indoor time-of-arrival (TOA)-based wireless localization systems is to mitigate non-line-of-sight (NLOS) propagation conditions, which degrade the overall positioning performance. The positive skewed non-Gaussian nature of TOA observations under LOS/NLOS conditions can be modeled as a heavy-tailed skew  $t$ -distributed measurement noise. The main goal of this article is to provide a robust Bayesian inference framework to deal with target localization under NLOS conditions. A key point is to take advantage of the conditionally Gaussian formulation of the skew  $t$ -distribution, thus being able to use computationally light Gaussian filtering and smoothing methods as the core of the new approach. The unknown non-Gaussian noise latent variables are marginalized using Monte Carlo sampling. Numerical results are provided to show the performance improvement of the proposed approach.

**Keywords:** Robust Bayesian inference, Gaussian filtering and smoothing, NLOS mitigation, Skew  $t$ -distributed measurement noise, Indoor localization, Monte Carlo integration

## 1 Introduction

The knowledge of position is ubiquitous in many applications and services, playing an important role. The widely diffused Global Navigation Satellite System (GNSS) offers a worldwide service coverage due to a network of dedicated satellites [1]. GNSS is recognized to be the *de facto* system in outdoor environments when it is available. Under the assumption that its reception is not obstructed or jammed [2–4], there is no doubt that GNSS is the main enabler for location-based services (LBS). One of such situations is indoor positioning and tracking, where satellite signals are hardly useful (unless extremely large integration times are considered). In indoor scenarios, a plethora of alternative and complementary technologies can be considered [1, 5, 6].

We are interested in a particular propagation phenomena encountered in most positioning technologies (both outdoor and indoor), known as non-line-of sight (NLOS). It is one of the most challenging problems for tracking. Particularly, when considering time-of-arrival (TOA) measurements as range estimates, the measured

distance can be severely degraded. These ranges are typically positively biased with respect to the true distances, therefore seen as outliers at the receiver. It is of interest to develop NLOS mitigation techniques, providing enhanced robustness to tracking methods based on TOA measurements [6].

In general, the problem under study concerns the derivation of new robust methods to solve the Bayesian filtering and smoothing problem in challenging applications such as the LOS/NLOS propagation conditions in indoor localization systems. The state-space models (SSM) of interest are expressed as

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{u}_k, \quad \mathbf{u}_k \sim \mathcal{N}(0, \mathbf{Q}_k), \quad (1)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k, \quad \mathbf{n}_k \sim \mathcal{ST}(\boldsymbol{\phi}_k), \quad (2)$$

where  $\mathbf{x}_k \in \mathbb{R}^{n_x}$  and  $\mathbf{y}_k \in \mathbb{R}^{n_y}$  are the hidden states of the system and measurements at time  $k$ .  $\mathbf{f}_{k-1}(\cdot)$  and  $\mathbf{h}_k(\cdot)$  are known to be the possibly nonlinear functions of the state; and both process and observation noises,  $\mathbf{u}_k$  and  $\mathbf{n}_k$ , assumed to be mutually independent. In real-life applications, we may not have a complete knowledge of the system conditions, thus the measurement noise statistics are assumed to be unknown to a certain extent. In contrast, we consider a known process noise covariance  $\mathbf{Q}_k$ . Regarding the measurement noise, we assume that it is

\*Correspondence: jvila@cttc.cat

<sup>1</sup>Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Castelldefels, Barcelona, Spain

Full list of author information is available at the end of the article

distributed according to a parametric heavy-tailed skew  $t$ -distribution,  $\mathbf{n}_k \sim \mathcal{ST}(\boldsymbol{\phi}_k)$ , with  $\boldsymbol{\phi}_k$  representing the set of possibly unknown parameters of the non-Gaussian distribution. The probability density function (pdf) of the univariate skew  $t$  distribution of interest can be written as [19]

$$\mathcal{ST}(z; \mu, \sigma^2, \lambda, \nu) = 2\mathcal{T}(z; \mu, \lambda^2 + \sigma^2, \nu) T(\tilde{z}; 0, 1, \nu + 1),$$

with  $\mu \in \mathbb{R}$ ,  $\sigma^2 \in \mathbb{R}^+$ ,  $\lambda \in \mathbb{R}$ , and  $\nu \in \mathbb{R}^+$ , referring to the distribution location, scale, skewness, and degrees of freedom, respectively.  $\mathcal{T}(z; \mu, \sigma^2, \nu)$  is the pdf of the Student's  $t$  distribution,

$$\mathcal{T}(z; \mu, \sigma^2, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sigma\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{(z-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}},$$

with  $\Gamma(\cdot)$  the gamma function.  $T(\tilde{z}; 0, 1, \nu)$  is the cumulative distribution function (CDF) of the Student's  $t$  distribution with  $\nu$  degrees of freedom and

$$\tilde{z} = \frac{(z-\mu)\lambda}{\sigma} \sqrt{\frac{\nu+1}{\nu(\lambda^2 + \sigma^2) + (z-\mu)^2}}.$$

Notice that the standard Student  $t$  distribution is  $\mathcal{T}(z; \mu, \sigma^2, \nu) = \mathcal{ST}(z; \mu, \sigma^2, \lambda = 0, \nu)$ , the skew normal pdf is  $\mathcal{SN}(z; \mu, \sigma^2, \lambda) = \mathcal{ST}(z; \mu, \sigma^2, \lambda, \nu \rightarrow \infty)$ , and  $\mathcal{N}(z; \mu, \sigma^2) = \mathcal{ST}(z; \mu, \sigma^2, \lambda = 0, \nu \rightarrow \infty)$  the normal distribution.

### 1.1 State-of-the-art

The skew  $t$  distribution has been recently shown to provide a reasonable fit to realistic indoor TOA measurements. For instance, characterizing range measures in NLOS conditions in ultra-wideband (UWB) localization [7] or in multipath channels when ranging is computed with long-term evolution (LTE) networks [8]. Interestingly, this distribution allows a Gaussian mean-scale mixture (GMSM) representation, which implies that the distribution can be reformulated as hierarchically (conditionally) Gaussian [9, 10]. Mathematically, if we have the skew  $t$ -distributed random variable  $z \sim \mathcal{ST}(z; \mu, \sigma^2, \lambda, \nu)$ , then we can write [41]

$$z|\gamma, \tau \sim \mathcal{N}(z; \mu + \lambda\gamma, \tau^{-1}\sigma^2), \quad (3)$$

with  $\tau \sim \mathcal{G}(\tau; \frac{\nu}{2}, \frac{\nu}{2})$ ,  $\gamma|\tau \sim \mathcal{N}_+(\gamma; 0, \tau^{-1})$ , and  $\mathcal{N}_+(\cdot)$  and  $\mathcal{G}(\cdot)$  as the positive truncated normal and gamma distributions. This is a key point in our problem formulation, because under the knowledge of the noise parameters (i.e.,  $\mu, \sigma^2, \lambda, \nu, \gamma$  and  $\tau$  in (3)), both the conditional marginal filtering and smoothing posterior distributions of the states,  $p(\mathbf{x}_k|\mathbf{y}_{1:k})$  and  $p(\mathbf{x}_k|\mathbf{y}_{1:N})$ , turn to be Gaussian and thus we are able to use computationally light Gaussian smoothing methods to infer the states of the system.

In the literature, some contributions dealing with conditionally Gaussian SSMs corrupted by both heavy-tailed symmetric and skewed noise distributions were proposed. A particle filter (PF) solution for linear SSMs in symmetric  $\alpha$ -stable ( $S\alpha S$ ) noise was presented in [11]. This idea was further explored in [12] for nonlinear systems and generalized to other symmetric distributions in [13]. The key idea was to take advantage of the conditionally Gaussian form and use a sigma-point Gaussian filter [14, 15] for the nonlinear state estimation. A robust filtering variational Bayesian (VB) approach was considered for linear systems in [16] and further extended to nonlinear SSMs in [17] considering a symmetric Student  $t$  measurement noise. But symmetric distributions may not always be appropriate to characterize the system noise. Recently, two interesting approaches to deal with linear SSMs under skewed noise were proposed, the first one uses a marginalized PF [18] and the other considers a VB solution [7, 19]. It is important to point out that (i) these contributions deal with either nonlinear systems corrupted by symmetric distributed noises or linear SSMs under skewed noise and (ii) the core of these methods use standard Bayesian filtering algorithms, then the smoothing problem needs to be further analyzed within this context.

Related to the problem under study, it is worth saying that several contributions deal with the filtering problem in nonlinear/non-Gaussian SSM under model uncertainty using sequential Monte Carlo (SMC) methods, for instance, joint state and parameter estimation solutions [20], model selection strategies using interacting parallel PFs [21, 22], or model information fusion within the SMC formulation [23]. The main drawback of SMC methods is their high computational complexity and the curse-of-dimensionality [24]. That is the reason why we propose to take advantage of the underlying conditional Gaussian nature of the problem and use more efficient methods in this context.

### 1.2 Contributions

The main contributions of the article, which generalize the preliminary results in [25], are summarized as:

- New Bayesian filtering and smoothing-based solutions for SSMs corrupted by parametric heavy-tailed skewed measurement noise
- Marginalization of the unknown non-Gaussian noise latent variables by Monte Carlo integration
- TOA-based robust target tracking, where the LOS/NLOS propagation is modeled using a skew  $t$ -distributed measurement noise. Whereas a Gaussian filter and smoother deals with the nonlinear state estimation problem, the time-varying skew  $t$  distribution parameters are marginalized via Monte Carlo sampling

The article is organized as follows: first, we provide a discussion on Gaussian filtering and smoothing in nonlinear/Gaussian systems, together with the sigma-point-based approximation of the multidimensional integrals in the conceptual solution, being computationally more efficient than SMC methods under the Gaussian assumption; then, we provide the conditionally Gaussian formulation of the measurement noise and a method to deal with the unknown non-Gaussian noise latent variables, and finally, we propose a NLOS indoor localization solution, based on the Gaussian smoother and the sequential noise latent variables marginalization. Numerical results are provided in realistic scenarios using UWB signals.

## 2 Gaussian filtering and smoothing

This section reviews the general filtering and smoothing solutions in the case of nonlinear/Gaussian systems, this material corresponds to Sections 2.1 and 2.2, respectively. Then, in Section 2.3, we provide the implementation details when sigma-points are used to solve the filtering/smoothing equations. Notice that when the system is linear/Gaussian, the optimal solutions are given by the standard Kalman filter (KF) [35] and Kalman smoother (KS) [36], and for general nonlinear/non-Gaussian systems, one should consider more sophisticated SMC techniques [29].

For the formal derivation of the Gaussian filter/smoothing, we assume that the measurement noise in (2) is zero-mean Gaussian with known covariance  $\mathbf{R}_k$ . Later, in Section 3, we discuss how the method can be used in the context of conditionally Gaussian models.

### 2.1 Bayesian Gaussian filtering

From a theoretical point of view, all necessary information to infer information of the unknown states resides in the marginal posterior distribution of the states,  $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ . Thus, the Bayesian filtering problem is one of evaluating this distribution. It can be recursively computed [26] in two steps: (1) *prediction* of  $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$  using the prior information and the previous filtering distribution and (2) *update* with new measurements  $\mathbf{y}_k$  to obtain.

The recursive solution provides an estimation framework that is optimal in the Bayesian sense, that is, the characterization of the posterior distribution allows us to compute the minimum mean-squared error (MMSE), the maximum a posteriori (MAP) or the median of the posterior (minimax) estimators, addressing optimality in many senses. The multidimensional integrals in the *prediction* and *update* steps are analytically intractable in the general case. Actually, there are few cases where the optimal Bayesian recursion can be analytically solved. This is the case of linear/Gaussian models, where the KF yields to the optimal solution [27]. In more general models, one

must resort to suboptimal algorithms. A plethora of methods can be found in the literature [28]. A popular tool are particle filters (PF) [29–32], a set of simulation-based methods which are applicable in nonlinear/non-Gaussian setups. Under the Gaussian assumption of interest, the quadrature KF (QKF) [14, 15, 33] and cubature KF (CKF) [34] are typically the methods of choice. In this case, the marginal predictive and posterior distributions are

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \Sigma_{x,k|k-1}), \tag{4}$$

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \Sigma_{x,k|k}). \tag{5}$$

In the prediction step, we compute the marginal predictive distribution mean and covariance as <sup>1</sup>

$$\hat{\mathbf{x}}_{k|k-1} = \int \mathbf{f}(\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1}, \tag{6}$$

$$\begin{aligned} \Sigma_{x,k|k-1} &= \int \mathbf{f}^2(\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1} \\ &\quad - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^T + \mathbf{Q}_k. \end{aligned} \tag{7}$$

In the update step, the mean and covariance of the marginal posterior are given by the KF Equations [35]

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}), \tag{8}$$

$$\Sigma_{x,k|k} = \Sigma_{x,k|k-1} - \mathbf{K}_k\Sigma_{y,k|k-1}\mathbf{K}_k^T, \tag{9}$$

where the Kalman gain is  $\mathbf{K}_k = \Sigma_{xy,k|k-1}\Sigma_{y,k|k-1}^{-1}$ . The predicted measurement and both innovation and cross-covariance matrices are computed as

$$\hat{\mathbf{y}}_{k|k-1} = \int \mathbf{h}(\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_k, \tag{10}$$

$$\begin{aligned} \Sigma_{y,k|k-1} &= \int \mathbf{h}^2(\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_k \\ &\quad - \hat{\mathbf{y}}_{k|k-1}\hat{\mathbf{y}}_{k|k-1}^T + \mathbf{R}_k, \end{aligned} \tag{11}$$

$$\begin{aligned} \Sigma_{xy,k|k-1} &= \int \mathbf{x}_k\mathbf{h}^T(\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_k \\ &\quad - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{y}}_{k|k-1}^T. \end{aligned} \tag{12}$$

The problem reduces to the approximation of these integrals.

### 2.2 Gaussian smoothing

In the previous section, we summarized the general Gaussian Bayesian filtering solution but sometimes it may be interesting to obtain an estimate of the smoothing posterior and not its filtering counterpart. In the problem

under study, we consider a *forward-backward* smoother formulation [36] to obtain the marginal smoothing posterior,  $p(\mathbf{x}_k|\mathbf{y}_{1:N})$ ,

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{y}_{1:N}) &= \int p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{y}_{1:N})d\mathbf{x}_{k+1} \\ &= \int p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:k})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:N})d\mathbf{x}_{k+1} \\ &= \underbrace{p(\mathbf{x}_k|\mathbf{y}_{1:k})}_{\text{filtering pdf}} \int \frac{p(\mathbf{x}_{k+1}|\mathbf{y}_{1:N})p(\mathbf{x}_{k+1}|\mathbf{x}_k)}{\underbrace{p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}_{\text{predictive pdf}}}d\mathbf{x}_{k+1}, \end{aligned}$$

where we used the state that is Markovian and then

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:N}) &= p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:k}) \\ &= \frac{p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{y}_{1:k})}{p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})} \\ &= \frac{p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k})}{p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}. \end{aligned}$$

The *forward-backward* smoothing [36] performs two filtering passes, that is, first a standard *forward filtering* from time  $k = 1$  to  $N$ , and then, the *backward filtering* from  $k = N$  to  $1$ , backwards in time. Notice that the predictive and filtering distributions may be obtained from the standard Bayesian filtering solution. At time  $k$ , if we consider that we know the filtering distribution,  $\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \Sigma_{x,k|k})$ , and the predictive distribution,  $\mathcal{N}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \Sigma_{x,k+1|k})$ , from the *forward filtering*, together with the smoothed density at  $k + 1$ ,  $\mathcal{N}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \Sigma_{x,k+1|N})$ , because the smoother is running backwards, then the analytical solution to the marginal smoothing posterior is obtained as follows: using the Markovian properties of states, we have that  $p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:N}) = p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:k})$ , and then we can obtain the conditional smoothing distribution of  $\mathbf{x}_k$  as

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:N}) = \frac{p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{y}_{1:k})}{p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}, \quad (13)$$

with

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{y}_{1:k}) &= p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k}), \\ p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k}) &= \int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k})d\mathbf{x}_k. \end{aligned}$$

The joint smoothing distribution  $p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{y}_{1:N})$  is

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:N})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:N}), \quad (14)$$

which can be used to obtain the smoothing distribution by marginalization over  $\mathbf{x}_{k+1}$ ,

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{y}_{1:N}) &= \int p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{1:N}) \\ &\quad \times p(\mathbf{x}_{k+1}|\mathbf{y}_{1:N})d\mathbf{x}_{k+1}. \end{aligned} \quad (15)$$

Under the Gaussian assumption, the problem is to recursively obtain the mean and covariance of the

Gaussian marginal smoothing posterior distribution, which is given by [37, 38]

$$p(\mathbf{x}_k|\mathbf{y}_{1:N}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|N}, \Sigma_{x,k|N}), \quad (16)$$

$$\hat{\mathbf{x}}_{k|N} = \hat{\mathbf{x}}_{k|k} + \mathbf{D}_k (\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k}), \quad (17)$$

$$\Sigma_{x,k|N} = \Sigma_{x,k|k} + \mathbf{D}_k (\Sigma_{x,k+1|N} - \Sigma_{x,k+1|k}) \mathbf{D}_k^\top,$$

with  $\mathbf{D}_k = \Sigma_{k,k+1|k} \Sigma_{x,k+1|k}^{-1}$  and  $\Sigma_{k,k+1|k}$  referring to the cross-covariance between  $\mathbf{x}_k$  and  $\mathbf{x}_{k+1}$ . Note that in practice we do not require the computation of the smoothing estimation error covariance,  $\Sigma_{x,k|N}$ , for the smoother recursion. However, it is useful in order to have a measure of the smoothing uncertainty. The smoother gain  $\mathbf{D}_k$  can be easily obtained from the standard forward filtering pass, therefore adding very few extra computation.

### 2.3 Sigma-point Gaussian filtering and smoothing

An appealing class of filters and smoothers within the nonlinear Gaussian framework are the sigma-point Gaussian filters (SPGF) [14, 15, 34, 39, 40] and smoothers (SPGS) [37, 38], a family of derivative-free algorithms which are based on a weighted sum of function values at specified (i.e., deterministic) points within the domain of integration, as opposite to the stochastic sampling performed by particle filtering methods. The idea is to use a set of so-called sigma-points to efficiently characterize the propagation of the normal distribution over the nonlinear system. In the sequel, we detail the formulation of such approximation and how it can be used to perform filtering or smoothing.

#### 2.3.1 Filtering

Consider a set of sigma-points,  $\{\boldsymbol{\xi}_i, \omega_i\}_{i=1, \dots, L_M}$ . Then, construct the transformed set which captures the mean and covariance of the posterior distribution,  $\mathbf{x}_{i,k-1|k-1} = \mathbf{S}_{x,k-1|k-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k-1|k-1}$ , with  $\Sigma_{x,k-1|k-1} = \mathbf{S}_{x,k-1|k-1} \mathbf{S}_{x,k-1|k-1}^\top$ . The integrals in the prediction step can be approximated as

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{f}(\mathbf{x}_{i,k-1|k-1}), \\ \Sigma_{x,k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{f}^2(\mathbf{x}_{i,k-1|k-1}) - (\hat{\mathbf{x}}_{k|k-1})^2 + \mathbf{Q}_k. \end{aligned} \quad (18)$$

In the following update step, first compute the transformed set to capture the mean and covariance of the

predictive marginal distribution,  $\mathbf{x}_{i,k|k-1} = \mathbf{S}_{x,k|k-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k|k-1}$ , with  $\Sigma_{x,k|k-1} = \mathbf{S}_{x,k|k-1} \mathbf{S}_{x,k|k-1}^\top$ . Then, we approximate the integrals of interest as,

$$\begin{aligned} \hat{\mathbf{y}}_{k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{h}(\mathbf{x}_{i,k|k-1}), \\ \Sigma_{y,k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{h}^2(\mathbf{x}_{i,k|k-1}) - (\hat{\mathbf{y}}_{k|k-1})^2 + \mathbf{R}_k, \\ \Sigma_{xy,k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{x}_{i,k|k-1} \mathbf{h}(\mathbf{x}_{i,k|k-1})^\top \\ &\quad - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{y}}_{k|k-1}^\top. \end{aligned} \quad (19)$$

### 2.3.2 Smoothing

The smoothed state (17) is obtained using the predicted filtering and smoothing states,  $\hat{\mathbf{x}}_{k+1|k}$  and  $\hat{\mathbf{x}}_{k+1|N}$ , respectively. Define again a set of sigma-points and weights,  $\{\boldsymbol{\xi}_i, \omega_i\}_{i=1, \dots, L_M}$ , and the transformed set which captures the corresponding mean and covariance,  $\mathbf{x}_{i,k|k} = \mathbf{S}_{x,k|k} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k|k}$ . Use this transformed sigma-points to estimate the predicted subspace state, its prediction error covariance, and the cross-covariance as

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= \sum_{i=1}^{L_M} \omega_i \mathbf{f}(\mathbf{x}_{i,k|k}), \\ \Sigma_{x,k+1|k} &= \sum_{i=1}^{L_M} \omega_i \mathbf{f}^2(\mathbf{x}_{i,k|k}) - (\hat{\mathbf{x}}_{k+1|k})^2 + \mathbf{Q}_k, \\ \Sigma_{k,k+1|k} &= \sum_{i=1}^{L_M} \omega_i \mathbf{x}_{i,k|k} \mathbf{f}(\mathbf{x}_{i,k|k}) - \hat{\mathbf{x}}_{k|k} \hat{\mathbf{x}}_{k+1|k}^\top. \end{aligned} \quad (20)$$

Finally, estimate the smoothed subspace and covariance as

$$\begin{aligned} \hat{\mathbf{x}}_{k|N} &= \hat{\mathbf{x}}_{k|k} + \mathbf{D}_k (\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k}), \\ \Sigma_{x,k|N} &= \Sigma_{x,k|k} + \mathbf{D}_k (\Sigma_{x,k+1|N} - \Sigma_{x,k+1|k}) \mathbf{D}_k^\top, \end{aligned} \quad (21)$$

with  $\mathbf{D}_k = \Sigma_{k,k+1|k} \Sigma_{x,k+1|k}^{-1}$ . Notice that the smoother gain can be embedded into the prediction step of the forward filtering, then only the last step is performed in the backward recursion. At time  $k = N$ , both filtering and smoothing estimates are the same, then the backward pass runs from time  $N - 1$  to 1. Compared to the filtering process, implementation of the smoothing solution only impacts in having additional steps 6 and 12 in Algorithm 1, where we use the notation  $\Sigma_x = \mathbf{S}_x \mathbf{S}_x^\top$  for the factorized covariances.

## 3 Hierarchically Gaussian measurement noise formulation

In the previous Section 2, we assumed a Gaussian measurement noise with known covariance matrix. But in

### Algorithm 1 Sigma-point Kalman filter/smoothing

**Require:**  $\hat{\mathbf{y}}_{1:N}$ ,  $\hat{\mathbf{x}}_0$ ,  $\mathbf{S}_{x,0|0}$ ,  $\mathbf{Q}_k$ ,  $\mathbf{R}_k$ ,  $\{\boldsymbol{\xi}_i, \omega_i\}_{i=1}^{L_M}$ .

- 1: Set  $k \leftarrow 1$
- 2: **Forward filtering** (from  $k = 1$  to  $N$ )
- 3: **Time update**
- 4: Factorize covariance and evaluate sigma-points:
 
$$\mathbf{x}_{i,k-1|k-1} = \mathbf{S}_{x,k-1|k-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k-1|k-1}, \quad i = 1, \dots, L_M$$
- 5: Estimate the predicted subspace state and covariance:
 
$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{f}(\mathbf{x}_{i,k-1|k-1})$$

$$\Sigma_{x,k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{f}^2(\mathbf{x}_{i,k-1|k-1}) - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{x}}_{k|k-1}^\top + \mathbf{Q}_k$$
- 6: Estimate the cross-covariance and smoother gain:
 
$$\Sigma_{k-1,k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{x}_{i,k-1|k-1}$$

$$\quad \times \mathbf{f}(\mathbf{x}_{i,k-1|k-1}) - \hat{\mathbf{x}}_{k-1|k-1} \hat{\mathbf{x}}_{k|k-1}^\top$$

$$\mathbf{D}_{k-1} = \Sigma_{k-1,k|k-1} (\Sigma_{x,k|k-1})^{-1}$$
- 7: **Measurement update**
- 8: Factorize covariance and evaluate the sigma-points:
 
$$\mathbf{x}_{i,k|k-1} = \mathbf{S}_{x,k|k-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k|k-1}, \quad i = 1, \dots, L_M$$
- 9: Predicted measurement, cross and innovation covariance:
 
$$\hat{\mathbf{y}}_{k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{h}(\mathbf{x}_{i,k-1|k-1})$$

$$\Sigma_{xy,k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{f}(\mathbf{x}_{i,k-1|k-1})$$

$$\quad \times \mathbf{h}(\mathbf{x}_{i,k-1|k-1})^\top - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{y}}_{k|k-1}^\top$$

$$\Sigma_{y,k|k-1} = \sum_{i=1}^{L_M} \omega_i \mathbf{h}^2(\mathbf{x}_{i,k-1|k-1}) - (\hat{\mathbf{y}}_{k|k-1})^2 + \mathbf{R}_k$$
- 10: Estimate the updated subspace state and error covariance:
 
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})$$

$$\Sigma_{x,k|k} = \Sigma_{x,k|k-1} - \mathbf{K}_k \Sigma_{y,k|k-1} (\mathbf{K}_k)^\top$$
 with  $\mathbf{K}_k = \Sigma_{xy,k|k-1} \Sigma_{y,k|k-1}^{-1}$ .
- 11: **Backward filtering** (from  $k = (N - 1)$  to 1)
- 12: Estimate the smoothed subspace and covariance:
 
$$\hat{\mathbf{x}}_{k|N} = \hat{\mathbf{x}}_{k|k} + \mathbf{D}_k (\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k})$$

$$\Sigma_{x,k|N} = \Sigma_{x,k|k} + \mathbf{D}_k (\Sigma_{x,k+1|N} - \Sigma_{x,k+1|k}) \mathbf{D}_k^\top.$$

challenging applications such as the NLOS propagation conditions of interest here, the Gaussian assumption does not hold and noise parameters may be unknown to a certain extent. In such scenarios, one may have outliers or impulsive behaviors that produce biased estimates, for instance, under NLOS conditions the receiver is likely to estimate distances to the anchors larger than the true ones [6]; therefore, we must account for more accurate observation models.

In general, these non-Gaussian behaviors can be effectively characterized by parametric heavy-tailed and positive-skewed noise distributions. It has been recently shown experimentally that TOA-based positioning under NLOS conditions [7] and multipath ranging error distributions in LTE networks [8] can be well approximated by a skew  $t$ -distribution [9]. Taking into account the problem at hand, we are interested in measurement models with independent observation components and measurement noise models where the noise components are independently univariate skew  $t$ -distributed

$$\mathbf{y}_k = [\mathcal{Y}_{k,1}, \dots, \mathcal{Y}_{k,n_y}]^\top, \quad (22)$$

$$y_{k,i} = \mathbf{h}_{k,i}(\mathbf{x}_k) + n_{k,i}, \quad (23)$$

$$\mathbf{n}_k = [n_{k,1}, \dots, n_{k,n_y}]^\top, \quad (24)$$

$$n_{k,i} \sim \mathcal{ST}(n_{k,i}; \mu, \sigma^2, \lambda, \nu), \quad (25)$$

with  $\mathcal{ST}(n_{k,i}; \mu, \sigma^2, \lambda, \nu)$  defined in Section 1.

A key point on the problem formulation is to take advantage of the hierarchically (conditionally) Gaussian formulation of the measurement noise distribution. The hierarchical Gaussian representation of the skew  $t$ -distribution is written as [41]

$$n_{k,i} | \gamma_{k,i}, \tau_{k,i} \sim \mathcal{N}(n_{k,i}; \mu + \lambda \gamma_{k,i}, \tau_{k,i}^{-1} \sigma^2), \quad (26)$$

$$\gamma_{k,i} | \tau_{k,i} \sim \mathcal{N}_+(\gamma_{k,i}; 0, \tau_{k,i}^{-1}), \quad (27)$$

$$\tau_{k,i} \sim \mathcal{G}(\tau_{k,i}; \frac{\nu}{2}, \frac{\nu}{2}). \quad (28)$$

While  $\tau_{k,i}$  controls the heavy-tailed behavior,  $\gamma_{k,i}$  controls the skewness of the distribution.

We can define the vector with  $2 \times n_y$  noise distribution latent variables,  $\boldsymbol{\phi}_k = \{\gamma_{k,i}, \tau_{k,i}\}_{i=1, \dots, n_y}$ , where we omit the dependence with respect to the hyperparameters (i.e.,  $\mu, \sigma^2, \lambda, \nu$ ) for the sake of clarity.

The measurement noise in (24) can be written as

$$\mathbf{n}_k | \boldsymbol{\phi}_k \sim \mathcal{N}(\mathbf{n}_k; \mathbf{m}_k(\boldsymbol{\phi}_k), \mathbf{R}_k(\boldsymbol{\phi}_k)), \quad (29)$$

where  $[\mathbf{m}_k(\boldsymbol{\phi}_k)]_i = \mu + \lambda \gamma_{k,i}$  and  $[\mathbf{R}_k(\boldsymbol{\phi}_k)]_{i,i} = \tau_{k,i}^{-1} \sigma^2$ .

The distribution hyperparameters are application dependent and typically assumed a priori known. The standard Gaussian filter/smoothing in charge of the state estimation assumes a zero-mean Gaussian measurement noise with known parameters. In the skew  $t$ -distributed case, at every time step, the filter requires an estimate of the corresponding mean and covariance,  $\mathbf{m}_k(\boldsymbol{\phi}_k)$  and  $\mathbf{R}_k(\boldsymbol{\phi}_k)$ , respectively. In the following, we consider the marginalization of the noise latent variables in the general filter/smoothing formulation.

#### 4 Noise latent variables marginalization

In the problem of interest, the measurement noise is conditionally Gaussian with unknown noise latent variables. Therefore, the filtering/smoothing formulation in Section 2 must be modified to take such uncertainty into account. We assume known measurement noise distribution hyperparameters, and thus, we want to marginalize the state estimation with respect to the noise latent variables,  $\gamma_{k,i}$  and  $\tau_{k,i}$ . We can write the marginalized posterior distribution as

$$p(\mathbf{x}_k | \mathbf{y}_{1:k}) = \int p(\mathbf{x}_k | \boldsymbol{\phi}_k, \mathbf{y}_{1:k}) p(\boldsymbol{\phi}_k | \mathbf{y}_{1:k}) d\boldsymbol{\phi}_k.$$

Notice that the measurement noise parameters only affect the computation of the innovation in the measurement update of the filter/smoothing. Within this context, the predicted measurement is reformulated as

$$\begin{aligned} \hat{\mathbf{y}}_{k|k-1} &= \int (\mathbf{h}(\mathbf{x}_k) + \mathbf{n}_k(\boldsymbol{\phi}_k)) p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) d\mathbf{x}_k, \\ &= \int \mathbf{h}(\mathbf{x}_k) p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) d\mathbf{x}_k \\ &\quad + \int \mathbf{n}_k(\boldsymbol{\phi}_k) p(\boldsymbol{\phi}_k | \mathbf{y}_{1:k-1}) d\boldsymbol{\phi}_k, \end{aligned} \quad (30)$$

where the first term corresponds to the *standard* Gaussian case, and the second is a marginalization of the noise term over the last available distribution (i.e., we write  $\mathbf{n}_k(\boldsymbol{\phi}_k)$  to explicitly emphasize the dependency on the noise latent variables,  $\boldsymbol{\phi}_k$ ). Using a similar formulation, we can rewrite the innovation covariance as

$$\begin{aligned} \boldsymbol{\Sigma}_{y,k|k-1} &= \int \mathbf{h}^2(\mathbf{x}_k) p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) d\mathbf{x}_k \\ &\quad - (\hat{\mathbf{y}}_{k|k-1})^2 + \int \mathbf{n}_k(\boldsymbol{\phi}_k) \mathbf{n}_k^\top(\boldsymbol{\phi}_k) p(\boldsymbol{\phi}_k | \mathbf{y}_{1:k-1}) d\boldsymbol{\phi}_k. \end{aligned} \quad (31)$$

For the marginalization of the noise latent variables, a key point is to obtain the posterior distributions of  $\gamma_{k,i}$  and  $\tau_{k,i}$ . The joint posterior is given by

$$\begin{aligned} p(\gamma_{k,i}, \tau_{k,i} | \mathbf{y}_{1:k,i}) &= p(\gamma_{k,i}, \tau_{k,i} | y_{k,i}) \\ &\propto p(y_{k,i} | \gamma_{k,i}, \tau_{k,i}) p(\gamma_{k,i} | \tau_{k,i}) p(\tau_{k,i}). \end{aligned} \quad (32)$$

As the observation components are assumed to be independent, the likelihood function is

$$y_{k,i} | \gamma_{k,i}, \tau_{k,i} \sim \mathcal{N}(y_{k,i}; \mathbf{h}_{k,i}(\mathbf{x}_k) + \mu + \lambda \gamma_{k,i}, \tau_{k,i}^{-1} \sigma^2). \quad (33)$$

We can define a normalized observation

$$\tilde{y}_{k,i} \triangleq \frac{y_{k,i} - \mathbf{h}_{k,i}(\hat{\mathbf{x}}_{k|k-1}) - \mu}{\sigma}, \quad (34)$$

and  $\tilde{\gamma}_{k,i} \triangleq \lambda \gamma_{k,i} / \sigma$ . Then, we have the normalized likelihood is given by

$$p(\tilde{y}_{k,i} | \gamma_{k,i}, \tau_{k,i}) = \mathcal{N}(\tilde{y}_{k,i}; \tilde{\gamma}_{k,i}, \tau_{k,i}^{-1}). \quad (35)$$

Using the conjugate nature of the prior distributions [42], it is possible to obtain the analytical solution for the posterior of  $\gamma_{k,i}$  and  $\tau_{k,i}$ . In this case, we have from (28) that the a priori distributions are

$$p(\tilde{\gamma}_{k,i} | \tau_{k,i}) = \mathcal{N}(\tilde{\gamma}_{k,i}; \gamma_{k-1}, (\kappa_{k-1} \tau_{k,i})^{-1}), \quad (36)$$

$$p(\tau_{k,i}) = \mathcal{G}(\tau_{k,i}; \alpha_{k-1}, \beta_{k-1}), \quad (37)$$

with  $\gamma_0 = 0$ ,  $\kappa_0 = \sigma^2 / \lambda^2$ , and  $\alpha_0 = \beta_0 = \nu / 2$ . We are interested in updating with the new measurements to get the posterior distributions

$$p(\tilde{y}_{k,i}|\tau_{k,i}, \tilde{y}_{k,i}) = \mathcal{N}(\tilde{y}_{k,i}; \gamma_k, (\kappa_k \tau_{k,i})^{-1}), \quad (38)$$

$$p(\tau_{k,i}|\tilde{y}_{k,i}) = \mathcal{G}(\tau_{k,i}; \alpha_k, \beta_k), \quad (39)$$

with

$$\gamma_k = \frac{\kappa_{k-1}\gamma_{k-1} + \tilde{y}_{k,i}}{\kappa_{k-1} + 1}, \quad (40)$$

$$\kappa_k = \kappa_{k-1} + 1, \quad (41)$$

$$\alpha_k = \alpha_{k-1} + \frac{1}{2}, \quad (42)$$

$$\beta_k = \beta_{k-1} + \frac{\kappa_{k-1}(\tilde{y}_{k,i} - \gamma_{k-1})^2}{2(\kappa_{k-1} + 1)}, \quad (43)$$

from basic conjugate analysis results. Interestingly, the posteriors at  $k$  in (38) and (39) can be used as the priors in  $k + 1$  instead of (36) and (37). In this way, the algorithm is learning the environment as it progresses over time. However, given the assumed model, it is more meaningful to reset the prior at each time instant instead of sequentially using the latest posterior. The reason is that measurements are assumed independent, so there is no benefit in carrying out information from one time instant to the other. Instead, under these conditions, we suggest to use the values  $\gamma_0$ ,  $\kappa_0$ ,  $\alpha_0$ , and  $\beta_0$  at  $k - 1$  before updating the distribution with  $\tilde{y}_{k,i}$ . Sequential use of the posterior will be interesting when the generative model is known to have some memory.

In [25], we proposed to use a point estimate for  $\gamma_{k,i}$  and  $\tau_{k,i}$  from a single observation using their posterior marginals. The corresponding modes of these distributions were used as point estimates  $\hat{\gamma}_{k,i} = \frac{|\tilde{y}_{k,i}| \sigma}{2 \lambda}$  and  $\hat{\tau}_{k,i} = \frac{\alpha - 1}{\beta}$ , where we took into account that  $\hat{\gamma}_{k,i} \in \mathbb{R}^+$  by construction. In this contribution, instead of using a point estimate, we consider a Monte Carlo-based marginalization drawing  $L$  samples from the joint posterior given by (38) and (39). Using these distributions, we propose to compute the two integrals of interest as

$$\begin{aligned} \tilde{\mathbf{m}}_k &= \int \mathbf{n}_k(\phi_k) p(\phi_k | \mathbf{y}_{1:k-1}) d\phi_k \\ &\approx \frac{1}{L} \sum_{j=1}^L \mathbf{m}_k(\phi_k^{(j)}), \\ \tilde{\mathbf{R}}_k &= \int \mathbf{n}_k(\phi_k) \mathbf{n}_k^\top(\phi_k) p(\phi_k | \mathbf{y}_{1:k-1}) d\phi_k \\ &\approx \frac{1}{L} \sum_{j=1}^L \left( \mathbf{R}_k(\phi_k^{(j)}) + \mathbf{m}_k(\phi_k^{(j)}) \mathbf{m}_k^\top(\phi_k^{(j)}) \right), \end{aligned}$$

with  $\phi_k^{(j)}$  being random samples drawn from the joint posterior distribution of the noise latent variables,  $\phi_k$ . In practice, this can be easily implemented by first drawing a sample from (39) and then, using that sample, draw from (38). These expressions can be further expanded as follows

$$\tilde{\mathbf{m}}_k \approx \begin{bmatrix} \frac{1}{L} \sum_{j=1}^L (\mu + \lambda \gamma_{k,1}^{(j)}) \\ \vdots \\ \frac{1}{L} \sum_{j=1}^L (\mu + \lambda \gamma_{k,n_y}^{(j)}) \end{bmatrix}, \quad (44)$$

where  $\gamma_{k,i}^{(j)}$  are random samples drawn from the posterior of  $\gamma_{k,i}$ , and  $\tilde{\mathbf{R}}_k$  is approximated by a diagonal matrix, where the  $p$ -th element of the diagonal is

$$\frac{1}{L} \sum_{j=1}^L \left( (\tau_{k,p}^{(j)})^{-1} \sigma^2 + (\mu + \lambda \gamma_{k,p}^{(j)})^2 \right) \quad (45)$$

Finally, we have that the marginalized Monte Carlo sigma-point Gaussian filter and smoother (MSPGF/S) proposed in this contribution is given by Algorithm 1 with step 9 modified as

$$\begin{aligned} \hat{\mathbf{y}}_{k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{h}(\mathbf{x}_{i,k-1|k-1}) + \tilde{\mathbf{m}}_k, \\ \Sigma_{y,k|k-1} &= \sum_{i=1}^{L_M} \omega_i \mathbf{h}^2(\mathbf{x}_{i,k-1|k-1}) - (\hat{\mathbf{y}}_{k|k-1})^2 + \tilde{\mathbf{R}}_k. \end{aligned} \quad (46)$$

A further improvement of standard SPGF/S schemes comes from the fact that the filter should preserve the properties of a covariance matrix, namely, its symmetry and positive-definiteness. In practice, however, due to lack of arithmetic precision, numerical errors may lead to a loss of these properties. To circumvent this problem, a square-root filter can be considered to propagate the square root of the covariance matrix instead of the covariance itself [33, 34]. We propose to use square-root cubature and quadrature Kalman filters/smothers (named SCKF/S and SQKF/S, respectively) [38, 43] as the core implementation of the new square-root MSPGF/S. These methods resort to cubature [34] and Gauss-Hermite quadrature rules [15] to approximate the integrals in the optimal solution. While the SCKF/S uses  $L_c = 2n_x$  sigma-points, in the SQKF/S we have  $L_q = \alpha^{n_x}$ , where  $\alpha$  determines the number of sigma-points per dimension, which is typically set to  $\alpha = 3$ . A straightforward solution to avoid the exponential computational complexity increase of the standard QKF in high-dimensional systems is the use of sparse-grid quadrature rules, which reduce the computational complexity with negligible penalty in numerical accuracy [44, 45].

## 5 Application to indoor localization

### 5.1 SSM for the TOA-based localization problem

To show the performance of the proposed approach, we consider a TOA-based localization problem, where a set of  $N$  anchor nodes at known locations,  $\mathbf{x}_{k,i} = [x_{k,i}, y_{k,i}]^\top$ , provide range information. If we define the state to be



inferred as position and velocity components of the target,  $\mathbf{p}_k \triangleq (x_k, y_k)^\top$  and  $\mathbf{v}_k \triangleq (\dot{x}_k, \dot{y}_k)^\top$ , respectively, the observed range from each node  $i$  to the target, is modeled as  $\hat{\rho}_{k,i} = \rho_i(\mathbf{x}_k) + n_{k,i}$ ,  $i \in \{1, \dots, N\}$ , with  $n_{k,i}$  denoting the ranging error and  $\rho_i(\mathbf{x}_k) \triangleq \rho_{k,i} = \|\mathbf{x}_k - \mathbf{x}_{k,i}\|$  the true distance from the  $i$ -th node to the target node at time  $k$ . The complete measurement equation is given by

$$\begin{aligned} \boldsymbol{\rho}_k &= [\rho_{k,1}, \dots, \rho_{k,N}]^\top \\ &= \underbrace{\begin{pmatrix} \|\mathbf{x}_k - \mathbf{x}_{k,1}\| \\ \vdots \\ \|\mathbf{x}_k - \mathbf{x}_{k,N}\| \end{pmatrix}}_{\mathbf{h}_k(\mathbf{x}_k)} + \underbrace{\begin{pmatrix} n_{k,1} \\ \vdots \\ n_{k,N} \end{pmatrix}}_{\mathbf{n}_k}. \end{aligned} \quad (47)$$

In standard localization applications, the measurement noise is nominally distributed according to  $\mathbf{n}_t \sim \mathcal{N}(\mathbf{n}_t; \mathbf{0}, \sigma^2 \cdot \mathbf{I}_N)$ , where  $\sigma$  depends on the technology used to obtain the ranging estimates. In the case of UWB devices, this is typically on the order of 0.1 to 1 meter. But the Gaussian distribution does not capture the NLOS propagation conditions [7]; thus, we must account for more accurate measurement models such as the skew  $t$ -distribution introduced in the previous sections. Using the noise  $\mathbf{n}_k$  defined in (29), the measurement equation is defined as

$$\boldsymbol{\rho}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k, \quad n_{k,i} \sim \mathcal{ST}(n_{k,i}; \mu, \sigma^2, \lambda, \nu). \quad (48)$$

Considering the state  $\mathbf{x}_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^\top$ , the process equation is defined as a linear constant acceleration model

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\mathbf{u}_k, \quad \mathbf{u}_k \sim \mathcal{N}(\mathbf{u}_k; \mathbf{0}, \mathbf{Q}), \quad (49)$$

with

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{G} = \begin{pmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{pmatrix}.$$

The Gaussian process noise  $\mathbf{u}_k \sim \mathcal{N}(\mathbf{u}_k; \mathbf{0}, \sigma_u^2 \mathbf{I}_2)$  models an acceleration of  $\sigma_u$  m/s<sup>2</sup>.

### 5.2 Numerical results

For the numerical evaluation of the proposed method, the root mean square error (RMSE) of position is used as the measure of performance, which is obtained from 1000 Monte Carlo runs. The new method was validated in a realistic scenario composed of  $N = 6$  anchor nodes, circularly deployed in a  $40 \times 40$  m<sup>2</sup> area, and considering  $\sigma_u = 0.03$  m/s<sup>2</sup>. We compare the tracking performance obtained with four methods:

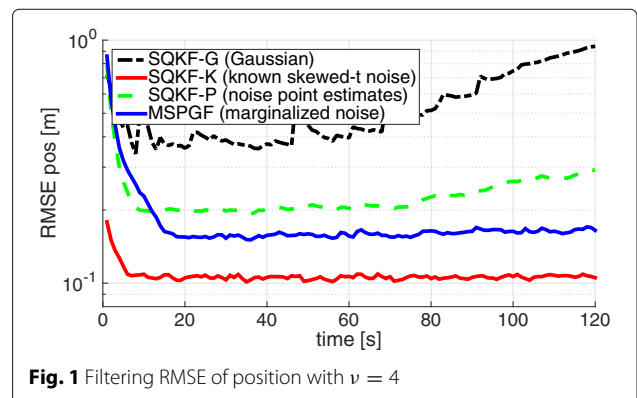
1. SQKF/S operating under the Gaussian assumption without accounting for the non-Gaussian nature of the measurement noise (SQKF/S-G).

2. SQKF/S using point estimates of the noise latent variables  $\boldsymbol{\psi}_{k,i}$  as proposed in [25] (SQKF/S-P).
3. New square-root SPGF/S-based solution with marginalized noise latent variables  $\boldsymbol{\phi}_k$  within the filter/smoothing formulation via Monte Carlo sampling (MSPGF/S).
4. A clairvoyant SQKF/S that knows exactly the realization of the latent variables  $\boldsymbol{\phi}_k$  at each instant  $k$  and thus can use  $\mathbf{m}_k(\boldsymbol{\phi}_k)$  and  $\mathbf{R}_k(\boldsymbol{\phi}_k)$ . This is the performance benchmark for the new methodology (SQKF/S-K).

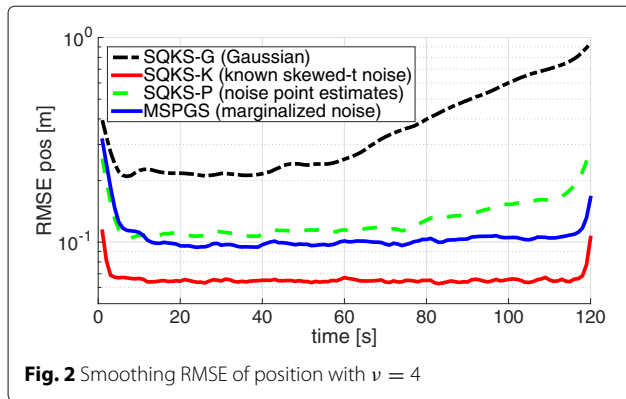
We also considered a sampling importance resampling PF with 81 particles (i.e., equivalent to the number of sigma-points in the SQKF/S), but as already shown in [46], the filter is in general not able to correctly localize the target (i.e., the filter diverges). Moreover, to obtain the same performance than the clairvoyant SQKF/S, we must consider a much larger number of particles. This is the reason why these results are not shown in the figures, since for the fair comparison in terms of number of particles the PF does not provide convergent result.

The proposed MSPGF/S can be implemented using cubature [34] and Gauss-Hermite approximations [15], then using respectively  $L_c = 2n_x = 8$  and  $L_q = \alpha^{n_x} = 81$  deterministic samples to approximate the integrals of the general solution. In the proposed indoor localization scenario, we tested both cubature and quadrature approximations, and the performance obtained was found strictly equivalent. In practice, the method of choice is the cubature-based solution, which has the lowest computational complexity.

Notice that all the methods consider known distribution hyperparameters, which are application dependent. We consider an UWB TOA-based indoor localization realistic scenario, with hyperparameters given in [7] and adjusted to match real data:  $\mu = -0.1$  m,  $\sigma = 0.3$  m, and  $\lambda = 0.6$  m and  $\nu = 4$ . The corresponding Gaussian approximation is given by  $\mu_G = 1.3$  and  $\sigma_G = 1.6$ .



**Fig. 1** Filtering RMSE of position with  $\nu = 4$



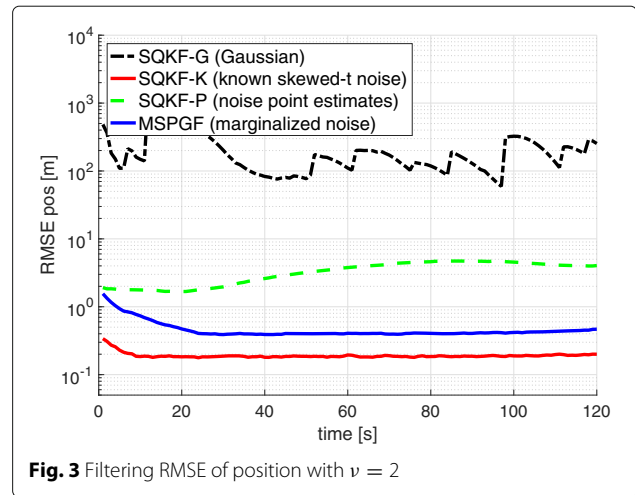
**Fig. 2** Smoothing RMSE of position with  $\nu = 4$

Figures 1 and 2 show the filtering and smoothing RMSE of position, respectively, for the different methods and considering  $L = 1000$  Monte Carlo samples for the MSPGF/S. In both cases, we obtained similar results. Although the clairvoyant filter/smoother (SQKF/S-K) with fully known measurement noise parameters outperforms the rest, we have a small performance loss with the proposed methodology considering unknown noise latent variables. The new marginalized approach is more robust and outperforms the SQKF/S-P using point estimates first proposed in the filtering context in [25]. The SQKF/S-G operating under the full Gaussian assumption, even if the parameters of the underlying Gaussian noise (i.e.,  $\mu_G$  and  $\sigma_G$ ) are correctly obtained to fit the real data, shows the worst performance. This is because this filter/smoother does not take into account the NLOS-induced outliers in the measurement noise. For the sake of completeness, we assess the impact of the Monte Carlo sample size in the MSPGF/S performance. The mean RMSE of position and velocity, for the different methods and several representative values of  $L$ , are given in Table 1. The performance of the proposed approach is not degraded when using a sample size as low as  $L = 50$  samples, therefore being possible to keep a low overall computational complexity.

Notice that the parameter  $\nu$  of the skew  $t$  distribution controls the tails of the distribution. Lower  $\nu$  implies

**Table 1** Mean RMSE of position and velocity versus Monte Carlo sample size  $L$  in the MSPGS

Smoother	RMSE position	RMSE velocity
SQKS-K (known noise)	0.0653	0.0362
MSPGS $L = 1000$	0.1017	0.0421
MSPGS $L = 500$	0.1020	0.0421
MSPGS $L = 100$	0.1033	0.0421
MSPGS $L = 50$	0.1033	0.0421
SQKS-P	0.1338	0.0452
SQKS-G	0.4202	0.0617

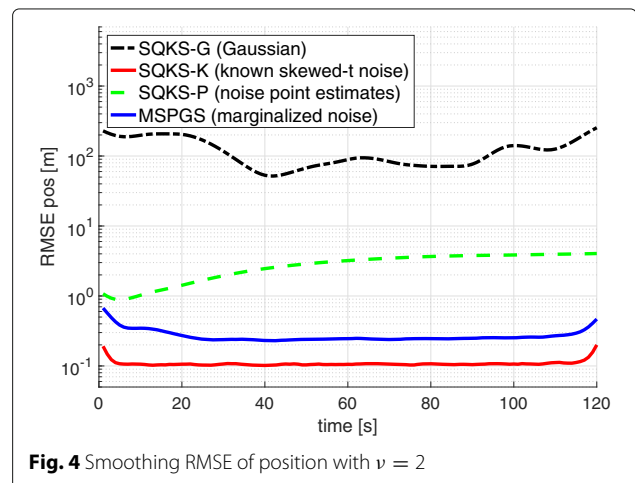


**Fig. 3** Filtering RMSE of position with  $\nu = 2$

heavier tails, thus more outliers and impulsive behaviors. To fully characterize the new method, the performance obtained in the UWB TOA-based localization scenario but now with  $\nu = 2$ , to induce more outliers in the measurements, is shown in Figs. 3 and 4. The proposed method correctly deals with the non-Gaussian noise and approaches the optimal solution. In this case, the performance given by the filter/smoother under the Gaussian assumption (SQKF/S-G) is really poor. This is because the underlying noise distribution is more heavy-tailed, then the Gaussian approximation is no longer valid.

### 6 Conclusions

This article presented a new Bayesian filtering and smoothing framework to deal with nonlinear systems corrupted by parametric heavy-tailed skew  $t$ -distributed measurement noise. The new method takes advantage of the conditionally Gaussian form of the skew  $t$ -distribution, which allows to use a computationally light



**Fig. 4** Smoothing RMSE of position with  $\nu = 2$

Gaussian filter and smoother to deal with the state estimation. The unknown non-Gaussian noise latent variables are marginalized from the general filtering/smoothing solution via Monte Carlo sampling. The performance of the new solution was evaluated in a representative TOA-based localization scenario, where the positive skewed behavior of NLOS propagation conditions is typically modeled using such non-Gaussian distributions.

## Endnote

<sup>1</sup> We write  $(\mathbf{x})^2$ ,  $(\mathbf{y})^2$ ,  $f^2(\cdot)$ , and  $\mathbf{h}^2(\cdot)$  as the shorthand for  $\mathbf{x}\mathbf{x}^\top$ ,  $\mathbf{y}\mathbf{y}^\top$ ,  $\mathbf{f}(\cdot)\mathbf{f}^\top(\cdot)$ , and  $\mathbf{h}(\cdot)\mathbf{h}^\top(\cdot)$ , respectively. We omitted the dependence with time of  $\mathbf{f}_{k-1}(\cdot)$  and  $\mathbf{h}_k(\cdot)$  for the sake of clarity.

## Acknowledgements

This work has been supported by the Spanish Ministry of Economy and Competitiveness through project TEC2015-69868-C2-2-R (ADVENTURE) and by the Government of Catalonia under Grant 2014-SGR-1567.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Castelldefels, Barcelona, Spain. <sup>2</sup>Northeastern University, 02115 Boston, MA, USA.

Received: 24 March 2017 Accepted: 18 August 2017

Published online: 29 August 2017

## References

- D Dardari, E Falletti, M Luise, *Satellite and terrestrial radio positioning techniques: a signal processing perspective*. (Elsevier Academic Press, Oxford, 2011)
- MG Amin, P Closas, A Broumandan, JL Volakis, Vulnerabilities, threats, and authentication in satellite-based navigation systems [scanning the issue]. *Proc. IEEE*. **104**(6), 1169–1173 (2016)
- JT Curran, M Navarro, M Anghileri, P Closas, S Pfletschinger, Coding aspects of secure GNSS receivers. *Proc. IEEE*. **104**(6), 1271–1287 (2016)
- R Ioannides, T Pany, G Gibbons, Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques. *Proc. IEEE*. **104**(6), 1174–1194 (2016)
- H Liu, H Darabi, P Banerjee, J Liu, Survey of wireless indoor positioning techniques and systems. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **37**(6), 1067–1080 (2007)
- D Dardari, P Closas, PM Djuric, Indoor tracking: theory, methods, and technologies. *IEEE Trans. Veh. Technol.* **64**(4), 1263–1278 (2015)
- H Nurminen, T Ardeshiri, Piché, F Gustafsson, in *Proc. of the 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. A NLOS-robust TOA positioning filter based on a skew-t measurement noise model (IEEE, Banff, AB, Canada, 2015)
- P Müller, et al, Statistical trilateration with skew-t distributed errors in LTE networks. *IEEE Trans. Wirel. Commun.* **15**(10), 7114–7127 (2016)
- T-I Lin, JC Lee, WJ Hsieh, Robust mixture modeling using the skew t distribution. *Stat. Comput.* **17**, 81–92 (2007)
- S Zozor, C Vignat, Some results on the denoising problem in the elliptically distributed context. *IEEE Trans. Sig. Process.* **58**(1), 134–150 (2010)
- MJ Lombardi, SJ Godsill, On-line Bayesian estimation of signals in symmetric alpha-stable noise. *IEEE Trans. Sig. Process.* **54**(2), 775–779 (2006)
- J Vilà-Valls, C Fernández-Prades, P Closas, JA Fernández-Rubio, in *Proc. of the European Signal Processing Conference, Eusipco 2011*. Bayesian filtering for nonlinear state-space models in symmetric  $\alpha$ -stable measurement noise (IEEE, Barcelona, 2011), pp. 674–678
- J Vilà-Valls, P Closas, C Fernández-Prades, JA Fernández-Rubio, in *Proc. of the European Signal Processing Conference (Eusipco)*. Nonlinear Bayesian filtering in the Gaussian scale mixture context (IEEE, Bucharest, 2012), pp. 529–533
- K Ito, K Xiong, Gaussian filters for nonlinear filtering problems. *IEEE Trans. Autom. Control*. **45**(5), 910–927 (2000)
- I Arasaratnam, S Haykin, RJ Elliot, Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature. *Proc. IEEE*. **95**(5), 953–977 (2007)
- G Agamennoni, JI Nieto, EM Nebot, Approximate inference in state-space models with heavy-tailed noise. *IEEE Trans. Sig. Process.* **60**(10), 5024–5037 (2012)
- R Piché, S Särkkä, J Hartikainen, in *Proc. of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. Recursive outlier-robust Student-t distribution (IEEE, Bucharest, 2012)
- S Saha, Noise robust online inference for linear dynamic systems (2015). <http://arxiv.org/abs/1504.05723>. Accessed 24 Aug 2017
- H Nurminen, T Ardeshiri, R Piché, F Gustafsson, Robust inference for state-space models with skewed measurement noise. *IEEE Sig. Process. Lett.* **22**(11), 1898–1902 (2015)
- N Chopin, PE Jacob, O Papaspiliopoulos, SMC2: an efficient algorithm for sequential analysis of state-space models. *J. R. Statist. Soc. B*. **75**(3), 397–426 (2013)
- L Martino, J Read, V Elvira, F Louzada, Cooperative parallel particle filters for on-line model selection and applications to urban mobility. *Digit. Sig. Process.* **60**, 172–185 (2017)
- CC Drovandi, J McGree, AN Pettitt, A sequential Monte Carlo algorithm to incorporate model uncertainty in Bayesian sequential design. *J. Comput. Graph. Stat.* **23**(1), 3–24 (2014)
- I Urteaga, MF Bugallo, PM Djuric, in *Proc. of the IEEE Statistical Signal Processing Workshop (SSP)*. Sequential Monte Carlo methods under model uncertainty (IEEE, Palma de Mallorca, 2016)
- F Daum, J Huang, in *Proc. of IEEE Aerospace Conference*. Curse of dimensionality and particle filters (IEEE, Big Sky, 2003)
- P Closas, J Vilà-Valls, in *Proc. of the IEEE Statistical Signal Processing Workshop (SSP)*. NLOS mitigation in TOA-based indoor localization by nonlinear filtering under skew t-distributed measurement noise (IEEE, Palma de Mallorca, 2016)
- H Sorenson, in *Recursive estimation for nonlinear dynamic systems*, ed. by JC Spall (Marcel Dekker, New York, 1988)
- RE Kalman, A new approach to linear filtering and prediction problems. *J. Basic Eng. Trans. ASME*. **82**(1), 35–45 (1960)
- Z Chen, Bayesian filtering: from Kalman filters to particle filters, and beyond (2003). Tech. Rep. Adaptive Syst. Lab. McMaster University. Ontario
- A Doucet, N de Freitas, N Gordon (eds.), *Sequential Monte Carlo Methods in Practice* (Springer, New York, 2001)
- S Arulampalam, S Maskell, N Gordon, T Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*. **50**(2), 174–188 (2002)
- PM Djuric, JH Kotecha, J Zhang, Y Huang, T Ghirmai, MF Bugallo, J Míguez, Particle filtering. *IEEE Signal Processing Magazine*. **20**(5), 19–38 (2003)
- B Ristic, S Arulampalam, N Gordon (eds.), *Beyond the Kalman Filter: Particle Filters for Tracking Applications* (Artech House, Boston, 2004)
- I Arasaratnam, S Haykin, Square-root quadrature Kalman filtering. *IEEE Trans. Sig. Process.* **56**(6), 2589–2593 (2008)
- I Arasaratnam, S Haykin, Cubature Kalman filters. *IEEE Trans. Autom. Control*. **54**(6), 1254–1269 (2009)
- B Anderson, JB Moore, *Optimal Filtering*. (Prentice-Hall, Englewood Cliffs, 1979)
- HE Rauch, F Tung, CT Striebel, Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **3**(8), 1445–1450 (1965)
- S Särkkä, Unscented Rauch-Tung-Striebel smoother. *IEEE Trans. Autom. Control*. **53**(3), 845–849 (2008)
- I Arasaratnam, S Haykin, Cubature Kalman smoothers. *Automatica*. **47**, 2245–2250 (2011)
- SJ Julier, JK Uhlmann, HF Durrant-Whyte, A new method for nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Autom. Control*. **45**(3), 472–482 (2000)

40. M Norgaard, NK Poulsen, O Ravn, New developments in state estimation of nonlinear systems. *Automatica*. **36**, 1627–1638 (2000)
41. T-I Lin, Robust mixture modeling using multivariate skew t distributions. *Stat. Comput.* **20**, 343–356 (2010)
42. JM Bernardo, AFM Smith, *Bayesian Theory*, vol. 405. (Wiley, New York, 2009)
43. S Särkkä, *Bayesian Filtering and Smoothing*. (Cambridge University Press, Cambridge, 2013)
44. P Closas, J Vilà-Valls, C Fernández-Prades, in *Proc. of the CAMSAP'15*. Computational complexity reduction techniques for quadrature Kalman filters (IEEE, Cancun, 2015)
45. B Jia, M Xin, Y Cheng, Sparse-grid quadrature nonlinear filtering. *Automatica*. **48**(2), 327–341 (2012)
46. J Vilà-Valls, P Closas, AF García-Fernández, Uncertainty exchange through multiple quadrature Kalman filtering. *IEEE Sig. Process. Lett.* **23**(12), 1825–1829 (2016)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---