NONLINEAR BAYESIAN FILTERING IN THE GAUSSIAN SCALE MIXTURE CONTEXT

Jordi Vilà-Valls[(1)], *Pau Closas*[(2)], *Carles Fernández-Prades*[(2)] *and Juan A. Fernandez-Rubio*[(1)]

[(1)] Universitat Politècnica de Catalunya (UPC), Jordi Girona 1-3, 08034, Barcelona, Spain.
[(2)] Centre Tec. de Telecom. de Catalunya (CTTC), Av. Friedrich Gauss 7, 08860, Castelldefels, Spain.
email: { jordi.vila-valls, juan.a.fernandez }@upc.edu, { carles.fernandez, pau.closas }@cttc.cat

ABSTRACT

In many real–life Bayesian estimation problems, it is appropriate to consider non-Gaussian noise distributions to model possible outliers or impulsive behaviors in the measurements. In this paper, we considered a nonlinear Bayesian filtering problem with a Gaussian process noise and a Gaussian scale mixture (GSM) distributed measurement noise. Both processes' statistics parameters are assumed unknown. Within this framework, we present a filtering method based on a sigma–point core that exploits GSM's product property and accounts for such heavier distribution tail and parameter uncertainty. Numerical results exhibit enhanced robustness against both outliers and a weak knowledge of the system with respect to state–of–the–art nonlinear Bayesian filters based on the Gaussian assumption, requiring much less computational load than standard Sequential Monte-Carlo methods and approaching theoretical bounds of performance.

Index Terms— Nonlinear Bayesian filtering, Gaussian Scale Mixtures, covariance estimation, sigma–point Kalman filters, Monte Carlo methods

1. INTRODUCTION

The problem under study concerns the derivation of efficient and robust methods to solve the recursive Bayesian filtering problem, which implies the on–line estimation of the time–varying unknown states of a system, using the incoming flow of information (observations) from the system, along with some prior statistical knowledge about the variations of such states. The standard Kalman filter (KF) provides the closed form solution to the optimal filtering problem in linear/Gaussian systems, assumptions that not always hold, reason why suboptimal techniques have to be used. A plethora of alternatives has been proposed in the last decade to solve the nonlinear estimation problem, among them, the family of sigma–point KFs (SPKF) [1, 2] within the Gaussian framework, and the family of Sequential Monte Carlo (SMC) methods [3] for arbitrary noise distributions. A limitation of these methods is that they assume some *a priori* knowledge of the noise statistics affecting the system (*i.e.* not only its distribution but its parameters).

In many real–life systems, Gaussian noise models do not apply and the noise statistics are unknown. In these scenarios the methods based on the standard Gaussian Kalman framework (KF, extended KF and SPKF) give poor performances and we cannot directly apply SMC methods because we have to estimate the states together with the noise statistics.

Heavy-tailed and elliptical distributions have been shown to be appropriate for a large number of applications in signal processing [4, 5] and other fields such economics, engineering or statistics [6]. In particular, these distributions are useful to model the existence of outliers or impulsive behaviors on the measurement model. For a recent review on the use of elliptical distributions in signal processing applications see [4] and references therein. An important subclass of the elliptical distributions family is the Gaussian scale mixtures (GSM), a.k.a. Scale Mixture of Normals (SMiN) [7], which include the Gaussian, the Student-t, the Laplacian and symmetric α-stable ($S\alpha S$) distributions, to name a few.

In the literature, some contributions already considered the robust Bayesian filtering problem in the context of GSM distributed noises. In [8], the author considered the case of $S\alpha S$ distributions and gave a solution based on Markov Chain Monte Carlo (MCMC) inference techniques, and [9] proposed a direct SMC solution to solve the Bayesian estimation problem for linear Time–Varying AutoRegressive (TVAR) models in the same context. In [10], we proposed a solution to deal with $S\alpha S$ measurement noise and unknown statistics for nonlinear multivariate state–space models, but a framework to deal with general nonlinear state–space models corrupted by GSM noise, to the authors' knowledge, is still missing.

In this contribution, the results presented in [10] are generalized, providing a framework to deal with GSM distributed measurement noise. Within this framework, we propose a robust Bayesian solution against both outliers and a weak knowledge of the system. Numerical results exhibit enhanced performances with respect to state-of-the-art non-

This work has been partially supported by project 2009 SGR-1236 AGAUR of the Catalan Government, by the Spanish Science and Technology Commission: CONSOLIDER CSD 2008-00010 (COMONSENS), TEC 2010-19171 (MOSAIC), and TEC 2008-02685 (NARRA) and by the European Commission in the framework of the COST Action IC0803 (RFCSET).

linear Bayesian filters based on the Gaussian assumption, requiring much less computational load than standard SMC methods and approaching theoretical bounds of performance. The paper is organized as follows: Section 2 introduces the GSM distributions and the system model, Section 3 proposes a robust Bayesian solution and numerical results for a radar target tracking problem are given in Section 4.

2. GAUSSIAN SCALE MIXTURE DISTRIBUTIONS AND THE SYSTEM MODEL

In this section, the GSM distributions, and how to express a GSM distributed random variable as conditionally Gaussian, are introduced. Then, the system model which considers GSM distributed measurement noise and the Bayesian estimation problem to be solved are given.

2.1. Gaussian scale mixture distributions

We can define a GSM distributed random vector $\boldsymbol{x}$ through its stochastic representation [4]:

$$\boldsymbol{x} \stackrel{d}{=} \sqrt{z}\boldsymbol{v}, \tag{1}$$

where $\stackrel{d}{=}$ means equality in distribution, $\boldsymbol{v}$ is a zero-mean standard Gaussian random vector and z is a positive scalar random variable, independent of $\boldsymbol{v}$, called the *generating variate* or *scale parameter*. The probability density function of $\boldsymbol{x}$ (when de probability density function (pdf) of the generating variate exists) can be obtained as

$$p_{\mathbf{x}}(\mathbf{x}) = \int_{\mathbb{R}_+} p_{\mathbf{x}|z}(\mathbf{x}|z)p_z(z)dz. \tag{2}$$

These equations describe different families of parametric distributions depending on the pdf of the generating variate, $p_z(z)$, usually referred to as mixing distribution. The pdf of the generating variate, z, might depend on a set of parameters $\boldsymbol{\phi}$, so we should write $p_z(z;\boldsymbol{\phi})$ (*e.g.* for a Student-t distribution $\phi = \nu$, the number of degrees of freedom). Some of these distributions and the resulting GSM pdf, for the univariate case, are listed in Table 1.

GSM pdf $p_x(x)$	Mixing distribution $p_z(z)$
Gaussian	Dirac
Laplacian	Exponential
$S\alpha S$	Positive skewed α-stable
K-distribution	Gamma
Student-t	Inverse Gamma
Normal-Inverse Gaussian	Inverse Gaussian

Table 1. Relation between the GSM distribution and the corresponding mixing distribution.

The pdf of $\boldsymbol{x}$ is specified by an ensemble of parameters $\boldsymbol{\psi}$ which depends on the generating variate and its parametrization $\boldsymbol{\phi}$, and may include a shift and a scaling (*e.g.*, for a $S\alpha S$, $\phi = \alpha$ and $\boldsymbol{\psi} = \{\delta, \gamma, \lambda, \alpha\}$, the shift, the scaling, the generating variate and the index of stability, respectively, see [10]). We note that within this context, the random vector $\boldsymbol{x}$ is Gaussian when conditioned to z, which is an important property and the key point of the Bayesian solution presented in Section 3. Then,

$$\boldsymbol{x}|z \sim \mathcal{N}(\mathbf{0}, z\mathbf{I}), \tag{3}$$

where we note that the covariance of the Gaussian distribution is controlled by the generating variate.

2.2. System model and the estimation problem

In this paper we are interested in nonlinear filtering problems where the process noise is Gaussian and the measurement noise is GSM distributed, both being additive. The heavy–tailed measurement noise accounts for possible outliers or impulsive behavior in the observations, giving a more general and flexible framework than considering the standard Gaussian case [11].

The assumed discrete state-space model is expressed as

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{v}_k, \tag{4}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k, \tag{5}$$

where $k \in \mathbb{Z}$ refers to discrete time instants, $\mathbf{x}_k \in \mathbb{R}^N$ and $\mathbf{y}_k \in \mathbb{R}^L$ are the states and the observations at time k. We assume that the components of the measurement $(y_{k,1}, ..., y_{k,L})$ are independent. $\mathbf{f}$ and $\mathbf{h}$ are the process and measurement functions, assumed known and nonlinear in a general case. Vectors $\mathbf{v} = \{\mathbf{v}_k, k \in \mathbb{Z}\}$ and $\mathbf{n} = \{\mathbf{n}_k, k \in \mathbb{Z}\}$ are the process (Gaussian) and observation (GSM) noises, which are mutually independent with unknown statistics (*i.e.*, in real–life systems we do not have complete knowledge of the system dynamics). The Gaussian process noise, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{\Sigma}_{v,k})$, is characterized by its covariance matrix, $\mathbf{\Sigma}_{v,k}$. We consider that the components of the measurement noise $(n_{k,1}, ..., n_{k,L})$ are independent, each one being GSM distributed. Using the conditionally Gaussian form of a GSM distribution we can write the measurement noise as $\mathbf{n}_k \sim \mathcal{N}(0, \mathbf{\Sigma}_{n,k})$, and because of the independence between components the covariance matrix is

$$\mathbf{\Sigma}_{n,k} = \mathrm{diag}\,(z_{k,1}, \ldots, z_{k,L}) \quad z_{k,i} \sim p_{z_i}(z_i; \boldsymbol{\phi}_i). \tag{6}$$

where the subscript i in the generating variate, z, and its parametrization accounts for possible different GSM noise distributions within the measurement noise.

The robust Bayesian filtering problem within the conditionally Gaussian form concerns the recursive estimation of the states $\mathbf{x}_k$ and the unknown parameters of the system, namely the process noise covariance matrix $\mathbf{\Sigma}_{v,k}$ and the parameters describing the GSM distribution $\boldsymbol{\psi}_k$, or equivalently the random covariance matrix $\mathbf{\Sigma}_{n,k}$. We denote as $\boldsymbol{\theta}_k$ the overall parameter vector containing both process and measurement noise parameters.

3. BAYESIAN SOLUTION

The solution to the robust Bayesian filtering problem for the state–space model defined in (4) and (5) is given by the joint *a posteriori* distribution $p(\mathbf{x}_k, \boldsymbol{\theta}_k | \mathbf{y}_{1:k})$, which contains all the information about the states and the model contained in the observations and the prior knowledge. Its characterization allows us to obtain an optimal estimate with respect to any criterion, for example, the Mininum Mean Square Error (MMSE) or the Maximum a Posteriori (MAP) estimates. This pdf can be rewritten as

$$p(\mathbf{x}_k, \boldsymbol{\theta}_k | \mathbf{y}_{1:k}) = p(\mathbf{x}_k | \boldsymbol{\theta}_k, \mathbf{y}_{1:k}) p(\boldsymbol{\theta}_k | \mathbf{y}_{1:k}). \quad (7)$$

A direct application of a SMC method to obtain the joint estimation of the states and the parameters of the model is unviable in general because the dimensionality of the problem is too large and the method would collapse. To overcome this problem we profit of the underlying Gaussian structure: under the knowledge of the noise statistics parameters $\boldsymbol{\Sigma}_{v,0:k}$ and $\boldsymbol{\psi}_{0:k}$, the state–space model is conditionally Gaussian and we can resort to standard Gaussian techniques to compute the *a posteriori* distribution $p(\mathbf{x}_k | \boldsymbol{\theta}_k, \mathbf{y}_{1:k})$ (what we call the *core method*, which could be any nonlinear Gaussian Bayesian filter). In addition, at each time step, we need estimates of the parameters $\boldsymbol{\theta}_k$, that we can obtain using different strategies. We propose to use two separate methods, to estimate both process and measurement noise parameters, which are coupled to the core method. The block diagram structure of the method is depicted in Fig. 1.

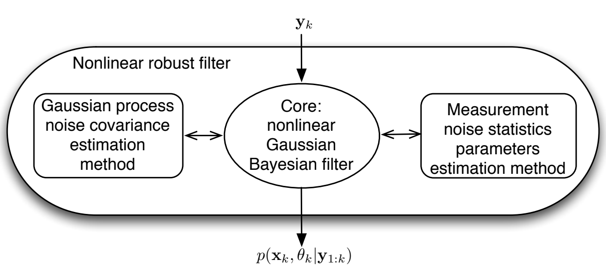


Fig. 1. Block diagram with the three coupled methods' structure

3.1. The core method

SPKFs are a family of powerful methods that use a deterministic sampling to approximate the integrals of the optimal Bayesian filter. The key idea is to compute the means and covariances used in the standard Kalman solution by propagating the sample set through the process and measurement nonlinear functions. A further refinement of sigma–point (SP) schemes comes from the fact that, when we propagate the covariance matrix through a nonlinear function, 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 alleviate this problem, a square-root filter is introduced to propagate the square-root of the covariance matrix instead of the covariance itself [12, 2]. We propose to use a square–root sigma–point Kalman filter using quadrature rules (SQKF) to estimate the conditionally Gaussian filtering density $p(\mathbf{x}_k | \boldsymbol{\theta}_k, \mathbf{y}_{1:k})$. The process and measurement noise covariance matrices are provided by two auxiliary estimation methods described hereafter. For a detailed structure of the SQKF, see [12].

3.2. Process noise covariance estimation

We propose to use a nonlinear version of the covariance matching method first introduced in [13]. The unbiased estimator of $\boldsymbol{\Sigma}_{v,k+1}$ at time k within the sigma–point formulation is

$$\hat{\boldsymbol{\Sigma}}_{v,k+1} = \frac{1}{k-1} \sum_{j=1}^{k} (\mathbf{q}_j - \hat{\mathbf{q}})(\mathbf{q}_j - \hat{\mathbf{q}})^T - \frac{1}{k} \sum_{j=1}^{k} \boldsymbol{\beta}_j. \quad (8)$$

where $\mathbf{q}_k = \hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1}$, $\hat{\mathbf{q}} = 1/k \sum_{i=1}^{k} \mathbf{q}_i$ and $\boldsymbol{\beta}_j = \sum_{i=1}^{M} \omega_i \tilde{\mathbf{x}}_{i,j|j-1} \tilde{\mathbf{x}}_{i,j|j-1}^T - \hat{\mathbf{x}}_{j|j-1} \hat{\mathbf{x}}_{j|j-1}^T - \boldsymbol{\Sigma}_{x,j|j}$ are obtained from the SP core method [10].

3.3. GSM measurement noise parameters estimation

As we stated in Section 2, the GSM noise distribution depends on a set of parameters $\boldsymbol{\psi}_k$, that may be time-varying in a general case. Because the components $y_{k,i}$ are supposed to be independent, we can estimate the subsets $\boldsymbol{\psi}_{k,i}$ independently for each component. So we can use L parallel filters, each one estimating one subset, instead of using a method to estimate the whole noise parameters vector at once, what might be computationally unaffordable.

First of all, we need to define the prior knowledge about the parameters to be estimated, which information is gathered in $p(\psi_{k,i})$. This distribution depends on the generating variate and the type of GSM distribution that we consider. From the SPKF we are able to compute the likelihood $p(y_{k,i} | \boldsymbol{\psi}_{0:k}, y_{1:k-1,i})$, which depends on the predicted measurement, $\hat{\mathbf{y}}_{k|k-1}$, and the innovation's covariance matrix, $\boldsymbol{\Sigma}_{y,k|k-1}$,

$$p(y_{k,i} | \boldsymbol{\phi}_{0:k,i}, y_{1:k-1,i}) = \mathcal{N}\left(\hat{\mathbf{y}}_{k|k-1,i}, \left[\boldsymbol{\Sigma}_{y,k|k-1}\right]_{ii}\right). \quad (9)$$

Using this likelihood and the prior density we are able to construct a Monte Carlo-type solution [14, 3] to estimate the subset $\boldsymbol{\psi}_{k,i}$ for each element of the observation, $y_{k,i}$.

3.4. Computational cost

As a final remark, it is worthwhile to point out the implicit reduction in the computational cost of the proposed method with respect to SMC methods directly applied to jointly estimate the states, $\mathbf{x}_k$, and the parameter vector $\boldsymbol{\theta}_k$ (for joint estimation a SMC approach is needed because of the non-Gaussian measurement noise). Our method also reduces the number of evaluations of the nonlinear functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$, which uses to be time-consuming, given that the number of SP required to estimate $\mathbf{x}$ with a given performance is much lower than the number of random particles needed to achieve the same performance in the joint estimation problem.

4. COMPUTER SIMULATIONS

In this section, in order to provide illustrative numerical results, the performance of the proposed method is shown in a radar target tracking example where the heavy–tailed measurement noise applies [15, 11]. In the presented application, a target was moving in a 2-D plane and was tracked by a radar whose measurements were range and azimuth, $\mathbf{y}_k = [r_k, \psi_k]^T$. The states to be tracked were position, velocity and acceleration of the target. These were respectively gathered in vector $\mathbf{x}_k = [p_{x,k}, p_{y,k}, v_{x,k}, v_{y,k}, a_{x,k}, a_{y,k}]^T$. Both the trajectory and measurements were modeled as

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{I} & T\times\mathbf{I} & T^2/2\times\mathbf{I} \\ \mathbf{0} & \mathbf{I} & T\times\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{x}_{k-1} + \mathbf{v}_k, \quad (10)$$

$$\mathbf{y}_k = \begin{pmatrix} \sqrt{p_{x,k}^2 + p_{y,k}^2} \\ \arctan(p_{y,k}/p_{x,k}) \end{pmatrix} + \mathbf{n}_k, \quad (11)$$

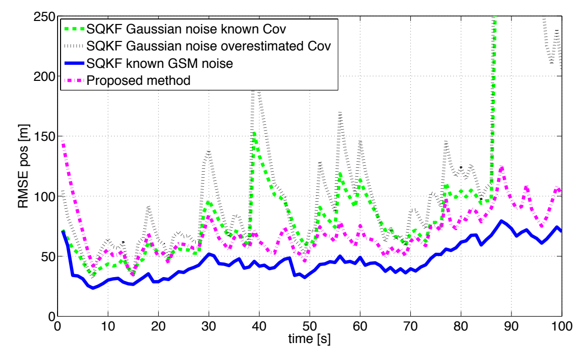
where T is the time–interval between measurements, set to 1 second. The Gaussian process noise was modeled as $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_v)$ and $\boldsymbol{\Sigma}_v = \mathrm{diag}(4, 4, 4, 4, 0.01, 0.01)$. Each component of the measurement noise was Student-t distributed, $n_{k,i} \sim \mathcal{T}(0, \sigma_i^2, \nu_i)$, where ν_i refers to the number of degrees of freedom of the Student-t distribution and σ_i^2 to a scale parameter. Using the GSM representation, the Student-t noise can be written as

$$n_{k,i}|\sigma_i^2, \nu_i, z_{k,i} \sim \mathcal{N}(0, z_{k,i}\sigma_i^2) \quad (12)$$

where $z_{k,i}|\nu \sim \mathcal{IG}(\nu_i/2, 2/\nu_i)$ and $\mathcal{IG}(\alpha, \beta)$ is an inverse gamma distribution with shape parameter α and scale parameter β. So for each measurement component and at each time step, we have to estimate the triad $\{\nu_i, \sigma_i, z_i\}$. Concerning the prior distributions of these parameters, ν_i and σ_i are considered static and $z_{k,i}$ inverse-gamma distributed.

In this application, two different heavy tailed scenarios were considered: $\nu_1 = \nu_2 = 2$ and $\nu_1 = \nu_2 = 1.5$. In both scenarios the following parametrization was used: $\sigma_1^2 = 100, \sigma_2^2 = 0.001$. We note that lower the parameter ν heavier the tails of the noise distribution and so stronger outliers. These parameters imply that in the Gaussian scenario ($\nu_1 \to \infty$ and $\nu_2 \to \infty$) the measurement noise covariance matrix would be $\mathbf{R} = \mathrm{diag}(100, 0.001)$, which will be used for comparison.

The initial state estimate was drawn from $\mathcal{N}(\hat{\boldsymbol{x}}_0, \boldsymbol{\Sigma}_{x,0})$ for each Monte Carlo trial, with $\hat{\boldsymbol{x}}_0 = [2000, 2000, 20, 20, 0, 0]$ and $\boldsymbol{\Sigma}_{x,k} = \mathrm{diag}(5\cdot 10^4, 5\cdot 10^4, 8, 8, 0.02, 0.02)$. We used 3 sigma-points/dimension so the method required $M = 3^N = 729$ points. The unknown process covariance matrix was initialized to $100\cdot\boldsymbol{\Sigma}_v$. We made 2000 independent Monte Carlo runs with 100 scans per run and we used the root-mean square error (RMSE) as the measure of performance. As the process noise covariance matrix is constant we use all the samples available in the covariance matching method (process noise covariance estimation method). Concerning the Monte Carlo method (measurement noise statistics parameters estimation method), the particles were drawn from the prior distributions: from $\mathcal{U}(1, 5)$ for ν_i, from $\mathcal{U}(0, 5\cdot\sigma_i)$ for the estimation of σ_i and from $\mathcal{IG}(\nu_{i,p}/2, 2/\nu_{i,p})$ for the estimation of $z_{k,i}$, where $\nu_{i,p}$ are the particles generated for ν_i. We used $N_p = 500$ particles.



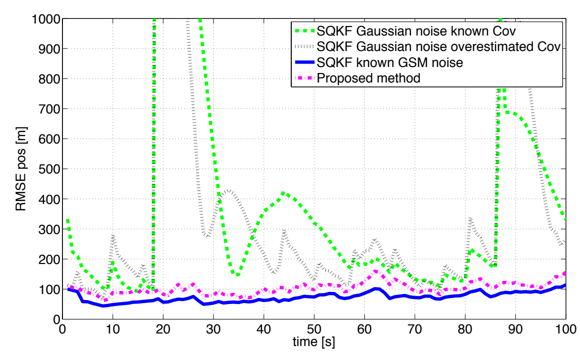


Fig. 2. Position RMSE for the robust Bayesian method and the SQKF with known statistics for both Gaussian and GSM scenarios, considering two cases: $\nu_1 = \nu_2 = 2$ (top) and $\nu_1 = \nu_2 = 1.5$ (bottom).

Fig. 2 plots the RMSE of position estimates obtained with different algorithms in two different scenarios: the first one considers $\nu_1 = \nu_2 = 2$ and the second one $\nu_1 = \nu_2 = 1.5$. The solid blue line plots the performance obtained with the SQKF with a full knowledge of both the process noise covariance matrix and the measurement GSM noise parameters. The results obtained in this case are a reference of the ultimate achievable performance with the robust Bayesian solution proposed in this paper.

The results obtained with the proposed method (dot-dashed magenta line) are really encouraging. In this example, the proposed filter is able to estimate the noise parameters, and deals correctly with the impulsive behavior of the measurement noise with limited performance degradation in scenarios where other Gaussian filters give worse results: dashed green and dotted black lines correspond to the performance obtained with the SQKF considering Gaussian noise with $\mathbf{\Sigma}_n = \boldsymbol{R}$ and known $\mathbf{\Sigma}_v$, and overestimated covariance matrices $100 \cdot \mathbf{\Sigma}_v$ and $\mathbf{\Sigma}_n = 10 \cdot \boldsymbol{R}$, respectively.

5. CONCLUSIONS

This paper presented a robust (adaptive) solution to the Bayesian filtering problem for nonlinear state-space models with Gaussian scale mixture (GSM) distributed measurement noise, what has been proven to be a more appropriate representation of the measurement than the standard Gaussian case in many real–life systems.

The performance of the proposed method was validated by computer simulation in a target tracking application. In this scenario, we saw that the proposed method attains good performance results dealing correctly with both outliers/impulsive behaviors in the measurement and unknown process and measurement noise statistics, while being computationally affordable when compared to standard methods. We considered a Student-t noise distribution, but this solution can be applied to any GSM noise distribution, providing a general framework for the GSM noise context.

Even if the method's performance was shown for a target tracking application, the proposed Bayesian solution is a powerful tool to deal with outliers/glitches or unknown noise environments in other localization applications such as integrated navigation systems (*i.e.* coupling Global Navigation Satellite Systems with inertial sensors) or indoor positioning systems using wireless sensor networks.

6. REFERENCES

[1] I. Arasaratnam, S. Haykin, and R. J. Elliot, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proc. of the IEEE*, vol. 95, no. 5, pp. 953–977, 2007.

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

[3] A. Doucet, N. De Freitas, and N. Gordon, Eds., *Sequential Monte Carlo methods in practice*, Springer, 2001.

[4] S. Zozor and C. Vignat, "Some results on the denoising problem in the elliptically distributed context," *IEEE Trans. on Signal Processing*, vol. 58, no. 1, pp. 134–150, Jan. 2010.

[5] C. L. Nikias and M. Shao, *Signal processing with alpha–stable distributions and applications*, John Wiley and Sons, 1995.

[6] R. J. Adler, R. E. Feldman, and M. S. Taqqu, Eds., *A practical guide to heavy tails: statistical techniques and applications*, Birkhaüser, 1998.

[7] A. F. Andrews and C. L. Mallows, "Scale mixtures of normal distributions," *Journal of the Royal Statistical Society. Series B*, vol. 36, no. 1, pp. 99–102, 1974.

[8] S. J. Godsill, "Inference in symmetric alpha-stable noise using MCMC and the slice sampler," in *Proc. IEEE ICASSP*, Istanbul, Turkey, June 2000, pp. 3806 – 3809.

[9] M. J. Lombardi and S. J. Godsill, "On–line Bayesian estimation of signals in symmetric alpha-stable noise," *IEEE Trans. on Signal Processing*, vol. 54, no. 2, pp. 775–779, February 2006.

[10] J. Vilà-Valls, C. Fernández-Prades, P. Closas, and J. A. Fernández-Rubio, "Bayesian filtering for nonlinear state–space models in symmetric α–stable measurement noise," in *Proc. of the European Signal Processing Conference, Eusipco 2011*, Barcelona, Spain, August 2011, pp. 674–678.

[11] D. Middleton, "Non–Gaussian noise models in signal processing for telecommunications: new methods and results for Class A and Class B noise models," *IEEE Trans. on Information Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[12] I. Arasaratnam and S. Haykin, "Square-root quadrature Kalman filtering," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2589–2593, June 2008.

[13] K. A. Myers and B. D. Tapley, "Adaptive sequential estimation with unknown noise statistics," *IEEE Trans. on Automatic Control*, vol. 21, no. 8, pp. 520–523, 1976.

[14] C. P. Robert and G. Casella, *Monte Carlo statistical methods*, Springer, 2nd edition, 2004.

[15] B. Borden and M. Mumford, "A statistical glint/radar cross section target model," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 19, no. 1, pp. 781–785, 1983.