# DECENTRALIZED INFORMATION FILTERING UNDER SKEW-LAPLACE NOISE

*Jordi Vilà-Valls*$^{(1)}$*, François Vincent*$^{(1)}$ *and Pau Closas*$^{(2)}$

$(1)$ University of Toulouse/ISAE-SUPAERO, DEOS/SCAN, Toulouse (France)
$(2)$ Dept. of Electrical & Computer Engineering, Northeastern University, Boston, MA (USA)
e-mail: {jordi.vila-valls, francois.vincent}@isae.fr, closas@northeastern.edu

**ABSTRACT**

Localization in large sensor networks requires decentralized computationally efficient filtering solutions. To model challenging indoor propagation conditions, including non-line-of-sight conditions and other channel variations, it may be necessary to consider non-Gaussian distributed errors. In this case, Gaussian filters cannot be considered as is and particle filters do not meet the system requirements on computational cost and/or available memory. In this article we explore decentralized Gaussian information filtering strategies under skew-Laplace errors, exploiting the hierarchically Gaussian formulation of such distribution. An illustrative example is considered to show the performance and support the discussion.

***Index Terms*—** Network localization, distributed filtering, object tracking, skew-Laplace distributed noise.

## 1. MOTIVATION & SIGNAL MODEL

### 1.1. RSS-based Localization State-space Model

We are interested in robust object localization exploiting received signal strength (RSS) measurements to improve the localization in challenging non-Gaussian indoor scenarios. An object is localized using a set of $N$ RSS sensors. Its 2D position/velocity are to be inferred, $\mathbf{x}_t = [p_{x,t}, p_{y,t}, v_{x,t}, v_{y,t}]^\top$, $\mathbf{p}_t = [p_{x,t}, p_{y,t}]^\top$, $\mathbf{v}_t = [v_{x,t}, v_{y,t}]^\top$,

$$\mathbf{x}_t = \underbrace{\begin{pmatrix} \mathbf{I}_2 & T_s \cdot \mathbf{I}_2 \\ \mathbf{0} & \mathbf{I}_2 \end{pmatrix}}_{\mathbf{F}} \mathbf{x}_{t-1} + \boldsymbol{\nu}_{t-1}, \ \boldsymbol{\nu}_{t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \quad (1)$$

with $\mathbf{Q} = \text{diag}(\sigma^2_{p_x}, \sigma^2_{p_y}, \sigma^2_{v_x}, \sigma^2_{v_y})$, $T_s$ the sampling period and $\mathbf{I}_2$ the 2x2 identity matrix. The $m$-th sensor RSS is

$$y_{m,t} = 10 \log_{10}\left(\frac{1}{|\mathbf{r}_m - \mathbf{p}_t|^2}\right) + n_{m,t},$$

with measurement noise typically considered to be Gaussian distributed $n_{m,t} \sim \mathcal{N}(n_{m,t}; 0, \sigma^2_m)$, with $\sigma^2_m$ depending on line-of-sight (LOS) or non-line-of-sight (NLOS) conditions, and known RSS sensor position $\mathbf{r}_m = [r_{x,m}, r_{y,m}]^\top$. But the Gaussian distribution does not account for possible channel variations between objects and RSS anchor nodes. For instance, in crowded indoor scenarios where several people may cross such propagation path, then a non-Gaussian distribution must be accounted for, $n_{m,t} \sim \mathcal{D}(n_{m,t}; \boldsymbol{\phi}_m)$, with $\boldsymbol{\phi}_m$ a known hyperparameters vector. The complete RSS measurement equation is then

$$\begin{aligned} \mathbf{y}_t &= [y_{1,t}, \cdots, y_{N,t}]^\top \\ &= \underbrace{\begin{pmatrix} 10 \log_{10}\left(\frac{1}{|\mathbf{r}_1 - \mathbf{p}_t|^2}\right) \\ \vdots \\ 10 \log_{10}\left(\frac{1}{|\mathbf{r}_N - \mathbf{p}_t|^2}\right) \end{pmatrix}}_{\mathbf{h}_t(\mathbf{x}_t)} + \underbrace{\begin{pmatrix} n_{1,t} \\ \vdots \\ n_{N,t} \end{pmatrix}}_{\mathbf{n}_t}. \quad (2) \end{aligned}$$

We can consider the following two approaches:

- Centralized Filtering: if a centralized localization approach is considered, the complete set of available observations $\mathbf{y}_t$ is used into the filtering method, where each noise components in $\mathbf{n}_t$ follow a non-Gaussian distribution. Then the state-space model (SSM) given by (1) and (2), is nonlinear/non-Gaussian.

- Decentralized Filtering: in a distributed approach a single sensor or a cluster of close sensors must perform the estimation task, without having access to the complete set of observations. This is typically done through a three step approach: i) a set of sensors in a cluster are able to communicate and perform the estimation task, ii) the estimated states are transmitted to a fusion center which refines these estimates by combining the output of the different clusters, and iii) the updated estimates are re-transmitted to the set of sensors/clusters to perform the subsequent estimation.

### 1.2. Rationale behind the Proposed Filtering Strategy

Because of the SSM of interest, the conventional solution would be to consider a Particle Filter (PF) approach [1]. But

This work has been partially supported by the DGA/AID (2018.60.0072.00.470.75.01), and NSF under awards CNS-1815349 and ECCS-1845833.

the PF is not suited for decentralized localization in large wireless sensor networks (WSN), which require cost-efficient algorithms whose implementation does not involve high computational complexity. On the other hand, standard nonlinear Gaussian filters [2, 3], and their decentralized versions [4, 5] assume Gaussian distributed errors, which is not the case in the indoor localization context of interest, where non-Gaussian distributions must be considered to properly characterize NLOS and other possible RSS channel variations. The proposed filtering strategy wants to: i) provide a decentralized solution, ii) use computationally efficient Gaussian filters, iii) while considering the possibly non-Gaussian measurement noise. To achieve such goal, we exploit the fact that noise distributions may be expressed in a conditionally Gaussian form (see Section 1.3). We provide a new decentralized collaborative multiple sensor fusion algorithm, which operates in large-dimensional non-Gaussian systems and reduces the computational load at each cluster of sensing nodes. This work further analyses the concepts introduced in [6] for centralized TOA localization under skew t-distributed noise.

### 1.3. Hierarchically Gaussian RSS Error Distribution

Several contributions deal with the characterization of RSS errors in mixed LOS/NLOS environments, see [7] and references therein. RSS errors under LOS and NLOS conditions are typically modeled as Gaussian distributed with different variances, $\sigma^2_{LOS} < \sigma^2_{NLOS}$, to account for different RSS channel variations. It is suggested in [8] that the fadings induced by people crossing a RSS link follow a skew-Laplace distribution, which is typically the case in crowded environments. In this article we adopt such RSS error distribution, which allows to avoid to distinguish between LOS and NLOS conditions. The univariate skew-Laplace pdf is given by [9]

$$\mathcal{SL}(z;\mu,\sigma,\lambda) = \frac{1}{2\sigma\alpha}\exp\left(\lambda\left(\frac{x-\mu}{\sigma^2}\right) - \alpha\left|\frac{x-\mu}{\sigma}\right|\right), \tag{3}$$

with hyperparameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$, being respectively the distribution location, scale and skewness, and $\alpha = \sqrt{1+\lambda^2/\sigma^2}$. Notice that this distribution accepts a Normal Variance-Mean Mixture (NVMM) representation. This implies that it can be reformulated as hierarchically Gaussian. Mathematically, $z \sim \mathcal{SL}(z;\mu,\sigma^2,\lambda)$ can be rewritten as [9]

$$z|\beta \sim \mathcal{N}\left(z;\mu+\beta\lambda,\beta\sigma^2\right), \tag{4}$$

with $\beta \sim \mathcal{G}(\beta;1,2)$ and $\mathcal{G}(\cdot)$ the gamma distribution. Notice that with $\lambda = 0$, this becomes the (symmetric) Laplace distribution $\mathcal{L}(z;\mu,\sigma) = \mathcal{SL}(z;\mu,\sigma,\lambda=0)$.

Under the knowledge of $\mu,\sigma,\lambda$ and $\beta$ in (4), $z$ is Gaussian, then $\mathcal{D}(n_{i,t};\boldsymbol{\phi}_i) = \mathcal{SL}(n_{i,t};\boldsymbol{\phi}_i)$ (i.e., $\boldsymbol{\phi}_i = [\mu_i,\sigma_i,\lambda_i]$) and using the NVMM formulation, the SSM becomes nonlinear/Gaussian. The marginal posterior distribution of the states, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ turns to be Gaussian and thus we are able to use computationally light Gaussian filtering methods.

## 2. METHODOLOGY

The nonlinear/non-Gaussian SSM of interest is written as

$$\mathbf{x}_t = \mathbf{f}_{t-1}(\mathbf{x}_{t-1}) + \boldsymbol{\nu}_{t-1}\ ,\boldsymbol{\nu}_{t-1} \sim \mathcal{N}(\boldsymbol{\nu}_{t-1};\mathbf{0},\mathbf{Q}_{t-1}), \tag{5}$$

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t) + \mathbf{n}_t, n_{i,t} \sim \mathcal{SL}(n_{i,t};\boldsymbol{\phi}_i), \tag{6}$$

where each measurement noise component is skew-Laplace distributed, being conditionally Gaussian from the formulation in Section 1.3. $\mathbf{x}_t \in \mathbb{R}^{n_x}$ are the hidden state of the system to be inferred from the available observations $\mathbf{y}_t \in \mathbb{R}^{n_y}$; $\mathbf{f}_{t-1}(\cdot)$ and $\mathbf{h}_t(\cdot)$ are known functions of the state; both noise sequences $\boldsymbol{\nu}_{t-1}$ and $\mathbf{n}_t$ are assumed to be independent. Considering a set of observations taken at $M$ different clusters of sensors, where the $j$-th cluster (with $N_j$ RSS sensors) observation model in $\mathbb{R}^{N_j}$ is given by a subset of $N_j$ RSS observations ($\mathbf{y}_{j,t} = \mathbf{h}_{j,t}(\mathbf{x}_t) + \mathbf{n}_{j,t},$ for $j = 1,\ldots,M$), the goal is that each cluster characterizes the posterior $p(\mathbf{x}_t|\mathbf{y}_{j,1:t})$.

### 2.1. Gaussian Information Filtering Background

The information filter (IF) is an algebraically equivalent form of the Kalman filter (KF), where instead of considering the state vector and its associated estimation error covariance, the filter propagates the so-called information vector and information matrix (i.e., the inverse of the covariance). The main advantage is in terms of information fusion, because the aggregation of information provided by different clusters of sensors is just a sum of individual information vectors [4, 5]. Considering linear/Gaussian systems, i.e., $\mathbf{f}_{t-1}(\mathbf{x}_{t-1}) = \mathbf{F}_{t-1}\mathbf{x}_{t-1}$, $\mathbf{h}_t(\mathbf{x}_t) = \mathbf{H}_t\mathbf{x}_t$ and $\mathbf{n}_t \sim \mathcal{N}(\mathbf{0},\mathbf{R}_t)$, the KF computes the predicted and updated state estimates at discrete time $t$ as $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_{t-1}\hat{\mathbf{x}}_{t-1|t-1}$ and $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}\right)$, with $\mathbf{K}_t$ the Kalman gain, $\boldsymbol{\Sigma}_{x,t|t-1}$ and $\boldsymbol{\Sigma}_{x,t|t}$ the prediction and estimation error covariance matrices, respectively. To reformulate the KF as an IF, we define the information vector and matrix as,

$$\hat{\mathbf{z}}_{t|t} = \boldsymbol{\Sigma}^{-1}_{x,t|t}\hat{\mathbf{x}}_{t|t} = \mathbf{Z}_{t|t}\hat{\mathbf{x}}_{t|t}\ ;\ \mathbf{Z}_{t|t} = \boldsymbol{\Sigma}^{-1}_{x,t|t}, \tag{7}$$

and then the standard KF recursion is rewritten as

$$\hat{\mathbf{z}}_{t|t-1} = \mathbf{L}_t\hat{\mathbf{z}}_{t-1|t-1} \tag{8}$$

$$\mathbf{Z}_{t|t-1} = \left(\mathbf{F}_{t-1}\mathbf{Z}^{-1}_{t-1|t-1}\mathbf{F}^{\top}_{t-1} + \mathbf{Q}_{t-1}\right)^{-1}, \tag{9}$$

$$\hat{\mathbf{z}}_{t|t} = \hat{\mathbf{z}}_{t|t-1} + \mathbf{i}_t\ ;\ \mathbf{Z}_{t|t} = \mathbf{Z}_{t|t-1} + \mathcal{I}_t. \tag{10}$$

with $\mathbf{L}_t = \mathbf{Z}_{t|t-1}\mathbf{F}_{t-1}\mathbf{Z}^{-1}_{t|t-1}$, and the information contributions to the updates $\mathbf{i}_t = \mathbf{H}^{\top}_t\mathbf{R}^{-1}_t\mathbf{y}_t$ and $\mathcal{I}_t = \mathbf{H}^{\top}_t\mathbf{R}^{-1}_t\mathbf{H}_t$. Considering a set of observations taken at $M$ different clusters of sensors, each cluster $j$ computes its own estimate and then the global estimate can be updated simply as

$$\hat{\mathbf{z}}_{t|t} = \hat{\mathbf{z}}_{t|t-1} + \sum_{j=1}^{M}\mathbf{i}_{j,t}\ ;\ \mathbf{Z}_{t|t} = \mathbf{Z}_{t|t-1} + \sum_{j=1}^{M}\mathcal{I}_{j,t}, \tag{11}$$

with $\mathbf{i}_{j,t} = \mathbf{H}_{j,t}^\top \mathbf{R}_{j,t}^{-1} \mathbf{y}_{j,t}$ and $\mathcal{I}_{j,t} = \mathbf{H}_{j,t}^\top \mathbf{R}_{j,t}^{-1} \mathbf{H}_{j,t}$. Notice that the filter complexity (e.g., inversion of matrices) is translated from the measurement update to the state prediction, which is substantially lower dimensional in WSN.

For nonlinear/Gaussian systems, deterministic sampling-based information filters (SPIFs) have been proposed [4, 5], where the prediction step can be implemented as in standard sigma-point filters [3], and the information contributions to construct the measurement update at cluster $j$ are given by

$$\mathbf{i}_{j,t} = \mathbf{Z}_{j,t|t-1} \mathbf{\Sigma}_{j,xy,t|t-1} \mathbf{R}_{j,t}^{-1} \left( \mathbf{y}_{j,t} - \hat{\mathbf{y}}_{j,t|t-1} + \mathbf{\Sigma}_{j,xy,t|t-1}^\top \hat{\mathbf{z}}_{j,t|t-1} \right)$$
$$\mathcal{I}_{j,t} = \mathbf{Z}_{j,t|t-1} \mathbf{\Sigma}_{j,xy,t|t-1} \mathbf{R}_{j,t}^{-1} \left( \mathbf{Z}_{j,t|t-1} \mathbf{\Sigma}_{j,xy,t|t-1} \right)^\top ,$$

where the measurement prediction and cross-covariance are $\hat{\mathbf{y}}_{j,t|t-1} \approx \sum_{i=1}^{L} \omega_i \mathbf{h}_{j,t}(\mathbf{x}_{i,t|t-1})$ and $\mathbf{\Sigma}_{j,xy,t|t-1} \approx \sum_{i=1}^{L} \omega_i \mathbf{x}_{i,t|t-1} \mathbf{h}_{j,t}^\top(\mathbf{x}_{i,t|t-1}) - \hat{\mathbf{x}}_{j,t|t-1} \hat{\mathbf{y}}_{j,t|t-1}^\top$, respectively, with $\{\boldsymbol{\xi}_i, \omega_i\}_{i=1,\dots,L}$ a set of sigma-points and weights, $\mathbf{x}_{i,t|t-1} = \mathbf{S}_{j,x,t|t-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{j,t|t-1}$, and $\mathbf{S}_{j,x,t|t-1}$ the square-root Cholesky factorization of $\mathbf{\Sigma}_{j,x,t|t-1}$. Notice that after computing the global estimate at the fusion center, the information vector and matrix are retransmitted to each cluster.

### 2.2. Decentralized IF under Skew-Laplace Noise

Considering the nonlinear/non-Gaussian SSM of interest (5)-(6), the proposed decentralized IF exploits the NVMM respresentation of the skew-Laplace distribution introduced in Section 1.3, which allows to resort to Gaussian filtering techniques. The SPIF core considers a Gaussian noise, then the filter requires at every time step an estimate of the noise random mean and variance. The key point is to marginalize the unknown non-Gaussian noise latent variable (i.e., $\beta$ in (4)) within the SPIF framework. For the marginalization we need to obtain an expression for the posterior distribution of the latent variable. Under Gaussianity the marginal predictive and posterior densities are given by

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) \approx \mathcal{N}\left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \mathbf{\Sigma}_{x,t|t-1}\right), \quad (12)$$
$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \mathcal{N}\left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}, \mathbf{\Sigma}_{x,t|t}\right), \quad (13)$$

but we have to take into account that the independent noise components are univariate skew-Laplace distributed,

$$\mathbf{n}_t = [n_{1,t}, ..., n_{N,t}]^\top \ ; \ n_{i,t} \sim \mathcal{SL}(n_{i,t}; \mu_i, \sigma_i, \lambda_i), \quad (14)$$
$$n_{i,t}|\beta_{i,t} \sim \mathcal{N}\left(n_{i,t}; \mu_i + \beta_{i,t}\lambda_i, \beta_{i,t}\sigma_i^2\right), \quad (15)$$
$$\beta_{i,t} \sim \mathcal{G}\left(\beta_{i,t}; 1, 2\right) \ ; \ \boldsymbol{\psi}_t = \{\beta_{i,t}\}|_{i=1,\dots,N}, \quad (16)$$
$$\mathbf{n}_t|\boldsymbol{\psi}_t \sim \mathcal{N}\left(\mathbf{n}_t; \mathbf{m}_t(\boldsymbol{\psi}_t), \mathbf{R}_t(\boldsymbol{\psi}_t)\right), \quad (17)$$
$$[\mathbf{m}_t(\boldsymbol{\psi}_t)]_i = \mu_i + \beta_{i,t}\lambda_i \ ; \ [\mathbf{R}_t(\boldsymbol{\psi}_t)]_{i,i} = \beta_{i,t}\sigma_i^2. \quad (18)$$

In this case, we can write the marginalized posterior as

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \int p(\mathbf{x}_t|\boldsymbol{\psi}_t, \mathbf{y}_{1:t}) p(\boldsymbol{\psi}_t|\mathbf{y}_{1:t}) d\boldsymbol{\psi}_t, \quad (19)$$

then the key point is to obtain the noise latent variables posterior, $p(\boldsymbol{\psi}_t|\mathbf{y}_{1:t})$. Observations are independent, then

$$p(\boldsymbol{\psi}_t|\mathbf{y}_{1:t}) = \prod_i p(\beta_{i,t}|y_{i,1:t}), \quad (20)$$
$$p(\beta_{i,t}|y_{i,1:t}) = p(\beta_{i,t}|y_{i,t}) \propto p(y_{i,t}|\beta_{i,t}) p(\beta_{i,t}), \quad (21)$$
$$y_{i,t}|\beta_{i,t} \sim \mathcal{N}(y_{i,t}; \mathbf{h}_{i,t}(\mathbf{x}_t) + \mu_i + \beta_{i,t}\lambda_i, \beta_{i,t}\sigma_i^2). \quad (22)$$

Recall that for a gamma distributed random variable $z \sim \mathcal{G}(z; m, s)$, where $z > 0$,

$$\mathcal{G}(z; m, s) = \frac{1}{\Gamma(m) s^m} z^{m-1} e^{-z/s}, \quad (23)$$

with shape and scale parameters $m$ and $s$, respectively, then $\mathbb{E}[z] = ms$ and $\mathrm{Var}[z] = ms^2$. Also, it is important to notice that the gamma distribution is a special case of the generalized inverse Gaussian distribution,

$$\mathcal{GIG}(z; p, a, b) = \frac{(a/b)^{p/2}}{2K_p(\sqrt{ab})} z^{p-1} e^{-(az+b/z)/2}, \quad (24)$$

with parameters $p \in \mathbb{R}$, $a > 0$, $b > 0$, $z > 0$ and $K_p(\cdot)$ a modified Bessel function. For $p > 0$, we obtain that

$$\mathcal{G}(z; m, s) = \mathcal{GIG}(z; m, 2/s, 0). \quad (25)$$

Interestingly, the generalized inverse Gaussian distribution is the conjugate prior to the Gaussian likelihood within the NVMM formulation, that is, prior, likelihood and posterior are related as

$$p(z|p, a, b) = \mathcal{GIG}(z; p, a, b),$$
$$p(x|z, c, d) = \mathcal{N}(x; c + dz, z),$$
$$p(z|x, p, a, b, c, d) = \mathcal{GIG}(z; p - 1/2, a + d^2, b + (x - c)^2),$$

then, for the particular case of the $\mathcal{G}(\cdot)$ mixing distribution,

$$p(z|m, s) = \mathcal{G}(z; m, s) = \mathcal{GIG}(z; m, 2/s, 0),$$
$$p(x|z, c, d) = \mathcal{N}(x; c + dz, z),$$
$$p(z|x, m, s, c, d) = \mathcal{GIG}(z; m - 1/2, 2/s + d^2, (x - c)^2).$$

Going back to our problem,

$$p(\beta_{i,t}) = \mathcal{G}(\beta_{i,t}; m, s) = \mathcal{GIG}(\beta_{i,t}; m, 2/s, 0), \quad (26)$$
$$p(y_{i,t}|\beta_{i,t}) = \mathcal{N}(y_{i,t}; \mathbf{h}_{i,t}(\mathbf{x}_t) + \mu_i + \beta_{i,t}\lambda_i, \beta_{i,t}\sigma_i^2). \quad (27)$$

We can define a normalized observation, $\tilde{y}_{i,t} \triangleq \frac{y_{i,t}}{\sigma_i}$, and the corresponding normalized likelihood is

$$p(\tilde{y}_{i,t}|\beta_{i,t}) = \mathcal{N}\left(\tilde{y}_{i,t}; \frac{\mathbf{h}_{i,t}(\mathbf{x}_t) + \mu_i}{\sigma_i} + \frac{\lambda_i}{\sigma_i}\beta_{i,t}, \beta_{i,t}\right) \quad (28)$$

The mixing variable posterior distribution is given by

$$p(\beta_{i,t}|y_{i,t}) = \mathcal{GIG}(\beta_{i,t}; p_{i,t}, a_{i,t}, b_{i,t}), \quad (29)$$

with

$$p_{i,t} = m - \frac{1}{2} = \frac{1}{2} \; ; \; a_{i,t} = \frac{2}{s} + \left(\frac{\lambda_i}{\sigma_i}\right)^2 = 1 + \left(\frac{\lambda_i}{\sigma_i}\right)^2$$

$$b_{i,t} = \left(\frac{y_{i,t}}{\sigma_i} - \frac{\mathbf{h}_{i,t}(\mathbf{x}_t) + \mu_i}{\sigma_i}\right)^2, \tag{30}$$

and the posterior mode can be computed as

$$\text{Mode}[\beta_{i,t}|y_{i,t}] = \frac{(p_{i,t}-1) + \sqrt{(p_{i,t}-1)^2 + a_{i,t} b_{i,t}}}{a_{i,t}},$$

from which we can obtain a point estimate of $\hat{\beta}_{i,t}$ at each time step replacing $\mathbf{x}_t$ by $\hat{\mathbf{x}}_{i,t|t-1}$, which can be used in the update step of each filter using

$$[\mathbf{m}_t(\boldsymbol{\psi}_t)]_i \approx \mu_i + \hat{\beta}_{i,t}\lambda_i \; ; \; [\mathbf{R}_t(\boldsymbol{\psi}_t)]_{i,i} \approx \hat{\beta}_{i,t}\sigma_i^2. \tag{31}$$

Notice that the final information contributions in (11) are obtained from these estimates as

$$\mathbf{i}_{j,t} = \mathbf{Z}_{j,t|t-1}\boldsymbol{\Sigma}_{j,xy,t|t-1}[\mathbf{R}_{j,t}(\boldsymbol{\psi}_t)]^{-1} \left(\mathbf{y}_{j,t} - \hat{\mathbf{y}}_{j,t|t-1} - \mathbf{m}_{j,t}(\boldsymbol{\psi}_t) + \boldsymbol{\Sigma}_{j,xy,t|t-1}^{\top}\hat{\mathbf{z}}_{j,t|t-1}\right), \tag{32}$$

$$\mathcal{I}_{j,t} = \mathbf{Z}_{j,t|t-1}\boldsymbol{\Sigma}_{j,xy,t|t-1}[\mathbf{R}_{j,t}(\boldsymbol{\psi}_t)]^{-1} \times \left(\mathbf{Z}_{j,t|t-1}\boldsymbol{\Sigma}_{j,xy,t|t-1}\right)^{\top}. \tag{33}$$

Instead of directly using (30) at each time step, we can use the posterior at $t-1$ as the prior at $t$, then (30) are used for the initial $\beta_{i,1}$, and the posterior update becomes

$$p_{i,t} = p_{i,t-1} - \frac{1}{2} \; ; \; a_{i,t} = a_{i,t-1} + \left(\frac{\lambda_i}{\sigma_i}\right)^2, \tag{34}$$

$$b_{i,t} = b_{i,t-1} + \left(\frac{y_{i,t}}{\sigma_i} - \frac{\mathbf{h}_{i,t}(\mathbf{x}_t) + \mu_i}{\sigma_i}\right)^2. \tag{35}$$

## 3. RESULTS

An illustrative 2D RSS object tracking example will be used to support the discussion and to show the performance improvement of the proposed algorithm, with respect to the standard centralized Cubature KF (CKF) [3] and decentralized Cubature IF (CIF) [5], considering a zero-mean Gaussian noise. As already stated in Section 1.1, an object is localized using a set of $N$ RSS sensors ($M$ clusters of $N_j$ sensors), for which we estimate its 2D position and velocity, $\mathbf{x}_t = [p_{x,t}, p_{y,t}, v_{x,t}, v_{y,t}]^{\top}$. The object constant velocity model is given in (1), and the observation model in (2). As a measure of performance, the root mean square error (RMSE) on the 2D position estimate is obtained from 100 Monte Carlo runs. We consider the following setup: $T_s = 1$s, $\sigma_{p_x}^2 = \sigma_{p_y}^2 = 0.025$ and $\sigma_{v_x}^2 = \sigma_{v_y}^2 = 0.1$. The filters are initialized at $\hat{\mathbf{x}}_0 \sim \mathcal{N}(\mathbf{x}_0, \boldsymbol{\Sigma}_0)$, $\boldsymbol{\Sigma}_0 = \text{diag}(10, 10, 1, 1)$.

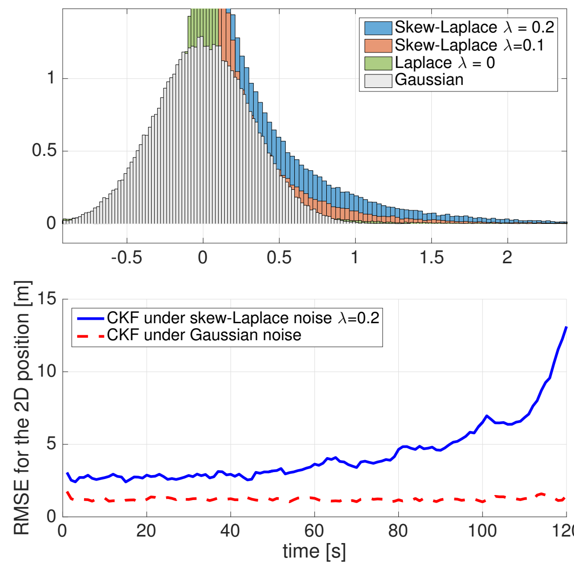


**Fig. 1**. (Top) Empirical pdf for a skew-Laplace distribution ($\lambda = 0.2, 0.1$), a Laplace distribution ($\lambda = 0$) and the corresponding Gaussian distribution ($\mu = 0, \sigma^2 = 0.1$). (Bottom) CKF performance considering Gaussian and a skew-Laplace noise ($\lambda = 0.2$).

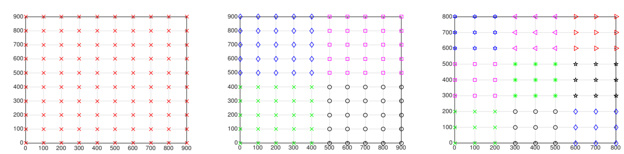

**Fig. 2**. 3 tested scenarios: (left) set of 100 sensors, (middle) 4 clusters of 25 sensors, and (right) 9 clusters of 9 sensors.

In the case of the CKF and CIF, these filters assume a zero-mean Gaussian noise with $\sigma_m^2 = 0.1$ for all sensors. For the skew-Laplace distribution ($\mu = 0, \sigma^2 = 0.1$, $\beta \sim \mathcal{G}(\beta; 1, 2)$), an example of the empirical pdf obtained for different values of the skewness parameter ($\lambda = 0, 0.1, 0.2$) is shown in Fig. 1 (top). We can clearly see that both Laplace and skew-Laplace distributions have much heavier tails. To further motivate the need for improved Gaussian filtering techniques under non-Gaussian noise, Fig. 1 (bottom) shows the RMSE obtained with the CKF for both Gaussian and a skew-Laplace noise with $\lambda = 0.2$.

The new distributed SL-CIF taking into account the skew-Laplace measurement noise (SL-CIF) is compared to the corresponding (Gaussian) CKF/CIF in the following scenarios (see Fig. 2): i) CKF using 100 sensors in a 900 × 900 m$^2$ grid, ii) CIF and SL-CIF using $M = 4$ clusters of 25 sensors, and iii) CIF and SL-CIF using $M = 9$ clusters of 9 sensors.

Notice that as a performance benchmark, we compare the results to the optimal CKF (CKF$_{\text{opt}}$), that is, a filter using the full set of sensors, and the true noise being Gaussian.

Figure 3 plots the results RMSE of the 2D position ob-

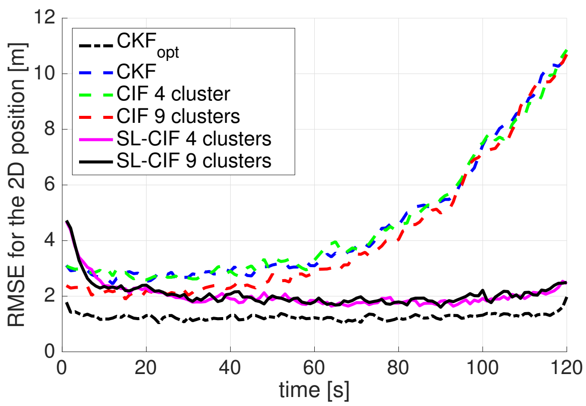


**Fig. 3**. RMSE for 2 targets (top) and 3 targets (bottom), considering the centralized MGF, and both decentralized DMGIF-1 and DMGIF-2.

tained for the three different localization scenarios using the Gaussian CKF/CIF and the new SL-CIF. Notice that for the distributed filters using different clusters of sensors, first each cluster process the information independently, and then a fusion center estimates the global estimate, which in turn is shared with the different filters for the subsequent filtering step. To avoid the communication feedback an alternative (not considered in this contribution) would be that different filters work independently from other clusters, and only the fusion center obtains a global estimate, which may slightly degrade the performance depending on the WSN geometry.

Regarding the results, first we can see that the performance obtained with Gaussian filters (i.e., CKF and CIF), not taking into account the non-Gaussian noise, clearly degrades with respect to the optimal Gaussian performance, and the filters diverge due to the heavy-tailed nature of the skew-Laplace noise. For the new SL-CIF, even if there is a small performance loss with respect to the optimal, the filter is able to mitigate the impact of the non-Gaussian noise within the Gaussian filter, which confirms the validity of the proposed formulation and the good behavior of the new methodology.

## 4. CONCLUSIONS

This article presented a new computationally light distributed sigma-point information filter able to cope with non-Gaussian skew-Laplace distributed measurement noise. This may be a promising solution for RSS-based localization in large sensor networks, where the skew-Laplace noise allows to correctly model challenging indoor propagation conditions. The key idea behind the new information filter is to exploit the hierarchically Gaussian formulation of the skew-Laplace distribution, which allows to resort to Gaussian filtering techniques, and the analytic expression of the mixing distribution posterior obtained by conjugate prior analysis, to cope with the underlying non-Gaussianity. An illustrative example was given to support the discussion and show the promising capabilities of such distributed filtering approach.

## 5. REFERENCES


[1] P. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 140, no. 2, pp. 19–38, Sept. 2003.

[2] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

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

[4] T. Vercauteren and X. Wang, "Decentralized sigma-point information filters for target tracking in collaborative sensor networks," *IEEE Trans. on Sig. Process.*, vol. 53, no. 8, pp. 2997–3009, Aug. 2005.

[5] I. Arasaratnam and K.P.B. Chandra, *Multisensor Data Fusion: From Algorithm and Architecture Design to Applications*, chapter Cubature Information Filters: Theory and Applications to Multisensor Fusion, CRC Press, 2015.

[6] J. Vilà-Valls and P. Closas, "NLOS mitigation in indoor localization by marginalized Monte Carlo Gaussian smoothing," *EURASIP Journal on Adv. in Sig. Process.*, vol. 62, Aug. 2017.

[7] L. Carlino et al., "Robust distributed cooperative RSS-based localization for directed graphs in mixed LoS/NLoS environments," *EURASIP Journal on Wireless Communications and Networking*, vol. 19, 2019.

[8] J. Wilson and N. Patwari, "A fade-level skew-Laplace signal strength model for device-free localization with wireless networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 947–958, June 2012.

[9] O. Arslan, "An alternative multivariate skew Laplace distribution: properties and estimation," *Statistical Papers*, vol. 51, no. 4, pp. 865–887, 2010.