A New Outlier-Robust Student's t based Gaussian Approximate Filter for Cooperative Localization

Yulong Huang, Yonggang Zhang, *Senior Member, IEEE,* Bo Xu, Zhemin Wu, Jonathon Chambers, *Fellow, IEEE*

***Abstract*—In this paper, a new outlier-robust Student's t based Gaussian approximate filter is proposed to address the heavy-tailed process and measurement noises induced by the outlier measurements of velocity and range in cooperative localization of autonomous underwater vehicles (AUVs). The state vector, scale matrices and degrees of freedom (dof) parameters are estimated based on the variational Bayesian approach by using the constructed Student's t based hierarchical Gaussian state-space model. The performances of the proposed filter and existing filters are tested in the cooperative localization of an AUV through a lake trial. Experimental results illustrate that the proposed filter has better localization accuracy and robustness than existing state-of-the-art outlier-robust filters.**

***Index Terms*—Autonomous underwater vehicles, cooperative localization, outlier, heavy-tailed noise, Student's t distribution, variational Bayesian, nonlinear filtering**

I. Introduction

LOCALIZATION of autonomous underwater vehicles (AUVs) has always been a challenging problem due to rapid attenuation of radio-frequency and global positioning system (GPS) signals [1], [2]. Inspired by the cooperative idea in [3], an acoustic range measurement based cooperative localization scheme has been proposed to solve this problem [4]. A few surface craft or AUVs that serve as leaders are equipped with high-accuracy navigation systems, and they aid the remaining AUVs with low-accuracy dead-reckoning (DR) systems by transmitting the relative range measurements through acoustic modems so that the localization errors are bounded [5], [6].

In the cooperative localization of AUVs, an important problem is estimating the positions of the AUVs. A filtering technique based on a state-space model is an effective way to complete this task, which can achieve statistically optimal state estimates [7]–[9]. A large number of nonlinear filters for cooperative localization of AUVs have been proposed, such as the extended Kalman filter [4], unscented Kalman filter (UKF) [10], and moving horizon estimation algorithm [6], [11]. In practical cooperative localization, outlier measurements of velocity and range may occur, which can induce heavy-tailed non-Gaussian process and measurement noises respectively. The outlier measurements of velocity occur when the Doppler velocity log (DVL) is physically misaligned with the body framework or when water lock occurs [12]. The outlier measurements of range are predominantly caused by the multiple acoustic propagation paths between source and receiver that are induced by the refraction of sound waves due to changes in sound speed with depth and reflections off the surface and sea bed [13]. However, these nonlinear filters mentioned above may fail when outlier measurements of velocity and range occur since they are specially designed for Gaussian noises so that they are sensitive to heavy-tailed non-Gaussian noises.

To solve the outlier problem, the Huber-based nonlinear Kalman filter (HNKF) has been derived based on the Huber technique and statistical linearized method [14], [15]. However, the influence function of the HNKF doesn't redescend, which may result in limited estimation accuracy. To cope with large outliers, the maximum correntropy criterion based Kalman filter (MCCKF) has been proposed by maximising the correntropy of the predicted error and residual [16]. However, there is a lack of theoretical basis to develop the estimation error covariance matrix, which degrades the estimation accuracy. A reasonable approach to improve the estimation accuracy is employing the Student's t distribution to model the heavy-tailed non-Gaussian noises. The Student's t distribution is a generalized Gaussian distribution but has heavier tails than the Gaussian distribution, which makes it more suitable for modelling the heavy-tailed non-Gaussian noise [17]. On one hand, a robust Student's t based nonlinear filter (RSTNF) has been derived by approximating the posterior probability density function (PDF) as a Student's t distribution [18], [19]. However, in the RSTNF, the growth of the degrees of freedom (dof) parameter must be prevented to preserve the heavy-tailed properties and closed Student's t-distributed form of the posterior PDF, which may degrade the filtering performance. On the other hand, a robust Student's t based Kalman filter (RSTKF) has been proposed by approximating the posterior PDF as a Gaussian distribution based on the variational Bayesian (VB) approach [20]. The performance of RSTKF depends heavily on the chosen scale matrices and dof parameters of the Student's t distributions. However, in practical cooperative localization, it is very difficult to determine the true scale matrices and dof parameters, which may degrade the performance of RSTKF.

In this paper, to better address the heavy-tailed process and measurement noises induced by the outlier measurements of

This work was supported by the National Natural Science Foundation of China under Grant No. 61371173 and the Fundamental Research Funds for the Central Universities of Harbin Engineering University No. HEUCF160401 and the Engineering and Physical Sciences Research Council (EPSRC) of the UK grant no. EP/K014307/1. Corresponding author is Y. G. Zhang.

Y. L. Huang, Y. G. Zhang and B. Xu are with the Department of Automation, Harbin Engineering University, Harbin 150001, China (e-mail: heuedu@163.com; zhangyg@hrbeu.edu.cn; xubocarter@sina.com).

Z. M. Wu is with the School of Electrical Engineering & Automation, Harbin Institute of Technology, Harbin 150080, China (e-mail: myemail-abc@163.com).

J. Chambers is with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK (e-mail: Jonathon.Chambers@newcastle.ac.uk).

velocity and range in cooperative localization of AUVs, the one-step predicted PDF and likelihood PDF are modelled as Student's t distributions, and the conjugate prior distributions of unknown scale matrices and degrees of freedom (dof) parameters are chosen as inverse Wishart and Gamma distributions respectively, based on which a new Student's t based hierarchical Gaussian state-space model is constructed. A new outlier-robust Student's t based Gaussian approximate (GA) filter is derived by using the constructed hierarchical Gaussian state-space model, where the state vector, scale matrices and dof parameters are estimated based on the VB approach. The performances of the proposed filter and existing filters are tested in the cooperative localization of an AUV through a lake trial. Experimental results illustrate that the proposed filter has better localization accuracy and robustness than existing state-of-the-art outlier-robust filters.

II. Theoretical framework

A. Problem formulation

We consider the master-slave mode based cooperative localization of AUVs, where the leader is used as a communication and navigation aid (CNA). Since the depths of the CNA and AUV can be observed by a precise pressure sensor, a three-dimensional (3-D) cooperative localization problem can be simplified to a two-dimensional (2-D) cooperative localization problem. The state-space model of the acoustic range measurement based cooperative localization system is formulated as [4]

$$\begin{cases} x_k = x_{k-1} + \Delta t(\hat{v}_k \cos\hat{\theta}_k + \hat{\omega}_k \sin\hat{\theta}_k) + w_{x,k-1} \\ y_k = y_{k-1} + \Delta t(\hat{v}_k \sin\hat{\theta}_k - \hat{\omega}_k \cos\hat{\theta}_k) + w_{y,k-1} \end{cases} \tag{1}$$

$$z_k = \sqrt{(x_k - x_k^r)^2 + (y_k - y_k^r)^2} + \delta_k \tag{2}$$

where x_k and y_k are respectively the east and north positions of the AUV at discrete time k, Δt is the sampling time, $\hat{\omega}_k$ and $\hat{v}_k$ are respectively the starboard and forward velocities in the body framework provided by the DVL, $\hat{\theta}_k$ is the heading measured by the compass, $\mathbf{x}_k = [x_k,\ y_k]^T$ denotes the position vector of the AUV, $\mathbf{w}_k = [w_{x,k}, w_{y,k}]^T$ denotes the process noise vector, z_k is the relative range between the CNA and AUV and it is is measured by the time of arrival method using an acoustic modem, $\mathbf{x}_k^r = [x_k^r,\ y_k^r]^T$ is the position of the CNA at time k provided periodically by the acoustic modem, and δ_k represents the measurement noise.

Using (1)-(2), a general discrete state-space model of the cooperative localization is formulated as

$$\begin{cases} \mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{u}_k + \mathbf{w}_{k-1} & \text{(Process equation)} \\ z_k = \mathbf{h}(\mathbf{x}_k) + \delta_k & \text{(Measurement equation)} \end{cases} \tag{3}$$

where the state transition matrix $\mathbf{F} = \mathbf{I}_2$, $\mathbf{I}_2$ denotes the 2-D identity matrix, control input $\mathbf{u}_k = [\Delta t(\hat{v}_k \cos\hat{\theta}_k + \hat{\omega}_k \sin\hat{\theta}_k),\ \ \Delta t(\hat{v}_k \sin\hat{\theta}_k - \hat{\omega}_k \cos\hat{\theta}_k)]^T$, and measurement function $\mathbf{h}(\mathbf{x}_k) = \sqrt{(x_k - x_k^r)^2 + (y_k - y_k^r)^2}$.

In practical cooperative localization, the outlier measurements of velocity and range may induce heavy-tailed non-Gaussian process and measurement noises respectively, which may degrade the performance of existing cooperative localization algorithms. To solve this problem, next a new outlier-robust Student's t based GA filter for cooperative localization will be proposed based on the VB approach.

B. Student's t based hierarchical Gaussian state-space model

Building upon our earlier work in [20], to address the heavy-tailed process and measurement noises, the one-step predicted PDF $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ and the likelihood PDF $p(z_k|\mathbf{x}_k)$ are modelled as Student's t distributions as follows

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \mathrm{St}(\mathbf{x}_k; \mu_k, \mathbf{\Sigma}_k, \sigma_k) = \int_0^{+\infty} \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{\Sigma}_k/\xi_k)\mathrm{G}(\xi_k; \frac{\sigma_k}{2}, \frac{\sigma_k}{2})d\xi_k \tag{4}$$

$$p(z_k|\mathbf{x}_k) = \mathrm{St}(z_k; \mathbf{h}(\mathbf{x}_k), R_k, \nu_k) = \int_0^{+\infty} \mathrm{N}(z_k; \mathbf{h}(\mathbf{x}_k), R_k/\lambda_k)\mathrm{G}(\lambda_k; \frac{\nu_k}{2}, \frac{\nu_k}{2})d\lambda_k \tag{5}$$

where $\mathbf{z}_{1:l} \triangleq \{z_k\}_{k=1}^l$ denotes the set of measurements from time 1 to l, and $\mathrm{St}(\cdot; \mu, \mathbf{\Sigma}, \nu)$ denotes a Student's t PDF with mean vector μ, scale matrix $\mathbf{\Sigma}$ and dof parameter ν, and $\mathrm{N}(\cdot; \mu, \mathbf{\Sigma})$ denotes a Gaussian PDF with mean vector μ and covariance matrix $\mathbf{\Sigma}$, and $\mathrm{G}(\cdot; \alpha, \beta)$ denotes a Gamma PDF with shape parameter α and rate parameter β, and ξ_k and λ_k are auxiliary random variables, and the mean vector μ_k of $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ is given by

$$\mu_k = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{u}_k \tag{6}$$

where $\hat{\mathbf{x}}_{k-1|k-1}$ is the filtering estimation at time $k-1$.

According to (4)-(5), the one-step predicted PDF $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ and the likelihood PDF $p(z_k|\mathbf{x}_k)$ can be rewritten in the following hierarchical Gaussian forms

$$\begin{cases} p(\mathbf{x}_k|\xi_k, \mathbf{z}_{1:k-1}) = \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{\Sigma}_k/\xi_k) \\ p(\xi_k) = \mathrm{G}(\xi_k; \frac{\sigma_k}{2}, \frac{\sigma_k}{2}) \\ p(z_k|\mathbf{x}_k, \lambda_k) = \mathrm{N}(z_k; \mathbf{h}(\mathbf{x}_k), R_k/\lambda_k) \\ p(\lambda_k) = \mathrm{G}(\lambda_k; \frac{\nu_k}{2}, \frac{\nu_k}{2}) \end{cases} \tag{7}$$

However, in practical cooperative localization, the scale matrices R_k and $\mathbf{\Sigma}_k$ and dof parameters ν_k and σ_k can't be determined based on the moment matching method since the true process and measurement noise covariance matrices are unknown as a result of the outlier measurements of velocity and range. To solve this problem, as a contribution in this paper, the scale matrices and dof parameters will be adaptively estimated based on the VB approach. A conjugate prior distribution should be chosen for the unknown parameter in the Bayesian inference because the conjugacy can guarantee that the posterior distribution has the same form as the prior distribution. In Bayesian statistics, the inverse Wishart distribution and Gamma distribution are the conjugate prior distributions for the covariance matrix of a Gaussian distribution and the dof parameter of a Gamma distribution respectively [20], [21], [23]. Since the scale matrix is proportional to the covariance matrix in terms of the hierarchical Gaussian form formulated in (7), the inverse Wishart distribution can also be used as the conjugate prior distribution for the scale matrix. Based on the

above discussions, the conjugate prior distributions for scale matrices $\mathbf{\Sigma}_k$ and R_k and dof parameters σ_k and ν_k are selected as follows

$$\begin{cases} p(\mathbf{\Sigma}_k) = \mathrm{IW}(\mathbf{\Sigma}_k; u_k, \mathbf{U}_k), & p(R_k) = \mathrm{IW}(R_k; t_k, T_k) \\ p(\sigma_k) = \mathrm{G}(\sigma_k; a_k, b_k), & p(\nu_k) = \mathrm{G}(\nu_k; c_k, d_k) \end{cases} \tag{8}$$

where $\mathrm{IW}(\cdot; u, \mathbf{U})$ denotes an inverse Wishart PDF with dof parameter u and inverse scale matrix $\mathbf{U}$.

To capture the prior information of scale matrices, the mean values of $\mathbf{\Sigma}_k$ and R_k are respectively set as the nominal predicted error covariance matrix $\mathbf{P}_{k|k-1}$ and nominal measurement noise covariance matrix $\bar{R}_k$, i.e.

$$\frac{\mathbf{U}_k}{u_k - n - 1} = \mathbf{P}_{k|k-1} \qquad \frac{T_k}{t_k - m - 1} = \bar{R}_k \tag{9}$$

where n and m are respectively the dimensions of the state vector and measurement, and $\mathbf{P}_{k|k-1}$ is given by

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \bar{\mathbf{Q}}_{k-1} \tag{10}$$

where $\mathbf{P}_{k-1|k-1}$ is the filtering estimation error covariance matrix at time $k-1$, and $\bar{\mathbf{Q}}_{k-1}$ is the nominal process noise covariance matrix.

Equations (6)-(10) constitute a Student's t based hierarchical Gaussian state-space model for cooperative localization. Next, the state vector, auxiliary random variables, scale matrices, and dof parameters, i.e., $\mathbf{\Theta}_k \triangleq \{\mathbf{x}_k, \xi_k, \lambda_k, \mathbf{\Sigma}_k, R_k, \sigma_k, \nu_k\}$, will be inferred based on the VB approach by using the constructed Student's t based hierarchical Gaussian state-space model.

C. A new outlier-robust Student's t based GA filter

Our aim is to calculate the joint posterior PDF $p(\mathbf{\Theta}_k|\mathbf{z}_{1:k})$. However, there is not an analytical solution for this posterior PDF based on the hierarchical Gaussian state-space model. Thus, to obtain an approximate solution, the VB approach is used to provide a free form factored approximate PDF for $p(\mathbf{\Theta}_k|\mathbf{z}_{1:k})$, i.e., [24]

$$p(\mathbf{\Theta}_k|\mathbf{z}_{1:k}) \approx q(\mathbf{x}_k)q(\xi_k)q(\lambda_k)q(\mathbf{\Sigma}_k)q(R_k)q(\sigma_k)q(\nu_k) \tag{11}$$

where $q(\cdot)$ is the approximate posterior PDF, and its optimal solution satisfies the following equation

$$\log q(\phi) = \mathrm{E}_{\mathbf{\Theta}_k^{(-\phi)}}[\log p(\mathbf{\Theta}_k, \mathbf{z}_{1:k})] + c_\phi \tag{12}$$

where $\mathrm{E}[\cdot]$ represents the expectation operation, and ϕ is an arbitrary element of $\mathbf{\Theta}$, and $\mathbf{\Theta}^{(-\phi)}$ is the set of all elements in $\mathbf{\Theta}$ except for ϕ, and c_ϕ denotes the constant with respect to variable ϕ. Since the variational parameters are coupled, we need to employ fixed-point iterations to solve (12), where the approximate posterior PDF $q(\phi)$ of the arbitrary element ϕ is updated as $q^{(i+1)}(\phi)$ at the $i+1$th iteration using the approximate posterior PDF $q^{(i)}(\mathbf{\Theta}^{(-\phi)})$ [24]. The iterations converge to a local optimum of (12).

Using the conditional independence properties of the hierarchical Gaussian state-space model, the joint PDF $p(\mathbf{\Theta}_k, \mathbf{z}_{1:k})$ can be factored as

$$\begin{aligned} &p(\mathbf{\Theta}_k, \mathbf{z}_{1:k}) = \mathrm{N}(z_k; \mathbf{h}(\mathbf{x}_k), R_k/\lambda_k)\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{\Sigma}_k/\xi_k) \\ &\mathrm{G}(\xi_k; \frac{\sigma_k}{2}, \frac{\sigma_k}{2})\mathrm{G}(\lambda_k; \frac{\nu_k}{2}, \frac{\nu_k}{2})\mathrm{IW}(\mathbf{\Sigma}_k; u_k, \mathbf{U}_k)\mathrm{IW}(R_k; t_k, T_k) \\ &\mathrm{G}(\sigma_k; a_k, b_k)\mathrm{G}(\nu_k; c_k, d_k)p(\mathbf{z}_{1:k-1}) \end{aligned} \tag{13}$$

Using (13), $\log p(\mathbf{\Theta}_k, \mathbf{z}_{1:k})$ is calculated as

$$\begin{aligned} &\log p(\mathbf{\Theta}_k, \mathbf{z}_{1:k}) = (\frac{m+\nu_k}{2} - 1)\log\lambda_k - \frac{\nu_k}{2}\lambda_k - \frac{\sigma_k}{2}\xi_k + \\ &(\frac{n+\sigma_k}{2} - 1)\log\xi_k - \frac{\lambda_k}{2}(z_k - \mathbf{h}(\mathbf{x}_k))^T R_k^{-1}(z_k - \mathbf{h}(\mathbf{x}_k)) \\ &-\frac{1}{2}\mathrm{tr}(T_k R_k^{-1}) - \frac{1}{2}(t_k + m + 2)\log|R_k| - \frac{1}{2}\mathrm{tr}(\mathbf{U}_k\mathbf{\Sigma}_k^{-1}) \\ &-\frac{1}{2}(n + u_k + 2)\log|\mathbf{\Sigma}_k| - \frac{\xi_k}{2}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\mathbf{\Sigma}_k^{-1} \times \\ &(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \frac{\sigma_k}{2}\log\frac{\sigma_k}{2} - \log\Gamma(\frac{\sigma_k}{2}) + (a_k - 1)\log\sigma_k \\ &-b_k\sigma_k + \frac{\nu_k}{2}\log\frac{\nu_k}{2} - \log\Gamma(\frac{\nu_k}{2}) + (c_k - 1)\log\nu_k - d_k\nu_k \\ &+c_{\mathbf{\Theta}} \end{aligned} \tag{14}$$

1) Computations of approximate posterior PDFs: Let $\phi = \mathbf{x}_k$ and using (14) in (12) yields

$$\begin{aligned} &\log q^{(i+1)}(\mathbf{x}_k) = -0.5\mathrm{E}^{(i)}[\lambda_k](z_k - \mathbf{h}(\mathbf{x}_k))^T\mathrm{E}^{(i)}[R_k^{-1}] \times \\ &(z_k - \mathbf{h}(\mathbf{x}_k)) - 0.5\mathrm{E}^{(i)}[\xi_k](\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\mathrm{E}^{(i)}[\mathbf{\Sigma}_k^{-1}] \times \\ &(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + c_{\mathbf{x}}. \end{aligned} \tag{15}$$

where $q^{(i+1)}(\cdot)$ is the approximation of PDF $q(\cdot)$ at the $i+1$th iteration, and $\mathrm{E}^{(i)}[\rho]$ is the expectation of variable ρ at the ith iteration, and $\mathrm{tr}(\cdot)$ denotes the trace operation.

The modified predicted error covariance matrix $\tilde{\mathbf{P}}_{k|k-1}^{(i)}$ and modified measurement noise covariance matrix $\tilde{R}_k^{(i)}$ are defined as follows

$$\tilde{\mathbf{P}}_{k|k-1}^{(i)} = \frac{\left\{\mathrm{E}^{(i)}[\mathbf{\Sigma}_k^{-1}]\right\}^{-1}}{\mathrm{E}^{(i)}[\xi_k]} \quad \tilde{R}_k^{(i)} = \frac{\left\{\mathrm{E}^{(i)}[R_k^{-1}]\right\}^{-1}}{\mathrm{E}^{(i)}[\lambda_k]} \tag{16}$$

then $q^{(i+1)}(\mathbf{x}_k)$ is calculated as

$$q^{(i+1)}(\mathbf{x}_k) = \frac{\mathrm{N}(z_k; \mathbf{h}(\mathbf{x}_k), \tilde{R}_k^{(i)})\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{P}}_{k|k-1}^{(i)})}{\int \mathrm{N}(z_k; \mathbf{h}(\mathbf{x}_k), \tilde{R}_k^{(i)})\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{P}}_{k|k-1}^{(i)})d\mathbf{x}_k} \tag{17}$$

It is seen from (17) that $q^{(i+1)}(\mathbf{x}_k)$ is not Gaussian due to the nonlinear measurement function $\mathbf{h}(\mathbf{x}_k)$. In this paper, $q^{(i+1)}(\mathbf{x}_k)$ is approximated as a Gaussian PDF based on the existing GA filter [22], i.e.,

$$q^{(i+1)}(\mathbf{x}_k) \approx \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(i+1)}, \mathbf{P}_{k|k}^{(i+1)}) \tag{18}$$

where the mean vector $\hat{\mathbf{x}}_{k|k}^{(i+1)}$ and covariance matrix $\mathbf{P}_{k|k}^{(i+1)}$ are given by

$$\begin{cases} \mathbf{K}_k^{(i+1)} = \mathbf{P}_{xz,k|k-1}^{(i+1)}(P_{zz,k|k-1}^{(i+1)})^{-1} \\ \hat{\mathbf{x}}_{k|k}^{(i+1)} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^{(i+1)}(z_k - \hat{z}_{k|k-1}^{(i+1)}) \\ \mathbf{P}_{k|k}^{(i+1)} = \tilde{\mathbf{P}}_{k|k-1}^{(i)} - \mathbf{K}_k^{(i+1)}P_{zz,k|k-1}^{(i+1)}(\mathbf{K}_k^{(i+1)})^T \end{cases} \tag{19}$$

where $\mathbf{K}_k^{(i+1)}$ denotes the modified Kalman gain, and parameters $\hat{z}_{k|k-1}^{(i+1)}$, $P_{zz,k|k-1}^{(i+1)}$, and $\mathbf{P}_{xz,k|k-1}^{(i+1)}$ are given by

$$\begin{cases} \hat{z}_{k|k-1}^{(i+1)} = \int \mathbf{h}(\mathbf{x}_k)\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{P}}_{k|k-1}^{(i)})d\mathbf{x}_k \\ P_{zz,k|k-1}^{(i+1)} = \int \mathbf{h}(\mathbf{x}_k)\mathbf{h}^T(\mathbf{x}_k)\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{P}}_{k|k-1}^{(i)})d\mathbf{x}_k - \\ \hat{z}_{k|k-1}^{(i+1)}(\hat{z}_{k|k-1}^{(i+1)})^T + \tilde{R}_k^{(i)} \\ \mathbf{P}_{xz,k|k-1}^{(i+1)} = \int \mathbf{x}_k\mathbf{h}^T(\mathbf{x}_k)\mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{P}}_{k|k-1}^{(i)})d\mathbf{x}_k - \\ \hat{\mathbf{x}}_{k|k-1}(\hat{z}_{k|k-1}^{(i+1)})^T \end{cases} \tag{20}$$

Let $\phi = \xi_k$ and using (14) in (12) yields

$$\log q^{(i+1)}(\xi_k) = (\frac{n + \mathrm{E}^{(i)}[\sigma_k]}{2} - 1)\log \xi_k - 0.5\left\{\mathrm{E}^{(i)}[\sigma_k] + \mathrm{tr}(\mathbf{A}_k^{(i+1)}\mathrm{E}^{(i)}[\mathbf{\Sigma}_k^{-1}])\right\}\xi_k + c_\xi \quad (21)$$

where $\mathbf{A}_k^{(i+1)}$ is given by

$$\mathbf{A}_k^{(i+1)} = \mathrm{E}^{(i+1)}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T] \quad (22)$$

Using (21), $q^{(i+1)}(\xi_k)$ is updated as a Gamma PDF with shape parameter α_k^{i+1} and rate parameter β_k^{i+1},

$$\begin{cases} q^{(i+1)}(\xi_k) = \mathrm{G}(\xi_k; \alpha_k^{i+1}, \beta_k^{i+1}) \\ \alpha_k^{i+1} = 0.5(n + \mathrm{E}^{(i)}[\sigma_k]) \\ \beta_k^{i+1} = 0.5\left\{\mathrm{E}^{(i)}[\sigma_k] + \mathrm{tr}(\mathbf{A}_k^{(i+1)}\mathrm{E}^{(i)}[\mathbf{\Sigma}_k^{-1}])\right\} \end{cases} \quad (23)$$

Let $\phi = \lambda_k$ and using (14) in (12) gives

$$\log q^{(i+1)}(\lambda_k) = (\frac{m + \mathrm{E}^{(i)}[\nu_k]}{2} - 1)\log \lambda_k - 0.5\left\{\mathrm{E}^{(i)}[\nu_k] + \mathrm{tr}(\mathbf{B}_k^{(i+1)}\mathrm{E}^{(i)}[R_k^{-1}])\right\}\lambda_k + c_\lambda \quad (24)$$

where $\mathbf{B}_k^{(i+1)}$ is given by

$$\mathbf{B}_k^{(i+1)} = \mathrm{E}^{(i+1)}[(z_k - \mathbf{h}(\mathbf{x}_k))(z_k - \mathbf{h}(\mathbf{x}_k))^T] \quad (25)$$

Using (24), $q^{(i+1)}(\lambda_k)$ is updated as a Gamma PDF with shape parameter γ_k^{i+1} and rate parameter η_k^{i+1},

$$\begin{cases} q^{(i+1)}(\lambda_k) = \mathrm{G}(\lambda_k; \gamma_k^{i+1}, \eta_k^{i+1}) \\ \gamma_k^{i+1} = 0.5(m + \mathrm{E}^{(i)}[\nu_k]) \\ \eta_k^{i+1} = 0.5\left\{\mathrm{E}^{(i)}[\nu_k] + \mathrm{tr}(\mathbf{B}_k^{(i+1)}\mathrm{E}^{(i)}[R_k^{-1}])\right\} \end{cases} \quad (26)$$

Let $\phi = \mathbf{\Sigma}_k$ and using (14) in (12) yields

$$\log q^{(i+1)}(\mathbf{\Sigma}_k) = -\frac{1}{2}(n + u_k + 2)\log|\mathbf{\Sigma}_k| - \frac{1}{2}\mathrm{tr}\left[(\mathbf{U}_k + \mathrm{E}^{(i+1)}[\xi_k]\mathbf{A}_k^{(i+1)})\mathbf{\Sigma}_k^{-1}\right] + c_{\mathbf{\Sigma}} \quad (27)$$

Exploiting (27), $q^{(i+1)}(\mathbf{\Sigma}_k)$ is updated as an inverse Wishart PDF with dof parameter $\hat{u}_k^{(i+1)}$ and inverse scale matrix $\hat{\mathbf{U}}_k^{(i+1)}$,

$$\begin{cases} q^{(i+1)}(\mathbf{\Sigma}_k) = \mathrm{IW}(\mathbf{\Sigma}_k; \hat{u}_k^{(i+1)}, \hat{\mathbf{U}}_k^{(i+1)}) \\ \hat{u}_k^{(i+1)} = u_k + 1 \quad \hat{\mathbf{U}}_k^{(i+1)} = \mathbf{U}_k + \mathrm{E}^{(i+1)}[\xi_k]\mathbf{A}_k^{(i+1)} \end{cases} \quad (28)$$

Let $\phi = R_k$ and using (14) in (12) obtains

$$\log q^{(i+1)}(R_k) = -\frac{1}{2}(m + t_k + 2)\log|R_k| - \frac{1}{2}\mathrm{tr}\left[(T_k + \mathrm{E}^{(i+1)}[\lambda_k]\mathbf{B}_k^{(i+1)})R_k^{-1}\right] + c_R \quad (29)$$

Using (29), $q^{(i+1)}(R_k)$ is updated as an inverse Wishart PDF with dof parameter $\hat{t}_k^{(i+1)}$ and inverse scale matrix $\hat{T}_k^{(i+1)}$,

$$\begin{cases} q^{(i+1)}(R_k) = \mathrm{IW}(R_k; \hat{t}_k^{(i+1)}, \hat{T}_k^{(i+1)}) \\ \hat{t}_k^{(i+1)} = t_k + 1 \quad \hat{T}_k^{(i+1)} = T_k + \mathrm{E}^{(i+1)}[\lambda_k]\mathbf{B}_k^{(i+1)} \end{cases} \quad (30)$$

Let $\phi = \sigma_k$ and using (14) in (12) gives

$$\log q^{(i+1)}(\sigma_k) = 0.5\mathrm{E}^{(i+1)}[\log \xi_k]\sigma_k - 0.5\mathrm{E}^{(i+1)}[\xi_k]\sigma_k + 0.5\sigma_k \log(0.5\sigma_k) - \log\Gamma(0.5\sigma_k) + (a_k - 1)\log\sigma_k - b_k\sigma_k \quad (31)$$

Using Stirling's approximation: $\log\Gamma(0.5\sigma_k) \approx (0.5\sigma_k - 0.5)\log(0.5\sigma_k) - 0.5\sigma_k$ in (31), thus we obtain

$$\log q^{(i+1)}(\sigma_k) = (a_k + 0.5 - 1)\log\sigma_k - \{b_k + 0.5\mathrm{E}^{(i+1)}[\xi_k] + 0.5 - 0.5\mathrm{E}^{(i+1)}[\log\xi_k]\}\sigma_k \quad (32)$$

Employing (32), $q^{(i+1)}(\sigma_k)$ is updated as a Gamma PDF with shape parameter $\hat{a}_k^{(i+1)}$ and rate parameter $\hat{b}_k^{(i+1)}$,

$$\begin{cases} q^{(i+1)}(\sigma_k) = \mathrm{G}(\sigma_k; \hat{a}_k^{i+1}, \hat{b}_k^{i+1}) \\ \hat{a}_k^{(i+1)} = a_k + 0.5 \\ \hat{b}_k^{(i+1)} = b_k + 0.5\mathrm{E}^{(i+1)}[\xi_k] + 0.5 - 0.5\mathrm{E}^{(i+1)}[\log\xi_k] \end{cases} \quad (33)$$

Let $\phi = \nu_k$ and using (14) in equation (12) yields

$$\log q^{(i+1)}(\nu_k) = 0.5\mathrm{E}^{(i+1)}[\log\lambda_k]\nu_k - 0.5\mathrm{E}^{(i+1)}[\lambda_k]\nu_k + 0.5\nu_k\log(0.5\nu_k) - \log\Gamma(0.5\nu_k) + (c_k - 1)\log\nu_k - d_k\nu_k \quad (34)$$

Again, utilising Stirling's approximation: $\log\Gamma(0.5\nu_k) \approx (0.5\nu_k - 0.5)\log(0.5\nu_k) - 0.5\nu_k$ in (34) gives

$$\log q^{(i+1)}(\nu_k) = (c_k + 0.5 - 1)\log\nu_k - \{d_k + 0.5\mathrm{E}^{(i+1)}[\lambda_k] + 0.5 - 0.5\mathrm{E}^{(i+1)}[\log\lambda_k]\}\nu_k \quad (35)$$

Using (35), $q^{(i+1)}(\nu_k)$ is updated as a Gamma PDF with shape parameter $\hat{c}_k^{(i+1)}$ and rate parameter $\hat{d}_k^{(i+1)}$, i.e.,

$$\begin{cases} q^{(i+1)}(\nu_k) = \mathrm{G}(\nu_k; \hat{c}_k^{i+1}, \hat{d}_k^{i+1}) \\ \hat{c}_k^{(i+1)} = c_k + 0.5 \\ \hat{d}_k^{(i+1)} = d_k + 0.5\mathrm{E}^{(i+1)}[\lambda_k] + 0.5 - 0.5\mathrm{E}^{(i+1)}[\log\lambda_k] \end{cases} \quad (36)$$

After fixed point iteration N, the required posterior PDF of the state vector is approximated as

$$q(\mathbf{x}_k) \approx \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(N)}, \mathbf{P}_{k|k}^{(N)}) = \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) \quad (37)$$

where $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are respectively the state estimation and corresponding estimation error covariance matrix.

2) Computation of expectations: Using equations (23), (26), (28), (30), (33) and (36), the required expectations $\mathrm{E}^{(i+1)}[\xi_k]$, $\mathrm{E}^{(i+1)}[\log\xi_k]$, $\mathrm{E}^{(i+1)}[\lambda_k]$, $\mathrm{E}^{(i+1)}[\log\lambda_k]$, $\mathrm{E}^{(i+1)}[\mathbf{\Sigma}_k^{-1}]$, $\mathrm{E}^{(i+1)}[R_k^{-1}]$, $\mathrm{E}^{(i+1)}[\sigma_k]$, and $\mathrm{E}^{(i+1)}[\nu_k]$ are calculated as

$$\begin{cases} \mathrm{E}^{(i+1)}[\xi_k] = \alpha_k^{i+1}/\beta_k^{i+1} \quad \mathrm{E}^{(i+1)}[\log\xi_k] = \psi(\alpha_k^{i+1}) - \log\beta_k^{i+1} \\ \mathrm{E}^{(i+1)}[\lambda_k] = \gamma_k^{i+1}/\eta_k^{i+1} \quad \mathrm{E}^{(i+1)}[\log\lambda_k] = \psi(\gamma_k^{i+1}) - \log\eta_k^{i+1} \\ \mathrm{E}^{(i+1)}[\mathbf{\Sigma}_k^{-1}] = (\hat{u}_k^{(i+1)} - n - 1)(\hat{\mathbf{U}}_k^{(i+1)})^{-1} \\ \mathrm{E}^{(i+1)}[R_k^{-1}] = (\hat{t}_k^{(i+1)} - m - 1)(\hat{T}_k^{(i+1)})^{-1} \\ \mathrm{E}^{(i+1)}[\sigma_k] = \hat{a}_k^{(i+1)}/\hat{b}_k^{(i+1)} \quad \mathrm{E}^{(i+1)}[\nu_k] = \hat{c}_k^{(i+1)}/\hat{d}_k^{(i+1)} \end{cases} \quad (38)$$

where $\psi(\cdot)$ denotes the digamma function [23].

Exploiting (18), the required expectation $\mathbf{A}_k^{(i+1)}$ and $\mathbf{B}_k^{(i+1)}$ are calculated as

$$\begin{cases} \mathbf{A}_k^{(i+1)} = \mathbf{P}_{k|k}^{(i+1)} + (\hat{\mathbf{x}}_{k|k}^{(i+1)} - \hat{\mathbf{x}}_{k|k-1})(\hat{\mathbf{x}}_{k|k}^{(i+1)} - \hat{\mathbf{x}}_{k|k-1})^T \\ \mathbf{B}_k^{(i+1)} = \int (z_k - \mathbf{h}(\mathbf{x}_k))(z_k - \mathbf{h}(\mathbf{x}_k))^T \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(i+1)}, \mathbf{P}_{k|k}^{(i+1)}) d\mathbf{x}_k \end{cases} \quad (39)$$

Algorithm 1: One time step of the proposed outlier-robust GA filter.

Inputs: $\hat{\mathbf{x}}_{k-1|k-1}$, $\mathbf{P}_{k-1|k-1}$, $\mathbf{F}$, $\mathbf{h}(\cdot)$, z_k, $\bar{\mathbf{Q}}_{k-1}$, $\bar{R}_k$, m, n, N, a_k, b_k, c_k, d_k, u_k, t_k

Time update:

1. Calculate μ_k and $\mathbf{P}_{k|k-1}$ using (6) and (10).

Measurement update:

2. Calculate $\mathbf{U}_k$ and T_k using (9), and initial expectations using (38), and initial $\tilde{\mathbf{P}}_{k|k-1}^{(0)}$ and $\tilde{R}_k^{(0)}$ using (16).

for $i = 0 : N-1$

3. Update $q^{(i+1)}(\mathbf{x}_k)$ using (17)-(20).

4. Calculate $\mathbf{A}_k^{(i+1)}$ and $\mathbf{B}_k^{(i+1)}$ using (39).

5. Update $q^{(i+1)}(\xi_k)$ and $q^{(i+1)}(\lambda_k)$ using (23) and (26).

6. Calculate $\mathrm{E}^{(i+1)}[\xi_k]$, $\mathrm{E}^{(i+1)}[\log \xi_k]$, $\mathrm{E}^{(i+1)}[\lambda_k]$ and $\mathrm{E}^{(i+1)}[\log \lambda_k]$ using (38).

7. Update $q^{(i+1)}(\mathbf{\Sigma}_k)$, $q^{(i+1)}(R_k)$, $q^{(i+1)}(\sigma_k)$, and $q^{(i+1)}(\nu_k)$ using (28), (30), (33) and (36).

8. Calculate $\mathrm{E}^{(i+1)}[\mathbf{\Sigma}_k^{-1}]$, $\mathrm{E}^{(i+1)}[R_k^{-1}]$, $\mathrm{E}^{(i+1)}[\sigma_k]$, and $\mathrm{E}^{(i+1)}[\nu_k]$ using (38).

9. Calculate $\tilde{\mathbf{P}}_{k|k-1}^{(i+1)}$ and $\tilde{R}_k^{(i+1)}$ using (16).

end for

10. $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k}^{(N)}, \quad \mathbf{P}_{k|k} = \mathbf{P}_{k|k}^{(N)}$

Outputs: $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$

The Gaussian weighted integral in (39) can't be analytically calculated due to nonlinear measurement function $\mathbf{h}(\mathbf{x}_k)$. Fortunately, existing Gaussian weighted integral rules can be utilised to approximate this integral, such as the unscented transform (UT) rule [10]. The proposed outlier-robust GA filter consists of (6)-(10), (16), (17)-(20), (23), (26), (28), (30), (33) and (36). The implementation pseudo-code for one time step of the proposed filter is summarized in Algorithm 1.

III. Lake-water field trials

A. Experimental setup and description

A lake trial was carried out to verify the effectiveness and superiority of the proposed algorithm. Three survey vessels were employed, where two vessels served as surface leaders known as CNAs and the other one acted as an AUV. The two leaders and AUV were all equipped with an acoustic modem ATM-885, and broadcasted information through the underwater acoustic modem. Fig. 1 illustrates the acoustic communication procedures between the two leaders and AUV in the test, in which "AC" and "CL" denote acoustic communication and cooperative localization respectively. It is seen from Fig. 1 that the AUV only communicated with one of the two leaders at every time. Thus, only a single leader served as the CNA at every time. The diagram of underwater acoustic communication between the AUV and CNA is shown in Fig. 2, and the communication procedures are as follows: (i) a request signal was first sent from the AUV to the CNA; (ii) after receiving the request signal, an acoustic data packet including the arrival time of the request signal and the reference position of the CNA were sent from the CNA to the AUV; (iii) upon

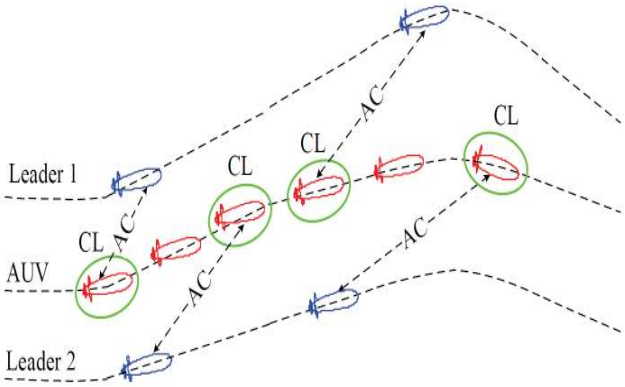

Fig. 1: The acoustic communication procedures between two leaders and AUV.

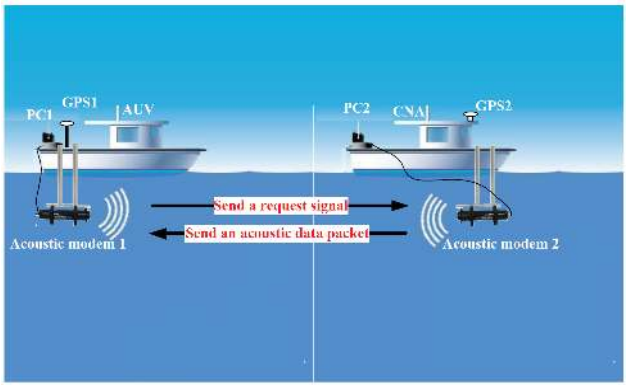

Fig. 2: The diagram of underwater acoustic communication between the AUV and CNA.

Fig. 3: The sensors and computer employed in the experiment.

reception of the acoustic data packet, the AUV can use the time of flight of the acoustic signal to determine the range from the CNA. The CNA was equipped with a GPS to provide a reference position for the AUV, and a GPS enabling the collection of true positions was also present on the AUV to provide a benchmark for cooperative localization. Moreover, a DVL providing velocity information and a compass obtaining a heading were installed in the AUV. The sensors and computer employed in the experiment are shown in Fig. 3, and the performance parameters of the sensors are listed in Table I.

TABLE I: The parameters of the utilised sensors.

Sensors	Index	Parameters
ATM-885	Working range	Up to 8000m
	Error rate	Less than 10^{-7}
GPS	Position accuracy	1.8m (CEP)
	Data update rate	10Hz
Compass	Heading accuracy	0.3°
DVL	Working range	$-150\text{m/s} - 200\text{m/s}$
	Measurement accuracy	$0.1\% - 0.3\%$

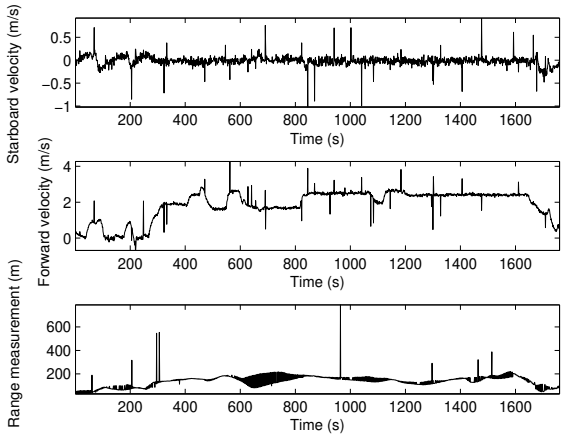

Fig. 4: The observed velocities and range measurement.

In the test, the observed velocity and range measurements are shown in Fig. 4. It is seen from Fig. 4 that there are some outlier measurements of velocity and range. The reasons why the outliers occur have been listed in the introduction. To illustrate the effect of outlier measurements on process and measurement noises, the noise values are calculated as follows

$$\begin{cases} \hat{w}_{x,k-1} = \hat{x}_k - \hat{x}_{k-1} - \Delta t(\hat{v}_k \cos\hat{\theta}_k + \hat{\omega}_k \sin\hat{\theta}_k) \\ \hat{w}_{y,k-1} = \hat{y}_k - \hat{y}_{k-1} - \Delta t(\hat{v}_k \sin\hat{\theta}_k - \hat{\omega}_k \cos\hat{\theta}_k) \\ \hat{\delta}_k = z_k - \sqrt{(\hat{x}_k - x_k^r)^2 + (\hat{y}_k - y_k^r)^2} \end{cases} \quad (40)$$

where $\hat{w}_{x,k-1}$ and $\hat{w}_{y,k-1}$ are respectively the approximate east and north position noise values, and $\hat{\delta}_k$ is the approximate measurement noise value, and $(\hat{x}_k, \hat{y}_k)$ is the position of the AUV at time k provided by the GPS, and the sampling time $\Delta t = 1\text{s}$ in the test. Using (40), a set of approximate process and measurement noise values can be obtained. Fig. 5 shows the probability density curves of process and measurement noises. It is seen from Fig. 5 that Gaussian distributions can't fit to the process and measurement noise values since the process and measurement noises have heavy-tailed distributions. Based on the above discussion, the lake trials can simulate the scenario of cooperative localization of an AUV, and can be used to verify the effectiveness and superiority of the proposed algorithm.

B. Comparison of outlier-robust filtering algorithms

Existing UKF [10], robust Student's t based unscented filter (RSTUF) [18], nonlinear regression Huber Kalman filter

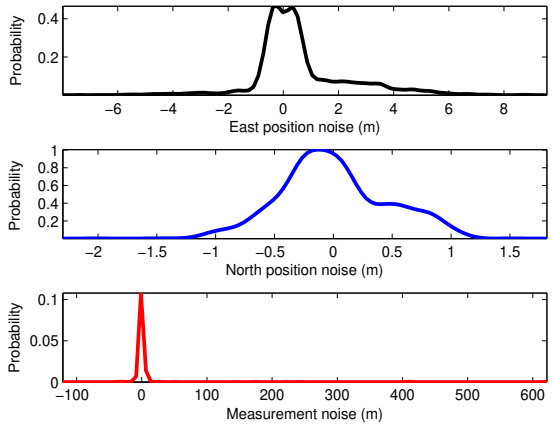

Fig. 5: Probability density curves of process position and measurement noises.

(NRHKF) [14], MCCKF [16], RSTKF [20], and the proposed filter are tested in the cooperative localization of the AUV. The initial state estimate $\hat{\mathbf{x}}_{0|0}$ is provided by GPS, and the initial state estimation error covariance matrix is set as $\mathbf{P}_{0|0} = \mathbf{I}_2$. The nominal process and measurement noise covariance matrices are respectively set as $\bar{\mathbf{Q}}_k = \text{diag}[(0.5\text{m})^2\ (0.5\text{m})^2]$ and $\bar{R}_k = 10\text{m}^2$. In the proposed filter, the prior parameters are set as $a_k = c_k = 5$, $b_k = d_k = 1$, $u_k = 8$ and $t_k = 7$, and the Gaussian weighted integral in (39) is calculated using the UT. In the RSTUF and RSTKF, the scale matrices of process and measurement noises are respectively set as $\bar{\mathbf{Q}}_k$ and $\bar{R}_k$, and the dof parameters are set as $\nu = 5$. Moreover, the tuning parameters of the NRHKF and RSTKF are set as $\gamma = 5$ and $\tau = 5$, and the kernel size of the MCCKF is chosen as $\sigma = 15$, and the numbers of iteration of the NRHKF, RSTKF and the proposed filter are set as $N = 10$. The proposed filter and existing filters are coded with MATLAB and run on a computer with Intel Core i7-3770 CPU at 3.40 GHz.

The localization error (LE) and averaged LE (ALE) are chosen as performance metrics, which are defined as follows

$$\begin{cases} \text{LE}(k) = \sqrt{(\hat{x}_k - \hat{x}_{k|k})^2 + (\hat{y}_k - \hat{y}_{k|k})^2} \\ \text{ALE} = \frac{1}{T}\sum\limits_{k=1}^{T} \sqrt{(\hat{x}_k - \hat{x}_{k|k})^2 + (\hat{y}_k - \hat{y}_{k|k})^2} \end{cases} \quad (41)$$

where $(\hat{x}_k, \hat{y}_k)$ is the reference position of the AUV at time k provided by GPS, $(\hat{x}_{k|k}, \hat{y}_{k|k})$ is the estimated position at time k, and $T = 1760\text{s}$ denotes the experimental time.

The LEs, ALEs and implementation times of the proposed filter and existing filters for a single step run when $N = 10$ are respectively shown in Fig. 6 and Tables II and III. The improvement of ALE (IALE) from the proposed filter as compared with existing filters when $N = 10$ is given in Table IV. The ALEs with different numbers of iterations $N = 1, 2, \ldots, 20$ are shown in Fig. 7. It is seen from Fig. 6 and Tables II and III that the proposed filter has smaller LE and ALE but needs greater implementation times than existing filters when $N = 10$. Also, we can see from Table IV that the ALE from the proposed filter is reduced by 37.9% at least as

TABLE II: ALEs of the proposed filter and existing filters when $N = 10$.

Filters	UKF	RSTUF	NRHKF	MCCKF	RSTKF	The proposed filter
ALE (m)	21.4	17.7	17.3	17.3	49.2	10.7

TABLE III: The implementation times of the proposed filter and existing filters for a single step run when $N = 10$.

Filters	UKF	RSTUF	NRHKF	MCCKF	RSTKF	The proposed filter
Time (ms)	0.167	0.120	0.556	0.074	1.688	1.740

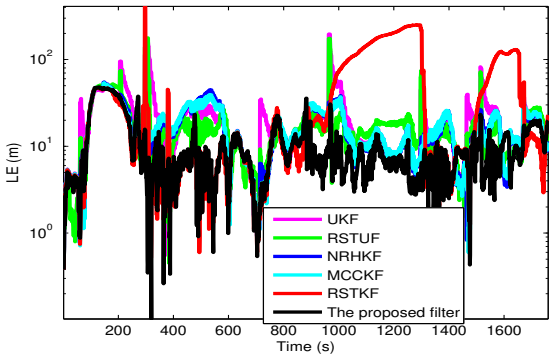

Fig. 6: LEs when $N = 10$.

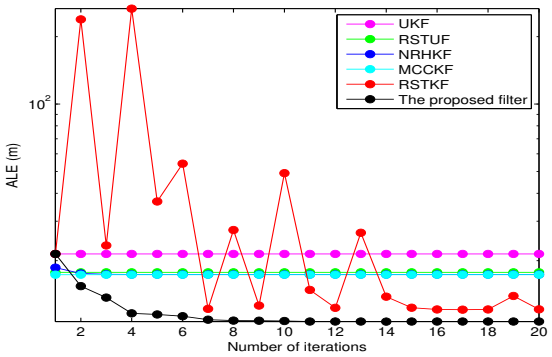

Fig. 7: ALEs when $N = 1, 2, \ldots, 20$.

TABLE IV: The improvement of ALE from the proposed filter as compared with existing filters when $N = 10$.

Filters	UKF	RSTUF	NRHKF	MCCKF	RSTKF
IALE	49.8%	39.3%	37.9%	37.9%	78.1%

compared with existing filters. It can be seen from Fig. 7 that the proposed filter has smaller ALE than existing filters when $N \geq 2$. Moreover, we can see from Fig. 7 that the ALE from the proposed filter converges when $N \geq 8$. Thus, the proposed filter has better localization accuracy but higher computational complexity than existing state-of-the-art filters.

Large values of LE appear frequently with the UKF since it is specially designed for Gaussian noise as a result it is sensitive to heavy-tailed non-Gaussian noises. The RSTUF has worse performance than the proposed filter since it needs to prevent the growth of the dof parameter artificially. The proposed filter has better performance than the existing NRHKF and MCCKF methods since the utilized Student's t approach in the proposed method can address heavy-tailed noises better than the Huber technique and maximum corr-entropy method. Large values of LE often appear with the RSTKF and the ALE of the RSTKF doesn't converge with respect to the number of iterations since it uses inaccurate scale matrices and dof parameters. Thus, the proposed filter has better robustness than existing state-of-the-art filters.

IV. Conclusion

In this paper, a new outlier-robust Student's t based GA filter for cooperative localization of AUVs was proposed. The state vector, scale matrices and dof parameters were estimated by using the constructed Student's t based hierarchical Gaussian state-space model and the VB approach. Experimental results showed that the proposed filter has better localization accuracy and robustness than existing state-of-the-art outlier-robust filters.

References

[1] L. Paull, S. Saeedi, M. Seto, and H. Li, "Sensor-driven online coverage planning for autonomous underwater vehicles," *IEEE-ASME Trans. Mech.*, vol. 18, no. 6, pp. 1827–1838, Dec. 2013.

[2] R. Costanzi, F. Fanelli, N. Monni, A. Ridolfi, and B. Allotta, "An Attitude Estimation Algorithm for Mobile Robots Under Unknown Magnetic Disturbances," *IEEE-ASME Trans. Mech.*, vol. 21, no. 4, pp. 1900–1911 Aug. 2016.

[3] P. Shi, and Q. K. Shen, "Cooperative control of multi-agent systems with unknown state-dependent controlling effects," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 3, pp.827–834, July 2015.

[4] M. F. Fallon, G. Papadopoulos, J. J. Leonard, and N. M. Patrikalakis, "Cooperative AUV navigation using a single maneuvering surface craft," *Int. J. Robot. Res.*, vol. 29, no. 12, pp. 1461–1474, Oct. 2010.

[5] H. L. Yu, K. Meier, M. Argyle, and R. W. Beard, "Cooperative path planning for target tracking in urban environments using unmanned air and ground vehicles," *IEEE-ASME Trans. Mech.*, vol. 20, no. 2, pp. 541–552, Apr. 2015.

[6] S. Wang, L. Chen, D. Gu, and H. Hu, "An optimization based Moving Horizon Estimation with application to localization of Autonomous Underwater Vehicles," *Robot. Auton. Syst.*, vol. 62, no. 10, pp. 1581–1596, Oct. 2014.

[7] P. Shi, F. B. Li, L. G. Wu and C. C. Lim, "Neural network-based passive filtering for delayed neutral-type semi-Markovian jump systems," *IEEE Trans. Neur. Net. Lear. Sys.*, DOI: 10.1109/TNNLS.2016.2573853, May 2016.

[8] P. Shi, X. J. Su and F. B. Li, "Dissipativity-based filtering for fuzzy switched systems with stochastic perturbation," *IEEE Trans. Autom. Control*, vol.61, no.6, pp.1694-1699, June 2016.

[9] P. Shi, Y. Q. Zhang, M. Chadli and R. K. Agarwal, "Mixed H-infinity and passive filtering for discrete fuzzy neural networks with stochastic jumps and time delays," *IEEE Trans. Neur. Net. Lear. Sys.*, vol.27, no.4, pp.903-909, Apr. 2016.

[10] B. Allotta, A. Caiti, R. Costanzi, F. Fanelli, E. Meli, and A. Ridolfi, "Development and online validation of an UKF-based navigation algorithm for AUVs," *IFAC-PapersOnLine*, vol. 49, no. 15, pp. 69–74, July 2016.

[11] W. Gao, J. Yang, J. Liu, H. Shi, and B. Xu, "Moving horizon estimation for cooperative localisation with communication delay," *J. Navigation*, vol. 68, no. 3, pp. 493–510, Mar. 2015.

[12] A. Vasilijevic, B. Borovic, and Z. Vukic, "Underwater vehicle localization with complementary filter: performance analysis in the shallow water environment," *J. Intell. Robot. Syst.*, vol. 68, no. 3-4, pp. 373–386, Dec. 2012.

[13] J. Vaganay, J. J. Leonard, and J. G. Bellingham, "Outlier rejection for autonomous acoustic navigation," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, pp. 2174–2181, Apr. 1996.

[14] C. D. Karlgaard, "Nonlinear regression Huber-Kalman filtering and fixed-interval smoothing," *J. Guid., Control, Dynam.*, vol. 38, no. 2, pp. 322–330, Feb. 2015.

[15] W. Gao, Y. L. Liu, and B. Xu, "Robust Huber-based iterated divided difference filtering with application to cooperative localization of autonomous underwater vehicles," *Sensors*, vol. 14, no. 12, pp. 24523–24542, Dec. 2014.

[16] R. Izanloo, S. A. Fakoorian, H. S. Yazdi, and D. Simon, "Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise," in Proceedings of 50th Annual Conference on Information Science and Systems, 2016.

[17] J. Loxam and T. Drummond, "Student mixture filter for robust, real-time visual tracking," *in Proceedings of 10th European Conference on Computer Vision: Part III*, 2008.

[18] Y. L. Huang, Y. G. Zhang, N. Li and J. Chambers, "Robust Student's t based nonlinear filter and smoother," *IEEE Trans. Aero. Elec. Sys.*, vol. 52, no. 5, pp. 2586–2596, Oct. 2016.

[19] Y. L. Huang, Y. G. Zhang, N. Li, S. M. Naqvi and J. Chambers, "A robust Student's t based cubature filter," in *19th International Conference on Information Fusion (FUSION)*, July 2016, pp. 9–16.

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

[21] A. O'Hagan, and J. J. Forster, *Kendall's Advanced Theory of Statistics: Bayesian Inference*. Arnold, 2004.

[22] I. Arasaratnam and S. Haykin, "Cubature Kalman filter," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254−1269, Jun 2009.

[23] Y. L. Huang, Y. G. Zhang, N. Li and J. Chambers, "A robust Gaussian approximate fixed-interval smoother for nonlinear systems with heavy-tailed process and measurement noises," *IEEE Signal Process. Lett.*, vol. 23, no. 4, pp. 468–472, Apr. 2016.

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