Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Learning-based Adaptive Estimation for AOA Target Tracking with Non-Gaussian White Noise

Sheng Xu[1,2], Yongsheng Ou[1,2,⋆] and Xinyu Wu[1,2]

Abstract—**In angle-of-arrival (AOA) target tracking, the target states are estimated using the noisy angular measurements. The non-Gaussian measurement noise is common in practical applications and will decrease the estimation accuracy significantly. This paper is to reduce the negative impacts of the non-Gaussian white noise (NGWN), which is modeled by multiple mixture Gaussian white noises, and guarantee the estimation accuracy. A learning-based adaptive extended Kalman filter (EKF) for AOA target tracking is developed in the NGWN environment. In the proposed method, the extreme learning machine (ELM) using a three-layer neural network is applied to identify the characters of measurement noise at each sampling time. Consequence, the EKF will consider the noise characters by the trained ELM, and make the corresponding real-time adjustments to improve the tracking performance. The effectiveness of the proposed method is verified by simulation examples.**

Index Terms—**Target tracking, angle-of-arrival (AOA) measurement, non-Gaussian white noise, extended Kalman filter (EKF), extreme learning machine (ELM).**

I. Introduction

Angle-of-arrival (AOA) target tracking attracts a lot of interests because it has been widely required in various applications, such as beacon search, sensor networks, search/rescue missions and sonar/radar systems [1] [2] [3] [4] [5]. In AOA target tracking, the target state is estimated using the noisy angular measurements. The geometrical diagram of tracking a mobile target with constant velocity using AOA measurements is shown in Fig. 1. To acquire the target state from the nonlinear AOA measurements, different estimation algorithms have been proposed. As the classical and optimal linear approach for target tracking in Gaussian noise environment, Kalman filter (KF) becomes very popular [6] [7]. For the nonlinear measurement, e.g., AOA measurements, the extended Kalman filter (EKF), unscented Kalman filter (UKF), pseudolinear Kalman filter (PLKF), pseudolinear estimator (PLE) and maximum-likelihood estimator (MLE) [8] [9] [10] [11] [12] [13] were developed and widely applied. Among these nonlinear estimation algorithms, the KF-based methods are iterative and have a lower computation cost compared with the batch methods, e.g., PLE and MLE. Consequence, in this paper, we select the KF-based method as the core estimation algorithm.

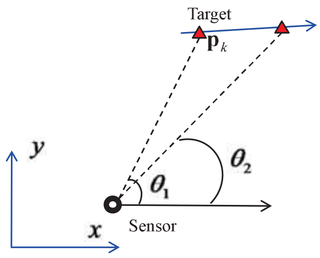

Fig. 1. Diagram of the AOA target localization. θ_1 and θ_2 are the angular measurements at different time indexes.

In many practical applications, the measurement noise is always non-Gaussian, which will degrade the performance of those above estimation methods significantly. In this paper, we only focus on the non-Gaussian white noise (NGWN) suppression problem, and thus the colorful noise is not considered. Consequence, a series of algorithms have been proposed to solve the negative impacts of the non-Gaussian problem [14] [15] [16] [17]. The sequential Monte Carlo (MC) method was applied for the nonlinear non-Gaussian target estimation problems. As a most popular MC method, the particle filter (PF), approximate the target state using a set of random samples which is not impacted by the noise distribution [15]. The computational complexity is determined by the selected particle number which could be very expensive. The authors of [17] proposed a variational Bayesian expectation maximization method for parameter estimation under unknown statistical non-Gaussian measurement noise. In addition, many recursive Kalman-based strategies were also designed. The interactive multiple model EKF (IMM-EKF), which is a recursive method, was also utilized to the non-Gaussian environment [18]. Different from the classical IMM-EKF [19], multiple mixture Gaussian noise models were combined and utilized to fitting the non-Gaussian noise characters. Besides, a Gaussian sum filter was developed in [20] for the state estimation with mixture Gaussian measurement noises. It also presented that a non-Gaussian white noise can be approximately modeled or fitted by mixture Gaussian models. In [21], Gaussian mixture Kalman filter was proposed and it was verified having a better performance than the IMM and PF algorithms in linear non-Gaussian estimation. Furthermore, robust Kalman filter strategy was also modified for non-Gaussian noise environment target tracking problems [22] [23].

Recently, the machine learning technique becomes popular

[1]CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, [2]Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China.
⋆Yongsheng Ou is the corresponding author (email: ys.ou@siat.ac.cn).
This research is funded by National Natural Science Foundation of China (61803363, U1613210), Guangdong Special Support Program (2017TX04X265), Science and Technology Planning Project of Guangdong Province (2019B090915002), Shenzhen Fundamental Research Program (JCYJ20170413165528221) and China Postdoctoral Science Foundation (2018M643254).

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

and is widely applied in different areas including the unknown parameter identification. Thus, it provides an appropriate selection for target tracking in non-Gaussian noise environment. Many machine learning algorithms, e.g., artificial neural networks, support vector machines (SVM) and Gaussian mixture models regression [24], [25], [26] [27], can be utilized to extract the non-Gaussian noise characteristics from the desired training data, i.e., demonstrations. In [25], a simple three-layer artificial neural network, which is called extreme learning machine (ELM), was developed and proven it has enough accuracy for identification problems. Consequence, when the characters of the non-Gaussian measurement noise is known, the estimation problem will be solved.

In this paper, we focus on developing a learning-based adaptive EKF for AOA target tracking with non-Gaussian white noise, and two contributions are made in this paper. First, a non-Gaussian measurement noise character identification strategy is proposed using the ELM. Besides, the corresponding demonstration data production method and training process are also designed. Second, the adjustment strategy of the adaptive EKF algorithm considering the real-time noise characters is developed. Finally, the effectiveness and the characteristics of the proposed method are demonstrated and discussed.

The paper is organized as follows. The AOA target tracking problem with non-Gaussian measurement noise is stated in Section II. Section III introduces the detailed EKF algorithm. The learning-based adaptive EKF is presented in Section IV. Section V proposes the strategies of demonstration data production and ELM training. The proposed method is verified with simulation examples in Section VI. The conclusion and discussion are drawn in Section VII.

II. Problem Formulation

In this paper, a static angle-of-arrival (AOA) sensor (e.g., radar sensor array), is applied to localize a mobile target with a constant velocity on the 2D plane. The unknown target state is defined as $\boldsymbol{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]^T$ where k denotes the discrete time index. $[\dot{x}_k, \dot{y}_k]$ are the constant velocities. The acceleration can be regarded as system model uncertainty.

The location of the static sensor is clearly known as $\boldsymbol{r} = [r_x, r_y]^T$. The target kinematic model can be written as

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{w}_k, \quad \boldsymbol{F}_{k-1} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where T is a constant sampling time-step between discrete-time instants $k-1$ and k. $\boldsymbol{w}_k$ is the process noise (e.g., the acceleration) which is assumed to be a zero-mean Gaussian noise with variance $\boldsymbol{Q}_k$. Assume the process noise is piecewise constant white acceleration errors [19] such that

$$\boldsymbol{Q}_k = \begin{bmatrix} q^x \boldsymbol{A}_k & \boldsymbol{0}_{2\times 2} \\ \boldsymbol{0}_{2\times 2} & q^y \boldsymbol{A}_k \end{bmatrix} \quad \text{and} \quad \boldsymbol{A}_k = \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix} \tag{2}$$

where q^x and q^y are the maximum acceleration values along the $x-$ and $y-$axes, and $\boldsymbol{F}_{k-1}$ is the target dynamical transition matrix.

To track the target, multiple AOA measurements at different time indices are required. The measurement model at time k can be written as

$$\tilde{\theta}_k = \tan^{-1}\frac{y_k - r_y}{x_k - r_x} + n_k \tag{3}$$

where $\tan^{-1}$ is the 4-quadrant arctangent and n_k is the additive non-Gaussian white noise (NGWN) at time k, which is independent of the process noise $\boldsymbol{w}_k$. Note that in the AOA tracking problem of Fig. 1, four unknown parameters (i.e., target initial position x_0, y_0 and the constant target velocity $\dot{x}_k$, $\dot{y}_k$) exist, and thus at least four different measurements (four equations like (3)) at various known time instants are required to guarantee the solution uniqueness. We use the Gaussian mixture model to describe the NGWN [17] as

$$n_k \propto \sum_{j=1}^{N} \alpha_j \eta_{jk}, \tag{4}$$

where α_j is the weight with $\sum_{j=1}^{N} \alpha_j = 1$, η_{jk} denotes the Gaussian noise with $\mathcal{N}(\mu_j, \sigma_j^2)$ at time k. Note that if one element with large weight and η_{jk} has large σ_j^2, n_k will be a glint noise. In addition, $\propto$ denotes that n_k is determined by $\sum_{j=1}^{N} \alpha_j \eta_{jk}$ but not exactly equal, and the exact n_k can be calculated by the Matlab "gmdistribution" function using the information of $\sum_{j=1}^{N} \alpha_j \eta_{jk}$.

In this paper, the key objective is to estimate the target real-time state using the AOA measurements with complex non-Gaussian noises.

III. Extended Kalman Filter

In the previous works, to estimate the target state, different statistic-based algorithms have been developed. Compared with the particle filter and unscented Kalmen filter [15] [8], extended Kalman filter (EKF), has a low computational cost and has been widely applied. Thus, we select the EKF as the core estimation algorithm.

The EKF includes three steps. First, the priori estimate and covariance updates are

$$\boldsymbol{x}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1|k-1} \tag{5a}$$
$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{Q}_{k-1}. \tag{5b}$$

Second, the calculations of the Kalman gain, residual are

$$\boldsymbol{H}_k = \begin{bmatrix} -\frac{\boldsymbol{x}_{k|k-1}(3)-r_y}{[\boldsymbol{x}_{k|k-1}(1)-r_x]^2+[\boldsymbol{x}_{k|k-1}(3)-r_y]^2} \\ 0 \\ \frac{\boldsymbol{x}_{k|k-1}(1)-r_x}{[\boldsymbol{x}_{k|k-1}(1)-r_x]^2+[\boldsymbol{x}_{k|k-1}(3)-r_y]^2} \\ 0 \end{bmatrix}^T \tag{6a}$$

$$\boldsymbol{J}_k = \hat{\boldsymbol{R}}_k + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^{jT} \tag{6b}$$
$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T/\boldsymbol{J}_k \tag{6c}$$
$$s_k = \tilde{\theta}_k - \tan^{-1}\frac{\boldsymbol{x}_{k|k-1}(3) - r_y}{\boldsymbol{x}_{k|k-1}(1) - r_x} - \hat{u}_k \tag{6d}$$

where $\boldsymbol{H}_k$ is the Jacobian matrix of $\tan^{-1}()$ respect to the priori estimate $\boldsymbol{x}_{k|k-1}$ because the true $\boldsymbol{x}_k$ is unknown. $\boldsymbol{J}_k$ is just an intermediate matrix, $\hat{u}_k$ and $\hat{\boldsymbol{R}}_k$ are the bias and measurement covariance matrix which describe the measurement noise bias and variance at time k.
Finally, the posterior estimate and covariance have

$$\boldsymbol{x}_{k|k} = \boldsymbol{x}_{k|k-1} + \boldsymbol{K}_k s_k \tag{7a}$$
$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{H}_k \boldsymbol{P}_{k|k-1}. \tag{7b}$$

Note that in the classical EKF, i.e., in the Gaussian measurement noise environment, $\hat{u}_k$ and $\hat{\boldsymbol{R}}_k$ in (6) are assumed as zero and a known constant matrix, respectively. When the measurement noise is non-Gaussian, at each time discrete the noise should have different bias and variance. Consequence, different statistic-based methods [28] [17] can be applied to estimate the bias $\hat{u}_k$ and variance $\hat{\boldsymbol{R}}_k$ of the real-time measurement noise, and thus the adaptive EKF (AEKF) has been developed. However, in the statistic-based methods, as the noise characters are calculated using the inaccurate estimates, the estimation performance is limited. In this paper, we will develop a learning-based strategy (a kind of data driven strategy) to realize the real-time noise character identification.

IV. ELM-BASED ADAPTIVE EXTENDED KALMAN FILTER

Apart from using statistical methods to compute the noise characteristics, a learning-based method will be developed in this section. The proposed method is to identify the accurate measurement bias and variance information directly from demonstration data. The learning-based method is a data driven strategy which requires a lot of non-Gaussian white noises with their characters and some special information produced in an EKF algorithm. More details will be introduced later in this section.

A three-layer neural network, i.e., extreme learning machine (ELM), is utilized to identify the real-time characteristics of the non-Gaussian white noise (NGWN). In the learning-based algorithm, the residual s_k in (6) is defined as the input of the ELM, and the noise real-time bias and variance are the outputs. Theoretically, the outputs can be calculated accurately by the neural network. In this process, a well-trained neural network is required and thus a demonstration data library is necessary for training the neural network.

The defined input and outputs of the ELM satisfy

$$[\hat{u}_k, \hat{\boldsymbol{R}}_k]^T = f(s_k) \tag{8a}$$

where $f(\cdot)$ is the neural network function. For the identification neural network, we have

$$\left[\hat{u}_k, \hat{\boldsymbol{R}}_k\right]^T = \sum_{l=1}^{N} \beta_l g\left(w_l s_k + a_l\right) \tag{9}$$

where $g(\cdot)$ is the activation function, N is the number of hidden nodes, $\boldsymbol{a} = [a_l, a_2, ..., a_N]$, $\boldsymbol{w} = [w_l, w_2, ..., w_N]$ and $\boldsymbol{\beta} = \begin{bmatrix} \beta_{b1}, & \beta_{b2}, & ..., & \beta_{bN} \\ \beta_{v1}, & \beta_{v2}, & ..., & \beta_{vN} \end{bmatrix}_{2\times N}$ are the hidden layer bias, input weight and output weight, respectively. β_{bl} and β_{vl} are the bias and variance weights. According to [25], $\boldsymbol{a}$ and $\boldsymbol{w}$ in (9) can be initialized with invariable random constants, and the identification performance will not be impacted. Thus, the training process of the ELM becomes simple. The number of the total data points in the demonstrations is M. Then, the output weight $\boldsymbol{\beta}$ can be calculated by solving the following least-square problem

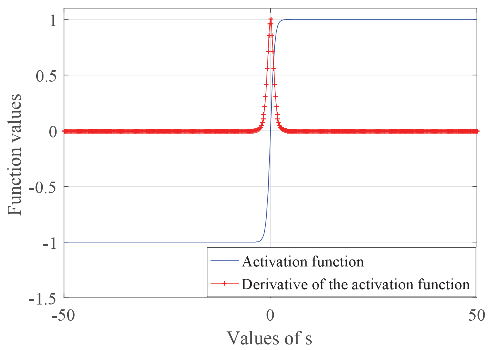

Fig. 2. The selected activation function and its derivative.

$$\min_{\boldsymbol{\beta}} \left\| \boldsymbol{D}\boldsymbol{\beta}^T - \boldsymbol{U} \right\| \tag{10a}$$

$$\text{where} \quad \boldsymbol{D} = \begin{bmatrix} g\left(w_1 s_1 + a_1\right) & \cdots & g\left(w_N s_1 + a_N\right) \\ \vdots & \ddots & \vdots \\ g\left(w_1 s_M + a_1\right) & \cdots & g\left(w_N s_M + a_N\right) \end{bmatrix}_{M\times N} \tag{10b}$$

$$\text{and} \quad \boldsymbol{U} = \begin{bmatrix} u^1, & u^2, & ..., & u^M \\ \boldsymbol{R}^1, & \boldsymbol{R}^2, & ..., & \boldsymbol{R}^M \end{bmatrix}^T_{M\times 2}. \tag{10c}$$

$\boldsymbol{D}$ is called as the hidden layer output matrix. The outputs of the M data points from the multiple demonstrations are denoted by $\boldsymbol{U}$, and the elements in $\boldsymbol{U}$ are accurate values rather than estimates. As $\boldsymbol{w}$ and $\boldsymbol{a}$ are constants, the solution of (10) becomes

$$\hat{\boldsymbol{\beta}} = \boldsymbol{D}^{+}\boldsymbol{U} \tag{11}$$

where $\boldsymbol{D}^{+}$ is the Moore-Penrose generalized inverse of $\boldsymbol{D}$ [25].

Furthermore, the selected activation function $g(\cdot)$ should be continuous and continuously differentiable and thus is has

$$g(0) = 0 \tag{12a}$$
$$\dot{g}(s) > 0, \quad s \neq 0. \tag{12b}$$

Consequently, we choose

$$g(s) = \frac{2}{1+\exp(-2s)} - 1 \tag{13}$$

as the activation function and the details are shown in Fig. 2

In summary, the proposed learning-based AEKF takes

$$\boldsymbol{x}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1|k-1} \tag{14a}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{Q}_{k-1} \tag{14b}$$

$$\boldsymbol{H}_k = \begin{bmatrix} -\frac{\boldsymbol{x}_{k|k-1}(3)-r_y}{[\boldsymbol{x}_{k|k-1}(1)-r_x]^2+[\boldsymbol{x}_{k|k-1}(3)-r_y]^2} \\ 0 \\ \frac{\boldsymbol{x}_{k|k-1}(1)-r_x}{[\boldsymbol{x}_{k|k-1}(1)-r_x]^2+[\boldsymbol{x}_{k|k-1}(3)-r_y]^2} \\ 0 \end{bmatrix}^T \tag{14c}$$

$$[\hat{u}_k, \hat{\boldsymbol{R}}_k]^T = \sum_{l=1}^{N}\boldsymbol{\beta}_l g\left(w_l s_k + a_l\right) \tag{14d}$$

$$\boldsymbol{J}_k = \hat{\boldsymbol{R}}_k + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^{jT} \tag{14e}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T/\boldsymbol{J}_k \tag{14f}$$

$$s_k = \tilde{\theta}_k - \tan^{-1}\frac{\boldsymbol{x}_{k|k-1}(3) - r_y}{\boldsymbol{x}_{k|k-1}(1) - r_x} - \hat{u}_k \tag{14g}$$

$$\boldsymbol{x}_{k|k} = \boldsymbol{x}_{k|k-1} + \boldsymbol{K}_k s_k \tag{14h}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}. \tag{14i}$$

The key part in the proposed algorithm is the trained ELM, i.e., (14d). Next, we will figure out how to guarantee the performance of the ELM through a well-designed training process.

V. Demonstration production and ELM training

In the demonstration production process, multiple AOA measurements with different clearly known non-Gaussian white noises (NGWNs) and a classical EKF are applied. More specifically, to train the ELM, the input and output data, i.e., u and $\boldsymbol{R}$ in (10), from the demonstrations are required and known.

First, we use the EKF to track a target using the AOA measurements with NGWN. Then, the target moving directions will be changed and the tracking process will be repeated. We will use the EKF to track eight targets with different moving directions to produce eight groups of demonstration data. Therefore, for example, if in each target estimation the total measurement updates is N, i.e., $k = 1, 2, ..., N$, the total data point number of the eight demonstrations becomes $N \times 8 = M$. In this demonstration data production process, the estimation residuals, s_k, are recorded and will be applied as the inputs of the ELM in (10). But note that in each demonstration, the inaccurate data at the beginning of the tracking are ignored because of EKF convergence problem [6]. In addition, Fig. 3 shows the true target moving trajectories in the 8 demonstrations.

Second, we will obtain the output data based on the actual added noises to train the ELM. The NGWNs [17] are produced by three mixture Gaussian models with equal weights. We set $\eta_{1k} \sim \mathcal{N}(\mu_1 = 1^o, \sigma_1^2 = 1^{o2})$, $\eta_{2k} \sim \mathcal{N}(\mu_2 = 2^o, \sigma_2^2 = 3^{o2})$ and $\eta_{3k} \sim \mathcal{N}(\mu_3 = 3^o, \sigma_3^2 = 6^{o2})$. Consequence, the noise value n_k is clearly known (can be produced by Matlab "gmdistribution" function) and thus the measurement bias and variance at each time instant can be acquired through modeling the real-time noise by multiple different Gaussian distributions. Therefore, the output $\boldsymbol{U}$ in (10) and (11) which will be utilized for the ELM training are acquired. Note that different strategies can be applied to model the noise with a more appropriate Gaussian distribution and calculate the bias and variance.

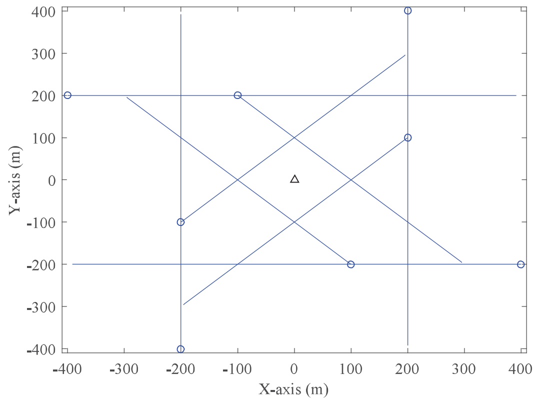

Fig. 3. True target trajectories of the eight demonstrations. The blue "○" marks the start position of a target trajectory, and the black "△" denotes the sensor located at $(0, 0)m$.

Based on the acquired input and output data, s_k and $\boldsymbol{U}$, the ELM in (9) can be obtained. In summary, as more different demonstrations using different target trajectories with appropriate noise character calculation methods, the identification performance of the trained ELM will be improved [25].

VI. Simulation studies

The effectiveness of the proposed method is verified by Matlab simulation examples. Three examples are provided, and the target moves from diffeent start positions $(150, 100)$m (example 1) and $(-200, -100)$m (example 2 and 3) with different velocities in these examples. Besides, the non-Gaussian measurement noises are also different. For comparison reason, the EKF and the proposed learning-based AEKF are applied to track the three different target trajectories, respectively. Note that the initial state and covariance matrix of the different methods are same. The measurement update is $L = 100$ and the time instant is $T = 1$s, which means the total simulation time is 200s. The static radar sensor locates at $(0, -200)$m. In the EKF, $\hat{u}_k = 0$ and $\hat{\boldsymbol{R}}_k = 5^{o2}$. Note that in the practical applications, the changes of $\hat{u}_k$ and $\hat{\boldsymbol{R}}_k$ between different continuous discrete times should not be too large to avoid divergence risk, and thus two thresholds are set for $\hat{u}_k, \hat{\boldsymbol{R}}_k$. Besides, the ELM will work after $k = 20$ to avoid the impacts of the inaccurate initials. In order to evaluate the estimation performance, the bias norms between the x- and y-axes estimates and true target states, and the trace of the covariance matrix, regarding as the estimated position and velocity mean-squared-error (MSE), are applied. Note that as the computational cost of the UKF or PF are larger than the EKF, the simulations and comparisons of using UKF or PF are not provided in this paper.

In example 1, the target velocity satisfies $v_x = 1$m/s, $v_y = -2$m/s. The non-Gaussian measurement noise is a mixture of three Gaussian noises, i.e., $\eta_1 \sim \mathcal{N}(1^o, 1^{o2})$, $\eta_2 \sim \mathcal{N}(3^o, 3^{o2})$, and $\eta_3 \sim \mathcal{N}(3.5^o, 3^{o2})$. The true target, estimated trajectories, sensor position and the corresponding performance are shown in Fig. 4. The direct applied EKF shows limited estimation performance while the proposed method performs better both in the bias and MSE results. At the beginning, the trained ELM was not used and thus the performance of the two methods are same. As the number of measurement data increases, the identification and corresponding estimation performance is improved. As the target derive from the sensor, the tracking performance will degenerate, which is a common phenomena in AOA target localization.

In example 2, the target velocity satisfies $v_x = 3$m/s, $v_y = 1$m/s. The measurement noise is the equal mixture of $\eta_1 \sim \mathcal{N}(1^o, 1^{o2})$, $\eta_2 \sim \mathcal{N}(2^o, 1^{o2})$ and $\eta_3 \sim \mathcal{N}(0^o, 3^{o2})$. The simulation result is shown in Fig. 5. The proposed method also shows a more accurate estimation result.

In example 3, the measurement noise is the equal mixture of $\eta_1 \sim \mathcal{N}(3^o, 1^{o2})$, $\eta_2 \sim \mathcal{N}(2^o, 1.5^{o2})$ and $\eta_3 \sim \mathcal{N}(0.5^o, 3^{o2})$. The simulation details are shown in Fig. 6. The true target moves with $v_x = -1$m/s, $v_y = 3$m/s. The results of the three examples verified the effectiveness of the proposed method.

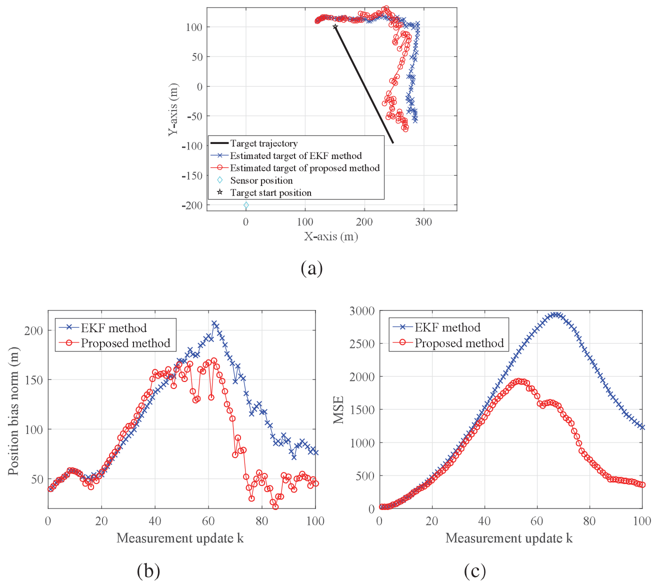

Fig. 4. Tracking performance of example 1 using the proposed algorithm, (a) Target and the estimated trajectories, (b) estimated bias norm performance, (c) MSE performance.

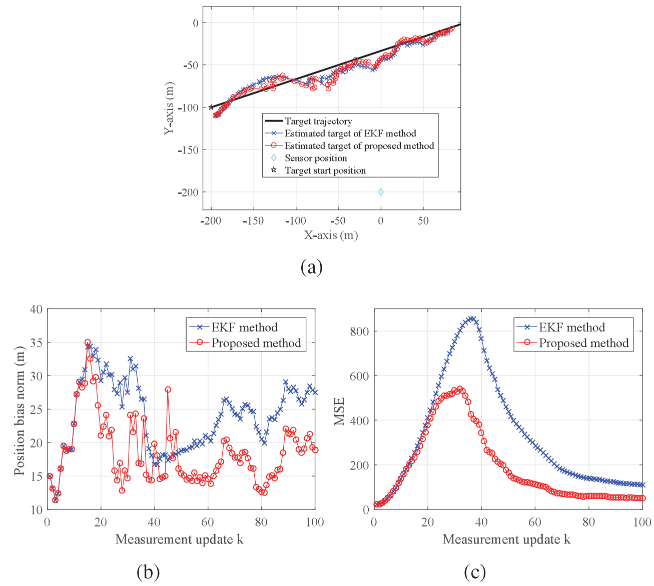

Fig. 5. Tracking performance of example 2 using the proposed algorithm, (a) Target and the estimated trajectories, (b) estimated bias norm performance, (c) MSE performance.

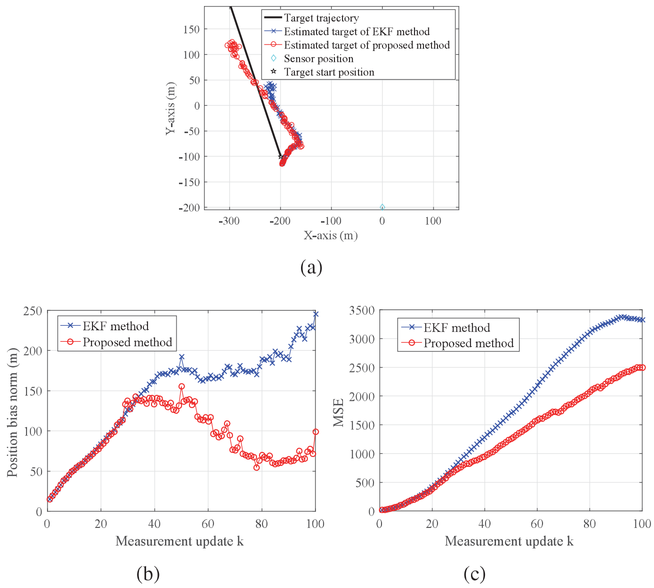

Fig. 6. Tracking performance of example 3 using the proposed algorithm, (a) Target and the estimated trajectories, (b) estimated bias norm performance, (c) MSE performance.

VII. Conclusion and Discussion

In this paper, a learning-based AEKF is proposed for AOA target tracking under non-Gaussian white noise. In the proposed method, the ELM using three-layer neural network is applied to identify the characters of the actual measurement noise at each sampling time. Consequence, the EKF will consider the noise characters and make the corresponding real-time adjustments. The effectiveness of the proposed method is verified by simulations with the comparison between the normal EKF and the proposed algorithms.

In the proposed learning-based AEKF, the training process is off-line, and the identification performance can be significantly improved by providing more data. The computational cost of the on-line estimation algorithm is similar to the EKF, which is lower than the statistic-based ones. More specifically, compared with the EKF, the proposed method has an extra step, i.e., calculating the bias and variance in (8) using the trained ELM. We just need to substitute the residuals into the parameter-fixed function and use some basic operations in this calculation. Thus, the proposed method has a low computational cost and good estimation performance under non-Gaussian noise environment at the expense of a complex off-line preparation, which includes data collection and the ELM training. Because we only focus on the final estimation results, the correction rate of the noise character identification was not evaluated, which can be definitely improved. We believe there still a lot of space for the estimation algorithm improvement, such as obtaining high equality demo data from various forms of target maneuvering, improving data collection and simplifying the training process. Furthermore, the proposed method can also be tried for the combination of the UKF or Particle filter under the non-Gaussian noises satisfy Cauchy, Poisson or chi-square distributions.

REFERENCES

[1] E. Olson, J. J. Leonard, and S. Teller, "Robust range-only beacon localization," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 4, pp. 949–958, 2006.

[2] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, 2005.

[3] M. A. Goodrich, B. S. Morse, D. Gerhardt, J. L. Cooper, M. Quigley, J. A. Adams, and C. Humphrey, "Supporting wilderness search and rescue using a camera-equipped mini UAV," *Journal of Field Robotics*, vol. 25, no. 1-2, pp. 89–110, 2008.

[4] T. Hanselmann, M. Morelande, B. Moran, and P. Sarunic, "Sensor scheduling for multiple target tracking and detection using passive measurements," in *International Conference on Information Fusion*, 2008, pp. 1–8.

[5] S. Xu and K. Doğançay, "Optimal sensor placement for 3-D angle-of-arrival target localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1196–1211, 2017.

[6] M. S. Grewal and A. P. Andrews, *Kalman filtering: theory and practice using MATLAB*. John Wiley & Sons, 2011.

[7] M. W. Khan, N. Salman, A. Ali, A. M. Khan, and A. H. Kemp, "A comparative study of target tracking with kalman filter, extended kalman filter and particle filter using received signal strength measurements," in *International Conference on Emerging Technologies*, 2016, pp. 1–6.

[8] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, March 2004.

[9] Y. Oshman and P. Davidson, "Optimization of observer trajectories for bearings-only target localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 892–902, 1999.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Pearson Education, 2011, vol. 1.

[11] S. Xu, K. Doğançay, and H. Hmam, "3D pseudolinear Kalman filter with own-ship path optimization for AOA target tracking," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Match, 2016, pp. 3136–3140.

[12] K. Doğançay, "3D pseudolinear target motion analysis from angle measurements," *Signal Processing, IEEE Transactions on*, vol. 63, no. 6, pp. 1570–1580, 2015.

[13] S. Xu, K. Doğançay, and H. Hmam, "Distributed pseudolinear estimation and UAV path optimization for 3D AOA target tracking," *Signal Processing*, vol. 133, pp. 64–78, 2017.

[14] C. J. Masreliez, "Approximate non-Gaussian filtering with linear state and observation relation," *IEEE Transactions on Automatic Control*, vol. 20, no. 1, pp. 107–110, 1975.

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

[16] D. Huang and H. Leung, "Maximum likelihood state estimation of semi-Markovian switching system in non-Gaussian measurement noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 133–146, Jan 2010.

[17] H. Zhu, H. Leung, and Z. He, "State estimation in unknown non-Gaussian measurement noise using variational Bayesian technique," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 49, no. 4, pp. 2601–2614, 2013.

[18] E. Daeipour and Y. Bar-Shalom, "Interacting multiple model approach for target tracking with glint noise," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 31, no. 2, pp. 706–715, 1995.

[19] Y. Bar-Shalom, *Multitarget-multisensor tracking: Applications and advances. Volume III*. Norwood, MA, Artech House, Inc., 2000.

[20] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, August 1972.

[21] I. Bilik and J. Tabrikian, "MMSE-based filtering in presence of non-Gaussian system and measurement noise," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 46, no. 3, pp. 1153–1170, 2010.

[22] 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 *2016 Annual Conference on Information Science and Systems (CISS)*, March 2016, pp. 500–505.

[23] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J. A. Chambers, "Robust Kalman filters based on Gaussian scale mixture distributions with application to target tracking," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–15, 2018.

[24] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks." *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.

[25] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[26] A. M. Andrew, "An introduction to support vector machines and other kernel based learning methods," *Kybernetes*, vol. 32, no. 1, pp. 1–28, 2001.

[27] S. Xu, Y. Ou, J. Duan, X. Wu, W. Feng, and M. Liu, "Robot trajectory tracking control using learning from demonstration method," *Neurocomputing*, vol. 338, pp. 249 – 261, Jan 2019.

[28] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, March 2000.