Distributed Target Localization with Inaccurate Collaborative Sensors in Multipath Environments

Yuan Feng[1,2], Qinsiwei Yan[2], Po-Hsuan Tseng[3], Ganlin Hao[2] and Nan Wu[2]
[1] China Research and Development Academy of Machinery Equipment, Beijing, 100089, China
[2] School of Information and Electronics, Beijing Institute of Technology, Beijing, 100081, China
[3] Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan
*Corresponding author: Nan Wu

Abstract

Location-aware networks are of great importance for both civil lives and military applications. Methods based on line-of-sight (LOS) measurements suffer sever performance loss in harsh environments such as indoor scenarios, where sensors can receive both LOS and non-line-of-sight (NLOS) measurements. In this paper, we propose a data association (DA) process based on the expectation maximization (EM) algorithm, which enables us to exploit multipath components (MPCs). By setting the mapping relationship between the measurements and scatters as a latent variable, coefficients of the Gaussian mixture model are estimated. Moreover, considering the misalignment of sensor position, we propose a space-alternating generalized expectation maximization (SAGE)-based algorithms to jointly update the target localization and sensor position information. A two dimensional (2-D) circularly symmetric Gaussian distribution is employed to approximate the probability density function of the sensor's position uncertainty via the minimization of the Kullback-Leibler divergence (KLD), which enables us to calculate the expectation step with low computational complexity. Moreover, a distributed implementation is derived based on the average consensus method to improve the scalability of the proposed algorithm. Simulation results demonstrate that the proposed centralized and distributed algorithms can perform close to the Monte Carlo-based method with much lower communication overhead and computational complexity.

1. Introduction

Wireless sensor networks (WSNs) consist a large number of power-limited sensors that use wireless transmission to communicate with neighbors [1]. They can collaboratively estimate the position and state of a target with limited power and computational capabilities. WSNs localization has been employed in many applications such as habitat monitoring, customers tracking, and control of appliances [2]-[5].

For line-of-sight (LOS) propagation scenarios, existing algorithms for cooperative localization [6], object tracking [7], [8] and simultaneous localization and mapping (SLAM) [9] have been well studied. However, for the applications in commercial shopping malls, urban canyons and jungles with various obstacles, these algorithms designed for LOS signals result in severe performance loss due to the multipath components (MPCs) in the received signal, which are a mixture of LOS and a bunch of non-line-of-sight (NLOS) components [10]-[12]. Therefore, studies on how to deal with MPCs in WSNs localization is interesting and challenging. An iterative method is proposed in [10] to estimate the MPCs' ranging probability density function (pdf) based on a Gaussian mixture model. However, the assumption that all the received measurements are a mixture from both LOS and NLOS paths may lead to performance loss, especially when sensors only receive one kind of measurement. A method based on detection and elimination of NLOS components in MPCs is proposed in [13]. Moreover, [14] and [15] use plentiful LOS components to replace NLOS measurements. However, the performance of these methods become very poor when sensors cannot receive enough LOS components. In [16], it is proved that single reflections of NLOS components contain additional position information which can effectively improve the localization accuracy. For convenience, we use NLOS to represent the single reflection in the remaining part.

Cooperative localization algorithms based on TOA measurements have been studied in [17], [18], in which the NLOS measurements are also taken into consideration. Given the map of layout, a TOA based localization algorithm that utilizes NLOS measurements is proposed in [19]. Although superior performance can be observed, the requirement of detailed map information is impractical. To solve this problem, methods based on the main scatter's number and heading angles are proposed in [20]-[23], which do not need layout information. Meanwhile, considering sensors and target may be noncooperative, an RF localization method is proposed in [22], using both the time-difference-of-arrival (TDOA) and angle-of-arrival (AOA) measurements, which makes great contributions to the availability of localization system. Nevertheless, this method relaxes layout constraint by a strong assumption that each sensor only receives one NLOS path with known data association (DA) information, which may have poor performance in complex multipath environments. Moreover, the above researches do not consider the misalignment or movement of sensors in WSNs, which may also impact the location performance [24]-[26]. In this paper, we first propose a method to solve the DA problem by giving the number of main scatters and their heading angles. Then we extend our algorithm by considering the position uncertainty of sensors.

Practically, sensors are often randomly distributed with limited power. In a centralized implementation, the raw measurements obtained by sensors have to be transmitted to a data fusion center (FC) which is probably far away. The stability of such centralized algorithms can not be guaranteed. Alternatively, a consensus-based [27]-[29] distributed implementation without FC is proposed in this paper.

In this paper, we focus on a 2-D target localization problem based on TDOA measurements with inaccurate collaborative sensors in a multipath environment. Expectation maximization (EM)-based DA process is proposed based on the Gaussian mixture model. We propose our centralized space-alternating generalized expectation maximization (SAGE) -based [31] algorithm to locate the target and remove the effects of inaccurate sensors [32], in which the minimization of Kullback-Leibler Divergence and Taylor expansions have been employed to reduce the complexity and derive a closed-form expression in the *E*-step. Finally, a consensus-based distributed implementation of the proposed algorithm is given. The main contributions of this paper can be summarized as follows:

- We propose a novel EM-based DA process using a Gaussian mixture model in multipath environments based on the TDOA and AOA measurements. Therefore, both LOS and NLOS components can be utilized in target localization.

- A centralized SAGE-based algorithm is proposed to jointly estimate the target position and update the position information of inaccurate sensors. The distribution of sensor uncertainty is approximated to a 2-D circularly symmetric Gaussian distribution via KLD minimization. Taylor expansion is employed to linearize the nonlinear confluent hypergeometric function. Accordingly, a closed-form expression can be derived in the *E*-step and the complexity of the proposed algorithm can be significantly reduced.

- To further improve the scalability of the target localization in complex scenarios, a beneficial distributed implementation of the proposed algorithm is given based on the average consensus.

The remainder of this paper is organized as follows. Section II introduces the system model and the problem at hand. Section III gives the proposed EM-based DA method on a Gaussian Mixture model. Section IV elaborates the centralized algorithm for joint considerations of localization and sensor position uncertainty and derives the distributed implementation of the proposed algorithm based on average consensus. Comprehensive simulations are shown in Section V. Finally, concluding remarks are made in section VI.

2. System Model

We consider a 2-D target localization in a multipath environment, as shown in **Fig. 1**. Assume that the number of scatters L and the orientations of scatters are known. There are a total of N sensors, in which each sensor may receive signals of a target from LOS and/or NLOS paths. The number of signal paths (i.e., measurements) received by sensor i (i.e., S_i) is denoted as M_i. Using ultra-wideband technologies, these signals can be distinguished. We define $i \in \mathcal{N}_{\text{LOS}}$ if S_i received signal from LOS path; otherwise, $i \in \mathcal{N}_{\text{\textbackslash LOS}}$. We denote $\mathcal{N} = \{1, 2, ...N\}$ as the set of the indexes of sensors, where $\mathcal{N} = \mathcal{N}_{\text{\textbackslash LOS}} \cup \mathcal{N}_{LOS}$.

Since the target is non-cooperative, all the sensors have to resort to the TDOA. Without loss of generality, the first measurement of sensor 1, i.e., S_1, is used as the reference. Then, the j-th TDOA measurement received by the i-th sensor $\Delta \hat{d}_{i,j}$ is given as follows

$$\Delta \hat{d}_{i,j} = g(\theta_{i,j}, \gamma_k)^T (\mathbf{q} - \bar{\mathbf{p}}_i) - g(\theta_{1,1}, \gamma_{1,1}')^T (\mathbf{q} - \mathbf{p}_1) + \tilde{n}_{i,j}, \qquad (1)$$

where $\mathbf{q} = [x_q \ y_q]^T$ is the ground-truth position of the target. $\bar{\mathbf{p}}_i = [\bar{x}_i \ \bar{y}_i]^T$ is the recorded position of S_i for $i \in \mathcal{N}$. $\theta_{i,j}$ is the AOA of the j-th signal path at S_i. γ_k is the orientation of the k-th scatter, which is associated to the measurement $\Delta \hat{d}_{i,j}$, for $k \in \{1,...,L\}$. $\gamma'_{1,1}$ is assumed to be known as the orientation of the scatter associated to sensor 1, which is explained and proved in [22]. $\tilde{n}_{i,j}$ is the TDOA measurement noise, which is a zero-mean Gaussian random variable with variance $\sigma_{i,j}^2$. The function $g(\theta_{i,j}, \gamma_k)$ is given by

$$g(\theta_{i,j}, \gamma_k) = \frac{1}{\cos(\theta_{i,j} - \gamma_k)} \begin{bmatrix} \cos \gamma_k \\ \sin \gamma_k \end{bmatrix}, \qquad (2)$$

which can be derived by projecting distance along the direction between the target and S_i. Besides, the AOA measurement $\hat{\theta}_{i,j}$ is contaminated by noise, which can be modeled as $\hat{\theta}_{i,j} = \theta_{i,j} + \eta_{i,j}$, where $\eta_{i,j}$ is the noise with uniform distribution.
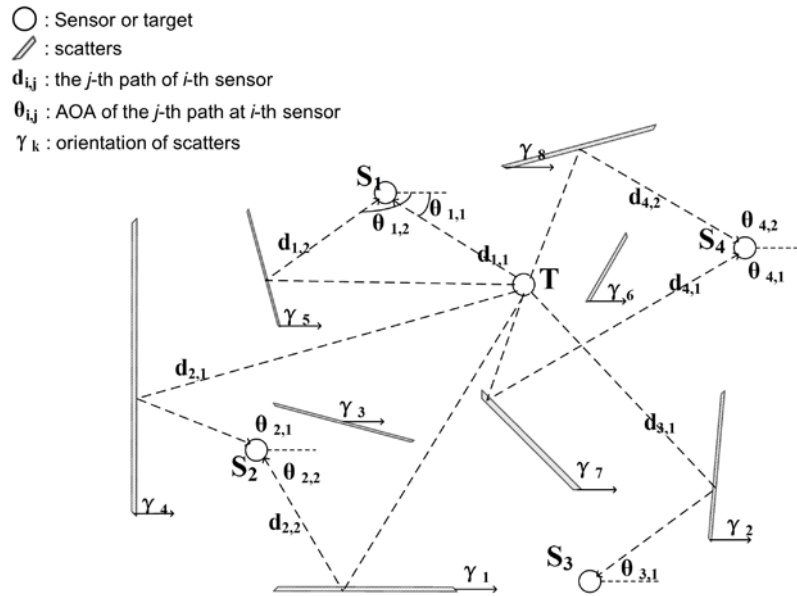

Fig. 1. An example of the NLOS and LOS paths from the target to different sensors.

In (1), besides the target position $\mathbf{q}$, the scatter index k associated to the measurement $\Delta \hat{d}_{i,j}$ can not be directly observed. Moreover, the ground-truth position of sensor $\mathbf{p}_i$ might not be available due to misalignment of sensors. We further model the two types of uncertainty in the following subsections.

2.1 The Uncertainty of Model Coefficients

Generally, the uncertainty of model coefficients of the range measurement includes two questions: (i) Is the ranging measurement coming from LOS or NLOS signal path? (ii) If a

NLOS path exists, which scatterer is the signal reflected from? In the data association problem, we treat the scatter index k in (1) as an unknown variable. The first measurement in each sensor has the probability that can come from the LOS path or the NLOS path from the scatter k, the first measurement ($j = 1$) of each sensor has $L+1$ submodels in Gaussian mixture model; for the remaining measurements, there are L submodels. We denote the number of the submodels $K_{i,j}$ from the j -th measurement of the i -th sensor as

$$K_{i,j} = \begin{cases} L+1, & j=1 \\ L, & j \neq 1 \end{cases}. \tag{3}$$

We define a latent variable for the mapping information $\rho_{i,j,k}$. When the j -th measurement comes from the k -th model, $\rho_{i,j,k} = 1$. Otherwise, $\rho_{i,j,k} = 0$. Since the measurement $\Delta \hat{d}_{i,j}$ in (1) follows Gaussian distribution, without other information, the likelihood function of each sensor becomes a Gaussian mixture model [33]. Here, we define $\boldsymbol{\rho} = \{\{\{\rho_{i,j,k}\}_{k=1}^{K_{i,j}}\}_{j=1}^{M_i}\}_{i=1}^{N}$ and $\boldsymbol{\alpha} = \{\{\{\alpha_{i,j,k}\}_{k=1}^{K_{i,j}}\}_{j=1}^{M_i}\}_{i=1}^{N}$, where M_i is the number of measurements at the i -th sensor. Then we have

$$p(\Delta \hat{d}_{i,j} \mid \mathbf{q}, \theta_{i,j}, \theta_{1,1}, \boldsymbol{\rho}, \boldsymbol{\alpha}) = \sum_{k=1}^{K_{i,j}} [\alpha_{i,j,k} \rho_{i,j,k} \Phi(\Delta \hat{d}_{i,j} \mid \mathbf{q}, \theta_{i,j}, \theta_{1,1})], \tag{4}$$

where $\alpha_{i,j,k}$ is the weight, $\alpha_{i,j,k} \geq 0, \sum_{k=1}^{K_{i,j}} \alpha_{i,j,k} = 1$. In (4), the first path includes $L+1$ possible modes, and the other path includes L possible modes at each sensor.

Moreover, we have

$$\Phi(\Delta \hat{d}_{i,j} \mid \mathbf{q}, \theta_{i,j}, \theta_{1,1}) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(\Delta \hat{d}_{i,j} - \mu_{i,j})^2}{2\sigma^2}), \tag{5}$$

with

$$\mu_{i,j} = g(\hat{\theta}_{i,j}, \gamma_k)^T (\mathbf{q} - \bar{\mathbf{p}}_i) - g(\hat{\theta}_{1,1}, \gamma_{1,1}')^T (\mathbf{q} - \mathbf{p}_1). \tag{6}$$

Here, for $j = 1$, the $(L+1)$ -th submodel corresponds to the measurement coming from the LOS path. Therefore, in this case, we treat $\gamma_{L+1} = \theta_{i,1}$. Based on the model in (4), the problem of the detection of LOS or NLOS path and that of the estimation of the signal reflected from which scatter are transferred into the estimation of weight $\alpha_{i,j,k}$, by introducing a latent variable $\rho_{i,j,k}$.

2.2 The Uncertainty of Sensor Positions

In practice, the position of sensors could be inaccurate due to the misalignment of sensors. We denote the position uncertainty of the i -th sensor as $\Delta \mathbf{p}_i = [\Delta x_i \ \Delta y_i]^T$. The ground-true position of sensor i is $\mathbf{p}_i = [x_i \ y_i]^T$, which can be written as $\mathbf{p}_i = \bar{\mathbf{p}}_i + \Delta \mathbf{p}_i$. Then, we have

$$\Delta\hat{d}_{i,j} = \mathrm{g}(\hat{\theta}_{i,j}, \gamma_k)^T (\mathbf{q} - \overline{\mathbf{p}}_i - \Delta\mathbf{p}_i) - \mathrm{g}(\hat{\theta}_{1,1}, \theta_{1,1})^T (\mathbf{q} - \mathbf{p}_1) + \tilde{n}_{i,j}. \tag{7}$$

The sensor's position uncertainty $\Delta\mathbf{p}_i$ is modeled as a 2-D Gaussian distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\Delta\mathbf{p}_i}$, i.e.,

$$p(\Delta\mathbf{p}_i) = \frac{1}{\sqrt{(2\pi)^2} \, |\boldsymbol{\Sigma}_{\Delta\mathbf{p}_i}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} \Delta\mathbf{p}_i^\mathrm{T} \boldsymbol{\Sigma}_{\Delta\mathbf{p}_i}^{-1} \Delta\mathbf{p}_i \right\}, \tag{8}$$

with a covariance matrix

$$\boldsymbol{\Sigma}_{\Delta\mathbf{p}_i} = \begin{bmatrix} \sigma^2_{\Delta x_i} & \omega\sigma_{\Delta x_i}\sigma_{\Delta y_i} \\ \omega\sigma_{\Delta x_i}\sigma_{\Delta y_i} & \sigma^2_{\Delta y_i} \end{bmatrix}, \tag{9}$$

where $\sigma_{\Delta x_i}$ and $\sigma_{\Delta y_i}$ are the standard deviations on x-axis and y-axis, respectively, and ω is the correlation between x-axis and y-axis.

3. Expectation Maximization-based Data Association

In this section, we first assume sensor positions are accurate, i.e., $\mathbf{p}_i = \overline{\mathbf{p}}_i$. In the proposed DA process, we are interested in parameters $\mathbf{x} = [\mathbf{q}, \boldsymbol{\theta}, \boldsymbol{\alpha}]^T$. The log-likelihood function is given by considering the mapping information $\rho_{i,j,k}$ as follows

$$\begin{aligned}
\ln p(\Delta\hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \boldsymbol{\rho} \,|\, \mathbf{x}) &= \ln p(\{\{\Delta\hat{d}_{i,j}, \hat{\theta}_{i,j}\}_{j=1}^{M_i}\}_{i=1}^{N}, \{\{\{\rho_{i,j,k}\}_{k=1}^{\mathrm{K}_{i,j}}\}_{j=1}^{M_i}\}_{i=1}^{N} \,|\, \mathbf{x}) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M_i} \ln p(\hat{\theta}_{i,j} \,|\, \theta_{i,j}) + \sum_{i=1}^{N} \sum_{j=1}^{M_i} \sum_{k=1}^{\mathrm{K}_{i,j}} \ln p(\rho_{i,j,k} \,|\, \mathbf{q}, \theta_{i,j}, \alpha_{i,j,k}) \\
&\quad + \sum_{i=1}^{N} \sum_{j=1}^{M_i} \sum_{k=1}^{\mathrm{K}_{i,j}} \ln p(\Delta\hat{d}_{i,j} \,|\, \mathbf{q}, \theta_{i,j}, \alpha_{i,j,k}, \rho_{i,j,k}),
\end{aligned} \tag{10}$$

where

$$p(\rho_{i,j,k} \,|\, \mathbf{q}, \theta_{i,j}, \alpha_{i,j,k}) \propto p(\mathbf{q} \,|\, \rho_{i,j,k}, \theta_{i,j}, \alpha_{i,j,k}) p(\rho_{i,j,k}). \tag{11}$$

Since the space of possible target location is restricted to the area spanned by the direction of the AOA at the sensor and the direction opposite to the angle of departure (AOD) at the target [22], we have $p(\mathbf{q} \,|\, \rho_{i,j,k}, \theta_{i,j}, \alpha_{i,j,k}) = 1$ in (11). As each model has a certain weight $\alpha_{i,j,k}$, the prior of $\rho_{i,j,k}$ can be written as

$$p(\rho_{i,j,k}) = (\alpha_{i,j,k})^{\rho_{i,j,k}} (1 - \alpha_{i,j,k})^{(1-\rho_{i,j,k})}. \tag{12}$$

1) *E*-step: The *E*-step is to account for the mapping information. Here, we have

$$\mathcal{Q}(\mathbf{x}, \mathbf{x}^{l-1}) = \sum_{\boldsymbol{\rho}} p(\boldsymbol{\rho} \,|\, \Delta\hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \mathbf{x}^{l-1}) \cdot \ln p(\Delta\hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \boldsymbol{\rho} \,|\, \mathbf{x}), \tag{13}$$

where l is the number of iteration and

$$p(\boldsymbol{\rho} \,|\, \Delta\hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \mathbf{x}^{l-1}) \propto p(\Delta\hat{\mathbf{d}} \,|\, \mathbf{x}^{l-1}, \boldsymbol{\rho}) \cdot p(\boldsymbol{\rho} \,|\, \mathbf{x}^{l-1}). \tag{14}$$

2) *M*-step: The *M*-step is to maximize $\mathcal{Q}(\mathbf{x}, \mathbf{x}^{l-1})$ with respect to $\mathbf{x}$, i.e.,

$$\mathbf{x}^l = \arg \max_{\mathbf{x}} \mathcal{Q}(\mathbf{x}, \mathbf{x}^{l-1}) . \qquad (15)$$

The *E*-step and *M*-step will repeat for $\mathrm{N}_{\mathrm{iter}}$ times until the estimation of target location $\mathbf{q}$ converges. Then, based on $\alpha_{i,j,k}^l$ for $k = 1, \cdots, \mathrm{K}_{\mathrm{i,j}}$, we can perform DA to determine the model coefficients. We set $\rho_{i,j,k^*} = 1$, with $k^* = \arg \max_{k} \alpha_{i,j,k}^l$; otherwise, $\rho_{i,j,k}$ is zero in the Gaussian mixture model.

4. Target Localization with Inaccurate Collaborative Sensors

The position of sensors may be inaccurate in practice due to the misalignment of the sensors. In this case, the position error $\{\Delta \mathbf{p}_i\}_{i=1}^N$ should be considered to improve the performance of target localization. However, when N is large, using the above EM algorithm directly will lead to huge computational complexity and low convergence rate. To solve this problem, we first approximate the posterior distribution of $\Delta \mathbf{p}$ by a 2-D circularly symmetric Gaussian distribution via Kullback-Leibler divergence minimization [34]. Then, we propose a SAGE framework to update the target position estimate $\mathbf{q}$. Since the sensor's processing ability is limited, we consider both the centralized and the distributed network implementations of the proposed algorithms in the following subsections. We will show that the distributed algorithm can attain the performance of the centralized algorithm by using only the local information at each sensor.

4.1 Centralized SAGE with Sensor Position Uncertainty

Based on the DA process, the uncertainty of model coefficients related to $\boldsymbol{\alpha}$ is solved. Here, we define the parameters of interest as $\mathbf{x} = [\mathbf{q}, \boldsymbol{\theta}]$. Since the measurements are independent of each other, the likelihood function can be expressed as

$$p(\Delta \hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \overline{\mathbf{p}} \mid \mathbf{x}, \Delta \mathbf{p}) = \prod_{i=1}^N \prod_{j=1}^{M_i} p(\Delta \hat{d}_{i,j}, \hat{\theta}_{i,j}, \overline{\mathbf{p}}_i \mid \mathbf{q}, \Delta \mathbf{p}_i, \theta_{i,j}) . \qquad (16)$$

Then, we have

$$\tilde{\mathcal{Q}}(\mathbf{x}, \mathbf{x}^{l-1}) = \int p(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \times \ln p(\Delta \hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \Delta \mathbf{p}_i, \overline{\mathbf{p}}_i \mid \mathbf{x}) d\Delta \mathbf{p}_i , \qquad (17)$$

where

$$\ln p(\Delta \hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}, \Delta \mathbf{p}_i, \overline{\mathbf{p}}_i \mid \mathbf{x}) = \sum_{i=1}^N \sum_{j=1}^{M_i} \ln p(\hat{\theta}_{i,j} \mid \theta_{i,j}) + \sum_{i=1}^N \sum_{j=1}^{M_i} \ln p(\Delta \hat{d}_{i,j}, \Delta \mathbf{p}_i, \overline{\mathbf{p}}_i \mid \mathbf{x}) . \qquad (18)$$

Since the first term in (18) does not depend on $\mathbf{x}$, it can be dropped during the maximization of $\tilde{\mathcal{Q}}(\mathbf{x}, \mathbf{x}^{l-1})$. Therefore, $\tilde{\mathcal{Q}}(\mathbf{x}, \mathbf{x}^{l-1})$ can be reformulated as

$$\tilde{\mathcal{Q}}(\mathbf{x}, \mathbf{x}^{l-1}) = \int p(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \times \ln p(\Delta \hat{\mathbf{d}}, \Delta \mathbf{p}_i, \overline{\mathbf{p}}_i \mid \mathbf{x}) d\Delta \mathbf{p}_i . \qquad (19)$$

Using Bayes' rule, $p(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1})$ in (19) can be expressed as

$$p(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \propto p(\Delta \mathbf{p}_i) p(\Delta \hat{\mathbf{d}} \mid \overline{\mathbf{p}}_i, \mathbf{x}, \Delta \mathbf{p}_i), \qquad (20)$$

where $p(\Delta\hat{\mathbf{d}} \mid \overline{\mathbf{p}}_i, \mathbf{x}, \Delta\mathbf{p}_i)$ is the likelihood function corresponding to the j-th measurement of S_i. Since $p(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1})$ is in general not Gaussian, no closed-form expression of (19) can be obtained. To solve this problem, we approximate $p(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1})$ by a 2-D circularly symmetric Gaussian distribution, with zero mean and variance $\sigma^2_{\Delta\mathbf{p}_i}$, i.e.,

$$f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) = \frac{1}{2\pi\sigma^2_{\Delta\mathbf{p}_i}} \exp(-\frac{\|\Delta\mathbf{p}_i\|^2}{2\sigma^2_{\Delta\mathbf{p}_i}}),$$ (21)

The parameters in (21) can be optimized by minimizing the KLD, i.e.,

$$D_{KL}(f \parallel p) = \int f(\Delta\mathbf{p}_i) \ln \frac{f(\Delta\mathbf{p}_i)}{p(\Delta\mathbf{p}_i)} d\Delta\mathbf{p}_i.$$ (22)

By substituting (20) and (21) into (22), we have (23). More details are shown in Appendix A.

$$\begin{aligned}
D_{KL}(f \parallel p) = &\int f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \ln f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) d\Delta\mathbf{p}_i \\
&- \int f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \ln p(\Delta\mathbf{p}_i) d\Delta\mathbf{p}_i \\
&- \int f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \times \ln \prod_{j \in M_i} p(\Delta\hat{d}_{i,j}, \hat{\theta}_{i,j} \mid \mathbf{q}, \overline{\mathbf{p}}_i, \Delta\mathbf{p}_i, \theta_{i,j}) d\Delta\mathbf{p}_i
\end{aligned}$$ (23)

1) NLOS Case: For sensors that do not have LOS component to the target, by combining (46)-(49) in Appendix A, we obtain (24).

$$\begin{aligned}
D_{KL}(f \parallel p) = &\mathcal{C} - \ln\sigma^2_{\Delta\mathbf{p}_i} + \frac{1}{2(1-\rho^2)}\left[\frac{\sigma^2_{\Delta\mathbf{p}_i} + \overline{x}_i^2}{\sigma^2_{\Delta x_i}} + \frac{\sigma^2_{\Delta\mathbf{p}_i} + \overline{y}_i^2}{\sigma^2_{\Delta y_i}} - \frac{2\overline{x}_i\overline{y}_i}{\sigma_{\Delta x_i}}\sigma_{\Delta y_i}\right] \\
&- \sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2}\left[2E_{i,j}\overline{x}_i + 2F_{i,j}\overline{y}_i - 2a_{i,j}b_{i,j}\overline{x}_i\overline{y}_i - a_{i,j}^2(\sigma^2_{\Delta\mathbf{p}_i} + \overline{x}_i^2) - b_{i,j}^2(\sigma^2_{\Delta x_i} + \overline{y}_i^2)\right]
\end{aligned}$$ (24)

where $\mathcal{C}$ is a constant.

Then the global $\Delta\mathbf{p}^l$ can be expressed as

$$\Delta\mathbf{p}^l = \arg\min_{\Delta p}\left\{D_{KL}(f(\Delta\mathbf{p} \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \parallel f(\Delta\mathbf{p} \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}))\right\}.$$ (25)

The solution of (25) can be obtained by setting the partial derivatives of $D_{KL}(f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}) \parallel f(\Delta\mathbf{p}_i \mid \Delta\hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{x}^{l-1}))$ with respect to $\overline{x}_i, \overline{y}_i$ and $\sigma^2_{\Delta\mathbf{p}_i}$ to zeros. After this, we have $\Delta\mathbf{p}^l = (\Delta x_i^l, \Delta y_i^l)$. And more details are shown in Appendix B. For the i-th sensor containing only NLOS components, we obtain (26).

$$\begin{aligned}
\tilde{\mathcal{Q}}_{i,\backslash LOS}(\mathbf{q}, \mathbf{q}^{l-1}) \propto &-\sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2}\left[(a_{i,j} - a_0)^2 q_x^2 + (b_{i,j} - b_0)^2 q_y^2 - 2(a_{i,j} - a_0)\mathrm{H}_{i,j}q_x \right. \\
&\left. - 2(b_{i,j} - b_0)\mathrm{H}_{i,j}q_y + 2\mathrm{K}_{i,j}q_x q_y\right]
\end{aligned}$$ (26)

2) LOS Case: Since the LOS component is much reliable than the NLOS ones, we drop the other MPCs in this case. Then, for sensor i we have $D_{KL}(f \parallel p)$ as

$$D_{KL}(f \parallel p) = -\ln \sigma_{\Delta \mathbf{p}_i}^2 + \frac{1}{2(1-\rho^2)} \Big[\frac{\sigma_{\Delta \mathbf{p}_i}^2 + \overline{x}_i^2}{\sigma_{\Delta x_i}^2} + \frac{\sigma_{\Delta \mathbf{p}_i}^2 + \overline{y}_i^2}{\sigma_{\Delta y_i}^2} - \frac{2\overline{x}_i \overline{y}_i}{\sigma_{\Delta x_i} \sigma_{\Delta y_i}} \Big]$$

$$- \frac{D_{i,j}}{\sigma_i^2} \Big(\parallel \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \parallel + \frac{\overline{x}_i - q_x^{l-1}}{\parallel \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \parallel} \overline{x}_i + \frac{\overline{y}_i - q_y^{l-1}}{\parallel \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \parallel} \overline{y}_i \Big) \tag{27}$$

$$+ \frac{1}{2\sigma_i^2} \big(2\sigma_{\Delta \mathbf{p}_i}^2 + \parallel \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \parallel^2 \big) + \mathcal{C}$$

Similar to the NLOS case, we have $\Delta \mathbf{p}^l = (\Delta x_i^l, \Delta y_i^l)$, which is shown in Appendix B. Then, we can obtain the closed form of (26) as (28),

$$\tilde{\mathcal{Q}}_{i,LOS}(\mathbf{q}, \mathbf{q}^{l-1}) \propto \ln \frac{p(\hat{\theta}_{i,j} \mid \theta_{i,j}) \alpha_{i,j,k}}{\sqrt{2\pi\sigma_i^2}} - \frac{1}{2\sigma_i^2} \Big[D_{i,j}^2(\mathbf{q}) - 2D_{i,j}(\mathbf{q}) \sigma_{\Delta \mathbf{p}_i} \sqrt{\frac{\pi}{2}}$$

$$\times {}_1 F_1 \Big(-\frac{1}{2}; 1; -\frac{\parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^{l-1} - \mathbf{q} \parallel^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big) + 2\sigma_{\Delta \mathbf{p}_i} + \parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^{l-1} - \mathbf{q} \parallel^2 \Big] \tag{28}$$

where ${}_1 F_1(a;b;c)$ is the first kind confluent hypergeometric function [35].

Since this hypergeometric function has a complex expression for analysis and calculation, we rely on the first order Taylor expansion around the previous estimation $\mathbf{q}^{l-1}$ to obtain an approximated estimation. We adopt similar operations to [32] but we use the first order Taylor series to estimate $\mathbf{q}^l = [q_x^l, q_y^l]$, i.e.,

$$f(\mathbf{q}) \approx f(\mathbf{q}^{l-1}) + f'_{q_x}(\mathbf{q}^{l-1})(q_x - q_x^{l-1}) + f'_{q_y}(\mathbf{q}^{l-1})(q_y - q_y^{l-1}), \tag{29}$$

where $f(\mathbf{q}) = {}_1 F_1 \Big(-\frac{1}{2}; 1; -\frac{\parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^l - \mathbf{q} \parallel^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big)$, then we have the following derivations [36].

$$f'_{q_x}(\mathbf{q}^{l-1}) = \frac{q_x - (\Delta x_i^l + \overline{x}_i)}{2\sigma_{\Delta \mathbf{p}_i}^2} {}_1 F_1 \Big(\frac{1}{2}; 2; -\frac{\parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^l - \mathbf{q} \parallel^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big) \tag{30}$$

$$f'_{q_y}(\mathbf{q}^{l-1}) = \frac{q_y - (\Delta y_i^l + \overline{y}_i)}{2\sigma_{\Delta \mathbf{p}_i}^2} {}_1 F_1 \Big(\frac{1}{2}; 2; -\frac{\parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^l - \mathbf{q} \parallel^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big) \tag{31}$$

Then (28) can be formulated as (32) by dropping the irrelevant terms.

$$\tilde{\mathcal{Q}}_{i,LOS}(\mathbf{q}, \mathbf{q}^{l-1}) \propto -\frac{1}{2\sigma_i^2} \Big[D_{i,j}^2(\mathbf{q}) - 2D_{i,j}(\mathbf{q}) \sigma_{\Delta \mathbf{p}_i} \sqrt{\frac{\pi}{2}} \times {}_1 F_1 \Big(-\frac{1}{2}; 1; -\frac{\parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^l - \mathbf{q} \parallel^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big)$$

$$+ \parallel \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i^l - \mathbf{q} \parallel^2 \Big] \tag{32}$$

By substituting (26) and (32) into (17) and dropping the irrelevant terms, we have

$$\tilde{\mathcal{Q}}(\mathbf{q}, \mathbf{q}^{l-1}) = \sum_{i \in \mathcal{N}_{\backslash LOS}} \tilde{\mathcal{Q}}_{i,\backslash LOS}(\mathbf{q}, \mathbf{q}^{l-1}) + \sum_{i \in \mathcal{N}_{LOS}} \tilde{\mathcal{Q}}_{i,LOS}(\mathbf{q}, \mathbf{q}^{l-1}). \tag{33}$$

4.2 Distributed SAGE with Sensor Position Uncertainty

The centralized method needs a fusion center which may involve high operating costs and increase the risk of failure if the fusion center break down [37]. To this ends, we propose a

distributed implementation of the proposed method. Specificially, an average consensus scheme is proposed to acquire the conditioned expectation of $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ locally at each sensor. As $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ are independent to each other, each sensor can calculate $\tilde{\mathcal{Q}}_i$ locally, i.e., $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1}) = \tilde{\mathcal{Q}}_{i, \backslash LOS}(\mathbf{q}, \mathbf{q}^{l-1})$ or $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1}) = \tilde{\mathcal{Q}}_{i, LOS}(\mathbf{q}, \mathbf{q}^{l-1})$. Then the global $\tilde{\mathcal{Q}}(\mathbf{q}, \mathbf{q}^{l-1})$ can be obtained as (33) to calculate the interested parameters in vector $[\mathrm{q}_x, \mathrm{q}_y]^T$. It is clear that average consensus aims for exchanging the conditioned expectation of the local log-likelihood functions $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ between neighboring sensors. For the proposed distributed algorithm, we use p to represent the iteration index of the consensus algorithm among sensors, and the update process can be written as

$$
\begin{aligned}
\tilde{\mathcal{Q}}_i^{(p)}(\mathbf{q}, \mathbf{q}^{l-1}) &= \tilde{\mathcal{Q}}_i^{(p-1)}(\mathbf{q}, \mathbf{q}^{l-1}) + \sum_{k \in \mathcal{P}_i} \xi_{ki} \left((\tilde{\mathcal{Q}}_k^{(p-1)}(\mathbf{q}, \mathbf{q}^{l-1}) - \tilde{\mathcal{Q}}_i^{(p-1)}(\mathbf{q}, \mathbf{q}^{l-1}) \right) \\
&= \xi_{ii} \tilde{\mathcal{Q}}_i^{(p-1)}(\mathbf{q}, \mathbf{q}^{l-1}) + \sum_{k \in \mathcal{P}_i} \xi_{ki} \tilde{\mathcal{Q}}_k^{(p-1)}(\mathbf{q}, \mathbf{q}^{l-1}),
\end{aligned}
\tag{34}
$$

which includes the metropolis update rate as

$$
\xi_{ki} = \xi_{ik} = \begin{cases} 1 / \max(|\mathcal{P}_i|, |\mathcal{P}_k|), & \text{for } i \neq k, \\ 1 - \sum_{k' \in \mathcal{P}_i} \xi_{k'i}, & \text{for } i = k, \end{cases}
$$

where $\mathcal{P}_i$ and $\mathcal{P}_k$ denote the index sets of neighboring sensor of the sensor i and k, and $|\mathcal{P}_i|$ and $|\mathcal{P}_k|$ stand for the numbers of neighboring sensors. After N'_{iter} iterations, all sensors achieve consensus locally, i.e., $\tilde{\mathcal{Q}}_i^{(\mathrm{N}'_{iter})}(\mathbf{q}, \mathbf{q}^{l-1}) = \tilde{\mathcal{Q}}_j^{(\mathrm{N}'_{iter})}(\mathbf{q}, \mathbf{q}^{l-1})$. Finally, all sensors are able to update positions information simultaneously, i.e.,

$$
\tilde{\mathcal{Q}}(\mathbf{q}, \mathbf{q}^{l-1}) = \mathrm{N} \cdot \tilde{\mathcal{Q}}_i^{(\mathrm{N}'_{iter})}(\mathbf{q}, \mathbf{q}^{l-1}).
\tag{35}
$$

The convergence of average consensus has been proved for connected sensor networks in [38].

As $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ is continuous and sensors communicates with each other by wireless channels, particle-based methods are typically used to broadcast the $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ among sensors. To obtain precise localization, a large amount of particles are required, which increases the complexity significantly. As a quadratic polynomial form exists in both (26) and (32), we can approximate $\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1})$ as follows. For sensor i which obtains LOS/NLOS signal, we can unify the expression as

$$
\tilde{\mathcal{Q}}_i(\mathbf{q}, \mathbf{q}^{l-1}) \approx -\frac{(q_x - \bar{m}_{x_i})^2}{\sigma_{x_i}^2} - \frac{(q_y - \bar{n}_{y_i})^2}{\sigma_{y_i}^2} - \bar{o}_{x_i y_i} q_x q_y.
\tag{36}
$$

According to the quadratic polynomial approximation, in the average consensus operation, only the mean and variance of $\mathbf{q}$ and the coefficient $\bar{o}_{x_i y_i}$ of $q_x q_y$ should be exchanged

among sensors, which reduces the communication overhead significantly.

Referring to the metropolis update rate, the local likelihood function at the p-th iteration of the consensus algorithm can be given as

$$\tilde{\mathcal{Q}}_i^{(p)}(\mathbf{q},\mathbf{q}^{l-1}) \approx -\frac{(q_x - m_i(p))^2}{\sigma_{x_i}^2(p)} - \frac{(q_y - n_i(p))^2}{\sigma_{y_i}^2(p)} - o_i(p)q_xq_y, \tag{37}$$

where $m_i^l(p), \sigma_{m_i^l}^2(p), n_i^l(p), \sigma_{n_i^l}^2(p)$ and $o_i^l(p)$ are updated as

$$m_i^l(p) = \sigma_{m_i^l}^2(p-1)\left(\frac{\xi_{ii}m_i^l(p-1)}{\sigma_{m_i^l}^2(t)} + \sum_{k\in\mathcal{N}_i}\frac{\xi_{ki}m_i^l(p-1)}{\sigma_{x_k}^2(p-1)}\right), \tag{38}$$

$$n_i^l(p) = \sigma_{n_i^l}^2(p-1)\left(\frac{\xi_{ii}n_i^l(p-1)}{\sigma_{n_i^l}^2(t)} + \sum_{k\in\mathcal{N}_i}\frac{\xi_{ki}n_i^l(p-1)}{\sigma_{y_k}^2(p-1)}\right), \tag{39}$$

$$\sigma_{m_i^l}^2(p) = \left(\frac{\xi_{ii}}{\sigma_{m_i^l}^2(p-1)} + \sum_{k\in\mathcal{N}_i}\frac{\xi_{ki}}{\sigma_{x_k}^2(p-1)}\right)^{-1}, \tag{40}$$

$$\sigma_{n_i^l}^2(p) = \left(\frac{\xi_{ii}}{\sigma_{n_i^l}^2(p-1)} + \sum_{k\in\mathcal{N}_i}\frac{\xi_{ki}}{\sigma_{y_k}^2(p-1)}\right)^{-1}, \tag{41}$$

$$o_i^l(p) = \xi_{ii}o_i^l(p-1) + \sum_{k\in\mathcal{N}_i}\xi_{ki}o_i^l(p-1). \tag{42}$$

1) NLOS case: For the sensors which obtain NLOS components only, we adjust the

$\tilde{\mathcal{Q}}_{i,\backslash LOS}(\mathbf{q},\mathbf{q}^{l-1})$'s expression and approximates it as follows

$$\tilde{\mathcal{Q}}_{i,\backslash LOS}(\mathbf{q},\mathbf{q}^{l-1}) \approx -\frac{(q_x - \hat{m}_i^l)^2}{2\sigma_{\hat{m}_i^l}^2} - \frac{(q_y - \hat{n}_i^l)^2}{2\sigma_{\hat{n}_i^l}^2} - \hat{o}_i^l q_xq_y, \tag{43}$$

where $\hat{m}_i^l$, $\hat{n}_i^l$ and $\hat{o}_i^l$ are expressed as $\hat{m}_i^l = \mathrm{I}_i/(a_{i,j} - a_0)$, $\hat{n}_i^l = \mathrm{H}_i/(b_{i,j} - b_0)$, $\sigma_{\hat{m}_i^l}^2 = \sigma_i^2/M_i(a_{i,j} - a_0)$, $\sigma_{\hat{n}_i^l}^2 = \sigma_i^2/M_i(b_{i,j} - b_0)$ and $\hat{o}_i^l = \dfrac{\mathrm{K}_i}{\sigma_i^2/M_i}$.

2) LOS case: For sensors which obtain LOS components, as the nonlinear term exists, we adopt the first order Taylor expansion in (32), which becomes

$$\tilde{\mathcal{Q}}_{i,LOS}(\mathbf{q},\mathbf{q}^{l-1}) \approx -\frac{(q_x - \bar{m}_i^l)^2}{2\sigma_{\bar{m}_i^l}^2} - \frac{(q_y - \bar{n}_i^l)^2}{2\sigma_{\bar{n}_i^l}^2} - \bar{o}_i^l q_xq_y. \tag{44}$$

For the i-th iteration, the local function $\tilde{\mathcal{Q}}_i(\mathbf{q},\mathbf{q}^{l-1})$ is initialized based on (43) and (44). After several iterations, each sensor gets the global function $\tilde{\mathcal{Q}}(\mathbf{q},\mathbf{q}^{l-1})$. We maximize $\tilde{\mathcal{Q}}(\mathbf{q},\mathbf{q}^{l-1})$ to obtain $\mathbf{q}$'s final estimate, which can be written as

$$\mathbf{q}^l = [m_i(\mathrm{N}'_{iter}), n_i(\mathrm{N}'_{iter})]^T. \tag{45}$$

5. Performance Evaluation

5.1 Simulation Setup

We evaluate the performance of our proposed DA algorithm and localization algorithms in both centralized and distributed implementation. Without loss of generality, we adopt $100 \times 100 m^2$ space with one target $\mathbf{q}$ and four sensors $S_1, \ldots, S_4$. The important parameters related to the simulations are summarized in **Table 1**. Sensors' positions are $\mathbf{p}_1 = [20\ 20]^T$, $\mathbf{p}_2 = [80\ 30]^T$, $\mathbf{p}_3 = [60\ 90]^T$, $\mathbf{p}_4 = [70\ 70]^T$ and the standard deviation $\sigma_{\Delta \mathbf{p}_i}$ in **Table 1** is assumed from $1m \sim 3m$. The number of scatters is set to $L = 8$ in **Fig. 1**. The corresponding orientations are $\gamma = [0^{\circ}, 86^{\circ}, 150^{\circ}, 90^{\circ}, 111^{\circ}, 55^{\circ}, 135^{\circ}, 11^{\circ}]$, which hold smooth surfaces and long enough to reflect signals compared with other scatters. Furthermore, the ground-truth AOAs are $\boldsymbol{\theta} = [-35^{\circ} - 135^{\circ}; 168^{\circ} - 60^{\circ}; 30^{\circ}; -155^{\circ} 152^{\circ}]$, which corresponding to S_1 to S_4, as illustrated in **Fig. 1**. In this simulation scenario, the number of measurements ($|\mathcal{M}_1| \sim |\mathcal{M}_4|$) that sensor nodes ($S_1 \sim S_4$) received are set to $[2, 2, 1, 2]^T$ respectively.

Table 1. Important Parameters

Parameter	Note	Value
L × W	Simulation space	$100m \times 100m$
$\mathbf{p}_1$	Reference node S_1	$[20m\ 20m]^T$
$\mathbf{q}$	Object node	$[40m\ 50m]^T$
$\alpha_{i,j,k}$	Weights of the submodel	$1/K_{i,j}$
$\Delta \mathbf{p}_i$	Sensor i's uncertainty	$\Delta \mathbf{p}_i \sim N(0, \sigma_{\Delta \mathbf{p}_i})$
$\sigma_{\Delta \hat{\mathbf{d}}}$	Std. deviation for $\Delta \hat{\mathbf{d}}$	1m
$\eta_{i,j}$	Std. deviation for AOA	$\eta_{i,j} \sim \mathrm{Unif}[-3^{\circ}, 3^{\circ}]$
$K_{i,j}$	The number of submodels for first MPC The number of submodels for other MPCs	$K_{1,j} = 9$ $K_{i,j} = 8, j > 1$

5.2 Localization Performance

We perform Monte Carlo simulations with 1,000 independent trials. A coarse localization of the proposed algorithm is evaluated according to the proposed DA method in section III, without considering sensor position uncertainty as in (10). The performance of the method in [10] is also plotted for comparison. It is seen in **Fig. 2** that the localization error of the method in [10] is significantly greater than that of the proposed algorithm with DA. This is due to the fact that [10] does not perform the DA process, which makes the assumption that all measurements are mixture of LOS and NLOS. On the other hand, without considering the uncertainty of sensor position, the values of each submodel's weights become inaccurate, which increases the risk of mismatch in the DA process. Accordingly, we can see in **Fig. 2** that

the localization performance of the proposed DA method becomes worse when increasing sensor position uncertainty. By comparing the results in **Fig. 2**, the proposed algorithm with small sensor position uncertainty outperforms the method in [10] and performs close to the ideal one, which demonstrates the effectiveness of DA process.
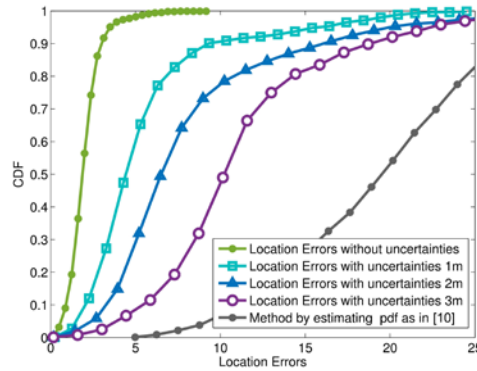

Fig. 2. Coarse localization performance with different uncertainty of sensor position.

We evaluate the localization performance by taking into account sensors' uncertainties with (33) and (45) in **Fig. 3**, which corresponds to the proposed centralized and distributed SAGE methods, respectively. Comparing the results in **Fig. 2** and **Fig. 3**, we can observe that the proposed SAGE algorithm shows superiors performance by jointly updating the target and the sensors' positions simultaneously. Moreover, it is seen in **Fig. 3** that the distributed implementation performs very close to the centralized method, which demonstrates the effectiveness of the quadratic polynomial approximation and the averaged consensus employed in the distributed implementation.
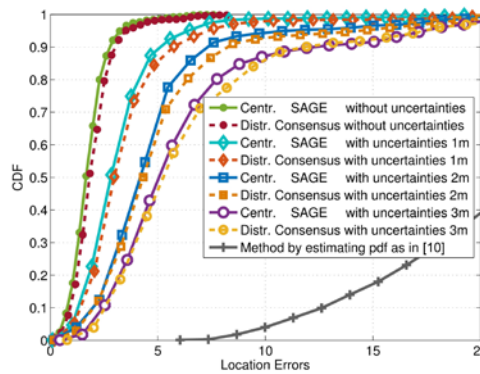

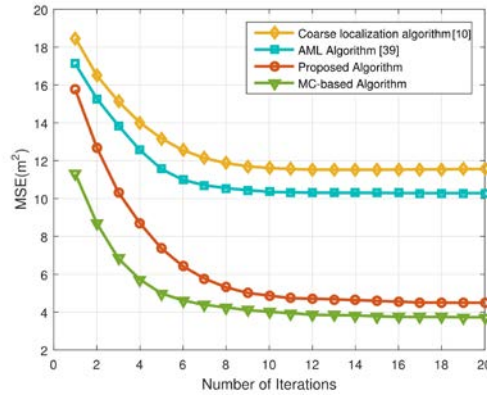**Fig. 3.** Localization performance of the proposed centralized and distributed implementations.

Fig. 4. The MSE performance of different estimators

In **Fig. 4**, the mean square error (MSE) performance of the proposed centralized algorithm is compared with the state-of-the-art methods, such as the coarse localization estimator, the approximate maximum-likelihood (AML) estimator [39], and a Monte Carlo (MC)-based estimator with 2000 particles, where we assume that the standard deviation of sensors position uncertainty is $\sigma_{\Delta \mathbf{p}_i} = 1\,\text{m}$. Without considering sensors uncertainties and DA process, the coarse localization estimator gives the worst performance. Since the AML estimator treats the sensors position uncertainties as Delta functions after DA process, it shows larger MSE than the proposed algorithm. The MC-based estimator takes samples from sensor i's position uncertainty's distribution, which transforms the intractable integration into the cumulative sum of finite terms. However, the superior performance of the MC-based estimator is at the cost of huge computational complexity. The proposed SAGE algorithm can perform close to the MC-based estimator with much lower complexity.
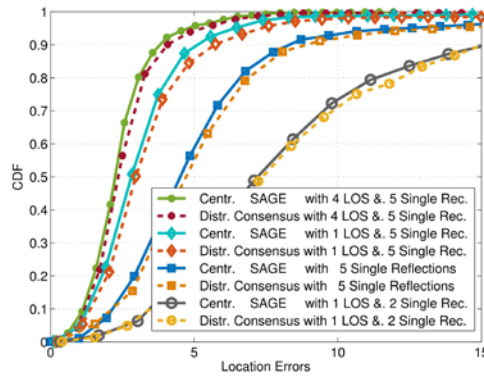

Fig. 5. Localization performance with different numbers of measurements combination.

Since the number of LOS and NLOS components may vary for different locations. In **Fig. 5**, we evaluate the impact of different number of LOS and NLOS components on the localization performance. We assume the standard deviation of sensors position uncertainty is $\sigma_{\Delta \mathbf{p}_i} = 1\,\text{m}$. The four positions (90 40), (40 50), (40 10), (80 20) are named as

- Case 1, a relatively ideal scenario: 4 LOS signals and 5 NLOS signals at (90 40).
- Case 2, a general situation of the complex scenario: 1 LOS signal and 5 NLOS signals at (40 50).

- Case 3, a severe NLOS situation: with 5 NLOS signals at (40 10).
- Case 4, an extreme scenario: 1 LOS signal and 2 NLOS signals at (80 20).

It is observed in **Fig. 5** that both the centralized and distributed implementation of the proposed algorithm can work with different combinations of measurements. The localization performance of Case 4 is worse than that of Case 3, although the former has 1 LOS measurement. This indicates that enough number of measurements are useful for localization and the utilization of the NLOS MPCs can be beneficial. Comparing Case 3 with Case 2, it is seen that the LOS signals can further improve the performance since the LOS signal has no AOA measurement error. When increasing the number of LOS measurements, e.g., from Case 2 to Case 1, localization performance can be improved. However, the improvement becomes marginal when giving sufficient number of NLOS signals.

6. Conclusion

In this paper, a novel expectation maximization (EM)-based data association process was proposed to solve the target localization problem in multipath environments. Considering the position uncertainty of sensors, we proposed a SAGE-based algorithm using TDOA and AOA measurements to jointly locate the target and update the position of sensors. The KLD minimization was employed to approximate the distribution of sensor uncertainty by a **2-D** circularly symmetric Gaussian distribution. Then, Taylor expansion was used to linearize the specific nonlinear term induced by the first kind confluent hypergeometric function in *E*-step. Accordingly, a closed-form expression was derived with low computational requirements. To improve the scalability of the proposed algorithms in complex scenarios, we further derived a distributed implementation based on the average consensus. Simulation results showed that the proposed algorithms performed close to the Monte Carlo-based algorithm with much lower communication overhead and complexity. The distributed implement achieved an accuracy close to the proposed centralized SAGE-based algorithm.

Appendix A

In this appendix, we give a further derivation of the KLD in (23). Here, we have

$$\int f(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}_i, \overline{\mathbf{p}}_i, \mathbf{q}^{l-1}) \ln f(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}_i, \overline{\mathbf{p}}_i, \mathbf{q}^{l-1}) d\Delta \mathbf{p}_i$$

$$= \int \frac{1}{2\pi\sigma_{\Delta \mathbf{p}_i}^2} \exp(-\frac{\| \Delta \mathbf{p}_i - \Delta \overline{\mathbf{p}}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2}) \times \Big(\ln \frac{1}{2\pi\sigma_{\Delta \mathbf{p}_i}^2} - \frac{\| \Delta \mathbf{p}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2} \Big) d\Delta \mathbf{p}_i = \ln \frac{1}{2\pi\sigma_{\Delta \mathbf{p}_i}^2} - 1. \tag{46}$$

The latter part can be derived as

$$\int f(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}_i, \overline{\mathbf{p}}_i, \mathbf{q}^{l-1}) \ln p(\Delta \mathbf{p}_i) d\Delta \mathbf{p}_i$$

$$= \int \frac{1}{2\pi\sigma_{\Delta \mathbf{p}_i}^2} \exp(-\frac{\| \Delta \mathbf{p}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2}) \times \Big\{ \ln \frac{1}{2\pi\sigma_{\Delta_{x_i}} \sigma_{\Delta_{y_i}} \sqrt{1-\rho^2}} - \frac{1}{2(1-\rho^2)}$$

$$\times \Big[\frac{\Delta x_i^2}{\sigma_{\Delta_{x_i}}^2} + \frac{\Delta y_i^2}{\sigma_{\Delta_{y_i}}^2} - \frac{2\rho\Delta x_i \Delta y_i}{\sigma_{\Delta_{x_i}} \sigma_{\Delta_{y_i}}} \Big] \Big\} d\Delta \mathbf{p}_i. \tag{47}$$

The last part can be formulated as

$$\int f(\Delta \mathbf{p}_i \mid \Delta \hat{\mathbf{d}}, \overline{\mathbf{p}}_i, \mathbf{q}^{l-1}) \ln \prod_{j \in M_i} p(\Delta \hat{d}_{i,j}, \hat{\theta}_{i,j} \mid \mathbf{q}^{l-1}, \overline{\mathbf{p}}_i, \Delta \mathbf{p}_i, \theta_{i,j}) d\Delta \mathbf{p}_i$$

$$= \int \frac{1}{2\pi \sigma_{\Delta \mathbf{p}_i}^2} \exp(-\frac{\| \Delta \mathbf{p}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2}) \times \sum_{j=1}^{M_i} \Big[\ln \frac{p(\hat{\theta}_{i,j} \mid \theta_{i,j}) \alpha_{i,j,k}^l}{\sqrt{2\pi \sigma_i^2}} - \frac{1}{2\sigma_i^2} \big(D_{i,j} - \Delta \hat{d}_{i,j}\big)^2 \Big] d\Delta \mathbf{p}_i, \tag{48}$$

where

$$D_{i,j} = \Delta \hat{d}_{i,j} + \mathrm{g}(\theta_{1,1}, \gamma_{1,1}')^T (\mathbf{q}^{l-1} - \mathbf{p}_1), \quad \Delta \mathbf{p}_i = (\Delta x_i, \Delta y_i), \quad \overline{\mathbf{p}}_i = (\overline{x}_i, \overline{y}_i), \quad \mathbf{q}^{l-1} = (q_x^{l-1}, q_y^{l-1}).$$

Depending on the LOS or NLOS components, we further express (48) in the following subsections.

1) NLOS case: For the NLOS components in sensor i, $d_{i,j}$ equals the item in (43). Then the integration in (48) can be calculated after some mathematical operations as

$$\sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2} \int \frac{1}{2\pi \sigma_{\Delta \mathbf{p}_i}^2} \exp(-\frac{\| \Delta \mathbf{p}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2})(2D_{i,j}\Delta \hat{d}_{i,j} - \Delta \hat{d}_{i,j}^2) d\Delta \mathbf{p}_i$$

$$= \sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2} \Big[2E_{i,j}\Delta x_i + 2F_{i,j}\Delta y_i - 2a_{i,j}b_{i,j}\Delta x_i \Delta y_i \tag{49}$$

$$- a_{i,j}^2 (\sigma_{\Delta \overline{\mathbf{p}}_i}^2 + \Delta x_i^2) - b_{i,j}^2 (\sigma_{\Delta \mathbf{p}_i}^2 + \Delta y_i^2) \Big] + \mathcal{C},$$

where $E_{i,j} = a_{i,j}^2 (x_{\mathbf{q}}^{l-1} - x_{\overline{\mathbf{p}}_i}^{l-1}) + a_{i,j}b_{i,j}(y_{\mathbf{q}}^{l-1} - y_{\overline{\mathbf{p}}_i}^{l-1}) - D_{i,j}$, $F_{i,j} = b_{i,j}^2 (y_{\mathbf{q}}^{l-1} - y_{\overline{\mathbf{p}}_i}^{l-1}) + a_{i,j}b_{i,j}(x_{\mathbf{q}}^{l-1}$ $- x_{\overline{\mathbf{p}}_i}^{l-1}) - D_{i,j}$.

2) LOS case: For sensor i which obtains LOS signal, we will process as (28). As the integral in (48) includes nonlinear terms, we use the approximation of the first-order Taylor expansion. So, the integration in (48) can be calculated after some mathematical operations as

$$\sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2} \int \frac{1}{2\pi \sigma_{\Delta \mathbf{p}_i}^2} \exp(-\frac{\| \Delta \mathbf{p}_i \|^2}{2\sigma_{\Delta \mathbf{p}_i}^2})(2D_{i,j}\Delta \hat{d}_{i,j} - \Delta \hat{d}_{i,j}^2) d\Delta \mathbf{p}_i$$

$$= \sum_{j=1}^{M_i} \Big[\frac{D_{i,j}}{\sigma_i^2} \Big(\| \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \| + \frac{\overline{x}_i - q_x'}{\| \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \|} \Delta x_i + \frac{\overline{y}_i - q_y^{l-1}}{\| \overline{\mathbf{p}}_i - \mathbf{q}^{l-1} \|} \Delta y_i \Big) \tag{50}$$

$$- \frac{1}{2\sigma_i^2} \big(2\sigma_{\Delta \mathbf{p}_i}^2 + \| \overline{\mathbf{p}}_i + \Delta \mathbf{p}_i - \mathbf{q}^{l-1} \|^2 \big) \Big]$$

Appendix B

In this appendix, we give the detail of $\Delta \mathbf{p}^l = (\Delta x_i^l, \Delta y_i^l)$,

$$\Delta x_i^l = \frac{\dfrac{\Delta y_i^{l-1}}{(1-\rho^2)\sigma_{\Delta x_i}\sigma_{\Delta y_i}} + \sum_{j=1}^{M_i}\dfrac{1}{\sigma_i^2}(E_{i,j} - a_{i,j}b_{i,j}\Delta y_i^{l-1})}{\dfrac{1}{(1-\rho^2)\sigma_{\Delta x_i}^2} - \dfrac{1}{\sigma_i^2}\sum_{j=1}^{M_i}a_{i,j}^2},$$

$$\Delta y_i^l = \frac{\dfrac{x_i^{l-1}}{(1-\rho^2)\sigma_{\Delta x_i}\sigma_{\Delta y_i}} + \sum_{j=1}^{M_i}\dfrac{1}{\sigma_i^2}(F_{i,j} - a_{i,j}b_{i,j}\Delta x_i^{l-1})}{\dfrac{1}{(1-\rho^2)\sigma_{\Delta y_i}^2} - \dfrac{1}{\sigma_i^2}\sum_{j=1}^{M_i}b_{i,j}^2},$$

$$\sigma_{\Delta \mathbf{p}_i} = \sqrt{\frac{2(1-\rho^2)}{(1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2\sum_{j=1}^{M_i}(a_{i,j}^2 + b_{i,j}^2) + \sigma_i^2(\sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2)}}\,\sigma_i\sigma_{\Delta x_i}\sigma_{\Delta y_i}, \tag{51}$$

where

$$\mathrm{H}_{i,j} = \Delta\hat{d}_{i,j} + a_{i,j}\overline{x}_i + b_{i,j}\overline{y}_i - a_0 x_0 - b_0 y_0, \tag{52}$$

$$\mathrm{K}_{i,j} = a_{i,j}b_{i,j} + a_0 b_{i,j} + a_{i,j}b_0 + a_0 b_0. \tag{53}$$

As for the LOS cases in the centrialized algorithm, we have $\Delta\mathbf{p}^l = (\Delta x_i^l, \Delta y_i^l)$ as

$$x_i^l = (D_{i,j} - \|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|) \times \Big[\Big(\frac{\sigma_i^2\sigma_{\Delta x_i}^2 + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}{\sigma_i^2(\sigma_i^2 + \sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2) + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}\Big)\frac{\overline{x}_i - q_x^{l-1}}{\|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|}$$

$$+ \Big(\frac{\rho\sigma_i^2\sigma_{\Delta x_i}\sigma_{\Delta y_i}}{\sigma_i^2(\sigma_i^2 + \sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2) + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}\Big)\frac{\overline{y}_i - q_y^{l-1}}{\|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|}\Big]$$

$$y_i^l = (D_{i,j} - \|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|) \times \Big[\Big(\frac{\rho\sigma_i^2\sigma_{\Delta x_i}\sigma_{\Delta y_i}}{\sigma_i^2(\sigma_i^2 + \sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2) + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}\Big)$$

$$\frac{\overline{x}_i - q_x^{l-1}}{\|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|} + \Big(\frac{\sigma_i^2\sigma_{\Delta y_i}^2 + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}{\sigma_i^2(\sigma_i^2 + \sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2) + (1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2}\Big)\frac{\overline{y}_i - q_y^{l-1}}{\|\overline{\mathbf{p}}_i - \mathbf{q}^{l-1}\|}\Big] \tag{54}$$

$$\sigma_{\Delta \mathbf{p}_i} = \sqrt{\frac{2(1-\rho^2)}{2(1-\rho^2)\sigma_{\Delta x_i}^2\sigma_{\Delta y_i}^2 + \sigma_i^2(\sigma_{\Delta x_i}^2 + \sigma_{\Delta y_i}^2)}}\,\sigma_i\sigma_{\Delta x_i}\sigma_{\Delta y_i}$$

References

[1] F. Zhao and L. J. Guibas, "Wireless Sensor Networks: An Information Processing Approach," *The Morgan Kaufmann Series in Networking*, pp. iii, 2004. Article (CrossRef Link).
[2] I. F. Akyildiz, W. Su, and Y. Sankarasubramaniam, "Wireless Sensor Networks: A survey," *Comput. Netw. J.*, vol. 38, pp. 393-422, Mar. 2002. Article (CrossRef Link).

[3]   K. Witrisal, P. Meissner, and E. Leitinger, "High-Accuracy Localization for Assisted Living: 5G
      systems will turn multipath channels from foe to friend," *IEEE Signal Processing Mag.*, vol. 33,
      no. 2, pp. 59-70, 2016. Article (CrossRef Link).
[4]   S. Gezici, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via
      ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal
      Processing Mag.*, vol. 22, no. 4, pp. 70-84, 2005. Article (CrossRef Link).
[5]   M. Shon, M. Jo, H. Choo, "An interactive cluster-based MDS localization scheme for multimedia
      information in wireless sensor networks," *Comput. Commun.*, vol. 35, no. 15, pp. 1921-1929, Sept.
      2012. Article (CrossRef Link).
[6]   H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc.
      IEEE*, vol. 97, no. 2, pp. 427-450, 2009. Article (CrossRef Link)
[7]   Y. Yu, "Consensus-Based Distributed Mixture Kalman Filter for Maneuvering Object Tracking in
      Wireless Sensor Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8669-8681, 2016.
      Article (CrossRef Link).
[8]   V. Tran-Quang, T. Ngo-Quynh, M . Jo, "A Lateration-localizing Algorithm for Energy-efficient
      Target Tracking in Wireless Sensor Networks," *Ad. Hoc. Sens. Wirel. Ne.,* vol. 34, no. 1-4, pp.
      191-220, Dec. 2016. Article (CrossRef Link).
[9]   G. Grisetti, R. KuMmerle, and C. Stachniss, "A tutorial on graph-based SLAM," *IEEE Intell.
      Transp. Syst. Mag.*, vol. 2, no. 4, pp. 31-43, 2010. Article (CrossRef Link).
[10] F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir, "TOA-based robust wireless geolocation and
      Cramr-Rao lower bound analysis in harsh LOS/NLOS environments," *IEEE Trans. Signal
      Processing.*, vol. 61, no.9, pp. 2243-2255, 2013. Article (CrossRef Link).
[11] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos and H. Vincent Poor, "Network Operation Strategies
      for Efficient Localization and Navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224-1254, July. 2018.
      Article (CrossRef Link).
[12] M. Z. Win, Y. Shen and W. Dai, "A Theoretical Foundation of Network Localization and
      Navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136-1165, July 2018. Article (CrossRef Link).
[13] C. K. Seow and S. Y. Tan, "Non-Line-of-Sight Localization in Multipath Environments," *IEEE
      Trans. Mob. Comput.*, vol. 7, no. 5, pp. 647-660, 2008. Article (CrossRef Link).
[14] L. Cong and W. Zhuang, "Non-line-of-sight error mitigation in TDOA mobile location," in *Proc.
      of Global Telecommunications Conference(GLOBECOM)*, vol.1, pp. 680-684, Nov. 2001.
      Article (CrossRef Link).
[15] L. Yi, S. G. Razul, Z. Lin, and C-M. See, "Individual aoameasurement detection algorithm for
      object tracking in mixed LOS/NLOS environments," in *Proc. of IEEE Int. Conf. Acoustics, Speech
      and Signal Processing*, pp. 3924-3928, 2013. Article (CrossRef Link).
[16] Y. Shen, M. Z. Win, "On the use of multipath geometry for wideband cooperative localization," in
      Proc. of Global Telecommunications Conference(GLOBECOM), pp. 1-6, 2009.
      Article (CrossRef Link).
[17] W. Yuan, N. Wu, B. Etzlinger, H. Wang and J. Kuang, "Cooperative Joint Localization and Clock
      Synchronization Based on Gaussian Message Passing in Asynchronous Wireless Networks,"
      IEEE Trans. Veh. Technol., vol. 65, no. 9, pp. 7258-7273, Sept. 2016. Article (CrossRef Link).
[18] W. Dai, Y. Shen and M. Z. Win, "Energy-Efficient Network Navigation Algorithms," *IEEE. J. Sel.
      Areas. Commun.*, vol. 33, no. 7, pp. 1418-1430, July. 2015. Article (CrossRef Link).
[19] E. Leitinger, P. Meissner, and C. Rüdisser, "Evaluation of Position-related Information in
      Multipath Components for Indoor Positioning," *IEEE J. Sel. Areas. Commun.*, vol. 33, no. 11, pp.
      2313-2328, 2015. Article (CrossRef Link).
[20] J. Li, J. Conan, and S. Pierre, "Mobile terminal location for MIMO communication systems,"
      IEEE Trans. Antennas Propagat., vol. 55, no. 8, pp. 2417-2420, 2007. Article (CrossRef Link).
[21] H. Miao, K. Yu, and M. Juntti, "Positioning for NLOS Propagation: Algorithm Derivations and
      CramerRao Bounds," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2568-2580, 2007.
      Article (CrossRef Link).

[22] W. Xu, F. Quitin, M. Leng, W. P. Tay and S. Razul, "Distributed Localization of a RF Target in NLOS Environments," *IEEE J. Sel. Areas Commun.* vol. 33, no. 7, pp. 1317-1330, July. 2015. Article (CrossRef Link).

[23] A. Athalye, V. Savic, and M. Bolic, "Novel Semi-Passive RFID System for Indoor Localization," *IEEE Sensors J.*, vol. 13, no. 2, pp. 528-537, 2013. Article (CrossRef Link).

[24] Y. Xiong, N. Wu, Y. Shen and M. Z. Win, "Cooperative Network Synchronization: Asymptotic Analysis," *IEEE Trans. Signal Processing.*, vol. 66, no. 3, pp. 757-772, Feb. 2018. Article (CrossRef Link).

[25] Y. Xiong, N. Wu, H. Wang and J. Kuang, "Cooperative Detection-Assisted Localization in Wireless Networks in the Presence of Ranging Outliers," *IEEE. Trans. Commun.*, vol. 65, no. 12, pp. 5165-5179, Dec. 2017. Article (CrossRef Link).

[26] T. Wang, Y. Shen, A. Conti and M. Z. Win, "Network Navigation With Scheduling: Error Evolution," *IEEE. Trans. Inf. Theory.*, vol. 63, no. 11, pp. 7509-7534, Nov. 2017. Article (CrossRef Link).

[27] R. Olfati-Saber, A. Fax and M. R. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215-233, 2007. Article (CrossRef Link).

[28] O. Hlinka, O. Sluciak, F. Hlawatsch, P. M. Djuric, and M.~Rupp, "Likelihood consensus and its application to distributed particle filtering," *IEEE Trans. Signal Processing.*, vol. 60, no. 8, pp. 4334-4349, 2012. Article (CrossRef Link).

[29] A. G. Dimakis, S. Kar, and J. M. F. Moura, "Gossip Algorithms for Distributed Signal Processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847-1864, 2010. Article (CrossRef Link).

[30] F. Yin, C. Fritsche, and D. Jin, "Cooperative Localization in WSNs Using Gaussian Mixture Modeling: Distributed ECM Algorithms," *IEEE Trans. Signal Processing.*, vol. 63, no. 6, pp. 1448-1463, 2015. Article (CrossRef Link).

[31] J. A. Fessler and A. O. Hero, "Space-Alternating Generalized Expectation-Maximization Algorithm," *IEEE Trans. Signal Processing.*, vol. 42, no. 10, pp. 2664-2677, 1994. Article (CrossRef Link).

[32] B. Li, N. Wu, and H. Wang, "Expectation-maximisation-based localisation using anchors with uncertainties in wireless sensor networks," *IET Commun.*, vol. 8, pp. 1977-1987, 2014. Article (CrossRef Link).

[33] C. Constantinopoulos, M. K. Titsias and A. Likas, "Bayesian feature and model selection for Gaussian mixture models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 6, pp. 1013-1018, June. 2006. Article (CrossRef Link).

[34] T. M. Cover and J. A. Thomas, "Elements of Information Theory," *1st ed. New York: Wiley*, 2001. Article (CrossRef Link).

[35] S. O. Rice, "Mathematical analysis of random noise," *The Bell System Technical Journal*, vol. 23, no. 3, pp. 282-332, July. 1944. Article (CrossRef Link).

[36] I. S. Gradshteyn and I. M. Ryzhik, "Table of integrals, series, and products," *Academic Press, London*, 1980. Article (CrossRef Link).

[37] W. Yuan, N. Wu, B. Etzlinger, Y. Li, C. Yan and L. Hanzo, "Expectation Maximization-Based Passive Localization Relying on Asynchronous Receivers: Centralized versus Distributed Implementations," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 668-681, 2019 Article (CrossRef Link).

[38] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in Sensor Networks," *Lecture Notes in Control and Information Sciences*, vol. 331, pp. 169-182, 2006. Article (CrossRef Link).

[39] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc., Ser.B*, vol. 39, no. 1, pp. 1-38, 1977. Article (CrossRef Link).

Yuan Feng received his B.S. and Ph.D. degrees from Beijing Institute of Technology (BIT), Beijing, China in 2003 and 2011, respectively. He is currently an Associate Research fellow with the China Research and Development Academy of Machinery Equipment. His research interests include signal processing in wireless communication networks and system engineering.

Qinsiwei Yan received the B.S. degree from Beijing Institute of Technology (BIT), Beijing, China, in 2017. He is currently working toward his Ph.D degree at School of Information and Electronics, BIT. His current research interests include the graph signal processing and network topology theory.

Po-Hsuan Tseng received the B.S. and Ph.D. degrees in communication engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 2005 and 2011, respectively. Since February 2017, he has been an associate professor with the Department of Electronic Engineering, National Taipei University of Technology (NTUT), Taipei, Taiwan, where he was an assistant professor from August 2012 to January 2017. From January 2010 to October 2010, he was a visiting researcher with the University of California at Davis. His research interests are in the areas of signal processing for networking and communications, including location estimation and tracking, cooperative localization, and mobile broadband wireless access system design. He is an honorary member of the Phi Tau Phi Scholastic Honor Society of R.O.C.

Ganlin Hao received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China and M.S. degree from Beijing Institute of Technology (BIT), Beijing, China, in 2013 and 2017, respectively. His research interests include the indoor localization and distributed signal processing.

Nan Wu received his B.S., M.S. and Ph.D. degrees from Beijing Institute of Technology (BIT), Beijing, China in 2003, 2005 and 2011, respectively. He is currently an Associate Professor with School of Information and Electronics, BIT. From Nov. 2008 to Nov. 2009, he was a visiting Ph.D. student with Department of Electrical Engineering, Pennsylvania State University, USA. He is the recipient of National Excellent Doctoral Dissertation Award by MOE of China in 2013. His research interests include signal processing in wireless communication networks. He serves as an associate editor for *IEEE Access*, *International Journal of Electronics and Communications*, *KSII Transactions on Internet and Information Systems* and *IEICE Transactions on Communications*.