

# Secure Localization and Velocity Estimation in Mobile IoT Networks with Malicious Attacks

Yunfei Li, Shaodan Ma, Guanghua Yang and Kai-Kit Wong

## Abstract

Secure localization and velocity estimation are of great importance in Internet of Things (IoT) applications and are particularly challenging in the presence of malicious attacks. The problem becomes even more challenging in practical scenarios in which attack information is unknown and anchor node location uncertainties occur due to node mobility and falsification of malicious nodes. This challenging problem is investigated in this paper. With reasonable assumptions on the attack model and uncertainties, the secure localization and velocity estimation problem is formulated as an intractable maximum a posterior (MAP) problem. A variational-message-passing (VMP) based algorithm is proposed to approximate the true posterior distribution iteratively and find the closed-form estimates of the location and velocity securely. The identification of malicious nodes is also achieved in the meantime. The convergence of the proposed VMP-based algorithm is also discussed. Numerical simulations are finally conducted and the results show the VMP-based joint localization and velocity estimation algorithm can approach the Bayesian Cramer Rao bound and is superior to other secure algorithms.

## Index Terms

Variational message passing (VMP), secure localization, velocity estimation, maximum a posterior estimation.

Y. Li is with Institute of Physical Internet, Jinan University, Zhuhai Campus, Zhuhai 519070, China, and the Department of Electrical and Computer Engineering, University of Macau, Taipa, Macao, China (email: yunfeili1989@gmail.com).

S. Ma is with the State Key Laboratory of Internet of Things for Smart City and the Department of Electrical and Computer Engineering, University of Macau, Taipa, Macao, China (e-mail: shaodanma@um.edu.mo).

G. Yang is with Institute of Physical Internet, Jinan University, Zhuhai Campus, Zhuhai 519070, China (e-mail: ghyang@jnu.edu.cn).

K. K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (e-mail: kai-kit.wong@ucl.ac.uk).

## I. INTRODUCTION

With the explosive growth of Internet of Things (IoT) devices and emerging intelligent applications, e.g., autonomous driving vehicles and smart logistics, etc., localization and velocity estimation of mobile target nodes have become more and more important [1, 2]. Usually they are treated as individual problems and are achieved based on location-related and velocity-related measurements, respectively. Specifically, the location-related measurements mainly include time-of-arrival (TOA), time-difference-of-arrival (TDOA), angle-of-arrival (AOA) and received signal strength (RSS) [3–5], while the velocity-related measurements can be obtained from Doppler shifts [6, 7] or synthetic aperture radars (SARs) [8, 9]. Significant progress has been made and quite a lot algorithms have been proposed in the literature. For example, regarding to target localization, least squares [10], Bayesian inference [11], convex optimization [12] and iterative algorithms [13] have been proposed and a survey on localization techniques can be found in [14]. On the other hand, various velocity estimation algorithms based on Doppler shifts and/or SARs measurements have been reported in [6, 7] and [8, 9], respectively.

Considering that the location information is also contained in Doppler shift measurements, Doppler shift measurements can be further exploited to improve the localization accuracy. Joint localization and velocity estimation based on both range-based and Doppler shift measurements is thus feasible and has been discussed in [15–18]. Among them, [18] introduced nuisance parameters to transform the nonlinear measurement equation into a pseudo linear one, and then presented a two-stage weighted least squares (WLS) algorithm for joint localization and velocity estimation based on both TDOA and frequency-difference-of-arrival (FDOA) measurements. The two-stage WLS algorithm was later extended to multiple input multiple output (MIMO) radar systems in [16]. By using Taylor expansion for linearization, another WLS-based joint localization and velocity estimation algorithm was proposed in [17]. When considering large noise scenarios, a bias-reduced nonlinear WLS method was presented in [15].

In all the aforementioned works, the range measurements were assumed collected securely. In other words, they were obtained from trusty IoT nodes. However, with the increasing computing capability of the IoT devices and the presence of malicious users, the range measurements may be attacked and manipulated intentionally to mislead the localization process and/or prevent from localizing the target node accurately [19, 20]. The malicious attacks, e.g., Sybil attack, wormhole

attack and replay attack, are severe threats to reliable and accurate localization [21]. They may cause serious damages to property or loss of life especially in location-critical applications, e.g., autonomous driving vehicles. To avoid this and improve localization accuracy in the presence of malicious attacks, secure localization becomes necessary and has attracted considerable research attentions recently.

Some secure results have been reported in [20–27]. Generally they can be classified into two categories, i.e., secure location verification and secure localization. Regarding to secure location verification, it mainly aims to identify the malicious nodes or detect the location spoofing attacks. By using distance bounding protocol, two secure location verification algorithms were proposed respectively in [22, 23]. Leveraging on verifiable multilateration, a drone-based secure location verification scheme was also developed in [24]. Another secure algorithm was proposed to detect malicious nodes based on locations clustering and consistency evaluation via sequential probability ratio test in [26]. In these algorithms, one or multiple verifier nodes which are completely trustworthy are needed to achieve secure location verification. Without the need of verifier nodes, a location spoofing detection algorithm was proposed in [27]. It achieved the spoofing detection based on generalized likelihood ratio test, by using two-way TOA and audibility information, and modeling the malicious delay measurements as Gaussian noise with known mean and variance.

After the malicious nodes are identified, their associated measurements can be either discarded or further exploited for secure localization. For the latter case, a secure localization algorithm based on maximum likelihood criterion was developed in [21]. Specifically, by assuming perfect identification of malicious attacks and modeling the range attacks as Gaussian distribution with known mean and variance, the secure localization problem was formulated as a maximization problem of mixture likelihood function of non-malicious and malicious measurements. On the other hand, without identifying the malicious nodes in advance, secure localization of the target node was investigated in [20]. By adopting the Gaussian range attack model, an iterative gradient descent algorithm was developed to localize the target node securely.

All the secure localization algorithms assume the locations of both non-malicious and malicious anchor nodes are perfectly known. This is difficult to achieve in mobile IoT networks where mobile nodes serve as anchors. It is even impossible to achieve in the presence of malicious anchor nodes due to their malice nature and capability in falsifying their locations. Moreover,

Doppler shift measurements which also contain location information are not exploited to enhance the performance of secure localization. To address these issues, we investigate the joint problem of secure localization and velocity estimation in mobile IoT networks with malicious attacks in this paper. To be practical, uncertainties in anchor locations and velocities are considered. Moreover, both the identification of the malicious nodes and the statistical information of the range attacks are assumed unknown. By assuming reasonable prior distributions on the attack parameters and the uncertainties, we formulate the joint problem as a maximum a posterior (MAP) problem. Due to the presence of unknown malicious attacks and uncertainties, the MAP problem is intractable. To solve the problem, we propose a variational-message-passing based algorithm for joint secure localization and velocity estimation. It approximates the intractable posterior probability iteratively by the product of variational distributions, each being the approximation of posterior probability with respect to individual unknown parameter. In each iteration, the variational distributions are derived in tractable and closed forms, which enables the estimation of the target location and velocity accurately after convergence. The identification of malicious nodes is also achieved as a byproduct. The convergence of the proposed algorithm is then discussed and proved. The superior performance in localization and velocity estimation is finally demonstrated by simulations.

The remainder of this paper is organized as follows. Section II introduces the mobile IoT networks with unknown malicious attacks and uncertainties in locations and velocities. The variational-message-passing based secure localization and velocity estimation algorithm and the convergence analysis are presented in Section III. Section IV validates the proposed algorithm through simulations and comparison with other algorithms. Conclusions are finally drawn in Section V.

## II. SYSTEM MODEL

We consider a wireless sensor network consisting of  $\mathcal{S} = \{1, 2, \dots, M\}$  mobile sensors and one mobile target node. In the network, some sensors are malicious nodes that send erroneous information to mislead the target node. In practice, the set of malicious nodes is unknown and denoted as  $\mathcal{S}_{\mathcal{M}}$ . The rest of the sensors are legitimate nodes and are denoted as  $\mathcal{S}_{\mathcal{L}}$ , where  $\mathcal{S}_{\mathcal{L}} \cup \mathcal{S}_{\mathcal{M}} = \mathcal{S}$  and  $\mathcal{S}_{\mathcal{L}} \cap \mathcal{S}_{\mathcal{M}} = \emptyset$ . The location of the  $i$ th,  $i \in \mathcal{S}$  sensor node is denoted as  $\mathbf{x}_i = [x_i, y_i]^T$ , while the location of the target node is denoted as  $\mathbf{x} = [x, y]^T$ . Due to the

mobility and localization error, the sensor location is coarsely known as  $\bar{\mathbf{x}}_i$  and its true location can be modeled as

$$\mathbf{x}_i = \bar{\mathbf{x}}_i + \boldsymbol{\omega}_i^x, \quad (1)$$

where  $\boldsymbol{\omega}_i^x$  denotes the location uncertainty following Gaussian distribution with zero mean and covariance matrix  $\Sigma_i^x$ . Similarly, the velocity of the  $i$ th sensor node is also measured/estimated as  $\bar{\mathbf{v}}_i$  and its true velocity is written as

$$\mathbf{v}_i = \bar{\mathbf{v}}_i + \boldsymbol{\omega}_i^v, \quad (2)$$

where  $\boldsymbol{\omega}_i^v$  represents a Gaussian velocity uncertainty with zero mean and covariance matrix  $\Sigma_i^v$ .

In the sensor network, sensors are communicated over a channel with carrier frequency of  $f_c$ . Due to the mobility of nodes and the mismatches in the nodes' oscillators, Doppler frequency shifts usually exist between the target node and the sensors [28]. The measurement of the Doppler shift between the  $i$ th sensor and the target node can be collected and written as

$$h_i = f_i + f + \psi_i + \eta_i, \quad (3)$$

where  $f$  and  $f_i$  denote the carrier frequency offsets at the target node and the  $i$ th sensor, respectively,  $\eta_i$  is the Gaussian measurement noise with zero-mean and variance  $\iota_i$ , and  $\psi_i$  is the Doppler shift caused by the mobility of the target node and the  $i$ th sensor and is given by [29–31]

$$\psi_i = \frac{(\mathbf{x}_i - \mathbf{x})^T (\mathbf{v}_i - \mathbf{v})}{\|\mathbf{x}_i - \mathbf{x}\| c} f_c, \quad (4)$$

in which  $c$  is the speed of light and  $\mathbf{v}$  denotes the velocity of the target node to be estimated.

Collecting the Doppler measurements associated with all the sensors together as a vector  $\mathbf{h} = [h_1 \dots, h_M]$ , the likelihood probability function can be written as

$$p(\mathbf{h} | \mathbf{f}_s, f, \mathbf{x}_s, \mathbf{x}, \mathbf{v}_s, \mathbf{v}) = \prod_{i=1}^M p(h_i | f_i, f, \mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v}), \quad (5)$$

where  $\mathbf{x}_s = [\mathbf{x}_1^T, \dots, \mathbf{x}_M^T]$  is the vector of all sensors locations,  $\mathbf{f}_s = [f_1, \dots, f_M]$  is the vector of the carrier frequency offsets for all sensors, and  $\mathbf{v}_s = [\mathbf{v}_1^T, \dots, \mathbf{v}_M^T]$  denotes the velocity vector for all the sensors.

On the other hand, the target node can also obtain the range measurements corresponding to the sensors. Usually the range measurements are obtained based on time of flight measurements. Due to the presence of the malicious nodes, the time of flight measurements may be intentionally

manipulated by introducing an extra delay, thus causing additional interfering term in the range measurement [21]. Considering the possible attacks from malicious nodes, the range measurement from the  $i$ th sensor can be modeled as

$$r_i = \underbrace{\|\mathbf{x}_i - \mathbf{x}\|}_{d_i} + \Xi_i + \varepsilon_i. \quad (6)$$

where  $\varepsilon_i$  is the measurement error following a zero-mean Gaussian distribution with variance of  $\Lambda_{\mathcal{L}}^{-1}$ , and  $\Xi_i$  characterizes the interfering term caused by the possible attack from the malicious node. When the sensor node is legitimate one,  $\Xi_i$  equals to zero. Following the attack models in [21, 27], the interfering term  $\Xi_i$  from the malicious node can be modeled as mean-shifted Gaussian interference with non-zero mean  $\mu_{\mathcal{M}}$  and variance  $\Lambda_{\mathcal{M}}^{-1}$ .

In practice, the number of malicious nodes and the identification of the malicious nodes are unknown. Without loss of generality, we assume that the sensor node is a legitimate node with a probability of  $\alpha_1$ , while it is a malicious node with a probability of  $\alpha_2$ . The probabilities are unknown with  $\sum_{l=1}^2 \alpha_l = 1$ . By combining the measurement error with the interfering term together as  $\tilde{\varepsilon}_i = \varepsilon_i + \Xi_i$ , we have

$$\tilde{\varepsilon}_i = \begin{cases} \mathcal{N}(0, \Lambda_{\mathcal{L}}^{-1}) & i \in \mathcal{S}_{\mathcal{L}} \text{ with probability of } \alpha_1, \\ \mathcal{N}(\mu, \Lambda^{-1}) & i \in \mathcal{S}_{\mathcal{M}} \text{ with probability of } \alpha_2. \end{cases} \quad (7)$$

with  $\mu = \mu_{\mathcal{M}}$ , and  $\Lambda^{-1} = \Lambda_{\mathcal{M}}^{-1} + \Lambda_{\mathcal{L}}^{-1}$ .

In order to differentiate the attack and secure measurements and ease the system modeling, an auxiliary indicator vector with binary elements  $\mathbf{y}_i = (y_{i,1}, y_{i,2})$  is introduced. Specifically, the indicator vector is defined as

$$\mathbf{y}_i = \begin{cases} (1, 0) & i \in \mathcal{S}_{\mathcal{L}}, \\ (0, 1) & i \in \mathcal{S}_{\mathcal{M}}, \end{cases}. \quad (8)$$

Given the probability of  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ , the indicator vector  $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M\}$  then follows the distribution as

$$p(\mathbf{y}|\boldsymbol{\alpha}) = \prod_{i=1}^M \prod_{l=1}^2 [\alpha_l]^{y_{i,l}}. \quad (9)$$

Considering the strong concealing capability of malicious nodes, the statistical information regarding to malicious nodes, i.e.,  $\boldsymbol{\alpha}$ ,  $\mu$  and  $\Lambda^{-1}$ , are all assumed unknown. But they can be modeled as random variables following certain distributions as reported in the literature [32–34].

Specifically, considering the practical constraint (i.e.,  $\sum_{l=1}^2 \alpha_l = 1$ ), the prior distribution of the probability  $\alpha$  can be modeled as a Dirichlet distribution given by [32]

$$p(\alpha|\bar{\lambda}) = \prod_{l=1}^2 \text{Dir}(2, \bar{\lambda}_l), \quad (10)$$

where  $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2)$  is the parameter of Dirichlet distribution. In regards to the covariance inverse  $\Lambda$ , it is assumed following Gamma distribution by following the conjugate prior principle [33] and the prior probability is given by

$$p(\Lambda|\bar{a}, \bar{b}) = \Gamma(\Lambda|\bar{a}, \bar{b}), \quad (11)$$

where  $\bar{a}$  and  $\bar{b}$  are the parameters of Gamma distribution. Finally, the prior distribution of the Gaussian mean  $\mu$  can also be assumed as a Gaussian distribution given by [34]

$$p(\mu|\bar{m}, (\bar{\gamma})^{-1}) = \mathcal{N}(\mu|\bar{m}, (\bar{\gamma})^{-1}), \quad (12)$$

with the mean and covariance of  $\bar{m}$  and  $(\bar{\gamma})^{-1}$ , respectively.

Now based on the range and error models (6) and (7), the likelihood function of the range measurements can be formulated as

$$p(\mathbf{r}|\mathbf{x}_s, \mathbf{x}, \boldsymbol{\theta}) = \prod_{i=1}^M p(r_i|\mathbf{x}_i, \mathbf{x}, \boldsymbol{\theta}) = \prod_{i=1}^M \mathcal{N}(r_i - d_i, \Lambda_{\mathcal{L}}^{-1})^{y_{i,1}} \mathcal{N}(r_i - d_i - \mu, \Lambda^{-1})^{y_{i,2}}, \quad (13)$$

where  $\boldsymbol{\theta} = [\mu, \Lambda, \alpha, \mathbf{y}]$  is the unknown parameter vector and  $\mathbf{r} = [r_1 \dots, r_M]$  is the range measurement vector of the target node.

### III. JOINT SECURE LOCALIZATION AND VELOCITY ESTIMATION

#### A. Probabilistic Model and Variational Message Passing

By collecting range and Doppler measurements  $\mathbf{r}$  and  $\mathbf{h}$  from all the sensor nodes and based on the maximum a posterior (MAP) criterion, the target location and velocity can be theoretically estimated as

$$[\mathbf{x}^{\text{MAP}}, \mathbf{v}^{\text{MAP}}] = \operatorname{argmax}_{\Phi_{\setminus \mathbf{x}, \mathbf{v}}} \int p(\Phi|\mathbf{r}, \mathbf{h}) d\Phi_{\setminus \mathbf{x}, \mathbf{v}}, \quad (14)$$

where  $\Phi = [\mathbf{f}_s, f, \mathbf{x}_s, \mathbf{x}, \mathbf{v}_s, \mathbf{v}, \boldsymbol{\theta}]$ ,  $\Phi_{\setminus \mathbf{x}, \mathbf{v}}$  denotes the set of variables by removing  $\mathbf{x}$  and  $\mathbf{v}$  from  $\Phi$ , and the posterior probability is given as

$$p(\Phi|\mathbf{r}, \mathbf{h}) \propto p(\mathbf{r}, \mathbf{h}|\Phi) p(\Phi) \quad (15)$$

with  $p(\Phi)$  being the prior probability and  $p(\mathbf{r}, \mathbf{h}|\Phi)$  being the joint likelihood function of the target node as

$$p(\mathbf{r}, \mathbf{h}|\Phi) = \prod_{i=1}^M p(r_i|\mathbf{x}_i, \mathbf{x}, \boldsymbol{\theta}) p(h_i|f_i, f, \mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v}). \quad (16)$$

Unfortunately, with the presence of malicious nodes, the uncertainties in the sensor locations and velocities, the unknown identification of the malicious nodes and the unknown statistics of the range errors, the posterior distribution in (15) is very complicated and the MAP estimation in (14) is intractable.

In this paper, we propose to find a tractable approximation  $q(\Phi)$  to the posterior distribution  $p(\Phi|\mathbf{r}, \mathbf{h})$  in (15) based on variational Bayesian framework. Specifically, considering the independence of the variables in the set  $\Phi$ , the variational distribution  $q(\Phi)$  can be factorized as

$$q(\Phi) = \prod_{\Phi_k \in \Phi} q(\Phi_k), \quad (17)$$

where  $\Phi_k$  denotes the  $k$ th variable in the set  $\Phi$ . To find a good approximation and inspired by the non-negative property of Kullback-Leibler (KL) divergence of any two distributions [34], the approximation is proposed by minimizing the KL divergence between the two distributions  $q(\Phi)$  and  $p(\Phi|\mathbf{r}, \mathbf{h})$  as

$$\text{KL}(q(\Phi) || p(\Phi|\mathbf{r}, \mathbf{h})) = -\mathbb{E}_{q(\Phi)} \left\{ \ln \frac{p(\Phi|\mathbf{r}, \mathbf{h})}{q(\Phi)} \right\} \geq 0, \quad (18)$$

where  $\mathbb{E}_{q(\Phi)}$  means the expectation with respect to  $q(\Phi)$  and the equality holds when  $q(\Phi) = p(\Phi|\mathbf{r}, \mathbf{h})$  as proved in [35]. Based on the factorization (17) and the KL divergence minimization (18), the approximated distribution  $q(\Phi_k)$  in fact can be regarded as the approximation of the corresponding posterior distribution  $p(\Phi_k|\mathbf{r}, \mathbf{h})$ . For example,  $q(\mathbf{v})$  is the approximation to the posterior distribution  $p(\mathbf{v}|\mathbf{r}, \mathbf{h})$ . Then the MAP estimation of each parameter  $\Phi_k$  can be achieved as

$$\Phi_k^{\text{MAP}} = \text{argmax} q(\Phi_k). \quad (19)$$

With the factorization (17) and by adopting alternative optimization method, the variational distribution can be iteratively approximated as [32]

$$q^{(\xi)}(\Phi_k) \propto \exp \left\{ \mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} [\ln p(\Phi, \mathbf{r}, \mathbf{h})] \right\}, \quad (20)$$



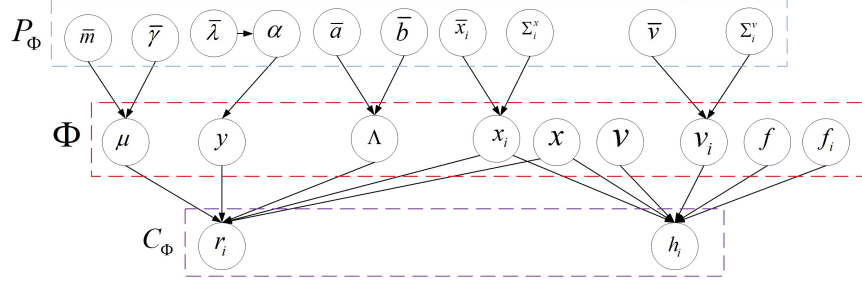


Fig. 1: The probabilistic graph of the joint probability

where  $q^{(\xi)}(\Phi_k)$  is the approximation in the  $\xi$ th iteration and  $p(\Phi, \mathbf{r}, \mathbf{h})$  is the joint probability. Considering the complexity in the joint probability  $p(\Phi, \mathbf{r}, \mathbf{h})$ , we further adopt a probabilistic graph model to analyze it. As shown in Fig. 1, the variables are denoted as nodes in the graph and the directed edges signify the dependence of the variables in terms of conditional probabilities. Moreover, according to the dependencies of the variables, the nodes excluding those in the set  $\Phi$  can be further classified as the parent nodes  $\mathcal{P}_{\Phi_k}$  and child nodes  $\mathcal{C}_{\Phi_k}$  with respect to  $\Phi_k$ . Based on the probabilistic graph model, the joint probability can be written as the product of the marginal probabilities of the nodes and the conditional probabilities among the nodes and their parent/child nodes, i.e.,  $p(\Phi_k | \mathcal{P}_{\Phi_k})$  and  $p(\mathcal{C}_{\Phi_k} | \Phi_k, \Phi_m)$  where  $\Phi_m$  is the other parent node of  $\mathcal{C}_{\Phi_k}$  if any with  $m \neq k$ . Taking the mean of the attack delay  $\mu$  as an example (i.e.,  $\Phi_k = \mu$ ), it contributes to the joint probability through the prior distribution  $p(\mu | \bar{m}, \bar{\gamma})$  and the likelihood distribution  $p(\mathbf{r}, \mathbf{h} | \mu, \Phi_{\setminus \mu})$  where  $\bar{m}$  and  $\bar{\gamma}$  are the parent nodes while  $\mathbf{r}$  and  $\mathbf{h}$  are the child nodes of  $\mu$ , respectively.

We notice that all the Dirichlet, Gamma and Gaussian distributions involved in our system model can be rewritten in the form of exponential functions. This fact inspires us to reformulate the two conditional distributions as [33]

$$p(\Phi_k | \mathcal{P}_{\Phi_k}) = \exp[\mathcal{H}_{\mathcal{P}}(\mathcal{P}_{\Phi_k}) \Omega(\Phi_k) + \mathbf{Z}(\mathcal{P}_{\Phi_k}) + f_{\Phi_k}(\Phi_k)], \quad (21)$$

$$p(\mathcal{C}_{\Phi_k} | \Phi_k, \Phi_m) = \exp[\mathcal{H}_{\mathcal{C}}(\Phi_m, \mathcal{C}_{\Phi_k}) \Omega(\Phi_k) + \Upsilon(\Phi_m, \mathcal{C}_{\Phi_k}) + f_{\mathcal{C}_{\Phi_k}}(\mathcal{C}_{\Phi_k})], \quad (22)$$

where  $\Upsilon(\cdot)$  and  $\mathbf{Z}(\cdot)$  can be regarded as the normalization functions associated with the corresponding variables in the conditions, and  $f_{\mathcal{C}_{\Phi_k}}(\mathcal{C}_{\Phi_k})$  and  $f_{\Phi_k}(\Phi_k)$  are the functions only associated with the corresponding variables. The expressions of the functions in (21) and (22) change with its associated variables and will be specifically given in the next subsection.

By plugging (21) and (22) into (20) and after some tedious manipulations, it yields

$$q^{(\xi+1)}(\Phi_k) \propto \exp \left\{ \left( \underbrace{\mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} [\mathcal{H}_{\mathcal{P}}(\mathcal{P}_{\Phi_k})] + \sum_{\mathcal{C}_{\Phi_k}} \mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} [\mathcal{H}_{\mathcal{C}}(\Phi_m, \mathcal{C}_{\Phi_k})]}_{\mathcal{H}_{\Phi_k}^{(\xi+1)}} \right) \Omega(\Phi_k) \right\} \exp(f_{\Phi_k}(\Phi_k)) \exp \left( \underbrace{\mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} \left( \mathbf{Z}(\mathcal{P}_{\Phi_k}) + \Upsilon(\Phi_m, \mathcal{C}_{\Phi_k}) + f_{\mathcal{C}_{\Phi_k}}(\mathcal{C}_{\Phi_k}) \right)}_{\text{constant term}} \right), \quad (23)$$

where the last exponential term is irrelevant to the variable  $\Phi_k$  and can be regarded as constant and ignored in the MAP estimation of  $\Phi_k$ .

Clearly from (23), in the update of the variational distribution of  $q^{(\xi+1)}(\Phi_k)$  in each iteration, only the term  $\mathcal{H}_{\Phi_k}^{(\xi+1)}$  needs to be updated. It can be termed as a natural parameter and further partitioned into two sub-terms as follows

$$\mathcal{H}_{\Phi_k}^{(\xi+1)} = \underbrace{\mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} [\mathcal{H}_{\mathcal{P}}(\mathcal{P}_{\Phi_k})]}_{\mathcal{M}_{\mathcal{P}_{\Phi_k} \rightarrow \Phi_k}^{(\xi)}} + \sum_{\mathcal{C}_{\Phi_k}} \underbrace{\mathbb{E}_{\prod_{j \neq k} q^{(\xi)}(\Phi_j)} [\mathcal{H}_{\mathcal{C}}(\Phi_m, \mathcal{C}_{\Phi_k})]}_{\mathcal{M}_{\mathcal{C}_{\Phi_k} \rightarrow \Phi_k}^{(\xi)}}, \quad (24)$$

The first subterm  $\mathcal{M}_{\mathcal{P}_{\Phi_k} \rightarrow \Phi_k}^{(\xi)}$  can be regarded as the message from the parent node  $\mathcal{P}_{\Phi_k}$ , while the second subterm is the sum of messages  $\mathcal{M}_{\mathcal{C}_{\Phi_k} \rightarrow \Phi_k}^{(\xi)}$  from the child nodes  $\mathcal{C}_{\Phi_k}$ .

With (24), we now can propose a variational message passing (VMP) algorithm to update the natural parameter  $\mathcal{H}_{\Phi_k}^{(\xi+1)}$ . Specifically, the natural parameter can be updated in an iterative message passing way. The messages are the variational expectations of natural parameters of neighboring nodes. Through the variational message passing algorithm, the natural parameters of the complicated posterior distributions can be approximated iteratively. In the considered network, the natural parameters of variational distributions  $q(\mathbf{x})$  and  $q(\mathbf{v})$  can be approximated iteratively by using the information from its neighboring nodes. Then the secure localization and velocity estimation can be finally achieved based on (19). Since the considered localization network is very complicated with unknown malicious nodes and uncertainties in the sensor locations and velocities, the derivations of the natural parameters  $\mathcal{H}_{\Phi_k}^{(\xi+1)}$  and the variational distribution  $q(\Phi_k)$  are non-trivial and still challenging. The details of the derivations are then given in the next subsection.

### B. VMP-Based Joint Localization and Velocity Estimation

Based on (24), the messages between unknown parameters are updated and passing through the parent nodes and child nodes of the variable nodes, and then enable the MAP estimation of all the parameters. In the following, the message update and the estimation of each parameter are presented one by one in details.

1) **Estimation of  $\mathbf{y}_i$ :** According to the probabilistic graph in Fig. 1, the parent and child nodes of the variable node  $\mathbf{y}_i$  are the nodes  $\boldsymbol{\alpha}$  and  $r_i$  respectively. The messages to the variable node  $\mathbf{y}_i$  include the message  $\mathcal{M}_{\boldsymbol{\alpha} \rightarrow \mathbf{y}_i}^{(\xi)}$  from the parent node  $\boldsymbol{\alpha}$  and the message  $\mathcal{M}_{r_i \rightarrow \mathbf{y}_i}^{(\xi)}$  from the child node  $r_i$ .

**Message  $\mathcal{M}_{\boldsymbol{\alpha} \rightarrow \mathbf{y}_i}^{(\xi)}$ :** The prior distribution  $p(\mathbf{y}_i | \boldsymbol{\alpha})$  in (9) can be rewritten as

$$p(\mathbf{y}_i | \boldsymbol{\alpha}) = \exp \left( \begin{bmatrix} \ln \alpha_1 \\ \ln \alpha_2 \end{bmatrix}^T \begin{bmatrix} y_{i,1} \\ y_{i,2} \end{bmatrix} \right). \quad (25)$$

By matching (25) with the form in (21), it follows that  $\mathcal{H}_{\mathcal{P}}(\mathbf{y}_i) = \begin{bmatrix} \ln \alpha_1 \\ \ln \alpha_2 \end{bmatrix}^T$  and  $\boldsymbol{\Omega}(\mathbf{y}_i) = \begin{bmatrix} y_{i,1} \\ y_{i,2} \end{bmatrix}$ .

Substituting (10) and (25) into (24), the message from the parent node  $\boldsymbol{\alpha}$  can be derived as

$$\mathcal{M}_{\boldsymbol{\alpha} \rightarrow \mathbf{y}_i}^{(\xi)} = \mathbb{E}_{q^{(\xi)}(\boldsymbol{\Phi}_{\setminus \mathbf{y}_i})} [\mathcal{H}_{\mathcal{P}}(\mathbf{y}_i)] = \begin{bmatrix} \varsigma(\lambda_1^{(\xi)}) - \varsigma(\sum_{l=1}^2 \lambda_l^{(\xi)}) \\ \varsigma(\lambda_2^{(\xi)}) - \varsigma(\sum_{l=1}^2 \lambda_l^{(\xi)}) \end{bmatrix}^T, \quad (26)$$

where  $\varsigma(\cdot)$  is the digamma function and  $\lambda_l^{(\xi)}$  is the parameter of the variational distribution  $q^{(\xi)}(\boldsymbol{\alpha})$ . By following the conjugate prior principle and with the prior distribution of  $\boldsymbol{\alpha}$  in (10), we have  $q^{(\xi)}(\boldsymbol{\alpha}) = \text{Dir}(\boldsymbol{\alpha} | \lambda_1^{(\xi)}, \lambda_2^{(\xi)})$ . The parameters of  $\lambda_l^{(\xi)}$  will be derived later in Section III-B3.

**Message  $\mathcal{M}_{r_i \rightarrow \mathbf{y}_i}^{(\xi)}$ :** According to (22), the likelihood function (13) can be factorized into

$$p(r_i | \mathbf{x}_i, \mathbf{x}, \boldsymbol{\theta}) = \exp \left( \begin{bmatrix} \ln \mathcal{N}(r_i - d_i, \Lambda_{\mathcal{L}}^{-1}) \\ \ln \mathcal{N}(r_i - d_i - \mu, \Lambda^{-1}) \end{bmatrix}^T \begin{bmatrix} y_{i,1} \\ y_{i,2} \end{bmatrix} \right), \quad (27)$$

$$\text{and } \mathcal{H}_{\mathcal{C}}(r_i) = \begin{bmatrix} \ln \mathcal{N}(r_i - d_i, \Lambda_{\mathcal{L}}^{-1}) \\ \ln \mathcal{N}(r_i - d_i - \mu, \Lambda^{-1}) \end{bmatrix}^T.$$

Plugging (27) into (24) and with the derivations in Appendix A, the message  $\mathcal{M}_{r_i \rightarrow \mathbf{y}_i}^{(\xi)}$  from  $r_i$  to  $\mathbf{y}_i$  can be obtained as

$$\mathcal{M}_{r_i \rightarrow \mathbf{y}_i}^{(\xi)} = \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})} [\mathcal{H}_{\mathcal{C}}(r_i)] = \begin{bmatrix} -\frac{\Lambda_{\mathcal{C}}}{2} \omega_{i,1}^{(\xi)} - \frac{1}{2} \ln 2\pi + \frac{1}{2} \ln \Lambda_{\mathcal{C}} \\ -\frac{\beta^{(\xi)}}{2} \omega_{i,2}^{(\xi)} - \frac{1}{2} \ln 2\pi + \frac{1}{2} \varphi(a^{(\xi)}) + \frac{1}{2} \ln(b^{(\xi)}) \end{bmatrix}^T \triangleq \begin{bmatrix} \vartheta_{i,1}^{(\xi)} \\ \vartheta_{i,2}^{(\xi)} \end{bmatrix}^T, \quad (28)$$

where  $\varphi(\cdot)$  is a Gamma function,  $\beta^{(\xi)} = a^{(\xi)}/b^{(\xi)}$ ,  $a^{(\xi)}$  and  $b^{(\xi)}$  are the parameters of the variational distribution  $q^{(\xi)}(\Lambda)$ , and  $\omega_{i,l}^{(\xi)}$  is defined in Appendix A. The parameters of  $a^{(\xi)}$  and  $b^{(\xi)}$  will be derived later in the subsection corresponding to the estimation of  $\Lambda$ .

**Natural Parameter  $\mathcal{H}_{\mathbf{y}_i}^{(\xi+1)}$ :** By substituting (26) and (28) into (24), it yields

$$\mathcal{H}_{\mathbf{y}_i}^{(\xi+1)} = \begin{bmatrix} \varsigma(\bar{\lambda}_1) - \varsigma\left(\sum_{l=1}^2 \bar{\lambda}_l\right) + \vartheta_{i,1}^{(\xi)} \\ \varsigma(\bar{\lambda}_2) - \varsigma\left(\sum_{l=1}^2 \bar{\lambda}_l\right) + \vartheta_{i,2}^{(\xi)} \end{bmatrix}^T \triangleq \begin{bmatrix} \ln \phi_{i,1}^{(\xi+1)} \\ \ln \phi_{i,2}^{(\xi+1)} \end{bmatrix}^T. \quad (29)$$

Plugging (29) and  $\Omega(\mathbf{y}_i)$  into (23), the variational distribution  $q^{(\xi+1)}(\mathbf{y}_i)$  is given by

$$q^{(\xi+1)}(\mathbf{y}_i) \propto \prod_{l=1}^2 e^{y_{i,l} \ln \phi_{i,l}^{(\xi+1)}}. \quad (30)$$

By normalizing  $\phi_{i,l}^{(\xi+1)}$ , we obtain

$$\kappa_{i,l}^{(\xi+1)} \triangleq \mathbb{E}_{q^{(\xi+1)}(\mathbf{y}_i)}(y_{i,l}) = \frac{\phi_{i,l}^{(\xi+1)}}{\sum_{l=1}^2 \phi_{i,l}^{(\xi+1)}}, \quad (31)$$

and (30) can be reformulated as

$$q^{(\xi+1)}(\mathbf{y}_i) = \prod_{l=1}^2 e^{y_{i,l} \ln \kappa_{i,l}^{(\xi+1)}}, \quad (32)$$

Considering  $\mathbf{y}_i$  is a vector with only binary elements and putting (32) into (19) for MAP estimation, the MAP estimate of  $\mathbf{y}_i$  in the  $(\xi + 1)$ th iteration can be obtained as

$$\mathbf{y}_i^{(\xi+1)} = \begin{cases} (1, 0) & \epsilon_{i,1}^{(\xi+1)} \geq 1 \\ (0, 1) & \epsilon_{i,1}^{(\xi+1)} < 1 \end{cases} \quad (33)$$

where  $\epsilon_{i,1}^{(\xi+1)}$  is the posterior ratio given as

$$\epsilon_{i,1}^{(\xi+1)} \triangleq \frac{q^{(\xi+1)}(y_{i,1} = 1)}{q^{(\xi+1)}(y_{i,1} = 0)} = e^{\ln \phi_{i,1}^{(\xi+1)}}. \quad (34)$$

Notice that the estimation of  $\mathbf{y}_i$  in fact can be regarded as the identification process of the malicious nodes.

2) **Estimation of  $\mu$** : Clearly from Fig. 1, the messages to the variable node  $\mu$  include the message  $\mathcal{M}_{r_i \rightarrow \mu}^{(\xi)}$  from the child node  $r_i$  and the messages  $\mathcal{M}_{\mathcal{P}_\mu \rightarrow \mu}^{(\xi)}$  from parent nodes  $\bar{m}$  and  $\bar{\gamma}$ .

**Message  $\mathcal{M}_{\mathcal{P}_\mu \rightarrow \mu}^{(\xi)}$** : By plugging (12) into (24), the message  $\mathcal{M}_{\mathcal{P}_\mu \rightarrow \mu}^{(\xi)}$  can be directly derived as

$$\mathcal{M}_{\mathcal{P}_\mu \rightarrow \mu}^{(\xi)} = \begin{bmatrix} \bar{\gamma} \bar{m} \\ -\frac{\bar{\gamma}}{2} \end{bmatrix}^T, \quad (35)$$

and  $\Omega(\mu) = \begin{bmatrix} \mu^2 \\ \mu \end{bmatrix}$ .

**Message  $\mathcal{M}_{r_i \rightarrow \mu}^{(\xi)}$** : Following similar derivations as (27) and (28), the message  $\mathcal{M}_{r_i \rightarrow \mu}^{(\xi)}$  can be given by

$$\mathcal{M}_{r_i \rightarrow \mu}^{(\xi)} = \begin{bmatrix} \beta^{(\xi)} (r_i - d_i^{(\xi)}) \kappa_{i,2}^{(\xi)} \\ -\frac{\beta^{(\xi)} \kappa_{i,2}^{(\xi)}}{2} \end{bmatrix}^T, \quad (36)$$

where  $d_i^{(\xi)} = \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i) = \|\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}\|$  as shown in Appendix A.

**Natural Parameter  $\mathcal{H}_\mu^{(\xi+1)}$** : By substituting (35) and (36) into (24), it yields

$$\mathcal{H}_\mu^{(\xi+1)} = \begin{bmatrix} \bar{\gamma} \bar{m} + \sum_{i=1}^M \beta^{(\xi)} (r_i - d_i^{(\xi)}) \kappa_{i,2}^{(\xi)} \\ -\frac{\bar{\gamma}}{2} - \sum_{i=1}^M \frac{\beta^{(\xi)} \kappa_{i,2}^{(\xi)}}{2} \end{bmatrix}^T. \quad (37)$$

By using the conjugate prior principle, the variational distribution  $q^{(\xi+1)}(\mu)$  is also a Gaussian distribution with mean and variance denoted as  $m^{(\xi+1)}$  and  $[\gamma^{(\xi+1)}]^{-1}$  respectively. Hence the natural parameter should follow

$$\mathcal{H}_\mu^{(\xi+1)} = \begin{bmatrix} m^{(\xi+1)} \gamma^{(\xi+1)} \\ -\frac{\gamma^{(\xi+1)}}{2} \end{bmatrix}^T. \quad (38)$$

Mapping (37) with (38) yields

$$\gamma^{(\xi+1)} = \bar{\gamma} + \beta^{(\xi)} \nu_2^{(\xi)}, \quad (39)$$

$$m^{(\xi+1)} = \frac{1}{\gamma^{(\xi+1)}} \left[ \bar{\gamma} \bar{m} + \sum_{i=1}^M \beta^{(\xi)} (r_i - d_i^{(\xi)}) \kappa_{i,2}^{(\xi)} \right] = \frac{1}{\gamma^{(\xi+1)}} \left[ \bar{\gamma} \bar{m} + \nu_2^{(\xi)} e_2^{(\xi)} \right], \quad (40)$$

where  $e_2^{(\xi)} = \frac{1}{\nu_2^{(\xi)}} \sum_{i=1}^M \beta^{(\xi)} (r_i - d_i^{(\xi)}) \kappa_{i,2}^{(\xi)}$  and  $\nu_2^{(\xi)} = \sum_{i=1}^M \kappa_{i,2}^{(\xi)}$ .

Hence, the MAP estimation of  $\mu$  in the  $(\xi + 1)$ th iteration can be given by

$$\mu^{(\xi+1)} = m^{(\xi+1)}. \quad (41)$$

3) **Estimation of  $\alpha$** : The messages to the variable node  $\alpha$  include the message  $\mathcal{M}_{\lambda \rightarrow \alpha}^{(\xi)}$  from its parent node  $\lambda$  and the message  $\mathcal{M}_{y \rightarrow \alpha}^{(\xi)}$  from its child node  $y$ .

**Message  $\mathcal{M}_{\lambda \rightarrow \alpha}^{(\xi)}$** : Plugging (10) into (24), the message  $\mathcal{M}_{\lambda \rightarrow \alpha}^{(\xi)}$  is obtained as

$$\mathcal{M}_{\lambda \rightarrow \alpha}^{(\xi)} = \begin{bmatrix} \bar{\lambda}_1 - 1 \\ \bar{\lambda}_2 - 1 \end{bmatrix}^T. \quad (42)$$

and  $\Omega(\alpha) = \begin{bmatrix} \ln \alpha_1 \\ \ln \alpha_2 \end{bmatrix}$ .

**Message  $\mathcal{M}_{y \rightarrow \alpha}^{(\xi)}$** : Using (31) and (24), the message  $\mathcal{M}_{y \rightarrow \alpha}^{(\xi)}$  is given by

$$\mathcal{M}_{y \rightarrow \alpha}^{(\xi)} = \sum_{i=1}^M \begin{bmatrix} \kappa_{i,1}^{(\xi)} \\ \kappa_{i,2}^{(\xi)} \end{bmatrix}^T \triangleq \begin{bmatrix} \nu_1^{(\xi)} \\ \nu_2^{(\xi)} \end{bmatrix}^T. \quad (43)$$

**Natural Parameter  $\mathcal{H}_\alpha^{(\xi+1)}$** : By plugging (42) and (43) into (24), we obtain the natural parameter as

$$\mathcal{H}_\alpha^{(\xi+1)} = \begin{bmatrix} \bar{\lambda}_1 + \nu_1^{(\xi)} - 1 \\ \bar{\lambda}_2 + \nu_2^{(\xi)} - 1 \end{bmatrix}^T \triangleq \begin{bmatrix} \lambda_1^{(\xi+1)} - 1 \\ \lambda_2^{(\xi+1)} - 1 \end{bmatrix}^T, \quad (44)$$

Following the conjugate prior principle, the variational distribution  $q^{(\xi+1)}(\alpha)$  is also a Dirichlet distribution denoted as  $Dir(\alpha | \lambda^{(\xi+1)})$  with  $\lambda^{(\xi+1)} = [\lambda_1^{(\xi+1)}, \lambda_2^{(\xi+1)}]$ . Hence, the MAP estimation of  $\alpha_l$  can be given by

$$\alpha_l^{(\xi+1)} = \frac{\lambda_l^{(\xi+1)}}{\sum_{l=1}^2 \lambda_l^{(\xi+1)}}. \quad (45)$$

4) **Estimation of  $\Lambda$** : The messages to  $\Lambda$  include the message  $\mathcal{M}_{\mathcal{P}_\Lambda \rightarrow \Lambda}^{(\xi)}$  from its parent nodes  $\bar{a}$  and  $\bar{b}$  and the message  $\mathcal{M}_{r_i \rightarrow \Lambda}^{(\xi)}$  from its child node. Following similar derivations, the messages can be given as

$$\mathcal{M}_{\mathcal{P}_\Lambda \rightarrow \Lambda}^{(\xi)} = \begin{bmatrix} -\frac{1}{\bar{b}} \\ \bar{a} - 1 \end{bmatrix}^T, \quad (46)$$

$$\mathcal{M}_{r_i \rightarrow \Lambda}^{(\xi)} = \begin{bmatrix} -\frac{\kappa_{i,2}^{(\xi)}}{2} \omega_{i,2}^{(\xi)} \\ \frac{\kappa_{i,2}^{(\xi)}}{2} \end{bmatrix}^T. \quad (47)$$

and  $\Omega(\Lambda) = \begin{bmatrix} \Lambda \\ \ln \Lambda \end{bmatrix}$ .

Substituting (46) and (47) into (24), the natural parameter follows

$$\mathcal{H}_\Lambda^{(\xi+1)} = \begin{bmatrix} -\frac{1}{b} - \frac{\zeta_2^{(\xi)}}{2} \\ \bar{a} + \frac{\nu_2^{(\xi)}}{2} - 1 \end{bmatrix}^T \quad (48)$$

where  $\zeta_2^{(\xi)} = \sum_{i=1}^M \kappa_{i,2}^{(\xi)} \omega_{i,2}^{(\xi)}$ .

Following the conjugate prior principle, the variational distribution  $q^{(\xi+1)}(\Lambda)$  is also a Gamma distribution denoted as  $\Gamma(\Lambda|a^{(\xi+1)}, b^{(\xi+1)})$ . The natural parameter thus should follow

$$\mathcal{H}_\Lambda^{(\xi+1)} = \begin{bmatrix} -\frac{1}{b^{(\xi+1)}} \\ a^{(\xi+1)} - 1 \end{bmatrix}^T, \quad (49)$$

With (48) and (49), we have

$$b^{(\xi+1)} = \left( \frac{1}{\bar{b}} + \frac{\zeta_2^{(\xi)}}{2} \right)^{-1}, \quad (50)$$

$$a^{(\xi+1)} = \bar{a} + \frac{\nu_2^{(\xi)}}{2}. \quad (51)$$

Then the MAP estimation is the posterior mean of  $\Lambda$  given by

$$\Lambda^{(\xi+1)} = \beta^{(\xi+1)} = a^{(\xi+1)}/b^{(\xi+1)}. \quad (52)$$

5) **Estimation of  $f$** : As shown in Fig. 1, the variable node  $f$  does not have a parent node. To enable the message passing algorithm, we assume a non-informative Gaussian distribution  $\mathcal{N}(0, \delta^f)$  with  $\delta^f$  approaching to infinity as the prior distribution of the carrier frequency offset  $f$  [36]. The two kinds of messages to the variable node  $f$  are then as follows.

**Message  $\mathcal{M}_{\mathcal{P}_f \rightarrow f}^{(\xi)}$** : By formatting the Gaussian prior into the form as (21) and then putting the prior into (24), we have the message  $\mathcal{M}_{\mathcal{P}_f \rightarrow f}^{(\xi)}$  as

$$\mathcal{M}_{\mathcal{P}_f \rightarrow f}^{(\xi)} = \frac{1}{\delta^f} \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T. \quad (53)$$

and  $\Omega(f) = \begin{bmatrix} f^2 \\ f \end{bmatrix}$ .

**Message  $\mathcal{M}_{h_i \rightarrow f}^{(\xi)}$** : Similarly, rewriting (5) into the form as (22) and substituting the result into (24), it yields

$$\mathcal{M}_{h_i \rightarrow f}^{(\xi)} = \frac{1}{l_i} \begin{bmatrix} 1 \\ \mathcal{D}_i \end{bmatrix}^T. \quad (54)$$

where  $\mathcal{D}_i = \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus f})} \left( h_i - f_i - \frac{(\mathbf{x}_i - \mathbf{x})^T (\mathbf{v}_i - \mathbf{v})}{cd_i} f_c \right)$  is obtained by using Taylor expansion to approximate the intractable term  $\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus f})} \left( \frac{\mathbf{x}_i - \mathbf{x}}{d_i} \right)$ . As derived in Appendix B, it can be approximated as  $\mathcal{D}_i \approx h_i - f_i^{(\xi)} - \frac{(\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})^T (\mathbf{v}_i^{(\xi)} - \mathbf{v}^{(\xi)})}{cd_i^{(\xi)}} f_c$ .

**Natural Parameter  $\mathcal{H}_f^{(\xi+1)}$ :** By plugging the message  $\mathcal{M}_{\mathcal{P}_f \rightarrow f}^{(\xi)}$  and the message  $\mathcal{M}_{h_i \rightarrow f}^{(\xi)}$  into (24), we can obtain the natural parameter  $\mathcal{H}_f^{(\xi+1)}$  as

$$\mathcal{H}_f^{(\xi+1)} = \begin{bmatrix} \frac{1}{\delta^f} + \sum_{i=1}^M \frac{1}{l_i} \\ \sum_{i=1}^M \frac{\mathcal{D}_i}{l_i} \end{bmatrix}^T \quad (55)$$

Since the variational distribution  $q^{(\xi+1)}(f)$  also follows a Gaussian distribution based on the conjugate prior principle, by denoting the mean and variance as  $f^{(\xi+1)}$  and  $\delta_f^{(\xi+1)}$  respectively, the natural parameter of  $q^{(\xi+1)}(f)$  can also be given as

$$\mathcal{H}_f^{(\xi+1)} = \begin{bmatrix} \frac{1}{\delta_f^{(\xi+1)}} \\ f^{(\xi+1)} \\ \frac{f^{(\xi+1)}}{\delta_f^{(\xi+1)}} \end{bmatrix}^T. \quad (56)$$

By combining (55) and (56), the posterior mean  $f^{(\xi+1)}$  then follows

$$f^{(\xi+1)} = \left( \frac{1}{\delta^f} + \sum_{i=1}^M \frac{1}{l_i} \right)^{-1} \left( \sum_{i=1}^M \frac{\mathcal{D}_i}{l_i} \right), \quad (57)$$

which is also the MAP estimation of  $f$  in the  $(\xi + 1)$ th iteration. Similarly, the carrier frequency offset of node  $i$  can also be formulated to follow  $\mathcal{N} \left( f_i | f_i^{(\xi+1)}, \left( \frac{1}{\delta_i^f} + \frac{1}{l_i} \right)^{-1} \right)$  with  $f_i^{(\xi+1)} = \left( \frac{1}{\delta_i^f} + \frac{1}{l_i} \right)^{-1} \frac{\mathcal{D}_i}{l_i}$  and a non-informative prior  $\mathcal{N} \left( 0, \delta_i^f \right)$ .

6) **Estimation of  $\mathbf{v}$ :** Following the similar approach as the estimation of  $f$ , we assume a non-informative Gaussian distribution  $\mathcal{N}(\mathbf{0}, \Sigma^v)$  with  $\|\Sigma^v\|_2 \rightarrow \infty$  as the prior distribution of the velocity  $\mathbf{v}$  [36]. The message  $\mathcal{M}_{\mathcal{P}_v \rightarrow v}^{(\xi)}$  from the parent node can then be written as

$$\mathcal{M}_{\mathcal{P}_v \rightarrow v}^{(\xi)} = \begin{bmatrix} (\Sigma^v)^{-1} \\ \mathbf{0} \end{bmatrix}^T. \quad (58)$$

$$\text{and } \Omega(\mathbf{v}) = \begin{bmatrix} \mathbf{v} \mathbf{v}^T \\ \mathbf{v} \end{bmatrix}.$$



On the other hand, by putting (5) into (24), the messages from the child node  $h_i$  to  $\mathbf{v}$  are given by

$$\mathcal{M}_{h_i \rightarrow \mathbf{v}}^{(\xi)} = \begin{bmatrix} \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \\ \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \tilde{\mathbf{v}}_i^{(\xi)} \end{bmatrix}^T, \quad (59)$$

where

$$\tilde{\Sigma}_i^{v,(\xi)} = \frac{f_c^2}{c^2 l_i} \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^2} \right) \approx \frac{f_c^2}{c^2 l_i} \frac{(\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})(\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})^T}{(d_i^{(\xi)})^2}, \quad (60)$$

$$\tilde{\mathbf{v}}_i^{(\xi)} \approx - \left( \left( h_i - f_i^{(\xi)} - f^{(\xi)} \right) \left( \frac{\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}}{d_i^{(\xi)}} \right) \right) \frac{f_c}{c l_i} + \tilde{\Sigma}_i^{v,(\xi)} \mathbf{v}_i^{(\xi)}. \quad (61)$$

and detailed approximation of (60) is given in Appendix C.

**Natural Parameter  $\mathcal{H}_v^{(\xi+1)}$ :** Given the posterior distribution  $q^{(\xi+1)}(\mathbf{v})$  as Gaussian distribution with mean and variance denoted as  $\mathbf{v}^{(\xi+1)}$  and  $\Sigma^{v,(\xi+1)}$  respectively by following the conjugate prior principle, the natural parameter of the variational distribution can be written as

$$\mathcal{H}_v^{(\xi+1)} = \begin{bmatrix} \left( \Sigma^{v,(\xi+1)} \right)^{-1} \\ \left( \Sigma^{v,(\xi+1)} \right)^{-1} \mathbf{v}^{(\xi+1)} \end{bmatrix}^T. \quad (62)$$

Putting (58) and (59) into (24) and mapping the result with (62), it yields

$$\mathbf{v}^{(\xi+1)} = \left( \Sigma^{v,(\xi+1)} \right)^{-1} \sum_{i=1}^M \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \tilde{\mathbf{v}}_i^{(\xi)}, \quad (63)$$

where

$$\Sigma^{v,(\xi+1)} = \left( (\Sigma^v)^{-1} + \sum_{i=1}^M \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \right)^{-1}. \quad (64)$$

The MAP estimation of  $\mathbf{v}$  is thus obtained as the posterior mean as (63). Similarly, the velocities of sensor nodes  $\mathbf{v}_i$  can also be given to follow  $\mathcal{N}(\mathbf{v}_i | \mathbf{v}_i^{(\xi+1)}, \Sigma_i^{v,(\xi+1)})$  with

$$\Sigma_i^{v,(\xi+1)} = \left( (\Sigma_i^v)^{-1} + \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \right)^{-1}, \quad (65)$$

$$\mathbf{v}_i^{(\xi+1)} = \left( \Sigma_i^{v,(\xi+1)} \right)^{-1} \left[ \left( \tilde{\Sigma}_i^{v,(\xi)} \right)^{-1} \tilde{\mathbf{v}}_i^{(\xi)} + (\Sigma_i^v)^{-1} \bar{\mathbf{v}}_i \right], \quad (66)$$

$$\tilde{\mathbf{v}}_i^{(\xi)} \approx \left( \left( h_i - f_i^{(\xi)} - f^{(\xi)} \right) \left( \frac{\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}}{d_i^{(\xi)}} \right) \right) \frac{f_c}{c l_i} + \tilde{\Sigma}_i^{v,(\xi)} \mathbf{v}_i^{(\xi)}. \quad (67)$$

7) **Estimation of  $\mathbf{x}$** : Similarly, we assume a non-informative Gaussian distribution as the prior distribution of the target location, that is,  $\mathcal{N}(\mathbf{x}|\bar{\mathbf{x}}, \Sigma^{\mathbf{x}})$  where  $\bar{\mathbf{x}}$  is a coarse estimate of  $\mathbf{x}$  and  $\Sigma^{\mathbf{x}}$  is the covariance matrix with a large value of  $\|\Sigma^{\mathbf{x}}\|_2$  [35]. With this prior assumption, the messages to the variable node  $\mathbf{x}$  include the message  $\mathcal{M}_{\mathcal{P}_{\bar{\mathbf{x}} \rightarrow \mathbf{x}}}^{(\xi)}$  from the parent nodes  $\bar{\mathbf{x}}$  and  $\Sigma^{\mathbf{x}}$ , and the message  $\mathcal{M}_{(r_i, h_i) \rightarrow \mathbf{x}}^{(\xi)}$  from the child nodes  $r_i$  and  $h_i$ .

**Message  $\mathcal{M}_{\mathcal{P}_{\bar{\mathbf{x}} \rightarrow \mathbf{x}}}^{(\xi)}$** : By plugging the non-informative Gaussian prior of  $\mathbf{x}$  into (24), the message from the parent nodes to  $\mathbf{x}$  is obtained as

$$\mathcal{M}_{\mathcal{P}_{\bar{\mathbf{x}} \rightarrow \mathbf{x}}}^{(\xi)} = \begin{bmatrix} (\Sigma^{\mathbf{x}})^{-1} \\ (\Sigma^{\mathbf{x}})^{-1} \bar{\mathbf{x}} \end{bmatrix}^T, \quad (68)$$

and  $\Omega(\mathbf{x}) = \begin{bmatrix} \mathbf{x}\mathbf{x}^T \\ \mathbf{x} \end{bmatrix}$ .

**Message  $\mathcal{M}_{(r_i, h_i) \rightarrow \mathbf{x}}^{(\xi)}$** : The range and Doppler measurements involve nonlinear terms  $d_i$  (6) and  $\psi_i$  (4) respectively. To ease the derivation, we first linearize them based on the  $\xi$ -th estimations of the locations and velocities, i.e.,  $(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)})$  and  $(\mathbf{v}_i^{(\xi)}, \mathbf{v}^{(\xi)})$ . As presented in Appendix A,  $d_i$  is approximated as

$$d_i \approx d_i^{(\xi)} + \boldsymbol{\rho}_i^T (\mathbf{x}_i - \mathbf{x}_i^{(\xi)} + \mathbf{x}^{(\xi)} - \mathbf{x}), \quad (69)$$

By denoting  $\psi_i$  (4) as  $\psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v})$ , the Doppler shift can be approximated based on first order Taylor expansion as

$$\begin{aligned} \psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v}) &= \psi_i(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)}, \mathbf{v}_i^{(\xi)}, \mathbf{v}^{(\xi)}) + \mathcal{F}_{\mathbf{x}_i}^{(\xi)} (\mathbf{x}_i - \mathbf{x}_i^{(\xi)}) + \mathcal{F}_{\mathbf{x}}^{(\xi)} (\mathbf{x} - \mathbf{x}^{(\xi)}) \\ &\quad + \mathcal{F}_{\mathbf{v}_i}^{(\xi)} (\mathbf{v}_i - \mathbf{v}_i^{(\xi)}) + \mathcal{F}_{\mathbf{v}}^{(\xi)} (\mathbf{v} - \mathbf{v}^{(\xi)}), \end{aligned} \quad (70)$$

where  $\mathcal{F}_{\mathbf{x}_i} = \left(\frac{\partial \psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v})}{\partial \mathbf{x}_i}\right)^T$ ,  $\mathcal{F}_{\mathbf{x}} = \left(\frac{\partial \psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v})}{\partial \mathbf{x}}\right)^T$ ,  $\mathcal{F}_{\mathbf{v}_i} = \left(\frac{\partial \psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v})}{\partial \mathbf{v}_i}\right)^T$  and  $\mathcal{F}_{\mathbf{v}} = \left(\frac{\partial \psi_i(\mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v})}{\partial \mathbf{v}}\right)^T$  are the first order derivatives and are respectively given by

$$\mathcal{F}_{\mathbf{x}} = -\mathcal{F}_{\mathbf{x}_i} = (\mathbf{v}_i^{(\xi)} - \mathbf{v}^{(\xi)})^T \frac{f_c}{c} \left( \frac{\mathbf{I}}{d_i} - \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^3} \right), \quad (71)$$

$$\mathcal{F}_{\mathbf{v}_i}^{(\xi)} = -\mathcal{F}_{\mathbf{v}}^{(\xi)} = \frac{(\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})^T}{\|\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}\| c} f_c, \quad (72)$$

and  $\mathcal{F}_{\mathbf{x}_i}^{(\xi)}$  and  $\mathcal{F}_{\mathbf{x}}^{(\xi)}$  are obtained by plugging the estimates  $\mathbf{x}_i^{(\xi)}$ ,  $\mathbf{x}^{(\xi)}$  into (71).

As  $r_i$  and  $h_i$  are the child nodes of  $\mathbf{x}$ , the message  $\mathcal{M}_{(r_i, h_i) \rightarrow \mathbf{x}}^{(\xi)}$  is the sum of messages  $\mathcal{M}_{r_i \rightarrow \mathbf{x}}^{(\xi)}$  and  $\mathcal{M}_{h_i \rightarrow \mathbf{x}}^{(\xi)}$ . Firstly, the range likelihood (13) can be reformulated as

$$p(\mathbf{r}|\mathbf{x}, \mathbf{x}_s, \boldsymbol{\theta}) = \prod_{i=1}^M p(r_i|\mathbf{x}, \mathbf{x}_s, \boldsymbol{\theta}) = \exp \left( \sum_{i=1}^M (y_{i,1} \ln \mathcal{N}(r_i - d_i, \Lambda_{\mathcal{L}}^{-1}) + y_{i,2} \ln \mathcal{N}(r_i - d_i - \mu, \Lambda^{-1})) \right). \quad (73)$$

By substituting the linearized distance (69) into (73), the range likelihood can be rewritten as

$$\begin{aligned} p(\mathbf{r}|\mathbf{x}, \mathbf{x}_s, \boldsymbol{\theta}) &= \exp \left( \sum_{i=1}^M (y_{i,1} \ln \mathcal{N}(r_i - d_i, \Lambda_{\mathcal{L}}^{-1}) + y_{i,2} \ln \mathcal{N}(r_i - d_i - \mu, \Lambda^{-1})) \right) \\ &= \exp \left( \text{tr} \left( \left[ \begin{array}{c} (y_{i,1} \Lambda_{\mathcal{L}}^{-1} + y_{i,2} \Lambda^{-1}) \mathbf{I} \\ (y_{i,1} \Lambda_{\mathcal{L}}^{-1} \boldsymbol{\eta}_{i,1}^{(\xi)} + y_{i,2} \Lambda^{-1} \boldsymbol{\eta}_{i,2}^{(\xi)}) \end{array} \right]^T \left[ \begin{array}{c} \mathbf{x} \mathbf{x}^T \\ \mathbf{x} \end{array} \right] \right) \right) + \mathcal{C}, \end{aligned} \quad (74)$$

where  $\boldsymbol{\eta}_{i,l}^{(\xi)} = \mathbf{x}_i^{(\xi)} - \boldsymbol{\rho}_i(r_i - \mu(l-1))$  and  $\mathcal{C}$  is a constant. Based on (22) and (24), the message from  $\mathbf{r}$  can be derived as

$$\mathcal{M}_{\mathbf{r} \rightarrow \mathbf{x}}^{(\xi)} = \sum_{i=1}^M \mathcal{M}_{r_i \rightarrow \mathbf{x}}^{(\xi)} = \sum_{i=1}^M \left[ \begin{array}{c} \boldsymbol{\Theta}_i^{(\xi)} \\ \sum_{l=1}^2 \tilde{\boldsymbol{\eta}}_{i,l}^{(\xi)} g_{i,l}^{(\xi)} \end{array} \right]^T = \sum_{i=1}^M \left[ \begin{array}{c} \boldsymbol{\Theta}_i^{(\xi)} \\ \mathcal{U}_i^{(\xi)} \end{array} \right]^T, \quad (75)$$

where  $\boldsymbol{\Theta}_i^{(\xi)} = \sum_{l=1}^2 g_{i,1}^{(\xi)} \mathbf{I}$ ,  $g_{i,1}^{(\xi)} = \kappa_{i,1}^{(\xi)} \Lambda_{\mathcal{L}}$ ,  $g_{i,2}^{(\xi)} = \kappa_{i,2}^{(\xi)} \beta^{(\xi)}$ ,  $\tilde{\boldsymbol{\eta}}_{i,l}^{(\xi)} = \mathbf{x}_i^{(\xi)} - \boldsymbol{\rho}_i(r_i - m^{(\xi)}(l-1))$  and  $\mathcal{U}_i^{(\xi)} = \sum_{l=1}^2 \tilde{\boldsymbol{\eta}}_{i,l}^{(\xi)} g_{i,l}^{(\xi)}$ .

Similarly, the Doppler likelihood (5) can be reformulated as

$$p(\mathbf{h}|\mathbf{f}_s, f, \mathbf{x}_s, \mathbf{x}, \mathbf{v}_s, \mathbf{v}) = \prod_{i=1}^M p(h_i|f_i, f, \mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v}) = \exp \left( \sum_{i=1}^M \ln \mathcal{N}(h_i - f_i - f - \psi_i, \iota_i) \right), \quad (76)$$

Plugging (70) into (76) leads to

$$p(h_i|f_i, f, \mathbf{x}_i, \mathbf{x}, \mathbf{v}_i, \mathbf{v}) = \exp \left( \text{tr} \left( \left[ \begin{array}{c} \mathcal{G}^{(\xi)} \\ \mathbf{R}_i \end{array} \right]^T \left[ \begin{array}{c} \mathbf{x} \mathbf{x}^T \\ \mathbf{x} \end{array} \right] \right) \right) + \mathcal{C}, \quad (77)$$

where

$$\mathcal{G}^{(\xi)} = \frac{1}{\iota_i} (\mathcal{F}_{\mathbf{x}}^{(\xi)})^T \mathcal{F}_{\mathbf{x}}^{(\xi)}, \quad (78)$$

$$\mathbf{R}_i = \frac{1}{\iota_i} \left( h_i - f - f_i - \psi_i \left( \mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)}, \mathbf{v}_i^{(\xi)}, \mathbf{v}^{(\xi)} \right) \right) \mathcal{F}_{\mathbf{x}}^{(\xi)}. \quad (79)$$

Then the message from  $\mathbf{h}$  can be given by

$$\mathcal{M}_{\mathbf{h} \rightarrow \mathbf{x}}^{(\xi)} = \sum_{i=1}^M \mathcal{M}_{h_i \rightarrow \mathbf{x}}^{(\xi)} = \sum_{i=1}^M \begin{bmatrix} \mathcal{G}_i^{(\xi)} \\ \mathcal{R}_i^{(\xi)} \end{bmatrix}^T, \quad (80)$$

where  $\mathcal{R}_i^{(\xi)} = \mathbb{E}_{q^{(\xi)}(f, f_i)}(\mathbf{R}_i) = \frac{1}{\iota_i} \left( h_i - f^{(\xi)} - f_i^{(\xi)} - \psi_i(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)}, \mathbf{v}_i^{(\xi)}, \mathbf{v}^{(\xi)}) \right) \mathcal{F}_{\mathbf{x}_i}^{(\xi)}$ .

By plugging (75) and (80) into (24), we obtain

$$\mathcal{M}_{(\mathbf{r}, \mathbf{h}) \rightarrow \mathbf{x}}^{(\xi)} = \sum_{i=1}^M \mathcal{M}_{(r_i, h_i) \rightarrow \mathbf{x}}^{(\xi)} = \underbrace{\sum_{i=1}^M \begin{bmatrix} \Theta_i^{(\xi)} \\ \mathcal{U}_i^{(\xi)} \end{bmatrix}^T}_{\text{TOA-related message}} + \underbrace{\sum_{i=1}^M \begin{bmatrix} \mathcal{G}_i^{(\xi)} \\ \mathcal{R}_i^{(\xi)} \end{bmatrix}^T}_{\text{FOA-related message}}. \quad (81)$$

**Natural Parameter  $\mathcal{H}_{\mathbf{x}}^{(\xi+1)}$ :** By substituting (68) and (81) into (24) and with some manipulations under the constraint of conjugate prior, we can obtain

$$\mathcal{H}_{\mathbf{x}}^{(\xi+1)} = \begin{bmatrix} \sum_{i=1}^M \left( \Theta_i^{(\xi)} + \mathcal{G}_i^{(\xi)} \right) + (\Sigma^{\mathbf{x}})^{-1} \\ \sum_{i=1}^M \left( \mathcal{U}_i^{(\xi)} + \mathcal{R}_i^{(\xi)} \right) + (\Sigma^{\mathbf{x}})^{-1} \bar{\mathbf{x}} \end{bmatrix}^T = \begin{bmatrix} (\Sigma^{(\xi+1)})^{-1} \\ (\mathbf{x}^{(\xi+1)})^T (\Sigma^{(\xi+1)})^{-1} \end{bmatrix}^T, \quad (82)$$

where  $\mathbf{x}^{(\xi+1)}$  and  $\Sigma^{(\xi+1)}$  directly follow as

$$\mathbf{x}^{(\xi+1)} = \left( \sum_{i=1}^M \left( \Theta_i^{(\xi)} + \mathcal{G}_i^{(\xi)} \right) + (\Sigma^{\mathbf{x}})^{-1} \right)^{-1} \left( \sum_{i=1}^M \left( \mathcal{U}_i^{(\xi)} + \mathcal{R}_i^{(\xi)} \right) + \bar{\mathbf{x}}^T (\Sigma^{\mathbf{x}})^{-1} \right), \quad (83)$$

$$\Sigma^{(\xi+1)} = \left( \sum_{i=1}^M \left( \Theta_i^{(\xi)} + \mathcal{G}_i^{(\xi)} \right) + (\Sigma^{\mathbf{x}})^{-1} \right)^{-1}. \quad (84)$$

Similarly to (56) and (62), the MAP estimate of the location  $\mathbf{x}$  is the posterior mean  $\mathbf{x}^{(\xi+1)}$  given in (83). The locations of sensor nodes can also be refined similarly to follow  $\mathcal{N}(\mathbf{x}_i | \mathbf{x}_i^{(\xi+1)}, \Sigma_i^{\mathbf{x}, (\xi+1)})$  with

$$\Sigma_i^{\mathbf{x}, (\xi+1)} = \left( \Theta_i^{(\xi)} + \mathcal{G}_i^{(\xi)} + (\Sigma_i^{\mathbf{x}})^{-1} \right)^{-1}, \quad (85)$$

$$\mathbf{x}_i^{(\xi+1)} = \left( \Sigma^{(\xi+1)} \right)^{-1} \left( \mathcal{U}_i^{(\xi)} + \mathcal{R}_i^{(\xi)} + (\Sigma_i^{\mathbf{x}})^{-1} \bar{\mathbf{x}}_i \right), \quad (86)$$

$$\mathcal{U}_i^{(\xi)} = \sum_{l=1}^2 \left( \mathbf{x}_i^{(\xi)} + \boldsymbol{\rho}_i(r_i + m^{(\xi)}(l-1)) \right) g_{i,l}^{(\xi)}, \quad (87)$$

$$\mathcal{R}_i^{(\xi)} = \frac{1}{\iota_i} \left( h_i - f^{(\xi)} - f_i^{(\xi)} - \psi_i(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)}, \mathbf{v}_i^{(\xi)}, \mathbf{v}^{(\xi)}) \right) \mathcal{F}_{\mathbf{x}_i}^{(\xi)}, \quad (88)$$

$$\mathcal{G}_i^{(\xi)} = \frac{1}{\iota_i} \left( \mathcal{F}_{\mathbf{x}_i}^{(\xi)} \right)^T \mathcal{F}_{\mathbf{x}_i}^{(\xi)}. \quad (89)$$

### C. Summary and Discussion

Our proposed algorithm is an iterative algorithm developed based on Bayesian criterion and variational message passing framework. By reformulating the prior and likelihood probability distributions as exponential family, messages can be calculated easily and only the natural parameters need to be updated in each iteration. As shown in (29), (38), (44), (49), (56), (62) and (82), the natural parameters only involve the mean and covariance of the corresponding variational distributions. In other words, only the means and covariance matrices need to be updated iteratively. Clearly from Fig. 1, the probabilistic graph is loop-free. As proved in [33, 34], under loop-free graph, the VMP-based algorithm can reduce the KL divergence in each iteration and thus the convergence of our proposed algorithm is guaranteed. After convergence, the target location and its velocity can be finally estimated as (83) and (63) respectively. In summary, the implementation of the proposed algorithm is shown in **Algorithm 1**.

## IV. SIMULATION RESULTS

In this section, we demonstrate the estimation performances of the proposed algorithm in different scenarios. We consider the 2D localization of one target node with the help of  $M = 10$  sensor nodes in the field of  $100m \times 100m$ . Among the 10 sensor nodes, some of them are malicious nodes. The number of malicious nodes are unknown and the expectation of attack probability is given by  $\mathbb{E}(\alpha_2) = \frac{\bar{\lambda}_2}{\bar{\lambda}_1 + \bar{\lambda}_2}$ . The hyper parameters are respectively given as  $\Lambda_{\mathcal{L}}^{-1} = 1$ ,  $\bar{m}_2 = 10$ ,  $\bar{\gamma} = 1$ ,  $\bar{a} = 10^2$ , and  $\bar{b} = 10^4$ . The location uncertainty of the sensor node is set as  $\Sigma_i^x = \delta^2 \mathbf{I} = 0.1 \mathbf{I}$  and the non-informative covariance matrix of target location is set to be  $\Sigma^x = 100 \mathbf{I}$ . The unnormalized CFO is generated in  $2\pi [-0.2, 0.2]$  for each node. The non-informative variances  $\delta_f$  and  $\Sigma_v$  are set as  $10^3$  and  $10^3 \mathbf{I}$  respectively. The mean and variance of sensor velocity are respectively given by  $\bar{v}_i = 20m/s \cdot \mathbf{1}^T$  and  $\Sigma_i^v = 0.5 \mathbf{I}$ . The threshold for convergence control is set as  $\mathcal{T} = 0.01$ . The simulation settings keep unaltered unless otherwise stated. The results shown in this paper are the average of the results over 1000 localization realizations. For comparison, the proposed algorithm referred as VMP-JLVE is compared to the following algorithms:

- WLS1: A weighted least square algorithm proposed in [17] which exploited the range and Doppler measurements to localize a target node with the aid of a set of non-malicious nodes with sensor node location uncertainties;

---

**Algorithm 1** VMP-Based Joint Localization and Velocity Estimation
 

---

- 1: Input the prior distributions  $p(\mathbf{v}_i)$  and  $p(\mathbf{x}_i)$ ,  $i \in \mathcal{S}$ , the non-informative Gaussian priors for the target velocity and the target location, and the prior information  $\bar{\lambda}_l, \bar{m}, \bar{\gamma}, \bar{a}, \bar{b}$ ;
  - 2: Collect the range and Doppler measurements  $\mathbf{r}$  and  $\mathbf{h}$ ;
  - 3: Set the iteration index  $\xi = 0$ ;
  - 4: **while**  $|\text{KL}(q^{(\xi+1)}(\Phi) || p(\Phi | \mathbf{r}, \mathbf{h})) - \text{KL}(q^{(\xi)}(\Phi) || p(\Phi | \mathbf{r}, \mathbf{h}))| > \mathcal{T}$  **do**
  - 5:    $\mathbf{y}_i$ : Calculate the messages from parent node  $\alpha$  and child node  $r_i$  according to (26) and (28) respectively;
  - 6:    $\mu$ : Calculate the messages from parent nodes  $\bar{m}$  and  $\bar{\gamma}$  and child node  $r_i$  according to (35) and (36) respectively;
  - 7:    $\alpha$ : Calculate the messages from parent node  $\lambda$  and child nodes  $\mathbf{y}$  according to (42) and (43) respectively;
  - 8:    $\Lambda$ : Calculate the messages from parent nodes  $\bar{a}$  and  $\bar{b}$  and child node  $r_i$  according to (46) and (47) respectively;
  - 9:    $f$ : Calculate the messages from parent node  $\delta_f$  and child node  $h_i$  according to (53) and (54) respectively;
  - 10:    $\mathbf{v}$ : Calculate the messages from parent nodes  $\Sigma_v$  and child node  $h_i$  according to (58) and (59) respectively;
  - 11:    $\mathbf{x}$ : Calculate the messages from parent nodes  $\bar{\mathbf{x}}$  and  $\Sigma_x$  and child nodes  $r_i$  and  $h_i$  according to (68), (75), (80) and (81) respectively;
  - 12:   Calculate the expectations of all variable nodes based on the updated natural parameters from (24);
  - 13:   Update the expectations as:  $\kappa_{i,l}^{(\xi+1)} = \frac{\phi_{i,l}^{(\xi+1)}}{\sum_{l=1}^2 \phi_{i,l}^{(\xi+1)}}$ ,  $\mu^{(\xi+1)} = m^{(\xi+1)}$ ,  $\alpha_l^{(\xi+1)} = \frac{\lambda_l^{(\xi+1)}}{\sum_{l=1}^2 \lambda_l^{(\xi+1)}}$ ,  $\Lambda^{(\xi+1)} = a^{(\xi+1)}/b^{(\xi+1)}$ ,  $f^{(\xi+1)}$ ,  $\mathbf{v}^{(\xi+1)}$ ,  $\mathbf{x}^{(\xi+1)}$  according to (31), (41), (45), (52), (57), (63) and (83) respectively;
  - 14:    $\xi = \xi + 1$ ;
  - 15: **end while**
  - 16: Output the location estimation  $\mathbf{x} = \mathbf{x}^{(\xi+1)}$  and velocity estimation  $\mathbf{v} = \mathbf{v}^{(\xi+1)}$  according to (83) and (63) respectively, as well as the estimates of the nuisance parameters  $\mathbf{y}$ ,  $\alpha$ ,  $\Lambda$  and  $\mu$ .
- 

- WLS2: Another weighted least square algorithm proposed in [18] which exploited the range and Doppler measurements to localize a target node with the aid of a set of non-malicious nodes with sensor node location uncertainties;
- SLOT: A secure localization algorithm proposed in [21] which assumed perfect identification of malicious nodes, perfect knowledge of attack parameters and perfect sensor node locations, and exploited only the range measurements for localization;
- ELSA: A secure detection and localization algorithm proposed in [27] which assumed perfect knowledge of attack parameters and perfect sensor node locations, and exploited only the range measurements for malicious node identification and target node localization;

- **BCRB**: The Bayesian Cramer Rao Bound for the localization performance. Given the variable vector  $\Phi = [\mathbf{x}, \mathbf{v}, \mathbf{v}_s, f, \mathbf{f}_s, \mathbf{x}_s, \boldsymbol{\theta}]$ , the Fisher information matrix  $J(\Phi)$  is defined as [35]

$$[J(\Phi)]_{m,n} = \underbrace{-\mathbb{E}_{\mathbf{r}, \mathbf{h}, \Phi} \left[ \frac{\partial^2 \ln p(\mathbf{r}, \mathbf{h} | \Phi)}{\partial \{\Phi\}_m \partial \{\Phi\}_n} \right]}_{\text{Likelihood FIM}} \underbrace{-\mathbb{E}_{\Phi} \left[ \frac{\partial^2 \ln p(\Phi)}{\partial \{\Phi\}_m \partial \{\Phi\}_n} \right]}_{\text{Prior FIM}}, \quad (90)$$

where  $\mathbb{E}_{\Phi, \mathbf{r}, \mathbf{h}}[\cdot]$  denotes the expectation with respect to the distribution  $p(\mathbf{r}, \mathbf{h}, \Phi)$ . By following the similar derivations in [35], the BCRB is given by

$$\text{BCRB} = \text{tr}(J^{-1}(\Phi)). \quad (91)$$

Notice that the WLS1 and WLS2 algorithms [17, 18] assumed perfect frequency synchronization in the network, i.e., the carrier frequency offsets ( $f$  and  $f_i$ ) are assumed as zero. This assumption however is not necessary in our proposed VMP-JLVE algorithm.

#### A. Malicious Nodes Identification

In the proposed VMP-JLVE-based algorithm, the identification of malicious nodes can be achieved through the MAP estimation of the auxiliary indicator vector  $\mathbf{y}_i$  based on (33). The identification performance of the proposed algorithm is compared with that of ELSA in [27] and the results are shown in Fig. 2. In the comparisons, two scenarios with various attack variances  $\mathbb{E}(\Lambda^{-1}) = \bar{b}/\bar{a} = \{100, 10\}$  are considered. It is clear that our proposed algorithm can provide much higher identification accuracy than the ELSA algorithm. Specifically, the identification accuracy of our algorithm could be 95%, while that of ELSA is only around 87% even with the perfect statistical information of the attack distribution, i.e., perfect information of  $\mu$  and  $\Lambda$ .

#### B. Localization and Velocity Estimation under No-Attack Environment

No attack environment is first considered to test the localization performance of our proposed algorithm. The results are shown in Fig. 3. Notice that in the WLS1 and WLS2 algorithms, the carrier frequency offsets in the nodes are ignored, while in the SLOT and ELSA algorithm, brutal-force grid searching is adopted for localization and the grid size is set as  $50 \times 50$  in the simulations. Clearly, our proposed VMP-JLVE algorithm performs the best among all the algorithms since MAP criterion is adopted and additional Doppler measurements are exploited for estimation. Its performance is close to the BCRB bound after convergence. In the no

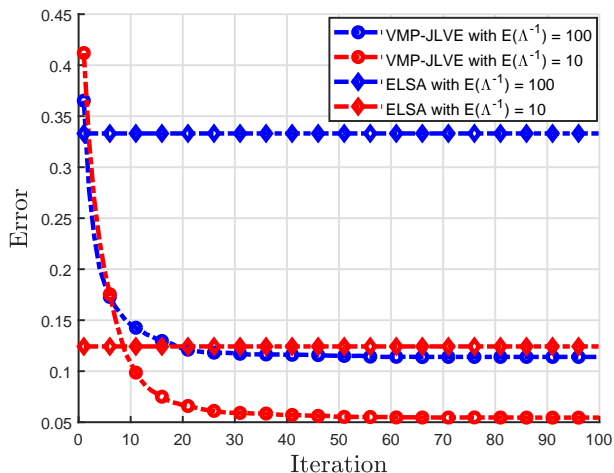


Fig. 2: Identification error under attack environments with  $\delta = 0.1$  and  $\mathbb{E}(\alpha_2) = 0.1$

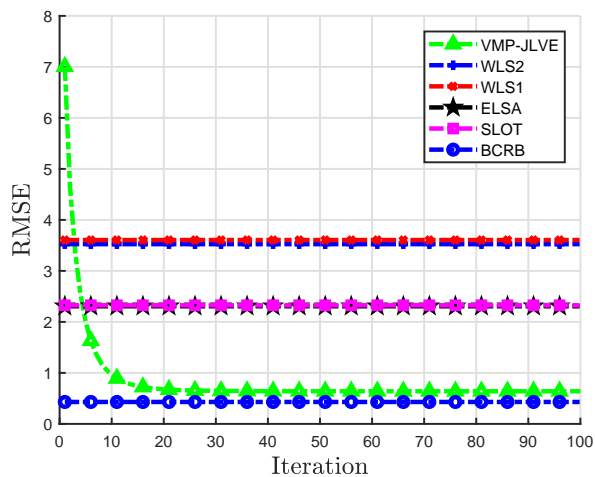


Fig. 3: Localization RMSEs under no-attack environment with  $\delta = 0.1$

attack environment, the WLS1 and WLS2 algorithm perform similarly as the worst due to the impact of ignored carrier frequency offsets. The velocity estimation performance of the proposed algorithm is then compared with that of WLS1 and WLS2, as shown in Fig. 4. Similarly, due to the frequency synchronization in (57), our proposed algorithm provides much higher velocity estimation accuracy than the WLS1 and WLS2 algorithms.



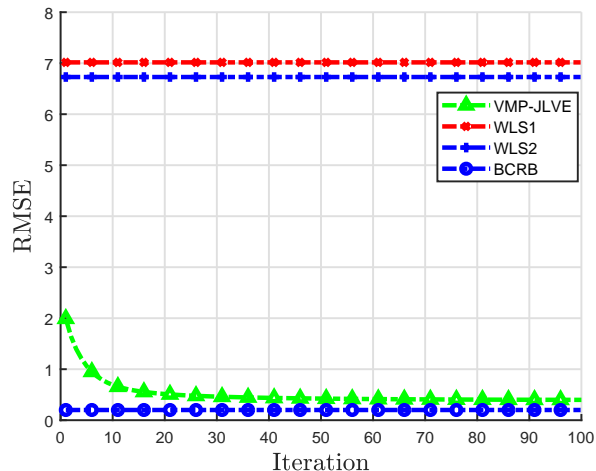


Fig. 4: Velocity RMSEs under no-attack environment with  $\delta = 0.1$

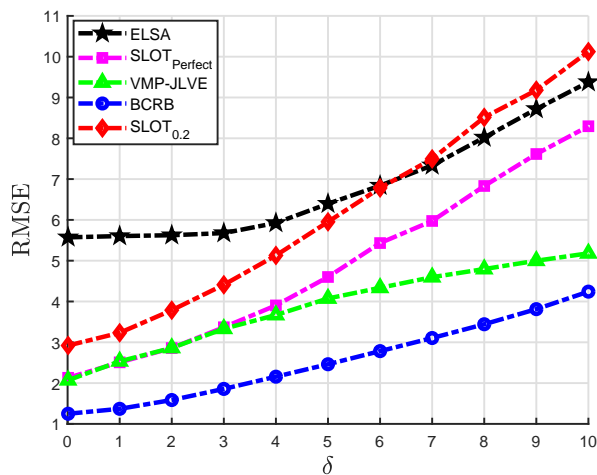


Fig. 5: Localization RMSEs with Different Sensor Node Uncertainty  $\delta$ ,  $\mathbb{E}(\Lambda^{-1}) = 10$  and  $\mathbb{E}(\alpha_2) = 0.2$

### C. Secure Localization under Attack Environment

Now attack environment is considered, and the impacts of sensor location uncertainty and the attack probability on secure localization algorithms are investigated. The results are shown in Fig. 5 and Fig. 6 respectively. In the simulations, two implementations of SLOT algorithm are tested. In original SLOT algorithm [21], perfect identification of malicious nodes is assumed.

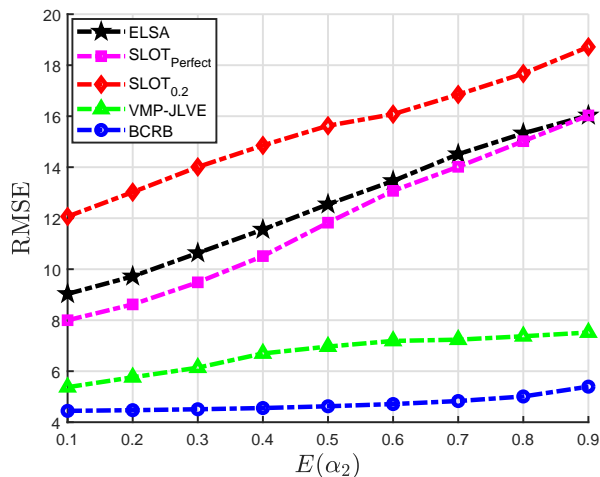


Fig. 6: Localization RMSEs with different attack ratios  $\mathbb{E}(\alpha_2)$ ,  $\delta = 10$  and  $\mathbb{E}^{-1}(\Lambda) = 100$

It is referred as  $\text{SLOT}_{\text{Perfect}}$ . Since the perfect identification of malicious nodes is difficult to achieve in practice, we also simulate another implementation of SLOT algorithm with 20% identification error. This implementation is referred as  $\text{SLOT}_{0.2}$ . As shown in Fig. 5, under 20% attack probability, i.e.,  $\mathbb{E}(\alpha_2) = 0.2$ , when the sensor location uncertainty  $\delta$  increases from 0 to 10 which covers the cases of perfect sensor location and large uncertainty of sensor location, the localization accuracy of our proposed VMP-JLVE-based algorithm slightly degrades. However, it still performs better than the ELSA and SLOT algorithms even with perfect identification of malicious nodes. Moreover, as shown in Fig. 6, when the attack probability  $\mathbb{E}(\alpha_2)$  increases from 0.1 to 0.9, only slight degradation of localization accuracy is observed for our proposed VMP-JLVE algorithm, while significant degradation occurs for both SLOT and ELSA algorithms. These results demonstrate the superiority and robustness of our proposed algorithm under attack environments. The localization superiority of our proposed algorithm is mainly coming from MAP estimation with additional prior information and Doppler measurements, when comparing to the SLOT and ELSA algorithms with maximum likelihood (ML) estimation and range measurements only.

## V. CONCLUSION

In this paper, we proposed a variational message passing algorithm for secure localization and velocity estimation in mobile sensor networks with multiple malicious nodes and sensor location

uncertainty. The proposed VMP-JLVE-based algorithm followed MAP criterion and solved the intractable MAP estimation problem by an iterative message passing algorithm. In each iteration, the variational distributions were updated in closed-forms, which enabled the estimation of the target location and velocity together with the identification of malicious nodes. Only the statistical information of the attacks from malicious nodes was assumed, which improved the practicability of our proposed algorithm. The superiority and robustness of our proposed VMP-JLVE-based algorithm were well validated by comparison with existing secure localization algorithms.

## APPENDIX A

In (28), the expectation of  $\mathcal{H}_C(r_i)$  involves expectations of two elements with respect to variables  $\mathbf{x}$ ,  $\mathbf{x}_i$ ,  $\mu$  and  $\Lambda$ . The expectation is given by

$$\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}[\mathcal{H}_C(r_i)] = \left[ \begin{array}{c} -\frac{\Lambda_C}{2} \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(r_i - d_i)^2 - \frac{1}{2} \ln 2\pi + \frac{1}{2} \ln \Lambda_C \\ -\frac{\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})(\Lambda)}}{2} \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(r_i - d_i - \mu)^2 - \frac{1}{2} \ln 2\pi + \frac{1}{2} \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(\ln \Lambda) \end{array} \right]. \quad (92)$$

The involved expectations are calculated individually as follows. At first, we have

$$\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(\Lambda) = \mathbb{E}_{q^{(\xi)}(\Lambda)}(\Lambda) = a^{(\xi)}/b^{(\xi)} \triangleq \beta^{(\xi)}, \quad (93)$$

where  $a^{(\xi)}$  and  $b^{(\xi)}$  are the parameters of the variational distribution  $q^{(\xi)}(\Lambda)$ . By following the conjugate prior principle and with the prior distribution of  $\Lambda$  in (11), we have  $q^{(\xi)}(\Lambda) = \Gamma(\Lambda|a^{(\xi)}, b^{(\xi)})$ . The parameters of  $a^{(\xi)}$  and  $b^{(\xi)}$  will be derived later in Section III-B4.

Using the results in [37], it yields

$$\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(\ln \Lambda) = \varphi(a^{(\xi)}) + \ln(b^{(\xi)}), \quad (94)$$

where  $\varphi(\cdot)$  is a Gamma function.

On the other hand, for the expectation terms  $\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(r_i - d_i - \mu)^2$  and  $\mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(r_i - d_i)^2$ , we define  $\omega_{i,l}^{(\xi)} \triangleq \mathbb{E}_{q^{(\xi)}(\Phi_{\setminus \mathbf{y}_i})}(r_i - d_i - \mu(l-1))^2$  with  $l=1$  corresponding to the non-attack case while  $l=2$  corresponding to the malicious attack case from sensor  $i$ . By expanding the quadratic term and taking expectations of variables with respect to  $\mathbf{x}$ ,  $\mathbf{x}_i$  and  $\mu$ , we can obtain

$$\omega_{i,l}^{(\xi)} = (r_i - m^{(\xi)}(l-1))^2 + \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i^2) - 2(r_i - m^{(\xi)}(l-1)) \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i) + [\gamma^{(\xi)}]^{-1}(l-1), \quad (95)$$

where  $m^{(\xi)}$  and  $[\gamma^{(\xi)}]^{-1}$  are the  $\xi$ -th mean and variance of the distribution  $\mathcal{N}(\mu|m^{(\xi)}, [\gamma^{(\xi)}]^{-1})$ . By following the conjugate prior principle and with the prior distribution of  $\mu$  in (12), we have  $q^{(\xi)}(\mu) = \mathcal{N}(\mu|m^{(\xi)}, [\gamma^{(\xi)}]^{-1})$ . The parameters of  $m^{(\xi)}$  and  $\gamma^{(\xi)}$  are derived in the corresponding estimation of  $\mu$  in Section III-B2. Similarly, we also have  $q^{(\xi)}(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{x}^{(\xi)}, \Sigma^{\mathbf{x},(\xi)})$  and  $q^{(\xi)}(\mathbf{x}_i) = \mathcal{N}(\mathbf{x}_i|\mathbf{x}_i^{(\xi)}, \Sigma_i^{\mathbf{x},(\xi)})$ , where  $\mathbf{x}_i^{(\xi)}$ ,  $\mathbf{x}^{(\xi)}$ ,  $\Sigma_i^{\mathbf{x},(\xi)}$  and  $\Sigma^{\mathbf{x},(\xi)}$  are the estimated locations and covariance matrices of the  $i$ th sensor node and target node respectively.

Due to the nonlinear expression in the distance  $d_i$ , we apply the Taylor expansion at  $(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)})$  for linearization, which is given by

$$d_i \approx d_i^{(\xi)} + \boldsymbol{\rho}_i^T (\mathbf{x}_i - \mathbf{x}_i^{(\xi)} + \mathbf{x}^{(\xi)} - \mathbf{x}), \quad (96)$$

where  $\boldsymbol{\rho}_i = \frac{\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}}{d_i^{(\xi)}}$ .

Hence the expectations  $\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i^2)$  and  $\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i)$  are respectively given by

$$\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i) = d_i^{(\xi)}, \quad (97)$$

and

$$\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})}(d_i^2) = (\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})^T (\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}) + \text{tr}(\Sigma_i^{\mathbf{x},(\xi)} + \Sigma^{\mathbf{x},(\xi)}). \quad (98)$$

By plugging (93), (94) and (95) into (92), we can obtain the result in (28).

## APPENDIX B

The term  $\mathcal{D}_i$  can be reformulated as

$$\mathcal{D}_i = h_i - \mathbb{E}_{q^{(\xi)}(f_i)}(f_i) - \frac{f_c}{c} \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{\mathbf{x}_i - \mathbf{x}}{d_i} \right)^T \mathbb{E}_{q^{(\xi)}(\mathbf{v}_i, \mathbf{v})} (\mathbf{v}_i - \mathbf{v}). \quad (99)$$

Clearly, the second term in (99) follows

$$\mathbb{E}_{q^{(\xi)}(f_i)}(f_i) = f_i^{(\xi)}. \quad (100)$$

where  $f_i^{(\xi)}$  is the posterior mean given in Section III-B5. Meanwhile, we have

$$\mathbb{E}_{q^{(\xi)}(\mathbf{v}_i, \mathbf{v})} (\mathbf{v}_i - \mathbf{v}) = \mathbf{v}_i^{(\xi)} - \mathbf{v}^{(\xi)}, \quad (101)$$

where  $\mathbf{v}_i^{(\xi)}$  and  $\mathbf{v}^{(\xi)}$  are the corresponding posterior means given in Section III-B6.

For the expectation term  $\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{\mathbf{x}_i - \mathbf{x}}{d_i} \right)^T$ , we apply Taylor expansion to approximate it. By denoting  $\mathcal{S}(\mathbf{x}_i, \mathbf{x}) = \frac{\mathbf{x}_i - \mathbf{x}}{d_i}$ , we linearize the nonlinear function  $\mathcal{S}(\mathbf{x}_i, \mathbf{x})$  with respect to  $(\mathbf{x}^{(\xi)}, \mathbf{x}_i^{(\xi)})$  as

$$\mathcal{S}(\mathbf{x}_i, \mathbf{x}) = \mathcal{S}(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)}) + \left. \frac{\partial \mathcal{S}(\mathbf{x}_i, \mathbf{x})}{\partial \mathbf{x}_i} \right|_{\substack{\mathbf{x}_i = \mathbf{x}_i^{(\xi)} \\ \mathbf{x} = \mathbf{x}^{(\xi)}}} (\mathbf{x}_i - \mathbf{x}_i^{(\xi)}) + \left. \frac{\partial \mathcal{S}(\mathbf{x}_i, \mathbf{x})}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}_i = \mathbf{x}_i^{(\xi)} \\ \mathbf{x} = \mathbf{x}^{(\xi)}}} (\mathbf{x} - \mathbf{x}^{(\xi)}), \quad (102)$$

where  $\frac{\partial \mathcal{S}(\mathbf{x}_i, \mathbf{x})}{\partial \mathbf{x}_i} = -\frac{\partial \mathcal{S}(\mathbf{x}_i, \mathbf{x})}{\partial \mathbf{x}} = \frac{\mathbf{I}}{d_i} - \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^3}$ . Hence  $\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{\mathbf{x}_i - \mathbf{x}}{d_i} \right) = \frac{\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}}{d_i^{(\xi)}}$ . Together with (100) and (101), the result of  $\mathcal{D}_i$  directly follows.

### APPENDIX C

The term  $\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^2} \right)$  is a symmetric matrix that involves expectations of nonlinear terms. We approximate each element in the matrix by using Taylor expansion at  $(\mathbf{x}_i^{(\xi)}, \mathbf{x}^{(\xi)})$ . Specifically,

$$\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^2} \right) = \mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \begin{bmatrix} \frac{(\mathbf{x}_i - \mathbf{x})^2}{d_i^2} & \frac{(\mathbf{x}_i - \mathbf{x})(y_i - y)}{d_i^2} \\ \frac{(\mathbf{x}_i - \mathbf{x})(y_i - y)}{d_i^2} & \frac{(y_i - y)^2}{d_i^2} \end{bmatrix}. \quad (103)$$

By taking the first diagonal element as an example and using results in [38], we can obtain

$$\begin{aligned} s(x_i, x) = \frac{(x_i - x)^2}{d_i^2} &\approx \frac{(x_i^{(\xi)} - x^{(\xi)})^2}{(d_i^{(\xi)})^2} + \nabla_{x_i} s(x_i^{(\xi)}, x^{(\xi)}) (x_i - x_i^{(\xi)}) + \nabla_x s(x_i^{(\xi)}, x^{(\xi)}) (x - x^{(\xi)}) \\ &\quad + \nabla_{y_i} s(x_i^{(\xi)}, x^{(\xi)}) (x_i - x_i^{(\xi)}) + \nabla_y s(x_i^{(\xi)}, x^{(\xi)}) (x - x^{(\xi)}), \end{aligned} \quad (104)$$

where  $\nabla_{x_i} s(x_i^{(\xi)}, x^{(\xi)})$ ,  $\nabla_x s(x_i^{(\xi)}, x^{(\xi)})$  and  $\nabla_{y_i} s(x_i^{(\xi)}, x^{(\xi)})$  and  $\nabla_y s(x_i^{(\xi)}, x^{(\xi)})$  are respectively the gradients at the point  $(x^{(\xi)}, y^{(\xi)})$  and  $(x_i^{(\xi)}, y_i^{(\xi)})$ . Hence we can obtain

$$\mathbb{E}_{q^{(\xi)}(\mathbf{x}_i, \mathbf{x})} \left( \frac{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T}{d_i^2} \right) \approx \frac{(\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)}) (\mathbf{x}_i^{(\xi)} - \mathbf{x}^{(\xi)})^T}{(d_i^{(\xi)})^2}. \quad (105)$$

### REFERENCES

- [1] B. Xue, L. Zhang, Y. Yu, and W. Zhu, "Locating the nodes from incomplete Euclidean distance matrix using Bayesian learning," *IEEE Access*, pp. 1–1, 2019.

- [2] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, “Efficient multisensor localization for the Internet of Things: Exploring a new class of scalable localization algorithms,” *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 153–167, Sept 2018.
- [3] Y. Sun, D. K. C. Ho, and Q. Wan, “Solution and analysis of TDOA localization of a near or distant source in closed-form,” *IEEE Transactions on Signal Processing*, pp. 1–1, 2018.
- [4] L. Wang and M. Zawodniok, “New theoretical limit analysis of LoS and RSS based positioning methods for Ricean fading channel in RF systems,” *IEEE Transactions on Mobile Computing*, pp. 1–1, 2018.
- [5] W. Wang, G. Wang, J. Zhang, and Y. Li, “Robust weighted least squares method for TOA-based localization under mixed LOS/NLOS conditions,” *IEEE Communications Letters*, vol. 21, no. 10, pp. 2226–2229, Oct 2017.
- [6] S. Brs, M. Izadi, C. Silvestre, A. Sanyal, and P. Oliveira, “Nonlinear observer for 3d rigid body motion estimation using Doppler measurements,” *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3580–3585, 2016.
- [7] Y. Li, X. Wang, and Z. Ding, “Multi-target position and velocity estimation using ofdm communication signals,” *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 1160–1174, 2020.
- [8] T. Sakamoto, A. Matsuoka, and H. Yomo, “Estimation of doppler velocities from sub-Nyquist ultra-wideband radar measurements,” *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8557–8565, 2016.
- [9] C. Tremblay-Darveau, R. Williams, P. S. Sheeran, L. Milot, M. Bruce, and P. N. Burns, “Concepts and tradeoffs in velocity estimation with plane-wave contrast-enhanced Doppler,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 11, pp. 1890–1905, 2016.
- [10] Y. Wang, Y. Wu, and Y. Shen, “Joint spatiotemporal multipath mitigation in large-scale array localization,” *IEEE Transactions on Signal Processing*, vol. 67, no. 3, pp. 783–797, 2019.
- [11] C. Geng, X. Yuan, and H. Huang, “Exploiting channel correlations for nlos toa localization with multivariate gaussian mixture models,” *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 70–73, 2020.
- [12] Y. Zou and H. Liu, “Semidefinite programming methods for alleviating clock synchroniza-

- tion bias and sensor position errors in tdoa localization,” *IEEE Signal Processing Letters*, vol. 27, pp. 241–245, 2020.
- [13] J. Yan, X. Zhang, X. Luo, Y. Wang, C. Chen, and X. Guan, “Asynchronous localization with mobility prediction for underwater acoustic sensor networks,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2543–2556, March 2018.
- [14] G. Han, J. Jiang, C. Zhang, T. Q. Duong, M. Guizani, and G. K. Karagiannidis, “A survey on mobile anchor node assisted localization in wireless sensor networks,” *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 2220–2243, 2016.
- [15] G. Wang, S. Cai, Y. Li, and N. Ansari, “A bias-reduced nonlinear WLS method for TDOA/FDOA-based source localization,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8603–8615, Oct 2016.
- [16] Y. Du and P. Wei, “An explicit solution for target localization in noncoherent distributed MIMO radar systems,” *IEEE Signal Processing Letters*, vol. 21, no. 9, pp. 1093–1097, Sep. 2014.
- [17] K. C. Ho, X. Lu, and L. Kovavisaruch, “Source localization using TDOA and FDOA measurements in the presence of receiver location errors: Analysis and solution,” *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 684–696, 2007.
- [18] M. Sun and K. C. Ho, “An asymptotically efficient estimator for TDOA and FDOA positioning of multiple disjoint sources in the presence of sensor location uncertainties,” *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3434–3440, July 2011.
- [19] A. Vempaty, O. Ozdemir, K. Agrawal, H. Chen, and P. K. Varshney, “Localization in wireless sensor networks: Byzantines and mitigation techniques,” *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1495–1508, March 2013.
- [20] R. Garg, A. L. Varna, and M. Wu, “An efficient gradient descent approach to secure localization in resource constrained wireless sensor networks,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 717–730, April 2012.
- [21] P. Zhang, S. G. Nagarajan, and I. Nevat, “Secure location of Things (SLOT): Mitigating localization spoofing attacks in the Internet of Things,” *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2199–2206, Dec 2017.
- [22] J. T. Chiang, J. J. Haas, J. Choi, and Y. Hu, “Secure location verification using simultaneous multilateration,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 584–

591, February 2012.

- [23] O. Abumansoor and A. Boukerche, “A secure cooperative approach for nonline-of-sight location verification in VANET,” *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 275–285, Jan 2012.
- [24] P. Perazzo, F. B. Sorbelli, M. Conti, G. Dini, and C. M. Pinotti, “Drone path planning for secure positioning and secure position verification,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 9, pp. 2478–2493, 2017.
- [25] M. U. Aftab, Y. Munir, A. Oluwasanmi, Z. Qin, M. H. Aziz, Zakria, N. T. Son, and V. D. Tran, “A hybrid access control model with dynamic COI for secure localization of satellite and IoT-based vehicles,” *IEEE Access*, vol. 8, pp. 24 196–24 208, 2020.
- [26] X. Liu, S. Su, F. Han, Y. Liu, and Z. Pan, “A range-based secure localization algorithm for wireless sensor networks,” *IEEE Sensors Journal*, vol. 19, no. 2, pp. 785–796, 2019.
- [27] J. Y. Koh, I. Nevat, D. Leong, and W. Wong, “Geo-spatial location spoofing detection for Internet of Things,” *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 971–978, Dec 2016.
- [28] J. Du and Y. Wu, “Network-wide distributed carrier frequency offsets estimation and compensation via Belief Propagation,” *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5868–5877, Dec 2013.
- [29] L. Deng, P. Wei, Z. Zhang, and H. Zhang, “Doppler frequency shift based source localization in presence of sensor location errors,” *IEEE Access*, vol. 6, pp. 59 752–59 760, 2018.
- [30] J. S. Picard and A. J. Weiss, “Direct position determination sensitivity to NLOS propagation effects on Doppler-shift,” *IEEE Transactions on Signal Processing*, vol. 67, no. 14, pp. 3870–3881, July 2019.
- [31] J. Jiang, G. Wang, and K. C. Ho, “Sensor network-based rigid body localization via semi-definite relaxation using arrival time and Doppler measurements,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1011–1025, Feb 2019.
- [32] N. Nasios and A. G. Bors, “Variational learning for Gaussian mixture models,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 4, pp. 849–862, 2006.
- [33] J. Winn and C. M. Bishop, “Variational message passing,” *Journal of Machine Learning Research*, vol. 6, no. Apr, pp. 661–694, 2005.
- [34] C. W. Fox and S. J. Roberts, “A tutorial on variational Bayesian inference,” *Artificial*



- intelligence review*, vol. 38, no. 2, pp. 85–95, 2012.
- [35] Y. Li, S. Ma, G. Yang, and K. Wong, “Robust localization for mixed LOS/NLOS environments with anchor uncertainties,” *IEEE Transactions on Communications*, pp. 1–1, 2020.
- [36] J. Du and Y. Wu, “Distributed clock skew and offset estimation in wireless sensor networks: Asynchronous algorithm and convergence analysis,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5908–5917, November 2013.
- [37] S. Y. Park and A. K. Bera, “Maximum entropy autoregressive conditional heteroskedasticity model,” *Journal of Econometrics*, vol. 150, no. 2, pp. 219–230, 2009.
- [38] W. J. Vetter, “Matrix calculus operations and Taylor expansions,” *SIAM review*, vol. 15, no. 2, pp. 352–369, 1973.