

Article

# The Cramér–Rao Bounds and Sensor Selection for Nonlinear Systems with Uncertain Observations

Zhiguo Wang , Xiaojing Shen \* , Ping Wang and Yunmin Zhu

School of Mathematics, Sichuan University, Chengdu 610064, China; wangzg315@126.com (Z.W.); yidianyuanzhidian@163.com (P.W.); ymzhu@scu.edu.cn (Y.Z.)

\* Correspondence: xiao23332@163.com; Tel.: +86-028-85412720

Received: 19 March 2018; Accepted: 2 April 2018; Published: 5 April 2018



**Abstract:** This paper considers the problems of the posterior Cramér–Rao bound and sensor selection for multi-sensor nonlinear systems with uncertain observations. In order to effectively overcome the difficulties caused by uncertainty, we investigate two methods to derive the posterior Cramér–Rao bound. The first method is based on the recursive formula of the Cramér–Rao bound and the Gaussian mixture model. Nevertheless, it needs to compute a complex integral based on the joint probability density function of the sensor measurements and the target state. The computation burden of this method is relatively high, especially in large sensor networks. Inspired by the idea of the expectation maximization algorithm, the second method is to introduce some 0–1 latent variables to deal with the Gaussian mixture model. Since the regular condition of the posterior Cramér–Rao bound is unsatisfied for the discrete uncertain system, we use some continuous variables to approximate the discrete latent variables. Then, a new Cramér–Rao bound can be achieved by a limiting process of the Cramér–Rao bound of the continuous system. It avoids the complex integral, which can reduce the computation burden. Based on the new posterior Cramér–Rao bound, the optimal solution of the sensor selection problem can be derived analytically. Thus, it can be used to deal with the sensor selection of a large-scale sensor networks. Two typical numerical examples verify the effectiveness of the proposed methods.

**Keywords:** Cramér–Rao bound; sensor selection; uncertain measurement; target tracking

## 1. Introduction

In practical problems, we always encounter some sensors have the uncertain measurement subjected to random interference, natural interruptions or sensor failures. Using the mode parameters without considering the uncertainty is unavailable, and there are a lot of researchers that have studied the state estimation with uncertain measurement, such as [1–4]. In this paper, we consider the uncertainty caused by occlusions, i.e., the sensors may not be able to observe the target when blocked by some obstacles [5]. For the linear dynamic systems involving uncertainty in [6,7], the authors use a Kalman filter to track the target. However, it is difficult to obtain the optimal estimation for a nonlinear uncertain dynamic system, but we are particular interested in measuring their efficiency. For this purpose, it is natural to compare a lower bound of the estimation error, which gives an indication of performance limitations. Moreover, it can be used to determine whether imposed performance requirements are realistic or not.

The most popular lower bound is the well-known Cramér–Rao bound (CRB). In time-invariant statistical models, the estimated parameter vector is usually assumed to be real-valued (non-random). The lower bound is given by the inverse of the Fisher information matrix. When we deal with the time-varying systems, the estimated parameter vector is modeled randomly. A lower bound that is analogous to the CRB for random parameters is derived in [8], and this bound is also known as the

Van Trees version of the CRB, or referred to as posterior CRB (PCRB). In fact, the underlying static random system needs to satisfy the regularity condition, which is absolute integrability of the first two derivatives of all related probability density functions. The first derivation of a sequential PCRB version applicable to discrete-time dynamic system filtering is done in [9] and then extended in [10–12]. The most general form of sequential PCRB for discrete-time nonlinear systems is presented in [13]. Together with the original static form of the CRB, these results serve as a basis for a large number of applications [14–16].

Most of the papers on PCRB are obtained without considering the uncertainty in the dynamic systems. When the sensors have uncertain measurements, we need to consider the influence of the uncertainty [17,18]. The CRB is presented in [19,20] to target tracking with detection probability smaller than one. If the uncertain measurement is prone to discretely-distributed faults, a Cramér–Rao-type bound is shown in [21]. Actually, the authors in [22,23] have considered uncertainty as the mixed Gaussian probabilistic model, where the sensor observation is assumed to contain only noise if the sensor cannot sense the target. Therefore, we hope to derive a recursive PCRB based on the uncertain model of the Gaussian mixture distribution.

Since the PCRB needs to compute the Fisher information, which is obtained by the derivatives of the log likelihood function of the Gaussian mixture model, and it is much more difficult than the case of a single Gaussian distribution. The reason is that the presence of the summation that occurs inside of the logarithm, and the PCRB of the Gaussian mixture model needs to compute the complex integral, which is with respect to the joint probability density function of the sensor measurements and the target state. These reasons motivate us to research another approach to derive the PCRB.

In large wireless sensor networks (WSNs), sensors are battery-powered devices with limited signal processing capabilities [24,25]. In such situations, it is inefficient to utilize all the sensors including the uninformative ones, which is hardly helpful to the tracking task but still consumes resources. This issue has been researched and shown via the development of sensor selection schemes, whose goal is to select the best non-redundant set of sensors for the tracking task while satisfying the resource constraints [26,27]. The previous research [28,29] on sensor selection assumes that the target tracking process does not have any interruptions. As the sensor observations are quite uncertain, we need to consider the sensor selection based on the proposed PCRB.

In this paper, we use two methods to derive the PCRB to effectively overcome the difficulties caused by uncertainty. The first method is based on the recursive formula of the Cramér–Rao bound and the Gaussian mixture model. Nevertheless, it needs to compute a complex integral based on the joint probability density function of the sensor measurements and the target state, which leads to the computation burden of this method being relatively high, especially in large sensor networks, so that it is not better using this PCRB as a measure criteria of the sensor selection. In order to reduce the computation burden and deal with the sensor selection of a large-scale sensor networks, our contributions are as follows:

- Inspired by the idea of the expectation maximization algorithm, we introduce some 0–1 latent variables to treat the Gaussian mixture model. Since the regular condition of the posterior Cramér–Rao bound is unsatisfied in the discrete uncertain system, we use some continuous variables to approximate the discrete latent variables, then a new Cramér–Rao bound can be achieved by a limiting process of the Cramér–Rao bound of the continuous system. The Cramér–Rao bound avoids the complex integral with a less computation burden.
- Based on the proposed posterior Cramér–Rao bound, the sensor selection problems for the nonlinear uncertain dynamic system can be efficiently solved, and the optimal solution of the sensor selection problem can be derived analytically. Thus, it can be used to deal with the sensor selection for the large-scale sensor networks.

The remainder of this paper is organized as follows. The system uncertain model is defined and the problem is formulated in Section 2. The PCRB for the dynamic system with uncertain observations is detailed and justified in Section 3. The optimal sensor selection with uncertain observations is shown

in Section 4. Two numerical examples are presented in Section 5. Finally, the conclusions are offered in the final section.

## 2. Problem Formulation

Consider the  $L$ -sensor nonlinear dynamic systems with the uncertain observations [5,30],

$$\begin{aligned} \mathbf{x}_k &= f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k, & (1) \\ \mathbf{y}_k^i &= \begin{cases} h_k^i(\mathbf{x}_k) + \mathbf{v}_k^i & \text{with probability } p_k^i, \\ \mathbf{v}_k^i & \text{with probability } 1 - p_k^i, \end{cases} & (2) \\ & i = 1, 2, \dots, L, \end{aligned}$$

where  $p_k^i$  is the sensing probability of sensor  $i$ ,  $\mathbf{x}_k$  is the state of system at time  $k$ ,  $\mathbf{y}_k^i$  is the measurement at the  $i$ th sensor,  $i = 1, \dots, L$ ,  $f_k(\mathbf{x}_{k-1})$  is the nonlinear state function, and  $h_k^i(\mathbf{x}_k)$  is the nonlinear measurement function of  $\mathbf{x}_k$  at the  $i$ th sensor.  $\mathbf{w}_k$  and  $\mathbf{v}_k^i$  are the state noise and the measurement noise, respectively, and they are mutually independent.  $\mathbf{v}_k^i$  is assumed to be independent across time steps and across sensors. The measurement information of the  $i$ th sensor is denoted by  $\mathbf{Y}_k^i = \{\mathbf{y}_1^i, \mathbf{y}_2^i, \dots, \mathbf{y}_k^i\}$ .

Assume that  $\mathbf{w}_k$  and  $\mathbf{v}_k^i$  are white Gaussian noise with  $\mathcal{N}(0, \mathbf{Q}_k)$  and  $\mathcal{N}(0, \mathbf{R}_k^i)$ ,  $i = 1, \dots, L$ , respectively, where,  $\mathbf{Q}_k$  and  $\mathbf{R}_k^i$  are the corresponding covariance matrices. We also assume that the initial state  $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{\Sigma}_0)$ , and, if  $\mathbf{x}_k$  is given, then the measurement  $\mathbf{y}_k^i$  follows the Gaussian distribution  $\mathcal{N}(h_k^i(\mathbf{x}_k), \mathbf{R}_k^i)$  with probability  $p_k^i$ , and follows the Gaussian distribution  $\mathcal{N}(0, \mathbf{R}_k^i)$  with probability  $1 - p_k^i$ , i.e.,

$$p(\mathbf{y}_k^i | \mathbf{x}_k) = p_k^i \mathcal{N}(h_k^i(\mathbf{x}_k), \mathbf{R}_k^i) + (1 - p_k^i) \mathcal{N}(0, \mathbf{R}_k^i). \quad (3)$$

Obviously, the conditional probability density function is a Gaussian mixture distribution, which is hard to calculate the PCRFB. This difficult problem motivates us to introduce a hidden state variable, which draws lessons from the idea of the expectation maximization (EM) algorithm [31].

We introduce the 0–1 hidden state variables  $I_k^i, i = 1, 2, \dots, L$ , which indicate whether the dynamic system has uncertainty. In other words, if  $I_k^i = 1$ , then  $\mathbf{y}_k^i = h_k^i(\mathbf{x}_k) + \mathbf{v}_k^i$ , and  $I_k^i = 0$ , then  $\mathbf{y}_k^i = \mathbf{v}_k^i$ . Now, we transfer the nonlinear systems (1) and (2) as follows:

$$\begin{cases} \mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k, \\ I_k^i = 0 \cdot I_{k-1}^i + \tilde{w}_k^i, \end{cases} \quad (4)$$

$$\mathbf{y}_k^i = I_k^i \cdot h_k^i(\mathbf{x}_k) + \mathbf{v}_k^i, \quad i = 1, 2, \dots, L. \quad (5)$$

Then, the compact form for Equations (4) and (5) can be written as follows:

$$\check{\mathbf{x}}_k = \mathbf{F}_k(\check{\mathbf{x}}_{k-1}) + \mathbf{W}_k, \quad (6)$$

$$\mathbf{y}_k^i = I_k^i \cdot h_k^i(\mathbf{x}_k) + \mathbf{v}_k^i, \quad i = 1, 2, \dots, L, \quad (7)$$

where  $\check{\mathbf{x}}_k = [\mathbf{x}_k^T \mathbf{I}_k^T]^T$ ,  $\mathbf{W}_k = [\mathbf{w}_k^T \tilde{w}_k^1, \dots, \tilde{w}_k^L]^T$ ,  $\mathbf{F}_k(\mathbf{X}_k) = [f_k^T(\mathbf{x}_k) \ 0]^T$ ,  $\mathbf{I}_k = [I_k^1, \dots, I_k^L]^T$ ,  $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ .  $\tilde{w}_k^i \sim \mathcal{B}(1, p_k^i)$ , which means a Bernoulli distribution with probability parameter  $p_k^i$ , if  $P(\tilde{w}_k^i = 1) = p_k^i$  and  $P(\tilde{w}_k^i = 0) = 1 - p_k^i$ . The process noise is independent of the uncertainty. Then, we assume  $\mathbf{w}_k$  and  $\tilde{w}_k^i, i = 1, 2, \dots, L$  are mutually independent.

Since the PCRFB is an important criterion of sensor selection, we derive two PCRFBs of the uncertain dynamic systems (1), (2), (6) and (7) in Sections 3 and 4, respectively. The former is accurate, but it is difficult to be computed. Thus, the latter is derived by introduced some hidden state variables, which avoids the complex integral and can reduce the computation burden. Finally, based on the second PCRFB, we hope to obtain the analytically optimal solution of the sensor selection problem, so that it can be applied to the large-scale sensor selection problem for the uncertain dynamic systems.

### 3. The Posterior Cramér–Rao Bound with Uncertain Observations

In this section, we mainly discuss two methods to calculate the PCRB of multiple sensors. The first method is based on the nonlinear dynamic system with uncertain observations (1) and (2) and Gaussian mixture model [5,13,15,32]. The other approach is based on the nonlinear dynamic system (6) and (7) motivated by the EM algorithm [33].

Let  $\theta$  be a  $r$ -dimensional estimated random parameter,  $\mathbf{z}$  represents a vector of measured data, let  $p(\mathbf{z}, \theta)$  be the joint probability density of the pair  $(\mathbf{z}, \theta)$ , and let  $g(\mathbf{z})$  be a function of  $\mathbf{z}$ , which is an estimate of  $\theta$ . Let  $\Delta$  and  $\nabla$  be operators of the first and second-order partial derivatives, respectively,

$$\begin{aligned} \nabla_{\eta} &= \left[ \frac{\partial}{\partial \eta_1}, \dots, \frac{\partial}{\partial \eta_L} \right], \\ \Delta_{\eta}^{\xi} &= \nabla_{\eta} \nabla_{\xi}^T. \end{aligned}$$

The PCRB on the estimate error has the form

$$\mathbf{P} = E \left\{ [g(\mathbf{z}) - \theta][g(\mathbf{z}) - \theta]^T \right\} \geq \mathbf{J}^{-1}, \tag{8}$$

where  $\mathbf{J} = E \left\{ -\Delta_{\Theta}^{\Theta} \log p_{\mathbf{z},\theta}(\mathbf{Z}, \Theta) \right\}$  is the (Fisher) information matrix denoted by Van Trees [8]. For example, if the posterior distribution of  $\theta$  conditioned on  $\mathbf{z}$  is Gaussian with mean  $\bar{\theta}_{\mathbf{z}}$  and a covariance matrix  $\Sigma_{\mathbf{z}}$ . Then, the information matrix (8) reads  $\mathbf{J} = E \{ \Sigma_{\mathbf{z}}^{-1} \}$ .

Assume now that the parameter  $\theta$  is decomposed into two parts as  $\theta = [\theta_{\alpha}^T, \theta_{\beta}^T]^T$ , and the information matrix  $\mathbf{J}$  is correspondingly divided into blocks

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\alpha\alpha} & \mathbf{J}_{\alpha\beta} \\ \mathbf{J}_{\beta\alpha} & \mathbf{J}_{\beta\beta} \end{bmatrix}.$$

Then, it can be easily shown that the covariance of estimation of  $\theta_{\beta}$  is lower bounded by the right-lower block of  $\mathbf{J}^{-1}$ , i.e.,

$$\begin{aligned} \mathbf{P}_{\beta} &= E \{ [g_{\beta}(\mathbf{x}) - \theta_{\beta}][g_{\beta}(\mathbf{x}) - \theta_{\beta}]^T \} \\ &\geq [\mathbf{J}_{\beta\beta} - \mathbf{J}_{\beta\alpha} \mathbf{J}_{\alpha\alpha}^{-1} \mathbf{J}_{\alpha\beta}]^{-1}, \end{aligned}$$

where we assume that  $\mathbf{J}_{\alpha\alpha}^{-1}$  exists. Denoted  $\mathbf{J}_{\beta} = \mathbf{J}_{\beta\beta} - \mathbf{J}_{\beta\alpha} \mathbf{J}_{\alpha\alpha}^{-1} \mathbf{J}_{\alpha\beta}$ , which is called the information submatrix for  $\beta$ .

Now, for nonlinear dynamic systems with uncertain observations (1) and (2), the following proposition gives a method to compute the information submatrix  $\mathbf{J}_k$  recursively.

**Proposition 1.** *The Fisher information submatrix  $\mathbf{J}_k$  for the estimating state vectors  $\{\mathbf{x}_k\}$  obeys the recursion:*

$$\mathbf{J}_{k+1} = \mathbf{D}_k^{22} - \mathbf{D}_k^{21} (\mathbf{J}_k + \mathbf{D}_k^{11})^{-1} \mathbf{D}_k^{12}, \tag{9}$$

$$\mathbf{J}_0 = E[-\Delta_{\mathbf{x}_0}^{\mathbf{x}_0} \log p(\mathbf{x}_0)], \tag{10}$$

with

$$\mathbf{D}_k^{11} = E[\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k) \mathbf{Q}_k^{-1} [\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k)]^T], \tag{11}$$

$$\mathbf{D}_k^{12} = -E[\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k) \mathbf{Q}_k^{-1}], \tag{12}$$

$$\mathbf{D}_k^{21} = (\mathbf{D}_k^{12})^T, \tag{13}$$

$$\mathbf{D}_k^{22} = \mathbf{Q}_k^{-1} + \sum_{i=1}^L E \{ -\Delta_{\mathbf{x}_{k+1}^i}^{\mathbf{x}_{k+1}^i} \log p(\mathbf{y}_{k+1}^i | \mathbf{x}_{k+1}) \}, \tag{14}$$

where

$$p(\mathbf{y}_{k+1}^i | \mathbf{x}_{k+1}) = p_k^i \mathcal{N}(h_k^i(\mathbf{x}_{k+1}), \mathbf{R}_{k+1}^i) + (1 - p_k^i) \mathcal{N}(0, \mathbf{R}_{k+1}^i). \quad (15)$$

**Proof.** Equations (1) and (2) together with  $p(\mathbf{x}_0)$  determine the joint probability distribution of  $\mathbf{X}_k = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k]$  and  $\mathbf{Y}_k = [\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k]$ , where  $\mathbf{y}_k = (\mathbf{y}_k^1, \mathbf{y}_k^2, \dots, \mathbf{y}_k^L)^T$ ,

$$\begin{aligned} p(\mathbf{X}_k, \mathbf{Y}_k) &= p(\mathbf{X}_{k-1}, \mathbf{Y}_{k-1}) \cdot p(\mathbf{x}_k | \mathbf{X}_{k-1}, \mathbf{Y}_{k-1}) \cdot p(\mathbf{y}_k | \mathbf{X}_k, \mathbf{Y}_{k-1}) \\ &= p(\mathbf{X}_{k-1}, \mathbf{Y}_{k-1}) \cdot p(\mathbf{x}_k | \mathbf{x}_{k-1}) \cdot p(\mathbf{y}_k | \mathbf{x}_k) \\ &= p(\mathbf{X}_{k-1}, \mathbf{Y}_{k-1}) \cdot p(\mathbf{x}_k | \mathbf{x}_{k-1}) \cdot \prod_{i=1}^L p(\mathbf{y}_k^i | \mathbf{x}_k). \end{aligned} \quad (16)$$

The conditional probability densities  $p(\mathbf{x}_k | \mathbf{x}_{k-1})$  and  $p(\mathbf{y}_k^i | \mathbf{x}_k)$  can be calculated by Equations (1) and (2), respectively. Denote  $p_k = p(\mathbf{X}_k, \mathbf{Y}_k)$ , by Equation (16), we can obtain the formula about  $p_{k+1}$  as follows:

$$p_{k+1} = p_k \cdot p(\mathbf{x}_{k+1} | \mathbf{x}_k) \cdot \prod_{i=1}^L p(\mathbf{y}_{k+1}^i | \mathbf{x}_{k+1}). \quad (17)$$

Therefore,

$$\log p_{k+1} = \log p_k + \log p(\mathbf{x}_{k+1} | \mathbf{x}_k) + \sum_{i=1}^L \log p(\mathbf{y}_{k+1}^i | \mathbf{x}_{k+1}). \quad (18)$$

If we divide  $\mathbf{X}_k$  into  $\mathbf{X}_k = [\mathbf{X}_{k-1}^T, \mathbf{x}_k^T]^T$ , then

$$\mathbf{J}(\mathbf{X}_k) = E\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p_k\} = \begin{bmatrix} E\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_{k-1}} \log p_k\} & E\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_k} \log p_k\} \\ E\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k-1}} \log p_k\} & E\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p_k\} \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \\ \mathbf{B}_k^T & \mathbf{C}_k \end{bmatrix}. \quad (19)$$

The information submatrix  $\mathbf{J}_k$  for  $\mathbf{x}_k$  can be obtained as follows:

$$\mathbf{J}_k = \mathbf{C}_k - \mathbf{B}_k^T \mathbf{A}_k^{-1} \mathbf{B}_k. \quad (20)$$

Moreover, let  $\mathbf{X}_{k+1} = [\mathbf{X}_{k-1}^T, \mathbf{x}_k^T, \mathbf{x}_{k+1}^T]^T$ , then the posterior information matrix for  $\mathbf{X}_{k+1}$  can be written as the following block form by Equation (18),

$$\mathbf{J}(\mathbf{X}_{k+1}) = \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k & 0 \\ \mathbf{B}_k^T & \mathbf{C}_k + \mathbf{D}_k^{11} & \mathbf{D}_k^{12} \\ 0 & \mathbf{D}_k^{21} & \mathbf{D}_k^{22} \end{bmatrix}, \quad (21)$$

where 0 stands for zero blocks of appropriate dimensions, and  $\mathbf{D}_k^{11}, \mathbf{D}_k^{12}, \mathbf{D}_k^{22}$  are calculated as follows:

$$\begin{aligned} \mathbf{D}_k^{11} &= E\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_{k+1} | \mathbf{x}_k)\}, \\ \mathbf{D}_k^{12} &= E\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k+1}} \log p(\mathbf{x}_{k+1} | \mathbf{x}_k)\} = (\mathbf{D}_k^{21})^T, \\ \mathbf{D}_k^{22} &= \sum_{i=1}^L E\{-\Delta_{\mathbf{x}_{k+1}}^{\mathbf{x}_{k+1}} \log p(\mathbf{y}_{k+1}^i | \mathbf{x}_{k+1})\}. \end{aligned}$$

Then, the information submatrix  $\mathbf{J}_{k+1}$  for  $\mathbf{x}_{k+1}$  can be computed as

$$\begin{aligned} \mathbf{J}_{k+1} &= \mathbf{D}_k^{22} - \begin{bmatrix} 0 & \mathbf{D}_k^{21} \end{bmatrix} \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \\ \mathbf{B}_k^T & \mathbf{C}_k + \mathbf{D}_k^{11} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{D}_k^{12} \end{bmatrix} \\ &= \mathbf{D}_k^{22} - \mathbf{D}_k^{21} [\mathbf{C}_k + \mathbf{D}_k^{11} - \mathbf{B}_k^T \mathbf{A}_k^{-1} \mathbf{B}_k]^{-1} \mathbf{D}_k^{12}. \end{aligned} \quad (22)$$

Based on the definition of  $\mathbf{J}_k$  in (20), we can obtain the desired recursion (9). Since the state noise and the measurement noise are Gaussian with zero mean and invertible covariance matrices  $\mathbf{Q}_k$  and  $\mathbf{R}_k^i$ ,  $i = 1, \dots, L$ , respectively. Moreover, the dynamic systems have the uncertain observations. From these assumptions and Equation (3), it follows that

$$\begin{aligned} -\log p(\mathbf{x}_{k+1}|\mathbf{x}_k) &= c_1 + \frac{1}{2} [\mathbf{x}_{k+1} - f_k(\mathbf{x}_k)]^T \mathbf{Q}_k^{-1} [\mathbf{x}_{k+1} - f_k(\mathbf{x}_k)] \\ -\log p(\mathbf{y}_{k+1}^i|\mathbf{x}_{k+1}) &= -\log \left( p_k \mathcal{N}(h_{k+1}^i(\mathbf{x}_{k+1}), \mathbf{R}_{k+1}^i) + (1 - p_k) \mathcal{N}(0, \mathbf{R}_{k+1}^i) \right), \end{aligned}$$

where  $c_1$  is a constant. Therefore,  $\mathbf{D}_k^{11}$ ,  $\mathbf{D}_k^{12}$ ,  $\mathbf{D}_k^{22}$  can be simplified to (11)–(14).  $\square$

From Equations (14) and (15), we see that the appearance of the summation inside of the logarithm, and the computation of  $\mathbf{D}_{22}^k$  is related to the joint probability density function of the sensor measurements  $\mathbf{y}_{k+1}$  and the target state  $\mathbf{x}_{k+1}$ , then  $\mathbf{D}_{22}^k$  is not easy to calculate. These reasons motivate us to study another approach to derive the PCRB.

Based on the equivalence between the systems (1)–(2) and (6)–(7), we can derive PCRB for the dynamic systems (6) and (7) by introducing a hidden variable  $\mathbf{I}_k$ , and the new PCRB may be easier to compute. Since the second derivation for the discrete augmented variable  $\mathbf{I}_k$  do not exist, then we bring in a continuous random variable  $\tilde{\mathbf{I}}_k$  to approximate the 0–1 variable  $\mathbf{I}_k$ . The augmented state vector  $\check{\mathbf{x}}_k = [\mathbf{x}_k, \mathbf{I}_k]^T$  has changed into  $\check{\mathbf{x}}_k = [\mathbf{x}_k, \tilde{\mathbf{I}}_k]^T$ . Therefore, the new system can be expressed as follows:

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k, \quad (23)$$

$$\tilde{\mathbf{I}}_k^i = 0 \cdot \tilde{\mathbf{I}}_{k-1}^i + \tilde{w}_k, \quad (24)$$

$$\mathbf{y}_k^i = \tilde{\mathbf{I}}_k^i \cdot h_k^i(\mathbf{x}_k) + v_k^i, i = 1, 2, \dots, L. \quad (25)$$

**Lemma 1** ([21]). *If  $\mathbf{I}_k^i \sim \mathcal{B}(1, p_k^i)$  and  $\tilde{\mathbf{I}}_k^i \sim p_k^i \mathcal{N}(1, \sigma^2) + (1 - p_k^i) \mathcal{N}(0, \sigma^2)$ ,  $i = 1, \dots, L$ , then the limit of  $\tilde{\mathbf{I}}_k$  is the state variable  $\mathbf{I}_k$  when  $\sigma \rightarrow 0$ , i.e.,  $\lim_{\sigma \rightarrow 0} \tilde{\mathbf{I}}_k = \mathbf{I}_k$ .*

Let  $\tilde{\mathbf{J}}_k$  represents the PCRB about the approximated augment vector  $\check{\mathbf{x}}_k$  of systems (23)–(25), respectively. Then, we can easily get the following conclusion:

**Lemma 2** ([21]). *Assume that  $\tilde{\mathbf{I}}_k^i \sim p_k^i \mathcal{N}(1, \sigma^2) + (1 - p_k^i) \mathcal{N}(0, \sigma^2)$ ,  $i = 1, \dots, L$ , then  $\mathbf{P}_k(\check{\mathbf{x}}_k) \geq \lim_{\sigma \rightarrow 0} \tilde{\mathbf{J}}_k^{-1} = \begin{bmatrix} \tilde{\mathbf{J}}_k^{-1} & 0 \\ 0 & 0 \end{bmatrix}$ , where  $\mathbf{P}_k(\check{\mathbf{x}}_k)$  denotes the estimation error covariance matrix about the vector  $\check{\mathbf{x}}_k$  and  $\tilde{\mathbf{J}}_k$  denotes the Fisher information submatrix about the vector  $\mathbf{x}_k$ .*

Based on Lemmas 1 and 2, for nonlinear dynamic system with the uncertain observations (6) and (7), it is easy to see that  $\tilde{\mathbf{J}}_k^{-1}$  can also represent a CRB for the estimation error covariance matrix of vector  $\mathbf{x}_k$ .

**Proposition 2.** At time  $k + 1$ , the Fisher information submatrix  $\bar{\mathbf{J}}_{k+1}$  of  $\mathbf{x}_{k+1}$  for the multi-sensor uncertain systems (6) and (7) is computed according to the following recursion:

$$\bar{\mathbf{J}}_{k+1} = \bar{\mathbf{D}}_k^{22} - \bar{\mathbf{D}}_k^{21}(\bar{\mathbf{J}}_k + \bar{\mathbf{D}}_k^{11})^{-1}\bar{\mathbf{D}}_k^{12}, \tag{26}$$

$$\bar{\mathbf{J}}_0 = \Sigma_0^{-1}, \tag{27}$$

with

$$\bar{\mathbf{D}}_k^{11} = E\{[\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k)]\mathbf{Q}_k^{-1}[\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k)]^T\}, \tag{28}$$

$$\bar{\mathbf{D}}_k^{12} = -E[\nabla_{\mathbf{x}_k} f_k^T(\mathbf{x}_k)]\mathbf{Q}_k^{-1}, \tag{29}$$

$$\bar{\mathbf{D}}_k^{21} = (\bar{\mathbf{D}}_k^{12})^T, \tag{30}$$

$$\bar{\mathbf{D}}_k^{22} = \mathbf{Q}_k^{-1} + \sum_{i=1}^L p_{k+1}^i E\{[\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]^T(\mathbf{R}_{k+1}^i)^{-1}[\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]\}. \tag{31}$$

**Proof.** According to Lemma 1 and the derivation of Proposition 1, the new augmented state vector  $\tilde{\mathbf{x}}_k$  has the following PCRB for systems (6) and (7):

$$\tilde{\mathbf{J}}_{k+1} = \tilde{\mathbf{D}}_k^{22} - \tilde{\mathbf{D}}_k^{21}(\tilde{\mathbf{J}}_k + \tilde{\mathbf{D}}_k^{11})^{-1}\tilde{\mathbf{D}}_k^{12}, \tag{32}$$

where  $\tilde{\mathbf{D}}_k^{11}$ ,  $\tilde{\mathbf{D}}_k^{12}$ ,  $\tilde{\mathbf{D}}_k^{22}$  are denoted as follows:

$$\tilde{\mathbf{D}}_k^{11} = E_{\tilde{\mathbf{x}}_k}\{-\Delta_{\tilde{\mathbf{x}}_k}^{\tilde{\mathbf{x}}_k} \log p(\tilde{\mathbf{x}}_{k+1}|\tilde{\mathbf{x}}_k)\},$$

$$\tilde{\mathbf{D}}_k^{12} = E_{\tilde{\mathbf{x}}_k}\{-\Delta_{\tilde{\mathbf{x}}_k}^{\tilde{\mathbf{x}}_{k+1}} \log p(\tilde{\mathbf{x}}_{k+1}|\tilde{\mathbf{x}}_k)\},$$

$$\tilde{\mathbf{D}}_k^{21} = (\tilde{\mathbf{D}}_k^{12})^T,$$

$$\tilde{\mathbf{D}}_k^{22} = E_{\tilde{\mathbf{x}}_{k+1}}\{-\Delta_{\tilde{\mathbf{x}}_{k+1}}^{\tilde{\mathbf{x}}_{k+1}} \log p(\tilde{\mathbf{x}}_{k+1}|\tilde{\mathbf{x}}_k)\} + \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}}\{-\Delta_{\tilde{\mathbf{x}}_{k+1}}^{\tilde{\mathbf{x}}_{k+1}} \log p(\mathbf{y}_{k+1}^i|\tilde{\mathbf{x}}_{k+1})\}.$$

In order to obtain the lower bound for  $\mathbf{x}_k$ , it is necessary for us to calculate the following probability densities, according to Equations (6) and (7),

$$\begin{aligned} -\log p(\tilde{\mathbf{x}}_{k+1}|\tilde{\mathbf{x}}_k) &= -\log p(\mathbf{x}_{k+1}|\mathbf{x}_k) \cdot p(\tilde{\mathbf{I}}_{k+1}) \\ &= c_3 + \frac{1}{2}[\mathbf{x}_{k+1} - f_k(\mathbf{x}_k)]^T \mathbf{Q}_k^{-1}[\mathbf{x}_{k+1} - f_k(\mathbf{x}_k)] - \log p(\tilde{\mathbf{I}}_{k+1}), \end{aligned} \tag{33}$$

where  $c_3$  is a constant, and the first equality follows from the independence and the second follows from (2). The another probability density is as follows:

$$-\log p(\mathbf{y}_{k+1}^i|\tilde{\mathbf{x}}_{k+1}) = c_4 + \frac{1}{2}[\mathbf{y}_{k+1}^i - \tilde{f}_{k+1}^i h_{k+1}^i(\mathbf{x}_{k+1})]^T (\mathbf{R}_{k+1}^i)^{-1}[\mathbf{y}_{k+1}^i - \tilde{f}_{k+1}^i h_{k+1}^i(\mathbf{x}_{k+1})], \tag{34}$$

where  $c_4$  is a constant. Since  $\tilde{\mathbf{x}}_k = [\mathbf{x}_k, \tilde{\mathbf{I}}_k]^T$ , we use Equations (33)–(34) and Lemma 1, and the suitable partitioned expressions for  $\tilde{\mathbf{D}}_k^{11}$ ,  $\tilde{\mathbf{D}}_k^{12}$ ,  $\tilde{\mathbf{D}}_k^{22}$  are obtained:

$$\tilde{\mathbf{D}}_k^{11} = \begin{bmatrix} \bar{\mathbf{D}}_k^{11} & 0 \\ 0 & 0 \end{bmatrix}, \tag{35}$$

$$\tilde{\mathbf{D}}_k^{12} = \begin{bmatrix} \bar{\mathbf{D}}_k^{12} & 0 \\ 0 & 0 \end{bmatrix}, \tag{36}$$

$$\tilde{\mathbf{D}}_k^{22} = \begin{bmatrix} \bar{\mathbf{D}}_k^{22} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}, \tag{37}$$

where  $\bar{\mathbf{D}}_k^{11}$ ,  $\bar{\mathbf{D}}_k^{12}$  are denoted in (28) and (29), while  $\mathbf{C}_{12}$ ,  $\mathbf{C}_{21}$  and  $\mathbf{C}_{22}$  are calculated as follows:

$$\begin{aligned} \mathbf{C}_{12} &= E_{\tilde{\mathbf{x}}_{k+1}} \left\{ -\Delta_{\tilde{\mathbf{x}}_{k+1}}^{\tilde{\mathbf{I}}_{k+1}^1} \log p(\tilde{\mathbf{x}}_{k+1} | \tilde{\mathbf{x}}_k) \right\} + \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ -\Delta_{\tilde{\mathbf{x}}_{k+1}}^{\tilde{\mathbf{I}}_{k+1}^i} \log p(\mathbf{y}_{k+1}^i | \tilde{\mathbf{x}}_{k+1}) \right\} \\ &= -\sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]^T (\mathbf{R}_{k+1}^i)^{-1} \mathbf{y}_{k+1}^i \right\} \\ &\quad + 2 \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ \tilde{\mathbf{I}}_{k+1}^i h_{k+1}^i(\mathbf{x}_{k+1}) (\mathbf{R}_{k+1}^i)^{-1} [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})] \right\} \\ &= (\mathbf{C}_{21})^T, \end{aligned}$$

$$\begin{aligned} \mathbf{C}_{22} &= \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ -\Delta_{\tilde{\mathbf{I}}_{k+1}^i}^{\tilde{\mathbf{I}}_{k+1}^i} \log p(\mathbf{y}_{k+1}^i | \tilde{\mathbf{x}}_{k+1}) \right\} + \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ -\Delta_{\tilde{\mathbf{I}}_{k+1}^i}^{\tilde{\mathbf{I}}_{k+1}^i} \log p(\tilde{\mathbf{I}}_{k+1}^i) \right\} \\ &= \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ (h_{k+1}^i(\mathbf{x}_{k+1}))^T (\mathbf{R}_{k+1}^i)^{-1} (h_{k+1}^i(\mathbf{x}_{k+1})) \right\} + \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ -\Delta_{\tilde{\mathbf{I}}_{k+1}^i}^{\tilde{\mathbf{I}}_{k+1}^i} \log p(\tilde{\mathbf{I}}_{k+1}^i) \right\}. \end{aligned}$$

If we divide  $\tilde{\mathbf{J}}_k$  as the following block matrix

$$\tilde{\mathbf{J}}_k = \begin{bmatrix} \tilde{\mathbf{J}}_k^{11} & \tilde{\mathbf{J}}_k^{12} \\ \tilde{\mathbf{J}}_k^{21} & \tilde{\mathbf{J}}_k^{22} \end{bmatrix}, \tag{38}$$

then according to (32) and (35)–(37), the value of  $\tilde{\mathbf{J}}_{k+1}$  is

$$\tilde{\mathbf{J}}_{k+1} = \begin{bmatrix} \bar{\mathbf{D}}_k^{22} + \bar{\mathbf{D}}_k^{21} (\bar{\mathbf{D}}_k^{11} + \tilde{\mathbf{J}}_k^{11} - \tilde{\mathbf{J}}_k^{12} (\tilde{\mathbf{J}}_k^{22})^{-1} \tilde{\mathbf{J}}_k^{21})^{-1} \bar{\mathbf{D}}_k^{12} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}. \tag{39}$$

Since the matrix  $\mathbf{C}_{22}$  is the function of  $\sigma$ , it is shown in [21] that

$$\lim_{\sigma \rightarrow 0} \tilde{\mathbf{J}}_{k+1}^{-1} = \begin{bmatrix} \left( \bar{\mathbf{D}}_k^{22} + \bar{\mathbf{D}}_k^{21} (\bar{\mathbf{D}}_k^{11} + \tilde{\mathbf{J}}_k^{11})^{-1} \bar{\mathbf{D}}_k^{12} \right)^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \tag{40}$$

Using Lemma 1, we can obtain (26), and the matrix  $\bar{\mathbf{D}}_k^{22}$  can be computed as

$$\begin{aligned} \bar{\mathbf{D}}_k^{22} &= \sum_{i=1}^L E_{\tilde{\mathbf{x}}_{k+1}} \left\{ \mathbf{I}_{k+1}^i \nabla_{\mathbf{x}_{k+1}} (h_{k+1}^i(\mathbf{x}_{k+1}))^T (\mathbf{R}_{k+1}^i)^{-1} \nabla_{\mathbf{x}_{k+1}} (h_{k+1}^i(\mathbf{x}_{k+1})) \mathbf{I}_{k+1}^i \right\} + \mathbf{Q}_k^{-1} \\ &= \mathbf{Q}_k^{-1} + \sum_{i=1}^L p_{k+1}^i E_{\tilde{\mathbf{x}}_{k+1}} \left\{ [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]^T (\mathbf{R}_{k+1}^i)^{-1} [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})] \right\}. \end{aligned}$$

□

**Remark 1.** Note that PCRBs derived in Propositions 1 and 2 have different forms. The first one is optimal. The second one is only approximately optimal with less computational burden. Since it may be approximated from above or below, which one is lower cannot be judged. The simulation in Section 6 shows that they are almost equal and the computational complexity of the approximate bound is less than that of the accurate bound.

**Remark 2.** In the case of  $p = 1$  and  $L = 1$ , the multi-sensor dynamic systems (1) and (2) has the certain observations. Obviously, the PCRB derived by the method in [13] is consistent with that derived in Proposition 2.

#### 4. Sensor Selection with Uncertain Observations

In large sensor networks, it is an important problem to manage the communication resources efficiently. The calculation of PCRB by Proposition 1 needs to use the joint probability density function of the sensor measurements and the target state, which leads to the computational burden being heavy, so that it is detrimental to be used as a measure criteria of the sensor selection. In this section, we consider the problem of sensor selection by Proposition 2.

For the nonlinear dynamic system at time  $k$ , assume that  $s$  sensors will be selected from  $L$  sensors by maximizing the Fisher information matrix, then they will send their measurements or local estimates to the fusion center. Finally, the fusion center makes the estimates for the state. In order to select the optimal  $s$  sensors, we need to introduce a selection vector  $\mathbf{s}_k = [s_k^1, \dots, s_k^L]^T \in \{0, 1\}^L$ . If the  $i$ th sensor is selected, let  $s_k^i = 1$ ; otherwise,  $s_k^i = 0$ ,  $i = 1, \dots, L$ . According to the derivation of the Fisher information matrix in Section 3, the selection vector modifies the log conditional probability density  $\log p(\mathbf{y}_k | \mathbf{x}_k)$  as [34]

$$\log \prod_{i=1}^L (p(\mathbf{y}_k^i | \mathbf{x}_k))^{s_k^i} = \sum_{i=1}^L s_k^i \log p(\mathbf{y}_k^i | \mathbf{x}_k). \quad (41)$$

In fact, the selected variable  $\mathbf{s}_k$  only has an effect on  $\bar{\mathbf{D}}_k^{22}$  of Proposition 2. Then,  $\bar{\mathbf{D}}_k^{22}$  can be written as

$$\bar{\mathbf{D}}_k^{22} = \mathbf{Q}_k^{-1} + \sum_{i=1}^L s_{k+1}^i p_{k+1}^i E \left\{ [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]^T (\mathbf{R}_{k+1}^i)^{-1} [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})] \right\}. \quad (42)$$

Therefore, the information matrix of  $\mathbf{x}_{k+1}$  is the function of the selected variable  $\mathbf{s}_k$ . Now, the sensor selection problem can be expressed as the following optimization problem:

$$\max_{\mathbf{s}_{k+1}} \text{tr}(\bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1})), \quad (43)$$

$$\text{s.t.} \quad \sum_{i=1}^L s_{k+1}^i = s, \quad (44)$$

$$s_{k+1}^i \in \{0, 1\}, \quad i = 1, \dots, L, \quad (45)$$

where “tr” means “trace”, which is the sum of squares of semiaxes lengths of the Fisher information matrix. “s.t.” means “subjected to”.

**Remark 3.** In fact, the objective function in (43) should be matrix  $\bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1})$ . Then, the problem (43)–(45) is a matrix optimization problem, which is considered in the sense that if  $\mathbf{s}_{k+1}^*$  is an optimal solution. Then, for an arbitrary feasible solution  $\mathbf{s}_{k+1}$ , we have  $\bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1}) \succeq \bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1}^*)$ , i.e.,  $\bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1}) - \bar{\mathbf{J}}_{k+1}(\mathbf{s}_{k+1}^*)$  is a positive semidefinite matrix. There are two reasons to choose trace function as the objective function. First, it is a linear function, which helps us to easily derive the optimal solution. Second, some researchers [26–28] have proved that it has many advantages to apply to sensor selection, such as, if the primal matrix optimization problem has an optimal solution and  $\bar{\mathbf{D}}_k^{12}$  in (29) is invertible, then the matrix optimization problem for sensor selection can be equivalently transformed to this convex optimization problem (43)–(45).

Let the information measure corresponding to the  $i$ -th sensor at  $k + 1$ -th time be denoted as

$$b_{k+1}^i = p_{k+1}^i \text{tr} \left( E \left\{ [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})]^T (\mathbf{R}_{k+1}^i)^{-1} [\nabla_{\mathbf{x}_{k+1}} h_{k+1}^i(\mathbf{x}_{k+1})] \right\} \right), \quad i = 1, \dots, L. \quad (46)$$

Let  $\{b_{k+1}^{r_1}, \dots, b_{k+1}^{r_L}\}$  denote  $\{b_{k+1}^1, \dots, b_{k+1}^L\}$  as rearrangement with descending order, i.e.,  $b_{k+1}^{r_1} \geq \dots \geq b_{k+1}^{r_L}$ . The optimal solution of the optimization problem (43)–(45) can be obtained by the following proposition.

**Proposition 3.** For multisensor nonlinear dynamic system with the uncertain observations (1) and (2), the optimal sensor selection scheme for the problem (43)–(45) is  $s_{k+1}^{r_1} = \dots = s_{k+1}^{r_s} = 1$  and  $s_{k+1}^{r_{s+1}} = \dots = s_{k+1}^{r_L} = 0$ .

**Proof.** Since  $\bar{\mathbf{D}}_k^{11}$  and  $\bar{\mathbf{D}}_k^{12}$  are not related to  $\mathbf{s}_k$ , based on Proposition 2, the optimization problem (43)–(45) can be equivalent to

$$\max_{\mathbf{s}_{k+1}} \sum_{i=1}^L b_{k+1}^i s_{k+1}^i, \quad (47)$$

$$\text{s.t. } \sum_{i=1}^L s_{k+1}^i = s, \quad (48)$$

$$s_{k+1}^i \in \{0, 1\}, i = 1, \dots, L, \quad (49)$$

where  $b_{k+1}^i$  is denoted by (46). According to  $b_{k+1}^{r_1} \geq \dots \geq b_{k+1}^{r_L}$ , and  $s_{k+1}^i, i = 1, \dots, L$  needs to satisfy (48) and (49), and we have

$$\sum_{i=1}^L b_{k+1}^i s_{k+1}^i \leq \sum_{i=1}^s b_{k+1}^{r_i}.$$

The equality holds with  $s_{k+1}^{r_1} = \dots = s_{k+1}^{r_s} = 1$  and  $s_{k+1}^{r_{s+1}} = \dots = s_{k+1}^{r_L} = 0$ . Thus, the optimal solution is got.  $\square$

## 5. Simulation

In this section, we provide two examples to compare the different PCRB by Proposition 1 with Proposition 2, and select the optimal sensors by Proposition 3.

*Example 1:* Consider an uncertain nonlinear dynamic system for the mobile robot. At time  $k$ , the mobile robot pose is described with the state vector  $\mathbf{x}_k = [x_k \ y_k \ \theta_k]$ , where  $x_k$  and  $y_k$  are the coordinates on a 2D plane relative to an external coordinate frame, and  $\theta_k$  is the heading angle. We use the control commands  $\mathbf{u}_k = [\Delta d_k, \Delta \theta_k]$  to determine the motion of the mobile robot, where  $\Delta d_k$  is the incremental distance robot (in meters) and  $\Delta \theta_k$  is the incremental change in heading angle (in degrees). The robot motion can be described as follows [35]:

$$\begin{aligned} f_{1,k} &= x_{k-1} + \Delta d_k \cos(\theta_{k-1} + \frac{1}{2} \Delta \theta_k), \\ f_{2,k} &= y_{k-1} + \Delta d_k \sin(\theta_{k-1} + \frac{1}{2} \Delta \theta_k), \\ f_{3,k} &= \theta_{k-1} + \Delta \theta_k, \end{aligned}$$

where  $\Delta d_k = 5, \Delta \theta_k = 5$ . The state equation is defined as  $f_k = [f_{1,k}, f_{2,k}, f_{3,k}]^T$ , and then the state model is

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k. \quad (50)$$

The measurement equation is

$$\mathbf{y}_k^i = \begin{cases} h^i(\mathbf{x}_k) + v_k^j, & \text{with probability } p_k^i, \\ v_k^j, & \text{with probability } 1 - p_k^i, \end{cases} \quad (51)$$

for  $i = 1, \dots, L, p_k^i = 0.8$ ,

where

$$h^i(\mathbf{x}_k) = \begin{bmatrix} \sqrt{(\mathbf{x}_k(1) - \mathbf{z}_k^i(1))^2 + (\mathbf{x}_k(2) - \mathbf{z}_k^i(2))^2} \\ \arctan\left(\frac{\mathbf{x}_k(2) - \mathbf{z}_k^i(2)}{\mathbf{x}_k(1) - \mathbf{z}_k^i(1)}\right) \end{bmatrix}.$$

$\mathbf{z}_k^i = [\mathbf{z}_k^i(1) \ \mathbf{z}_k^i(2)]^T$  is the position of the  $i$ th sensor. In the simulation, we consider the WSN shown in Figure 1, which has  $L = 6 \times 6 = 36$  sensors deployed in the area  $100 \times 100 \text{ m}^2$  [5]. The noise covariances are set as  $\mathbf{Q}_k = \text{diag}([0.1^2, 0.1^2, 3^2])$ ,  $\mathbf{R}_k^i = \text{diag}([1^2, 1^2])$ .

In the example, the initial state of the robot starts is  $[8, 8, 1]$  and the initial covariance matrix is  $\mathbf{P}_0 = \text{diag}([10, 10, 2])$  [35]. The sampling length is assigned to  $flag = 50$ . Here, the number of Monte Carlo (MC) simulation is  $MC = 200$ .

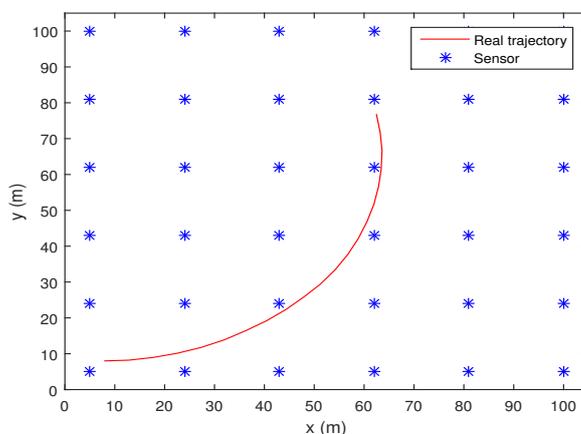


Figure 1. The trajectory of the mobile robot and the location of the  $L$  sensors.

The following simulation results include three parts: the first part is about the trajectory of the mobile robot and PCRB of the state estimation, the second part is about the average computation time, and the third part is about the PCRB with different sensing probability  $p$ .

- Figure 1 shows the trajectory of the mobile robot and the location of the  $L$  sensors. Figures 2 and 3 show that the PCRB of position along the  $x$ - and  $y$ -directions based on Proposition 1 and Proposition 2, respectively. From Figures 2 and 3, we can see the different PCRBs are shown to converge to the same values. However, the PCRB changes so much in the first seconds, and there are two possible reasons. First, the dynamic system is nonlinear. It may cause the algorithm to require some time to be convergent. Second, the initial variance may not be given better, such that it is far away from the convergence point.
- The average computation time of calculating PCRB based on Proposition 1 and Proposition 2 is presented in Figure 4. From Figure 4, obviously, when the number of the sensors increases, the computational complexity of Proposition 1 is much higher than that of Proposition 2, and the average computation time of PCRB by Proposition 2 increases slowly. The reason may be that the expression of PCRB based on Proposition 2 has a more concise form where the  $\bar{D}_k^{22}$  is easier to compute. Thus, Proposition 2 is more suitable for the sensor selection in the large-scale sensor networks.
- In Figure 5, the average PCRB of 20 time steps is plotted as a function of number of sensors. It shows that the PCRB obtained by Proposition 1 is the same as that based on Proposition 2. The larger  $p$  is, the smaller the number of required sensors. The reason may be that the sensors can take more observation information, when the sensor probability  $p$  is larger.

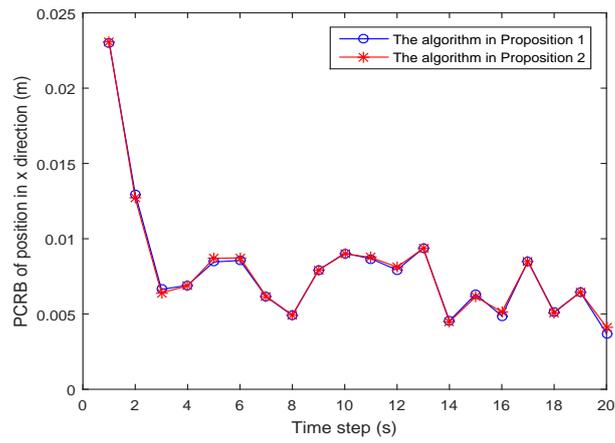


Figure 2. The PCRb of position in the x-direction is plotted as a function of time steps.

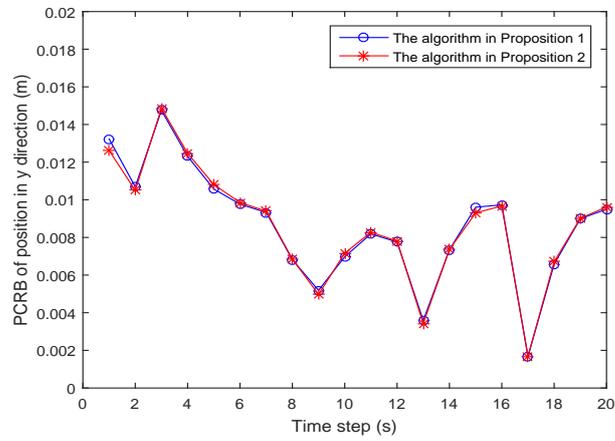


Figure 3. The PCRb of position in the y-direction is plotted as a function of time steps.

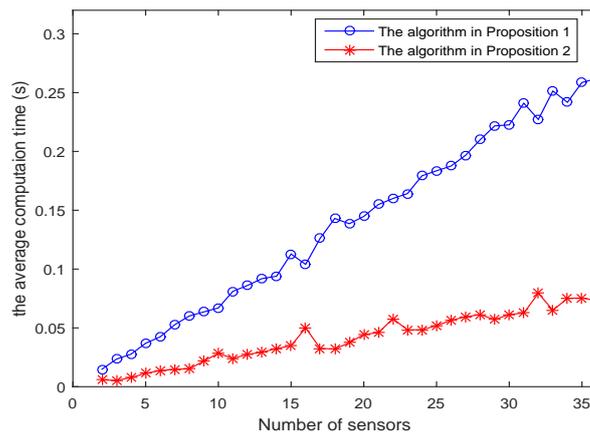
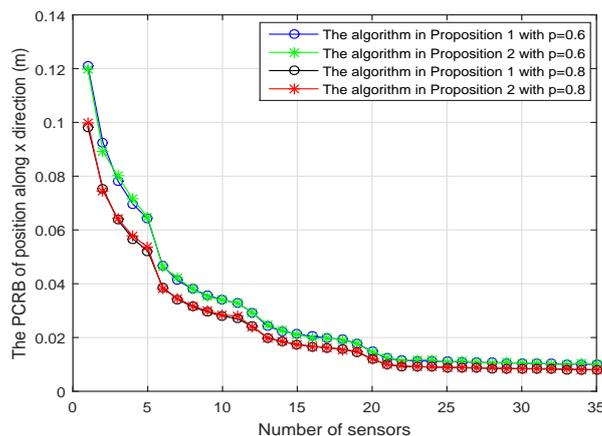


Figure 4. The average computation time of PCRb is plotted as function of number of sensors.



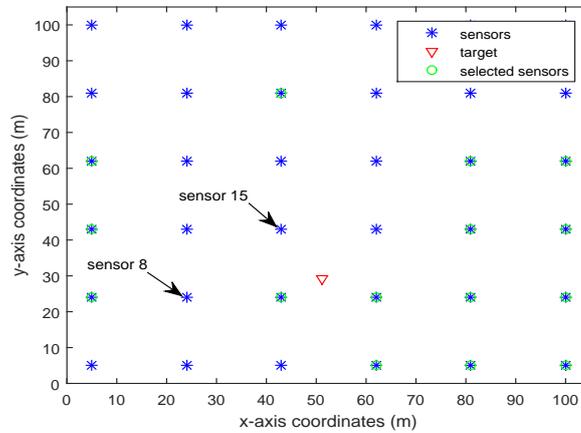
**Figure 5.** The average PCRB of 20 time steps is plotted as function of number of sensors with the different probability  $p$ .

*Example 2:* In order to manage the communication resources efficiently in large wireless sensor networks, we need to select some appropriate sensors. Thus, let us consider the above dynamic system (50) and (51) and the WSNs [35]. In general, if the sensors are close to the target, which may have higher sensing probabilities compared to other sensors in the WSN, then it is highly likely to select those sensors, owing to being both closer to target and higher sensing probability. Here, we consider a relatively difficult case that the sensors around the target have relatively low sensing probabilities. Then, we compare our algorithm in Proposition 3 with the recent two methods given in [5,28].

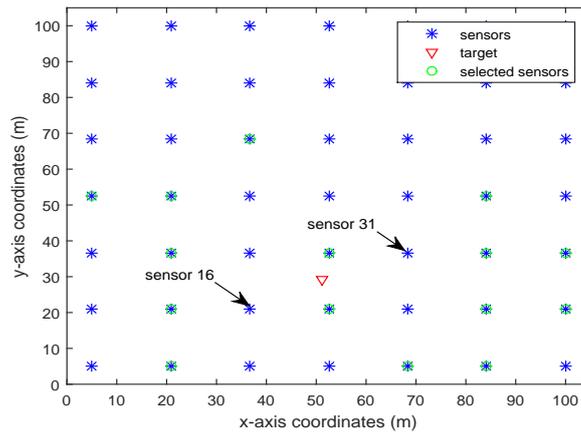
In this example, we present two cases with different number of sensors in WSN. Firstly, we consider  $L = 36$  and  $s_k = 15$ . Moreover, let  $p_k^i = 0.1$ ,  $i = 7, 8, 9, 10, 13, 14, 15, 16, 20, 21, 22, 23, 24$ , and the other sensing probabilities are between 0.8 and 1. Secondly, we also consider  $L = 49$  and  $s_k = 15$ . Moreover, let  $p_k^i = 0.1$ ,  $i = 14, \dots, 18, 22, \dots, 26, 30, \dots, 34$ , and the other sensing probabilities are between 0.8 and 1. The following simulation results contain three parts. The first part is about the sensor selection in the application of wireless sensor network, the second part is about the mean squared error based on the selected sensors, and the third part is about the computation time.

- Figures 6 and 7 present the location of  $L = 36$  and  $L = 49$  sensors, respectively. The target is showed at the time 10 s, and we use our algorithm in Proposition 3 to select the optimal  $s_k = 15$  sensors. When the uncertainty in the dynamic system is ignored, the recent method in [28] can be used to select the required sensors, and the results are shown in Figures 8 and 9. Comparing Figure 6 with Figure 8, some sensors are close to the target, such as sensor 8 and sensor 15, but they are not selected in Figure 6. In Figure 8, they are selected and the only closer sensors can be selected. The reason is that the sensing probability of sensor 8 and sensor 15 is very low, and they may be not given us much useful information, although they are close to the target. Comparing Figure 7 with Figure 9, it has a similar phenomenon, such as sensor 16 and 31 not being selected in Figure 7, but they are selected in Figure 9.
- In Figures 10 and 11, the mean squared errors of position in  $x$ - and  $y$ -directions are plotted for the algorithm given in Proposition 3 and the algorithms in [5,28]. It shows that our algorithm can derive the best performance. The reason is that our algorithm considers the influence of uncertain observation, and the optimal selected sensors are obtained. Although the algorithm in [5] considers the uncertain observation, it is difficult to obtain the optimal selected sensors, since it involves relaxing the variable  $\{0, 1\}$  to the interval  $[0, 1]$ . From Figures 12 and 13, we can also see that the proposed method also performs best in the case of  $L = 49$ , thus the performance of the new method is stable with the increase of the number of sensors.
- The computation times of obtaining PCRB are plotted in Figures 14 and 15 for the three algorithms, respectively. Figures 14 and 15 show that the computation time of the method in Proposition 3 is

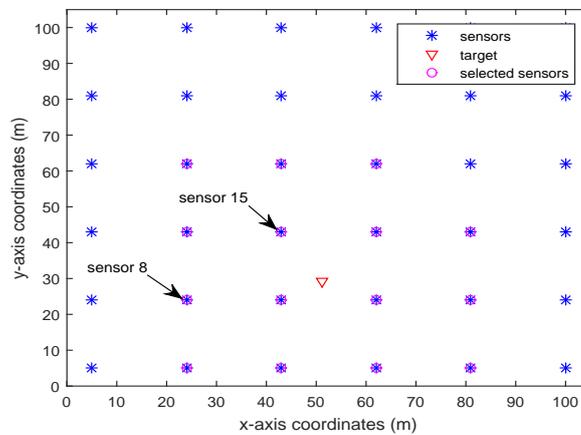
much smaller than that of the other two methods. The reason is that the method in Proposition 3 is an analytical solution. Therefore, the proposed algorithm in Proposition 3 is more suitable for the large sensor networks.



**Figure 6.**  $L = 36$  sensors placement and selected sensors  $s_k = 15$  based on the algorithm in Proposition 3.



**Figure 7.**  $L = 49$  sensors placement and selected sensors  $s_k = 15$  based on the algorithm in Proposition 3.



**Figure 8.**  $L = 36$  sensors placement and selected sensors  $s_k = 15$  based on the algorithm in [28].

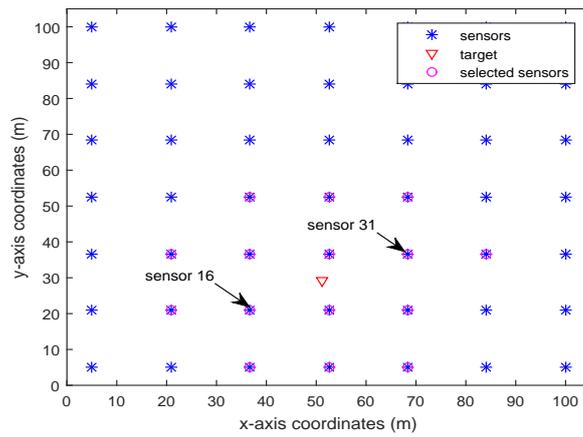


Figure 9.  $L = 49$  sensors placement and selected sensors  $s_k = 15$  based on the algorithm in [28].

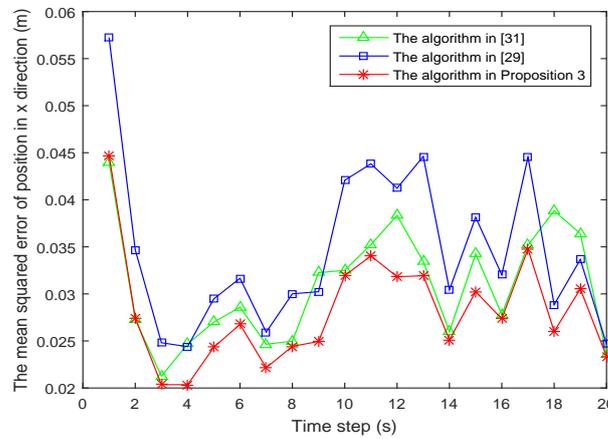


Figure 10. The mean squared error of position in the  $x$ -direction is plotted as function of time steps with  $L = 36$  sensors.

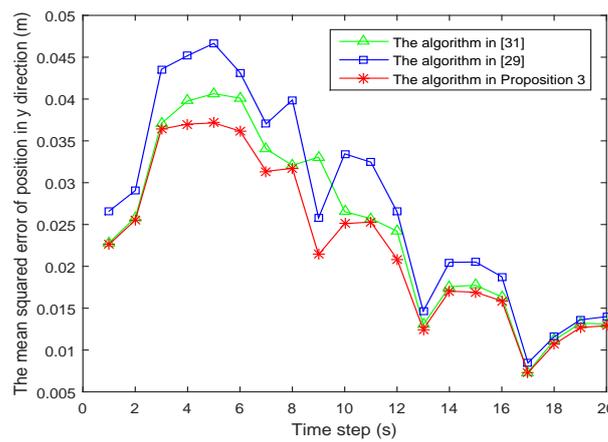
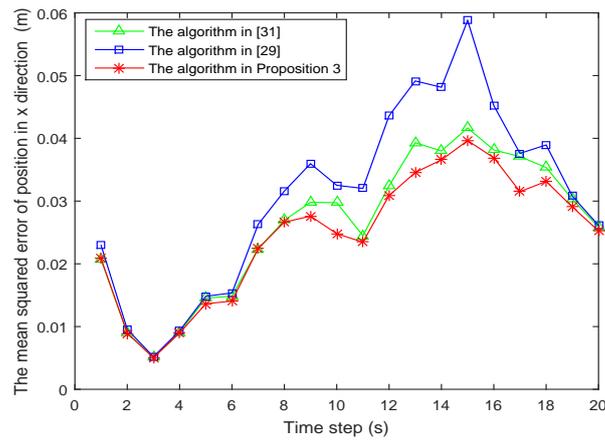
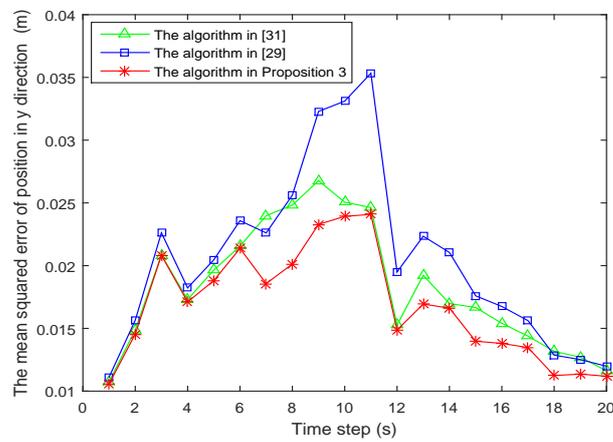


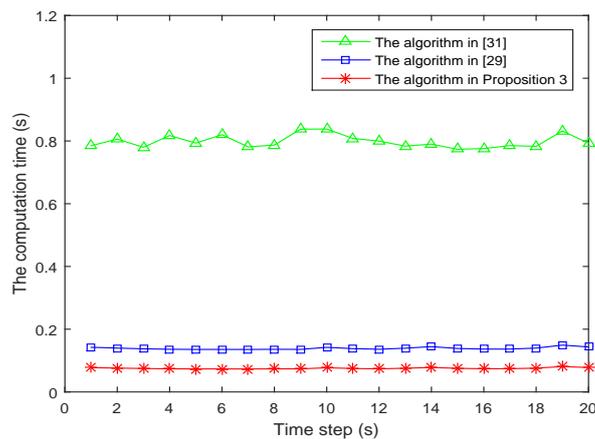
Figure 11. The mean squared error of position in the  $y$ -direction is plotted as function of time steps with  $L = 36$  sensors.



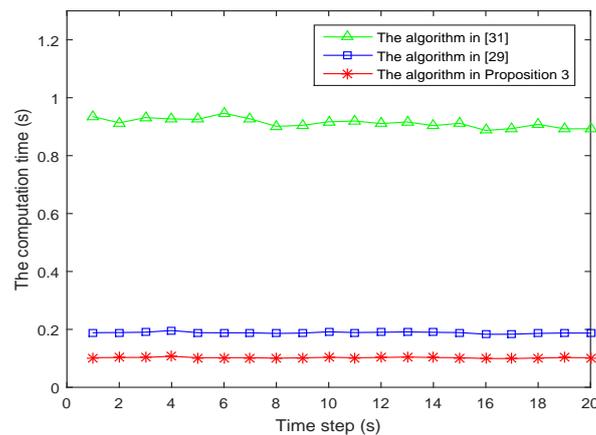
**Figure 12.** The mean squared error of position in the  $x$ -direction is plotted as function of time steps with  $L = 49$  sensors.



**Figure 13.** The mean squared error of position in the  $y$ -direction is plotted as function of time steps with  $L = 49$  sensors.



**Figure 14.** The computation time of obtaining PCRB is plotted as function of time steps with  $L = 36$  sensors.



**Figure 15.** The computation time of obtaining PCRB is plotted as function of time steps with  $L = 49$  sensors.

## 6. Conclusions

This paper has proposed two methods to derive the PCRB to effectively overcome the difficulties caused by uncertainty. The first method is based on the recursive formula of the Cramér–Rao bound and the Gaussian mixture model. Nevertheless, it needs to compute a complex integral based on the joint probability density function of the sensor measurements and the target state. The computational burden of this method is relatively high, especially in large sensor networks. Inspired by the idea of the expectation maximization algorithm, the second method is to introduce some 0–1 latent variables to treat the Gaussian mixture model. Since the regular condition of the posterior Cramér–Rao bound is unsatisfied for the discrete uncertain system, we use some continuous variables to approximate the discrete latent variables. Then, a new Cramér–Rao bound can be achieved by a limiting process of the Cramér–Rao bound of the continuous system. It avoids the complex integral, which can reduce the computation burden. Thus, the sensor selection problems for the nonlinear uncertain dynamic system with linear equality or inequality constraints can be efficiently solved, and the optimal solution of the sensor selection problem can be derived analytically. Thus, it can be used to deal with the sensor selection of large-scale sensor networks. Two typical numerical examples verify the effectiveness of the proposed methods.

**Acknowledgments:** This work was supported in part by the NSFC No. 61673282, the open research funds of BACC-STAFDL of China under Grant No. 2015afdl010 and the PCSIRT16R53.

**Author Contributions:** Zhiguo Wang and Xiaojing Shen proposed the algorithm; Ping Wang performed the experiments; Yunmin Zhu contributed analysis tools.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Nahi, N.E. Optimal recursive estimation with uncertain observation. *IEEE Trans. Inf. Theory* **1969**, *15*, 457–462.
2. Hadidi, M.T.; Schwartz, S.C. Linear recursive state estimators under uncertain observations. *IEEE Trans. Autom. Control* **1979**, *24*, 944–948.
3. Lin, H.; Sun, S. State estimation for a class of non-uniform sampling systems with missing measurements. *Sensors* **2016**, *16*, 1155.
4. Luo, Y.; Zhu, Y.; Luo, D.; Zhou, J.; Song, E.; Wang, D. Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications. *Sensors* **2008**, *8*, 8086–8103.
5. Cao, N.; Choi, S.; Masazade, E.; Varshney, P.K. Sensor selection for target tracking in wireless sensor networks with uncertainty. *IEEE Trans. Signal Process.* **2016**, *64*, 5191–5204.

6. Costa, O.L.V.; Guerra, S. Stationary filter for linear minimum mean square error estimator of discrete-time Markovian jump systems. *IEEE Trans. Autom. Control* **2002**, *47*, 1351–1356.
7. Sinopoli, B.; Guerra, S.; Schenato, L.; Franceschetti, M.; Poolla, K.; Jordan, M.I.; Sastry, S.S. Kalman filtering with intermittent observations. *IEEE Trans. Autom. Control* **2004**, *49*, 1453–1464.
8. Van Trees, H.L. *Detection, Estimation, and Modulation Theory, Part I*; Wiley: New York, NY, USA, 1968.
9. Bobrovsky, B.Z.; Zakai, M. A lower bound on the estimation error for Markov processes. *IEEE Trans. Autom. Control* **1975**, *20*, 785–788.
10. Taylor, J. The Cramér–Rao estimation error lower bound computation for deterministic nonlinear systems. *IEEE Trans. Autom. Control* **1979**, *24*, 343–344.
11. Galdos, J.I. A Cramér–Rao bound for multidimensional discrete-time dynamical systems. *IEEE Trans. Autom. Control* **1980**, *25*, 117–119.
12. Chang, C.B. Two lower bounds on the covariance for nonlinear estimation problems. *IEEE Trans. Autom. Control* **1981**, *26*, 1294–1297.
13. Tichavský, P.; Muravchik, C.H.; Nehorai, A. Posterior Cramér–Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.* **1998**, *46*, 1386–1396.
14. Kirubarajan, T.; Bar-Shalom, Y. Low observable target motion analysis using amplitude information. *IEEE Trans. Aerosp. Electron. Syst.* **1996**, *32*, 1367–1384.
15. Zheng, Y.; Ozdemir, O.; Niu, R.; Varshney, P.K. New conditional posterior Cramér–Rao lower bounds for nonlinear sequential Bayesian estimation. *IEEE Trans. Signal Process.* **2012**, *60*, 5549–5556.
16. Šimandl, M.; Královec, J.; Tichavský, P. Filtering, predictive, and smoothing Cramér–Rao bounds for discrete-time nonlinear dynamic systems. *Automatica* **2001**, *37*, 1703–1716.
17. Hernandez, M.L.; Marrs, A.D.; Gordon, N.J.; Maskell, S.R.; Reed, C.M. Cramér–Rao bounds for non-linear filtering with measurement origin uncertainty. In Proceeding of the 5th International Conference on Information Fusion, Annapolis, MD, USA, 8–11 July 2002.
18. Zhang, X.; Willett, P.; Bar-Shalom, Y. The Cramer-Rao bound for dynamic target tracking with measurement origin uncertainty. In Proceeding of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, USA, 10–13 December 2002.
19. Farina, A.; Ristic, B.; Timmoneri, L. Cramér–Rao bound for nonlinear filtering with  $P_d < 1$  and its application to target tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 1916–1924.
20. Hernandez, M.; Ristic, B.; Farina, A. A comparison of two Cramér–Rao bounds for nonlinear filtering with  $P_d < 1$ . *IEEE Trans. Signal Process.* **2004**, *52*, 2361–2370.
21. Rapoport, I.; Oshman, Y. A Cramér–Rao-type estimation lower bound for systems with measurement faults. *IEEE Trans. Autom. Control* **2005**, *50*, 1234–1245.
22. Hounkpevi, F.O.; Yaz, E.E. Robust minimum variance linear state estimators for multiple sensors with different failure rates. *Automatica* **2007**, *43*, 1274–1280.
23. Zhang, H.; Shi, Y.; Mehr, A.S. Robust weighted  $H_\infty$  filtering for networked systems with intermittent measurements of multiple sensors. *Internat J. Adapt. Control Signal Process.* **2011**, *25*, 313–330.
24. Yick, J.; Mukherjee, B.; Ghosal, D. Wireless sensor network survey. *Comput. Netw.* **2008**, *52*, 2292–2330.
25. Gungor, V.C.; Hancke, G.P. Industrial wireless sensor networks: challenges, design principles, and technical approaches. *IEEE Trans. Ind. Electron.* **2009**, *56*, 4258–4265.
26. Shen, X.; Varshney, P.K. Sensor selection based on generalized information gain for target tracking in large sensor networks. *IEEE Trans. Signal Process.* **2014**, *62*, 363–375.
27. Joshi, S.; Boyd, S. Sensor selection via convex optimization. *IEEE Trans. Signal Process.* **2009**, *57*, 451–462.
28. Shen, X.; Liu, S.; Varshney, P.K. Sensor selection for nonlinear systems in large sensor networks. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 2664–2678.
29. Liu, S.; Fardad, M.; Masazade, E.; Varshney, P.K. Optimal periodic sensor scheduling in networks of dynamical systems. *IEEE Trans. Signal Process.* **2014**, *62*, 3055–3068.
30. Wang, P.; Wang, Z.G.; Shen, X.J.; Zhu, Y.M. The estimation fusion and Cramér–Rao bounds for nonlinear systems with uncertain observations. In Proceeding of the 20th International Conference on Information Fusion, Xi’an, China, 10–13 July 2017.
31. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
32. Yang, X.J.; Niu, R.X. Sparsity-Promoting sensor selection for nonlinear target tracking with quantized data. In Proceeding of the 20th International Conference on Information Fusion, Xi’an, China, 10–13 July 2017.

33. Bilmes, J.A. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *Int. Comput. Sci. Inst.* **1998**, *4*, 126.
34. Chepuri, S.P.; Leus, G. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Trans. Signal Process.* **2015**, *63*, 684–698.
35. Pak, J.M.; Ahn, C.K.; Shmaliy, Y.S.; Lim, M.T. Improving reliability of particle filter-based localization in wireless sensor networks via hybrid particle/FIR filtering. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1089–1098.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).