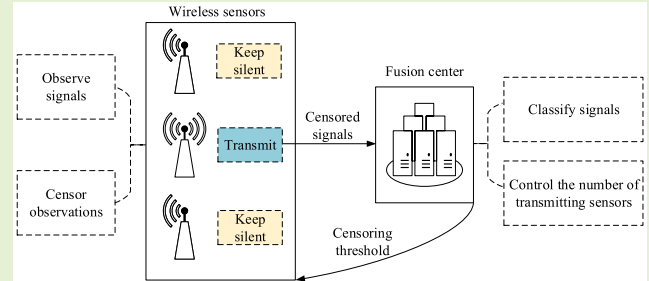


# A Censoring Scheme for Multiclassification in Wireless Sensor Networks

Xiangsen Chen<sup>1</sup>, Graduate Student Member, IEEE, Wenbo Xu<sup>2</sup>, Member, IEEE,  
and Yue Wang<sup>3</sup>, Senior Member, IEEE

**Abstract**—Censoring has been widely applied in wireless sensor networks (WSNs) as an effective method to achieve a balance between energy consumption and the quality of the observed signals. However, most recent studies focus on the censoring schemes applied in binary hypothesis problems (i.e., binary classification and detection problems). To expand the application of censoring in WSNs, we propose a censoring scheme for multiclassification problems in this article. Sensors in this scheme only transmit observations deemed informative enough for classification, where the decision region of whether to transmit is derived based on log likelihood ratios (LLRs). By analyzing the relationship between the communication rate of the WSN and the censoring threshold, we design an adaptive strategy in the censoring scheme so that the censoring threshold can be adjusted according to the communication rate. We further derive the theoretical lower bound of the classification accuracy, which is formulated via the Chernoff distance among different signals. The performance superiority of the censored signals compared with the original ones without censoring is revealed in the form of the theoretical lower bound, verified by experimental results on WSN applications where our proposed censoring scheme allows significant communication saving without the sacrifice of performance.

**Index Terms**—Censoring strategy, compressive sensing (CS), multiclassification, performance bound, wireless sensor networks (WSNs).



## NOMENCLATURE

| Notation                     | Description                                                     |
|------------------------------|-----------------------------------------------------------------|
| $i, N$                       | Scalar.                                                         |
| $\mathbf{s}, \mathbf{v}$     | Vector.                                                         |
| $\mathbf{H}$                 | Matrix.                                                         |
| $\mathcal{L}, \mathcal{D}$   | Set.                                                            |
| $[a, b)$                     | Interval, which means $\{x \mid a \leq x < b\}$ .               |
| $\mathcal{H}_i$              | Hypothesis.                                                     |
| $\mathcal{N}(\mu, \sigma^2)$ | Gaussian distribution with mean $\mu$ and variance $\sigma^2$ . |
| $\mathcal{C}$                | Chernoff distance.                                              |
| $p(\cdot)$                   | Probability density function (pdf) of a variable.               |
| $P(\cdot)$                   | Probability of an event.                                        |

$|a|$  Absolute value of  $a$ .  
 $|\Sigma|$  Determinant of the square matrix  $\Sigma$ .

## I. INTRODUCTION

RECENTLY, people witness an obvious growth of wireless sensor networks (WSNs) owing to the proliferation of the Internet of Things (IoT) in many applications, e.g., environment, industry, military, and health [1], [2], [3], [4]. WSNs can be generally divided into centralized WSNs and decentralized WSNs according to the presence or absence of the fusion center (FC). Decentralized WSNs are usually deployed in target tracking scenarios, where the signal source moves in different sensing ranges of the sensors [5], [6]. By contrast, centralized WSNs are common in traditional signal processing scenarios, e.g., signal detection [7], [8], classification [9], [10], and reconstruction [11], [12]. This article focuses on the centralized WSN, which consists of a group of sensor nodes (SNs) and an FC. These spatially distributed SNs observe signals and transmit these observations to FC by wireless communication, and they are required to be as small and lightweight as possible with a simple structure in order to facilitate the deployment.

However, sensors in many practical applications sample signals at high sampling rates and generate large volumes of data,

Manuscript received 9 April 2023; accepted 1 May 2023. Date of publication 19 May 2023; date of current version 29 June 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61871050, in part by the U.S. National Science Foundation under Grant 2136202, and in part by the Virginia Research Investment Fund under Grant CCI-223996. The associate editor coordinating the review of this article and approving it for publication was Dr. Liansheng Liu. (Corresponding author: Wenbo Xu.)

Xiangsen Chen and Wenbo Xu are with the Key Laboratory of Universal Wireless Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: xuwb@bupt.edu.cn).

Yue Wang is with the Department of Computer Science, Georgia State University, Atlanta, GA 30303 USA (e-mail: wangyuegsu@gmail.com).

Digital Object Identifier 10.1109/JSEN.2023.3276418

which causes a great burden on the communication resources of the sensors. To avoid excessive consumption of communication resources caused by directly transmitting originally observed signals, compressive sensing (CS) has been adopted in WSNs owing to the inherent sparsity in the data collected by sensors [13], [14], [15]. With a properly designed sensing matrix and reconstruction algorithm, sensors efficiently sample signals with much lower sampling rates than the Nyquist sampling rate, and FC can still obtain accurate reconstructions. In such a way, transmitting compressed signals acquired by CS reduces the consumption of communication resources between the FC and SNs.

To further improve the communication efficiency, apart from applying CS in signal transmission, some methods have emerged to transmit only a part of the compressed observations on the premise of ensuring as little performance loss as possible. The existing improved schemes can be divided into two categories, i.e., sensor selection and sensor censoring. The former is a technology that FC determines which sensors can transmit data at a given time [16], [17], whereas in the latter one, the sensor censoring technology hands over the selection task to the sensor to decide whether the observed signal is informative to be transmitted to the FC. Comparatively, sensor censoring is more flexible since sensors can determine their own communication states, i.e., transmitting observations to FC or keeping silent based on the current state, and the extra feedback from FC is avoided [18], [19], [20].

Censoring schemes have been applied in many scenarios in WSNs [21], [22], [23]. In various schemes, local likelihood ratios (LRs) and log likelihood ratios (LLRs) are usually calculated to measure the signal quality. For example, [24] proposes a censoring-based change-point-detection scheme based on LLR, which is utilized to measure the probability that a system is abnormal. Wu et al. [25] propose a ternary censoring rule with three states: directly transmitting the observed signal, transmitting a one-bit hard decision, and keeping silent. Rago et al. [26] propose a simple LR-based censoring rule for the signal detection problem, which establishes a relationship between the censoring threshold and the communication rate of WSNs. However, the above-mentioned works are formulated and solved as simple binary hypothesis problems for either on-off signal detection or binary classification tasks, where the decision region reduces to a simplified single interval and, thus, can be calculated easily [27]. They cannot be straightforwardly extended to the practical but more complicated multihypotheses scenarios. To fill such a gap, we are motivated to design a censoring scheme for multiclassification problems. The main contributions of this article are summarized as follows.

- 1) We propose a WSN censoring scheme for practical multiclassification problems, in which sensors decide locally whether to transmit observations. This scheme provides the censoring rule, which is determined in the offline stage based on a censoring threshold to avoid the energy consumption of calculating LLRs online for each signal. In addition, we design a binary search scheme to find a suitable censoring threshold corresponding to the

preset communication rate, which is determined by the actual transmission ability of the WSN.

- 2) To evaluate the effectiveness of the proposed scheme, we analyze the theoretical performance of the proposed scheme in terms of the lower bound of the classification accuracy. By comparing the theoretical lower bound of the originally observed signals and censored signals, we prove the performance advantage of our censoring scheme.
- 3) We provide the experiments of our proposed scheme in both the binary classification scenario and the multiclassification scenario. The results show that when compared with the conventional schemes, the proposed censoring scheme effectively reduces the communication costs without compromising the classification performance.

The rest of this article is organized as follows. Section II illustrates the system model and the definition of LLR. Our censoring scheme is proposed afterward in Section III. The theoretical analysis, i.e., the theoretical lower bound of the classification accuracy, is provided in Section IV. Simulation results are shown in Section V to verify our analysis. Finally, Section VI draws a conclusion to this article.

The notations used in this article are listed in the Nomenclature for the convenience of the following description.

## II. SIGNAL MODEL AND LLR

This section is devoted to describing the signal model and the definition of LLR.

### A. Signal Model

The multiclassification problem in the WSN aims to classify an observed signal that belongs to one of the  $L$  hypotheses, where the probability of the  $i$ th hypothesis is denoted as  $P_i$ . The original signal  $\mathbf{x} \in \mathbb{R}^{N \times 1}$  under each hypothesis is formulated as follows:

$$\begin{aligned} \mathcal{H}_1 : \mathbf{x} &= \mathbf{s}_1 + \mathbf{v} \\ \mathcal{H}_2 : \mathbf{x} &= \mathbf{s}_2 + \mathbf{v} \\ &\vdots \\ \mathcal{H}_L : \mathbf{x} &= \mathbf{s}_L + \mathbf{v} \end{aligned} \quad (1)$$

where  $\mathbf{s}_i \in \mathbb{R}^{N \times 1}$  is the source signal under the  $i$ th hypothesis,  $i$  belongs to the index set of hypotheses  $\mathcal{L} = \{1, 2, \dots, L\}$ , and  $\mathbf{v} \in \mathbb{R}^{N \times 1}$  is the additive white Gaussian noise (AWGN).

Suppose that the WSN consists of  $M$  sensors and an FC. The signal observation of the  $m$ th sensor is formulated as

$$y_m = \mathbf{h}_m^T \mathbf{x} = \mathbf{h}_m^T (\mathbf{s}_i + \mathbf{v}) \quad (2)$$

where  $y_m$  is the compressed signal and  $\mathbf{h}_m \in \mathbb{R}^{N \times 1}$  is the sensing vector of the  $m$ th sensor.

We assume that the entries of  $\mathbf{s}_i$  are independent random Gaussian variables, i.e.,  $\mathbf{s}_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$  and the noise  $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_N)$ , where  $\mathbf{I}_N$  is the  $N \times N$  identity matrix. Then, we obtain the distribution of the observed signal  $y_m$  and the vector  $\mathbf{y} = (y_1, y_2, \dots, y_M)^T$  consisting of all observations

under the  $i$ th hypothesis as follows:

$$\begin{aligned} y_m | \mathcal{H}_i &\sim \mathcal{N}(\mathbf{h}_m^T \boldsymbol{\mu}_i, \mathbf{h}_m^T (\boldsymbol{\Sigma}_i + \sigma_v^2 \mathbf{I}_N) \mathbf{h}_m) \\ \mathbf{y} | \mathcal{H}_i &\sim \mathcal{N}(\mathbf{H}^T \boldsymbol{\mu}_i, \mathbf{H}^T (\boldsymbol{\Sigma}_i + \sigma_v^2 \mathbf{I}_N) \mathbf{H}) \end{aligned} \quad (3)$$

where  $\mathbf{H} \triangleq (\mathbf{h}_1, \dots, \mathbf{h}_M)^T$  is the equivalent sensing matrix of the whole WSN. For the convenience of illustration, the mean  $\mathbf{h}_m^T \boldsymbol{\mu}_i$  and the variance  $\mathbf{h}_m^T (\boldsymbol{\Sigma}_i + \sigma_v^2 \mathbf{I}_N) \mathbf{h}_m$  of the  $m$ th signal in (3) are denoted as  $\mu_{mi}$  and  $\sigma_{mi}^2$ , respectively. Besides, we denote the mean vector of  $\mathbf{y} | \mathcal{H}_i$  as  $\boldsymbol{\mu}_{yi} \triangleq \mathbf{H}^T \boldsymbol{\mu}_i$  and the covariance matrix as  $\boldsymbol{\Sigma}_{yi} \triangleq \mathbf{H}^T (\boldsymbol{\Sigma}_i + \sigma_v^2 \mathbf{I}_N) \mathbf{H}$ .

### B. Log Likelihood Ratio

The LLR is usually employed to measure the difference between two pdfs, which is defined as follows:

$$\text{LLR}_{ij}(y_m) = \ln \left( \frac{p(y_m | \mathcal{H}_i)}{p(y_m | \mathcal{H}_j)} \right) \quad (4)$$

where  $p(y_m | \mathcal{H}_i)$  is the conditional pdf of the observed signal  $y_m$  under the  $\mathcal{H}_i$  hypothesis. According to (3), the LLR is given by

$$\text{LLR}_{ij}(y_m) = \ln \left( \frac{\sigma_{mj}}{\sigma_{mi}} \right) + \frac{(y_m - \mu_{mj})^2}{2\sigma_{mj}^2} - \frac{(y_m - \mu_{mi})^2}{2\sigma_{mi}^2}. \quad (5)$$

## III. PROPOSED SCHEME

In this section, we present a censoring scheme for multiclassification. We illustrate the censoring rule of our scheme, the decision region of discarding less informative observations, the selection of the censoring threshold, and the improvement of the classifier in FC.

### A. Censoring Rule for Multiclassification

The communication state of sensors is denoted as  $\mathbf{u} = (u_1, u_2, \dots, u_M)$ , where  $u_m \in \{0, 1\}$  represents the state of the  $m$ th sensor, such that

$$\begin{cases} u_m = 1, & \text{transmitting} \\ u_m = 0, & \text{keeping silent.} \end{cases} \quad (6)$$

The purpose of censoring is to transmit informative observations and discard less informative ones, where the latter contributes less to the classification. Most existing censoring schemes [19], [24], [25], [26] first define the uninformative decision region of each SN and then adopt the following general rule:

$$\begin{cases} u_m = 0, & y_m \in \mathcal{D}_m \\ u_m = 1, & y_m \notin \mathcal{D}_m \end{cases} \quad (7)$$

where  $\mathcal{D}_m$  is the uninformative decision region of the  $m$ th SN. To guarantee efficient censoring, the first important step is to calculate suitable decision regions. Though recent works have made some achievements in this aspect [19], [26], [27], they only focus on the binary classification problem. For example, [27, Th. 1] shows that the optimal uninformative decision

region at each SN of the binary classification problem is in the form of

$$\mathcal{D}_m = \{y_m | t_1 \leq LR(y_m) \leq t_2\} \quad (8)$$

where  $LR(y_m) = (p(y_m | \mathcal{H}_1)) / (p(y_m | \mathcal{H}_0))$ ;  $\mathcal{H}_0$  and  $\mathcal{H}_1$  represent the absence and presence of the source signal, respectively; and  $t_1$  and  $t_2$  are the lower and upper thresholds of this single interval, respectively. In addition, [19, Th. 2.2] further shows  $t_1$  can be simplified to zero when  $P_0$ , i.e., the prior probability of  $\mathcal{H}_0$ , is larger than a special value. This condition is usually satisfied in the detection problem (i.e., a specific binary classification problem, where  $P_0$  is generally larger than  $P_1$ ), and  $t_1 = 0$  is set by [26]. However, this condition cannot be always satisfied in the general binary classification problem. In this case, a more feasible censoring decision region is still the one in (8) with  $t_1 \neq 0$ . To fairly constrain the conditional probability of  $y_m$  under either  $\mathcal{H}_0$  or  $\mathcal{H}_1$ , a feasible simplification of (8) is to set  $t_1 = (1/t_2)$ . Then, we obtain the following simplified decision region:

$$\mathcal{D}_m = \{y_m | |LLR_{01}(y_m)| \leq t\} \quad (9)$$

where  $t = |\ln t_2| = |\ln t_1|$ . By doing so, the region with two thresholds is reduced to the one with only one threshold  $t$ .

To extend the censoring rule to the multiclassification scenario, it is obvious that the transmitted observations should be informative for all hypotheses. Thus, the decision region of censoring should consider all LLRs, i.e.,  $\text{LLR}_{ij}(y_m)$ ,  $\forall i, j \in \mathcal{L}, i \neq j$ . Each LLR corresponds to a decision subregion  $\mathcal{D}(m; i, j)$ , so the whole decision region of the  $m$ th SN should be the union of all the decision subregions

$$\begin{aligned} \mathcal{D}_m &= \bigcup_{\substack{i \neq j \\ \forall i, j \in \mathcal{L}}} \mathcal{D}(m; i, j) \\ \mathcal{D}(m; i, j) &\triangleq \{y_m | |\text{LLR}_{ij}(y_m)| \leq t\}. \end{aligned} \quad (10)$$

### B. Calculation of Decision Region

Note that the LLR in (5) means that we have to calculate the LLR for each observation. However, it is inappropriate to implement such tedious calculations on each sensor, owing to the high complexity and limited energy. Therefore, it is more desirable to determine the decision region in the offline stage. Combining (5) and (10), the decision region is calculated by the following inequation:

$$\left| \ln \left( \frac{\sigma_{mj}}{\sigma_{mi}} \right) + \frac{(y_m - \mu_{mj})^2}{2\sigma_{mj}^2} - \frac{(y_m - \mu_{mi})^2}{2\sigma_{mi}^2} \right| < t. \quad (11)$$

The left-hand side of this inequation is denoted as  $|G(y_m)|$ . To determine whether  $y_m$  should be discarded, we discuss (11) in the following three cases.

The first case corresponds to  $\sigma_{mi}^2 = \sigma_{mj}^2$ , and  $G(y_m)$  can be reduced to an affine function, i.e.,  $G(y_m) = ((2y - \mu_{mi} - \mu_{mj})(\mu_{mi} - \mu_{mj})) / (2\sigma_{mi}^2)$ . We denote this case as Case 1, which is illustrated in Fig. 1(a) with the settings shown in the figure. The solution of inequation (11) in Case 1 is given by

$$y_m \in [-a_1 - b_1, -a_1 + b_1] \quad (12)$$

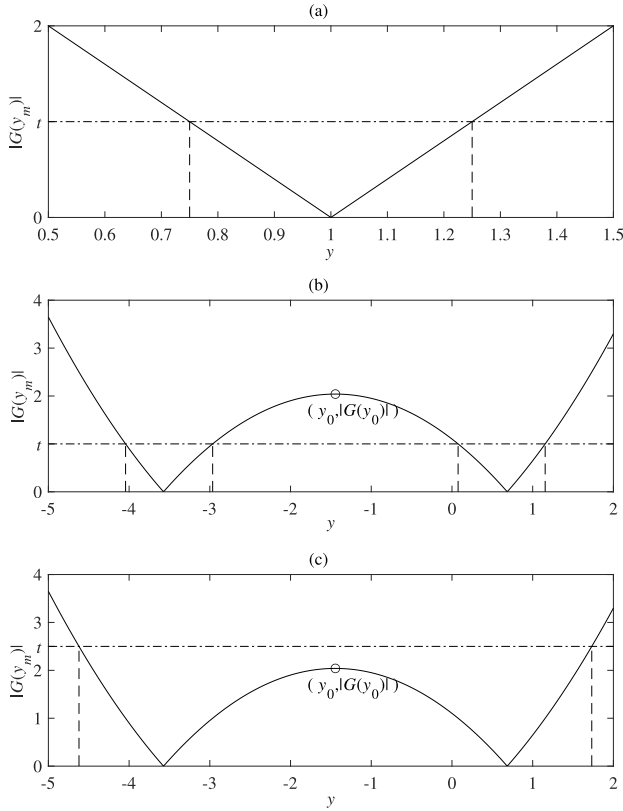


Fig. 1. Three cases of calculating decision regions.

where  $a_1 = (\mu_{mi} + \mu_{mj})/2$  and  $b_1 = (\sigma_{mi}^2 t)/(\mu_{mj} - \mu_{mi})$ .

When  $\sigma_{mi}^2 \neq \sigma_{mj}^2$ , the solution of inequation (11) can be further divided into two cases, which are shown, respectively, in Fig. 1(b) and (c). Case 2 occurs when  $t$  is less than the extreme value of  $G(y_m)$ , denoted as  $G(y_0)$  given by

$$G(y_0) = \left| -\frac{(\mu_{mi} - \mu_{mj})^2}{2(\sigma_{mi}^2 - \sigma_{mj}^2)} + \ln\left(\frac{\sigma_{mi}}{\sigma_{mj}}\right) \right| \quad (13)$$

where  $y_0$  is the corresponding extreme point. As shown in Fig. 1(b), the solution of (11) in this case contains two intervals, which is formulated in (14), as shown at the bottom of the next page, where  $a_2$ ,  $b_2$ , and  $b_3$  are given in the equation (14).

$$\begin{cases} a_2 = -\frac{\mu_{mj}\sigma_{mi}^2 - \mu_{mi}\sigma_{mj}^2}{\sigma_{mj}^2 - \sigma_{mi}^2} \\ b_2 = \frac{\sigma_{mi}^2\sigma_{mj}^2(\mu_{mi} - \mu_{mj})^2}{(\sigma_{mj}^2 - \sigma_{mi}^2)^2} + \frac{2\sigma_{mi}^2\sigma_{mj}^2}{\sigma_{mj}^2 - \sigma_{mi}^2} \left( \ln \frac{\sigma_{mj}}{\sigma_{mi}} - t \right) \\ b_3 = \frac{\sigma_{mi}^2\sigma_{mj}^2(\mu_{mi} - \mu_{mj})^2}{(\sigma_{mj}^2 - \sigma_{mi}^2)^2} + \frac{2\sigma_{mi}^2\sigma_{mj}^2}{\sigma_{mj}^2 - \sigma_{mi}^2} \left( \ln \frac{\sigma_{mj}}{\sigma_{mi}} + t \right). \end{cases} \quad (15)$$

The last case, i.e., Case 3, where  $t$  is larger than  $G(y_0)$ , is shown in Fig. 1(c). It is obvious that the solution of (11) in

this case is a single interval, which is formulated as follows:

$$\begin{cases} y_m \in [a_2 - \sqrt{b_2}, a_2 + \sqrt{b_2}], & \sigma_{mi}^2 > \sigma_{mj}^2 \\ y_m \in [a_2 - \sqrt{b_3}, a_2 + \sqrt{b_3}], & \sigma_{mi}^2 < \sigma_{mj}^2. \end{cases} \quad (16)$$

According to the above arguments, we conclude the solution cases of (11) as follows:

$$\begin{cases} \sigma_{mi}^2 = \sigma_{mj}^2, & \text{Case 1} \\ \sigma_{mi}^2 \neq \sigma_{mj}^2 \text{ and } t < G(y_0), & \text{Case 2} \\ \sigma_{mi}^2 \neq \sigma_{mj}^2 \text{ and } t \geq G(y_0), & \text{Case 3} \end{cases} \quad (17)$$

and the decision regions of censoring are expressed, respectively, in (12), (14), and (16). Based on the above illustrations, the steps of calculating the decision region of one SN are summarized in **Scheme 1**.

---

**Scheme 1** Computing the Decision Region

---

**Require:**  $\mu_{yi}$ ,  $\Sigma_{yi}$  and the censoring threshold  $t$

**Ensure:** decision regions of all SNs

```

1: for  $i = 1$  to  $L - 1$  do
2:   for  $j = i + 1$  to  $L$  do
3:     switch case based on (17)
4:       case 1: Compute  $\mathcal{D}(m; i, j)$  from (12)
5:       case 2: Compute  $\mathcal{D}(m; i, j)$  from (14)
6:       case 3: Compute  $\mathcal{D}(m; i, j)$  from (16)
7:     end switch
8:   end for
9: end for
10:  $\mathcal{D}_m = \bigcup_{\substack{i \neq j \\ \forall i, j \in \mathcal{L}}} \mathcal{D}(m; i, j)$ 
11: return  $\mathcal{D}_m$ 

```

---

For the convenience of the following description, we rewrite the decision region of the  $m$ th sensor as the union of  $N_m$  single intervals:  $\mathcal{D}_{m,1}, \dots, \mathcal{D}_{m,N_m}$ , which have no intersections with each other

$$\mathcal{D}_m = \bigcup_{i=1}^{N_m} \mathcal{D}_{m,i} \quad \forall i, j \in \{1, \dots, N_m\}, \quad \mathcal{D}_{m,i} \cap \mathcal{D}_{m,j} = \emptyset. \quad (18)$$

In the above equation, each single interval is denoted as  $\mathcal{D}_{m,i} = [c_{mi}, d_{mi}]$ , where  $c_{mi}$  and  $d_{mi}$  are the lower and upper bounds of this interval, respectively.

### C. Relationship Between the Censoring Threshold and the Communication Rate

It is noted that the decision region derived in Section III-B depends on the censoring threshold  $t$ . In this section, we determine the censoring threshold  $t$  according to the communication rate  $\varepsilon$ , which is defined as the ratio of sensors maintaining communications with FC in a unit time interval. Following the definition in [26], the communication rate of the whole WSN system is given by

$$\begin{aligned} \varepsilon &= \frac{1}{M} \sum_{m=1}^M P(u_m = 1) \\ &= \frac{1}{M} \sum_{m=1}^M \left\{ \sum_{i=1}^L P_i P(u_m = 1 | \mathcal{H}_i) \right\}. \end{aligned} \quad (19)$$

The conditional probability of the  $m$ th sensor transmitting signals under the  $i$ th hypothesis is given in (20), as shown at the bottom of the next page. Observing from (12), (14), and (16), the lower bound  $c_{mj}$  and the upper bound  $d_{mj}$  of the decision regions are the functions of  $t$ . Thus, we express them as  $c_{mj}(t)$  and  $d_{mj}(t)$ , respectively. Combining (19) and (20), we derive the expression for the communication rate with respect to  $t$  in (21), as shown at the bottom of the next page.

It is difficult to derive the inverse function of  $t$  with respect to  $\varepsilon$  from (21) owing to the summation in this equation. We adopt the binary search scheme to search for  $t$  corresponding to a preset communication rate  $\varepsilon_0$  and then give the determined decision region, as shown in **Scheme 2**. Specifically, this scheme aims to search for  $t$  within a possible search space  $[t_{\text{left}}, t_{\text{right}}]$  and stops if  $|\varepsilon_{\text{mid}} - \varepsilon_0| \leq \delta$ , where  $\varepsilon_{\text{mid}}$  is the communication rate corresponding to the midpoint  $t_{\text{mid}}$  between  $t_{\text{left}}$  and  $t_{\text{right}}$ , and  $\delta$  is the searching precision. It is obvious that  $\varepsilon = 1|_{t=0}$  and  $\lim_{t \rightarrow \infty} \varepsilon = 0$ . Thus,  $t_{\text{left}}$  and  $t_{\text{right}}$  can be initialized to 0 and a large value, respectively. However, a too-large value of  $t_{\text{right}}$  affects the convergence of the searching scheme, so we introduce an iterative step with a given search step length  $t_s$  to initialize  $t_{\text{right}}$ . That is,  $t_{\text{right}}$  is first set to 0 and then increases by  $t_s$  in each iteration, until  $\varepsilon_{\text{right}} \leq \varepsilon_0$ , as shown in **steps 2–6**. In iteration search **steps 7–18**, the scheme iteratively computes  $\varepsilon_{\text{mid}}$  and updates the search space by comparing the values of  $\varepsilon_{\text{mid}}$  and  $\varepsilon_0$ . By such iteration steps,  $\varepsilon_{\text{mid}}$  keeps approaching the preset  $\varepsilon_0$  until the iteration stops when the difference between them is smaller than  $\delta$ .

#### D. Enhanced Classifier

With the censoring scheme proposed in Section III-B, only informative observations are retained and transmitted to FC. As a result, the distribution of observations received by FC is different from (3), which corresponds to the case that all observations are transmitted without censoring. Thus, it is essential to redesign the classifier to fit the new distribution. Note that the censored observations contain those that are not in the decision region  $\mathcal{D}_m$ . The distribution of the censored observations for the  $i$ th hypothesis, which is denoted as  $p_c(y_m|\mathcal{H}_i)$ , is given by

$$p_c(y_m|\mathcal{H}_i) = \begin{cases} \frac{p_o(y_m|\mathcal{H}_i)}{P(u_m = 1|\mathcal{H}_i)}, & y_m \notin \mathcal{D}_m \\ 0, & y_m \in \mathcal{D}_m \end{cases} \quad (22)$$

where  $p_o(y_m|\mathcal{H}_i)$  is the pdf of the original observations, which is given by (3). When the sensing matrix is fixed and the distribution of signals is known, the optimal classifier minimizing the classification error is the maximum a posteriori (MAP) classifier, which is expressed as follows:

$$\hat{\mathcal{H}} = \arg \max_{1 \leq i \leq L} p(\mathcal{H}_i|\mathbf{y}) = \arg \max_{1 \leq i \leq L} p(\mathbf{y}|\mathcal{H}_i) P_i. \quad (23)$$

#### Scheme 2 Binary Search for Determining Decision Regions

**Require:**  $\mu_{yi}$ ,  $\Sigma_{yi}$ , the preset communication rate  $\varepsilon_0$ , the searching precision  $\delta$ , and the step length  $t_s$

**Ensure:** Determined decision regions

```

1: Initialization:  $t_{\text{left}}, t_{\text{right}} = 0$ 
2: repeat
3:    $t_{\text{right}} = t_{\text{right}} + t_s$ 
4:    $\forall m$ , call Scheme 1 with  $t_{\text{right}}$  to compute  $\mathcal{D}_m$ 
5:   Compute  $\varepsilon_{\text{right}}$  from (21) based on  $\mathcal{D}_1, \dots, \mathcal{D}_M$ 
6: until  $\varepsilon_{\text{right}} \leq \varepsilon_0$ 
7: loop
8:    $t_{\text{mid}} = (t_{\text{left}} + t_{\text{right}})/2$ 
9:    $\forall m$ , call Scheme 1 with  $t_{\text{mid}}$  to compute  $\mathcal{D}_m$ 
10:  Compute  $\varepsilon_{\text{mid}}$  from (21) based on  $\mathcal{D}_1, \dots, \mathcal{D}_M$ 
11:  if  $|\varepsilon_{\text{mid}} - \varepsilon_0| \leq \delta$  then
12:    break
13:  else if  $\varepsilon_{\text{mid}} > \varepsilon_0$  then
14:     $t_{\text{left}} = t_{\text{mid}}$ 
15:  else
16:     $t_{\text{right}} = t_{\text{mid}}$ 
17:  end if
18: end loop
19: return  $\mathcal{D}_1, \dots, \mathcal{D}_M$ 

```

By combining (23) and (22), our improved MAP classifier for censored observations is given in (24), as shown at the bottom of the next page.

Based on these discussions, the detailed procedures of the WSN adapting to the censoring scheme are shown in **Scheme 3**. The complete procedures contain offline and online stages. In the offline stage, the means and variances of observations are calculated, and **Scheme 2** is used to determine the decision regions. Then, the FC notifies each decision region to the corresponding SN. In the online stage, SNs continuously observe signals and censor the observations based on the decision regions. The FC receives the censored observations transmitted from SNs and completes the classification task by using the enhanced classifier in (24).

To describe how our scheme works, we give an illustrated example of a WSN adopting our censoring scheme consisting of an FC and four SNs. Assume that the decision regions of these SNs are  $[-1, 0.5]$ ,  $[0.3, 1.5]$ ,  $[1, 1.9]$ , and  $[-0.2, 1.3]$ , which are calculated according to **Scheme 2** in the offline stage. In an observation period of the online stage, the SNs observe the source signal  $\mathbf{s}$  and obtain their respective compressed observations: 0.2, -0.9, 1.3, and 1.8. By censoring,  $0.2 \in [-1, 0.5]$  and  $1.3 \in [1, 1.9]$ , whereas  $-0.9 \notin [0.3, 1.5]$  and  $1.8 \notin [-0.2, 1.3]$ . Therefore, SNs 1 and 3 keep silent; SNs 2 and 4 transmit the symbols with the values -0.9 and 1.8 to FC, respectively, in this observation period. After FC receives -0.9 and 1.8, it uses them to determine which hypothesis holds according to (24).

$$\begin{cases} y_m \in [a_2 - \sqrt{b_3}, a_2 - \sqrt{b_2}] \cup [a_2 + \sqrt{b_2}, a_2 + \sqrt{b_3}], & \sigma_{mi}^2 < \sigma_{mj}^2 \\ y_m \in [a_2 - \sqrt{b_2}, a_2 - \sqrt{b_3}] \cup [a_2 + \sqrt{b_3}, a_2 + \sqrt{b_2}], & \sigma_{mi}^2 > \sigma_{mj}^2 \end{cases} \quad (14)$$

**Scheme 3** Proposed Censoring Scheme

**Require:** Distribution parameters of all types of signals,  $\varepsilon_0$ ,  $\delta$ , and  $t_s$

**Ensure:** Classification result of the signal

**Offline stage at FC:**

- 1: Compute  $\mu_{yi} = \mathbf{H}^T \mu_i$  and  $\Sigma_{yi} = \mathbf{H}^T (\Sigma_i + \sigma_v^2 \mathbf{I}_N) \mathbf{H}$
- 2: Call **Scheme 2** to determine decision regions
- 3: Notify decision regions to SNs

**Online stage:**

*At the  $m$ -th sensor:*

- 4: Sample signals
- 5: **if**  $y_m \notin \mathcal{D}_m$  **then**
- 6:   Transmit  $y_m$  to FC
- 7: **else**
- 8:   Keep silent
- 9: **end if**

*At FC:*

- 10: Receive censored signals
- 11: Classify the signal based on (24)
- 12: **return** the classification result  $\hat{\mathcal{H}}$

## IV. THEORETICAL ANALYSIS

In this section, we derive a performance lower bound of the multiclassification when utilizing the censoring scheme in Section III. Afterward, the performance superiority of the censoring scheme is proved by comparing the lower bound of the classification performance for noncensored signals and censored signals.

## A. Chernoff Distance

The Chernoff distance is utilized as a metric to deduce the theoretical lower bound of signal classification [28]. The Chernoff distance between two pdfs is defined as follows:

$$\begin{aligned} \mathcal{C}(p(\mathbf{x}|\mathcal{H}_i), p(\mathbf{x}|\mathcal{H}_j)) &\triangleq \max_{0 \leq t \leq 1} \tilde{\mathcal{C}}(t; p(\mathbf{x}|\mathcal{H}_i), p(\mathbf{x}|\mathcal{H}_j)) \\ &= -\ln \left\{ \int p(\mathbf{x}|\mathcal{H}_i)^{1-t_0} p(\mathbf{x}|\mathcal{H}_j)^{t_0} d\mathbf{x} \right\} \end{aligned} \quad (25)$$

where  $t_0 = \arg \max_{0 \leq t \leq 1} \tilde{\mathcal{C}}(t; p(\mathbf{x}|\mathcal{H}_i), p(\mathbf{x}|\mathcal{H}_j))$ . Since it is difficult to obtain an analytical solution in (25), we consider the special case of the Chernoff distance, where  $t_0 = 0.5$  [29].

To calculate the Chernoff distance, we first calculate the definite integral part in (25). For the convenience of description, we define a function  $F_{m,ij}(a, b)$  in (26), as shown at the bottom of the next page, where  $\mu_{Fm} = (\mu_{mi}\sigma_{mj}^2 + \mu_{mj}\sigma_{mi}^2)/(\sigma_{mi}^2 + \sigma_{mj}^2)$  and  $\sigma_{Fm} = ((2\sigma_{mi}^2\sigma_{mj}^2)/(\sigma_{mi}^2 + \sigma_{mj}^2))^{1/2}$ .

The Chernoff distances of the received signals in the original scheme and the censoring scheme are given in (27), as shown at the bottom of the next page, where  $\mathcal{C}_{o,ij}$  is the Chernoff distance of the original signals and  $\mathcal{C}_{c,ij}$  is that of the censored signals, respectively.

## B. Theoretical Lower Bound of Multiclassification Accuracy

Based on the derivation of Chernoff distance, we further derive the theoretical lower bound of multiclassification accuracy. Suppose that  $\mathbf{H}\mathbf{H}^T = \mathbf{I}$ . [29] shows that the upper bound of error probability between two hypotheses is formulated in (28), as shown at the bottom of the next page. By assuming equiprobable hypotheses, the classification accuracy of the multiclassification problem is lower bounded by

$$P_{acc} = 1 - P_{err} \geq 1 - \sum_{i=1}^L \sum_{j=1, j \neq i}^L P(\hat{\mathcal{H}} = \mathcal{H}_i | \mathcal{H}_j) P_j. \quad (29)$$

We provide some numerical experiments of (29) in Fig. 2 to show the superiority of our censoring scheme. In this figure, “Original” and “Censored” represent the schemes with non-censored observations and censored observations, respectively. From this figure, the lower bound of classification performance increases with the increase in  $M$ , which shows that increasing the number of observations can improve the performance of classification. The gap of lower bounds between original observations and censored observations demonstrates the

$$\begin{aligned} P(u_m = 1 | \mathcal{H}_i) &= 1 - \int_{\mathcal{D}_m} P(y_m | \mathcal{H}_i) dy \\ &= 1 - \sum_{j=1}^{N_m} \int_{\mathcal{D}_{m,j}} P(y_m | \mathcal{H}_i) dy \\ &= 1 - \frac{1}{2} \sum_{j=1}^{N_m} \left[ \operatorname{erf} \left( \frac{d_{mj} - \mu_{mi}}{\sqrt{2\sigma_{mi}^2}} \right) - \operatorname{erf} \left( \frac{c_{mj} - \mu_{mi}}{\sqrt{2\sigma_{mi}^2}} \right) \right] \end{aligned} \quad (20)$$

$$\varepsilon = 1 - \frac{1}{2M} \sum_{m=1}^M \sum_{i=1}^L \sum_{j=1}^{N_m} P_i \left[ \operatorname{erf} \left( \frac{d_{mj}(t) - \mu_{mi}}{\sqrt{2\sigma_{mi}^2}} \right) - \operatorname{erf} \left( \frac{c_{mj}(t) - \mu_{mi}}{\sqrt{2\sigma_{mi}^2}} \right) \right] \quad (21)$$

$$\hat{\mathcal{H}} = \arg \min_{1 \leq i \leq L} \left\{ (\mathbf{y} - \mu_{yi})^T \Sigma_{yi}^{-1} (\mathbf{y} - \mu_{yi}) + \ln(|\Sigma_{yi}|) + 2 \ln P(u_m = 1 | \mathcal{H}_i) - 2 \ln P_i \right\} \quad (24)$$

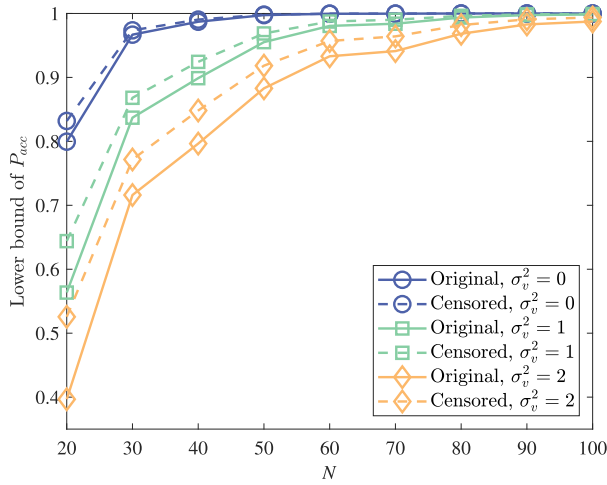


Fig. 2. Theoretical lower bounds of classification accuracy.

superiority of the latter because it transmits more informative observations than the original scheme.

## V. EXPERIMENTS

This section presents the experimental results of the proposed scheme to demonstrate its superiority. The experiments consider both the binary classification scenario and the multiclassification scenario.

The sensing matrices of these WSNs are generated by the QR decomposition of a Gaussian random matrix

$$\begin{aligned} \mathbf{Q} &= \mathbf{Q}R(\mathbf{K}) \\ \mathbf{H} &= \mathbf{Q}_{1:M} \end{aligned} \quad (30)$$

where  $\mathbf{K}$  is an  $N \times N$  random matrix, whose entries follow the normal distribution,  $\mathbf{Q}R(\cdot)$  represents QR decomposition,  $\mathbf{Q}$  is the orthogonal matrix obtained by QR decomposition, and  $\mathbf{H}$  is the sensing matrix consisting of the first  $M$  rows of  $\mathbf{Q}$ . In doing so, the sensing matrix satisfies  $\mathbf{H}\mathbf{H}^T = \mathbf{I}$ .

### A. Results in the Binary Classification Scenario

Our scheme focuses on censoring in the multiclassification scenario; however, no equivalent schemes are available in the current literature for comparison. To verify the advantage of our scheme, we compare it with the censoring scheme in [26], where only two classes are considered. The problem is classifying two hypotheses, i.e., the absence and presence of the sparse signal. The sparse signal with  $N = 1000$  and the sparsity of 20 is generated with a fixed support set, whose nonzero entries follow  $\mathcal{N}(0, \sigma_s^2)$ , and the noise follows the i.i.d. standard normal distribution. To ensure the fairness of the comparison, we replace the locally most powerful test (LMPT) classifier in [26] with the MAP classifier in (24). In the following comparison, the WSNs adopting our

$$\begin{aligned} F_{m,ij}(a, b) &\triangleq \int_a^b p_o(y_m|\mathcal{H}_i)^{\frac{1}{2}} p_o(y_m|\mathcal{H}_j)^{\frac{1}{2}} dy \\ &= \sqrt{\frac{2\sigma_{mi}\sigma_{mj}}{\sigma_{mi}^2 + \sigma_{mj}^2}} \exp\left(-\frac{(\mu_{mi} - \mu_{mj})^2}{4(\sigma_{mi}^2 + \sigma_{mj}^2)}\right) \times \frac{1}{2} \left[ \operatorname{erf}\left(\frac{b - \mu_{Fm}}{\sqrt{2}\sigma_{Fm}}\right) - \operatorname{erf}\left(\frac{a - \mu_{Fm}}{\sqrt{2}\sigma_{Fm}}\right) \right] \end{aligned} \quad (26)$$

$$\begin{aligned} \mathcal{C}_{o,ij} &\triangleq \mathcal{C}(p_o(\mathbf{y}|\mathcal{H}_i), p_o(\mathbf{y}|\mathcal{H}_j)) \\ &= -\ln \left\{ \prod_{m=1}^M \int_{-\infty}^{+\infty} p(y_m|\mathcal{H}_i)^{\frac{1}{2}} p(y_m|\mathcal{H}_j)^{\frac{1}{2}} dy_m \right\} \\ &= -\sum_{m=1}^M \ln \{F_{m,ij}(-\infty, +\infty)\} \\ \mathcal{C}_{c,ij} &\triangleq \mathcal{C}(p_c(\mathbf{y}|\mathcal{H}_i), p_c(\mathbf{y}|\mathcal{H}_j)) \\ &= -\ln \left\{ \prod_{m=1}^M \left[ \int_{-\infty}^{+\infty} p(y_m|\mathcal{H}_i)^{\frac{1}{2}} p(y_m|\mathcal{H}_j)^{\frac{1}{2}} dy_m - \sum_{l=1}^{N_m} \int_{c_{ml}}^{d_{ml}} p(y_m|\mathcal{H}_i)^{\frac{1}{2}} p(y_m|\mathcal{H}_j)^{\frac{1}{2}} dy_m \right] \right\} \\ &= -\sum_{m=1}^M \ln \left\{ \frac{F_{m,ij}(-\infty, +\infty) - \sum_{l=1}^{N_m} F_{m,ij}(c_{ml}, d_{ml})}{\sqrt{P(u_m = 1|\mathcal{H}_i)P(u_m = 1|\mathcal{H}_j)}} \right\} \end{aligned} \quad (27)$$

$$\begin{aligned} \Pr(\hat{\mathcal{H}} = \mathcal{H}_i|\mathcal{H}_j) &\leq \frac{1}{2} \exp(-\mathcal{C}_{c,ij}) \\ &\leq \frac{1}{2} \prod_{m=1}^M \left\{ \frac{F_{m,ij}(-\infty, +\infty) - \sum_{l=1}^{N_m} F_{m,ij}(c_{ml}, d_{ml})}{\sqrt{P(u_m = 1|\mathcal{H}_i)P(u_m = 1|\mathcal{H}_j)}} \right\} \end{aligned} \quad (28)$$

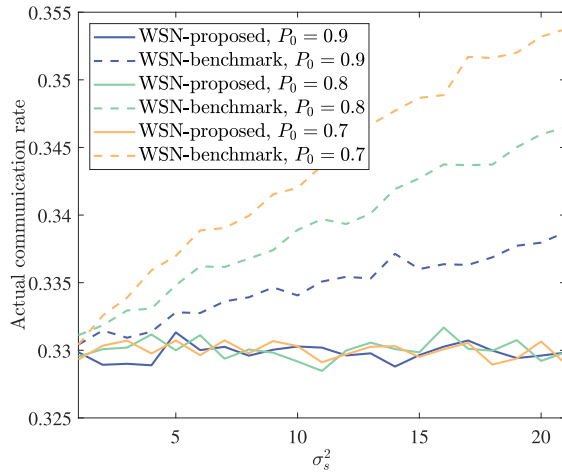


Fig. 3. Actual communication rate of the WSN-proposed and the WSN-benchmark with varying signal variances.

censoring scheme and the comparison scheme are denoted as WSN-proposed and WSN-benchmark, respectively. We set the number of SNs as  $M = 100$  and the preset communication rate as  $\varepsilon_0 = 0.33$  of both WSNs to ensure fairness, where the latter is determined by the actual transmission ability of the communication system. In addition,  $\delta = 0.1$ , and  $t_s = 1$  in the proposed scheme.

First, we calculate the actual communication rate  $\hat{\varepsilon}$  (the statistical average of the numbers of the transmitting SNs in 10 000 Monte Carlo trials) of the two schemes when the preset communication rate  $\varepsilon_0$  is fixed at 0.33, as shown in Fig. 3. Observed from this figure, the actual communication rate of WSN-benchmark increases as  $\sigma_s^2$  increases and gradually deviates from the preset communication rate 0.33. It is due to the reason that the sensors in the comparison scheme are more inclined to transmit observations when  $P_1$  is larger according to the censoring rule in [26]. This phenomenon is confirmed in this figure, where a smaller  $P_0$  corresponds to a larger actual communication rate of WSN-benchmark. As for our proposed scheme, the actual communication rate is almost identical to the preset communication rate no matter of the varying  $P_0$  and  $\sigma_s^2$ , which demonstrates that the adaptability and stability of our scheme are both superior to those of the comparison scheme.

Next, we depict the classification accuracy curves with varying  $\sigma_s^2$  in Fig. 4 by averaging over 10 000 Monte Carlo trials. From this figure, WSN-proposed performs better than WSN-benchmark on the classification accuracy when  $\sigma_s^2$  is not very large. It can be also noted that the classification accuracy of WSN-proposed is slightly worse than that of WSN-benchmark when  $\sigma_s^2$  is large. It is due to the reason that the actual communication rate of WSN-benchmark increases as  $\sigma_s^2$  increases, which means more observations are received for classification at FC in WSN-benchmark. However, such an advantage does not exist in a practical system, where the communication rate is generally limited.

### B. Results in the Multiclassification Scenario

In this section, we present the results in the multiclassification scenario. The following four WSNs will be considered in this section.

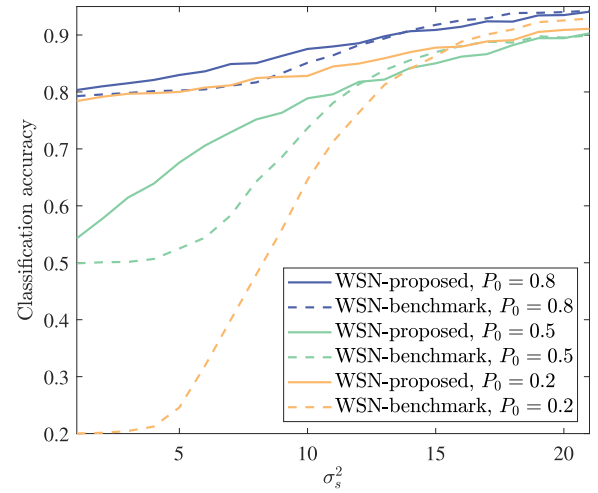


Fig. 4. Classification accuracy of the WSN-proposed and the WSN-benchmark with varying signal variances.

TABLE I  
DISTRIBUTION PARAMETERS OF SYNTHETIC DENSE SIGNALS

|          | Signal 1 | Signal 2 | Signal 3 | Signal 4 | Signal 5  |
|----------|----------|----------|----------|----------|-----------|
| Mean     | 0        | 1        | 2        | 3        | 4         |
| Variance | 1        | 1.1      | 0.9      | 1.12     | 0.95      |
|          | Signal 6 | Signal 7 | Signal 8 | Signal 9 | Signal 10 |
| Mean     | 5        | 6        | 7        | 8        | 9         |
| Variance | 0.88     | 0.93     | 1.05     | 0.92     | 1.15      |

- 1) *WSN-ori1*: the original WSN without censoring including  $M$  sensors.
- 2) *WSN-cen*: the WSN with our proposed censoring scheme including  $M$  sensors. The preset communication rate of this WSN is  $\varepsilon_0$ , and the auxiliary parameters of the scheme are set to be  $\delta = 0.1$ , and  $t_s = 1$ .
- 3) *WSN-sel*: the WSN adopts the random selection scheme. It has  $\varepsilon_0 M$  transmitting sensors, which are randomly selected from  $M$  sensors in each round of simulation.
- 4) *WSN-ori2*: the original WSN without censoring including fixed  $\varepsilon_0 M$  sensors.

These WSNs experiment with synthetic dense signals, synthetic sparse signals, and the MNIST dataset so that we can evaluate the performance of our scheme in various signal conditions. All these datasets contain ten types of signals. For the synthetic dense signal, the length of the signal is set to 100, and each element of the signal follows the same Gaussian distribution, whose means and variances are shown in Table I. For the synthetic sparse signal, the length is 1000, and the sparsity is 100. Different types of sparse signals have different support sets, and the nonzero elements of the signal follow the Gaussian distribution, whose means and variances are shown in Table II. For the MNIST dataset, the samples are first normalized. Then, suppose the pixels of each digit in the dataset follow the mixture Gaussian distribution [30]. The training set in MNIST is utilized to estimate the means and variances of pixels in each digit, and the test set is used to show the classification accuracy.

We first discuss the classification results with different energies of noise of three datasets, as shown in Fig. 5. The

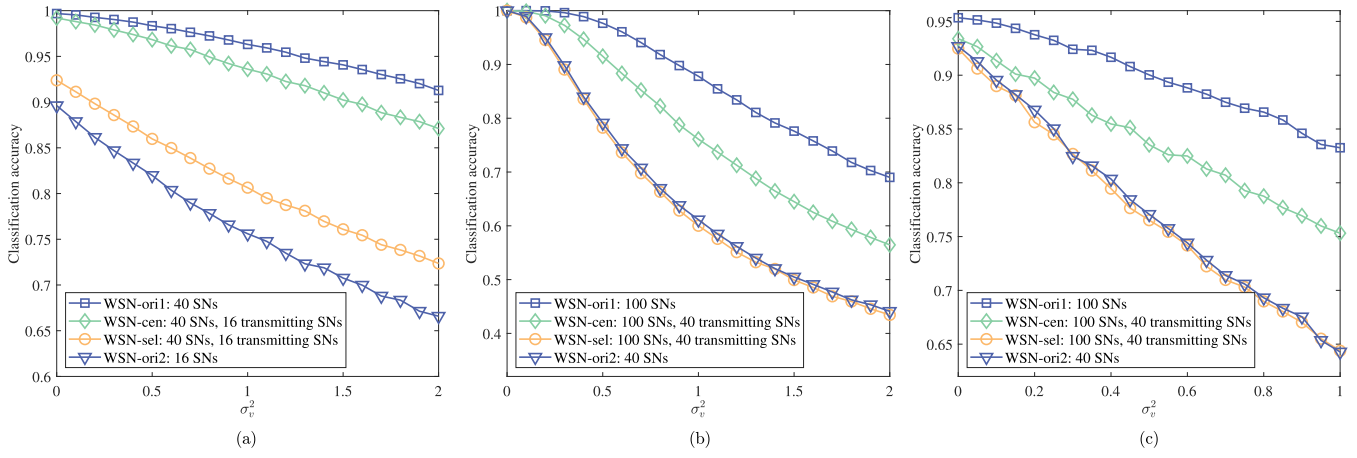


Fig. 5. Classification accuracy with different energies of noise.

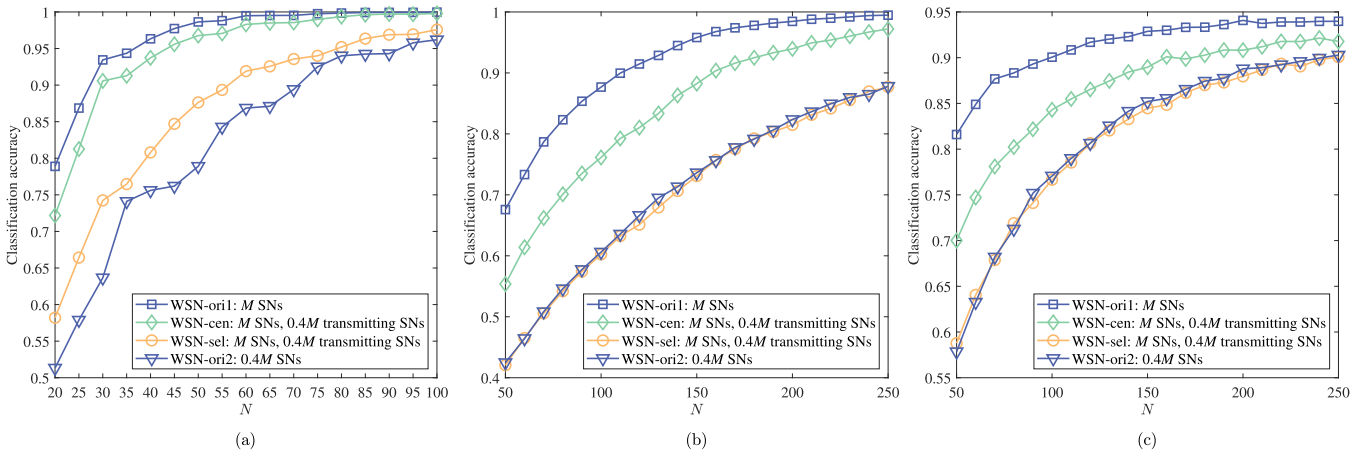


Fig. 6. Classification accuracy with different numbers of SNs.

TABLE II  
DISTRIBUTION PARAMETERS OF THE NONZERO ENTRIES IN  
SYNTHETIC SPARSE SIGNALS

|          | Signal 1 | Signal 2 | Signal 3 | Signal 4 | Signal 5  |
|----------|----------|----------|----------|----------|-----------|
| Mean     | 1        | 1        | 1        | 1        | 1         |
| Variance | 1        | 1.1      | 0.9      | 1.12     | 0.95      |
|          | Signal 6 | Signal 7 | Signal 8 | Signal 9 | Signal 10 |
| Mean     | 1        | 1        | 1        | 1        | 1         |
| Variance | 0.88     | 0.93     | 1.05     | 0.92     | 1.15      |

results of synthetic dense signals, synthetic sparse signals, and MNIST are shown in Fig. 5(a)–(c), respectively. As shown in this figure, the classification accuracy of all schemes decreases with the increase of  $\sigma_v^2$ . All three subfigures show that WSN-cen performs significantly better than WSN-sel and WSN-ori2, which means the WSN with our censoring scheme transmits more informative observations than the other two schemes to FC. As expected, the performance of WSN-ori1 is better than WSN-cen. This is due to the fact that all SNs transmit observation in the former, whereas in the latter, only 40% SNs are activated. Meanwhile, the gap between WSN-cen and WSN-sel, and that between WSN-cen and WSN-ori2 increase with the increase in  $\sigma_v^2$  in the experiments of all three datasets. This indicates that the effect of censoring is better when the signal is harder to classify. That means our censoring

scheme has superiority in the low signal-to-noise ratio (SNR) condition.

For the result of synthetic dense signals in Fig. 5(a), WSN-ori2 performs worse than WSN-sel, but in Fig. 5(b) and (c), the performance of these two WSNs is similar. That is because the sensing vectors on different SNs are different in terms of the ability to distinguish signals. Take the first and second SNs for example. If  $|\mathbf{h}_1^T \boldsymbol{\mu}_i - \mathbf{h}_1^T \boldsymbol{\mu}_j|$  is larger than  $|\mathbf{h}_2^T \boldsymbol{\mu}_i - \mathbf{h}_2^T \boldsymbol{\mu}_j|$ , the observations from the first SN possess better classification ability than those from the second SN. Such a phenomenon implies the problem that the sensing vectors of a WSN may be unfortunately unsatisfactory to distinguish signals. Note that such a problem is more likely to occur when the number of SNs is smaller, which explains the relatively poor performance of WSN-ori2 in Fig. 5(a). In contrast, WSN-cen and WSN-sel avoid this problem due to the randomness of activated SNs in each round of simulation. We will soon observe that the aforementioned problem vanishes in the subsequent experiments of synthetic sparse signals and MNIST because of the large number of SNs.

Then, we focus on the results with different numbers of SNs, as shown in Fig. 6. It can be observed that the classification accuracy of all schemes increases with the increase in  $M$ , and the degree of performance improvement slows down with the increase in  $M$ . This implies that the performance

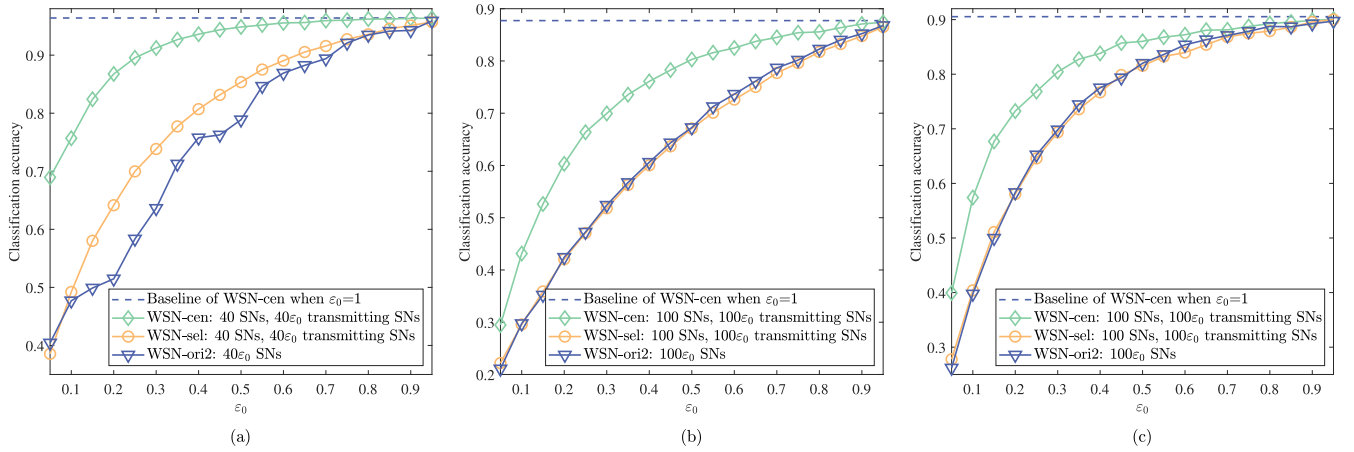


Fig. 7. Classification accuracy with different communication rates.

of each WSN converges an upper bound when  $M$  is large enough. Notably, the performance gap between WSN-ori1 and WSN-cen narrows with the increase in  $M$ , especially narrows to 0 in Fig. 6(a). This means WSN with our censoring scheme can save a large amount of transmitting energy consumption but with only little performance loss when  $M$  is large.

The performance curves in Fig. 6(a) are not as smooth as those in Fig. 6(b) and (c). This phenomenon is caused by the similar reason mentioned in the fourth paragraph of this subsection for Fig. 6 that different SNs have different abilities to distinguish signals. Especially, the smaller the  $M$  is, the greater the influence of a single sensing vector on the performance of the WSN. The numbers of SNs in Fig. 6(a) are smaller than Fig. 6(b) and (c), and thus, the phenomenon of unsmooth curves appears.

At last, the results with different communication rates are shown in Fig. 7. This figure shows that, with the increase in the communication rate, the performance of all three schemes first improves rapidly and then slowly approaches the baseline of WSN-cen when the communication rate is set to 1. The performance degradation in the low communication rate case indicates that the performance loss caused by a significant decrease in observations cannot be compensated by the advantage of informative observations. We can observe that the performance gaps between WSN-cen and WSN-sel, and WSN-cen and WSN-ori2 first increase and then decrease. Meanwhile, the increasing degree of classification accuracy of WSN-cen slows down when  $\varepsilon_0 > 0.5$ . This indicates that the communication rate should be set approximately by considering the balance between performance requirements and energy-saving requirements in practice.

## VI. CONCLUSION

This article proposes a WSN censoring scheme for multiclassification. This scheme gives a decision region of whether to transmit observations based on LLRs and a censoring threshold. Sensors compute this decision region offline and censor signals efficiently based on this region in the online stage. We further theoretically analyze the Chernoff distances between the signals of different types for the censored scheme and the noncensored scheme, which demonstrate the effectiveness of our proposed scheme. Experimental results in

the binary classification scenario and the multiclassification scenario demonstrate that the proposed scheme achieves the desired performance of multiclassification comparable with the conventional scheme without censoring while significantly reducing the communication costs.

## REFERENCES

- [1] A. Vena, N. Samat, B. Sorli, and J. Podlecki, "Ultralow power and compact backscatter wireless device and its SDR-based reading system for environment monitoring in UHF band," *IEEE Sensors Lett.*, vol. 5, no. 5, pp. 1–4, May 2021.
- [2] P. López Díez, I. Gabilondo, E. Alarcón, and F. Moll, "Mechanical energy harvesting taxonomy for industrial environments: Application to the railway industry," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2696–2706, Jul. 2020.
- [3] S. Verma, S. Zeadally, S. Kaur, and A. K. Sharma, "Intelligent and secure clustering in Wireless Sensor Network (WSN)-based intelligent transportation systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 13473–13481, Aug. 2022.
- [4] E. M. Ar-Reyouchi, K. Ghomid, D. Ar-Reyouchi, S. Rattal, R. Yahiaoui, and O. Elmazria, "Protocol wireless medical sensor networks in IoT for the efficiency of healthcare," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10693–10704, Jul. 2022.
- [5] T. Liang, Y. Lin, L. Shi, J. Li, Y. Zhang, and Y. Qian, "Distributed vehicle tracking in wireless sensor network: A fully decentralized multiagent reinforcement learning approach," *IEEE Sensors Lett.*, vol. 5, no. 1, pp. 1–4, Jan. 2021.
- [6] G. Li, G. Li, and Y. He, "Distributed GGIW-CPHD-based extended target tracking over a sensor network," *IEEE Signal Process. Lett.*, vol. 29, pp. 842–846, 2022.
- [7] X. Wang, G. Li, and P. K. Varshney, "Detection of sparse signals in sensor networks via locally most powerful tests," *IEEE Signal Process. Lett.*, vol. 25, no. 9, pp. 1418–1422, Sep. 2018.
- [8] A. Mohammadi, D. Ciuonzo, A. Khazaei, and P. S. Rossi, "Generalized locally most powerful tests for distributed sparse signal detection," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 8, pp. 528–542, 2022.
- [9] M. Liu, K. Yang, N. Zhao, Y. Chen, H. Song, and F. Gong, "Intelligent signal classification in industrial distributed wireless sensor networks based industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4946–4956, Jul. 2021.
- [10] J. Guo, R. G. Raj, D. J. Love, and C. G. Brinton, "Nonparametric decentralized detection and sparse sensor selection via multi-sensor online kernel scalar quantization," *IEEE Trans. Signal Process.*, vol. 70, pp. 2593–2608, 2022.
- [11] K. Sekar, K. S. Devi, and P. Srinivasan, "Compressed tensor completion: A robust technique for fast and efficient data reconstruction in wireless sensor networks," *IEEE Sensors J.*, vol. 22, no. 11, pp. 10794–10807, Jun. 2022.
- [12] J. Feng, F. Chen, and H. Chen, "Data reconstruction coverage based on graph signal processing for wireless sensor networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 48–52, Jan. 2022.

- [13] S. Li, L. D. Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things," *IEEE Trans. Ind. Inform.*, vol. 9, no. 4, pp. 2177–2186, Nov. 2013.
- [14] M. Razzaque and S. Dobson, "Energy-efficient sensing in wireless sensor networks using compressed sensing," *Sensors*, vol. 14, no. 2, pp. 2822–2859, Feb. 2014.
- [15] Y. Qie, C. Hao, and P. Song, "Wireless transmission method for large data based on hierarchical compressed sensing and sparse decomposition," *Sensors*, vol. 20, no. 24, p. 7146, Dec. 2020.
- [16] T. Nagata, T. Nonomura, K. Nakai, K. Yamada, Y. Saito, and S. Ono, "Data-driven sparse sensor selection based on A-optimal design of experiment with ADMM," *IEEE Sensors J.*, vol. 21, no. 13, pp. 15248–15257, Jul. 2021.
- [17] F. Guan, W.-W. Cui, L.-F. Li, and J. Wu, "A comprehensive evaluation method of sensor selection for PHM based on grey clustering," *Sensors*, vol. 20, no. 6, p. 1710, Mar. 2020.
- [18] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2668–2673, Nov. 2005.
- [19] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Decentralized detection with censoring sensors," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1362–1373, Apr. 2008.
- [20] K. A. Flanigan and J. P. Lynch, "Optimal event-based policy for remote parameter estimation in wireless sensing architectures under resource constraints," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5293–5304, Jul. 2022.
- [21] P. Xu, Y. Wang, X. Chen, and Z. Tian, "COKE: Communication-censored decentralized kernel learning," *J. Mach. Learn. Res.*, vol. 22, no. 196, pp. 1–35, 2021.
- [22] W. Li, Y. Liu, Z. Tian, and Q. Ling, "Communication-censored linearized ADMM for decentralized consensus optimization," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 6, pp. 18–34, 2020.
- [23] L. Yang, H. Zhu, H. Wang, K. Kang, and H. Qian, "Data censoring with network lifetime constraint in wireless sensor networks," *Digit. Signal Process.*, vol. 92, pp. 73–81, Sep. 2019.
- [24] Y. Gu, Y. Jiao, X. Xu, and Q. Yu, "Request-response and censoring-based energy-efficient decentralized change-point detection with IoT applications," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6771–6788, Apr. 2021.
- [25] J. Wu, M. Yang, and T. Wang, "Energy-efficient sensor censoring for compressive distributed sparse signal recovery," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2137–2152, May 2018.
- [26] C. Li, G. Li, and P. K. Varshney, "Distributed detection of sparse signals with censoring sensors via locally most powerful test," *IEEE Signal Process. Lett.*, vol. 27, pp. 346–350, 2020.
- [27] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.
- [28] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Berlin, Germany: Springer, 2013.
- [29] T. Wimalajeewa, H. Chen, and P. K. Varshney, "Performance limits of compressive sensing-based signal classification," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2758–2770, Jun. 2012.
- [30] I. Ralasic, A. Tafro, and D. Sersic, "Statistical compressive sensing for efficient signal reconstruction and classification," in *Proc. 4th Int. Conf. Frontiers Signal Process. (ICFSP)*, Sep. 2018, pp. 44–49.



**Xiangsen Chen** (Graduate Student Member, IEEE) received the B.S. degree in information engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2020, where he is currently pursuing the Ph.D. degree with the Key Laboratory of Universal Wireless Communications.

His research interests include signal processing, wireless sensor networks, and censoring techniques.



**Wenbo Xu** (Member, IEEE) received the B.S. degree from the School of Information Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2005, and the Ph.D. degree from the School of Information and Communication Engineering, BUPT, in 2010.

Since 2010, she has been with the BUPT, where she is currently a Professor with the School of Artificial Intelligence. Her current research interests include sparse signal processing, noncooperative communication, machine learning, and signal processing in wireless networks.



**Yue Wang** (Senior Member, IEEE) received the Ph.D. degree in communication and information system from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, in 2011.

He is currently an Assistant Professor with the Department of Computer Science, Georgia State University, Atlanta, GA, USA. Previously, he was a Research Assistant Professor with the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA, USA. His general interests include machine learning, signal processing, wireless communications, and their applications in cyber-physical systems. His specific research focuses on compressive sensing, massive MIMO, millimeter-wave communications, wideband spectrum sensing, cognitive radios, DoA estimation, high-dimensional data analysis, and distributed optimization and learning.