Distributed Quantized Detection of Sparse Signals Under Byzantine Attacks

Chen Quan, Yunghsiang S. Han, *Fellow, IEEE*, Baocheng Geng and Pramod K. Varshney, *Life Fellow, IEEE*

***Abstract*—This paper investigates distributed detection of sparse stochastic signals with quantized measurements under Byzantine attacks, where sensors may send falsified data to the Fusion Center (FC) to degrade system performance. Here, the Bernoulli-Gaussian (BG) distribution is used to model sparse stochastic signals. Several detectors with significantly improved detection performance are proposed by incorporating estimates of attack parameters into the detection process. In the case of unknown sparsity degree and attack parameters, we propose the generalized likelihood ratio test with reference sensors (GLRTRS) as well as the locally most powerful test with reference sensors (LMPTRS). Our simulation results show that these detectors outperform the LMPT and GLRT detectors designed in attack-free environments and achieve detection performance close to the benchmark likelihood ratio test (LRT) detector. In the case of unknown sparsity degree and known fraction of Byzantine nodes in the network, we further propose enhanced LMPTRS (E-LMPTRS) and enhanced GLRTRS (E-GLRTRS) detectors by filtering out potential malicious sensors in the network, resulting in improved detection performance compared to GLRTRS and LMPTRS detectors.**

***Index Terms*—Byzantine attacks, wireless sensor networks, distributed detection, compressed sensing.**

I. Introduction

With the development of compressive sensing (CS) [1]–[4] in recent years, the sensors in sensor networks often send low-dimensional compressed measurements to the Fusion Center (FC) instead of high-dimensional sparse data, thereby improving bandwidth efficiency and reducing the communication overhead. A high-dimensional signal is sparse when only a few entries in the signal are non-zero, and others are zeros. Under the CS framework, the reconstruction and the detection of sparse signals have received considerable attention. In this paper, we are interested in detecting compressed sparse signals.

The problem of compressed sparse signal detection in sensor networks has been studied in the literature [5]–[13]. In these studies, the recovery of sparse signals was not necessarily required. In [5]–[7], partly or completely reconstructed sparse signals are required to derive the test statistics for sparse signal detection, while in [8]–[12], the test statistics are directly derived from compressed measurements to perform sparse signal detection. In [5] and [6], the authors proposed orthogonal matching pursuit algorithms to detect the presence of a sparse signal based on single measurement vectors and multiple measurement vectors, respectively, by estimating only a fraction of the support set of a sparse signal. In [8], the Bernoulli-Gaussian (BG) distribution was utilized to model the random sparsity of sparse signals, and the generalized likelihood ratio test (GLRT) was proposed to address the unknown degree of sparsity. Note that under the BG model (which is widely used to model the sparsity of signals [8], [14]–[17]), the sparse signal has zero sparsity degree if the signal is absent, but a nonzero sparsity degree that approaches zero if the signal is present. Due to this property, parameter testing based on the sparsity degree can be employed for sparse signal detection by formulating the problem as a one-sided and close hypothesis testing problem. In [9], instead of GLRT, a method based on the locally most powerful test (LMPT), which is a popular tool for the problems of one-sided and close hypothesis testing, was proposed for detecting sparse signals in sensor networks. The test statistic of the LMPT detector was directly derived from the compressed measurements without any signal recovery. The detectors proposed in [5], [6], [8], [9] assume that the raw signals are transmitted within the network. However, due to limited bandwidth constraints in practical scenarios, it is necessary to consider the case where only quantized data is transmitted over sensor networks. To satisfy this requirement, many studies have been conducted on the design of sparse signal detectors based on quantized data [7], [10]–[13], [18], [19].

A two-stage detector based on the GLRT, where sparse signal recovery is integrated into the detection framework, was proposed in [7] for sparse signal detection from 1-bit CS-based measurements. However, due to substantial information loss caused by 1-bit quantization, there is a noticeable performance gap compared to the clairvoyant detector based on analog measurements [18]. To address this issue, the authors in [13] proposed a quantized LMPT detector that enables the system to achieve detection performance comparable to a clairvoyant LMPT detector by selecting a reasonable number of reference sensors. The work was extended in [10] to consider generalized Gaussian noise. Additionally, [11] proposed an improved-1-bit LMPT detector that optimizes the quantization process and reduces the required number of sensor nodes to compensate for the performance loss caused by 1-bit quantization. The authors of [12] proposed a computationally-efficient generalized LMPT detector for the detection of distributed sparse signals when non-ideal reporting channels between the sensors and the FC are considered. In [19], the authors

C. Quan and P. K. Varshney are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: {chquan,varshney}@syr.edu).

Y. S. Han is with the Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, China (e-mail: yunghsiangh@gmail.com).

B. Geng is with the Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL 35294 USA (e-mail: bgeng@uab.edu).

proposed an energy-efficient censoring-based LMPT detector in clustered sensor networks to address the excessively high energy consumption caused by data transmission in existing centralized LMPT detectors. In this scheme, the cluster head sensors and the ordinary sensors only transmit data that is sufficiently informative to the FC.

In this paper, we address the resilience and detection performance of GLRT-based detectors [7], [8] and LMPT-based detectors [9]–[12] in the presence of Byzantine attacks, where one or more sensors in the network may get compromised and may send falsified data to the FC to degrade the detection performance of the system [20]–[27]. Unlike previous studies that focused on attack-free environments, we investigate the impact of compromised sensors and falsified data on the detection performance, and enhance the resilience of the detectors. More specifically, we consider the GLRT-based and LMPT-based detectors with unknown random sparse signals operating under Byzantine attacks. The random unknown sparse signals are still characterized by the BG distribution as in [7]–[12], [14], [15]. When such a system is under Byzantine attacks, two factors need to be taken into account: the unknown sparsity of the signal and the presence of unidentified attacks. We assume that the Byzantines do not have perfect knowledge about the actual state of the phenomenon of interest and attack based on their local decisions, and we also assume that the system does not have perfect knowledge about the attack strategy. Under such assumptions, we evaluate the performance of the GLRT-based and the LMPT-based detectors. The simulation results show that the detectors are vulnerable to Byzantine attacks because their performance degrades.

To improve the resilience of the system in the presence of Byzantine attacks, intuitively, we need more information about the attack parameters. In this work, we develop a framework for estimating unknown parameters that are inspired by the works in [28], [29], where supervised machine learning was utilized as quality of transmission estimator for optical transport networks. In [28] and [29], a fraction of the total data is used to obtain a sufficiently accurate estimate of the unknown underlying parameters. Correspondingly, a subset of the sensors in this work is randomly selected, with their decisions serving as training samples for estimating the unknown attack parameters in the network. We introduce the notion of reference sensors to represent those sensors whose local decisions serve as training samples in our problem and propose the generalized likelihood ratio test with reference sensors (GLRTRS) and the locally most powerful test with reference sensors (LMPTRS) with adaptive thresholds, given that the sparsity degree and the attack parameter are unknown. The proposed detectors allow us to yield excellent system performance. When the fraction of Byzantines in the networks is assumed to be known, we propose enhanced LMPTRS (E-LMPTRS) and enhanced GLRTRS (E-GLRTRS) detectors which can further improve the detection performance of the system. The main contributions of this work are summarized as follows.

- We perform a comprehensive performance analysis of existing GLRT-based and LMPT-based detectors in the presence of Byzantine attacks. Our analysis and simulation results reveal the degree to which both detectors are vulnerable to attacks.
- We propose a novel approach to design resilient GLRT and LMPT based detectors by considering the potential existence of adversarial Byzantine attacks. Specifically, we integrate the estimation of attack parameters into the detection process.
- Given that the sparsity degree and the attack parameters (i.e., the fraction of Byzantine nodes and the probability that Byzantines flip local decisions) are unknown, we propose GLRTRS and LMPTRS detectors with adaptive thresholds. Our simulation results indicate that both GLRTRS and LMPTRS detectors are resilient to Byzantine attacks. They can achieve detection performance close to that of the benchmark likelihood ratios test (LRT) detector, which has perfect knowledge of the sparsity degree and attack parameters.
- When the fraction of Byzantines in the networks is assumed to be known, we propose E-GLRTRS and E-LMPTRS detectors, which further improve the detection performance of the system by filtering out potential malicious sensors. Our simulation results show that the proposed enhanced detectors outperform LMPTRS and GLRTRS detectors.

The paper is organized as follows. We present our system model in Section II. The performance of GLRT and quantized LMPT detectors under Byzantine attacks is evaluated in Section III. The resilient GLRTRS, LMPTRS, E-GLRTRS, and E-LMPTRS detectors with adaptive thresholds are proposed in Section IV. We present our simulation results in Section V and conclude in Section VI.

Notation: Throughout this paper, we use bold lowercase letters for vectors (e.g., $\mathbf{x_i}$, $\mathbf{h_i}$) and normal font letters for scalars (e.g., n_i). For a vector $\mathbf{x}_i$, we use $x_{i,m}$ to denote its m-th element. The function $Q(\cdot)$ denotes the tail distribution function of the standard normal distribution. The function $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution. The function $\mathbb{E}(\cdot)$ denotes the expected value function, and $\mathrm{Var}(\cdot)$ denotes the variance function. $(\cdot)^T$ denotes the transpose operation. The function $I(a,b)$ is an indicator function that returns 1 if a equals b and returns 0 otherwise. $x \sim \mathcal{N}(\mu, \sigma)$ represent the case where x follow Gaussian distribution with mean μ and variance σ.

II. System model

Consider the binary hypothesis testing problem of detecting sparse signals where hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$ indicate the presence and absence of the sparse signal, respectively. We consider a distributed network consisting of one fusion center (FC) and N sensors that observe the signals that share the joint sparsity pattern[1] as shown in Fig. 1. Let y_i be the received observation at sensor $i \in \{1, 2, \ldots, N\}$. We assume that all the observations are independent and identically distributed (i.i.d)

[1] Joint sparsity pattern indicates that non-zero elements of all the signals occur at the same locations, and the sparsity pattern is the same across all signals. This assumption of joint sparsity pattern can be readily observed in the field of compressed sensing, e.g., [30]–[33].

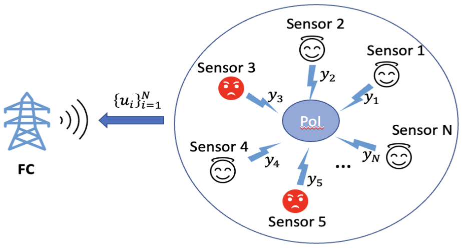

Fig. 1: System model of distributed network. The red sensors are malicious.

conditioned on the hypotheses. For sensor i, the observation y_i is modeled as

$$y_i = \begin{cases} n_i & \text{under } \mathcal{H}_0 \\ \mathbf{h_i}^T \mathbf{x_i} + n_i & \text{under } \mathcal{H}_1, \end{cases} \tag{1}$$

where $\mathbf{x_i} \in \Re^{M\times 1}$ is the sparse signal received by sensor i, $\mathbf{h_i} \in \Re^{M\times 1}$ is the channel gain of sensor i, which is modeled as a random vector to account for the variability and uncertainty in the communication channel, and n_i is Gaussian noise with zero mean and variance σ_n^2. Based on the received compressed measurements $\{y_i\}_{i=1}^N$ from all the sensors, the FC makes a global decision about the absence or presence of the sparse signals.

We adopt the BG distribution introduced in [7]–[12], [14], [15] to model the sparse signals where the joint sparsity pattern is shared among all the signals observed by the sensors. The locations of nonzero coefficients in x_i are assumed to be the same across all the sensors. Let $\mathbf{s} \in \Re^{M\times 1}$ describe the joint sparsity pattern of $\{\mathbf{x}_i\}_{i=1}^N$, where

$$\begin{cases} s_m = 1, & \text{for } \{x_{i,m} \neq 0, i = 1, 2, \ldots, N\} \\ s_m = 0, & \text{for } \{x_{i,m} = 0, i = 1, 2, \ldots, N\} \end{cases} \tag{2}$$

for $m = 1, 2, \ldots, M$. $\{s_m\}_{m=1}^M$ are assumed to be i.i.d. Bernoulli random variables with a common parameter p ($p \to 0^+$), where $P(s_m = 1) = p$ and $P(s_m = 0) = 1 - p$. In other words, p represents the sparsity degree of the sparse signal $\mathbf{x}_i$ for $\forall i \in \{1, 2, \ldots, N\}$. Moreover, each element of $\mathbf{x}_i$ is assumed to follow an i.i.d. Gaussian distribution $\mathcal{N}(0, \sigma_x^2)$ [34]. Therefore, the BG distribution is imposed on $x_{i,m}$ as

$$x_{i,m} \sim p\mathcal{N}(0, \sigma_x^2) + (1-p)\delta(x_{i,m}), \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta function. Due to the limited bandwidth, the sensors send their quantized observations instead of raw observations $\{y_i\}_{i=1}^N$ to the FC. We assume that a fraction α of the total N sensors, namely, αN sensors, are compromised by the Byzantines. We also assume that the compromised sensors are uniformly distributed in the network. In other words, a sensor i can be honest (H) with probability $1-\alpha$ or Byzantine (B) with probability α. The Byzantines may intentionally send falsified local decisions to the FC with an attack probability, i.e., the probability that Byzantines flip their decision. The fraction of Byzantines α and the probability that Byzantines flip their decision, P_A, are considered attack parameters. Note that the fusion rule is assumed not to be altered by Byzantine nodes.[2] Let $\mathbf{z_i}$ denote the actual quantized observation at sensor $i \in \{1, 2, \ldots, N\}$. The q-bit quantizer at the i^{th} sensor is defined as

$$\mathbf{z_i} = \begin{cases} \mathbf{v_1} & \tau_{i,0} \leq y_i \leq \tau_{i,1} \\ \mathbf{v_2} & \tau_{i,1} \leq y_i \leq \tau_{i,2} \\ \vdots & \vdots \\ \mathbf{v_{2^q}} & \tau_{i,2^q-1} \leq y_i \leq \tau_{i,2^q}, \end{cases} \tag{4}$$

where $\mathbf{v_k}$ is the binary code word with $\mathbf{v_k} \in \{0,1\}^q$ that represents the quantized observation and $\{\tau_{i,l}, l = 0, 1, 2, \ldots, 2^q\}$ are the quantization thresholds. For example, given $q = 2$, we have $\mathbf{v_1} = 00$, $\mathbf{v_2} = 01$, $\mathbf{v_3} = 10$ and $\mathbf{v_4} = 11$. Let $\mathbf{u_i}$ be the binary vector sent to the FC, which represents one of the possible quantizer observations $\{\mathbf{v_k} : k = 1, \ldots 2^q\}$. $\mathbf{u_i}$ can also be interpreted as a (soft) decision. If sensor i is honest, we have $P(\mathbf{u_i} = \mathbf{z_i}|i = H) = 1$, otherwise we have $P(\mathbf{u_i} \neq \mathbf{z_i}|i = B) = P_A$. Here, the probability density function (PDF) of the local decision $\mathbf{u_i}$ if i is honest is given as

$$\begin{aligned} P(\mathbf{u_i}|i = H, \mathcal{H}_h) =& P(\mathbf{z_i}|i = H, \mathcal{H}_h) \\ =& \prod_{j=1}^{2^q} P(\mathbf{z_i} = \mathbf{v_j}|i = H, \mathcal{H}_h)^{I(\mathbf{z_i}, \mathbf{v_i})} \end{aligned} \tag{5}$$

for $h = 0, 1$, where

$$P(\mathbf{z_i} = \mathbf{v_j}|i = H, \mathcal{H}_h) = P(\tau_{i,j-1} \leq y_i \leq \tau_{i,j}|i = H, H_h) \tag{6}$$

based on (4) and $I(\mathbf{z_i}, \mathbf{v_i})$ is an indicator function that returns 1 if $\mathbf{z_i}$ is element-wise equal equal to $\mathbf{v_i}$ and returns 0 otherwise. In (5), we need to know the PDF of y_i, for $i = 1, 2, \ldots, N$. According to [38], both $y_i|\mathcal{H}_0$ and $y_i|\mathcal{H}_1$ follow Gaussian distributions as shown in (7), where $\beta_{i,0}^2 = \sigma_n^2$, $\beta_{i,1}^2 = \sigma_n^2 + p\sigma_x^2||\mathbf{h}_i||_2^2$ and $b \overset{a}{\sim} f(b)$ means variable b asymptotically follows PDF $f(b)$.

$$y_i|\mathcal{H}_0 \sim \mathcal{N}(0, \beta_{i,0}^2) \tag{7a}$$

$$y_i|\mathcal{H}_1 \overset{a}{\sim} \mathcal{N}(0, \beta_{i,1}^2), \tag{7b}$$

The proof of (7b) is provided in [[38], Appendix B], where the Lyapounov Central Limit Theorem (CLT) is utilized to derive the results. Let $A_{i,j,h}$ represent the probability that y_i falls within the range of $[\tau_{i,j-1}, \tau_{i,j}]$ when sensor i is honest under hypothesis $\mathcal{H}_h$, i.e., $P(\tau_{i,j-1} \leq y_i \leq \tau_{i,j}|i = H, \mathcal{H}_h)$. Then $A_{i,j,h}$ is given by

$$A_{i,j,h} = Q(\frac{\tau_{i,j-1}}{\beta_{i,h}}) - Q(\frac{\tau_{i,j}}{\beta_{i,h}}) \tag{8}$$

for $h = 0, 1$, where $Q(\cdot)$ denotes the tail distribution function of the standard normal distribution. If sensor i is Byzantine, $\mathbf{u_i}$ does not have to be equal to $\mathbf{z_i}$. The attack model for Byzantine nodes is illustrated in Fig. 2. According to the chain rule, the PDF of local decision $\mathbf{u_i}$ is given as (11), where

$$P(\mathbf{u_i} = \mathbf{v_j}|\mathbf{u_i} = \mathbf{z_i}, \mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) = \begin{cases} 1 & j = k \\ 0 & j \neq k, \end{cases} \tag{9}$$

[2] This assumption aligns with some related works such as [20], [35]–[37].

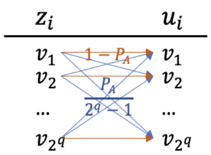

Fig. 2: Attack model for a Byzantine node i. With a probability of $P_A/(2^q-1)$, each Byzantine node decides to send a soft decision that differs from the one it believes to be correct. With probability $1-P_A$, the Byzantine nodes send the soft decision that they believe to be correct.

$$P(\mathbf{u_i}\!=\!\mathbf{v_j}|\mathbf{u_i}\!\neq\!\mathbf{z_i},\mathbf{z_i}\!=\!\mathbf{v_k},i\!=\!B,\mathcal{H}_h)\!=\!\begin{cases}0 & j\!=\!k\\ \frac{1}{2^q-1} & j\!\neq\!k,\end{cases} \quad (10)$$

$P(\mathbf{u_i}\neq\mathbf{z_i}|\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)=P_A$, $P(\mathbf{u_i}=\mathbf{z_i}|\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)=1-P_A$ and $P(\mathbf{z_i}=\mathbf{v_k}|i=B,\mathcal{H}_h)=Q(\frac{\tau_{i,k-1}}{\beta_{i,h}})-Q(\frac{\tau_{i,k}}{\beta_{i,h}})$ for $h=0,1$. Note that (9) and (10) are equivalent to $I(i,k)$ and $\frac{1-I(i,k)}{2^q-1}$, respectively. Hence, (11) can be rewritten as

$$\begin{aligned}
&P(\mathbf{u_i}|i=B,\mathcal{H}_h)=\prod_{j=1}^{2^q}P(\mathbf{u_i}=\mathbf{v_j}|i=B,\mathcal{H}_h)^{I(\mathbf{u_i},\mathbf{v_j})}\\
&=\prod_{j=1}^{2^q}\left\{\sum_{k=1}^{2^q}A_{i,k,h}\left[(1-P_A)I(j,k)+\frac{P_A(1-I(i,k))}{2^q-1}\right]\right\}^{I(\mathbf{u_i},\mathbf{v_j})}\\
&=\prod_{j=1}^{2^q}\left\{\sum_{k=1}^{2^q}A_{i,k,h}\left[(1-P_A-\frac{P_A}{2^q-1})I(j,k)+\frac{P_A}{2^q-1}\right]\right\}^{I(\mathbf{u_i},\mathbf{v_j})}\\
&=\prod_{j=1}^{2^q}\left\{A_{i,j,h}(1-P_A)+(1-A_{i,j,h})\frac{P_A}{2^q-1}\right\}^{I(\mathbf{u_i},\mathbf{v_j})}. \qquad (12)
\end{aligned}$$

Due to the statistical independence of the local decisions $\{u_1,u_2,\ldots,u_N\}$, we have

$$P(\mathbf{U}|\mathcal{H}_h)\!=\!\prod_{i=1}^{N}\prod_{j=1}^{2^q}\left[\sum_{X=B,H}P(\mathbf{u_i}\!=\!\mathbf{v_j}|i\!=\!X,\mathcal{H}_h)P(i\!=\!X)\right]^{I(\mathbf{u_i},\mathbf{v_j})} \quad (13)$$

for $h=0,1$.

III. GLRT and Quantized LMPT Detectors

In this section, we start with a brief review of the GLRT and the quantized LMPT detectors where all the sensors are assumed to be honest so that they send uncorrupted decisions to the FC, i.e., $\mathbf{u_i}=\mathbf{z_i}$. Then, the performance of the GLRT and the quantized LMPT detectors under Byzantine attacks is evaluated. The sparse signals here are characterized by the BG model. Under the BG model, the problem of distributed detection of sparse stochastic signals can be formulated as a problem of one-sided and close hypothesis testing which is given as

$$\begin{cases}\mathcal{H}_0: & p=0\\ \mathcal{H}_1: & p\to 0^+.\end{cases} \quad (14)$$

A. Fusion Rule for GLRT and Quantized LMPT Detectors with Honest Sensors

1) GLRT Detector: The fusion rule of the GLRT detector is given by

$$\frac{\max_p P(\mathbf{U}|\mathcal{H}_1;p)}{P(\mathbf{U}|\mathcal{H}_0;p=0)}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\lambda', \quad (15)$$

We can obtain the estimated sparsity degree $\hat{p}$ via maximum-likelihood estimation (MLE) which is given as $\hat{p}=\arg\max_p P(\mathbf{U}|\mathcal{H}_1;p)$. By replacing p by $\hat{p}$ in (15) and taking the logarithm of both sides of (15), the fusion rule can be expressed as

$$\Gamma_{GLRT}=\sum_{i=1}^{N}\sum_{j=1}^{2^q}I(\mathbf{z_i}=\mathbf{v_j})G_{i,j}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\lambda_1, \quad (16)$$

where $G_{i,j}=\hat{A}_{i,j,1}-\hat{A}_{i,j,0}$, $\hat{A}_{i,j,1}=Q(\frac{\tau_{i,j-1}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}})-Q(\frac{\tau_{i,j}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}})$ and $\hat{A}_{i,j,0}=A_{i,j,0}$.

2) Quantized LMPT Detector: Since the sparsity degree p is positive and close to zero under $\mathcal{H}_1$, and $p=0$ under $\mathcal{H}_0$, the problem of distributed detection of sparse stochastic signals can be performed via locally most powerful tests as shown in [10]. Firstly, the logarithm form of the LRT, which is given by

$$lnP(\mathbf{U}|\mathcal{H}_1;p)-lnP(\mathbf{U}|\mathcal{H}_0)\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}ln(p_0/p_1), \quad (17)$$

is considered for decision-making at the FC, where $P(\mathbf{U}|\mathcal{H}_h)=\prod_{i=1}^{N}P(\mathbf{u_i}|\mathcal{H}_h,i=H)$ and $P(\mathcal{H}_h)=p_h$ for $h=0,1$. Due to the fact that the sparsity degree p is close to zero, the first-order Taylor's series expansion of $lnP(\mathbf{U}|\mathcal{H}_1;p)$ around zero is given as

$$lnP(\mathbf{U}|\mathcal{H}_1;p)\!=\!lnP(\mathbf{U}|\mathcal{H}_1;p\!=\!0)\!+\!p\left(\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p}\right)_{p=0}. \quad (18)$$

By substituting (18) in (17), the test statistic of the quantized LMPT detector is given by

$$\left(\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p}\right)_{p=0}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\frac{ln(p_0/p_1)}{p}=\lambda_2, \quad (19)$$

where

$$\begin{aligned}\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p}&=\sum_{i=1}^{N}\frac{\partial lnP(\mathbf{u_i}|\mathcal{H}_1,i=H;p)}{\partial p}\\ &=\sum_{i=1}^{N}\sum_{j=1}^{2^q}w_{i,j}I(\mathbf{u_i}=\mathbf{v_j}) \qquad (20)\end{aligned}$$

and $w_{i,j}=\frac{\sigma_x^2||h_i||_2^2}{2\beta_{i,1}^3}\left[\tau_{i,j-1}\Phi(\frac{\tau_{i,j-1}}{\beta_{i,1}})-\tau_{i,j}\Phi(\frac{\tau_{i,j}}{\beta_{i,1}})\right]A_{i,j,1}^{-1}$. Here, $\Phi(\cdot)$ denotes the CDF of the standard normal distribution. Hence, the decision rule is given as

$$\Gamma_{LMPT}=\sum_{i=1}^{N}\sum_{j=1}^{2^q}I(\mathbf{u_i}=\mathbf{v_j})\widetilde{w}_{i,j}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\lambda_2, \quad (21)$$

where $\widetilde{w}_{i,j}=(w_{i,j})_{p=0}$.

$$
\begin{aligned}
&P(\mathbf{u_i}|i=B,\mathcal{H}_h)=\prod_{j=1}^{2^q}P(\mathbf{u_i}=\mathbf{v_j}|i=B,\mathcal{H}_h)^{I(\mathbf{u_i},\mathbf{v_j})}\\
&=\prod_{j=1}^{2^q}[\sum_{k=1}^{2^q}P(\mathbf{z_i}=\mathbf{v_k}|i=B,\mathcal{H}_h)P(\mathbf{u_i}=\mathbf{z_i}|\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)P(\mathbf{u_i}=\mathbf{v_j}|\mathbf{u_i}=\mathbf{z_i},\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)\\
&\quad+P(\mathbf{z_i}=\mathbf{v_k}|i=B,\mathcal{H}_h)P(\mathbf{u_i}\neq\mathbf{z_i}|\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)P(\mathbf{u_i}=\mathbf{v_j}|\mathbf{u_i}\neq\mathbf{z_i},\mathbf{z_i}=\mathbf{v_k},i=B,\mathcal{H}_h)]^{I(\mathbf{u_i},\mathbf{v_j})}
\end{aligned}
\tag{11}
$$

B. Performance Analysis of the GLRT and the Quantized LMPT Detectors in the Presence of Byzantines

In this subsection, we evaluate the detection performance of the GLRT and the quantized LMPT detectors in the presence of Byzantines. We also derive the optimal attack strategy of the Byzantines.

Let $L=\sum_{i=1}^{N}L_i$ denote the global statistic for the fusion rule given in (16) or (21), where $L_i=\sum_{j=1}^{2^q}I(\mathbf{u_i}=\mathbf{v_j})d_{i,j}$ and $d_{i,j}\in\{\widetilde{w}_{i,j},g_{i,j}\}$. According to the Lyapunov CLT, L approximately follows a Gaussian distribution with mean $E(\sum_{i=1}^{N}L_i)$ and variance $Var(\sum_{i=1}^{N}L_i)$ when N is sufficiently large. Under both hypotheses, $E(L)$ and $Var(L)$ are given as

$$
\begin{aligned}
E(L|\mathcal{H}_h)&=\sum_{i=1}^{N}E(L_i|\mathcal{H}_h)=\sum_{i=1}^{N}E\left(\sum_{j=1}^{2^q}I(\mathbf{u_i}=\mathbf{v_j})d_{i,j}\right)\\
&=\sum_{i=1}^{N}\sum_{j=1}^{2^q}P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h)d_{i,j}\\
&=\sum_{i=1}^{N}\sum_{j=1}^{2^q}[P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h,i=H)(1-\alpha)\\
&\quad+P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h,i=B)\alpha]d_{i,j}
\end{aligned}
\tag{22}
$$

and

$$
\begin{aligned}
Var(L|\mathcal{H}_h)&=\sum_{i=1}^{N}Var(L_i|\mathcal{H}_h)=\sum_{i=1}^{N}\left[E\left(L_i^2|\mathcal{H}_h\right)-E(L_i|\mathcal{H}_h)^2\right]\\
&=\sum_{i=1}^{N}E\left[\left(\sum_{j=1}^{2^q}I(\mathbf{u_i}=\mathbf{v_j})d_{i,j}\right)^2\right]-E(L|\mathcal{H}_h)^2\\
&=\sum_{i=1}^{N}\sum_{j=1}^{2^q}P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h)d_{i,j}^2-E(L|\mathcal{H}_h)^2\\
&=\sum_{i=1}^{N}\sum_{j=1}^{2^q}[P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h,i=H)(1-\alpha)\\
&\quad+P(\mathbf{u_i}=\mathbf{v_j}|\mathcal{H}_h,i=B)\alpha]d_{i,j}^2-E(L|\mathcal{H}_h)^2,
\end{aligned}
\tag{23}
$$

respectively. Using the expression in (22) and (23), the probabilities of detection and false alarm can be calculated as

$$
P_d=P(L>\lambda|\mathcal{H}_1)=Q\left(\frac{\lambda-E(L|\mathcal{H}_1)}{\sqrt{Var(L|\mathcal{H}_1)}}\right)
\tag{24}
$$

and

$$
P_f=P(L>\lambda|\mathcal{H}_0)=Q\left(\frac{\lambda-E(L|\mathcal{H}_0)}{\sqrt{Var(L|\mathcal{H}_0)}}\right),
\tag{25}
$$

respectively, where $\lambda\in\{\lambda_1,\lambda_2\}$.

Next, we investigate the optimal attack strategy that can be adopted by Byzantines. From the attackers' perspective, the optimal strategy is to render the system blind, aiming to achieve a probability of detection equal to 1/2. To determine the optimal attack strategy, we utilize the deflection coefficient, which provides a simple and yet effective measure of the global probability of detection. The deflection coefficient is defined as $D_f=\frac{(E(L|\mathcal{H}_1)-E(L|\mathcal{H}_0))^2}{Var(L|\mathcal{H}_1)}$. Thus, to blind the FC, Byzantines need to strategically design the attack parameters so that $D_f=0$, i.e., $E(L|\mathcal{H}_1)=E(L|\mathcal{H}_0)$. By utilizing (22), we can obtain

$$
\alpha P_A=\frac{\sum_{i=1}^{N}\sum_{j=1}^{2^q}(A_{i,j,1}-A_{i,j,0})d_{i,j}}{\sum_{i=1}^{N}\sum_{j=1}^{2^q}\left[\frac{1}{2^q-1}+(1-\frac{1}{2^q-1})(A_{i,j,1}-A_{i,j,0})\right]d_{i,j}}.
\tag{26}
$$

When αP_A equals the right-hand side of (26), the attackers are able to blind the FC. From the simulation results presented later in Sec. V, both the GLRT and the quantized LMPT detectors are very vulnerable to Byzantine attacks, even if the attack parameter P_A is very small. A possible explanation is that, since detectors make their decisions based on observations with the same mean and slightly different variances under the two hypotheses, it is easy for them to make incorrect decisions in the presence of Byzantines.

IV. Resilient fusion rule

In order to improve the resilience of the detector, we attempt to elicit some additional information regarding the attack parameters from the local decisions of some sensors and incorporate it into the design of the fusion rule. In general, a detector's performance improves as additional information is obtained, e.g., sparsity degree p, the fraction of Byzantines α, and attack probability P_A. Intuitively, a GLRT detector can be designed, which takes both the unknown sparsity degree and the unknown attack parameters into consideration, as shown in (27).

$$
\frac{\max_{p,P_A,\alpha}P(\mathbf{U}|\mathcal{H}_1;p)}{\max_{P_A,\alpha}P(\mathbf{U}|\mathcal{H}_0;p=0)}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\lambda''.
\tag{27}
$$

If we assume that the sparse signals are weak and the number of sensors is large, the MLE attains its asymptotic PDF, and an appropriate threshold λ'' can be found based on the asymptotic detection performance of the GLRT detectors (see Sec. 6.5 in [39]). However, sparse signals need not be weak. In that case, it is not tractable to obtain an appropriate threshold value λ''. Moreover, the presence of nuisance parameters P_A and α results in a degradation of the detection performance of GLRT detectors.

To overcome these problems, as alluded to earlier, we randomly select a fraction of the sensors as reference sensors from the set of all sensors and estimate unknown parameters (i.e., α, P_A and p) in two steps. In the first step, nuisance attack parameters are estimated based on the local decisions coming from the reference sensors. In the second step, the estimated attack parameters are utilized to estimate the unknown sparsity degree p based on the local decisions from the remaining sensors. The proposed GLRTRS detector is based on the above parameter estimates. As the LMPT-based detector does not require the knowledge of the sparsity degree p, the only estimation occurs in the first step, which is the estimation of the nuisance attack parameters. Later in this section, we will provide details about the proposed GLRTRS and LMPTRS detectors.

Since we carry out the entire estimation process in two steps, we would like to minimize the performance loss caused by partitioning the estimation process. Let us take the GLRT detector presented in (27) as an example. Suppose we want to partition the entire estimation process into two steps, as described above. In that case, we want to ensure that the performance degradation caused by the unknown sparsity degree p is negligible while estimating the attack parameters. In other words, the two pairs of estimated attack parameters we obtain, which are $\{\alpha_{H_1}, P_{A,H_1}\} = \arg\max_{\alpha, P_A} P(\mathbf{U}|\mathcal{H}_1, p, \alpha, P_A)$ and $\{\alpha_{H_0}, P_{A,H_0}\} = \arg\max_{\alpha, P_A} P(\mathbf{U}|\mathcal{H}_0, p = 0, \alpha, P_A)$, should be very close to each other. To complete this task, we introduce reference sensors to assist us. We randomly select a set of reference sensors from the set of all the sensors to estimate the unknown nuisance attack parameters P_A and α.[3] At the reference sensors, we employ different predefined thresholds so that the decisions of the reference sensors satisfy Assumption 1 below.

Assumption 1: The probability $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_h)$ (or $Pr(\mathbf{z_i} = v_1|\mathcal{H}_h)$) is approximately equal to 1 for $h = 0, 1$.

Note that the condition in Assumption 1 can be attained when reference sensors send v_{2^q} (or v_{2^1}) with a probability that is close to 1, regardless of the underlying true hypothesis $\mathcal{H}_h$. To satisfy Assumption 1, one of the simplest methods is to either set $\widetilde{\tau}_{j,2^q-1} \ll \tau_{i,1}$ or $\tau_{i,2^q} \ll \widetilde{\tau}_{j,1}$. This is because the limit $\lim_{\widetilde{\tau}_{j,2^q-1} \to -\infty} Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_h) = 1$ (or $\lim_{\widetilde{\tau}_{j,1} \to +\infty} Pr(\mathbf{z_i} = v_1|\mathcal{H}_h) = 1$) always holds.[4] Therefore, if Assumption 1 is satisfied, it is highly likely that the reference sensors will continue to send the same decision regardless of the true underlying hypothesis. It allows us to ensure that the performance degradation caused by the unknown sparsity degree p is negligible while the attack parameters are being estimated.

In the following subsections, we consider two cases: (i) The sparsity degree p and the attack parameters $\{\alpha, P_A\}$ are all unknown; (ii) α is known, but sparsity degree p and attack probability P_A are unknown.

A. Networks with Unknown p, α and P_A

Two detectors are proposed in this subsection: the GLRTRS detector that requires the estimation of unknown parameter p, and the LMPTRS detector that does not require the estimation of p.

a) GLRTRS detector: According to (13), we are able to obtain

$$P(\mathbf{U}|\mathcal{H}_h) = \prod_{i=1}^{N}\prod_{j=1}^{2^q}\left[A_{i,j,h} + x\left(\frac{1}{2^q-1} - A_{i,j,h} - \frac{A_{i,j,h}}{2^q-1}\right)\right]^{I(\mathbf{u_i},\mathbf{v_j})} \tag{28}$$

where $x = \alpha P_A$. For convenience, instead of considering the two attack parameters α and P_A separately, we consider a single attack parameter x. The problem of distributed detection of a sparse stochastic signal can be formulated as

$$\begin{cases} \mathcal{H}_0: & p = 0, 0 \le x \le 1 \\ \mathcal{H}_1: & p \to 0^+, 0 \le x \le 1 \end{cases}. \tag{29}$$

The fusion rule of the GLRTRS detector is given by

$$\frac{\max_p \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, \hat{x})}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p = 0, \hat{x})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda, \tag{30}$$

where N_{ref} is the number of reference sensors and they are labelled as $1, 2, 3 \ldots, N_{ref}$. The estimate of the unknown attack parameter x, i.e., $\hat{x}$ is made via MLE based on the reference sensors data. Here, the estimated attack parameter x is given as

$$x_{H_h} = \arg\max_{x} P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) \tag{31}$$

for $h = 0, 1$. $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ in (31) is the joint pmf of local decisions coming from the reference sensors and it is given as

$$\begin{aligned} &P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) \\ &= \prod_{i=1}^{N_{ref}}\prod_{j=1}^{2^q}\left[\sum_{X=B,H} P(\mathbf{u_i} = \mathbf{v_j}|i = X, \mathcal{H}_h)P(i = X)\right]^{I(\mathbf{u_i},\mathbf{v_j})} \\ &= \prod_{i=1}^{N_{ref}}\prod_{j=1}^{2^q}\left[C_{i,j,h} + x\left(\frac{1}{2^q-1} - C_{i,j,h} - \frac{1}{2^q-1}C_{i,j,h}\right)\right]^{I(\mathbf{u_i}=\mathbf{v_j})} \end{aligned} \tag{32}$$

for $h = 0, 1$, where $C_{i,j,h} = Q(\frac{\widetilde{\tau}_{i,j-1}}{\beta_{i,h}}) - Q(\frac{\widetilde{\tau}_{i,j}}{\beta_{i,h}})$.

Note that if Assumption 1 holds employed at any q-bit quantizer of reference sensors, i.e., $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1) \approx Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0) \approx 1$ for any reference sensor i, the absolute value of

[3] Since we have assumed that α fraction of Byzantine nodes are uniformly distributed in the network, there are α fraction of Byzantine nodes within both the set of reference sensors and remaining sensors.

[4] Based upon (7), the observation y_i for $i \in \{1, 2, \ldots, N\}$ has zero mean and different variances that are related to sparsity degree p given different hypotheses. Since a sparse signal is considered in the paper for which the sparsity degree p tends to 0, it is possible to design reasonable quantizer thresholds for reference nodes. A reasonable quantizer threshold refers to a quantizer threshold that is not excessively large or small. From experiments, it has been shown that $\tau_{i,1} - \widetilde{\tau}_{j,2^q-1} = 6$ (or $\widetilde{\tau}_{j,1} - \tau_{i,2^q} = 6$) is sufficient to satisfy Assumption 1 for the reference sensors.

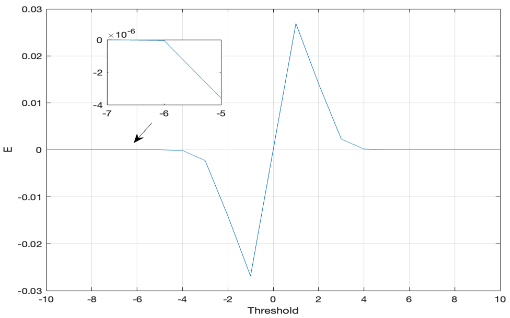

Fig. 3: E versus $\widetilde{\tau}_{j,2^q-1}$ given $p = 0.05$, $\sigma_x^2 = 5$, $\sigma_n^2 = 5$, $q = 1$ and $||\mathbf{h_i}||_2 = 1$ for all i.

$\widetilde{\tau}_{j,2^q-1}$ will be sufficiently large, thus, the difference between the probabilities $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1)$ and $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0)$ will be really small. Let $E_i = Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1) - Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0)$ denote the difference between the probabilities of local decisions under $\mathcal{H}_1$ and $\mathcal{H}_0$ for any reference sensor i. According to Eq. (7), we have $E_i = Q(\frac{\widetilde{\tau}_{j,2^q-1}}{\beta_{i,1}}) - Q(\frac{\widetilde{\tau}_{j,2^q-1}}{\beta_{i,0}})$. The values of E_i as a function of $\widetilde{\tau}_{j,2^q-1}$ are shown in Fig. 3. We can observe that for a sufficiently large (or small) value of $\widetilde{\tau}_{j,2^q-1}$, for example, $\widetilde{\tau}_{j,2^q-1} = -6$, E becomes significantly small, with $E < 10^{-6}$.

Based on the above discussion, we can easily derive $P(u_i|\mathcal{H}_h, p, x) \approx (1-x)^{I(\mathbf{u_i}=\mathbf{v}_{2^q})} \prod_{j=1}^{2^q-1} (\frac{x}{2^q-1})^{I(\mathbf{u_i}=\mathbf{v_j})}$ for any reference sensor i. So the difference between the estimated x under different hypotheses will be significantly small and can be assumed negligible, i.e., $x_{H_0} \approx x_{H_1}$. This result is employed in the following theorem stating that the estimator considered in (31) is an efficient MLE when Assumption 1 is satisfied.

Theorem 4.1: The MLE of the unknown attack parameter x based on the data from the reference sensors is unbiased, and it attains the Cramér–Rao lower bound (CRLB) of the problem, which equals $\frac{(1-x)x}{N_{ref}}$.

Proof: Please see Appendix A. ■

By replacing $\hat{x}$ by x_{H_1} in $P(\mathbf{u_i}|\mathcal{H}_1, p, x_{H_1})$ and $\hat{x}$ by x_{H_0} in $P(\mathbf{u_i}|\mathcal{H}_1, p = 0, x_{H_0})$ in (30), the fusion rule can be reformulated as

$$\frac{\max_p \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_{H_1})}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p = 0, x_{H_0})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa, \tag{33}$$

where $P(\mathbf{u_i}|\mathcal{H}_h, p, x_{H_h}) = \prod_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, p, x_{H_h})$. Since x_{H_0} is approximately the same as x_{H_1}, i.e., $x_{H_0} \approx x_{H_1}$, choosing x_{H_0} or x_{H_1} as the estimated x under both hypotheses, or choosing the average of x_{H_0} and x_{H_1} as the estimated x under both hypotheses are all acceptable options. Here, we opt to replace both x_{H_1} and x_{H_0} in (33) with their averaged estimate $x_H = \frac{x_{H_1}+x_{H_0}}{2}$. The fusion rule then can be simplified as follows:

$$\frac{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_H)}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p = 0, x_H)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa, \tag{34}$$

where κ is the threshold to be set in order to ensure the desired probability of false alarm PFA. Next, we calculate the estimated sparsity degree $\hat{p}$, which is given as $\hat{p} = \arg\max_p \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_H)$. Upon taking the logarithm of both sides of (34), the simplified fusion rule is given as

$$\Gamma_{GLRTRS} = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) F_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa', \tag{35}$$

where $\kappa' = \log(\kappa)$, $F_{i,j} = f_{i,j,1} - f_{i,j,0}$, $f_{i,j,h} = \hat{A}_{i,j,h} + x_H \left(\frac{1}{2^q-1} - \hat{A}_{i,j,h} - \frac{1}{2^q-1}\hat{A}_{i,j,h} \right)$, $\hat{A}_{i,j,1} = Q(\frac{\tau_{i,j-1}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}}) - Q(\frac{\tau_{i,j}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}})$ and $\hat{A}_{i,j,0} = A_{i,j,0}$. Assume that $N - N_{Nef}$ is sufficiently large, the global statistic Γ_{GLRTRS} then follows a Gaussian distribution with mean

$$E(\Gamma_{GLRTRS}|H_h) = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} F_{i,j} P(\mathbf{u_i} = \mathbf{v_j}|H_h, x_H, p) \tag{36}$$

and variance

$$\begin{aligned} Var(\Gamma_{GLRTRS}|H_h) = & \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} F_{i,j}^2 P(\mathbf{u_i} = \mathbf{v_j}|H_h, x_H, p) \\ & - E^2(\Gamma_{GLRTRS}|H_h) \end{aligned} \tag{37}$$

for $h = 0, 1$. With (36) and (37), the probabilities of detection and false alarm are respectively given as

$$P_d = Q\left(\frac{\kappa' - E(\Gamma_{GLRTRS}|H_1)}{\sqrt{Var(\Gamma_{GLRTRS}|H_1)}} \right), \tag{38}$$

$$P_f = Q\left(\frac{\kappa' - E(\Gamma_{GLRTRS}|H_0)}{\sqrt{Var(\Gamma_{GLRTRS}|H_0)}} \right). \tag{39}$$

For a given false alarm PFA, we can obtain the suboptimal adaptive threshold used by the FC as shown in (40).[5]

$$\kappa' = Q^{-1}(PFA)\sqrt{Var(\Gamma_{GLRTRS}|H_0)} + E(\Gamma_{GLRTRS}|H_0) \tag{40}$$

b) LMPTRS detector: Similarly, after we obtain the estimated attack parameter x_H, the test statistic of the proposed LMPTRS detector can be expressed as

$$\left(\frac{\partial ln P(\mathbf{U}|\mathcal{H}_1, p, x_H)}{\partial p} \right)_{p=0} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{ln(p_0/p_1)}{p}, \tag{41}$$

where

$$\begin{aligned} \frac{\partial ln P(\mathbf{U}|\mathcal{H}_1, p, x_H)}{\partial p} &= \sum_{i=1}^{N} \frac{\partial ln P(\mathbf{u_i}|\mathcal{H}_1, p, x_H)}{\partial p} \\ &= \sum_{i=1}^{N} \sum_{j=1}^{2^q} \frac{\sigma_x^2 ||h_i||_2^2 I(\mathbf{u_i} = \mathbf{v_j})}{2(p\sigma_x^2 ||h_i||_2^2 + \sigma_n^2)^{\frac{3}{2}}} \left[\tau_{i,j-1} \Phi(\frac{\tau_{i,j-1}}{\sqrt{p\sigma_x^2 ||h_i||_2^2 + \sigma_n^2}}) \right. \\ &\quad \left. - \tau_{i,j} \Phi(\frac{\tau_{i,j}}{\sqrt{p\sigma_x^2 ||h_i||_2^2 + \sigma_n^2}}) \right] \frac{1 - x_H - x_H A_{i,j,1}}{A_{i,j,1} + x_H(1 - x_H - x_H A_{i,j,1})} \\ &= \sum_{i=1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) g_{i,j}. \end{aligned} \tag{42}$$

[5] Since we obtain the adaptive threshold based on the estimated attack parameter, it is a suboptimal threshold that approximately satisfies a desired false alarm.

The fusion rule can be reformulated as

$$\Gamma_{LMPTRS}=\sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})\widetilde{g}_{i,j} \overset{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \gamma', \tag{43}$$

where $\gamma' = \frac{ln(p_0/p_1)}{p}$ and $\widetilde{g}_{i,j} = (g_{i,j})_{p=0}$. Like the one employed earlier, we can derive the threshold γ' in (43) for a given false alarm PFA. We can obtain that $\gamma' = Q^{-1}(PFA)\sqrt{Var(\Gamma_{LMPTRS}|H_0)} + E(\Gamma_{LMPTRS}|H_0)$, where $E(\Gamma_{LMPTRS}|H_0) = \sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q}\widetilde{w}_{i,j}P(\mathbf{u_i} = \mathbf{v_j}|H_0, x_H, p = 0)$ and $Var(\Gamma_{LMPTRS}|H_0) = \sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q}\widetilde{w}_{i,j}^2 P(\mathbf{u_i} = \mathbf{v_j}|H_0, x_H, p = 0) - E^2(\Gamma_{LMPTRS}|H_0)$.

B. Networks with Known α, Unknown p and Unknown P_A

When it is assumed that we know the fraction of Byzantine nodes α in the network, we can obtain more accurate information and achieve better detection performance. In this subsection, the GLRTRS and the LMPTRS detectors are further enhanced by introducing a local decision filter at the FC, which allows us to select sensors that are more likely to be honest. The proposed enhanced detectors are referred to as the E-GLRTRS and the E-LMPTRS detectors.

Upon receiving local decisions $\{\mathbf{U}(1),\ldots,\mathbf{U}(t)\}$ until time step t, where $\mathbf{U}(t) = \{\mathbf{u_1}(t),\ldots,\mathbf{u_N}(\mathbf{t})\}$, each sensor's statistical behavior is used to filter local decisions. The local decision filter distinguishes malicious nodes from honest nodes at time t by the following

$$\sum_{j=1}^{2^q}|R_j-\widetilde{p}_t(\mathbf{u_i}=\mathbf{v_j})| \overset{b_i(t)=1}{\underset{b_i(t)=0}{\gtrless}} \tau, \forall i\in\{N_{ref}+1,\ldots,N\}, \tag{44}$$

where $R_j = min(P(\mathbf{u_i}=\mathbf{v_j}|i=H,\mathcal{H}_1), P(\mathbf{u_i}=\mathbf{v_j}|i=H,\mathcal{H}_0))$ is a benchmark value to filter out the potential malicious sensors[6] and $b_i(t)$ represents the behavioral identity of sensor i at time t. If $b_i(t) = 1$, the sensor i is regarded as an honest node; otherwise, it is regarded as a potential Byzantine node. $\widetilde{p}_t(\mathbf{u_i} = \mathbf{v_j})$ is the empirical probability of $\mathbf{u_i} = \mathbf{v_j}$ until time step t according to the history of local decisions and it is given as

$$\widetilde{p}_t(\mathbf{u_i}=\mathbf{v_j}) = \frac{\sum_{q=1}^{t} I(\mathbf{u_i}(q), \mathbf{v_j})}{t}, \tag{45}$$

where $\mathbf{u_i}(q)$ is the $\mathbf{u_i}$ at time step q. The left side of (44) measures the deviation of the empirical probability of $\mathbf{u_i} = \mathbf{v_j}$ from the benchmark value R_j. Sensors are potential Byzantine nodes if the deviation exceeds a predefined threshold τ. Based on the behavioral identity of all the sensors $\{b_i(t)\}_{i=1}^{N}$ at time step t, we can obtain the fusion rules of enhanced detectors. Note that both GLRTRS and LMPTRS have the form

$$\sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})W_{i,j} \overset{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \eta, \tag{46}$$

where $(W_{i,j},\eta)\in\{(\widetilde{g}_{i,j},\gamma'),(F_{i,j},\kappa')\}$. Hence, the enhanced fusion rule at time step t is given by

$$\Gamma_E(t) = \sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q} b_i(t)I(\mathbf{u_i}(t)=\mathbf{v_j})W_{i,j}(t) \overset{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \eta(t). \tag{47}$$

Let $\alpha_t(t)$ and $P_A(t)$ denote the probability that a sensor is a Byzantine node and the probability that a Byzantine node attacks at time step t, respectively, and let α be the initial value of α_t. We first obtain the estimated attack probability $\hat{p}_A(0) = x_H(0)/\alpha$ at time $t = 0$ as initial value of $\hat{P}_A$, where $x_H(0) = \frac{x_{H_1}(0)+x_{H_0}(0)}{2}$ and $x_{H_h}(0)$ is given in (31) for $h = 0, 1$. After filtering the possible Byzantine nodes, the value of α_t at time step $t = 0$ is updated according to $\{b_i(0)\}_{i=N_{ref}+1}^{N}$. The updating rule is given as

$$\alpha_t(0) = \alpha - \frac{\sum_{i=N_{ref}+1}^{N} b_i(0)}{N - N_{ref}}. \tag{48}$$

At the next time step, the updated $\alpha_t(0)$ is employed as the new prior to estimate $\hat{p}_A(2)$ and $\hat{P}_A(1) = \frac{\sum_{i=0}^{1}\hat{p}_A(i)}{2}$. The value of α_t is also updated at time step $t = 1$ according to $\{b_i(1)\}_{i=N_{ref}+1}^{N}$ in the same manner as (48), i.e., $\alpha(1) = \alpha(0) - \frac{\sum_{i=N_{ref}+1}^{N} b_i(1)}{N-N_{ref}}$, and becomes the new prior at the next time step. Thus, at time step t, $\alpha_t(t-1) = \alpha_t(t-2) - \frac{\sum_{i=N_{ref}+1}^{N} b_i(t-1)}{N-N_{ref}}$ is utilized to obtain $\hat{P}_A(t) = \frac{\sum_{i=0}^{t}\hat{p}_A(i)}{t+1}$. By replacing x_H and $F_{i,j}$ with $X_H(t) = \hat{P}_A(t)\alpha_t(t-1)$ and $b_i(t)W_{i,j}$, respectively, in (36) and (37), we can obtain $E(\Gamma_E(t)|H_h)$ and $Var(\Gamma_E(t)|H_h)$. Similarly, for a given false alarm PFA, we can obtain the threshold used by the FC at time step t, which is given as $\eta(t) = Q^{-1}(PFA)\sqrt{Var(\Gamma_E(t)|H_0)} + E(\Gamma_E(t)|H_0)$. To compare the detectors over all of the scenarios we consider, we provide a summary table shown in Table I.

V. Simulation results

In this section, we present the simulation results to evaluate the performance of the proposed detectors in the presence of Byzantine attacks and compare them with the quantized LMPT-based detector (proposed in [10]) and the commonly used GLRT-based detector. Via simulations, we analyze the performance of the proposed schemes in terms of the probability of error in the system. The channel gains $\{\mathbf{h_i}\}_{i=1}^{N}$ are all assumed to be sampled from normal distribution with a homogeneous scenario so that $||\mathbf{h_i}||_2 = 1, \forall i$ as described in [10]. Table II presents the parameter settings for reference. Unless otherwise noted, we assume the number of sensors N to be 280. When reference sensors are employed, we employ $N_{ref} = 80$ out of 280 sensors as reference sensors, except when we evaluate system performance as a function of N_{ref}.

TABLE II: Summary of parameter settings.

	N	N_{ref}	σ_n^2	σ_x^2	$\|\|\mathbf{h_i}\|\|_2$
value	280	80	1	5	1
	α	PFA	π_1	μ_w	p
value	0.3	0.4	0.5	0	0.05

[6] Note that based upon (7), the observation $y_i, \forall i \in \{1, 2, \ldots, N\}$ has zero mean and different variances that are related to the sparsity degree p given different hypotheses. Regardless of the quantizer thresholds that have been chosen, sensors tend to transmit the same decisions with slightly different probabilities based upon different hypotheses, i.e, $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_1)$ and $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_0)$ are slightly different. The simplest method of choosing R_j is to take the minimum value between $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_1)$ and $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_0)$.

TABLE I: Summary of GLRT-based and LMPT-based detectors under different scenarios.

unknown $\{P_A, \alpha, p\}$	known α and unknown $\{P_A, p\}$
GLRTRS: $\sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) F_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa'$	E-GLRTRS: $\sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} b_i(t) I(\mathbf{u_i}(t) = \mathbf{v_j}) F_{i,j}(t) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa'(t)$
LMPTRS: $\sum_{i=1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) \widetilde{g}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma'$	E-LMPTRS: $\sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} b_i(t) I(\mathbf{u_i}(t) = \mathbf{v_j}) \widetilde{g}_{i,j}(t) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma'(t)$
commonly used GLRT-based detector: $\sum_{i=1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) G_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_1$	
LMPT-based detector [10]: $\sum_{i=1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) \widetilde{w}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_2$	

In Fig. 4, we demonstrate the error probabilities of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$, the GLRT detector, and the proposed GLRTRS detector. Two different quantizers are employed, i.e., $q = 1$ and $q = 2$. The error probability of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$ shown in Fig. 4 is used as the benchmark to assess the performance of the proposed detectors. It can be observed that the GLRT detector is extremely vulnerable to attacks for both one-bit quantization and multilevel quantization, and a small fraction of Byzantine nodes α with a small attack parameter P_A are sufficient to break down the entire system. However, the proposed GLRTRS detector can obtain an error probability close to that of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$. We can observe from Fig. 4 that in the cases of $q = 1$ and $q = 2$, the GLRTRS detector outperforms the commonly used GLRT-based detector in the presence of attacks, with a performance close to the benchmark LRT detector. Note that the GLRTRS detector uses only 200 sensors for detection purposes and exhibits performance close to the benchmark detector that uses 280 sensors for detection purposes. Hence, when no attacks are present, the commonly used GLRT-based detector performs slightly better. The number of quantization levels also affects the performance of the GLRTRS detector. As shown in Fig. 4, with an increase in q, the error probability of the proposed GLRTRS detector further decreases due to the reduction of performance losses caused by quantization. From Fig. 4, we can also observe that the difference between the benchmark error probability and the error probability of the proposed GLRTRS detector is larger when the value of q increases. It is because the GLRTRS detector is a sub-optimal detector, while the benchmark LRT detector is an optimal one.

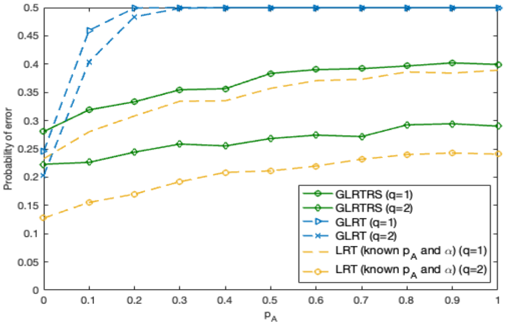

Fig. 4: Comparison of Pe for the GLRTRS, LRT and GLRT detectors.

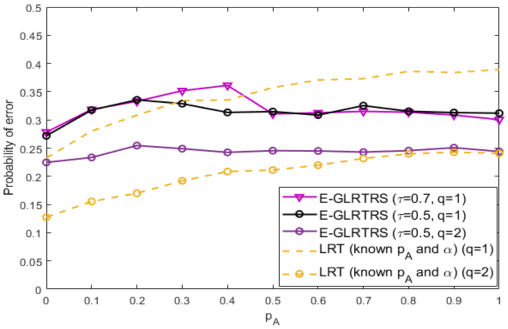

Fig. 5: Pe versus P_A when different values of q and the different values of threshold τ are utilized for the E-GLRTRS detectors.

If we assume that the fraction of Byzantine nodes α is known to the system, The error probability of the system can be further reduced by employing the E-GLRTRS detector. As shown in Fig. 5, the error probability of the E-GLRTRS detector decreases with an appropriately designed threshold τ compared to the GLRTRS detector. We can filter out different numbers of potential Byzantine nodes with different values of the threshold τ in (44). A potential Byzantine node can be either an actual Byzantine or a falsely identified one. It is obvious that a smaller threshold results in greater false filtering, while a larger threshold results in greater miss filtering. False filtering implies that honest nodes are falsely filtered out, whereas miss filtering implies that malicious nodes remain unfiltered. Both false filtering and miss filtering result in degrading the system's performance. Therefore, the system will likely perform better if the threshold τ is set appropriately. As shown in Fig. 5, $\tau = 0.5$ is more appropriate than $\tau = 0.7$. It can be observed that when $\tau = 0.5$, $q = 1$ and $P_A > 0.3$, the E-GLRTRS detector outperforms the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$. This is because the E-GLRTRS detector filters out potential Byzantine nodes and utilizes the rest of the sensors for detection. In contrast, the benchmark LRT detector utilizes all the sensors for detection purposes. Although the E-GLRTRS detector is inferior to the benchmark LRT detector when $q = 1$ and $P_A < 0.3$, the difference in error probabilities is not too significant.

In Fig. 6, the error probability and the convergence rate of the GLRTRS detector with different number of reference nodes are presented. The number of sensors used for detection purposes in the GLRTRS detectors with different values of

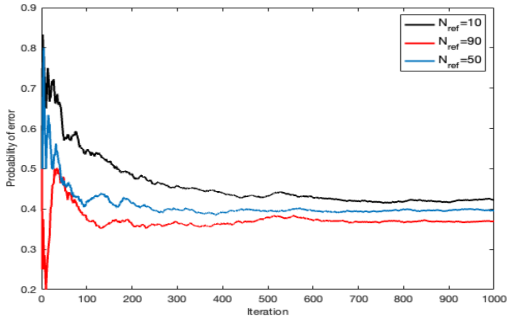

Fig. 6: Pe versus the number of iterations when different values of N_{ref} are utilized for the GLRTRS detector.

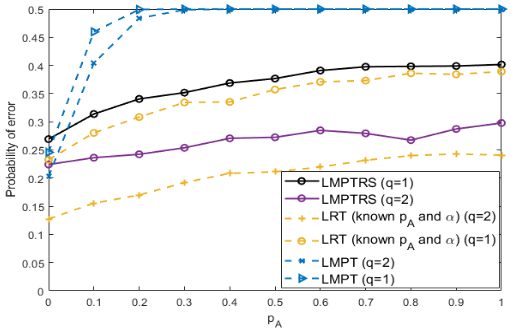

Fig. 7: Comparison of Pe for the LMPTRS, LRT and quantized LMPT detectors.

N_{ref} are equal to 200, i.e., $N - N_{ref} = 200$. It can be observed that the convergence rate is faster, and the error probability is lower when more reference nodes are used.

Fig. 7 shows the error probabilities of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$, the quantized LMPT detector (proposed in [10]) and the proposed LMPTRS detector for $q = 1$ and $q = 2$, respectively. We can observe that the quantized LMPT detector proposed in [10] is also extremely vulnerable to attacks for both one-bit and multilevel quantization when all the p, P_A and α are unknown. However, it can be observed that when $q = 1$, the proposed LMPTRS detector is capable of obtaining an error probability close to the benchmark error probability that is obtained by employing the LRT detector with perfect knowledge of the attack parameters $\{P_A, \alpha, p\}$. Similar to the conclusion we obtained from Fig. 4, the LMPTRS detector outperforms the quantized LMPT detector proposed in [10] in the presence of attacks. The error probability of the proposed LMPTRS detector decreases with increasing q, and a higher value of q increases the difference between the benchmark error probability and the proposed LMPTRS detector error probability. It is also possible to further reduce the error probability of the system by assuming that the fraction of Byzantine nodes α is known to the system. As shown in Fig. 8, the E-LMPTRS detector outperforms both the quantized LMPT detector and the benchmark LRT detector with perfect knowledge of the attack parameters by filtering potential Byzantine nodes when $q = 1$. When q increases (e.g., $q = 2$), the E-LMPTRS detector still outperforms the quantized LMPT detector. In Fig. 9, we demonstrate the performance of our proposed detectors, which were originally designed for the simple Gaussian case, in the presence of one realization of generalized Gaussian noise. The noise here is assumed to follow the Laplace distribution, which is a special case of the generalized Gaussian distribution with parameter $\beta = 1$. We also note that according to [10], all types of generalized Gaussian distributed high-dimensional sparse signals asymptotically follow Gaussian distributions. We can observe that our proposed detector exhibits a certain level of resilience to the Byzantine attack when the tail of the distribution is not heavy.

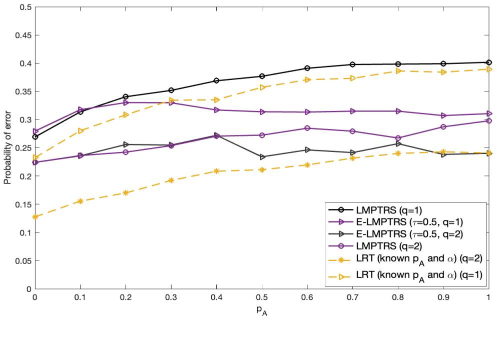

Fig. 8: Pe versus P_A when different values of q are utilized for the LMPTRS and the E-LMPTRS detectors.

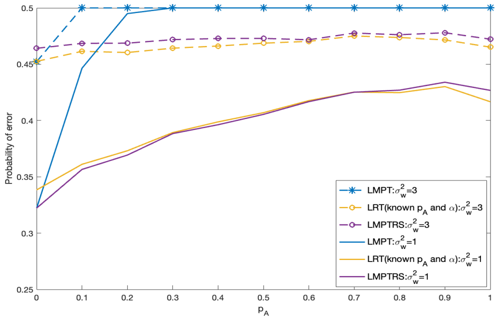

Fig. 9: P_e versus P_A for benchmark LRT, LMPT and LMPTRS detectors under Laplace distributed noise. The noise has a mean of $\mu_w = 0$ and a variance of σ_w^2 with probability of false alarm ($PFA = 0.4$). The sparse signals are assumed to asymptotically follow Gaussian distribution with mean 0 and variance $p\sigma_x^2||\mathbf{h}_i||_2^2$.

VI. Conclusion

The distributed detection problem of sparse stochastic signals with quantized measurements in the presence of Byzantine attacks was investigated in this paper. The sparse stochastic signals were characterized by their sparsity degrees, and the BG distribution was utilized to model sparsity. We proposed the LMPTRS and GLRTRS detectors with adaptive thresholds, given that the sparsity degree p and the attack parameters, i.e., α and P_A are unknown. The simulation results showed that the LMPTRS and GLRTRS detectors outperformed the LMPT detector under attack and achieved detection performance close to the benchmark LRT detector with perfect knowledge of the attack parameters and sparsity degree p. When the fraction of Byzantines α in the networks is assumed to be known, the E-LMPTRS and E-GLRTRS detectors were proposed to further improve the detection performance of the system by filtering out potential malicious sensors. Simulation results showed that the proposed enhanced detectors outperform LMPTRS and GLRTRS detectors.

In this work, the predefined quantizer thresholds we utilized come from [9]. In the future, we intend to consider the optimization of the predefined quantizer thresholds for our proposed detectors and the design of resilient quantized LMPT detector under noisy channels.

Appendix A
Proof of Theorem 4.1

We first consider the scenario where sensors send binary decisions to the FC, i.e., $q = 1$. After that, we consider the system where sensors send q-bit decisions to the FC ($q \geq 2$). Here, we only consider the assumption that $\widetilde{\tau}_{j,2^q} \ll \tau_{i,1}$. Nevertheless, we can reach similar conclusions if we assume $\tau_{i,2^q} \ll \widetilde{\tau}_{j,1}$.

1) When sensors send binary decisions (q=1): The joint pmf of local decisions coming from the reference sensors under hypothesis $\mathcal{H}_h$ is given as $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) = \prod_{i=1}^{N_{ref}} (1-x)^{\mathbf{u_i}} x^{1-\mathbf{u_i}}$ for $h = 0, 1$. Take the logarithm of both sides, we have

$$\begin{aligned}\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) =& \sum_{i=1}^{N_{ref}} [\mathbf{u_i} \log(1-x) + (1-\mathbf{u_i}) \log x] \\ =& Y \log(1-x) + (N_{ref} - Y) \log x, \end{aligned} \tag{49}$$

where $Y = \sum_{i=1}^{N_{ref}} \mathbf{u_i}$. Let $\frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x} = 0$, we are able to obtain the estimated attack parameter $\hat{x}_h$ under hypothesis $\mathcal{H}_h$ which maximizes $\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ and the estimated attack parameter $\hat{x}_h$ is given as $\hat{x}_h = 1 - \frac{Y}{N_{ref}}$.

In order to evaluate the estimator performance, it should be noted that it is unbiased since

$$E[\hat{x}_h] = 1 - \frac{1}{N_{ref}} \sum_{i=1}^{N_{ref}} E[\mathbf{u_i}] = x \tag{50}$$

The variance of the estimator is given as

$$\begin{aligned} E[\hat{x}_h] =& E[\hat{x}_h^2] - E^2[\hat{x}_h] = E\left[\left(1 - \frac{Y}{N_{ref}}\right)^2\right] - x^2 \\ =& 1 - x^2 - \frac{2}{N_{ref}} E[Y] + \frac{1}{N_{ref}^2} E[Y^2] \\ =& \frac{(1-x)x}{N_{ref}} \end{aligned} \tag{51}$$

To evaluate the performance of the estimator, the CRLB can be calculated which is $-\frac{1}{E[\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)/\partial x^2]}$. Taking the second derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to x, we have $\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x^2} = \sum_{i=1}^{N_{ref}} \left[-\frac{\mathbf{u_i}}{(1-x)^2} - \frac{1-\mathbf{u_i}}{x^2}\right]$. Subsequently, taking the expectation of the above equation, we have

$$\begin{aligned} E\left[\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x^2}\right] =& \sum_{i=1}^{N_{ref}} E\left[\frac{\partial^2 P(\mathbf{u_i}|\mathcal{H}_h, p, x)}{\partial x^2}\right] \\ =& -\frac{N_{ref}}{(1-x)x}. \end{aligned} \tag{52}$$

Therefore, the CRLB is $\frac{(1-x)x}{N_{ref}}$ which is the same as (51). This indicates that the proposed estimator attains the CRLB; that is, it is an efficient estimator when sensors in the network send binary decisions.

2) When sensors send q-bit decisions ($q \geq 2$): The joint pmf of local decisions coming from the reference sensors under hypothesis $\mathcal{H}_h$ is given as $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) = \prod_{i=1}^{N_{ref}} (1-x)^{I(\mathbf{u_i} = \mathbf{v}_{2^q})} \prod_{j=1}^{2^q-1} (\frac{x}{2^q-1})^{I(\mathbf{u_i} = \mathbf{v_j})}$ for $h = 0, 1$. Take the logarithm of both sides, we have

$$\begin{aligned} &\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) \\ &= \sum_{i=1}^{N_{ref}} I(\mathbf{u_i} = 2^q) \log(1-x) + \sum_{j=1}^{2^q-1} I(\mathbf{u_i} = \mathbf{v_j}) \log(\frac{x}{2^q-1}), \end{aligned} \tag{53}$$

Taking the first derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to x, we have

$$\begin{aligned} \frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x} =& \sum_{i=1}^{N_{ref}} \frac{-1}{1-x} I(\mathbf{u_i} = 2^q) + \sum_{j=1}^{2^q-1} \frac{1}{x} I(\mathbf{u_i} = \mathbf{v_j}) \\ =& \frac{-Y_1}{1-x} + \frac{N_{ref} - Y_1}{x} \end{aligned} \tag{54}$$

where $Y_1 = \sum_{i=1}^{N_{ref}} I(\mathbf{u_i} = v_{2^q})$. Let $\frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x} = 0$, we are able to obtain the estimated attack parameter $\hat{x}$ which maximizes $\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$. The estimated attack parameter $\hat{x}_h$ under hypothesis $\mathcal{H}_h$ is given as $\hat{x}_h = 1 - \frac{Y_1}{N_{ref}}$.

In order to evaluate the estimator performance, it should be noted that it is unbiased since

$$E[\hat{x}_h] = 1 - \frac{1}{N_{ref}} E[Y_1] = x \tag{55}$$

Similarly, the variance of the estimator is given as

$$\begin{aligned} E[\hat{x}_h] =& E[\hat{x}_h^2] - E^2[\hat{x}_h] = E\left[\left(1 - \frac{Y_1}{N_{ref}}\right)^2\right] - x^2 \\ =& 1 - x^2 - \frac{2}{N_{ref}} E[Y_1] + \frac{1}{N_{ref}^2} E[Y_1^2] \\ =& \frac{(1-x)x}{N_{ref}} \end{aligned} \tag{56}$$

To evaluate the performance of the estimator, the CRLB can be calculated which is $-\frac{1}{E[\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)/\partial x^2]}$. Taking the second derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to p, we have

$$\begin{aligned} \frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x^2} =& \sum_{i=1}^{N_{ref}} -\frac{I(\mathbf{u_i} = 2^q)}{(1-x)^2} - \sum_{i=1}^{2^q-1} \frac{I(\mathbf{u_i} = \mathbf{v_j})}{x^2} \\ =& \sum_{i=1}^{N_{ref}} - \frac{I(\mathbf{u_i} = 2^q)}{(1-x)^2} - \frac{1 - I(\mathbf{u_i} = 2^q)}{x^2} \end{aligned} \tag{57}$$

Subsequently, taking the expectation of the above equation, we have

$$\begin{aligned} E\left[\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h,p,x)}{\partial x^2}\right] &= \sum_{i=1}^{N_{ref}} E\left[\frac{\partial^2 P(\mathbf{u_i}|\mathcal{H}_h,p,x)}{\partial x^2}\right] \\ &= \sum_{i=1}^{N_{ref}} -\frac{1}{(1-x)^2}(1-x) - \frac{1}{x^2}x \\ &= -\frac{N_{ref}}{(1-x)x} \end{aligned} \tag{58}$$

Therefore, the CRLB is $\frac{(1-x)x}{N_{ref}}$ which is the same as (56). This indicates that the proposed estimator attains the CRLB; that is, it is an efficient estimator when sensors in the network send q-bits decisions. This completes our proof.

References

[1] M. Fornasier and H. Rauhut, “Compressive sensing.” *Handbook of mathematical methods in imaging*, vol. 1, pp. 187–229, 2015.

[2] D.-g. Zhang, T. Zhang, J. Zhang, Y. Dong, and X.-d. Zhang, “A kind of effective data aggregating method based on compressive sensing for wireless sensor network,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 1–15, 2018.

[3] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[4] D. Ciuonzo, S. H. Javadi, A. Mohammadi, and P. S. Rossi, “Bandwidth-constrained decentralized detection of an unknown vector signal via multisensor fusion,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 744–758, 2020.

[5] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, “Sparse signal detection from incoherent projections,” in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 3. IEEE, 2006, pp. III–III.

[6] T. Wimalajeewa and P. K. Varshney, “Sparse signal detection with compressive measurements via partial support set estimation,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 46–60, 2016.

[7] H. Zayyani, F. Haddadi, and M. Korki, “Double detector for sparse signal detection from one-bit compressed sensing measurements,” *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1637–1641, 2016.

[8] A. Hariri and M. Babaie-Zadeh, “Compressive detection of sparse signals in additive white gaussian noise without signal reconstruction,” *Signal Processing*, vol. 131, pp. 376–385, 2017.

[9] X. Wang, G. Li, and P. K. Varshney, “Detection of sparse signals in sensor networks via locally most powerful tests,” *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1418–1422, 2018.

[10] X. Wang, G. Li, C. Quan, and P. K. Varshney, “Distributed detection of sparse stochastic signals with quantized measurements: The generalized gaussian case,” *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4886–4898, 2019.

[11] C. Li, Y. He, X. Wang, G. Li, and P. K. Varshney, “Distributed detection of sparse stochastic signals via fusion of 1-bit local likelihood ratios,” *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1738–1742, 2019.

[12] A. Mohammadi, D. Ciuonzo, A. Khazaee, and P. S. Rossi, “Generalized locally most powerful tests for distributed sparse signal detection,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 528–542, 2022.

[13] X. Wang, G. Li, and P. K. Varshney, “Detection of sparse stochastic signals with quantized measurements in sensor networks,” *IEEE Transactions on Signal Processing*, vol. 67, no. 8, pp. 2210–2220, 2019.

[14] M. Korki, J. Zhang, C. Zhang, and H. Zayyani, “Iterative bayesian reconstruction of non-iid block-sparse signals,” *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3297–3307, 2016.

[15] C. Soussen, J. Idier, D. Brie, and J. Duan, “From bernoulli–gaussian deconvolution to sparse signal restoration,” *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4572–4584, 2011.

[16] M. Korki, J. Zhang, C. Zhang, and H. Zayyani, “Block-sparse impulsive noise reduction in ofdm systems—a novel iterative bayesian approach,” *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 271–284, 2015.

[17] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, “An iterative bayesian algorithm for sparse component analysis in presence of noise,” *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4378–4390, 2009.

[18] F. Gao, L. Guo, H. Li, J. Liu, and J. Fang, “Quantizer design for distributed glrt detection of weak signal in wireless sensor networks,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2032–2042, 2014.

[19] C. Li, G. Li, and P. K. Varshney, “Distributed detection of sparse signals with censoring sensors in clustered sensor networks,” *Information Fusion*, vol. 83, pp. 1–18, 2022.

[20] S. Marano, V. Matta, and L. Tong, “Distributed detection in the presence of byzantine attacks,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 16–29, 2008.

[21] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, “Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks,” *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 774–786, 2010.

[22] A. Vempaty, B. Kailkhura, and P. K. Varshney, *Secure Networked Inference with Unreliable Data Sources*. Springer, 2018.

[23] C. Quan, B. Geng, Y. Han, and P. K. Varshney, “Enhanced audit bit based distributed Bayesian detection in the presence of strategic attacks,” *IEEE Transactions on Signal and Information Processing over Networks*, 2022.

[24] H.-Y. Lin, P.-N. Chen, Y. S. Han, and P. K. Varshney, “Minimum Byzantine effort for blinding distributed detection in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 647–661, 2020.

[25] J. Wu, T. Song, Y. Yu, C. Wang, and J. Hu, “Generalized Byzantine attack and defense in cooperative spectrum sensing for cognitive radio networks,” *IEEE Access*, vol. 6, pp. 53 272–53 286, 2018.

[26] Y. Liu and C. Li, “Secure distributed estimation over wireless sensor networks under attacks,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1815–1831, 2018.

[27] Y. Fu and Z. He, “Entropy-based weighted decision combining for collaborative spectrum sensing over Byzantine attack,” *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1528–1532, 2019.

[28] I. Sartzetakis, K. K. Christodoulopoulos, and E. M. Varvarigos, “Accurate quality of transmission estimation with machine learning,” *Journal of Optical Communications and Networking*, vol. 11, no. 3, pp. 140–150, 2019.

[29] I. Sartzetakis, K. Christodoulopoulos, and E. Varvarigos, “Improving qot estimation accuracy through active monitoring,” in *2017 19th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2017, pp. 1–4.

[30] Y. Cui, S. Li, and W. Zhang, “Jointly sparse signal recovery and support recovery via deep learning with applications in mimo-based grant-free random access,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 788–803, 2021.

[31] H. Palangi, R. Ward, and L. Deng, “Distributed compressive sensing: A deep learning approach,” *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4504–4518, 2016.

[32] S. Kafle, B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, “Decentralized joint sparsity pattern recovery using 1-bit compressive sensing,” in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2016, pp. 1354–1358.

[33] Z. Qin, J. Fan, Y. Liu, Y. Gao, and G. Y. Li, “Sparse representation for wireless communications: A compressive sensing approach,” *IEEE Signal Processing Magazine*, vol. 35, no. 3, pp. 40–58, 2018.

[34] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *IEEE transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[35] W. Hashlamoun, S. Brahma, and P. K. Varshney, “Mitigation of byzantine attacks on distributed detection systems using audit bits,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 18–32, 2017.

[36] J. Wu, P. Li, Y. Chen, J. Tang, C. Wei, L. Xia, and T. Song, “Analysis of byzantine attack strategy for cooperative spectrum sensing,” *IEEE Communications Letters*, vol. 24, no. 8, pp. 1631–1635, 2020.

[37] J. Wu, Y. Yu, H. Zhu, T. Song, and J. Hu, “Cost-benefit tradeoff of byzantine attack in cooperative spectrum sensing,” *IEEE Systems Journal*, vol. 14, no. 2, pp. 2532–2543, 2020.

[38] T. Wimalajeewa and P. K. Varshney, “Compressive sensing-based detection with multimodal dependent data,” *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 627–640, 2017.

[39] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.