On Ordered Transmission based Distributed Gaussian Shift-in-Mean Detection under Byzantine Attacks

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

***Abstract*—The ordered transmission based (OT-based) schemes reduce the number of transmissions needed in a distributed detection network without any loss in the probability of error performance. In this paper, we investigate the performance of a conventional OT-based system in the presence of additive Byzantine attacks in Gaussian shift in mean problems. In this work, by launching additive Byzantine attacks, attackers are able to alter the order as well as the data for the binary hypothesis testing problem. We also determine the optimal attack strategy for the Byzantine sensors. Furthermore, we analyze a communication efficient OT-based (CEOT-based) scheme in the presence of additive Byzantine attacks. We obtain the probabilities of error for both the OT-based system and the CEOT-based system under attack and evaluate the number of transmissions they save. We also derive analytical bounds for the number of transmissions saved in both systems under attack. Simulation results show that the additive Byzantine attacks have significant impact on the number of transmissions saved even when the signal strength is sufficiently large. A comparison of detection performance between the conventional OT-based system and the CEOT-based system reveals that the CEOT-based system is more robust to additive Byzantine attacks.**

***Index Terms*—Ordered transmissions, Byzantine attacks, wireless sensor networks, distributed detection.**

I. Introduction

ENERGY-EFFICIENCY is an important aspect to consider while designing a wireless sensor network (WSN) with prolonged lifetime [1]. Several notable schemes have been proposed to improve energy efficiency by reducing the number of transmissions in the networks [2]–[5]. In this paper, we consider one such scheme called the ordered transmission based (OT-based) scheme [5] in the distributed setup. In the conventional OT-based scheme, all the sensors in the network transmit in decreasing order of their respective absolute values of the log-likelihood ratios (LLRs). Specifically, the starting time of transmission at each sensor is proportional to the inverse of the absolute value of its LLR. Hence, the more informative sensors (sensors with larger magnitudes of the value of LLR) transmit earlier than the less informative ones (sensors with smaller magnitudes of the value of LLR). When the fusion center (FC) has received enough observations to make the final decision of desired quality, the FC broadcasts a stop signal to stop the sensors from further transmitting. The sensors that have not yet transmitted their observations reset their timers for the next decision interval after they receive the stop signal. For simplicity, in the rest of the paper, the OT-based scheme refers to the conventional OT-based scheme.

The OT-based scheme for distributed networks was first proposed in [5], where only informative sensors in the network transmitted their LLRs to the FC instead of sending raw data. The concept was extended to an ordering approach for a class of noncoherent signal detection problems where the LLRs at each sensor could only take nonnegative values in [6]. The authors in [7] demonstrated that a single observation was sufficient to make a final decision for an OT-based system with a large number of sensors. In [8], sequential detection along with OT was considered for cooperative spectrum sensing to obtain fast and reliable decisions regarding primary user activities over the spectrum. The sequential test was run at the FC with a constraint on the maximum number of sensors that reported their LLRs. This constraint was incorporated using the OT-based scheme. Furthermore, the authors in [9] considered the quickest change detection problem to detect the change in the distribution of independent observations by proposing a new approach where the transmissions from the sensors were ordered and stopped when sufficient information was accumulated at the FC. The authors showed that the proposed approach achieved the optimal performance equivalent to the centralized cumulative sum (CUSUM) algorithm with less sensor transmissions. In [10], the OT-based scheme is employed in the quickest change detection problem with dependent observations to reduce communications without increasing detection delays. The dependence among sensor observations is characterized using a decomposable graphical model (DGM). The authors showed that the proposed algorithm is able to achieve identical performance to the non-OT based scheme in terms of the worst-case average detection delay. In order to eliminate some sensor-to-FC uplink communications, the authors in [11] considered an ordered gradient approach where the timer at each sensor was set inversely proportional to the magnitude of the gradient of the loss function. This resultant gradient-based approach achieved the same order of convergence rate as the gradient descent

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

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

approach for nonconvex smooth loss functions. Also, the work in [12] considered an OT-based algorithm for the discretized estimation problem with a latency constraint. The authors showed that the proposed algorithm can greatly reduce latency without loss of estimation accuracy. In [13], the OT-based scheme was incorporated along with energy harvesting in the WSNs in order to improve energy efficiency of the sensors. A correlation-aware OT-based scheme was proposed in [14] where spatial correlation between the sensors was considered. The OT-based framework was applied to determine a shift in the mean and covariance, and the decision rule was proposed accordingly. In the CEOT-based scheme, informative sensors transmit binary decisions to the FC, improving communication efficiency in the distributed setup, rather than sending raw LLR values.In [15], the authors proposed a communication-efficient OT based (CEOT-based) scheme in which informative sensors send binary decisions, rather than sending raw LLR values, to the FC in order to improve the communication efficiency. The above works show that the OT-based schemes in the Gaussian shift in mean problem are capable of efficiently reducing the number of transmissions needed for decision-making while maintaining the same inference performance.

However, due to the large number of low-cost sensors and the vulnerability of WSNs to failures and adversarial attacks, the robustness of the OT-based and CEOT-based systems under attack is an important aspect to consider. Typically, when the WSNs are under adversarial attacks, one or more sensors may get compromised and may send falsified data to the FC to degrade the detection performance of the system [16], [17]. This Byzantine threat model has been extensively studied in [17]–[23]. Unlike the previous OT-based systems that only considered honest sensors in the networks [5]–[15], [24], we aim to evaluate the performance of the OT-based system via both the detection performance and the number of transmissions saved in the presence of Byzantine sensors. In our previous work [25], we have investigated the effect of data falsification attacks, where the local decisions are altered by Byzantine sensors, on the detection performance and the number of transmissions saved for the CEOT-based system. In this paper, we consider additive Byzantine attack models. Under the OT-based framework, the Byzantine sensors can alter not only their decisions but also the order in which decisions are transmitted by altering their LLRs. Furthermore, we investigate the effect of different additive Byzantine attacks on the performance of both OT-based and CEOT-based systems. The following are our major contributions:

- We investigate the performance of the OT-based distributed detection system in the presence of two kinds of additive Byzantine attacks. The first type involves shifting the mean of the actual observations, which is referred to as the mean-shift attack model. The second type involves shifting both the mean and the variance of the actual observations, known as the mean-variance-shift attack model. We also determine the optimal attack strategy for such Byzantine sensors. The attack strategy is determined by utilizing the deflection coefficient (DC) as a surrogate for the probability of error. Also, we evaluate the performance of CEOT-based distributed detection systems in the presence of such additive Byzantine attacks.
- We show that the detection performance of the two systems remains the same whether ordering is considered or not in the presence of additive Byzantine attacks.
- We derive the probabilities of error for the OT-based and CEOT-based systems under additive Byzantine attacks. We also derive an upper bound (UB) and a lower bound (LB) on the number of transmissions saved in the network for both systems.
- We evaluate the number of transmissions saved in the OT-based system numerically via the Monte Carlo approach in the presence of Byzantine sensors. Our simulation results show that the optimal attack strategy that maximizes the probability of error, also maximizes the number of transmissions needed when utilizing the OT-based scheme. Furthermore, a comparison between the OT-based system and the CEOT-based system is made. We observe that the CEOT-based system is more robust to additive Byzantine attacks since the impact of additive Byzantine attacks on the CEOT-based system is reduced by the quantization of data.

The paper is organized as follows. We present our system model in Section II. We evaluate the performance of the OT-based system under additive Byzantine attacks and derive the bounds for the number of transmissions saved in the OT-based system in Section III. The performance of the CEOT-based system under additive Byzantine attacks is investigated and the bounds for the number of transmissions saved are derived in Section IV. We present our numerical results in Section V and conclude in Section VI.

II. System Model

In this section, we consider a binary hypothesis testing problem where hypothesis $\mathcal{H}_1$ indicates the presence of the signal and $\mathcal{H}_0$ indicates the absence of the signal. We consider a distributed network consisting of N sensors and one FC. Furthermore, the OT-based scheme is considered to reduce the number of transmissions in the network. 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) conditioned on the hypotheses. For sensor i, the observation y_i is modeled as

$$y_i = \begin{cases} n_i & \text{under } \mathcal{H}_0 \\ s + n_i & \text{under } \mathcal{H}_1, \end{cases} \tag{1}$$

where s is non-negative and it is the signal strength at each sensor, and n_i is the Gaussian noise with zero mean and variance σ^2. We assume that s and n_i are independent. Note that y_i is Gaussian with mean s and variance σ^2 under hypothesis $\mathcal{H}_1$, and is Gaussian with mean 0 and variance σ^2 under hypothesis $\mathcal{H}_0$.

Next, we review two OT-based schemes where all the sensors are assumed to be honest. One is the OT-based scheme proposed in [5] where the local sensors send their LLRs to the FC. The other is the CEOT-based scheme proposed in [15] where the local sensors transmit binary decisions to the FC.

A. Network with OT-based Scheme

Let L_i denote the LLR for sensor i given by $L_i = \log\left(\frac{f_{Y_i}(y_i|\mathcal{H}_1)}{f_{Y_i}(y_i|\mathcal{H}_0)}\right)$, where $f_{Y_i}(y_i|\mathcal{H}_h)$ is the probability density function (PDF) of y_i given hypothesis $\mathcal{H}_h$, for $h = 0, 1$. The prior probabilities of hypotheses $\mathcal{H}_h$ are $P(\mathcal{H}_h) = \pi_h$, for $h \in \{0, 1\}$. Recall that the LLR-based optimal Bayesian hypothesis test at the FC for an unordered system is given by $\sum_{i=1}^{N} L_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda = \log\left(\frac{\pi_0}{\pi_1}\right)$, where λ is the threshold used by the FC. In the OT-based system, the sensor transmissions are ordered based on the magnitude of their respective LLRs. We denote the magnitude of the ordered transmissions as $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$, where $|L_{[i]}|$ indicates the i^{th} largest LLR. Hence, the sensor with LLR $L_{[1]}$ transmits first, the sensor with LLR $L_{[2]}$ transmits second, and so on. The optimal decision rule of the FC [5] becomes

$$\begin{cases} \sum_{i=1}^{k} L_{[i]} > \lambda + n_{UT}|L_{[k]}| & \text{decide } \mathcal{H}_1 \\ \sum_{i=1}^{k} L_{[i]} < \lambda - n_{UT}|L_{[k]}| & \text{decide } \mathcal{H}_0, \end{cases} \tag{2}$$

for an OT-based system, where n_{UT} is the number of sensors that have not yet transmitted at time k. The FC waits for the next transmission if it can not make the decision with desired accuracy. In this work, we assume that both the sensors and the FC are aware of the relationship between the transmission time t of the sensors and the corresponding magnitude of their LLRs, i.e, $t \propto 1/|L_i|, \forall i \in 1, 2 \ldots, N$. Note that the following assumption was also made in [5] for their analysis.

Assumption 1: $Pr(L_i > 0|\mathcal{H}_1) \to 1$ and $Pr(L_i < 0|\mathcal{H}_0) \to 1$ when s is sufficiently large.

Intuitively, the assumption states that the true hypothesis can be decided easily based on L_i for any sensor i if the distance (dependent on s) between the distributions of the observations of sensor i occurring under the two hypotheses becomes large.

B. Network with CEOT-based Scheme

Based on the local observations $\{y_i\}_{i=1}^{N}$, each sensor $i \in \{1, \ldots, N\}$ makes a binary decision $u_i \in \{0, 1\}$ regarding the true hypothesis using the LLR test $L_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \log\left(\frac{\pi_0}{\pi_1}\right)$ [26] and u_i is the local decision of sensor i. We assume that the sensor transmissions are still ordered based on the magnitude of their LLRs. Recall that the magnitudes of the LLRs are ordered as $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$. Then, the sensors transmit their local decisions to the FC in the order determined by their magnitude-ordered LLRs, i.e., in the order of $u_{[1]}, u_{[2]}, \ldots, u_{[N]}$, where $u_{[k]}$ is the local decision of the sensor with k^{th} largest LLR.[1]

The optimal decision rule [15] is given by

$$\begin{cases} \sum_{i=1}^{k} u_{[i]} \geq T & \text{decide } \mathcal{H}_1 \\ \sum_{i=1}^{k} u_{[i]} < T - (N - k) & \text{decide } \mathcal{H}_0, \end{cases} \tag{3}$$

which follows the T out of N counting rule. Here, we consider the majority rule where $T = N/2 + 1$. The following assumption is made in [15] for the CEOT-based scheme similar to the OT-based scheme made in [5].

Assumption 2: $Pr(u_i = 1|\mathcal{H}_1) \to 1$ and $Pr(u_i = 0|\mathcal{H}_0) \to 1$ when s is sufficiently large.

Remark. Note that large s is key to proving the result that the average number of transmissions saved by utilizing both the OT-based scheme and the CEOT-based scheme is lower bounded by $N/2$ (see [5, Theorem 2] and [15]). However, when s is small or when there are Byzantine sensors in the system, Assumptions 1 and 2 are no longer valid.

C. Mean-shift Attack Model

Next, we discuss the mean-shift attack model adopted by Byzantine sensors. We assume that the Byzantine sensors falsify data by controlling the attack strength in both the OT-based system and the CEOT-based system. In this work, only passive systems are considered, i.e. the system is unaware of the presence of attackers. We assume that the Byzantine sensors have perfect knowledge of the underlying true hypothesis. Admittedly, it is hard to realize in practice but it is useful to consider this case as it provides the impact of Byzantines in the worst case. Note that we want to analyze the ability of the Byzantine sensors to affect the decision at the FC. Specifically, we want to analyze from the attacker's perspective by determining the most effective attack strategy for the Byzantine sensors. We also assume that each sensor can be a Byzantine with probability α. Furthermore, the falsified observation $\tilde{y}_i$ for Byzantine node i is given by

$$\tilde{y}_i = \begin{cases} s + n_i - D & \text{under } \mathcal{H}_1 \\ n_i + D & \text{under } \mathcal{H}_0, \end{cases} \tag{4}$$

where D is the attack strength and it is a non-negative constant value. The above attack strategy adopted by Byzantine nodes is equivalent to launching attacks by generating falsified observations from another distribution, and it is commonly used in the literature [17], [27]–[30]. Here, the distribution used by Byzantine nodes to generate falsified observations is obtained by shifting the mean of the actual distribution with a constant value. This attack strategy can easily be extended to a more general case, i.e, mean-variance-shift attack strategy, where the the attack strength is a random variable. Details will be discussed in the following section. Note that for an honest node, the observation is y_i as given in (1). Hence in our setup, a sensor i can be honest (H) or Byzantine (B). However, Assumptions 1 and 2 made in [5] and [15], respectively, are no longer valid.

Remark. Note that both the sensors and the FC are aware of the relationship between the transmission time t of the sensors and the corresponding magnitude of their LLRs, i.e, $t \propto 1/|L_i|, \forall i \in 1, 2 \ldots, N$. If an attacker deviates from the ordered-transmission protocol, they introduce an additional dimension of adversarial behavior. This non-compliance makes their malicious actions more conspicuous and susceptible to identification by the FC. In other words, it increases the possibility of being easily detected by the system as being malicious. Here, we assume that the Byzantines follow the ordered-transmission protocol. By making this assumption, we

[1]Note that the magnitude-ordered LLRs do not imply that local decisions are also magnitude-ordered, i.e., $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$ does not imply $u_{[1]} > u_{[2]} > \ldots > u_{[N]}$.

are assuming a more challenging situation where attackers attempt to hide their malicious actions within the prescribed protocol.

In the following sections, we analyze the detection performance of the OT-based system and the CEOT-based system when confronted with Byzantine sensors employing two attack strategies. The first strategy involves the sensors adopting a mean-shift attack, while the second strategy involves the sensors employing a more general attack.

III. OT-BASED SYSTEM WITH BYZANTINES

In this section, we first present an important Lemma that enables us to evaluate the detection performance of the OT-based scheme in the presence of Byzantine sensors adopting mean-shift attack strategy. Next, we obtain the expression for the number of transmissions saved in the network which is significantly impacted by the presence of Byzantine sensors. Furthermore, we also determine the optimal attack strategy for Byzantine sensors by using Deflection Coefficient (DC). Finally, we discuss the performance of the OT-based scheme under mean-variance-shift attacks.

A. Detection Performance under Mean-shift Attack

We begin our analysis of the detection performance of the OT-based scheme in the presence of Byzantine sensors by first presenting the following Lemma which states that the OT-based system can achieve the same detection performance as the one without ordering.[2]

Lemma III.1. [3] *Under the optimum Bayesian decision rule, the detection performance remains the same whether or not the system uses the OT-based scheme in the presence of Byzantine sensors.*

Proof: The proof is relegated to Appendix A. ∎

Thus, based on above Lemma, we can obtain the detection performance of the OT-based system by evaluating the detection performance of the system without ordering. For the system without ordering, we have $L_i = \frac{2y_i s - s^2}{2\sigma^2}$ when sensor i is honest ($i = H$). When sensor i is Byzantine ($i = B$), the LLR is given as

$$L_i = \begin{cases} \frac{2(y_i - D)s - s^2}{2\sigma^2} & \text{under } \mathcal{H}_1 \\ \frac{2(y_i + D)s - s^2}{2\sigma^2} & \text{under } \mathcal{H}_0 \end{cases} \tag{5}$$

Hence, if sensor $i = H$, the PDF of L_i conditioned on hypothesis $\mathcal{H}_h$ follows Gaussian distribution with mean μ_h and variance σ_h^2 for $h = 0, 1$, where $\mu_1 = \frac{s^2}{2\sigma^2}$, $\mu_0 = \frac{-s^2}{2\sigma^2}$, $\sigma_1^2 = \sigma_0^2 = \frac{s^2}{\sigma^2} \triangleq \beta$. Furthermore, if sensor $i = B$, the PDF of L_i conditioned on hypothesis $\mathcal{H}_h$ follows Gaussian distribution with mean η_h and variance ν_h^2 for $h = 0, 1$, where $\eta_0 = \frac{s^2 - 2Ds}{2\sigma^2}$, $\eta_1 = \frac{2Ds - s^2}{2\sigma^2}$, $\nu_0^2 = \nu_1^2 = \frac{s^2}{\sigma^2} \triangleq \beta$. Therefore, the PDF of L_i given hypothesis $\mathcal{H}_h$ is expressed as

$$\begin{aligned} f_L(l_i|\mathcal{H}_h) &= \alpha f_L(l_i|\mathcal{H}_h, i = B) + (1-\alpha) f_L(l_i|\mathcal{H}_h, i = H) \\ &= \alpha \mathcal{N}(\eta_h, \nu_h^2) + (1-\alpha)\mathcal{N}(\mu_h, \sigma_h^2), \end{aligned} \tag{6}$$

for $h = 0, 1$. Here, α denotes the probability of a node being Byzantine. Let $\mathcal{K} = \{A_1, \ldots A_t, \ldots, A_{2^N}\}$ denote the power set that contains all possible subsets of set $\{1, \ldots, N\}$ and A_t be the t^{th} subset of the combination of honest sensors. Also, $|A_t|$ is the cardinality of set A_t. Let $Z = \sum_{i=1}^{N} L_{[i]}$ denote the global test statistic and $f(Z|\mathcal{H}_h)$ denote the Gaussian mixture with PDF given by $f(Z|\mathcal{H}_h) = \sum_{t=1}^{2^N} (1-\alpha)^{m_t} \alpha^{N-m_t} \mathcal{N}((\mu_h)_{A_t}, N\beta)$ for $h = 0, 1$, where $(\mu_h)_{A_t} = \mu_h |A_t| + \eta_h (N - |A_t|)$ and m_t denotes the cardinality of set A_t, i.e., $m_t = |A_t|$.

Therefore, the detection performance can be evaluated in terms of the probability of detection P_d^{FC} and the probability of false alarm P_f^{FC} of the FC given as $P_d^{FC} = \sum_{t=1}^{2^N} (1-\alpha)^{N-m_t} \alpha^{m_t} Q\left(\frac{\lambda - (\mu_1)_{A_t}}{\sqrt{N\beta}}\right)$ and $P_f^{FC} = \sum_{t=1}^{2^N} (1-\alpha)^{N-m_t} \alpha^{m_t} Q\left(\frac{\lambda - (\mu_0)_{A_t}}{\sqrt{N\beta}}\right)$ by following steps that are similar to those outlined in [29], where $Q(.)$ is the tail distribution function of the standard normal distribution.

B. Average Number of Transmissions Saved for the OT-based System under Mean-shift Attack

We consider the effect of Byzantine sensors on the number of transmissions saved for the OT-based scheme. When the system is under attack, we derive an expression for the average number of transmissions $\bar{N}_t$ in the following theorem. Let k^* denote the minimum number of transmissions needed to make a final decision with desired accuracy. Let $F_{|L|}(l_i|\mathcal{H}_h)$ be the cumulative distribution function (CDF) of $|L_i|$ for $h = 0, 1$ provided as

$$\begin{aligned} F_{|L|}(l_i|\mathcal{H}_h) =& \alpha \left(Q\left(\frac{-l_i - \eta_h}{\nu_h}\right) - Q\left(\frac{l_i - \eta_h}{\nu_h}\right)\right) \\ &+ (1-\alpha)\left(Q\left(\frac{-l_i - \mu_h}{\sigma_h}\right) - Q\left(\frac{l_i - \mu_h}{\sigma_h}\right)\right). \end{aligned} \tag{7}$$

Theorem III.2. *The average number of transmissions $\bar{N}_t$ is given as*

$$\bar{N}_t = \sum_{k=1}^{N} \pi_1 Pr(k^* \geq k|\mathcal{H}_1) + \pi_0 Pr(k^* \geq k|\mathcal{H}_0) \tag{8}$$

where

$$\begin{aligned} &Pr(k^* \geq k|\mathcal{H}_h) \\ =& E_{\mathbf{L}_{k-1}}\left[F_{|L|}(L_{k-1}|\mathcal{H}_h)^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}} \frac{N!}{(N-k+1)!}\right], \end{aligned} \tag{9}$$

for $h = 0, 1$. The indicator function $\mathbf{1}_{\{\mathcal{J}\}}$ is 1 when $\mathbf{L}_{k-1} = \{L_1, L_2 \ldots, L_{k-1}\}$ is in the region $\mathcal{J}$, and 0 otherwise. Here, $\mathcal{J}$ is a hyperplane with $k-1$ dimensions formed by the

[2]Note that both the OT-based and unordered systems have the same probability of error in the presence of Byzantine sensors. However, the number of transmissions saved is significantly impacted by the presence of Byzantine sensors for the OT-based system as discussed later.

[3]Note that this paper presents analytical proofs for Lemma III.1 and Lemma IV.1 in the presence of Byzantines, which have not been presented in previous works on OT-based frameworks.

$$\bar{N_s}^U = \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr\left(|L_{[k]}| \leq \frac{g_U - \lambda}{N-k} | \mathcal{H}_h \right) + Pr\left(|L_{[k]}| \leq \frac{\lambda - g_L}{N-k} | \mathcal{H}_h \right) - Pr\left(|L_{[k]}| \leq \min\left(\frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k} \right) | \mathcal{H}_h \right) \right] \quad (13)$$

$$\bar{N_s}^L = \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr\left(|L_{[k]}| < \frac{g_L - \lambda}{(N-k)} | \mathcal{H}_h \right) + Pr\left(|L_{[k]}| < \frac{\lambda - g_U}{(N-k)} | \mathcal{H}_h \right) \right] \quad (14)$$

intersection of three hyperplanes, $\mathcal{J} = \mathcal{L} \bigcap \mathcal{U} \bigcap \mathcal{D}$, which are given below

$$\mathcal{L} = \left\{ \mathbf{L}_{k-1} : \sum_{i=1}^{k-1} L_i \leq \lambda + (N-k+1)|L_{k-1}| \right\} \quad (10)$$

$$\mathcal{U} = \left\{ \mathbf{L}_{k-1} : \sum_{i=1}^{k-1} L_i \geq \lambda - (N-k+1)|L_{k-1}| \right\} \quad (11)$$

$$\mathcal{D} = \{ \mathbf{L}_{k-1} : |L_1| > |L_2| > \cdots > |L_{k-1}| \}. \quad (12)$$

Proof: Please see Appendix B. ∎

Note that the set $\mathcal{L}$ is the set of $\mathbf{L}_{k-1}$ such that the FC can not decide hypothesis $\mathcal{H}_1$. Also, the set $\mathcal{U}$ is the set of $\mathbf{L}_{k-1}$ such that the FC can not decide hypothesis $\mathcal{H}_0$. Furthermore, the set $\mathcal{D}$ is the set of $\mathbf{L}_{k-1}$ such that $L_1, L_2, \ldots, L_{k-1}$ are ordered in magnitude. For a given k, we evaluate (9) numerically using the Monte Carlo approach as the following. We generate M i.i.d. realizations of $L_1, L_2, \ldots, L_{k-1}$, where the PDF of L_i is given in (6) for $\forall i \in \{1, 2, \ldots, k-1\}$. From our experiments, we observe that when N increases, the number of samples M needed to get an accurate evaluation of (9) significantly increases.

Next, we derive the upper bound (UB) and the lower bound (LB) for the number of the transmissions saved by the OT-based scheme under Byzantine attack in the following Theorem. Let $\bar{N_s}^U$ and $\bar{N_s}^L$ denote the UB and the LB of the average number of transmissions saved.

Theorem III.3. *When N is sufficiently large, the average number of transmissions saved $\bar{N}_s$ can be bounded as $\bar{N_s}^L \leq \bar{N}_s \leq \bar{N_s}^U$, where $\bar{N_s}^U$ and $\bar{N_s}^L$ are given in* (13) *and* (14)*, respectively. Furthermore, we have $Pr\left(|L_{[k]}| < W | \mathcal{H}_h\right) = \int_{-W}^{W} f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h) \mathrm{d}l_{[k]}$ for $W \in \{ \frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k}, \min(\frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k}), \frac{g_L - \lambda}{N-k}, \frac{\lambda - g_U}{N-k} \}$ and $f_{|L_{[k]}|}(|l_{[k]}||\mathcal{H}_h)$ is shown in* (49)*. We have $g_L = -[\sum (c_i - \bar{c})^2 N \zeta_h^2]^{\frac{1}{2}} + k\delta_h$ and $g_U = [\sum (c_i - \bar{c})^2 N \zeta_h^2]^{\frac{1}{2}} + k\delta_h$, where δ_h and ζ_h^2 are shown in* (46)*. Here, $\bar{c} = \frac{\sum_{i=1}^{N} c_i}{N}$ where $c_i = 1$ if $i \leq k$ and $c_i = 0$ if $i > k$.*

Proof: Please see Appendix C. ∎

Next, we discuss the optimal attack strategy for Byzantine sensors and later (see Section V) show the effect of Byzantine sensors that utilize the optimal attack strategy on the OT-based system.

C. Optimal Mean-shift Attack Strategy

From the analysis in Sec. III-A, we can obtain the error probability of the system which is given by $P_e = \pi_1(1 - P_d^{FC}) + \pi_0 P_f^{FC}$. However, $|\mathcal{K}|$ grows exponentially as N increases. Therefore, it is intractable to evaluate the system performance using P_e. Hence, we utilize the DC [31] as a surrogate to determine the best attack strategy. By minimizing DC, P_e is maximized.

The DC is defined as $D(\tilde{Z}) = \frac{(\mathbb{E}(\tilde{Z}|H_1) - \mathbb{E}(\tilde{Z}|H_0))^2}{Var(\tilde{Z}|H_0)}$. For the system without ordering, let $\tilde{Z} = \sum_{i=1}^{N} L_i$ denote the global statistic. Therefore, we have $\mathbb{E}(\tilde{Z}|H_1) = -\mathbb{E}(\tilde{Z}|H_0) = N\frac{s^2 - 2Ds\alpha}{2\sigma^2}$. From Lemma III.1 and the above discussion, to maximize the probability of error of the system with ordering, we can minimize the DC of the system without ordering. For a specific value of α, the value of D which minimizes DC is the optimal attack strength D^*. Since the DC is always non-negative, the optimal strategy for the Byzantine sensors is to make $D(\tilde{Z}) = 0$. From the definition of DC, when $\mathbb{E}(\tilde{Z}|H_1) = \mathbb{E}(\tilde{Z}|H_0)$, we have $D(\tilde{Z}) = 0$. Hence, for a given α, the optimal attack strength D^* is given by

$$D^* = \frac{s}{2\alpha}, \quad (15)$$

which is the minimum attack strength to blind the FC, i.e., to make the probability of error equal to 1/2.

D. Mean-variance-shift Attack Model

The mean-shift attack strategy can be easily extended to a more general case, i.e., mean-variance-shift attack strategy, where the signal is perturbed by random noise. For the sake of simplicity in performance analysis, we consider the scenario where the actual data is perturbed by Gaussian noise.

Recall that the mean-shift attack strategy assumes that Byzantines falsify their observations with constant values D and $-D$ as shown in (4). Consequently, the LLR for Byzantine sensor i can be expressed as

$$L_i = \begin{cases} \frac{2(y_i - D)s - s^2}{2\sigma^2} = L_{i,true} + f_1(D) & \text{under } \mathcal{H}_1 \\ \frac{2(y_i + D)s - s^2}{2\sigma^2} = L_{i,true} + f_0(D) & \text{under } \mathcal{H}_0, \end{cases} \quad (16)$$

where $f_1(D) = -\frac{Ds}{\sigma^2}$, $f_0(D) = \frac{Ds}{\sigma^2}$ and $L_{i,true} = \frac{2y_i s - s^2}{2\sigma^2}$ is the actual value of sensor i's LLR. We can easily observe that Byzantines falsify their actual observations y with D and $-D$, which can be equivalently viewed as falsifying their actual LLRs with constant values $f_1(D)$ and $f_0(D)$. In this case, the falsified LLRs are generated from another Gaussian distribution with a different mean and the same variance. If we assume a more general attack strategy, where the actual observations are perturbed by Gaussian noise, both the mean and variance of the Byzantines' LLRs will be altered.

Assuming that the actual LLR of compromised sensor $i \in \{1, 2, \ldots, N\}$ is perturbed by a random noise component that follows a Gaussian distribution, the perturbed LLR is given by:

$$L_i = \begin{cases} L_{i,true} + n_{1,i,w} & \text{under } \mathcal{H}_1 \\ L_{i,true} + n_{0,i,w} & \text{under } \mathcal{H}_0, \end{cases} \quad (17)$$

where $n_{1,i,w}$ represents the perturbation noise under hypothesis $\mathcal{H}_1$ that follows a Gaussian distribution with mean $f_1(D)$ and variance σ_w^2, and $n_{0,i,w}$ represents the perturbation noise

under hypothesis $\mathcal{H}_0$ that follows a Gaussian distribution with mean $f_0(D)$ and variance σ_w^2. It is easy to obtain that the falsified L_i of sensor i follows a Gaussian distribution with its mean given by

$$E[L_i|\mathcal{H}_h] = \mu_h + f_h(D), \tag{18}$$

where $\mu_1 = \frac{s^2}{2\sigma^2}$ and $\mu_0 = \frac{-s^2}{2\sigma^2}$, and the variance given by

$$Var[L_i|\mathcal{H}_h] = \frac{s^2}{\sigma^2} + \sigma_w^2 \triangleq \nu_{h,w}^2 \tag{19}$$

for $h = 0, 1$. To evaluate the performance of the system under such general attacks, we only need to replace ν_0^2 and ν_1^2 with $\nu_{0,w}^2$ and $\nu_{1,w}^2$, respectively, in all the above discussions.

IV. CEOT-based System with Byzantines

In this section, we evaluate the performance of the CEOT-based system in the presence of Byzantine sensors by analyzing the detection performance and the number of transmissions saved in the network. Similar to the previous section, we show that the detection performance is the same whether we use ordering or not in the presence of Byzantine sensors for the CEOT-based scheme. Note that the performance analysis of the CEOT-based system with Gaussian noise perturbed signals can be obtained by following a similar approach as in the previous section. For the sake of brevity, the detailed discussion and analysis are omitted in this section.

A. Detection Performance

We begin our analysis of the detection performance of the CEOT-based system in the presence of Byzantine sensors by first presenting the following Lemma which states that the CEOT-based system can achieve the same detection performance as the one without ordering.

Lemma IV.1. *Under the optimum Bayesian decision rule, the detection performance of the CEOT-based system with and without ordering is the same in the presence of Byzantines.*

Proof: The proof follows similar procedure as the proof of Lemma III.1. Hence, due to space constraints, the proof of Lemma IV.1 is omitted. ∎

Hence, based on the above Lemma, we can obtain the detection performance of the CEOT-based system by evaluating the detection performance of the distributed system without ordering. For the system without ordering, if the sensor is honest $i = H$, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed as $\pi_{1,h}^H = P(u_i = 1|\mathcal{H}_h, i = H) = Q\left(\frac{\lambda-\mu_h}{\sigma_h}\right)$ and $\pi_{0,h}^H = P(u_i = 0|\mathcal{H}_h, i = H) = 1 - \pi_{1,h}^H$, respectively, for $h = 0, 1$. If the sensor is Byzantine $i = B$, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed as $\pi_{1,h}^B = P(u_i = 1|\mathcal{H}_h, i = B) = Q\left(\frac{\lambda-\eta_h}{\nu_h}\right)$ and $\pi_{0,h}^B = P(u_i = 0|\mathcal{H}_h, i = B) = 1-\pi_{1,h}^B$, respectively, for $h = 0, 1$. [4] Therefore, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed as

$$\pi_{1,h} = P(u_i = 1|\mathcal{H}_h) = \alpha\pi_{1,h}^B + (1-\alpha)\pi_{1,h}^H,$$

and $\pi_{0,h} = P(u_i = 0|\mathcal{H}_h) = \alpha\pi_{0,h}^B + (1-\alpha)\pi_{0,h}^H$, respectively, for $h = 0, 1$.

The fusion rule of the distributed system without ordering is given as $\sum_{i=1}^N u_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T$, by noting that we can consider the unordered scheme and taking $k = N$ in (3). Based on the fusion rule of unordered system, the detection performance can be evaluated in terms of the probability of detection $P_{d,CEOT}^{FC}$ and the probability of false alarm $P_{f,CEOT}^{FC}$ given as $P_{d,CEOT}^{FC} = \sum_{i=T}^N \binom{N}{i} \pi_{1,1}^i \pi_{0,1}^{N-i}$ and $P_{f,CEOT}^{FC} = \sum_{i=T}^N \binom{N}{i} \pi_{1,0}^i \pi_{0,0}^{N-i}$.

B. Average Number of Transmissions Saved for the CEOT-based System under Attack

Next, we consider the effect of additive Byzantine attacks on the number of transmissions saved for the CEOT-based scheme. When the system is under attack, we derive the UB and the LB for the average number of transmissions saved for the CEOT-based scheme in the theorem later in this section. Let $\bar{N}_{s,CEOT}$ denote the average number of transmissions saved in the CEOT-based scheme given as

$$\bar{N}_{s,CEOT} = E(N - k^*) = \sum_{k=1}^{N}(N-k)Pr(k^* = k) = \sum_{k=1}^{N-1} Pr(k^* \leq k), \tag{20}$$

where k^* denotes the minimum number of transmissions needed to make a final decision with desired accuracy. However, the computation of $Pr(k^* \leq k)$ is intractable. Hence, we derive the UB and LB of $\bar{N}_{s,CEOT}$ by considering the best case and the worst case scenarios for the number of transmissions saved in the network in the presence of Byzantines, respectively. The information of global statistic of the distributed system without ordering, which is given as $\Gamma = \sum_{i=1}^N u_i$, is utilized to derive both LB and UB. It is easy to conclude that $\Gamma < T$ means that there exists a k^* such that $\sum_{i=1}^{k^*} u_{[i]} < T - (N - k^*)$ and $\Gamma \geq T$ means that there exists a k^* such that $\sum_{i=1}^{k^*} u_{[i]} \geq T$ according to Lemma IV.1. In order to find the LB and UB of $\bar{N}_{s,CEOT}$, we consider the worst and best cases as follows.

When we consider the worst case, we try to find the maximum k^* needed to make a final decision for a given set of local decisions $\{u_i\}_{i=1}^N$. Therefore, the worst case given $\Gamma < T$ would be that the magnitude of local decisions are ordered in descending order expressed as

$$|z_{[1]}| \geq |z_{[2]}| \cdots \geq |z_{[N]}|, \tag{21}$$

where $z_{[k]} \in \{0, 1\}$, for $\forall k \in \{1, 2, \ldots, N\}$ is the k^{th} largest local decision.[5] This is due to the fact that $\Gamma < T$ implies that the unordered system (i.e., the system where the FC receives all local decisions) has more '0' decisions than '1' decisions[6], and the detection performance of the unordered system is the same as the ordered system, as stated in Lemma IV.1. The worst case scenario would occur if the magnitudes of local decisions are ordered in descending order. Similarly, the

[4] When we consider Gaussian noise perturbed signals, we only need to replace ν_h^2 with $\nu_{h,w}^2$ for $h = 0, 1$.

[5] Note that $z_{[k]}$ is not the same as $u_{[k]}$. The values $u_{[1]}, u_{[2]}, \ldots, u_{[N]}$ are ordered based on the magnitude of their LLRs, while $z_{[1]}, z_{[2]}, \ldots, z_{[N]}$ are ordered based on the magnitude of local decisions $\{u_i\}_{i=1}^N$.

[6] More specifically, the number of '1's should be smaller than T.

$$\bar{N}^{U}_{s,CEOT} = \sum_{h=0}^{1}\sum_{k=1}^{N-1} \pi_h \left[P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) P(\Gamma \geq T|\mathcal{H}_h) + P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right) P(\Gamma < T|\mathcal{H}_h)\right], \tag{23}$$

$$\bar{N}^{L}_{s,CEOT} = \sum_{h=0}^{1}\sum_{k=1}^{N-1} \pi_h \left[P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) P(\Gamma \geq T|\mathcal{H}_h) + P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right) P(\Gamma < T|\mathcal{H}_h)\right], \tag{24}$$

worst case given $\Gamma \geq T$ would be that the magnitude of local decisions are ordered in ascending order expressed as

$$|z_{(1)}| \leq |z_{(2)}| \cdots \leq |z_{(N)}|, \tag{22}$$

where $z_{(k)} \in \{0, 1\}$ for $\forall k \in \{1, 2, \ldots, N\}$ is the k^{th} smallest local decision.

Similar to the above discussion, for the best case, we try to find the minimum k^* needed to make a final decision for a given set of local decisions $\{u_i\}_{i=1}^{N}$. The best case given $\Gamma < T$ would be that the magnitude of local decisions are ordered in ascending order as shown in (22), The best case given $\Gamma \geq T$ would be that the magnitude of local decisions are ordered in descending order as shown in (21). Based on the above analysis, we have the following theorem.

Theorem IV.2. *The average number of transmissions saved $\bar{N}_{s,CEOT}$ can be bounded as $\bar{N}^{L}_{s,CEOT} \leq \bar{N}_{s,CEOT} \leq \bar{N}^{U}_{s,CEOT}$. Here, the upper bound $\bar{N}^{U}_{s,CEOT}$ and the lower bound $\bar{N}^{L}_{s,CEOT}$ are given in* (23) *and* (24)*, respectively, where $P(\Gamma \geq T|\mathcal{H}_h) = \sum_{i=T}^{N} \binom{N}{i} \pi_{1,h}^{i} \pi_{0,h}^{N-i}$, $P(\Gamma < T|\mathcal{H}_h) = 1 - P(\Gamma \geq T|\mathcal{H}_h)$, and*

$$P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{N-T} \binom{N}{i} \pi_{0,h}^{i} \pi_{1,h}^{N-i}, \tag{25}$$

$$P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{min(N-T,k-T)} \binom{N}{i} \pi_{0,h}^{i} \pi_{1,h}^{N-i}, \tag{26}$$

when $k \geq T$, and

$$P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{T-1} \binom{N}{i} \pi_{1,h}^{i} \pi_{0,h}^{N-i}, \tag{27}$$

$$P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{min(T-1,k-(N-T+1))} \binom{N}{i} \pi_{1,h}^{i} \pi_{0,h}^{N-i}, \tag{28}$$

when $k > N - T$. Otherwise, $P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right)$, $P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right)$, $P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right)$ and $P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right)$ are equal to 0. Here, k_0^* and k_1^* denote the minimum number of transmissions needed to make a final decision for descending and ascending ordered local decisions, respectively.

Proof: Please see Appendix D. ∎

V. Simulation Results

In this section, we present the numerical results. We set the channel noise variance $\sigma^2 = 1$ and the prior probabilities $\pi_1 = \pi_0 = 0.5$. The detection performance in Fig. 1 and the actual average number of transmissions saved in Figs. 3, 5, 7, 8 are obtained via Monte Carlo method with 10^4 trials and the average number of transmissions saved in Figs. 2 and 4 are obtained via Monte Carlo method with 10^7 trials. In order to obtain an accurate evaluation of the average number of transmissions saved in the network as shown in Figs. 2 and 4, we need to significantly increase the number of trials as the number of sensors increases. Note that the other parameters like the perturbation noise variance σ_w^2, signal strength s, attack strength D, total number of sensors N, and the probability of each sensor being Byzantine α required for the simulations are included in the respective captions of the figures.

Comparison of OT-based and CEOT-based systems: Fig. 1 shows the effect of additive Byzantine attacks on the detection performance of the OT-based system and the CEOT-based system. In Fig. 1, we compare the probability of error of the CEOT-based system to the OT-based system and we observe that the CEOT-based system is more robust to additive Byzantine attacks with the same attack parameters. This is due to the fact that the global statistic of an OT-based system is a summation of LLRs, and some of these could be falsified to very large values when D/s is large. In this case, a large deviation is generated from the actual summation of LLRs. However, the global statistic of the CEOT-based system is the summation of quantized LLRs. Although some Byzantine nodes may falsify data, it is unlikely to lead to a significant deviation in the sum of quantized LLR values, even if D/s is large. Hence, D has less negative impact on the probability of error of the CEOT-based system than the OT-based system.

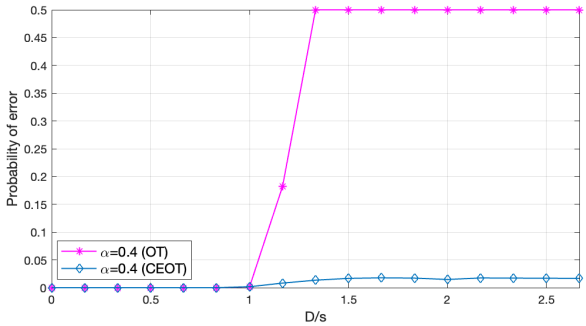


Fig. 1: P_e as a function of D/s in the CEOT-based system and the OT-based system when $s = 3$ and $N = 300$.

Effect of additive Byzantine attacks on $\bar{N}_s/N$ in the OT-based system: Figs. 2 and 3 show the effect of additive Byzantine attacks on the average percentage of savings for the OT-based system. Fig. 2 presents the average percentage of saving $\bar{N}_s/N$ in an OT-based system as a function of D/s for different values of α. Initially, $\bar{N}_s/N$ decreases when D/s increases. However, when D/s increases further, the FC starts to make wrong decisions and the number of transmissions needed to make the final decision starts to decrease and the savings start to increase. We compare the results obtained via simulation using the Monte Carlo method with our analysis

using (8), and observe a good match. In Fig. 3, we observe

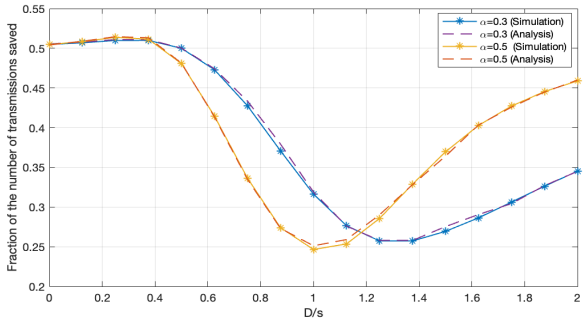


Fig. 2: Comparison of $\bar{N}_s/N$ as a function of D/s for different values of α when $s = 4$ and $N = 10$ in the OT-based system.

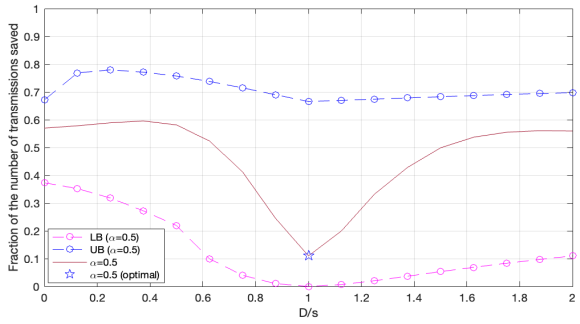
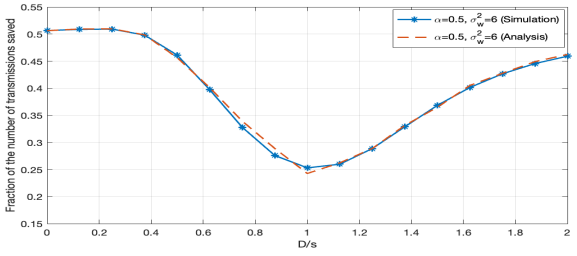


Fig. 3: Benchmarking UB and LB for $\bar{N}_s/N$ as a function of D/s for $\alpha = 0.5$ when $s = 4$ and $N = 300$ in the OT-based system.

that the attack strength D^* obtained in (15) for the OT-based system is near the point where the average percentage of savings is minimum. Compared with the OT-based system when no Byzantines are present, i.e., D=0, the system in the presence of attacks needs more transmissions to make a final decision. Therefore, the attack strength D^* from (15) not only blinds the FC but also leads to a smaller average percentage of savings. Fig. 3 also shows the UB and LB for the average percentage of savings as a function of D/s in an OT-based system. We observe that both the LB obtained in (14) and the UB obtained in (13) show a similar trend as that of the average percentage of saving, i.e., the UB and LB track the change in actual average number of transmissions that have been saved. Compared to the UB, the LB performs better in tracking the changes, which enables us to infer what the optimal attack strategy for the attacker is, i.e., what is the value of D that the attacker will employ to cause the greatest damage to the system. As for the UB, it provides more information regarding the maximum number of average transmissions saved in the network as well as alerts about the existence of outliers. For example, if the average number of transmissions saved is larger than the maximum value of UB, we can determine that there are potential outliers and they deviate far from the actual data, i.e., the attackers use an extremely large value of D.

Figs. 4 and 5 illustrate the impact of mean-variance-shift attacks on the average percentage of savings for the OT-

(a)



(b)

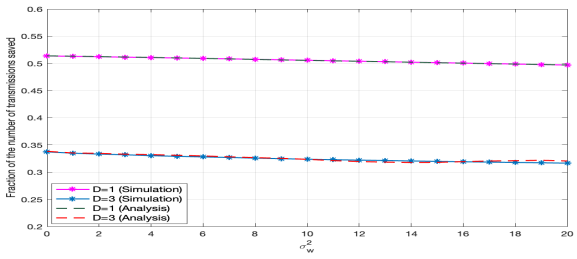


Fig. 4: (a) $\bar{N}_s/N$ as a function of D/s for $\sigma_w^2 = 6$; and (b) $\bar{N}_s/N$ as a function of σ_w^2 when $\alpha = 0.5$, $s = 4$, and $N = 10$ in the OT-based system under mean-variance-shift attacks.

based system. Fig. 4 shows that the error probability obtained via simulation using the Monte Carlo method and our error probability analysis have a good match. Fig. 5 shows the UB and LB we obtained that show a similar trend as that of the average percentage of saving. As we can observe, Figs. 4 (a) and 5 (a) demonstrate very similar trends as Figs. 2 and 3 when the mean of perturbation noise changes. In Figs. 4 (b) and 5 (b), we can observe that the values of the variance of the perturbation noise do not significantly affect the average number of transmissions saved. This phenomenon may arise from the fact that the change in variance value only affects the extent to which the noise samples deviate from the mean. Consequently, samples of perturbation noise might fall below or exceed the mean perturbation noise value. Given that the FC's global statistic is the summation of received LLRs, the overall perturbation to this statistic corresponds to the accumulation of perturbation noise from malicious nodes. The average perturbation for each malicious node will tend to the mean of the perturbation noise as the perturbation noise samples below the mean value will balance out the samples above the mean value. Therefore, the change in variance does not have as significant an effect as the change in mean on the number of transmissions saved in the network.

When we consider the case where the sensor observations follow an exponential distribution $f(y) = \frac{1}{\lambda}e^{-\frac{y}{\lambda}}$(a non-Gaussian distribution) with $\lambda = 2$ under hypothesis $\mathcal{H}_0$ and $\lambda = 8$ under hypothesis $\mathcal{H}_1$, we can observe a trend in Fig. 6 similar to that shown in Fig. 2 regarding the fraction of the number of transmissions saved as a function of D/s.

Effect of additive Byzantine attacks on $\bar{N}_s/N$ in the CEOT-based system: Figs. 7, and 8 show the effect of additive Byzantine attacks on the average percentage of savings for

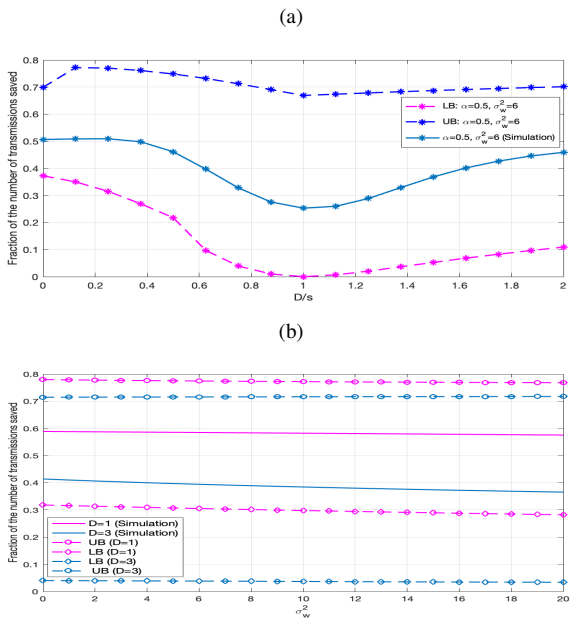


Fig. 5: Benchmarking UBs and LBs for (a) $\bar{N}_s/N$ as a function of D/s; and (b) $\bar{N}_s/N$ as a function of σ_w^2 when $\alpha = 0.5$, $s = 4$ and $N = 300$ in the OT-based system under the mean-variance-shift attacks.

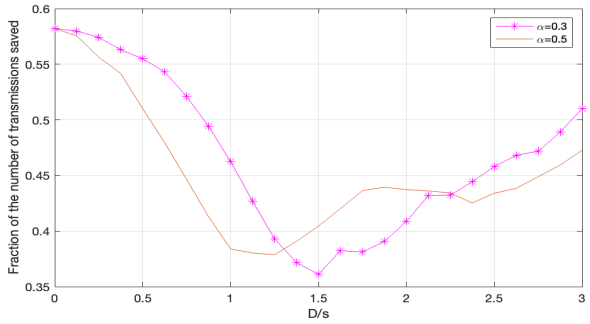


Fig. 6: $\bar{N}_s/N$ as a function of D/s in the OT-based system for different values α when $\lambda = 2$ under hypothesis $\mathcal{H}_0$, $\lambda = 8$ under hypothesis $\mathcal{H}_1$ for exponentially distributed observations and $N = 50$.

the CEOT-based system. Fig. 7 shows the UBs obtained in (23) and LBs obtained in (24) for the average percentage of saving $\bar{N}_{s,CEOT}/N$ as a function of D/s for different values of α. We observe that Byzantine sensors have more negative impact on the final decision making process with an increasing D/s. However, the additive Byzantine attacks have limited negative impact on the number of transmissions saved in the CEOT-based system compared to the OT-based system. When D/s is large enough, the first several local decisions received by the FC are most likely from Byzantine sensors which is the worst case scenario in terms of the performance for the system. With further increase of D/s, the impact of Byzantines on the number of transmissions saved in the network does not further increase since the LLRs are quantized, which limits the negative impact of Byzantine sensors on the system. In Fig. 8(a), the average percentage of saving, the UB and the LB are shown for a system with a relatively weak signal $s = 3$. Furthermore, Fig. 8(b) shows the the plots for a system with a relatively strong signal $s = 6$. By comparing Fig. 8(a) and Fig. 8(b), we observe that the LB gets tighter when we increase the signal strength s. This is reasonable due to the facts that i) Assumption 2 always works for honest sensors when s is sufficiently large; ii) the first several local decisions received by the FC are most likely from Byzantine sensors when D/s is sufficiently large. The above two reasons make the error probability of the CEOT-based system with a sufficiently large D/s approach the LB of the error probability we obtained in Theorem IV.2.

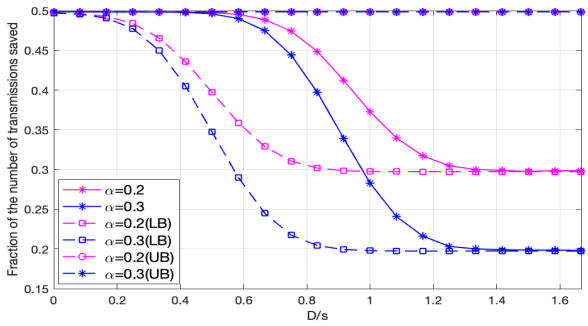


Fig. 7: Benchmarking UBs and LBs for $\bar{N}_{s,CEOT}/N$ as a function of D/s with different values of α when $s = 6$ and $N = 300$ in the CEOT-based system.

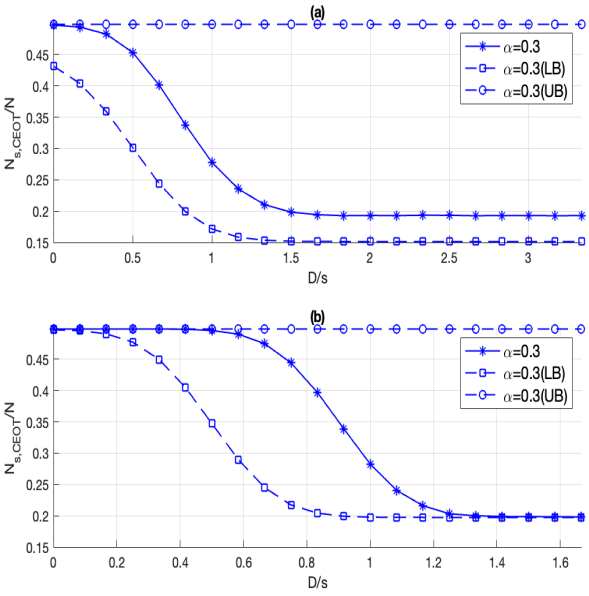


Fig. 8: Benchmarking UBs and LBs for $\bar{N}_{s,CEOT}/N$ as a function of D/s (a) when $s = 3$ and $N = 300$; (b) when $s = 6$ and $N = 300$ in the CEOT-based system.

VI. Conclusion

In this paper, we investigated the effect of additive Byzantine attacks on the performance of both the OT-based system

and the CEOT-based system for a binary hypothesis testing problem in distributed networks. We derived the error probabilities for both systems under attack and the number of transmissions saved. We also obtained the upper and the lower bounds on the number of transmissions saved for both systems under attack. The optimal attack strategy was investigated based on the deflection coefficient for the OT-based system. The simulation results showed that the Byzantine sensors can both maximize the probability of error and significantly increase the number of transmissions needed to make the final decision when they adopt the optimal attack strategy. A comparison of detection performance between the OT-based system and the CEOT-based system showed that the CEOT-based system is more robust, irrespective of the attacker's strategy. In the future, we intend to consider a more general assumption that Byzantines do not have the true information about the state of nature and they act based on their decisions.

APPENDIX A
PROOF OF LEMMA III.1

According to the fusion rule given in (2), we can infer that when inequality $\sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda$ holds, the FC can decide $\mathcal{H}_1$ based on the first k received transmissions. Similarly, when inequality $\sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda$ holds, the FC can decide $\mathcal{H}_0$ based on the first k received transmissions. The minimum value of k that satisfies either of the inequalities in (2), i.e., the minimum number of transmissions required to make a decision, is denoted as

$$k_{min} = \begin{cases} k_U^* & \text{when the FC decides } \mathcal{H}_0 \\ k_L^* & \text{when the FC decides } \mathcal{H}_1, \end{cases} \tag{29}$$

where $k_U^* = \underset{1\le k\le N}{\arg\min} \left\{ \sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda \right\}$ and $k_L^* = \underset{1\le k\le N}{\arg\min} \left\{ \sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda \right\}$[7] denote the minimum number of transmissions required to decide $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Under $\mathcal{H}_0$ ($k_{min} = k_U^*$), we have

$$Z_U = \sum_{i=1}^{k_{min}} L_{[i]} + (N-k_{min})|L_{[k_{min}]}| \ge \sum_{i=1}^{N} L_{[i]} = Z, \tag{30}$$

and under $\mathcal{H}_1$ ($k_{min} = k_L^*$), we have

$$Z_L = \sum_{i=1}^{k_{min}} L_{[i]} - (N-k_{min})|L_{[k_{min}]}| \le \sum_{i=1}^{N} L_{[i]} = Z. \tag{31}$$

This is because of the fact that $|L_{[1]}| > |L_{[2]}| > \cdots > |L_{[N]}|$). Note that $k_{min} = k_U^*$ is equivalent to $Z_U < \lambda$, and $k_{min} = k_L^*$ is equivalent to $Z_L > \lambda$. Based on (30) and (31), we can easily infer that $Pr(Z < \lambda | Z_U < \lambda) = 1$ and $Pr(Z > \lambda | Z_L > \lambda) = 1$, respectively. On the other hand, since

$$Z > \lambda \Leftrightarrow \sum_{i=1}^{k} L_{[i]} + \sum_{i=k+1}^{N} L_{[i]} > \lambda \tag{32a}$$

$$\Longrightarrow \sum_{i=1}^{k} L_{[i]} > \lambda - \sum_{i=k+1}^{N} L_{[i]} \ge \lambda - (N-k)|L_{[k]}| \tag{32b}$$

$$\Longrightarrow \sum_{i=1}^{k} L_{[i]} > \lambda - (N-k)|L_{[k]}| \tag{32c}$$

holds $\forall k$, from the definition of k_U^*, it becomes evident that the FC is unable to make a decision $\mathcal{H}_0$ for any value of k. So if $Z > \lambda$, we have $Pr(k_{min} = k_U^*) = 0$ and $Pr(k_{min} = k_L^*) = 1$, i.e., $Pr(Z_U < \lambda) = 0$ and $Pr(Z_L > \lambda) = 1$. It can be concluded that $Pr(Z_L > \lambda | Z > \lambda, \mathcal{H}_j) = 1$. Following a similar procedure, we can also obtain $Pr(Z_U < \lambda | Z < \lambda, \mathcal{H}_j) = 1$.

Based on the above analysis, we can calculate $Pr(Z_L > \lambda | \mathcal{H}_j)$ according to Bayesian rule given as

$$\begin{aligned} Pr(Z_L > \lambda | \mathcal{H}_j) &= \frac{Pr(Z_L > \lambda | Z > \lambda, \mathcal{H}_j) Pr(Z > \lambda | \mathcal{H}_j)}{Pr(Z > \lambda | Z_L > \lambda, \mathcal{H}_j)} \\ &= Pr(Z > \lambda | \mathcal{H}_j). \end{aligned} \tag{33}$$

Similarly, we obtain $Pr(Z_U < \lambda | \mathcal{H}_j) = Pr(Z < \lambda | \mathcal{H}_j)$. Hence, the probability of error of the the OT-based system is given as

$$\begin{aligned} P_e^{(OT)} &= \pi_0 Pr(Z_L > \lambda | \mathcal{H}_0) + \pi_1 Pr(Z_U < \lambda | \mathcal{H}_1) \\ &= \pi_0 Pr(Z > \lambda | \mathcal{H}_0) + \pi_1 Pr(Z < \lambda | \mathcal{H}_1) = P_e^{(opt)}, \end{aligned} \tag{34}$$

where $P_e^{(opt)}$ is the error probability of the unordered system.

APPENDIX B
PROOF OF THEOREM III.2

Let $\bar{N}_t$ denote the average number of transmissions in the network. $\bar{N}_t$ is given as

$$\bar{N}_t = E(k^*) = \sum_{k=1}^{N} k Pr(k^* = k) = \sum_{k=1}^{N} Pr(k^* \ge k) \tag{35a}$$

$$= \sum_{k=1}^{N} Pr(k^* \ge k | \mathcal{H}_0)\pi_0 + Pr(k^* \ge k | \mathcal{H}_1)\pi_1, \tag{35b}$$

where $Pr(k^* \ge k)$ is the probability that at least k transmissions in the network are needed to make the final decision. Note that k^* is the minimum number of observations/transmissions required to make a decision. k can be considered as the number of observations that have already been received by the FC. The global statistic at the FC is given by $\sum_{i=1}^{k} L_{[i]}$, where $\sum_{i=1}^{k} L_{[i]}$ represents the accumulated LLRs up to the k^{th} transmission at the FC. Next Lemma helps us to obtain the probability of the event that at least k transmissions are required to make the final decision.

Lemma B.1. *The FC can not decide $\mathcal{H}_1$ or $\mathcal{H}_0$ until the FC has received at least k transmissions if $\sum_{i=1}^{k-1} L_{[i]}$ satisfies*

[7]Please note that if there is no $k \in \{1, 2, \ldots, N\}$ that satisfies the condition $\sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda$ (or if there is no $k \in \{1, 2, \ldots, N\}$ that satisfies $\sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda$), we define $k_U^* = \arg\min \emptyset = 0$ (or $k_L^* = \arg\min \emptyset = 0$).

both $\sum_{i=1}^{k-1} L_{[i]} \leq \lambda + (N-k+1)|L_{[k-1]}|$ *and* $\sum_{i=1}^{k-1} L_{[i]} \geq \lambda - (N-k+1)|L_{[k-1]}|$.

Proof: When the FC received the first $(k-1)$ LLRs, i.e, $[L_{[1]}, L_{[2]}, \ldots, L_{[k-1]}]$, we discuss the cases that the FC can not decide $\mathcal{H}_1$ and the FC can not decide $\mathcal{H}_0$.

Recall that $|L_{[1]}| \geq |L_{[2]}| \cdots \geq |L_{[N]}|$, we have $Z \leq \sum_{i=1}^{k-1} L_{[i]} + (N-k+1)|L_{[k-1]}| = \eta_U$. Obviously, the FC is not able to decide $\mathcal{H}_0$ when $\eta_U > \lambda$. Moreover, (36) shows that if the FC doesn't decide $\mathcal{H}_0$ after receiving the first $(k-1)$ LLRs, it can't decide $\mathcal{H}_0$ after receiving the first $(k-2)$ observations.

$$\begin{aligned}\eta_U =& \sum_{i=1}^{k-1} L_{[i]} + (N-k+1)|L_{[k-1]}| && \text{(36a)}\\ =& \sum_{i=1}^{k-2} L_{[i]} + (N-k+2)|L_{[k-2]}| + (N-k+1) \\ & \times (|L_{[k-1]}| - |L_{[k-2]}|) + (L_{[k-1]} - |L_{[k-2]}|). && \text{(36b)}\end{aligned}$$

As $|L_{[k-1]}| \leq |L_{[k-2]}|$ and $L_{[k-1]} \leq |L_{[k-1]}| \leq |L_{[k-2]}|$, we have $|L_{[k-1]}| - |L_{[k-2]}| \leq 0$ and $L_{[k-1]} - |L_{[k-2]}| \leq 0$ in (36b). Hence, we can obtain that (36b)$> \lambda$ implies $\sum_{i=1}^{k-2} L_{[i]} + (N-k+2)|L_{[k-2]}| > \lambda$. Following the similar procedure as shown in (36), we are able to conclude that if the FC can't decide $\mathcal{H}_0$ after receiving the first $(k-1)$ LLRs, it can't decide $\mathcal{H}_0$ after receiving 0 or 1 or $\ldots$, or $(k-2)$ observations.

we can obtain that $\eta_L = \sum_{i=1}^{k-1} L_{[i]} - (N-k+1)|L_{[k-1]}| \leq Z$ after the FC has received the first $(k-1)$ LLRs. Obviously, the FC can not decide $\mathcal{H}_1$ when $\eta_L < \lambda$. Following the similar procedure as shown in (36), we can prove that if the FC can't decide $\mathcal{H}_1$ after receiving the first $(k-1)$ largest LLRs, it can't decide $\mathcal{H}_1$ after receiving 0 or 1 or $\ldots$, or $(k-2)$ observations. The proof for this is similar as above and is skipped. ∎

To evaluate $Pr(k^* \geq k|\mathcal{H}_h)$, we have

$$\begin{aligned}& Pr(k^* \geq k|\mathcal{H}_h) \\ =& \int_{\mathbf{l}_{k-1} \in \mathcal{J}} f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h) \mathrm{d}l_1 \ldots \mathrm{d}l_{k-1}, && \text{(37)}\end{aligned}$$

where $f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h)$ is the joint pdf of $l_{[1]}, l_{[2]}, \ldots, l_{[k-1]}$ given $\mathcal{H}_h$ for $h = 0, 1$. According to [32], the joint pdf of $l_{[1]}, l_{[2]}, \ldots, l_{[k-1]}$ given $\mathcal{H}_h$ is given as

$$\begin{aligned}& f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h) \\ =& \frac{N!}{(N-k+1)!} \left[\prod_{i=1}^{k-1} f_L(l_i|\mathcal{H}_h)\right] \left[F_{|L|}(l_{k-1}|\mathcal{H}_h)\right]^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}} && \text{(38)}\end{aligned}$$

where $\mathcal{J} = \mathcal{L} \bigcap \mathcal{U} \bigcap \mathcal{D}$ is the intersection of hyperplanes $\mathcal{L}$, $\mathcal{U}$ and $\mathcal{D}$, and $F_{|L_k|}(l_k|\mathcal{H}_h)$ is the cdf of $|L_k|$ for $h = 0, 1$. By substituting (38) in (37) and utilizing the law of total expectation, (37) can be rewritten as

$$\begin{aligned}& Pr(k^* \geq k|\mathcal{H}_h) \\ =& E_{\mathbf{L}_{k-1}} \left[\frac{N!}{(N-k+1)!} \left[F_{|L|}(l_{k-1}|\mathcal{H}_h)\right)\right]^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}}\right] && \text{(39)}\end{aligned}$$

for $h = 0, 1$, where $F_{|L|}(l_{k-1}|\mathcal{H}_h))$ is given in (7). Note that the Byzantines affect the average number of transmissions by affecting attack parameters (α, D) in $F_{|L|}(L_{k-1}|\mathcal{H}_h)$.[8]

APPENDIX C
PROOF OF THEOREM III.3

Let $\bar{N}_s$ denote the average number of transmissions saved in the network given as

$$\begin{aligned}\bar{N}_s =& \sum_{k=1}^{N} (N-k) Pr(k^* = k) = \sum_{k=1}^{N-1} Pr(k^* \leq k) && \text{(40a)}\\ =& \sum_{k=1}^{N-1} Pr(k^* \leq k|\mathcal{H}_0)\pi_0 + Pr(k^* \leq k|\mathcal{H}_1)\pi_1. && \text{(40b)}\end{aligned}$$

Next, we use the following lemma from [32, Chapter 5] to prove Theorem 2.

Lemma C.1. *According to Cauchy-Schwarz inequality, we have*

$$|\sum c_i (L_{[i]} - \bar{L})| \leq [\sum (c_i - \bar{c})^2 (N-1) v]^{\frac{1}{2}} \qquad (41)$$

in terms of empirical mean $\bar{L}$ and empirical variance v for any constants $\{c_i\}_{i=1}^N$. If c_i is non-increasing when i increases, the bound is sharp.

From Lemma C.1, we have $|\sum_{i=1}^{k} L_{[i]} - k\bar{L}| \leq [\sum (c_i - \bar{c})^2 (N-1) v]^{\frac{1}{2}}$ if we let $c_1 = c_2 = \cdots = c_k = 1$ and $c_{k+1} = \cdots = c_N = 0$. Hence, the LB and the UB of $\sum_{i=1}^{k} L_{[i]}$ are given by $g_L \leq \sum_{i=1}^{k} L_{[i]} \leq g_U$, where $g_L = -[\sum (c_i - \bar{c})^2 (N-1) v]^{\frac{1}{2}} + k\bar{L}$ and $g_U = [\sum (c_i - \bar{c})^2 (N-1) v]^{\frac{1}{2}} + k\bar{L}$.

a) LB of $\bar{N}_s$: When the FC decides $\mathcal{H}_1$ in at most k transmissions given hypothesis $\mathcal{H}_h$, we have

$$Pr(k^* \leq k|\mathcal{H}_h) = Pr(\sum_{i=1}^{k} L_{[k]} > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h). \quad (42)$$

for $h = 0, 1$. It is easy to show that $g_L > \lambda + (N-k)|L_{[k]}|$ implies $\sum_{i=1}^{k} L_{[i]} > \lambda + (N-k)|L_{[k]}|$. Hence, from (42), we get

$$Pr(k^* \leq k|\mathcal{H}_h) \geq Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h) \quad (43)$$

Similarly, when the FC decides $\mathcal{H}_0$ in at most k transmissions given hypothesis $\mathcal{H}_h$, we get

$$Pr(k^* \leq k|\mathcal{H}_h) \geq Pr(g_U < \lambda - (N-k)|L_{[k]}||\mathcal{H}_h) \quad (44)$$

The inequality in (44) is true due to the fact that $g_U < \lambda - (N-k)|L_{[k]}|$ implies $\sum_{i=1}^{k} L_{[i]} < \lambda - (N-k)|L_{[k]}|$. Substituting $Pr(k^* \leq k|\mathcal{H}_0)$ and $Pr(k^* \leq k|\mathcal{H}_1)$ in (40) with their LBs $Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h)$ and $Pr(g_U < \lambda - (N-k)|L_{[k]}||\mathcal{H}_h)$, respectively, we get

$$\begin{aligned}\bar{N}_s \geq & \sum_{k=1}^{N-1} \sum_{h=0}^{1} \pi_h \big[Pr(g_L > \lambda + n_{UT}|L_{[k]}||\mathcal{H}_h) \\ & + Pr(g_U < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h)\big] && \text{(45)}\end{aligned}$$

[8] D affects η_1 and η_0 in $F_{|L|}(L_{k-1}|\mathcal{H}_h)$.

where $n_{UT} = N - k$. A Monte Carlo approach can be utilized to evaluate $Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h)$ and $Pr(g_U > \lambda-(N-k)|L_{[k]}||\mathcal{H}_h)$. We generate M_2 realizations of $L_{[1]}, L_{[2]}, \ldots, L_{[N]}$ so that the empirical mean $\bar{L}$ and the empirical variance v can be calculated. When M_2 is sufficiently large, $\bar{L}$ approaches the population mean. The population mean and the population variance under $\mathcal{H}_h$ are, respectively, expressed as

$$\delta_h = E[L_i|\mathcal{H}_h] = \alpha\eta_h + (1-\alpha)\mu_h, \quad \zeta_h^2 = E[L_i^2|\mathcal{H}_h] - \delta_h^2, \quad (46)$$

where $E[L_i^2|\mathcal{H}_h] = \alpha E[L_i^2|\mathcal{H}_h, i = B] + (1-\alpha)E[L_i^2|\mathcal{H}_h, i = H] = \beta + \alpha\eta_h^2 + (1-\alpha)\mu_h^2$ for $h = 0, 1$. Substituting the parameters $(\bar{L}, v)$ in (45) with parameters $(\delta_h, \frac{N}{N-1}\zeta_h^2)$ under $\mathcal{H}_h$ for $h = 0, 1$ yields

$$\bar{N}_s \geq \sum_{k=1}^{N-1}\sum_{h=0}^{1}\pi_h \left[Pr\left(|L_{[k]}| < \frac{g_L - \lambda}{(N-k)}|\mathcal{H}_h\right) + Pr\left(|L_{[k]}| < \frac{\lambda - g_U}{(N-k)}|\mathcal{H}_h\right)\right], \quad (47)$$

where $Pr\left(|L_{[k]}| < r|\mathcal{H}_h\right) = \int_0^r f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h)\mathrm{d}l_{[k]}$ for $r \in \{\frac{g_L-\lambda}{(N-k)}, \frac{\lambda-g_U}{(N-k)}\}$. It is given in closed form as [32]

$$f_{L_{[k]}}(l_{[k]}|\mathcal{H}_h) = N f_L(l_{[k]}|\mathcal{H}_h)\binom{N-1}{k-1} \times F_L(l_{[k]}|\mathcal{H}_h)^{(N-k)}(1 - F_L(l_{[k]}|\mathcal{H}_h))^{(k-1)}. \quad (48)$$

Hence, the pdf of $f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h)$ is given by

$$f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h) = \frac{\mathrm{d}Pr(|L_{[k]}| \leq l_{[k]})}{\mathrm{d}l_{[k]}} = \begin{cases} f_{L_{[k]}}(l_{[k]}|\mathcal{H}_h) - f_{L_{[k]}}(-l_{[k]}|\mathcal{H}_h) & \text{if } l_{[k]} \geq 0 \\ 0 & \text{if } l_{[k]} < 0 \end{cases} \quad (49)$$

Substituting (49) in (47), we are able to obtain the lower bound of the number of transmissions saved.

b) UB of $\bar{N}_s$: It is easy to show that $\sum_{i=1}^{k} L_{[k]} > \lambda + (N-k)|L_{[k]}|$ implies $g_U > \lambda + (N-k)|L_{[k]}|$. Hence, from (42), we get

$$Pr(g_U > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h) \geq Pr(k^* \leq k|\mathcal{H}_h) \quad (50)$$

Similarly, due to the fact that $\sum_{i=1}^{k} L_{[i]} < \lambda - (N-k)|L_{[k]}|$ implies $g_L < \lambda - (N-k)|L_{[k]}|$, we can also get

$$Pr(g_L < \lambda - (N-k)|L_{[k]}||\mathcal{H}_h) \geq Pr(k^* \leq k|\mathcal{H}_h). \quad (51)$$

Hence, we have

$$\bar{N}_s \leq \sum_{k=1}^{N-1}\sum_{h=0}^{1} Pr\big(g_U > \lambda + n_{UT}|L_{[k]}| \text{ or } g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h\big)\pi_h, \quad (52)$$

where $n_{UT} = N - k$ and

$$\begin{aligned} &Pr(g_U > \lambda + n_{UT}|L_{[k]}| \text{ or } g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\ &= Pr(g_U > \lambda + n_{UT}|L_{[k]}||\mathcal{H}_h) + Pr(g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\ &\quad - Pr(g_U > \lambda + n_{UT}|L_{[k]}| \text{ and } g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\ &= Pr\left(|L_{[k]}| \leq \frac{g_U - \lambda}{N-k}|\mathcal{H}_h\right) + Pr\left(|L_{[k]}| \leq \frac{\lambda - g_L}{N-k}|\mathcal{H}_h\right) \\ &\quad - Pr\left(|L_{[k]}| \leq \min\left(\frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k}\right)|\mathcal{H}_h\right). \end{aligned} \quad (53)$$

Following the similar procedure when we obtain the LB of $\bar{N}_s$, we can get the UB of $\bar{N}_s$. Then, we can obtain the UB and the LB in Theorem III.2.

APPENDIX D
PROOF OF THEOREM IV.2

According to Equation (20), $\bar{N}_{s,CEOT}$ is given as

$$\begin{aligned} \bar{N}_{s,CEOT} &= \sum_{k=1}^{N-1} Pr(k^* \leq k) \\ &= \sum_{k=1}^{N-1} Pr(k^* \leq k|\Gamma < T)Pr(\Gamma < T) \\ &\quad + Pr(k^* \leq k|\Gamma \geq T)Pr(\Gamma \geq T). \end{aligned} \quad (54)$$

a) LB of $\bar{N}_{s,CEOT}$: Recall that k_0^* and k_1^* denote the minimum number of transmissions needed to make a final decision for descending and ascending ordered local decisions, respectively. It is easy to show that $k_1^* \leq k$ implies $k^* \leq k$ given $\Gamma \geq T$ and $k_0^* \leq k$ implies $k^* \leq k$ given $\Gamma < T$. Hence, we have

$$Pr(k^* \leq k|\Gamma \geq T) \geq Pr(k_1^* \leq k|\Gamma \geq T), \quad (55)$$

$$Pr(k^* \leq k|\Gamma < T) \geq Pr(k_0^* \leq k|\Gamma < T). \quad (56)$$

Substituting $Pr(k^* \leq k|\Gamma < T)$ and $Pr(k^* \leq k|\Gamma \geq T)$ in (54) with their LBs $Pr(k_0^* \leq k|\Gamma < T)$ and $Pr(k_1^* \leq k|\Gamma \geq T)$, respectively, we get

$$\begin{aligned} \bar{N}_{s,CEOT} \geq &\sum_{h=0}^{1}\sum_{k=1}^{N-1} P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right)P(\Gamma \geq T|\mathcal{H}_h)\pi_h \\ &+ \sum_{k=1}^{N-1} P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right)P(\Gamma < T|\mathcal{H}_h)\pi_h. \end{aligned} \quad (57)$$

Since $z_{(i)}$ and $z_{[i]}$ for $\forall i \in \{1, 2, \ldots, N\}$ are non-negative, we have $0 \leq \sum_{i=1}^{k_1^*} z_{(i)} \leq k_1^*$ and $0 \leq \sum_{i=1}^{k_0^*} z_{[i]} \leq k_0^*$. For the fusion rule of equivalent worst case given $\Gamma \geq T$, which is given as

$$\sum_{i=1}^{k_1^*} z_{(i)} \geq T \quad \text{decides } \mathcal{H}_1, \quad (58)$$

where $k_1^* \geq T$ is needed to make a decision $\mathcal{H}_1$.

For the fusion rule of equivalent worst case given $\Gamma < T$ given as

$$\sum_{i=1}^{k_0^*} z_{[i]} < T - (N - k_0^*) \quad \text{decides } \mathcal{H}_0, \quad (59)$$

where $k_0^* > N - T$ is needed to make a decision $\mathcal{H}_0$. Hence, it is obvious that the FC can not make decision $\mathcal{H}_0$ given $\Gamma < T$ when $k_0^* \leq N - T$ and the FC can not make decision $\mathcal{H}_1$ given $\Gamma \geq T$ when $k < T$. Hence, we have

$$\sum_{k=1}^{T-1} P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = \sum_{k=1}^{N-T} P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right) = 0. \quad (60)$$

As shown in (22), the magnitude of local decisions are ordered in an ascending order, i.e., $|z_{(1)}| \leq |z_{(2)}|, \ldots, \leq |z_{(N)}|$, when

we consider the equivalent worst case given $\Gamma \geq T$. It is apparent that $\Gamma \geq T$ implies that the distributed system without ordering would make a decision of $\mathcal{H}_1$. According to Lemma IV.1, the detection performance of the CEOT-based system is the same as that of the distributed system without ordering. We can easily conclude that $k^* \leq k_1^*$ is always satisfied, which indicates that the minimum number of transmissions required to make a decision for equivalent worst case given $\Gamma \geq T$ is always greater than or equal to the actual minimum number of transmissions required. Since at most $min(N-T, k-T)$ 0s are required when $\Gamma \geq T$ for the unordered distributed system, we have

$$P\left(k_1^* \leq k | \Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{min(N-T,k-T)} \binom{N}{i} \pi_{0,h}^i \pi_{1,h}^{N-i}, \quad (61)$$

when $k \geq T$.

Similarly, as shown in (21), the magnitude of local decisions are ordered in a descending order, i.e., $z_{[1]} \geq z_{[2]}, \ldots, \geq z_{[N]}$, when we consider the equivalent worst case given $\Gamma < T$. Here, $\Gamma < T$ implies that the distributed system without ordering would make a decision of $\mathcal{H}_0$. According to Lemma IV.1, we can also easily conclude that $k^* \leq k_0^*$ is always satisfied, which means that the minimum number of transmissions required to make a decision for equivalent worst case given $\Gamma < T$ is always greater than or equal to the true minimum number of transmissions required. Since at most $min(T-1, k-(N-T+1))$ 1s are required when $\Gamma < T$ for the unordered distributed system, we have

$$P\left(k_0^* \leq k | \Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{min(T-1,k-(N-T+1))} \binom{N}{i} \pi_{1,h}^i \pi_{0,h}^{N-i} \quad (62)$$

if $k > N-T$.

b) UB of $\bar{N}_{s,CEOT}$: By substituting k_1^* in (58) with k_0^*, we can obtain the fusion rule of equivalent best case given $\Gamma \geq T$ where $k_0^* \geq T$ is needed to make a decision $\mathcal{H}_1$. Similarly, by substituting k_0^* in (59) with k_1^*, we can obtain the fusion rule of equivalent best case given $\Gamma < T$ where $k_1^* > N-T$ is needed to make a decision $\mathcal{H}_0$. It is easy to show that $k^* \leq k$ implies $k_1^* \leq k$ given $\Gamma < T$ and $k^* \leq k$ implies $k_0^* \leq k$ given $\Gamma \geq T$. Hence, we get

$$Pr(k_1^* \leq k | \Gamma < T) \geq Pr(k^* \leq k | \Gamma < T), \quad (63)$$
$$Pr(k_0^* \leq k | \Gamma \geq T) \geq Pr(k^* \leq k | \Gamma \geq T). \quad (64)$$

Substituting $Pr(k^* \leq k | \Gamma < T)$ and $Pr(k^* \leq k | \Gamma \geq T)$ in (54) with their UBs $Pr(k_1^* \leq k | \Gamma < T)$ and $Pr(k_0^* \leq k | \Gamma \geq T)$, respectively, we get

$$\begin{aligned}\bar{N}_{s,CEOT} \leq & \sum_{h=0}^{1} \sum_{k=1}^{N-1} P\left(k_1^* \leq k | \Gamma < T, \mathcal{H}_h\right) P(\Gamma < T | \mathcal{H}_h) \pi_h \\ & + \sum_{k=1}^{N-1} P\left(k_0^* \leq k | \Gamma \geq T, \mathcal{H}_h\right) P(\Gamma \geq T | \mathcal{H}_h) \pi_h. \end{aligned} \quad (65)$$

Following the similar procedure, we have

$$\sum_{k=1}^{N-T} P\left(k_1^* \leq k | \Gamma < T, \mathcal{H}_h\right) = \sum_{k=1}^{T-1} P\left(k_0^* \leq k | \Gamma \geq T, \mathcal{H}_h\right) = 0. \quad (66)$$

According to Lemma IV.1, it is apparent that $k_1^* \leq k^*$ is always satisfied if $\Gamma \geq T$, i.e., the minimum number of transmissions required to make a decision for equivalent best case given $\Gamma \geq T$ is always less than or equal to the actual minimum number of transmissions required. Similarly, we can also conclude that $k_0^* \leq k^*$ is always satisfied if $\Gamma < T$. Following a similar procedure to derive the LB, we obtain

$$P\left(k_1^* \leq k | \Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{T-1} \binom{N}{i} \pi_{1,h}^i \pi_{0,h}^{N-i}, \quad (67)$$

when $k > N-T$ and

$$P\left(k_0^* \leq k | \Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{N-T} \binom{N}{i} \pi_{0,h}^i \pi_{1,h}^{N-i}, \quad (68)$$

when $k \geq T$. Then, we obtain the UB, which is given by substituting (66), (67) and (68) in (65), and the LB, which is given by substituting (60), (61) and (62) in (57), in Theorem IV.2.

REFERENCES

[1] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, 2012.

[2] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, 1996.

[3] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, vol. 3. IEEE, 2003, pp. 1713–1723.

[4] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011.

[5] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3229–3235, 2008.

[6] Z. N. Rawas, Q. He, and R. S. Blum, "Energy-efficient noncoherent signal detection for networked sensors using ordered transmissions," in *2011 45th Annual Conference on Information Sciences and Systems*. IEEE, 2011, pp. 1–5.

[7] P. Braca, S. Marano, and V. Matta, "Single-transmission distributed detection via order statistics," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2042–2048, 2011.

[8] L. Hesham, A. Sultan, M. Nafie, and F. Digham, "Distributed spectrum sensing with sequential ordered transmissions to a cognitive fusion center," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2524–2538, 2012.

[9] Y. Chen, R. S. Blum, and B. M. Sadler, "Optimal quickest change detection in sensor networks using ordered transmissions," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.

[10] ——, "Ordering for communication-efficient quickest change detection in a decomposable graphical model," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4710–4723, 2021.

[11] Y. Chen, B. M. Sadler, and R. S. Blum, "Ordered gradient approach for communication-efficient distributed learning," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.

[12] N. Bessis, "A model to manage smart devices in mobile sensing applications," Ph.D. dissertation, Edge Hill University, UK, 2021.

[13] S. S. Gupta, S. K. Pallapothu, and N. B. Mehta, "Ordered transmissions for energy-efficient detection in energy harvesting wireless sensor networks," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2525–2537, 2020.

[14] S. S. Gupta and N. B. Mehta, "Ordered transmissions schemes for detection in spatially correlated wireless sensor networks," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1565–1577, 2020.

[15] N. Sriranga, K. G. Nagananda, R. S. Blum, A. Saucan, and P. K. Varshney, "Energy-efficient decision fusion for distributed detection in wireless sensor networks," in *2018 21st International conference on information fusion (FUSION)*. IEEE, 2018, pp. 1541–1547.

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

[17] 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.

[18] C.-Y. Wei, P.-N. Chen, Y. S. Han, and P. K. Varshney, "Local threshold design for target localization using error correcting codes in wireless sensor networks in the presence of Byzantine attacks," *IEEE transactions on information forensics and security*, vol. 12, no. 7, pp. 1571–1584, 2017.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] D. Dolev, "The Byzantine generals strike again," *Journal of algorithms*, vol. 3, no. 1, pp. 14–30, 1982.

[25] C. Quan, N. Sriranga, H. Yang, Y. S. Han, B. Geng, and P. K. Varshney, "Efficient ordered-transmission based distributed detection under data falsification attacks," *IEEE Signal Processing Letters*, vol. 30, pp. 145–149, 2023.

[26] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals and Systems*, vol. 1, no. 2, pp. 167–182, 1988.

[27] N. Nguyen-Thanh and I. Koo, "An enhanced cooperative spectrum sensing scheme based on evidence theory and reliability source evaluation in cognitive radio context," *IEEE Communications Letters*, vol. 13, no. 7, pp. 492–494, 2009.

[28] L. Zhang, Q. Wu, G. Ding, S. Feng, and J. Wang, "Performance analysis of probabilistic soft SSDF attack in cooperative spectrum sensing," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–9, 2014.

[29] B. Kailkhura, S. Brahma, and P. Varshney, "On the performance analysis of data fusion schemes with Byzantines," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 7411–7415.

[30] B. Kailkhura, S. Brahma, and P. K. Varshney, "Data falsification attacks on consensus-based detection systems," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 145–158, 2016.

[31] H. L. Van Trees, *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons, 2004.

[32] H. A. David and H. N. Nagaraja, *Order statistics*. John Wiley & Sons, 2004.