Minimum Byzantine Effort for Blinding Distributed Detection in Wireless Sensor Networks

Hsuan-Yin Lin, *Senior Member, IEEE,* Po-Ning Chen, *Senior Member, IEEE,* Yunghsiang S. Han, *Fellow, IEEE,* and Pramod K. Varshney, *Life Fellow, IEEE*

***Abstract*—In this work, we consider the general problem of distributed detection in the presence of Byzantines using wireless sensor networks. Instead of attempting to mitigate Byzantine attacks as a system designer, we investigate the issue from the perspective of a Byzantine attacker. The probability for each individual sensor to be compromised (compromising probability) required to blind the system operation is adopted as the attack measure. Under the system setting that the fusion center (FC) declares the most likely hypothesis to be true based on the M-ary data from N local sensors, a Byzantine attack policy that can blind the FC with the minimum compromising probability for each individual sensor is derived under the assumption that the Byzantine attacker knows the statistics of the local outputs. The closed-form expression for a blind-achieving Byzantine transition probability that is used to alter the statistics of the local outputs of compromised sensors is also established. Our results indicate that the statistics of the local outputs is essential for the minimization of an attacker's effort.**

***Index Terms*—Distributed detection, distributed inference network security, wireless sensor networks, Byzantine attacks**

I. Introduction

WIRELESS sensor networks (WSNs) have been studied for well over a couple of decades [1]–[5]. For applications employing WSNs, distributed inference plays an essential role and hence its design is one of the key problems that has been investigated extensively [6]–[10]. In a WSN, simple inexpensive sensors are deployed to observe a phenomenon of interest (POI) θ that is drawn from a finite set Θ of size L. Due to limited resources at each sensor, the observed local data is processed into an M-ary symbol for transmission to the fusion center (FC). A quantization is thus required at each sensor to convert the observed data into one of the M symbols. The sensors send the resulting M-ary outputs to the FC, where a global inference regarding the POI is made.

In applications involving inference over WSNs, an important issue is the robustness of global inference against hostile actions. A recent research direction regarding this issue has revolved around how to protect the global inference made by the FC when a fraction of sensors are compromised [11]–[16]. These compromised sensors are usually called *Byzantine sensors*, while the remaining sensors are called *Honest sensors*.[1] In their problem formulations, some researchers assume that only binary data are transmitted by the local sensors [18], and others restrict their focus on binary hypothesis testing [11]–[16]. Some prior works such as [11] have dealt with non-binary transmissions from the local sensors but consider only asymptotic blindness of the FC as the number of sensors grows without bound. In certain scenarios, coordination among sensors is also introduced [19].

Different from the above mentioned works, we attempt to approach the issue of inference robustness in a more practical non-asymptotic scenario, where the number of sensors N is finite. Furthermore, a general (possibly, non-binary) hypothesis test regarding a POI is performed. We assume that coordination among sensors is not feasible, hence sensors make their local decisions based only on their local observations. Such a general L-ary hypothesis testing problem formulation over a WSN of finite size is more applicable to practical scenarios than a system with only binary decisions or very large (infinite) number of sensors.

In 2014, an optimal Byzantine attack policy for distributed inference sensor networks was proposed in [17], where the attacker does not have any knowledge (or partial knowledge at best) about the true state of the POI, or quantization thresholds of the sensors, or their statistics. In the absence of such information, hostile actions at local sensors are restricted only to the modifications of the M-ary symbols transmitted to the FC. This surely limits the capability of a Byzantine attacker, who can only carry out the so-called "man-in-the-middle" attack by performing data falsification.

In this paper, we significantly extend the work presented in [17] by assuming that a Byzantine attacker is quite capable and is either endowed with the knowledge of the statistics of

H.-Y. Lin was with the Inst. of Comm. Eng., Nat'l Chiao Tung Univ., Hsinchu 30010, Taiwan, and is now with Simula UiB, N-5008, Bergen, Norway (email: hsuan-yin.lin@ieee.org)

P.-N. Chen is with the Inst. of Comm. Eng., Nat'l Chiao Tung Univ., Hsinchu 30010, Taiwan (email: poning@faculty.nctu.edu.tw).

Y. S. Han is with the School of Elec. Eng. & Intelligentization, Dongguan Univ. of Tech., Dongguan, 523808, P. R. China (E-mail: yunghsiangh@gmail.com).

P. K. Varshney is with the Dept. of Elec. Eng. and Comp. Science, Syracuse Univ., Syracuse, NY 13244 USA (email: varshney@ecs.syr.edu).

This work was presented in part at the *2016 IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, July 10–15, 2016.

This work has been supported by the Minister of Science and Technology of Taiwan (MOST 103-2911-I-011-515/104-2911-I-011-503), the National Natural Science Foundation of China (No. 61671007), Start Fund of Dongguan University of Technology (KCYXM2017025), and the Army Research Office of the USA (W911NF-14-1-0339). A part of the work was completed during a visit to the Center for Advanced Security Research Darmstadt (CASED), Technische Universität Darmstadt, Germany, during 2014-2016.

[1] Byzantine typically means that the compromised sensors know what the *Honest sensors* know as they are part of the inner circle. By following what has been assumed in [11], [16], [17], *Byzantine sensors* in this work refer to those sensors that have a complete knowledge of the statistics of local outputs (which is exactly the statistics of local outputs of the *Honest sensors* under the assumption of independent and identically distributed observations), but the *Byzantine sensors* do not know what the *Honest sensors* have observed.

local outputs or has the capability to learn it. In addition, the attacker is assumed to be aware of or can have a sufficiently good estimate of the POI θ, and it knows the true operational probability for each sensor to be independently compromised α. As a result, an attacker can devise an $M \times M$ Byzantine transition probability matrix $\mathbb{P}^{(\theta,\alpha)}$ with respect to each $\theta \in \Theta$ and $\alpha \in [0,1]$, according to which the M-ary local data of each compromised sensor is statistically converted. We then show analytically that with this additional knowledge, the FC can be blinded if, and only if, the probability that an individual sensor is compromised by an attacker is no less than a certain minimum value α^*_{blind}.[2] The achievability argument involves an effective statistical Byzantine attack policy $\mathbf{P}^{(\alpha)}:=\{\mathbb{P}^{(\theta,\alpha)}\}_{\theta\in\Theta}$ that can blind the FC with $\alpha \geq \alpha^*_{\text{blind}}$.

After determining the blind-achieving Byzantine transition probability policy $\mathbf{P}^{(\alpha)}$, practical concerns about the assumptions of perfect knowledge on α and θ are addressed below. First, the *actual* sensor compromising probability α may be *unknown* to an attacker at the time the Byzantine transition probability $\mathbf{P}^{(\alpha)}$ is embedded onto a compromised sensor. It brings up the question regarding what value of α should be used in this situation. We answer the question by establishing the expression for the probability of detection at the FC as a function of α_{B}, which is the value used in the embedded Byzantine transition probability matrices $\mathbf{P}^{(\alpha_{\text{B}})}$ when the knowledge of the true sensor compromising probability α is unavailable. We found that the probability of detection can be expressed as a function of the ratio α/α_{B}. Further numerical examination indicates that the probability of detection decreases as the ratio α/α_{B} increases. This leads to a useful rule of thumb that subject to the achievability of system blindness (i.e., $\alpha_{\text{B}} \geq \alpha^*_{\text{blind}}$), maximizing the global detection error probability generally conforms to minimizing α_{B} subject to $\alpha_{\text{B}} \geq \alpha^*_{\text{blind}}$, and hence an attacker shall set $\alpha_{\text{B}} = \alpha^*_{\text{blind}}$. This policy not only maximizes the global detection error probability according to empirical experiments, but reduces the operational effort of an attacker since no dynamic adjustment of α-value for $\mathbf{P}^{(\alpha)}$ is necessary as long as the true sensor compromising probability α varies between α^*_{blind} and 1.

Secondly, the perfect knowledge of POI θ may be too optimistic for an attacker. In practice, a compromised sensor may perform an initial estimate of the POI, denoted as θ_{B}, possibly obtained through cooperation among a small number of nearby compromised sensors. We will show in our numerical experiments in Section IV-B that a rough estimate of the particular phenomenon θ_{B}, applied to the optimal $\mathbb{P}^{(\theta_{\text{B}},\alpha_{\text{B}})}$ we derive in later sections, is adequate to force the FC to suffer a near-blinding performance. In particular, this design brings an advantage that varying $\mathbb{P}^{(\theta_{\text{B}},\alpha_{\text{B}})}$ according to a rough estimate of $\theta = \theta_{\text{B}}$ at local compromised sensors mitigates possible identification of compromised sensors by examining their statistics by an anti-attack scheme.

[2]This work adopts a more general definition of system blindness than the one used in [17]. Specifically, blindness in [17] meant that the FC receives *equi-probable* outputs from local sensors and hence can only make a blind random guess regarding the POI. In this work, the notion of blindness is generalized, which only requires that the local outputs are made independent of the POI. It is obvious that requiring the FC to receive equiprobable outputs is a sufficient condition for making the local outputs and the POI independent.

It should be noted that the results obtained in this paper are extensions of our previous preliminary study [20], where it was shown that a Byzantine attacker can statistically modify the M-ary local data at compromised sensors such that the receptions at the FC become independent of the POI. How to minimize the sensor compromising probability α among all attack policies (i.e., to determine α^*_{blind}), however, was not solved in [20]. Also, a Byzantine transition probability $\mathbf{P}^{(\alpha)}$ that can blind the system for a given sensor compromising probability $\alpha \geq \alpha^*_{\text{blind}}$ was not derived in [20]. In this paper, both of the above two unsolved issues have been addressed. Our results indicate that a proper choice of the target statistics for M-ary local transmissions can significantly reduce the sensor compromising probability required to blind the system, as compared to the one that requires identical statistics for receptions at the FC as in [20]. We summarize the main contributions of the paper in the following.

- Determination of the exact expression of the minimum sensor compromising probability α^*_{blind} as a function of the statistics of the local quantization outputs (cf. Theorem 2).
- Identification of a blind-achieving Byzantine transition policy $\mathbf{P}^{(\alpha)} := \{\mathbb{P}^{(\theta,\alpha)}\}_{\theta\in\Theta}$ for all $\alpha \geq \alpha^*_{\text{blind}}$ (cf. Theorems 1 and 3).
- Empirical confirmation of the sufficiency of setting the operational $\alpha_{\text{B}} = \alpha^*_{\text{blind}}$ in order to force an error performance worse than a random guess when embedding $\mathbf{P}^{(\alpha)}$ onto a compromised sensor (cf. Section III-C).

The rest of the paper is organized as follows. The system model and problem formulation are introduced in Section II. Determination of the minimum sensor compromising probability α^*_{blind} that makes the POI and the local outputs statistically independent is presented in Section III. Numerical results are given in Section IV and conclusion is drawn in Section V.

II. System Model and Problem Formulation

A. Setup and Assumptions

Consider a WSN that is designed to estimate a particular POI θ as shown in Fig. 1(a). Assume that θ is randomly drawn from a finite set Θ with cardinality $|\Theta| = L$, according to a known probability mass function (pmf). The local sensors acquire conditionally independent and identically distributed (i.i.d.) observations $\boldsymbol{r} = (r_1, r_2, \ldots, r_N)$ given θ, where r_i is the local observation of the ith sensor. We consider a model that each sensor is independently compromised with an identical probability α by a Byzantine attacker, such scenario is also used in [11], [17]. These *Byzantine sensors* transmit falsified data to the FC in order to deteriorate the global decision of the WSN, and the FC is assumed to be unaware of the presence of these Byzantine attacks.

Due to limited local resources such as energy and bandwidth, a local decision rule at the ith sensor, $i \in \mathcal{N}:=\{1,2,\ldots,N\}$, converts r_i to one of the M symbols, denoted as $u_i \in \mathcal{M}:=\{1,2,\ldots,M\}$, before conveying it to the FC. Since the transmitted symbol may be compromised by a Byzantine attacker and hence could be different from u_i, we denote by v_i the symbol that is actually transmitted by

the ith sensor. Accordingly, if sensor i is an *Honest sensor*, we have $v_i = u_i$; otherwise, the ith sensor modifies $u_i = \ell$ to $v_i = m$ with probability $p_{\ell,m}^{(\theta)}$ as depicted in Fig. 1(b). As a result, the so-called Byzantine transition probability can be modeled via a row-stochastic matrix as follows:[3]

$$\mathbb{P}^{(\theta)} := \begin{bmatrix} p_{1,1}^{(\theta)} & p_{1,2}^{(\theta)} & \cdots & p_{1,M}^{(\theta)} \\ p_{2,1}^{(\theta)} & p_{2,2}^{(\theta)} & \cdots & p_{2,M}^{(\theta)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M,1}^{(\theta)} & p_{M,2}^{(\theta)} & \cdots & p_{M,M}^{(\theta)} \end{bmatrix}.$$

In our design, each compromised sensor will have an initial estimate of the POI $\theta_{\rm B}$ and will use $\mathbb{P}^{(\theta_{\rm B})}$ to attack the system. To facilitate the analysis, we make the strong assumption $\theta_{\rm B} = \theta$ in our derivation of blind-achieving $\mathbb{P}^{(\theta)}$, i.e., we assume that the Byzantine attacker knows the exact value of θ. Note that such a strong assumption has also been made in [11], where the authors investigated security problems in distributed detection networks by assuming that the true hypothesis is known to compromised sensors. Although this assumption is too optimistic to be effective in reality, it can be regarded as an ideal benchmark to shoot for from attacker's standpoint. We will examine this assumption in our numerical experiments in Section IV-B and will show that a rough estimate of the particular phenomenon $\theta_{\rm B}$, together with the optimal $\mathbb{P}^{(\theta)}$ we derive in later sections, is adequate to force the FC to suffer a near-blinding performance.

When deriving the $\mathbb{P}^{(\theta)}$ that can blind the FC, a Byzantine attacker is assumed to know the probability mass function (pmf) of u_i, which is denoted as $c_m^{(\theta)} := \Pr(u_i = m|\theta)$. Without loss of generality, we assume

$$\min_{m\in\mathcal{M}} \max_{\theta\in\Theta} c_m^{(\theta)} > 0, \tag{1}$$

since the attacker can exclude the mth row and the mth column from $\mathbb{P}^{(\theta)}$ when $\max_{\theta\in\Theta} c_m^{(\theta)} = 0$ and design a $\mathbb{P}^{(\theta)}$ of smaller size to blind the FC.

Additionally, we denote the transition probability matrix of the discrete noisy link between a sensor and the FC by:

$$\mathbb{Q} := \begin{bmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,M} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ q_{M,1} & q_{M,2} & \cdots & q_{M,M} \end{bmatrix},$$

where $q_{\ell,m}$ is the probability of $v_i = \ell$ being converted to symbol $z_i = m$ during the noisy transmission. It is reasonable to assume that the noisy link is independent of the phenomenon θ and hence $\mathbb{Q}$ remains invariant when the value of θ varies. From elementary probability theory, the vector $\boldsymbol{c}^{(\theta)} := [c_1^{(\theta)}\ c_2^{(\theta)}\ \cdots\ c_M^{(\theta)}]^{\intercal}$ and the two matrices $\mathbb{P}^{(\theta)}$ and $\mathbb{Q}$ must satisfy

$$\begin{cases} \mathbf{1}^{\intercal}\boldsymbol{c}^{(\theta)} = 1 & \text{with } 0 \le c_m^{(\theta)} \le 1 \text{ for } m \in \mathcal{M}; \\ \mathbb{P}^{(\theta)}\mathbf{1} = \mathbf{1} & \text{with } 0 \le p_{\ell,m}^{(\theta)} \le 1 \text{ for } \ell, m \in \mathcal{M}; \\ \mathbb{Q}\,\mathbf{1} = \mathbf{1} & \text{with } 0 \le q_{\ell,m} \le 1 \text{ for } \ell, m \in \mathcal{M}, \end{cases}$$

[3]Note that the optimal Byzantine transition probability that we derive in later sections such as the one in Theorem 3 will also be a function of the sensor compromising probability α. In this section, we use $\mathbb{P}^{(\theta)}$ simply to denote a general design that can be arbitrary, not necessarily dependent on α, but is only a function of the POI θ.

where superscript "$^{\intercal}$" is the matrix transpose operation and $\mathbf{1}$ is the $M \times 1$ all-one column vector.

With the above setting, the conditional probability of receiving z_i from the ith sensor, given that θ is the true phenomenon, can be obtained as follows:

$$\begin{aligned} \Pr(z_i = m\,|\,\theta) &= \sum_{j=1}^{M} q_{j,m} \Pr(v_i = j\,|\,\theta) \\ &= \sum_{j=1}^{M} q_{j,m} \Big(\alpha \Pr(v_i = j\,|\,i = \text{Byzantine}, \theta) \\ &\qquad + (1-\alpha) \Pr(v_i = j\,|\,i = \text{Honest}, \theta) \Big) \\ &= \alpha \sum_{j=1}^{M} q_{j,m} \sum_{\ell=1}^{M} \Pr(v_i = j\,|\,u_i = \ell, \theta) \cdot \Pr(u_i = \ell\,|\,\theta) \\ &\qquad + (1-\alpha) \sum_{j=1}^{M} q_{j,m} \Pr(u_i = j\,|\,\theta) \\ &= \alpha \left(\sum_{j=1}^{M} q_{j,m} \sum_{\ell=1}^{M} p_{\ell,j}^{(\theta)} c_\ell^{(\theta)} - \sum_{j=1}^{M} q_{j,m} c_j^{(\theta)} \right) + \sum_{j=1}^{M} q_{j,m} c_j^{(\theta)}. \end{aligned} \tag{2}$$

Moreover, since each sensor is independently compromised with probability α, and the noisy links are assumed independent, we have

$$\Pr(\boldsymbol{z} = \boldsymbol{m}\,|\,\theta) = \prod_{i=1}^{N} \Pr(z_i = m_i\,|\,\theta), \tag{3}$$

where $\boldsymbol{z} = [z_1\ z_2\ \cdots z_N]^{\intercal}$ and $\boldsymbol{m} = [m_1\ m_2\ \cdots\ m_N]^{\intercal}$.

B. Problem Formulation

A Byzantine attack is targeted to blind the FC with the least amount of effort, i.e., with the minimum sensor compromising probability α, such that the observation $\Pr(\boldsymbol{z} = \boldsymbol{m}\,|\,\theta)$ at the FC and θ become independent, which can be characterized as:

$$\Pr(\boldsymbol{z} = \boldsymbol{m}\,|\,\theta) = \Pr(\boldsymbol{z} = \boldsymbol{m}). \tag{4}$$

From (3), we have the following lemma.

Lemma 1: Eq. (4) holds if, and only if, for every $1 \le i \le N$,

$$\Pr(z_i = m\,|\,\theta) = \Pr(z_i = m) = b_m, \quad \forall\, m \in \mathcal{M}, \tag{5}$$

for some pmf $\boldsymbol{b} = [b_1\ b_2\ \cdots\ b_M]^{\intercal}$. In other words, (4) can be guaranteed by sensor-wise independence between z_i and θ.

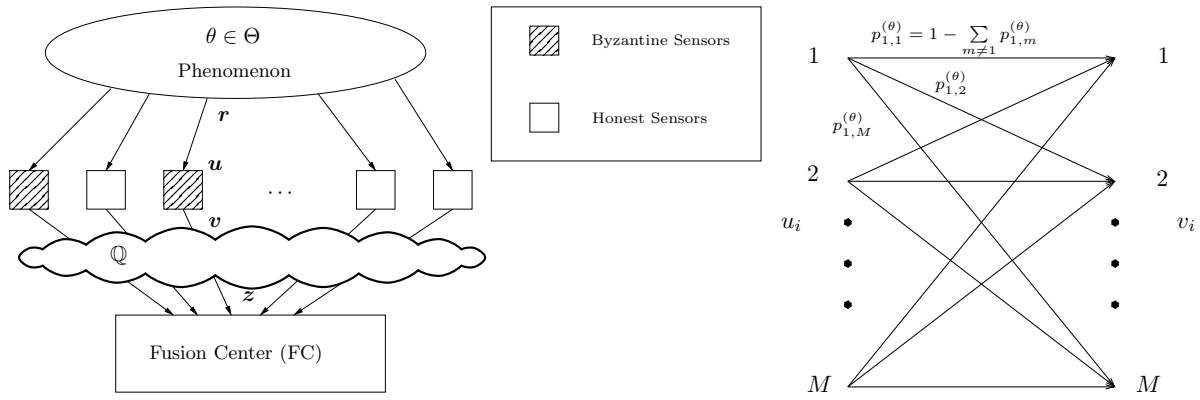
Proof: Eq. (5) implying (4) holds straightforwardly from (3). Conversely, (2) shows that $\Pr(z_i = m\,|\,\theta) = \Pr(z = m\,|\,\theta)$ has nothing to do with the index of the sensor; hence, based on the validity of (4), we can obtain from (3) that:

$$\Pr(\boldsymbol{z} = m\mathbf{1}) = \prod_{i=1}^{N} \Pr(z = m\,|\,\theta) = \left(\Pr(z = m\,|\,\theta)\right)^N.$$

Accordingly, $\Pr(z = m\,|\,\theta)$ is not a function of θ. ■

With the objective of making z_i and θ independent, together with (2), the problem that this paper focuses on is to find the minimum α, subject to

$$\sum_{j=1}^{M} q_{j,m} c_j^{(\theta)} - b_m = \alpha \left(\sum_{j=1}^{M} q_{j,m} c_j^{(\theta)} - \sum_{j=1}^{M} q_{j,m} \sum_{\ell=1}^{M} p_{\ell,j}^{(\theta)} c_\ell^{(\theta)} \right)$$



(a) WSN model (b) Byzantine attack model

Fig. 1: Byzantine attacks for distributed detection network

for all $m \in \mathcal{M}$ and all $\theta \in \Theta$, over all attacker's choices of Byzantine transition probabilities $\mathbf{P} = \{\mathbb{P}^{(\theta)}\}_{\theta\in\Theta}$ and pmf $\boldsymbol{b}$. In matrix form, the above equations can be expressed as:

$$\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)} - \boldsymbol{b} = \alpha\,\mathbb{Q}^{\mathsf{T}}\big(\mathbb{I} - (\mathbb{P}^{(\theta)})^{\mathsf{T}}\big)\boldsymbol{c}^{(\theta)}, \quad \forall\,\theta \in \Theta, \tag{6}$$

where $\mathbb{I}$ is the $M \times M$ identity matrix.

It can be seen from (6) that two statistical modifications $\mathbb{P}^{(\theta)}$ and $\mathbb{Q}$ are applied to the quantization outputs of *Byzantine sensors*. The former is specifically designed to blind the FC, while the latter is due to channel noise. The actual impact of the two is somehow combined, which might complicate the determination of the minimum sensor compromising probability α^*_{blind} among all solutions α for (6). We, however, found that as long as $\mathbb{Q}$ admits an inverse and $\boldsymbol{d} = (\mathbb{Q}^{\mathsf{T}})^{-1}\boldsymbol{b}$ is a pmf whenever $\boldsymbol{b}$ is a pmf, which holds in many channels of practical interest (e.g. modulo-M additive noise channels), the channel noise no longer impacts the determination of α^*_{blind}, and (6) can be equivalently simplified to

$$\boldsymbol{c}^{(\theta)} - \boldsymbol{d} = \alpha\big(\mathbb{I} - (\mathbb{P}^{(\theta)})^{\mathsf{T}}\big)\boldsymbol{c}^{(\theta)}, \quad \forall\,\theta \in \Theta \tag{7}$$

with

$$\Pr(v_i = m\,|\,\theta) = \Pr(v_i = m) = d_m, \quad \forall\, m \in \mathcal{M},$$

and $\boldsymbol{b} = \mathbb{Q}^{\mathsf{T}}\boldsymbol{d} = \mathbb{Q}^{\mathsf{T}}[d_1\ d_2\ \cdots\ d_M]^{\mathsf{T}}$. As a result, every α that satisfies (6) must fulfill (7). This justifies the analysis in the next section, which ignores the impact of $\mathbb{Q}$ and determines α^*_{blind} simply based on (7).

In situations when $\mathbb{Q}$ is not invertible,[4] the set of all solutions α for (7) may become a proper subset of the solution set for (6). In such a case, the introduction of channel noise can help save the Byzantine effort and further reduce α^*_{blind} as shown in the following example.

[4]An operational implication of the invertibility of $\mathbb{Q}$ in our system setting can be seen from the relation of $\boldsymbol{b} = \mathbb{Q}^{\mathsf{T}}\boldsymbol{d}$. When $\mathbb{Q}$ does not admit an inverse, the mapping from $\boldsymbol{d}$ to $\boldsymbol{b}$ becomes surjective and so does the mapping from $\boldsymbol{c}^{(\theta)}$ to $\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)}$. Thus, distinct $\boldsymbol{c}^{(\theta)}$ for two POIs may result in the same output statistics after passing through the wireless link; in such a case, the FC can no longer distinguish the two POIs and hence becomes blind even without Byzantine attack. This is precisely what has been encountered in Example 1.

Example 1: Suppose $\mathbb{Q} = \boldsymbol{x}\boldsymbol{w}^{\mathsf{T}}$ is a rank-one matrix. Then, for $\theta \in \Theta$,

$$\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)} = \boldsymbol{w}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{c}^{(\theta)} = \boldsymbol{w}\frac{\mathbf{1}^{\mathsf{T}}}{(\boldsymbol{w}^{\mathsf{T}}\mathbf{1})}\boldsymbol{c}^{(\theta)} = \boldsymbol{w}\frac{(\mathbf{1}^{\mathsf{T}}\boldsymbol{c}^{(\theta)})}{(\boldsymbol{w}^{\mathsf{T}}\mathbf{1})} = \frac{\boldsymbol{w}}{(\boldsymbol{w}^{\mathsf{T}}\mathbf{1})},$$

where the second equality follows from $\mathbf{1} = \mathbb{Q}\mathbf{1} = \boldsymbol{x}(\boldsymbol{w}^{\mathsf{T}}\mathbf{1})$. Hence, choosing $\boldsymbol{b} = \frac{\boldsymbol{w}}{(\boldsymbol{w}^{\mathsf{T}}\mathbf{1})}$ immediately gives that the minimum α that satisfies (6) is zero, even if the minimum α that fulfills (7) is strictly positive (cf. Theorem 2). □

For convenience, we denote by $\alpha(\boldsymbol{d}, \mathbf{P})$ all the solutions of α satisfying (7) for given $\boldsymbol{d}$ and $\mathbf{P}$. Then, the global minimum α^*_{blind} can be approached via the following two minimization problems:

$$\text{either} \quad \min_{\boldsymbol{d}\in\mathcal{D}} \min_{\mathbf{P}\in\mathcal{P}} \alpha(\boldsymbol{d}, \mathbf{P}) \quad \text{or} \quad \min_{\mathbf{P}\in\mathcal{P}} \min_{\boldsymbol{d}\in\mathcal{D}} \alpha(\boldsymbol{d}, \mathbf{P}),$$

where $\mathcal{P}$ is the set that is exhaustive over all legitimate $\mathbf{P}$, and $\mathcal{D}$ consists of all $M \times 1$ pmf vectors. We choose the former approach due to analytical convenience.[5] Specifically, given any target distribution $\boldsymbol{d}$, the smallest

$$\alpha_{\text{blind}}(\boldsymbol{d}) := \min_{\mathbf{P}\in\mathcal{P}} \alpha(\boldsymbol{d}, \mathbf{P})$$

and its corresponding minimizer $\mathbf{P}^*(\boldsymbol{d})$ are first determined. We then determine

$$\alpha^*_{\text{blind}} = \min_{\boldsymbol{d}\in\mathcal{D}} \alpha_{\text{blind}}(\boldsymbol{d})$$

as well as the minimizers $\boldsymbol{d}^*$ and $\mathbf{P}^* := \mathbf{P}^*(\boldsymbol{d}^*)$ that can achieve α^*_{blind}.

[5]It can be noted from Theorem 1 that $\alpha_{\text{blind}}(\boldsymbol{d})$ always exists for any given $\boldsymbol{d}$. However, the solution set for $\alpha_{\text{blind}}(\mathbf{P}) := \min_{\boldsymbol{d}\in\mathcal{D}} \alpha(\boldsymbol{d}, \mathbf{P})$ can be empty for an improperly chosen $\mathbf{P}$. For example, if $\mathbb{P}^{(\theta)} = \mathbb{R}$ for all $\theta \in \Theta$ and $\boldsymbol{c}^{(\theta)}$ distinct for different θ, then (7) requires

$$\boldsymbol{c}^{(\theta)} = \big((1-\alpha)\mathbb{I} + \alpha\mathbb{R}^{\mathsf{T}}\big)^{-1}\boldsymbol{d} \quad \forall\,\theta \in \Theta$$

which cannot be equated when $(1-\alpha)\mathbb{I} + \alpha\mathbb{R}^{\mathsf{T}}$ admits an inverse for every $\alpha \in [0, 1]$. Thus, to approach α^*_{blind} via the minimization $\min_{\mathbf{P}\in\mathcal{P}} \min_{\boldsymbol{d}\in\mathcal{D}} \alpha(\boldsymbol{d}, \mathbf{P})$ needs to specify those $\mathbf{P}$ such that $\alpha_{\text{blind}}(\mathbf{P})$ exists, which may be a challenging task.

C. System Implication of the Attack Measure: Sensor Compromising Probability

In this subsection, we clarify the system implication of two analogous but distinct attack measures, i.e., the minimum compromising probability for an individual sensor and the minimum ratio of compromised sensors to all the sensors.

Similar to [17], each sensor being independently compromised with an identical probability α (referred to as α-setting in the sequel) is considered in our mathematical formulation. This is in contrast to the notion of compromising exactly β fraction of sensors (referred to as β-setting in the sequel). Although the two settings may be indistinguishable as the number of sensors goes to infinity, the difference between them is not negligible when the size of the sensor network is finite. As in [17], here we employ the α-setting rather than the β-setting.

As a result of the adoption of the α-setting, the mathematical problem to be solved has a single-sensor formulation. Specifically, under the assumption that each sensor is independently compromised with an identical probability α, a Byzantine transition probability $\mathbf{P}^{(\alpha)}$ that statistically alters the local outputs of compromised sensors can be devised by an attacker. The minimum sensor compromising probability α^*_{blind} required by an attacker to blind the FC, as well as the corresponding blind-achieving Byzantine transition probability $\mathbf{P}^{(\alpha^*_{\text{blind}})}$, are accordingly obtained.

Our simulation results confirm that when we set $\alpha = \beta$, the resulting average global detection error probabilities of the two settings for a system with only ten sensors are actually close to each other (cf. Figs. 3 and 4). This indicates that the parameters that respectively govern the two settings (i.e., α and β) not only have the same asymptotical impact on their respective global detection error probabilities but exhibit little difference for a system of finite size.

III. A Blind-Achieving Byzantine Attack

In this section, $\alpha_{\text{blind}}(\boldsymbol{d}) = \min_{\mathbf{P}\in\mathcal{P}} \alpha(\boldsymbol{d}, \mathbf{P})$ will be established in Subsection III-A. What follows is the derivation of $\alpha^*_{\text{blind}} = \min_{\boldsymbol{d}\in\mathcal{D}} \alpha_{\text{blind}}(\boldsymbol{d})$ in Subsection III-B. Finally, an attack policy that targets the maximization of the probability of misdetection will be given in Subsection III-C.

A. Minimum Sensor Compromising Probability for given Local Output Statistics

Before presenting the derivation of $\alpha_{\text{blind}}(\boldsymbol{d})$, some preliminary analysis is necessary. For a given pmf $\boldsymbol{d}$, we have that

$$\sum_{m=1}^{M} (c_m^{(\theta)} - d_m) = \sum_{m=1}^{M} c_m^{(\theta)} - \sum_{m=1}^{M} d_m = 1 - 1 = 0$$

for every $\theta \in \Theta$. As a result, $\max_{m\in\mathcal{M}}\{c_m^{(\theta)} - d_m\} \geq 0$, $\forall\, \theta \in \Theta$. Corresponding to each $\theta \in \Theta$, we divide the index set $\mathcal{M} = \{1, 2, \ldots, M\}$ into two groups. The first group $\mathcal{M}_1^{(\theta)}$ contains those indices m that satisfy $\max\{c_m^{(\theta)} - d_m, 0\} = 0$ (equivalently, $c_m^{(\theta)} \leq d_m$), while the remaining indices belong to the second group $\mathcal{M}_2^{(\theta)}$.

It can be verified that $\max_{m\in\mathcal{M}}\{c_m^{(\theta)} - d_m\} > 0$ implies $1 \leq |\mathcal{M}_1^{(\theta)}| < M$, ensuring that none of the two groups are empty.[6] If $\max_{m\in\mathcal{M}}\{c_m^{(\theta)} - d_m\} = 0$, we have $\mathcal{M}_1^{(\theta)} = \mathcal{M}$ and hence $\mathcal{M}_2^{(\theta)}$ is empty.

In the case that $\mathcal{M}_2^{(\theta)} \neq \emptyset$, we define

$$e_m^{(\theta)} := 1 - \frac{d_m}{c_m^{(\theta)}}, \quad m \in \mathcal{M}_2^{(\theta)}. \tag{8}$$

Note that for those m's in a non-empty $\mathcal{M}_2^{(\theta)}$, $\max\{c_m^{(\theta)} - d_m, 0\} > 0$ implies the denominator $c_m^{(\theta)} > 0$ and $0 \leq d_m < c_m^{(\theta)}$, and hence $e_m^{(\theta)}$ is well defined and positive.

Theorem 1: If a Byzantine attacker learns the conditional pmf of the local output $\boldsymbol{c}^{(\theta)}$ for every $\theta \in \Theta$, then for given $\boldsymbol{d} \in \mathcal{D}$,

$$\alpha_{\text{blind}}(\boldsymbol{d}) = \max_{\theta\in\Theta} \max_{m\in\mathcal{M}:\, c_m^{(\theta)}>0} \left\{1 - \frac{d_m}{c_m^{(\theta)}}\right\}. \tag{9}$$

Furthermore, (9) can be achieved by $\mathbf{P}^* = \{\mathbb{P}^{(\theta)*}\}_{\theta\in\Theta}$ with its elements defined as follows:

$$p_{\ell,j}^{(\theta)*} = \begin{cases} 1, & j = \ell \in \mathcal{M}_1^{(\theta)}; \\ 1 - \dfrac{e_\ell^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})}, & j = \ell \in \mathcal{M}_2^{(\theta)}; \\ \dfrac{(d_j - c_j^{(\theta)})}{\sum\limits_{m\in\mathcal{M}_1^{(\theta)}} (d_m - c_m^{(\theta)})} \left(\dfrac{e_\ell^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})}\right), & \\ & j \in \mathcal{M}_1^{(\theta)} \ \&\ \ell \in \mathcal{M}_2^{(\theta)}; \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

when $\mathcal{M}_2^{(\theta)}$ is non-empty, and

$$\mathbb{P}^{(\theta)*} = \mathbb{I}$$

when $\mathcal{M}_2^{(\theta)}$ is empty.

Proof: We defer the proof to Appendix A for better readability. ∎

Note that when $\mathcal{M}_2^{(\theta)}$ is non-empty, we have $\sum_{m\in\mathcal{M}_1^{(\theta)}}(d_m - c_m^{(\theta)}) = \sum_{m\in\mathcal{M}_2^{(\theta)}}(c_m^{(\theta)} - d_m) > 0$ and the denominator $\alpha_{\text{blind}}(\boldsymbol{d}) \geq e_{k(\theta)}^{(\theta)} > 0$, where

$$k(\theta) := \arg \max_{m\in\mathcal{M}_2^{(\theta)}} e_m^{(\theta)}. \tag{11}$$

Hence, the $\mathbb{P}^{(\theta)*}$ in (10) is well defined.

Two remarks are made based on the above theorem. First, the pmf $\boldsymbol{d}$ in (7) can be regarded as the target distribution that a Byzantine attacker intends to maliciously change $\boldsymbol{c}^{(\theta)}$ to. Thus, for those m's in $\mathcal{M}_1^{(\theta)}$, the value of $c_m^{(\theta)}$ should be increased, while $c_m^{(\theta)}$ must be scaled down for $m \in \mathcal{M}_2^{(\theta)}$.

[6] $\mathcal{M}_1^{(\theta)} = \mathcal{M}$ implies $c_m^{(\theta)} \leq d_m$ for all $m \in \mathcal{M}$; and hence, $\max_{m\in\mathcal{M}}\{c_m^{(\theta)} - d_m\} = 0$. By this, we can equivalently infer that $\max_{m\in\mathcal{M}}\{c_m^{(\theta)} - d_m\} > 0$ implies $|\mathcal{M}_1^{(\theta)}| < M$.

Second, the minimizer $\mathbb{P}^{(\theta)*}$ is not unique. We can relax the assignment of $p_{\ell,j}^{(\theta)*}$ in (10) for $j \in \mathcal{M}_1^{(\theta)}$ and $\ell \in \mathcal{M}_2^{(\theta)}$ to any values satisfying

$$\sum_{\ell \in \mathcal{M}_2^{(\theta)}} p_{\ell,j}^{(\theta)*} c_\ell^{(\theta)} = \frac{d_j - c_j^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} \quad \text{for } j \in \mathcal{M}_1^{(\theta)} \tag{12}$$

and

$$\sum_{j \in \mathcal{M}_1^{(\theta)}} p_{\ell,j}^{(\theta)*} = \frac{e_\ell^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} \quad \text{for } \ell \in \mathcal{M}_2^{(\theta)}. \tag{13}$$

This gives extra freedom to an attacker in the choice of the blind-achieving Byzantine transition probability.

A design of the blind-achieving Byzantine transition probability $\mathbf{P}^*$ is illustrated in the following example.

Example 2: Consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$, and assume that the local data is quantized into 4-ary symbols. Let the statistics of local quantization outputs given each hypothesis be given by $\boldsymbol{c}^{(\theta_1)} = \left[\frac{1}{10}, \frac{1}{5}, \frac{3}{10}, \frac{2}{5}\right]^\intercal$, $\boldsymbol{c}^{(\theta_2)} = \left[\frac{3}{5}, \frac{1}{5}, \frac{1}{10}, \frac{1}{10}\right]^\intercal$, and $\boldsymbol{c}^{(\theta_3)} = \left[\frac{1}{10}, \frac{1}{5}, \frac{1}{5}, \frac{1}{2}\right]^\intercal$.

Then, setting the target distribution $\boldsymbol{d}$ to be the uniform distribution, i.e.,

$$\boldsymbol{d} = \boldsymbol{d}_{\text{uni}} := \tfrac{1}{M}\mathbf{1} = \left[\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}\right]^\intercal \text{ with } M = 4,$$

we obtain $\mathcal{M}_1^{(\theta_1)} = \{1,2\}$, $\mathcal{M}_2^{(\theta_1)} = \{3,4\}$, $\boldsymbol{e}^{(\theta_1)} = \left[-, -, \frac{1}{6}, \frac{3}{8}\right]^\intercal$; $\mathcal{M}_1^{(\theta_2)} = \{2,3,4\}$, $\mathcal{M}_2^{(\theta_2)} = \{1\}$, $\boldsymbol{e}^{(\theta_2)} = \left[\frac{7}{12}, -, -, -\right]^\intercal$; and $\mathcal{M}_1^{(\theta_3)} = \{1,2,3\}$, $\mathcal{M}_2^{(\theta_3)} = \{4\}$, $\boldsymbol{e}^{(\theta_3)} = \left[-, -, -, \frac{1}{2}\right]^\intercal$. Theorem 1 implies that

$$\begin{aligned}\alpha_{\text{blind}}(\boldsymbol{d}_{\text{uni}}) &= \max_{\theta \in \Theta} \max_{m \in \mathcal{M}:\, c_m^{(\theta)} > 0} \left\{1 - \frac{d_m}{c_m^{(\theta)}}\right\} \\ &= \max_{\theta \in \Theta} \max_{m \in \mathcal{M}_2^{(\theta)}} e_m^{(\theta)} = \frac{7}{12}.\end{aligned}$$

A set of minimizers that achieve $\alpha_{\text{blind}}(\boldsymbol{d}_{\text{uni}})$, according to Theorem 1, is respectively given by

$$\mathbb{P}^{(\theta_1)*} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{3}{14} & \frac{1}{14} & \frac{5}{7} & 0 \\ \frac{27}{56} & \frac{9}{56} & 0 & \frac{5}{14} \end{bmatrix}, \tag{14}$$

$$\mathbb{P}^{(\theta_2)*} = \begin{bmatrix} 0 & \frac{1}{7} & \frac{3}{7} & \frac{3}{7} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbb{P}^{(\theta_3)*} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{18}{35} & \frac{6}{35} & \frac{6}{35} & \frac{1}{7} \end{bmatrix}.$$

As we have remarked previously, the minimizers that achieve $\alpha_{\text{blind}}(\boldsymbol{d}_{\text{uni}})$ are not unique. For example, (12) and (13) jointly imply that (14) can be substituted by

$$\mathbb{P}^{(\theta_1)*} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{5}{7} & 0 \\ \frac{15}{28} & \frac{3}{28} & 0 & \frac{5}{14} \end{bmatrix}. \tag{15}$$

Since both (14) and (15) satisfy (7) with identical α, it can be inferred similarly from the later analysis in (19) and (20) in Subsection III-C that both result in the same global detection error probability, and therefore, an attacker can simply select one of the two to embed onto those compromised sensors. □

B. Minimum Sensor Compromising Probability for System Blindness

In this subsection, we determine the minimizer $\boldsymbol{d}^* \in \mathcal{D}$ that achieves α_{blind}^*.

Theorem 2: If the Byzantine attacker has the knowledge of the conditional pmf of the local output $\boldsymbol{c}^{(\theta)}$ for all $\theta \in \Theta$, then

$$\alpha_{\text{blind}}^* = 1 - \frac{1}{\sum_{m=1}^M \max_{\theta \in \Theta} c_m^{(\theta)}}. \tag{16}$$

Furthermore, (16) can be achieved by $\boldsymbol{d}^*$ with its components defined as:

$$d_m^* := \frac{\max_{\theta \in \Theta} c_m^{(\theta)}}{\sum_{\ell=1}^M \max_{\theta \in \Theta} c_\ell^{(\theta)}} \quad \text{for } m \in \mathcal{M}.$$

Proof: See Appendix B. ■

The exact expression of α_{blind}^* in Theorem 2 indicates that the minimum Byzantine effort grows as $\sum_{m=1}^M \max_{\theta \in \Theta} c_m^{(\theta)}$ increases. It can be derived that under equal prior probability on $\theta \in \Theta$,

$$\begin{aligned}\sum_{m=1}^M \max_{\theta \in \Theta} c_m^{(\theta)} &= \sum_{m=1}^M \max_{\theta \in \Theta} \Pr(u_i = m | \theta) \\ &= \sum_{m=1}^M \frac{\Pr(u_i = m)}{\Pr(\theta)} \max_{\theta \in \Theta} \Pr(u_i = m | \theta) \\ &= \sum_{m=1}^M L \cdot \Pr(u_i = m) \max_{\theta \in \Theta} \Pr(\theta | u_i = m) \\ &= L \cdot P_{\text{c-single}}\,,\end{aligned}$$

and

$$P_{\text{c-single}} := \sum_{m=1}^M \Pr(u_i = m) \max_{\theta \in \Theta} \Pr(\theta | u_i = m)$$

is the probability of correct detection of θ based on one single observation u_i. This leads to $\alpha_{\text{blind}}^* = 1 - \frac{1}{L \cdot P_{\text{c-single}}}$. Consequently, if the local output u_i knows more about θ (in the sense of a higher $P_{\text{c-single}}$), the attacker requires more effort to blind the system. On the other hand, if the cardinality of the POI (i.e., L) is larger, a higher Byzantine effort for system blindness is required. In addition, it is interesting to note that the minimizer d_m^* can be shown to be proportional to $\max_{\theta \in \Theta} c_\ell^{(\theta)}$, i.e.,

$$\frac{d_m^*}{\max_{\theta \in \Theta} c_m^{(\theta)}} = 1 - \alpha_{\text{blind}}^* = \frac{1}{L \cdot P_{\text{c-single}}}$$

for every $m \in \mathcal{M}$.

From Theorem 2, we can also obtain that the minimum blind-achieving sensor compromising probability α_{blind}^* always satisfies

$$0 \le \alpha_{\text{blind}}^* \le 1 - \frac{1}{M}. \tag{17}$$

Note that $1 - \frac{1}{M}$ is the minimum blind-achieving sensor compromising probability obtained in [17, Thm. 1]. Our result, however, shows that α_{blind}^* is always no larger than $1 - \frac{1}{M}$. This is an anticipated result since extra knowledge of the statistics

of local decisions should help reduce the Byzantine effort. In fact, we can refine the upper bound in (17) down to

$$0 \le \alpha^*_{\text{blind}} \le 1 - \frac{1}{\min\{M, L\}}$$

by noting that

$$\sum_{m=1}^{M} \max_{\theta\in\Theta} c_m^{(\theta)} = \sum_{\theta\in\Theta} \sum_{m\in\mathcal{A}^{(\theta)}} c_m^{(\theta)} \le \sum_{\theta\in\Theta} 1 = L,$$

where $\{\mathcal{A}^{(\theta)}\}_{\theta\in\Theta}$ are disjoint partitions of $\mathcal{M}$ such that $c_m^{(\theta)} = \max_{\theta'\in\Theta} c_m^{(\theta')}$ for $m \in \mathcal{A}^{(\theta)}$.

In the special case of binary hypothesis testing, where $\Theta = \{\theta_1, \theta_2\}$, the determination of the minimum Byzantine effort that can blind the FC is equivalent to finding the smallest *blinding power* such that the resultant Kullback-Leibler divergence between the two conditional pmfs $\Pr(v_i = m|\theta = \theta_1)$ and $\Pr(v_i = m|\theta = \theta_2)$ is zero as considered in [11]. In such a case, the α^*_{blind} that we establish in Theorem 2 can be written as

$$\begin{aligned}\alpha^*_{\text{blind}} &= 1 - \frac{1}{\sum_{m\in\mathcal{M}^{(\theta_1)}} c_m^{(\theta_1)} + \sum_{m\in\mathcal{M}^{(\theta_2)}} c_m^{(\theta_2)}} \\ &= 1 - \frac{1}{1 + \sum_{m\in\mathcal{M}^{(\theta_1)}} [c_m^{(\theta_1)} - c_m^{(\theta_2)}]} \\ &= \frac{\sum_{m\in\mathcal{M}^{(\theta_1)}} \left[c_m^{(\theta_1)} - c_m^{(\theta_2)}\right]}{1 + \sum_{m\in\mathcal{M}^{(\theta_1)}} \left[c_m^{(\theta_1)} - c_m^{(\theta_2)}\right]},\end{aligned}$$

which points to the same result as in [11, Thm. 1].

We next continue Example 2 for the illustration of $\boldsymbol{d}^*$.

Example 3: It is obvious that the setting in Example 2 gives $\left[\max_{\theta\in\Theta} c_1^{(\theta)}, \max_{\theta\in\Theta} c_2^{(\theta)}, \max_{\theta\in\Theta} c_3^{(\theta)}, \max_{\theta\in\Theta} c_4^{(\theta)}\right]^{\intercal} = \left[\frac{3}{5}, \frac{1}{5}, \frac{3}{10}, \frac{1}{2}\right]^{\intercal}$ and

$$\boldsymbol{d}^* = \begin{bmatrix} (\frac{3}{5})/(\frac{3}{5}+\frac{1}{5}+\frac{3}{10}+\frac{1}{2}) \\ (\frac{1}{5})/(\frac{3}{5}+\frac{1}{5}+\frac{3}{10}+\frac{1}{2}) \\ (\frac{3}{10})/(\frac{3}{5}+\frac{1}{5}+\frac{3}{10}+\frac{1}{2}) \\ (\frac{1}{2})/(\frac{3}{5}+\frac{1}{5}+\frac{3}{10}+\frac{1}{2}) \end{bmatrix} = \begin{bmatrix} \frac{3}{8} \\ \frac{1}{8} \\ \frac{3}{16} \\ \frac{5}{16} \end{bmatrix}.$$

Hence, we have

$$\begin{aligned}\alpha^*_{\text{blind}} &= \alpha_{\text{blind}}(\boldsymbol{d}^*) \\ &= \max_{\theta\in\Theta} \max_{m\in\mathcal{M}:\, c_m^{(\theta)}>0} \left\{1 - \frac{d^*_m}{c_m^{(\theta)}}\right\} = 1 - \frac{\frac{3}{8}}{\frac{3}{5}} = \frac{3}{8}.\end{aligned}$$

In comparison with $\alpha_{\text{blind}}(\boldsymbol{d}_{\text{uni}}) = \frac{7}{12}$, we conclude (conveniently in terms of the β-setting perspective with $\beta = \alpha$) that an attacker needs to compromise only 9 out of 24 sensors on average for system blindness, instead of 14 out of 24, if a better $\boldsymbol{d} = \boldsymbol{d}^*$ rather than $\boldsymbol{d} = \boldsymbol{d}_{\text{uni}}$ is employed. □

C. Byzantine Transition Probability and its Resulting Global Detection Error Probability

In the previous two subsections, the minimum Byzantine effort α^*_{blind} required for system blindness was determined. It implied that it is impossible to blind the system if $0 \le \alpha < \alpha^*_{\text{blind}}$. Furthermore, when $\alpha = \alpha^*_{\text{blind}}$, a blind-achieving Byzantine transition probability $\mathbf{P}^* = \{\mathbb{P}^{(\theta)*}\}_{\theta\in\Theta}$ was given in Theorem 1. However, it should be pointed out that the $\mathbf{P}^*$ in Theorem 1 cannot make the local quantization output v_i and θ statistically independent for $\alpha > \alpha^*_{\text{blind}}$. An extension of $\mathbf{P}^*$ that can blind the system for a given $\alpha > \alpha^*_{\text{blind}}$ is provided in the next theorem.

Theorem 3: For a fixed target distribution of local outputs $\boldsymbol{d}$ and also for an α satisfying $\alpha_{\text{blind}}(\boldsymbol{d}) \le \alpha \le 1$, the pair of attack parameters $(\alpha, \mathbf{P}^{(\alpha)}) = (\alpha, \{\mathbb{P}^{(\theta,\alpha)}\}_{\theta\in\Theta})$ satisfy (7), where the elements of $M \times M$ matrix $\mathbb{P}^{(\theta,\alpha)}$ are given by

$$p^{(\theta,\alpha)}_{\ell,j} = \begin{cases} 1, & j = \ell \in \mathcal{M}_1^{(\theta)}; \\ 1 - \dfrac{e_\ell^{(\theta)}}{\alpha}, & j = \ell \in \mathcal{M}_2^{(\theta)}; \\ \dfrac{(d_j - c_j^{(\theta)})}{\sum\limits_{m\in\mathcal{M}_1^{(\theta)}} (d_m - c_m^{(\theta)})} \left(\dfrac{e_\ell^{(\theta)}}{\alpha}\right), & \\ & j \in \mathcal{M}_1^{(\theta)} \ \&\ \ell \in \mathcal{M}_2^{(\theta)}; \\ 0, & \text{otherwise}, \end{cases} \tag{18}$$

when $\mathcal{M}_2^{(\theta)}$ is non-empty, and $\mathbb{P}^{(\theta,\alpha)} = \mathbb{I}$ when $\mathcal{M}_2^{(\theta)}$ is empty.

Proof: By taking $(\alpha, \mathbf{P}^{(\alpha)})$ into (7), the theorem can be proved via a similar procedure as in Step 2 of the proof of Theorem 1. ■

The above theorem shows that the attacker shall adjust the Byzantine transition probability matrix $\mathbf{P}^{(\alpha)}$ (from u_i to v_i) according to the *true* sensor compromising probability α and manipulate it to lie within the range $\alpha_{\text{blind}}(\boldsymbol{d}) < \alpha \le 1$. In particular, from the β-setting equivalence, the true sensor compromising probability should match the ratio of the number of compromised sensors and the total number of sensors. This adjustment, however, might not be a feasible attack option for certain scenarios such as when the true sensor compromising probability varies in time or is even unknown. In these situations, the *designed* sensor compromising probability, denoted as α_{B}, could be different from the *true* sensor compromising probability α. As a result, how to select the designed parameter α_{B} at the time a Byzantine transition probability $\mathbf{P}^{(\alpha_{\text{B}})}$ is embedded onto a compromised sensor becomes an essential concern from attackers' standpoint. With the new objective in mind, we derive below the probability of successful detection at the FC due to *managed* Byzantine transition probability $\mathbf{P}^{(\alpha_{\text{B}})}$ and *true* sensor compromising probability α for a target distribution of local outputs $\boldsymbol{d}$.

For notational convenience, denote the corresponding $\Pr(v_i = m|\theta)$ and $\Pr(z_i = m|\theta)$ as $a_m^{(\theta,\alpha_{\text{B}},\alpha)}$ and $o_m^{(\theta,\alpha_{\text{B}},\alpha)}$, respectively. Define

$$\boldsymbol{a}^{(\theta,\alpha_{\text{B}},\alpha)} := \left[a_1^{(\theta,\alpha_{\text{B}},\alpha)} \ a_2^{(\theta,\alpha_{\text{B}},\alpha)} \ \cdots \ a_M^{(\theta,\alpha_{\text{B}},\alpha)}\right]^{\intercal}$$

and

$$\boldsymbol{o}^{(\theta,\alpha_{\text{B}},\alpha)} := \left[o_1^{(\theta,\alpha_{\text{B}},\alpha)} \ o_2^{(\theta,\alpha_{\text{B}},\alpha)} \ \cdots \ o_M^{(\theta,\alpha_{\text{B}},\alpha)}\right]^{\intercal}.$$

Then, we obtain from Theorem 3 that the Byzantine transition probability $\mathbf{P}^{(\alpha_{\rm B})}$ satisfies $\alpha_{\rm B}\big(\mathbb{I}-(\mathbb{P}^{(\theta,\alpha_{\rm B})})^{\mathsf{T}}\big)\boldsymbol{c}^{(\theta)}=\boldsymbol{c}^{(\theta)}-\boldsymbol{d}$, $\forall\,\theta\in\Theta$, which implies that for all $\theta\in\Theta$,

$$\begin{aligned}\boldsymbol{a}^{(\theta,\alpha_{\rm B},\alpha)}&=\Big[(1-\alpha)\mathbb{I}+\alpha\big(\mathbb{P}^{(\theta,\alpha_{\rm B})}\big)^{\mathsf{T}}\Big]\boldsymbol{c}^{(\theta)}\\&=\boldsymbol{c}^{(\theta)}-\alpha\big(\mathbb{I}-\big(\mathbb{P}^{(\theta,\alpha_{\rm B})}\big)^{\mathsf{T}}\big)\boldsymbol{c}^{(\theta)}\\&=\left(1-\frac{\alpha}{\alpha_{\rm B}}\right)\boldsymbol{c}^{(\theta)}+\frac{\alpha}{\alpha_{\rm B}}\boldsymbol{d}\\&=\boldsymbol{c}^{(\theta)}+\frac{\alpha}{\alpha_{\rm B}}\Big(\boldsymbol{d}-\boldsymbol{c}^{(\theta)}\Big),\end{aligned}\tag{19}$$

and

$$\begin{aligned}\boldsymbol{o}^{(\theta,\alpha_{\rm B},\alpha)}&=\mathbb{Q}^{\mathsf{T}}\boldsymbol{a}^{(\theta,\alpha_{\rm B},\alpha)}\\&=\left(1-\frac{\alpha}{\alpha_{\rm B}}\right)\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)}+\frac{\alpha}{\alpha_{\rm B}}\mathbb{Q}^{\mathsf{T}}\boldsymbol{d}\\&=\mathbb{Q}^{\mathsf{T}}\left[\boldsymbol{c}^{(\theta)}+\frac{\alpha}{\alpha_{\rm B}}\Big(\boldsymbol{d}-\boldsymbol{c}^{(\theta)}\Big)\right].\end{aligned}\tag{20}$$

Unaware of the Byzantine attack, the FC makes its global decision according to $\boldsymbol{o}^{(\theta,\alpha_{\rm B},0)}=\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)}$ (i.e., $\alpha=0$). Hence, under equal prior probability on the POI, the global decision rule at the FC can be written as

$$\hat{\theta}(\boldsymbol{z})=\arg\max_{\theta\in\Theta}\prod_{i=1}^{N}o_{z_i}^{(\theta,\alpha_{\rm B},0)}=\arg\max_{\theta\in\Theta}\prod_{i=1}^{N}\sum_{\ell=1}^{M}q_{\ell,z_i}c_{\ell}^{(\theta)}.$$

The probability of detection at the FC when θ is uniformly distributed is thus given by

$$P_{\rm c}=\frac{1}{L}\sum_{\theta\in\Theta}\sum_{\boldsymbol{z}\in\mathcal{O}^{(\theta)}}\prod_{i=1}^{N}o_{z_i}^{(\theta,\alpha_{\rm B},\alpha)}\tag{21}$$

$$=\frac{1}{L}\sum_{\theta\in\Theta}\sum_{\boldsymbol{z}\in\mathcal{O}^{(\theta)}}\prod_{i=1}^{N}\left(\sum_{\ell=1}^{M}q_{\ell,z_i}\left[c_{\ell}^{(\theta)}+\frac{\alpha}{\alpha_{\rm B}}\Big(d_{\ell}-c_{\ell}^{(\theta)}\Big)\right]\right),\tag{22}$$

where $\{\mathcal{O}^{(\theta)}\}_{\theta\in\Theta}$ are disjoint partitions on $\mathcal{M}^N=\mathcal{M}\times\cdots\times\mathcal{M}$ such that

$$\boldsymbol{z}\in\mathcal{O}^{(\theta)}\Rightarrow\prod_{i=1}^{N}\left(\sum_{\ell=1}^{M}q_{\ell,z_i}c_{\ell}^{(\theta)}\right)\geq\max_{\theta'\in\Theta}\prod_{i=1}^{N}\left(\sum_{\ell=1}^{M}q_{\ell,z_i}c_{\ell}^{(\theta')}\right).\tag{23}$$

The above expression shows that the probability of detection $P_{\rm c}=P_{\rm c}(\alpha/\alpha_{\rm B})$ depends only on the ratio of $\alpha/\alpha_{\rm B}$ for a given pmf $\boldsymbol{d}$. As expected, when $\alpha_{\rm B}=\alpha$, the expression is reduced to a random guess since

$$\begin{aligned}P_{\rm c}(1)&=\frac{1}{L}\sum_{\theta\in\Theta}\sum_{\boldsymbol{z}\in\mathcal{O}^{(\theta)}}\prod_{i=1}^{N}\left(\sum_{\ell=1}^{M}q_{\ell,z_i}d_{\ell}\right)\\&=\frac{1}{L}\sum_{\boldsymbol{z}\in\mathcal{M}^N}\prod_{i=1}^{N}\left(\sum_{\ell=1}^{M}q_{\ell,z_i}d_{\ell}\right)=\frac{1}{L},\end{aligned}\tag{24}$$

where (24) holds because $\mathbf{1}^{\mathsf{T}}\big(\mathbb{Q}^{\mathsf{T}}\boldsymbol{d}\big)=\big(\mathbb{Q}\mathbf{1}\big)^{\mathsf{T}}\boldsymbol{d}=\mathbf{1}^{\mathsf{T}}\boldsymbol{d}=1$.

Our numerical experiments under $\boldsymbol{d}=\boldsymbol{d}^*$ indicate that the larger the ratio is, the larger the global detection error probability is. Thus, without sacrificing the possibility of blinding the FC, a capable attacker would lower $\alpha_{\rm B}$ whenever possible. This justifies the choice of fixing $\alpha_{\rm B}=\alpha^*_{\rm blind}$, which is the minimum value attainable for $\mathbf{P}^{(\alpha_{\rm B})}$ that can either blind the FC (if $\alpha=\alpha^*_{\rm blind}$), or force an error performance worse than a random guess (if $\alpha^*_{\rm blind}<\alpha\leq 1$).

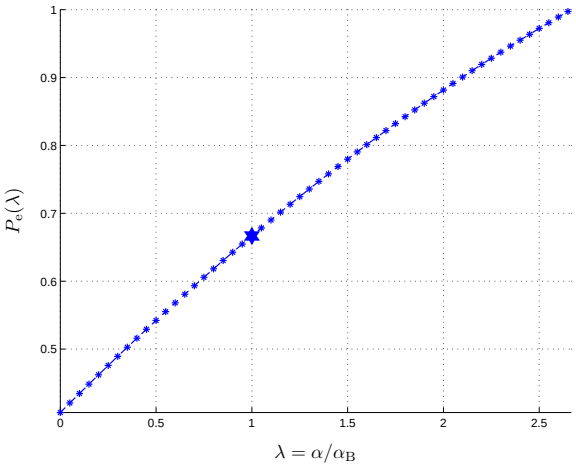


Fig. 2: Global detection error probability $P_{\rm e}(\lambda)$ as a function of $\lambda{:=}\alpha/\alpha_{\rm B}$ under $\boldsymbol{d}=\boldsymbol{d}^*$ and $\mathbb{Q}=\mathbb{I}$. The pentagram marks the point that $P_{\rm e}(\lambda)=1-\frac{1}{3}=\frac{2}{3}$.

Example 4: We continue Example 3 with $N=2$. Assume noiseless wireless links between the sensors and the FC, i.e., $\mathbb{Q}=\mathbb{I}$. Unaware of a Byzantine attack, the FC will partition $\mathcal{M}^2=\mathcal{M}\times\mathcal{M}$ into

$$\begin{aligned}\mathcal{O}^{(\theta_1)}&=\left\{\begin{bmatrix}2\\3\end{bmatrix},\begin{bmatrix}3\\2\end{bmatrix},\begin{bmatrix}3\\3\end{bmatrix},\begin{bmatrix}3\\4\end{bmatrix},\begin{bmatrix}4\\3\end{bmatrix}\right\},\\\mathcal{O}^{(\theta_2)}&=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}1\\3\end{bmatrix},\begin{bmatrix}1\\4\end{bmatrix},\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}3\\1\end{bmatrix},\begin{bmatrix}4\\1\end{bmatrix}\right\},\\\mathcal{O}^{(\theta_3)}&=\left\{\begin{bmatrix}2\\2\end{bmatrix},\begin{bmatrix}2\\4\end{bmatrix},\begin{bmatrix}4\\2\end{bmatrix},\begin{bmatrix}4\\4\end{bmatrix}\right\}.\end{aligned}$$

Setting $\boldsymbol{d}=\boldsymbol{d}^*$ and $\alpha_{\rm B}=\alpha^*_{\rm blind}=\frac{3}{8}$ from Example 3, we obtain that for $0\leq\lambda{:=}\alpha/\alpha_{\rm B}\leq\frac{8}{3}$,

$$\begin{aligned}P_{\rm c}(\lambda)&=\frac{1}{3}\sum_{\theta\in\Theta}\sum_{\boldsymbol{z}\in\mathcal{O}^{(\theta)}}\prod_{i=1}^{2}\big[c_{z_i}^{(\theta)}+\lambda\cdot(d^*_{z_i}-c_{z_i}^{(\theta)})\big]\\&=\frac{1}{3}\sum_{\theta\in\Theta}\sum_{\boldsymbol{z}\in\mathcal{O}^{(\theta)}}\big[c_{z_1}^{(\theta)}c_{z_2}^{(\theta)}+\lambda^2\cdot(d^*_{z_1}-c_{z_1}^{(\theta)})(d^*_{z_2}-c_{z_2}^{(\theta)})\\&\qquad+\lambda\cdot(d^*_{z_1}-c_{z_1}^{(\theta)})c_{z_2}^{(\theta)}+\lambda\cdot(d^*_{z_2}-c_{z_2}^{(\theta)})c_{z_1}^{(\theta)}\big].\end{aligned}$$

We then observe from Fig. 2 that $P_{\rm e}(\lambda)=1-P_{\rm c}(\lambda)$ is equal to $\frac{2}{3}$ when $\lambda=1$, exceeds $\frac{2}{3}$ as λ grows beyond 1, and reaches 1 when $\lambda=1/\alpha^*_{\rm blind}=\frac{8}{3}$. □

IV. Numerical Results

In this section, two scenarios will be investigated. Specifically, Subsection IV-A examines the Byzantine attack policy proposed in Subsection III-C, and verifies the closeness between the average global detection error probabilities of the α-setting and the β-setting. Subsection IV-B examines how an imperfect Byzantine estimate regarding the POI affects the attack performance.

A. Optimal Sensor Compromising Probability based on Perfect Estimate of the POI

We now examine the performance deterioration of global detection due to the proposed Byzantine attack policy in Subsection III-C with $\alpha_{\rm B} = \alpha^*_{\rm blind}$, subject to a perfect knowledge of the POI.

The system model in [17] is employed and summarized as follows. Let the local observation of the ith sensor be modeled by $r_i = \theta + s_i$, $i \in \mathcal{N}$, where $\theta \in \Theta = \{-\mu, \mu\}$ is an antipodally modulated signal to be estimated, and $\{s_i\}_{i=1}^N$ is an independent sequence of random variables having the same Gaussian distribution with mean zero and variance $\sigma^2_{\rm sen}$. The conditional pmf of u_i given θ follows the simple threshold quantizer as $u_i = m$ if $\eta_{m-1} < r_i \le \eta_m$ with

$$\eta_m := \begin{cases} -\infty, & m = 0; \\ A_M \cdot (2m - M), & 1 \le m < M; \\ \infty, & m = M, \end{cases}$$

where $A_M := \frac{A}{(M-3)\cdot \mathbf{1}\{M>2\}+1}$, and A is an *overloading* parameter [21]. As a result, for $m \in \mathcal{M}$,

$$\begin{aligned} c_m^{(\theta)} &= \Pr(u_i = m \,|\, \theta) \\ &= \Pr(\eta_{m-1} < \theta + s_i \le \eta_m) \\ &= \Pr(\eta_{m-1} - \theta < s_i \le \eta_m - \theta) \\ &= \Phi\left(\frac{\eta_m - \theta}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{\eta_{m-1} - \theta}{\sigma_{\rm sen}}\right), \end{aligned} \tag{25}$$

where Φ is the standard normal cumulative distribution function (cdf).

The transition probability matrix $\mathbb{Q}$ of the discrete noisy link between the sensors and the FC is given by $q_{\ell,m} = \Pr(z_i = m | v_i = \ell) = \Pr(\eta_{m-1} < y_i \le \eta_m | v_i = \ell)$, where

$$y_i = A_M \cdot (2v_i - M - 1) + n_i \text{ for } v_i = 1, 2, \ldots, M, \tag{26}$$

and $\{n_i\}_{i=1}^N$ is an independent sequence of random variables with each n_i being Gaussian distributed with mean zero and variance σ^2. Note that the factor A_M is multiplied onto $(2v_i - M - 1)$ in (26) so that we can conveniently re-use thresholds η_{m-1} and η_m to define $q_{\ell,m}$. As an example of the transition probability of z_i given v_i under $M = 4$ and $A = 2$, we have

$$\mathbb{Q} = \begin{bmatrix} 1-\epsilon_1 & \epsilon_1 - \epsilon_3 & \epsilon_3 - \epsilon_5 & \epsilon_5 \\ \epsilon_1 & 1 - 2\epsilon_1 & \epsilon_1 - \epsilon_3 & \epsilon_3 \\ \epsilon_3 & \epsilon_1 - \epsilon_3 & 1 - 2\epsilon_1 & \epsilon_1 \\ \epsilon_5 & \epsilon_3 - \epsilon_5 & \epsilon_1 - \epsilon_3 & 1 - \epsilon_1 \end{bmatrix}, \tag{27}$$

where $\epsilon_k := \Phi(-k/\sigma)$ for $k \in \{1, 2, \ldots\}$.

Under the above setting, we first examine the situation, where an attacker attempts to blind the FC by targeting a local output pmf $\boldsymbol{d}^\diamond := (\mathbb{Q}^{\mathsf{T}})^{-1}\boldsymbol{b}_{\rm uni}$ with $\boldsymbol{b}_{\rm uni} = \frac{1}{M}\mathbf{1}$ being a uniform distribution as in [17]. The Byzantine attacker thus devises the Byzantine probability matrix $\mathbf{P}^{(\alpha_{\rm B})}$ according to (18), parameterized with $\alpha_{\rm B} = \alpha_{\rm blind}(\boldsymbol{d}^\diamond)$. We then illustrate the global detection error probability $P_{\rm e}$ derived in (22) under $N = 10$, $\mu = 1$, $\sigma^2_{\rm sen} = 1$, $\sigma^2 = 4$ and $A = 2$ as a function of the true sensor compromising probability α in Fig. 3. For comparison, we also plot the global detection error probability under the β-setting with $\beta = K/N$, where the first K sensors are compromised.

Three observations are made. First, it can be observed from Fig. 3 that the global detection error probabilities of the α-setting and the β-setting do not coincide with each other when the size of the sensor network is finite. However, their difference is quite small, confirming that the performance of a β-setting system can be approximated by that of an α-setting system. Second, $\alpha_{\rm blind}(\boldsymbol{d}^\diamond)$ increases from 0.4057 to 0.5508 as M grows from 2 to 4. Thus, a higher quantization resolution increases the blinding effort of a Byzantine attacker if uniform $\boldsymbol{b}_{\rm uni}$ is the one to be achieved (as concluded in [17]). Third, further increase in M such as $M = 8$ will lead to a $\boldsymbol{d}^\diamond = (\mathbb{Q}^{\mathsf{T}})^{-1}\boldsymbol{b}_{\rm uni}$ that contains negative components and hence is no longer a legitimate pmf. Note that the authors of [17] consider only the *doubly stochastic* noisy wireless links that satisfy both $\mathbb{Q}\,\mathbf{1} = \mathbf{1}$ and $\mathbb{Q}^{\mathsf{T}}\,\mathbf{1} = \mathbf{1}$, which is clearly violated by (27).[7]

Next, we replace $\boldsymbol{d} = (\mathbb{Q}^{\mathsf{T}})^{-1}\boldsymbol{b}_{\rm uni}$ by $\boldsymbol{d} = \boldsymbol{d}^*$ in Theorem 2 and summarize the results in Fig. 4. For simplicity, the global detection error probabilities for the β-setting are given only for $M = 16$. We observe from Fig. 4 that by taking $\boldsymbol{d} = \boldsymbol{d}^*$, all four curves equal $1/2$ at $\alpha = \alpha^*_{\rm blind} = 1 - \frac{1}{2\Phi(1)} \approx 0.4057$. This observation can be analytically verified as follows. From the fact that $\eta_m = -\eta_{M-m}$ for $0 \le m \le M$, it can be derived from (25) that for $1 \le m \le M$,

$$\begin{aligned} c_m^{(\mu)} &= \Phi\left(\frac{\eta_m - \mu}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{\eta_{m-1} - \mu}{\sigma_{\rm sen}}\right) \\ &= \Phi\left(\frac{-\eta_{m-1} + \mu}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{-\eta_m + \mu}{\sigma_{\rm sen}}\right) \\ &= \Phi\left(\frac{\eta_{M-m+1} + \mu}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{\eta_{M-m} + \mu}{\sigma_{\rm sen}}\right) \\ &= c_{M-m+1}^{(-\mu)} \end{aligned}$$

and for $1 \le m \le M/2$ with M even,

$$\begin{aligned} c_m^{(-\mu)} &= \Phi\left(\frac{\eta_m + \mu}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{\eta_{m-1} + \mu}{\sigma_{\rm sen}}\right) \\ &> \Phi\left(\frac{-\eta_{m-1} + \mu}{\sigma_{\rm sen}}\right) - \Phi\left(\frac{-\eta_m + \mu}{\sigma_{\rm sen}}\right) \\ &= c_{M-m+1}^{(-\mu)}. \end{aligned}$$

We thus have

$$\begin{aligned} \sum_{m=1}^{M} \max_{\theta\in\Theta} c_m^{(\theta)} &= \sum_{m=1}^{M} \max\left\{c_m^{(-\mu)}, c_m^{(\mu)}\right\} \\ &= \sum_{m=1}^{M} \max\left\{c_m^{(-\mu)}, c_{M-m+1}^{(-\mu)}\right\} \end{aligned}$$

[7] One may notice an anti-intuitive result in Fig. 3 that increasing the resolution from 1 bit to 2 bits actually degrades the global detection error probability at $\alpha = 0$. Note that as the partitions in (23) are optimal without Byzantine attack, the global detection error at $\alpha = 0$ is optimal. This anti-intuitive result is indeed due to the fact that the channel transition probability under $M = 4$ as given in (27) is not a refinement of $\mathbb{Q} = \left[\begin{smallmatrix} 1-\epsilon_2 & \epsilon_2 \\ \epsilon_2 & 1-\epsilon_2 \end{smallmatrix}\right]$ under $M = 2$. When replacing $\mathbb{Q}$ by an identity matrix, a larger M shall give a lower global detection error at $\alpha = 0$ as anticipated.

When $\alpha > 0$, the partitions in (23) are no longer optimal with respect to α; thus, there is no guarantee that a higher resolution renders a smaller global detection error probability even if identity $\mathbb{Q}$ is used.

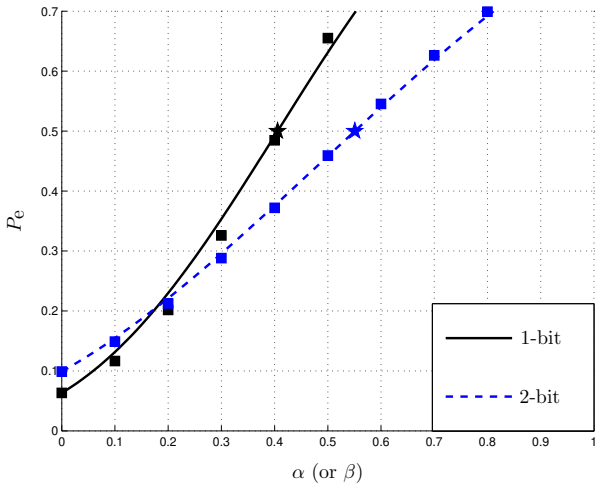


Fig. 3: Global detection error probability P_e as a function of α under $\boldsymbol{d} = \boldsymbol{d}^\diamond$ and $\alpha_B = \alpha_{\text{blind}}(\boldsymbol{d}^\diamond)$, where $\boldsymbol{d}^\diamond := (\mathbb{Q}^\intercal)^{-1}\boldsymbol{b}_{\text{uni}}$ with $\boldsymbol{b}_{\text{uni}} = \frac{1}{M}\mathbf{1}$. The setting of this simulation is that $N = 10$, $\mu = 1$, $\sigma^2_{\text{sen}} = 1$, $\sigma^2 = 4$ and $A = 2$. The two pentagrams mark the values of $\alpha_{\text{blind}}(\boldsymbol{d}^\diamond)$ for $M = 2$ (i.e., 1 bit) and $M = 4$ (i.e., 2 bits). As a reference, square points mark the resulting global detection error probability for the β-setting.

$$
\begin{aligned}
&= 2\sum_{m=1}^{M/2} \max\Big\{c_m^{(-\mu)}, c_{M-m+1}^{(-\mu)}\Big\} \\
&= 2\sum_{m=1}^{M/2} c_m^{(-\mu)} \\
&= 2\cdot \Pr[s_i < \mu] = 2\cdot\Phi(\sqrt{\gamma_{\text{sen}}}), \qquad (28)
\end{aligned}
$$

where $\gamma_{\text{sen}} := \mu^2/\sigma^2_{\text{sen}}$ is the sensing signal-to-noise ratio. Hence, $\alpha^*_{\text{blind}} = 1 - \frac{1}{2\cdot\Phi(\sqrt{\gamma_{\text{sen}}})}$, which is independent of M. This result indicates that a higher quantization resolution does not necessarily increase the blinding effort of a Byzantine attacker if an elaborate design of $\boldsymbol{d} = \boldsymbol{d}^*$ is adopted.

Note that the constant α^*_{blind} resulting from the binary state space $\{-\mu, \mu\}$ is actually a special case, and α^*_{blind} in general grows slightly as the local quantization resolution M increases. An example is given in Table I, where the state space of $\Theta = \{-3\mu, -\mu, \mu, 3\mu\}$ results in a mildly growing α^*_{blind} with respect to increasing local quantization resolution from 3 to 5. In spite of a larger α^*_{blind} as obtained in Table I, in comparison with the binary state space, the table shows that blinding the system is always possible, i.e., an $\alpha^*_{\text{blind}} \le 1$ can always be obtained. However, if a Byzantine attacker targets for uniform $\boldsymbol{b}_{\text{uni}}$, system blindness becomes unattainable when $\log_2(M) \ge 3$.

We next examine how the sensing signal-to-noise ratio γ_{sen} and the overloading parameter A affect the minimum blinding effort α^*_{blind}. Under $\Theta = \{-3\mu, -\mu, \mu, 3\mu\}$, we derive similar to (28) that

$$
\begin{aligned}
\sum_{m=1}^{M} \max_{\theta\in\Theta} c_m^{(\theta)} &= \sum_{m=1}^{M} \max\Big\{c_m^{(-3\mu)}, c_m^{(-\mu)}, c_m^{(\mu)}, c_m^{(3\mu)}\Big\} \\
&= \sum_{m=1}^{M} \max\Big\{c_m^{(-3\mu)}, c_m^{(-\mu)}, c_{M-m+1}^{(-\mu)}, c_{M-m+1}^{(-3\mu)}\Big\} \\
&= 2\sum_{m=1}^{M/2} \max\Big\{c_m^{(-3\mu)}, c_m^{(-\mu)}, c_{M-m+1}^{(-\mu)}, c_{M-m+1}^{(-3\mu)}\Big\} \\
&= 2\sum_{m=1}^{M/2} \max\Big\{c_m^{(-3\mu)}, c_m^{(-\mu)}\Big\} \\
&= 2c_1^{(-3\mu)} + 2\sum_{m=2}^{M/2} \max\Big\{c_m^{(-3\mu)}, c_m^{(-\mu)}\Big\}.
\end{aligned}
$$

For $M = 4$, we obtain

$$
\begin{aligned}
&\sum_{m=1}^{4} \max_{\theta\in\Theta} c_m^{(\theta)} = 2\cdot c_1^{(-3\mu)} + 2\cdot \max\Big\{c_2^{(-3\mu)}, c_2^{(-\mu)}\Big\} \\
&= 2\cdot \max\Big\{\Phi\Big(\tfrac{-A+3\mu}{\sigma_{\text{sen}}}\Big) + \Phi\Big(\tfrac{\mu}{\sigma_{\text{sen}}}\Big) - \Phi\Big(\tfrac{-A+\mu}{\sigma_{\text{sen}}}\Big), \Phi\Big(\tfrac{3\mu}{\sigma_{\text{sen}}}\Big)\Big\} \\
&= \begin{cases} 2\cdot\Big[\Phi\Big(\tfrac{-A+3\mu}{\sigma_{\text{sen}}}\Big) + \Phi\Big(\tfrac{\mu}{\sigma_{\text{sen}}}\Big) - \Phi\Big(\tfrac{-A+\mu}{\sigma_{\text{sen}}}\Big)\Big], & A < 4\mu; \\ 2\cdot\Phi\Big(\tfrac{3\mu}{\sigma_{\text{sen}}}\Big), & A \ge 4\mu. \end{cases} \qquad (29)
\end{aligned}
$$

With $\alpha^*_{\text{blind}} = 1 - 1/\sum_{m=1}^{4} \max_{\theta\in\Theta} c_m^{(\theta)}$, we note as anticipated that a higher sensing signal-to-noise ratio $\gamma_{\text{sen}} = \mu^2/\sigma^2_{\text{sen}}$ requires a higher Byzantine effort for system blindness. However, under $A \ge 4\mu$, α^*_{blind} can only approach $1/2$ as γ_{sen} grows large, which is contrary to what has been shown in Table I, in which α^*_{blind} tends to reach $1 - \frac{1}{\min\{L,M\}} = \frac{3}{4}$ by increasing M. It can be further derived from (29) that $\sum_{m=1}^{4} \max_{\theta\in\Theta} c_m^{(\theta)}$ is maximized by taking $A = 2\mu$, which leads to $\alpha^*_{\text{blind}} = 1 - \frac{1}{2[3\Phi(\gamma_{\text{sen}})-1]} \to \frac{3}{4}$ as γ_{sen} goes to infinity. As a result, an attacker would prefer a system with a larger A and a lower quantization resolution.

TABLE I: Minimum blinding effort α^*_{blind} for $\Theta = \{-3\mu, -\mu, \mu, 3\mu\}$. The system setting is the same as in Fig. 4 except $A = 4$.

Number of quantization bits	α^*_{blind}	$\alpha_{\text{blind}}(\boldsymbol{d}^\diamond)$
2	0.4993	0.7017
3	0.6489	Unattainable
4	0.6677	Unattainable
5	0.6710	Unattainable

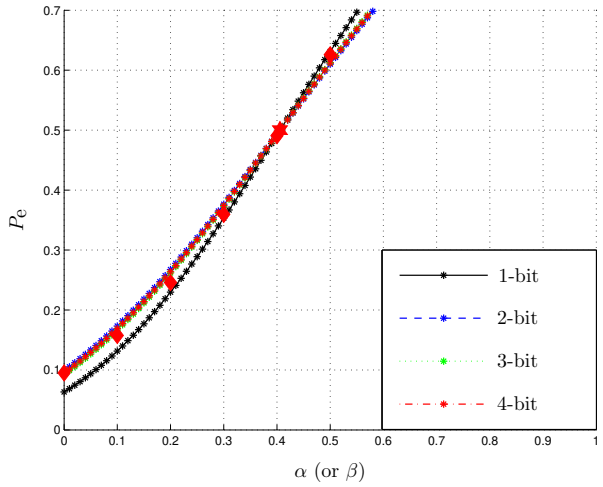
B. Near-Optimal Sensor Compromising Probability based on Imprecisely Estimated POI

We now examine the impact of an imprecisely estimated POI on the attack performance.

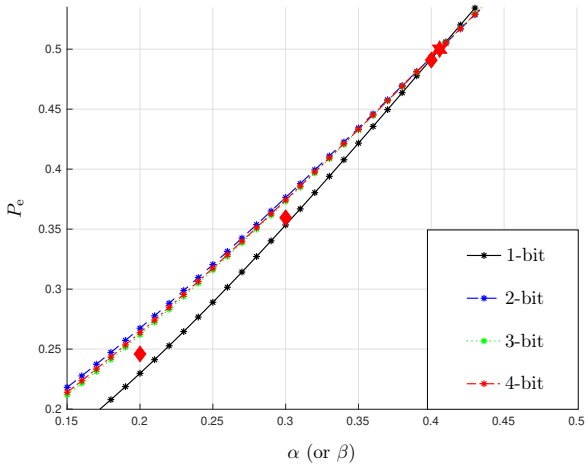
If a group of N_B compromised sensors can exchange local quantization outputs $\{u_i\}_{i=1}^{N_B}$ through, e.g., a separate secret channel and jointly make an estimate of the POI, then the detection probability of error is equal to

$$
P_{\text{e, Byzantine}} = 1 - \frac{1}{L}\sum_{\boldsymbol{u}\in\mathcal{M}^{N_B}} \max_{\theta\in\Theta} \prod_{i=1}^{N_B} c_{u_i}^{(\theta)}. \qquad (30)
$$

We list $P_{\text{e, Byzantine}}$ in Table II under the same setting as in Fig. 4 for N_B less than ten, and notice that the Byzantine



(a)



(b) A zoomed in version from $\alpha = 0.15$ to $\alpha = 0.5$

Fig. 4: Global detection error probability P_e as a function of α under $\boldsymbol{d} = \boldsymbol{d}^*$ and $\alpha_B = \alpha^*_{\text{blind}}$. The setting of this simulation is that $N = 10$, $\mu = 1$, $\sigma^2_{\text{sen}} = 1$, $\sigma^2 = 4$ and $A = 2$. For all local quantization resolutions of $M = 2$, 4, 8 and 16, we obtain $\alpha^*_{\text{blind}} = 1 - \frac{1}{2\Phi(1)} \approx 0.4057$, which is marked by a red hexagram. As a reference, diamond points mark the resulting detection error probability for the β-setting with $M = 16$.

TABLE II: Byzantine detection error probability $P_{\text{e, Byzantine}}$ as a function of the number of cooperative compromised sensors N_B and the local quantization resolution M

N_B	$P_{\text{e, Byzantine}}$			
	Number of quantization bits			
	1	2	3	4
1	0.1587	0.1587	0.1587	0.1587
2	0.1587	0.1346	0.0863	0.0801
3	0.0675	0.0576	0.0436	0.0421
4	0.0675	0.0494	0.0258	0.0234
5	0.0310	0.0230	0.0139	0.0130
6	0.0310	0.0197	0.0084	0.0074
7	0.0148	0.0095	0.0046	0.0042
8	0.0148	0.0082	0.0029	0.0025
9	0.0072	0.0041	0.0016	0.0014

detection error probability $P_{\text{e, Byzantine}}$ can be made below 0.05 through cooperation among just four compromised sensors when $M \geq 4$. The same Byzantine detection error probability level can be reached for $M \geq 2$ when only five compromised sensors form a cooperative group. Notably, $P_{\text{e, Byzantine}}$ is independent of the number of sensors N but only a function of the number of cooperative compromised sensors N_B (see the exact expression in (30)); hence, when N is further increased, the attacker can attain the same level of $P_{\text{e, Byzantine}}$ with a smaller fraction of cooperative compromised sensors to all the sensors.

In order to evaluate how the accuracy of Byzantine estimate affects the attack performance, we substitute (30) by a simple model as follows:

$$\Pr\left(\hat{\theta}_{\text{Byzantine}} = \theta' \,\middle|\, \theta\right) = \begin{cases} 1-(L-1)\delta, & \theta' = \theta; \\ \delta, & \theta' \neq \theta, \end{cases}$$

where $\delta > 0$ is regarded as a rough approximation of the probability that the compromised sensors guess wrongly on the true POI θ.

We then re-derive (19) by incorporating the above POI estimation model as:

$$\begin{aligned} \boldsymbol{a}^{(\theta,\alpha_B,\alpha)} &= \Big[(1-\alpha)\mathbb{I} + \alpha\big(1-(L-1)\delta\big)\big(\mathbb{P}^{(\theta,\alpha_B)}\big)^{\mathsf{T}} \\ &\quad + \sum_{\theta'\in\Theta,\theta'\neq\theta} \alpha\delta\big(\mathbb{P}^{(\theta',\alpha_B)}\big)^{\mathsf{T}}\Big]\boldsymbol{c}^{(\theta)} \\ &= \boldsymbol{c}^{(\theta)} - \alpha\big(1-(L-1)\delta\big)\big[\mathbb{I} - \big(\mathbb{P}^{(\theta,\alpha_B)}\big)^{\mathsf{T}}\big]\boldsymbol{c}^{(\theta)} \\ &\quad - \sum_{\theta'\in\Theta,\theta'\neq\theta} \alpha\delta\big[\mathbb{I} - \big(\mathbb{P}^{(\theta',\alpha_B)}\big)^{\mathsf{T}}\big]\boldsymbol{c}^{(\theta)} \\ &= \boldsymbol{c}^{(\theta)} - \alpha\big(1-(L-1)\delta\big)\frac{\big(\boldsymbol{c}^{(\theta)} - \boldsymbol{d}\big)}{\alpha_B} \\ &\quad - \sum_{\theta'\in\Theta,\theta'\neq\theta} \alpha\delta\frac{\big(\boldsymbol{c}^{(\theta')} - \boldsymbol{d}\big)}{\alpha_B} \\ &= \left[1 - \big(1-(L-1)\delta\big)\frac{\alpha}{\alpha_B}\right]\boldsymbol{c}^{(\theta)} \\ &\quad - \sum_{\theta'\in\Theta,\theta'\neq\theta} \delta\frac{\alpha}{\alpha_B}\boldsymbol{c}^{(\theta')} + \frac{\alpha}{\alpha_B}\boldsymbol{d}, \end{aligned}$$

and the conditional distribution of z_i given θ is

$$\begin{aligned} \boldsymbol{o}^{(\theta,\alpha_B,\alpha)} &= \mathbb{Q}^{\mathsf{T}}\boldsymbol{a}^{(\theta,\alpha_B,\alpha)} \\ &= \left[1 - \frac{\alpha}{\alpha_B} + (L-1)\delta\frac{\alpha}{\alpha_B}\right]\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta)} \\ &\quad - \sum_{\theta'\in\Theta,\theta'\neq\theta} \delta\frac{\alpha}{\alpha_B}\mathbb{Q}^{\mathsf{T}}\boldsymbol{c}^{(\theta')} + \frac{\alpha}{\alpha_B}\boldsymbol{b}. \end{aligned} \quad (31)$$

The global detection error probability at the FC can thus be obtained via the formulation of (21), which according to (31) becomes a function of Byzantine estimation error δ as shown in Table III. Specifically, under the same setting as in Fig. 4,

TABLE III: Global detection error probability $P_{\rm e}$ as a function of the Byzantine estimate error δ on the phenomenon θ under $\alpha_{\rm B} = \alpha^*_{\rm blind} \approx 0.4057$ for two different sensor compromising probabilities $\alpha = \alpha^*_{\rm blind}$ and $\alpha = 0.5$

δ	Number of quantization bits	$P_{\rm e}$ for $\alpha = \alpha^*_{\rm blind}$	$P_{\rm e}$ for $\alpha = 0.5$
0.1	1	0.3866	0.4919
	2	0.4049	0.4933
	3	0.4025	0.4931
	4	0.4032	0.4931
0.05	1	0.4428	0.5624
	2	0.4522	0.5521
	3	0.4509	0.5534
	4	0.4513	0.5530
0.01	1	0.4885	0.6177
	2	0.4904	0.5986
	3	0.4902	0.6010
	4	0.4902	0.6002

we observe from Table III that an imperfect Byzantine estimate does mitigate the deterioration effect of a Byzantine attack, and setting $\alpha = \alpha^*_{\rm blind}$ can no longer blind the system; however, the global detection error probability at the FC is still very close to $1 - \frac{1}{L} = \frac{1}{2}$. From the attacker viewpoint, a remedy could be to raise α from $\alpha^*_{\rm blind} \approx 0.4057$ to, e.g., 0.5, in which case a Byzantine estimation error δ of 0.1 can still degrade the global detection error down to a random-guess performance.

We remark at the end that $\alpha^*_{\rm blind}$ obtained in Fig. 4 seems to hint that the attacker needs to compromise at least half of the sensors because the smallest integer larger than $N \times \alpha^*_{\rm blind} = 10 \times \alpha^*_{\rm blind}$ is 5. However, such a high need for the fraction of compromised sensors is due to the strong requirement of blinding the system. When targeting only a certain degree of deterioration of global detection performance instead of blinding the system, an attacker can still adopt the statistical attacking scheme proposed in Subsection III-C, and expect that a fairly serious deterioration can be reached with an α smaller than $\alpha^*_{\rm blind}$. This expectation can be supported by Fig. 4, where the curves of global detection error probabilities at the FC follow closely a direct trace connecting the point at $\alpha = 0$ and the point with coordinate $(\alpha, P_{\rm e}) = (\alpha^*_{\rm blind}, \frac{1}{2})$. Since $\alpha^*_{\rm blind}$ is the minimum sensor compromising probability attainable by any statistical attacking scheme that is required to result in $P_{\rm e} = \frac{1}{2}$, it may be justified to state that the global detection error probability at the FC is close to a certain worst value by our attacking scheme at any α between 0 and $\alpha^*_{\rm blind}$.

V. Conclusion

An optimal Byzantine attack policy for WSNs with M-ary data from a finite of number of local sensors was derived under the assumption that an attacker can acquire the statistics of local outputs. The closed-form expression of the minimum sensor compromising probability $\alpha^*_{\rm blind}$ to achieve the perfect blindness of global detection was obtained. While most of the early works on robustness of distributed detection focused on the binary hypothesis testing problem due to analytical convenience, our results can be well applied for an arbitrary finite number of hypotheses. In addition, we have generalized the notion of system blindness, which simply requires the independence between the receptions from sensors and the POI; as a result of this generalization, the optimal Byzantine attack policy we proposed works for any noisy link as long as the transition matrix of the noisy link $\mathbb{Q}$ admits an inverse. In situations when $\mathbb{Q}$ is not invertible, the minimum sensor compromising probability is often smaller than the $\alpha^*_{\rm blind}$ from Theorem 2. In particular, when $M = 2$, the minimum sensor compromising probability is reduced to zero for any non-invertible $\mathbb{Q}$. For $M > 2$, however, the minimum sensor compromising probability subject to a non-invertible $\mathbb{Q}$ shall become a function of $\mathbb{Q}$ and may not exhibit a simple expression as that in Theorem 2, and its determination could be an interesting theoretical challenge.

Appendix A
Proof of Theorem 1

The theorem is actually intended to prove

$$\alpha_{\rm blind}(\boldsymbol{d}) = \max_{\theta\in\Theta\,:\,\mathcal{M}_2^{(\theta)}\neq\emptyset} e^{(\theta)}_{k(\theta)}$$

where $e^{(\theta)}_m$ and $k(\theta)$ are respectively defined in (8) and (11). Note that it is reasonable to assume that

$$\mathcal{M}_2^{(\theta)} \neq \emptyset \text{ for at least one } \theta \in \Theta; \tag{32}$$

otherwise, we have $\boldsymbol{c}^{(\theta)} = \boldsymbol{d}$ for every $\theta \in \Theta$, which implies trivially $\alpha_{\rm blind}(\boldsymbol{d}) = 0$.

With the validity of (32), the theorem can be proved in two steps. The first step shows that every α satisfying $\boldsymbol{c}^{(\theta)} - \boldsymbol{d} = \alpha\big(\mathbb{I} - (\mathbb{P}^{(\theta)})^{\intercal}\big)\boldsymbol{c}^{(\theta)}$ for some $\mathbb{P}^{(\theta)}$ must be no less than $e^{(\theta)}_{k(\theta)}$ if $\mathcal{M}_2^{(\theta)} \neq \emptyset$. The second step gives the specific choice of $\mathbb{P}^{(\theta)*}$ that verifies the achievability of the claimed $\alpha_{\rm blind}(\boldsymbol{d})$.

Step 1: $\alpha_{\rm blind}(\boldsymbol{d}) \geq e^{(\theta)}_{k(\theta)}$ for every $\theta \in \Theta$ with $\mathcal{M}_2^{(\theta)} \neq \emptyset$: We know from (7) that

$$\begin{aligned} c^{(\theta)}_{k(\theta)} - d_{k(\theta)} &= \alpha\left(c^{(\theta)}_{k(\theta)} - \sum_{m\in\mathcal{M}} p^{(\theta)}_{m,k(\theta)} c^{(\theta)}_m\right) \\ &= \alpha\left[\left(1 - p^{(\theta)}_{k(\theta),k(\theta)}\right)c^{(\theta)}_{k(\theta)} - \sum_{m\in\mathcal{M}\setminus\{k(\theta)\}} p^{(\theta)}_{m,k(\theta)} c^{(\theta)}_m\right]. \end{aligned} \tag{33}$$

This implies

$$\begin{aligned} \alpha &\geq \alpha\left(1 - p^{(\theta)}_{k(\theta),k(\theta)}\right) \\ &= 1 - \frac{d_{k(\theta)}}{c^{(\theta)}_{k(\theta)}} + \alpha \underbrace{\sum_{m\in\mathcal{M}\setminus\{k(\theta)\}} p^{(\theta)}_{m,k(\theta)} \frac{c^{(\theta)}_m}{c^{(\theta)}_{k(\theta)}}}_{\geq 0} \qquad (34) \\ &\geq 1 - \frac{d_{k(\theta)}}{c^{(\theta)}_{k(\theta)}} = e^{(\theta)}_{k(\theta)}, \qquad (35) \end{aligned}$$

where the equality in (34) follows from (33).

Step 2: $\alpha_{blind}(\boldsymbol{d}) = \max_{\theta\in\Theta\,:\,\mathcal{M}_2^{(\theta)}\neq\emptyset} e_{k(\theta)}^{(\theta)}$*:* We now show that the specified $\mathbb{P}^{(\theta)*}$ validates (7) with $\alpha = \alpha_{\text{blind}}(\boldsymbol{d})$.

When $\mathcal{M}_2^{(\theta)}$ is empty, we have $\boldsymbol{c}^{(\theta)} = \boldsymbol{d}$. Thus, taking $\mathbb{P}^{(\theta)*} = \mathbb{I}$ and $\alpha = \alpha_{\text{blind}}(\boldsymbol{d})$ trivially validates (7).

Next, subject to that $\mathcal{M}_2^{(\theta)}$ is non-empty, we re-write (7) as $\alpha\big(\mathbb{P}^{(\theta)}\big)^{\mathsf{T}}\boldsymbol{c}^{(\theta)} = \alpha\boldsymbol{c}^{(\theta)} + \boldsymbol{d} - \boldsymbol{c}^{(\theta)}$, which is equivalent to

$$p_{j,j}^{(\theta)} c_j^{(\theta)} + \sum_{\ell=1,\ell\neq j}^{M} p_{\ell,j}^{(\theta)} c_\ell^{(\theta)} = c_j^{(\theta)} + \frac{d_j - c_j^{(\theta)}}{\alpha}$$

for $j \in \mathcal{M}$; here, the action of dividing by α is justified as we have already proved in (35) that $\alpha \geq e_{k(\theta)}^{(\theta)}$, and $e_{k(\theta)}^{(\theta)} > 0$ for non-empty $\mathcal{M}_2^{(\theta)}$.

We thus take $\alpha = \alpha_{\text{blind}}(\boldsymbol{d})$ and $\mathbb{P}^{(\theta)} = \mathbb{P}^{(\theta)*}$ into the above equation and obtain:

$$\begin{cases} p_{j,j}^{(\theta)*} c_j^{(\theta)} + \displaystyle\sum_{\ell\in\mathcal{M}_1^{(\theta)}\setminus\{j\}} p_{\ell,j}^{(\theta)*} c_\ell^{(\theta)} + \sum_{\ell\in\mathcal{M}_2^{(\theta)}} p_{\ell,j}^{(\theta)*} c_\ell^{(\theta)} \\ \qquad\qquad = c_j^{(\theta)} + \dfrac{d_j - c_j^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} \text{ for } j \in \mathcal{M}_1^{(\theta)}; \\ p_{j,j}^{(\theta)*} + \displaystyle\sum_{\ell\in\mathcal{M}\setminus\{j\}} p_{\ell,j}^{(\theta)*} \frac{c_\ell^{(\theta)}}{c_j^{(\theta)}} = 1 - \frac{e_j^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} \text{ for } j \in \mathcal{M}_2^{(\theta)}. \end{cases} \tag{36}$$

It remains to show that the elements of $\mathbb{P}^{(\theta)*}$ satisfy the two sets of equations in (36).

From (10), we obtain that

$$\begin{cases} p_{j,j}^{(\theta)*} = 1 - \dfrac{e_j^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} & \text{for } j \in \mathcal{M}_2^{(\theta)}; \\ p_{\ell,j}^{(\theta)*} = 0 & \text{for } \ell \in \mathcal{M} \text{ and } j \in \mathcal{M}_2^{(\theta)} \text{ and } \ell \neq j, \end{cases}$$

which immediately validates the second set of equations in (36). The first set of equations in (36) can be confirmed by noting that:

$$\begin{cases} p_{j,j}^{(\theta)*} = 1 & \text{for } j \in \mathcal{M}_1^{(\theta)}; \\ p_{\ell,j}^{(\theta)*} = 0 & \text{for } \ell \in \mathcal{M}_1^{(\theta)} \text{ and } j \in \mathcal{M}_1^{(\theta)} \text{ and } \ell \neq j; \\ \displaystyle\sum_{\ell\in\mathcal{M}_2^{(\theta)}} p_{\ell,j}^{(\theta)*} c_\ell^{(\theta)} = \dfrac{d_j - c_j^{(\theta)}}{\alpha_{\text{blind}}(\boldsymbol{d})} & \text{for } j \in \mathcal{M}_1^{(\theta)}. \end{cases}$$

Appendix B
Proof of Theorem 2

For notational convenience, we define

$$t(\boldsymbol{d}) := 1 - \alpha_{\text{blind}}(\boldsymbol{d}) = \min_{\theta\in\Theta} \min_{m\in\mathcal{M}\,:\,c_m^{(\theta)}>0} \left\{\frac{d_m}{c_m^{(\theta)}}\right\},$$

and

$$t^* := \frac{1}{\sum_{m=1}^{M} \max_{\theta\in\Theta} c_m^{(\theta)}} = \frac{1}{\sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} c_m^{(\theta)}},$$

where $\{\mathcal{A}^{(\theta)}\}_{\theta\in\Theta}$ are L disjoint partitions of $\mathcal{M}$ (possibly with some empty partitions) such that $c_m^{(\theta)} = \max_{\theta'\in\Theta} c_m^{(\theta')}$ for $m \in \mathcal{A}^{(\theta)}$.

The theorem can be proved in two steps. We first show that $t(\boldsymbol{d}) \leq t^*$ for all $\boldsymbol{d} \in \mathcal{D}$. We next provide a specific $\boldsymbol{d}^*$ that validates $t(\boldsymbol{d}^*) = t^*$.

We now prove the first step by contradiction. Suppose that there exists a $\boldsymbol{d}^\circ \in \mathcal{D}$ satisfying $t(\boldsymbol{d}^\circ) > t^*$. Then, for every $\theta \in \Theta$ and for every $m \in \mathcal{M}$ with $c_m^{(\theta)} > 0$,

$$\frac{d_m^\circ}{c_m^{(\theta)}} > t^* = \frac{1}{\sum_{\theta\in\Theta}\sum_{m\in\mathcal{M}^{(\theta)}} c_m^{(\theta)}}.$$

Equivalently, for every $\theta \in \Theta$ and for every $m \in \mathcal{M}$ with $c_m^{(\theta)} > 0$,

$$d_m^\circ > \frac{c_m^{(\theta)}}{\sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} c_m^{(\theta)}}. \tag{37}$$

Noting that $c_m^{(\theta)} = \max_{\theta'\in\Theta} c_m^{(\theta')} > 0$ for $m \in \mathcal{A}^{(\theta)}$ according to the assumption in (1), we derive from (37) that

$$\begin{aligned} \sum_{m=1}^{M} d_m^\circ &= \sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} d_m^\circ \\ &> \sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} \frac{c_m^{(\theta)}}{\sum_{\theta\in\Theta}\sum_{\ell\in\mathcal{A}^{(\theta)}} c_\ell^{(\theta)}} \\ &= \frac{\sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} c_m^{(\theta)}}{\sum_{\theta\in\Theta}\sum_{\ell\in\mathcal{A}^{(\theta)}} c_\ell^{(\theta)}} = 1, \end{aligned}$$

which contradicts the fact that $\boldsymbol{d}^\circ$ is a pmf in $\mathcal{D}$. This finishes the proof of the first step.

It remains to provide a $\boldsymbol{d}^*$ such that $t(\boldsymbol{d}^*) = t^*$. Define

$$d_m^* := c_m^{(\theta)} \cdot t^* \quad \text{for } \theta \in \Theta \text{ and } m \in \mathcal{A}^{(\theta)}$$

and verify that

$$\sum_{m=1}^{M} d_m^* = \sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} d_m^* = t^* \sum_{\theta\in\Theta}\sum_{m\in\mathcal{A}^{(\theta)}} c_m^{(\theta)} = 1.$$

Then,

$$\begin{aligned} t(\boldsymbol{d}^*) &= \min_{\theta\in\Theta} \min_{m\in\mathcal{M}\,:\,c_m^{(\theta)}>0} \left\{\frac{d_m^*}{c_m^{(\theta)}}\right\} \\ &= \min_{\theta\in\Theta} \min\left\{ \min_{m\in\mathcal{A}^{(\theta)}} \left\{\frac{d_m^*}{c_m^{(\theta)}}\right\}, \min_{m\notin\mathcal{A}^{(\theta)}} \left\{\frac{d_m^*}{c_m^{(\theta)}}\right\} \right\} \\ &= \min_{\theta\in\Theta} \min\left\{ t^*, \min_{m\in\bigcup_{\theta'\in\Theta:\theta'\neq\theta}\mathcal{A}^{(\theta')}} \left\{\frac{c_m^{(\theta')}}{c_m^{(\theta)}} t^*\right\} \right\} \\ &= \min_{\theta\in\Theta} t^* = t^* \end{aligned} \tag{38}$$

where (38) holds since for $m \in \mathcal{A}^{(\theta')}$ and $\theta' \neq \theta$, $c_m^{(\theta')} = \max_{\theta''\in\Theta} c_m^{(\theta'')} \geq c_m^{(\theta)}$. Thus, $\max_{\boldsymbol{d}\in\mathcal{D}} t(\boldsymbol{d}) = t(\boldsymbol{d}^*) = t^*$.

References

[1] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[2] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors II. Advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan 1997.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Netw.*, vol. 38, no. 4, pp. 393–422, 2002.
[4] R. R. Brooks, P. Ramanathan, and A. M. Sayeed, "Distributed target classification and tracking in sensor networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1163–1171, 2003.
[5] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philos. Trans. Roy. Soc. London A, Math. Phys. Eng. Sci.*, vol. 370, no. 1958, pp. 100–117, Jan. 2012.
[6] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. New York: JAI Press, 1993, vol. 2, pp. 297–344.
[7] P. K. Varshney, *Distributed Detection and Data Fusion*. New York, NY, USA: Springer-Verlag, 1997.
[8] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Proc.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
[9] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Proc. Mag.*, vol. 23, no. 4, pp. 16–26, Jul. 2006.
[10] J.-F. Chamberland and V. V. Veeravalli, "Wireless sensors in distributed detection applications," *IEEE Signal Proc. Mag.*, vol. 24, no. 3, pp. 16–25, May 2007.
[11] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of Byzantine attacks," *IEEE Trans. Signal Proc.*, vol. 57, no. 1, pp. 16–29, Jan. 2009.
[12] 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 Trans. Signal Proc.*, vol. 59, no. 2, pp. 774–786, 2011.
[13] A. Vempaty, K. Agrawal, H. Chen, and P. Varshney, "Adaptive learning of Byzantines' behavior in cooperative spectrum sensing," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Cancun, Quintana Roo, Mexico, Mar. 28–31, 2011, pp. 1310–1315.
[14] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, "Distributed detection in tree topologies with Byzantines," *IEEE Trans. Signal Proc.*, vol. 62, no. 12, pp. 3208–3219, Jun. 2014.
[15] E. Soltanmohammadi, M. Orooji, and M. Naraghi-Pour, "Decentralized hypothesis testing in wireless sensor networks in the presence of misbehaving nodes," *IEEE Trans. Inf. Forens. & Secur.*, vol. 8, no. 1, pp. 205–215, Jan. 2013.
[16] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with Byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, pp. 65–75, Sep. 2013.
[17] V. S. S. Nadendla, Y. S. Han, and P. K. Varshney, "Distributed inference with M-ary quantized data in the presence of Byzantine attacks," *IEEE Trans. Signal Proc.*, vol. 62, no. 10, pp. 2681–2695, May 2014.
[18] C. Yao, P.-N. Chen, T.-Y. Wang, Y. S. Han, and P. K. Varshney, "Performance analysis and code design for minimum Hamming distance fusion in wireless sensor networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1716–1734, May 2007.
[19] M. J. Neely, "Distributed stochastic optimization via correlated scheduling," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 759–772, Apr. 2016.
[20] P.-N. Chen, Y. S. Han, H.-Y. Lin, and P. K. Varshney, "Optimal Byzantine attack for distributed inference with M-ary quantized data," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 10–15, 2016, pp. 2474–2478.
[21] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, 4th ed. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2007.

Hsuan-Yin Lin (S'09–M'13–SM'19) received his B.S. major degree in electrical engineering and minor degree in mathematics from National Tsing-Hua University (NTHU), Taiwan, in 2007, and his M.S. degree and Ph.D. degree in electrical and computer engineering from National Chiao Tung University (NCTU), Taiwan, in 2008 and 2013, respectively. During January to October 2012, Dr. Lin was a visiting scholar of the Information Theory and Coding (ITC) Group at the Department of Information and Communication Technologies of Universitat Pompeu Fabra, Barcelona, Spain. From late 2014 to 2016, Dr. Lin was a visiting scholar at CYSEC, TU Darmstadt, Germany, and from December 2016 to 2018, he was a postdoctoral research fellow at Simula UiB, Bergen, Norway. Currently, he is a research scientist at Simula UiB, Bergen, Norway. In 2014, Dr. Hsuan-Yin Lin was awarded the Honor Membership of the Phi Tau Phi Scholastic Honor Society of the Republic of China (Taiwan) and the New Partnership Program for the Connection to the Top Labs in the World (subsidized by the Ministry of Science and Technology, Taiwan). His research interests include finite blocklength information theory, privacy-preserving technologies, coding in distributed storage systems, quantum error correcting codes, inference security and target localization in wireless sensor networks, and scheduling in millimeter-wave cellular networks.

Po-Ning Chen (S'93–M'95–SM'01) was born in Taipei in 1963. He received the B.S. and M.S. degrees in electrical engineering from National Tsing-Hua University (NTHU), Taiwan, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 1994.

From 1985 to 1987, he was with NTHU Image Processing Laboratory, where he worked on the recognition of printed Chinese characters. In 1989, he was with Star Tech. Inc., where he focused on the development of fingerprint recognition systems. Upon the reception of his Ph.D. degree in 1994, he jointed Wan Ta Technology Inc. as a vice general manager, manufacturing products for retail point-of-sale systems. In 1995, he became a research staff at Advanced Technology Center (ATC), Computer and Communication Laboratories (CCL), Industrial Technology Research Institute (ITRI) in Taiwan, where he led a project on Java-based Network Managements. Since 1996, he has been an associate professor with the Department of Communications Engineering, National Chiao-Tung University (NCTU), Taiwan, and became a full professor in 2001. He was elected to be the Chair of the IEEE Communications Society (ComSoc) Taipei Chapter in 2006, and the IEEE ComSoc Taipei Chapter received the IEEE ComSoc Chapter Achievement Awards (CAA) and IEEE ComSoc Chapter of the Year (CoY) in the following year. During 2007–2009, he has served as the chairman of the Department of Communications Engineering, NCTU. During 2012–2015, he was the Associate Chief Director of NCTU Microelectronics and Information Systems Research Center, Taiwan. In 2017, he became the Associate Dean of College of Electrical and Computer Engineering, NCTU.

Dr. Chen received the Annual Research Awards from Taiwan National Science Council five years in a row since 1996. His notes for the course of Communication Networks Laboratory have been awarded as the Annual Best Teaching Materials for Communications Education by the Ministry of Education (MoS), Taiwan, in 1998. He received the Award for Junior Research Investigators, Academia Sinica, in 2000. He has been selected as NCTU Outstanding Tutor Teacher in 2002, 2013, and 2014. He was also the recipient of the Distinguished Teaching Award from the College of Electrical and Computer Engineering, NCTU, in 2003 and 2014. His research interests lie in information and coding theory, large deviations theory, distributed detection, and sensor networks.

Yunghsiang S. Han (S'90–M'93–SM'08–F'11) was born in Taipei, Taiwan, 1962. He received B.Sc. and M.Sc. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and a Ph.D. degree from the School of Computer and Information Science, Syracuse University, Syracuse, NY, in 1993. He was from 1986 to 1988 a lecturer at Ming-Hsin Engineering College, Hsinchu, Taiwan. He was a teaching assistant from 1989 to 1992, and a research associate in the School of Computer and Information Science, Syracuse University from 1992 to 1993. He was, from 1993 to 1997, an Associate Professor in the Department of Electronic Engineering at Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. He was with the Department of Computer Science and Information Engineering at National Chi Nan University, Nantou, Taiwan from 1997 to 2004. He was promoted to Professor in 1998. He was a visiting scholar in the Department of Electrical Engineering at University of Hawaii at Manoa, HI from June to October 2001, the SUPRIA visiting research scholar in the Department of Electrical Engineering and Computer Science and CASE center at Syracuse University, NY from September 2002 to January 2004 and July 2012 to June 2013, and the visiting scholar in the Department of Electrical and Computer Engineering at University of Texas at Austin, TX from August 2008 to June 2009. He was with the Graduate Institute of Communication Engineering at National Taipei University, Taipei, Taiwan from August 2004 to July 2010. From August 2010 to January 2017, he was with the Department of Electrical Engineering at National Taiwan University of Science and Technology as Chair Professor. Now he is with School of Electrical Engineering & Intelligentization at Dongguan University of Technology, China. He is also a Chair Professor at National Taipei University from February 2015. His research interests are in error-control coding, wireless networks, and security.

Dr. Han was a winner of the 1994 Syracuse University Doctoral Prize and a Fellow of IEEE. One of his papers won the prestigious 2013 ACM CCS Test-of-Time Award in cybersecurity.

Pramod K. Varshney (S'72–M'77–SM'82–F'97–LF'18) was born in Allahabad, India, in 1952. He received the B.S. (Hons.) degree in electrical engineering and computer science and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana–Champaign, USA, in 1972, 1974, and 1976, respectively, and the Doctor of Engineering (Hons.) degree from Drexel University, Philadelphia, PA, USA, in 2014. Since 1976, he has been with Syracuse University, Syracuse, NY, USA, where he is currently a Distinguished Professor with the electrical engineering and computer science and the Director with the Center for Advanced Systems and Engineering. As the Director of CASE, he is responsible for technology transition of university expertise to make economic impact in the high-tech economy of New York. He served as an Associate Chair of the department from 1993 to 1996. He is also an Adjunct Professor of radiology with Upstate Medical University, Syracuse, NY, USA. He has published extensively. He is the author of Distributed Detection and Data Fusion (New York, NY, USA: Springer-Verlag, 1997). He was a James Scholar, a Bronze Tablet Senior, and a Fellow while at the University of Illinois. His current research interests include distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, physical layer security, and machine learning.

He is a member of Tau Beta Pi. He was a recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. In 1997, he was elected to the grade of Fellow of the IEEE for his contributions in the area of distributed detection and data fusion. In 2000, he was also the recipient of the Third Millennium Medal from the IEEE and the Chancellor's Citation for exceptional academic achievement at Syracuse University, the IEEE 2012 Judith A. Resnik Award, the ECE Distinguished Alumni Award from the University of Illinois in 2015, and the ISIF's Yaakov Bar-Shalom Award for a Lifetime of Excellence in Information Fusion in 2018. He was the Guest Editor of the Special Issue on Data Fusion of the IEEE PROCEEDINGS in 1997. He is on the Editorial Board of the Journal of Advances in Information Fusion and has served on the Editorial Boards of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING MAGAZINE. He was the President of International Society of Information Fusion in 2001.