<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *IEEE Journal on Selected Topics in Signal Processing*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Li, Z., Oechtering, T. (2015)

Privacy-Aware Distributed Bayesian Detection

IEEE Journal on Selected Topics in Signal Processing, 9(7): 1345-1357

<https://doi.org/10.1109/JSTSP.2015.2429123>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-165346>

Privacy-Aware Distributed Bayesian Detection

Zuxing Li, *Student Member, IEEE*, and Tobias J. Oechtering, *Senior Member, IEEE*

Abstract—We study the eavesdropping problem in the remotely distributed sensing of a privacy-sensible hypothesis from the Bayesian detection perspective. We consider a parallel distributed detection network where remote decision makers independently make local decisions defined on finite domains and forward them to the fusion center which makes the final decision. An eavesdropper is assumed to intercept a specific set of local decisions to make also a guess on the hypothesis. We propose a novel Bayesian detection-operational privacy metric given by the minimal achievable Bayesian risk of the eavesdropper. Further, we introduce two privacy-aware distributed Bayesian detection formulations, namely the privacy-constrained distributed Bayesian detection problem and the privacy-concerned distributed Bayesian detection problem where the detection performance is optimized under a privacy guarantee constraint and a weighted sum objective of the detection performance and privacy risk is minimized respectively. For an optimal privacy-aware distributed Bayesian detection design, the optimal decision strategy of employing a deterministic likelihood test or a randomized strategy thereof is identified. Further, it is shown that equivalent problems of different formulations always exist and lead to the same optimal privacy-aware distributed Bayesian detection design. The results are illustrated and discussed by numerical examples. The idea of privacy-aware distributed Bayesian detection design provides a novel solution to realize future trustworthy Internet of Things applications.

Index Terms—Internet of Things, parallel distributed detection, person-by-person optimality, physical-layer secrecy.

I. INTRODUCTION

The idea of Internet of Things (IoT) envisions a world where everything will be connected. The growing information flows make the life more convenient but also increase the risk of exposing our privacy. Therefore, the privacy issue of IoT attracts much attention nowadays. In [1], it summarized and classified privacy threats as internal attacks or external attacks. An external attack is caused by unauthorized participants of the system and can be made in a passive or an active way. Generally, eavesdropping, jamming, interfering, and power exhausting are external threats. On the contrary, compromised authorized nodes would be the major internal threats. Compared with external attacks, internal threats are more difficult to detect and raise more serious security challenges, e.g., stealing secrets from the encrypted data, generating wrong information, reporting normal nodes as compromised nodes, breaching routing, and colliding with other nodes.

The work was supported in part by the Strategic Research Agenda Program, Information and Communication Technology - The Next Generation (SRA ICT - TNG) through the Swedish Government and in part by the Swedish Research Council (VR) under Grant E0628201.

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Zuxing Li and Tobias J. Oechtering are with the School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, 100 44, Sweden. e-mail: zuxing@kth.se; oech@kth.se.

In face of a variety of privacy threats, a large amount of fruitful works have been done. As a conventional security tool, cryptography was reported in some papers [2] to protect the private data. Although it is an effective method, cryptography is not always applicable because of its demands of high computation capability, high power consumption, and complicated key management. Some studies investigated the routing security issue in multi-hop systems, e.g., the reputation-based scheme [3] and broadcast authentication [4].

All these aforementioned technologies employ additive security functionality blocks on the upper layers. Another idea is to realize *inherent* privacy directly in the physical-layer design and leads to a privacy-per-design approach. In [5], Shannon first pointed out that perfectly secure data transmission can be realized if the key is as long as the encrypted message. The study on the wire-tap channel [6] derives its secrecy capacity. Based on the theory of wire-tap channel, people have developed privacy schemes, such as artificial noise [7] and cooperative jamming [8]. Recently, secure data compression in source coding also attracts much attention [9].

Most physical-layer secrecy works focus on the data transmission phase. However, the physical layer of an IoT application can be modeled in multiple ways. Local data processing is the core function of many proposed IoT applications, e.g., smart house and personal health-monitoring systems. In a smart house, sensors collect raw data, process it, and forward local results to a terminal which makes the final conclusion for the user to monitor the house. Such a local data processing can be modeled as a distributed detection system.

Distributed detection problems have been widely studied. In [10], [11], they summarized classic problems in different formulations, such as Bayesian, Neyman-Pearson, or information-theoretic criterion, parallel, serial, or tree topology, one-shot or sequential processing, hard or soft decision, and so on. Extensions of the basic distributed detection concepts were reviewed in [12].

When the privacy is taken into account, these established theories and results have to be modified. There are two pioneering works [13], [14] making efforts in this direction. In [13], it assumed that the eavesdropper is only interested in the state of data transmission between remote decision makers and the fusion node. However, an eavesdropper in reality can be more aggressive and also aim to intercept the transmitted information. In a more practical eavesdropping scenario, with an asymptotic assumption of an infinite number of remote decision makers, and using an information-theoretic privacy metric, it has been proved in [14] that a likelihood-ratio test (LRT) is an asymptotically optimal decision strategy of a remote decision maker under a constraint on the eavesdropper's performance. Whereas, the authors did not provide a hint to design the asymptotically optimal privacy-constrained system.

Some recent researches devised privacy schemes based on the assumption that the eavesdropper is uninformed of certain parameters. In [15], [16], they studied stochastic encryption methods where sensors intentionally generate errors to confuse the eavesdropper which is uninformed of the statistic of the induced errors. Similarly, works [17], [18] proposed channel aware encryption methods to design the transmission scheme between the sensors and fusion node based on their channel states which are not known by the eavesdropper. Besides the passive eavesdropping problems, active adversaries are also studied. Among them, the problem of Byzantines in the distributed detection network gains more attention recently [19], [20]. In [21], we studied the privacy-constrained parallel distributed Neyman-Pearson detection problem where the detection performance and eavesdropping privacy risk are evaluated by the Neyman-Pearson detection-operational metrics. For an optimal privacy-constrained distributed Neyman-Pearson detection design, it is sufficient to consider deterministic LRTs for the remote decision makers while a randomized fusion strategy might be needed.

The privacy issue is also addressed in the distributed estimation context. Some works employed the similar method of stochastic encryption, e.g., a stochastic cipher was utilized in [22] to protect privacy in a network where both the fusion center and eavesdropper make maximum-likelihood estimations. In [23], the optimal power allocation scheme was studied in a decentralized minimum mean square error estimation susceptible to eavesdropping.

Besides privacy problems in distributed detections and estimations, the statistical inference risks have been discussed in several other contexts. In particular, research on the differential privacy has gained much interest recently. The differential privacy problem [24] was studied to guarantee the statistical privacy by sanitizing mechanisms when an adversary observes two neighbor databases. In [25], the degradation of the differential privacy level under adaptive interactions was characterized. In [26], for any statistical estimator and input distribution satisfying a regularity condition, it was proved that there is a differentially private estimator with the same asymptotic output distribution. In [27], the methods were developed to approximate a filter by its differentially private version. A survey of the differential privacy in machine learning was given in [28]. Relations between different formalisms for statistical inference privacy were discussed in [29].

In this work, we study the physical-layer secrecy problem of a distributed detection¹ network susceptible to an eavesdropper, which leads to a privacy-aware system design. Two privacy-aware formulations are considered here: the privacy-constrained distributed detection and privacy-concerned distributed detection. Different from the asymptotically optimal design in [14], we attempt to design the optimal privacy-aware system from the Bayesian detection-operational perspective. The eavesdropper in this study is assumed to be informed of the distributed detection system and therefore can raise severer privacy challenges than the eavesdropping problems

Fig. 1. The model of a parallel distributed detection network with an eavesdropper. It consists of an m -ary hypothesis H , n remote decision makers $\{DM_1, \dots, DM_n\}$, a fusion center FC, and an eavesdropper EVE. Each remote decision maker DM_i ($i \in \{1, \dots, n\}$) has a continuous random observation Y_i which is conditionally independent from other remote decision makers' observations given h . Based on its observation, each remote decision maker independently makes its m_i -ary local decision U_i and transmits the local decision to the fusion center through a rate-limited error-free link. Based on local decisions of all remote decision makers, the fusion center makes the final m -ary fusion decision U_F to infer the hypothesis. An eavesdropper is assumed to intercept a set of local decisions and makes its own m -ary decision U_E to infer the hypothesis.

studied in [15]–[18]. There are three questions, which we want to answer in this paper. First, what is a reasonable privacy metric in a Bayesian detection formulation? Second, can we characterize or classify the decision strategies in the optimal privacy-aware distributed detection design? Third, how can we obtain the optimal privacy-aware distributed detection design? This work is extended from [30] where we considered a privacy-constrained distributed Bayesian detection with particular problem settings of a binary hypothesis, binary decisions, and only one intercepted remote decision maker.

The remaining of this paper is organized as follows: Section II describes the studied model of a parallel distributed detection network with an eavesdropper; Section III formulates and discusses the privacy-aware distributed Bayesian detection problems; Section IV illustrates the trade-off between the detection performance and privacy risk in numerical examples; and the conclusion is given in Section V. In the following, we will represent a random variable by a capital letter, its realization by the corresponding lower-case letter, and its definition domain set by the corresponding calligraphic letter.

II. DISTRIBUTED DETECTION SUSCEPTIBLE TO AN EAVESDROPPER

In this work, we study the eavesdropping problem in the distributed detection network and develop a concept of privacy-aware distributed detection design. In the following, we will show the concept of privacy-aware distributed detection design by considering the parallel model shown in Figure 1.

The parallel model consists of an m -ary hypothesis H , n remote decision makers $\{DM_i : i \in \{1, \dots, n\}\}$, a fusion center FC, and an eavesdropper EVE. The hypothesis H is defined on the set $\mathcal{H} = \{0, \dots, m-1\}$ and is generated according to the prior probability $p_H(h)$. Each remote decision maker DM_i independently makes an m_i -ary local decision

¹We use the term detection including also m -ary hypothesis testing for $m > 2$, which is also known as classification problem.

U_i defined on the set $\mathcal{U}_i = \{0, \dots, m_i - 1\}$ based on its continuous² random observation Y_i which is defined on the set $\mathcal{Y}_i$. Here, we assume that any likelihood $f_{Y_i|H}(y_i|h)$ contains no point masses of probability and all local observations are conditionally independent given the hypothesis,

$$f_{Y_1, \dots, Y_n|H}(y_1, \dots, y_n|h) = \prod_{i=1}^n f_{Y_i|H}(y_i|h).$$

All local decisions are transmitted to the fusion center through error-free rate-limited links. Based on these local decisions, the fusion center makes the final m -ary fusion decision U_F to infer the hypothesis H . It is assumed that the eavesdropper intercepts a set of local decisions $\{U_i : i \in \mathcal{I}, \mathcal{I} \subseteq \{1, \dots, n\}\}$ and makes its own m -ary decision U_E to infer the hypothesis as well. U_F and U_E are defined on the set $\mathcal{H}$, i.e., $\mathcal{U}_F = \mathcal{U}_E = \mathcal{H}$. In this work, we allow independent randomized decision strategies [11] of DM_i , fusion center, and eavesdropper and denote them by ϕ_i , ϕ_F , and ϕ_E respectively as

$$\begin{aligned} U_i &= \phi_i(y_i), \\ U_F &= \phi_F(u_1, \dots, u_n), \\ U_E &= \phi_E(\{u_i : i \in \mathcal{I}\}). \end{aligned} \quad (1)$$

Note that the observations and decisions in (1) are denoted by lower-case letters and capital letters respectively. It means that each node can make a random decision given a deterministic observation based on a conditional p.m.f. $p_{U_i|Y_i}(u_i|y_i)$, $p_{U_F|U_1, \dots, U_n}(u_F|u_1, \dots, u_n)$, or $p_{U_E|\{U_i : i \in \mathcal{I}\}}(u_E|\{u_i : i \in \mathcal{I}\})$. We use Φ_i , Φ_F , and Φ_E to denote sets of all randomized decision strategies of DM_i , fusion center, and eavesdropper, i.e., $\phi_i \in \Phi_i$, $\phi_F \in \Phi_F$, and $\phi_E \in \Phi_E$, and define their subsets Φ_i^D , Φ_F^D , and Φ_E^D , each of which consists of all deterministic decision strategies of a node. Denote the numbers of deterministic strategies in Φ_F^D and Φ_E^D by d_F and d_E . For a node, any randomized decision strategy is realized by a random implementation of deterministic strategies. In addition, random implementations of deterministic decision strategies by different nodes are statistically independent [11].

III. PRIVACY-AWARE DISTRIBUTED BAYESIAN DETECTION PROBLEMS

In this section, we will answer the three questions in Section I by proposing a novel Bayesian detection-operational privacy metric, formulating and studying the privacy-aware distributed Bayesian detection problems, and extending the standard algorithm to obtain the optimal privacy-aware distributed Bayesian detection designs.

To develop the Bayesian framework, important results and concepts of the distributed Bayesian detection problem are briefly presented first. Denote the non-negative detection cost of the fusion center to make a decision u_F when the hypothesis presents h by $c_F(u_F, h)$. In the Bayesian formulation, it is assumed that the hypothesis prior probability $p_H(h)$, likelihoods of the local observations $\{f_{Y_i|H}(y_i|h) : i \in \{1, \dots, n\}\}$, and detection cost assignment of the fusion center are known. If

there is no eavesdropping threat in Figure 1, the distributed Bayesian detection problem is to find the optimal design $\{\phi_1^*, \dots, \phi_n^*, \phi_F^*\}$ which minimizes the Bayesian risk of the fusion center r_F as

$$\min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} r_F, \quad (2)$$

where the Bayesian risk r_F is given as follows

$$r_F = \sum_{u_F \in \mathcal{U}_F, h \in \mathcal{H}} c_F(u_F, h) p_{U_F, H}(u_F, h). \quad (3)$$

In [11], the author shows that it is sufficient to consider deterministic likelihood tests for $\{\phi_1^*, \dots, \phi_n^*, \phi_F^*\}$ in the form of

$$\phi_i^*(y_i) = \arg \min_{u_i \in \mathcal{U}_i} \sum_{h \in \mathcal{H}} a_i^{u_i, h} f_{Y_i|H}(y_i|h), \quad (4)$$

$$\begin{aligned} &\phi_F^*(u_1, \dots, u_n) \\ &= \arg \min_{u_F \in \mathcal{U}_F} \sum_{h \in \mathcal{H}} a_F^{u_F, h} p_{U_1, \dots, U_n|H}(u_1, \dots, u_n|h), \end{aligned} \quad (5)$$

where $a_i^{u_i, h}$, $a_F^{u_F, h}$ are constant coefficients. In the particular case of binary hypothesis and decisions ($m = m_i = 2, \forall i$), likelihood tests in (4) and (5) reduce to LRTs.

Define $p_i^{u_i, h} = p_{U_i|H}(u_i|h)$. The operation of any randomized decision strategy of a remote decision maker DM_i can be represented by a set of corresponding conditional probabilities as

$$(p_i^{1,0}, \dots, p_i^{m_i-1,0}, \dots, p_i^{1,m-1}, \dots, p_i^{m_i-1,m-1}),$$

which can be also seen as a point $\mathbf{p}_i$ in the $(m_i m - m)$ -dimension operation region $\mathcal{R}_i$. In the following, we will always use $\partial \cdot$ to denote a boundary of a set, e.g., the boundary of $\mathcal{R}_i$ is denoted by $\partial \mathcal{R}_i$. Properties of $\mathcal{R}_i$ are listed as follows:

- 1) $\mathcal{R}_i$ is a convex set.
- 2) Any point on $\partial \mathcal{R}_i$ can be achieved by a deterministic likelihood test in the form of (4).

The first property is self-evident since randomized decision strategies are allowed. The second property can be proved by the same method of using Lagrange multipliers as shown in [10].

Remark 1: Based on these properties, any point within $\mathcal{R}_i$ can be achieved by a randomized decision strategy of two³ deterministic likelihood tests.

Figure 2 illustrates an operation region of DM_i if H and U_i are binary random variables ($m = m_i = 2$). In the particular case of binary H and U_i , the detection and false-alarm probabilities are defined as $p_i^D = p_i^{1,1} = p_{U_i|H}(1|1)$ and $p_i^F = p_i^{1,0} = p_{U_i|H}(1|0)$.

The method of person-by-person optimization (PBPO) [31] is a standard approach to design the parallel distributed Bayesian detection network although it does not guarantee the global optimality. Consider to optimize ϕ_i given other decision and fusion strategies in the problem:

$$\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F) = \arg \min_{\phi_i \in \Phi_i} r_F. \quad (6)$$

²The concepts can be also developed for discrete random observations. But then, continuity of a remote decision maker's operation region boundary requires randomized decision strategies.

³In the case of having a discrete random observation, any point within $\mathcal{R}_i$ can be achieved by a randomized decision strategy of $m_i m - m + 1$ deterministic strategies based on the Carathéodory's Theorem.

Fig. 2. Illustration of $\mathcal{R}_i$ if H and U_i are binary random variables, $Y_i|h = 0 \sim \mathcal{N}(0, 1)$, and $Y_i|h = 1 \sim \mathcal{N}(1, 1)$.

It is sufficient to consider a deterministic likelihood test in the form of (4) for $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ where each constant coefficient $a_i^{u_i, h}$ is determined by fixed decision and fusion strategies as shown in the appendix. Another problem is to optimize ϕ_F given decision strategies of all remote decision makers as

$$\phi_F^\Delta(\phi_1, \dots, \phi_n) = \arg \min_{\phi_F \in \Phi_F} r_F. \quad (7)$$

Similarly, it is sufficient to consider a deterministic likelihood test in the form of (5) for $\phi_F^\Delta(\phi_1, \dots, \phi_n)$ where each constant coefficient $a_F^{u_F, h}$ is determined by fixed decision strategies of remote decision makers as shown in the appendix. The standard PBPO algorithm is to approach the optimal distributed Bayesian detection design by iteratively refining decision and fusion strategies one by one. When refining ϕ_i , the person-by-person optimal decision strategy is the deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$. When refining ϕ_F , the person-by-person optimal fusion strategy is the deterministic likelihood test $\phi_F^\Delta(\phi_1, \dots, \phi_n)$. Given an initial system design, this algorithm always converges since the Bayesian risk r_F is monotonically decreased when a decision strategy is refined. Since it does not necessarily converge to the globally optimal distributed Bayesian detection design, the algorithm has to be implemented with a large number of initial system designs to approach the optimal distributed Bayesian detection design in practice.

A. Eavesdropper and Privacy Metric

1) *Informed and Greedy Eavesdropper*: Suppose that there is an eavesdropper in the system. The objective of the eavesdropper is to infer the hypothesis which results in the privacy risk of information leakage. Since the behavior of the eavesdropper may be unknown, we make assumptions on it based on our own privacy interests. This will lead to a concrete system privacy level which the system designer can ensure.

- 1) DM-FC links in the studied distributed detection network can be open-access or protected, i.e., some links are accessible for the eavesdropper and others are not. The set $\mathcal{I}$ consists of the indices of remote decision makers whose local decisions are intercepted by the eavesdropper.

- 2) The eavesdropper is informed of the distributed detection system, i.e., the hypothesis prior probability $p_H(h)$, likelihoods of the local observations $\{f_{Y_i|H}(y_i|h) : i \in \{1, \dots, n\}\}$, detection cost assignment of the fusion center $c_F(u_F, h)$, and all decision strategies are known by the eavesdropper.
- 3) We are interested in the Bayesian risk of the eavesdropper r_E . Therefore, we assign a non-negative detection cost of the eavesdropper $c_E(u_E, h)$ in the case that the eavesdropper makes a decision u_E when the hypothesis h is presented. Then,

$$r_E = \sum_{u_E \in \mathcal{U}_E, h \in \mathcal{H}} c_E(u_E, h) p_{U_E, H}(u_E, h). \quad (8)$$

In addition, the eavesdropper is assumed to be greedy to infer the hypothesis, i.e., the informed eavesdropper always employs the optimal decision strategy ϕ_E^* which minimizes its own Bayesian risk.

2) *Privacy Metric*: The idea of privacy-aware distributed Bayesian detection design addresses both objectives: the detection performance and privacy risk. The detection performance is measured by the Bayesian risk r_F . Based on our privacy interest and assumptions of the informed and greedy eavesdropper, we propose a novel Bayesian detection-operational privacy metric as follows

$$r_E^{\min} = \min_{\phi_E \in \Phi_E} r_E, \quad (9)$$

which represents the minimal achievable Bayesian risk of the eavesdropper. There are two properties of this privacy metric:

- 1) Since a randomized strategy does not help to improve the Bayesian detection performance [11], the optimal decision strategy of the eavesdropper $\phi_E^* = \arg \min_{\phi_E \in \Phi_E} r_E$ is always deterministic, i.e., $\phi_E^* \in \Phi_E^D$.
- 2) There is an upper bound γ on $r_E^{\min}$:

$$r_E^{\min} \leq \gamma = \min_{u_E \in \mathcal{U}_E} \sum_{h \in \mathcal{H}} c_E(u_E, h) p_H(h), \quad (10)$$

where each term $\sum_{h \in \mathcal{H}} c_E(u_E, h) p_H(h)$ is the Bayesian risk of the eavesdropper to always make a decision $u_E \in \mathcal{U}_E$ regardless of any intercepted local decisions.

The first property implies that we only need to discuss deterministic decision strategies for the eavesdropper. Since the intercepted local decisions of the eavesdropper and its own decision are discrete random variables, there are $d_E = m(\prod_{i \in \mathcal{I}} m_i)$ deterministic decision strategies as the optimal decision strategy candidates of the eavesdropper. The second property shows that the highest privacy level of the system is equivalent to let the eavesdropper infer the hypothesis based on its knowledge of the hypothesis prior probability only.

In the following, we show two privacy-aware formulations, the privacy-constrained design and privacy-concerned design. The privacy-constrained problem is formulated to optimize the detection performance under a privacy guarantee constraint. The privacy-concerned problem minimizes a weighted sum objective of the detection performance and privacy risk. Then, we answer the other two research questions raised in Section I.

B. Privacy-Constrained Distributed Bayesian Detection

We propose the privacy-constrained design which optimizes the detection performance under a constraint to guarantee a certain system privacy level.

1) *Privacy-Constrained Problem Formulation:* Using the proposed Bayesian detection-operational performance and privacy metrics, we formulate the privacy-constrained distributed Bayesian detection problem as

$$\begin{aligned} \min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} \quad & r_F \\ \text{s.t.} \quad & r_E^{\min} \geq \beta, \end{aligned} \quad (11)$$

where β represents the guaranteed system privacy level. Because of the upper bound of $r_E^{\min}$ shown in (10), the privacy-constrained optimization problem is feasible iff $\beta \leq \gamma$.

Compared with the distributed Bayesian detection problem (2), the proposed privacy-constrained problem has the same main objective to optimize the detection performance with an additional operational privacy constraint condition imposed.

2) *Properties of the Optimal Privacy-Constrained Design:* To answer the second research question, properties of the optimal privacy-constrained Bayesian design need to be identified. However, it is difficult to solve (11) and to obtain the optimal privacy-constrained decision strategies $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\}$ directly. Alternatively, we can derive properties of the optimal privacy-constrained design by exploiting the person-by-person optimality argument which states that properties of the person-by-person optimal privacy-constrained strategy are necessary to be satisfied by the optimal privacy-constrained strategy. That is because an optimal privacy-constrained strategy must be the person-by-person optimal privacy-constrained strategy given other decision strategies in the optimal privacy-constrained distributed Bayesian detection design.

Proposition 1: For the fusion center, it is sufficient to consider a deterministic likelihood test in the form of (5) for ϕ_F^{P*} in the optimal privacy-constrained distributed Bayesian detection design.

Proof: Suppose that the decision strategies of remote decision makers are fixed. Let us consider the following optimization problem

$$\begin{aligned} \min_{\phi_F \in \Phi_F} \quad & r_F \\ \text{s.t.} \quad & r_E^{\min} \geq \beta. \end{aligned} \quad (12)$$

Since the final fusion decision U_F is not intercepted, the optimization of ϕ_F is not constrained by the privacy condition. In this case, the above problem is equivalent to (7) and it is sufficient to consider the deterministic likelihood test $\phi_F^{\Delta}(\phi_1, \dots, \phi_n)$ for the person-by-person optimal privacy-constrained fusion strategy here. Therefore, we conclude the property of ϕ_F^{P*} as stated in Proposition 1. ■

Remark 2: Since the observations and decision of the fusion center are discrete random variables, there are $d_F = m(\prod_{i=1}^n m_i)$ deterministic fusion strategies. In practice, a simple method to obtain ϕ_F^{P*} is brute-force search of the optimal deterministic fusion strategy which minimizes r_F through all deterministic fusion strategies in Φ_F^D .

Proposition 2: For a remote decision maker DM_i whose local decision U_i is not intercepted, i.e., $i \notin \mathcal{I}$, it is sufficient

to consider a deterministic likelihood test in the form of (4) for ϕ_i^{P*} in the optimal privacy-constrained distributed Bayesian detection design.

Proof: Given other decision and fusion strategies, consider the following optimization problem of ϕ_i .

$$\begin{aligned} \text{For } i \notin \mathcal{I}: \quad & \min_{\phi_i \in \Phi_i} r_F \\ \text{s.t.} \quad & r_E^{\min} \geq \beta. \end{aligned} \quad (13)$$

Since the local decision U_i is not intercepted, the optimization of ϕ_i is not constrained by the privacy condition. In this case, the above problem is equivalent to (6) and it is sufficient to consider the deterministic likelihood test $\phi_i^{\Delta}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ for the person-by-person optimal privacy-constrained decision strategy here. Then, we have the property of ϕ_i^{P*} with $i \notin \mathcal{I}$ as stated in Proposition 2. ■

In the following, we discuss the property of ϕ_i^{P*} with $i \in \mathcal{I}$. Since the local decision U_i is intercepted by the eavesdropper, the privacy constraint condition takes effect in the following optimization problem of ϕ_i when all other decision and fusion strategies are fixed.

$$\begin{aligned} \text{For } i \in \mathcal{I}: \quad & \min_{\phi_i \in \Phi_i} r_F \\ \text{s.t.} \quad & r_E^{\min} \geq \beta. \end{aligned} \quad (14)$$

As discussed before, any randomized decision strategy ϕ_i can be represented by its corresponding conditional probabilities as a point $\mathbf{p}_i$ in the $(m_i m - m)$ -dimension convex operation region $\mathcal{R}_i$. In the problem (14), the Bayesian risk r_F is an affine function of $\mathbf{p}_i$ as

$$\begin{aligned} r_F &= b_i^{1,0} p_i^{1,0} + \dots + b_i^{m_i-1, m-1} p_i^{m_i-1, m-1} + b_i \\ &= \mathbf{b}_i \mathbf{p}_i^{\top} + b_i, \end{aligned} \quad (15)$$

where $\mathbf{b}_i = (b_i^{1,0}, \dots, b_i^{m_i-1, m-1})$ and b_i are constant coefficients determined by the fixed decision and fusion strategies, detection cost assignment of the fusion center, and prior probability as shown in the appendix. The privacy constraint condition $r_E^{\min} \geq \beta$ means $r_E \geq \beta$ regardless of the employed decision strategy ϕ_E , i.e., $r_E(\phi_E) \geq \beta, \forall \phi_E$. The previous analysis has shown that we only need to discuss $d_E = m(\prod_{i \in \mathcal{I}} m_i)$ deterministic decision strategies for the eavesdropper. Therefore, the privacy constraint can be further rewritten as $r_E(\phi_E^{k_E}) \geq \beta, \forall \phi_E^{k_E}$, where $\phi_E^{k_E} \in \Phi_E^D$ denotes a deterministic decision strategy of the eavesdropper and $k_E \in \{1, \dots, d_E\}$. Given a deterministic strategy $\phi_E^{k_E}$, r_E is also an affine function of $\mathbf{p}_i$ as

$$\begin{aligned} r_E &= b_i^{k_E, 1,0} p_i^{1,0} + \dots + b_i^{k_E, m_i-1, m-1} p_i^{m_i-1, m-1} + b_i^{k_E} \\ &= \mathbf{b}_i^{k_E} \mathbf{p}_i^{\top} + b_i^{k_E}, \end{aligned} \quad (16)$$

where $\mathbf{b}_i^{k_E} = (b_i^{k_E, 1,0}, \dots, b_i^{k_E, m_i-1, m-1})$ and $b_i^{k_E}$ are constant coefficients determined by the given $\phi_E^{k_E}$, fixed decision and fusion strategies, detection cost assignment of the eavesdropper, and prior probability as shown in the appendix. Then, the problem in (14) with $i \in \mathcal{I}$ can be equivalently represented as

$$\begin{aligned} \min_{\mathbf{p}_i \in \mathcal{R}_i} \quad & \mathbf{b}_i \mathbf{p}_i^{\top} + b_i \\ \text{s.t.} \quad & \mathbf{b}_i^{k_E} \mathbf{p}_i^{\top} + b_i^{k_E} \geq \beta, \forall \phi_E^{k_E}, k_E \in \{1, \dots, d_E\}. \end{aligned} \quad (17)$$

Fig. 3. Illustration of $\mathcal{R}_i^P$ of an intercepted remote decision maker DM_i if H and U_i are binary random variables, $Y_i|h = 0 \sim \mathcal{N}(0, 1)$, $Y_i|h = 1 \sim \mathcal{N}(1, 1)$, and there are four privacy constraint regions taking effect.

Instead of solving the optimal decision strategy of the problem (14), we are going to find the corresponding optimal point which minimizes the affine objective function and satisfies the set of affine privacy inequalities. Define a privacy constraint region $\mathcal{P}_i^{k_E}$ and its privacy constraint boundary $\partial\mathcal{P}_i^{k_E}$ as

$$\begin{aligned}\mathcal{P}_i^{k_E} &= \left\{ \mathbf{p}_i : \mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E} \geq \beta \right\}, \\ \partial\mathcal{P}_i^{k_E} &= \left\{ \mathbf{p}_i : \mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E} = \beta \right\}.\end{aligned}\quad (18)$$

Then, the privacy-constrained operation region $\mathcal{R}_i^P$ of DM_i is defined as

$$\mathcal{R}_i^P = \left\{ \bigcap_{k_E=1}^{d_E} \mathcal{P}_i^{k_E} \right\} \cap \mathcal{R}_i. \quad (19)$$

Since $\mathcal{R}_i$ and $\mathcal{P}_i^{k_E}$ are convex sets, $\mathcal{R}_i^P$ is a convex set [32]. An example of $\mathcal{R}_i^P$ is illustrated in Figure 3. The problem (17) is to find the optimal point which minimizes the affine objective function $\mathbf{b}_i \mathbf{p}_i^\top + b_i$ in the convex privacy-constrained operation region $\mathcal{R}_i^P$. Therefore, we have the following result.

Theorem 1: According to [32], the optimal point of the problem (17) is on the boundary $\partial\mathcal{R}_i^P$ and has a supporting hyperplane in parallel with the objective function hyperplane $\mathbf{b}_i \mathbf{p}_i^\top + b_i$.

The boundary $\partial\mathcal{R}_i^P$ consists of two parts: $\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i$ and $\bigcup_{k_E=1}^{d_E} \left\{ \partial\mathcal{R}_i^P \cap \partial\mathcal{P}_i^{k_E} \right\}$. If the optimal point of the problem (17) is on $\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i$, the second property of $\mathcal{R}_i$ implies that the optimal decision strategy of the problem (14) can be a deterministic likelihood test in the form of (4). Otherwise, we need to find the optimal point which minimizes the affine objective function on the hyperplanes $\bigcup_{k_E=1}^{d_E} \left\{ \partial\mathcal{R}_i^P \cap \partial\mathcal{P}_i^{k_E} \right\}$ or sufficiently in the intersections of hyperplanes. In this case, the optimal point of the problem (17) is in the set

$$\bigcup_{k_E, k'_E \in \{1, \dots, d_E\}, k_E \neq k'_E} \left\{ \partial\mathcal{R}_i^P \cap \partial\mathcal{P}_i^{k_E} \cap \partial\mathcal{P}_i^{k'_E} \right\}. \quad (20)$$

Such an optimal intersection point can be achieved by a randomized strategy of two deterministic likelihood tests in the form of (4) which is the optimal decision strategy for the problem (14). Based on the above analysis of the person-by-person optimal privacy-constrained decision strategy of an

intercepted remote decision maker DM_i , we can make the following conclusion about ϕ_i^{P*} with $i \in \mathcal{I}$.

Theorem 2: For a remote decision maker DM_i whose local decision U_i is intercepted, i.e., $i \in \mathcal{I}$, ϕ_i^{P*} can be a deterministic likelihood test in the form of (4) or a randomized strategy of two deterministic likelihood tests in the form of (4) in the optimal privacy-constrained distributed Bayesian detection design.

Remark 3: When the optimal decision strategy of the problem (14) is a randomized decision strategy of two deterministic likelihood tests in the form (4), this optimal randomized strategy is determined by the privacy constraint condition only, i.e., it achieves a point to be an intersection of two privacy constraint hyperplanes.

3) *Extended PBPO Algorithm:* The standard PBPO algorithm is used to approach the optimal distributed Bayesian detection design without a privacy objective. Here, we extend the standard PBPO algorithm to design the optimal privacy-constrained distributed Bayesian detection system.

Given decision strategies of all remote decision makers, Remark 2 suggests that we can find the person-by-person optimal privacy-constrained fusion strategy by searching through all deterministic strategies of the fusion center.

Consider a remote decision maker DM_i whose local decision U_i is not intercepted, i.e., $i \notin \mathcal{I}$. Given other decision and fusion strategies, it is sufficient to consider the output deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ of the standard PBPO algorithm for the person-by-person optimal privacy-constrained decision strategy of the remote decision maker DM_i .

Next, consider an intercepted remote decision maker DM_i , i.e., $i \in \mathcal{I}$. Given decision and fusion strategies of all other nodes, it has been shown that the person-by-person optimal privacy-constrained decision strategy of DM_i can be a deterministic likelihood test in the form of (4) or a randomized strategy of two deterministic likelihood tests in the form of (4). In the following, we discuss the extended algorithm to refine ϕ_i of the intercepted remote decision maker DM_i in two cases.

- 1) If the point achieved by the output deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ of the standard PBPO algorithm is on the boundary $\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i$, the deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ is the person-by-person optimal privacy-constrained decision strategy of the intercepted DM_i .
- 2) Otherwise, we search the person-by-person optimal privacy-constrained decision strategy of the intercepted DM_i through deterministic likelihood tests which are in the form of (4) and achieve points in the set $\bigcup_{k_E=1}^{d_E} \left\{ \partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i \cap \partial\mathcal{P}_i^{k_E} \right\}$ and randomized strategies of two deterministic likelihood tests in the form of (4) which achieve points in the hyperplane intersection set (20). Linear programming methods [33] can be used to search through the privacy constraint hyperplane intersections.

Proof:

- 1) In the first case, the output deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ of the standard PBPO algo-

rithm achieves the optimal point in the operation region $\mathcal{R}_i$. Since the optimal point is in the subset $\mathcal{R}_i^P$, it is also optimal in $\mathcal{R}_i^P$. Therefore, $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ is the person-by-person optimal privacy-constrained decision strategy of the intercepted DM_i.

- 2) In the second case, the point achieved by the deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ is on $\partial\mathcal{R}_i \setminus \{\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i\}$ and has a supporting hyperplane in parallel with the objective function hyperplane $\mathbf{b}_i \mathbf{p}_i^\top + b_i$. Since the likelihood $f_{Y_i|H}(y_i|h)$ is assumed to contain no point masses of probability, there is only one another point on $\partial\mathcal{R}_i$ which has a supporting hyperplane in parallel with $\mathbf{b}_i \mathbf{p}_i^\top + b_i$ but maximizes the objective function. For the problem (17), the optimal point to minimize the affine objective function is on $\partial\mathcal{R}_i^P$ and has a supporting hyperplane in parallel with $\mathbf{b}_i \mathbf{p}_i^\top + b_i$. Thus, points in $\{\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i\} \setminus \bigcup_{k_E=1}^{d_E} \{\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i \cap \partial\mathcal{P}_i^{k_E}\}$ cannot be the optimal point candidates. For the other part of $\partial\mathcal{R}_i^P$ on the hyperplanes $\bigcup_{k_E=1}^{d_E} \{\partial\mathcal{R}_i^P \cap \partial\mathcal{P}_i^{k_E}\}$, it is sufficient to consider the intersections of the privacy constraint hyperplanes in the set (20) as the optimal point candidates. ■

Algorithm 1 is the extended PBPO algorithm to design the optimal privacy-constrained distributed Bayesian detection network.

C. Privacy-Concerned Distributed Bayesian Detection

The privacy-constrained problem provides a guarantee of an absolute system privacy level and optimizes the detection performance. A different privacy-aware idea in the proposed privacy-concerned formulation is to address both objectives of improving the detection performance and suppressing the privacy risk with different weights.

1) *Privacy-Concerned Problem Formulation:* Using the proposed Bayesian detection-operational metrics (3) and (9), the privacy-concerned distributed Bayesian detection problem is formulated as follows

$$\min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} \lambda r_F - (1 - \lambda) r_E^{\min}, \quad (21)$$

where $\lambda \in (0, 1]$ represents the concern weight of minimizing the Bayesian risk r_F and $1 - \lambda$ represents the concern weight of maximizing $r_E^{\min}$ in the system design. In this general formulation, λ cannot be set to be 0. Otherwise, when $\mathcal{I} \neq \{1, \dots, n\}$, the privacy-concerned optimization problem deals with the privacy metric $r_E^{\min}$ which only depends on decision strategies of intercepted remote decision makers. When λ decreases, we concern more about suppressing the privacy risk. On the contrary, we concern more about improving the detection performance with increasing λ .

2) *Properties of the Optimal Privacy-Concerned Design:* Similar as the study of the privacy-constrained distributed Bayesian detection problem, it is difficult to solve (21) and to obtain the optimal privacy-concerned decision strategies $\{\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}\}$ directly. Again, we derive properties of the optimal privacy-concerned distributed Bayesian detection

Algorithm 1 The Extended PBPO Algorithm to Optimize the Privacy-Constrained Distributed Bayesian Detection Design

- 1: generate a sufficiently large number of initial decision strategy sets of $\{\phi_1, \dots, \phi_n\}$
- 2: given an initial decision strategy set
- 3: **while** r_F is not minimized **do**
- 4: given $\{\phi_1, \dots, \phi_n\}$ such that $r_E^{\min} \geq \beta$
- 5: $\phi_F^{\text{temp}} \leftarrow$
search the optimal strategy through all deterministic fusion strategies in Φ_F^D
- 6: $\phi_F \leftarrow \phi_F^{\text{temp}}$
- 7: **for** $i \in \{1, \dots, n\}$ **do**
- 8: given $\{\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F\}$
- 9: $\phi_i^{\text{temp}} \leftarrow \phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$
- 10: **if** $i \notin \mathcal{I}$ **then**
 $\phi_i \leftarrow \phi_i^{\text{temp}}$
- 11: **else if** $i \in \mathcal{I}$ and $\mathbf{p}_i$ achieved by $\phi_i^{\text{temp}} \in \{\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i\}$ **then**
 $\phi_i \leftarrow \phi_i^{\text{temp}}$
- 12: **else**
 $\phi_i^{\text{temp}} \leftarrow$
search the optimal strategy through deterministic likelihood tests in the form of (4) which achieve points in $\bigcup_{k_E=1}^{d_E} \{\partial\mathcal{R}_i^P \cap \partial\mathcal{R}_i \cap \partial\mathcal{P}_i^{k_E}\}$ and randomized strategies of two deterministic likelihood tests in the form of (4) which achieve points in $\bigcup_{k_E, k'_E \in \{1, \dots, d_E\}, k_E \neq k'_E} \{\partial\mathcal{R}_i^P \cap \partial\mathcal{P}_i^{k_E} \cap \partial\mathcal{P}_i^{k'_E}\}$
 $\phi_i \leftarrow \phi_i^{\text{temp}}$
- 13: $\phi_i \leftarrow \phi_i^{\text{temp}}$
- 14: **end if**
- 15: **end for**
- 16: **end while**
- 17: compare obtained distributed detection designs corresponding to all initial decision strategy sets
- 18: **Output:** the optimal privacy-constrained distributed Bayesian detection design $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\}$

design based on the person-by-person optimality argument that properties of a person-by-person optimal privacy-concerned strategy are necessary to be satisfied by the corresponding optimal privacy-concerned strategy.

Proposition 3: For the fusion center, it is sufficient to consider a deterministic likelihood test in the form of (5) for ϕ_F^{C*} in the optimal privacy-concerned distributed Bayesian detection design.

Proof: Suppose that the decision strategies of remote decision makers are fixed and consider the following optimization problem

$$\min_{\phi_F \in \Phi_F} \lambda r_F - (1 - \lambda) r_E^{\min}. \quad (22)$$

Since the final fusion decision U_F is not intercepted, $r_E^{\min}$ in (22) has been determined by the fixed decision strategies of the intercepted remote decision makers. In this case, the above problem is equivalent to (7) and it is sufficient to consider the deterministic likelihood test $\phi_F^\Delta(\phi_1, \dots, \phi_n)$ for the person-by-person optimal privacy-concerned fusion strategy here.

Based on the person-by-person optimality argument, we obtain the property of ϕ_F^{C*} in Proposition 3. ■

Remark 4: Same as stated in Remark 2, we can find ϕ_F^{C*} by searching the optimal deterministic fusion strategy which minimizes the objective function $\lambda r_F - (1 - \lambda)r_E^{\min}$ through all deterministic fusion strategies.

Proposition 4: For a remote decision maker DM_i whose local decision U_i is not intercepted, i.e., $i \notin \mathcal{I}$, it is sufficient to consider a deterministic likelihood test in the form of (4) for ϕ_i^{C*} in the optimal privacy-concerned distributed Bayesian detection design.

Proof: Given other decision and fusion strategies, consider the following optimization problem.

$$\text{For } i \notin \mathcal{I} : \min_{\phi_i \in \Phi_i} \lambda r_F - (1 - \lambda)r_E^{\min}. \quad (23)$$

Since the local decision U_i is not intercepted, $r_E^{\min}$ in (23) has been determined by the fixed decision strategies of the intercepted remote decision makers. In this case, the above problem is equivalent to (6) and it is sufficient to consider the deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ for the person-by-person optimal privacy-concerned strategy here. Then, we obtain the property of ϕ_i^{C*} with $i \notin \mathcal{I}$ stated in Proposition 4. ■

In the following, we discuss the property of ϕ_i^{C*} of the intercepted remote decision maker DM_i , i.e., $i \in \mathcal{I}$. Given decision and fusion strategies of all other remote decision makers and the fusion center, we study the following optimization problem.

$$\text{For } i \in \mathcal{I} : \min_{\phi_i \in \Phi_i} \lambda r_F - (1 - \lambda)r_E^{\min}. \quad (24)$$

Note that ϕ_i can be represented as a point $\mathbf{p}_i$ in $\mathcal{R}_i$. The Bayesian risk r_F can be represented as the affine function of $\mathbf{p}_i$ in (15). Given a deterministic decision strategy $\phi_E^{k_E}$, r_E can be represented as the affine function of $\mathbf{p}_i$ in (16). Next, define a privacy concern halfspace $\mathcal{C}_i^{k_E, k'_E}$ and its privacy concern hyperplane $\partial \mathcal{C}_i^{k_E, k'_E}$ as

$$\begin{aligned} \mathcal{C}_i^{k_E, k'_E} &= \left\{ \mathbf{p}_i : \mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E} \leq \mathbf{b}_i^{k'_E} \mathbf{p}_i^\top + b_i^{k'_E} \right\}, \\ \partial \mathcal{C}_i^{k_E, k'_E} &= \left\{ \mathbf{p}_i : \mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E} = \mathbf{b}_i^{k'_E} \mathbf{p}_i^\top + b_i^{k'_E} \right\}, \end{aligned} \quad (25)$$

where $k_E \neq k'_E$. Then, a privacy concern region $\mathcal{R}_i^{k_E}$ is defined as

$$\mathcal{R}_i^{k_E} = \left\{ \bigcap_{k'_E \in \{1, \dots, d_E\}, k'_E \neq k_E} \mathcal{C}_i^{k_E, k'_E} \right\} \cap \mathcal{R}_i. \quad (26)$$

If we employ any decision strategy ϕ_i which achieves a point in $\mathcal{R}_i^{k_E}$, the optimal decision strategy of the eavesdropper ϕ_E^* is the deterministic strategy $\phi_E^{k_E}$. $\mathcal{R}_i^{k_E}$ has the following properties:

- 1) $\mathcal{R}_i^{k_E}$ is a convex set.
- 2) $\bigcup_{k_E=1}^{d_E} \mathcal{R}_i^{k_E} = \mathcal{R}_i$.
- 3) $\forall k_E, k'_E \in \{1, \dots, d_E\}$ with $k_E \neq k'_E$, $\{\mathcal{R}_i^{k_E} \cap \mathcal{R}_i^{k'_E}\} \subseteq \partial \mathcal{C}_i^{k_E, k'_E}$.
- 4) The boundary $\partial \mathcal{R}_i^{k_E}$ consists of two parts: $\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i$ and $\bigcup_{k'_E \in \{1, \dots, d_E\}, k'_E \neq k_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{C}_i^{k_E, k'_E}\}$.

Fig. 4. Illustration of three convex privacy concern regions of $\mathcal{R}_i$ with $i \in \mathcal{I}$ if H and U_i are binary random variables, $Y_i|h=0 \sim \mathcal{N}(0, 1)$, and $Y_i|h=1 \sim \mathcal{N}(1, 1)$.

$$5) \bigcup_{k_E=1}^{d_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i\} = \partial \mathcal{R}_i.$$

Therefore, the operation region $\mathcal{R}_i$ of an intercepted remote decision maker DM_i can be divided into convex privacy concern regions. Each privacy concern region consists of points achieved by the decision strategies which let the eavesdropper employ the same optimal deterministic strategy. Figure 4 illustrates the divisions of $\mathcal{R}_i$. Based on the result, we can rewrite the problem (24) as

$$\min_{k_E \in \{1, \dots, d_E\}} \min_{\mathbf{p}_i \in \mathcal{R}_i^{k_E}} \lambda(\mathbf{b}_i \mathbf{p}_i^\top + b_i) - (1 - \lambda)(\mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E}). \quad (27)$$

Note that the inner optimization of (27) is to minimize an affine function over a convex set. Therefore, the optimal point of the inner minimization is on the boundary $\partial \mathcal{R}_i^{k_E}$. If the inner optimal point is on $\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i$, it can be achieved by a deterministic likelihood test in the form of (4) based on the second property of $\mathcal{R}_i$. Otherwise, the inner optimal point is on the hyperplanes $\bigcup_{k'_E \in \{1, \dots, d_E\}, k'_E \neq k_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{C}_i^{k_E, k'_E}\}$ or sufficiently in the intersections of the privacy concern hyperplanes, i.e., the inner optimal point is in the set

$$\bigcup_{\substack{k'_E, k''_E \in \{1, \dots, d_E\} \\ k'_E \neq k_E, k''_E \neq k_E, k'_E \neq k''_E}} \left\{ \partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{C}_i^{k_E, k'_E} \cap \partial \mathcal{C}_i^{k_E, k''_E} \right\}. \quad (28)$$

In this case, the inner optimal point can be achieved by a randomized strategy of two deterministic likelihood tests in the form of (4). Since the inner optimal point is an optimal candidate of the outer minimization, we can conclude that the optimal point of the problem (27) can be on the boundary $\partial \mathcal{R}_i = \bigcup_{k_E=1}^{d_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i\}$ or in the intersection set

$$\bigcup_{\substack{k_E, k'_E, k''_E \in \{1, \dots, d_E\} \\ k'_E \neq k_E, k''_E \neq k_E, k'_E \neq k''_E}} \left\{ \partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{C}_i^{k_E, k'_E} \cap \partial \mathcal{C}_i^{k_E, k''_E} \right\}. \quad (29)$$

Therefore, the person-by-person optimal privacy-concerned strategy of (24) can be a deterministic likelihood test in the form of (4) or a randomized decision strategy of two deterministic likelihood tests in the form of (4). Based on the person-by-person optimality argument, we have the following conclusion about ϕ_i^{C*} of the intercepted remote decision maker DM_i .

Theorem 3: For a remote decision maker DM_i whose local decision U_i is intercepted, i.e., $i \in \mathcal{I}$, ϕ_i^{C*} can be a deterministic likelihood test in the form of (4) or a randomized decision strategy of two deterministic likelihood tests in the form of (4) in the optimal privacy-concerned distributed Bayesian detection design.

3) *Extended PBPO Algorithm:* Note that the objective function of the privacy-concerned distributed Bayesian detection problem (21) is different from that of the distributed Bayesian detection problem (2). Given a deterministic decision strategy of the eavesdropper $\phi_E^{k_E}$ and fixing other strategies but ϕ_i with $i \in \mathcal{I}$, a reformed person-by-person optimal decision strategy of an intercepted DM_i $\phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$ is given by

$$\text{For } i \in \mathcal{I}: \quad \phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E}) \\ = \arg \min_{\phi_i \in \Phi_i} \lambda r_F - (1 - \lambda) r_E, \quad (30)$$

where r_E is the Bayesian risk when the eavesdropper employs $\phi_E^{k_E}$. It is sufficient to consider a deterministic likelihood test in the form of (4) for $\phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$ where each constant coefficient $a_i^{u_i, h}$ is determined by the fixed decision and fusion strategies and the given $\phi_E^{k_E}$. The deterministic likelihood test $\phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$ achieves a point on $\partial \mathcal{R}_i$ with a supporting hyperplane in parallel with the privacy-concerned objective function hyperplane $\lambda(\mathbf{b}_i \mathbf{p}_i^\top + b_i) - (1 - \lambda)(\mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E})$.

Given decision strategies of all remote decision makers, Remark 4 shows that we can find the person-by-person optimal privacy-concerned fusion strategy by searching through all deterministic strategies of the fusion center.

Consider a remote decision maker DM_i whose local decision U_i is not intercepted, i.e., $i \notin \mathcal{I}$. Given other decision and fusion strategies, it is sufficient to consider the output deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$ of the standard PBPO algorithm for the person-by-person optimal privacy-concerned decision strategy of the remote decision maker DM_i with $i \notin \mathcal{I}$.

Next, consider an intercepted remote decision maker DM_i , i.e., $i \in \mathcal{I}$. Given decision and fusion strategies of all other nodes, it has been proved that the person-by-person optimal privacy-concerned strategy of the intercepted DM_i can be a deterministic likelihood test in the form of (4) or a randomized strategy of two deterministic likelihood tests in the form of (4). The extended PBPO algorithm searches the person-by-person optimal privacy-concerned decision strategy among candidates of all deterministic eavesdropping decision strategy cases. For a case of $\phi_E^{k_E}$, we discuss the extended algorithm to find the candidate in two cases.

- 1) If the point achieved by the deterministic likelihood test $\phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$ is on the boundary $\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i$, the deterministic likelihood test $\phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$ is the candidate of the person-by-person optimal privacy-concerned decision strategy of the intercepted DM_i .
- 2) Otherwise, we search the candidate of the person-by-person optimal privacy-concerned decision strategy of the intercepted DM_i through deterministic likelihood

tests which are in the form of (4) and achieve points in the set $\bigcup_{k'_E \in \{1, \dots, d_E\}, k'_E \neq k_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i \cap \partial \mathcal{C}_i^{k_E, k'_E}\}$ and randomized strategies of two deterministic likelihood tests in the form of (4) which achieve points in the set (28).

The proof here is similar to that of Algorithm 1. We need to consider the privacy concern region $\mathcal{R}_i^{k_E}$ and objective function hyperplane $\lambda(\mathbf{b}_i \mathbf{p}_i^\top + b_i) - (1 - \lambda)(\mathbf{b}_i^{k_E} \mathbf{p}_i^\top + b_i^{k_E})$ instead of the privacy-constrained region $\mathcal{R}_i^P$ and $\mathbf{b}_i \mathbf{p}_i^\top + b_i$.

Algorithm 2 is the extended PBPO algorithm to design the optimal privacy-concerned distributed Bayesian detection network.

Algorithm 2 The Extended PBPO Algorithm to Optimize the Privacy-Concerned Distributed Bayesian Detection Design

- 1: generate a sufficiently large number of initial decision strategy sets of $\{\phi_1, \dots, \phi_n\}$
 - 2: given an initial decision strategy set
 - 3: **while** $\lambda r_F - (1 - \lambda) r_E^{\min}$ is not minimized **do**
 - 4: given $\{\phi_1, \dots, \phi_n\}$
 - 5: $\phi_F^{\text{temp}} \leftarrow$
 search the optimal strategy through all deterministic fusion strategies
 - 6: $\phi_F \leftarrow \phi_F^{\text{temp}}$
 - 7: **for** $i \in \{1, \dots, n\}$ **do**
 - 8: given $\{\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F\}$
 - 9: **if** $i \notin \mathcal{I}$ **then**
 - 10: $\phi_i^{\text{temp}} \leftarrow \phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$
 - 11: $\phi_i \leftarrow \phi_i^{\text{temp}}$
 - 12: **else**
 - 13: **for** $k_E \in \{1, \dots, d_E\}$ **do**
 - 14: $\phi_i^{\text{temp}, k_E} \leftarrow \phi_i^{\Delta, k_E}(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F, \phi_E^{k_E})$
 - 15: **if** $\mathbf{p}_i$ achieved by $\phi_i^{\text{temp}, k_E} \in \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i\}$ **then**
 - 16: $\phi_i^{\text{candidate}, k_E} \leftarrow \phi_i^{\text{temp}, k_E}$
 - 17: **else**
 - 18: $\phi_i^{\text{temp}, k_E} \leftarrow$
 search the optimal strategy which minimizes the privacy-concerned objective $\lambda r_F - (1 - \lambda) r_E$ given $\phi_E^{k_E}$ through deterministic likelihood tests in the form of (4) which achieve points in the set $\bigcup_{k'_E \in \{1, \dots, d_E\}, k'_E \neq k_E} \{\partial \mathcal{R}_i^{k_E} \cap \partial \mathcal{R}_i \cap \partial \mathcal{C}_i^{k_E, k'_E}\}$ and randomized strategies of two deterministic likelihood tests in the form of (4) which achieve points in the set (28)
 - 19: $\phi_i^{\text{candidate}, k_E} \leftarrow \phi_i^{\text{temp}, k_E}$
 - 20: **end if**
 - 21: **end for**
 - 22: $\phi_i \leftarrow \arg \min_{\phi_i^{\text{candidate}, k_E}, k_E \in \{1, \dots, d_E\}} \lambda r_F - (1 - \lambda) r_E$
 - 23: **end if**
 - 24: **end for**
 - 25: **end while**
 - 26: compare obtained distributed detection designs corresponding to all initial decision strategy sets
- Output:** the optimal privacy-concerned distributed Bayesian detection design $\{\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}\}$

D. Equivalent Privacy-Aware Problems

We have shown two privacy-aware distributed Bayesian detection problem formulations. They address both the objectives of improving the detection performance and suppressing the privacy risk in different ways. In the next theorems, we show that both problems are equivalent, i.e., we can always formulate privacy-constrained and privacy-concerned problems which lead to the same optimal privacy-aware distributed Bayesian detection design.

Theorem 4: Given a privacy-concerned problem

$$\min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} \lambda r_F - (1 - \lambda) r_E^{\min}$$

with $\lambda \in (0, 1]$. Let $\{\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}\}$ be an optimal privacy-concerned distributed Bayesian detection design with privacy risk $r_E^{\min}(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*})$. Then, $\{\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}\}$ is also an optimal design for the following privacy-constrained distributed Bayesian detection problem:

$$\begin{aligned} \min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} & r_F \\ \text{s.t.} & r_E^{\min} \geq r_E^{\min}(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}). \end{aligned}$$

Proof: Given any distributed detection design $\{\phi_1, \dots, \phi_n, \phi_F\}$, the inequality $\lambda r_F(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) - (1 - \lambda) r_E^{\min}(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) \leq \lambda r_F(\phi_1, \dots, \phi_n, \phi_F) - (1 - \lambda) r_E^{\min}(\phi_1, \dots, \phi_n, \phi_F)$ always holds. Next, consider a distributed detection design which satisfies the privacy constraint condition, i.e., $r_E^{\min}(\phi_1, \dots, \phi_n, \phi_F) \geq r_E^{\min}(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*})$. Then,

$$\begin{aligned} & \lambda r_F(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) - (1 - \lambda) r_E^{\min}(\phi_1, \dots, \phi_n, \phi_F) \\ & \leq \lambda r_F(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) - (1 - \lambda) r_E^{\min}(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) \\ & \leq \lambda r_F(\phi_1, \dots, \phi_n, \phi_F) - (1 - \lambda) r_E^{\min}(\phi_1, \dots, \phi_n, \phi_F). \end{aligned}$$

Therefore, $r_F(\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}) \leq r_F(\phi_1, \dots, \phi_n, \phi_F)$ for all distributed detection designs which satisfy the privacy constraint. That means the optimal privacy-concerned design $\{\phi_1^{C*}, \dots, \phi_n^{C*}, \phi_F^{C*}\}$ is also an optimal privacy-constrained design for the equivalent privacy-constrained problem. ■

Theorem 5: Given a privacy-constrained problem

$$\begin{aligned} \min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} & r_F \\ \text{s.t.} & r_E^{\min} \geq \beta \end{aligned}$$

with $\beta \leq \gamma$ (otherwise, the privacy-constrained problem is not feasible). Let $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\}$ be an optimal privacy-constrained distributed Bayesian detection design with privacy risk $r_E^{\min}(\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*})$ and α^* be the optimal dual Lagrange variable for the constraint $r_E^{\min} \geq \beta$. Then, $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\}$ is also an optimal design for the following privacy-concerned distributed Bayesian detection problem:

$$\min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} \lambda r_F - (1 - \lambda) r_E^{\min}$$

$$\text{with } \lambda = \begin{cases} 1 & , \text{ if } \alpha^* = 0 \\ \frac{1}{1 + \alpha^*} & , \text{ if } \alpha^* > 0 \end{cases}$$

Proof: Let $L(\phi_1, \dots, \phi_n, \phi_F, \alpha) = r_F - \alpha(r_E^{\min} - \beta)$ with $\alpha \geq 0$ denote the Lagrange function for the given privacy-constrained problem and $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}, \alpha^*\}$ be an optimum. Since the Karush-Kuhn-Tucker conditions are necessary for the optimum, the optimal solution has to satisfy the complementary slackness condition $\alpha^*(r_E^{\min}(\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}) - \beta) = 0$. If $\alpha^* = 0$, then we can pick $\lambda = 1$ and it follows that $\{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\}$ is an optimal design of the corresponding privacy-concerned problem. If $\alpha^* > 0$, we must have $r_E^{\min}(\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}) = \beta$ based on the complementary slackness condition. Then, we pick $\lambda = \frac{1}{1 + \alpha^*}$ and it follows that

$$\begin{aligned} & \{\phi_1^{P*}, \dots, \phi_n^{P*}, \phi_F^{P*}\} \\ & = \arg \min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} r_F - \frac{1 - \lambda}{\lambda} (r_E^{\min} - \beta) \\ & = \arg \min_{\{\phi_1, \dots, \phi_n, \phi_F\} \in (\Phi_1 \times \dots \times \Phi_n \times \Phi_F)} \lambda r_F - (1 - \lambda) r_E^{\min}. \end{aligned}$$

Equivalent privacy-aware problems of different formulations are based on different ways to address the objectives of privacy risk and detection performance. Theorems 4 and 5 imply that any optimal privacy-aware distributed Bayesian detection design can be interpreted in both the privacy-constrained and privacy-concerned optimality senses. ■

IV. NUMERICAL STUDY

In this section, we show the optimal privacy-aware distributed Bayesian detection designs of a simple parallel model. Consider the system model in Figure 1 with a binary hypothesis H ($m = 2$), two remote decision makers ($n = 2$), and binary decisions U_1, U_2, U_F, U_E ($m = m_1 = m_2 = 2$). Conditionally independent observations of the remote decision makers are generated according to $Y_1|h = 0, Y_2|h = 0 \sim \mathcal{N}(0, 1)$ and $Y_1|h = 1, Y_2|h = 1 \sim \mathcal{N}(1, 1)$. The non-negative detection costs of the fusion center are assigned as $c_F(0, 0) = c_F(1, 1) = 0$ and $c_F(0, 1) = c_F(1, 0) = 1$. Then, the Bayesian risk r_F reduces to the average error probability of the fusion decision to infer the hypothesis.

A. Powerful Eavesdropper

Here, we consider the case that both DM-FC links are open-access and assume a powerful eavesdropper who intercepts all local decisions, i.e., $\mathcal{I} = \{1, 2\}$. Two detection cost assignments of the eavesdropper are considered. The first case assumes that the eavesdropper concerns more about its probability of miss $p_E^M = 1 - p_E^D = p_{U_E|H}(0|1)$ and assigns the non-negative detection costs of the eavesdropper as $c_E(0, 0) = c_E(1, 1) = 0$, $c_E(0, 1) = 2$, and $c_E(1, 0) = 1$. In the second case, we assume that the eavesdropper has the same detection cost assignment as the fusion center such that $c_E(0, 0) = c_E(1, 1) = 0$ and $c_E(0, 1) = c_E(1, 0) = 1$. Next, set the prior probability $p_H(0) = p_H(1) = 0.5$.

Figures 5 and 6 compare the optimal privacy-constrained designs and optimal privacy-concerned designs respectively. The numerical results show the trade-off of the detection performance and privacy risk. In Figure 5, the system privacy

Fig. 5. The detection performance r_F and privacy risk $r_E^{\min}$ of the optimal privacy-constrained distributed Bayesian detection designs against different privacy level guarantees β and detection cost assignments of the eavesdropper. Here, the eavesdropper is powerful to intercept all local decisions. The prior probability is set as $p_H(0) = p_H(1) = 0.5$. For the second case of detection cost assignment of the eavesdropper, $r_F = r_E^{\min}$.

is strengthened ($r_E^{\min}$ increases) along with worse detection performance (r_F increases) for each case of detection cost assignment of the eavesdropper. In Figure 6, the detection performance is improved (r_F decreases) along with weaker system privacy level ($r_E^{\min}$ decreases). For each detection cost assignment of the eavesdropper, setting $\beta = 0$ in the privacy-constrained problem and $\lambda = 1$ in the privacy-concerned problem will lead to the same optimal design which minimizes the Bayesian risk r_F without considering the eavesdropping risk $r_E^{\min}$. In Figure 5, the optimal privacy-constrained design is the optimal design without considering the privacy risk when β is smaller than a certain threshold. That is because $r_E^{\min}$ of the optimal distributed detection design without considering the privacy risk satisfies the privacy constraint condition when the privacy guarantee parameter β is small. For the second detection cost assignment of the eavesdropper, the two metrics r_F and $r_E^{\min}$ of an optimal privacy-constrained or optimal privacy-concerned design are always equal, i.e., the red square line fully overlaps the pink circle line in Figures 5 and 6. That is because the fusion center and eavesdropper have the same observations, detection cost assignment, and therefore minimize the same Bayesian risk function. Based on this property, we can explain the observation that the optimal privacy-constrained distributed Bayesian detection design always has $r_F = r_E^{\min} = \max_{\{\phi_1, \phi_2\} \in (\Phi_1 \times \Phi_2)} r_E^{\min}$ when $\lambda < 0.5$ and $r_F = r_E^{\min} = \min_{\{\phi_1, \phi_2, \phi_F\} \in (\Phi_1 \times \Phi_2 \times \Phi_F)} r_F$ when $\lambda > 0.5$ for the second detection cost assignment of the eavesdropper in Figure 6.

B. A Single Intercepted Local Decision

Here, the eavesdropper is supposed to intercept the DM₁-FC link. We assume that the eavesdropper has the same detection cost assignment as the fusion center. We consider to provide the maximum privacy guarantee in the optimal privacy-constrained distributed Bayesian detection design, i.e., $\beta = \gamma$. Since the maximum privacy constraint condition requires $r_E^{\min} \geq \beta = \gamma$ and $r_E^{\min} \leq \gamma$, the optimal privacy-constrained distributed Bayesian detection design will always have $r_E^{\min} = \gamma$. Note that the maximum privacy guarantee can

Fig. 6. The detection performance r_F and privacy risk $r_E^{\min}$ of the optimal privacy-concerned distributed Bayesian detection designs against different concern weights λ and detection cost assignments of the eavesdropper. Here, the eavesdropper is powerful to intercept all local decisions. The prior probability is set as $p_H(0) = p_H(1) = 0.5$. For the second case of detection cost assignment of the eavesdropper, $r_F = r_E^{\min}$.

Fig. 7. The detection performance r_F and privacy risk $r_E^{\min}$ of the optimal privacy-constrained distributed Bayesian detection design against different prior probabilities $p_H(0)$. Here, the eavesdropper is assumed to intercept U_1 and has the same detection cost assignment as the fusion center. The maximum privacy guarantee is set as $\beta = \gamma$. The detection performance of the optimal single-branch design, which also realizes the maximum privacy guarantee, is shown as a benchmark.

be also realized by cutting off DM₁-FC link and forcing the eavesdropper to infer the hypothesis by its knowledge of the prior probability only.

Figure 7 shows the detection performance and privacy risk of the optimal privacy-constrained distributed detection designs. As a benchmark, the detection performance of the optimal single-branch design, which cuts off the intercepted DM₁-FC link and optimizes the other secure link to minimize the Bayesian risk r_F , is also shown. Comparing the detection performances of the optimal privacy-constrained design and optimal single-branch design, we can conclude that the intercepted remote decision maker can be optimally designed to transmit data which is useless for the eavesdropper but is useful to make the fusion decision jointly with the information from the other optimally designed remote decision maker.

C. Soft Local Decision

Consider the same privacy-constrained distributed detection problem under the maximum privacy constraint and allow soft local decisions for remote decision makers. Here, we set

TABLE I
DETECTION PERFORMANCE IMPROVEMENT BY USING A SOFT LOCAL DECISION U_2

$p_H(0)$	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
r_F^H	0.048954	0.093829	0.133764	0.168622	0.198404	0.223128	0.243819	0.26347	0.284013	0.308538
r_F^S	0.048918	0.093473	0.132643	0.166293	0.1945	0.218629	0.240656	0.261948	0.28367	0.308538
$\frac{r_F^H - r_F^S}{r_F^H}$ (%)	0.074	0.379	0.838	1.381	1.968	2.016	1.297	0.578	0.121	0

$m = m_1 = 2$ and $m_2 = 3$, i.e., U_1 is a binary local decision while U_2 is a 3-ary local decision. In Table I, it shows the detection performance of the optimal privacy-constrained designs of making binary local decisions only (denoted by r_F^H) and making a 3-ary local decision U_2 (denoted by r_F^S). Compared with the numerical results in the case of binary decisions, the optimal privacy-constrained distributed detection design has a better or the same detection performance and r_F can be reduced by up to about 2% when DM_2 makes a 3-ary local decision. That is because more or at least the same information can be transmitted by a soft local decision than a hard binary decision.

V. CONCLUSION

In this paper, the privacy problem of a distributed detection network vulnerable to an informed and greedy eavesdropper is studied. In the eavesdropping scenario, the detection performance and privacy risk are the two objectives in the distributed detection design. We propose the idea of privacy-aware distributed detection design to address both of the objectives, which leads to a privacy-per-design approach. Here, a novel Bayesian detection-operational privacy metric is proposed to evaluate the privacy risk of information leakage caused by the informed and greedy eavesdropper. Employing the Bayesian detection-operational metrics, two privacy-aware distributed detection problems are formulated: the privacy-constrained distributed Bayesian detection problem and privacy-concerned distributed Bayesian detection problem. The privacy-constrained problem optimizes the detection performance under a privacy constraint to guarantee a certain privacy level. The privacy-concerned problem addresses both objectives with different weights in the distributed Bayesian detection design. If the optimal privacy-aware decision strategy of an intercepted remote decision maker is not characterized by an intersection of privacy constraint hyperplanes or privacy concern hyperplanes, our study shows that it is sufficient to consider a deterministic likelihood test for the optimal privacy-aware strategy. The obtained properties of the optimal privacy-aware designs enable us to extend the standard PBPO algorithm to solve the privacy-aware distributed Bayesian detection problems and to obtain the corresponding optimal detection designs. Interestingly, we can always formulate equivalent privacy-constrained and privacy-concerned problems which lead to the same optimal privacy-aware distributed Bayesian detection design. The numerical results show that the decision strategies of intercepted remote decision makers can be smartly designed to make local decisions which are useful for the fusion decision but useless for the eavesdropper. We

believe that the privacy-aware distributed detection design is an interesting approach for trustworthy IoT systems in future.

APPENDIX

Constant coefficients in the deterministic likelihood test $\phi_i^\Delta(\dots, \phi_{i-1}, \phi_{i+1}, \dots, \phi_F)$:

$$\begin{aligned} & a_i^{u_i, h} \\ &= \sum_{\{ \dots, u_{i-1}, u_{i+1}, \dots, u_F \} \in (\dots \times \mathcal{U}_{i-1} \times \mathcal{U}_{i+1} \times \dots \times \mathcal{U}_F)} \{ c_F(u_F, h) \\ & \times p_{U_F | \dots, U_{i-1}, U_i, U_{i+1}, \dots}(u_F | \dots, u_{i-1}, u_i, u_{i+1}, \dots) \\ & \times \dots \times p_{U_{i-1} | H}(u_{i-1} | h) \times p_{U_{i+1} | H}(u_{i+1} | h) \times \dots \\ & \times p_H(h) \}. \end{aligned} \quad (31)$$

Constant coefficients in the deterministic likelihood test $\phi_F^\Delta(\phi_1, \dots, \phi_n)$:

$$a_F^{u_F, h} = c_F(u_F, h) p_H(h). \quad (32)$$

Constant coefficients in the affine function (15):

$$b_i^{u_i, h} = a_i^{u_i, h} - a_i^{0, h}, (u_i \neq 0), \quad (33)$$

$$b_i = \sum_{h \in \mathcal{H}} a_i^{0, h}. \quad (34)$$

Constant coefficients in the affine function (16):

$$\begin{aligned} & b_i^{k_E, u_i, h} \\ &= \sum_{\{ \{ u_j : j \in \mathcal{I}, j \neq i \}, u_F \} \in (\{ \times_{j \in \mathcal{I}, j \neq i} \mathcal{U}_j \} \times \mathcal{U}_F)} \{ c_E(u_E, h) \\ & \times \{ p_{U_E | \{ U_j : j \in \mathcal{I}, j \neq i \}, U_i}(u_E | \{ u_j : j \in \mathcal{I}, j \neq i \}, u_i) \\ & - p_{U_E | \{ U_j : j \in \mathcal{I}, j \neq i \}, U_i}(u_E | \{ u_j : j \in \mathcal{I}, j \neq i \}, 0) \} \\ & \times \prod_{j \in \mathcal{I}, j \neq i} p_{U_j | H}(u_j | h) \times p_H(h) \}, (u_i \neq 0), \\ & b_i^{k_E} \\ &= \sum_{\{ \{ u_j : j \in \mathcal{I}, j \neq i \}, u_F \} \in (\{ \times_{j \in \mathcal{I}, j \neq i} \mathcal{U}_j \} \times \mathcal{U}_F), h \in \mathcal{H}} \{ c_E(u_E, h) \\ & \times p_{U_E | \{ U_j : j \in \mathcal{I}, j \neq i \}, U_i}(u_E | \{ u_j : j \in \mathcal{I}, j \neq i \}, 0) \\ & \times \prod_{j \in \mathcal{I}, j \neq i} p_{U_j | H}(u_j | h) \times p_H(h) \}, \end{aligned} \quad (35)$$

where $p_{U_E | \{ U_j : j \in \mathcal{I}, j \neq i \}, U_i}(u_E | \{ u_j : j \in \mathcal{I}, j \neq i \}, u_i)$ has been determined by the given deterministic eavesdropping strategy $\phi_E^{k_E}$.

REFERENCES

- [1] A. K. Pathan, H.-W. Lee, and C. S. Hong, "Security in wireless sensor networks: Issues and challenges," in *Proceedings of ICACT 2006*, vol. 2, 2006, pp. 1043–1048.
- [2] S. Schmidt, H. Krah, S. Fischer, and D. Wätjen, "A security architecture for mobile wireless sensor networks," in *Proceedings of the First European Conference on Security in Ad-hoc and Sensor Networks*. Springer-Verlag, 2005, pp. 166–177.
- [3] Z. Cao, J. Hu, Z. Chen, M. Xu, and X. Zhou, "Feedback: Towards dynamic behavior and secure routing for wireless sensor networks," in *Proceedings of AINA 2006*, vol. 2, 2006, pp. 160–164.
- [4] D. Liu, P. Ning, S. Zhu, and S. Jajodia, "Practical broadcast authentication in sensor networks," in *Proceedings of MobiQuitous 2005*, 2005, pp. 118–129.
- [5] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.
- [6] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [7] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 6, pp. 2180–2189, 2008.
- [8] E. Tekin, "The Gaussian multiple access wire-tap channel: Wireless secrecy and cooperative jamming," in *Proceedings of ITA 2007*, 2007, pp. 404–413.
- [9] K. Kittichokechai, T. J. Oechtering, and M. Skoglund, "Secure source coding with action-dependent side information," in *Proceedings of ISIT 2011*, 2011, pp. 1678–1682.
- [10] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer-Verlag New York, Inc., 1996.
- [11] J. N. Tsitsiklis, "Decentralized detection," in *Proceedings of Advances in Statistical Signal Processing*. JAI Press, 1993, pp. 297–344.
- [12] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II - Advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.
- [13] S. Marano, V. Matta, and P. K. Willett, "Distributed detection with censoring sensors under physical layer secrecy," *Signal Processing, IEEE Transactions on*, vol. 57, no. 5, pp. 1976–1986, 2009.
- [14] V. S. S. Nadendla, H. Chen, and P. K. Varshney, "Secure distributed detection in the presence of eavesdroppers," in *Proceedings of ASIOMAR 2010*, 2010, pp. 1437–1441.
- [15] V. Nadendla, "Secure distributed detection in wireless sensor networks via encryption of sensor decision," Master's thesis, Louisiana State University, 2009.
- [16] R. Soosahabi, M. Naraghi-Pour, D. Perkins, and M. A. Bayoumi, "Optimal probabilistic encryption for secure detection in wireless sensor networks," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 3, pp. 375–385, 2014.
- [17] H. Jeon, D. Hwang, J. Choi, H. Lee, and J. Ha, "Secure type-based multiple access," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 3, pp. 763–774, 2011.
- [18] H. Jeon, J. Choi, S. W. McLaughlin, and J. Ha, "Channel aware encryption and decision fusion for wireless sensor networks," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 4, pp. 619–625, 2013.
- [19] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, "Optimal distributed detection in the presence of Byzantines," in *Proceedings of ICASSP 2013*, 2013, pp. 2925–2929.
- [20] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of Byzantine attacks," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 16–29, 2009.
- [21] Z. Li, T. J. Oechtering, and J. Jaldén, "Parallel distributed Neyman-Pearson detection with privacy constraints," in *Proceedings of ICC 2014 Workshop*, 2014, pp. 765–770.
- [22] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *Information Forensics and Security, IEEE Transactions on*, vol. 3, no. 2, pp. 273–289, 2008.
- [23] X. Guo, A. S. Leong, and S. Dey, "Physical layer security in decentralized estimation systems," available online: http://people.eng.unimelb.edu.au/asleong/secure_estimation_journal.pdf.
- [24] C. Dwork, "Differential privacy," in *Proceedings of ICALP 2006*. Springer, 2006, pp. 1–12.
- [25] S. Oh and P. Viswanath, "The composition theorem for differential privacy," eprint arXiv:1311.0776.
- [26] A. Smith, "Privacy-preserving statistical estimation with optimal convergence rates," in *Proceedings of the Forty-Third Annual ACM Symposium on the Theory of Computing*, 2011, pp. 813–822.
- [27] J. Le Ny and G. J. Pappas, "Differentially private filtering," *Automatic Control, IEEE Transactions on*, vol. 59, no. 2, pp. 341–354, 2014.
- [28] Z. Ji, Z. C. Lipton, and C. Elkan, "Differential privacy and machine learning: A survey and review," eprint arXiv:1412.7584.
- [29] R. Foygel Barber and J. C. Duchi, "Privacy and statistical risk: Formalisms and minimax bounds," eprint arXiv:1412.4451.
- [30] Z. Li, T. J. Oechtering, and K. Kittichokechai, "Parallel distributed Bayesian detection with privacy constraints," in *Proceedings of ICC 2014*, 2014, pp. 2178–2183.
- [31] I. Y. Hoballah and P. K. Varshney, "Distributed Bayesian signal detection," *Information Theory, IEEE Transactions on*, vol. 35, no. 5, pp. 995–1000, 1989.
- [32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [33] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.

Zuxing Li (S'14) received his B.Eng. degree in Information Security in 2009 from Shanghai Jiao Tong University, P. R. China, his M.Sc. degrees in Electrical Engineering in 2013 from Technical University of Catalonia, Spain and KTH Royal Institute of Technology, Sweden. He is currently working toward the Ph.D. degree with the Communication Theory Lab. at KTH Royal Institute of Technology. His research interests include information security, distributed detection, information theory, and networked control.

Tobias J. Oechtering (S'01-M'08-SM'12) received his Dipl.-Ing degree in Electrical Engineering and Information Technology in 2002 from RWTH Aachen University, Germany, his Dr.-Ing degree in Electrical Engineering in 2007 from Technische Universität Berlin, Germany, and his Docent degree in Communication Theory in 2012 from KTH Royal Institute of Technology, Sweden. Between 2002 and 2008, he has been with Technische Universität Berlin and Fraunhofer German-Sino Lab. for Mobile Communications, Berlin, Germany. In 2008, he joined the Communication Theory Lab. at KTH Royal Institute of Technology and has been an Associate Professor since May 2013. Presently, he is serving as an editor for IEEE Communications Letters. Dr. Oechtering received the "Förderpreis 2009" from the Vodafone Foundation. His research interests include physical-layer secrecy, statistical signal processing, communication and information theory, as well as communication for networked control.