Cloud-Cluster Architecture for Detection in Intermittently Connected Sensor Networks

Michal Yemini, Stephanie Gil, and Andrea J. Goldsmith

Abstract

We consider a centralized detection problem where sensors experience noisy measurements and intermittent connectivity to a centralized fusion center. The sensors may collaborate locally within predefined sensor clusters and fuse their noisy sensor data to reach a common local estimate of the detected event in each cluster. The connectivity of each sensor cluster is intermittent and depends on the available communication opportunities of the sensors to the fusion center. Upon receiving the estimates from all the connected sensor clusters the fusion center fuses the received estimates to make a final determination regarding the occurrence of the event across the deployment area. We refer to this hybrid communication scheme as a *cloud-cluster* architecture. We propose a method for optimizing the decision rule for each cluster and analyzing the expected detection performance resulting from our hybrid scheme. Our method is tractable and addresses the high computational complexity caused by heterogeneous sensors' and clusters' detection quality, heterogeneity in their communication opportunities, and non-convexity of the loss function. Our analysis shows that clustering the sensors provides resilience to noise in the case of low sensor communication probability with the cloud. For larger clusters, a steep improvement in detection performance is possible even for a low communication probability by using our cloud-cluster architecture.

I. Introduction

The next generation of wireless infrastructure enables cloud connectivity, and with it, powerful centralized decision making based on sensor data. However, cloud connectivity of sensors cannot

M. Yemini and A. J. Goldsmith are with the Electrical and Computer Engineering Department, Princeton University, Princeton, NJ, 08544 USA, e-mails: myemini@princeton.edu; goldsmith@princeton.edu.

S. Gil is with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, Email: sgil@seas.harvard.edu.

A summary of the results presented in this paper was presented in the IEEE Global Communications Conference 2020 [1].

This work was supported partially by the AFOSR award #002484665 and partially by the NSF CAREER award #2114733.

be guaranteed at all times, particularly for sensors operating over mmWave frequency bands or in complex and potentially remote environments. Thus, a new paradigm that takes intermittent connectivity of sensors into account is needed. Currently, the analysis for sensor networks often adopts one of two extremes: i) a centralized architecture that is fully connected, or ii) a distributed architecture, such as peer-to-cloud, where connectivity is intermittent. In reality, a fully centralized case where sensors convey information directly to the cloud is faulty since connectivity to the cloud is intermittent. Alternatively, a distributed architecture where a sensor conveys its information directly to all of its neighbors to reach a common estimate distributively is not always feasible [2] as this also suffers from long convergence time in large networks. Therefore, adopting either of these extremes can be problematic when the assumption of a continuously connected system is not practical, and alternatively, requiring fully distributed communication leads to an overly conservative system.

The best way to fuse noisy data between sensors locally, and communicate this information to the cloud on an intermittent and sporadic basis, requires a trade-off between accuracy and reliability of transmission. Failure to correctly consolidate noisy information may sacrifice accuracy. Alternatively, requiring raw sensory data to be submitted over the cloud may suffer from either poor reliability due to sparse connectivity or high scheduling overhead due to a high connectivity requirement. Thus, the question of how the communication infrastructure affects resilience to noise and the decision making abilities of sensors presents a knowledge gap in our understanding of the vulnerability of multi-robot decision making systems in real world environments. Design questions for how to resolve this trade-off requires a formal analysis which is the subject of this work.

A more realistic scenario for multi-sensor systems operating in environments with limited connectivity is that they will have access to a combination of these two communication alternatives, a *hybrid* local (i.e., clustered) network *and* a (sporadically available) cloud network. We call this a *cloud-cluster* communication architecture. Such hybrid communication architectures give rise to important questions such as, 1) *how should the data be fused* at a local level in order to achieve the best global decision making ability at the cloud? and 2) what is the optimal size for the sensor clusters that would provide some *resilience to sensor noise and sporadic connectivity of sensors to the cloud?* Answering these questions would allow us the necessary insight to best optimize a cloud-cluster communication architecture for multi-sensor decision making.

This paper investigates the best architecture to achieve reliable prediction in the case of multiple sensors detecting an event of interest in the environment. We employ a hybrid architecture where clusters of sensors pre-process their noisy observations, sending a compressed lower-dimensional aggregate observation to the cloud according to the probabilistic availability of the link. We develop a parameterized understanding of the trade-offs involved between architectures; either using larger clusters of sensors approaching a cluster-based (distributed) communication scheme, or, using smaller clusters of sensors approaching a cloud based (centralized) communication scheme. We show that the cloud-cluster architecture can drastically improve resilience to noise when communication to the cloud is sporadic such as in real-world environments. We quantify the sensing noise of an individual sensor by its missed detection and false alarm probabilities, and its intermittent connectivity to the cloud by a Bernoulli random variable.

Paper contributions: The main contribution of our work is an analysis of hybrid cloud-cluster communication architectures to support multi-sensor decision making at the cloud when connectivity is intermittent. Understanding the optimal cloud-cluster communication architecture for multi-sensor systems allows us to *optimally use communication opportunities* to the cloud and to control the size of sensor clusters in a way that improves the quality of fused sensor data received at the cloud. We show that this lever is a powerful tool that can be used to arrive at improved decision capabilities for the sensors, while combating intermittent connectivity and noise in the sensing abilities of the individual sensors.

A. Related Work

There has been much work in the area of determining analytical rules for event detection in clustered sensor networks. In particular, the works [3]–[11] consider clustered sensor networks as a network organization scheme to reduce the communication overhead to the fusion center (FC). Sensor networks are often characterized by extreme power and communication constraints and thus the objective in decentralized detection for these systems is to perform well, in their ability to detect an event, while transmitting the smallest number of bits possible. While these works make a significant contribution to our understanding of the clustered sensor networks, they do not consider the sporadic nature of the intermittent connectivity of sensors systems. This aspect of the problem is very important, for example, in mmWave communication systems [12]–[14] that are vulnerable to temporary blockages, also known as outages. When a channel is blocked, no information can be passed through it, as its capacity is zero. These blockages occur with positive and non-negligible probability as is modeled in [15]–[17] and they become more frequent as the distance between the transmitter and receiver grows. Connectivity is also a common problem in mobile robotic systems (see [18]–[21]), where robot location affects both the robot connectivity to the FC, and its event-detection probability. To the best of our knowledge, minimizing the expected loss function of cloud-cluster sensor networks where sensors are intermittently connected to the cloud was not previously investigated. In this work we show that, using recently improved concentration inequalities, we can approximate the expected loss function caused by detection errors. We note that like prior works [3]–[10], we do not address the problem of optimizing sensor placement, or how to cluster existing sensors, but rather analyze the performance of existing system architectures.

Another related body of works analyzes the effect of the communication channel on the detection performance [22]–[26]. These works study the effect of the quality of the communication channel, available side information and transmission power constraints on the distortion of the signals that are sent to the FC by the sensors. Our work considers a starkly different setup where channels from sensors to the FC may be blocked, thereby causing intermittent connectivity. In this case no information can be received by the FC from senors with blocked channels. Our system architecture aims at improving connectivity to the FC using sensor clustering with optimized decision rules.

B. Paper Organization

The rest of the paper is organized as follows: Section II presents the system model and problem formulation. Section III analyzes the optimal cloud-cluster decision rules. Sections IV and V include approximations to the optimal decision rules when they are intractable. In particular Section IV presents system analysis and optimization for a homogeneous system setup, whereas Section V includes tractable analysis and decision rules for heterogeneous setups. Section VI presents numerical results. Finally, Section VII concludes the paper.

II. System Model and Problem Formulation

This section presents the system model we study in this work and the technical details of the associated system optimization problem we study. We provide a list of notable parameters used in this paper in Table I.

TABLE I: Table of Notable Parameters

Notation	Description
N	number of sensors
s_i	sensor i
p_{com,s_i}	communication probability of sensor i to the FC
P_{FA,s_i}	false alarm probability of sensor i
P_{MD,s_i}	missed detection probability of sensor i
t_i	indicator for the event that sensor i can communicates with the FC
y_i	binary measurement of sensor s_i
N_c	number of clusters
$\mathcal{C}_j$	the jth cluster of sensors
$n_{\mathcal{C}_j}$	number of sensors in cluster $\mathcal{C}_j$
$p_{\text{com},\mathcal{C}_j}$	communication probability of cluster j to the FC
τ_j	indicator for the event that cluster $\mathcal{C}_j$ can communicates with the FC
z_j	binary decision of cluster $\mathcal{C}_j$
$P_{\text{FA},\mathcal{C}_j}$	false alarm probability of cluster j
$P_{\text{MD},\mathcal{C}_j}$	missed detection probability of cluster j
γ_j	decision threshold at cluster $\mathcal{C}_j$
P_{FA}	false alarm probability of the FC
P_{MD}	missed detection probability of the FC
γ	decision threshold at the FC

A. System Model

We consider a team of multiple sensors indexed by i, $i \in \{1, \dots, N\}$, that are deployed to sense the environment and determine if the event of interest has occurred. We assume that the sensors are noisy and their ability to detect the event is captured for sensor i, by the probabilities P_{MD,s_i} of missed detection and P_{FA,s_i} of false alarm. Suppose that there are two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, the first occurs with probability $p_0 = 1 - p_1$ and the second with probability p_1. We denote the random variable that symbolizes the correct hypothesis by Ξ, where $\Xi \in \{0, 1\}$. We assume for each sensor i that the measured bit y_i may be swapped with the following probabilities

$$P_{\text{FA},s_i} = \Pr(y_i = 1 | \Xi = 0),$$

$$P_{\text{MD},s_i} = \Pr(y_i = 0 | \Xi = 1),$$

where $P_{\text{FA},s_i}, P_{\text{MD},s_i} \in (0, 0.5)$ without loss of generality. We allow for heterogeneity in each sensor's ability to detect the event of interest. In practice these can arise due to characteristics such as the quality of their sensors and their proximity to the measured event. The sensors have intermittent connectivity to a centralized cloud server, or *FC*. This intermittent connectivity is modeled by a binary random variable t_i that is equal to 1 if sensor s_i can communicate with the FC and 0 otherwise. We denote by p_{com,s_i} the probability that sensor s_i can communicate with the cloud (or FC), that is, $p_{\text{com},s_i} = \Pr(t_i = 1)$. Upon obtaining a communication link to the cloud server, a sensor will transmit sensed information from its cluster of sensors to the cloud.

Definition 1 (Cloud Architecture): In a cloud architecture all sensors transmit raw sensor data, y_i, to the cloud whenever a communication opportunity to the cloud exists. Connectivity to the

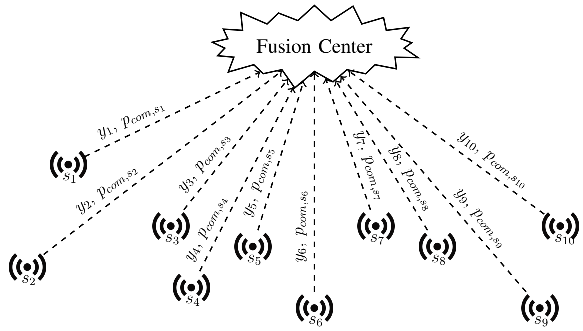


Fig. 1: Multi-robot system with cloud architecture.

cloud is provided as a probability p_{com,s_i}. The favorable case that $p_{\text{com},s_i} = 1$ for all i is equivalent to the classical centralized case since here all sensors have continual access to the cloud which in turn has access to all sensed measurements for event detection.

In a *cloud* approach, depicted in Fig. 1, the FC has the objective of determining whether the event has occurred after observing the measurements y_i of all communicating sensors. The FC gathers the information it receives from the sensors, and aims at estimating the correct hypothesis by minimizing the following expected loss function:

$$E(L) = Pr(\Xi = 0)P_{FA}L_{10} + Pr(\Xi = 1)P_{MD}L_{01}, \tag{1}$$

where L_{10} is the loss caused by false alarm, L_{01} is the loss caused by missed detection, and P_{FA} and P_{MD} are the false alarm and missed detection probabilities resulting from the FC detection decision, respectively. In the cloud approach, the FC may suffer from loss of connectivity to many sensors when connectivity is low. On other hand, high connectivity incurs high communication overhead such as scheduling that is undesirable. To reduce the communication overhead at the FC and also improve network connectivity, we propose an alternative approach to overcome these issues.

B. Problem Formulation

We study a different communication architecture where the sensors in the system are clustered into teams, and the sensors in each of these teams communicate with one another to arrive at a joint decision. This decision is then forwarded to the FC by a member of the cluster that can communicate with the FC. In this way, a cluster's decision can be forwarded to the FC if at least one sensor in the cluster can communicate with the FC. Upon receiving the processed measurement from the clusters, the FC estimates the correct hypothesis by minimizing (1) over

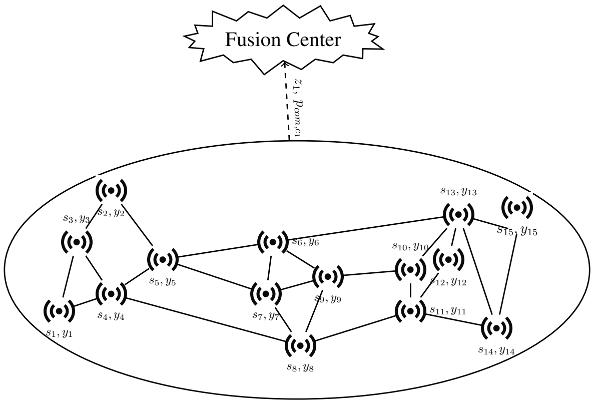


Fig. 2: Multi-robot system with cluster architecture.

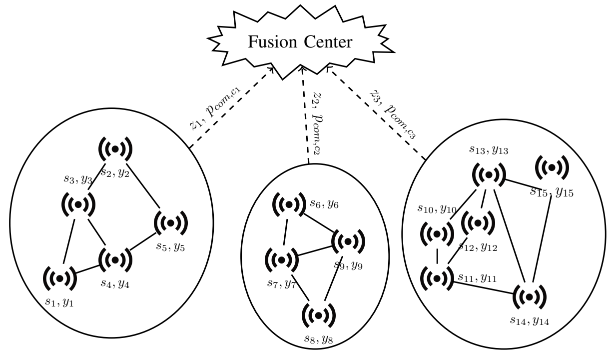


Fig. 3: Multi-robot system with cloud-cluster architecture.

all sensor clusters. We call this hybrid design of the sensor communication architecture a cloud-cluster architecture. Along these lines we study three main communication architectures and analytically study the performance of each as a function of probability of connectivity to the cloud and sensor noise (see Fig. 3):

Definition 2 (Cluster Architecture): In a cluster architecture, depicted in Fig. 2, all sensors have a fully connected local network and form a cluster where data is fused at a local level before being transmitted to the cloud. Connectivity to the cloud exists if any sensor s_i can communicate with the cloud. In this case the fused sensor data is transmitted to the cloud by the sensor s_i.

Definition 3 (Cloud-cluster Architecture): The cloud-cluster architecture is a hybrid between a *cloud* and a *cluster* architecture where sensors are divided into several *clusters*. It is assumed that sensors within a cluster are fully connected and can communicate locally. The number of clusters in the system can range from 1 (cluster architecture) to N (cloud architecture) and is often determined by the problem setting, i.e. sensors operating in the same room of a building would constitute a cluster. Sensed data by sensors operating in a cluster is fused at a local level before being transmitted to the cloud. Connectivity to the cloud exists for each cluster if there is a sensor in the cluster that can communicate with the cloud. In this case the fused sensor data for that cluster is transmitted to the cloud.

Since the cloud architecture and the cluster architecture are special cases of the cloud-cluster architecture, our analysis is presented for the case of a cloud-cluster architecture.

We consider a hybrid cloud-cluster system that is composed of N_c clusters, denoted by $\mathcal{C}_1, \ldots, \mathcal{C}_{N_c}$.

Definition 4 (Cluster connectivity): A cluster $\mathcal{C}_j$ communicates with the FC if at least one of the sensors within the cluster can communicate with the FC.

Let τ_j be a binary random variable that is equal to one if *cluster* $\mathcal{C}_j$ is communicating with the FC and zero otherwise and denote $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{N_c})$.

Every sensor cluster $\mathcal{C}_j$ communicating with the cloud sends a pre-processed value z_j that captures the observations of all sensors in cluster j. If cluster $\mathcal{C}_j$ cannot communicate with the FC z_j will take an arbitrary predefined deterministic value. We denote the vector of the pre-processed values by $\boldsymbol{z} = (z_1, \ldots, z_{N_c})$. The FC at the cloud determines its final decision of whether an event has occurred or not by using the optimal decision rule to minimize (1). This optimal decision rule[1] is to choose hypothesis $\mathcal{H}_1$ if:

$$\frac{\Pr(\boldsymbol{z}|H_1, \boldsymbol{\tau})}{\Pr(\boldsymbol{z}|H_0, \boldsymbol{\tau})} \geq \frac{L_{10}p_0}{L_{01}p_1} \tag{2}$$

and $\mathcal{H}_0$ otherwise.

We investigate the following questions:

1) how the data $\boldsymbol{z}$ is pre-processed at the cluster layer to reduce the expected loss at the FC,
2) how intermittent communication with the cloud impacts performance,
3) how the estimates of missed detection and false alarm probabilities are impacted by the communication architecture (i.e., the number of clusters and the number of sensors per cluster).

III. System Analysis and Optimization

In this section we optimize the decision at the cluster level and the FC. Additionally, we obtain the expected number of clusters that can communicate with the FC under the cloud-cluster architecture.

A. Cloud-Cluster Communication

Our cloud-cluster architecture is aimed at improving connectivity to FC when the probabilities p_{com,s_i} are small, and reducing scheduling and communication overheads when the probabilities p_{com,s_i} approach 1. We assume that the sensors are clustered into N_c groups. As stated in Definition 4, a cluster of sensors communicates with the FC if one of the sensors comprising the cluster sees a communication opportunity to the FC. Each cluster estimates the hypothesis and sends its estimation to the FC provided there is a communication opportunity to the FC.

[1] We refer the reader to [27, Chapter 3] for a primer on detection theory and hypothesis testing.

B. Communication probability of clusters and the expected number of communicating clusters

As we wrote before, a cluster of sensors $\mathcal{C}_j$ communicates with the FC if at least one of the sensors that comprises it can communicate with the FC. It follows that the probability that the cluster $\mathcal{C}_j$ can communicate with the FC, i.e., $\tau_j = 1$, is:

$$p_{\text{com},\mathcal{C}_j} = 1 - \prod_{i:s_i \in \mathcal{C}_j} (1 - p_{\text{com},s_i}). \tag{3}$$

Let $n_{\mathcal{C}_j}$ be the number of sensors in cluster $\mathcal{C}_j$. We can see that as we increase the number of sensors to the clusters, $p_{\text{com},\mathcal{C}_j}$ increases. Therefore, $p_{\text{com},\mathcal{C}_j}$ is maximized in the cluster architecture where $n_{\mathcal{C}_j} = N$. On the other hand, $p_{\text{com},\mathcal{C}_j}$ is minimized in the cloud architecture where $n_{\mathcal{C}_j} = 1$. Additionally, as we increase the probability that a sensor can communicate with the FC, $p_{\text{com},\mathcal{C}_j}$ is increased. Denote $(x_j)_{j=1}^{N} \triangleq (x_1, \ldots, x_N)$. From (3) we can calculate the following expected number of communicating clusters:

$$\eta \left(N_c, (\mathcal{C}_j)_{j=1}^{N_c}, (p_{\text{com},s_i})_{i=1}^{N}\right) = N_c - \sum_{j=1}^{N_c} \prod_{i:s_i \in \mathcal{C}_j} (1 - p_{\text{com},s_i}). \tag{4}$$

The optimization of the term (4) is beyond the scope of this paper since we assume a given clustering. Nonetheless, a closer look at the term (4) provides the following key observations. First, the expected number of communicating clusters is affected by three factors, namely, the number of clusters, the number of sensors in each cluster and the probability of connectivity to the FC. Second, the function η is monotonically increasing with p_{com,s_i}. However, the relationship between N_c, $|\mathcal{C}|_j$ and η given a fixed number of sensors N is more intriguing. Considering, for example, the homogeneous case where $|\mathcal{C}|_j = N/N_c$ and $p_{\text{com},s_i} = p_{\text{com},s}$ we have that:

$$\eta = N_c \cdot \left(1 - (1 - p_{\text{com},s})^{N/N_c}\right)$$

Therefore, for small values of $p_{\text{com},s}$ decreasing the number of clusters N_c increases η instead of decreasing it; this behavior is observed until the probability $(1 - p_{\text{com},s})^{N/N_c}$ becomes sufficiently small. When $p_{\text{com},s}$ is large, decreasing the number of clusters N_c decreases η; in this scenario clustering reduces the scheduling overhead at the FC.

C. Estimations in clusters

While the objective in the FC is to minimize (1) directly, the objective in the cluster level is to find the optimal trade-off between the probabilities of false alarm and missed detection. That is, the minimum probability of missed-detection that can be obtained for each value of the false alarm probability. By the Neyman-Pearson Lemma [27, Chapter 3] the optimal trade-off can be found by using the following likelihood ratio test with a desired threshold γ_j:

$$\frac{\Pr\left((y_i)_{i:s_i\in\mathcal{C}_j}|H_1\right)}{\Pr\left((y_i)_{i:s_i\in\mathcal{C}_j}|H_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_j. \tag{5}$$

In case of equality a random decision is made where hypothesis $\mathcal{H}_1$ is chosen with probability p_j and hypothesis $\mathcal{H}_0$ is chosen with probability $1-p_j$, where p_j is an additional parameter to be optimized.

Let

$$w_{1,s_i} = \ln\left(\frac{1-P_{\text{MD},s_i}}{P_{\text{FA},s_i}}\right), \qquad w_{0,s_i} = \ln\left(\frac{1-P_{\text{FA},s_i}}{P_{\text{MD},s_i}}\right). \tag{6}$$

We can rewrite the likelihood ratio test (5) for decision in cluster $\mathcal{C}_j$ as follows:

$$\sum_{i:s_i\in\mathcal{C}_j} \left[w_{1,s_i}y_i - w_{0,s_i}(1-y_i)\right] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_j. \tag{7}$$

In case of equality a random decision is made where hypothesis $\mathcal{H}_1$ is chosen with probability p_j and hypothesis $\mathcal{H}_0$ is chosen with probability $1-p_j$.

Denote, $P_{\text{FA},\mathcal{C}_j} = \Pr(z_j = 1|\mathcal{H}_0)$ and $P_{\text{MD},\mathcal{C}_j} = \Pr(z_j = 0|\mathcal{H}_1)$. Then the choice of threshold γ_j and tiebreak probability p_j results in the following detection error probabilities:

$$\begin{aligned}
P_{\text{FA},\mathcal{C}_j} &= \Pr\left(\sum_{i:s_i\in\mathcal{C}_j}\left[w_{1,s_i}y_i - w_{0,s_i}(1-y_i)\right] > \gamma_j|\mathcal{H}_0\right) \\
&\quad + p_j \Pr\left(\sum_{i:s_i\in\mathcal{C}_j}\left[w_{1,s_i}y_i - w_{0,s_i}(1-y_i)\right] = \gamma_j|\mathcal{H}_0\right), \\
P_{\text{MD},\mathcal{C}_j} &= \Pr\left(\sum_{i:s_i\in\mathcal{C}_j}\left[w_{1,s_i}y_i - w_{0,s_i}(1-y_i)\right] < \gamma_j|\mathcal{H}_1\right) \\
&\quad + (1-p_j) \Pr\left(\sum_{i:s_i\in\mathcal{C}_j}\left[w_{1,s_i}y_i - w_{0,s_i}(1-y_i)\right] = \gamma_j|\mathcal{H}_1\right).
\end{aligned} \tag{8}$$

Generally, as we discuss in Section III-E, the calculation of the probabilities $P_{\text{FA},\mathcal{C}_j}$ and $P_{\text{MD},\mathcal{C}_j}$ is intractable except for special cases such as the homogeneous case analyzed in Section IV. Therefore, our calculations for the general case, presented in Section V, rely on concentration inequalities to approximate $P_{\text{FA},\mathcal{C}_j}$ and $P_{\text{MD},\mathcal{C}_j}$.

The threshold γ_j and the probability p_j are parameters that the system architecture aims at optimizing to reduce the expected loss at the FC. Denote

$$\ell_{\min,j} = -\sum_{i:s_i\in\mathcal{C}_j} w_{0,s_i}, \qquad \ell_{\max,j} = \sum_{i:s_i\in\mathcal{C}_j} w_{1,s_i}. \tag{9}$$

The threshold γ_j can be optimized by searching over the interval $\mathcal{L}_j = [\ell_{\min,j}, \ell_{\max,j}]$ to minimize (1). Additionally, the probability p_j can be optimized by searching over the interval $[0,1]$. We note that the thresholds γ_j and probabilities p_j that dictate the clusters' decisions do not depend on the set of clusters whose measurements are successfully received and fused at the FC, using the decision rule (2). This choice obviates the need to optimize the thresholds γ_j and the probabilities p_j for all the possible 2^{N_c} combinations of communicating clusters. It also reduces the communication overhead that is caused by detecting the set of clusters that can communicate with the FC and sending this information back to the clusters for the correct choice of the γ_j and p_j every time the FC makes a detection decision.

D. FC Final Decision

Suppose that the cluster $\mathcal{C}_j$ is communicating with the FC and denote the data it sends to the FC by z_i. The optimal decision rule that minimizes (1) is choosing hypothesis $\mathcal{H}_1$ whenever (2) holds and hypothesis $\mathcal{H}_0$ otherwise.

Let

$$w_{1,\mathcal{C}_j} = \ln\left(\frac{1-P_{\text{MD},\mathcal{C}_j}}{P_{\text{FA},\mathcal{C}_j}}\right), \qquad w_{0,\mathcal{C}_j} = \ln\left(\frac{1-P_{\text{FA},\mathcal{C}_j}}{P_{\text{MD},\mathcal{C}_j}}\right). \tag{10}$$

The rule (2) can be written as:

$$\sum_{j=1}^{N_c} \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right] \geq \ln\left(\frac{L_{10}p_0}{L_{01}p_1}\right) = \gamma.$$

Note that in the case of equality, the expected loss due to detection error is equal for both the false alarm and missed-detection errors. Thus, in the case of equality we may choose hypothesis $\mathcal{H}_1$ arbitrarily since both hypotheses lead to the same loss.

Thus the sensing quality at the FC for a particular realization of the identity of communicating clusters can be written as

$$P_{\text{FA}}(\boldsymbol{\tau}) = \Pr\left(\sum_{j=1}^{N_c} \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right] \geq \gamma | \mathcal{H}_0, \boldsymbol{\tau}\right),$$

$$P_{\text{MD}}(\boldsymbol{\tau}) = \Pr\left(\sum_{j=1}^{N_c} \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right] < \gamma | \mathcal{H}_1, \boldsymbol{\tau}\right).$$

The probability of that particular realization of the identity of communicating clusters is

$$P(\boldsymbol{\tau}) = \prod_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}^{\tau_j} (1 - p_{\text{com},\mathcal{C}_j})^{1-\tau_j}. \tag{11}$$

This results in the following sensing probabilities

$$P_{\text{FA}} = \Pr\left(\sum_{j=1}^{N_c} \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right] \geq \gamma | \mathcal{H}_0\right) = \sum_{\boldsymbol{\tau}\in\{0,1\}^N} P(\boldsymbol{\tau}) P_{\text{FA}}(\boldsymbol{\tau}),$$

$$P_{\text{MD}} = \Pr\left(\sum_{j=1}^{N_c} \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right] < \gamma | \mathcal{H}_1\right) = \sum_{\boldsymbol{\tau}\in\{0,1\}^N} P(\boldsymbol{\tau}) P_{\text{MD}}(\boldsymbol{\tau}). \tag{12}$$

E. The Threshold Optimization Problem

Recall that $E(L) = Pr(\Xi = 0)P_{FA}L_{10} + Pr(\Xi = 1)P_{MD}L_{01}$. The global optimization problem resulting from the cloud-cluster architecture is as follows:

$$\min_{\{p_j\}_{j=1}^{N_c}, \{\gamma_j\}_{j=1}^{N_c}} E(L), \tag{13}$$

where P_{FA} and P_{MD} are defined as (12).

The complexity of calculating the optimal values p_j, γ_j, γ is high for the following reasons: first, the function $E(L)$ is not necessarily convex, thus the complexity can be exponential in the number of variables, i.e., exponential in $2N_c + 1$. Additionally, currently no close form method is known to calculate (8) and (12) efficiently since the coefficient are heterogeneous irrational numbers. We refer the reader to [28] for the case were the coefficients are rational numbers, additionally, the case of homogeneous coefficients is tractable as well. It follows that the overall complexity of optimizing $E(L)$ can be exponential in $2N_c + 1 + \max\{\max_j\{|\mathcal{C}_j|\}, N_c\}$, where the last term in the addition follows from the calculation of (8) and (12).

IV. Tractable Decision Optimization in Homogeneous Systems with Equal Thresholds

This section considers a special case of our system model that is homogeneous, i.e., all the clusters comprises an equal number of sensors and all sensors are homogeneous, i.e., $P_{\text{FA},s_i} = P_{\text{FA},s}$, $P_{\text{MD},s_i} = P_{\text{MD},s}$, $P_{\text{com},s_i} = P_{\text{com},s}$ for all $i \in \{1, \ldots, N\}$. In this case, $w_{1,s_i} = w_{1,s}$ and $w_{0,s_i} = w_{0,s}$ for all $i \in \{1, \ldots, N\}$.

For this setup, we consider the possibly suboptimal equal thresholds γ_j and probabilities p_j of the clusters. That is, $\gamma_j = \tilde{\gamma}_{\mathcal{C}}$ and $p_j = p_{\mathcal{C}}$ for all $i \in \{1, \ldots, N\}$, so that the calculation of the expected loss is tractable. Denote $w_{1,s} = \ln\left(\frac{1-P_{\text{MD},s}}{P_{\text{FA},s}}\right)$ and $w_{0,s} = \ln\left(\frac{1-P_{\text{FA},s}}{P_{\text{MD},s}}\right)$ and let

$$\gamma_{\mathcal{C}} = \frac{\tilde{\gamma}_{\mathcal{C}} + w_{0,s}}{w_{1,s} + w_{0,s}}.$$

Recall that $P_{\text{FA},s_i}, P_{\text{MD},s_i} \in (0, 0.5)$, therefore, $w_{0,s} > 0$ and $w_{1,s} > 0$. Under the assumptions of a homogeneous system and equal thresholds, we can rewrite (8) as

$$\begin{aligned} P_{\text{FA},\mathcal{C}_j} &= \Pr\left(\sum_{i:s_i \in \mathcal{C}_j} y_i > \gamma_{\mathcal{C}} | \mathcal{H}_0\right) + p_{\mathcal{C}} \Pr\left(\sum_{i:s_i \in \mathcal{C}_j} y_i = \gamma_{\mathcal{C}} | \mathcal{H}_0\right) \\ P_{\text{MD},\mathcal{C}_j} &= \Pr\left(\sum_{i:s_i \in \mathcal{C}_j} y_i < \gamma_{\mathcal{C}} | \mathcal{H}_1\right) + (1 - p_{\mathcal{C}}) \Pr\left(\sum_{i:s_i \in \mathcal{C}_j} y_i = \gamma_{\mathcal{C}} | \mathcal{H}_1\right). \end{aligned} \tag{14}$$

We can calculate the terms in (14) efficiently for each $\gamma_{\mathcal{C}}$ since the term $\sum_{i:s_i \in \mathcal{C}_j} y_i$ is distributed according to a binomial distribution for all $j \in \{1, \ldots, N_c\}$.

The equal decision rules in the clusters create homogeneous clusters, i.e., $P_{\text{FA},\mathcal{C}_j} = P_{\text{FA},\mathcal{C}}$ and $P_{\text{MD},\mathcal{C}_j} = P_{\text{MD},\mathcal{C}}$ for all $j \in \{1, \ldots, N_c\}$. Let $\mathbf{1}$ denote the N-dimensional row vector with all entries equal to 1. Then, $P_{\text{FA}}(\boldsymbol{\tau}_1) = P_{\text{FA}}(\boldsymbol{\tau}_2)$ and $P_{\text{MD}}(\boldsymbol{\tau}_1) = P_{\text{MD}}(\boldsymbol{\tau}_2)$ for all $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \{0,1\}^{N_c}$ such that $\boldsymbol{\tau}_1 \mathbf{1}^T = \boldsymbol{\tau}_2 \mathbf{1}^T$ where $(\cdot)^T$ denotes the transpose operator. Denote

$$\begin{aligned} P_{\text{FA},k} &= \Pr\left(\sum_{j=1}^{k} z_j \geq \frac{\gamma + w_{0,\mathcal{C}_j}}{w_{1,\mathcal{C}_j} + w_{0,\mathcal{C}_j}} | \mathcal{H}_0, \boldsymbol{\tau}\mathbf{1}^T = k\right), \\ P_{\text{MD},k} &= \Pr\left(\sum_{j=1}^{k} z_j < \frac{\gamma + w_{0,\mathcal{C}_j}}{w_{1,\mathcal{C}_j} + w_{0,\mathcal{C}_j}} | \mathcal{H}_1, \boldsymbol{\tau}\mathbf{1}^T = k\right). \end{aligned}$$

Note that due to the homogeneity of the setup, the identity of the communicating clusters does not affect the probabilities $P_{\text{FA},k}$ and $P_{\text{MD},k}$.

Furthermore, due to the homogeneity of the clusters we have that

$$P_{\text{com},\mathcal{C}_j} = P_{\text{com},\mathcal{C}} = 1 - (P_{\text{com},s})^{N/N_c} \tag{15}$$

for all $j \in \{1, \ldots, N_c\}$. Now, by (12) for each pair $(p_{\mathcal{C}}, \gamma_{\mathcal{C}})$ we have that

$$\begin{aligned} P_{\text{FA}} &= \sum_{k=0}^{N_c} \Pr\left(\boldsymbol{\tau}\mathbf{1}^T = k\right) P_{\text{FA},k}, \\ P_{\text{MD}} &= \sum_{k=0}^{N_c} \Pr\left(\boldsymbol{\tau}\mathbf{1}^T = k\right) P_{\text{MD},k}, \end{aligned} \tag{16}$$

where $\boldsymbol{\tau}\mathbf{1}^T$ is a binomial random variable with N_c experiments, each with probability of success $p_{\text{com},\mathcal{C}}$. Therefore, the problem (13) can be upper bounded by

$$\min_{p_{\mathcal{C}},\gamma_{\mathcal{C}},\gamma} E(L), \tag{17}$$

under the homogeneity assumptions included in this section.

Recall that in a homogeneous setup all the clusters include an equal number of sensors. Therefore, the number of sensors in each clusters is N/N_c. Algo. 1 depicts the resulting algorithm. It follows from (14) that the optimal value of $\gamma_{\mathcal{C}}$, under the homogeneity assumptions, is in the set $\{0, 1, \ldots, N/N_c\}$. Additionally, we perform a line search in the interval $[0, 1]$ to optimize the probability $p_{\mathcal{C}}$.

Algorithm 1 Optimization for homogeneous setup and equal cluster thresholds setup

1: Input: A set of N_c homogeneous clusters, each comprises n/N_c homogeneous sensors $\mathcal{C}_1, \ldots, \mathcal{C}_{N_c}$;
2: Input: $r_p \in \mathbb{N}_+$
3: Set $P_{\text{FA},s_i} = P_{\text{FA},s}$, $P_{\text{MD},s_i} = P_{\text{MD},s}$, $P_{\text{com},s_i} = P_{\text{com},s}$ for all $i \in \{1, \ldots, N\}$;
4: Set $d_p = 1/r_p$;
5: Set $\Gamma_{\mathcal{C}} = \{0, 1, \ldots, N/N_c\}$ and set $\Gamma_p = \{0, d_p, 2d_p, \ldots, 1\}$;
6: Set $P_{\text{FA},\mathcal{C}}$ and $P_{\text{MD},\mathcal{C}}$ as (14).
7: Solve

$$(p_{\mathcal{C}}, \gamma_{\mathcal{C}}) = \arg \min_{p_{\mathcal{C}} \in \Gamma_p, \gamma_{\mathcal{C}} \in \Gamma_{\mathcal{C}}} E(L);$$

8: Set $p_j = p_{\mathcal{C}}$ and $\gamma_j = \gamma_{\mathcal{C}} \cdot (w_{1,s} + w_{0,s}) - w_{0,s}$ for all $j \in \{1, \ldots, N_c\}$;

V. Tractable Decision Optimization in Heterogeneous Systems

This section optimizes the decision thresholds for heterogeneous systems at the cluster level using the Gauss-Seidel iterative method[2] which iteratively reduces the expected loss function at the FC. In the case that the terms (8) and (12) are intractable we approximate them via concentration inequalities. Algo. 2 depicts the optimization scheme we develop in this section. Additionally, we propose several initial values for Algo. 2 that we compare numerically in Section VI. We note that for the sake of clarity of presentation we present proofs and analytical analysis in Appendices A-E. Finally, hereafter we denote $\{x_j\}_{j=1}^N \triangleq \{x_1, \ldots, x_N\}$.

[2]The Gauss-Seidel iterative approach is considered in a relation to sensor network optimization in [4].

Algorithm 2 Optimization for heterogeneous setup

Input: A set of clusters of sensors $\mathcal{C}_1, \dots, \mathcal{C}_{N_c}$;
Inputs: $\{\gamma_j^{(0)}\}_{j=1}^{N_c}, \{p_j^{(0)}\}_{j=1}^{N_c}$;
Inputs: $\{\ell_{\min,j}\}_{j=1}^{N_c}$, $\{\ell_{\max,j}\}_{j=1}^{N_c}$, and $r_\gamma, r_p \in \mathbb{N}_+$;
Inputs: $\overline{\delta}_\gamma > 0$, $\overline{\delta}_p > 0$, $T > 0$, $m_s > 0$,$m_{\mathcal{C}} > 0$;
Set $\delta_\gamma^{(0)} = 2\overline{\delta}_\gamma$, $\delta_p^{(0)} = 2\overline{\delta}_p$, and $\Delta_{\gamma_j} = 2\overline{\delta}_\gamma$ and $\Delta_{p_j} = 2\overline{\delta}_p$ for all $j \in \{1, \dots, N_c\}$;
Set $d_j = (\ell_{\max,j} - \ell_{\min,j})/r_\gamma$ for all $j \in \{1, \dots, N_c\}$ and set $d_p = 1/r_p$;
Set $\Gamma_j = \{\ell_{\min,j}, \ell_{\min,j} + d_j, \ell_{\min,j} + 2d_j, \dots, \ell_{\max,j}\}$ and set $\Gamma_p = \{0, d_p, 2d_p, \dots, 1\}$;
Set $t = 0$, $j = 0$,
while $t < T$ **do**
  while $\delta_\gamma^{(t)} > \overline{\delta}_\gamma$ or $\delta_p^{(t)} > \overline{\delta}_p$ **do**
    Set $t = t + 1$;
    Set $j = \max\{\text{mod}(j + 1, N_c), 1\}$;
    Set $\gamma_k = \gamma_k^{(t-1)}$ and $p_k = p_k^{(t-1)}$ for all $k \in \{1, \dots, N_c\}$ such that $k \neq j$;
    if $n_{\mathcal{C}_j} > m_s$ and $N_c > m_{\mathcal{C}}$ **then**
      Substitute $P_{\text{FA},\mathcal{C}_j}$ by its estimate $U\left(n_{\mathcal{C}_j}, \alpha_{\text{FA},j}, M_{\text{FA},j}, \sigma^2_{\text{FA},j}\right)$ in the calculation of $E(L)$.
      Substitute $P_{\text{MD},\mathcal{C}_j}$ by its estimate $U\left(n_{\mathcal{C}_j}, \alpha_{\text{MD},j}, M_{\text{MD},j}, \sigma^2_{\text{MD},j}\right)$ in the calculation of $E(L)$.
      Substitute P_{FA} by its estimate $U\left(N_c, \alpha_{\text{FA}}, M_{\text{FA}}, \sigma^2_{\text{FA}}\right)$ in the calculation of $E(L)$.
      Substitute $P_{\text{MD},\mathcal{C}_j}$ by its estimate $U\left(N_c, \alpha_{\text{MD}}, M_{\text{MD}}, \sigma^2_{\text{MD}}\right)$ in the calculation of $E(L)$.
      Set $p_j^{(t)} = 1$ and $\gamma_j^{(t)} = \min_{\gamma_j \in \Gamma_j} \overline{E}(L)$, where $\overline{E}(L)$ is calculated by using the estimation for the terms $P_{\text{FA},\mathcal{C}_j}, P_{\text{MD},\mathcal{C}_j}, P_{\text{FA}}$ and P_{MD} in the calculation of $E(L)$;
    else if $n_{\mathcal{C}_j} \leq m_s$ and $N_c > m_{\mathcal{C}}$ **then**
      Substitute P_{FA} by its estimate $U\left(N_c, \alpha_{\text{FA}}, M_{\text{FA}}, \sigma^2_{\text{FA}}\right)$ in the calculation of $E(L)$.
      Substitute $P_{\text{MD},\mathcal{C}_j}$ by its estimate $U\left(N_c, \alpha_{\text{MD}}, M_{\text{MD}}, \sigma^2_{\text{MD}}\right)$ in the calculation of $E(L)$.
      Set $(\gamma_j^{(t)}, p_j^{(t)}) = \min_{\gamma_j \in \Gamma_j, p_j \in \Gamma_p} \overline{E}(L)$, where $\overline{E}(L)$ is calculated by using the estimation for the terms P_{FA} and P_{MD} in the calculation of $E(L)$;
    else if $n_{\mathcal{C}_j} > m_s$ and $N_c \leq m_{\mathcal{C}}$ **then**
      Substitute $P_{\text{FA},\mathcal{C}_j}$ by its estimate $U\left(n_{\mathcal{C}_j}, \alpha_{\text{FA},j}, M_{\text{FA},j}, \sigma^2_{\text{FA},j}\right)$ in the calculation of $E(L)$.
      Substitute $P_{\text{MD},\mathcal{C}_j}$ by its estimate $U\left(n_{\mathcal{C}_j}, \alpha_{\text{MD},j}, M_{\text{MD},j}, \sigma^2_{\text{MD},j}\right)$ in the calculation of $E(L)$.
      Set $p_j^{(t)} = 1$ and $\gamma_j^{(t)} = \min_{\gamma_j \in \Gamma_j} \overline{E}(L)$, where $\overline{E}(L)$ is calculated by using the estimation for the terms $P_{\text{FA},\mathcal{C}_j}$ and $P_{\text{MD},\mathcal{C}_j}$ in the calculation of $E(L)$;
    else
      Set $(\gamma_j^{(t)}, p_j^{(t)}) = \min_{\gamma_j \in \Gamma_j, p_j \in \Gamma_p} E(L)$;
    end if
    Set $\Delta_{\gamma_j} = |\gamma_j^{(t)} - \gamma_j^{(t-1)}|$ and set $\delta_\gamma^{(t)} = \max\{\Delta_{\gamma_k}\}_{k=1}^{N_c}$;
    Set $\Delta_{p_j} = |p_j^{(t)} - p_j^{(t-1)}|$ and set $\delta_p^{(t)} = \max\{\Delta_{p_k}\}_{k=1}^{N_c}$;
  end while
end while

A. From grid search to line search

We overcome the non-convexity of the objective function of (13) with respect to γ_j and p_j by optimizing these variables using a combination of the Gauss-Seidel iterative method with a line search at each iteration. Starting from chosen initial values for γ_j and p_j, this method optimizes the thresholds iteratively until convergence, one cluster at a time, while fixing the decision thresholds of all the other clusters. At each iteration a line search is performed over a predefined bounded interval to minimize the overall expected loss. We propose four different initial values for γ_j and p_j in Section V-C.

B. Approximating (8) *and* (12) *via concentration inequalities*

Now, we explore optimizing the thresholds γ_j via concentration inequalities, specifically, the improved Bennet's inequality that is stated in Theorem 2, Appendix A. We note that it is possible to approximate the detection error probability using the normal approximation. However, it yields smaller approximate probabilities than the true ones, which we want to upper bound, when the false alarm and missed detection probabilities are small. Therefore, it is not suitable to use in the estimation of the loss function at the FC when the clusters are large. Thus, for the clarity of presentation, we use the improved Bennet's inequality in our analysis, which upper bounds the desired probability in all scenarios.

Next, we present the notations we use in this section. Let $W(\cdot)$ denote the Lambert W function. Denote

$$U(n, \alpha, M, \sigma^2) \triangleq \exp\left[-\frac{\Lambda\alpha}{M} + n\ln\left(1 + \frac{\sigma^2}{M^2}\left(e^{\Lambda} - 1 - \Lambda\right)\right)\right], \tag{18}$$

where

$$A = \frac{M^2}{\sigma^2} + \frac{nM}{\alpha} - 1, \qquad B = \frac{nM}{\alpha} - 1, \qquad \Lambda = A - W(Be^{A}). \tag{19}$$

We separate the concentration inequalities analysis into two scenarios, both of which are intractable on their own.

1) Large number of sensors in cluster j ($n_{\mathcal{C}_j} \gg 1$): In this case we approximate the false alarm and missed detection probabilities of the decision of cluster j by applying the improved Bennet's inequality as follows.

Proposition 1: Let

$$\alpha_{\mathrm{FA},j} = \gamma_j - \sum_{i:s_i\in\mathcal{C}_j} (P_{\mathrm{FA},s_i} w_{1,s_i} - (1 - P_{\mathrm{FA},s_i}) w_{0,s_i}),$$

$$\sigma^2_{\text{FA},j} = \frac{1}{n_{\mathcal{C}_j}} \sum_{i:s_i \in \mathcal{C}_j} P_{\text{FA},s_i}(1 - P_{\text{FA},s_i})(w_{1,s_i} + w_{0,s_i})^2,$$

and $M_{\text{FA},j} = \max_{i:s_i \in \mathcal{C}_j}\{m_{\text{FA},i}\}$ where

$$m_{\text{FA},i} = \max\{|(1 - P_{\text{FA},s_i})(w_{1,s_i} + w_{0,s_i})|, |P_{\text{FA},s_i}(w_{1,s_i} + w_{0,s_i})|\}.$$

Then,

$$P_{\text{FA},\mathcal{C}_j} \leq U\left(n_{\mathcal{C}_j}, \alpha_{\text{FA},j}, M_{\text{FA},j}, \sigma^2_{\text{FA},j}\right), \tag{20}$$

for every γ_j such that $0 \leq \gamma_j - \sum_{i:s_i \in \mathcal{C}_j}(P_{\text{FA},s_i} w_{1,s_i} - (1 - P_{\text{FA},s_i})w_{0,s_i}) < n_{\mathcal{C}_j} \cdot M_{\text{FA},j}$.

Proposition 2: Denote

$$\alpha_{\text{MD},j} = \sum_{i:s_i \in \mathcal{C}_j}((1 - P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i} w_{0,s_i}) - \gamma_j,$$
$$\sigma^2_{\text{MD},j} = \frac{1}{n_{\mathcal{C}_j}} \sum_{i:s_i \in \mathcal{C}_j} P_{\text{MD},s_i}(1 - P_{\text{MD},s_i})(w_{1,s_i} + w_{0,s_i})^2,$$

and $M_{\text{MD},j} = \max_{i:s_i \in \mathcal{C}_j}\{m_{\text{MD},i}\}$ where

$$m_{\text{MD},i} = \max\{|P_{\text{MD},s_i}(w_{1,s_i} + w_{0,s_i})|, |(1 - P_{\text{MD},s_i})(w_{1,s_i} + w_{0,s_i})|\}.$$

Then,

$$P_{\text{MD},j} \leq U\left(n_{\mathcal{C}_j}, \alpha_{\text{MD},j}, M_{\text{MD},j}, \sigma^2_{\text{MD},j}\right), \tag{21}$$

for every γ_j such that $0 \leq \sum_{i:s_i \in \mathcal{C}_j}((1 - P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i} w_{0,s_i}) - \gamma_j < n_{\mathcal{C}_j} M_{\text{MD},j}$.
We prove Proposition 1 and Proposition 2 in Appendix B and Appendix C, respectively.

2) Large number of clusters ($N_c \gg 1$): In this case we approximate the false alarm and missed detection probabilities of the decision of the FC by Propositions 3 and 4 that are achieved by applying the improved Bennet's inequality.

Proposition 3: Let

$$\alpha_{FA} = \gamma - \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j})w_{0,\mathcal{C}_j}\right],$$
$$\sigma^2_{\text{FA}} = \frac{1}{N_c}\sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w^2_{1,\mathcal{C}_j} + (1 - P_{\text{FA},\mathcal{C}_j})w^2_{0,\mathcal{C}_j}\right]$$
$$- \frac{1}{N_c}\sum_{j=1}^{N_c} p^2_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j})w_{0,\mathcal{C}_j}\right]^2$$

and $M_{\text{FA}} = \max_{j\in\{1,\dots,N_c\}}\{m_{\text{FA},j}\}$ where

$$m_{\text{FA},j} = \max\{|w_{1,\mathcal{C}_j} - p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]|, \\ |w_{0,\mathcal{C}_j} + p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]|\}.$$

Then

$$P_{\text{FA}} \leq U\left(N_c, \alpha_{\text{FA}}, M_{\text{FA}}, \sigma^2_{\text{FA}}\right), \tag{22}$$

for every γ such that $0 \leq \gamma - \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right] < N_c \cdot M_{\text{FA}}$.

Proposition 4: Denote

$$\alpha_{MD} = \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right] - \gamma,$$

$$\sigma^2_{\text{MD}} = \frac{1}{N_c}\sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right] \\ - \frac{1}{N_c}\sum_{j=1}^{N_c} p^2_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w_{1,\mathcal{C}_j} - P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}\right]^2$$

and $M_{\text{MD}} = \max_{j\in\{1,\dots,N_c\}}\{m_{\text{MD},j}\}$ where

$$m_{\text{MD},j} = \max\{|w_{1,\mathcal{C}_j} - p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right]|, \\ |w_{0,\mathcal{C}_j} + p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right]|\}.$$

Then

$$P_{\text{MD}} \leq U\left(N_c, \alpha_{\text{MD}}, M_{\text{MD}}, \sigma^2_{\text{MD}}\right), \tag{23}$$

for every γ such that $0 \leq \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right] - \gamma < N_c M_{\text{MD}}$. We prove Proposition 3 and Proposition 4 in Appendix D and Appendix E, respectively.

Using the probability approximations in Propositions 1-4 we can evaluate and minimize the expected loss function to optimize the quality of detection even when the exact calculations are intractable.

C. Initial Inputs to Algorithm 2

Since Algo. 2 uses the Gauss-Seidel iterative algorithm it is required to provide it with the initial values $\{\gamma_j^{(0)}\}_{j=1}^{N_c}, \{p_j^{(0)}\}_{j=1}^{N_c}$. We consider the following four initial values:

1) For each cluster $\mathcal{C}_j$ the choice of $\gamma_j^{(0)}$ and $p_j^{(0)}$ is found using the equal threshold solution as in Algo. 1 under the assumption that there is N_c clusters that are identical to cluster $\mathcal{C}_j$, i.e. they include the same number of sensors as cluster $\mathcal{C}_j$ with the same probabilities of false alarm, missed-detection and communication to the cloud as the sensors in cluster $\mathcal{C}_j$. The probabilities $P_{\text{FA},\mathcal{C}}$ and $P_{\text{MD},\mathcal{C}}$ are calculated using the approximations we presented in V-B if they are intractabe.
2) Middle point of the intervals $[\ell_{\min,j}, \ell_{\max,j}]$ and $[0,1]$, respectively. That is,

$$\gamma_j^{(0)} = \frac{\ell_{\min,j} + \ell_{\max,j}}{2}, \qquad p_j^{(0)} = 0.5. \tag{24}$$

3) $\gamma_j^{(0)} = \ell_{\min,j}$ and $p_j^{(0)} = 1$, that is, $P_{\text{FA},\mathcal{C}_j} = 1$, $P_{\text{MD},\mathcal{C}_j} = 0$.
4) $\gamma_j^{(0)} = \ell_{\max,j}$ and $p_j^{(0)} = 0$, that is, $P_{\text{FA},\mathcal{C}_j} = 0$, $P_{\text{MD},\mathcal{C}_j} = 1$.

VI. Numerical Results

This section presents numerical results in which we evaluate the performance of the proposed cloud-cluster architecture. We consider a system with the following characteristics: 500 sensors, to evaluate both the actual and approximate performance, $p(\Xi = 1) = 0.65$, $L_{01} = 200$ and $L_{10} = 100$. To evaluate the performance of the proposed approach we compare two systems: a homogeneous one in which $p_{\text{FA},s_i} = 0.2$, $p_{\text{MD},s_i} = 0.35$ for all the sensors in the network, and a heterogeneous system in which for each sensor i we have that $p_{\text{FA},s_i} \sim U([0.16, 0.24])$ and $p_{\text{MD},s_i} \sim U([0.28, 0.42])$, that is, both the false alarm and missed detection probabilities of each sensor has a random deviation of 20% from their values in the homogeneous system. In the heterogeneous setup we average the expected loss of each realization of the false alarm and missed detection probabilities over 250 realizations. Additionally, in each grid search that we perform for optimizing γ_j we use 50 points per sensor, i.e., a total of $r_\gamma = 50 \times n_{\mathcal{C}_j}$ points. Finally, the line search resolution for the variable p_j is 0.01, that is, $r_p = 100$.

First, we evaluate in Fig. 4 the communication probability of a cluster to the cloud as a function of the number of sensors it comprises for three values of individual sensor communication probability, $P_{\text{com},s_i} = 0.05, 0.25, 0.5$. Fig. 4 validates that the communication probability of a cluster grows monotonically with the number of sensors it includes. Additionally, it shows that for higher values of P_{com,s_i} the increase in communication probability occurs and saturates faster than for lower values of P_{com,s_i}.

Figs. 5-6 evaluate the approximate loss that each of the initial inputs of Algo. 2 that we present in Section V-C yields. Comparing the four initial thresholds for Algo. 2, we can see that the first

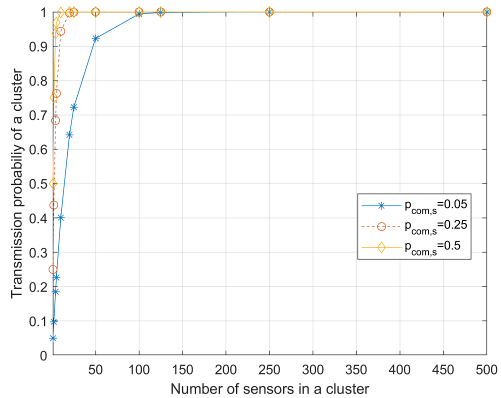
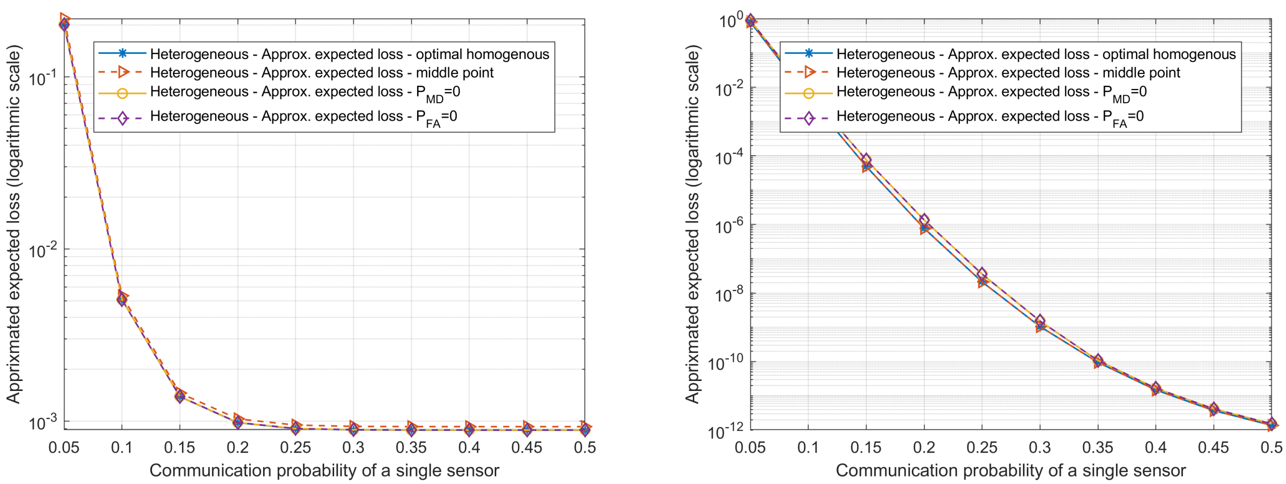


Fig. 4: Communication probability to the cloud as a function of the number of sensors it includes.



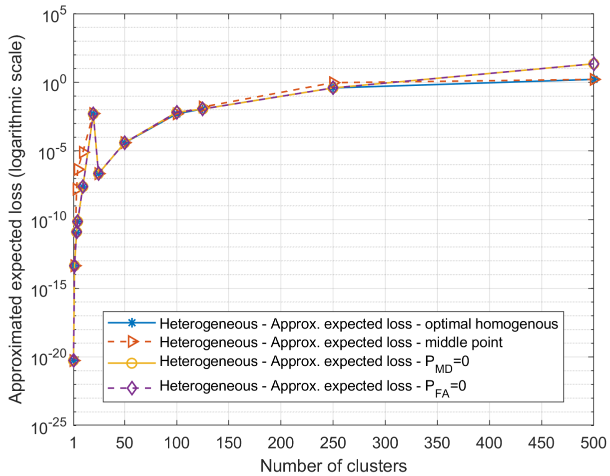
(a) A system with 20 clusters, each including 50 sensors.

(b) A system with 100 clusters, each including 5 sensors.

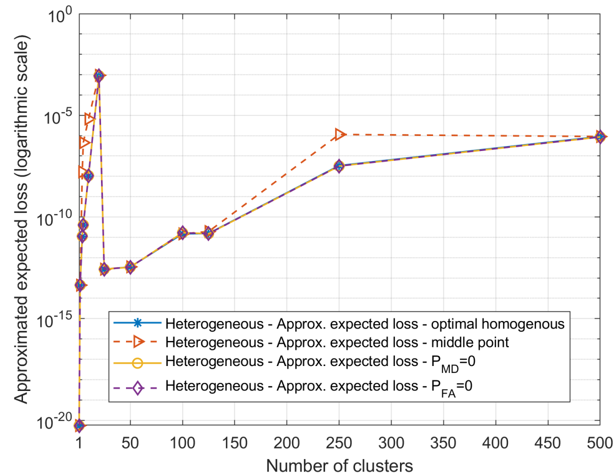
Fig. 5: The expected loss as a function of the communication probability of each sensor, for each of the initial thresholds presented in Section V-C. The approximated expected loss values resulting from the different initial thresholds are similar. Nevertheless, there is a small but persistent advantage for the "optimal homogeneous" initial threshold that minimizes the expected loss function assuming identical clusters.

initial threshold that we propose in Section V-C, which chooses for each cluster the threshold that minimizes the expected loss function assuming identical clusters, is consistently on-par or outperforms the other three initial threshold values we propose in Section V-C.

To evaluate the exact performance achieved by thresholds that are optimized using the approximations that we present in Section V, we use a homogeneous setup with equal cluster size as a tractable setup for which we can calculate the expected loss exactly. We then compare the exact calculation to its approximation that is calculated using Eqs. (20)-(23). In the heterogeneous



(a) A system with sensor communication probability $p_{\text{com},s_i} = 0.1$.



(b) A system with sensor communication probability $p_{\text{com},s_i} = 0.4$.

Fig. 6: The expected loss as a function of the number of the equal sized clusters, for each of the initial thresholds presented in Section V-C. Similarly to Fig. 5 the "optimal homogeneous" initial threshold which minimizes the expected loss function assuming identical clusters consistently outperforms or is on-par with the other three candidates.

setup we choose the initial threshold γ_j for each cluster $\mathcal{C}_j$ using the first initial threshold that we propose in SectionV-C. In the homogeneous setup we optimize the system by using Algo. 1. Additionally, in both the heterogeneous setup and the approximate calculation in the homogeneous setup we use the approximate missed detection and false alarm probabilities to approximate $P_{\text{FA},\mathcal{C}_j}$ and $P_{\text{MD},\mathcal{C}_j}$ presented in Section V-B if $n_{\mathcal{C}_j} > 20$. Additionally, we use the approximate missed detection and false alarm probabilities to approximate P_{FA} and P_{MD}, i.e., the error probabilities at the FC, presented in Section V-B if $N_c > 10$. Otherwise we use exact calculations.

Figs. 7-8 depict the expected loss as a function of the sensor communication probability $p_{\text{com},s}$ for various values of N_c (the number of clusters). Figs. 9-10 depict the expected loss as a function of the number of clusters N_c that comprise the system for various values of sensor communication probabilities $p_{\text{com},s}$. Each of the Figs. 7-10 includes five lines also denoted in the legends. These are defined as:

Expected loss - exact calculation: the expected loss of the homogeneous system using exact calculations in Algo. 1.

Expected loss - majority: the expected loss of the homogeneous system in which each cluster makes a majority rule decision where $\gamma_j = \lfloor n_{\mathcal{C}_j}/2 \rfloor + 1$. The expected loss is calculated exactly.

Expected loss - γ_j calculated using approximation: the exact expected loss that the choice γ_j yields, where γ_j is optimized using the concentration inequalities depicted in Section V-B in Algo. 1 instead of the exact calculation of the loss function.

Approximate expected loss - homogeneous: the approximate expected loss that is calculated using the concentration inequalities depicted in Section V-B in Algo. 1 instead of the exact calculation of the loss function.

Approximate expected loss - heterogeneous: the approximate expected loss that is calculated using Algo. 2 with the first initial threshold that is proposed in SectionV-C.

Figs. 7-8 show that when the number of clusters is large (i.e., each cluster consists of a small number of sensors), the improvement in performance of a highly connected system compared with that of a sparsely connected system is much more significant than the contrasting scenario of a system with a small number of clusters. Additionally, Figs. 7-8 confirm that optimizing the thresholds γ_j using concentration inequalities yield an actual expected loss that is on par with that of optimizing γ_j using exact calculations. Additionally, Figs. 7-8 depict the gap between the approximate loss function and the exact one for the homogeneous setup and show that our use of the improved Bennet's inequality results in a good approximation for the expected loss function. Therefore, while the heterogeneous setup is not tractable we can expect that our use of the improved Bennet's inequality results in a good approximation for the expected loss function for the heterogeneous setup as well. Finally, Figs. 7-8 shows the large gain that optimizing the threshold values provides instead of choosing a majority decision rule.

Figs. 9-10 show that when the communication probabilities of sensors to the FC are low, we observe a monotonic decrease in the loss function as we decrease the number of clusters in the exact loss function. This is also observed for the approximate loss function with the exception of a small increase when the systems is composed of 4 clusters. When the communication probabilities of sensors to the FC are higher, clustering may actually increase the expected loss. This follows because of the single bit compression that occurs in the clusters' single bit decisions. That is, there is a trade-off between the error probabilities of the decisions in clusters and that of the FC. Increasing the number of clusters reduces the number of measurements that the clusters use to make their decisions, and also reduces the communication probability to the FC since clusters include less sensors and thus reduced chances of seeing an opportunity to access the cloud. However, if the communication probability is high, increasing the number of clusters can

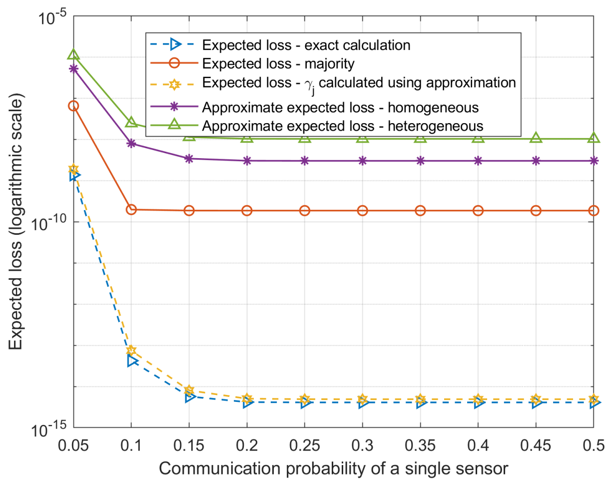


Fig. 7: The expected loss function of the communication probability of each sensor for a system with 10 clusters, each including 50 sensors. For cloud-cluster architectures we attain a dramatic improvement in performance due to clustering if sensor communication probability to the cloud is at least 0.15.

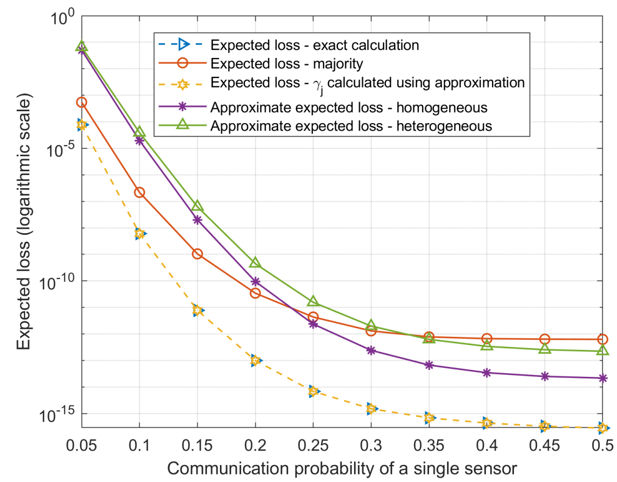


Fig. 8: The expected loss function vs. the communication probability of each sensor for a system with 50 clusters, each including 10 sensors. For small size clusters, approaching a distributed architecture, higher probability of communication to the cloud is required for better performance.

result in the FC having more measurements to rely on upon making its final decision.

VII. Conclusion

We consider multi-sensor systems that operate in environments where cloud connectivity is available intermittently. We provide an analytical study of the tradeoffs between different information exchange architectures to support an event detection task. Our results show that if cloud connectivity is reliable, directing sensors to share their sensed values to the cloud for event detection at a centralized fusion center will always perform best. However, in the more likely scenario where cloud connectivity is intermittent, clustering sensors into local neighborhoods where their sensed values are processed and then sent to the cloud during sporadic communication opportunities performs best. In particular, our results give insight into the optimal cluster sizes needed to achieve minimum detection loss at the cloud even in the face of noisy sensor data and intermittent communication. Future work can use the results presented here to optimize the locations of sensors such that they attain the recommended cluster sizes for best detection performance over the environment.

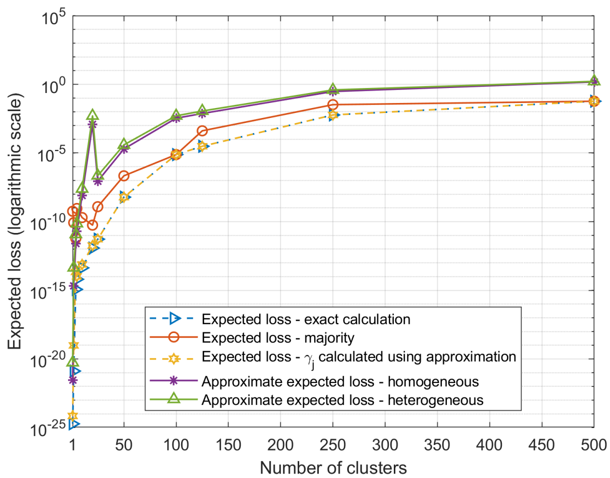


Fig. 9: The expected loss function of the number of equal size clusters N_c for $p_{\text{com},s_i} = 0.1$. Since connectivity to the FC is low, reducing the number of clusters (more sensors per cluster) increases the chances of communication to the cloud and improves the overall performance.

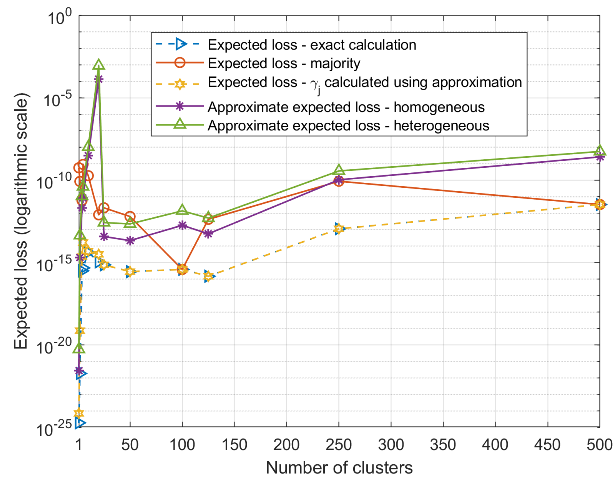


Fig. 10: The expected loss function of the number of equal N_c size clusters for $p_{\text{com},s_i} = 0.5$. When connectivity of sensors to the cloud is high, smaller clusters are favored for improving multi-sensor system performance since sensor fusion at the cluster level can be thought of as a form of lossy compression.

APPENDIX A
PRIMER ON CONCENTRATION INEQUALITIES

We first provide a primer on key concentration inequality results that we will use for the development of our analysis. Since we consider a heterogeneous setup in which the false alarm and missed detection probabilities may vary, we cannot use the concentration inequality [29] for the binomial distribution. Instead we use an improved Bennett's inequality which is known to outperform both Bernstein and Hoeffding's inequalities as well as the Bennet's inequality [30].

Theorem 1 (Bennet's inequality [30]): Assume that $x_1 \ldots, x_n$ are independent random variables and $E(x_i) = 0$, $E(x_i^2) = \sigma_i^2$ and $|x_i| < M$ almost surely. Then, for any $0 \leq t < nM$

$$\Pr\left(\sum_{i=1}^{n} x_i \geq \alpha\right) \leq \exp\left(-\frac{n\sigma^2}{M^2} h\left(\frac{\alpha M}{n\sigma^2}\right)\right), \tag{25}$$

where $h(x) = (1+x)\ln(1+x) - x$ and $n\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$.

Theorem 2 (The improved Bennet's inequality [31]): Assume that $x_1 \ldots, x_n$ are independent random variables and $E(x_i) = 0$, $E(x_i^2) = \sigma_i^2$ and $|x_i| < M$ almost surely. Additionally, let $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} \sigma_i^2$ and

$$A = \frac{M^2}{\sigma^2} + \frac{nM}{\alpha} - 1, \qquad B = \frac{nM}{\alpha} - 1, \qquad \Lambda = A - W(Be^A), \tag{26}$$

where $W(\cdot)$ is the Lambert W function. Denote

$$U(n,\alpha,M,\sigma^2) \triangleq \exp\left[-\frac{\Lambda\alpha}{M} + n\ln\left(1+\frac{\sigma^2}{M^2}\left(e^{\Lambda}-1-\Lambda\right)\right)\right]. \tag{27}$$

Then for any $0 \le \alpha < nM$

$$\Pr\left(\sum_{i=1}^{n} x_i \ge \alpha\right) \le U(n,\alpha,M,\sigma^2).$$

APPENDIX B
PROOF OF PROPOSITION 1

Denote $\tilde{y}_i = w_{1,s_i} y_i - w_{0,s_i}(1-y_i)$. We can upper bound the false alarm probability in (8) by

$$P_{\text{FA},\mathcal{C}_j} \le \Pr\left(\sum_{i:s_i\in\mathcal{C}_j} \left[\tilde{y}_i - E\left(\tilde{y}_i|\mathcal{H}_0\right)\right] \ge \gamma_j - \sum_{i:s_i\in\mathcal{C}_j} E\left(\tilde{y}_i\middle|\mathcal{H}_0\right)|\mathcal{H}_0\right).$$

Furthermore,

$$\begin{aligned} E\left(\tilde{y}_i|\mathcal{H}_0\right) &= P_{\text{FA},s_i} w_{1,s_i} - (1-P_{\text{FA},s_i}) w_{0,s_i}, \\ E(\tilde{y}_i^2|\mathcal{H}_0) &= P_{\text{FA},s_i} w_{1,s_i}^2 + (1-P_{\text{FA},s_i}) w_{0,s_i}^2. \end{aligned} \tag{28}$$

It follows that

$$\begin{aligned} \sigma_{\text{FA},s_i}^2 &\triangleq \text{var}\left(\tilde{y}_i - E\left(\tilde{y}_i|\mathcal{H}_0\right)|\mathcal{H}_0\right) = \text{var}\left(\tilde{y}_i|\mathcal{H}_0\right) = E\left(\tilde{y}_i^2|\mathcal{H}_0\right) - \left[E\left(\tilde{y}_i|\mathcal{H}_0\right)\right]^2 \\ &= P_{\text{FA},s_i}(1-P_{\text{FA},s_i})(w_{1,s_i}+w_{0,s_i})^2. \end{aligned} \tag{29}$$

Now, we can use Theorem 2 to upper bound the false alarm probability of the decision of cluster j by substituting

$$\begin{aligned} x_i &= \tilde{y}_i - E\left(\tilde{y}_i|\mathcal{H}_0\right) = \tilde{y}_i - P_{\text{FA},s_i} w_{1,s_i} + (1-P_{\text{FA},s_i}) w_{0,s_i} \\ \alpha_{\text{FA},j} &= \gamma_j - \sum_{i:s_i\in\mathcal{C}_j} E(\tilde{y}_i|\mathcal{H}_0) = \gamma_j - \sum_{i:s_i\in\mathcal{C}_j} (P_{\text{FA},s_i} w_{1,s_i} - (1-P_{\text{FA},s_i}) w_{0,s_i}). \end{aligned}$$

In this case,

$$\sigma_{\text{FA},j}^2 = \frac{1}{n_{\mathcal{C}_j}} \sum_{i:s_i\in\mathcal{C}_j} P_{\text{FA},s_i}(1-P_{\text{FA},s_i})(w_{1,s_i}+w_{0,s_i})^2, \tag{30}$$

and $M_{\text{FA},j} = \max_{i:s_i\in\mathcal{C}_j}\{m_{\text{FA},i}\}$ where

$$\begin{aligned} m_{\text{FA},i} &= \max\left\{\left|w_{1,s_i} - E\left(\tilde{y}_i|\mathcal{H}_0\right)\right|, \left|w_{0,s_i} + E\left(\tilde{y}_i|\mathcal{H}_0\right)\right|\right\} \\ &= \max\{|(1-P_{\text{FA},s_i})(w_{1,s_i}+w_{0,s_i})|, |P_{\text{FA},s_i}(w_{1,s_i}+w_{0,s_i})|\}. \end{aligned}$$

We denote the resulting constants defined in Theorem 2 by $A_{\text{FA},j}$, $B_{\text{FA},j}$ and $\Lambda_{\text{FA},j}$. Thus, by the improved Bennett's inequality we have that

$$P_{\text{FA},\mathcal{C}_j} \leq U\left(n_{\mathcal{C}_j}, \alpha_{\text{FA},j}, M_{\text{FA},j}, \sigma^2_{\text{FA},j}\right),$$

for every γ_j such that $0 \leq \gamma_j - \sum_{i:s_i\in\mathcal{C}_j}(P_{\text{FA},s_i}w_{1,s_i} - (1-P_{\text{FA},s_i})w_{0,s_i}) < n_{\mathcal{C}_j} \cdot M_{\text{FA},j}$.

APPENDIX C
PROOF OF PROPOSITION 2

Similarly to the proof presented in Appendix B, we can use Theorem 2 to upper bound the missed detection probability of cluster j. Recall that $\tilde{y}_i = w_{1,s_i}y_i - w_{0,s_i}(1-y_i)$. We upper bound the missed detection probability, $P_{\text{MD},\mathcal{C}_j}$, in (8) as follows

$$P_{\text{MD},\mathcal{C}_j} \leq \Pr\left(\sum_{i:s_i\in\mathcal{C}_j}\left[E\left(\tilde{y}_i|\mathcal{H}_1\right) - \tilde{y}_i\right] \geq \sum_{i:s_i\in\mathcal{C}_j} E\left(\tilde{y}_i|\mathcal{H}_1\right) - \gamma_j \middle| \mathcal{H}_1\right).$$

Furthermore,

$$E\left(\tilde{y}_i|\mathcal{H}_1\right) = (1-P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i}w_{0,s_i},$$
$$E(\tilde{y}_i^2|\mathcal{H}_1) = (1-P_{\text{MD},s_i})w^2_{1,s_i} + P_{\text{MD},s_i}w^2_{0,s_i}.$$

It follows that

$$\begin{aligned}\sigma^2_{\text{MD},s_i} &\triangleq \text{var}\left(E\left(\tilde{y}_i|\mathcal{H}_1\right) - \tilde{y}_i|\mathcal{H}_1\right) = \text{var}\left(\tilde{y}_i|\mathcal{H}_1\right) = E\left(\tilde{y}_i^2|\mathcal{H}_1\right) - \left[E\left(\tilde{y}_i|\mathcal{H}_1\right)\right]^2 \\ &= P_{\text{MD},s_i}(1-P_{\text{MD},s_i})(w_{1,s_i}+w_{0,s_i})^2.\end{aligned}$$

Now, we can use Theorem 2 to upper bound the missed detection probability of the decision of cluster j by substituting

$$x_i = E\left(\tilde{y}_i|\mathcal{H}_1\right) - \tilde{y}_i = (1-P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i}w_{0,s_i} - \tilde{y}_i,$$
$$\alpha_{\text{MD},j} = \sum_{i:s_i\in\mathcal{C}_j} E(\tilde{y}_i|\mathcal{H}_1) - \gamma_j = \sum_{i:s_i\in\mathcal{C}_j}((1-P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i}w_{0,s_i}) - \gamma_j.$$

In this case,

$$\sigma^2_{\text{MD},j} = \frac{1}{n_{\mathcal{C}_j}}\sum_{i:s_i\in\mathcal{C}_j} P_{\text{MD},s_i}(1-P_{\text{MD},s_i})(w_{1,s_i}+w_{0,s_i})^2,$$

and $M_{\text{MD},j} = \max_{i:s_i\in\mathcal{C}_j}\{m_{\text{MD},i}\}$ where

$$m_{\text{MD},i} = \max\left\{\left|w_{1,s_i} - E\left(\tilde{y}_i|\mathcal{H}_1\right)\right|, \left|w_{0,s_i} + E\left(\tilde{y}_i|\mathcal{H}_1\right)\right|\right\}$$

$$= \max\{|P_{\text{MD},s_i}(w_{1,s_i} + w_{0,s_i})|, |(1 - P_{\text{MD},s_i})(w_{1,s_i} + w_{0,s_i})|\}.$$

We denote the resulting constants defined in Theorem 2 by $A_{\text{MD},j}$, $B_{\text{MD},j}$ and $\Lambda_{\text{MD},j}$. By the improved Bennet's inequality we have that

$$P_{\text{MD},j} \leq U\left(n_{\mathcal{C}_j}, \alpha_{\text{MD},j}, M_{\text{MD},j}, \sigma^2_{\text{MD},j}\right),$$

for every γ_j such that $0 \leq \sum_{i:s_i \in \mathcal{C}_j}((1 - P_{\text{MD},s_i})w_{1,s_i} - P_{\text{MD},s_i} w_{0,s_i}) - \gamma_j < n_{\mathcal{C}_j} M_{\text{MD},j}$.

APPENDIX D
PROOF OF PROPOSITION 3

Denote $\tilde{z}_j = \tau_j \left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1 - z_j)\right]$. We rewrite the false alarm probability in (12) as

$$P_{\text{FA}} = \Pr\left(\sum_{j=1}^{N_c} [\tilde{z}_j - E(\tilde{z}_j|\mathcal{H}_0)] \geq \gamma - \sum_{\sum_{j=1}^{N_c}} E(\tilde{z}_j|\mathcal{H}_0) \middle| \mathcal{H}_0\right).$$

By the law of total expectation on τ_j.

$$E\left(\tilde{z}_j|\mathcal{H}_0\right) = p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right],$$
$$E(\tilde{z}_j^2|\mathcal{H}_0) = p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w^2_{1,\mathcal{C}_j} + (1 - P_{\text{FA},\mathcal{C}_j}) w^2_{0,\mathcal{C}_j}\right],$$

It follows that

$$\sigma^2_{\text{FA},\mathcal{C}_j} \triangleq \text{var}\left(\tilde{z}_j - E\left(\tilde{z}_j|\mathcal{H}_0\right)|\mathcal{H}_0\right) = \text{var}\left(\tilde{z}_j|\mathcal{H}_0\right) = E\left(\tilde{z}_j^2|\mathcal{H}_0\right) - \left[E\left(\tilde{z}_j|\mathcal{H}_0\right)\right]^2$$
$$= p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w^2_{1,\mathcal{C}_j} + (1 - P_{\text{FA},\mathcal{C}_j}) w^2_{0,\mathcal{C}_j}\right] - p^2_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]^2.$$

We use Theorem 2 to upper bound the false alarm probability of the final decision of the FC by substituting j with i in Theorem 2 and

$$x_j = \tilde{z}_j - E\left(\tilde{z}_j|\mathcal{H}_0\right) = \tilde{z}_j - p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right],$$
$$\alpha_{FA} = \gamma - \sum_{j=1}^{N_c} E(\tilde{z}_j|\mathcal{H}_0) = \gamma - \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right].$$

In this case,

$$\sigma^2_{\text{FA}} = \frac{1}{N_c}\left[\sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w^2_{1,\mathcal{C}_j} + (1 - P_{\text{FA},\mathcal{C}_j}) w^2_{0,\mathcal{C}_j}\right] - \sum_{j=1}^{N_c} p^2_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1 - P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]^2\right]$$

and $M_{\text{FA}} = \max_{j \in \{1,\ldots,N_c\}}\{m_{\text{FA},j}\}$ where

$$m_{\text{FA},j} = \max\left\{\left|w_{1,\mathcal{C}_j} - E\left(\tilde{z}_j|\mathcal{H}_0\right)\right|, \left|w_{0,\mathcal{C}_j} + E\left(\tilde{z}_j|\mathcal{H}_0\right)\right|\right\}$$

$$= \max\{|w_{1,\mathcal{C}_j} - p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]|, \\ |w_{0,\mathcal{C}_j} + p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right]|\}.$$

We denote the resulting constants defined in Theorem 2 by A_{FA}, B_{FA} and Λ_{FA}. It follows from the improved Bennett's inequality that

$$P_{\text{FA}} \leq U\left(N_c, \alpha_{\text{FA}}, M_{\text{FA}}, \sigma^2_{\text{FA}}\right),$$

for every γ such that $0 \leq \gamma - \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[P_{\text{FA},\mathcal{C}_j} w_{1,\mathcal{C}_j} - (1-P_{\text{FA},\mathcal{C}_j}) w_{0,\mathcal{C}_j}\right] < N_c \cdot M_{\text{FA}}$.

APPENDIX E
PROOF OF PROPOSITION 4

Similarly to the proof presented in Appendix D, we can use Theorem 2 to upper bound the missed detection probability of the final decision of the FC. Recall that $\tilde{z}_j = \tau_j\left[w_{1,\mathcal{C}_j} z_j - w_{0,\mathcal{C}_j}(1-z_j)\right]$. We can rewrite the missed detection probability in (12) as

$$P_{\text{MD}} = \Pr\left(\sum_{j=1}^{N_c}\left[E\left(\tilde{z}_j|\mathcal{H}_1\right) - \tilde{z}_j\right] > \sum_{j=1}^{N_c} E\left(\tilde{z}_j|\mathcal{H}_1\right) - \gamma_j \middle| \mathcal{H}_1\right).$$

By the law of total expectation on τ_j

$$E\left(\tilde{z}_j|\mathcal{H}_1\right) = p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w_{1,\mathcal{C}_j} - P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}\right],$$
$$E(\tilde{z}_j^2|\mathcal{H}_1) = p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right].$$

It follows that

$$\sigma^2_{\text{MD},\mathcal{C}_j} \triangleq \text{var}\left(E\left(\tilde{z}_j|\mathcal{H}_1\right) - \tilde{z}_j|\mathcal{H}_1\right) = \text{var}\left(\tilde{z}_j|\mathcal{H}_1\right) = E\left(\tilde{z}_j^2|\mathcal{H}_1\right) - \left[E\left(\tilde{z}_j|\mathcal{H}_1\right)\right]^2 \\ = p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right] - p^2_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w_{1,\mathcal{C}_j} - P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}\right]^2.$$

We use Theorem 2 we upper bound the missed detection probability of the final decision of the FC by substituting j with i in Theorem 2 and

$$x_j = E\left(\tilde{z}_j|\mathcal{H}_1\right) - \tilde{z}_j = p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w_{1,\mathcal{C}_j} - P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}\right] - \tilde{z}_j,$$
$$\alpha_{MD} = \sum_{j=1}^{N_c} E(\tilde{z}_j|\mathcal{H}_1) - \gamma = \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right] - \gamma.$$

In this case,

$$\sigma^2_{\text{MD}} = \frac{1}{N_c}\left[\sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j}) w^2_{1,\mathcal{C}_j} + P_{\text{MD},\mathcal{C}_j} w^2_{0,\mathcal{C}_j}\right]\right.$$

$$-\sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}^2 \left[(1-P_{\text{MD},\mathcal{C}_j})w_{1,\mathcal{C}_j} - P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}\right]^2 \Bigg],$$

and $M_{\text{MD}} = \max_{j\in\{1,\ldots,N_c\}}\{m_{\text{MD},j}\}$ where

$$\begin{aligned} m_{\text{MD},j} &= \max\left\{\left|w_{1,\mathcal{C}_j} - E\left(\tilde{z}_j|\mathcal{H}_1\right)\right|, \left|w_{0,\mathcal{C}_j} + E\left(\tilde{z}_j|\mathcal{H}_1\right)\right|\right\} \\ &= \max\Big\{|w_{1,\mathcal{C}_j} - p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j})w_{1,\mathcal{C}_j}^2 + P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}^2\right]|, \\ &\qquad |w_{0,\mathcal{C}_j} + p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j})w_{1,\mathcal{C}_j}^2 + P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}^2\right]|\Big\}. \end{aligned}$$

We denote the resulting constants defined in Theorem 2 by A_{MD}, B_{MD} and Λ_{MD}. By the improved Bennet's inequality we have that

$$P_{\text{MD}} \leq U\left(N_c, \alpha_{\text{MD}}, M_{\text{MD}}, \sigma_{\text{MD}}^2\right),$$

for every γ such that $0 \leq \sum_{j=1}^{N_c} p_{\text{com},\mathcal{C}_j}\left[(1-P_{\text{MD},\mathcal{C}_j})w_{1,\mathcal{C}_j}^2 + P_{\text{MD},\mathcal{C}_j} w_{0,\mathcal{C}_j}^2\right] - \gamma < N_c M_{\text{MD}}$.

REFERENCES

[1] M. Yemini, S. Gil, and A. Goldsmith, "Exploiting local and cloud sensor fusion in intermittently connected sensor networks," in *2020 IEEE Global Communications Conference (GLOBECOM 2020)*, 2020, pp. 1–7.

[2] J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Distributed anonymous discrete function computation," *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2276–2289, 2011.

[3] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 501–510, July 1981.

[4] J. N. Tsitsiklis, "Decentralized detection," in *In Advances in Statistical Signal Processing*. JAI Press, 1993, pp. 297–344.

[5] N. Katenka, E. Levina, and G. Michailidis, "Local vote decision fusion for target detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 329–338, Jan 2008.

[6] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Data fusion trees for detection: Does architecture matter?" *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4155–4168, Sep. 2008.

[7] ——, "Bayesian detection in bounded height tree networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4042–4051, Oct 2009.

[8] G. Ferrari, M. Martalo, and R. Pagliari, "Decentralized detection in clustered sensor networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 959–973, April 2011.

[9] S. A. Aldalahmeh1, M. Ghogho, D. McLernon, and E. Nurellari, "Optimal fusion rule for distributed detection in clustered wireless sensor networks," *EURASIP J. Adv. Signal Process.*, vol. 5, 2016.

[10] S. A. Aldalahmeh, S. O. Al-Jazzar, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Fusion rules for distributed detection in clustered wireless sensor networks with imperfect channels," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 585–597, Sep. 2019.

[11] M. Shirazi and A. Vosoughi, "On distributed estimation in hierarchical power constrained wireless sensor networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 442–459, 2020.

[12] F. Khan and Z. Pi, "mmWave mobile broadband (mmb): Unleashing the 3–300ghz spectrum," in *34th IEEE Sarnoff Symposium*, May 2011, pp. 1–6.

[13] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, June 2011.

[14] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: Potentials and challenges,” *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[15] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[16] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, E. Aryafar, S. Yeh, N. Himayat, S. Andreev, and Y. Koucheryavy, “Analysis of human-body blockage in urban millimeter-wave cellular communications,” in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–7.

[17] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy, “On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 124–10 138, Nov 2017.

[18] Y. Yan and Y. Mostofi, “Co-optimization of communication and motion planning of a robotic operation under resource constraints and in fading environments,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1562–1572, April 2013.

[19] M. M. Zavlanos, M. B. Egerstedt, and G. J. Pappas, “Graph-theoretic connectivity control of mobile robot networks,” *Proc. IEEE*, vol. 99, no. 9, pp. 1525–1540, Sep. 2011.

[20] N. Michael, M. M. Zavlanos, V. Kumar, and G. J. Pappas, “Maintaining connectivity in mobile robot networks,” in *Experimental Robotics*, 2009.

[21] S. Gil, S. Kumar, D. Katabi, and D. Rus, “Adaptive communication in multi-robot systems using directionality of signal strength,” *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 946–968, 2015.

[22] Biao Chen and P. K. Willett, “On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 693–699, 2005.

[23] Ruixin Niu, Biao Chen, and P. K. Varshney, “Fusion of decisions transmitted over rayleigh fading channels in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, 2006.

[24] K. Cohen and A. Leshem, “Energy-efficient detection in wireless sensor networks using likelihood ratio and channel state information,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1671–1683, 2011.

[25] D. Ciuonzo, G. Romano, and P. S. Rossi, “Channel-aware decision fusion in distributed mimo wireless sensor networks: Decode-and-fuse vs. decode-then-fuse,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2976–2985, 2012.

[26] I. Nevat, G. W. Peters, and I. B. Collings, “Distributed detection in sensor networks over fading channels with multiple antennas at the fusion centre,” *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 671–683, 2014.

[27] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[28] M. Zhang, Y. Hong, and N. Balakrishnan, “The generalized poisson-binomial distribution and the computation of its distribution function,” *Journal of Statistical Computation and Simulation*, vol. 88, no. 8, pp. 1515–1527, 2018.

[29] R. Arratia and L. Gordon, “Tutorial on large deviations for the binomial distribution,” *Bulletin of Mathematical Biology*, vol. 51, no. 1, pp. 125 – 131, 1989.

[30] G. Bennett, “Probability inequalities for the sum of independent random variables,” *Journal of the American Statistical Association*, vol. 57, no. 297, pp. 33–45, 1962.

[31] S. Zheng, “An improved Bennett’s inequality,” *Communications in Statistics - Theory and Methods*, vol. 47, no. 17, pp. 4152–4159, 2018.